Аналитическая Геометрия
и
Линейная Алгебра

Конспект по 1 курсу специальности «прикладная математика»

(лектор А. В. Филипцов)

Оглавление

Часть I

Аналитическая геометрия

Глава 1

Метод координат

1.1 Величина направленного отрезка. Теорема Шаля. Декартова система координат на прямой.

• ***Отрезком*** *называется часть прямой, ограниченная двумя точками. Отрезок называется* **направленным,** *если указано, какая из граничных точек является начальной и какая конечной (Обозначения: $\overline{AB}$ – направленный отрезок; $|\overline{AB}|$ – длина направленного отрезка).*

• *Отрезок называется* ***нулевым***, *если его начальная и конечная точки совпадают.*

Длина нулевого направленного отрезка равна нулю.

• *Прямая линия с указанным на ней направлением называется* ***осью***.

• *Пусть дана некоторая ось Δ, и на ней расположен направленный отрезок $\overline{AB}$,* ***величиной*** *которого на оси Δ называется число, равное:*

$$AB = \begin{cases} |\overline{AB}|, \ \overline{AB} \uparrow\uparrow \Delta, \\ -|\overline{AB}|, \ \overline{AB} \uparrow\downarrow \Delta; \end{cases}$$

Свойства величины направленного отрезка:

1. *Величина AB противоположна величине BA.*
2. *Длина $|\overline{AB}|$ равна модулю величин AB.*

Теорема Шаля. *При любом расположении трех точек A, B и C на оси справедливо равенство $AB + BC = AC$.*

♦

1. Пусть точка B лежит между точками A и C, тогда направленные отрезки $\overline{AB}$, $\overline{BC}$ и $\overline{AC}$ сонаправлены. При этом $|\overline{AB}| + |\overline{BC}| = |\overline{AC}|$.
   (a) Если направляющие оси совпадают с направлениями отрезков, то $|\overline{AB}| = AB$, $|\overline{BC}| = BC$, $|\overline{AC}| = AC$. Следовательно, $AB + BC = AC$.
   (b) Если направляющие оси не совпадают с направлениями отрезков, то $|\overline{AB}| = -AB$, $|\overline{BC}| = -BC$, $|\overline{AC}| = -AC$. Следовательно, $-AB + -BC = -AC \Rightarrow AB + BC = AC$.

2. Если точка C лежит между точками A и B, то, по доказанному выше, $AC + CB = AB \Rightarrow AC = AB - CB = AB + BC$.

3. Если точка A лежит между точками B и C, то $BA + AC = BC \Rightarrow AC = -BA + BC = AB + BC$.

Теорема остается справедливой, даже если некоторые точки совпадают. ⊠

Пусть на прямой Δ дана точка O, заданы единица масштаба и определенное направление.

• *Говорят, что на прямой* Δ *задана* ***декартовая система координат****. При этом ось* Δ *называется* ***координатной осью****, а точка* O *–* ***началом координат****.*

Пусть точка A — некоторая фиксированная точка на прямой Δ.

• ***Координатой точки*** A *называется число, равное величине* OA *направленного отрезка* $\overline{OA}$*. Обозначение:* X_A.

Точки, расположенные по одну сторону от начала координат, имеют одинаковые по знаку координаты. Полуось, на которой все точки имеют положительные координаты, называют положительной полуосью. Аналогично с отрицательной полуосью.

Теорема. *Для любых двух точек* A *и* B *на* Δ *выполняется:* $AB = X_B - X_A$, $|\overline{AB}| = |X_B - X_A|$

⧫ Рассмотрим величину направленного отрезка $\overline{AB}$. Тогда, по теореме Шаля, $AB = AO + OB = OB - OA = X_B - X_A$ ⊠

Пусть на оси Δ заданы 2 различные точки: A с координатой X_A и B с координатой X_B.

• *Говорят, что точка* C ***делит*** *направленный отрезок* AB *в отношении* λ*, если* $\lambda = \dfrac{AC}{CB}$*. При этом число* λ *называется* ***простым отношением*** *трех точек* A, C, B.

1. Если точка C лежит между точками A и B, то направляющие отрезки AC и CB сонаправлены. Следовательно, знак у них одинаковый и $\lambda > 0$. При этом говорят, что точка C делит направленный отрезок AB внутренним образом.

2. Если точка C лежит вне отрезка AB, то говорят, что она делит отрезок AB внешним образом. При этом отрезки AC и CB противоположно направлены (не сонаправлены) и, следовательно, $\lambda < 0$.

3. Если $\lambda = 0$, то $AC = 0$, следовательно, $|\overline{AC}| = 0$ и точки A и C совпадают.

4. Если точка C лежит за концом отрезка AB, то $\left|\dfrac{AC}{CB}\right| = \dfrac{|AC|}{|CB|} = \dfrac{|\overline{AC}|}{|\overline{CB}|} < 1 \Rightarrow \lambda < -1$ $(|\overline{AC}| > |\overline{BC}|)$.

5. Если точка C лежит перед началом отрезка AB, то $|\overline{AC}| < |\overline{BC}|$, тогда $-1 < \lambda < 0$.

Заметим, что $\lambda \neq -1$, так как в этом случае $-1 = \frac{AC}{CB}$. Следовательно, $AC = -CB \Rightarrow AC = CB \Rightarrow A$ и B совпадают, что противоречит выбору A и B.

Теорема. *Координата точки C, делящей направленный отрезок AB в отношении λ равна*

$$X_C = \frac{X_A + \lambda X_B}{1+\lambda}$$

♦ Из определения деления отрезка в данном отношении следует:

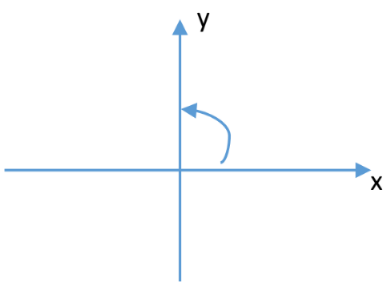
$$\lambda = \frac{AC}{CB} = \frac{X_C - X_A}{X_B - X_C} \Rightarrow (X_B - X_C)\cdot\lambda = X_C - X_A \Rightarrow \lambda X_B + X_A = X_C + \lambda X_C = X_C(1+\lambda) \Rightarrow$$

$$X_C = \frac{X_A + \lambda X_B}{1+\lambda}$$ ⊠

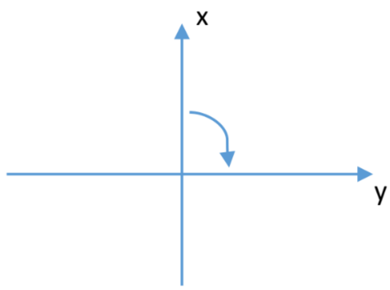
1.2 Декартова прямоугольная система координат на плоскости, в пространстве (ДПСК).

• ***Декартовой прямоугольной системой координат на плоскости*** *называются две взаимно перпендикулярные координатные оси с общим началом координат. Ось Ox называется осью* ***абсцисс****, Oy — осью* ***ординат****.*

• *Система координат называется* ***правой****, если кратчайший поворот от оси X к оси Y против часовой стрелки, иначе* ***левой****.*



Правая ДПСК



Левая ДПСК

Пусть M — некоторая точка плоскости, M_x и M_y — проекции на оси Ox и Oy соответственно.

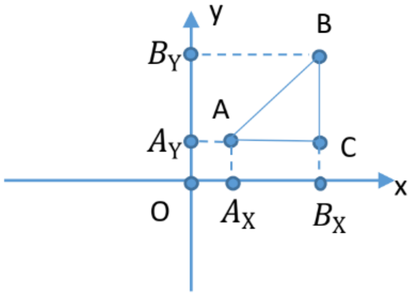
• ***Декартовыми прямоугольными координатами*** *точки $M(X_M, Y_M)$ называются величины направленных отрезков $\overline{OM_x}$ и $\overline{OM_y}$ на координатных осях, то есть $X_M = OM_x$, $Y_M = OM_y$.*

Пусть $A(X_A, Y_A)$ и $B(X_B, Y_B)$ — две различные точки на плоскости.

Теорема. *Длина направленного отрезка $\overline{AB}$ равна:*

$$|\overline{AB}| = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$$

♦ Построим проекции точек A и B на координатные оси. Обозначим через C точку пересечения проекций.



По теореме Пифагора $|AB|^2 = |AC|^2 + |BC|^2$. Тогда $|\overline{AB}| = \sqrt{|\overline{A_xB_x}|^2 + |\overline{A_yB_y}|^2} =$
$= \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2}$.

Рассмотрим случай, когда отрезок AB параллелен Ox (или AB параллелен Oy). Тогда $Y_A = Y_B$ и $\sqrt{(X_A - X_B)^2 + \underbrace{(Y_A - Y_B)^2}_{=0}} = |X_A - X_B| = |\overline{AB}| = |\overline{A_xB_x}|$. ⊠

Проведем через точки A и B некоторую ось. Тогда имеет смысл понятие «деление отрезка в данном отношении».

Теорема. *Координаты точки C, делящей $\overline{AB}$ в отношении λ, равны:*

$$X_C = \frac{X_A + \lambda X_B}{1+\lambda}, \; Y_C = \frac{Y_A + \lambda Y_B}{1+\lambda}$$

⧫ Построим проекции точек A, B и C на ось Ox. Заметим, что при любом расположении точки C точка C_x будет иметь такой же характер расположения, как и точка C (если C лежит между A и B, то и C_x будет лежать между A_x и B_x). Следовательно, $\frac{|\overline{AC}|}{|\overline{CB}|}$ и $\frac{\overline{A_xC_x}}{\overline{C_xB_x}}$ имеют один знак.

По теореме Фалеса длина $\frac{|\overline{A_xC_x}|}{|\overline{C_xB_x}|} = \frac{|\overline{AC}|}{|\overline{CB}|} \Rightarrow \left|\frac{A_xC_x}{C_xB_x}\right| = \left|\frac{AC}{CB}\right|$. Следовательно, $\frac{A_xC_x}{C_xB_x} = \frac{AC}{CB} = \lambda$ и точка C_x делит направленный отрезок $\overline{A_xB_x}$ также в отношении λ. Следовательно, $X_C = \frac{X_A + \lambda X_B}{1+\lambda}$.

Для ординаты точки C равенство доказывается аналогично при рассмотрении случая, когда прямая AB параллельна одной из осей. ⊠

• ***ДПСК в пространстве*** *образуют три взаимно перпендикулярные координатные оси с общим началом координат. Третья ось Oz называется осью* ***аппликатой***.

• *Система координат называется* ***правой****, если, глядя со стороны направления Oz, кратчайший поворот от оси Ox к оси Oy осуществляется против часовой стрелки, иначе —* ***левой***.

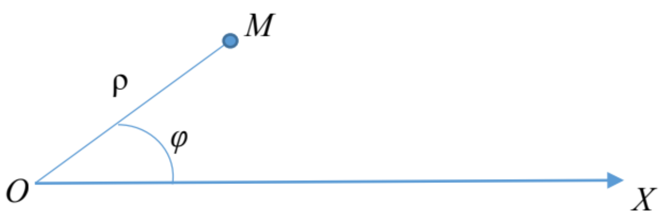
Пусть $A(X_A, Y_A, Z_A)$ и $B(X_B, Y_B, Z_B)$ — две различные точки пространства. Тогда

1. $|\overline{AB}| = \sqrt{(X_A - X_B)^2 + (Y_A - Y_B)^2 + (Z_A - Z_B)^2}$
2. Если C делит $\overline{AB}$ в отношении λ, то

$$X_C = \frac{X_A + \lambda X_B}{1+\lambda}, \; Y_C = \frac{Y_A + \lambda Y_B}{1+\lambda}, \; Z_C = \frac{Z_A + \lambda Z_B}{1+\lambda}.$$

1.3 Полярные, цилиндрические и сферические координаты. Общая декартова система координат.

• ***Полярную систему координат*** *образуют точка, выходящий из неё луч и единичный отрезок. При этом точка O называется* ***полюсом****, а луч Ox — полярной осью.*



Пусть M — произвольная точка плоскости.

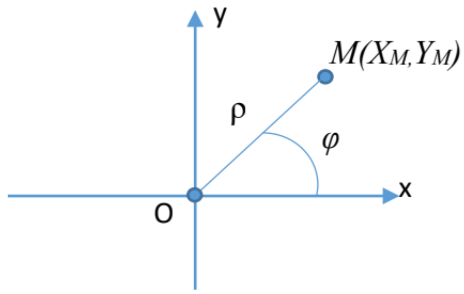
• ***Полярными координатами точки*** *M называются два числа ρ, φ, где ρ — длина OM ($\rho = |\overline{OM}|$), а угол φ равен углу, на который необходимо повернуть полярную ось вокруг точки O против часовой стрелки до совмещения ее с отрезком OM.*

• *Число ρ называется* ***полярным радиусом*** *точки. Число φ —* ***полярным углом*** *точки.*

Заметим, что $\rho \geqslant 0$, а φ определён с точностью до $2\pi k$, где $k \in \mathbb{Z}$. Точка O имеет $\rho = 0$, а φ для неё не определён (ему можно приписать любое направление).

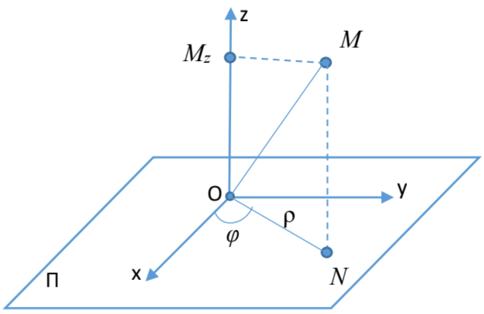
Значение φ, принадлежащее $[0; 2\pi)$, называется главным значением φ.
Полярной осью называют точку и выходящий из неё луч.

Покажем связь между декартовой и полярной системами координат.
Расположим ДПСК и ПСК следующим образом: полюс поместим в начало координат, а полярную ось совместим с положительной полуосью Ox. Расположенные таким образом ДПСК и ПСК называют **соответствующими друг другу**.



Пусть M произвольная точка плоскости. Тогда полярные координаты точки M равны:
$X_M = OM_x = |OM| \cdot \cos\varphi = \rho \cdot \cos\varphi$,
$Y_M = OM_y = |OM| \cdot \sin\varphi = \rho \cdot \sin\varphi$, где $\rho = \sqrt{X_M^2 + Y_M^2}$, а угол φ можно найти, решив следующую систему: $\begin{cases} \cos\varphi = \dfrac{X_M}{\rho} = \dfrac{X_M}{\sqrt{X_M^2 + Y_M^2}}, \\ \\ \sin\varphi = \dfrac{Y_M}{\rho} = \dfrac{Y_M}{\sqrt{X_M^2 + Y_M^2}}; \end{cases}$

Цилиндрическая и сферическая системы координат в пространстве вводятся следующим образом: в пространстве зафиксируем некоторую плоскость П. Построим на ней полярную систему координат, а так же координатную ось Oz с началом в точке O и перпендикулярную плоскости П.



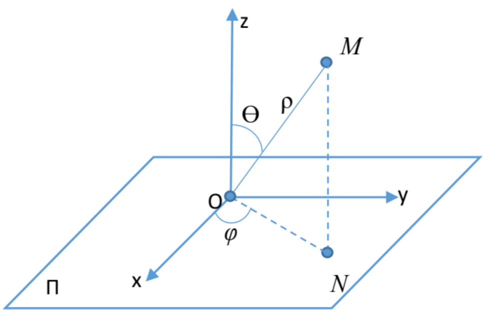
И пусть точка M — произвольная точка пространства, а точка N — её проекция на плоскость П.

• ***Цилиндрическими координатами точки*** *M называются числа ρ, φ, z, где ρ и φ — полярные координаты точки N относительно построенной полярной системы координат, а z — величина направленного отрезка $\overline{OM_z}$ ($z = OM_z$).*

• ***Сферическими координатами точки*** *M называются три числа ρ, φ, Θ, где ρ — длина отрезка $|\overline{OM}|$, φ — полярный радиус точки N относительно построенной полярной системы координат, а Θ — угол между направленным отрезком $\overline{OM}$ и осью Oz.*

Из определений следует, что $\rho \geqslant 0$ и для цилиндрических, и для сферических координат, $\Theta \in [0; \pi]$, а φ определён с точностью до $2\pi n, n \in \mathbb{Z}$.

Построим в пространстве ДПСК следующим образом: начало координат совпадает с точкой O, ось Oz совместим с построенной ранее осью Oz, а положительное направление оси Ox — с полярной осью.

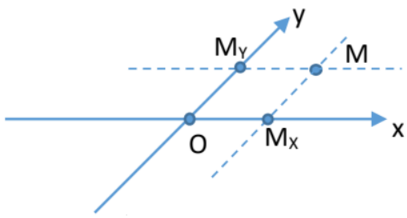


Тогда точка M имеет следующие координаты в системах:

$$\text{цилиндрическая: } \begin{cases} X_M = \rho \cdot \cos\varphi, \\ Y_M = \rho \cdot \sin\varphi, \\ Z_M = z; \end{cases} \quad \text{сферическая: } \begin{cases} X_M = \rho \cdot \cos\varphi \cdot \sin\Theta, \\ Y_M = \rho \cdot \sin\varphi \cdot \sin\Theta, \\ Z_M = \rho \cdot \cos\Theta. \end{cases}$$

• ***Общую декартову систему координат (аффинную систему координат)*** *на плоскости образуют две пересекающие координатные оси с общим началом координат.*

Пусть точка M — произвольная точка плоскости. Построим через точку M прямые, параллельные координатным осям.



• ***Общими декартовыми координатами точки*** *M называются величины направленных отрезков $\overline{OM_x}$ и $\overline{OM_y}$.*

Глава 2

Векторы

2.1 Понятие вектора. Линейные операции над векторами.

• ***Вектором*** *называется множество всех одинаково направленных отрезков равной длины. Обозначение:* a, b, c *или* $\vec{a}, \vec{b}, \vec{c}$.

Пусть $\overline{AB}$ — направленный отрезок. Тогда существует единственный вектор, которому принадлежит этот отрезок. То есть любой направленный отрезок $\overline{AB}$ однозначно определяет вектор, который обозначается как $\overrightarrow{AB}$.

Пусть заданы вектор $\vec{a}$ и некоторая точка A. Тогда существует единственная точка B, такая, что $a = \overrightarrow{AB}$.

• *Операция построения такой точки B называется* ***откладыванием*** *вектора $\vec{a}$ от точки A.*

• ***Длиной вектора*** *$\vec{a}$ называется длина любого из направленных отрезков среди его образующих. Обозначение: $|\vec{a}|$.*

• *Вектор длины 0 называется* ***нулевым****. Обозначение: $\vec{0}$. Вектор длины 1 называется* ***единичным*** *или* ***ортом****.*

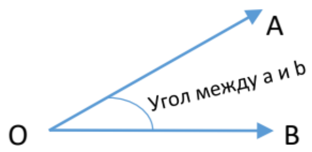
• *Векторы называются* ***коллинеарными****, если образующие их отрезки параллельны некоторой прямой.*

• *Векторы называются* ***компланарными****, если образующие их отрезки параллельны некоторой плоскости.*

Заметим, что нулевой вектор и коллинеарен любому вектору, и компланарен любым двум другим векторам.

Пусть a и b — 2 ненулевых вектора, отложенных от некоторой точки O. То есть найдём такие точки A, и B, что $a = \overrightarrow{OA}, b = \overrightarrow{OB}$.

• ***Углом между векторами*** *a и b называертся угол между направленными отрезками $\overline{OA}$ и $\overline{OB}$.*

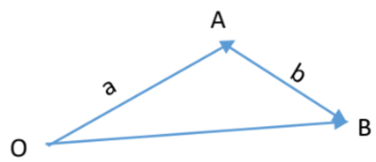


Из определения следует, что угол между двумя векторами $\in [0; \pi]$.

Линейные операции над векторами:

- сложение;
- умножение на число.

Пусть a и b — два вектора. Возьмем некоторую точку O и отложим от нее вектор $a = \overrightarrow{OA}$. Затем от точки A отложим вектор $b = \overrightarrow{AB}$.



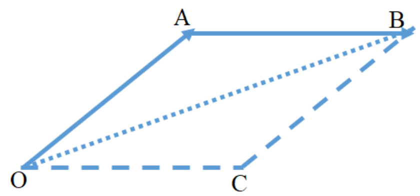
• ***Суммой векторов*** *a и b будет называться вектор $\overrightarrow{OB}$. Обозначение: $a + b$.*

• ***Произведением вектора*** *a на действительное число A называется вектор, длина которого равна $|A| \cdot |a|$, а направление совпадает с направлением вектора a, если $A > 0$, и противоположно направлению вектора a, если $A < 0$. Обозначение: $A \cdot a$.*

Свойства линейных операций:

1. $a + b = b + a$.

   ♦ Пусть $a = \overrightarrow{OA}$, $b = \overrightarrow{AB}$. На направленных отрезках $\overline{OA}$ и $\overline{AB}$ построим параллелограмм $OABC$.

   Заметим, что векторы $\overrightarrow{CB}$ и $\overrightarrow{OA}$ сонаправлены и $|\overline{CB}| = |\overline{OA}|$. Следовательно, $\overrightarrow{CB} = \overrightarrow{OA} = a$. Путем аналогичных рассуждений получим $\overrightarrow{OC} = \overrightarrow{AB} = b$.
   По определению суммы векторов, вектор $\overrightarrow{OB}$ можно, с одной стороны, рассматривать как сумму векторов $\overrightarrow{OA} + \overrightarrow{AB}$, а с другой стороны, как сумму $\overrightarrow{OC} + \overrightarrow{CB}$. Следовательно, $\overrightarrow{OB} = \begin{cases} \overrightarrow{OA} + \overrightarrow{AB} = a + b, \\ \overrightarrow{OC} + \overrightarrow{CB} = \overrightarrow{AB} + \overrightarrow{OA} = b + a; \end{cases} \Rightarrow a + b = b + a.$ ⊠

2. $(a + b) + c = a + (b + c)$.

3. $\vec{a} + \vec{0} = \vec{a}, \quad \forall \vec{a}$.

4. $\alpha(a+b) = \alpha a + \alpha b$,
   $(\alpha+\beta)a = \alpha a + \beta a$,
   $\alpha(\beta a) = (\alpha\beta)a$, где a, b — векторы, α, β — числа.

• *Вектор b называется* ***противоположным*** *вектору a, если его длина совпадает с длиной вектора a, а направление противоположно.*

Из определения следует, что

1. $-1 \cdot a = -a$, где вектор $(-a)$ — противоположный вектору a;

2. $1 \cdot a = a$.

• ***Разностью векторов*** *a и b называется вектор, прибавив который к вектору b, получим вектор a, то есть вектор $c = a - b$, где $c + b = a$.*

Правило вычитания: *Чтобы из вектора a вычесть вектор b, необходимо к вектору a прибавить вектор, противоположный вектору b, то есть $a - b = a + (-b)$.*

♦ $(a + (-b)) + b = a + (b + (-b)) = a + \vec{0} = a$. ⊠

2.2 Координаты вектора.

Пусть $a_1, \ldots, a_k$ — произвольная система векторов. $\alpha_1, \ldots, \alpha_k$ — некоторые действительные числа.

• *Вектор $\alpha_1 a_1 + \ldots + \alpha_k a_k$ называется линейной комбинацией векторов $a_1, \ldots, a_k$ с коэффициентами $\alpha_1, \ldots, \alpha_k$.*

Если вектор представлен в виде линейной комбинации системы векторов (то есть $b = \alpha_1 a_1 + \ldots + \alpha_k a_k$), то говорят, что данный вектор разложим по этим векторам системы.

• ***Базисом на прямой*** *называется любой ненулевой вектор, параллельный этой прямой.*

• ***Базисом на плоскости*** *называется упорядоченная пара неколлинеарных векторов.*

• ***Базисом в пространстве*** *называется упорядоченная тройка некомпланарных векторов.*

Нулевой вектор не может быть элементом базиса.

Теорема.

1. *Любой вектор, параллельный прямой, может быть единственным образом разложен по базису этой прямой.*

2. *Любой вектор, параллельный плоскости, может быть единственным образом разложен по базису на этой плоскости.*

3. *Любой вектор в пространстве может быть единственным образом по базису в пространстве.*

♦ Покажем существование разложения для каждого из пунктов.

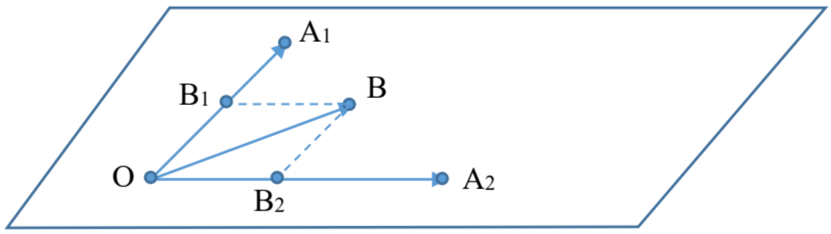
1. Пусть Δ — некоторая прямая, a — базис на ней, b — вектор, параллельный Δ. Тогда $a \parallel b \Rightarrow a$ и b коллинеарны.

   Пусть e_a, e_b — единичные векторы, параллельные векторам a и b соответственно. Так как $e_a = \dfrac{1}{|a|}a$, $e_b = \dfrac{1}{|b|}b$ и a и b коллинеарны, то и e_a и e_b также коллинеарны. Следовательно, они либо противоположные, либо равные.

   Если они равные, то $e_a = e_b \Rightarrow \dfrac{1}{|a|}a = \dfrac{1}{|b|}b \Rightarrow b = \dfrac{|b|}{|a|}a =$[замена $\dfrac{|b|}{|a|}$ на l]$= la$

   Если они противоположные, то $e_a = -e_b \Rightarrow \dfrac{1}{|a|}a = -\dfrac{1}{|b|}b \Rightarrow b = -\dfrac{|b|}{|a|}a =$[замена $-\dfrac{|b|}{|a|}$ на l]$= la \Rightarrow$ вектор b разложим по базису a.

2. Пусть Π — некоторая плоскость, a_1 и a_2 — базис на плоскости, $b \parallel \Pi \Rightarrow a_1,\ a_2,\ b$ — компланарны. Отложим векторы базиса от некоторой точки O на плоскости Π, то есть $a_1 = \overrightarrow{OA_1}, a_2 = \overrightarrow{OA_2}, b = \overrightarrow{OB}$.

   Построим прямую, проходящую через точку B, параллельную OA_2. Пусть B_1 — точка пересечения построенной прямой и прямой OA_1. Аналогично на OA_2 построим точку B_2.

   По определению суммы векторов $b = \overrightarrow{OB} = \overrightarrow{OB_1} + \overrightarrow{OB_2}$. Тогда $OB_1 \parallel OA_1$, а ненулевой вектор $\overrightarrow{OA_1}$ является на этой прямой базисом. Следовательно, по доказанному выше в пункте 1, $\overrightarrow{OB_1} = l_1 \cdot \overrightarrow{OA_1} = l_1 \cdot a_1$. По аналогичным рассуждениям получим $\overrightarrow{OB_2} = l_2 \cdot a_2 \Rightarrow b = l_1 \cdot a_1 + l_2 \cdot a_2$.

3. Доказательство данного пункта проводится по аналогичной схеме с пунктом 2. Покажем единственность: рассмотрим самый сложный вариант, в пространстве. На плоскости и прямой доказательства будут аналогичными.

   Пусть $a_1,\ a_2,\ a_3$ — базис в пространстве, а b — некоторый вектор пространства. И пусть существуют 2 различных разложения: $\begin{cases} b = l_1a_1 = l_2a_2 + l_3a_3, \\ b = \beta_1a_1 + \beta_2a_2 + \beta_3a_3; \end{cases} \Rightarrow$ [вычтем из первого второе] $\Rightarrow 0 = (l_1 - \beta_1)a_1 + (l_2 - \beta_2)a_2 + (l_3 - \beta_3)a_3$.

   Пусть $l_i - \beta_i \neq 0 \Rightarrow a_1 = -\dfrac{l_2 - \beta_2}{l_1 - \beta_1}a_2 - \dfrac{l_3 - \beta_3}{l_1 - \beta_1}a_3 \Rightarrow a_1,\ a_2,\ a_3$ компланарные, что является противоречием с тем, что базис в пространстве — это 3 некомпланарных вектора $\Rightarrow l_i - \beta_i = 0 \Rightarrow$ два разложения совпадают.

$\boxtimes$

Пусть a — некоторый вектор в пространстве. e_1, e_2, e_3 — некоторый базис в пространстве. По теореме выше вектор a единственным образом представим в виде $l_1e_1 + l_2e_2 + l_3e_3$.

• *Коэффициенты разложения вектора a по базису e_1, e_2, e_3 называются* ***координатами вектора*** *a* ***в базисе*** *e_1, e_2, e_3. Обозначение: $a(l_1, l_2, l_3)$.*

Теорема. *При умножении вектора на число все его координаты умножаются на это число. При сложении векторов их соответствующие координаты складываются.*

♦ Пусть a, b — произвольные векторы, e_1, e_2, e_3 — базис. (l_1, l_2, l_3) и $(\beta_1, \beta_2, \beta_3)$ — координаты векторов a и b в базисе e_1, e_2, e_3 соответственно. Тогда
$l \cdot a = l \cdot (l_1e_1 + l_2e_2 + l_3e_3) = l \cdot l_1e_1 + l \cdot l_2e_2 + l \cdot l_3e_3$;
$a + b = (l_1e_1 + l_2e_2 + l_3e_3) + (\beta_1e_1 + \beta_2e_2 + \beta_3e_3) = (l_1 + \beta_1)e_1 + (l_2 + \beta_2)e_2 + (l_3 + \beta_3)e_3$.
Получили векторы с координатами $la(l \cdot l_1, l \cdot l_2, l \cdot l_3),\ a + b(l_1 + \beta_1, l_2 + \beta_2, l_3 + \beta_3)$. $\boxtimes$

Следствие (критерий коллинеарности векторов). *Два ненулевых вектора a и b коллинеарны $\Longleftrightarrow$ их координаты пропорциональны.*

♦ $\Rightarrow$) Пусть $a(l_1, l_2, l_3), b(\beta_1, \beta_2, \beta_3)$ — 2 ненулевых коллинеарных вектора $\Rightarrow$ вектор a является базисом на прямой $\Rightarrow$ b представим в виде: $b = la \Rightarrow b$ имеет координаты $(l \cdot l_1, l \cdot l_2, l \cdot l_3) \Rightarrow$ координаты пропорциональны.

$\Leftarrow$) Пусть координаты пропорциональны: $l \cdot l_i = \beta_i \Rightarrow b = \beta_1e_1 + \beta_2e_2 + \beta_3e_3 = l \cdot l_1e_1 + l \cdot l_2e_2 + l \cdot l_3e_3 = l \cdot (l_1e_1 + l_2e_2 + l_3e_3) = la \Rightarrow$ вектор a либо $\uparrow\uparrow b$, либо $\uparrow\downarrow b \Rightarrow a$ коллинеарен b. $\boxtimes$

2.3 Декартовы координаты вектора.

• *Рассмотрим ДПСК $Oxyz$, и пусть i, j, k — единичные векторы, сонаправленные с осями Ox, Oy, Oz. Такие векторы называют* ***направляющими ортами*** *осей координат.*

• *Векторы i, j, k некомпланарны, следовательно, они образуют базис в пространстве. Данный базис является* ***ортонормированным базисом****.*

Пусть a — произвольный вектор пространства. Тогда по первой теореме из предыдущего параграфа $a = x \cdot i + y \cdot j + z \cdot k$, где (x, y, z) — декартовы координаты вектора a. То есть декартовы координаты вектора — это коэффициенты при разложении этого вектора по ортонормированному базису.

Рассмотрим некоторый вектор a и некоторую ось Δ. Отложим вектор a от точки A и пусть $a = \overrightarrow{AB}$. Построим проекции A и B на ось Δ и пусть это будут точки A_1 и B_1.

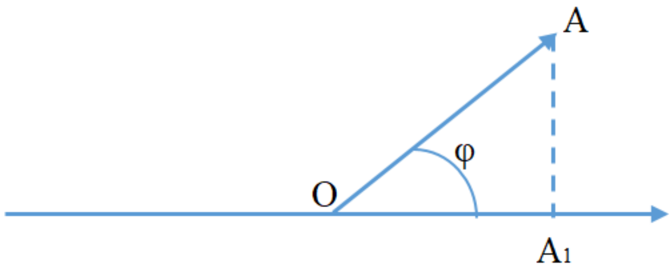
• *Полученный в результате направленный отрезок $\overline{A_1B_1}$ определяет вектор $\overrightarrow{A_1B_1}$, который называется* ***проекцией*** *вектора a на ось Δ.*

• *Величина направленного отрезка $\overline{A_1B_1}$ называется* ***величиной проекции*** *вектора a на ось Δ. Обозначение: $ПР_\Delta a$*

Свойства величин проекции вектора на ось:

1. $\text{ПР}_\Delta a = |a| \cdot \cos\varphi$, где φ — угол между вектором a и осью Δ.

   ♦ Отложим от некоторой точки O на оси Δ вектор a (то есть $a = \overrightarrow{OA}$). A_1 — проекция точки A. Тогда $\text{ПР}_\Delta a = OA_1$.

   Если угол φ острый, то $\text{ПР}_\Delta a = |\overline{OA_1}| \Rightarrow \text{ПР}_\Delta a = |\overline{OA}| \cdot \cos(\varphi) = |a| \cdot \cos(\varphi)$.

   Если угол φ тупой, то $\text{ПР}_\Delta a = -|\overline{OA_1}| = -|\overline{OA}| \cdot \cos(\pi - \varphi) = |\overline{OA}| \cdot \cos(\varphi) = |a| \cdot \cos(\varphi)$ ⊠

2. $\text{ПР}_\Delta (la) = l \cdot \text{ПР}_\Delta a$.

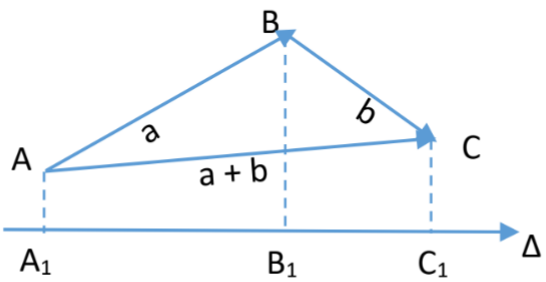
   ♦ По первому свойству, $\text{ПР}_\Delta (la) = |la| \cdot \cos(\varphi_1) = |l| \cdot |a| \cdot \cos(\varphi_1)$, где φ_1 — угол между Δ и la.

   Если $l > 0$, то $|l| = l$ и φ_1 является углом между Δ и a (равен углу φ). Следовательно, $\text{ПР}_\Delta (la) = l \cdot \text{ПР}_\Delta a$.

   Если $l < 0$, то $|l| = -l$ и φ_1 равен углу $\pi - \varphi$. $\cos(\pi - \varphi) = -\cos(\varphi)$. Следовательно, $\text{ПР}_\Delta (la) = l \cdot \text{ПР}_\Delta a$. ⊠

3. $\text{ПР}_\Delta (a + b) = \text{ПР}_\Delta a + \text{ПР}_\Delta b$.

   ♦ Пусть $a = \overrightarrow{AB}, b = \overrightarrow{BC}$ и $a + b = \overrightarrow{AC}$. Построим на ось Δ проекции точек A, B и C (обозначим как A_1, B_1 и C_1 соответственно).

   Заметим, что

   (a) $\text{ПР}_\Delta (a + b) = A_1C_1$,
   (b) $\text{ПР}_\Delta a = A_1B_1$,
   (c) $\text{ПР}_\Delta b = B_1C_1$.

   По теореме Шаля, $A_1C_1 = A_1B_1 + B_1C_1 \Rightarrow \text{ПР}_\Delta (a + b) = \text{ПР}_\Delta a + \text{ПР}_\Delta b$. ⊠

Теорема. *Декартовы координаты вектора равны величинам проекции этого вектора на координатные оси.*

♦ Пусть a — некоторый вектор пространства с координатами x, y, z. Отложим a от начала координат, то есть $= OM$, а затем построим проекции точки M на оси Ox, Oy, Oz. Тогда $\overrightarrow{OM} = \overrightarrow{OM_x} + \overrightarrow{OM_y} + \overrightarrow{OM_z}$. Так как $OM_x \parallel i \parallel Ox$ и, при этом, вектор i ненулевой, он является базисом и $\overrightarrow{OM_x}$ представим в виде: $\overrightarrow{OM_x} = l \cdot i$.

Из определения произведения вектора на число, $|\overrightarrow{OM_x}| = |l| \cdot |i|$. Следовательно, $|l| = \dfrac{|\overrightarrow{OM_x}}{|i|} = [|i| = 1] = |\overrightarrow{OM_x}| = |\overline{OM_x}|$, а по знаку $l = \begin{cases} |OM_x|, \overline{OM_x} \uparrow\uparrow Ox, \\ -|OM_x|, \overline{OM_x} \uparrow\downarrow Ox. \end{cases}$ $\Rightarrow$ по определению проекции, $l = OM_x = \text{ПР}_{Ox} \cdot OM = \text{ПР}_{Ox} \cdot a$.

Аналогично $OM_y = \text{ПР}_{Oy} \cdot a, OM_z = \text{ПР}_{Oz} \cdot a \Rightarrow a = \text{ПР}_{Ox} \cdot i + \text{ПР}_{Oy} \cdot j + \text{ПР}_{Oz} \cdot k$. А так как разложение по базису единственно, то $\text{ПР}_{Ox} \cdot a = x, \text{ПР}_{Oy} \cdot a = y, \text{ПР}_{Oz} \cdot a = z$. ⊠

Следствие. *Длина вектора a с координатами (x, y, z) равна* $\sqrt{x^2 + y^2 + z^2}$.

♦ Из параллелепипеда, построенного на отрезках OM_x, OM_y, OM_z, получаем

$$|a| = |\overline{OM}| = \sqrt{|\overline{OM_x}|^2 + |\overline{OM_y}|^2 + |\overline{OM_z}|^2} = \sqrt{x^2 + y^2 + z^2}.$$

⊠ Обозначим через α, β и γ углы между вектор a и векторами i, j и k соответственно.

• *Числа* $\cos(\alpha), \cos(\beta)$ *и* $\cos(\gamma)$ *называются* ***направляющими косинусами вектора*** a.

Следствие. $\cos^2(\alpha) + \cos^2(\beta) + \cos^2(\gamma) = 1$.

♦ По теореме, $x = \text{ПР}_{Ox} \cdot a = |a| \cdot \cos(a, Ox) = |a| \cdot \cos(a, i) = |a| \cdot \cos(\alpha) \Rightarrow \cos(\alpha) = \dfrac{x}{|a|}$.

Аналогично найдем $\cos(\beta) = \dfrac{y}{|a|}$, $\cos(\gamma) = \dfrac{z}{|a|} \Rightarrow$

$$\cos^2(\alpha) + \cos^2(\beta) + cos^2(\gamma) = \frac{x^2 + y^2 + z^2}{|a|^2} = 1.$$

⊠ • *Единичный вектор, сонаправленный с a, называется* ***ортом*** *вектора a, и имеет координаты* $e_a = (\cos(\alpha), \cos(\beta), \cos(\gamma))$.

• *Пусть M — некоторая точка пространства. Вектор OM называется* ***радиус-вектором*** *точки M. Обозначение: r_m.*

Теорема. *Радиус-вектор точки M с координатами (x_m, y_m, z_m) имеет такие же координаты (x_m, y_m, z_m).*

♦ Разложим r_m по базису $i, j, k : r_m = x \cdot i + y \cdot j + z \cdot k$. По первой теореме параграфа, $x_m = OM_x \Rightarrow x = x_m$. Аналогично $y = y_m$, $z = z_m$. ⊠

Следствие. *Пусть $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$ — две различные точки пространства. Тогда вектор M_1M_2 имеет координаты $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$.*

♦ $OM_1 + M_1M_2 = OM_2 \Rightarrow M_1M_2 = OM_2 - OM1 = r_{m_2} - r_{m_1} \Rightarrow$ координаты вектора M_1M_2 равны $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$. ⊠

2.4 Скалярное произведение векторов.

• ***Скалярным произведением*** *векторов a и b называется число, равное произведению длин этих векторов на косинус угла между ними. Обозначение: $(a, b) = |a| \cdot |b| \cdot \cos(a, b)$.*

Свойства скалярного произведения:

1. $a \cdot b = b \cdot a$.

   ♦ Доказательство следует из определения скалярного произведения. ⊠

2. $\forall a, b \neq 0 \quad a \cdot b = |a| \cdot \text{ПР}_a \cdot b = |b| \cdot \text{ПР}_b \cdot a$.

   ♦ $a \cdot b = |a| \cdot |b| \cdot \cos(a, b) =$ [по свойствам величины проекции $= \text{ПР}_a \cdot b = |b| \cdot \cos(a, b)$] $= |a| \cdot \text{ПР}_a \cdot b$ ⊠

3. $(\lambda \cdot a) \cdot b = \lambda \cdot (a \cdot b) = a \cdot (\lambda \cdot b)$.

   ♦ $(\lambda \cdot a) \cdot b = |b| \cdot \text{ПР}_b \cdot (\lambda \cdot a) = |b| \cdot \lambda \cdot \text{ПР}_b \cdot a = \lambda \cdot (a \cdot b)$. ⊠

4. $(a + b) \cdot c = a \cdot c + b \cdot c, \quad a \cdot (b + c) = a \cdot b + a \cdot c$.

   ♦ $(a+b) \cdot c = |c| \cdot \text{ПР}_c \cdot (a+b) = |c| \cdot (\text{ПР}_c \cdot a + \text{ПР}_c \cdot b) = |c| \cdot \text{ПР}_c \cdot a + |c| \cdot \text{ПР}_c \cdot b = a \cdot c + b \cdot c$. ⊠

5. $a \cdot a \geqslant 0$, *причем* $a \cdot a = 0 \iff a$ *— нулевой вектор.*

   ♦ $a \cdot a = |a| \cdot |a| \cdot \underbrace{\cos(a, a)}_{=1} = |a|^2 \geqslant 0$, причем $|a|^2 = 0 \iff |a| = 0$, то есть вектор a — нулевой. ⊠

6. **Критерий перпендикулярности векторов**.
   Два ненулевых вектора a и b перпендикулярны $\iff$ их скалярное произведение равно 0.

   ♦ $\Rightarrow$) Пусть $a \perp b \Rightarrow \angle\,(a, b) = \dfrac{\pi}{2} \Rightarrow a \cdot b = |a| \cdot |b| \cdot \cos\dfrac{\pi}{2} = 0$.

   $\Leftarrow$) Пусть $a \cdot b = 0 \Rightarrow |a| \cdot |b| \cdot \cos(a, b) = 0 \Rightarrow \cos(a, b) = 0 \Rightarrow \angle\,(a, b) = \dfrac{\pi}{2}$. ⊠

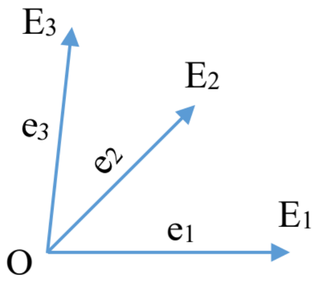
Теорема. *Скалярное произведение векторов a и b, имеющих декартовы координаты $a(x_1, y_1, z_1)$, $b(x_2, y_2, z_2)$, равно $a \cdot b = x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2$.*

♦ По определению декартовых координат, $a = x_1 \cdot i + y_1 \cdot j + z_1 \cdot k$, $b = x_2 \cdot i + y_2 \cdot j + z_2 \cdot k$, где i, j, k — ортонормированный базис.
$a \cdot b = (x_1 \cdot i + y_1 \cdot j + z_1 \cdot k) \cdot (x_2 \cdot i + y_2 \cdot j + z_2 \cdot k) = x_1 \cdot x_2 \cdot i \cdot i + x_1 \cdot y_2 \cdot i \cdot j + x_1 \cdot z_2 \cdot i \cdot k + \ldots =$ [$i \cdot j = j \cdot k = i \cdot k = 0$, так как они друг с другом перпендикулярны, а $i \cdot i = j \cdot j = k \cdot k = 1$, так как орты, а $\cos = 1$] $= x_1 \cdot x_2 + y_1 \cdot y_2 + z_1 \cdot z_2$. ⊠

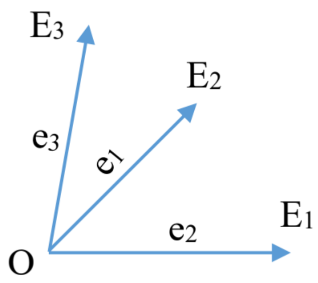
2.5 Векторное произведение векторов.

Рассмотрим упорядоченную тройку некомпланарных векторов e_1, e_2, e_3, отложим их от некоторой точки O и пусть $e_1 = \overrightarrow{OE_1}$, $e_2 = \overrightarrow{OE_2}$, $e_3 = \overrightarrow{OE_3}$.

• *Если при наблюдении с конца направленного отрезка OE_3 кратчайший поворот от $\overrightarrow{OE_1}$ к $\overrightarrow{OE_2}$ осуществляется против часовой стрелки, то тройка векторов e_1, e_2, e_3 называется **правой**, в противном случае **левой**. Данное свойство также называется **ориентацией** векторов.*



правая тройка



левая тройка

Из векторов a, b, c можно составить 6 упорядоченных троек:

$$a, b, c \qquad b, c, a \qquad c, a, b \text{ — правая тройка;}$$
$$a, c, b \qquad b, a, c \qquad c, b, a \text{ — левая тройка.}$$

Из определения ориентации тройки векторов следует, что первые три тройки имеют одну ориентацию, а вторые три противоположную им ориентацию.

• ***Векторным произведением*** *вектора a на вектор b называется вектор, обозначаемый $a \times b$ или $[a, b]$, который удовлетворяет следующим условиям:*

1. $|[a, b]| = |a| \cdot |b| \cdot \sin(a, b)$;
2. $[a, b] \perp a$, $[a, b] \perp b$;
3. *упорядоченная тройка $a, b, [a, b]$ является правой.*

Свойства векторного произведения векторов:

1. *Длина векторного произведения векторов a и b равна площади параллелограмма, построенного на векторах a и b, отложенных от одной точки.*

   ⧫ Следует из первого пункта определения. ⊠

2. *Два ненулевых вектора a и b коллинеарны $\iff$ их векторное произведение равно нулевому вектору.*

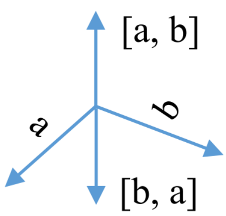
   ⧫ $\Rightarrow$) Пусть векторы a и b коллинеарны, тогда угол между ними в радианах либо 0, либо $\pi \Rightarrow sin(a, b) = 0 \Rightarrow |[a, b]| = |a| \cdot |b| \cdot 0 = 0$, получили нулевой вектор.

   $\Leftarrow$) Пусть векторное произведение $[a, b] = \overrightarrow{0} \Rightarrow |[a, b]| = 0 \Rightarrow |a| \cdot |b| \cdot \sin(a, b) = 0$ $(|a|, |b| \neq 0) \Rightarrow \sin(a, b) = 0 \Rightarrow \angle(a, b) = 0$ или $\angle(a, b) = \pi \Rightarrow a, b$ коллениарны. ⊠

3. $[a, b] = -[b, a]$.

   ⧫ Если векторы a и b коллениарны, то утверждение очевидно. Так как $[a, b] = \overrightarrow{0}$ по второму свойству и, следовательно, $[b, a] = \overrightarrow{0}$.

   Пусть векторы a и b неколлениарны, тогда векторы $[a, b]$ и $[b, a]$ коллениарны, так как каждый из них одновременно перпендикулярен векторам a и b, и, при этом, длины их равны.



Следовательно, возможны два варианта:

(a) $[a, b] = [b, a]$;

(b) $[a, b] = -[b, a]$.

Пусть $[a, b] = [b, a]$. По определению векторного произведения, тройка векторов $(a, b, [a, b])$ правая $\Rightarrow$ $(a, b, [b, a])$ тоже правая, но, по определению векторного произведения, тройка векторов $(b, a, [b, a])$ также правая. Тогда, глядя со стороны направления вектора b на a, кратчайший поворот против часовой стрелки осуществляется, с одной стороны, от a к b, а с другой — от b к a, что является противоречием $\Rightarrow$ $[a, b] = -[b, a]$. ⊠

4. $[\lambda \cdot a, b] = \lambda \cdot [a, b]$, $[a, \lambda \cdot b] = \lambda \cdot [a, b] \quad \lambda \in \mathbb{R}$.

♦ Если векторы a и b коллениарны или $\lambda = 0$, то утверждение очевидно.

Пусть $\lambda \neq 0$ и векторы a и b неколлениарны. Тогда покажем равенство длин для обоих случаев:
$|\lambda[a, b]| = |\lambda| \cdot |[a, b]| = |\lambda| \cdot |a| \cdot |b| \cdot \sin(a, b)$.
$|[\lambda a, b]| = |\lambda a| \cdot |b| \cdot \sin(\lambda a, b) = \begin{cases} |\lambda| \cdot |a| \cdot |b| \cdot \sin(a, b), \ \lambda > 0; \\ |\lambda| \cdot |a| \cdot |b| \cdot \sin(\pi - \angle(a, b)), \ \lambda < 0. \end{cases} = |\lambda| \cdot |a| \cdot |b| \cdot \sin(a, b)$.

Покажем направления. $[\lambda a, b]$ и $\lambda[a, b]$ перпендикулярны и a, и b $\Rightarrow$ $[\lambda a, b] \parallel \lambda[a, b]$.
При $\lambda > 0$ векторное произведение $[a, b]$ сонаправлено с обоими векторами a, b.
При $\lambda < 0$ векторное произведение $[a, b]$ имеет противоположное направление с векторами a, b $\Rightarrow$

$$[\lambda a, b] \uparrow\downarrow \lambda[a, b] \Rightarrow [a, \lambda b] = -[\lambda b, a] = -\lambda[b, a] = \lambda[a, b].$$

⊠

5. *a)* $[a + b, c] = [a, c] + [b, c]$;

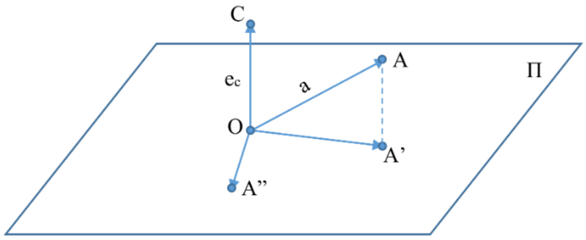
   b) $[a, b + c] = [a, b] + [a, c]$.

♦

(a) Если один из векторов a, b, c нулевой, то утверждение очевидно.

Пусть векторы a, b, c ненулевые. Обозначим через e_c орт вектора c, тогда $e_c = \dfrac{1}{|c|} \cdot c \Rightarrow |e_c| = 1$.
Построим векторное произведение вектора a на e_c следующим образом:

1) отложим векторы a и e_c от некоторой точки O. Тогда $a = \overrightarrow{OA}$, $e_c = \overrightarrow{OC}$;
2) проведем через точку O плоскость перпендикулярно отрезку OC;
3) построим отрезок OA_1 — проекцию отрезка OA на плоскость Π, а затем повернем этот отрезок в плоскости Π вокруг точки O по часовой стрелке на угол $\dfrac{\pi}{2}$ и получим отрезок OA_2.



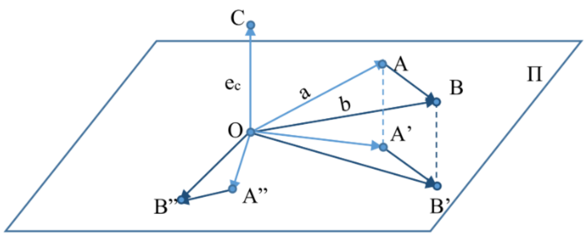
Докажем, что $\overrightarrow{OA_2} = [\overrightarrow{OA}, \overrightarrow{OC}] = [a, e_c]$.

1) Длина: $|\overrightarrow{OA_2}| = |\overline{OA_2}| = |\overline{OA_1}| = |\overline{OA}| \cdot \cos\angle AOA_1 = |\overline{OA}| \cdot \cos(\pm(\dfrac{\pi}{2} - \angle(a, e_c)) = |a| \cdot |e_c| \cdot \sin(a, e_c) = |[a, e_c]|$.

2) Направление: $\overline{OA_2} \perp \overline{OC}$, так как $\overline{OA_2}$ лежит в плоскости, перпендикулярной отрезку OC, $\overline{OA_2} \perp \overline{OA} \Rightarrow \overline{OA_2} \perp e_c$, $\overline{OA_2} \perp a$ по теореме о трех перпендикулярах.

3) Тройка векторов $\overrightarrow{OA}$, $\overrightarrow{OC}$, $\overrightarrow{OA_2}$ правая по построению.

Таким образом мы показали, что $\overrightarrow{OA_2} = [a, e_c]$.
Отложим вектор b от точки A. Пусть $b = \overrightarrow{AB}$. Построим треугольник OA_1B_1 — проекцию треугольника OAB на плоскость П, а затем повернем этот треугольник по часовой стрелке вокруг точки O на угол $\dfrac{\pi}{2}$ и получим треугольник OA_2B_2.



По определению суммы векторов, $\overrightarrow{OB_2} = \overrightarrow{OA_2} + \overrightarrow{A_2B_2}$. По доказанному выше, $\overrightarrow{OB_2} = [\overrightarrow{OB}, \overrightarrow{OC}] = [a + b, e_c]$, $\overrightarrow{A_2B_2} = [\overrightarrow{AB}, \overrightarrow{OC}] = [b, e_c]$
Следовательно, из полученного выше имеем $[a + b, e_c] = [a, e_c] + [b, e_c]$.
Домножим полученное равенство на длину вектора c и используем четвертое свойство. Получим $[a + b, c] = [a, c] + [b, c]$.

(b) $[a, b + c] =$ -$[b + c, a] = -([b, a] + [c, a]) = [a, b] + [a, c]$.

⊠

- *Квадратная таблица чисел вида* $\begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix}$ *называется* ***матрицей второго порядка***.
- ***Определителем*** *матрицы второго порядка называется число вида*

$$\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix} = \alpha_1\beta_2 - \alpha_2\beta_1.$$

Теорема. *Пусть векторы a и b в правой ДПСК имеют координаты $a(X_1, Y_1, Z_1)$ и $b(X_2, Y_2, Z_2)$, тогда вектор $[a, b]$ имеет координаты*

$$[a, b] = \left(\begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}, - \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}, \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \right).$$

♦ Из определения декартовых координат вектора следует, что $a = X_1 \cdot i + Y_1 \cdot j + Z_1 \cdot k$, $b = X_2 \cdot i + Y_2 \cdot j + Z_2 \cdot k$. Тогда
$[a, b] = [X_1 \cdot i + Y_1 \cdot j + Z_1 \cdot k,\ X_2 \cdot i + Y_2 \cdot j + Z_2 \cdot k] = X_1 \cdot X_2 \cdot [i, i] + X_1 \cdot Y_2 \cdot [i, j] +$
$+ X_1 \cdot Z_2 \cdot [i, k] + Y_1 \cdot X_2 \cdot [i, j] + Y_1 \cdot Z_2 \cdot [j, k] + Z_1 \cdot X_2 \cdot [k, i] + Z_1 \cdot Y_2 \cdot [k, j] + Z_1 \cdot Z_2 \cdot [k, k] + Y_1 \cdot Y_2 \cdot [j, j]$.

$[i, i] = \overrightarrow{0}$, $[j, j] = \overrightarrow{0}$, $[k, k] = \overrightarrow{0}$, так как векторы параллельны сами себе, следовательно, применимо второе свойство. $[i, j] = k$, $[i, k] = -j$, $[j, i] = -k$, $[j, k] = i$, $[k, i] = j$, $[k, j] = -i$.

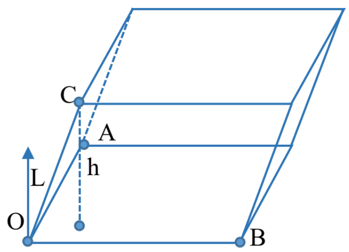
Подставим и получим $X_1 \cdot Y_2 \cdot k + X_1 \cdot Z_2 \cdot (-j) + Y_1 \cdot X_2 \cdot (-k) + Y_1 \cdot Z_2 \cdot i + Z_1 \cdot X_2 \cdot j + Z_1 \cdot Y_2 \cdot (-i)$ $= (Y_1 \cdot Z_2 - Z_1 \cdot Y_2) \cdot i$ - $(X_1 \cdot Z_2 - Z_1 \cdot X_2) \cdot j + (X_1 \cdot Y_2 - Y_1 \cdot X_2) \cdot k =$
$= \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix} \cdot i - \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix} \cdot j + \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \cdot k$. ⊠

2.6 Смешанное произведение векторов.

• ***Смешанным произведением трех векторов a, b, c*** *называется число* $abc = [a, b] \cdot c$.

Теорема. *Пусть V — объём параллелепипеда, построенного на отложенных от одной точки трех некомпланарных векторах a, b, c. Тогда* $abc = \begin{cases} V, \text{ где } a,b,c \text{ — правая тройка;} \\ -V, \text{ где } a,b,c \text{ — левая тройка.} \end{cases}$

♦ Пусть векторы a, b и c отложены от точки O. Тогда $a = \overrightarrow{OA}$, $b = \overrightarrow{OB}$ и $c = \overrightarrow{OC}$. Так как векторы a, b, c некомпланарны, то векторы a и b неколлениарны.



Обозначим через S площадь параллелограмма, построенного на отрезках OA и OB. Тогда $S = |\overrightarrow{OA}| \cdot |\overrightarrow{OB}| \cdot sin\angle(BOA) = |a| \cdot |b| \cdot \sin(a, b) = |[a, b]|$.

Обозначим через e орт $[a, b]$. Тогда тройка векторов a, b, c — правая и при этом
$[a, b] = |[a, b]| \cdot e = S \cdot e \Rightarrow abc = [a, b] \cdot c = (S \cdot e) \cdot c = S \cdot (e \cdot c) = S \cdot |e| \cdot \text{ПР}_e c = S \cdot \text{ПР}_e c$.

Пусть h — высота рассматриваемого параллелепипеда, опущенная из вершины C на плоскость AOB, тогда $|\text{ПР}_e c| = h$,
но при этом $\begin{cases} \text{ПР}_e c > 0, \text{ если } a, b, c \text{ — правая тройка векторов;} \\ \text{ПР}_e c < 0, \text{ если } a, b, c \text{ — левая тройка векторов.} \end{cases} \Rightarrow$
$abc = \begin{cases} abc = S \cdot h, \text{ если } a, b, c \text{ — правая тройка;} \\ abc = -S \cdot h, \text{ если } a, b, c \text{ — левая тройка.} \end{cases} \Rightarrow V = |abc|$ и условие теоремы выполняется. ⊠

Свойства смешанного произведения:

1. *Векторы a, b, c компланарны $\Longleftrightarrow$ их смешанное произведение $abc = 0$.*

   ♦ $\Rightarrow$) Пусть векторы a, b, c компланарны.

1) $a \parallel b \Rightarrow [a,b] = 0 \Rightarrow abc = [a,b] \cdot c = 0$;

2) $a \nparallel b \Rightarrow [a,b] \perp a, [a,b] \perp b \Rightarrow [a,b] \perp c \Rightarrow abc = [a,b] \cdot c = 0$.

$\Leftarrow$) Пусть $abc = 0$, и предположим, что векторы a, b, c некомпланарны. Тогда, отложив их от одной точки, можно построить параллелепипед с $V = |abc| \neq 0$, что противоречит тому, что $abc = 0$. Значит a, b, c компланарны. ⊠

2. $abc = bca = cab = -acb = -bac = -cba$.

   ♦ Если векторы a, b, c компланарны, то все произведения равны 0.

   Если векторы a, b, c некомпланарны, то модули рассматриваемых смешанных произведений будут равны, а знак зависит от ориентации соответствующих троек.

   Первые три смешанных произведений имеют одну ориентацию, а последние три противоположную ориентацию. ⊠

3. $(\alpha a)bc = a(\alpha b)c = ab(\alpha c) = \alpha(abc)$.

   ♦ $ab(\alpha c) = [a,b](\alpha c) = \alpha([a,b]c) = \alpha(abc)$. ⊠

4. $(a+b) \cdot c \cdot d = acd + bcd$;
   $a \cdot (b+c) \cdot d = abd + acd$;
   $a \cdot b \cdot (c+d) = abc + abd$.

   ♦ $(a+b) \cdot c \cdot d = [a+b, c] \cdot d = ([a,c] + [b,c]) \cdot d = [a,c] \cdot d + [b,c] \cdot d = acd + bcd$. ⊠

• *Квадратная таблица чисел вида* $\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix}$ *называется* ***матрицей третьего порядка***.

• ***Определителем*** *матрицы третьего порядка является число вида*

$$\begin{vmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix} = = \alpha_1 \begin{vmatrix} \beta_1 & \beta_3 \\ \gamma_2 & \gamma_3 \end{vmatrix} - \alpha_2 \begin{vmatrix} \beta_1 & \beta_3 \\ \gamma_1 & \gamma_3 \end{vmatrix} + \alpha_3 \begin{vmatrix} \beta_1 & \beta_2 \\ \gamma_1 & \gamma_2 \end{vmatrix}.$$

Теорема. *Пусть векторы a, b, c в правой ДПСК имеют координаты $a(x_1, y_1, z_1)$, $b(x_2, y_2, z_2)$, $c(x_3, y_3, z_3)$, тогда* $abc = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$.

♦ $abc = bca = [b,c] \cdot a = \left(\begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix}, -\begin{vmatrix} X_2 & Z_2 \\ X_3 & Z_3 \end{vmatrix}, \begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{vmatrix} \right) \cdot (x_1, y_1, z_1) =$

$= x_1 \begin{vmatrix} Y_2 & Z_2 \\ Y_3 & Z_3 \end{vmatrix} - y_1 \begin{vmatrix} X_2 & Z_2 \\ X_3 & Z_3 \end{vmatrix} + z_1 \begin{vmatrix} X_2 & Y_2 \\ X_3 & Y_3 \end{vmatrix} = \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}$. ⊠

2.7 Преобразование декартовых прямоугольных координат на плоскости и в пространстве.

Пусть на плоскости заданы две ДПСК: Oxy, $O'x'y'$ и пусть некоторая точка M плоскости в первой ДПСК имеет координаты (x, y), а во второй (x', y'). Выразим координаты (x, y)

через (x', y').

Пусть i, j, i', j' — направляющие орты осей координат $Ox, Oy, O'x', O'y'$ соответственно. Так как точка M в ДПСК Oxy имеет координаты (x, y), то вектор $\overrightarrow{OM}$ представим в виде $\overrightarrow{OM} = x \cdot i + y \cdot j$. Аналогично $\overrightarrow{O'M} = x' \cdot i' + y' \cdot j'$.

Пусть точка O' в исходной ДПСК имеет координаты (a, b), тогда $\overrightarrow{OO'} = a \cdot i + b \cdot j$.

Так как любой вектор можно разложить по базису на плоскости, то векторы i', j' можно разложить по базисам i, j на плоскости, то есть представить в виде $i' = \alpha_1 \cdot i + \alpha_2 \cdot j$, $j' = \beta_1 \cdot i + \beta_2 \cdot j$.

Определим коэффициенты этих разложений. Для этого умножим каждое равенство на векторы i, j:
$i' \cdot i = \alpha_1 \cdot i \cdot i + \alpha_2 \cdot j \cdot i = \alpha_1 \Rightarrow \alpha_1 = i' \cdot i = |i'| \cdot |i| \cdot \cos(i, i') = \cos(i, i')$
$i' \cdot j = \alpha_1 \cdot i \cdot j + \alpha_2 \cdot j \cdot j \Rightarrow \alpha_2 = i' \cdot j = |i'| \cdot |j| \cdot \cos(j, i')$
$j' \cdot i = \beta_1 \cdot i \cdot i + \beta_2 \cdot j \cdot i = \beta_1 \Rightarrow \beta_1 = j' \cdot i = \cos(i, j')$
$j' \cdot j = \beta_1 \cdot i \cdot j + \beta_2 \cdot j \cdot j = \beta_2 \Rightarrow \beta_2 = j' \cdot j = \cos(j, j')$

Из определения суммы векторов следует, что $\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M}$. Тогда из разложения по базису этих векторов получим

$$x \cdot i + y \cdot j = a \cdot i + b \cdot j + x' \cdot (\alpha_1 \cdot i + \alpha_2 \cdot j) + y' \cdot (\beta_1 \cdot i + \beta_2 \cdot j);$$
$$x \cdot i + y \cdot j = (a + x' \cdot \alpha_1 + y' \cdot \beta_1) \cdot i + (b + x' \cdot \alpha_2 + y' \cdot \beta_2) \cdot j.$$

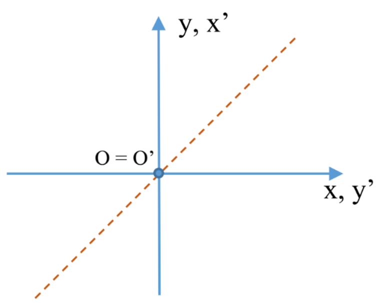
Так как разложение любого вектора по базису единственно, то $\begin{cases} x = a + x' \cdot \alpha_1 + y' \cdot \beta_1, \\ y = b + x' \cdot \alpha_2 + y' \cdot \beta_2. \end{cases} \Rightarrow$

$$\begin{cases} x = a + x' \cdot \cos(i, i') + y' \cdot \cos(i, j'), \\ y = b + x' \cdot \cos(j, i') + y' \cdot \cos(j, j'). \end{cases} \tag{2.7.1}$$

• *Полученная формула* (2.7.1) *называется* ***формулой преобразования декартовых координат точки***.

Рассмотрим некоторые частные случаи.

1. **Преобразование симметрии относительно прямой** $(y = x)$.
   Пусть точки $O(x, y)$ и $O'(x', y')$ — начала двух ДПСК разной ориентации. И пусть они совпадают $(O = O')$.

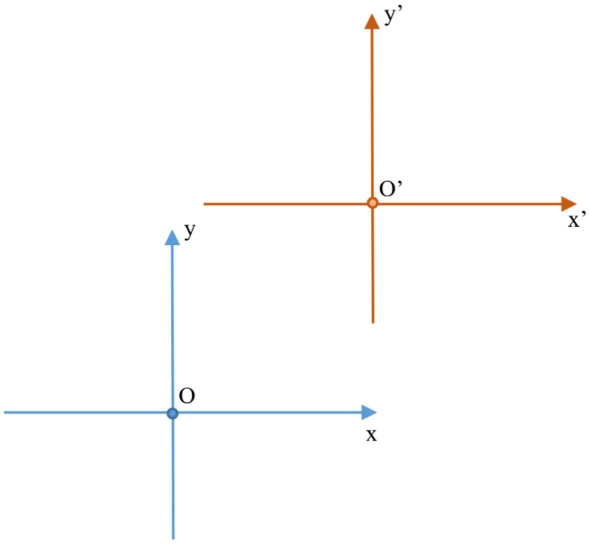


Тогда получаем $\angle(i, i') = \angle(i, j) = \dfrac{\pi}{2}$, $\angle(j, j') = \angle(j, i) = \dfrac{\pi}{2}$, $\angle(i, j') = \angle(j, i') = 0$. Отсюда имеем

$$\begin{cases} x = y', \\ y = x'. \end{cases} \tag{2.7.2}$$

• *Полученная формула* (2.7.2) *называется* ***формулой преобразования симметрии относительно прямой*** $y = x$.

2. **Параллельный перенос координатных осей.**
Пусть начало новой ДПСК располагается в точке $O'(a, b)$ исходной ДПСК.

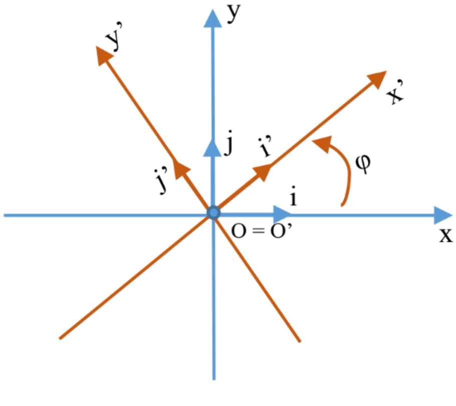


При параллельном переносе координатных осей их орты совпадают, то есть $i = i'$, $j = j'$. Тогда

$$\begin{cases} x = a + x' \cdot \cos 0 + y' \cdot \cos \dfrac{\pi}{2}, \\ y = b + x' \cdot \cos \dfrac{\pi}{2} + y' \cdot \sin \dfrac{\pi}{2}. \end{cases} \Rightarrow \begin{cases} x = a + x', \\ y = b + y'. \end{cases} \tag{2.7.3}$$

• *Полученная формула* (2.7.3) *называется* ***формулой параллельного переноса***.

3. **Преобразование поворота.**
Пусть ДПСК $O'x'y'$ получена из правой ДПСК Oxy путем поворота вокрут точки O против часовой стрелки на угол φ.



И пусть начала координат равны ($O' = O$). Тогда
$\cos(i, i') = \cos\varphi$,
$\cos(i, j') = \cos(\varphi + \dfrac{\pi}{2})$,
$\cos(j, i') = \cos(\dfrac{\pi}{2} - \varphi)$,
$\cos(j, j') = \cos\varphi$.
Из уравнений выше следует система

$$\begin{cases} x = x'\cos\varphi - y'\sin\varphi, \\ y = x'\sin\varphi + y'\cos\varphi. \end{cases} \tag{2.7.4}$$

• *Полученная формула* (2.7.4) *называется* ***формулой преобразования поворота.***

Пусть точка M в ДПСК $Oxyz$ имеет координаты (x, y, z), а в ДПСК $O'x'y'z'$ — (x', y', z'), и пусть точка O имеет координаты $O(a, b, c)$ в ДПСК $Oxyz$. Тогда справедлива система

$$\begin{cases} x = a + x'\cos(i, i') + y'\cos(i, j') + z'\cos(i, k'), \\ y = b + x'\cos(j, i') + y'\cos(j, j') + z'\cos(j, k'), \\ z = c + x'\cos(k, i') + y'\cos(k, j') + z'\cos(k, k'). \end{cases} \tag{2.7.5}$$

• *Полученная формула* (2.7.5) *называется* ***формулой преобразования декартовых координат в пространстве.***

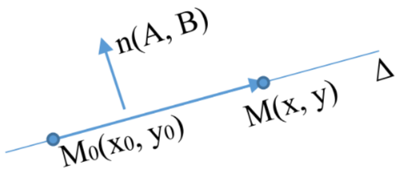
Глава 3

Уравнения прямой на плоскости

3.1 Общее уравнение прямой. Уравнение прямой в отрезках.

• ***Нормальным вектором*** *прямой на плоскости называется ненулевой вектор, перпендикулярный данной прямой.*

Рассмотрим некоторую прямую Δ. Пусть она проходит через некоторую точку $M_0(x_0, y_0)$ и она перпендикулярна ненулевому вектору $n(A, B)$.



Так как вектор n ненулевой, то $A^2 + B^2 \neq 0$. Если некоторая точка $M(x, y)$ лежит на прямой Δ, то $\overrightarrow{M_0M} \perp n$ и наоборот, если точка M не лежит на прямой Δ, то $\overrightarrow{M_0M} \not\perp n$.

Таким образом, точка $M \in \Delta \iff \overrightarrow{M_0M} \perp n$.

• *Уравнение* $\overrightarrow{M_0M} \cdot n = 0$ *называется* ***уравнением в векторной форме прямой на плоскости****, проходящей через точку* M_0 *перпендикулярно вектору* n.

Вектор $\overrightarrow{M_0M}$ в ДПСК имеет координаты $(x - x_0, y - y_0)$. Тогда

$$A(x - x_0) + B(y - y_0) = 0. \tag{3.1.1}$$

• *Уравнение* (3.1.1) *называется* ***уравнением в координатной форме прямой на плоскости****, проходящей через точку с координатами* (x_0, y_0) *перпендикулярно вектору* n *с координатами* (A, B).

Теорема. *Любая прямая на плоскости в ДПСК может быть задана уравнением вида*

$$Ax + By + C = 0, \quad A^2 + B^2 \neq 0. \tag{3.1.2}$$

В ДПСК любое уравнение вида (3.1.2) *определяет прямую и только.*

♦

1. Пусть произвольная прямая Δ проходит через точку $M_0(x_0, y_0)$ перпендикулярно вектору $n(A, B)$. Преобразуем уравнение (3.1.1) следующим образом: $Ax + By + \underbrace{(-Ax_0 - By_0)}_{C} = 0$. Тогда уравнение прямой примет вид (3.1.2), причем A и B — координаты нормального вектора. Следовательно, $A^2 + B^2 \neq 0$.

2. Так как коэффициенты A и B не обращаются в 0 одновременно, то уравнение (3.1.2) всегда имеет решение. Пусть (x_0, y_0) — решение уравнения (3.1.2). Тогда $Ax_0 + By_0 + C = 0$.

   Вычтем полученное равенство из уравнения (3.1.2) и получим $A(x - x_0) + B(y - y_0) = 0$ — уравнение прямой, проходящей через точку M_0. Значит уравнение (3.1.2) является уравнением прямой.

⊠

• *Уравнение* (3.1.2) *называется* ***общим уравнением прямой на плоскости***. *Причем коэффициенты A и B являются координатами нормального вектора n.*

Теорема. *Если два общих уравнения* $\begin{matrix} A_1x + B_1y + C_1 = 0, \\ A_2x + B_2y + C_2 = 0 \end{matrix}$ *определяют одну прямую, то существует действительное число λ такое, что $A_1 = \lambda \cdot A_2$, $B_1 = \lambda \cdot B_2$, $C_1 = \lambda \cdot C_2$.*

⧫ Так как оба уравнения описывают одну прямую, то векторы $n_1(A_1, B_1)$ и $n_2(A_2, B_2)$ являются нормальными векторами одной прямой $\Rightarrow$ они коллениарны и существует λ такое, что $A_1 = \lambda \cdot A_2$, а $B_1 = \lambda \cdot B_2$. Пусть прямая проходит через точку $M_0(x_0, y_0)$, тогда подставим и получим следующие уравнения:

$$\begin{matrix} A_1x_0 + B_2y_0 + C_1 = 0, \\ A_2x_0 + B_2y_0 + C_2 = 0. \end{matrix}$$

Отнимем от первого уравнения второе домноженное на λ и получим
$\underbrace{(A_1 - \lambda A_2)}_{=0} \cdot x_0 + \underbrace{(B_1 - \lambda B_2)}_{=0} \cdot y_0 + (C_1 - \lambda \cdot C_2) = 0 \Rightarrow C_1 - \lambda C_2 = 0 \Rightarrow C_1 = \lambda C_2$. ⊠

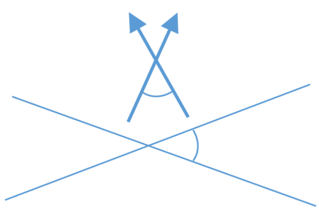
Рассмотрим взаимное расположение двух прямых. Пусть $A_1x + B_1y + C_1 = 0$ и $A_2x + B_2y + C_2 = 0$ — уравнения этих двух прямых.

1. Прямые совпадают, если $\exists \lambda \in \mathbb{R} : A_1 = \lambda \cdot A_2,\ B_1 = \lambda \cdot B_2,\ C_1 = \lambda \cdot C_2$;

2. Прямые параллельны, если $\exists \lambda \in \mathbb{R} : A_1 = \lambda \cdot A_2,\ B_1 = \lambda \cdot B_2,\ C_1 \neq \lambda \cdot C_2$;

3. Прямые пересекаются, если $\nexists \lambda \in \mathbb{R} : A_1 = \lambda \cdot A_2,\ B_1 = \lambda \cdot B_2$. Угол пересечения равен углу между нормальными векторами этих прямых, то есть

$$\cos(\Delta_1, \Delta_2) = \frac{|n_1 \cdot n_2|}{|n_1| \cdot |n_2|} = \frac{|A_1A_2 + B_1B_2|}{\sqrt{A_1^2 + B_1^2} \cdot \sqrt{A_2^2 + B_2^2}}.$$



• *Общее уравнение прямой называется* ***полным****, если все его коэффициенты* A, B *и* $C \neq 0$. *В противном случае уравнение называется* ***неполным****.*

Рассмотрим виды неполных уравнений. Пусть Δ — прямая.

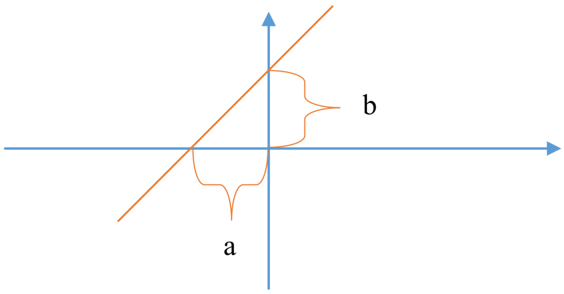
1. $C = 0 \Rightarrow O(0,0) \in \Delta$, прямая проходит через начало координат;
2. $B = 0 \Rightarrow \Delta \parallel Oy$;
3. $A = 0 \Rightarrow \Delta \parallel Ox$;
4. $B = 0, C = 0 \Rightarrow \Delta = Oy$;
5. $A = 0, C = 0 \Rightarrow \Delta = Ox$.

Рассмотрим полное уравнение $Ax + By + C = 0$. Прямая, описанная этим уравнением, не проходит через начало координат и не параллельна ни одной из координатных осей. Преобразуем уравнение следующим образом:
$$Ax + By = -C \Rightarrow \frac{Ax}{-C} + \frac{By}{-C} = 1 \Rightarrow \frac{x}{\frac{-C}{A}} + \frac{y}{\frac{-C}{B}} = 1 \Rightarrow \left[\frac{-C}{A} = a, \quad \frac{-C}{B} = b\right] \Rightarrow$$

$$\frac{x}{a} + \frac{y}{b} = 1. \tag{3.1.3}$$

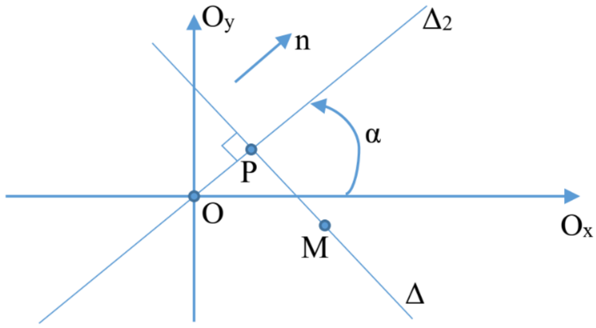
• *Уравнение* (3.1.3) *называется* ***уравнением прямой в отрезках****.*

Так как точки пересечения этой прямой с координатными осями имеют координаты $(a, 0)$ и $(0, b)$, то a и b — величины направленных отрезков, отсекаемых прямой по координатным осям от начала координат.



3.2 Нормальное уравнение прямой.

Рассмотрим некоторую прямую Δ. Построим прямую Δ_2, проходящую через начало координат и перпендикулярную прямой Δ. Пусть P — точка пересечения прямых Δ и Δ_2. Обозначим через n единичный вектор, перпендикулярный прямой Δ и сонаправленный с вектором $\overrightarrow{OP}$.



Если точки O и P совпадают, то n — произвольный единичный вектор, перпендикулярный прямой Δ. Обозначим через p расстояние между точками O и P, то есть $p = |\overline{OP}|$. Через α обозначим угол, на который надо повернуть ось Ox против часовой стрелки, чтобы ее направление совпало с направлением нормального вектора n. Тогда вектор n имеет координаты $n(\cos\alpha, \cos\beta)$.

Пусть $M(x, y)$ — произвольная точка плоскости. $M \in \Delta \iff \overrightarrow{PM} \perp n \iff \overrightarrow{PM} \cdot n = 0$.

По определению суммы векторов $\overrightarrow{PM} = \overrightarrow{PO} + \overrightarrow{OM} = \overrightarrow{OM} - \overrightarrow{OP}$, тогда подставим и получим $(\overrightarrow{OM} - \overrightarrow{OP}) \cdot n = 0 \Rightarrow \overrightarrow{OM} \cdot n - \overrightarrow{OP} \cdot n = 0$. Распишем уменьшаемое и вычитаемое:
$\overrightarrow{OM} \cdot n = (x, y) \cdot (\cos\alpha, \sin\alpha) = x\cos\alpha + y\sin\alpha$.
$\overrightarrow{OP} \cdot n = |\overline{OP}| \cdot |n| \cdot \cos\underbrace{(\overrightarrow{OP}, n)}_{=0} = p \cdot 1 \cdot 1 = p \Rightarrow$ получаем уравнение

$$x\cos\alpha + y\sin\alpha - p = 0. \tag{3.2.1}$$

• *Уравнение* (3.2.1) *называется нормальным* ***нормальным уравнением прямой****.*

Пусть $M_0(x_0, y_0)$ — произвольная точка плоскости, а точка $M_1(x_1, y_1)$ — ее проекция на прямую Δ. Обозначим через $\rho(M_0, \Delta)$ — расстояние от точки M_0 до прямой Δ. Тогда $\rho(M_0, \Delta) = |\overline{M_1 M_0}|$.

• ***Отклонением*** *точки M_0 от прямой Δ называется число*

$$\delta(M_0, \Delta) = \begin{cases} \rho(M_0, \Delta), & \overrightarrow{M_1 M_0} \uparrow\uparrow n, \\ -\rho(M_0, \Delta), & \overrightarrow{M_1 M_0} \uparrow\downarrow n. \end{cases}$$

Из определения отклонения следует, что $\rho(M_0, \Delta) = |\delta(M_0, \Delta)|$ и $\delta(M_0, \Delta) = \text{ПР}_n\, \overrightarrow{M_1 M_0}$.

Теорема. *Если $x\cos\alpha + y\sin\alpha - p = 0$ — нормальное уравнение прямой, то отклонение точки $M_0(x_0, y_0)$ от прямой Δ следующее: $\delta(M_0, \Delta) = x_0\cos\alpha + y_0\sin\alpha - p$.*

♦ $\delta(M_0, \Delta) = \text{ПР}_n \overrightarrow{M_1 M_0} = [\ \overrightarrow{M_1 M_0} = \overrightarrow{M_1 O} + \overrightarrow{OM_0} = \overrightarrow{OM_0} - \overrightarrow{OM_1}\ ] \Rightarrow \text{ПР}_n(\overrightarrow{OM_0} - \overrightarrow{OM_1}) =$
$= \text{ПР}_n \overrightarrow{OM_0} - \text{ПР}_n \overrightarrow{OM_1}$. Распишем обе проекции по отдельности:

$\text{ПР}_n \overrightarrow{OM_0} = |\overrightarrow{OM_0}| \cdot \cos(n, \overrightarrow{OM_0}) \cdot 1 = \overrightarrow{OM_0} \cdot n = (x_0, y_0)(\cos\alpha, \sin\alpha) = x_0\cos\alpha + y_0\sin\alpha$.

$\text{ПР}_n \overrightarrow{OM_1} = p$ (по определению величины проекции). Тогда, подставив вместо проекций полученные выражения, имеем $\delta(M_0, \Delta) = \text{ПР}_n \overrightarrow{OM_0} - \text{ПР}_n \overrightarrow{OM_1} = x_0 cos\alpha + y_0 sin\alpha - p$. ⊠

Построим нормальное уравнение прямой с помощью общего уравнения.
Пусть прямая Δ удовлетворяет общему уравнению

$$Ax + By + C = 0. \tag{3.2.2}$$

И пусть уравнение (3.2.1) — нормальное уравнение прямой Δ. Так как уравнения эти уравнения определяют одну прямую, то их коэффициенты пропорциональны, то есть $\exists \lambda \in \mathbb{R}$, $\cos\alpha = \lambda \cdot A$, $\sin\alpha = \lambda \cdot B$, $-p = \lambda \cdot C$. Из первых двух равенств имеем $1 = \cos^2\alpha + \sin^2\alpha = \lambda^2(A^2 + B^2) \Rightarrow \lambda = \pm \dfrac{1}{\sqrt{A^2 + B^2}}$, причем из полученного соотношения следует, что знак λ противоположен знаку C.

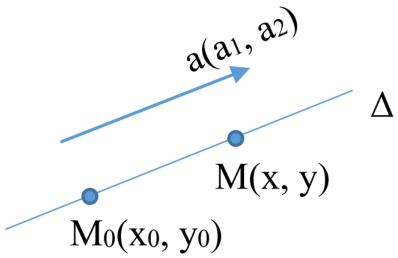
• *Полученное таким образом число* λ *называется* ***нормирующим множителем*** *для общего уравнения* (3.2.2)*, так как, домножив это уравнение на* λ*, получим нормальное уравнение.*

3.3 Параметрическое и каноническое уравнения прямой. Уравнение прямой с угловым коэффициентом.

До этого мы задавали прямую с помощью нормального вектора. Теперь зададим её другим способом.

• *Ненулевой вектор, параллельный прямой, называется* ***направляющим вектором прямой****.*

Рассмотрим некоторую прямую Δ и пусть $a(a_1, a_2)$ — ее направляющий вектор, $M_0(x_0, y_0)$ — точка на прямой. Тогда точка $M(x, y) \in \Delta \iff \overrightarrow{M_0M} \parallel a$.



Следовательно, $\exists t \in R$ такое, что $\overrightarrow{M_0M} = ta$. $\overrightarrow{M_0M} = \overrightarrow{M_0O} + \overrightarrow{OM} = \overrightarrow{OM} - \overrightarrow{OM_0} = r_m - r_{m_0}$.
Получаем $r_m - r_{m_0} = ta$. Тогда

$$r_m = r_{m_0} + ta \tag{3.3.1}$$

• *Уравнение* (3.3.1) *называется* ***параметрическим уравнением в векторной форме****.*

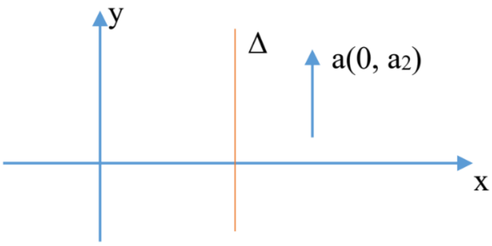
Запишем это уравнение через координаты:

$$\begin{cases} x = x_0 + t \cdot a_1, \\ y = y_0 + t \cdot a_2. \end{cases} \tag{3.3.2}$$

• *Уравнение* (3.3.2) *называется* ***параметрическим уравнением прямой в координатной форме****.*

Если прямая не параллельна ни одной из координатных осей, то $a_1 \neq 0$ и $a_2 \neq 0$.
Тогда $t = \dfrac{x - x_0}{a_1}$, $t = \dfrac{y - y_0}{a_2} \Rightarrow$
$$\frac{x - x_0}{a_1} = \frac{y - y_0}{a_2}. \tag{3.3.3}$$
• *Уравнение* (3.3.3) *называется* ***каноническим уравнением прямой***.

Если прямая параллельна одной из координатных осей, то уравнение прямой также может быть записано в каноническом виде. При этом равенство нулю в знаменателе означает не деление на нуль, а равенство нулю числителя для всех точек прямой. Рассмотрим прямую, изображенную на графике:



Ее параметрическое уравнение имеет вид: $\begin{cases} x = x_0, \\ y = y_0 + a_2 \cdot t. \end{cases}$
Тогда каноническое уравнение этой прямой имеет вид: $\dfrac{x - x_0}{0} = \dfrac{y - y_0}{a_2}$.

Если прямая Δ проходит через две различные точки $M_0(x_0, y_0)$ и $M_1(x_1, y_1)$, то в качестве направляющего вектора прямой можно взять вектор $\overrightarrow{M_0M_1}\ (x_1 - x_0, y_1 - y_0)$. Тогда уравнение прямой примет вид:
$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0}. \tag{3.3.4}$$
• *Уравнение* (3.3.4) *называется* ***уравнением прямой, проходящей через 2 точки***.

Пусть прямая Δ не параллельна оси Oy, и пусть $a(a_1, a_2)$ — ее направляющий вектор $(a_1 \neq 0)$ и точка $M_0(x_0, y_0) \in \Delta$.

• ***Углом наклона прямой*** Δ *к оси* Ox *называется угол, на который необходимо повернуть ось* Ox *против часовой стрелки, чтобы она стала параллельна прямой* Δ.

• *Тангенс угла наклона прямой к оси* Ox *называется* ***угловым коэффициентом прямой***.

То есть если φ — угол наклона, а k — угловой коэффициент, то $k = \operatorname{tg} \varphi$.

Теорема. *Если вектор* $a(a_1, a_2)$ *— направляющий вектор прямой* Δ, *то* $k = \dfrac{a_2}{a_1}$.

♦

1. Пусть направляющий вектор a прямой Δ направлен так, что $\angle(a, i) = \varphi$. Тогда
   $a_1 = \text{ПР}_{Ox} a = \text{ПР}_i a = |a| \cdot \cos(a, i) = |a| \cdot \cos \varphi$.
   $a_2 = \text{ПР}_{Oy} a = \text{ПР}_j a = |a| \cdot \cos(a, j) = |a| \cdot \cos(\pm(\frac{\pi}{2} - \varphi)) = |a| \cdot \sin \varphi$.
   Тогда $\dfrac{a_2}{a_1} = \dfrac{|a| \cdot \sin \varphi}{|a| \cdot \cos \varphi} = \operatorname{tg} \varphi = k$.

2. Пусть вектор a направлен так, что $\angle(a, i) \neq \varphi$. Тогда вектор $b = -a$ также является направляющим вектором и $\angle(b, i) = \varphi$. По доказанному выше, $k = \dfrac{b_2}{b_1} = \dfrac{-a_2}{-a_1} = \dfrac{a_2}{a_1}$.

$\boxtimes$

Пусть $\dfrac{x - x_0}{a_1} = \dfrac{y - y_0}{a_2}$ — каноническое уравнение прямой Δ. Тогда, домножим все на a_2 и получим $y - y_0 = \dfrac{a_2}{a_1}(x - x_0) = k(x - x_0) \Rightarrow$

$$y - y_0 = k(x - x_0). \tag{3.3.5}$$

• *Уравнение* (3.3.5) *называется* ***уравнением с угловым коэффициентом прямой, проходящей через точку с координатами*** (x_0, y_0).

Обозначим $b = y_0 - kx_0$. Тогда уравнение прямой имеет вид

$$y = kx + b. \tag{3.3.6}$$

• *Уравнение* (3.3.6) *называется* ***уравнением прямой с угловым коэффициентом****.*

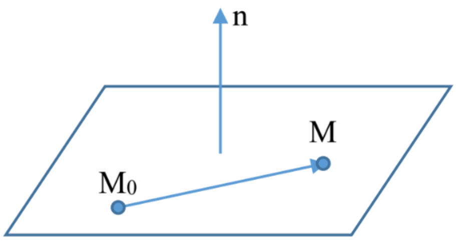
Глава 4

Уравнения плоскости и прямой в пространстве

4.1 Уравнение плоскости.

• ***Нормальным вектором плоскости*** *называется ненулевой вектор, перпендикулярный плоскости.*

Рассмотрим некоторую плоскость Π, проходящую через точку $M_0(x_0, y_0, z_0)$ перпендикулярно вектору $n(A, B, C)$.



Точка $M(x, y, z) \in \Pi \iff \overrightarrow{M_0M} \perp n \iff \overrightarrow{M_0M} \cdot n = 0$. Таким образом получаем уравнение

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \tag{4.1.1}$$

• *Уравнение* (4.1.1) *называется* ***уравнением плоскости, проходящей через точку*** $M_0(x_0, y_0, z_0)$ ***перпендикулярно вектору*** $n(A, B, C)$.

Обозначим $D = -Ax_0 - By_0 - Cz_0$, тогда уравнение (4.1.1) примет вид

$$Ax + By + Cz + D = 0 \quad A^2 + B^2 + C^2 \neq 0. \tag{4.1.2}$$

• *Уравнение* (4.1.2) *называется* ***общим уравнением плоскости.***

Теорема. *Любая плоскость в пространстве может быть задана уравнением вида* (4.1.2). *Любое уравнение вида* (4.1.2) *определяет плоскость.*

♦ Доказательство аналогично подобной теореме для случая с прямой. ⊠

Теорема. *Если два общих уравнения* $\begin{matrix} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0 \end{matrix}$ *определяют одну плоскость, то существует действительное число* λ *такое, что*

$$A_1 = \lambda A_2,\ B_1 = \lambda B_2,\ C_1 = \lambda C_2,\ D_1 = \lambda D_2.$$

♦ Доказательство аналогично подобной теореме для случая с прямой. ⊠

Пусть даны два нормальных вектора $n_1(A_1, B_1, C_1)$ и $n_2(A_2, B_2, C_2)$. Тогда косинус угла между плоскостями равен косинусу угла между нормальными векторами к плоскостям:

$$\cos(n_1, n_2) = \frac{n_1 \cdot n_2}{|n_1| \cdot |n_2|}.$$

• *Общее уравнение плоскости называется* ***полным****, если все его коэффициенты отличны от нуля, в противном случае называется* ***неполным****.*

Неполные уравнения описывают плоскости, параллельные координатным осям и (или) проходящие через начало координат.

Пусть $Ax + By + Cz + D = 0$ — полное общее уравнение плоскости. Перенесем D с противоположным знаком в правую часть уравнения и разделим на него же. Тогда получим

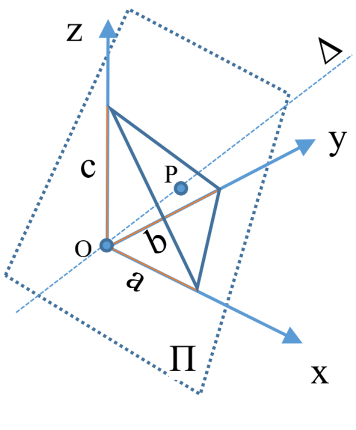
$$\frac{Ax}{-D} + \frac{By}{-D} + \frac{Cz}{-D} = 1 \Rightarrow \frac{x}{\frac{-D}{A}} + \frac{y}{\frac{-D}{B}} + \frac{z}{\frac{-D}{C}} = 1 \Rightarrow \left[\frac{-D}{A} = a, \frac{-D}{B} = b, \frac{-D}{C} = c\right] \Rightarrow$$

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \qquad (4.1.3)$$

• *Уравнение* (4.1.3) *называется* ***уравнением плоскости в отрезках****.*

Числа a, b, c являются величинами направленных отрезков, отсекаемых плоскостью по координатным осям от начала координат.

Пусть Π — произвольная плоскость. Построим прямую Δ, проходящую через начало координат перпендикулярно плоскости Π. Обозначим через P точку пересечения прямой Δ и плоскости Π. Через p обозначим $|\overline{OP}|$, а через n — единичный вектор, перпендикулярный плоскости Π и сонаправленный с вектором $\overrightarrow{OP}$. Если точки O и P совпадают, то направление единичного нормального вектора n выбираем произвольно.



Так как n — единичный вектор, то его координаты $n(cos\alpha, cos\beta, cos\gamma)$, где α, β, γ — углы между n и осями координат Ox, Oy, Oz соответственно.

Точка $M(x, y, z) \in \Pi \iff x \cdot \cos\alpha + y \cdot \cos\beta + z \cdot \cos\gamma = p$. Перенесем p в левую сторону и получим

$$x \cdot \cos\alpha + y \cdot \cos\beta + z \cdot \cos\gamma - p = 0. \qquad (4.1.4)$$

• *Уравнение* (4.1.4) *называется* ***нормальным уравнением плоскости***.

Нормальное уравнение можно построить из общего уравнения, домножив его на нормирующий множитель $\lambda = \pm \dfrac{1}{\sqrt{A^2+B^2+C^2}}$, где знак λ противоположен знаку D.

Пусть $M(x,y,z)$ — некоторая произвольная точка пространства, M_1 — ее проекция на плоскость Π, тогда $\rho(M,\Pi) = |\overline{M_1M}|$ — расстояние от точки M до плоскости Π.

• ***Отклонением*** $\delta(M, \Pi)$ *называется число, равное*

$$\delta(M,\Pi) = \begin{cases} \rho(M,\Pi), \overrightarrow{M_1M} \uparrow\uparrow n, \\ -\rho(M,\Pi), \overrightarrow{M_1M} \uparrow\downarrow n. \end{cases}$$

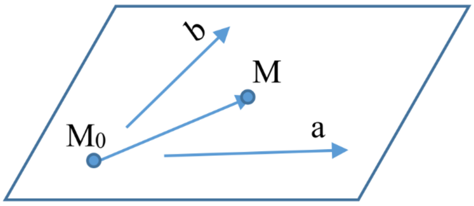
Теорема. *Если* $x\cdot\cos\alpha + y\cdot\cos\beta + z\cdot\cos\gamma - p = 0$ *— нормальное уравнение плоскости П, то отклонение точки* $M_0(x_0,y_0,z_0)$ *от плоскости П равняется:*

$$\delta(M_0, \Pi) = x_0\cdot\cos\alpha + y_0\cdot\cos\beta + z_0\cdot\cos\gamma - p.$$

⧫ Доказательство аналогично подобной теореме для случая с прямой. ⊠

Заметим, что $\rho(M_0,\Pi) = |x_0\cdot\cos\alpha + y_0\cdot\cos\beta + z_0\cdot\cos\gamma - p|$.

Пусть плоскость Π проходит через точку $M_0(x_0,y_0,z_0)$ параллельно двум неколлинеарным векторам $a(a_1,a_2,a_3)$ и $b(b_1,b_2,b_3)$.



Тогда $M(x,y,z) \in \Pi \iff \overrightarrow{M_0M}, a, b$ — компланарные $\iff \overrightarrow{M_0M}\cdot a\cdot b = 0 \iff$

$$(r_m - r_{m_0})\cdot a\cdot b = 0. \tag{4.1.5}$$

• *Уравнение* (4.1.5) *называется* ***уравнением плоскости, проходящей через точку*** M_0***, параллельно неколлениарным векторам*** a ***и*** b ***в векторной форме***.

Так же это уравнение можно записать через определитель:

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = 0. \tag{4.1.6}$$

• *Уравнение* (4.1.6) *называется* ***уравнением плоскости, проходящей через точку*** M_0***, параллельно неколлениарным векторам*** a ***и*** b.

Если плоскость Π проходит через три точки, не лежащие на одной прямой $M_0(x_0,y_0,z_0)$, $M_1(x_1,y_1,z_1)$, $M_2(x_2,y_2,z_2)$, то в качестве неколлинеарных векторов, параллельных плоскости, можно взять $\overrightarrow{M_0M}$ и $\overrightarrow{M_0M_2}$ и тогда уравнение примет вид:

$$\begin{vmatrix} x-x_0 & y-y_0 & z-z_0 \\ x_1-x_0 & y_1-y_0 & z_1-z_0 \\ x_2-x_0 & y_2-y_0 & z_2-z_0 \end{vmatrix} = 0. \tag{4.1.7}$$

• *Уравнение* (4.1.7) *называется* ***уравнением плоскости, проходящей через три точки***.

Так как векторы a и b неколлинеарны, то они на плоскости Π образуют базис. Следовательно, любой вектор, параллельный плоскости, в том числе и $\overrightarrow{M_0M}$, если точка M принадлежит плоскости Π, может быть разложен по этому базису, то есть представим в виде $\overrightarrow{M_0M} = t \cdot a + s \cdot b$, где $t, s \in \mathbb{R}$. Тогда

$$r_m = r_{m_0} + t \cdot a + s \cdot b. \tag{4.1.8}$$

• *Уравнение* (4.1.8) *называется* ***параметрическим уравнением плоскости в векторной форме***.

Распишем векторы по координатам и получим

$$\begin{cases} x = x_0 + t \cdot a_1 + s \cdot b_1, \\ y = y_0 + t \cdot a_2 + s \cdot b_2, \\ z = z_0 + t \cdot a_3 + s \cdot b_3. \end{cases} \tag{4.1.9}$$

• *Уравнение* (4.1.9) *называется* ***параметрическим уравнением плоскости в координатной форме***.

4.2 Уравнение прямой в пространстве.

• *Ненулевой вектор, параллельный прямой, называется* ***направляющим вектором*** *прямой.*

Рассмотрим некоторую прямую Δ и пусть $a(a_1, a_2, a_3)$ — ее направляющий вектор, $M_0(x_0, y_0, z_0)$ — точка на прямой Δ. Тогда точка $M(x, y, z) \in \Delta \Longleftrightarrow \overrightarrow{M_0M} \parallel a \Longleftrightarrow$ $\exists t \in \mathbb{R} : \overrightarrow{M_0M} = t \cdot a$. Отсюда получаем

$$r_m = r_{m_0} + t \cdot a. \tag{4.2.1}$$

• *Уравнение* (4.2.1) *называется* ***параметрическим уравнением прямой в векторной форме***.

Распишем это уравнения через координаты векторов и получим

$$\begin{cases} x = x_0 + t \cdot a_1, \\ y = y_0 + t \cdot a_2, \\ z = z_0 + t \cdot a_3. \end{cases} \tag{4.2.2}$$

• *Уравнение* (4.2.2) *называется* ***параметрическим уравнением прямой в координатной форме***.

Выразив параметр t из каждого параметрического уравнения получим

$$\frac{x - x_0}{a_1} = \frac{y - y_0}{a_2} = \frac{z - z_0}{a_3}. \tag{4.2.3}$$

• *Уравнение* (4.2.3) *называется* ***каноническим уравнением прямой***.

Если прямая Δ проходит через две различные точки $M_0(x_0, y_0, z_0)$ и $M_1(x_1, y_1, z_1)$, то в качестве направляющего вектора можно взять вектор $\overrightarrow{M_0M_1}$. Подставляем координаты в каноническое уравнение и получаем

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = \frac{z - z_0}{z_1 - z_0}. \tag{4.2.4}$$

• *Уравнение* (4.2.4) *называется* ***уравнением прямой по двум точкам.***

Пусть $A_1x+B_1y+C_1z+D_1 = 0$ и $A_2x+B_2y+C_2z+D_2 = 0$ — две различные плоскости, проходящие через прямую Δ, тогда точка $M(x, y, z) \in \Delta \iff$ ее координаты удовлетворяют каждому уравнению плоскости. Следовательно,

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \tag{4.2.5}$$

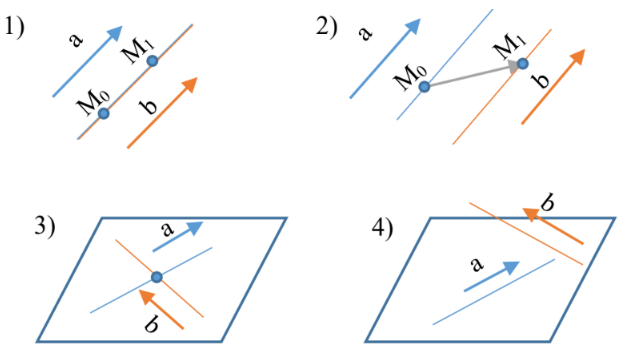
• *Уравнение* (4.2.5) *называется* ***уравнением прямой, заданной как линия пересечения двух плоскостей (общим уравнением прямой).***

Рассмотрим взаимное расположение прямых в пространстве. Возьмём две прямые Δ_1 и Δ_2, проходящие через точки M_0 и M_1, описанные каноническими уравненениями

$$\frac{x - x_0}{a_1} = \frac{y - y_0}{a_2} = \frac{z - z_0}{a_3} \text{ и } \frac{x - x_1}{b_1} = \frac{y - y_1}{b_2} = \frac{z - z_1}{b_3}$$

соответственно. При этом $a(a_1, a_2, a_3)$ и $b(b_1, b_2, b_3)$ — два направляющих вектора соответственно.

1. прямые совпадают, если $a \parallel b \parallel \overrightarrow{M_0M_1} \Rightarrow \dfrac{a_1}{b_1} = \dfrac{a_2}{b_2} = \dfrac{a_3}{b_3}$;
2. прямые параллельны, если $a \parallel b$, $\overrightarrow{M_0M_1} \nparallel a, b$;
3. прямые пересекаются, если $a \nparallel b \nparallel \overrightarrow{M_0M_1}$, но при этом $\overrightarrow{M_0M_1} \cdot a \cdot b = 0$ (то есть векторы $a, b, \overrightarrow{M_0M_1}$ компланарны);
4. прямые скрещивающиеся, если $a \nparallel b \nparallel \overrightarrow{M_0M_1}$, но при этом $\overrightarrow{M_0M_1} \cdot a \cdot b \neq 0$ (то есть $a, b, \overrightarrow{M_0M_1}$ не компланарны).



Заметим, что угол между двумя прямыми равен углу между их направляющими векторами, то есть

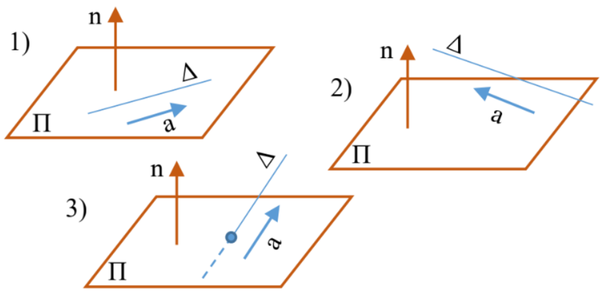
$$cos(\Delta_1, \Delta_2) = \frac{|a \cdot b|}{|a| \cdot |b|}.$$

Рассмотрим взаимное расположение прямой и плоскости. Пусть прямая задана каноническим уравнением $\dfrac{x-x_0}{a_1}=\dfrac{y-y_0}{a_2}=\dfrac{z-z_0}{a_3}$, а плоскость задана общим уравнением $Ax+By+Cz+D=0$.

1. Если прямая лежит на плоскости, то $a \perp n \iff a \cdot n = 0$, а сами прямая и плоскость имеют одну общую точку.

2. Прямая параллельна плоскости, если $a \cdot n = 0$, а общих точек у них нет.

3. Прямая и плоскость пересекаются, если $a \cdot n \neq 0$.



При этом угол между прямой и плоскостью равен углу между нормальным вектором плоскости и направляющим вектором прямой.

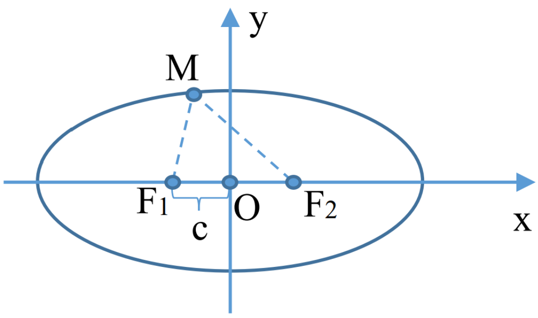
$$sin(\Delta, \Pi) = \frac{|a \cdot n|}{|a| \cdot |n|}.$$

Глава 5

Линии и поверхности второго порядка

5.1 Эллипс.

• ***Эллипсом*** *называется множество всех точек плоскости, сумма расстояний от которых до данных двух точек* F_1 *и* F_2 *есть величина постоянная. При этом точки* F_1 *и* F_2 *называются* ***фокусами эллипса***.



Выведем формулу: пусть F_1 и F_2 – фокусы, $|\overline{F_1F_2}| = 2c$.

Построим ДПСК, чтобы Ox проходила через фокусы так, чтобы направление вектора $\overrightarrow{F_1F_2}$ совпадало с направлением оси x, а начало координат располагалось в центе $\overline{F_1F_2}$. Тогда фокусы F_1 и F_2 имеют координаты $(-c, 0)$ и $(c, 0)$ соответственно. И пусть $2a$ —- сумма расстояний от точки эллипса $M(x, y)$ до фокусов. Заметим, что $2a > 2c \Rightarrow a > c$. Отрезки MF_1 и MF_2 назовём **фокальными радиусами точки** M.

По определению $|\overline{MF_1}| + |\overline{MF_2}| = 2a \Rightarrow$

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a. \tag{5.1.1}$$

Преобразуем уравнение (5.1.1): $\sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y^2} \Leftrightarrow$ (возведём в квадрат) $(x+c)^2 + y^2 = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2 \Leftrightarrow$ (перенесём выражение с корнем в левую часть, всё остальное в правую и раскроем квадраты) $4a\sqrt{(x-c)^2 + y^2} = 4a^2 + x^2 + c^2 - 2xc - x^2 - c^2 - 2xc + y^2 - y^2 \Leftrightarrow$ (уберём подобные)
$4a\sqrt{(x-c)^2 + y^2 = 4a^2 - 4xc} \Leftrightarrow$ (сократим на 4) $a\sqrt{(x-c)^2 + y^2 = a2 - xc} \Leftrightarrow$ (снова возведём в квадрат) $a^2((x-c)^2 + y^2) = (a^2 - xc)^2 \Leftrightarrow$ (раскроем скобки) $a^2x^2 - 2a^2xc + a^2c^2 + a^2y^2 = a^4 - 2a^2xc + x^2c^2 \Leftrightarrow$ (сгруппируем x и y слева, a и c справа и уберём подобные) $(a^2 - c^2)c^2 + a^2y^2 = a^2(a^2 - c^2)$.
Обозначим $b = \sqrt{a^2 - c^2}$ (так как $a > c$). Заметим, что $b < a$. Тогда $b^2x^2 + a^2y^2 = a^2b^2$.
Разделим обе части уравнения на a^2b^2:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{5.1.2}$$

• *Уравнение* (5.1.2) *называется* ***каноническим уравнением эллипса***.

Любая точка, удовлетворяющая уравнению (5.1.1) (то есть принадлежащая эллипсу), удовлетворяет уравнению (5.1.2). Покажем, что любая точка, удовлетворяющая уравнению (5.1.2), будет принадлежать эллипсу.

Пусть $M_1(x_1, y_1)$ удовлетворяет (5.1.2) $\Rightarrow$

$$\frac{x_1^2}{a^2} + \frac{y_1^2}{b^2} = 1. \tag{5.1.3}$$

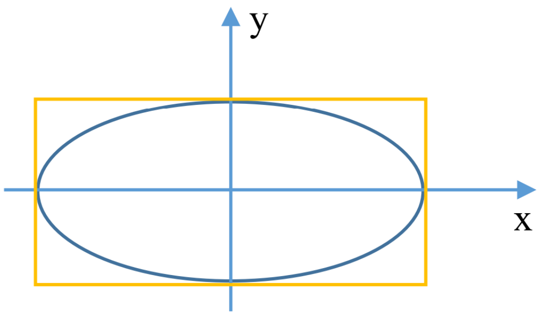
$|\overline{M_1F_1}| = \sqrt{(x_1+c)^2+y^2} = [\text{из (5.1.3) } y^2 = b^2(1-\frac{x_1^2}{a^2})] = \sqrt{(x_1+c)^2+b^2(1-\frac{x_1^2}{a^2})} = [b^2 = a^2 - c^2] = \sqrt{(x_1+c)^2+(a^2-c^2)(1-\frac{x_1^2}{a^2})} = \sqrt{x_1^2+2x_1c+c^2+a^2 = c^2 - x_1^2 + \frac{c^2x_1^2}{a^2}} = \sqrt{a^2+2x_1c+\frac{c^2x_1^2}{a^2}} = \sqrt{(a+\frac{c}{a}x_1)^2} = |a+\frac{c}{a}x_1|$. Аналогично $|\overline{M_1F_2}| = |a-\frac{c}{a}x_1|$.

Раскроем модули, оценив подмодульные выражения. Из (5.1.3) следует, что $\dfrac{x_1^2}{a^2} \leqslant 1 \Rightarrow -1 \leqslant \dfrac{x_1}{a} \leqslant 1 \Rightarrow -a < -c < \dfrac{cx_1}{a} < c < a \Rightarrow |a+\dfrac{c}{a}x_1| = a+\dfrac{c}{a}x_1, |a-\dfrac{c}{a}x_1| = a-\dfrac{c}{a}x_1 \Rightarrow |\overline{M_1F_1}|+|\overline{M_1F_2}| = a+\dfrac{c}{a}x_1+a-\dfrac{c}{a}x_1 = 2a \Rightarrow$ точка M_1 удовлетворяет уравнению (5.1.1) $\Rightarrow M_1$ — точка, принадлежащая эллипсу $\Rightarrow$ уравнение (5.1.2) является каноническим уравнением эллипса, где $b < a$.

В случае, если $b > a$, каноническое уравнение имеет вид $\dfrac{x_1^2}{b^2} + \dfrac{y_1^2}{a^2} = 1$, а фокусы будут лежать не на Ox, а на Oy.

Исследование формы эллипса по его каноническому уравнению:

$\dfrac{x_1^2}{a^2}$ и $\dfrac{y_1^2}{b^2} \leqslant 1 \Rightarrow |x_1| \leqslant |a|$ и $|y_1| \leqslant |b|$, то есть эллипс — линия ограниченная и расположенная внутри прямоугольника, которая ограничена прямыми $x = \pm a, y = \pm b$.



В уравнении эллипса обе координаты имеют чётные степени, следовательно, если точка $M_1(x, y)$ принадлежит эллипсу, то и точки $M_2(-x, y), M_3(x, -y), M_4(-x, -y)$ также будут принадлежать эллипсу.

Эллипс симметричен относительно начала координат и осей.

• *Центр симметрии называется* ***центром эллипса.***

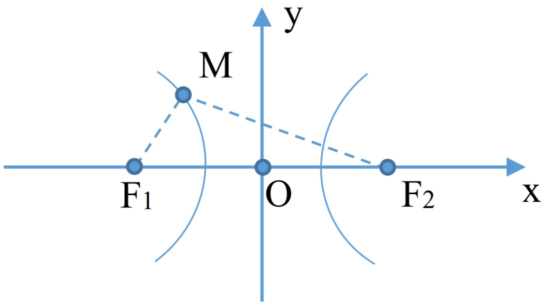
• *Прямая, проходящая через фокусы, называется* ***большой осью эллипса****, перпендикулярная ей и проходящая через центр –* ***малая ось.*** *Точки пересечения осей эллипса* *с эллипсом (Ox в точках $A_1(a,0)$ и $A_2(-a,0)$, Oy в точках $B_1(0,b)$ и $B_2(0,-b)$) называются* ***вершинами эллипса****. Отрезки OA_1, OA_2, OB_1, OB_2 –* ***полуоси*** *(большой и малой соответственно) эллипса с длинами a и b.*

Рассмотрим поведение эллипса в первой четверти ДПСК: $\begin{cases} 0 < x \leqslant a \\ 0 < y \leqslant b \end{cases}$

$y = b\sqrt{1-\frac{x^2}{a^2}}$ (из канонического уравнения) $\Rightarrow y = \dfrac{b}{a}(a^2-x^2)^{\frac{1}{2}} \Rightarrow y' = \dfrac{b}{a}\cdot\dfrac{1}{2}(a^2-x^2)^{-\frac{1}{2}}(-2x) = -\dfrac{b}{a}x(a^2-x^2)^{-\frac{1}{2}} < 0$, то есть функция убывает. $y'' = -\dfrac{b}{a}\cdot\left((a^2-x^2)^{-\frac{1}{2}}x + \dfrac{x(2x)(a^2-x^2)^{-\frac{3}{2}}}{2}\right) = -\dfrac{ab}{(a^2-x^2)^{\frac{3}{2}}} < 0$, то есть график функции вогнут.

5.2 Гипербола.

• ***Гиперболой*** *называется множество точек плоскости, модуль разности расстояний от которых до заранее заданных точек F_1 и F_2, называемых* ***фокусами гиперболы****, есть величина постоянная и меньшая, чем расстояние между фокусами..*



Выведем формулу: пусть F_1 и F_2 – фокусы, $|\overline{F_1F_2}| = 2c$.

Построим ДПСК: ось Ox проходит через фокусы и сонаправлена с $\overrightarrow{F_1F_2}$, а цент между фокусами — начало координат. Обозначим модуль разности точек гиперболы до фокусов за $2a \Rightarrow 2a < 2c \Rightarrow a < c$.

Пусть $M(x,y)$ — произвольная точка гиперболы. Как и в эллипсе, отрезки MF_1 иMF_2 будем называть фокальными радиусами.

По определению $||\overline{MF_1}| - |\overline{MF_2}|| = 2a \Rightarrow$

$$|\sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2}| = 2a. \tag{5.2.1}$$

• *Уравнение* (5.2.1) *называется* ***уравнением гиперболы****.*

Преобразуем это уравнение. Раскроем модуль и получим: $\sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2} = \pm 2a \Leftrightarrow \sqrt{(x+c)^2+y^2} = \pm 2a + \sqrt{(x-c)^2+y^2} \Leftrightarrow$ (возведём в квадрат и раскроем квадраты сумм и разности) $x^2+c^2+2xc+y^2 = 4a^2+x^2+c^2-2xc+y^2 \pm 4a\sqrt{(x-c)^2+y^2} \Leftrightarrow$ (приведём подобные) $\pm a\sqrt{(x-c)^2+y^2} = -a^2+xc \Leftrightarrow$ (снова возведём в квадрат и раскроем квадраты сумм и разности) $a^2(x^2+c^2-2xc+y^2) = a^4+x^2c^2-2a^2xc \Leftrightarrow$ (снова приводим подобные, раскрыв скобку слева, после чего группируем по x и y справа, а без них слева) $a^2(c^2-a^2) = x^2(c^2-a^2) - a^2y^2$. Так как $c > a \Rightarrow c^2 - a^2 > 0 \Rightarrow$ обозначим $b = \sqrt{c^2-a^2} \Rightarrow x^2b^2 - y^2a^2 = a^2b^2$. a и $b \neq 0 \Rightarrow$ делим на a^2b^2 и получаем

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{5.2.2}$$

Любая точка, удовлетворяющая уравнению (5.2.1), то есть принадлежащая гиперболе, удовлетворяет и уравнению (5.2.2). Покажем, что любая точка, удовлетворяющая уравнению (5.2.2), будет принадлежать гиперболе.

Пусть точка $M_1(x_1, y_1)$ удовлетворяет уравнению (5.2.2)$\Rightarrow$

$$\frac{x_1^2}{a^2} - \frac{y_1^2}{b^2} = 1. \tag{5.2.3}$$

$|\overline{M_1F_1}| = \sqrt{(x_1+c)^2 + y^2} =$ [из (5.2.3) $y^2 = b^2(\frac{x_1^2}{a^2} - 1)$] $= \sqrt{(x_1+c)^2 + b^2(\frac{x_1^2}{a^2} - 1)} =$ $\sqrt{(x_1+c)^2 + (c^2-a^2)(\frac{x_1^2}{a^2} - 1)} = \sqrt{x_1^2 + 2x_1c + c^2 + a^2 - c^2 - x_1^2 + \frac{c^2x_1^2}{a^2}} = \sqrt{(a + \frac{c}{a}x_1)^2} =$ $|a + \frac{c}{a}x_1|$. Аналогично $|\overline{M_1F_2}| = |a - \frac{c}{a}x_1|$.
Из (5.2.3) следует, что $\frac{x_1^2}{a^2} \geqslant 1 \Rightarrow x_1 \geqslant a$ или $x_1 \leqslant -a$.

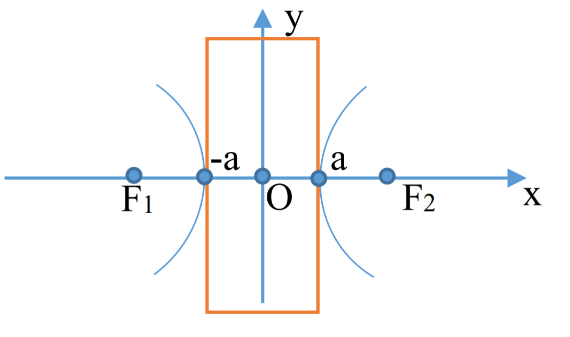
1. Пусть $x_1 \geqslant a \Rightarrow \frac{cx_1}{a} > c > a \Rightarrow |\overline{M_1F_1}| = a + \frac{c}{a}x_1, |\overline{M_1F_2}| = -a + \frac{c}{a}x_1 \Rightarrow |\overline{M_1F_1}| - |\overline{M_1F_2}| = 2a$.
2. Пусть $x_1 \leqslant -a \Rightarrow \frac{cx_1}{a} < -c < -a < 0 \Rightarrow |\overline{M_1F_1}| = -a - \frac{c}{a}x_1, |\overline{M_1F_2}| = a - \frac{c}{a}x_1 \Rightarrow |\overline{M_1F_1}| - |\overline{M_1F_2}| = -2a \Rightarrow ||\overline{M_1F_1}| - |\overline{M_1F_2}|| = 2a$.

Следовательно, точка M_1 принадлежит гиперболе. Тогда

- *Уравнение (5.2.2) называется* ***каноническим уравнением гиперболы****.*

Исследование формы гиперболы по её каноническому уравнению:

Из (5.2.2) следует, что $\frac{x_1^2}{a^2} \geqslant 1 \Rightarrow |x| \geqslant a \Rightarrow$ в полосе между прямыми $-a$ и a точек гиперболы нет.
Гиперболы пересекается с осью Ox в точках $A_1(a, 0)$ и $A_2(-a, 0)$.



Так как все степени четные, то гипербола симметрична относительно осей Ox, Oy и точки $O(0,0)$.

• *Центр симметрии гиперболы (центр между F_1 и F_2) называется* ***центром гиперболы****. Ось симметрии, проходящая через фокусы, называется* ***действительной (фокальной) осью****, а перпендикулярная ей ось, проходящая через центр гиперболы —* ***мнимой осью****. Точки пересечения гиперболы с действительной осью являются её* ***вершинами****. Величины a, b называются соответственно* ***действительными и мнимыми полуосями****. Если $a = b$, то гиперболу называют* ***равносторонней****.*

Исследуем формулу гиперболы в первой четверти, выразив y из канонического уравнения:
$y = \dfrac{b}{a}(x^2 - a^2)^{\frac{1}{2}}$.
$y' = \dfrac{1}{a} \cdot x(x^2 - a^2)^{\dfrac{1}{2}} > 0$, то есть функция возрастает;
$y'' = \dfrac{1}{a}(x^2 - a^2)^{-\frac{3}{2}}(-a^2) < 0$, то есть график функции вогнут.

Асимптота гиперболы в 1-й четверти: $y = kx + l$.
$k = \lim\limits_{x\to\infty} \dfrac{y}{x} = \lim\limits_{x\to\infty} \dfrac{b}{a} \cdot \dfrac{\sqrt{x^2 - a^2}}{x} = \dfrac{b}{a}$;
$l = \lim\limits_{x\to\infty} x\dfrac{b}{a}y = 0 \Rightarrow y = \dfrac{b}{a}x$.

5.3 Эксцентриситет и директрисы эллипса и гиперболы.

Рассмотрим эллипс, у которого фокусное расстояние равно $2c$, а сумма расстояний от точек эллипса до фокусов $2a$. Тогда каноническое уравнение эллипса имеет вид $\dfrac{x_1^2}{a^2} + \dfrac{y_1^2}{b^2} = 1$, где $b^2 = a^2 - c^2$.

Рассмотрим гиперболу, у которой фокусное расстояние равно $2c$, а модуль разности расстояний от точек гиперболы до фокусов $2a$. Тогда каноническое уравнение гиперболы имеет вид $\dfrac{x_1^2}{a^2} - \dfrac{y_1^2}{b^2} = 1$, где $b^2 = c^2 - a^2$.

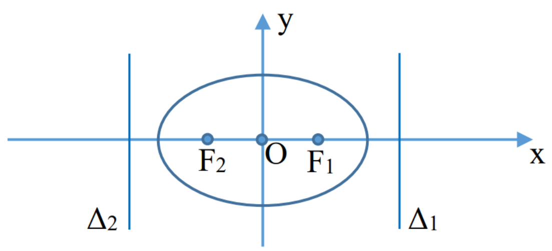
• ***Эксцентриситетом эллипса (гиперболы)*** *называется величина, равная $\varepsilon = \dfrac{c}{a}$.*

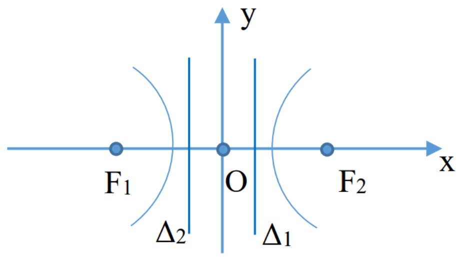
Заметим, что для эллипса $a > c \Rightarrow 0 < \varepsilon < 1$, для гиперболы $a < c \Rightarrow \varepsilon > 1$.

• *Прямые Δ_1 и Δ_2, проходящие перпендикулярно фокальной оси эллипса (гиперболы) на расстоянии равном $\dfrac{a}{\varepsilon}$ от центра эллипса (гиперболы), называются* ***директрисами эллипса (гиперболы)****.*

Если эллипс и гипербола заданы каноническим уравнением, то уравнение директрисы имеет вид $x = \pm\frac{a}{\varepsilon}$.

Директрисы расположены по разные стороны от Oy. Фокус и директриса, расположенные по одну сторону от Oy, называются **соответствующими друг другу**.





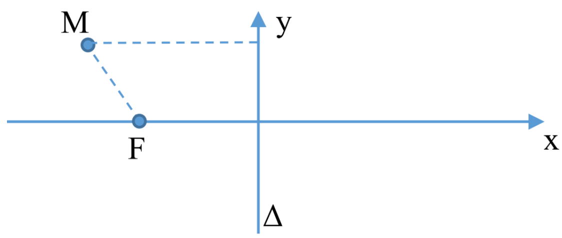
Теорема (основное свойство директрисы). *Отношение расстояния от произвольной точки эллипса (гиперболы) до фокуса F_i к расстоянию этой точки до соответствующей этому фокусу директрисы равна эксцентриситету эллипса (гиперболы). То есть*

$$\frac{p(M_0, F_i)}{p(M_0, \Delta_i)} = \varepsilon.$$

♦ Пусть точка $M_0(x_0, y_0)$ принадлежит эллипсу $\dfrac{x_1^2}{a^2} + \dfrac{y_1^2}{b^2} = 1$. Тогда $|\overline{M_0F_1}| = |a - \dfrac{c}{a}x_0| = |a - \varepsilon x_0|$ и $|\overline{M_0F_2}| = |a + \varepsilon x_0|$. Следовательно, так как $\Delta_1 = x + \dfrac{a}{\varepsilon}$, $\Delta_2 = -x - \dfrac{a}{\varepsilon}$, то $\rho(M_0, \Delta_1) = |x_0 - \dfrac{a}{\varepsilon}|$, $\rho(M_0, \Delta_2) = |x_0 + \dfrac{a}{\varepsilon}|$. Получаем $\dfrac{\rho(M_0, F_i)}{\rho(M_0, \Delta_i)} = \dfrac{|a - \varepsilon x_0|}{|x_0 - \frac{a}{\varepsilon}|} = \dfrac{|a - \varepsilon x_0|}{|\frac{a - \varepsilon x_0}{\varepsilon}|} = |\varepsilon| = \varepsilon$. Аналогично и во втором случае и для гиперболы. ⊠

Теорема. *Пусть в плоскости задана некоторая прямая Δ и точка F, не лежащая на ней. Множество точек плоскости, для которых отношение μ равняется отношению расстояния до F к расстоянию до Δ и отлично от 1, является эллипсом, если $\mu < 1$, и гиперболой, если $\mu > 1$. При этом точка F будет являться фокусом, Δ — директрисой, а μ — эксцентриситетом.*

♦ Построим ДПСК так, чтобы Oy совпадала с Δ, а Ox проходила через F.



Точка F будет иметь координаты $(k, 0)$, а $p(F, \Delta) = |k|$.

Пусть $M(x, y)$ — некоторая точка, удовлетворяющая отношению: $\mu = \dfrac{p(M, F)}{p(M, \Delta)}$. Преобразуем данное неравенство: $\dfrac{\sqrt{(x-k)^2 + y^2}}{|x|} = \mu \Rightarrow \sqrt{(x-k)^2 + y^2} = \mu|x| \Rightarrow$ (возведём в квадрат и раскроем скобки) $x^2 + k^2 - 2kx + y^2 = \mu^2 x^2 \Rightarrow x^2(1 - \mu^2) + k^2 - 2kx + y^2 = 0 \Rightarrow$ $x^2(1 - \mu^2) - 2kx + y^2 = -k^2 \Rightarrow$ (добавим в обе части неравенства выражение $\frac{k^2}{1-\mu^2}$)
$x^2(1 - \mu^2) - 2kx + \dfrac{k^2}{1 - \mu^2} + y^2 = \dfrac{k^2}{1 - u^2} - k^2 \Rightarrow$ (в левой части во всех слагаемых, за исключением y^2, выносим за скобки $(1 - \mu^2)$)
$(1 - \mu^2)(x^2 - \dfrac{2kx}{1 - \mu^2} + \dfrac{k^2}{(1 - \mu^2)^2}) + y^2 = \dfrac{k^2}{1 - \mu^2} - k^2 \Rightarrow$ (слева получился квадрат разности, а справа приведём к общему знаменателю)

$(1-\mu^2)(x^2 - \dfrac{k}{1-\mu^2})^2 + y^2 = \dfrac{k^2\mu^2}{1-\mu^2} \Rightarrow$ (поделим обе части на $1-\mu^2$)
$(x^2 - \dfrac{k}{1-\mu^2})^2 + \dfrac{y^2}{1-\mu^2} = \dfrac{k^2\mu^2}{(1-\mu^2)^2}$.

Выполним замену $a = |\dfrac{k\mu}{1-\mu^2}|$ и выполним преобразование параллельного переноса ДПСК:

$$\begin{cases} x' = x - \dfrac{k}{1-\mu^2}, \\ y' = y. \end{cases}$$

Тогда, поделив на правую часть, получим

$$\frac{(x')^2}{a^2} + \frac{(y')^2}{a^2(1-\mu^2)} = 1. \qquad (5.3.1)$$

Если $\mu < 1$, то уравнение (5.3.1) является каноническим уравнением эллипса, а если $\mu > 1$ — гиперболы.

Покажем, что μ — эксцентриситет. Если уравнение (5.3.1) задает эллипс, то $\varepsilon = \dfrac{c}{a} = [b^2 = a^2 - c^2 \Rightarrow c^2 = a^2 - b^2] = \dfrac{\sqrt{a^2-b^2}}{a} = [b^2 = a^2(1-\mu^2)$ из $(1)] = \dfrac{\sqrt{a^2 - a^2(1-\mu^2)}}{a} = \mu$.
Если гипербола, то $b^2 = c^2 - a^2 \Rightarrow c^2 = a^2 + b^2$ и аналогично.

$x' = \pm\dfrac{a}{\varepsilon} \Rightarrow x = \pm\dfrac{a}{\varepsilon} + \dfrac{k}{1-\mu^2} = \pm\dfrac{|\frac{k\mu}{1-\mu^2}|}{\varepsilon} + \dfrac{k}{1-\mu^2} \Rightarrow$ одно из уравнений имеет вид $x = 0$. В ДПСК $O'x'y'$ точка F будет иметь координаты $(\pm c, 0)$, директрисы будут иметь уравнение $\pm\dfrac{a}{\varepsilon} \Rightarrow$ выполнив обратный перенос ДПСК получим, что одна из директрис совпадает с прямой Δ и один из фокусов совпадает с точкой F. $\boxtimes$

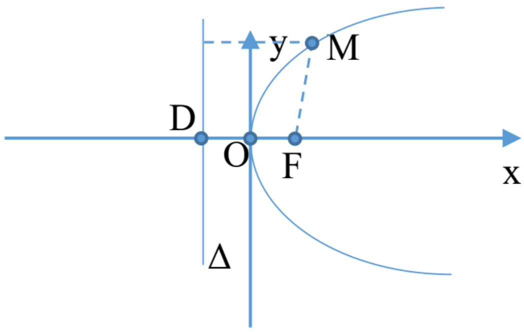
После доказательства теоремы определения эллипса и гиперболы можно дать по-новому.

• ***Эллипсом (гиперболой)*** *называется множество точек на плоскости, для которых отношения расстояния до заданной точки, называемой* ***фокусом****, к расстоянию до заданной прямой, называемой* ***директрисой****, постоянно меньше единицы (больше единицы).*

5.4 Парабола.

• ***Параболой*** *называется множество точек плоскости, каждая из которых равноудалена от точки называемой* ***фокусом*** *и от прямой называемой* ***директрисой****.*

Выведем уравнение параболы. Построим ДПСК так, чтобы ось Ox проходила через фокус F перпендикулярно прямой Δ. И пусть D — точка пересечения прямой Δ и оси Ox. Направление Ox выберем так, чтобы оно совпадало с направлением вектора $\overrightarrow{DF}$. Начало координат будет располагаться в центре между точками D и F.



Обозначим за p расстояние $\rho(F,\Delta)$. Тогда точка F будет иметь координаты $(\frac{p}{2},0)$ и координата x точки D равна $-\frac{p}{2}$. Следовательно, получим

$$x+\frac{p}{2}=0 \quad (x-\frac{p}{2}=0). \tag{5.4.1}$$

• *Уравнение* (5.4.1) *называется* ***нормальным уравнением параболы***.

Пусть $M(x,y)$ — произвольная точка параболы. Тогда $\rho(M,F)=\rho(M,\Delta) \Rightarrow \sqrt{(x-\frac{p}{2})^2+y^2} = |-x-\frac{p}{2}|$. Возведем в квадрат, раскроем скобки и получим $x^2+\frac{p^2}{2}-px+y^2=x^2\frac{p^2}{2}+px \Rightarrow$

$$y^2=2px. \tag{5.4.2}$$

• *Уравнение* (5.4.2) *называется* ***каноническим уравнением параболы***. *При этом параметр p называется* ***фокальным*** *параметром.*

Исследование формы параболы по её каноническому уравнению:

Пусть $x \geqslant 0$. Тогда парабола находится в I и IV четвертях. Так как уравнение содержит y в чётной степени, то парабола симметрична относительно оси Ox.

• *Ось симметрии называется* ***осью*** *параболы. Точка $O(0,0)$ пересечения оси параболы и самой параболы называется* ***вершиной*** *параболы.*

Рассмотрим поведение параболы в I четверти, то есть при $x>0, y>0$:
$y=\sqrt{2px}$. Тогда $y'=\dfrac{\sqrt{2p}}{2}x^{-\frac{1}{2}}>0$, то есть функция возрастает. И $y''=-\dfrac{\sqrt{2p}}{4}x^{-\frac{3}{2}}<0$, то есть функция выпукла. Асимптот нет, так как $k=\lim\limits_{x\to\infty}\dfrac{\sqrt{2px}}{x}=0$, а $b=\infty$.

5.5 Линии второго порядка.

• ***Алгебраической линией второго порядка*** *называется множество точек плоскости, координаты которых удовлетворяют уравнению*

$$Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0, \tag{5.5.1}$$

где коэффициенты A, B и C не обращаются в нуль одновременно. В случае, если уравнению (5.5.1) *не удовлетворяет ни одна из точек, то это уравнение определяет* ***мнимую линию***.

Определим, какие линии будут являться линиями второго порядка.

Пусть $B \neq 0$. Построим ДПСК $O'x'y'$, которая получена из ДПСК Oxy поворотом против часовой стрелки вокруг начала координат на угол φ. Тогда $\begin{cases} x = x'cos\varphi - y'sin\varphi, \\ y = x'sin\varphi + y'cos\varphi. \end{cases}$.
Заменим и получим $A(x'cos\varphi - y'sin\varphi)^2 + 2B(x'cos\varphi - y'sin\varphi) \cdot (x'cos\varphi + y'sin\varphi) + C(x'cos\varphi + y'sin\varphi)^2 + 2D(x'cos\varphi - y'sin\varphi) + 2E(x'cos\varphi + y'sin\varphi) + F = 0$.
Ракскрыв скобки и приведя подобные, получим

$$A'(x')^2 + 2Bx'y' + C'(y')^2 + 2D'x' + 2E'y' + F = 0. \tag{5.5.2}$$

Покажем, что коэффициенты A', B' и C' не обращаются одновременно в нуль:
$A' = A \cdot cos^2\varphi + 2B \cdot cos\varphi \cdot sin\varphi + C \cdot sin^2\varphi$;
$B' = -A \cdot cos\varphi \cdot sin\varphi + B \cdot (cos^2\varphi - sin^2\varphi) + C \cdot sin\varphi \cdot cos\varphi$;
$C' = A \cdot sin^2\varphi - 2B \cdot cos\varphi \cdot sin\varphi + C \cdot cos^2\varphi$.
Предположим, что A', B' и C' равны нулю. Сложив первое и третье уравнения, получим $A + C = 0$. Следовательно, $C = -A$. Из второго уравнения, приведя по формуле синуса и косинуса двойного угла, получим $0 = -A \cdot sin(2\varphi) + B \cdot cos(2\varphi)$, а из первого уравнения этой же заменой получим $0 = A \cdot cos(2\varphi) + B \cdot sin(2\varphi)$. Следовательно, получим систему

$$\begin{cases} A \cdot cos(2\varphi) + B \cdot sin(2\varphi) = 0, \\ -A \cdot sin(2\varphi) + B \cdot cos(2\varphi) = 0. \end{cases}$$

Определитель матрицы системы равен 1. В таком случае система будет иметь единственное решение $A = B = 0 = C$, что является противоречием с тем, что уравнение (5.5.1) — уравнение линии второго порядка. Значит коэффициенты A', B' и C' не обращаются в нуль одновременно.

Подберем угол φ так, чтобы коэффициент B' стал равен нулю. То есть $-A \cdot cos\varphi \cdot sin\varphi + B(cos^2\varphi - sin^2\varphi) + C \cdot sin\varphi \cdot cos\varphi = 0 \Leftrightarrow \dfrac{C - A}{2} \cdot sin(2\varphi) + B \cdot cos(2\varphi) = 0 \Leftrightarrow \dfrac{A - C}{2} \cdot sin(2\varphi) = B \cdot cos(2\varphi)$. Если $A - C \neq 0$, то $tg(2\varphi) = \dfrac{2B}{A - C}$. Если же $A - C = 0$, то $B \neq 0$. Следовательно, $cos2(\varphi) = 0 \Rightarrow 2\varphi = \dfrac{\pi}{2} \Rightarrow \varphi = \dfrac{\pi}{4}$.

В результате получим ДПСК, в которой уравнение линий второго порядка имеет вид

$$A'(x')^2 + C'(y')^2 + 2D'x' + 2E'y' + F = 0. \tag{5.5.3}$$

1. Пусть коэффициенты A' и C' ненулевые. Преобразуем уравнение следующим образом: добавим и вычтем числа $\dfrac{(D')^2}{A'}$ и $\dfrac{(E')^2}{C'}$. Везде, где есть x' или D' вынесем коэффициент A' за скобки. Аналогично, где y' или E' — C' за скобки. Обозначим за F' выражение $F' = F - \dfrac{(D')^2}{A'} - \dfrac{(E')^2}{C'}$. В итоге получим $A'\Big((x')^2 + 2\dfrac{D'x'}{A'} + (\dfrac{D'}{A'})^2\Big) + C'\Big((y')^2 + 2\dfrac{E'y'}{C'} + (\dfrac{E'}{C'})^2\Big) + F' = 0 \iff$ (свернем квадраты суммы) $A'(x' + \dfrac{D'}{A'})^2 + B'(y' + \dfrac{E'}{C'})^2 + F' = 0$. Выполним параллельный перенос ДПСК: $\begin{cases} x'' = x' + \dfrac{D'}{A'}, \\ y'' = y' + \dfrac{E'}{C'}. \end{cases}$

   Отсюда получим
   $$A'(x'')^2 + C'(y'')^2 = -F'. \tag{5.5.4}$$

   Здесь рассмотрим два случая:

(a) Пусть $F' \neq 0$. Тогда поделим уравнение (5.5.4) на $-F'$, перенесем A' и C' в знаменатели и получим $\frac{(x'')^2}{\frac{-F}{A'}} + \frac{(y'')^2}{\frac{-F}{C'}} = 1$.

Если знаменатели положительные, то уравнение примет вид $\frac{(x'')^2}{a^2} + \frac{(y'')^2}{b^2} = 1$, что является **каноническим уравнением эллипса**.

Если знаменатели отрицательные, то уравнение примет вид $\frac{(x'')^2}{a^2} + \frac{(y'')^2}{b^2} = -1$, что является **каноническим уравнением мнимого эллипса**.

(b) Пусть $F' = 0$. Тогда поделим обе части уравнения (5.5.4) на $A'C'$ и получим уравнение $\frac{(x'')^2}{C'} + \frac{(y'')^2}{A'} = 0$.

Если знаменатели разных знаков, то уравнение примет вид $\frac{(x'')^2}{a^2} - \frac{(y'')^2}{b^2} = 0$, что является **каноническим уравнением пары пересекающиъся прямых**.

Если знаменатели имеют один знак, то уравнение примет вид $\frac{(x'')^2}{a^2} + \frac{(y'')^2}{b^2} = 0$, что является **каноническим уравнением пары мнимых пересекающихся прямых**. Заметим, что точка $(0, 0)$ удовлетворяет этому уравнению.

2. Пусть $A' \neq 0, C' = 0$ (аналогично будет и наоборот). Уравнение (5.5.3) примет вид $A'((x')^2 + 2\frac{D'x'}{A'} + (\frac{D'}{A'})^2) + 2E'y' + F - \frac{(D')^2}{A'} = 0$. Аналогично с пунктом 1 рассмотрим 2 случая:

(a) Пусть $E' \neq 0$. Тогда вынесем во всех слагаемых, кроме тех, у которых есть множитель A', за скобку $2E'$. Получим

$$A'(x' + \frac{D'}{A'})^2 + 2E'(y' + \frac{F}{2E'} + \frac{(D')^2}{2E'A'}) = 0. \tag{5.5.5}$$

Выполним преобразование симметрии и получим уравнение вида $(y'')^2 = 2px''$, что является **каноническим уравнением параболы**.

(b) Пусть $E' = 0$. Тогда уравнение (5.5.5) примет вид $A'(x' + \frac{D'}{A'})^2 + F - \frac{(D')^2}{A'} = 0$. Выполним преобразование переноса $\begin{cases} x'' = x' + \frac{D'}{A'}, \\ y'' = y'; \end{cases}$ и заменим $F - \frac{(D')^2}{A'}$ на F''. В результате этих действий получим $(x'')^2 + \frac{F''}{A'} = 0$.

Если $\frac{F''}{A'} < 0$, то $(x'')^2 - a^2 = 0$, что является **каноническим уравнением пары параллельных прямых**.

Если $\frac{F''}{A'} = 0$, то $(x'')^2$, что является **каноническим уравнением пары совпадающих прямых**.

А если $\frac{F''}{A'} > 0$, то $(x'')^2 + a^2 = 0$, что является **каноническим уравнением пары мнимых параллельных прямых**.

Таким образом нами была доказана следующая теорема.

Теорема. *Для любой линии второго порядка существует ДПСК, в которой данная линия задается одним из следующих канонических уравнений:*

1. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ — *эллипс;*

2. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$ — *мнимый эллипс;*

3. $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ — *гипербола;*

4. $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$ — *пара пересекающихся прямых;*

5. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$ — *пара мнимых пересекающихся прямых;*

6. $y^2 = 2px$ — *парабола;*

7. $x^2 - a^2 = 0$ — *пара параллельных прямых;*

8. $x^2 + a^2 = 0$ — *пара мнимых параллельных прямых;*

9. $x^2 = 0$ — *пара совпадающих прямых.*

5.6 Поверхности второго порядка.

• ***Поверхностью второго порядка*** *называется множество точек пространства, декартовые координаты которых удовлетворяют уравнению следующего вида:*

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0. \qquad (5.6.1)$$

Если уравнению (5.6.1) не удовлетворяет ни одна точка, то поверхность мнимая.

Теорема. *Для любой поверхности второго порядка существует пространтсвенная ДПСК в которой эта поверхность определена одним из следующих канонических уравнений:*

1. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ — *эллипсоид;*

2. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = -1$ — *мнимый эллипсоид;*

3. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$ — *однополостной гиперболоид;*

4. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1$ — *двуполостной гиперболоид;*

5. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0$ — *конус второго порядка;*

6. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 0$ — *мнимый конус второго порядка;*

7. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 2z$ — *эллиптический параболоид;*

8. $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 2z$ — *гиперболический параболоид;*

9. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ — *эллиптический цилиндр;*

10. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$ — *мнимый эллиптический цилиндр;*

11. $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ — *гиперболический цилиндр;*

12. $y^2 = 2px$ — *параболоический цилиндр;*

13. $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$ — *пара пересекающихся плоскостей;*

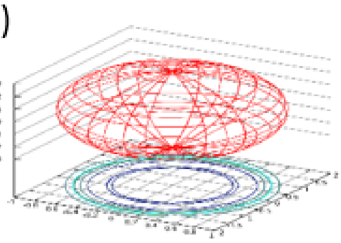
14. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$ — *пара мнимых пересекающихся плоскостей;*

15. $x^2 - a^2 = 0$ — *пара параллельных плоскостей;*

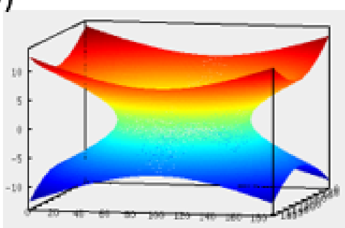
16. $x^2 + a^2 = 0$ — *пара мнимых параллельных плоскостей;*

17. $x^2 = 0$ — *пара совпадающих плоскостей.*

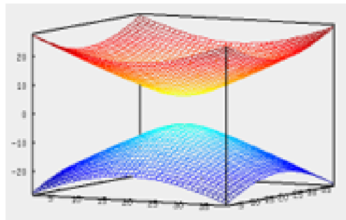
(1)



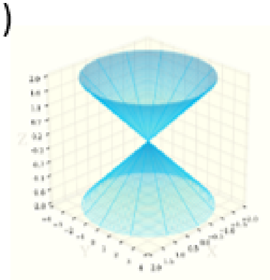
(3)



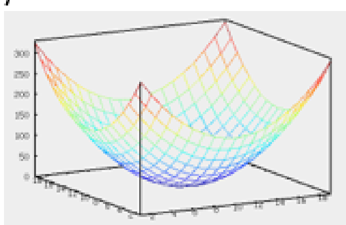
(4)



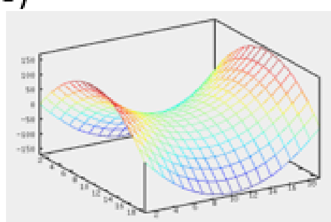
(5)



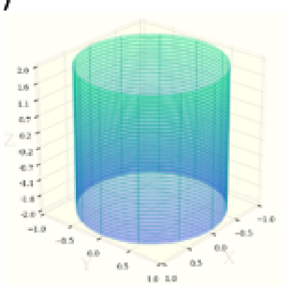
(7)



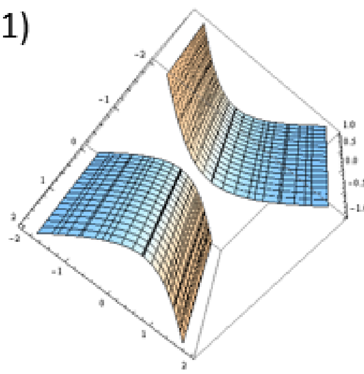
(8)



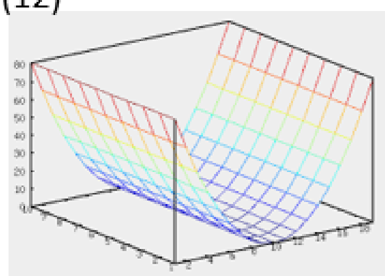
(9)



(11)



(12)



5.7 Цилиндрические и конические поверхности. Линейчатые поверхности. Поверхности вращения.

• ***Цилиндрической поверхностью*** *называется объединение всех параллельных прямых, которые пересекают некоторую линию L, при этом линия L называется* ***направляющей*** *цилиндрической поверхности, а прямые, из которых состоит поверхность, называются* ***образующими****.*

Пусть линия L задается как линия пересечения двух плоскостей, тогда она определена системой двух уравнений

$$\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0. \end{cases}$$

И пусть образующие цилиндрической поверхности имеют направляющий вектор $a(a_1, a_2, a_3)$. Точка $M(x, y, z)$ принадлежит цилиндрической поверхности $\Longleftrightarrow$ существует точка $M_0(x_0, y_0, z_0)$ такая, что

1. $M_0 \in L$;
2. $\overline{M_0M} \parallel a$.

Тогда получаем систему

$$\begin{cases} F_1(x_0, y_0, z_0) = 0, \\ F_2(x_0, y_0, z_0) = 0, \\ \dfrac{x - x_0}{a_1} = \dfrac{y - y_0}{a_2} = \dfrac{z - z_0}{a_3}. \end{cases}$$

Исключив из этой системы x_0, y_0 и z_0, получим **уравнение цилиндрической поверхности**.

• ***Канонической поверхностью*** *называется объединение всех прямых, которые проходят через некоторую точку P и пересекают некоторую линию L. При этом линия L называется* ***направляющей*** *канонической поверхности, а точка P —* ***вершиной*** *канонической поверхности.*

Пусть точка P имеет координаты (x_1, y_1, z_1). Точка $M(x, y, z)$ принадлежит канонической поверхности $\Longleftrightarrow$ существует точка $M_0(x_0, y_0, z_0)$ такая, что

1. $M_0 \in L$;
2. $\overline{PM} \parallel \overline{PM_0}$.

Тогда получаем систему

$$\begin{cases} F_1(x_0, y_0, z_0) = 0, \\ F_2(x_0, y_0, z_0) = 0, \\ \dfrac{x - x_1}{x_0 - x_1} = \dfrac{y - y_1}{y_0 - y_1} = \dfrac{z - z_1}{z_0 - z_1}. \end{cases}$$

Исключив из этой системы x_0, y_0 и z_0, получим **уравнение канонической повехности**.

• *Поверхность называется* ***линейчатой****, если она образована движением прямой. При этом прямые, из которых состоит линейчатая поверхность, называются прямолинейными образующими.*

Цилиндрическая и каноническая поверхности являются линейчатыми поверхностями.

Среди поверхностей второго порядка прямолинейными образующими обладают однополостный гиперболоид и гиперболический параболоид. Докажем это. Для этого рассмотрим однополостный гиперболоид. Тогда выполняется $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1$. Перенесем $\dfrac{y^2}{b^2}$ в правую часть и распишем разность квадратов: $\left(\dfrac{x}{a}-\dfrac{z}{c}\right)\left(\dfrac{x}{a}+\dfrac{z}{c}\right)=\left(1-\dfrac{y}{b}\right)\left(1+\dfrac{y}{b}\right)$. Составим систему. При фиксированных k эта система определяет прямую. Причем если точка M лежит на этой прямой, то она лежит и на однополостном гиперболоиде. Следовательно, все прямые целиком лежат на однополостном гиперболоиде.

$$\begin{cases} \dfrac{x}{a}-\dfrac{z}{c}=\left(1-\dfrac{y}{b}\right)\cdot k, \\ \dfrac{x}{a}+\dfrac{z}{c}=\left(1+\dfrac{y}{b}\right)\cdot \dfrac{1}{k}; \end{cases}, k\in\mathbb{R}. \tag{5.7.1}$$

Покажем обратное. Пусть $M_0(x_0,y_0,z_0)$ — произвольная точка однополостного гиперболоида. Тогда $\dfrac{x_0^2}{a^2}+\dfrac{y_0^2}{b^2}-\dfrac{z_0^2}{c^2}=1$. Найдем k_0 из уравнений:

$$\begin{cases} \dfrac{x_0}{a}-\dfrac{z_0}{c}=\left(1-\dfrac{y_0}{b}\right)\cdot k, \\ \dfrac{x_0}{a}+\dfrac{z_0}{c}=\left(1+\dfrac{y_0}{b}\right)\cdot \dfrac{1}{k}. \end{cases}$$

Следовательно, точка M_0 лежит на прямой (5.7.1). Аналогично монжо показать, что однополостный гиперболоид покрывается еще одним семейством прямых:

$$\begin{cases} \dfrac{x}{a}-\dfrac{z}{c}=\left(1+\dfrac{y}{b}\right)\cdot k, \\ \dfrac{x}{a}+\dfrac{z}{c}=\left(1-\dfrac{y}{b}\right)\cdot \dfrac{1}{k}. \end{cases}$$

Аналогично можно показать, что гиперболический параболоид также покрыт двумя семействами прямолинейных образующих. Для него выполняется $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=2z$. Распишем разность квадратов и получим $\left(\dfrac{x}{a}-\dfrac{y}{b}\right)\left(\dfrac{x}{a}+\dfrac{y}{b}\right)=2z$. Тогда

$$\begin{cases} \dfrac{x}{a}-\dfrac{y}{b}=2kz, \\ \dfrac{x}{a}+\dfrac{y}{b}=\dfrac{1}{k}; \end{cases} \text{— первое семейство;} \quad \begin{cases} \dfrac{x}{a}-\dfrac{y}{b}=\dfrac{1}{k}, \\ \dfrac{x}{a}+\dfrac{y}{b}=2kz; \end{cases} \text{— второе семейство;}$$

• ***Поверхностью вращения*** *линии L вокруг прямой Δ называется объединение окружностей, которые*

1. *пересекают линию L;*
2. *имеют центры на прямой Δ;*
3. *лежат в плоскости, перпендикулярной Δ.*

Рассмотрим поверхность вращения линии L такой, что $\begin{cases} F_1(x,y,z)=0, \\ F_2(x,y,z)=0; \end{cases}$ вокруг прямой Δ такой, что $\dfrac{x-x_1}{a_1}=\dfrac{y-y_1}{a_2}=\dfrac{z-z_1}{a_3}$. Точка M принадлежит поверхности вращения $\Longleftrightarrow$ существует точка $M_0(x_0,y_0,z_0)$ такая, что

1. $M_0 \in L$;
2. $\overrightarrow{M_0M} \perp a(a_1,a_2,a_3)$;
3. $\rho(M_0,\Delta)=\rho(M,\Delta)$.

Тогда получаем систему

$$\begin{cases} F_1(x_0,y_0,z_0)=0, \\ F_2(x_0,y_0,z_0)=0, \\ a_1\cdot(x-x_0)+a_2\cdot(y-y_0)+a_3\cdot(z-z_0)=0, \\ \rho(M_0,\Delta)=\rho(M,\Delta). \end{cases}$$

Исключив из этой системы x_0, y_0 и z_0, получим **уравнение поверхности вращения**.

Часть II

Основы высшей алгебры

Глава 6

Комплексные числа

6.1 Понятие комплексного числа. Арифметические операции.

• ***Комплексным числом*** *называется выражение вида* $z = a + bi$, *где* $a, b \in \mathbb{R}$ *(действительные числа), а* i *— символ, называемый мнимой единицей. При этом число* a *называется* ***действительной частью*** z *(Обозначение:* $a = Re(z)$*), а число* b *—* ***мнимой частью*** z *(Обозначение:* $b = Im(z)$*).*

Множество комплексных чисел обозначается символом $\mathbb{C}$.

Если $b = 0, a \neq 0$, то число $a+0i$ считается **совпадающим с действительными числом** a, то есть $a + 0i = a$.
Если $b \neq 0, a \neq 0$, то число называется **мнимым**. Если при $a = 0$, то число $0 + bi$ называется **чисто мнимым** и обозначается через bi, то есть $(0 + bi = bi)$.
Пусть $z_1 = a_1+b_1i$, $z_2 = a_2+b_2i$ — два комплексных числа. Комплексные числа z_1 и z_2 **равны**, если равны их действительные и мнимые части, то есть $z_1 = z_2 \Longleftrightarrow a_1 = a_2,\ b_1 = b_2$.

• ***Суммой комлексных чисел*** z_1 *и* z_2 *называется число:*

$$z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i$$

Свойства сложения комплексных чисел:

1. $z_1 + z_2 = z_2 + z_1$

   ♦ $z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i = (a_2 + a_1) + (b_2 + b_1)i = z_2 + z_1$ ⊠

2. $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$

Операция сложения пораждает операцию вычитания:

• ***Разностью комплексных чисел*** z_1 *и* z_2 *называется число:*

$$z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)i$$

• ***Произведением комплексных чисел*** z_1 *и* z_2 *называется число:*

$$z_1 \cdot z_2 = (a_1 \cdot a_2 - b_1 \cdot b_2) + (a_1 \cdot b_2 + b_1 \cdot a_2)i$$

Свойства произведения комплексных чисел:

1. $z_1 \cdot z_2 = z_2 \cdot z_1$
2. $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$
3. $(z_1 + z_2) \cdot z_3 = z_1 \cdot z_3 + z_2 \cdot z_3$

Из определения произведения комплексных чисел следует, что $i^2 = (0 + 1i)(0 + 1i) = -1 + 0i = -1$. То есть во множестве $\mathbb{C}$ уравнение $x^2 + 1 = 0$ имеет решение $x = i$.

Если числа z_1 и z_2 действительные, то операции сложения и умножения этих чисел совпадают с операциями сложения и умножения действительных чисел.
Пусть $z = a + bi$ — комплексное число.

• *Комплексное число* $\overline{z} = a - bi$ *называется* ***сопряжённым для комплексного числа*** z.

Из определения следует, что $\overline{z} = z$. Следовательно, можно говорить о паре сопряжённых друг другу комплексных чисел.

Операция умножения комплексных чисел пораждает операцию деления:

• ***Частным от деления комплексных чисел*** z_1 *и* z_2 *называется число* z_3 *такое, что* $z_2 \cdot z_3 = z_1$.

Тогда $\dfrac{z_1}{z_2} = \dfrac{a_1 + b_1 i}{a_2 + b_2 i} = [\text{домножим на сопряженное}] = \dfrac{(a_1 + b_1 i)(a_2 - b_2 i)}{(a_2 + b_2 i)(a_2 - b_2 i)} = \dfrac{a_1 \cdot a_2 + b_1 \cdot b_2}{a_2^2 + b_2^2} + \dfrac{b_1 \cdot a_2 - a_1 \cdot b_2}{a_2^2 + b_2^2} i$.

Свойства сопряжённых комплексных чисел:

1. $z + \overline{z} = 2 \cdot Re(z)$.
   $z - \overline{z} = 2 \cdot Im(z) \cdot i$.

   ♦ Если $z = a + bi$, то $z + \overline{z} = (a + bi) + (a - bi) = (a + a) + (b - b)i = 2a$. ⊠

2. $z = \overline{z} \Longleftrightarrow z \in \mathbb{R}$.

   ♦ $\Rightarrow$) Пусть $z = \overline{z}$, тогда $z + \overline{z} = 0 = 2 \cdot Im(z) \cdot i \Rightarrow Im(z) = 0 \Rightarrow z \in \mathbb{R}$.
   $\Leftarrow$) Пусть $z \in \mathbb{R}$. Следовательно, $z = a + 0i \Rightarrow \overline{z} = a - 0i = a = z$. ⊠

3. $\overline{z_1} + \overline{z_2} = \overline{z_1 + z_2}$; $\overline{z_1} - \overline{z_2} = \overline{z_1 - z_2}$.

   ♦ Пусть $z_1 = (a_1 + b_1 i)$, $z_2 = a_2 + b_2 i \Rightarrow z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i$. Тогда $\overline{z_1} + \overline{z_2} = (a_1 - b_1 i) + (a_2 - b_2 i) = (a_1 + a_2) - (b_1 + b_2)i = \overline{z_1 + z_2}$. ⊠

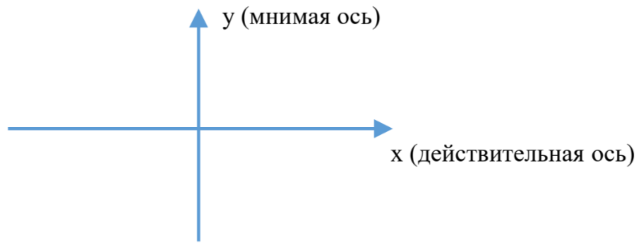
4. $\overline{z_1} \cdot \overline{z_2} = \overline{z_1 \cdot z_2}$.

   $\dfrac{\overline{z_1}}{\overline{z_2}} = \overline{\left(\dfrac{z_1}{z_2}\right)}$.

6.2 Комплексная плоскость. Тригонометрическая и экспоненциальная форма записи комплексного числа.

Пусть на плоскости задана ДПСК Oxy. Каждому комплексному числу вида $z = a + bi$ поставим в соответствие координаты (a, b). С другой стороны, каждой точке с координатами (a, b) мы можем поставить в соответствие комплексное число $z = a + bi$.

• *Плоскость, точки которой рассматриваются как комплексные числа, называется* ***комплексной плоскостью****. При этом ось Ох называется* ***действительной осью****, а ось Oy —* ***мнимой осью****, так как на ней располагаются чисто мнимые числа.*



Построим на комплексной плоскости соответствующую полярную систему координат. И пусть $z = a + bi$ — некоторое комплексное число. Тогда этому числу на комплексной плоскости соответствует точка, имеющая декартовы координаты (a, b) и полярные координаты (ρ, φ).

Следовательно, $\begin{cases} a = \rho \cdot cos\varphi, \\ b = \rho \cdot sin\varphi; \end{cases}$, где $\rho = \sqrt{a^2 + b^2}$, а угол φ можно найти, решив систему:

$$\begin{cases} cos\varphi = \dfrac{a}{\sqrt{a^2 + b^2}}, \\ \\ sin\varphi = \dfrac{b}{\sqrt{a^2 + b^2}}; \end{cases}$$

Тогда комплексное число z представимо в виде: $z = a + bi = \rho \cdot cos\varphi + i \cdot \rho \cdot sin\varphi = \rho(cos\varphi + i \cdot sin\varphi)$.

• *Форма вида $z = \rho(cos\varphi + i \cdot sin\varphi)$ называется* ***тригонометрической формой записи комплексного числа****. В свою очередь, форма вида $z = a + bi$ называется* ***алгебраической формой записи комплексного числа.***

• *Полярный радиус ρ называется* ***модулем комплексного числа****. А полярный угол φ называтся* ***аргументом комплексного числа*** *(Обозначения: $|z|$ и $Arg(z)$ соответственно).*

Аргумент комплексного числа определён с точностью до слагаемого $2\pi k$. Значение аргумента, принадлежащее полуотрезку $(\pi; \pi]$, называют **главным** значением аргумента.

Свойства значения аргумента:

1. $|z| = |\overline{z}|, Arg(z) = -Arg(\overline{z})$.

   ⧫ Доказательство следует оз определения числа z. ⊠

2. $|z|^2 = z \cdot \overline{z}$.

♦ Пусть $z = a + bi$, тогда $\overline{z} = a - bi$.
$z \cdot \overline{z} = (a + bi)(a - bi) = a^2 - (bi)^2 = a^2 + b^2 = |z|^2$. ⊠

3. *Модуль разности комплексных чисел равен расстоянию между точками на комплексной плоскости, соответствующих этим числам, то есть*

$$|z_1 - z_2| = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2}.$$

♦ Пусть $\begin{cases} z_1 = a_1 + b_1 i, \\ z_2 = a_2 + b_2 i; \end{cases} \Rightarrow z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)i \Rightarrow |z_1 - z_2| = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2} = \rho(z_1, z_2)$. ⊠

4. $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$.
$Arg(z_1 \cdot z_2) = Arg(z_1) + Arg(z_2)$.

♦ Запишем числа z_1 и z_2 в тригонометрической форме: $z_1 = \rho_1(cos\varphi_1 + i \cdot sin\varphi_1)$, $z_2 = \rho_2(cos\varphi_2 + i \cdot sin\varphi_2)$. Отсюда получаем, что $|z_1| = \rho_1$, $|z_2| = \rho_2$, $Arg(z_1) = \varphi_1$, $Arg(z_2) = \varphi_2$.
Тогда произведение чисел z_1 и z_2 в тригонометрической форме имеет вид: $z_1 \cdot z_2 = \rho_1\rho_2((cos\varphi_1 cos\varphi_2 - sin\varphi_1 sin\varphi_2) + i(cos\varphi_1 sin\varphi_2 + sin\varphi_1 cos\varphi_2)) = \rho_1\rho_2(cos(\varphi_1 + \varphi_2) + i \cdot sin(\varphi_1 + \varphi_2))$.
Из полученного следует, что $|z_1 \cdot z_2| = \rho_1\rho_2 = |z_1| \cdot |z_2|$, а $Arg(z_1 \cdot z_2) = \varphi_1 + \varphi_2 = Arg(z_1) + Arg(z_2)$. ⊠

5. $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}, Arg(\dfrac{z_1}{z_2}) = Arg(z_1) - Arg(z_2)$.

♦ Доказательство проводится по аналогии с предыдущим пунктом. ⊠

Следствие. $z^{-1} = \dfrac{1}{z} = \dfrac{1(cos\varphi + i \cdot sin\varphi)}{\rho(cos\varphi + i \cdot sin\varphi)} = \dfrac{1}{\rho}(cos(-\varphi) + i \cdot sin(\varphi))$.

6. **Формула Маувра**
Если $z = \rho(cos\varphi + i \cdot sin\varphi)$, *то* $\forall\ n \in \mathbb{Z}\ z^n = \rho^n(cos(n\varphi) + i \cdot sin(n\varphi))$.

♦ Доказательство следует по индукции из пункта 4 и следствия. ⊠

• $e^{i\varphi} = cos\varphi + i \cdot sin\varphi$ — ***формула Эйлера***.
• $z = \rho e^{i\varphi}$ — ***экспоненциальная форма записи комплексного числа***.

6.3 Извлечение корня из комплексного числа.

Пусть z — некоторое комплексное число.

• ***Корнем n-ой степени из комплексного числа*** *z называется число z_0 такое, что z_0 в степени n равно Самому комплексному числу z. То есть $z_0^n = z$.*

Теорема. *Извлечение корня n-ой степени из комплексного числа $z = \rho(cos\varphi + i \cdot sin\varphi)$ всегда возможно и, при $z \neq 0$, даёт ровно n различных значений:*

$$z_k = \sqrt[n]{\rho}\left(cos\left(\frac{\varphi + 2\pi k}{n}\right) + i \cdot sin\left(\frac{\varphi + 2\pi k}{n}\right)\right),\ k = 0, 1, \ldots, n - 1,$$

где $\sqrt[n]{\rho}$ — действительное положительное число, n-ая степень которого равна ρ.

♦ Пусть существует корень n-ой степени из числа z и он равен $z_0 = \rho_0(cos\varphi_0 + i \cdot sin\varphi_0)$. Тогда по формуле Маувра $z_0^n = \rho_0^n(cos(n\varphi_0) + i \cdot sin(n\varphi_0)) = \rho(cos(\varphi) + i \cdot sin(\varphi))$.

Если комплексные числа равны, то их модули равны, а аргументы могут отличаться на $2\pi k$, где $k \in \mathbb{Z}$. Следовательно, $\rho_0^n = \rho$, $n\varphi_0 = \varphi + 2\pi k \Rightarrow \varphi_0 = \dfrac{\varphi + 2\pi k}{n}, k \in \mathbb{Z} \Rightarrow z_k$ является корнем n-ой степени числа z.

Покажем, что при $k \in [0, n-1]$ числа z_k различные:

Пусть существуют 2 числа p и q из отрезка $[0, n-1]$ такие, что $z_p = z_q$, и пусть $0 \leqslant p < q < n \Rightarrow \exists l \in \mathbb{N} : \dfrac{\varphi + 2\pi q}{n} = \dfrac{\varphi + 2\pi p}{n} + 2\pi l \Rightarrow 2\pi l = \dfrac{2\pi(q-p)}{n}$. А так как $q < n$ и $p < n$, то $\dfrac{q-p}{n} < 1$. Следовательно, $l < 1$, что является противоречием с тем, что $l \in \mathbb{N}$.

Покажем, что при $k \notin [0, n-1]$ новых значений z_k мы не получим:

Разделим k на n с остатком. Пусть p и q — частное и остаток при делении соответственно. Тогда $z = p \cdot n + q, 0 \leqslant q < n$.
Тогда $Arg(z_k) = \varphi_k = \dfrac{\varphi + 2\pi k}{n} = \dfrac{\varphi + 2\pi(pn + q)}{n} = \dfrac{\varphi + 2\pi q}{n} + 2\pi p \Rightarrow z_k = z_q^n \Rightarrow$ новых корней мы не получим. ⊠

Глава 7

Алгебраические структуры

7.1 Бинарные отношения.

Пусть X, Y — два непустых множества.

• ***Декартовым произведением*** *множества X на множество Y называется множество всех упорядоченных пар (x, y), где $x \in X$ и $y \in Y$. Обозначается $X \times Y$. То есть*

$$X \times Y = \{(x, y) \mid x \in X, \, y \in Y\}.$$

• *Декартово произведение множества X на себя называется **декартовым квадратом** множества X. То есть*

$$X^2 = X \times X.$$

• ***Бинарным отношением,*** *заданным на множестве X, называется подмножество декартового квадрата множества X. Обозначается символом σ (сигма).*

Если σ — бинарное отношение на множестве X ($\sigma \subseteq X^2$) и пара $(x, y) \in \sigma$, то говорят, что элемент x находится в отношении σ с элементом y. Записывается $x \; \sigma \; y$.

• *Бинарное отношение σ, заданное на множестве X, называется:*

1. ***рефлексивным***, *если $\forall x \in X \, : x\sigma x$*
2. ***транзитивным***, *если из соотношений $x\sigma y$, $y\sigma z \Rightarrow x\sigma z$*
3. ***симметричным***, *если из $x\sigma y \Rightarrow y\sigma x$*
4. ***антисимметричным***, *если из $x\sigma y$, $y\sigma x \Rightarrow x = y$*

• ***Бинарным отношением эквивалентности*** *называется рефлексивное, транзитивное, симметричное бинарное отношение. Обозначается $\sim$.*

Пусть на множестве X задано бинарное отношение эквивалентности и пусть x — некоторый элемент множества X.

• *Подмножество множества X, состоящее из всех элементов, эквивалентных элементу x, называется **классом эквивалентности**, содержащим этот x. Обозначается $\overline{x}$. То есть*

$$\overline{x} = \{x' \in X \mid x' \sim x\}.$$

• *Любой элемент множества $\overline{x}$ называется* ***представителем*** *класса $\overline{x}$.*

• *Если множество X представимо в виде объединения попарно непересекающихся подмножеств, то говорят, что множество X* ***разбивается*** *на эти подмножества.*

Теорема. *Пусть X — некоторое множество, на котором задано отношение эквивалентности, тогда оно разбивается на классы эквивалентности по этому бинарному отношению.*

♦ Множество X всегда можно представить в виде $X = \bigcup\limits_{x \in X} \underbrace{\{x\}}_{\subset\ \overline{x}} \subseteq \bigcup\limits_{x \in X} \overline{x} \subset X \Rightarrow X = \bigcup\limits_{x \in X} \overline{x}$

Покажем, что два 2 класса эквивалентности $\overline{x_1}$ и $\overline{x_2}$ или совпадают, или не переекаются.

Пусть $\overline{x_1} \cap \overline{x_2} \neq \varnothing \Rightarrow \exists\ x \in \overline{x_1} \cap \overline{x_2} \Rightarrow x \in \overline{x_1},\ x \in \overline{x_2} \Rightarrow x \sim x_1,\ x \sim x_2 \overset{\text{симм.}}{\Rightarrow} x_1 \sim x,\ x \sim x_2 \overset{\text{транз.}}{\Rightarrow} x_1 \sim x_2 \Rightarrow \forall x \in \overline{x_1},\ x \sim x_1,\ x_1 \sim x_2 \overset{\text{транз.}}{\Rightarrow} x \sim x_2 \Rightarrow x \in \overline{x_2} \Rightarrow \overline{x_1} \subseteq \overline{x_2}$.

С другой стороны, так как бинарное отношение эквивалентности симметрично, то из этого следует, что $x_1 \sim x_2 \Rightarrow x_2 \sim x_1 \Rightarrow \overline{x_2} \subseteq \overline{x_1} \Rightarrow \overline{x_1} = \overline{x_2}$. ⊠

• *Бинарное отношение, заданное на множестве X, называется* ***отношением порядка****, если оно рефлексивно, транзитивно и антисимметрично. Обозначается $\leq$.*

Если $\forall x, y \in X$ или $x \leq y$, или $y \leq x$, то множество X называется **линейно упорядоченным**. Но пара (x, y) может и не находиться в бинарном отношении порядка. Если такая пара существует, то X — **частично упорядоченное множество.**

Пусть на множестве X задано бинарное отношение порядка.

• ***Наибольшим элементом*** *множества X называется такой элемент $n \in X$, что $\forall x \in X : x \leq n$.*

• ***Максимальным элементом*** *множества X называется такой элемент $m \in X$, что $\nexists\ x \in X,\ x \neq m : m \leq x$.*

Замечание: Максимальных элементов во множестве может быть несколько, а наибольший элемент, если он существует, только один.

7.2 Отображения.

Пусть X, Y — некоторые непустые множества.

• *Говорят, что определено однозначное* ***отображение*** *множества X на множество Y, если каждому элементу из множества X поставим в соответствие только один элемент из множества Y. Обозначение: $f\colon X \to Y$.*

Заметим, что $f \subset X \times Y$. А также $\forall x \in X\ \exists! y \in Y : (x, y) \in f$.

• *Если отображение f элемент x ставит в соответствие элементу y, то говорят, что y —* ***образ*** *элемента x, а x —* ***прообраз*** *элемента y.*
Обозначение образа: $\{f(x) \mid \forall x \in X\} = ImX = f(x)$.
Обозначение прообраза: $f(x) = y$ или $f\colon x \mapsto y$.

Из определения отображения следует, что $f(x) \subseteq Y$.

• *Отображение $f : X \to Y$ называется **инъективным**, если $\forall x_1, x_2 \in X$, $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.*

• *Отображение $f : X \to Y$ называется **сюръективным**, если $\forall y \in Y \ \exists x \in X : f(x) = y$.*

• *Отображение $f : X \to Y$ называется **биективным**, если оно инъективно и сюръективно (взаимнооднозначное соответствие).*

• *Два отображения $f\colon X \to Y$ и $g\colon X' \to Y'$ называются **равными**, если $X = X'$, $Y = Y'$ и $\forall x \in X \ f(x) = g(x)$.*

Пусть $f\colon X \to Y$ — некоторое отображение. $X' \subseteq X$. Тогда

• *Отображение $g\colon X' \to Y'$ называется **ограничением отображения** f на множестве X', если $\forall x \in X' : g(x) = f(x)$. Обозначение: $f|_{X'}$. Само же отображение $f : X \to Y$ называется **продолжением отображения** g на множестве X.*

Пусть $f\colon X \to Y$ и $g\colon Y \to Z$. Тогда

• *Отображение $g \circ f : X \to Z$ которое работает так, что $\forall x \in X \ (g \circ f)(x) = g(f(x))$, называется **композицией отображений** g и f.*

Из определения следует, что композиция g и f отображена.

Теорема. *Пусть $f\colon X \to Y$, $g\colon Y \to Z$ и $h = g \circ f$. Тогда*

1. *если f и g инъективны, то h инъективно;*
2. *если f и g сюръективны, то h сюръективно;*
3. *если f и g биективны, то h биективно.*

♦

1. Пусть f и g — инъекции, тогда $\forall x_1, x_2 \in X$, $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$, $g(f(x_1)) \neq g(f(x_2)) \Rightarrow h$ инъективно, так как $(g \circ f)(x_1) = g(f(x_1)) = h(x_1)$, $(g \circ f)(x_2) = g(f(x_2)) = h(x_2)$, $h(x_1) \neq h(x_2)$.
2. Пусть f и g — сюръекции и пусть z_0 — произвольный элемент множества Z. Так как g — сюръекция, то для z_0 $\exists y_0 \in Y$ такой, что $g(y_0) = z_0$. А так как f — сюръекция, то для этого элемента y_0 $\exists x_0 \in X$ такой, что $f(x_0) = y_0 \Rightarrow \forall z_0 \in Z \ \exists x_0 : h(x_0) = (g \circ f)(x_0) = g(f(x_0)) = g(y_0) = z_0 \Rightarrow h$ — сюръекция.
3. Биективность вытекает из двух предыдущих пунктов.

⊠

• *Отображение $e_x : X \to X$ такое, что $\forall x \in X \ e_x(x) = x$ называется **тождественным отображением** множества X.*

Пусть $f : X \to Y$ — некоторое отображение.

• *Отображение $g : Y \to X$ называется **обратным** для отображения f, если $g \circ f = e_x$, $f \circ g = e_y$.*

Теорема. *Отображение $f : X \to Y$ имеет обратное $\Longleftrightarrow$ оно биективно.*

⧫ $\Rightarrow$) Пусть отображение f: $X \to Y$ имеет обратное отображение $g : Y \to X$. Покажем, что f инъективно и сюръективно.

1. Инъективность.
   Возьмем произвольные $x_1, x_2 \in X$ такие, что $x_1 \neq x_2$. Покажем, что $f(x_1) \neq f(x_2)$.
   От противного. Предположим, что $f(x_1) = f(x_2)$.
   Возьмем элемент x_1, $x_1 = e_x(x_1) = (g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2) = e(x_2) = x_2$, что противоречит выбранным x_1 и $x_2 \Rightarrow f(x_1) \neq f(x_2) \Rightarrow f$ — инъекция.

2. Сюръективность.
   $\forall y_0 \in Y$ $y_0 = e_y(y_0) = (f \circ g)(y_0) = f(g(y_0))$. Тогда $\exists x_0 \in X : x_0 = g(y_0)$, $f(x_0) = y_0$ и f — сюръекция

Из этих двух пунктов получаем, что f — биекция.

$\Leftarrow$) Пусть отображение f: $X \to Y$ биективное. Тогда $\forall y_0 \in Y$ $\exists!$ $x_0 \in X$ [существование следует из сюръекции, а единственность из инъекции]: $f(x_0) = y_0 \Rightarrow$ существует отображение g: $Y \Rightarrow X$ такое, что $g(y_0) = x_0$, то есть обратное для f.
Тогда $\forall y_0 \in Y$ $(f \circ g)(y_0) = f(g(y_0)) = f(x_0) = y_0 = e_y(y_0) \Rightarrow e_y = f \circ g$.
$\forall x_0 \in X$ $(g \circ f)(x_0) = g(f(x_0)) = g(y_0) = x_0 = e_x(x_0) \Rightarrow e_x = g \circ f$. ⊠

Теорема. *Пусть f: $X \to Y$, g: $Y \to Z$, h: $Z \to U$, тогда $h \circ (g \circ f) = (h \circ g) \circ f$*

⧫ $g \circ f : X \to Z$, тогда $h \circ (g \circ f) : X \to U$.
$h \circ g : Y \to U$, тогда $h \circ (g \circ f) : X \to U$.
Следовательно, множества совпадают. Значит справедливы равенства $\forall x \in X$
$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x)))$,
$((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x)))$. ⊠

7.3 Бинарная алгебраическая операция.

Пусть X — некоторое непустое множество.

• *Отображение $f : X^2 \to X$, ставящее каждой упорядоченной паре $(a, b) \in X^2$ в соответствие элемент $c \in X$ называется **бинарной алгебраической операцией.** Обозначается $f : (a, b) \mapsto c$ или $f(a, b) = c$.*

Для произвольной алгебраической операции $a * b = c$, где $*$ — символ символ (они могут выглядеть по-разному: $+, -, *$ и так далее).

• *Алгебраическая операция $*$ называется*

1. ***ассоциативной,** если $(a * b) * c = a * (b * c) \quad \forall a, b, c \in X$;*

2. ***коммутативной,** если $a * b = b * a \quad \forall a, b \in X$.*

• *Элемент $n \in X$ называется* ***нейтральным*** *элементом в X относительно операции $*$, если $n * x = x * n = x \quad \forall x \in X$.*

Теорема. *В множестве X относительно алгебраической операции $*$ может существовать не более одного нейтрального элемента.*

⧫ От противного. Пусть в X относительно операции $*$ существуют два нейтральных элемента n_1 и n_2. Тогда $n_1 * n_2 = n_1$, а $n_2 * n_1 = n_2 \Rightarrow n_1 = n_2$, так как алгебраическая операция — однозначное отображение. ⊠

Пусть во множестве X определена операция $*$, относительно которой в X существует нейтральный элемент.

• *Элемент $a' \in X$ называется* ***симметричным*** *для элемента $a \in X$ относительно операции $*$, если $a' * a = a * a' = n$.*

Теорема. *Если на множестве X алгебраическая операция $*$ ассоциативна и существует нейтралный элемент, то для любого элемента $a \in X$ может существовать не более одного симметричного элемента.*

⧫ От противного. Пусть b_1, b_2 — два элемента, симметричных элементу a относительно операции $*$. Тогда $b_1 * a = a * b_1 = n$, $b_2 * a = a * b_2 = n \Rightarrow b_1 = b_1 * n = b_1 * (a * b_2) = (b_1 * a) * b_2 = n * b_2 = b_2 \Rightarrow b_1 = b_2 \Rightarrow$ симметричный элемент, если он есть, единственный. ⊠

Рассмотрим некоторые операции и их характеристику:

Операция	Аддитивная запись	Мультипликативная запись
Знак	$+$	$\cdot$
Название	сложение	умножение
Результат	сумма	произведение
Нейтральный элемент	0	1
Симметричный элемент	$-a$ (противоположный)	a^{-1} (обратный)

• *Множество X с определенной на нем алгебраической операцией $*$ является* ***алгебраической структурой****. Обозначается $(X, *)$.*

Замечание: Современная алгебра занимается изучением алгебраических структур.

7.4 Группа.

• *Непустое множество G с заданной на нем алгебраической операцией $*$ называется* ***группой***, если

1. *операция $*$ ассоциативна*;
2. *в G $\exists$ нейтральный элемент относительно операции $*$*;
3. *в G $\forall a$ $\exists$ симметричный ему элемент a' относительно операции $*$.*

Обозначение: $(G, *)$.

• *Если при этом операция $*$ является коммутативной, то группа называется* ***абелевой.***

• *Если группа содержит конечное количество элементов, то группа называется **конечной.***

• *Количество элементов в конечной группе называется **порядком** группы.*

Свойства групп:

1. *В группе существует единственный нейтральный элемент и для любого элемента группы в ней существует единственный симметричный*

   ⧫ Доказательство следует из теорем о единственности нейтрального и симметричного элементов. ⊠

2. *Пусть a, b — произвольные элементы группы $(G, *)$. a'— элемент, симметричный элементу a относительно $*$. Тогда уравнения*

$$a * x = b \qquad\qquad x * a = b$$

   имеют единственные решения соответственно

$$x_0 = a' * b \qquad\qquad x_0 = b * a'.$$

   ⧫ Доказательство проведем для первого уравнения, для второго доказательство аналогично. Покажем, что $x_0 = a' * b$ — решение уравнения $a * x = b$. Для проверки подставим: $a * x_0 = a * (a' * b) = (a * a') * b = n * b = b$. Следовательно, $a' * b$ — решение. Покажем, что это решение единственное.
   От противного. Предположим, что уравнение $a * x = b$ имеет еще одно некоторое решение x_1. Тогда $a * x_1 = b \Rightarrow x_0 = a' * b = a' * (a * x_1) = (a' * a) * x_1 = n * x_1 = x_1$, $x_0 = x_1 \Rightarrow$ решение единственное. ⊠

3. **Закон сокращения в группе:**

   - если $a * b = a * c$, то $b = c$;
   - если $b * a = c * a$, то $b = c$.

   ⧫ Пусть a' — элемент, симметричный элементу a относительно операции $*$. Тогда, $b = n * b = (a' * a) * b = a' * (a * b) = [a * b = a * c] = a' * (a * c) = (a' * a) * c = n * c = c$. ⊠

Пусть $(G, *)$ — некоторая группа.

• *Непустое подмножество множества G называется **подгруппой** группы G, если она сама является группой относительно той же операции $*$.*

Лемма. *Пусть $(H, *)$ — подгруппа группы $(G, *)$. Тогда нейтральный элемент n_G совпадает с нейтральным элементом n_H.*

⧫ Так как n_H — нейтральный элемент в группе H, то $n_H * n_H = n_H$. Так как $n_H \in H \subseteq G$, то в G существует единственный симметричный ему элемент n'_H. То есть $n_H * n'_H = n_G$. Тогда $n_G = n_H * n'_H = (n_H * n_H) * n'_H = n_H * (n_H * n'_H) = n_H * n_G = n_H$. ⊠

Теорема. *Пусть $(G, *)$ — группа. Непустое подмножество H множества G является подгруппой группы $G \iff$ выполняются следующие условия:*

1. *операция $*$ определена в H, то есть $\forall a, b \in H \quad a * b \in H$.*

2. $\forall a \in H$ симметричный ему в G элемент $a'_G \in H$.

♦ $\Rightarrow$)

1. Пусть $(H, *)$ — подгруппа группы $(G, *) \Rightarrow (H, *)$ сама является группой относительно операции $*$ $\Rightarrow$ операция $*$ в H определена.

2. Покажем, что $\forall a \in H$, $a'_G \in H$.
   Так как $a \in H$ и $(H, *)$ — группа, то в H существует единственный симметричный для a элемент a'_H, то есть $a * a'_H = a'_H * a = n_H = [$ по лемме $] = n_G \Rightarrow$ элемент a'_H является симметричным для элемента a в группе G, но по первому свойству группы элемент имеет только единственный сииметричный ему элемент. Следовательно, $a'_G = a'_H$, причем $a'_H \in H$.

$\Leftarrow$) Пусть H — непустое подмножество множества G. $(H, *)$ — группа.

1. Операция $*$ в H определена по первому условию достаточности;

2. Операция $*$ в H ассоциативна, так как она ассоциативна в G;

3. Нейтральный элемент группы G (n_G) является нейтральным и во множестве H, так как $\forall a \in H$ $\exists a'$: $a * a' = n_G \in H$;

4. Элемент $a'_G \in H$ и является симметричным для a в H.

Следовательно, $(H, *)$ — группа, а так как $H \subseteq G$, то $(H, *)$ — подгруппа группы $(G, *)$. ⊠

7.5 Кольцо.

• *Непустое множество K с определенными на нем операциями сложения и умножения называется **кольцом**, если*

1. $(K, +)$ — *абелева группа*;

2. *операция умножения дистрибутивна относительно операции сложения, то есть* $\forall a, b, c \in K$
   $a \cdot (b + c) = a \cdot b + a \cdot c$;
   $(b + c) \cdot a = b \cdot a + c \cdot a$.

• *Абелева группа $(K, +)$ называется **аддитивной группой** кольца K.*

• *Если умножение коммутативно в K, то кольцо называется **коммутативным**. Если умножение ассоциативно — **ассоциативным**.*

Свойства колец:

1. $a \cdot 0 = 0 \cdot a = 0 \quad \forall a \in K$.

   ♦ $a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 = a \cdot 0 + 0 \Rightarrow a \cdot 0 = 0$ по закону сокращения в группе $(K, +)$. ⊠
   Обратное утверждение для данного свойства, вообще говоря, неверно.

• *Если существуют числа $a, b \in K$ такие, что $a \cdot b = 0$, при этом $a \neq 0$, $b \neq 0$, то эти числа называются **делителями нуля**.*

2. $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$.

   ♦ $(-a) \cdot b + a \cdot b = b((-a) + a) = 0 \cdot b = 0$. Следовательно, элемент $(-a) \cdot b$ является симметричным относительно сложения для элемента $a \cdot b$. ⊠

3. *Умножение дистрибутивно относительно вычитания, то есть* $a - b = a + (-b)$.

   ♦ $(a - b) \cdot c = (a + (-b)) \cdot c = a \cdot c + (-b) \cdot c = a \cdot c + (-b \cdot c) = a \cdot c - b \cdot c$. ⊠

4. *Кольцо может иметь не более одной единицы. Если кольцо состоит из более, чем одного элемента, то этот элемент не является нулем.*

   ♦ Если в кольце существует нейтральный элемент, то он единственный. Следовательно, если единица — нейтральный элемент для операции умножения, то эта единица единственная.

   Пусть кольцо K состоит из более, чем одного элемента, и имеет единицу равную нулю ($1 = 0$), тогда $\forall a \in K$ $a \cdot 1 = a \cdot 0 = 0$. Следовательно, $K = \{a\}$, то есть в кольце K будет существовать только 1 элемент, что является противоречием. ⊠

Пусть K — кольцо с единицей.

• *Если для $a \in K$ существует $a^{-1} \in K$, то a называется **обратимым элементом** кольца K, а элемент a^{-1} — **обратным элементом** для a.*

Теорема. *В ассоциативном кольце с единицей множество K^* всех обратимых элементов кольца K является группой относительно операции умножения $(K^*, \cdot)$.*

♦ Докажем все аксиомы определения.

1. Так как $1 \cdot 1 = 1$, то элемент 1 обратимый. Следовательно, множество K^* непустое.

2. Пусть $a, b \in K^*$. Тогда $\exists a^{-1}, b^{-1} \in K$ и $b^{-1}a^{-1} \in K$ (так как в кольце определена операция умножения).
   $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = [a^{-1}a = n = 1] = b^{-1} \cdot 1 \cdot b = b^{-1}b = 1$. Аналогично $(ab)(b^{-1}a^{-1}) = 1$. Следовательно, элемент $b^{-1}a^{-1}$ обратный для ab и $ab \in K^*$.

3. Умножение в K^* ассоциативно, так как оно ассоциативно в K.

4. 1 — нейтральный относительно умножения элемент.

5. $\forall a \in K^* \exists a^{-1} \in K : aa^{-1} = 1$. Следовательно, элемент a обратимый для a^{-1}, $a \in K^*$. Тогда в K^* $\forall a \quad \exists a^{-1} : aa^{-1} = 1$.

⊠

• *Группа обратимых элементов ассоциативного кольца с единицей $(K^*, \cdot)$ называется **мультипликативной группой** кольца K.*

• *Подмножество L кольца K называется **подкольцом** кольца K, если $(L, +, \cdot)$ само является кольцом относительно тех же операций, что и K.*

Теорема. *Непустое подмножество $L \subseteq K$, если K — кольцо, является подкольцом $\Longleftrightarrow$ выполняются следующие условия:*

1. *операции $+$ и $\cdot$ определены в L, то есть $\forall a, b \quad \in L$ $a + b \in L, a \cdot b \in L$;*

2. *$\forall a \in L \quad (-a) \in L$.*

♦ $\Rightarrow$) Пусть $(L, +, \cdot)$ — подкольцо кольца $(K, +, \cdot)$. Тогда $(L, +, \cdot)$ само является кольцом. Следовательно

1. операции определены;

2. $\forall a \in L \quad (-a) \in L$, так как $(L, +)$ — абелева группа.

$\Leftarrow$) Из пунктов 1 и 2 достаточности следует, что $(L, +)$ — подгруппа группы $(K, +)$, и она является абелевой группой. Дистрибутивность относительно умножения в L следует из дистрибутивности в K, так как $L \subseteq K$. Следовательно, так как все аксиомы кольца выполняются в $(L, +, \cdot)$, то оно является кольцом, а так как $L \subseteq K$, то $(L, +, \cdot)$ является подкольцом кольца $(K, +, \cdot)$. $\boxtimes$

7.6 Поле.

• *Множество P, содержащее не менее двух элементов, с заданными на нем операциями сложения и умножения называется **полем**, если*

1. $(P, +)$ — *абелева группа*;

2. $(P \backslash \{0\}, \cdot)$ — *абелева группа*;

3. *умножение дистрибутивно относительно сложения, то есть $\forall a, b, c \in P$*
   $(a + b) \cdot c = a \cdot c + b \cdot c$.

То есть поле — обязательно кольцо ассоциативное, коммутативное, с единицей, каждый ненулевой элемент которого обратим.

Свойства полей:

1. *В поле нет делителей нуля.*

   ♦ Пусть для некоторых $a, b \in P$ выполняется $a \cdot b = 0$. Предположим, что $a \neq 0$. Следовательно, $\exists a^{-1} \in P : 0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b) = (a^{-1} \cdot a) \cdot b = 1 \cdot b = b$, то есть $b = 0$. $\boxtimes$

2. **Закон сокращения:** *Если $a \neq 0$ и $a \cdot b = a \cdot c$, то $b = c \quad \forall b, c \in P$.*

3. *Если $a \neq 0$, то уравнение $a \cdot x = b$ имеет единственное решение $a^{-1} \cdot b \quad \forall b \in P$.*

   ♦ Если $b \neq 0$ и $c \neq 0$, то утверждения 2 и 3 справедливы по свойствам группы $(P \backslash \{0\}, \cdot)$.
   Если $b = 0$ и $c = 0$, то утверждения 2 и 3 справедливы, так как в поле нет делителей нуля. $\boxtimes$

• *Непустое подмножество F поля $(P, +, \cdot)$ называется **подполем** поля P, если F является полем относительно тех же операций, что и поле P. В свою очередь, поле P называется **расширением** поля F.*

Теорема. *Непустое подмножество F поля P, содержащее не менее двух элементов, является подполем поля P, если*

1. *в F определенв операции сложения и умножения, то есть $\forall a, b \in F \quad (a+b), \ (a \cdot b) \ \in \ F$;*

2. *$\forall a \in F \quad (-a) \in F$;*

3. $\forall a \in F, a \neq 0 \quad a^{-1} \in F.$

⧫ Доказательство следует из критерия подгрупп. Аналогично схожей теореме для подкольца. ⊠

• *Число* $n \in \mathbb{N}$*, которое обозначает минимальное количество единиц в поле* P*, дающих в сумме нуль (то есть* $\underbrace{1 + 1 + \ldots + 1}_{n} = 0$*), называется* ***характеристикой*** *поля.*

Если такого числа не существует, характеристику поля считают равной нулю.

Свойства характеристики поля:

1. *Все конечные поля имеют положительную характеристику. Бесконечные поля могут иметь как нулевую, так и положительную характеристику.*

   ⧫ Пусть поле положительное и конечное. Тогда $\exists k, m : \underbrace{1 + 1 + \ldots + 1}_{k} = \underbrace{1 + 1 + \ldots + 1}_{m}$. Пусть $m > k$, тогда $0 = \underbrace{1 + 1 + \ldots + 1}_{m-k}$. Следовательно, $\exists n : \underbrace{1 + 1 + \ldots + 1}_{n} = 0$, то есть характеристика поля положительная. ⊠

2. *Если число* $n \in \mathbb{N}$ *является характеристикой поля* P*, то число* n *простое.*

   ⧫ Пусть $n = pk$, где $p, k \in \mathbb{R}$. Тогда $\underbrace{1 + 1 + \ldots + 1}_{n} = 0$, то есть $\underbrace{(1 + 1 + \ldots + 1)}_{p} \cdot$ $\cdot \underbrace{(1 + 1 + \ldots + 1)}_{k} = 0$. Следовательно, из первого свойства или первый множитель равен нулю, или второй множитель равен нулю, что является противоречием с тем, что число n минимальное. Значит число n простое. ⊠

Глава 8

Кольцо многочленов

8.1 Определение и простейшие свойства многочленов.

Пусть P — некоторое поле.

• ***Многочленом над полем*** P *называется выражение вида*

$$a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0,$$

где a_i — элементы поля P, а x — символ. При этом элементы a_i называются ***коэффициентами многочлена****, символ x –* ***переменной многочлена****, а отдельные выражения $a_i x^i$ –* ***членами многочлена****. Обозначение многочлена: $f(x)$*

Множество всех многочленов над полем P с переменной x обозначается как $P[x]$.

• *Число k называется* ***степенью многочлена****, если $a_k \neq 0$, а коэффициенты при более высоких степенях x равны нулю, если они есть. Обозначение: $deg(f(x))$. Коэффициент a_k при этом называется* ***старшим коэффициентом многочлена*** *$f(x)$, а коэффициент a_0 —* ***свободным членом*** *многочлена $f(x)$.*

• *Если все коэффициенты равны нулю, то многочлен называется* ***нулевым.***

Степень нулевого многочлена не определена. Любой элемент поля P можно рассматривать как многочлен нулевой степени.

• *Два многочлена называются* ***равными****, если равны их степени и коэффициенты при одинаковых степенях переменной.*

Рассмотрим два многочлена из $P[x]$: $f(x) = a_n x^n + \ldots + a_1 x + a_0$ и $g(x) = b_k x^k + \ldots + b_1 x + b_0$. Пусть $a_n \neq 0$ и $b_k \neq 0$, то есть $deg(f(x)) = n, deg(g(x))k$, и пусть $n \geqslant k$ (для определённости).

• ***Суммой многочленов*** *$f(x)$ и $g(x)$ называется многочлен*

$$(f + g)(x) = c_n x^n + \ldots + c_1 x + c_0,$$

где

$$c_i = \begin{cases} a_i + b_i,\ 0 \leqslant i \leqslant k, \\ a_i, k + 1 \leqslant i \leqslant n. \end{cases}$$

Из определения суммы многочленов следует, что $deg(f{+}g)(x) \leqslant \max\{deg(f(x)), deg(g(x))\}$.

• ***Произведением многочленов*** $f(x)$ ***и*** $g(x)$ *называется многочлен*

$$(f \cdot g)(x) = d_{n+k}x^{n+k} + \ldots + d_1x + d_0,$$

где $d_i = \sum_{i=s+l} a_s b_l$, $i = \overline{0, n+k}$.

Из определения произведения многочленов следует, что $deg(f(x) \cdot g(x)) = deg(f(x)) + deg(g(x))$.

Теорема. *Множество* $P[x]$ *является ассоциативным, коммутативным кольцом с единицей относительно операций сложения и умножения многочленов.*

⧫ Из определения суммы и произведения следует, что результат операций тоже является многочленом, следовательно, эти операции определены в $P[x]$.
Операция + коммутативна, так как $(f+g)(x) = (a_n+b_n)x^n + \ldots + (a_1+b_1)x + (a_0+b_0) = (b_n+a_n)x_n + \ldots + (b_1+a_1)x + (b_0+a_0) = (g+f)(x)$. Остальные аксиомы кольца доказываются аналогично.
При этом нейтральным элементом для сложения является нулевой многочлен, а для умножения — многочлен $f(x) = 1$.
Симметричным относительно сложения для многочлена $f(x)$ является многочлен $-f(x)$ (все коэффициенты с противоположным знаком). ⊠

Из определения сложения и умножения многочленов и последней теоремы следует, что многочлен есть сумма его членов, а каждый его член – произведение $a_i x^i$. Следовательно, порядок членов в многочлене можно менять произвольно, члены с нулевым коэффициентом можно произвольно опускать и вводить в многочлен.

• *Кольцо* $(P[x], +, \cdot)$ *называется* ***кольцом многочленов*** *над полем* P *от одной переменной* x.

Свойства кольца многочленов:

1. *В кольце многочленов* $P[x]$ *нет делителей нуля.*

   ⧫ Из определения произведения многочленов следует, что старший коэффициент произведения ненулевых членов равен произведению ненулевых старших коэффициентов множителей, которые, в свою очередь, являются элементами поля P. А так как в поле делителей нуля нет, то произведение ненулевых многочленов будет являться ненулевым многочленом. ⊠

2. *Ненулевой многочлен* $f(x)$ *обратим в кольце многочленов* $\Longleftrightarrow deg(f(x)) = 0$.

   ⧫ $\Rightarrow$) Пусть $f(x)$ обратим, то есть $\exists f^{-1}(x) \in P[x] : f(x) \cdot f^{-1}(x) = 1 \Rightarrow$
   $\begin{cases} deg(f(x) \cdot f^{-1}(x)) = deg(1) = 0, \\ deg(f(x) \cdot f^{-1}(x)) = deg(f(x)) + deg(f^{-1}(x)) = 0. \end{cases}$ Так как степень больше нуля быть не может, то $deg(f) = 0$.

   $\Leftarrow$) Пусть $deg(f) = 0$. Тогда $f(x) = a_0 \in P \backslash \{0\}$, а для элемента $a \exists a_0^{-1} \in P \Rightarrow \exists g(x) = a_0^{-1} \in P[x]$. При этом $f(x) \cdot g(x) = a \cdot a^{-1} = 1$. Значит $g(x) = f^{-1}(x)$. ⊠

Из второго свойства становится понятным, что сокращать многочлен можно только на константу.

3. **Закон сокращения:** *Если* $h(x) \neq 0$ *и* $f(x) \cdot h(x) = g(x) \cdot h(x)$*, то* $f(x) = g(x)$.

   ♦ Если $f(x)h(x) = g(x)h(x)$, то $f(x)h(x) - g(x)h(x) = 0$. Тогда $h(x)(f(x) - g(x)) = 0$, а так как по первому свойству делителей нуля в кольце многочленов нет и $h(x) \neq 0$ по условию, то $f(x) - g(x) = 0$ и $f(x) = g(x)$. ⊠

8.2 Деление с остатком в кольце многочленов.

Пусть P — некоторое поле.

Теорема о делении с остатком. *Для любых двух многочленов* $f(x)$ *и* $g(x) \in P[x]$*, причем* $g(x) \neq 0$*, существует единственная пара многочленов* $q(x)$ *и* $r(x)$ *такая, что*

$$f(x) = g(x) \cdot q(x) + r(x), \text{ где } r(x) = 0 \text{ или } deg(r(x)) < deg(g(x)).$$

♦

1. Существование.

   Если $f(x) = 0$ или $deg(f(x)) < deg(g(x))$, то утверждение верно при $q(x) = 0$ и $r(x) = f(x)$, так как $f(x) = 0 \cdot g(x) + f(x)$.

   Пусть $deg(f(x)) \geqslant deg(g(x))$, $f(x) = a_n x^n + \ldots + a_1 x + a_0$ и $g(x) = b_k x^k + \ldots + b_1 x + b_0$. При этом $a_n \neq 0$ и $b_n \neq 0 \Rightarrow deg(f(x)) = n, deg(g(x)) = k$, и пусть $n \geqslant k$.

   Рассмотрим многочлен $\dfrac{a_n}{b_x} \cdot x^{n-k} \cdot g(x)$. Его степень равна n, а старший коэффициент равен a_n. Тогда, составив многочлен $f_1(x) = f(x) - \frac{a_n}{b_x} \cdot x^{n-k} \cdot g(x)$, получим, что $f_1(x)$ или нулевой, или $deg(f_1(x)) = k_1 < n = deg(f(x))$.

   Если $k_1 \geqslant k$, то обозначим старший коэффициент $f_1(x)$ за $C^{(1)}$. Построим многочлен $f_2(x) = f_1(x) - \dfrac{C^{(1)}}{b_k} \cdot x^{k_1 - k} \cdot g(x)$. Он также является либо нулевым многочленом, либо $deg(f_2(x)) = k_2 < k_1 = deg(f_1(x))$.

   Аналогично, если $k_2 \geqslant k$, построим $f_3(x)$. Так как на каждом шаге степень $f_i(x)$ понижается, но при этом у ненулевых многочленов степень неотрицательная, продолжим этот процесс до тех пор, пока на $s+1$ шаге не получим многочлен $f_{s+1}(x) = f_s(x) - \dfrac{C^{(s)}}{b_k} \cdot x^{k_s - k} \cdot g(x)$, который либо нулевой, либо степени меньшей, чем $k = deg(g(x))$.

   Сложим полученные равенства: $f_1(x) + \ldots + f_s(x) + f_{s+1}(x) = f(x) + f_1(x) + \ldots + f_s(x) - g(x)(\dfrac{a_n}{b_k}x^{n-k} + \ldots + \dfrac{C^{(s)}}{b_k}x^{k_s-k}) \Rightarrow$ [вычтем с обеих сторон $f_1(x) + \ldots + f_s(x)$] $\Rightarrow\ f_{s+1}(x) = f(x) - g(x)(\dfrac{a_n}{b_k}x^{n-k} + \ldots + \dfrac{C^{(s)}}{b_k}x^{k_s-k})$.

   Пусть $f_{s+1}(x) = r(x), (\frac{a_n}{b_k}x^{n-k} + \ldots + \frac{C^{(s)}}{b_k}x^{k_{s-k}}) = q(x)$. Выразив $f(x)$, получим $f(x) = g(x) \cdot q(x) + r(x)$. При этом многочлен $r(x)$ либо нулевой, либо со степенью меньшей, чем $g(x)$.

2. Единственность.

   От противного. Пусть для $f(x)$ и $g(x)$ существует два различных разложения:

   (a) $f(x) = q_1(x) \cdot g(x) + r_1(x)$;

   (b) $f(x) = q_2(x) \cdot g(x) + r_2(x)$,

   $r_1(x) \neq 0, r_2(x) \neq 0$ или $deg(r_1(x)) < deg(g(x))$, $deg(r_2(x)) < deg(g(x))$.

   Вычтем из первого второе: $0 = (q_1(x) - q_2(x))g(x) + (r_1(x) - r_2(x)) \Rightarrow (q_2(x) - q_1(x))g(x) = r_1(x) - r_2(x)$.

   Если $q_2(x) - q_1(x) \neq 0$, то $r_1(x) - r_2(x) \neq 0$ (вспомним, что $g(x)$ по условию теоремы тоже не равно нулю). Тогда так как равные многочлены имеют равные степени по определению равенства многочленов, то $deg((q_2(x)-q_1(x))g(x)) = deg(q_2(x)-q_1(x)) + deg(g(x)) = deg(r_1(x)-r_2(x)) \geqslant deg(g(x))$ и $deg(r_1(x)-r_2(x)) \leqslant max\{deg(r_1), deg(r_2)\} \leqslant deg(g(x))$ — противоречие. Тогда $q_2(x) - q_1(x) = 0 \Rightarrow r_1(x) - r_2(x) = 0 \Rightarrow$ разложения совпадают.

$\boxtimes$

• *Многочлен $q(x)$ в теореме называется* ***частным*** *при делении многочлена $f(x)$ на $g(x)$, а многочлен $r(x)$ —* ***остатком****.*

8.3 Делимость многочленов. НОД. Алгоритм Евклида.

• *Говорят, что многочлен $g(x)$* ***делит*** *многочлен $f(x)$ (или $f(x)$* ***делится*** *на $g(x)$), если остаток при делении $f(x)$ на $g(x)$ равен нулю. Обозначения: $g(x)|f(x)$ или $f(x) \vdots g(x)$.*

Свойства делимости многочленов:

1. $g(x)|f(x) \Longleftrightarrow \exists\varphi(x) \in P[x] : f(x) = \varphi(x) \cdot g(x)$.

   ♦ $\Rightarrow$) Пусть $g(x)|f(x)$. Тогда по теореме о делении с остатком $\exists! q(x), r(x) : f(x) = g(x) \cdot q(x) + r(x)$, при этом, так как $g(x)|f(x)$, то $r(x) = 0$, а многочлен $q(x)$ обозначим как $\varphi(x)$. Тогда $f(x) = \varphi(x) \cdot g(x)$.

   $\Leftarrow$) Пусть $\exists\varphi(x) : f(x) = g(x) \cdot \varphi(x)$. Тогда $f(x) = g(x) \cdot \psi(x) + 0 =$ [по аналогичной теореме] $= g(x) \cdot q(x) + r(x)$, где $r(x) = 0$. По теореме о делении с остатком такое представление $f(x)$ единственное и $g(x)|f(x)$. $\boxtimes$

2. *Если $h(x)|g(x)$, а $g(x)|f(x)$, то $h(x)|f(x)$.*

   ♦ Если $g(x)|f(x)$, то $\exists\varphi_1(x) : f(x) = g(x) \cdot \varphi_1(x)$;
   Если $h(x)|g(x)$, то $\exists\varphi_2(x) : h(x) = f(x) \cdot \varphi_2(x) = [f(x) = g(x) \cdot \varphi_1(x)] = g(x) \cdot \varphi_1(x) \cdot \varphi_2(x) = g(x) \cdot (\varphi_1(x) \cdot \varphi_2(x)) = [\varphi_1(x) \cdot \varphi_2(x) = \varphi(x)] = g(x) \cdot \varphi(x)$.
   Следовательно, по свойству 1 $h(x)|f(x)$. $\boxtimes$

3. *Если $h(x)|f(x)$, $h(x)|g(x)$, то $h(x)|(f(x) \pm g(x))$.*

   ♦ Пусть $h(x)|f(x)$. Тогда $\exists\varphi_1(x) : f(x) = h(x) \cdot \varphi_1(x)$. Если $h(x)|g(x)$, то $\exists\varphi_2(x) : g(x) = h(x) \cdot \varphi_2(x)$. Из полученных равенств следует, что $f(x) \pm g(x) = h(x) \cdot \varphi_1(x) \pm h(x) \cdot \varphi_2(x) = h(x) \cdot (\varphi_1(x) \pm \varphi_2(x)) = [\varphi_1(x) \pm \varphi_2(x) = \varphi(x)] = h(x) \cdot \varphi(x) \Rightarrow h(x)|(f(x) \pm g(x))$. $\boxtimes$

4. *Если $g(x)|f(x)$, то $\forall\, h(x) \in P[x]:\ g(x)|f(x)\cdot h(x)$.*

   ♦ Пусть $g(x)|f(x)$. Тогда $\exists\varphi(x) : f(x) = g(x)\cdot\varphi(x) \Rightarrow f(x)\cdot h(x) = g(x)\cdot(\varphi(x)\cdot h(x)) = [\varphi(x)\cdot h(x) = \varphi_1(x)] = g(x)\cdot\varphi_1(x) \Rightarrow g(x)|f(x)\cdot h(x)$. ⊠

5. *Если $g(x)|f(x)$, то $\forall a \in P, \neq 0 \quad (a\cdot g(x))|f(x)$.*

   ♦ Пусть $g(x)|f(x)$. Тогда $\exists\varphi(x) : f(x) = g(x)\cdot\varphi(x) = \varphi(x)\cdot 1\cdot g(x) = (\varphi(x)\cdot a^{-1})\cdot(a\cdot g(x)) = \varphi(x)\cdot(a\cdot a^{-1})\cdot g(x) = [\varphi(x)\cdot a^{-1} = \varphi_1(x)] = \varphi_1(x)\cdot(a\cdot g(x)) \Rightarrow (a\cdot g(x))|f(x)$. ⊠

6. *Два ненулевых многочлена $f(x)$ и $g(x)$ делят друг друга $\Longleftrightarrow \exists a \in P, \neq 0 : f(x) = a\cdot g(x)$.*

   ♦ $\Rightarrow$) Многочлены $g(x)$ и $f(x)$ делят друг друга, то есть $g(x)|f(x)$, тогда $\exists\,\varphi(x)$: $f(x) = \varphi(x)\cdot g(x)$. Следовательно, $deg(f(x)) = deg(\varphi(x)) + deg(g(x)) \geqslant deg(g(x))$.
   По аналогичным рассуждениям $deg(g(x)) \geqslant deg(f(x))$.
   Так как многочлены делят друг друга, то $deg(g(x)) = deg(f(x)) \Rightarrow deg(\varphi(x)) = 0 \Rightarrow \exists a \in P, \neq 0 : \varphi(x) = a \Rightarrow f(x) = a\cdot g(x)$.
   $\Leftarrow$) Пусть $f(x) = a\cdot g(x)$. Тогда $a = \varphi(x) \Longrightarrow f(x) = \varphi(x)\cdot g(x) \Rightarrow g(x)|f(x)$ по первому свойству.
   С другой стороны, так как $a \neq 0$, то $\exists a^{-1} : g(x) = a^{-1}\cdot f(x) = [a^{-1} = \varphi_1(x)] = \varphi_1(x)\cdot f(x)$. Следовательно, $f(x)|g(x)$. ⊠

• *Многочлен $d(x)$ называется **общим делителем** $f(x)$ и $g(x)$, если он делит каждый из этих многочленов. Многочлены нулевой степени являются общими делителями для любых многочленов $f(x)$ и $g(x)$. Если других общих делителей нет, то многочлены $f(x)$ и $g(x)$ называюся **взаимно простыми**.*

• ***Наибольшим общим делителем (НОД)** многочленов $f(x)$ и $g(x)$ называется их общий делитель, который делится на любой другой общий делитель этих многочленов.*

Теорема. *НОД любых двух ненулевых многочленов $f(x)$ и $g(x)$ всегда существует и определён с точностью до умножения на ненулевой постоянный элемент поля P.*

♦

1. Существование.
   Поделим $f(x)$ на $g(x)$ с остатком, и пусть $q_1(x)$ и $r_1(x)$ — частное и остаток при этом делении соответственно, то есть $f(x) = q_1(x)\cdot g(x) + r_1(x)$. Если $r_1(x) \neq 0$, то $deg(r_1(x)) < deg(g(x))$, тогда разделим $g(x)$ на $r_1(x)$ с остатком, и пусть $q_2(x)$ и $r_2(x)$ — частное и остаток при этом делении. Если $r_2(x) \neq 0$, то $deg(r_2(x)) < deg(r_1(x))$, тогда разделим $r_1(x)$ на $r_2(x)$ с остатком.
   Так как при этих делениях степень остатков понижаются, то через конечное число делений получим остаток $r_{n-1}(x) = 0$. В результате процесса деления получим равенства
   $$\begin{cases} f(x) = q_1(x)g(x) + r_1(x), \\ g(x) = q_2(x)r_1(x) + r_2(x), \\ r_1(x) = q_3(x)r_2(x) + r_3(x), \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ r_{n-3}(x) = q_{n-1}(x)r_{n-2}(x) + r_{n-1}(x), \\ r_{n-2}(x) = q_n(x)r_{n-1}(x) + r_n(x), \\ r_{n-1}(x) = q_{n+1}(x)r_n(x); \end{cases} \tag{8.3.1}$$

Покажем, что многочлен $r_n(x)$ является общим делителем $f(x)$ и $g(x)$. Из последнего равенства следует, что $r_n(x)|r_{n-1}(x)$ (по первому свойству). Тогда из предыдущего равенства $r_n(x)|r_{n-2}(x)$ (по свойствам 3 и 4). Аналогично $r_n(x)|r_{n-3}(x)$.
Продолжая рассуждения аналогичным образом получим, что $r_n(x)|g(x)$ и $r_n(x)|f(x)$, то есть $r_n(x)$ — их общий делитель.
Покажем, что $r_n(x)$ делится на любой другой общий делитель $f(x)$ и $g(x)$. Пусть $h(x)$ — общий делитель многочленов $f(x)$ и $g(x)$. Из первого равенства следует, что $r_1(x) = f(x) - q_1(x) \cdot g(x)$, из которого по свойству делимости $h(x)|r_1(x)$. Тогда из второго равенства $r_2(x) = g(x) - q_2(x) \cdot r_1(x)$ и аналогично $h(x)|r_2(x)$. Продолжая рассуждать аналогичным образом получим, что $d(x)|r_n(x)$. Значит $r_n(x)$ — НОД $f(x)$ и $g(x)$.

2. Покажем, что НОД определен с точностью до ненулевого постоянного множителя.
Пусть $d(x)$ — НОД $f(x)$ и $g(x)$. Тогда $\forall a \neq 0 \in P : ad(x)|f(x)$, $ad(x)|g(x)$ (по пятому свойству). Тогда $a \cdot d(x)$ — общий делитель.
Покажем, что $ad(x)$ — НОД. Если $h(x)$ — произвольный общий делитель, то по четвертому свойству делимости $h(x)|(a \cdot d(x))$. Следовательно, $a \cdot d(x)$ — НОД $f(x)$ и $g(x)$.
Покажем, что других НОДов не существует. Пусть $d_1(x)$ и $d_2(x)$ — НОД $f(x)$ и $g(x)$. Тогда $d_1(x)|d_2(x)$ и $d_2(x)|d_1(x)$ (по определению НОД), то есть они делят друг друга. Тогда по шестому свойству делимости $\exists a \neq 0 \in P : d_1(x) = a \cdot d2(x)$. Значит НОД определён с точностью до умножения на ненулевой постоянный многочлен.

⊠

• *Процесс поиска НОД, описанный при доказательстве теоремы, называется* ***Алгоритмом Евклида*** *(равенства* (8.3.1)*).*

8.4 Теорема о разложении НОД. Критерий взаимной простоты многочленов.

Теорема. *Пусть $d(x)$ — НОД многочленов $f(x)$ и $g(x)$. Тогда существуют многочлены $\varphi(x)$ и $\psi(x)$ такие, что $d(x) = \varphi(x) \cdot f(x) + \psi(x) \cdot g(x)$.*

♦ Если $g(x)|f(x)$. Тогда многочлен $g(x)$ является НОДом $f(x)$ и $g(x) \Rightarrow \exists a \in P, \neq 0 :$ $d(x) = a \cdot g(x) = 0 \cdot f(x) + a \cdot g(x) = [0 = \varphi(x), a = \psi(x)] = \varphi(x) \cdot f(x) + \psi(x) \cdot g(x)$. Следовательно, получаем необходимое разложение.

Пусть $g(x)$ не делит $f(x)$, и пусть $r_n(x)$ — последний ненулевой остаток в Алгоритме Евклида для многочленов $f(x)$ и $g(x)$. Он может отличаться от $d(x)$ на ненулевую константу: $d(x) = a \cdot r_n(x)$.

Из предпоследнего равенства алгоритма Евклида $r_n(x) = r_{n-2}(x) - q_n(x) \cdot r_{n-1}(x)$. Пусть $\varphi_1(x) = 1$, $\psi_1(x) = -q_n(x)$. Тогда $r_n(x) = \varphi_1(x) \cdot r_{s-2}(x) + \psi_1(x) \cdot r_{n-1}(x)$. Используя предыдущее равенство, $r_{n-1}(x) = r_{n-3}(x) - q_{n-2}(x) \cdot r_{n-2}(x) \Rightarrow r_n = \varphi_1(x) \cdot r_{n-2}(x) + \psi_1(x) \cdot (r_{n-3}(x) - q_{n-1}(x) \cdot r_{n-2}(x)) = \psi_1(x) \cdot r_{n-3}(x) + (\varphi_1(x) - \psi_1(x) \cdot q_{n-1}(x)) \cdot r_{n-2}(x)$. Заменим $\psi_1(x) = \varphi_2(x)$, $\varphi_1(x) - \psi_1(x) \cdot q_{n-1}(x) = \psi_2(x) \Rightarrow r_n(x) = \varphi_2(x) \cdot r_{n-3}(x) + \psi_2(x) \cdot r_{n-2}(x)$.

Продолжая рассуждения аналогичным образом (используя равенства Алгоритма Евклида) получим, что $\exists \varphi_n(x)$ и $\psi_n(x) : r_n(x) = \varphi_s(x) \cdot f(x) + \psi_n(x) \cdot g(x)$. А тогда так как $d(x) = a \cdot r_n(x)$, то $d(x) = a \cdot \varphi_n(x) \cdot f(x) + a \cdot \psi_n(x) \cdot g(x) = [a \cdot \varphi_n(x) = \varphi(x), a \cdot \psi_n(x) = \psi(x)] = \varphi(x) \cdot f(x) + \psi(x) \cdot g(x)$. ⊠

Следствие (критерий взаимной простоты многочленов). *Многочлены $f(x)$ и $g(x)$ взаимно простые* $\Longleftrightarrow \exists\varphi(x)$ *и* $\psi(x) : \varphi(x)\cdot f(x)+\psi(x)\cdot g(x)=1$.

♦ $\Rightarrow$) Если многочлен $f(x)$ и $g(x)$ взаимно простые, то многочлен 1 является их НОДом. Следовательно, по теореме выше $\exists\varphi(x)$ и $\psi(x) : 1=\varphi(x)\cdot f(x)+\psi(x)\cdot g(x)$.

$\Leftarrow$ Если $d(x)$ — общий делитель $f(x)$ и $g(x)$, то по третьему свойству делимости он делит $\varphi(x)\cdot f(x)+\psi(x)\cdot g(x)$. Следовательно, он делит 1, которая так же является общим делителем многочленов $f(x)$ и $g(x)$. Значит все общие делители — многочлены нулевой степени и $f(x)$ и $g(x)$ взаимно простые. ⊠

Свойства взаимно простых многочленов:

1. *Если многочлен $h(x)$ взаимно прост с каждым из многочленов $f(x)$ и $g(x)$, то он взаимно прост и с их произведением $f(x)g(x)$.*

   ♦ Так как $f(x)$ и $h(x)$ взаимно просты, то $\exists\varphi(x)$ и $\psi(x) : \varphi(x)\cdot f(x)+\psi(x)\cdot h(x)=1$. Домножим это равенство на $g(x)$ и получим $g(x)=\varphi(x)\cdot(f(x)\cdot g(x))+\psi(x)\cdot(h(x)\cdot g(x))$.

   Пойдём от обратного: пусть многочлены $f(x)g(x)$ и $h(x)$ не взаимно просты. Тогда $\exists d(x), deg(d(x))>0$, который по второму свойству делимости делит многочлены $h(x)$ и $f(x)g(x)$. Следовательно, $d(x)|g(x)$ и $d(x)$ — НОД многочленов $g(x)$ и $h(x)$ ненулевой степени, то есть $h(x)$ и $g(x)$ не взаимно просты, что является противоречием. Значит многочлены $f(x)g(x)$ и $h(x)$ взаимно просты. ⊠

2. *Если многочлен $h(x)$ делит произведение $f(x)g(x)$, но многочлены $h(x)$ и $f(x)$ взаимно простые, то $h(x)$ делит $g(x)$.*

   ♦ Так как многочлены $f(x)$ и $h(x)$ взаимно просты, то $\exists\varphi(x)$ и $\psi(x) : \varphi(x)\cdot f(x)+\psi(x)\cdot h(x)=1$. Домножим это равенство на $g(x)$ и получим $g(x)=\varphi(x)\cdot(f(x)\cdot g(x))+\psi(x)\cdot(h(x)\cdot g(x))$. Так как оба слагаемых в правой части равенства делятся на $h(x)$, то $h(x)|g(x)$. ⊠

3. *Если $g(x)$ делит $f(x)$ и $h(x)$ делит $f(x)$ и при этом $g(x)$ взаимно прост с $h(x)$, то произведение $g(x)\cdot h(x)$ также делит многочлен $f(x)$.*

   ♦ Так как $g(x)|f(x)$, то $\exists\varphi(x) : f(x)=\varphi(x)\cdot g(x)$. Так как $h(x)|f(x)$, то $h(x)|(\varphi(x)\cdot g(x))$. Следовательно, по второму свойству взаимно простых многочленов $h(x)|\psi(x)\Rightarrow \exists\psi(x) : \varphi(x)=\psi(x)\cdot h(x)\Rightarrow f(x)=\psi(x)\cdot(h(x)\cdot g(x))\Rightarrow(h(x)\cdot g(x))|f(x)$. ⊠

8.5 Неприводимые многочлены.

• *Многочлен $f(x)\in P[x]$ положительной степени называется* ***неприводимым*** *над полем P, если его нельзя представить в виде произведения многочленов из $P[x]$, степень которых меньше, чем $deg(f(x))$. В противном случае многочлен называется* ***приводимым.***

Свойства неприводимых многочленов:

1. *Любой многочлен первой степени неприводим.*

   ♦ Если многочлен приводим и находится в первой степени, то его можно представить в виде произведения многочленов нулевой степени, но произведение многочленов нулевой степени является многочленом нулевой степени. Получили противоречие. ⊠

2. *Если многочлен $f(x)$ неприводим над полем P, то для любого ненулевого многочлена $g(x) \in P[x]$ либо $f(x)$ и $g(x)$ взаимно простые, либо $f(x)|g(x)$.*

   ♦ Пусть многочлен $d(x)$ — НОД многочленов $f(x)$ и $g(x)$. Тогда $d(x)|f(x)$ и $\exists\varphi(x)$: $f(x) = \varphi(x) \cdot d(x)$, но так как многочлен $f(x)$ неприводим, то либо $deg(f(x)) = deg(\varphi(x)) \Rightarrow deg(d(x)) = 0 \Rightarrow f(x)$ и $g(x)$ взаимно просты, либо $deg(f(x)) = deg(d(x)) \Rightarrow deg(\varphi(x)) = 0 \Rightarrow \exists a \in P, \neq 0 : f(x) = a \cdot d(x) \Rightarrow d(x) = a^{-1}f(x) \Rightarrow f(x)|d(x) \Rightarrow f(x)$ — НОД $f(x)$ и $g(x) \Rightarrow f(x)|g(x)$. ⊠

3. Если неприводимый многочлен $h(x)$ делит произведение $f(x)g(x)$, то хотя бы один из сомножителей делится на $h(x)$.

   ♦ Пусть $f(x)$ не делится на $h(x)$. Тогда $f(x)$ и $h(x)$ по второму свойству неприводимых многочленов взаимно просты, следовательно, по второму свойству взаимно простых многочленов $h(x)|g(x)$. ⊠

Теорема (основная теорема теории делимости многочленов). *Любой многочлен положительной степени $f(x) \in P[x]$ представим в виде произведения неприводимых многочленов*

$$f(x) = p_1(x) \cdot \ldots \cdot p_k(x),$$

причём это разложение единственно с точностью до множитедей нулевой степени.

♦

1. Существование.
   Если $f(x)$ неприводим, то разложение имеет вид $f(x) = f(x)$. Если же он приводим, то его можно представить в виде произведения многочленов меньшей степени. Если же они приводимы, то их так же можно представить в виде произведения многочленов меньшей степени. Этот процесс не может быть бесконечным, так как степень многочлена не может быть меньше 1. В результате получим разложение многочлена на неприводимые множители.

2. Единственность.
   Пусть $f(x)$ имеет ещё одно разложение на неприводимые многочлены, то есть $f(x) = q_1(x) \cdot \ldots \cdot q_s(x)$. Тогда $p_1(x) \cdot \ldots \cdot p_k(x) = q_1(x) \cdot \ldots \cdot q_s(x)$, и пусть для определённости $s \geqslant k$.
   Из равенства следует, что произведение $q_1(x) \cdot \ldots \cdot q_s(x)$ делится на $p_1(x)$, следовательно, так как каждый $p_1(x)$ – неприводимый многочлен, то $\exists q_i(x)$, который делится на $p_1(x)$ (по свойству 3 неприводимых многочленов). Не нарушая общности, перенумеруем эти $q_i(x)$ так, чтобы это был многочлен $q_1(x)$. Тогда $\exists\varphi(x) : q_1(x) = \varphi(x) \cdot p_1(x)$, но $q_1(x)$ так же неприводим. Тогда либо $deg(p_1(x)) = 0$, либо $deg(\varphi(x)) = 0$, то есть $\exists a_1 \in P : q_1(x) = a_1 \cdot p_1(x)$. Подставив это преставление в правую часть равенства, получим $p_1(x) \cdot \ldots \cdot p_k(x) = a_1 \cdot p_1(x) \cdot q_2(x) \cdot \ldots \cdot q_s(x)$. Сократим на $p_1(x)$ и получим $p_2(x) \cdot \ldots \cdot p_k(x) = a_1 \cdot q_2(x) \cdot \ldots \cdot q_s(x)$.
   Рассуждая аналогичным образом, проведём эту операцию ещё $(k-1)$ раз и получим равенство $a_1 \cdot a_2 \cdot \ldots \cdot a_k \cdot q_{k+1}(x) \cdot \ldots \cdot q_s(x) = 1$. Так как многочлены $q_i(x)$ неприводимы, то их степени не превосходят единицы, следовательно, последнее равенство возможно только при $k = s$ (так как у равных многочленов равные степени).
   Таким образом, второе разложение отличается от первого на константу $a_1 \cdot \ldots \cdot a_k$, то есть $f(x) = q_1(x) \cdot \ldots \cdot q_k(x)$, где $q_i(x) = a_i \cdot p_i(x)$.

⊠

Следствие. *Любой многочлен положительной степени $f(x) \in P[x]$ единственным образом представим в виде произведения*

$$f(x) = a \cdot h_1(x) \cdot \ldots \cdot h_k(x),$$

где a — старший коэффициент многочлена $f(x)$, а $h_i(x)$ — неприводимые многочлены со старшим коэффициентом 1.

♦ Любой многочлен $p_i(x)$ представим в виде произведения $p_i(x) = a_i \cdot h_i(x)$, где a_i — старший коэффициент многочлена, а $h_i(x)$ — неприводимый многочлен со старшим коэффициентом 1. Если $h_i(x)$ приводимый, то многочлен $p_i(x)$ также приводимый. Обозначим $a = a_1 \cdot \ldots \cdot a_k$ и получим искомое разложение. ⊠

• *Неприводимые многочлены $h_i(x)$ из следствия не обязательно различны. Если $h_i(x)$ входит в это разложение k раз, то он называется **k-кратным множителем для многочлена** $f(x)$. Если $k = 1$, то $h_i(x)$ называется **простым множителем.***

Таким образом, последнее разложение можно представить в виде

$$f(x) = a \cdot h_1(x)^{k_1} \cdot \ldots \cdot h_s(x)^{k_s},$$

где a — старший коэффициент многочлена $f(x)$, $h_i(x)$ — различные неприводимые многочлены со старшим коэффициентом 1, k_i — их кратности.

• *При этом многочлены $h_i(x)^{k_i}$ называются **элементарными делителями** многочлена $f(x)$.*

8.6 Корни многочлена.

Рассмотрим многочлен

$$f(x) = a_n x^n + \ldots + a_1 x + a_0 \in P[x].$$

• *Пусть c — произвольный элемент поля P. Элемент*

$$f(c) = a_n c^n + \ldots + a_1 c + a_0 \in P$$

*называется **значением многочлена** $f(x)$ при $x = c$.*

Из определения значения многочлена следует, что каждому элементу $c \in P$ можно поставить в соответствие элемент $f(c)$, так как любой многочлен $f(x) \in P[x]$ определяет отображение $f : P \to P$.

Свойства значения многочлена:

1. *Если $f(x) = g(x)$, то $\forall c \in P$ $f(c) = g(c)$. Обратное утверждение, вообще говоря, неверно.*

2. *Если $\varphi(x) = f(x) + g(x)$, то $\forall c \in P$ $\varphi(c) = f(c) + g(c)$.*

3. *Если $\varphi(x) = f(x) \cdot g(x)$, то $\forall c \in P$ $\varphi(c) = f(c) \cdot g(c)$.*

• *Элемент $c \in P$ называется **корнем многочлена** $f(x)$, если $f(c) = 0$.*

Теорема Безу. *Остаток при делении многочлена $f(x) \in P[x]$ на многочлен $(x-c)$ равен $f(c)$.*

♦ Разделим $f(x)$ на $(x-c)$ с остатком: $f(x) = q(x)\cdot(x-c)+r(x)$, где $r(x)$ или равен нулю, или его степень меньше, чем степень многочлена $x - c$ (то есть равна нулю). СЛедовательно, $r(x)$ — постоянная, то есть $\exists a \in P : r(x) = a$. Тогда $f(c) = q(c)\cdot(c-c)+a \Rightarrow f(c) = a = r(x)$. ⊠

Следствие. *Элемент $c \in P$ является корнем многочлена $f(x) \in P[x] \Longleftrightarrow (x - c)|f(x)$.*

♦ $\Rightarrow$) Пусть элемент c — корень. Тогда $f(c) = 0$. По тоереме Безу $r(x) = 0$ и $f(x) = q(x)(x - c)$. Следовательно, $(x - c)|f(x)$.

$\Leftarrow$) Пусть $(x - c)|f(x)$. Тогда $f(x) = \varphi(x)(x - c) + 0 \Rightarrow r(x) = 0$. Но по теореме Безу $f(x) = f(c) = 0$. Следовательно, c — корень многочлена $f(x)$. ⊠

По теореме о делении с остатком в конце многочленов любой многочлен $f(x)$ можно разделить с остатком нам многочлен $(x - c)$. Одним из способов является **схема Горнера**.

Схема Горнера:

Пусть $f(x) = a_nx^n + \ldots + a_1x + a_0$ и пусть

$$f(x) = q(x)(x - c) + r. \tag{8.6.1}$$

Тогда $deg(q(x)) = deg(f(x)) - 1 = n$, То есть $q(x) = b_{n-1}x^{n-1} + \ldots + b_1x + b_0$. Приравняем коэффициенты в обеих частях равенства (8.6.1):

$$\begin{cases} x^n : \ a_n = b_{n-1}, \\ x^{n-1} : \ a_{n-1} = b_{n-2} - c \cdot b_{n-1}, \\ x^{n-2} : \ a_{n-2} = b_{n-3} - c \cdot b_{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^1 : \ a_1 = b_0 - c \cdot b_1, \\ x^0 : \ a_0 = r - c \cdot b_0; \end{cases} \Rightarrow \begin{cases} b_{n-1} = a_n, \\ b_{n-2} = a_{n-1} + c \cdot b_{n-1}, \\ b_{n-3} = a_{n-2} + c \cdot b_{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ b_0 = a_1 + c \cdot b_1, \\ r = a_0 + c \cdot b_0; \end{cases} \text{— Схема Горнера}$$

a_i	a_n	a_{n-1}	a_{n-2}	$\cdots$	a_1	a_0
c	$b_{n-1} = a_n$	$b_{n-2} = a_{n-1} + c \cdot b_{n-1}$	$b_{n-3} = a_{n-2} + c \cdot b_{n-2}$	$\cdots$	$b_0 = a_1 + c \cdot b_1$	$r = a_0 + c \cdot b_0$

• *Пусть элемент $c \in P$ является корнем многочлена $f(x)$. Число $k \in \mathbb{N}$ называется* ***кратностью корня*** *c, если многочлен $f(x)$ делится на многочлен $(x-c)^k$ и не делится $(x - c)^{k+1}$. Если кратность корня равна 1, то корень называется* ***простым****, в проътивном случае —* ***кратным****.*

Из свойств делимости следует, что кратность корня не превосходит степени многочлена.

Пусть $f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$.

• ***Производной*** *многочлена $f(x)$ называется многочлен*

$$f'(x) = n \cdot a_nx^{n-1} + (n - 1) \cdot a_{n-1}x^{n-2} + \ldots + 2 \cdot a_2x + a_1.$$

Свойства производной:

1. $(f+g)'(x) = f'(x) + g'(x)$.
2. $(f \cdot g)'(x) = f'(x) \cdot g(x) + g'(x) \cdot f(x)$.
3. $(f^k)'(x) = k \cdot f^{k-1}(x) \cdot f(x)$.

Теорема. *Пусть P — поле нулевой характеристики. Если элемент $c \in P$ является корнем многочлена $f(x) \in P[x]$ кратности k, то при $k > 1$ элемент c является $(k-1)$-кратным корнем производной $f'(x)$, а при $k = 1$ элемент c не является корнем для $f'(x)$.*

♦ Пусть c — корень кратности k. Тогжа $(x-c)^k$ делит $f(x)$ и $(x-c)^{k+1}$ не делит $f(x)$, то есть $\exists \varphi(x) : f(x) = \varphi(x) \cdot (x-c)^k$, где $\varphi(x)$ не делится на $(x-c)$. Тогда по теореме Безу $q(c) \neq 0$. Следовательно, $f'(x) = k \cdot \varphi(x) \cdot (x-c)^{k-1} + \varphi'(x) \cdot (x-c)^k = (x-c)^{k-1}(k \cdot \varphi(x) + \varphi'(x) \cdot (x-c)) \Rightarrow (x-c)^{k-1} | f'(x)$.

Если $(x-c)^k | f'(x)$, то c — корень кратности k, тогда $(x-c)^k | (x-c)^{k-1}(k \cdot \varphi(x) + \varphi'(x) \cdot (x-c)) \Rightarrow (x-c) | (k \cdot \varphi(x) + \varphi'(x) \cdot (x-c)) \Rightarrow c$ — корень $(k \cdot \varphi(x) + \varphi'(x) \cdot (x-c)$. Но при подстановке c в данный многочлен получаем $k \cdot \varphi(c) + \varphi'(c) \cdot (c-c) = k \cdot q(c) = 0$. Вспомнив начало доказательства $[q(c) \neq 0]$, получаем противоречие с тем, что $(x-c)^k | f'(x)$. Значит c — корень производной $f'(x)$ кратности $k-1$. ⊠

8.7 Количество корней многочлена.

Теорема. *Любой многочлен $f(x) \in P[x]$ положительной степени $(n \geqslant 1)$ может иметь не более n корней, если каждый из его корней считать столько раз, какова его кратность.*

♦ Предположим, что многочлен $f(x)$ имеет различные корни $C_1, \ldots, C_s \in P$ кратности $k_1, \ldots, k_s \in \mathbb{N}$ соответственно.
Покажем, что $k_1 + \ldots + k_s \leqslant n$. Так как C_1 — корень многочлена $f(x)$ кратности k_1, то $(x-C_1)^k_1 | f(x)$. Следовательно, $\exists f_1(x) \in P[x] : f(x) = (x-C_1)^k_1 \cdot f_1(x)$.

Так как C_2 — корень $f(x)$ кратности k_2, то $(x-C_2)^k_2 | f(x)$, следовательно, $(x-C_2)^k_2 | (x-C_1)^k_1 \cdot f1(x)$.
А так как $(C_1 - C_2)^k_1 \neq 0$ (потому что $C_1 \neq C_2$), то многочлен $(x-C_1)^k_1$ взаимно прост с многочленом $(x-C_2)^k_2$. Тогда по второму свойству взаимно простых многочленов $(x-C_2)^k_2 | f_1(x)$, следовательно, $\exists f_2(x) \in P[x] : f_1(x) = (x-C_2)k_2 \cdot f_2(x) \Rightarrow f(x) = (x-C_1)^k_1 \cdot (x-C_2)^k_2 \cdot f_2(x)$.
Продолжая аналогичные рассуждения, на s-м шаге получим, что $\exists f_s(x) \in P[x] : f(x) = (x-C_1)^k_1 \cdot (x-C_2)^k_2 \cdot \ldots \cdot (x-C_s)^k_s \cdot f_s(x)$. Следовательно, так как у равных многочленов степени равны, то $n = deg(f_s(x)) + k_1 + \ldots + k_s \geqslant k_1 + \ldots + k_s$, и $f(x)$ имеет не более n корней, если каждый из корней считать столько раз, какова его кратность. ⊠

Следствие. *Любой многочлен $f(x) \in P[x]$ положительной степени $(n \geqslant 1)$ имеет не более n* ***различных*** *корней.*

Следствие (Формула Виета). *Если многочлен*

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in P[x]$$

имеет с учетом кратности ровно n корней $c_1, \dots, c_n$, то справедливы равенства

$$\begin{cases} a_{n-1} = -(c_1 + \ldots + c_n); \\ a_{n-2} = (c_1 \cdot c_2 + c_1 \cdot c_3 + \ldots + c_{n-1} c_n) = \sum\limits_{i \neq j} c_i \cdot c_j; \\ a_{n-3} = -(c_1 \cdot c_2 \cdot c_3 + \ldots + c_{n-2} \cdot c_{n-1} \cdot c_n) = -\sum\limits_{i \neq j \neq k} c_i \cdot c_j \cdot c_k; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_1 = (-1)^{n-1}(c_1 \cdot \ldots \cdot c_{n-1} + \ldots + c_2 \cdot \ldots \cdot c_n); \\ a_0 = (-1)^n c_1 \cdot \ldots \cdot c_n. \end{cases}$$

♦ Так как многочлен $f(x)$ с учётом кратности имеет ровно n корней, то, как и при доказательстве теоремы выше, его можно представить в виде $f(x) = (x - c_1) \cdot \ldots \cdot (x - c_n) \cdot f_n(x)$. $0 \leqslant deg(f_n(x)) = n - n = 0$. Тогда $\exists a \in P, \neq 0 : f_n(x) = a$. Приравнивая коэффициенты при соответствующих степенях переменной x у многочленов слева и справа, получим формулу Виета. ⊠

Теорема. *Если два многочлена $f(x)$, $g(x)$, степени которых не превосходят n, принимают одинаковые значения при $n+1$ различном значении переменной, то они равны, то есть $f(x) = g(x)$.*

♦ Пусть $deg(f(x)) \leqslant n$, $deg(g(x)) \leqslant n$ и $C_1, \ldots, C_{n+1} \in P$ — попарно различные элементы, для которых выполняется равенство $f(c_i) = g(c_i)$.
Рассмотрим многочлен $\varphi(x) = f(x) - g(x)$. Тогда или $\varphi(x) = 0$, или $deg(\varphi(x)) < max\{deg(f(x)), deg(g($
n. При этом $\varphi(c_i) = f(c_i) - g(c_i) = 0$. То есть если многочлен $\varphi(x)$ ненулевой, то он имеет $n+1$ различных корней, но его степень не превышает n, что является противоречием с теоремой. Тогда $\varphi(x)$ — нулевой многочлен и $f(x) = g(x)$. ⊠

Теорема. *Для любых элементов $b_1, \ldots, b_{n+1} \in P$ и попарно различных элементов $c_1, \ldots, c_{n+1} \in P$ существует единтсвенный многочлен $f(x)$, $deg(f(x)) \leqslant n$, для которого выполняется равенство $f(c_i) = b_i$, $i = \overline{1, n-1}$.*

♦ Построим многочлен

$$f(x) = \sum_{i=1}^{n+1} b_i \varphi_i(x), \tag{8.7.1}$$

где

$$\varphi_i(x) = \frac{(x - c_1)(x - c_2)\ldots(x - c_{i-1})(x - c_{i+1})\ldots(x - c_{n+1})}{(c_i - c_1)(c_i - c_2)\ldots(c_i - c_{i-1}(c_i - c_{i+1})\ldots(c_i - c_{n+1})}.$$

Каждое выражение $\varphi_i(x)$ является многочленом со степенью, не превышающей n. Следовательно, их сумма — многочлен степени не выше n, то есть $deg(f(x)) \leqslant n$ и при этом: $\varphi_i(c_j) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \Rightarrow f(c_j) = \sum_{i=1}^{n+1} b_i \varphi_i(c_j) =$[элемент суммы не равен нулю только при $i = j$]$= b_i \varphi_i(c_i) = b_i =$[так как $i = j$]$= b_j$.
Единственность же такого многочлена доказана в предыдущей теореме. ⊠

• *Многочлен* (8.7.1) *называется* ***интерполяционным многочленом Лагранжа.***

8.8 Основная теорема алгебры комплексных чисел. Неприводимость многочленов над полем действительных и комплексных чисел.

Основная теорема алгебры. *Любой многочлен над полем комплексных чисел ($f(x) \in \mathbb{C}[x]$) положительной степени имеет хотя бы один корень.*

Следствие. *Многочлен $f(x) \in \mathbb{C}[x]$ неприводим над $\mathbb{C} \iff deg(f(x)) = 1$.*

♦ $\Rightarrow$) Пусть многочлен $f(x)$ неприводим. Следовательно, $deg(f(x)) \geqslant 1$. По основной теореме алгебры существует корень c, а по теореме Безу $(x-c)|f(x)$. Тогда $\exists\varphi(x)$: $f(x) = \varphi(x)\cdot(x-c)$, но многочлен $f(x)$ неприводим. Следовательно, $deg(\varphi(x)) = 0$ и $deg(f(x)) = deg((x-c)) = 1$.

$\Leftarrow$) Из первого свойства неприводимости многочлен первой степени всегда неприводим. ⊠

Следствие. *Для любого многочлена $f(x) \in \mathbb{C}[x]$ степени $n \geqslant 1$ существуют числа $c_1, \ldots, c_n \in \mathbb{C}$ такие, что*

$$f(x) = a(x-c_1)\cdot\ldots\cdot(x-c_n). \tag{8.8.1}$$

♦ Из следствия основной теоремы делимости многочленов любой многочлен можно представить в виде произведения старшего коэффициента и неприводимых многочленов над $\mathbb{C}$ со старшим коэффициентом 1. ⊠

Следствие. *Любой многочлен $f(x) \in \mathbb{C}[x]$ положительной степени n имеет с учётом кратности ровно n корней над полем $\mathbb{C}$, если каждый из корней считать столько раз, какова его кратность.*

♦ Объединив одинаковые многочлены в разложении (8.8.1), получим $f(x) = a\cdot(x-c_1)^{k_1}\cdot\ldots\cdot(x-c_j)^{k_j}$, где $c_i \neq c_j$ для $i \neq j \Rightarrow (x-c_i)^{k_1}|f(x)$. Покажем, что многочлен $f(x)$ не делится на $(x-c_1)^{k_1+1}$ (пойдём от противного) и поделим $f(x)$ на $(x-c_1)^{k_1}$, получив многочлен $\varphi(x) = (x-c_2)\cdot\ldots\cdot(x-c_j)^{k_j}$. Подставив вместо x значение c_1, получим $\varphi(c) = (c_1-c_2)\cdot\ldots\cdot(c_1-c_j)^{k_j} \neq 0$ (так как $c_i \neq c_j$). Следовательно, c_1 — корень кратности k_1.
Аналогично показываем, что остальные корни c_i кратностей k_i. Так как Многочлены слева и справа в разложении равны, то их степени также равны. Значит $n = k_1 + \ldots + k_j$. ⊠

Теорема. *Если мнимое число z кратности k является корнем многочлена $f(x)$ с действительными коффициентами, то сопряженное число $\overline{z}$ также является корнем этого многочлена кратности k.*

♦ $f(\overline{z}) = a_n\overline{z}^n + \ldots + a_1\overline{z} + a_0 = \overline{a_n z^n} + \ldots + \overline{z} + \overline{a_0} = \overline{a_n z^n + \ldots + a_1 z + a_0} = \overline{f(z)} = \overline{0} = 0$.
Следовательно, z — корень.
Кратность докажем от противного. Пусть $\overline{z}$ — корень многочлена $f(x)$ кратности $m < k$ – кратность Z. Тогда $f(x) = (x-z)^k(x-\overline{z})^m\varphi(x) = ((x-z)(x-\overline{z}))^m(x-z)^{k-m}\varphi(x)$. Рассмотрим многочлен $g(x) = (x-z)(x-\overline{z}) = x^2 - x(z+\overline{z}) + z\cdot\overline{z}$. Заметим, что по свойствам сопряжённого комплексного числа коэффициенты $g(x)$ — действительные числа, то есть $g(x) \in \mathbb{R}[x]$.
Пусть $h(x) \in \mathbb{R}[x] = (x-Z)^{k-m}\varphi(x)$ — частное при делении $f(x)$ на $g(x)$. Тогда по теореме о делении с остатком в кольце многочленов $f(x) \in \mathbb{R}[x]$ и при этом z — его корень. Но по доказаному выше $\overline{z}$ — тоже его корень. Однако $(\overline{z}-z)^{k-m}\varphi(\overline{z}) \neq 0$, получили противоречие, так как z — корень. Значит $k = m$. ⊠

Следствие. *Любой многочлен с действительными коэффициентами имеет чётное число мнимых корней, которые попарно сопряжены между собой.*

Следствие. *Любой многочлен с действительными коэффициентами нечётной степени имеет хотя бы один действительный корень.*

Теорема. *Многочлен $f(x) \in \mathbb{R}[x]$ неприводим над полем $\mathbb{R} \iff$ он либо линейный (первой степени), либо квадратный, не имеющий действительных корней.*

♦ $\Rightarrow$) Пусть многочлен $f(x)$ неприводим. Так как $f(x) \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$, то у него существует корень $c \in \mathbb{C} \Rightarrow f(x) = (x-c) \cdot \varphi(x)$. Если $c \in \mathbb{R}$, то $\varphi(x)$ является частным при делении $f(x)$ на $(x-c)$ в кольце $\mathbb{R}[x]$, то есть $\varphi(x) \in \mathbb{R}[x]$. Тогда так как $f(x)$ неприводим, то $deg(\varphi(x)) = 0$, $deg(f(x)) = 1$. В противном случае c — мнимое число и его сопряжённое тоже корень, то есть $f(x) = (x-c)(x-\overline{c})\varphi(x)$.
Пусть $g(x) = (x-c)(x-\overline{c}) = x^2 - x(c+\overline{c}) + c \cdot \overline{c}$. Тогда у многочлена $g(x)$ действительные коэффициенты, то есть $g(x) \in \mathbb{R}[x]$. Тогда $\varphi(x) \in \mathbb{R}[x]$, $deg(\varphi(x)) = 0$. Следовательно, $f(x) = a \cdot (x-c)(x-\overline{c})$. Значит $f(x)$ — многочлен второй степени и не имеет действительных корней.

$\Leftarrow$) Многочлен первой степени неприводим по первому свойству неприводимости многочленов. Пусть $f(x)$ — квадратный многочлен, не имеющий действительных корней, и предположим, что он приводим, то есть представим в виде $f(x) = f_1(x) \cdot f_2(x)$, причём $deg(f_1(x)) = deg(f_2(x)) = 1$. Тогда многочлен $f_1(x)$ имеет вид $ax+b(a \neq 0)$, следовательно, многочлен f_1 имеет корень $c = -\frac{b}{a} \in \mathbb{R} \Rightarrow f(-\frac{b}{a}) = f_1(-\frac{b}{a}) \cdot f_2(-\frac{b}{a}) = 0$, что является противоречием с тем, что многочлен $f(x)$ не имеет действительных корней. Значит многочлен $f(x)$ неприводим. ⊠

Глава 9

Матрицы и определители

9.1 Матрицы. Основные виды матриц.

Пусть P — некоторое поле.

• *Прямоугольная таблица вида*

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

*составленная из элементов $\alpha_{ij} \in P$ называется **матрицей над полем** P.*

• *Если она состоит из m строк и n столбцов, то ее называют **матрицей размера** $m \times n$ или $m \times n$**-матрицей**. Если число строк и столбцов матрицы одинаково и равно n, то матрица называется **квадратной матрицей порядка** n.*

Множество всех матриц размера $m \times n$ над полем P обозначим через $P_{m,n}$.

• *Прямоугольная матрица, состоящая из одной строки, называется **матрицей-строкой** или **строкой**. Прямоугольная матрица, состоящая из ондого столбца, называется **матрицей-столбцом** или **столбцом**.*

• *Матрица размера $m \times n$, все элементы которой равны нулевому элементу поля P, называется **нулевой** (Обозначение: O или $O_{m,n}$).*

Пусть $A = (\alpha_{ij}) \in P_{n,n}$ — квадратная матрица порядка n.

• *Часть матрицы, определяемая элементами $a_{11}, a_{22}, \dots, a_{nn}$ называется **главной диагональю матрицы**. Часть матрицы, определяемая элементами $a_{n1}, a_{(n-1)2}, \dots, a_{1n}$ называется **побочной диагональю матрицы**.*

• *Квадратная матрица называется **диагональной**, если все ее элементы, не стоящие на главной диагонали, равны нулю. Причем, если равны нулю элементы ниже главной диагонали, то матрица называется **верхней треугольной**, а если равны нулю элементы выше главной диагонали, то матрица называется **нижней треугольной**.*

Пусть целые числа $m_0, m_1, \ldots, m_l, n_0, n_1, \ldots, n_k$ удовлетворяют соотношениям

$$0 = m_0 < m_1 < \cdots < m_l = m,$$

$$0 = n_0 < n_1 < \cdots < n_k = n.$$

Тогда матрица $A = (\alpha_{ij}) \in P_{m,n}$ представима в виде

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1k} \\ A_{21} & A_{22} & \ldots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{l1} & A_{l2} & \ldots & A_{lk} \end{pmatrix},$$

где A_{sp} — матрица (α_{ij}).

• *Такое представление матрицы A называется* ***блочным разбиением*** *(или* ***клеточным разбиением****). При этом матрицы A_{sp} называются* ***блоками*** *матрицы A (****клетками****).*

• *Матрица, имеющая блочное разбиение с квадратными блоками A_{ii} на главной диагонали и нулевыми блоками вне главной диагонали, называется* ***блочнодиагональной*** *или* ***квазидиагональной****. Матрица называется* ***блочнотреугольной****, если ее диагональные блоки — квадратные, а блоки, расположенные по одну сторону от главной диагонали, нулевые. Причем, если нулевыми являются блоки ниже главной диагонали, то матрица называется* ***верхней блочнотреугольной****, а если нулевыми являются блоки выше главной диагонали, то матрица называется* ***нижней блочнотреугольной****.*

• *Две матрица называются* ***равными****, если равны их размеры и элементы, стоящие на одних и тех же местах.*

9.2 Линейные операции над матрицами. Умножение матриц. Транспонирование.

• *Матрица $C = (c_{ij}) \in P_{m,n}$ называется* ***суммой матриц*** *$A = (a_{ij})$, $B = (b_{ij}) \in P_{m,n}$, если $c_{ij} = a_{ij} + b_{ij} \quad \forall i, j$ (Обозначение: $A + B$).*

Свойства сложения матриц:

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $\forall A \in P_{m,n} \quad A + O_{m,n} = O_{m,n} + A = A$;
4. $\forall A = (a_{ij}) \in P_{m,n}\ \exists B = (b_{ij}) \in P_{m,n} : A + B = B + A = O_{m,n}$.

   ⧫ Доказательства всех свойств следуют из того, что элементы матрицы являются элементами поля. ⊠

• *Если матрица B симметрична относительно сложения для матрицы A, то она называется* ***противоположной*** *для матрицы A (Обозначение: $-A$).*

Из свойств сложения матриц следует, что множество $P_{m,n}$ матриц одного размера образует абелеву группу относительно сложения.

• *Матрица $C = (c_{ij}) \in P_{m,n}$ называется* ***произведением матрицы*** *$A = (a_{ij}) \in P_{m,n}$* ***на элемент*** *$\alpha \in P$* ***(произведением матрицы на скаляр)****, если $c_{ij} = \alpha a_{ij} \quad \forall i, j$.*

Свойства произведения матрицы на элемент поля:

1. $\forall A = (a_{ij}) \in P_{m,n} \quad 1 \cdot A = A,\ (-1) \cdot A = -A$;

2. $\forall \alpha, \beta \in P,\ A, B \in P_{m,n}$
   $\alpha(\beta A) = (\alpha\beta)A$;
   $(\alpha + \beta)A = \alpha A + \beta A$;
   $\alpha(A + B) = \alpha A + \alpha B$.

• *Операции сложения матриц и умножения матрицы на элемент поля называются* ***линейными операциями****.*

• *Матрица $C = (c_{ij}) \in P_{m,k}$ называется* ***произведением матрицы*** *$A = (a_{ij}) \in P_{m,n}$* ***на матрицу*** *$B = (b_{ij}) \in P_{n,k}$, если $c_{ij} = a_{i1}b_{1j} + \ldots + a_{in}b_{nj} \quad \forall i, j$ (Обозначение: $A \cdot B$ или AB).*

Заметим, что произведение матриц некоммутативно, то есть $A \cdot B \neq B \cdot A$.

Свойства произведения матриц:

1. $(AB)C = A(BC)$;

2. $(A + B)C = AC + BC$;
   $C(A + B) = CA + CB$;

3. $\forall \alpha \in P \quad \alpha(AB) = A(\alpha B)$;

4. $\forall A = (a_{ij}) \in P_{m,n} \quad E_m \cdot A = A \cdot E_n = A$.

   ♦ Возьмем $A = (a_{ij}) \in P_{m,n}, E_m \in P_{m,m}$. Тогда $E_m A \in P_{m,n}$, следовательно, размер произведения совпадает с размером матрицы A. Обозначим за e_{ij} элементы матрицы E_m, тогда $e_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j; \end{cases}$. Отсюда получаем, что элементы произведения в i-ой строке и j-ом столбце равны $\sum_{s=1}^{n} e_{is}a_{sj} = e_{ii}a_{ij} = a_{ij}$. ⊠

5. $\forall A = (a_{ij}) \in P_{m,n}$
   $O_{k,m}A = O_{k,m}$;
   $AO_{n,k} = O_{m,k}$.

Из свойств сложения и умножения матриц следует, что множество квадратных матрицы $P_{n,n}$ одного порядка является ассоциативным кольцом с единицей относительно операций ссложения и умножения матриц. Это кольцо содержит делители нуля.

• *Матрица $C = (c_{ij}) \in P_{n,m}$ называется* ***транспонированной к матрице*** *$A = (a_{ij}) \in P_{m,n}$, если $c_{ij} = a_{ji} \quad \forall i, j$ (Обозначение: A^T).*

Свойства транспонирования матриц:

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

2. $(A+B)^T = A^T + B^T$;

3. $\forall \alpha \in P \quad (\alpha A)^T = \alpha A^T$;

4. $(AB)^T = B^T A^T$;

   ⧫ Так как $A \in P_{m,n}, B \in P_{n,k}$, то $AB \in P_{m,k}$, $(AB)^T \in P_{k,m}$, $A^T \in P_{n,m}$, $B^T \in P_{k,n}$, $B^T A^T \in P_{k,m}$.
   Пусть $A = (a_{ij})$, $B = (b_{ij})$, $AB = (c_{ij})$, $A^T = (a'_{ij})$, $B^T = (b'_{ij})$, $(AB)^T = (c'_{ij})$, $B^T A^T = (d_{ij})$.
   Тогда $d_{ij} = \sum_{n}^{s=1} b'_{is} a'_{sj} = \sum_{s=1}^{n} a_{js} b_{si} = c_{ji} = c'_{ij}$. Отсюда получаем, что $(AB)^T = B^T A^T$, так как равны соответствующие элементы и размеры матриц. ⊠

5. $(ABC)^T = C^T (AB)^T = C^T B^T A^T$.

• *Квадратная матрица называется* ***симметрической****, если $A = A^T (a_{ij} = a_{ji})$.*

9.3 Перестановки (Вводный).

• *Пусть $Z_n = \{1, 2, \ldots, n\} \subset \mathbb{Z}$. Конечная последовательность длины n, составленная из различных элементов множества Z_n называется* ***перестановкой множества*** *Z_n. (Обозначение: S_n — множество всех перестановок Z_n)*

Если a некоторый элемент множества S_n, то за a_i обозначается i-й элемент перестановки a, то есть $a = (a_1, \ldots, a_n)$.

Теорема. *Множество S_n состоит из $n!$ элементов.*

⧫ Так как в перестановке a каждый элемент a_i различен, то в качестве a_1 мы можем взять любой из n элементов множества Z_n. В качестве a_2 любой из оставшихся $n-1$ элементов множества Z_n и так далее. Итого количество возможных выборов чисел $a_1, \ldots, a_2 = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1 = n!$. ⊠

• *Пусть $a = (a_1, \ldots, a_n) \in S_n$ — некоторая перестановка. Говорят, что числа a_i и a_j образуют* ***инверсию****, если $i < j$ и $a_i > a_j$. Перестановка называется* ***четной****, если ее элементы составляют четное число инверсий, и* ***нечетной****, если нечетное число инверсий.*

• *Преобразование перестановки, при котором в перестановке меняются местами два элемента, при этом остальные элементы останутся на своих местах, называется* ***транспозицией****.*

Теорема. *Всякая транспозиция меняет чётность перестановки.*

⧫ Рассмотрим 2 случая: транспозицию рядом стоящих элементов и транспозицию элементов, не стоящих рядом.

1. Если поменять 2 стоящих рядом элементов местами, то количество инверсий изменится на 1, следовательно, чётность перестановки поменяется на противоположную.

2. Пусть между двумя элементами, которые мы меняем местами, находится s элементов. Тогда можно провести транспозицию этих элементов, меня элементы, стоящие рядом, местами, начиная первого из данных двух. Данная процедура будет проделана $2s+1$ раз (s раз первый элемент поменяется местами с теми, что лежат между, и 1 раз с тем, с которым мы хотели поменять его первоначально, потом второй s раз идёт на исходную позицию первого). Следовательно, так как при каждой перестановке по первому пункту чётность будет меняться $2s+1$ раз (нечётное число), то чётность перестановки так же поменяется.

⊠

Теорема. *При $n > 1$ множество S_n содержит равное количество чётных и нечётных перестановок ($\frac{n!}{2}$).*

⧫ Возьмём множество всех чётных перестановок и поменяем первые два элемента местами. Таким образом получим столько же нечётных перестановок. Значит количество нечётных перестановок не может быть меньше, чем чётных. Аналогичным образом поступаем с нечётными и получаем, что количество чётных так же не может быть меньше, чем нечётных. Следовательно, их количества равны. А так как всего перестановок $n!$, то число чётных и нечётных перестановок равно $\frac{n!}{2}$. ⊠

Теорема. *При $n > 1$ от любой перестановки из множества S_n можно перейти к любой другой перестановке из S_n применив несколько транспозиций.*

9.4 Определитель матрицы. Свойства определителей.

Пусть $A = (a_{ij}) \in P_{n,n}$ — квадратная матрица порядка n.

• ***Определителем матрицы*** *A называется элемент поля P*

$$detA = \sum_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n},$$

где $\varepsilon(\alpha)$ — количество инверсий в перестановке $\alpha = (\alpha_1, \dots, \alpha_n)$.
(Обозначения: $|A| = det(A)$.)

Из определения следует, что определитель треугольной матрицы равен произведению элементов главной диагонали матрицы, определитель единичной матрицы равен единице, нулевой — нулю.

Свойства определителя:

1. *Если в матрице A поменять местами две строки, то получим матрицу B, для которой $detB = -detA$.*

   ⧫ Пусть B получена из матрицы A путём перестановки s-й и k-й строки местами (для определённости пусть $s < k$). Тогда: $a_{sj} = b_{kj}$, $a_{kj} = b_{sj}$ $\forall s, k$; $a_{ij} = b_{ij}$ $\forall i \neq k,s$.
   $det(A) = \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots a_{s\alpha_S} \dots a_{k\alpha_k} \dots a_{n\alpha_n} =$
   $= \sum\limits_{\alpha \in S_n} (-1)(-1)^{\varepsilon((a_1,\dots,a_k,\dots,a_s,\dots,a_n))} b_{1\alpha_1} b_{2\alpha_2} \dots b_{s\alpha_s} \dots b_{k\alpha_k} \dots b_{n\alpha_n} = -det(B)$. ⊠

2. *Если матрица имеет две одинаковые строки, то её определитель равен 0.*

   ⧫ Пусть A имеет одинаковую s-ую и k-ую строки. Разобьём S_n на пары, перестановки которых отличны лишь элементами, стоящие на s-м и k-м месте. В каждой паре одна из перестановок чётная, другая нечётная, а произведения соответствующих этим перестановкам элементов будут равны $\Rightarrow$ сумма слагаемых определителя, соответствующих этим перестановкам $= 0 \Rightarrow det(A) = 0$. ⊠

3. *Если матрица B получена из матрицы A умножением элементов некоторой строки на $\lambda \in P$, то $detB = \lambda detA$.*

   ⧫ Пусть B получена из A умножением k-й строки на λ. Тогда:
   $a_{ij} = b_{ij}\, \forall i \neq k, b_{kj} = \lambda a_{kj} \Rightarrow det(B) = \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} b_{1\alpha_1} b_{2\alpha_2} \dots b_{k\alpha_k} \dots bn\alpha_n =$
   $= \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots \lambda a_{k\alpha_k} \dots a_{n\alpha_n} = \lambda \cdot det(A).$ ⊠

   Следствие. *Если матрица содержит нулевую строку, то её определитель равен нулю. ($detA = 0$).*

   ⧫ Элементы нулевой строки обнуляют слагаемые в сумме/элементы нулевой строки имеют общий множитель равный нулю. ⊠

   Следствие. *Если матрица имеет две пропорциональные строки, то её определитель равен нулю. ($detA = 0$).*

   ⧫ Если одну из этих строк умножить на коэффициент пропорциональности, то получим матрицу с одинаковыми строками, определитель которой по свойству 2 равен нулю. ⊠

   Следствие. $\forall A \in P_{n,n}\ det(\lambda A) = \lambda^n A$

   ⧫ Из доказательства третьего свойства определителя следует, что множитель λ будет у каждого элемента строки. ⊠

4. $$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a'_{k1} + a''_{k1} & \dots & a'_{kn} + a''_{kn} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a'_{k1} & \dots & a'_{kn} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a''_{k1} & \dots & a''_{kn} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

   ⧫ $det(A) = \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots (a_{k_1\alpha_{k_1}} + a_{k_2\alpha_{k_2}}) \dots a_{n\alpha_n} =$
   $= \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots a_{k_1\alpha_{k_1}} \dots a_{n\alpha_n} + \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots a_{k_2\alpha_{k_2}} \dots a_{n\alpha_n} =$
   $= det(A_1) + det(A_2).$ ⊠

5. *Определитель матрицы не изменится, если к элементам одной её строки прибавить соответствующие элементы другой строки, умноженные на произвольный элемент поля P.*

   ⧫ Пусть A' — матрица, полученная из матрицы A, добавлением к s-й строки k-ой, умноженную на λ, а B — матрица, у которой на месте s-й строки стоит k-я, умноженная на λ, тогда по свойству 4 определителя: $det(A') = det(A) + \lambda \cdot det(B) =$ [так как s-я и k-я строки у B совпадают, $det(B) = 0$] $= det(A) \Rightarrow$ определитель не изменится. ⊠

6. *Определитель матрицы при транспонировании не меняется:* $detA^T = detA$.

   ♦ Пусть $A = (a_{ij})$, $A^T = (a_{ij}) = (b_{ij})$. $det(A) = \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_n n}$.

   Попарными перестановками переупорядочим множители по первому индексу, получив множители $b_{1\alpha_1} b_{2\alpha_2} \dots b_{n\alpha_n}$ и перестановку $\beta = (\beta_1, \dots, \beta_n)$. Каждая перестановка множителей порождает транспозицию в перестановке, состоящей из индексов $\Rightarrow$ переупорядочивание элементов в произведении влечёт за собой цепочку транспозиций, переводящих перестановки.

   $a = (a_1, \dots, a_n) \to (1, \dots, n) \to \beta(\beta_1, \dots, \beta_n) \Rightarrow$ чётность перестановок a и β совпадает $\Rightarrow (-1)^{\varepsilon(a)} = (-1)^{\varepsilon(\beta)} \Rightarrow det(A) = \sum\limits_{\beta \in S_n} (-1)^{\varepsilon(\beta)} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_n n} = det(A^T)$. ⊠

9.5 Теорема Лапласа. Теорема о разложении определителя по элементам строки.

• *Рассмотрим матрицу* $A = (a_{ij}) \in P_{n,n}$. *Выберем в ней* k *различных строк с номерами* $i_1, \dots, i_k$ *и* k *различных столбцов с номерами* $j_1, \dots, j_k$. *Элементы матрицы* A*, стоящие на пересечении выбранных строк и столбцов составляют квадратную матрицу* k*-ого порядка. Определитель* M *этой матрицы называется* ***минором*** k***-ого порядка, расположенным в строках с номерами*** $i_1, \dots, i_k$ ***и столбцах с номерами*** $j_1, \dots, j_k$. *Элементы матрицы* A*, стоящие на пересечении оставшихся строк и столбцов, также образуюи квадратную матрицу порядка* $n - k$. *Определитель* M' *этой матрицы называется* ***дополнительным минором к минору M***.

• ***Алгебраическим дополнением минора*** M *называется произведение дополнительного к нему минора на* $(-1)^{S_M}$*, где* S_M *— сумма номеров строк и столбцов матрицы* A *в которых расположен минор* M*, то есть* $A_M = M' \cdot (-1)^{i_1 + \dots + i_k + j_1 + \dots + j_k}$.

Любой элемент a_{ij} матрицы A можно рассматривать как минор 1-ого порядка, расположенный в i-ой строке и j-ом столбце. Тогда дополнительный к нему минор M_{ij} - определитель матрицы, которая получается из матрицы A вычеркиванием i-ой строки и j-ого столбца, а алгебраическое дополнение элемента a_{ij} имеет значение $A_{ij} = (-1)^{i+j} M_{ij}$.

Теорема о разложении определителя матрицы по элементам строки. *Определитель матрицы равен суммы всех произведений элементов произвольной строки матрицы на их алгеибраические дополнения:*

$$detA = a_{k1}A_{k1} + \dots + a_{kn}A_{kn} \quad \forall k = 1, \dots, n.$$

♦ По четвертому свойству определителя матрицы представим определитель A в виде суммы n определителей, где у каждого j-го определителя все элементы в i-й строке и не j-м столбце равны нулю, а оставшийся элемент равен a_{ij}.

То есть $det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n-1} & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n-1} & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{i2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{vmatrix} + \dots +$

$+ \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n-1} & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{vmatrix}$.

Обозначим каждый определитель как $\Delta_1, \Delta_2, \ldots, \Delta_n$. Рассмотрим определитель Δ_j. Меняем по очереди строку i со строками $i+1, i+2, \ldots, n$, поставив её на последнее место, а затем столбец j со строками $j+1, \ldots, n$ поочередно, поставив его также на последнее место. Обозначим полученную матрицу за $B = (b_{ij}$. Определитель этой матрицы отличается от Δ_j на $(-1)^{n-i+n-j} = (-1)^{i+j}$. С другой стороны, так как в полученной матрице в n-ой строке расположен лишь один ненулевой элемент $= a_{ij}$, то $det(B) =$
$= \sum\limits_{\alpha \in S_n} (-1)^{\varepsilon(\alpha)} b_{1\alpha_1} \cdot \ldots \cdot b_{n-1\alpha_{n-1}} b_{n\alpha_n} = (\sum\limits_{\alpha \in S_{n-1}} (-1)^{\varepsilon(\alpha)} b_{1\alpha_1} \cdot \ldots \cdot b_{n-1\alpha_{n-1}}) b_{n\alpha_n} = a_{ij} M_{ij} =$
$= (-1)^{i+j} \Delta_j \Rightarrow det(A) = \sum\limits_{j=1}^{n} \Delta_j = \sum\limits_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}.$ ⊠

Теорема Лапласа. *Пусть $A \in P_{n,n}$. Выделим в матрице A любые k строк. Определитель матрицы A равен сумме всех произведений миноров k-ого порядка, расположенных в выделенных строках, на их алгебраические дополнения.*

Замечание. *Так как при транспонировании определитель матрицы не меняется, то утвержденияй, аналогичные теореме о разложении определителя матрицы по элементам строки и теореме Лапласа, можно сформулировать и для столбцов.*

9.6 Определитель блочнотреугольной матрицы. Определитель произведения матриц.

Лемма (следствие из теоремы Лапласа). *Определитель блочнотреугольной матрицы равен произведению определителей её диагональных блоков.*

♦ Рассмотрим матрицу $A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$, где $B \in P_{k,k}, C \in p_{n-k,n-k}$ — её диагональные блоки. Определитель матрицы A по теореме Лапласса разложим по первым k строкам: $det(A) = det(B)(-1)^{1+2+\ldots+k+1+\ldots+k} \cdot det(C) = det(B) \cdot det(C)$. Следствие будет доказано, если применить индукцию по количеству диагональных блоков. ⊠

Заметим, что так как при транспонировании определиель матрицы не изменяется, то следствие будет верно как и для верхней, так и для нижней блочнотреугольной матрицы.

Теорема. *Определитель произведения двух квадратных матриц равен произведению определителей множителей: $detAB = detA \cdot detB$.*

♦ Пусть $A = (a_{ij}), B = (b_{ij}) \in P_{n,n}$. Рассмотрим матрицу $D = \begin{pmatrix} A & 0_{n,n} \\ -E_n & B \end{pmatrix} \in P_{2n,2n}$. $det(D) = det(A) \cdot det(B)$ (по следствию из теоремы Лапласа). Преобразуем матрицу D: к первой строке прибавим $n+1$-ую строку, домноженную на a_{11}, ко второй строке $n+2$-ую строку, домноженную на a_{12} и так далее, к n-ой строке добавим $2n$-ую строку, домноженную на a_{1n}. В результате преобразования первые n элементов первой строки обнулятся, а остальные n элементов $c_{1i} = 0 + a_{11}b_{1i} + \ldots + a_{1n}b_{ni}$ являются элементами матрицы AB, стоящими в первой строке и в i-ом столбце. Аналогично преобразуем $n-1$ строку и получим матрицу вида $D' = \begin{pmatrix} 0_{n,n} & AB \\ -E_n & B \end{pmatrix}$, причём преобразования, с помощью которых из матрицы D была получена матрица D', не меняют определитель, следовательно, $det(D) = det(D')$. По теореме Лапласса вычислим определитель матрицы D', разложив его по первым n строкам: $det(D') = det(AB)(-1)^{1+2+\ldots+2n} det(-E_n) = det(AB)(-1)^{n(2n+1)}(-1)^n = det(AB)(-1)^{n(2n+2)} = det(AB) \Rightarrow det(A) \cdot det(B) = det(AB)$. ⊠

9.7 Обратная матрица.

Пусть $A = (\alpha_{ij}) \in P_{n,n}$ — квадратная матрица порядка n.

• *Матрица $X \in P_{n,n}$ называется* ***обратной*** *к матрице A, если $AX = XA = E_n$ (Обозначение: A^{-1}).*

• *Пусть A_{ij} — алгебраические дополнения элементов α_{ij}. Матрица* $\widetilde{A} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$ *называется* ***присоединенной*** *к матрице A.*

Лемма. $A \cdot \widetilde{A} = \widetilde{A} \cdot A = diag(detA, \dots, detA)$.

⧫ Пусть $A\widetilde{A} = (c_{ij})$. Тогда $c_{ij} = a_{i1} \cdot A_{i1} + \ldots + a_{in} \cdot A_{in}$. Это же значение имеет определитель матрицы, которая получена из матрицы A заменой элементов j-ой строки на элементы i-ой строки.

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{i1} \cdot A_{j1} + \ldots + a_{in} \cdot A_{jn} = \begin{cases} detA, \ i = j, \\ 0, \ i \neq j; \end{cases} \quad \Rightarrow A \cdot \widetilde{A} = diag(detA, \dots, detA)$$

⊠

• *Квадратная матрица называется* ***невырожденной****, если ее определитель отличен от нуля.*

Теорема. *Квадратная матрица имеет обратную $\Longleftrightarrow$ она невырожденная.*

⧫ $\Rightarrow$) Пусть матрица A имеет обратную матрицу A^{-1}. Тогда $A \cdot A^{-1} = E$ и $det(A \cdot A^{-1}) = detA \cdot detA^{-1} = detE = 1$. То есть $detA \neq 0$.

$\Leftarrow$) Пусть $detA \neq 0$. Построим матрицу $\dfrac{1}{detA}\widetilde{A}$. Тогда $A \cdot \dfrac{1}{detA}\widetilde{A} = \dfrac{1}{detA} A \cdot \widetilde{A} = \dfrac{1}{detA} detA \cdot E = E$. То есть матрица $\dfrac{1}{detA}\widetilde{A}$ обратная для матрицы A. ⊠

Свойства обратных матриц:

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

2. $(A^{-1})^T = (A^T)^{-1}$.

3. $detA^{-1} = (detA)^{-1}$.

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

⧫ Проведем доказательство для четвертого случая. Для остальных случае доказательство аналогично. $ABB^{-1}A^{-1} = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E \Rightarrow (AB)^{-1} = B^{-1}A^{-1}$. ⊠

9.8 Метод Крамера решения линейных невырожденных систем.

• ***Системой m линейных уравнений с n неизвестными*** $x_1, \ldots, x_n$ ***над полем P*** *называется система выражений вида*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m; \end{cases} \qquad a_{ij}, b_i \in P. \tag{9.8.1}$$

Элементы a_{ij} называются коэффициентами системы, элементы b_i — свободными членами. Если все свободные члены равны 0, то система называется ***однородной****, в противном случае —* ***неодонродной****.*

• *Последовательность $(\gamma_1, \ldots, \gamma_n) \in P^n$ элементов поля P называется* ***решением системы*** (9.8.1)*, если каждое уравнение обращается в верное равенство после замены неизвестных x_i элементами γ_i.*

• *Система называется* ***совместной****, если она имеет хотя бы одно решение. В противном случае система называется* ***несовместной****.*

• *Две системы линейных уравнений над одним полем P называются* ***равносильными****, если они имеют одно и то же множество решений.*

• *Матрица $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \in P_{m,n}$, составленная из коэффициентов системы, называется* ***матрицей системы****, столбец $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in P_{m,1}$ —* ***столбцом свободных членов****, матрица $\widetilde{A} = \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right) \in P_{m,n}$ —* ***расширенной матрицей системы****, столбец $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in P_{m,1}$ — столбцом неизвестных.*

Тогда систему можно записать в матричном виде как $AX = B$, а последовательность $(\gamma_1, \ldots, \gamma_n) \in P^n$ будет являться решением системы, если столбец $\Gamma = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \in P_{n,1}$ является решением матричного уравнения $AX = B$.

• *Система линейных уравнений называется* ***невырожденной****, если матрица ее системы квадратная и невырожденная.*

Теорема. *Невырожденная система n линейных уравнений с n неизвестными имеет единственное решение $(x_1, \ldots, x_n) = (\frac{\Delta_1}{\Delta}, \ldots, \frac{\Delta_n}{\Delta})$, где Δ — определитель матрицы A*

системы, Δ_i — определитель матрицы A_i, которая получается из матрицы A системы заменой i-го столбца столбцом свободных членов B.

♦ Разложим определитель Δ_i по элементам i-го столбца: $\Delta_i = b_1A_{1i} + \ldots + b_nA_{ni}$. Так как матрица A невырожденная, то существует обратная ей матрица A^{-1}. Домножим матричное уравнение $AX = B$ на матрицу A^{-1} слева и получим $A^{-1}AX = A^{-1}B \Longleftrightarrow EX = A^{-1}B \Longleftrightarrow X = A^{-1}B = \dfrac{1}{det(A)}\widetilde{A}B = \dfrac{1}{det(A)}\begin{pmatrix} A_{11} & \ldots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \ldots & A_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} =$

$$\begin{pmatrix} \dfrac{1}{det(A)}(b_1A_{11} + \ldots + b_nA_{1n}) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{1}{det(A)}(b_1A_{n1} + \ldots + b_nA_{nn}) \end{pmatrix} = \begin{pmatrix} \dfrac{\Delta_1}{\Delta} \\ \vdots \\ \dfrac{\Delta_n}{\Delta} \end{pmatrix}.$$ ⊠

Сделаем вывод: если матрица системы невырожденная, то система имеет единственное решение.

9.9 Метод Гаусса решения линейных систем.

Рассмотрим линейную систему уравнений над полем P

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

Метод Гаусса основан на двух принципах:

1. Множество решений системы не изменится, если одно из уравнений домножить на $a \in P\backslash\{0\}$;

2. Множество решений системы не изменится, если к одному из уравнений прибавить другое, домноженное на $a \in P$.

• *Такие преобразования называются* ***элементарными преобразованиями*** *линейной системы.*

Элементарные преобразования линейной системы не меняют множество решений.

Заметим, что с помощью преобразования 2 можно менять уравнения местами.

Если $a_{11} = 0$, то найдём уравнение, где коэффициент при x_1 ненулевой, и поменяем местами с первым. Если таких нет, то система явно не содержит переменную x_1. Следовательно, на месте x_1 может быть любой элемент поля P. Тогда рассмотрим систему как систему с $n-1$ неизвестными.

Пусть $a_{11} \neq 0$. Исключим из всех уравнений кроме первого x_1. Для этого из второго вычтем первое, домножив на $\frac{a_{21}}{a_{11}}$, из третьего первое, домножив на $\frac{a_{31}}{a_{11}}$ и так далее. В результате получим систему, где x_1 присутствует только в первом уравнении (и, возможно, x_2 и так далее вплоть до x_k).

Пусть k_1 — наименьший индекс переменной, присутствующей в какой-либо строке кроме первой. Поменяем эту строку со второй местами. Применим предыдущие операции относительно x_{k_1} ко всем строкам кроме первой.

Продолжим операции и остановимся на моменте, когда все последующие уравнения будут содержать коэффициенты 0, то есть на системе вида

$$\begin{cases} a_{11}x_1 + \ldots + a_{11}x_1 = b_1, \\ a_{2k_2}x_{k_2} + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{sk_s}x_{k_s} + \ldots + a_{sn}x_n = b_s, \\ 0 = b_{s+1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0 = b_m. \end{cases}$$

• *Такая система уравнений называется* ***ступенчатой****.*

Если полученная система имеет уравнения вида $0 = b_i$, где $b_i \neq 0$, то система несовместна. Если $b_i = 0$, то это уравнение вида $0 = 0$, которому удовлетворяют любые элементы поля P. Тогда, исключив такие уравнения из системы, получим систему равносильную исходной.

• *Незивестные $x_1, x_{k2}, \ldots, x_{ks}$ называются* ***главными*** *неизвестными, остальные —* ***свободными*** *неизвестными.*

Выразим все главные коэффициенты через свободные. Для этого из первых $s-1$ уравнений исключим переменную x_{ks}, из первых $s-2$ — x_{ks-1} и так далее. В итоге получим систему, которая в каждом уравнении имеет только одну неизвестную. Разделим каждое уравнение на коэффициент при главной неизвестной и получим решение системы: $x_{ks} = \frac{b_s}{a_{sk_s}}$; $x_{ks-1} = \frac{b_s - a_{sk_{s-1}}x_s}{a_{sk_{s-1}}}$ и так далее.

• *Матрица, из столбцов которой можно составить единичную, называется* ***частично-мономиальной****. Система, матрица которой является частично мономиальной, называется* ***частично-мономиальной****.*

Любая линейная система равносильна или несовместной системе, или частично-мономиальной, которые могут быть получены с помощью элементарных преобразований.

Элементарные преобразования удобно производить не над системой, а над её расширенной матрицей. Элементарные преобразования системы соответствуют следующим элементарным преобразованиям строк расширенной матрицы:

1. Умножение строки матрицы на ненулевой элемент поля P;
2. Прибавление к строке другой строки, домноженной на произвольный элемент поля P.

• *Матрицы называются* ***эквивалентными****, если одна из них может быть получена из другой путём конечного числа элементарных преобразований строк.*

Методом Гаусса можно вычислить обратную матрицу. Пусть $A \in P_{n,n}$. Рассмотрим матричное уравнение $AX = E$, которое имеет единственное решение $X = A^{-1}$. $X^{(i)}$ — i-й столбец матрицы X, который является решением уравнения $AX^{(i)} = E^{(i)}$ ($E^{(i)}$ — i-й столбец матрицы E). Следовательно, $X^{(i)}$ можно получить, решив систему с расширенной матрицей $(A|E^{(i)})$.

Заметим, что все рассматриваемые системы имеют одну и ту же левую часть, а все шаги алгоритма Гаусса определены именно в ней. Значит все системы можно решать одновременно: $(A|E) \sim (E|A^{-1})$.

Часть III

Линейная алгебра

Глава 10

Векторное пространство

10.1 Определение и простейшие свойства векторного пространства.

• *Пусть P — некоторое поле. Непустое множество V называется* ***векторным пространством*** *над полем P, если на V задана бинарная операция $V \times V \to V$, обычно называемая* ***сложением****, и операция $P \times V \to V$, называемая* ***умножением элемента множества V на элемент поля P****, для которых выполняются следующие условия:*

1. $(V, +)$ *— абелева группа.*

2. $\forall\ \alpha,\ \beta \in P$, *и* $\forall\ a,\ b \in V$:

   - $a \cdot 1 = a$
   - $(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$
   - $\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$
   - $\alpha \cdot (\beta \cdot a) = (\alpha \cdot \beta) \cdot a$

При этом элементы множества V называются ***векторами****, а элементы поля P —* ***скалярами****.*

• *Векторное пространство над полем $\mathbb{R}$ называется* ***действительным****, а над полем $\mathbb{C}$ —* ***комплексным****.*

Простейшие свойства векторных пространств:

1. *Свойства абелевой группы:*

   - В V $\exists!\ 0_v$ и $\forall a \in V\ \exists!\ (-a)$;
   - $a + x = b$ имеет единственное решение $b + (-a) = b - a = x$;
   - $a + c = b + c \Rightarrow a = b$.

2. $0_p \cdot a = 0_v$, $\forall\ a \in V$,
   $\alpha \cdot 0_v = 0_v$, $\forall\ \alpha \in P$.

   ♦ $0_pa + 0_pa = (0_p + 0_p)a = 0_pa + 0_v = 0_v$. ⊠

3. *Если* $\alpha \cdot a = 0_v$, *то* $\left[\begin{aligned} &a = 0_v, \\ &\alpha = 0_p. \end{aligned}\right.$

♦ Пусть $\alpha \in P, \neq 0_p \Rightarrow \exists\ \alpha^{-1} \in P : \alpha \cdot \alpha^{-1} = 1$
$a \cdot 1 = (\alpha^{-1} \cdot \alpha) \cdot a = \alpha^{-1} \cdot (\alpha \cdot a) = \alpha^{-1} \cdot 0_v = 0_v$ ⊠

4. $\alpha(-a) = (-\alpha)a = -(\alpha a)$.

   ♦ $\alpha(-a) + \alpha a = \alpha(a + (-a)) = \alpha 0_v = 0_v \Rightarrow (-\alpha a)$ является противоположным для (αa). ⊠

5. $\alpha(a - b) = \alpha a - \alpha b$;
   $a(\alpha - \beta) = \alpha a - \beta a$.

   ♦ $\alpha(a - b) = \alpha(a + (-b)) = \alpha a + \alpha(-b) = \alpha a + (-(\alpha b)) = \alpha a - \alpha b$. ⊠

6. $\forall \alpha, \beta \in P, \forall a \in V, \neq 0$
   Если $\alpha \cdot a = \beta \cdot a$, $a \neq 0_v$, *то* $\alpha = \beta$*;*
   Если $\alpha \cdot a = \alpha \cdot b$, $\alpha \neq 0_v$, *то* $a = b$.

   ♦ Пусть $\alpha a = \beta a \Rightarrow -\alpha a = -\beta a$.
   $0_v = \alpha a + (-\alpha a) = \alpha a + (-\beta a) = \alpha a - \beta a = (\alpha - \beta)a \Rightarrow \alpha - \beta = 0 \Rightarrow \alpha = \beta$ ⊠

10.2 Линейная зависимость и независимость системы векторов.

Пусть V — векторное пространство над полем P. Пусть $A(a_1, \dots, a_n) \in V$ — некоторая конечная система векторов (то есть упорядоченная последовательность необязательно различных элементов из V).

• *Пусть* $\alpha_1, \dots, \alpha_n \in P$*. Тогда вектор* $\alpha_1 a_1 + \ldots + \alpha_n a_n$ *называется* ***линейной комбинацией*** *векторов системы* A *с коэффициентами* $\alpha_1, \dots, \alpha_n$*. Множество всех линейных комбинаций векторов системы* A *называется* ***линейной оболочкой*** *системы* A *(Обозначение:* $L(A) = L(a_1, \dots, a_n)$*).*

• *Если некоторый вектор* b *является линейной комбинацией векторов системы* A*, то есть* $b \in L(A)$*, то говорят, что* b ***линейно выражается*** *через систему* A*.*
Если каждый вектор из $B(b_1, \dots, b_n)$ *линейно выражается через систему , то система* B ***линейно выражается*** *через систему* A*.*

• *Если система* A *линейно выражается через систему* B *и наоборот,* B *через* A*, то системы* A *и* B ***эквивалентны****. Обозначение:* $A \sim B$*.*

Лемма. *Если система векторов* $B(b_1, \dots, b_n)$ *линейно выражается через систему* $A(a_1, \dots, a_n)$*, то* $L(B) \subseteq L(A)$*.*

♦ Пусть b — произвольный вектор из $L(B)$. Тогда b представим в виде: $b = \beta_1 b_1 + \beta_2 b_2 + \dots + \beta_k b_k$. Так как B линейно выражается через A, то каждый вектор из B представим в виде:

$$\begin{cases} b_1 = \alpha_{11}a_1 + \ldots + \alpha_{1n}a_n, \\ b_2 = \alpha_{21}a_1 + \ldots + \alpha_{2n}a_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_k = \alpha_{k1}a_1 + \ldots + \alpha_{kn}a_n; \end{cases}$$

Следовательно, подставим: $b = \beta_1(\alpha_{11}a_1 + \ldots + \alpha_{1n}a_n) + \beta_2(\alpha_{21}a_1 + \ldots + \alpha_{2n}a_n) + \ldots + \beta_k(\alpha_{k1}a_1 + \ldots + \alpha_{kn}a_n) = a_1(\alpha_{11}\beta_1 + \ldots + \alpha_{k1}\beta_k) + a_2(\alpha_{12}\beta_1 + \ldots + \alpha_{k2}\beta_k) + \ldots + a_n(\alpha_{1n}\beta_1 + \ldots + \alpha_{kn}\beta_k)$ — линейная комбинация $\Rightarrow b \in L(A) \Rightarrow L(B) \subseteq L(A)$. ⊠

Следствие. *Системы $A \sim B \iff L(A) = L(B)$.*

Следствие. *Если вектор $b \in A$ линейно выражается через остальные векторы системы A, то $L(A) = L(A \setminus \{b\})$.*

• *Линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_n a_n$ называется **тривиальной**, если все коэффициенты $\alpha_i = 0$, иначе **нетривиальной**.*

Тривиальная комбинация любых векторов всегда равна нулевому вектору.

• *Если существует нетривиальная линейная комбинация векторов системы A, равная нулевому вектору, то такая система векторов называется **линейно зависимой**.*

То есть конечная система векторов A является линейно зависимой, если существуют скаляры $\alpha_1, \ldots, \alpha_n$, не обращающиеся в нуль одновременно, такие, что $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0_v$.

• *Если только тривиальная линейная комбинация векторов равна нулевому вектору, то система векторов называется **линейно независимой**.*

То есть система линейно независима, если из $\alpha_1 a_1 + \ldots + \alpha_n a_n$ следует, что $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0_p$.

Свойства линейной зависимости и независимости системы векторов:

1. *Если система векторов линейно независима, то любая её подсистема также линейно независима. Если какая-либо подсистема системы векторов линейно зависимая, то и вся система также линейно зависима.*

   ⧫ Пусть подсистема $B_1(a_1, \ldots, a_k)$ системы $A(a_1, \ldots, a_n)$ линейно зависима. Тогда существует нетривиальная линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_k a_k = 0_v$. Следовательно, линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_k a_k + 0 \cdot a_{k+1} + \ldots + 0 \cdot a_n$ является нетривиальной и равной нулевому вектору. ⊠

2. (a) *Система, состоящая из одного вектора, линейно зависимая $\iff$ этот вектор нулевой.*

   (b) *Система, состоящая более чем из одного вектора, линейно зависима $\iff$ один из векторов этой системы линейно выражается через* другие.

   ⧫

   (a) $\Rightarrow$) Пусть система линейно зависима. Тогда $\exists\ \alpha \in P, \neq 0 : \alpha a = 0_v \Rightarrow a = 0_v$.
   $\Leftarrow$) Пусть $A = (0_v) \Rightarrow \forall \alpha_1 \in P, \neq 0\ \alpha_1 0_v = 0_v \Rightarrow$ система линейно зависимая.

(b) $\Rightarrow$) Пусть $A = (a_1, \ldots, a_n)$ линейно зависима. То есть существует нетривиальная линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0_v$. Пусть $\exists \alpha_k \in P, \neq 0 \Rightarrow \exists \alpha_k^{-1} : a_k = -\dfrac{\alpha_1}{\alpha_k} a_1 - \ldots - \dfrac{\alpha_{k-1}}{\alpha_k} a_{k-1} - \dfrac{\alpha_{k+1}}{\alpha_k} a_{k+1} - \ldots - \dfrac{\alpha_n}{\alpha_k} a_n \Rightarrow a_k$ линейно выражается через остальные векторы.
$\Leftarrow$) Пусть a_k линейно выражается через остальные векторы системы A. Тогда a_k имеет вид:
$a_k = \alpha_1 a_1 + \ldots + \alpha_{k-1} a_{k-1} + \alpha_{k+1} a_{k+1} + \ldots + \alpha_n a_n$. Тогда линейная комбинация $\alpha_1 a_1 + \ldots + (-1) \cdot a_k + \ldots + \alpha_n a_n$ является нетривиальной и равной 0_v. Следовательно, A линейно зависимая.

⊠

Следствие. *Система, содержащая нулевой вектор, всегда линейно зависимая.*

Следствие. *Система, содержащая равные векторы, всегда линейно зависима.*

3. *Если система $A(a_1, \ldots, a_n)$ линейно независимая, а система векторов $B(a_1, \ldots, a_n, b)$ линейно зависимая, то вектор b линейно выражается через систему A.*

   ⧫ Так как B линейно зависима, то $\exists\ \alpha_1, \ldots, \alpha_n, \beta$, не обращающиеся в нуль одновременно, такие, что $\alpha_1 a_1 + \ldots + \alpha_n a_n + \beta b = 0_v$. Если $\beta = 0$, то линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0_v$ является нетривиальной. Следовательно, то, что система A линейно зависимая, — противоречие. Значит $\beta \neq 0 \Rightarrow b = \dfrac{\alpha_1}{\beta} a_1 + \ldots + \dfrac{\alpha_n}{\beta} a_n$. ⊠

10.3 Базис и размерность векторных пространств.

• *Бесконечная система векторов называется **линейно независимой**, если независима любая её конечная подсистема, и **линейно зависимой**, если существует конечная зависимая подсистема. Векторное пространство называется **бесконечномерным**, если в нём существует бесконечная линейно независимая система векторов. И **конечномерным**, если в ней все линейно независимые системы векторов конечны.*

• *Система векторов $B(b_1, \ldots, b_n)$ векторного пространства V называется **базисом пространства** V, если:*

- *B — линейно независимая система.*
- *Линейная оболочка $L(B) = V$, то есть любой вектор из V линейно выражается через систему B.*

В курсе линейной алгебры изучаются только конечномерные векторные пространства.

Теорема. *В конечномерном пространстве линейно независимая система векторов либо является базисом, либо может быть дополнена до базиса.*

⧫ Пусть $A(a_1, \ldots, a_n)$ — линейно независимая система векторов. Если A не базис, то $\exists$ вектор b, который не выражается через A.
Рассмотрим систему $(a_1, \ldots, a_n, b)$. Если она линейно зависимая, то по свойствам линейной зависимости b выражается через A, что является противоречием. Следовательно, $(a_1, \ldots, a_n, b)$ линейно независимая.
Если система A не является базисом, то существует такой вектор, присоединив который к этой системе, мы получим линейно независимую систему векторов. Так как пространство

конечномерно, то бесконечно присоединять векторы к системе с сохранением линейной независимости невозможно. Следовательно, через конечное число шагов процесс остановится, и мы получим линейно независимую систему, через которую выражается любой вектор пространства, то есть базис. При этом A — его подсистема. ⊠

Следствие. *В ненулевом конечномерном векторном пространстве существует конечный базис.*

♦ Если $a \neq 0_v$, то система векторов A линейно независима и может быть дополнена до базиса. ⊠

Теорема. *Все базисы ненулевого конечномерного векторного пространства состоят из одного и того же числа векторов.*

♦ Пусть $A(a_1,\ \dots,\ a_k)$, $B(b_1, \dots, b_n)$ — базисы пространства V и пусть $k \geqslant n$.

Рассмотрим систему $(a_1, b_1, \dots, b_n)$. Так как B — базис, то вектор a_1 линейно выражается через B. Тогда $L(a_1, b_1, \dots, b_n) = L(B) = V$.
Так как a_1 линейно выражается через B, то $(a_1, b_1, \dots, b_n)$ линейно зависима, то есть существует нетривиальная линейная комбинация векторов этой системы равная нулевому вектору: $\alpha_1 a_1 + \beta_1 b_1 + \ldots + \beta_n b_n = 0_v$. Если все $b_i = 0$, то $\alpha_1 a_1 = 0_v$. А так как a_1 — вектор из базиса (значит ненулевой), то $\alpha_1 = 0$. Получаем противоречие с нетривиальностью.
Следовательно, $\exists\ \beta_i \neq 0$. Пусть $\beta_1 \neq 0$ (иначе β_i можно перенумеровать). Тогда b_1 линейно выражается через $(a_1, b_2, \dots, b_n) \Rightarrow L(a_1, b_1, b_2, \dots, b_n) = L(a_1, b_2, \dots, b_n) = V$.
Продолжая подобные рассуждения ещё $(n-1)$ раз, получим, что $L(a_1, b_2, \dots, b_n) =$
$L(a_1, a_2, \dots, b_n) = \ldots = L(a_1, a_2, \dots, a_n) = V \Rightarrow$ векторы $a_{n+1}, \dots, a_k$ линейно выражаются через векторы $a_1, \dots, a_n \Rightarrow$ система векторов A линейно зависимая, что является противоречием с тем, что система A — базис $\Rightarrow k = n$. ⊠

• *Количество векторов в базисе конечномерного векторного пространства называется* ***размерностью векторного пространства*** *(Обозначение:* $\dim V$*). Векторное пространство размерности* n *называется* ***n-мерным****. Нулевое пространство принято считать* ***нульмерным****.*

Теорема. *Пусть* V *—* n*-мерное векторное пространство. Тогда*

1. *система векторов, состоящая более чем из* n *векторов, линейно заивисима.*
2. *линейно независимая система векторов, состоящая из* n *векторов, является базисом.*
3. *линейно независимая система векторов, состоящая менее чем из* n *векторов, может быть дополнена до базиса.*

♦ Любая линейно независимая система векторов либо базис, либо может быть дополнена до базиса, но в n-мерном векторном пространстве все базисы состоят из n векторов. Следовательно:

1. в пространстве V нет линейно независимых систем векторов, состоящих более чем из n векторов.
2. если система содержит n векторов, то любое её дополнение будет линейно зависимым.
3. линейно независимая система векторов, состоящая из менее чем n векторов, не является базисом.

⊠

Из этой теоремы следует, что базис — максимально независимая система векторов.

10.4 Координаты вектора. Изоморфизм векторных пространств.

Пусть V — n-мерное векторное пространство над полем P, $A(a_1, \dots, a_n)$ — его базис. Тогда произвольный вектор $b \in V$ линейно выражается через систему A, то есть представим в виде линейной комбинации векторов системы A:

$$b = \alpha_1 a_1 + \ldots + \alpha_n a_n,\ \alpha_1, \ldots, \alpha_n \in P.$$

• *Представление вектора b в виде линейной комбинации векторов базиса A называется* ***разложением*** *вектора b по базису A, коэффициенты разложения $\alpha_1, \ldots, \alpha_n$ называются* ***координатами*** *вектора b в базисе A.*

Так как A — упорядоченная система векторов, то координаты вектора b в базисе A являются элементами пространства P^n.

• *Столбец, составленный из координат вектора называется* ***координатным столбцом.***

Обозначим $X_b = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ координатный столбец вектора b в базисе A. Тогда разложение $b = \alpha_1 a_1 + \ldots + \alpha_n a_n$ может быть записано в виде: $b = AX_b$.

Теорема. *Координаты вектора в заданном базисе определяются однозначно.*

⧫ Пусть вектор b имеет 2 разложения в базисе A: $b = \alpha_1 a_1 + \ldots + \alpha_n a_n = \beta_1 a_1 + \ldots + \beta_n a_n$. Отнимем от первого разложения второе и получим: $0_v = (\alpha_1 - \beta_1)a_1 + \ldots + (\alpha_n - \beta_n)a_n$. Так как система A является базисом, то A линейно независима. Следовательно, она допускает лишь тривиальную линейную комбинацию: $\alpha_i - \beta_i = 0 \Rightarrow \alpha_i = \beta_i$ ⊠

Теорема. *Координатный столбец сумыы векторов равен сумме координатных стобцов слагаемых. При умножении вектора на скаляр его координатный столбец умножается на этот скаляр: $\forall a, b \in V$, $\forall \alpha \in P$*

$$X_{a+b} = X_a + X_b$$
$$X_{\alpha a} = \alpha X_a.$$

⧫ Пусть $X_a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, $X_b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ — координатные столбцы векторов a и b в базисе A.

Тогда $\begin{cases} a = \alpha_1 a_1 + \ldots + \alpha_n a_n \\ b = \beta_1 b_1 + \ldots + \beta_n b_n \end{cases} \Rightarrow \begin{cases} a + b = (\alpha_1 + \beta_1)a_1 + \ldots + (\alpha_n + \beta_n)a_n \\ \alpha a = \alpha\alpha_1 a_1 + \ldots + \alpha\alpha_n a_n \end{cases} \Rightarrow$

$X_{a+b} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} = X_a + X_b;$

$X_{\alpha a} = \begin{pmatrix} \alpha\alpha_1 \\ \vdots \\ \alpha\alpha_n \end{pmatrix} = \alpha X_a.$ ⊠

Следствие. *Векторы линейно зависимы $\Longleftrightarrow$ линейно зависимы их координатные столбцы.*

♦ Если существует нетривиальная линейная комбинация векторов, равная 0_v, то линейная комбинация координатных столбцов с теми же коэффициентами является нетривиальной и равной нулевому столбцу. ⊠

Изоморфизм

• *Пусть V, U — векторные пространства над одним полем P. Отображение $\varphi : V \to U$ называется изоморфизмом пространства V на пространство U, если*

1. φ — *биекция*

2. $\varphi(a+b) = \varphi(a) + \varphi(b), \quad \forall\ a, b \in V$
   $\varphi(\alpha a) = \alpha\varphi(a), \quad \forall\ a \in V,\ \alpha \in P.$

• *Если существует изоморфизм пространства V на пространство U, то V называется* ***изоморфным*** *пространству U (Обозначение: $V \cong U$).*

Свойство: *Бинарное отношение изоморфности на множестве векторных пространств является отношением эквивалентности.*

Лемма. *Если* $\dim V = n \geqslant 1$*, то* $V \cong P^n$*.*

♦ Из первой и второй теорем параграфа следует, что отображение, ставящее в соответствие каждому вектору его координатный столбец в некотором базисе, является изоморфизмом. ⊠

Следствие (Критерий изоморфности). *Конечномерные векторные пространства над одним и тем же полем изоморфны $\Longleftrightarrow$ их размерности равны.*

♦ $\Rightarrow$) Пусть $V \cong U$. Тогда $\exists\ \varphi : V \to U$. Базис V — линейно независимая система векторов. Изоморфность φ его отображает в линейно независимую систему U. Следовательно, число векторов в базисе U не меньше, чем в базисе V, то есть $\dim V \leqslant \dim U$. Обратное утверждение следует из того, что отображение φ^{-1} также является изоморфизмом.

$\Leftarrow$) Если $\dim U = \dim V = n$, то по лемме $U \cong P^n$ и $V \cong P^n \Rightarrow U \cong V$ по свойству изоморфизма. ⊠

10.5 Подпространство.

Пусть V — векторное пространство над полем P.

• *Подмножество U пространства V называется* ***подпространством*** *пространства V, если оно само является векторным пространством над тем же полем и относительно тех же операций, что и V.*

Теорема. *Непустое подмножество U векторного пространства V является подпространством $\Longleftrightarrow$ в U определены те же операции, что и в V, то есть:*

1. $a + b \in U,\ \forall\ a, b \in U$

2. $\alpha a \in U,\ \forall\ a \in U,\ \alpha \in P$

♦ $\Rightarrow$) Если U — подпространство V, то оно само является векторным пространством, а значит в нём операции определены.
$\Leftarrow$) Пусть в U определенены линейные операции. Тогда $\forall a \in U \ \exists(-a) = (-1) \cdot a \Rightarrow (-a) \in U$, а так как в U определены операции сложения, то $(U, +)$ — подгруппа $V \Rightarrow (U, +)$ — абелева группа.
Вторая группа аксиом определения векторного пространства справедлива для векторов из U, так как она справедлива для всех векторов из V. ⊠

Следствие. *Для любого подпространства U векторного пространства V $0_v \in U$*

♦ Так как U — подпространство, то U непустое множество. Следовательно, $\exists\ a \in U : 0_v = 0_p \cdot a \in U$ ⊠

Следствие. *Для любого непустого множества $A \subset V$ линейная оболочка $L(A)$ является подпространством.*

♦ Если $a, b \in L(A)$, то
$$\begin{cases} a = \alpha_1 a_1 + \ldots + \alpha_n a_n \\ b = \beta_1 a_1 + \ldots + \beta_n a_n \end{cases} \Rightarrow a + b = (\alpha_1 + \beta_1)a_1 + \ldots + (\alpha_n + \beta_n)a_n;$$
$\alpha a = \alpha\alpha_1 a_1 + \ldots + \alpha\alpha_n a_n$. Следовательно, $a + b,\ \alpha a \in L(A)$ ⊠

Следствие. *Размерность линейной оболочки $L(A)$ системы векторов A равна максимальному числу линейнонезависимых векторов в системе A*

♦ Пусть $A(a_1, \ldots, a_n)$ имеет k линейно независимых векторов $a_1, \ldots, a_k$, а векторы $a_{k+1}, \ldots, a_n$ линейно заивисимые. Тогда $a_{k+1}, \ldots, a_n$ линейно выражаются через $a_1, \ldots, a_k$. Следовательно, $L(A)$ совпадает с $L(a_1, \ldots, a_k)$, а значит все векторы из $L(A)$ линейно выражаются через $a_1, \ldots, a_k$, то есть $a_1, \ldots, a_k$ — линейно независимая подсистема, через которую выражаются все векторы из $L(A)$. Из этого можно сделать вывод, что подсистема $a_1, \ldots, a_k$ — базис $L(A)$. Следовательно $\dim L(A) = k$. Если A линейно независима, то она сама является линейно независимой системой, через которую выражаются все векторы из $L(A)$, то есть базисом. ⊠

Теорема о монотонности размерности. *Для любого подпространства U n-мерного векторного пространства V справедливо неравенство* $\dim U \leqslant n$, *причем если* $\dim U = n$, *то* $U = V$

♦ Базис $A(a_1, \ldots, a_k)$ подпространства U является линейно независимой системой векторов пространства V. Следовательно, $k = \dim U \leqslant n$, так как V не может быть линейно независимой системой, содержащей векторов больше, чем n. Если $\dim U = n$, то A — базис V, а значит $V = L(A) = U$ ⊠

10.6 Ранг системы векторов.

• *Пусть A — некоторая система векторов, B — её подсистема. Система B называется* ***базисом системы*** *A, если*

1. *система B линейно независима;*
2. *каждый вектор системы A линейно выражается через подсистему B.*

Свойства базиса систем векторов:

1. *Если B — базис системы векторов A, то $B \sim A$.*

   ⧫ Система B линейно выражается через систему A, так как B — подсистема A. В свою очередь, A линейно выражается через B, так как B — базис A. ⊠

2. *Если B — базис системы векторов A, то $L(A) = L(B)$.*

3. *Базис системы векторов является базисом линейной оболочки этой системы векторов.*

4. *Все базисы системы векторов состоят из одного и того же числа векторов.*

• *Число векторов в базисе системы векторов называется* ***рангом*** *системы векторов (Обозначение:* $\operatorname{rank} A$*). Если система содержит лишь нулевые векторы, то ранг системы будем считать равным нулю.*

Свойства ранга системы векторов:

1. *Для любой системы векторов* $\operatorname{rank} A = \dim L(A)$.

   ⧫ По третьему свойству базисов системы векторов. ⊠

2. *Если система векторов B линейно выражается через систему векторов A, то* $\operatorname{rank} B \leqslant \operatorname{rank} A$.

   ⧫ Если система векторов B линейно выражается через систему A, то $L(B)$ содержится в $L(A)$. Тогда, по теореме о монотонности размерности, $\dim L(B) \leqslant \dim L(A)$. Следовательно, $\operatorname{rank} B \leqslant \operatorname{rank} A$ ⊠

• ***Элементарными преобразованиями*** *системы векторов называются:*

1. *Домножение любого вектора системы на ненул. скаляр.*

2. *Сложение двух векторов системы, один из которых домножен на скаляр.*

Теорема. *Элементарные преобразования системы векторов не меняют её ранг.*

⧫ Пусть A — система векторов, и пусть система B получена из A заменой некоторого вектора a_1 на $\alpha a_1 \neq 0$, а система C — заменой a_1 на $a_1 + \alpha a_2$. Система A линейно выражается через B, поскольку $a_1 = \alpha^{-1}(\alpha a_1)$, а B — через A. Следовательно, эти системы эквивалентны и их ранги равны.
Система C линейно выражается через A, а A — через C, поскольку $a_1 = (a_1 + \alpha a_2) + (-\alpha) a_2$. Следовательно, A и C эквивалентны и их ранги равны. ⊠

10.7 Ранг матрицы.

Пусть $A = (\alpha_{ij}) \in P_{m,n}$ — $m \times n$-матрица над полем P, то есть

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

Обозначим $A_i = \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{mi} \end{pmatrix}$ — i-ый столбец матрицы A.

• ***Рангом*** *матрицы A называется ранг системы столбцов матрицы A, то есть* $\operatorname{rank} A = \operatorname{rank}(A_1, \ldots, A_n)$

• ***Базисным минором*** *матрицы A назывваается такой её минор M, для которого:*

1. $M \neq 0$;
2. *все миноры, порядок которых больше порядка минора M, равны нулю.*

Теорема о базисном миноре. *Столбцы матрицы, в которых расположен базисный минор, образуют базис системы столбцов матриц.*

⧫ Пусть M — базисный минор системы A. Так как порядок векторов в системе не влияет на зависимость, переупорядочим столбцы так, чтобы M распологался в первых k-столбцах. Возьмём строки $i_1, \ldots, i_k$. Таким образом M имеет вид:

$$M = \begin{vmatrix} \alpha_{i_1 1} & \ldots & \alpha_{i_1 k} \\ \vdots & \ddots & \vdots \\ \alpha_{i_k 1} & \ldots & \alpha_{i_k k} \end{vmatrix}$$

Покажем, что первые k столбцов A образуют базис системы столбцов: 1) Покажем линейную независимость от противного. Пусть система $(A_1, \ldots, A_k)$ линейно зависима. Тогда $\exists$ нетривиальная линейная комбинация $\beta_1 A_1 + \ldots + \beta_k A_k = 0_v$. Распишем

$$\begin{cases} \beta_1 \alpha_{11} + \ldots + \beta_n \alpha_{1k} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_1 \alpha_{m1} + \ldots + \beta_n \alpha_{mk} = 0; \end{cases}$$

Выберем из этой системы уравнения с номерами $i_1, \ldots, i_k$. Получим новую систему, матрица которой квадратная, а ее определитель не равен нулю. Следовательно, по принципу Крамера, $\beta_i = 0 \ \forall i \Rightarrow$ получаем противоречие с нетривиальностью исходной линейной комбинации $\Rightarrow$ столбцы $(A_1, \ldots, A_k)$ линейно независимы.

Покажем, что $A_p = A_{k+1}, \ldots, A_n, \ p = k+1, \ldots, n$ линейно выражается через $A_1, \ldots, A_k$. Для этого рассмотрим определители вида

$$M = \begin{vmatrix} \alpha_{i_1 1} & \ldots & \alpha_{i_1 k} & \alpha_{i_1 p} \\ \vdots & \ddots & \vdots & \vdots \\ \alpha_{i_k 1} & \ldots & \alpha_{i_k k} & \alpha_{i_k p} \\ \alpha_{s1} & \ldots & \alpha_{sk} & \alpha_{sp} \end{vmatrix} = 0,$$

которые получаются в ходе окаймления минора M элементами s-ой строки и p-ого столбца. Если $s \in \{i_1, \ldots, i_k\}$, то $det M_s = 0$, так как содержит две одинаковые строки.
Если $s \notin \{i_1, \ldots, i_k\}$, то M_s — минор порядка $k+1$, то есть порядка большего чем порядок базисного минора. Следовательно, $M_s = 0 \ \forall s$.

Разложим определители M_s по элементам последней строки. Заметим, что алгебраические дополнения элементов последней строки не зависят от выбора s в определителе M_s. Обозначим их через $D_1, \ldots, D_k, D_{k+1}$, причем $D_{k+1} = M \neq 0$. Тогда $\alpha_{s1} D_1 + \ldots + \alpha_{sk} D_k + \alpha_{sp} M = 0 = M_s$. Следовательно $A_1 D_1 + \ldots + A_k D_k + A_p M = 0 \Rightarrow A_p = -\frac{1}{M}(A_1 D_1 + \ldots + A_k D_k)$. То есть столбцы A_p линейно выражаются через столбцы $A_1, \ldots, A_k$. И, следовательно, $(A_1, \ldots, A_k)$ — базис системы столбцов матрицы A. ⊠

Следствие. *Ранг матрицы равен порядку ее базисного минора.*

♦ Количество векторов в базисе системы столбцов совпадает с порядком базисного минора. ⊠

Из доказательства теоремы следует, что ранг матрицы можно вычислить с помощью метода окаймляющих миноров.

Следствие. *Ранг матрицы не меняется при транспонировании и элементарных преобразованиях строк и столбцов.*

♦ При транспонировании и при элементарных преобразованиях определитель матрицы не изменяется. ⊠

Следствие (Критерий равенства определителя нулю). *Определитель квадратной матрицы равен нулю $\iff$ ее столбцы линейно зависимы.*

♦ $\Rightarrow$) Пусть $detA = 0$. Тогда этот определитель не является базисным минором матрицы $A \Rightarrow$ порядок базисного минора меньше порядка матрицы, следовательно, количество векторов в базисе системы столбцов меньше, чем количество столбцов в матрице.

$\Leftarrow$) Если ранг матрицы A меньше ее порядка, то определитель матрицы A не является базисным минором, следовательно, он равен нулю. ⊠

Следствие. *Элементарные преобразования строк и столбцов матрицы не меняют ее ранга.*

Теорема. *Ранг произведения матриц не превосходит ранга каждого из сомножителей. Если один из сомножителей — невырожденная матрица, то ранг произведения матриц равен рангу второго сомножителя.*

♦ Пусть матрица $A = (\alpha_{ij}) \in P_{m,n}$, $B = (\beta ij) \in P_{n,k}$ и $AB = (\gamma_{ij}) \in P_{m,k}$; Тогда s-ый столбец матрицы AB имеет вид

$$\begin{pmatrix} \gamma_{1s} \\ \vdots \\ \gamma_{ms} \end{pmatrix} = \begin{pmatrix} \alpha_{11}\beta_{1s} + \ldots + \alpha_{1n}\beta_{ns} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}\beta_{1s} + \ldots + \alpha_{mn}\beta_{ns} \end{pmatrix} = \beta_{1s} \underbrace{\begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{pmatrix}}_{A_1} + \ldots + \beta_{ns} \underbrace{\begin{pmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{pmatrix}}_{A_n}$$ — линейная комбинация столбцов матрицы A. Следовательно все столбцы матрицы AB линейно выражаются через столбцы матрицы A. Следовательно, $\operatorname{rank}(AB) \leqslant \operatorname{rank} A$
С другой стороны, $\operatorname{rank}(AB) = \operatorname{rank}(AB)^T = \operatorname{rank}(B^T A^T) \leqslant \operatorname{rank} B^T = \operatorname{rank} B$.

Если матрица A невырожденная, то существует обратная матрица $A^{-1} : B = A^{-1}(AB) \Rightarrow \operatorname{rank} B \leqslant \operatorname{rank}(AB) \Rightarrow \operatorname{rank}(AB) = \operatorname{rank} B$. ⊠

10.8 Матрица перехода от базиса к системе векторов. Преобразование координат векторов.

Пусть V — n-мерное векторное пространство над полем P, $A = (a_1, \ldots, a_n)$ — базис пространства V, $B = (b_1, \ldots, b_n)$ — некоторая система векторов. Разложим вектор b_i системы

B по базису $A: b_i = \alpha_{1i}a_1 + \ldots + \alpha_{ni}a_n$. Обозначим через $B_i = \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{ni} \end{pmatrix}$ — координатные столбцы векторов b_i в базисе A.

• *Матрица $S_{A\to B} = [B_1, \ldots, B_k]$, составленная из координатных столбцов векторов системы B в базисе A, называется* ***матрицей перехода*** *от базиса A к системе векторов B.*

Если системы векторов A и B формально записать в виде однострочных матриц $\overline{A} = (a_1, \ldots, a_n), \overline{B} = (b_1, \ldots, b_n)$, то $\overline{B} = \overline{A}S_{A\to B}$

Свойства матриц перехода:

1. $\operatorname{rank} S_{A\to B} = \operatorname{rank} B$.

   ♦ Система векторов линейно независима тогда и только тогда, когда линейно независима система координатных столбцов этих векторов. Следовательно, координатные столбцы векторов базисы системы B образуют базис системы координатных столбцов векторов системы B. Но матрица $S_{A\to B}$ и состоит из этих координатных столбцов. Значит, $\operatorname{rank} S_{A\to B} = \operatorname{rank} B$. ⊠

2. *Система векторов B является базисом пространства $V \longleftrightarrow$ матрица перехода от некоторого базиса A пространства V к системе B квадратная и невырожденная.*

   ♦ $\Rightarrow$) Если система B — базис пространства V, то количество векторов в нем совпадает с количеством векторов в базисе A. Следовательно, $S_{A\to B}$ является квадратной. Так как B — линейно независимая система векторов, то координатные столбцы векторов системы B также линейно независимы. Следовательно, $\operatorname{rank} S_{A\to B} = n \Rightarrow detS_{A\to B} \neq 0$ по следствию из теоремы о базисном миноре.

   $\Leftarrow$) Пусть $S_{A\to B}$ невырожденная и квадратная. Она квадратная, значит, количество векторов в A и B одинаково. А так как она невырожденная, то $\operatorname{rank} S_{A\to B} = n \Rightarrow$ столбцы $S_{A\to B}$ линейно независимы. Следовательно, B линейно независима и является базисом пространства. ⊠

3. *Пусть A, B, C — базисы пространства V, $S_{A\to B}$ $S_{A\to C}$ $S_{B\to C}$ — матрицы перехода. Тогда*

   (a) $S_{B\to A} = (S_{A\to B})^{-1}$

   (b) $S_{A\to C} = S_{A\to B}S_{B\to C}$

   ♦

   (a) Из определения матрицы перехода следует, что $\overline{B} = \overline{A}S_{A\to B}$. Так как $S_{A\to B}$ — матрица перехода от базиса A к базису B, то она невырожденная. Следовательно, $\exists(S_{A\to B})^{-1} : \overline{B}(S_{A\to B})^{-1} = \overline{A}$. Столбцы $(S_{A\to B})^{-1}$ — координатные столбцы векторов системы A в базисе B. Тогда матрицы $(S_{A\to B})^{-1}$ является матрицей перехода от базиса B к базису A.

   (b) Из определения матрицы перехода следует, что $\overline{C} = \overline{B}S_{B\to C}, \overline{B} = \overline{A}S_{A\to B} \Rightarrow \overline{C} = \overline{A}(S_{A\to B} \cdot S_{B\to C}) \Rightarrow S_{A\to B} \cdot S_{B\to C}$ является матрицей перехода от базиса A к базису C.

⊠

Теорема. *Пусть x — некоторый вектор пространства V, X_A и X_B — его координатные столбцы в базисах A и B соответственно, $S_{A\to B}$ — матрица перехода от базиса A к базису B. Тогда $X_A = S_{A\to B}X_B$*

⧫ Запишем в матричном виде разложение вектора x: $x = \overline{A}X_A = \overline{B}X_B$, но $\overline{B} = \overline{A}S_{A\to B}$. Следовательно, $\overline{A}X_A = \overline{A}S_{A\to B}X_B \Rightarrow \overline{A}(X_A - S_{A\to B}X_B) = 0_v \Rightarrow X_A - S_{A\to B}X_B$ является координатным столбцом нулевого вектора в базисе A. А так как нулевой вектор в любом базисе имеет нулевой координатный столбец, то $X_A - S_{A\to B}X_B = 0$ ⊠

10.9 Сумма и пересечение подпространств.

Пусть V — векторное пространство над полем P. И пусть U_1 и U_2 — подпространства V.

• ***Пересечением*** *подпространств U_1 и U_2 называется множество всех векторов, принадлежащих каждому из этих множеств (Обозначение: $U_1 \cap U_2$).*

• ***Суммой*** *подпространств называется множество векторов $a \in V$, представимых в виде: a_1+a_2, где $a_1 \in U_1,\ a_2 \in U_2$ (Обозначение: U_1+U_2). Если каждый вектор $a \in U_1+U_2$ представим в виде $a_1+a_2,\ a_1 \in U_1, a_2 \in U_2$ лишь единственным образом, то сумма U_1+U_2 называется* ***прямой*** *(Обозначение: $U_1 \oplus U_2$).*

Свойства суммы и пересечения подпространств:

1. *Пересечение и сумма подпространств являются подпространствами.*

   ⧫

   (a) Пересечение: Пусть $a, b \in U_1 \cap U_2 \Rightarrow a, b \in U_1,\ a, b \in U_2 \Rightarrow a+b, \alpha a \in U_1,\ a+b, \alpha a \in U_2 \Rightarrow a+b, \alpha a \in U_1 \cap U_2$. Значит в пересечении операции определены. Следовательно, $U_1 \cap U_2$ — подпространство по критерию подпространств.

   (b) Сумма: Пусть $a, b \in U_1 + U_2$.
   $\begin{cases} a = a_1 + a_2, \\ b = b_1 + b_2; \end{cases} \Rightarrow a_1, b_1 \in U_1,\ a_2, b_2 \in U_2 \Rightarrow$
   $a + b = \underset{\in U_1}{(a_1 + b_1)} + \underset{\in U_2}{(a_2 + b_2)} \Rightarrow a + b \in U_1 + U_2$
   $\alpha a = \underset{\in U_2}{\alpha a_1} + \underset{\in U_2}{\alpha a_2} \Rightarrow \alpha a \in U_1 + U_2$
   Так как U_1, U_2 — подпространства, то и $U_1 + U_2$ — подпространство.

   ⊠

2. $U_1 + U_2 = L(U_1 \cup U_2)$

   ⧫ Пусть $x \in U_1 + U_2 \Rightarrow x = \underset{\in U_1}{x_1} + \underset{\in U_2}{x_2} : x_1 \in U_1,\ x_2 \in U_2;\ x_1, x_2 \in U_1 \cup U_2 \Rightarrow$
   $x \in L(U_1 \cup U_2) \Rightarrow x_1 + x_2 \in L(U_1 \cup U_2) \Rightarrow U_1 + U_2 \subseteq L(U_1 \cup U_2)$

   Пусть $x \in L(U_1 \cup U_2) \Rightarrow$ он представим в виде: $x = \underbrace{\alpha_1 a_1 + \ldots + \alpha_n a_n}_{a_i \in U_1} + \underbrace{\beta_1 b_1 + \ldots + \beta_k b_k}_{b_i \in U_2} \Rightarrow$
   так как U_1 и U_2 — подпространства $\Rightarrow x \in U_1 + U_2 \Rightarrow L(U_1 \cup U_2) \subseteq U_1 + U_2$ ⊠

3. *Сумма $U_1 + U_2$ — прямая $\Longleftrightarrow U_1 \cap U_2 = \{0_v\}$*

   ⧫ $\Rightarrow$) Пусть сумма $U_1 + U_2$, — прямая. Предположим, что $U_1 \cap U_2 \neq \{0_v\}$. То есть, $\exists\, a \in U_1 \cap U_2,\ a \neq 0_v$. Так как $U_1 \cap U_2$ — подпространство, то $-a = (-1)a \in U_1 \cap U_2 \Rightarrow$ нулевой вектор, принадлежащий $U_1 + U_2$, представим в виде:
   $0_v = \underset{\in U_1}{0_v} + \underset{\in U_2}{0_v}$,
   $0_v = \underset{\in U_1}{a} + \underset{\in U_2}{(-a)}$.
   Мы получили два различных разложения нулевого вектора — противоречие с тем, что сумма прямая.

   $\Leftarrow$) Предположим, что $\exists\, x \in U_1 + U_2 : x = a_1 + a_2,\ x = b_1 + b_2$, где $a_1, b_1 \in U_1,\ a_2, b_2 \in U_2$
   Тогда $\underset{\in U_1}{(a_1 - b_1)} + \underset{\in U_2}{(a_2 - b_2)} = 0_v$.
   Так как U_1 — подпространство, то $(a_1 - b_1) \in U_1$. С другой стороны, $(b_1 - a_1) = (a_2 - b_2) \Rightarrow (a_1 - b_1) \in U_1 \Rightarrow (a_1 - b_1) \in U_1 \cap U_2 \Rightarrow a_1 - b_1 = 0_v \Rightarrow b_2 - a_2 = 0_v \Rightarrow a_1 = b_1,\ a_2 = b_2$. ⊠

Теорема (Формула Грассмана). *Если U_1, U_2 — подпространства векторного пространства V над полем P, то*

$$\dim(U_1 + U_2) = \dim U_1 \ + \ \dim U_2 \ - \ \dim(U_1 \cap U_2).$$

⧫ Пусть $C(c_1, \ldots, c_m)$ — базис пространства $U_1 \cap U_2$, тогда C — линейно независимая система векторов пространства $U_1 \cap U_2$, значит её можно дополнить до базисов каждого из подпространств U_1 и U_2.
Пусть $A(c_1, \ldots, c_m, a_1, \ldots, a_k)$ — базис U_1, а $B(c_1, \ldots, c_m, b_1, \ldots, b_s)$ — базис U_2.

Покажем, что $D(c, \ldots, c_m, a_1, \ldots, a_k, b_1, \ldots, b_s)$ является базисом пространства $U_1 + U_2$.

Докажем выражаемость:
Любой вектор $x \in U_1 + U_2$ представим в виде: $x = \underset{\in U_1}{x_1} + \underset{\in U_2}{x_2}$. Так как $x_1 \in U_1$, то он линейно выражается через систему A, соответственно $x_2 \in U_2$ линейно выражается через систему B. Следовательно x — линейная комбинация векторов системы D.

Докажем линейную независимость:
Пусть D — линейно зависимая система векторов. Тогда существует нетривиальная линейная комбинация, равная нулевому вектору:
$\gamma_1 c_1 + \ldots + \gamma_m c_m + \alpha_1 a_1 + \ldots + \alpha_k a_k + \beta_1 b_1 + \ldots + \beta_s b_s = 0_v$
$\gamma_1 c_1 + \ldots + \gamma_m c_m + \alpha_1 a_1 + \ldots + \alpha_k a_k = d$, где $d = -\beta_1 b_1 - \ldots - \beta_s b_s$.
Следовательно, d линейно выражается через систему векторов A и через систему векторов $B \Rightarrow d \in U_1 \cap U_2 \Rightarrow d$ линейно выражается через базис пересечения, то есть через систему $C : d = \delta_1 c_1 + .. + \delta_m c_m + 0 \cdot b_1 + \ldots + 0 \cdot b_s$. С другой стороны, справедливо разложение $d = 0 \cdot c_1 + \ldots + 0 \cdot c_n - \beta_1 b_1 - \ldots - \beta_s b_s$.
То есть d имеет два разложения по базису $B \Rightarrow \delta_i = \beta_i = 0\ \forall i \Rightarrow d = 0_v$ — это нетривиальная линейная комбинация векторов системы A равная $0_v \Rightarrow$ противоречие с тем, что система векторов A — базис $\Rightarrow$ система D — базис пространства $U_1 + U_2 \Rightarrow \dim(U_1 + U_2) = m + k + s = (m + k) + (m + s) - m = \dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$ ⊠

Следствие. *Если $W = U_1 \oplus U_2$, $A = (a_1, \ldots, a_k)$ — базис U_1, $B(b_1, \ldots, b_s)$ — базис U_2, то $(a_1, \ldots, a_k, b_1, \ldots, b_s)$ — базис $U_1 \oplus U_2$.*

10.10 Пространство решений линейной однородной системы.

Рассмотрим линейную однородную систему уравнений над полем P

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \ldots + \alpha_{mn}x_n = 0. \end{cases} \tag{10.10.1}$$

Обозначим через A матрицу системы (10.10.1), $A_1, \ldots, A_n$ — столбцы матрицы A, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ — столбец неизвестных. Тогда система (10.10.1) в матричном виде может быть записана следующим образом:

$$AX = 0 \text{ или } A_1x_1 + \ldots + A_nx_n = 0.$$

Так как однородная система линейных уравнений всегда имеет нулевое решение, то множество решений системы (10.10.1) непусто, а так как каждое решение является упорядоченной последовательностью элементов поля P, то есть элементом из P^n, то множество решений системы — непустое подмножество пространства P^n.

Теорема. *Множество решений линейной однородной системы уравнений над полем P с n неизвестными является подпространством пространства P^n размерности $n - \operatorname{rank} A$, где A — матрица системы.*

♦ Пусть L — множество решений линейной однородной системы уравнений (10.10.1). Тогда $L \subseteq P^n, L \neq \varnothing$.
Покажем, что L — подпространство пространства P^n. Пусть $b(\beta_1, \ldots, \beta_n) \in P^n$, $c(\gamma_1, \ldots, \gamma_n) \in P^n$ — решения системы (10.10.1). Тогда подставим решения в систему: $A_1\beta_1 + \ldots + A_n\beta_n = 0$, $A_1\gamma_1 + \ldots + A_n\gamma_n = 0$. Следовательно,

$$A_1(\beta_1 + \gamma_1) + \ldots + A_n(\beta_n + \gamma_n) = 0,$$

$$A_1(\alpha\beta_1) + \ldots + A_n(\alpha\beta_n) = 0.$$

Тогда $b + c = (\beta_1 + \gamma_1, \ldots, \beta_n + \gamma_n)$ и $\alpha b = (\alpha\beta_1, \ldots, \alpha\beta_n)$ — решения системы (10.10.1), то есть $b + c, \alpha b \in L$. Следовательно, L — подпространство P^n.
Покажем, что $\dim L = n - \operatorname{rank} A$.

1. Если $\operatorname{rank} A = 0$, то матрица A ненулевая. Следовательно, любой элемент пространства P^n является решением системы (10.10.1) и $L = P^n$. Тогда $\dim L = \dim P^n = n = n - 0 = n - \operatorname{rank} A$.

2. Пусть $\operatorname{rank} A = n$. Тогда все столбцы матрицы A линейно независимы. Следовательно, $A_1x_1 + \ldots + A_nx_n = 0$ справедливо лишь в случае, когда $x_1 = \ldots = x_n = 0 \Rightarrow L = \{0\} \Rightarrow \dim L = 0 = n - n = n - \operatorname{rank} A$.

3. Пусть $\operatorname{rank} A = k$, $0 < k < n$. Тогда базис системы столбцов матрицы A состоит из k столбцов. Без ограничения общности будем считать, что базис базис расположен в

первых k столбцах (в противном случае перенумеруем неизвестные). Тогда остальные $n-k$ столбцов линейно выражаются через первые k столбцов. То есть существуют $\beta_{i,j}$ такие, что $\begin{cases} \beta_{11}A_1 + \ldots + \beta_{1k}A_k + A_{k+1} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_{(n-k)1}A_1 + \ldots + \beta_{(n-k)k}A_k + A_n = 0. \end{cases}$ Тогда последовательности $\begin{cases} b_1(\beta_{11}, \ldots, \beta_{1k}, 1, 0, \ldots 0), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{n-k}(\beta_{(n-k)1}, \ldots, \beta_{(n-k)k}, 0, 0, \ldots 1); \end{cases}$ являются решениями системы (10.10.1).

Покажем, что последовательности $b_1, b_2, \ldots, b_{n-k}$ являются базисом пространства решений L. Запишем эти последовательности в виде матрицу по строкам:

$$\begin{pmatrix} \beta_{11} & \ldots & \beta_{1n} & 1 & 0 & \ldots & 0 \\ \beta_{21} & \ldots & \beta_{2n} & 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{k1} & \ldots & \beta_{kn} & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Последние $n-k$ столбцов образуют единичную матрицу. А так как ее определитель ненулевой, то он является базисным минором этой матрицы. Тогда строки, элементы которых образуют этот базисный минор, то есть $b_1, b_2, \ldots, b_{n-k}$, являются линейно независимыми.

Покажем, что любое решение системы (10.10.1) линейно выражается через эти последовательности. Пусть $c = (\gamma_1, \ldots, \gamma_n) \in L$ — произвольное решение системы (10.10.1). Построим последовательность $d = c - \gamma_{k+1}b_1 - \gamma_{k+2}b_2 - \ldots - \gamma_n b_{n-k}$, у которой все компоненты начиная с $k+1$ нулевые, то есть $d = (\delta_1, \ldots, \delta_k, 0, 0, \ldots, 0)$. А так как последовательность d является линейной комбинацией решений системы (10.10.1), то она также является решением системы (10.10.1). Следовательно, $A_1\delta_1 + \ldots + A_k\delta_k = 0$. Поскольку столбцы $A_1, \ldots, A_k$ линейно независимы, то $\delta_1 = \cdots = \delta_k = 0$, то есть d — нулевая последовательность (нулевой элемент поля P^n). Тогда $c = \gamma_{k+1}b_1 - \gamma_{k+2}b_2 - \ldots - \gamma_n b_{n-k}$ и последовательности $b_1, \ldots, b_{n-k}$ являются базисом пространства L. Получается, $\dim L = n - k = n - \operatorname{rank} A$.

⊠

Следствие. *Однородная система линейных уравнений имеет единственное решение $\Longleftrightarrow$ ранг матрицы системы равен числу неизвестных.*

⧫ Система имеет ненулевое решение $\Rightarrow \dim L > 0 \Rightarrow n - \operatorname{rank} A > 0 \Rightarrow n > \operatorname{rank} A$ ⊠

• *Базис пространства решений линейной однородной системы уравнений называется* ***фундаментальной системой решений****.*

Теорема. *Для любого подпространства L пространства P^n существует линейная однородная система уравнений над полем P с n неизвестными, множество решений которой совпадает с L.*

⧫ Для начала рассмотрим тривиальные случаи.

1. $L = P^n \Rightarrow$ система, все коэффициенты которой равны нулю, имеет множество решений, совпадающее с L;

2. $L = \{0_{P^n}\} \Rightarrow$ из принципа Крамера любая невырожденная система имеет множество решений, совпадающее с L;

Пусть $\dim L = k, 0 < k < n$. И пусть

$$\begin{cases} a_1(\alpha_{11}, \ldots, \alpha_{1n}), \\ \ldots\ldots\ldots\ldots\ldots \\ a_k(\alpha_{k1}, \ldots, \alpha_{kn}); \end{cases} \quad \text{— базис } L.$$

Составим систему

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{k1}x_1 + \ldots + \alpha_{kn}x_n = 0. \end{cases} \tag{10.10.2}$$

Так как векторы $a_1, \ldots, a_k$ линейно независимы, то ранг матрицы системы (10.10.2) равен k, а пространство решений системы (10.10.2) имеет размерность $n - k$. Пусть

$$\begin{cases} b_1(\beta_{11}, \ldots, \beta_{1n}), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{n-k}(\beta_{(n-k)1}, \ldots, \beta_{(n-k)n}); \end{cases} \quad \text{— фундаментальная системы решений системы (10.10.2).}$$

Построим систему

$$\begin{cases} \beta_{11}y_1 + \ldots + \beta_{1n}y_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_{(n-k)1}y_1 + \ldots + \beta_{(n-k)n}y_n = 0. \end{cases} \tag{10.10.3}$$

Покажем, что пространство решений системы (10.10.3) совпадает с пространством L. Последовательности $a_1, \ldots, a_k$ являются решениями системы (10.10.3), следовательно, любая линейная комбинация этих последовательностей также является решением системы (10.10.3), и любой элемент подпространства L является пространством решений (10.10.3). Тогда L — подпространство пространства решений системы (10.10.3).
Так как система векторов $(b_1, \ldots, b_{n-k})$ линейно независимая, то ранг матрицы системы (10.10.3) равен $n - k \Rightarrow$ пространство решений системы (10.10.3) имеет размерность $n - (n - k) = k \Rightarrow$ по теореме о монотонности размерности, L — пространство решений системы (10.10.3). ⊠

10.11 Критерий совместности линейных неоднородных систем. Структура множества решений линейной неоднородной системы уравнений.

Рассмотрим линейную неоднородную систему уравнений над полем P

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = b_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \ldots + \alpha_{mn}x_n = b_m. \end{cases} \tag{10.11.1}$$

Обозначим через A матрицу системы (10.11.1), $\widetilde{A}$ — расширенную матрицу системы (10.11.1), $A_1, \ldots, A_n$ — столбцы матрицы A, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ — столбец неизвестных, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ — столбец свободных членов. Тогда система (10.11.1) в матричном виде может быть записана следующим образом:

$$AX = B \text{ или } A_1x_1 + \ldots + A_nx_n = B.$$

Теорема Кронекера-Капелли (Критерий совместности линейных неоднородных систем). *Система линейных уравнений совместна* $\Longleftrightarrow$ *ранг матрицы системы равен рангу расширенной матрицы, то есть* $\operatorname{rank} A = \operatorname{rank} \widetilde{A}$.

♦ $\Rightarrow$) Если система (10.11.1) совместна, существуют скаляры $(\gamma_1, \ldots, \gamma_n) \in P$ такие, что $\gamma_1 A_1 + \ldots + \gamma_n A_n = B$. Следовательно, столбец B линейно выражается через столбцы $A_1, \ldots, A_n \Rightarrow L(A_1, \ldots, A_n, B) = L(A_1, \ldots, A_n) \Rightarrow \dim L(A_1, \ldots, A_n, B) = \dim L(A_1, \ldots, A_n) \Rightarrow \operatorname{rank}(A_1, \ldots, A_n, B) = \operatorname{rank}(A_1, \ldots, A_n) \Rightarrow \operatorname{rank} \widetilde{A} = \operatorname{rank} A$.

$\Leftarrow$) Пусть $\operatorname{rank} \widetilde{A} = \operatorname{rank} A$. Тогда $\operatorname{rank}(A_1, \ldots, A_n, B) = \operatorname{rank}(A_1, \ldots, A_n) \Rightarrow \dim L(A_1, \ldots, A_n, B) = \dim L(A_1, \ldots, A_n)$. Заметим, что $L(A_1, \ldots, A_n) \subseteq L(A_1, \ldots, A_n, B)$, по теореме о монотонности размерности $L(A_1, \ldots, A_n) = L(A_1, \ldots, A_n, B) \Rightarrow B \in L(A_1, \ldots, A_n)$, то есть существуют скаляры γ_i такие, что $B = \gamma_1 A_1 + \ldots + \gamma_n A_n \Rightarrow (\gamma_1, \ldots, \gamma_n)$ — решения системы (10.11.1). ⊠

• *Однородная система линейных уравнений $AX = 0$ с той же матрицей, что и система (10.11.1), называется **приведенной** для неоднородной системы (10.11.1).*

Теорема. *Если $x = (x_1, \ldots, x_n) \in P^n$ — решение системы (10.11.1), а L_o — пространство решений приведенной системы для системы (10.11.1), то множество решений системы (10.11.1) совпадает с множеством $L_H = \{x\} + L_o = \{x + y \mid y \in L_o\}$.*

♦ Пусть $z = (z_1, \ldots, z_n) \in L_H$. Тогда построим последовательность $y = (y_1, \ldots, y_n) \in L_o$ такую, что $y = z - x$. При этом $A_1y_1 + \ldots + A_ny_n = A_1(z_1 - x_1) + \ldots + A_n(z_n - x_n) = \underbrace{(A_1z_1 + \ldots + A_nz_n)}_{B} - \underbrace{(A_1x_1 + \ldots + A_nx_n)}_{B} = 0 \Rightarrow y \in L_o \Rightarrow z = x + y \Rightarrow L_H \subseteq \{x\} + L_o$.

Пусть $z = (z_1, \ldots, z_n) \in \{x\} + L_o$. Построим последовательность $z = x + y$. Тогда $A_1z_1 + \cdots + A_nz_n = A_1(x_1 + y_1) + \ldots + A_n(x_n + y_n) = \underbrace{(A_1x_1 + \ldots + A_nx_n)}_{B} + \underbrace{(A_1y_1 + \ldots + A_ny_n)}_{0} = B \Rightarrow z$ — решение системы (10.11.1), то есть $z \in L_H \Rightarrow \{x\} + L_o \subseteq L_H$. ⊠

Следствие. *Система линейных уравнений имеет единственное решение* $\Longleftrightarrow$ *ранги матрицы системы и расширенной матрицы системы равны числу неизвестных.*

♦ $\Rightarrow$) Если система (10.11.1) имеет единственное решение, то множество L_o состоит из одного элемента. А так как L_o — подпространство пространства P^n, то $L_o = \{0_{P^n}\} \Rightarrow \dim L_o = n - \operatorname{rank} A = 0 \Rightarrow n = \operatorname{rank} A = \operatorname{rank} \widetilde{A}$ ($\operatorname{rank} A = \operatorname{rank} \widetilde{A}$ по теореме Кронекера-Капелли).

$\Leftarrow$) От противного. Пусть $\operatorname{rank} A = \operatorname{rank} \widetilde{A} = n$ и пусть система (10.11.1) имеет два решения $x, y \in L_H \Rightarrow L_H = \{x\} + L_o = \{y\} + L_o$. Так как $\operatorname{rank} A = n$, то $\dim L_o = n - \operatorname{rank} A = 0 \Rightarrow L_o = \{0_v\} \Rightarrow \{x\} + L_o = \{x\} + \{0\} = \{x\}$ и $\{y\} + L_o = \{y\} + \{0\} = \{y\} \Rightarrow \{x\} = \{y\} \Rightarrow x = y$. ⊠

Глава 11

Линейные операторы

11.1 Определение и простейшие свойства линейных операторов.

• *Пусть V, U — векторные пространства над полем P. Отображение $f : V \to U$ называется* ***линейным оператором*** *(линейным отображением) пространства V в пространство U, если*

1. $f(a + b) = f(a) + f(b), \quad \forall a, b \in V$
2. $f(\alpha a) = \alpha f(a), \quad \forall a \in V,\ \forall \alpha \in P$

• *Линейный оператор пространства V в себя называется* ***линейным оператором*** *или* ***линейным преобразованием*** *пространства V. Множество линейных операторов пространства V (Обозначается: $End(V)$, эндоморфизм).*

Простейшие свойства линейных операторов:

1. $f(0_v) = 0_v$

   $f(-a) = -f(a), \quad \forall a \in V$

   ⧫ Следует из аксиомы 2 при $\alpha = 0$ и $\alpha = -1$. ⊠

2. *Линейный оператор линейно зависимую систему отображает в линейно зависимую.*

   ⧫ Пусть система $A(a_1, \dots, a_n)$ линейно зависимая. Тогда существует нетривиальная $\alpha_1 a_1 + \ldots + \alpha_n a_n = 0_v$, тогда $f(A) = (f(a_1), \ldots f(a_n))$ с теми же коэффициентами тоже нетривиальная и, при этом, $\alpha_1 f(a_1) + \ldots + \alpha_n f(a_n) = f(\alpha_1 a_1) + \ldots + f(\alpha_n a_n) = f(\alpha_1 a_1 + \ldots + \alpha_n a_n) = f(0_v) = 0_v \Rightarrow f(A) = (f(a_1), \ldots, f(a_n))$ линейно зависима. ⊠

Следствие. *Для любой системы векторов A и линейного оператора f* $\operatorname{rank} f(A) \leqslant \operatorname{rank} A$.

⧫ Максимальное количество линейно независимых векторов системы $f(A)$ не может быть больше, чем максимальное количество линейно независимых векторов из A. ⊠

3. *Если f — линейный оператор, $A(a_1, \ldots, a_n)$ — система векторов, то $f(L(A)) = L(f(A))$*

   ⧫ Следует из равенства $f(\alpha_1 a_1 + \ldots + \alpha_n a_n) = \alpha_1 f(a_1) + \ldots + \alpha_n f(a_n) \quad \forall \alpha_1 \in P$. ⊠

4. *Если U — подпространство пространства V, то $f(U)$ — также подпространство пространства V, причем $\dim f(U) \leqslant \dim U$.*

♦

(a) Покажем, что $f(U)$ — подпространство.
Если U — подпространство пространства V, то в нём существует базис A и $U = L(A)$. Тогда $f(U) = f(L(A)) = L(f(A))$, следовательно, $f(U)$ — подпространство.

(b) Докажем неравенство размерности.
$\dim f(U) = \dim L(f(A)) = \operatorname{rank} f(A) \leqslant \operatorname{rank} A = \dim L(A) = \dim U$.

⊠

11.2 Матрица линейного оператора.

Теорема. *Пусть V — векторное пространство над полем P, $A(a_1, \ldots, a_n)$ — базис пространства V, $B(b_1, \ldots, b_n)$ — некоторая система векторов из V. Тогда существует единственный линейный оператор $f : V \to V$, для которого $f(a_i) = b_i$.*

♦

1. Существование:

   Так как система векторов A — базис пространства V, то любой вектор $c \in V$ можно представить в виде $c = \gamma_1 a_1 + \ldots + \gamma_n a_n$, где γ_i — координаты вектора c в базисе A.

   Покажем, что отображение $f(c) = \gamma_1 b_1 + \ldots + \gamma_n b_n$ является единственным линейным оператором пространства V, для которого выполняется $f(a_i) = b_i \quad \forall i$.
   Вектор a_i в базисе A имеет разложение $a_i = 0 \cdot a_1 + \ldots + 0 \cdot a_{i-1} + 1 \cdot a_i + 0 \cdot a_{i+1} + \ldots + 0 \cdot a_n$.
   Тогда $f(a_i) = 0 \cdot b_1 + \ldots + 0 \cdot b_{i-1} + 1 \cdot b_i + 0 \cdot b_{i+1} + \ldots + 0 \cdot b_n = b_i$.

   Покажем, что отображение f — линейное. Пусть произвольные векторы векторы c, d пространства V в базисе A имеют координаты $(\gamma_1, \ldots, \gamma_n)$ и $(\delta_1, \ldots, \delta_n)$ соответственно. Тогда $f(c + d) = f((\gamma_1 a_1 + \ldots + \gamma_n a_n) + (\delta_1 a_1 + \ldots + \delta_n a_n)) = (\gamma_1 + \delta_1)b_1 + \ldots + (\gamma_n + \delta_n)b_n = (\gamma_1 b_1 + \ldots + \gamma_n b_n) + (\delta_1 b_1 + \ldots + \delta_n b_n) = f(c) + f(d)$. Случай $f(\alpha c) = \alpha f(c)$ доказывается аналогично.

2. Единственность:

   Пусть $g : V \to V$ — линейный оператор пространства V, для которого выполняется $g(a_i) = b_i \quad \forall i$. Тогда $\forall c \in V : c = \gamma_1 a_1 + \ldots + \gamma_n a_n$ будет справедливо $g(c) = g(\gamma_1 a_1 + \ldots + \gamma_n a_n) = \gamma_1 g(a_1) + \ldots + \gamma_n g(a_n) = \gamma_1 b_1 + \ldots + \gamma_n b_n = f(c)$. Следовательно, f и g — одно и то же отображение.

⊠

Из теоремы следует, что действие линейного оператора пространства V полностью определяется действием этого оператора на базис пространства V.

• *Пусть $A(a_1, \ldots, a_n)$ — базис пространства V, f — линейный оператор пространства V. Матрица, составленная из координатных столбцов векторов $f(a_i)$ в базисе A, называется* ***матрицей линейного оператора*** f ***в базисе*** A *(обозначение: M_f, M_f^A). Таким* *образом, если обозначить через A_i координатные столбцы векторов $f(a_i)$ в базисе A, то* $M_f^A = [A_1, \dots, A_n] \in P_{n,n}$

Из определения следует, что матрица линейного оператора f в базисе A является матрицей перехода от базиса A к системе векторов $f(A)$, то есть $M_f^A = S_{A \to f(A)}$.

Теорема. *Пусть M_f — матрица линейного оператора $f : V \to V$ в базисе A, X_a и $X_{f(a)}$ — координатные столбцы векторов a и $f(a)$ в базисе A. Тогда $X_{f(a)} = M_f^A X_a$.*

⧫ Если базис A пространства V записать в виде $\overline{A}(a_1, \dots, a_n)$, то произвольный вектор a имеет разложение $a = \alpha_1 a_1 + \ldots + \alpha_n a_n = \overline{A} X_a$ по базису A. Тогда $f(a) = \overline{A} X_{f(a)}$.

А из опеределения матрицы перехода следует, что $\overline{f(A)} = \overline{A} S_{A \to f(A)} = \overline{A} M_f^A$. Так как f — линейный оператор, то $f(a) = f(\alpha_1 a_1 + \ldots + \alpha_n a_n) = \alpha_1 f(a_1) + \ldots + \alpha_n f(a_n) = \overline{f(A)} X_A$.

Следовательно, из уравнения полученного ранее и этого уравнения имеем $\overline{A} X_{f(a)} = \overline{f(A)} X_A \Rightarrow (\overline{A} \cdot M_f^A) X_A = \overline{A}(M_f^A X_A)$. А так как разложение по базису единственно, то $X_{f(a)} = M_f X_a$, что и требовалось доказать. ⊠

11.3 Пространство линейных операторов.

Пусть V — n-мерное векторное пространство над полем P, а f и g — линейные операторы пространства V.

• ***Суммой*** *линейных операторов f и g называется отображение $(f + g) : V \to V$, которое каждому вектору $a \in V$ ставит в соответствие вектор $f(a) + g(a) \in V$, то есть*

$$(f + g)(a) = f(a) + g(a), \quad \forall\ a \in V.$$

• ***Произведением оператора** f **на скаляр** α называется отображение $(\alpha f) : V \to V$, которое каждому вектору $a \in V$ ставит в соответствие вектор $\alpha f(a) \in V$, то есть*

$$(\alpha f)(a) = \alpha f(a), \quad \forall\ a \in V.$$

• ***Композицией операторов** f и g называется отображение $(g \circ f) : V \to V$, которое каждому вектору $a \in V$ ставит в соответствие вектор $g(f(a)) \in V$, то есть*

$$(g \circ f)(a) = g(f(a)), \quad \forall\ a \in V.$$

Теорема. *Если f и g — линейные оператор пространства V, то отображения $(f + g), (\alpha f), (g \circ f)$ также являются линейными операторами пространства V, причем для их матриц в некотором фиксированном базисе справедливы равенства:*

$$M_{f+g} = M_f + M_g, \quad M_{\alpha f} = \alpha M_f, \quad M_{g \circ f} = M_g M_f.$$

⧫ Доказательство проведем для случая $g \circ f$. Для остальных случаев доказательство аналогично.

1. Линейность. $\forall a, b \in V$, $\forall \alpha \in P$
   $(g \circ f)(a + b) = g(f(a + b)) = g(f(a) + f(b)) = g(f(a)) + g(f(b)) = (g \circ f)(a) + (g \circ f)(b)$.
   $(g \circ f)(\alpha a) = g(f(\alpha a)) = g(\alpha f(a)) = \alpha(g(f(a))) = \alpha(g \circ f)(a)$.

2. Покажем, что $M_{g\circ f} = M_g M_f$.
   Пусть $M_f = (\alpha_{ij})$ и $M_g = (\beta_{ij})$ — матрицы операторов f и g в базисе $A(a_1, \ldots, a_n)$.
   Тогда $f(a_j) = \alpha_{1j}a_1 + \ldots + \alpha_{nj}a_n,\ g(a_j) = \beta_{1j}a_1 + \ldots + \beta_{nj}a_n$.
   $(g \circ f)(a_j) = g(f(a_j)) = g(\alpha_{1j}a_1 + \ldots + \alpha_{nj}a_n) = \alpha_{1j}g(a_1) + \ldots + \alpha_{nj}g(a_n) = \alpha_{1j}(\beta_{11}a_1 + \ldots + \beta_{n1}a_n) + \ldots + \alpha_{nj}(\beta_{1n}a_1 + \ldots + \beta_{nn}a_n) = (\alpha_{1j}\beta_{11} + \ldots + \alpha_{nj}\beta_{1n})a_1 + \ldots + (\alpha_{1j}\beta_{n1} + \ldots + \alpha_{nj}\beta_{nn})a_n \Rightarrow j$-ые столбцы матриц $M_{g\circ f}$ и $M_g M_f$ одни и те же.

⊠

Следствие. *Множество $End(V)$ всех линейных операторов пространства V является векторным пространством над полем P относительно операций сложения линейных операторов и умножения линейных операторов на скаляр.*

⧫ Из теоремы следует, что операции сложения и умножения на скаляр определены в $End(V)$, остальные же аксиомы векторного пространства проверяются аналогично. ⊠

Теорема. *Отображение $\varphi : End(V) \to P_{n,n}$, ставящее в соответствие каждому линейному оператору $f \in End(V)$ его матрицу в некотором фиксированном базисе A, то есть $\varphi(f) = M_f^A$, является изоморфизмом.*

⧫

1. Докажем биективность:
   (a) Докажем инъективность:
   От противного. Пусть $f \neq g$ и предположим, что $\varphi(f) = \varphi(g)$, то есть $M_f^A = M_g^A$. Тогда любой вектор $a \in V$ представим в виде $a = \alpha_1 a_1 + \ldots + \alpha_n a_n \Rightarrow f(a) = f(\alpha_1 a_1 + \ldots + \alpha_n a_n) = f(\alpha_1 a_1) + \ldots + f(\alpha_n a_n) = \alpha_1 f(a_1) + \ldots + \alpha_n f(a_n) = [M_f^A = M_g^A \Rightarrow f(a_j) = g(a_j)] = \alpha_1 g(a_1) + \ldots + \alpha_n g(a_n) = g(\alpha_1 a_1 + \ldots + \alpha_n a_n) = g(a)$, что противоречит тому, что $f \neq g$. Значит $f = g$ и φ — инъекция.
   (b) Докажем сюръективность:
   Покажем, что для любой матрицы $C = (\gamma_{ij}) \in P_{n,n}$ существует единственный линейный оператор $f \in End(V)$ такой, что $M_f^A = \varphi(f) = C$. Пусть $A(a_1, \ldots, a_n)$ — базис пространства V. Так как для произвольной системы векторов $B(b_1, \ldots, b_n)$ существует единственный линейный оператор, который векторы a_i базиса A отображает в векторы $b_i = \gamma_{1i}a_1 + \ldots + \gamma_{ni}a_n \quad \forall i$, то матрица этого линейного оператора в базисе A совпадает с матрицей C. Следовательно, φ — сюръекция. А значит и биекция.
2. Докажем линейность: $\forall f, g \in End(V),\ \forall \alpha \in P$
   $\varphi(f + g) = M_f + M_g = \varphi(f) + \varphi(g)$,
   $\varphi(\alpha f) = M_{\alpha f} = \alpha\varphi_f \Rightarrow$ отображение линейно.

⊠

11.4 Ранг и дефект линейного оператора.

Пусть $f : V \to V$ — линейный оператор пространства V над полем P.

• *Множество $\{f(a) \mid a \in V\}$ называется* ***образом*** *линейного оператора f. Обозначение: $Im\ f,\ f(V)$.*

• *Множество $\{a \in V \mid f(a) = 0_v\}$ называется* ***ядром*** *линейного оператора f. Обозначение: $Ker\ f$.*

Лемма. *Образ и ядро линейного оператора V являются подпространствами пространства V.*

♦

1. Для любого подпространства U пространства V множество $f(U)$ также является подпространством пространства V (по свойству 4). Так как V является подпространством самого себя, то $f(V)$ — подпространство пространства V.

2. Так как линейный оператор нулевой вектор отображает в нулевой, то $0_v \in Ker\ f$, следовательно, $Ker\ f \neq \varnothing$.
   Покажем, что множество $Ker\ f$ замкнуто относительно линейных операций. Пусть произвольные векторы $a, b \in Ker\ f,\ \alpha \in P$. Тогда
   $f(a+b) = f(a) + f(b) = 0_v + 0_v = 0_v$,
   $f(\alpha a) = \alpha f(a) = \alpha \cdot 0_v = 0_v$.
   Следовательно, во множестве $Ker\ f$ определены операции сложения векторов и умножения вектора на скаляр, и $Ker\ f$ — подпространство пространства V.

⊠

• *Размерность образа линейного оператора называется* ***рангом*** *оператора. Обозначение:* $\operatorname{rank} f$.

• *Размерность ядра линейного оператора называется* ***дефектом*** *оператора. Обозначение: $def f$.*

Теорема. *Линейный оператор инъективен $\Longleftrightarrow$ его ядро — нулевое подпространство.*

♦ ⇒) Пусть линейный оператор $f : V \to V$ инъективен. Так как любой линейный оператор нулевой вектор отображает в нулевой, то других векторов, отображенных инъективно в нулевой, нет. Следовательно, $Kerf = \{0_v\}$.

⇐) Пусть $Kerf = \{0_v\}$. Возьмем произвольные векторы $a,\ b \in V$ такие, что $a \neq b,\ f(a) = f(b)$. Тогда $f(a) - f(b) = 0_v \Rightarrow f(a-b) = 0_v \Rightarrow$ [по определению ядра] $\Rightarrow \underset{\neq 0_v}{a-b} \in Kerf \Rightarrow$ получаем противоречие с тем, что $a \neq b$. Значит $a = b$ и f — инъекция. ⊠

Пусть $f : V \to V$ — линейный оператор, M_f — матрица этого оператора в базисе A.

Теорема. *Ранг линейного оператора равен рангу матрицы оператора:* $\operatorname{rank} f = \operatorname{rank} M_f$. *Дефект линейного оператора пространства V равен разности размерности пространства V и линейного оператора: $def f = \dim V - \operatorname{rank} f$.*

♦

1. $\operatorname{rank} f =$ [по определению] $= \dim f(V) =$ [так как A — базис] $= \dim f(L(A)) = \dim L(f(A)) = \operatorname{rank} f(A)$. Так как система векторов линейно независима $\Longleftrightarrow$ линейно независима система координатных столбцов этих векторов, то максимальное число линейно независимых векторов в системе $f(A)$ совпадает с максимальным числом линейно независимых координатных столбцов векторов системы $f(A)$, то есть столбцов матрицы M_f, а значит $\operatorname{rank} f(A) = \operatorname{rank} M_f$.

2. Пусть $a \in V$ — некоторый вектор, X_a — его координатный столбец в базисе A, тогда $f(a)$ имеет координатный столбец равный $M_f X_a$.
   Вектор $a \in Ker\ f \Longleftrightarrow f(a) = 0_v \Longleftrightarrow$ координатный столбец X_a является решением

матричного уравнения $M_f X = 0 \iff M_f X_a = 0$. Значит ядро состоит из векторов, координатные столбцы которых совпадают с пространством решений системы $M_f X = 0$, которое имеет размерность $\dim(Ker\ f) = def = f = \dim V - \operatorname{rank} M_f = \dim V - \operatorname{rank} f$.

⊠

Следствие. *Линейный оператор пространства V инъективен $\iff$ он сюръективен.*

♦ $\Rightarrow$) Пусть линейный оператор f инъективен. Тогда $Kef\ f = \{0_v\} \Rightarrow def\ f = 0 \Rightarrow \operatorname{rank} f = \dim V - def\ f = \dim V$. Тогда, по теореме о монотонности размерности, подпространство $f(V)$ совпадает с пространством V. Следовательно, $f(V) = V$ и $\forall a \in V \exists b : f(a) = b$, значит оператор сюръективен.

$\Leftarrow$) Пусть линейный оператор f сюръективен. Тогда $f(V) = V \Rightarrow \operatorname{rank} f = \dim V \Rightarrow def\ f = 0 \Rightarrow Ker\ f = \{0_v\} \Rightarrow$ оператор f инъективен. ⊠

11.5 Изменение матрицы линейного оператора при замене базиса. Подобные матрицы.

• *Пусть A и B — квадратные матрицы одного порядка над полем P. Матрица A называется* ***подобной*** *матрице B над полем P, если существует невыоржденная матрица S над полем P такая, что $A = S^{-1}BS$. При этом матрица S называется матрицей,* ***трансформирующей*** *B в A.*

Свойства подобных матриц:

1. *Бинарное отношение подобия матриц является отношением эквивалентности.*

   ♦

   (a) рефлексивность: A подобна A, так как $A = E^{-1}AE$.

   (b) симметричность: пусть A подобна B. Тогда существует невырожденная матрица S такая, что под. $A = S^{-1}BS$. Обозначим $S' = S^{-1}$. Так как S невырожденная, то S' также невырожденная. Следовательно, $B = SAS^{-1} = (S')^{-1}AS' \Rightarrow B$ подобна A.

   (c) транзитивность: пусть B подобна A, C подобна B. Тогда существуют невырожденные матрицы S_1, S_2 такие, что $B = S_1^{-1}AS_1, C = S_2^{-1}BS_2 \Rightarrow C = S_2^{-1}(S^{-1}AS)S_2 = (S_2^{-1}S_1^{-1})A(S_1S_2) = (S_1S_2)^{-1}A(S_1S_2)$. $detS_1S_2 = detS_1 \cdot detS_2 \neq 0 \Rightarrow C$ подобна A.

   ⊠

2. *Ранги подобных матриц равны.*

   ♦ Пусть A подобна $B \Rightarrow$ существует невырожденная матрица S такая, что $A = S^{-1}BS \Rightarrow \operatorname{rank} A = \operatorname{rank}(S^{-1}BS) =$ [так как S невырожденная] $= \operatorname{rank}(S^{-1}B) =$ [так как S^{-1} невырожденная] $= \operatorname{rank} B$. ⊠

3. *Определители подобных матриц равны.*

   ♦ Пусть A и B подобны. Тогда существует невырожденная матрица S такая, что $A = S^{-1}BS \Rightarrow detA = detS^{-1} \cdot detB \cdot detS = \dfrac{1}{detS} \cdot detB \cdot detS = detB$. ⊠

Пусть V векторное пространство над полем P, A и B — базисы пространства V, f — линейный оператор пространства V.

Теорема. *Матрицы M^A и M^B линейного оператора f в базисах A и B подобны, причем матрица S, трансформирующая M^A и M^B, является матрицей перехода от базиса A к базису B:*

$$M^B = (S_{A\to B})^{-1} M^A S_{A\to B} = S_{B\to A} M^A S_{A\to B}.$$

♦ Пусть $x \in V$ — произвольный вектор и X_A и X_B — его координатные столбцы в базисах A и B соответственно. Тогда $X_A = S_{A\to B} X_B = S X_B$.

Аналогичное равенство справедливо для $f(x) \in V$, который в базисах A и B имеет столбцы $M^A f X_A$ и $M_f^B X_B \Rightarrow M^A X_A = S M^B X_B \Rightarrow M^A S X_B = S M^B X_B$. Так как полученное соотношение верно для всех векторов из пространства V, то оно, в частности, верно для b_1 (первого вектора из системы B), который в базисе B имеет координатный столбец $\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow$ первые столбцы $M^A S$ и $S M^B$ равны. Продолжая рассуждения аналогичным образом и используя остальные векторы из базиса B, получим равенства остальных столбцов, а значит $M^A S = S M^B$. Так как S — матрица перехода от базиса к базису, то она невырожденная $\Rightarrow M^B = S^{-1} M^A S$, где $S = S_{A\to B}$. ⊠

11.6 Инвариантное подпространство.

• *Пусть f — линейный оператор векторного пространства V над полем P. Подпространство $U \subseteq V$ называется* ***инвариантным*** *относительно f, если $f(U) \subseteq U$.*

Свойства инвариантных подпространств:

1. *Сумма и пересечение подпространств, инвариантных относительно линейного оператора f, также инвариантны относительно f.*

   ♦ Пусть U_1, U_2 инвариантны относительно оператора $f \Rightarrow f(U_1) \subseteq U_1, f(U_2) \subseteq U_2$.

   (a) $\forall a \in U_1 + U_2 \quad a = \underset{\in U_1}{a_1} + \underset{\in U_2}{a_2} \Rightarrow f(a) = f(a_1 + a_2) = \underset{\in U_1}{f(a_1)} + \underset{\in U_2}{f(a_2)} \Rightarrow$ так как U_1 и U_2 инвариантны относительно оператора f, то $f(a) \in U_1 + U_2$, то есть $U_1 + U_2$ инвариантно относительно f.

   (b) $\forall a \in U_1 \cap U_2 \Rightarrow a \in U_1,\ a \in U_2 \Rightarrow$ [так как U_1 и U_2 инвариантны относительно оператора f] $\Rightarrow f(a) \in U_1,\ f(a) \in U_2 \ \Rightarrow f(a) \in U_1 \cap U_2 \Rightarrow U_1 \cap U_2$ инвариантно относительно f.

   ⊠

2. *Если подпространство U инвариантны относительно линейных операторов f и g, то U инвариантно относительно операторов $f + g, \alpha f, f \circ g$.*

   ♦ Если $a \in U$, то $f(a) \in U,\ g(a) \in U$, так как U инвариантно относительно f и $g \Rightarrow$
   $(f + g)(a) = f(a) + g(a) \in U$,
   $(\alpha f)(a) = \alpha f(a) \in U$,
   $(f \circ g)(a) = f(g(a)) \in U$. ⊠

• *Пусть U — инвариантное относительно линейного оператора f подпространство пространства V. Тогда отображение $f|_U : U \to U$ такое, что $f|_U(a) = f(a), \forall a \in U$, является линейным оператором пространства U и называется* ***ограничением*** *оператора f на подпространство U.*

Теорема. *Если векторное пространство V имеет нетривиальное инвариантное относительно линейного оператора f подпространство, то существует базис пространства V, в котором матрица оператора f имеет блочнотреугольный вид.*

⧫ Пусть V — n-мерное векторное пространство, U — нетривиальное относительно линейного оператора f подпространство и $\dim U = k,\ 0 < k < n$. Построим систему векторов $A_1(a_1, \dots, a_k)$, являющуюся базисом пространства U. Дополним систему A_1 до базиса пространства $V : A = (\underbrace{a_1, \dots, a_k}_{U}, a_{k+1}, \dots, a_n)$ и построим матрицу M_f следующим образом: разложим вектор $f(a_i)$ по базису A, так как $(a_1, \dots, a_k)$ принадлежит инвариантному относительно f подпространству U, то $f(a_1), \dots, f(a_k) \in U \Rightarrow$ разложения этих векторов по базису A содержат лишь векторы из системы A_1:

$$\begin{cases} f(a_1) = \alpha_{11}a_1 + \ldots + \alpha_{k1}a_k, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ f(a_k) = \alpha_{1k}a_1 + \ldots + \alpha_{kk}a_k, \\ f(a_{k+1}) = \alpha_{1k+1}a_1 + \ldots + \alpha_{kk+1}a_k + \ldots + \alpha_{nk+1}a_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ f(a_n) = \alpha_{1n}a_1 + \ldots + \alpha_{kn}a_k + \ldots + \alpha_{nn}a_n. \end{cases}$$

$$M_f^A = \left(\begin{array}{ccc|ccc} \alpha_{11} & \dots & \alpha_{1k} & \alpha_{1,k+1} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{k1} & \dots & \alpha_{kk} & \alpha_{k,k+1} & \dots & \alpha_{kn} \\ \hline 0 & \dots & 0 & \alpha_{k,k+1} & \dots & \alpha_{kn} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{n,k+1} & \dots & \alpha_{nn} \end{array}\right).$$

⊠

Замечание: Блок $\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1k} \\ \vdots & \ddots & \vdots \\ \alpha_{k1} & \dots & \alpha_{kk} \end{pmatrix}$ является матрицей линейного оператора $f|_U$ в базисе $(a_1, \dots, a_k)$.

Следствие. *Если векторное пространство V является прямой суммой инвариантных относительно линейного оператора f подпространств, то существует базис пространства V, в котором матрица оператора f имеет блочнодиагональный вид.*

⧫ Пусть $V = U_1 \oplus U_2$, U_1 и U_2 — инвариантные относительно оператора f подпространства пространства V.
Если $A_1,\ A_2$ — базисы $U_1,\ U_2$ соответственно, то, по следствию из теоремы Грассмана, система векторов (A_1, A_2) является базисом пространства V. При этом матрица линейного оператора f блочнодиагональная, так как $f(A_1)$ раскладывается по базису A_1, а $f(A_2)$ — по базису A_2. ⊠

Следствие. *Если векторное пространство V является прямой суммой одномерных инвариантных относительно линейного оператора f подпространств, то существует базис пространства V, в котором матрица оператора f является диагональной.*

♦ Доказательство проводится аналогично предыдущему следствию. ⊠

• *Линейный оператор f векторного пространства V называется* ***оператором простой структуры****, если существует базис пространства V, в котором матрица оператора f имеет диагональный вид.*

• *Матрица называется* ***диагонализируемой****, если она имеет подобную диагональную матрицу.*

11.7 Одномерное инвариантное подпространство. Собственные векторы и собственные значения линейного оператора.

Пусть f — линейный оператор векторного пространства V над полем P.

• *Если для некоторого ненулевого вектора $a \in V$ справедливо равенство $f(a) = \lambda_0 a$, где $\lambda_0 \in P$, то скаляр λ_0 называется* ***собственным значением*** *оператора f, а вектор a — собственным вектором оператора f, соответствующим собственному значению λ_0.*

Лемма. *Собственные векторы, соответствующие попарно различным собственным значениям, линейно независимы.*

♦ От противного. Пусть $(a_1, \dots, a_k)$ — система векторов, состоящая из собственных векторов оператора f, соотвестветствующих попарно различным собственным значениям $\lambda_1, \dots, \lambda_k$, и пусть она линейно зависима. Тогда базис этой системы векторов состит менее чем из k векторов, и, без ограничения общности, система векторов $A(a_1, \dots, a_s)$ будет являться базисом пространства при условии, что $s < k$. Тогда вектор a_{s+1} линейно выражается через систему A, то есть представим в виде
$a_{s+1} = \alpha_1 a_1 + \dots + \alpha_s a_s$. Домножим на λ_{s+1} и получим

$$\lambda_{s+1} a_{s+1} = \alpha_1 \lambda_{s+1} a_1 + \dots + \alpha_s \lambda_{s+1} a_s. \tag{11.7.1}$$

С другой стороны, $\lambda_{s+1} a_{s+1} = f(a_{s+1}) = f(\alpha_1 a_1 + \dots + \alpha_s a_s) = \alpha_1 f(a_1) + \dots + \alpha_s f(a_s) \Rightarrow$

$$\lambda_{s+1} a_{s+1} = \alpha_1 \lambda_1 a_1 + \dots + \alpha_s \lambda_s a_s. \tag{11.7.2}$$

Вычтем из уравнения (11.7.1) уравнение (11.7.2) и получим $0_v = \alpha_1(\lambda_{s+1} - \lambda_1)a_1 + \dots + \alpha_s(\lambda_{s+1} - \lambda_s)a_s$. Так как система A линейно независима, то коэффициенты линейной комбинации нулевые, то есть $\alpha_i(\lambda_{s+1} - \lambda_1) = 0 \quad \forall i$. А так как все собственные значения λ_i попарно различны, то $(\lambda_{s+1} - \lambda_i) \neq 0 \Rightarrow \alpha_i = 0 \quad \forall i = \overline{1, s} \Rightarrow a_{s+1} = 0_v$, что противоречит тому, что вектор a_{s+1} собственный. ⊠

Следствие. *Линейный оператор n-мерного векторного пространства не может иметь больше чем n различных собственных значений.*

Теорема. *Одномерное подпространство векторного пространства V является инвариантным относительно линейного оператора $f \iff$ все его ненулевые векторы являются собственными векторами оператора f.*

♦ $\Rightarrow$) Пусть одномерное векторное пространство U инвариантно относительно линейного оператора f. Тогда $\forall\ a \in U,\ a \neq 0 \quad f(a) \in U$. Но так как подпространство U одномерное, то ненулевой вектор a является базисом этого подпространства. Следовательно любой вектор из U, в том числе и сам вектор a, представим в виде линейной комбинации $f(a) = \lambda a$, то есть вектор a является собственным векторов оператора f.

$\Leftarrow$) Пусть прозвольный вектор $a \neq 0$ и $a \in U$, $f(a) = \lambda a$. Так как U — подпространство, то $\lambda a \in U \Rightarrow f(a) \in U$. Заметим, что $f(0_v) = 0_v$, $f(0_v) \in U \Rightarrow \forall a \in U \quad f(a) \in U \Rightarrow f(U)$ является подпространством пространства U, а значит пространство U инвариантно относительно оператора f. ⊠

Теорема. *Линейный оператор векторного пространства V является оператором простой структуры $\iff$ существует базис пространства V, состоящий из собственных векторов оператора.*

♦ $\Rightarrow$) Пусть f — оператор простой структуры. Тогда существует базис $A(a_1, \ldots, a_n)$ пространства V, в котором матрица линейного оператора диагональна и имеет вид

$$M_f^A = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Тогда
$f(a_1) = \lambda_1 a_1 + 0 \cdot a_2 + \ldots + 0 \cdot a_n = \lambda_1 a_1,$
$f(a_2) = 0 \cdot a_1 + \lambda a_2 + \ldots + 0 \cdot a_n = \lambda_2 a_2,$
....
$f(a_n) = 0 \cdot a_1 + 0 \cdot a_2 + \ldots + \lambda_n a_n = \lambda_n a_n.$
Следовательно, $f(a_i) = \lambda_i a_i$, причем, так как векторы a_i ненулевые, поскольку система A является базисом, они являются собственными векторами оператора f.

$\Leftarrow$) Пусть система векторов $A(a_1, \ldots, a_n)$ является базисом пространства V, и пусть она состоит из собственных векторов линейного оператора f, соответствующих попарно различным собственным значениям λ_i. Тогда $f(a_i) = \lambda_i a_i \Rightarrow M_f^A = diag(\lambda_1, \ldots, \lambda_n) \Rightarrow f$ — оператор простой структуры. ⊠

Следствие. *Линейный оператор n-мерного векторного пространства, имеющий n различных собственных значений, является оператором простой структуры.*

♦ Если линейный оператор имеет n различных собственных значений, то он имеет n собственных векторов, соответствующих этим собственным значениям. Тогда по лемме эти векторы линейно независимы. Значит они образуют базис этого пространства. Следовательно, f — оператор простой структуры. ⊠

11.8 Алгебраическая кратность собственного значения.

Пусть f — линейный оператор n-мерного векторного пространства V над полем P.

Теорема.

1. *Вектор $x \in V$ является собственным вектором линейного оператора f, соответствующим собственному значению $\lambda_0 \iff$ его координатный столбец X в некотором базисе является ненулевым решением уравнения $(M_f - \lambda_0 E)X = 0$, где M_f — матрица оператора f в том же базисе.*

2. *Скаляр* $\lambda_0 \in P$ *является собственным значением линейного оператора* $f \iff det(M_f - \lambda_0 E) = 0$.

♦

1. Ненулевой вектор x является собственным вектором линейного оператора f, соответствующим собственному значению λ_0, если $f(x) = \lambda_0 x$. Два вектора равны $\Leftrightarrow$ равны их координатные столбцы. Если X — это координатный столбец вектора x, то $\lambda_0 x$ имеет координатный столбец $\lambda_0 X$, а $f(x)$ — $M_f X$. Тогда $\lambda_0 X = M_f X \Leftrightarrow M_f X - \lambda_0 X = 0 \Leftrightarrow M_f X - \lambda_0 E X = 0 \Leftrightarrow (M_f - \lambda_0 E)X = 0$.

2. Скаляр λ_0 является собственным значением линейного оператора $f \iff$ существует собственный вектор, соответствующий этому значению. А значит существует ненулевое решение матричного уравнения $(M_f - \lambda_0 E)X = 0$, что возможно лишь в случае, когда $det(M_f - \lambda E)X = 0$.

⊠

• *Пусть* A *— квадратная матрица порядка* n *над полем* P, λ *— переменная. Матрица* $A - \lambda E$ *называется* ***характеристической матрицей****, определитель* $det(A - \lambda E)$ *—* ***характеристическим многочленом*** *матрицы* A*, уравнение* $det(A - \lambda E) = 0$ *—* ***характеристическим уравнением*** *матрицы* A.

Свойства характ мн-ена:

1. *Степень характеристического многочлена матрицы равна порядку матрицы.*

2. *Если матрица* A *имеет порядок* n*, то старший коэффициент характеристического многочлена матрицы* A *равен* $(-1)^n$.

• *Сумма элементов главной диагонали матрицы* A *называется* ***следом матрицы*** A. *Обозначение:* $SpA = (\alpha_{11} + \ldots + \alpha_{nn})$.

3. *Коэффициент при* λ^{n-1} *равен* $(-1)^{n-1} SpA$.

4. *Свободный член равен* $detA$.

Лемма. *Характеристические многочлены подобных матриц равны.*

♦ Если матрицы A и B подобны, то существует невырожденная матрица S такая, что $A = S^{-1}BS$. Тогда $A - \lambda E = S^{-1}BS - \lambda S^{-1}ES = S^{-1}(B - \lambda E)S \Rightarrow det(A - \lambda E) = detS^{-1}det(B - \lambda E)detS = det(B - \lambda E)$. ⊠

Следствие. *Следы подобных матриц равны.*

Следствие. *Определители подобных матриц равны.*

Так как матрицы линейного оператора в различных базисах подобны, то их характеристические многочлены равны и, следовательно, зависят от самого оператора и не зависят от выбора базиса.

• *Характеристический многочлен матрицы линейного оператора называется* ***характеристическим многочленом оператора****.*

• *Из леммы следует, что собственное значение линейного оператора является корнем характеристического уравнения оператора. Его кратность называется* ***алгебраической***

кратностью *собственного значения.*

Так как степень характеристического многочлена линейного оператора пространства V равна $\dim V$, то сумма алгебраических кратностей собственных значений линейного оператора не превышает размерности пространства V.

11.9 Геометрическая кратность собственного значения.

Пусть f — линейный оператор n-мерного векторного пространства V над полем P.

Лемма. *Если $a_1,\ldots,a_k$ — собственные векторы линейного оператора f, соответствующие одному собственному значению λ_0, то любая ненулевая линейная комбинация этих векторов также является собственным вектором, соответствутющим собственному значению λ_0.*

♦ $f(\alpha_1a_1+\ldots+\alpha_na_k) = \alpha_1f(a_1)+\ldots+\alpha_nf(a_k) = \alpha_1\lambda_0a_1+\ldots+\alpha_k\lambda_0a_k = \lambda_0(\alpha_1a_1+\ldots+\alpha_na_k)$ ⊠

Следствие. *Множество, состоящее из нулевого вектора и всех собственных векторов линейного оператора f, соответствующих одному собственному значению λ_0, является подпросторанством пространства V.*

• *Подпространство, состоящее из нулевого вектора и всех собственных векторов линейного оператора, соответствующих одному собственному значению λ_0, называется* ***собственным подпространством*** *оператора, соответствующим собственному значению λ_0. Обозначение: $L(\lambda_0)$.*

• ***Геометрической кратностью*** *собственного значения линейного оператора называется размерность собственного подпространства, соответствующего этому собственному значению.*

Свойства геометрической кратности собственного значения:

1. *Геометрическая кратность r_0 собственного значения λ_0 равна*

$$r_0 = \dim V - \operatorname{rank}(M_f - \lambda_0E).$$

♦ Вектор X принадлежит собственному подпространству $L(\lambda_0) \Longleftrightarrow$ его координатный столбец является решением матричного уравнения $(M_f - \lambda_0E)X = 0 \Rightarrow$ размерность $L(\lambda_0)$ равна размерности пространства решений линейной системы $(M_f - \lambda_0E)X = 0$, то есть равна $\dim V - \operatorname{rank}(M_f - \lambda_0E)$. ⊠

2. *Геометрическая кратность собственного значения линейного оператора не превышает его алгебраической кратности.*

♦ Пусть собственное значение λ_0 имеет геометрическую кратность равную r_0 и пусть система векторов $(a_1,\ldots,a_r)$ — базис собственного подпространства $L(\lambda_0)$. Дополним его до базиса всего пространства V и построим матрицу M_f. Так как векторы $(a_1,\ldots,a_r)$ — собственные векторы оператора f, соответствующие собственному значению λ_0, то $f(a_i) = \lambda_0a_i,\ i = \overline{1,r}$. Тогда матрица M_f имеет вид

$$M_f = \left(\begin{array}{cccc|c} \lambda_0 & 0 & \ldots & 0 & \\ 0 & \lambda_0 & \ddots & 0 & B \\ 0 & 0 & \ldots & \lambda_0 & \\ \hline 0 & 0 & \ldots & 0 & C \end{array}\right) = \left(\begin{array}{c|c} \lambda_0E_r & B \\ \hline 0 & C \end{array}\right),$$ строим характеристичекий

многочлен: $det(M_f - \lambda E) = det((\lambda_0 - \lambda)E_r)det(C - \lambda E) = (\lambda - \lambda_0)^r det(C - \lambda E) \Rightarrow$ алгебраическая кратность собственного значения λ_0 не меньше r. ⊠

Следствие. *Сумма геометрических кратностей собственных занчений линейного оператора пространства V не превосходит размерности пространства V.*

⧫ $\sum\limits_i r_i \leqslant \sum\limits_i k_i \leqslant n = \dim V$ ⊠ ⊠

Теорема. *Линейный оператор векторного пространства V является оператором простой структуры $\Longleftrightarrow$ сумма геометрических кратностей собственных значений оператора равна размерности пространства V.*

⧫ $\Rightarrow$) Пусть f — операторо простой структуры. Тогда существует базис $A(a_1, \dots, a_n)$ пространства V, состоящий из собственных векторов этого оператора. Обозначим через $\lambda_1, \dots, \lambda_k$ собственные значения оператора f, а через $s_1, \dots, s_k$ — количество собственных векторов в базисе A, соответствующих этим собственным значениям. Тогда $s_1 + \dots + s_k = \dim V = n$. Так как количество векторов в линейно независимой системе векторов не может превышать размерности пространства, то количество собственных векторов в базисе A, соответствующих собственному значению λ_i, не превышает размерности собственного подпространства $L(\lambda_i))$, то есть $s_i \leqslant r_i$, где r_i — геометрическая кратность собственного значения $\lambda_i \Rightarrow$ [по предыдущему следствию] $s_1 + \dots + s_k \leqslant r_1 + \dots + r_k \leqslant n \Rightarrow r_1 + \dots + r_k \leqslant n \Rightarrow r_1 + \dots + r_k = n$.

$\Leftarrow$) Пусть сумма геометрических кратностей собственных значений линейного оператора f равна n. Тогда объединение базисов собственных подпространств линейного оператора состоит из n векторов, которые являются собственным векторами и линейно независимы, так как соответствуют различным собственным значениям $\Rightarrow$ построенная система векторов является базисом пространства V, состоящим из собственных векторов $f \Rightarrow$ оператор f — оператор простой структуры. ⊠

Глава 12

Полиномиальные матрицы

12.1 Эквивалентность полиномиальных матриц. Каноническая форма полиномиальной матрицы.

• ***Полиномиальным матрицами*** *называются матрицы, элементами которых являются многочлены над некоторым полем P.*

В дальнейшем будем рассматривать лишь квадратные полиномиальные матрицы с переменной λ вида:

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & \dots & a_{1n}(\lambda) \\ \vdots & \ddots & \vdots \\ a_{n1}(\lambda) & \dots & a_{nn}(\lambda) \end{pmatrix}.$$

• ***Элементарными преобразованиями*** *строк полиномиальной матрицы называются следующие два преобразования:*

1. *умножение строки полиномиальной матрицы на ненулевой элемент поля P;*
2. *прибавление к строке полиномиальной матрицы другой ее строки, умноженной на произвольный многочлен $f(\lambda) \in P[\lambda]$.*

Элементарные преобразования столбцов полиномиальной матрицы определяются аналогично.

• *Если матрица $B(\lambda)$ может быть получена из матрицы $A(\lambda)$ в результате применения к ней конечного числа элементарных преобразований строк и столбцов, то говорят, что матрица $B(\lambda)$* ***эквивалентна*** *матрице $A(\lambda)$. (Обозначение: $B(\lambda) \sim A(\lambda)$)*

Эквививалентость полиномиальных матриц рефлексивна, транзитивна и симметрична.

• *Диагональная матрица $K(\lambda) = diag(f_1(\lambda), \dots, f_n(\lambda))$ называется* ***канонической****, если*

1. *ненулевые многочлены $f_i(\lambda)$ имеют старший коэффициент 1;*
2. *каждый диагональный многочлен $f_i(\lambda)$, $i < n$, является делителем следующего элемента $f_{i+1}(\lambda)$.*

Теорема. *Для любой полиномиальной матрицы существует эквивалентная каноническая матрица.*

♦ Пусть $A(\lambda)$ — полиномиальная матрица порядка n. Если матрица $A(\lambda)$ нулевая, то она каноническая матрица. Пусть матрица $A(\lambda)$ ненулевая. Докажем существование эквивалентной канонической матрицы по индукции по порядку матрицы n.

Пусть $n = 1$. Тогда $A(\lambda) = (a(\lambda))$ (матрица состоящая из одного элемента). Умножим $A(\lambda)$ на элемент поля P, обратный старшему коэффициенту многочлена $a(\lambda)$ и получим каноническую, эквивалентную исходной.

Пусть $n > 1$ и пусть для любой матрицы порядка $n-1$ существует эквивалентная каноническая. Покажем, что матрица $A(\lambda)$ может быть приведена элементарными преобразованиями к канонической матрице:
Пусть S — класс эквивалентости, содержащий матрицы, эквивалентные матрице $A(\lambda)$. Возьмём из S матрицу $B(\lambda)$ такую, что

1. $b_{11}(\lambda) \neq 0$;
2. старший коэффициент b_{11} равен 1;
3. степень многочлена $b_{11}(\lambda)$ является наименьшей среди степеней всех многочленов, являющихся элементами, образующими выбранную матрицу из S.

Разделим многочлен $b_{12}(\lambda)$ с остатком на $b_{11}(\lambda)$, то есть $b_{12}(\lambda) = b_{11}(\lambda)q(\lambda) + r(\lambda)$, где $\deg\ r(\lambda) < \deg\ b_{11}(\lambda)$ или $r(\lambda) = 0$. Затем ко второму столбцу $B(\lambda)$ прибавим первый столбец, умноженный на $(-q(\lambda))$. Получим матрицу $B_1(\lambda)$, у которой элемент первой строки второго столбца равен $r(\lambda)$. Преобразование является элементарным $\Rightarrow B_1(\lambda) \sim B(\lambda) \Rightarrow B_1(\lambda) \in S$.
Если $r(\lambda) \neq 0$, то в S существует матрица $B_1(\lambda)$, у которой один из элементов имеет степень меньшую, чем степень $b_{11}(\lambda)$, что противоречит тому, что $\deg(b_{11}(\lambda))$ наименьшая в $S \Rightarrow r(\lambda) = 0 \Rightarrow$ первый элемент второго столбца B_1 равен 0.
Проведем аналогичные преобразования для остальных элементов первой строки и первого столбца матрицы $B(\lambda) \Rightarrow$ получим матрицу

$$\widetilde{B}(\lambda) = \left(\begin{array}{c|ccc} b_{11}(\lambda) & 0 & \dots & 0 \\ \hline 0 & c_{22}(\lambda) & \dots & c_{2n}(\lambda) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \end{array}\right) = \left(\begin{array}{c|c} b_{11}(\lambda) & 0 \\ \hline 0 & C(\lambda) \end{array}\right).$$

Матрица $C(\lambda)$ — полиномиальная матрица порядка $n-1 \Rightarrow$ по индуктивному предположению имеет эквивалентную каноническую матрицу $\widetilde{C}(\lambda) = diag(\widetilde{C}_2(\lambda), \dots, \widetilde{C}_n(\lambda))$.
Если эти преобразования применить к матрице $\widetilde{B}(\lambda)$, то они не будут менять элементы первой строки и первого столбца и приведут к диагональной матрице
$K(\lambda) = diag(b_{11}(\lambda), \widetilde{C}_2(\lambda), \dots, \widetilde{C}_n(\lambda))$. Покажем, что она является канонической.
Старшие коэффициенты ненулевых многочленов матрицы $A(\lambda)$ равны 1 и $\forall i \quad \widetilde{C}_i(\lambda)$ делит $\widetilde{C}_{i+1}(\lambda)$. Покажем, что $b_{11}(\lambda)$ делит $\widetilde{C}_1(\lambda)$.
Прибавим к первой строке вторую. Преобразование элементарное, поэтому эта матрица $\in S$. По доказанному выше, для любой матрицы из S, первый элемент которой b_{11}, все элементы первой строки делятся на $b_{11} \Rightarrow \widetilde{C}_2(\lambda)|b_{11} \Rightarrow \widetilde{C}_2(\lambda)$ — каноническая матрица. А значит и матрица $K(\lambda)$ каноническая. ⊠

• *Каноническая матрица $K(\lambda)$, эквивалентная матрице $A(\lambda)$, называется* ***канонической формой*** *матрицы $A(\lambda)$.*

12.2 Единственность канонической формы. Система НОД миноров полиномиальных матриц.

• *Пусть ранг полиномиальной матрицы $A(\lambda)$ равен r. Тогда все миноры матрицы $A(\lambda)$ порядка выше r равны нулю, а среди миноров порядка $k \in \{1, \dots, r\}$ существует ненулевой. Система многочленов $d_1(\lambda)$, ..., $d_n(\lambda)$ называется* ***системой НОД миноров*** *матрицы $A(\lambda)$, если при $k \in \{1, \dots, r\}$ многочлен $d_k(\lambda)$ равен НОД всех миноров k-ого порядка матрицы $A(\lambda)$ со старшим коэффициентом 1, а при $k \in \{r+1, \dots, n\}$ $d_k(\lambda) = 0$.*

Лемма. *Системы НОД миноров эквивалетных полиномиальных матриц совпадают.*

⧫ Покажем, что НОД миноры не изменяются при элементарных преобразованиях строк:

1. Пусть i-ая строка матрицы $A(\lambda)$ умножена на ненулевой скаляр α. Тогда миноры, в которые эта строка не входит, не изменяются, а миноры, в которые входят элементы этой строки, умножаются на α. Но НОД системы многочленов не меняется, если некоторые из этих многочленов умножаются на ненулевую константу.

2. Пусть матрица $B(\lambda)$ получена из матрицы $A(\lambda)$ прибавлением к i-ой строке элементов j-ой строки, умноженных на многочлен $c(\lambda)$. При этом миноры, в которые не входят элементы i-ой строки, не изменяются. Не меняются также миноры, в которые входят одновременно и элементы i-ой строки, и элементы j-ой. Рассмотрим миноры, в которые входят элементы i-ой строки, но не входят элементы j-ой, то есть

$$|a_{i1}(\lambda) + c(\lambda)a_{j1}(\lambda) + \ldots + a_{in}(\lambda) + a_{jn}(\lambda)c(\lambda)| =$$
$$= |\underbrace{a_{i1}(\lambda) \ldots a_{in}(\lambda)}_{M_1}| + c(\lambda)|\underbrace{a_{j1}(\lambda) + \ldots + a_{jn}(\lambda)}_{M_2}| = M_1 + c(\lambda)M_2.$$

   Заметим, что определители M_1 и M_2 с точностью до знака являются минорами матрицы $A(\lambda)$. Следовательно, все миноры матрицы $B(\lambda)$ либо равны минорам матрицы $A(\lambda)$, либо линейно выражаются через них.

   Пусть $d_k^A(\lambda)$ и $d_k^B(\lambda)$ — k-тые многочлены в системе НОД миноров матриц $A(\lambda)$ и $B(\lambda)$ соответственно. И пусть $\operatorname{rank} A(\lambda) = r$. Тогда если $k \in \{r+1, \dots, n\}$, то $d_k^A(\lambda) = 0 \Rightarrow$ все миноры k-того порядка матрицы $A(\lambda)$ равны $0 \Rightarrow$ все миноры k-того порядка матрицы $B(\lambda)$ также равны $0 \Rightarrow d_k^B(\lambda) = 0$.

   Пусть $k \in \{1, \dots, r\}$. Тогда многочлен $d_k^A(\lambda)$ — НОД миноров k-того порядка матрицы $A(\lambda) \Rightarrow$ он делит все миноры k-того порядка матрицы $A(\lambda) \Rightarrow$ он делит все миноры k-того порядка и матрицы $B(\lambda) \Rightarrow d_k^A(\lambda)$ делит НОД миноров k-того порядка матрицы B, то есть $d_k^A(\lambda) \mid d_k^B(\lambda)$.

   Заметим, что матрица $A(\lambda)$ может быть получена из матрицы $B(\lambda)$ с помощью обратного элементарного преобразования $\Rightarrow d_k^B(\lambda) \mid d_k^A(\lambda)$, и так как их старшие коэффициенты равны 1, то $d_k^B(\lambda) = d_k^A(\lambda)$.

⊠

Теорема. *Каноническая форма полиномиальной матрицы определяется однозначно.*

♦ Пусть $A(\lambda)$ — произвольная полиномиальная матрица, $\operatorname{rank} A(\lambda) = r$, $d_1(\lambda), \ldots, d_n(\lambda)$ — система НОД миноров матрицы $A(\lambda)$ и матрица $K(\lambda) = diag(f_1(\lambda), \ldots, f_n(\lambda))$ — ее некоторая каноническая форма. Тогда $K(\lambda) \sim A(\lambda) \Rightarrow$ их системы НОД миноров совпадают

$$\begin{aligned}
&d_1(\lambda) = f_1(\lambda),\\
&d_2(\lambda) = f_1(\lambda) f_2(\lambda),\\
&d_3(\lambda) = f_1(\lambda) f_2(\lambda) f_3(\lambda),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&d_r(\lambda) = f_1(\lambda) f_2(\lambda) \ldots f_r(\lambda),
\end{aligned}$$

Так как $\operatorname{rank} A(\lambda) = r$, то $d_{r+1} = \ldots = d_n(\lambda) = 0$, следовательно, $f_{r+1}(\lambda) = \ldots = f_n(\lambda) = 0$. Так как в противном случае у матрицы $K(\lambda)$ существует ненулевой минор порядка выше n.
Таким образом, диагональные элементы матрицы $K(\lambda)$ однозначно определены системой НОД миноров матрицы $A(\lambda)$

$$\begin{cases}
f_1(\lambda) = d_1(\lambda),\\
f_2(\lambda) = \dfrac{d_2(\lambda)}{d_1(\lambda)},\\
f_3(\lambda) = \dfrac{d_3(\lambda)}{d_2(\lambda)},\\
\ldots\ldots\ldots\ldots\ldots\\
f_r(\lambda) = \dfrac{d_r(\lambda)}{d_{r-1}(\lambda)}.
\end{cases}$$

⊠

• *Диагональные элементы $f_1(\lambda), \ldots, f_n(\lambda)$ канонической формы $K(\lambda)$ полиномиальной матрицы $A(\lambda)$ называются* ***инвариантными множителями*** *матрицы $A(\lambda)$.*

Полученные в ходе доказательства системы формул устанавливают связь между системой НОД миноров и системой инвариантных множителей полиномиальной матрицы ранга r.

Следствие (критерий эквивалентности полиномиальнымъ матриц). *Две полиномиальные матрицы эквивалентны $\Longleftrightarrow$ их системы инвариантных множителей совпадают.*

Следствие (критерий эквивалентности полиномиальнымъ матриц). *Две полиномиальные матрицы эквивалентны $\Longleftrightarrow$ их системы НОД миноров совпадают.*

12.3 Система элементарных делителей полиномиальной матрицы.

• *Любой многочлен $f(\lambda) \in P[\lambda]$ единственным образом представим в виде*

$$f(\lambda) = a \cdot (g_1(\lambda))^{k_1}\ (g_2(\lambda))^{k_2} \ldots (g_s(\lambda))^{k_s},$$

где $a \in P$ — старший коэффициент многочлена $f(\lambda)$, $g_i(\lambda)$ — различные неприводимые многочлены со старшим коэффициентом 1. Многочлены $(g_i(\lambda))^{k_i}$ называются ***элементарными делителями*** *многочлена $f(\lambda)$).*

• ***Системой элементарных делителей*** *полиномиальной матрицы называется совокупность элементарных делителей всех непостоянных инвариантных множителей полиномиальной матрицы. При этом каждый элементарный делитель включается в эту систему столько раз, сколько он является элементарным делителем инвариантных множителей.*

Теорема (критерий эквивалентности полиномиальных матриц). *Две полиномиальные матрицы эквивалентны $\Longleftrightarrow$ совпадают их системы элементарных делителей и ранги.*

⧫ $\Rightarrow$) Если две полиномиальные матрицы одного порядка эквивалентны, то

1. Их системы НОД миноров совпадают, следовательно, совпадают и их ранги.

2. Их инвариантные множители совпадают.

Значит, их системы элементарных делителей совпадают.

$\Leftarrow$) Пусть полиномиальная матрица $A(\lambda)$ имеет порядок n, ранг r и систему элементарных делителей равную S. Покажем, что можно однозначно построить $K(\lambda)$: Если $r < n$, то $f_{r+1} = \ldots = f_n(\lambda) = 0$.

Пусть система S состоит из степеней неприводимых многочленов $p_1(\lambda), \ldots, p_s(\lambda)$. Так как многочлен $f_r(\lambda)$ делится на каждый из предыдущих многочленов $f_i(\lambda)$, то он равен произведению всех многочленов $f_i(\lambda)$ с максимальной степенью среди всех имеющихся в системе S. Удалив из системы S элементарные делители, порождаемые многочленом $f_r(\lambda)$, получим новую систему многочленов S_1, с помощью которой аналогичным образом восстановим многочлены $f_{r-1}(\lambda)$ и так далее.

В результате построения мы исчерпаем всю систему S, так как произведение многочленов системы элементарных делителей равно произведению ненулевых инвариантных множителей, при этом возможны два случая:

1. На последнем шаге находим $f_1(\lambda)$ и построение закончено;

2. На последнем шаге получаем $f_j(\lambda)$, где $j > 1$. Тогда $f_1(\lambda) = \ldots = f_{j-1}(\lambda) = 1$.

⊠

Теорема. *Система элементарных делителей диагональной полиномиальной матрицы равна объединению систем элементарных делителей ее диагональных элементов. При этом каждый элементарный делитель включается в эту систему столько раз, сколько раз он является элементарным делителем диагональных многочленов.*

⧫ Рассмотрим диагональную полиномиальную матрицу $A(\lambda)$ порядка n такую, что $A(\lambda) = diag\{g_1(\lambda), \ldots, g_n(\lambda)\}$. Без ограничения общности считаем, что старшие коэффициенты у ненулевых многочленов $g_i(\lambda)$ равны 1 и $\begin{cases} g_i(\lambda) \neq 0, \; i = \overline{1, r}, \\ g_i(\lambda) = 0, \; i = \overline{r+1, n}. \end{cases}$

Тогда $\operatorname{rank} A(\lambda) = r$. Пусть $f_1(\lambda), \ldots, f_n(\lambda)$ — система инвариантных множителей матрицы $A(\lambda)$ и $d_1(\lambda), \ldots, d_n(\lambda)$ — система НОД миноров матрицы $A(\lambda)$. Так как $\operatorname{rank} A(\lambda) = r$, то $d_i(\lambda) = f_i(\lambda) = 0 \; \forall i = \overline{r+1, n}$. А так как $A(\lambda)$ имеет лишь один ненулевой минор r-ого порядка, то $d_r(\lambda) = g_1(\lambda) \cdot \dots \cdot g_r(\lambda) = f_1(\lambda) \cdot \dots \cdot f_r(\lambda) \Rightarrow$ элементарные делители $g_i(\lambda)$ и $f_i(\lambda)$ являются степенями одних и тех же неприводимых многочленов $p_1(\lambda), \ldots, p_s(\lambda)$. Разложим $g_i(\lambda)$ по $p_i(\lambda)$:

$$g_1(\lambda) = (p_1(\lambda))^{\alpha_{11}} \cdot (p_2(\lambda))^{\alpha_{12}} \cdot \ldots \cdot (p_k(\lambda))^{\alpha_{1k}},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$g_r(\lambda) = (p_1(\lambda))^{\alpha_{r1}} \cdot (p_2(\lambda))^{\alpha_{r2}} \cdot \ldots \cdot (p_k(\lambda))^{\alpha_{rk}}.$$

Каждую из систем чисел $(\alpha_{1i}, \ldots, \alpha_{ri})$ переобозначим следующим образом:
Через β_{1i} обозначим минимульное среди них число, β_{i2} — следующее в порядке возрастания и так далее, а через β_{ir} — максимальное среди всех чисел. Тогда

$$d_1(\lambda) = (p_1(\lambda))^{\beta_{11}}(p_2(\lambda))^{\beta_{12}} \cdot \ldots \cdot (p_k(\lambda))^{\beta_{1k}},$$
$$d_2(\lambda) = (p_1(\lambda))^{\beta_{11}+\beta_{21}}(p_2(\lambda))^{\beta_{12}+\beta_{22}} \cdot \ldots \cdot (p_k(\lambda))^{\beta_{1k}+\beta_{2k}},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Тогда

$$f_1(\lambda) = d_1(\lambda) = (p_1(\lambda))^{\beta_{11}} \cdot \ldots \cdot (p_k(\lambda))^{\beta_{1k}},$$
$$f_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)} = (p_1(\lambda))^{\beta_{21}} \cdot \ldots \cdot (p_k(\lambda))^{\beta_{2k}},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f_r(\lambda) = \frac{d_r(\lambda)}{d_{r-1}(\lambda)} = (p_1(\lambda))^{\beta_{r_1}} \cdot \ldots \cdot (p_k(\lambda))^{\beta_{r_k}}.$$

Значит, система элементарных делителей инвариантных множителей матрицы $A(\lambda)$ состоит из тех же многочленов, что и объединение СЭД ее диагональных элементов. ⊠

Теорема. *Система элементарных делителей блочнодиагональной полиномиальной матрицы равна объединению систем элементарных делителей ее диагональных блоков.*

⧫ Пусть $A(\lambda)$ — блочнодиагональная матрица вида $A(\lambda) = diag(A_1(\lambda), \ldots, A_s(\lambda))$, где $A_i(\lambda)$ — полиномиальная квадратная матрица порядка n_i. Каждая матрица $A_i(\lambda)$ имеет единственную каноническую форму $K_i(\lambda) = diag(f_{1i}(\lambda), \ldots, f_{n_i i}(\lambda))$. Тогда система элементарных делителей блока $A_i(\lambda)$ равна объединению систем элементарных делителей многочленов $f_{ij}(\lambda)$. Элементарные преобразования матриц $A_i(\lambda)$ можно рассматривать, как элементарные преобразования матрицы $A(\lambda)$, затрагивающие лишь строки и столбцы, в которых расположен блок $A_i(\lambda)$. Значит, $A(\lambda) \sim diag(f_{11}(\lambda), \ldots, f_{n_1 n}(\lambda), f_{12}(\lambda), \ldots, f_{n_s s}(\lambda))$. Следовательно, система элементарных делителей ее диагональных элементов, с одной стороны, является объединением систем элементарных делителей диагональных блоков $A_i(\lambda)$, а с другой стороны, по предыдущей теореме, системой элементарных делителей всей матрицы $A(\lambda)$. ⊠

12.4 Унимодулярные матрицы.

• *Полиномиальная матрица называется* ***унимодулярной****, если её определитель равен отличному от нуля скаляру.*

Свойства унимодулярных матриц:

1. *Полиномиальная матрица унимодулярна* $\iff$ *все её инвариантные множители равны 1, то есть она эквивалентна единичной матрице.*

   ⧫ Полиномиальная матрица унимодулярна $\iff d_n(\lambda) = 1 \iff f_1(\lambda) \ldots f_n(\lambda) = 1 \iff f_i(\lambda) = 1 \quad \forall i.$ ⊠

2. *Полиномиальная матрица унимодулярна* $\iff$ *она обратима во множестве полиномиальных матриц.*

♦ $\Rightarrow$) Пусть полиномиальная матрица $A(\lambda)$ унимодулярна. Тогда $det A(\lambda) \in P, \neq 0 \Rightarrow \dfrac{1}{det A} \in P, \neq 0$. Так как матрица $A(\lambda)$ полиномиальная, ее алгебраические дополнения являются многочленами. Следовательно, присоединенная матрица $\widetilde{A}(\lambda)$ также является полиномиальной $\Rightarrow \dfrac{1}{det A}\widetilde{A}(\lambda)$ является полиномиальной матрицей и обратной матрице $A(\lambda)$.

$\Leftarrow$) Пусть матрица $A(\lambda)$ имеет обратную полиномиальную матрицу $A^{-1}(\lambda)$. Тогда $A^{-1}(\lambda) \cdot A(\lambda) = E \Rightarrow det A(\lambda) det A^{-1}(\lambda) = 1$. Так как матрицы $A(\lambda), A^{-1}(\lambda)$ являются полиномиальными, их определители — многочлены $\Rightarrow det A(\lambda)$ — ненулевой элемент поля P, то есть матрица $A(\lambda)$ унимодулярна. ⊠

• *Полиномиальные матрицы вида*

$$S_i(\alpha) = \begin{array}{cc} & \begin{array}{ccccccc} & & & i & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & \alpha & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \end{array}, \quad \alpha \in P \backslash \{0\},$$

$$S_{i,j}(f(\lambda)) = \begin{array}{cc} & \begin{array}{ccccccc} & & j & & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \\ i \\ \\ \\ \end{array} & \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & f(\lambda) & \dots & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}, f(\lambda) \in P[\lambda]$$

называются ***элементарными***.

Элементарные преобразования строк (столбцов) полиномиальной матрицы эквивалентны умножению матрицы слева (справа) на соответствующую элементарную матрицу.

Теорема (критерий эквивалентности полиномиальных матриц). *Две полиномиальные матрицы $A(\lambda)$ и $B(\lambda)$ эквивалентны $\Longleftrightarrow$ существуют унимодулярные матрицы $U(\lambda)$ и $V(\lambda)$ того же порядка такие, что*

$$B(\lambda) = U(\lambda)A(\lambda)V(\lambda).$$

♦ Матрицы $A(\lambda)$ и $B(\lambda)$ эквивалентны $\Longleftrightarrow$ существует конечное число элементарных преобразований строк и столбцов, переводящих матрицу $A(\lambda)$ в $B(\lambda)$. Каждое элементарное преобразование эквивалентно домножению матрицы слева (справа) на соответствующую элементарную матрицу. Но каждая элементарная матрица является унимодулярной, и произведение унимодулярных матриц также является унимодулярной матрицей, то есть

$$\underbrace{U_1 \dots U_s}_{U(\lambda)} A(\lambda) \underbrace{V_1 \dots V_k}_{V(\lambda)} = B(\lambda).$$

⊠

Следствие. *Полиномиальная матрица унимодулярна $\Longleftrightarrow$ она представима в виде произведения элементарных матриц.*

♦ Матрица $A(\lambda)$ является унимодулярной матрицей $\Longleftrightarrow$ $A(\lambda) \sim E \Longleftrightarrow \exists U(\lambda), V(\lambda)$: $A(\lambda) = U(\lambda)EV(\lambda) = U(\lambda)V(\lambda)$, причем матрицы $U(\lambda)$ и $V(\lambda)$ — произведения элементарных матриц. ⊠

12.5 Матричные многочлены.

• ***Матричным многочленом*** *с переменной λ называется выражение вида*

$$A(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0 = \sum_{i=0}^{m} A_i\lambda^i,$$

где A_i — квадратные матрицы одного и того же порядка над полем P.

• *Порядок матриц A_i называется* ***порядком*** *матричного многочлена $A(\lambda)$.*

• *Если $A_m \neq 0$, то число m называется* ***степенью*** *матричного многочлена.*

Любая полиномиальная матрица представима в виде матричного многочлена, и, наоборот, любой матричный многочлен представим в виде полиномиальной матрицы.

• *Матричный многочлен называется* ***регулярным****, если определитель его мтаршего коэффициента не равен нулю, то есть $det A_m \neq 0$.*

• *Два матричных многочлена называются* ***равными****, если матрицы, стоящие в этих многочленах при одинаковых степенях переменной λ_i равны.*

Пусть $A(\lambda) = \sum\limits_{i=0}^{m} A_i\lambda^i$ и $B(\lambda) = \sum\limits_{i=0}^{k} B_i\lambda^i$, при $m \geqslant k$ — матричные многочлены.

• ***Суммой*** *матричных многочленов $A(\lambda)$ и $B(\lambda)$ называется многочлен*

$$(A+B)(\lambda) = \sum_{i=0}^{m}(A_i + B_i)\lambda^i.$$

• ***Произведением*** *матричных многочленов $A(\lambda)$ и $B(\lambda)$ называется многочлен*

$$(AB)(\lambda) = \sum_{i=0}^{m}\sum_{j=0}^{k}(A_iB_j)\lambda^{i+j}.$$

Степень суммы матричных многочленов не превосходит максимальной из степеней слагаемых. Степень произведения матричных многочленов не превосходит суммы степеней множителей. Если один из множителей является регулярным многочленом, то степень произведения равна сумме степеней множителей.

Теорема. *Для произвольного матричного многочлена $A(\lambda)$ и регулярного многочлена $B(\lambda)$*

1. *существует единственная пара многочленов $Q(\lambda)$ и $R(\lambda)$ таких, что*

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda),$$

причем или $R(\lambda) = 0$, или $\deg R(\lambda) < \deg B(\lambda)$;

2. *существует единственная пара многочленов* $Q'(\lambda), R'(\lambda)$ *таких, что*

$$A(\lambda) = B(\lambda)Q'(\lambda) + R'(\lambda),$$

причем $R'(\lambda) = 0$ *или* $\deg R'(\lambda) < \deg B(\lambda)$*;*

• *Матричные многочлены* $Q(\lambda)$ *и* $R(\lambda)$ *в представлении* $A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda)$ *называются соответственно* ***частным и остатком при делении справа*** *матричного многочлена* $A(\lambda)$ *на регулярный многочлен* $B(\lambda)$*. Аналогично многочлены* $Q'(\lambda), R'(\lambda)$ *называются* ***частным и остатком при делении слева*** *многочлена* $A(\lambda)$ *на* $B(\lambda)$*.*

Пусть $A(\lambda) = \sum\limits_{i=0}^{k} A_i\lambda^i$ — матричный многочлен порядка n, S — квадратная матрица над полем P порядка n.

• *Матрица* $A_R(S) = \sum\limits_{i=0}^{m} A_iS^i$ *называется* ***правым значением****, а* $A_L(S) = \sum\limits_{i=0}^{m} S^iA_i$ — ***левым значением*** *многочлена* $A(\lambda)$ *при замене* λ *матрицей* S*.*

Свойства правых значений:

1. $A(\lambda) = B(\lambda) \Rightarrow A_R(S) = B_R(S), \forall S$;
2. $(A + B)_R(S) = A(S)_R + B(S)_R, \forall S$ *(для произведения неверно).*

Вообще говоря, утверждение $(AB)_R(S) \neq A_R(S)B_R(S)$ неверно только в частном случае, так как $\sum\limits_{i=0}^{m}\sum\limits_{j=0}^{k} A_iB_iS^{i+j} \neq \sum\limits_{i=0}^{m}\sum\limits_{j=0}^{k} A_iS^iB_jS^j$.

Обобщенная теорема Безу. *Остаток при делении справа (слева) матричного многочлена* $A(\lambda)$ *на многочлен* $(\lambda E - C)$ *равен правому(левому) значению* $A(\lambda)$ *при* $\lambda = C$*.*

⧫ Так как многочлен $(\lambda E - C)$ является регулярным, то существуют такие $Q(\lambda), R(\lambda)$, что $A(\lambda) = Q(\lambda)(\lambda E - C) + R(\lambda)$, где или $R(\lambda) = 0$, или $\deg R(\lambda) < \deg(\lambda E - C) = 1$, то есть матрица $R(\lambda) \in P_{n,n}$.
Пусть $\deg A(\lambda) = m > 0$. Тогда $\deg Q(\lambda) = m - 1$, так как $\deg(Q(\lambda)(\lambda E - C)) = \deg A(\lambda)$. Отсюда получаем

$$Q(\lambda) = Q_{m-1}\lambda^{m-1} + \ldots + Q_1\lambda + Q_0,$$
$$A(\lambda) = A_m\lambda^m + \ldots + A_1\lambda + A_0 = (Q_{m-1}\lambda^{m-1} + \ldots + Q_1\lambda + Q_0)(\lambda E - C) + R.$$

Приравниваем коэффициенты при соответствующих степенях λ:

$$\begin{cases} \lambda^m : A_m = Q_{m-1}, \\ \lambda^{m-1} : A_{m-1} = -Q_{m-1}C + Q_{m-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda : A_1 = -Q_1C + Q_0, \\ \lambda^0 : A_0 = -Q_0C + R; \end{cases} \text{, выразим значение } R:$$

$R = A_0 + Q_0C = A_0 + (A_1 + Q_1C)C = A_0 + A_1C + Q_1C^2 = \ldots = A_0 + A_1C + A_2C^2 + \ldots + Q_{m-1}C^m = [Q_{m-1} = A_m] = A_0 + A_1C + A_2C^2 + \ldots + A_mC^m = A_R(C)$. ⊠

12.6 Критерий подобия матриц.

Теорема (критерий подобия матриц). *Две квадратные матрицы над полем P одного порядка подобны $\Longleftrightarrow$ их характеристические матрицы эквивалентны.*

♦ $\Rightarrow$) Пусть квадратные матрицы матрицы A и B одного порядка n подобны, то есть существует матрица S такая, что $A = S^{-1}BS$. Тогда $A - \lambda E = S^{-1}BS - \lambda S^{-1}ES = S^{-1}(B - \lambda E)S$. Так как S — невырожденная унимодулярная матрица над полем P, то и матрица S^{-1} также является полиномиальной унимодулярной матрицей. По теореме об унимодулярных матрицах, $(A - \lambda E) \sim (B - \lambda E)$.

$\Leftarrow$) Пусть $(B - \lambda E) \sim (A - \lambda E)$. Тогда, по теореме об унимодулярных матрицах, существуют унимодулярные матрицы $U(\lambda), V(\lambda)$ такие, что $(B - \lambda E) = U(\lambda) \cdot (A - \lambda E) \cdot V(\lambda)$. Так как матрица $U(\lambda)$ унимодулярна, то, по свойству существования обратной во множестве полиномиальной матрицы, существует полиномиальная матрица $U^{-1}(\lambda)$ такая, что $U^{-1}(\lambda) \cdot (B - \lambda E) = (A - \lambda E) \cdot V(\lambda)$.

Разделим $U^{-1}(\lambda)$ и $V(\lambda)$ на регулярные $(A - \lambda E), (B - \lambda E)$ слева и справа соответственно, то есть представим в виде:

$$\begin{cases} U^{-1}(\lambda) = (A - \lambda E) \cdot U_1(\lambda) + U_2, \\ V(\lambda) = V_1(\lambda) \cdot (B - \lambda E) + V_2. \end{cases}$$

Заметим, что U_2, V_2 — квадратные матрицы порядка n, так как их степень должна быть нулевой. Тогда подставим и получим

$$((A - \lambda E)U_1(\lambda) + U_2)(B - \lambda E) = (A - \lambda E)(V_1(B - \lambda E) + V_2),$$

$$\underbrace{(A - \lambda E)}_{\deg=1}(U_1(\lambda) - V_1(\lambda))\underbrace{(B - \lambda E)}_{\deg=1} = \underbrace{(A - \lambda E)V_2 - U_2(B - \lambda E)}_{const};$$

Если $U_1 - V_1 \neq 0$, то степень слева $\geqslant 2$, а справа $\leqslant 1$, что является противоречием. Следовательно, $U_1 - V_1 = 0$, значит значение справа равно нулю $\Rightarrow (A - \lambda E)V_2 - U_2(B - \lambda E) = 0 \Leftrightarrow (A - \lambda E)V_2 = U_2(B - \lambda E)$. Приравняем коэффициенты при соответствующих степенях:

$$\text{при } \lambda^1 : -EV_2 = U_2(-E) \Rightarrow V_2 = U_2,$$
$$\text{при } \lambda^0 : AV_2 = U_2B;$$

Покажем, что матрица U_2 невырожденная. Для этого разделим $U(\lambda)$ слева на $(B - \lambda E)$, то есть представим в виде $U(\lambda) = (B - \lambda E)U_3(\lambda) + U_4(\lambda)$ и домножим слева на $U^{-1}(\lambda)$. Тогда получим $U^{-1}(\lambda)U(\lambda) = U^{-1}(B - \lambda E)U_3(\lambda) + U^{-1}(\lambda)U_4(\lambda) \Rightarrow$
$E = (A - \lambda E)V(\lambda)U_3(\lambda) + ((A - \lambda E)U_1(\lambda) + U_2)U_4(\lambda) \Rightarrow E = \underbrace{(A - \lambda E)}_{\deg=1}\underbrace{(VU_3 + U_1U_4)}_{Q} + U_2U_4$.

Если $Q \neq 0$, то слева $\deg E = 0$, а справа степень $\geqslant 1$, что является противоречием. Значит, $Q = 0 \Rightarrow E = U_2U_4 \Rightarrow detU_2 \neq 0 \Rightarrow \exists U_2^{-1} \Rightarrow$ из того, что $AV_2 = U_2B$, получаем $B = U_2^{-1}AU_2$, то есть матрица A подобна матрице B. ⊠

12.7 Минимальный многочлен матрицы.

• *Пусть*

$$f(\lambda) = \alpha_m\lambda^m + \alpha_{m-1}\lambda^{m-1} + \ldots + \alpha_1\lambda + \alpha_0$$

— некоторый многочлен над полем P, A — квадратная матрица порядка n над полем P. Матрица

$$f(A) = \alpha_m A^m + \alpha_{m-1} A^{m-1} + \ldots + \alpha_1 A + \alpha_0 E$$

называется ***значением*** *многочлена f при $\lambda = A$.*

Свойства значений многочленов:

1. $(f+g)(A) = f(A) + g(A)$;
2. $(f \cdot g)(A) = f(A) \cdot g(A)$.

   ♦ Так как нет других матриц кроме матрицы A. ⊠

• *Ненулевой многочлен $f(\lambda)$ называется* ***аннулирующим*** *многочленом матрицы $A \in P_{n,n}$, если $f(A) = 0_{n,n}$.*

Так как $P_{n,n}$ — векторное пространство размерности n^2, то система матрицы $A^{n^2}, A^{n^2-1}, \ldots, A^1, A^0$ является линейно зависимой. Следовательно, существует нетривиальная линейная комбинация $\alpha_{n^2} A^{n^2} + \ldots + \alpha_1 A^1 + \alpha_0 A^0 = 0_{n,n}$. Отсюда получаем, что любой многочлен с теми же коэффициентами является ануллирующим для матрицы A и для любой матрицы A существует аннулирующий многочлен.

• ***Минимальным многочленом*** *называется аннулирующий многочлен, старший коэффициент которого равен 1, а степень минимальная из встепеней всех ненулевых аннулирующих многочленов матрицы.*

Свойства минимального многочлена:

1. *Минимальный многочлен определён однозначно.*

   ♦ Пусть матрица A имеет два различных минимальных многочлена: $m_1(\lambda)$, $m_2(\lambda)$. Рассмотрим многочлен $m(\lambda) = m_1(\lambda) - m_2(\lambda)$.

   (a) $m(\lambda) \neq 0$;

   (b) $m(\lambda)$ — аннулирующий многочлен для матрицы A, так как $m(A) = m_1(A) - m_2(A) = 0_{n,n} - 0_{n,n} = 0_{n,n}$;

   (c) Так как старшие коэффициенты $m_1(\lambda)$ и $m_2(\lambda)$ равны 1 и $\deg\ m_1(\lambda) = \deg\ m_2(\lambda)$, то степень $m(\lambda)$ меньше степеней $m_1(\lambda)$и $m_2(\lambda)$. Получаем противоречие с тем, что степени $m_1(\lambda)$и $m_2(\lambda)$ наименьшие среди всех степеней аннулирующий многочленов. Следовательно, $m_1(\lambda) = m_2(\lambda)$.

   ⊠

2. *Многочлен является аннулирующим многочленом матрицы $\Longleftrightarrow$ он делится на минимальный многочлен этой матрицы.*

   ♦ $\Rightarrow$) Пусть многочлен $f(\lambda)$ аннулирующий, а $m(\lambda)$ — минимальный многочлен. Разделим $f(\lambda)$ на $m(\lambda)$ с остатком:
   $f(\lambda) = m(\lambda) \cdot q(\lambda) + r(\lambda)$, где $r(\lambda) = 0$ или $\deg(r(\lambda)) < \deg(m(\lambda))$.
   Тогда $\underbrace{f(A)}_{0} = \underbrace{m(A)}_{0} \cdot q(A) + r(A) \Rightarrow r(A) = 0$. Если $r(\lambda) \neq 0$, то многочлен $r(\lambda)$ аннулирующий, степень которого меньше степени $m(\lambda)$, что является противоречием. Значит, $r(\lambda) = 0$, то есть $m(\lambda)$ делит $f(\lambda)$.

$\Leftarrow$) Любой многочлен $f(\lambda)$ представимый в виде $f(\lambda) = m(\lambda)q(\lambda)$ является аннулирующим для матрицы A, так как $f(A) = m(A)q(A) = 0$, то есть многочлен $f(\lambda)$ является аннулирующим. ⊠

Теорема. *Минимальный многочлен матрицы равен последнему инвариантному множителю ее характеристической матрицы.*

⧫ Пусть $d_1(\lambda), \ldots, d_n(\lambda)$ и $f_1(\lambda), \ldots, f_n(\lambda)$ — системы НОД миноров и инвариантных множителей матрицы $A - \lambda E$ соответственно, $d_i(\lambda) \neq 0 \ \forall i = \overline{1,n}$. Обозначим через $B(\lambda)$ матрицу, присоединенную к матрице $A - \lambda E$, то есть $(A-\lambda E)^{-1} = \dfrac{1}{det(A-\lambda E)} \cdot B(\lambda)$. Тогда по свойству присоединенной матрицы

$$B(\lambda)(A - \lambda E) = det(A - \lambda E)E_n. \tag{12.7.1}$$

Матрица $B(\lambda)$ по определению состоит из алгебраических дополнений элементов матрицы $A-\lambda E$, которые с точностью до знака являются минорами $(n-1)$-го порядка этой матрицы. Следовательно, матрица $B(\lambda)$ представима в виде $B(\lambda) = d_{n-1}(\lambda)C(\lambda)$, где $C(\lambda)$ — матрица НОД элементов, которые равны 1, так как $d_{n-1}(\lambda)$ — НОД. С другой стороны, $det(A - \lambda E) = (-1)^n d_n(\lambda)$ и при этом $f_n(\lambda) = \dfrac{d_n(\lambda)}{d_{n-1}(\lambda)}$. Следовательно, $d_n(\lambda) = d_{n-1}(\lambda)f_n(\lambda)$ и $det(A - \lambda E) = (-1)^n d_{n-1}(\lambda)f_n(\lambda)$.
Полученные равенства подставим в (12.7.1):

$$d_{n-1}(\lambda)C(\lambda)(A - \lambda E) = (-1)^n d_{n-1}(\lambda)f_n(\lambda)E_{n,n}.$$

Так как ранг матрицы $A - \lambda E$ равен порядку этой матрицы (потому что ее определитель равен характеристическому многочлену), то многочлен $d_{n-1}(\lambda)$ ненулевой. Следовательно, мы можем сократить на данный многочлен и получить $C(\lambda)(A - \lambda E) = (-1)f_n(\lambda)E_{n,n}$. Тогда

$$f_n(\lambda)E_n = (-1)^n C(\lambda)(A - \lambda E). \tag{12.7.2}$$

Пусть $f_n(\lambda) = \alpha_k\lambda^k + \ldots + \alpha_1\lambda + \alpha_0$. Тогда $f_n(\lambda)E = (\alpha_k E)\lambda^k + \ldots + (\alpha_1 E)\lambda + \lambda_0 E$ — матричный многочлен. И из уравнения (12.7.2) следует, что $f_n(\lambda)E$ делится справа на матричный многочлен $A - \lambda E$ (так как он равен произведению матричных многочленов). Тогда по обобщенной теореме Безу его правое значение при $\lambda = A$ равно нулю, то есть $f_n(\lambda) = \alpha_k\lambda^k + \ldots + \alpha_1\lambda + \alpha_0 = \alpha_k A^k + \ldots + \alpha_1 A + \alpha_0 E$. Следовательно, многочлен $f_n(\lambda)$ аннулирующий.
Покажем, что многочлен $f_n(\lambda)$ минимальный. Так как он аннулирующий, то он делится на минимальный многочлен $m(\lambda)$, то есть представим в виде $f_n(\lambda) = q(\lambda)m(\lambda)$, причем, так как $f_n(\lambda)$ и $m(\lambda)$ — многочлены со старшим коэффициентом 1, то многочлен $q(\lambda)$ также имеет старший кожффициент равный 1.
Так как многочлен $m(\lambda)$ аннулирующий, то $m(A) = 0$. Следовательно, $m(A)E = 0$ и матричный многочлен $m(\lambda)E$ имеет правое значение равное нулю. Тогда по обобщенной теореме Безу $A - \lambda E$ делит $m(\lambda)E$, то есть существует $D(\lambda)$ такое, что $m(\lambda)E = D(\lambda)(A - \lambda E) \Rightarrow f_n(\lambda)E = q(\lambda)m(\lambda)E = q(\lambda)D(\lambda)(A - \lambda E)$. Отсюда следует, что многочлен $q(\lambda)D(\lambda)$ является частным при делении $f_n(\lambda)$ справа на $(A - \lambda E)$, и все элементы частного делятся на $q(\lambda)$. Но из (12.7.2) следует, что частное при делении $f_n(\lambda)E$ справа на $(A - \lambda E)$ равно $(-1)^n C(\lambda)$. Тогда многочлен $q(\lambda)$ делит каждый элемент матрицы $C(\lambda)$. Но НОД элементов матрицы $C(\lambda)$ равен 1, поэтому $q(\lambda) = const$, а так как старший коэффициент $q(\lambda)$ равен 1, то и $q(\lambda) = 1$. Следовательно, $f_n(\lambda) = m(\lambda)$ и многочлен $f_n(\lambda)$ является минимальным.
⊠

Следствие (теорема Гамильтона-Кэли). *Характеристический многочлен матрицы является аннулирующим для нее.*

♦ Так как при элементарных преобразованиях определитель матрицы может измениться лишь на постоянный множитель, то $det(A-\lambda E) = det(\mathrm{diag}(f_1(\lambda), \dots, f_n(\lambda)))$ и $det(A-\lambda E)$ может отличаться от $f_1(\lambda), \dots, f_n(\lambda)$ лишь на постоянный множитель. Следовательно, все многочлены $f_i(\lambda)$, в том числе и $f_n(\lambda)$ делят многочлен $det(A-\lambda E)$. А так как $f_n(\lambda)$ — минимальный многочлен, то многочлен $det(A-\lambda E)$ является аннулирующим многочленом матрицы A. ⊠

Следствие. *Минимальные многочлены подобных матриц равны.*

♦ Матрицы A и B подобны $\iff (A-\lambda E) \sim (B-\lambda E) \iff$ системы инвариантным множителей матрицы A и B совпадают, а также совпадают и минимальные многочлены. ⊠

12.8 Жорданова нормальная форма матрицы.

• *Квадратная матрица над полем P порядка k вида*

$$J_k(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{pmatrix}$$

— называется ***жордановой клеткой****. Блочнодиагональная матрица называется* ***жордановой матрицей****, если ее диагональные блоки — жордановы клетки.*

Лемма. *Система элементарных делителей характеристической матрицы жордановой клетки $J_k(\lambda_0)$ состоит из одного многочлена $(\lambda-\lambda_0)^k$.*

♦Построим характеристическую матрицу жордановой клетки.

$$J_k(\lambda_0) - \lambda E = \begin{pmatrix} \lambda_0-\lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_0-\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0-\lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_0-\lambda \end{pmatrix}.$$

Система НОД миноров имеет вид: $d_1(\lambda) = 1, d_2(\lambda) = 1, \dots, d_{k-1}(\lambda) = 1, d_k(\lambda) = det(J_k(\lambda_0) - \lambda E) = (\lambda-\lambda_0)^k$ (или $(\lambda_0-\lambda)^k$ в зависимости от четности).

Система инвариантных множителей имеет вид: $f_1 = d_1 = 1, f_2 = \dfrac{d_2}{d_1} = 1, \dots, f_{k-1} = \dfrac{d_{k-1}}{d_{k-2}} = 1, f_k(\lambda) = \dfrac{d_k}{d_{k-1}} = (\lambda-\lambda_0)^k$ — система элементарных делителей . ⊠

Следствие. *Система элементарных делителей характеристической матрицы жордановой матрицы $J = \mathrm{diag}[J_{k_1}(\alpha_1), J_{k_2}(\alpha_2) \dots J_{k_s}(\alpha_s)]$ состоит из многочленов $(\lambda-\lambda_i)^{k_i}$ с учетом их повторения.*

♦ Система элементарных делителей блочнодиагональной матрицы есть объединение систем элементарных делителей ее диагональных блоков с учетом повторения. А характеристическая матрица жордановой матрицы является является блочнодиагональной матрицей, у которой по диагонали стоят характеристические матрицы жордановых клеток: $J_k(\lambda) - \lambda E = \mathrm{diag}\{J_{k_1}(\lambda_1) - \lambda E, \dots, J_{k_s}(\lambda_s) - \lambda E\}$. ⊠

Теорема. *Квадратная матрица $A \in P_{n,n}$ имеет подобную жорданову матрицу $\Longleftrightarrow$ ее характеристический многочлен разложим над полем P в произведение многочленов первой степени.*

⧫ $\Rightarrow$) Если матрица A имеет подобную матрицу J, то их характеристические многочлены равны, то есть $det(A - \lambda E) = det(J - \lambda E)$. Но так как матрица J треугольная, то матрица $(J - \lambda E)$ тоже треугольная. Следовательно, характеристический многочлен матрицы J равен произведению диагональных элементов матрицы $(J - \lambda E)$, которые имеют вид $(\lambda_i - \lambda)$. Тогжа и характеристический многочлен матрицы A есть произведение многочленов $(\lambda_i - \lambda)$.

$\Leftarrow$) Пусть характеристический многочлен матрицы A разложим в виде произведения многочленов первой степени. Так как при элементарных преобразованиях определитель может измениться лишь на постоянный множитель, то характеристический многочлен матрицы A равен с точностью до знака произведению инвариантных множителей матрицы $A - \lambda E$. Следовательно, каждый инвариантный множитель матрицы $A - \lambda E$ также разложим в произведение многочленов первой степени. Тогда система элементарных делителей матрицы $A - \lambda E$ состоит из многочленов $(\lambda - \lambda_i)^{k_i}$, причем $k_1 + \dots + k_s = n$. Заметим, что такую же систему элементарных делителей имеет характеристическая матрица жордановой матрицы J, и при этом порядок матрицы J равен порядку матрицы A. Исходя из этого ранги матриц $A - \lambda E$ и $J - \lambda E$ совпадают.
Таким образом, матрицы $A - \lambda E$ и $J - \lambda E$ эквивалентны, так как их ранги и системы элементарных делителей совпадают. А значит матрицы A и J подобны. ⊠

Следствие. *Любая квадратная матрица над полем $\mathbb{C}$ имеет подобную жорданову матрицу.*

• *Жорданова матрица, подобная матрице A, называется **жордановой нормальной формой** матрицы A.*

Теорема. *Две жордановы матрицы подобны $\Longleftrightarrow$ они состоят из одних и тех же жордановых клеток.*

⧫ J_1 подобна $J_2 \Longleftrightarrow J_1 - \lambda E \sim J_2 - \lambda E \Longleftrightarrow$ равны их ранги и системы элементарных делителей $\Longleftrightarrow$ они состоят из одних и тех же жордановых клеток. ⊠

Следствие. *Жорданова нормальная форма матрицы определена с точностью до порядка следования диагональных блоков.*

12.9 Теорема о количестве клеток жордановой нормальной формы матрицы.

Теорема. *Пусть A — квадратная матрица над полем P, J — ее жорданова нормальная форма. Тогда*

1. *диагональными элементами матрицы J являются собственные значения матрицы A, то есть все жордановы клетки матрицы J соответствуют собственным значениям матрицы A;*
2. *сумма порядков жордановых клеток матрицы J, соответсвующих собственному значению λ_0, равна алгебраической кратности собственного значения λ_0;*

3. *количество $l(\lambda_0)$ жордановых клеток матрицы J, соответствующих собственному значению λ_0, равна геометрической кратности собственного значения λ_0:*

$$l(\lambda_0) = n - \operatorname{rank}(A - \lambda_0 E);$$

4. *количество $l(\lambda_0)$ жордановых клеток матрицы J порядка n, соответствующих собственному значению λ_0, равно*

$$l_r(\lambda_0) = \operatorname{rank}(A - \lambda_0 E)^{r-1} - 2 \cdot \operatorname{rank}(A - \lambda_0 E)^r + \operatorname{rank}(A - \lambda_0 E)^{r+1}.$$

♦

1. Так как матрица A подобна матрице J, то их характеристические многочлены равны, то есть $det(A - \lambda E) = det(J - \lambda E)$. Следовательно, они имеют одни и те же корни такие, что корни характеристического многочлена матрицы A — собственные значения матрицы A, а корни характеристического многочлена J совпадают с диагональными элементами J, так как матрица J треугольная. А значит все жордановы клетки матрицы J соответствуют собственным значениям матрицы A.

2. Если λ_0 — корень характеристического уравнения $A - \lambda E$ кратности k, то λ_0 является корнем и характеристического уравнения матрицы J той же кратности. Следовательно, ровно k диагональных элементов равны λ_0. Но количество диагональных равных λ_0 элементов у J совпадает с суммой порядков жордановых клеток, соответствующих значению λ_0.

3. Рассмотрим жорданову клетку

$$J_k(\alpha) = \begin{pmatrix} \alpha & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \alpha & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \alpha & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \alpha & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \alpha \end{pmatrix}$$

и построим матрицу

$$J_k(\alpha) - \lambda_0 E = \begin{pmatrix} \alpha - \lambda_0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \alpha - \lambda_0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \alpha - \lambda_0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \alpha - \lambda_0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \alpha - \lambda_0 \end{pmatrix}.$$

Отсюда следует, что $\operatorname{rank}(J_k(\alpha) - \lambda_0 E) = \begin{cases} k, & \alpha \neq \lambda_0, \\ k - 1, & \alpha = \lambda_0. \end{cases}$

Так как матрица J блочнодиагональная, то ее ранг равен сумме рангов ее диагональных блоков. Значит ранг матрицы $J - \lambda_0 E$ отличается от порядка этой матрицы на столько единиц, сколько жордановых клеток соответствует значению λ_0. Тогда количество клеток l матрицы J, соответствуюзих значению λ_0, равно $l = n - \operatorname{rank}(J - \lambda_0 E)$. Так как матрица A подобна матрице J, то существует невырожденная матрица S такая, что $J = S^{-1}AS \Rightarrow J - \lambda_0 E = S^{-1}AS - \lambda_0 S^{-1}ES$, то есть матрицы $J - \lambda_0 E$ и $A - \lambda_0 E$ подобны и их ранги равны. Тогда количество клеток $l = n - \operatorname{rank}(A - \lambda_0 E)$ и равно геометрической кратности собственного значения λ_0.

4. Рассмотрим матрицу $(J_k(\alpha)-\lambda_0 E)^s$ такую, что если $\lambda_0 \neq \alpha$, то $det(J_k(\alpha)-\lambda_0 E) \neq 0 \Rightarrow det(J_k(\alpha)-\lambda_0 E)^s \Rightarrow$ это базисный минор и $\operatorname{rank}(J_k(\alpha)-\lambda_0 E) = k \quad \forall s \in \mathbb{N}$. А в случае, если $\alpha = \lambda_0$, получаем матрицу

$$J_k(\lambda_0)-\lambda_0 E = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

для которой $\operatorname{rank}(J_k(\lambda_0)-\lambda_0 E) = k-1$. Возведем эту матрицу в квадрат и получим матрицу

$$(J_k(\lambda_0)-\lambda_0 E)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

для которой $\operatorname{rank}(J_k(\lambda_0)-\lambda_0 E)^2 = k-2$. Проводя аналогичные рассуждения, придем к тому, что

$$\operatorname{rank}(J_k(\lambda_0)-\lambda_0 E)^s = \begin{cases} k-s, \ k > s, \\ 0, \ k \leqslant s. \end{cases}$$

Количество клеток J, соответствующих собственному значению λ_0 и имеющих порядок r и больше, равно $\operatorname{rank}(J-\lambda_0 E)^{r-1} - \operatorname{rank}(J-\lambda_0 E)^r$. Аналогично количество клеток J, соответствующих собственному значению λ_0 и имеющих размер $r+1$ и больше, равно $\operatorname{rank}(J-\lambda_0 E)^r - \operatorname{rank}(J-\lambda_0 E)^{r+1}$. Следовательно, количество клеток, имеющих порядок r, равно $l_r(\lambda_0) = \operatorname{rank}(J-\lambda_0 E)^{r-1} - 2\cdot\operatorname{rank}(J-\lambda_0 E)^r + \operatorname{rank}(J-\lambda_0 E)^{r+1}$. Так как матрицы A и J подобны, то по пункту 3 получаем, что матрицы $A-\lambda_0 E$ и $J-\lambda_0 E$ также подобны, то есть существует невырожденная матрица S такая, что $J-\lambda_0 E = S^{-1}(A-\lambda_0 E)S \Rightarrow (J-\lambda_0 E)^r = S^{-1}\cdot(A-\lambda_0 E)\cdot S\cdot S^{-1}\cdot(A-\lambda_0 E)\cdot S\cdot\ldots\cdot S^{-1}\cdot(A-\lambda_0 E)\cdot S = S^{-1}(A-\lambda_0 E)^r S$. Следовательно, так как матрицы $J-\lambda_0 E$ и $A-\lambda_0 E$ подобны и их ранги равны, то $l_r(\lambda_0) = \operatorname{rank}(A-\lambda_0 E)^{r-1} - 2\cdot\operatorname{rank}(A-\lambda_0 E)^r + \operatorname{rank}(A-\lambda_0 E)^{r+1}$.

$\boxtimes$

12.10 Жорданов базис.

Пусть V — векторное пространство над полем P, f — линейный оператор пространства V, M — матрица оператора f в некотором базисе E (не обязательно каноническом).

• ***Жордановым базисом*** *для линейного оператора f называется базис пространства V, в котором матрица M этого оператора f является жордановой.*

• *Система векторов $(a_1, \dots, a_k)$ называется* ***жордановой цепочкой*** *оператора f, соответствующей собственному значению λ_0, если вектор b_1 является собственным век-*

тором оператора f и

$$\begin{aligned}
&f(a_1) = \lambda_0 a_1,\\
&f(a_2) = \lambda_0 a_2 + a_1,\\
&f(a_3) = \lambda_0 a_3 + a_2,\\
&\dots\dots\dots\dots\dots\dots\\
&f(a_n) = \lambda_0 a_k + a_{k-1}.
\end{aligned}$$

*При этом вектор a_1 называется **началом цепочки**, а векторы $a_2, \dots, a_k$ называются соответственно первым, вторым и так далее векторами, **присоединенными** к вектору a_1. Если начало цепочки является собственным вектором, соответствующим собственному значению λ_0, то и цепочка называется соответствующей собственному значению λ_0.*

Свойства жордановых цепочек:

1. *Пусть $(a_1, \dots, a_k)$ — жорданова цепочка оператора f, соответствующая собственному значению $\lambda_0, X_1, \dots, X_k$ — координатные столбцы векторов a_i в базисе E Тогда*

$$(M - \lambda_0 E)X_i = \begin{cases} X_{i-1}, i > 1,\\ 0, i = 1.\end{cases}$$

   ⧫ Пусть $i = 1$. Тогда $f(a_1) = \lambda_0(a_1) \iff M_f X_1 = \lambda_0 X_1 \iff M_f X_1 - \lambda_0 X_1 = 0 \iff (M_f - \lambda_0)X_1 = 0$.

   Пусть $i > 1$. Тогда $f(a_i) = \lambda_0 a_i + a_{i-1} \iff MX_i = \lambda_0 X_i + X_{i-1} \iff MX_i - \lambda_0 X_i = X_{i-1} \iff (M - \lambda_0)X_i = X_{i-1}$. ⊠

Следствие. $(M - \lambda_0 E)^s X_i = \begin{cases} X_{i-s}, i > s,\\ 0, i \leqslant s.\end{cases}$

⧫ $(M - \lambda_0 E)^s X_i = (M - \lambda_0 E)^{s-1}(M - \lambda_0 E)X_i = (M - \lambda_0 E)^{s-1} X_{i-1}$. ⊠

2. *Жорданова цепочка является линейнонезависимой системой векторов.*

   ⧫ От противного:

   Пусть система $A(a_1, \dots, a_k)$ является жордановой цепочкой и предположим, что она линейно зависима. Тогда существует нетривиальная линейная комбинация равная нулевому вектору: $\alpha_1 a_1 + \ldots + \alpha_k a_k = 0_v$. Пусть X_i — координатные столбцы векторов a_i в базисе M и α_s — ненулевой коэффициент с наибольшим индексом, то есть отбросим все коэффициенты с индексом большим, чем s. Возьмем линейную комбинацию $\alpha_1 a_1 + \ldots + \alpha_s x_s = 0$ и домножим ее слева на $(M - \lambda_0 E)^{s-1} \cdot (a_1 + \ldots + \alpha_s x_s) = 0$, то есть $(M - \lambda_0 E)^{s-1}$. Тогда все векторы до последнего станут нулевыми по первому свойству, а последний вектор будет равен единице, то есть $\alpha_s x_1 = 0$, но так как вектор a_1 является началом жордановой цепочки, то он ненулевой, то есть $x_1 \neq 0$, значит $\alpha_s = 0$, что является противоречием. Значит жорданова цепочка линейно независимая. ⊠

3. *Система векторов, состоящая из жордановых цепочек линейного оператора, является линейно независимой $\iff$ линейно независимы входящие в нее собственные векторы этого оператора.*

$\blacklozenge$ $\Rightarrow$) Если система векторов, состоящая из жордановых цепочек, линейно независимая, то ее подсистема, состоящая из первых векторов (начал цепочек), также линейно независимая.

$\Leftarrow$) Для упрощения будем рассматривать систему (A, B), где $A(a_1, \ldots, a_{k_1})$, $B(b_1, \ldots, b_{k_2})$ — жордановы цепочки, и пусть система (a_1, b_1) линейно независима.
Предположим, что система (A, B) линейно зависима. Тогда существует нетривиальная линейная комбинация $\alpha_1 a_1 + \ldots + \alpha_{k_1} a_{k_1} + \beta_1 a_1 + \ldots + \beta_{k_2} a_{k_2} = 0_v$. Обозначим через X_i и Y_i координатные столбцы векторов a_i и b_i соответственно. Тогда

$$\underbrace{\alpha_1 x_1 + \ldots + \alpha_{k_1} x_{k_1}}_{X} + \underbrace{\beta_1 y_1 + \ldots + \beta_{k_2} y_{k_2}}_{Y} = 0_v. \qquad (12.10.1)$$

Рассмотрим два случая:

(a) Пусть системы A и B соответствуют одному и тому же собственному значению λ_0. Тогда обозначим $s_1 = max\{i \mid \alpha_i \neq 0\}, s_2 = max\{i \mid \beta_i \neq 0\}, s = max\{s_1, s_2\}$. Домножим уравнение (12.10.1) слева на $(M - \lambda_0 E)^{s-1}$ и получим систему $\begin{cases} \alpha_{s_1} x_1 = 0, \ s_1 > s_2, \\ \beta_{s_2} y_1 = 0, \ s_2 > s_1, \\ \alpha_s x_1 + \beta_s y_1 = 0, \ s_1 = s_2 = s; \end{cases}$, которая является нетривиальной линейной комбинацией координатных столбцов a_1 и b_1, что противоречит их линейной независимости.

(b) Пусть системы A и B соотвествуют различных собственным значениям λ_1 и λ_2. Так как система (12.10.1) нетривиальная, то среди собственных значений λ_i существуют ненулевые, так как в противном случае мы имеели бы нетривиальную линейную комбинацию системы B равную нулю, что противоречит второму свойству.
Домножим уравнение (12.10.1) слева на $(M - \lambda_2 E)^{k_2}$ и получим $(M - \lambda_2 E)^{k_2} X + \underbrace{(M - \lambda_2 E)^{k_2} Y}_{=0} = 0 \Rightarrow$

$$(M - \lambda_2 E)^{k_2} X = 0. \qquad (12.10.2)$$

Так как $\lambda_1 \neq \lambda_2$, то $(\lambda - \lambda_1)^{k_1}$ и $(\lambda - \lambda_2)^{k_2}$ — взаимно простые многочлены. Тогда по критерию взаимно простых многочленов

$$\exists \varphi(\lambda), \psi(\lambda) : \varphi(\lambda)(\lambda - \lambda_1)^{k_1} + \psi(\lambda)(\lambda - \lambda_2)^{k_2} = 1.$$

Вычислим значения многочленов в обеих частях равенства при $\lambda = M$:

$$\varphi(M)(M - \lambda_1 E)^{k_1} + \psi(M)(M - \lambda_2 E)^{k_2} = E.$$

Домножим это уравнение справа на X:

$$\underbrace{\varphi(M)(M - \lambda_1 E)^{k_1} X}_{=0} + \underbrace{\psi(M)(M - \lambda_2 E)^{k_2} X}_{=0} = X.$$

Следовательно, $0 = X$, X является нетривиальной линейной комбинацией координатных столбцов векторов жордановой цепочки равной 0_v, и система A является линейно зависимой, что противоречит с условием.

$\boxtimes$

Теорема. *Базис пространства V является жордановым для оператора $f \iff$ он состоит из жордановых цепочек.*

$\blacklozenge \Leftarrow$) Пусть базис $B(B_1, \ldots, B_s)$ пространства V состоит из жордановых цепочек $B_i(b_{i1}, \ldots, b_{ik_i})$, соответствующих собственному значению λ_i. Тогда

$$\begin{gathered} f(b_{i1}) = \lambda_i b_{i1}, \\ f(b_{i2}) = \lambda_i b_{i2} + b_{i1}, \\ f(b_{i3}) = \lambda_i b_{i3} + b_{i2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{gathered}$$

Пометим данную цепочку уравнений через (12.10.3). Построим матрицу оператора f в базисе B:

$$M_f^B = \left(\begin{array}{ccccc|c|c} \lambda_1 & 1 & 0 & \ldots & 0 & & \\ 0 & \lambda_1 & 1 & \ldots & 0 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 & \ldots \\ 0 & 0 & 0 & \ldots & 1 & & \\ 0 & 0 & 0 & \ldots & \lambda_1 & & \\ \hline & & 0 & & & J_{k_2} & \ldots \\ \hline & & \ldots & & & \ldots & J_{k_s} \end{array}\right) =$$

$$= \operatorname{diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \ldots, J_{k_s}(\lambda_s)). \tag{12.10.4}$$

Следовательно, получили жорданову матрицу (12.10.4). $\Rightarrow$) Если матрица оператора f в базисе B является жордановой вида (12.10.4), то для векторов базиса B справедливо равенство (12.10.3). Следовательно, базис B состоит из жордановых цепочек. ⊠

Следствие. *Количество жордановых цепочек в жордановом базисе оператора f, соответствующих собственному значению λ_0, совпадает с количеством жордановых клеток, соответствующих λ_0, а длины цепочек — с порядками этих клеток.*

Следствие. *Все жордановы базисы оператора состоят из одного и того же числа цепочек одной и той же длины.*

Следствие. *Матрица S, трансформирующая матрицу M в ее жорданову матрицу J, является матрицей перехода от базиса E к жорданову базису B.*

12.11 Нормальные формы Фробениуса матрицы.

Пусть $f(\lambda)$ — многочлен положительной степени над полем P вида

$$f(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \alpha_{n-2}\lambda^{n-2} + \ldots + \alpha_0.$$

Матрица

$$F = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -\alpha_0 \\ 1 & 0 & 0 & \ldots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \ldots & 0 & -\alpha_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & -\alpha_{n-2} \\ 0 & 0 & 0 & \ldots & 1 & -\alpha_{n-1} \end{pmatrix}$$

*называется **сопровождающей** матрицей для многочлена $f(\lambda)$.*

• *Блочнодиагональная матрица $F = \operatorname{diag}[F_1, \ldots, F_k]$ называется **матрицей Фробениуса**, если диагональные блоки F_i являются сопровождающими матрицами для многочленов $f_i(\lambda)$, каждый из которых делит последующий.*

Лемма. *Если F — сопровождающая матрица для многочлена $f(\lambda)$, то характеристическая матрица $(F - \lambda E)$ иммет систему инвариантных множителей $1, \ldots, 1, f(\lambda)$.*

♦ Построим матрицу $F - \lambda E$:

$$F - \lambda E = \begin{pmatrix} -\lambda & 0 & 0 & \ldots & 0 & -\alpha_0 \\ 1 & -\lambda & 0 & \ldots & 0 & -\alpha_1 \\ 0 & 1 & -\lambda & \ldots & 0 & -\alpha_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -\lambda & -\alpha_{n-2} \\ 0 & 0 & 0 & \ldots & 1 & -\alpha_{n-1} - \lambda \end{pmatrix}.$$

Система НОД миноров этой матрицы следующая: $d_1(\lambda) = 1, d_2(\lambda) = 1, \ldots, d_{k-1}(\lambda) = 1$.

Тогда $det(F - \lambda E) = (-\alpha_{k-1} - \lambda)(-1)^{2k}(-\lambda)^{k-1} + (-\alpha_{k-2})(-1)^{2k-1}(-\lambda)^k + \ldots + (-\alpha_1)(-1)^{k+2}(-\lambda) + (-\alpha_0)(-1)^{k+1} = (-1)^k(\lambda^k + \alpha_{n-1}\lambda^{n-1} + \alpha_{n-2}\lambda^{n-2} + \ldots + \alpha_1\lambda + \alpha_0) = (-1)^k f(\lambda)$.

Система инвариатных множителей имеет следующий вид: $f_1(\lambda) = 1, f_2(\lambda) = 1, \ldots, f_{k-1}(\lambda) = 1, f_k(\lambda) = f(\lambda)$. ⊠

Следствие. *Если $F = \operatorname{diag}[F_1, \ldots, F_s]$ — матрица Фробениуса, сопровождающая систему многочленов $f_1(\lambda), \ldots, f_k(\lambda)$, то матрица $F - \lambda E$ имеет систему инвариантных множителей $1, \ldots, 1, f_1(\lambda), \ldots, f_k(\lambda)$*

♦ Матрица $(F - \lambda E)$ является диагональной и имеет вид $\operatorname{diag}[F_1 - \lambda E, \ldots, F_s - \lambda E]$. С помощью элементарных преобразований приведем ее к виду $\operatorname{diag}(1, \ldots, 1, f_1(\lambda), 1, \ldots, 1, f_2(\lambda), \ldots, 1, \ldots, 1, f_s(\lambda))$, которую, в свою очередь, можно привести к виду $\operatorname{diag}(1, \ldots, 1, f_1(\lambda), \ldots, f_s(\lambda))$. Так как матрица F — матрица Фробениуса, то любой элемент $f_i(\lambda)$ делит элемент $f_{i+1}(\lambda)$. Следовательно, полученная матрица является канонической и ее диагональ состоит из инвариантных множителей. ⊠

Теорема. *Для любой квадратной матрицы над произвольным полем существует и только одна подобная матрица Фробениуса.*

♦

1. Существование.
   Пусть A — квадратная матрица поряка n, и пусть матрица $(A - \lambda E)$ имеет систему инвариантных множителей $1, \ldots, 1, f_1(\lambda), \ldots, f_s(\lambda)$. Построим блочнодиагональную матрицу $F = \operatorname{diag}[F_1, \ldots, F_s]$, где F_i — сопровождающие матрицы для многочленов $f_i(\lambda)$. Так как многочлены $f_i(\lambda)$ являются инвариантными множителями матрицы $A - \lambda E$, то многочлены $f_i(\lambda)$ делят многочлены $f_{i+1}(\lambda)$. Следовательно, F — матрица Фробениуса.
   Так как произведение инвариантных множителей с точностью до знака совпадает с характеристическим многочленом, то сумма степеней всех $f_i(\lambda)$ равна степени характеристического многочлена матрицы A. Тогда матрица F имеет тот же порядок, что и матрица A. И по следиствию из леммы матрица $(F - \lambda E)$ имеет ту же систему

инвариантных множителей, что и матрица $(A-\lambda E) \Rightarrow (A-\lambda E) \sim (F-\lambda E) \Rightarrow$ матрицы F и A подобны.

2. Единственность.
   Пусть матрица A имеет две подобные матрицы Фробениуса F_1 и F_2. Если матрицы F_1 и F_2 подобны, то $(F_1-\lambda) \sim (F_2-\lambda E)$. Тогда они имеют одну и ту же систему инвариантных множителей и F_1 и F_2 сопровождают одну и ту же систему многочленов. Следовательно, они состоят из одних и тех же блоков, а значит матрицы F_1 и F_2 равны.

⊠

• *Матрица Фробениуса, подобная A, называется* ***нормальной формой Фробениуса*** *(естественной нормальной формой) матрицы A.*

Лемма. *Пусть $p(\lambda)$ — неприводимый многочлен, H — сопровождающая матрица для многочлена $(p(\lambda))^k$. Тогда система элементарных делителей матрицы $H-\lambda E$ состоит из одного многочлена $(p(\lambda))^k$.*

⧫ Так как матрица H является сопровождающей для многочлена $(p(\lambda))^k$, то система инвариантных множителей имеет вид $1,\dots,1,(p(\lambda))^k$. Следовательно, система элементарных делителей матрицы $H-\lambda E$ совпадает с системой элементарных делителей многочлена $(p(\lambda))^k$. А так как многочлен $p(\lambda)$ неприводимый, то он и образует систему элементарных делителей самого себя. ⊠

Теорема. *Пусть A — некоторая матрица над полем P, $h_1(\lambda),\dots,h_k(\lambda)$ — система элементарных делителей матрицы $A-\lambda E$. Тогда матрица A подобная блочнодиагональной матрице $H=\mathrm{diag}[H_1,\dots,H_k]$, где H_i — сопровождающая матрица для многочлена $h_i(\lambda)$.*

⧫ Матрица $H-\lambda E$ является блочнодиагональной. Следовательно, ее система элементарных делителей является объединением систем элементарных делителей ее диагональных блоков. Но каждый блок H_i — сопряженная матрица элементарного делителя $h_i(\lambda)$, который является степенью неприводимого многочлена, а значит блоки $H_i-\lambda E$ по лемме имеют систему элементарных делителей $h_i(\lambda)$. Значит, матрицы $H-\lambda E$ и $A-\lambda E$ имеют одинаковую систему элементарных делителей, и ранги этих матриц равны их порядку. Тогда по критерию эквивалентности матриц $(H-\lambda E) \sim (A-\lambda E)$, и матрицы H и A подобны. ⊠

• *Блочногодиагональная матрица H, подобная матрице A, называется* ***второй нормальной формой Фробениуса*** *(квазиестественной нормальной формой) матрицы A, если ее диагональные клетки являются сопровождающими матрицами для элементарных делителей матрицы $A-\lambda E$.*

Вторая нормальная форма Фробениуса определяется с точностью до порядка следования диагональных клеток.

Глава 13

Квадратичные формы

13.1 Эквивалентность квадратичных форм.

Пусть P — или поле **действительных** чисел, или поле **комлексных** чисел.

• ***Квадратичной формой*** *от n переменных $x_1 \dots x_n$ над полем P называется многочлен от этих переменных с коэффициентами из поля P, каждый член которого относительно переменных имеет вторую степень, то есть*

$$f(x_1, \dots, x_n) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j, \quad \alpha_{ij} \in P$$

Любую действительную, комплексную квадратичную форму можно представить в симметрическом виде, то есть в виде квадратичной формы, коэффициенты которой удовлетворяют условию $\alpha_{ij} = \alpha_{ji}$.

• *Матрицу коэффициентов $A(\alpha_{ij})$ симметрического вида квадратичной формы называют* ***матрицей квадратичной формы****.* ***Рангом квадратичной формы*** *называется ранг матрицы квадратичной формы.*

Матрица квадратичной формы всегда симметрическая.

Если переменные x_i записать в столбец $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ то квадратичная форма в матричном виде может быть записана следующим образом: $f(X) = X^T AX$.

• *Пусть $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ и $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ — две системы переменных. Замена переменных x_i на y_i по формулам вида:*

$$\begin{cases} x_1 = s_{11}y_1 + \ldots + s_{1n}y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = s_{n1}y_1 + \ldots + s_{nn}y_n; \end{cases} \qquad s_{ij} \in P,$$

называется ***линейным преобразованием переменных*** *над полем P. Матрица $S(s_{ij})$ называется* ***матрицей преобразования****. Преобразование называется* ***невырожденным****, если $det S \neq 0$.*

Линейное преобразование переменных в матричном виде можно записать следующим образом: $X = SY$.

• *Квадратичная форма $f(X) = f(x_1, \ldots, x_n)$ называется* ***эквивалентной*** *квадратичной форме $g(Y) = g(y_1, \ldots, y_n)$, если существует линейное невырожденное преобразование переменных $X = SY$ переводящее $f(X)$ в $g(Y)$, то есть такое, что $f(SY) = g(Y)$.*

Лемма. *Эквивалентность квадратичных форм рефлексивна, симметрична, транзитивна.*

♦

1. Рефлексивность: $f(X) \sim f(X)$, так как линейное невырожденное преобразование $X = EX$ переводит квадратичную форму в эту же квадратичную форму.

2. Симметричность: Пусть $f(X) \sim g(X)$, то есть существует невырожденное линейное преобразование $X = SY : f(SY) = g(Y)$. Так как матрица S — невырожденная, то существует $S^{-1} : Y = S^{-1}X$ — линейное преобразование и при этом $g(S^{-1}X) = f(S(S^{-1}X)) = f(X)$, то есть $g(X) \sim f(X)$.

3. Транзитивность: Пусть $f(X) \sim g(Y), \quad g(Y) \sim h(Z)$. Тогда существуют линейные преобразования $X = S_1Y,\ Y = S_2Z : f(S_1Y) = g(Y),\ g(S_1Z) = h(Z) \Rightarrow h(Z) = g(S_2Z) = f(S_1(S_2Z)) = f((S_1S_2)Z)$, то есть $X = (S_1S_2)Z$ переводит квадратичную форму f в h и так как матрицы S_1 и S_2 невырожденные, то матрица S_1S_2 также невырожденная, следовательно, $f(X) \sim h(Z)$.

⊠

• *Квадратные матрицы A и B называются* ***конгруэнтными***, *если существует невырожденная матрица S такая, что $B = S^TAS$.*

Теорема. *Квадратичные формы $f(X)$ и $g(Y)$ эквивалентны $\Longleftrightarrow$ их матрицы конгруэнтны.*

♦ Пусть A_f и A_g — матрицы квадратичных форм $f(X)$ и $g(Y)$ соответственно. Тогда $f(X) = X^TA_fX,\ g(Y) = Y^TA_gY$.

Квадратичные формы $f(X)$ и $g(Y)$ эквивалентны $\Longleftrightarrow \exists S, detS \neq 0 : X = SY,\ f(SY) = g(Y)$, то есть $(SY)^TA_f(SY) = Y^TA_gY \Rightarrow Y^T(S^TA_fS)Y = Y^TA_gY$. (На Y^T и Y сокращать нельзя)

Покажем, что $S^TA_fS = A_g$. Для этого покажем, что они симметрические. Матрица A_g симметрическая по определению. $(S^TA_fS)^T = [A_f$ — симметрическая$] = S^TA_f^T(S^T)^T = S^TA_fS$ — также симметрическая $\Rightarrow A_g$ и S^TAS — матрицы одной и той же квадратичной формы $\Rightarrow A_g = S^TAS$. ⊠

Следствие. *Ранги эквивалентных квадратичных форм равны.*

13.2 Канонический вид квадратичной формы.

• *Квадратичная форма называется* ***канонической***, *если ее матрица диагональная, то есть каноническая квадратичная форма имеет вид $\alpha_{11}x_1^2 + \alpha_{22}x_2^2 + \ldots + \alpha_{nn}x_n^2$.*

Теорема. *Для любой квадратичной формы существует эквивалентная каноническая квадратичная форма.*

⧫ Докажем по индукции по количеству переменных в квадратичной форме.

1. Пусть $n = 1$. Тогда квадратичная форма каноническая.

2. Пусть теорема верна для меньше n переменных. Покажем для квадратичной формы $f(x_1, \ldots, x_n) = \sum\limits_{i,j=1}^{n} \alpha_{ij} x_i x_j$, где $\alpha_{ij} = \alpha_{ji}$. Для этого рассмотрим два случая:

   (a) среди коэффициентов α_{ij} есть ненулевые. Тогда с помощью перенумерации переменных получим квадратичную форму, у которой $\alpha_{11} \neq 0$. Учитываем, что перенумерация — невырожденное линейное преобразование, матрица которого мономиальная, значит, невырожденная. Следовательно, получаем эквивалентную квадратичную форму $f(x_1, \ldots, x_n) = \alpha_{11} x_1^2 + \alpha_{12} x_1 x_2 + \ldots + \alpha_{1n} x_1 x_n + \alpha_{21} x_2 x_1 + \ldots + \alpha_{n1} x_n x_1 + \sum\limits_{i,j=2}^{n} \alpha_{ij} x_i x_j = \alpha_{11}(x_1^2 + 2\dfrac{\alpha_{12}}{\alpha_{11}} x_1 x_2 + \ldots + 2\dfrac{\alpha_{1n}}{\alpha_{11}} x_1 x_n + (\dfrac{\alpha_{12}}{\alpha_{11}} x_2 + \ldots + \dfrac{\alpha_{1n}}{\alpha_{11}} x_n)^2) - \alpha_{11}(\dfrac{\alpha_{12}}{\alpha_{11}} x_2 + \ldots + \dfrac{\alpha_{1n}}{\alpha_{11}} x_n)^2 + \sum\limits_{i,j=2}^{n} \alpha_{ij} x_i x_j =$
   $\alpha_{11}(\dfrac{\alpha_{12}}{\alpha_{11}} x_2 + \ldots + \dfrac{\alpha_{1n}}{\alpha_{11}} x_n)^2 + g(x_2, \ldots, x_n)$, где g — квадратичная форма, зависящая от $n-1$ переменных.
   По индуктивному предположению для квадратичной формы g существует эквивалентная квадратичная форма $\widetilde{g}(y_1, \ldots, y_n) = \beta_{22} y_2^2 + \ldots + \beta_{nn} y_n^2$. А значит существует невырожденное преобразование

$$\begin{cases} y_2 = S_{22} x_2 + \ldots + S_{2n} x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = S_{n2} x_2 + \ldots + S_{nn} x_n; \end{cases},$$

   которое переводит $\widetilde{g}$ в g. Тогда преобразование

$$\begin{cases} y_1 = x_1 + \frac{\alpha_{12}}{\alpha_{11}} x_2 + \ldots + \frac{\alpha_{1n}}{\alpha_{11}} x_n, \\ y_2 = S_{22} x_2 + \ldots + S_{2n} x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = S_{n2} x_2 + \ldots + S_{nn} x_n; \end{cases}$$

   линейное и невырожденное. Построим матрицу преобразования:

$$det \left(\begin{array}{c|ccc} 1 & \dfrac{\alpha_{12}}{\alpha_{11}} & \ldots & \dfrac{\alpha_{1n}}{\alpha_{11}} \\ \hline & S_{22} & \ldots & S_{2n} \\ 0 & \vdots & \ddots & \vdots \\ & S_{n2} & \ldots & S_{nn} \end{array} \right) = 1 \cdot det S_{n-1} \neq 0.$$

   Причем, применив это преобразование к квадратичной форме, получим
   $f(y_1, \ldots, y_n) = \alpha_{11} y_1^2 + \beta_{22} y_2^2 + \ldots + \beta_{nn} y_n^2 \Rightarrow f \sim g$.

   (b) Пусть среди диагональных элементов a_{ii} нет ненулевых. И пусть все остальные элементы также нулевые. Тогда квадратичная форма каноническая.
   Пусть среди всех коэффициентов $\exists \alpha_{ij} \neq 0$, $i \neq j$. Перенумеруем так, чтобы этот

коэффициент был равен α_{12}. Применим следующее невырожденное преобразование

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_i = y_i, \quad i = \overline{3,n}; \end{cases},$$

при этом $2\alpha_{12}x_1x_2 = 2\alpha_{12}(y_1-y_2)(y_1+y_2) = 2\alpha_{12}y_1^2 - 2\alpha_{12}y_2^2$, и для него существует эквивалентная квадратичная форма. Так как мы получили ненулевой угловой коэффициент, то можем перейти к случаю (а).

⊠

• ***Каноническим видом*** *квадратичной формы называется эквивалентная ей каноническая форма.*

13.3 Формула Якоби канонического вида квадратичной формы.

Рассмотрим квадратичную форму $f(x_1, \dots, x_n) = \sum\limits_{i,j=1}^{n} \alpha_{ij}x_ix_j$, где $\alpha_{ij} = \alpha_{ji}$. Матрица $A = (\alpha_{ij})$ является матрицей квадратичной.

• ***Угловыми минорами*** *матрицы* $A(\alpha_{ij})$ *называются миноры*

$$\Delta_1 = \alpha_{11},\ \Delta_2 = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}, \dots, \Delta_k = \begin{vmatrix} \alpha_{11} & \dots & \alpha_{1k} \\ \vdots & \ddots & \vdots \\ \alpha_{k1} & \dots & \alpha_{kk} \end{vmatrix}, \dots, \Delta_n = detA.$$

Теорема (Формула Якоби). *Если матрица квадратичной формы* $f(x_1, \dots, x_n)$ *имеет ненулевые угловые миноры, то*

$$f \sim \frac{1}{\Delta_1}y_1^2 + \frac{\Delta_1}{\Delta_2}y_2^2 + \dots + \frac{\Delta_{n-1}}{\Delta_n}y_n^2$$

⧫ Так как $\Delta_1 \neq 0$, то $\alpha_{11} \neq 0$, то квадратичная форма f представим в виде $f(x_1, \dots, x_n) = \alpha_{11}x_1^2 + 2\alpha_{12}x_1x_2 + \dots + 2\alpha_{1n}x_1x_n + \sum\limits_{i,j=2}^{n} \alpha_i x_i x_j = \dfrac{1}{\alpha_{11}}(\alpha_{11}^2x_1^2 + 2x_1\alpha_{11} \cdot$
$(\alpha_{12}x_2 + \dots + \alpha_{1n}x_n) + (\alpha_{12}x_2 + \dots + \alpha_{1n}x_n)^2) \underbrace{- \dfrac{1}{\alpha_{11}}(\alpha_{12}x_2 + \dots + \alpha_{1n}x_n)^2 + \sum\limits_{i,j=2}^{n} \alpha_i x_i x_j}_{g(x_2,\dots,x_n)} =$

$\dfrac{1}{\alpha_{11}}(\alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n)^2 + g(x_2, \dots, x_n)$, где $g(x_2, \dots, x_n)$ — квадратичная форма.

Обозначим матрицу квадратичной формы g через $B(\beta_{ij}), \forall i,j = \overline{2,n}$. Вычислим элементы $\beta_{ij} : g(x_2, \dots, x_n) = f - \dfrac{1}{\alpha_{11}}(\alpha_{11}x_1 + \dots + \alpha_{1n}x_n)^2 \Rightarrow \beta_{ij} = \alpha_{ij} - \dfrac{\alpha_{1i}\alpha_{1j}}{\alpha_{11}}$.

С другой стороны, если к первому столбцу матрицы A применить алгоритм Гаусса, то есть получить из матрицы A матрицу $A' = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ 0 & \alpha'_{22} & \dots & \alpha'_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & \alpha'_{n2} & \dots & \alpha'_{nn} \end{pmatrix}$, где элементы матрицы

A' следующие: $\alpha'_{ij} = \alpha_{ij} + \alpha_{1j}(-\frac{\alpha_{i1}}{\alpha_{11}}) = \alpha_{ij} - \frac{\alpha_{1j}\alpha_{i1}}{\alpha_{11}} = \alpha_{ij} - \frac{\alpha_{1j}\alpha_{1i}}{\alpha_{11}} = \beta_{ij}$, то матрица A' имеет вид $A' = \left(\begin{array}{c|c} \alpha_{11} & \alpha_{12}\dots\alpha_{1n} \\ \hline 0 & B \end{array}\right)$

Таким образом, процесс выделения полных квадратов в квадратичной форме f совпадает с первым шагом метода Гаусса. При этом элементы первой строки матрицы A' совпадают с коэффициентами при переменных x_i в выделенном квадрате. Элемент, обратный первому элементу A', совпадает с коэффициентом при выделенном квадрате, а остальные — совпадают с элементами квадратичной формы g. Заметим, что преобразования не меняют угловых миноров: $\Delta_2 = \alpha_{11}\alpha'_{22} \neq 0 \Rightarrow \alpha'_{22} \neq 0 \Rightarrow$ мы можем применять следующие шаги метода Гаусса. Спустя $n-1$ шаг мы получим матрицу $C(\gamma_{ij})$, которая является верхнетреугольной $\Rightarrow$ она невырожденная $\Rightarrow f(x_1,\dots,x_n) = \frac{1}{\gamma_{11}}\underbrace{(\gamma_{11}x_1 + \ldots + \gamma_{1n}x_n)}_{y_1}{}^2 + \ldots + \frac{1}{\gamma_{nn}}(\gamma_{nn}x_n)^2 \Rightarrow$ квадратичная форма f может быть получена из $\widetilde{f}(y_1,\dots,y_n) = \frac{1}{\gamma_{11}}y_1^2 + ... + \frac{1}{\gamma_{nn}}y_n^2$ при помощи преобразований: $\begin{cases} y_1 = \gamma_{11}x_1 + \ldots + \gamma_{1n}x_n, \\ y_2 = \gamma_{22}x_2 + \ldots + \gamma_{2n}x_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ y_n = \qquad\qquad \gamma_{nn}x_n; \end{cases}$

При этом матрица этого преобразования совпадает с матрицей C, и, следовательно, невырожденная $\Rightarrow$ преобразование также невырожденное $\Rightarrow f \sim \widetilde{f}$. Выражаем коээфициенты $\widetilde{f}$ через элементы матрицы A. Матрица C получена из матрицы A путем преобразований, не изменяя угловые миноры $\Rightarrow$

$$\Delta_1 = \gamma_{11}, \quad \Delta_2 = \gamma_{11}\gamma_{22}, \quad \dots, \quad \Delta_n = \gamma_{11}\gamma_{22}\dots\gamma_{nn};$$
$$\frac{1}{\gamma_{11}} = \frac{1}{\Delta_1}, \quad \frac{1}{\gamma_{22}} = \frac{\gamma_{11}}{\Delta_2} = \frac{\Delta_1}{\Delta_2}, \quad \dots, \quad \frac{1}{\gamma_{nn}} = \frac{\Delta_{n-1}}{\Delta_n}.$$

⊠

13.4 Нормальный вид квадратичной формы.

• *Комплексная каноническая квадратичная форма $f(x_1,\dots,x_n)$ называется* ***нормальной****, если она имеет вид*

$$x_1^2 + \ldots + x_r^2 + 0\cdot x_{r+1}^2 + \ldots + 0\cdot x_n^2, \quad 0 \leqslant r \leqslant n.$$

Теорема. *Для любой комплексной квадратичной формы существует единственная эквивалентная нормальная квадратичная форма.*

♦

1. Существование. Для любой квадратичной формы существует эквивалентная ей каноническая форма $\alpha_1x_1^2 + \ldots + \alpha_nx_n^2$. Перенумеруем переменные так, чтобы α_i было ненулевым для $i = \overline{1,r}$ и $\alpha_i = 0$ для $i = \overline{r+1,n}$. А затем применим к квадратичной форме линейное преобразование: $\begin{cases} y_i = \sqrt{\alpha_i}x_i, \ i = \overline{1,r}, \\ y_i = x_i, \ i = \overline{r+1,n}; \end{cases}$, причем матрица этого преобразования невырожденная. В результате получили эквивалетную нормальную квадратичную форму.

2. Единственность. Пусть каноническая квадратичная форма имеет две нормальные формы:$\begin{cases} f(x_1,\ldots,x_n) = x_1^2 + \ldots + x_r^2, \\ g(x_1,\ldots,x_n) = x_1^2 + \ldots + x_q^2; \end{cases}$. Тогда $f \sim g \Rightarrow \operatorname{rank} f = \operatorname{rank} g \Rightarrow r = q \Rightarrow$ квадратичные формы $f(x_1,\ldots,x_n)$ и $g(x_1,\ldots,x_n)$ совпадают.

⊠

Следствие. *Комплексные квадратичные формы эквивалентны $\iff$ их ранги равны.*

• *Действительная каноническая квадратичная форма $f(x_1,\ldots,x_n)$ называется **нормальной**, если она имеет вид*

$$x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2 + 0 \cdot x_{r+1}^2 + \ldots + 0 \cdot x_n^2, \quad 0 \leqslant p \leqslant r \leqslant n.$$

Теорема (закон инерции действительных квадратичных форм). *Для любой действительной квадратичной формы существует единственная эквивалентная нормальная квадратичная форма.*

⧫

1. Существование. Любую дествительную квадратичную форму невырожденным линейным преобразованием переменных можно привести к эквивалентной канонической квадратичной форме вида $\alpha_1 x_1^2 + \ldots + \alpha_n x_n^2$. Перенумеруем в канонической квадратичной форме переменные следующим образзом: $\begin{cases} \alpha_i > 0 & i = \overline{1,p}, \\ \alpha_i < 0 & i = \overline{p+1,r}, \\ \alpha_i = 0 & i = \overline{r+1,n}; \end{cases}$. Выполнив еще одно невырожденное линейное преобразование, получим $\begin{cases} y_i = \sqrt{|\alpha_i|}x_i & i = \overline{1,r}, \\ y_i = x_i & i = \overline{r+1,n}; \end{cases}$, что и является нормальной квадратичной формой.

2. Единственность. Пусть квадратичная форма имеет две различные эквивалентные нормальные формы $f(x)$ и $g(y)$. Тогда $f(x) \sim g(y) \Rightarrow \operatorname{rank} f = \operatorname{rank} g \Rightarrow$ количество ненулевых коэффициентов одинаково, обозначим его с помощью r.

$$\begin{cases} f(x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2, \\ g(y) = y_1^2 + \ldots + y_q^2 - y_{q+1}^2 - \ldots - y_r^2. \end{cases}$$

Предположим, что $p \neq q$, и пусть $p < q$. Так как $f \sim g$, то существует линейное преобразование $X = SY$, где матрица $S = (S_{ij}) \in \mathbb{R}_{n,n}$ и является невырожденной. Тогда $f(SY) = g(Y)$. Построим линейную однородную систему уравнений

$$\begin{cases} S_{11}y_1 + \ldots + S_{1q}y_q = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ S_{p1}y_1 + \ldots + S_{pq}y_q = 0; \end{cases}$$

Так как $p < q$ то ранг матрицы меньше числа неизвестных. Следовательно, существуют ненулевые решения $(\tilde{y}_1,\ldots,\tilde{y}_q)$. Построим столбец этих решений $\tilde{Y} = \begin{pmatrix} \tilde{y}_1 \\ \ldots \\ \tilde{y}_q \\ \ldots \\ 0 \end{pmatrix}$

и вычислим $\tilde{X} = S\tilde{Y} = \begin{pmatrix} S_{11}\tilde{y}_1 + \ldots + S_{1q}\tilde{y}_q \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ S_{n1}\tilde{y}_1 + \ldots + S_{nq}\tilde{y}_q \end{pmatrix}$. Тогда $\tilde{x}_1 = \ldots = \tilde{x}_p = 0$;

Так как $f(SY) = g(Y)$, то $f(S\tilde{Y}) = g(\tilde{Y}) \Rightarrow f(\tilde{X}) = g(\tilde{Y}) \Rightarrow f(\tilde{X}) = \underbrace{-x_{p+1}^2 - \ldots - x_r^2}_{\leqslant 0} = g(\tilde{Y}) = \underbrace{y_1^2 + \ldots + y_q^2}_{>0}$, что является противоречием, так как y_i — ненулевое решение системы уравнений, следовательно, $p = q$.

⊠

• ***Нормальным видом квадратичной формы*** *называется эквивалентная ей нормальная квадратичная форма.*

• *Число положительных и отрицательных коэффициентов нормального вида действительной квадратичной формы называются соответственно* ***положительным*** *и* ***отрицательным индексом инерции квадратичной формы****. Разность положительного и отрицательного индексов инерции называется* ***сигнатурой*** *квадратичной формы.*

13.5 Знакоопределенные действительные квадратичные формы.

• *Действительная квадратичная форма $f(x_1, \ldots, x_n)$ называется* ***положительноопределенной (отрицательноопределнной)****, если для любых чисел $\gamma_1, \ldots, \gamma_n, \gamma_i \in \mathbb{R}$, не обращающихся одновременно в 0, $f(\gamma_1, \ldots, \gamma_n) > 0$ ($f(\gamma_1, \ldots, \gamma_n) < 0$).*

Теорема. *Квадратичная форма $f(x_1, \ldots, x_n)$ является положительноопределенной $\Longleftrightarrow$ она эквивалентна нормальной квадратичной форме*

$$g(y_1, \ldots, y_n) = y_1^2 + \ldots + y_n^2.$$

⧫ $\Rightarrow$) Пусть $f(x)$ — положительноопределенная квадратичная форма, и она эквивалентна нормальной квадратичной форме $g(y_1, \ldots, y_n) = \alpha_1 y_1^2 + \ldots + \alpha_n y_n^2$. Предположим, что $\alpha_n \neq 1$, то есть $\alpha_n = -1$ или $\alpha_n = 0$. Тогда для $e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ выполняется $g(e_n) = \alpha_n \leqslant 0$.

Так как $f \sim g$, то существует линейное невырожденное преобразование $X = SY$ такое, что $f(SY) = g(Y) \Rightarrow f(Se_n) = g(e_n) \leqslant 0$, но так как $Se_n = \begin{pmatrix} S_{1n} \\ \vdots \\ S_{nn} \end{pmatrix}$ (последний столбец матрицы S) и матрица S невырожденная, то столбец Se_n ненулевой, что является противоречием с тем, что квадратичная форма f положительноопределенная.

$\Leftarrow$) Пусть $f \sim g(y_1, \ldots, y_n) = y_1^2 + \ldots + y_n^2$. Тогда существует линейное невырожденное преобразование $Y = SX$ такое, что $g(SX) = f(X)$. Квадратичная форма $g(y)$ положительноопределенная, следовательно, для любого ненулевого столбца $\Gamma \in \mathbb{R}$ $g(S\Gamma) \geqslant 0$, причем $g(S\Gamma) = 0 \Leftrightarrow S\Gamma = 0 \Rightarrow \Gamma = S^{-1} \cdot 0 = 0$. Таким образом, $\forall \Gamma \in \mathbb{R}^n \backslash \{0\}$ $f(\Gamma) = g(S\Gamma) > 0 \Rightarrow$ квадратичная форма f является положительноопределенной. ⊠

Следствие. *Все положительноопределенные квадратичные формы являются эквивалентными.*

Следствие. *Матрица положительноопределенной квадратичной формы имеет положительный опеределитель.*

♦ Так как положительноопределенная квадратичная форма $f(x)$ эквивалентна нормальной квадратичной форме, то матрица A_f этой квадратичной формы конгруэентна матрице E, то есть существует невырожденная матрица S такая, что $A_f = S^T E S = S^T S \Rightarrow det A_f = det S^T S = det S^T \cdot det S = (det S)^2 > 0$. ⊠

Критерий Сильвестра. *Действительная квадратитчная форма является положительноопределенной* $\Longleftrightarrow$ *все угловые миноры ее матрицы положительны.*

♦ $\Rightarrow$) Пусть квадратичная форма $f(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j$ является положительноопределенной. Тогда квадратичная форма $f_m(x_1, \ldots, x_m) = \sum_{i,j=1}^{m} \alpha_{ij} x_i x_j$, $1 \leqslant m \leqslant n$ также является положительноопределенной. Так как если бы существовала последовательность $(\gamma_1, \ldots, \gamma_m) \neq 0$, для которой $f_m(\gamma_1, \ldots, \gamma_m) \leqslant 0$, то последовательность $(\gamma_1, \ldots, \gamma_m, 0, \ldots, 0)$ также была бы ненулевая и $f(\gamma_1, \ldots, \gamma_m, 0, \ldots, 0) = f_m(\gamma_1, \ldots, \gamma_m) \leqslant 0 \Rightarrow$ матрицы квадратичных форм имеют положительный определеитель, а определитель f_m равен угловому минору Δ_m.

$\Leftarrow$) Если все угловые миноры Δ_i положительны, то по формуле Якоби эта квадратичная форма эквивалентна канонической квадратичной форме с положительными коэффициентами, которая являтся положительноопределенной. Следовательно, квадратичная форма $f(x)$ также положительноопределенная. ⊠

Следствие. *Действительная квадратичная форма является отрицательноопределенной* $\Longleftrightarrow$ *все угловые миноры ее матрицы четного порядка положительны, нечетного порядка — отрицательны.*

• *Минор матрицы называется* ***главным****, если номера строк и номера столбцов, в которых он расположен, совпадают.*

Следствие. *Все главные миноры матрицы положительноопределенной квадратичной формы положительны.*

Глава 14

Евклидовы и унитарные пространства

14.1 Определение и простейшие свойства евклидовых пространств.

• *Пусть V — векторное пространство над полем $\mathbb{R}$. Отображение $V \times V \to \mathbb{R}$, ставящее в соответствие упорядоченной паре векторов $(a, b) \in V$ число $ab \in \mathbb{R}$, называется* ***скалярным произведением****, если $\forall \alpha \in \mathbb{R}, a, b, c \in V$*

1. $ab = ba$,
2. $(\alpha a)b = \alpha(ab)$,
3. $(a + b)c = ac + bc$,
4. $\forall a \neq 0_v, a \cdot a > 0$

• *Действительное векторное пространство с определенным на нем скалярным произведением называется* ***евклидовым пространством****.*

Простейшие свойства скалярного произведения:

1. $a(\alpha b) = \alpha(ab) \quad \forall a, b \in V, \alpha \in \mathbb{R}$
2. $a(b + c) = ab + ac \quad \forall a, b, c \in V$
3. $a \cdot 0_v = 0_v \cdot a = 0_{\mathbb{R}} \quad \forall a \in V$

   ♦ $a \cdot 0_v + a \cdot 0_v = a(0_v + 0_v) = a \cdot 0_v \Rightarrow a \cdot 0_v = 0$ ⊠

• *Пусть $B(b_1, \ldots, b_n)$ — некоторая система векторов евклидова пространства. Матрица* $G_B = \begin{pmatrix} b_1b_1 & b_1b_2 & \ldots & b_1b_n \\ b_2b_1 & b_2b_2 & \ldots & b_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ b_nb_1 & b_nb_2 & \ldots & b_nb_n \end{pmatrix} \in \mathbb{R}_{n,n}$*, составленная из скалярных произведений векторов системы B, называется* ***матрицей Грама*** *системы B.*

Из определения скалярного произведения следует, что матрица Грама является симметрической.

• *Если B — базис пространства V, то матрица G_B называется* ***матрицей скалярного произведения*** *в базисе B.*

Теорема. *Если G_B — матрица скалярного произведения пространства V в базисе B, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ — координатные столбцы векторов $x, y \in V$ в базисе B, то $xy = X^T G_B Y$.*

♦ $xy = (\sum\limits_{i=1}^{n} x_i b_i)(\sum\limits_{j=1}^{n} y_j b_j) = \sum\limits_{i,j=1}^{n} (x_i b_i)(y_j b_j) = \sum\limits_{i,j=1}^{n} x_i y_j (b_i b_j) = \sum\limits_{i=1}^{n} x_i (\sum\limits_{j=1}^{n} b_i b_j) y_j = X^T G_B Y$ ⊠

Следствие. *Матрица скалярного произведения является матрицей положительно определённой квадратичной формы.*

♦ Так как $\forall a \neq 0 \quad a \cdot a > 0$, то и координатный столбец вектора a также ненулевой. Следовательно, $a \cdot a = X_a^T G_B X_a > 0 \Rightarrow$ квадратичная форма $f(x_1, \ldots, x_n) = X^T G_B X > 0$, где $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, положительноопределенная. А так как матрица G_B симметрическая, то она является матрицей квадратичной формы. ⊠

Следствие. *Определитель матрицы скалярного произведения положителен.*

♦ Так как матрица скалярного произведения является матрицей положительно определенной квадратичной формы, то по критерию Сильвестра все угловые миноры положительны, в том числе и определитель матрицы. ⊠

Теорема (Критерий Грама линейной зависимости системы векторов). *Система векторов $B(b_1, \ldots, b_k)$ линейно зависима $\Longleftrightarrow$ определитель её матрицы Грама G_B равен нулю.*

♦ $\Rightarrow$) Пусть система векторов B линейно зависима $\Rightarrow$ существует тривиальная линейная комбинация $\alpha_1 b_1 + \ldots + \alpha_k b_k = 0_v$. Домножим ее справа на $b_1, \ldots, b_k$. Тогда

$$\begin{cases} \alpha_1 b_1 b_1 + \ldots + \alpha_k b_1 b_k = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_1 b_k b_1 + \ldots + \alpha_k b_k b_k = 0; \end{cases} \Rightarrow \begin{pmatrix} b_1 b_1 & \ldots & b_1 b_k \\ \vdots & \ddots & \vdots \\ b_k b_1 & \ldots & b_k b_k \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow G_B X = 0.$$

Линейная однородная система уравнений имеет ненулевое решение $\Rightarrow$ определитель матрицы Грама нулевой.

$\Leftarrow$) Пусть $G_B = 0$. Пойдем от противного. Если система векторов $B(b_1, \ldots, b_n)$ линейно независима, то она является базисом линейной оболочки системы B, а линейная оболочка $L(B)$ — подпространство V, то есть сама является подпространством. Следовательно, G_B — матрица скалярного произведения в пространстве $L(B) \Rightarrow det G_B > 0$ — противоречие. ⊠

Теорема. *Если G_A, G_B — матрицы скалярного произведения в базисах A, B соответственно, то $G_B = S^T G_A S$, где $S = S_{A \to B}$.*

♦ Пусть $x, y \in V$ — произвольные векторы. X_A, X_B, Y_A, Y_B — координатные столбцы соответственно в базисах A и B базисах, $X_A = S X_B$, $Y_A = S Y_B$. Тогда $xy = (X_A)^T G_A Y_A = (S X_B)^T G_A (S Y_B) = X_B^T (S^T G_A S) Y_B \Rightarrow$ элемент, стоящий в i-ой строке и j-ом столбце, $b_i b_j = S^T G_A S$. По определению матрицы Грама, элемент, стоящий в i-ой строке и j-ом столбце, также равен $b_i b_j$. Следовательно, $S^T G_A S = G_B$. ⊠

14.2 Ортогональные векторы.

Пусть V — евклидово пространство.

• *Векторы a и b называются **ортогональными**, если их скалярное произведение равно нулю. Система векторов $B(b_1, \dots, b_n)$ называется **ортогональной**, если все её векторы попарно ортогональны, то есть $b_i b_j = 0$ для всех $i \neq j$.*

Матрица Грама ортогональной системы векторов диагональная, причем на диагонали расположены положительные действительные числа.

Лемма. *Ортогональная система ненулевых векторов линейно независима.*

⧫ Поскольку матрица ортогональной системы векторов Грама диагональная, то её определитель больше нуля. Следовательно, по критерию Грама, система векторов является линейно независимой. ⊠

Теорема. *В любом евклидовом пространстве существует ортогональный базис.*

⧫ Пусть система векторов $B(b_1, \dots, b_n)$ является базисом пространства V. Построим $A(a_1, \dots, a_n)$ следующим образом:

1. $a_1 = b_1 \neq 0$;
2. $a_2 = b_2 + \alpha a_1$; $a_2 \cdot a_1 = 0$
   Тогда $\underbrace{a_2 a_1}_{0} = a_1 b_2 + \alpha \underbrace{(a_1 a_1)}_{0} \Rightarrow \alpha = -\dfrac{(b_2, a_1)}{(a_1, a_1)}$.
   Причем, так как $a_1 = b_1$ и коэффициент при $b_2 = 1$, то вектор a_2 является нетривиальной линейной комбинацией двух векторов системы B, следовательно, $a_2 \neq 0_v$.
3. $a_3 = b_3 + \alpha_1 a_1 + \alpha_2 a_2$; $a_1 \cdot a_3 = a_2 \cdot a_3 = 0$
   Тогда $\alpha_1 = -\dfrac{(b_3, a_1)}{(a_1, a_1)}$, $\alpha_2 = -\dfrac{(b_3, a_2)}{(a_2, a_2)}$.
   Причем, так как a_1 и a_2 линейно выражаются через b_1 и b_2 и коэффициент при $b_3 = 1$, то вектор a_3 является нетривиальной линейной комбинацией векторов b_1, b_2 и b_3, следовательно, $a_3 \neq 0_v$.

Продолжая рассуждения подобным образом, на k-ом шаге получим: $a_k = b_k + \sum\limits_{i=1}^{k-1} \alpha_i a_i$, $\alpha_i = -\dfrac{(b_k, a_i)}{(a_i, a_1)} \Rightarrow$ получим ортогональную систему векторов A, состоящую из ненулевых векторов $\Rightarrow$ она линейно независима по доказзанной выше лемме $\Rightarrow A$ — ортогональный базис и в нём столько же векторов, сколько и в базисе B. ⊠

• *Алгоритм построения ортогонального базиса из теоремы выше называется **процессом ортогонализации Грама-Шмидта**.*

Так как построенная на k-ом шаге процесса ортогонализации система векторов $(a_1, \dots, a_n)$ ортогональна и эквивалентна $(b_1, \dots, b_n)$, то в результате применения процесса ортогонализации к линейно зависимой системе векторов, на некотором шаге получим нулевой вектор.

Следствие. *Любую ортогональную систему ненулевых векторов можно дополнить до ортогонального базиса.*

♦ Если ортогональная система векторов линейно независима, то её можно дополнить до базиса пространства. Если к этому базису применить процесс ортогонализации, то первые векторы построенной системы будут совпадать с исходной ортогональной системой. Следовательно, полученный ортогональный базис является дополнением до ортогонального базиса исходной ортогональной системы векторов. ⊠

14.3 Длина вектора. Ортонормированный базис.

Пусть V — евклидово пространство.

• ***Длиной*** *вектора* $a \in V$ *называется число* $|a| = \sqrt{a \cdot a} \geqslant 0$.

Свойства длины вектора:

1. $|a| = 0 \iff a = 0_v$.

   ♦ Следует из определения скалярного произведения и его простейших свойств. ⊠

2. $|\alpha a| = |\alpha| \cdot |a|, \quad \forall a, b \in V, \alpha \in \mathbb{R}$.

   ♦ $|\alpha a|^2 = (\alpha \cdot a) \cdot (\alpha \cdot a) = (\alpha \cdot \alpha) \cdot (a \cdot a) = |\alpha|^2 \cdot |a|^2$. Извлечем корень и получим $|\alpha a| = |\alpha| \cdot |a|$. ⊠

3. **Неравенство Коши-Буняковского:** $|ab| \leqslant |a||b|$, *причем* $|ab| = |a||b| \iff a, b$ *— линейно зависимы.*

   ♦ Для любых векторов a, b пространства V матрица Грама имеет неотрицательный определитель. Причем определитель матрицы Грама равен нулю в том и только в том случае, когда векторы a, b линейно зависимы. Тогда $detG = \begin{vmatrix} aa & ab \\ ba & bb \end{vmatrix} = (a \cdot a) \cdot (b \cdot b) - (a \cdot b)^2 = |a|^2 \cdot |b|^2 - |a \cdot b|^2 \geqslant 0 \Rightarrow |ab|^2 \leqslant |a|^2|b|^2$. Извлечем корень и получим $|ab| \leqslant |a||b|$, причем равенство выполняется при условии, что векторы a и b линейно зависимы. ⊠

4. **Неравенство треугольника:** $|a + b| \leqslant |a| + |b|, \quad \forall a, b \in V$.

   ♦ $|a+b|^2 = (a+b) \cdot (a+b) = a \cdot a + b \cdot b + 2ab \leqslant |a|^2 + |b|^2 + 2|a \cdot b| \leqslant |a|^2 + |b|^2 + 2|a| \cdot |b| = (|a| + |b|)^2$. Извлечем корень и получим $|a + b| \leqslant |a| + |b|$. ⊠

5. **Теорема Пифагора:** *Если векторы* a *и* b *ортогональны, то* $|a + b|^2 = |a|^2 + |b|^2$.

   ♦ Пусть векторы a и b ортогональны. Тогда их скалярное произведение равно нулю. Следовательно, $|a + b|^2 = (a + b) \cdot (a + b) = a \cdot a + b \cdot b + 2ab = |a|^2 + |b|^2$. ⊠

Из неравенства Коши-Буняковского следует

• *В евклидовом пространстве* ***углом*** *между векторами* a *и* b *называется величина* $\varphi = arccos\dfrac{ab}{|a| \cdot |b|} \in [0, \pi]$.

• *Вектор называется* ***нормированным****, если его длина равна 1.*

• *Система векторов называется* ***ортонормированной****, если она ортогональна и каждый ее вектор нормированный.*

Матрица Грама ортонормированной системы векторов является единичной. Пусть B —

ортнормированный базис и x, y — произвольные векторы пространства V с координатными столбцами $X = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, Y = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ в базисе B. Тогда $xy = X^T G_B Y = [G_B = E] = X^T Y = \alpha_1\beta_1 + \ldots + \alpha_n\beta_n$.

Теорема. *В любом евклидовом пространстве существует ортонормированный базис.*

♦ В любом евклидовом пространстве V существует ортогональный базис $B(b_1, \ldots, b_n)$. Пронормируем каждый вектор b_i этого базиса, то есть домножим на $\dfrac{1}{|b_i|}$, и получим систему векторов $B' = (b'_1, \ldots, b'_n)$, для векторов которой справедливо следующее

$$b'_i b'_j = \frac{1}{|b_i|}b_i \cdot \frac{1}{|b_j|}b_j = \frac{1}{|b_i||b_j|}(b_i, b_j) = \begin{cases} \dfrac{1}{|b_i||b_j|}(b_i, b_j) = 1, \ i = j, \\ \dfrac{1}{|b_i||b_j|} \cdot 0 = 0, \ i \neq j; \end{cases} \Rightarrow \text{система векторов } B'$$

ортонормированная и линейно независимая, так как является ортогональной системой ненулевых векторов. При этом количество векторов в этой системе равно количеству векторов в базисе B пространства. Следовательно, эта система сама является базисом. ⊠

• *Действительная матрица $S \in \mathbb{R}_{n,n}$ называется* ***ортогональной****, если $S^T S = SS^T = E$ $(S^{-1} = S^T)$.*

Теорема. *Матрица перехода $S_{A\to B}$ от ортонормированного базиса A к ортонормированному базису B ортогональна.*

♦ Пусть G_A, G_B — матрицы скалярного произведения в ортонормированных базисах A и B соответственно. Тогда $G_B = S^T G_A S$, где $S = S_{A\to B}$. А так как базисы ортонормированные, то $G_A = G_B = E$. И получаем $E = S^T E S = S^T S$. ⊠

14.4 Ортогональное дополнение подпространства.

• *Пусть V — евклидово пространство, U — его подпространство. Множество всех векторов пространства V, ортогональных каждому вектору из U, называется* ***ортогональным дополнением*** *подпространства U (Обозначение: $U^\perp$).*

Теорема. *Ортогональное дополнение любого подпространства U векторного пространства V также является подпространством, причем $U \cap U^\perp = \{0_v\}$.*

♦ Так как нулевой вектор ортогонален любому вектору пространства V, то он ортогонален и любому вектору подпространства U. Следовательно, $0_v \in U^\perp$, значит $U^\perp$ — непустое множество.

Пусть векторы $x, y \in U^\perp \Rightarrow \forall a \in U \quad ax = ay = 0 \Rightarrow \begin{cases} a(x+y) = ax + ay = 0 + 0 = 0, \\ a(\alpha x) = \alpha(ax) = \alpha \cdot 0 = 0; \end{cases} \Rightarrow$ $(x + y), \alpha x \in U^\perp \Rightarrow U^\perp$ — подпространство V.

Покажем, что $U \cap U^\perp = \{0_v\}$. Пусть вектор $x \in U \cap U^\perp \Rightarrow x \in U, \ x \in U^\perp \Rightarrow$ вектор x ортогонален самому себе $\Rightarrow x \cdot x = 0 \Rightarrow x = 0_v$. ⊠

Теорема об ортогональном разложении пространства. *Для любого подпространства U векторное пространство V представимо в виде $V = U \oplus U^\perp$*

♦ Пусть U — тривиальное подпространство.

1. Если $U = \{0_v\}$, то $U^\perp = V$, так как нулевой вектор ортогонален любому вектору.

2. Если $U = V$, то $U^\perp = \{0_v\}$, так как $U \cap U^\perp = \{0_v\} \Rightarrow V = V + \{0_v\} = U + U^\perp$.

Пусть U — нетривиальное подпространство. Тогда в нем существует ортонормированный базис $B(b_1, \ldots, b_k)$. Дополним его до ортогонального базиса пространства V и получим систему векторов $(b_1, \ldots, b_k, a_1, \ldots, a_s)$. Тогда, так как любой вектор из линейной оболочки $L(a_1, \ldots, a_s)$ ортогонален любому вектору вектору из линейной оболочки $L(b_1, \ldots, b_k)$, то $(\sum\limits_{i=1}^{s} \alpha_i a_i)(\sum\limits_{j=1}^{k} \beta_j b_j) = \sum\limits_{i=1}^{s}\sum\limits_{j=1}^{k}(\alpha_i\beta_j)(a_i b_j) = 0 \Rightarrow L(a_1, \ldots, a_s) \subseteq U^\perp$. А так как $(b_1, \ldots, b_k, a_1, \ldots, a_s)$ — базис, то любой вектор $x \in V$ представим в виде $x = \underbrace{\beta_1 b_1 + \ldots + \beta_k b_k}_{\in U} + \underbrace{\alpha_1 a_1 + \ldots + \alpha_s a_s}_{\in U^\perp} \Rightarrow U + U^\perp = V$, а так как $U \cap U^\perp = \{0_v\}$, то сумма подпространств прямая. ⊠

Следствие. $\dim U + \dim U^\perp = \dim V$

♦ Доказательство следует из формулы Грассмана. ⊠

Следствие. $(U^\perp)^\perp = U$.

♦ Так как предыдущее следствие справедливо для любого подпространства, то оно српаведливо и для $U^\perp$. Тогда, по предыдущему следствию, $\begin{cases} \dim U = \dim V - \dim U^\perp \\ (\dim U^\perp)^\perp = \dim V - \dim U^\perp \end{cases} \Rightarrow$ $\dim U = \dim(U^\perp)^\perp$, причем любой вектор $a \in U$ ортогонален любому вектору $b \in U^\perp \Rightarrow U \subseteq (U^\perp)^\perp \Rightarrow$ по теореме о монотонности размерности, $U = (U^\perp)^\perp$ ⊠

• *Так как для любого подпространства U векторное пространство V представимо в виде $V = U \oplus U^\perp$, то любой вектор $a \in V$ представим единственным образом в виде $a = a_1 + a_2$, где $a_1 \in U$, $a_2 \in U^\perp$. Вектор a_1 называется* ***ортогональной проекцией*** *вектора a на подпространство U, вектор a_2 —* ***ортогональной составляющей*** *вектора a относительно подпространства U.*

14.5 Ортогональный оператор.

Пусть V — евклидово пространство.

• *Линейный оператор $f : V \to V$ называется* ***ортогональным****, если $f(x) \cdot f(y) = x \cdot y$ для всех $x, y \in V$.*

Свойства ортогонального оператора:

1. *Линейный оператор пространства V является ортогональным $\iff |f(x)| = |x|$ для всех $x \in V$.*

   ♦ $\Rightarrow$) Пусть f – ортогональный оператор $\Rightarrow |f(x)|^2 = f(x) \cdot f(x) = x \cdot x = |x|^2 \Rightarrow |f(x)| = |x|$.

   $\Leftarrow$) Пусть f сохраняет длины всех векторов $\Rightarrow \forall x, y \in V \ |x + y| = |f(x + y)| = |f(x) + f(y)|$.
   $|x + y|^2 = (x + y)(x + y) = xx + 2xy + yy = |x|^2 + 2xy + |y|^2$.
   $|f(x + y)|^2 = |f(x)^2| + 2f(x)f(y) + |f(y)^2| = |x|^2 + |y|^2 + 2f(x)f(y) \Rightarrow xy = f(x)f(y)$. ⊠

2. *Модуль собственного значения ортогонального оператора равен 1.*

   ♦ Пусть x_0 – собственный вектор, соответствующий собственному значению λ_0, то есть $f(x_0) = \lambda_0 x_0 \Rightarrow |x_0| = |f(x_0)| = |\lambda_0 x_0| = |\lambda_0| \cdot |x_0|$. Так как вектор x_0 собственный, то он ненулевой $\Rightarrow |x_0| \neq 0 \Rightarrow |\lambda_0| = 1$. ⊠

3. *Ортогональный оператор биективен.*

   ♦ $\forall x \in Ker\ f \quad f(x) = 0_v \Rightarrow |x| = |f(x)| = |0_v| = 0 \Rightarrow x = 0_v \Rightarrow Ker\ f = \{0_v\} \Rightarrow f$ — иньекция $\Rightarrow f$ — сюрьекция, а, значит, и биекция. ⊠

Теорема. *Линейный оператор является ортогональным $\Longleftrightarrow$ он ортонормированный базис отображает в ортонормированный.*

♦ $\Rightarrow$) Так как f ортогональный оператор, он ортонормированный базис $B(b_1, \ldots, b_n)$ отобразит в ортонормированную систему векторов. А так как ортогональный оператор сохраняет длины и скалярные произведения, то система векторов $f(B)$ состоит из попарно ортогональных ненулевых векторов, а, значит, является линейно независимой системой и содержит столько же векторов, сколько и базис. Следовательно, система сама является базисом.

$\Leftarrow$) Пусть оператор f переводит ортонормированный базис $B(b_1, \ldots, b_n)$ в ортонормированный базис $f(B) = (f(b_1), \ldots, f(b_n))$, и пусть для произвольных векторов $x, y \in V$ X, Y — координатные столбцы в базисе B. Тогда $xy = X^T G_B Y = X^T Y$.

Если $X = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, то $x = \alpha_1 b_1 + \ldots + \alpha_n b_n \Rightarrow f(x) = \alpha_1 f(b_1) + \ldots + \alpha_n f(b_n) \Rightarrow$ вектор $f(x)$ в базисе $f(B)$ имеет координатный столбец равный X. Аналогично для Y. Следовательно, $f(x)f(y) = X^T G_{f(B)} Y = X^T Y = xy$. ⊠

Теорема. *Линейный оператор является ортогональным $\Longleftrightarrow$ его матрица в ортонормированном базисе ортогональная.*

♦ $\Rightarrow$) Пусть $B(b_1, \ldots, b_n)$ — ортонормированный базис пространства V, $A = M_f^B$, A_i — столбцы матрицы A. Тогда A_i – координатный столбец вектора $f(b_i)$ в базисе $B \Rightarrow f(b_i)f(b_j) = A_i^T G_B A_j = A_i^T A_j$, то есть равно элементу матрицы $A^T A$, стоящему в i-й строке и j-м столбце.

С другой стороны, так как оператор f ортогональный, то $f(b_i)f(b_j) = \begin{cases} 0, & \text{если } i \neq j, \\ 1, & \text{если } i = j; \end{cases} \Rightarrow$ $A^T A = E$, то есть матрица ортогональная.

$\Leftarrow$) Пусть матрица M_f^B оператора f в ортонормированном базисе B равна ортогональной матрице A. Так как базис B ортонормированный, то $G_B = E$.
Пусть $x, y \in V$ — произвольные векторы с координатными столбцами X и Y в базисе B соответственно. Тогда $xy = X^T G_B Y = X^T Y$.
С другой стороны, векторы $f(x)$ и $f(y)$ в базисе B имеют координатные столбцы AX и AY соответственно $\Rightarrow f(x)f(y) = (AX)^T G_B AY = X^T (A^T A) Y = [A^T A = E] = X^T Y = xy$. ⊠

14.6 Самосопряженный оператор.

Пусть V — евклидово пространство.

• *Линейный оператор $f : V \to V$ называется* ***самосопряжённым****, если $f(x) \cdot y = x \cdot f(y)$ для всех векторов $x, y \in V$.*

Теорема. *Линейный оператор является самосопряженным $\iff$ его матрица в ортонормированном базисе симметрическая.*

⧫ $\Rightarrow$) Пусть $B(b_1, \ldots, b_n)$ — ортонормированный базис пространства V, а матрица $A(\alpha_{ij})$ равна матрице оператора f в базисе B, то есть M_f^B. Тогда $f(b_i) = \alpha_{1i}b_1 + \ldots + \alpha_{ni}b_n$. Из данного равенства имеем
$\begin{cases} f(b_i) \cdot b_j = (\alpha_{1i}b_1 + \ldots + \alpha_{ni}b_n) \cdot b_j = \alpha_{ji}, \\ f(b_j) \cdot b_i = (\alpha_{1j}b_1 + \ldots + \alpha_{nj}b_n) \cdot b_i = \alpha_{ij}; \end{cases}$ $\Rightarrow \alpha_{ij} = \alpha_{ji}$, так как оператор f самосопряженный.

$\Leftarrow$) Пусть матрица A оператора оператора f в базисе B является симметрической. Тогда $G_B = E \Rightarrow \forall x, y \in V$ X, Y — координатные столбцы векторов x и y в базисе B соответственно. Следовательно, $f(x) \cdot y = (AX)^T G_B Y = X^T A^T Y = X^T G_B (AY) = x \cdot f(y)$.
⊠

Свойства самосопряженных операторов:

1. *Характеристический многочлен самосопряженного оператора имеет с учетом кратности столько действительных корней, какова его степень.*

   Следствие. *Характеристический многочлен действительной симметрической матрицы порядка n имеет с учетом кратности ровно n действительных корней.*

2. *Собственные векторы самосопряженного оператора, соответствующие различным собственным значениям, ортогональны.*

   ⧫ Пусть a_1, a_2 — собственные векторы самосопряженного оператора f, соответствующие собственным значениям λ_1 и λ_2 таким, что $\lambda_1 \neq \lambda_2$.
   Тогда $f(a_1) = \lambda_1 a_1$ и $f(a_2) = \lambda_2 a_2$. При этом
   $\begin{cases} f(a_1) \cdot a_2 = \lambda_1 \cdot (a_1 \cdot a_2), \\ a_1 \cdot f(a_2) = \lambda_2 \cdot (a_1 \cdot a_2); \end{cases}$ $\Rightarrow$ так как оператор f самосопряженный, то $f(a_1) \cdot a_2 = a_1 \cdot f(a_2) \Rightarrow \lambda_1 \cdot (a_1 \cdot a_2) = \lambda_2 \cdot (a_1 \cdot a_2) \Rightarrow \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \cdot (a_1 \cdot a_2) = 0 \Rightarrow a_1 \cdot a_2 = 0$.
   Следовательно, векторы a_1 и a_2 ортогональны. ⊠

3. *Если f — самосопряженный оператор пространства V, U — инвариантное относительно f подпространство, то подпространство $U^\perp$ также инвариантно относительно f.*

   ⧫ Покажем, что $\forall x \in U^\perp$ $f(x) \in U^\perp$.
   Так как U — инвариантное относительно оператора f подпространство, то $\forall y \in U$ $f(y) \in U \Rightarrow x \cdot f(y) = 0_v \Rightarrow f(x) \cdot y = 0_v$, так как оператор f самосопряженный. Следовательно, $f(x) \in U^\perp$. ⊠

Теорема. *Для любого самосопряженного оператора f существует ортонормированный базис пространства V, состоящий из собственных векторов оператора f.*

⧫ Пусть f — самосопряженный оператор, и $\lambda_1, \ldots, \lambda_s$ — различные собственные значения оператора f. Обозначим через $L(\lambda_i)$ собственные подпространства оператора f, соответствующие собственным значениям λ_i. Построим в каждом из них ортонормированный базис B_i, а затем построим систему векторов $B(B_1, \ldots, B_s)$. Так как собственные векторы самосопряженного оператора, соответствующие попарно различным собственным значениям, ортогональны, то система B ортонормированная. Значит, система B линейно независимая, так как является ортогональной системой ненулевых векторов.

Покажем, что B — базис пространства V. От противного. Пусть система B не является базисом $\Rightarrow L(B) \neq V \Rightarrow \dim(L(B)) = k < n$. При этом $L(B) \oplus (L(B))^{\perp} = V \Rightarrow \dim(L(B))^{\perp} \geqslant 1$.

Покажем, что подпространство $L(B)$ инвариантно относительно оператора f. Пусть $x \in L(B)$. Тогда $x = \beta_1 b_1 + \ldots + \beta_k b_k$, а $f(x) = \beta_1 f(b_1) + \ldots + \beta_k f(b_k) = \beta_1\lambda_1 b_1 + \ldots + \beta_k \lambda_k b_k$ — линейная комбинация, а, значит, принадлежит линейной оболочке $L(B) \Rightarrow L(B)$ — инвариантное относительно оператора f подпространство пространства $V \Rightarrow$ подпространство $(L(B))^{\perp}$ также инвариантно относительно f.

Значит, существует ограничение $f|_{(L(B))^{\perp}}$, которое также является самосопряженным оператором. Следовательно, у него существует собственное значение λ_0, а, значит, и собственный вектор x_0, соответствующий данному собственному значению. Тогда $f(x_0) = f|_{(L(B))^{\perp}}(x_0) = \lambda_0 x_0 \Rightarrow x_0$ — собственный вектор и для оператора f, который принадлежит $L(B)$, а это противоречит тому, что подпространство $L(B)$ содержит все собственные векторы. Следовательно, система B является базисом пространства V. ⊠

14.7 Приведение действительной квадратичной формы к каноническому виду ортогональным преобразованием переменных.

Пусть

$$f(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j, \quad \alpha_{i,j} \in \mathbb{R}, \quad \alpha_{ij} = \alpha_{ji},$$

или

$$f(X) = X^T A X, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A(\alpha_{ij}) \in \mathbb{R}_{n,n}, \quad \alpha_{ij} = \alpha_{ji},$$

— некоторая действительная квадратичная форма.

• *Линейное преобразование переменных $X = SY$ называется* ***ортогональным****, если матрица S ортогональная, то есть $S^T S = E$.*

Так как ортогональная матрица является невырожденной, то и ортогональное преобразование тоже невырожденное.

Лемма. *Для любой действительной симметричесокй матрицы A существует ортогональная матрица S такая, что матрица $S^T A S$ является диагональной, причем ее диагональные элементы равны собственным значениям матрицы A.*

⧫ Так как отображение, ставящее в соответствие каждому линейному оператору пространства $\mathbb{R}$ его матрицу в базисе , является изоморфизмом, то $\exists$ линейный оператор $f \in \mathbb{R}^n$: $M_f^E = A$. Если базис E ортонормированный, то f является самосопряженным оператором, так как матрица A симметрическая. Следовательно, существует ортонормированный базис B, состоящий из собственных векторов $\Rightarrow M_f^B$ — диагональная матрица, причем диагональными элементами являются собственные значения оператора f. Так как M_f^B и A являются матрицами одного линейного оператора в разных базисах, они подобны, то есть $\exists S : M_f^B = S^{-1}M_f^E S = S^{-1}AS$, где $S = S_{E\to B}$. А так как базисы E, B ортонормированы, то матрица S ортогональна и $M_f^B = S^T AS$. ⊠

Теорема. *Для любой действительной квадратичной формы $f(X)$ существует линейное ортогональное преобразование $X = SY$, переводящее f в каноническую квадратичную форму $\lambda_1 y_1^2 + \ldots \lambda_n y_n^2$, причем числа λ_i являются собственными значениями матрицы квадратичной формы f.*

♦ Пусть A — симметрическая матрица квадратичной формы f. Тогда $f(X) = X^T AX \Rightarrow$ существует ортогональная матрица S такая, что $S^T AS = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = D$.
Применим замену $X = SY$ с матрицей S к квадратичной форме f и получиам $f(SY) = (SY)^T A(SY) = Y^T(S^T AS)Y = Y^T DY = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$. ⊠

Следствие (критерий положительной определенности квадратичной формы). *Действительная квадратичная форма является положительно определенной $\iff$ все собственные значения ее матрицы положительны.*

14.8 Полярное разложение линейного оператора.

Лемма. *Для любой действительной невырожденной матрицы A квадратичная форма $f(X) = X^T(A^T A)X$ является положительно определенной.*

♦ Поскольку матрица A невырожденная, то линейная однородная система уравнений $AX = 0$ имеет лишь нулевое решение. Следовательно, для любого ненулевого столбца X_0 столбец $AX_0 = \begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}$ является ненулевым. Тогда $f(X_0) = (X_0^T A^T)(AX_0) = (AX_0)^T(AX_0) = (\gamma_1, \ldots, \gamma_n)\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix} = \gamma_1^2 + \ldots + \gamma_n^2 > 0$, поскольку все γ_i не обращаются в ноль одновременно. Следовательно, квадратичная форма является положительноопределенной. ⊠

Следствие. *Для любой квадратной действительной невырожденной матрицы A собственные значения матрицы $A^T A$ положительные.*

♦ Поскольку $(A^T A)^T = A^T A$, то есть при транспонировании матрица $A^T A$ не изменилась, то матрица A симметрическая. Следовательно, она является матрицей квадратичной формы $f(X) = X^T(A^T A)X$, которая по предыщуей лемме является положительноопределенной. Тогда, по критерию положительной определенности квадратичной формы, все собственные значения матрицы $A^T A$ положительны. ⊠

Лемма. *Для любой квадратной невырожденной матрицы A существует симметрическая матрица C с положительными собственными значениями такая, что $A^T A = C^2$*

♦ Поскольку матрица $A^T A$ симметрическая, существует ортогональная матрица S такая, что матрица $S^T A^T AS$ является диагональной матрицей, содержащей все собственные значения матрицы $A^T A$ на диагонали.

Так как квадратичная форма с матрицей $A^T A$ является положительноопределенной по предыдущему следствию, то все собственные значения положительны. Тогда существует матрица $D = \operatorname{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$, причем $D^2 = S^T A^T A S$. Следовательно, $A^T A = (S^T)^{-1} D^2 S^{-1}$, и так как матрица S ортогональная, то есть $S^{-1} = S^T$, то $A^T A = (S^{-1})^{-1} D^2 S^T = SDEDS^T = (SDS^T)(SDS^T)$.

Обозначим через C матрицу SDS^T, тогда получим $A^T A = C^2$, при этом C — симметрическая матрица, так как $C^T = (SDS^T)^T = SD^T S^T =$ [диагональ матрицы D не изменяется при транспонировании] $= SDS^T = C$. А так как матрица S ортогональная, то $C = SDS^T = SDS^{-1}$. Следовательно, матрицы C и D подобны и имеют одинаковые собственные значения равные $\sqrt{\lambda_i} > 0$. ⊠

Теорема. *Любая действительная невырожденная квадратная матрица представима в виде произведения отрогональной матрицы и симметрической матрицы с положительными собственными значениями.*

⧫ Пусть A — произвольная действительная невырожденная матрица. Тогда существует симметрическая матрица C с положительными собственными значениями такая, что $C^2 = A^T A$.

Так как матрица C невырожденная, существует обратная ей невырожденная матрица C^{-1}. Построим матрицу $U = AC^{-1} \Rightarrow A = UC$. Покажем, что матрица U ортогональная: $U^T U = (AC^{-1})^T AC^{-1} = (C^{-1})^T A^T A C^{-1} = (C^{-1})^T C^2 C^{-1} = (C^{-1})^T C =$ [так как матрица C симметрическая] $= (C^{-1})^T C^T = (CC^{-1})^T = E$. ⊠

Следствие. *Любой биективный линейный оператор евклидового пространства представим в виде композиции ортогонального оператора и самосопряженного оператора с положительными собственными значениями.*

• *Представление линейного оператора евклидового пространства в виде композиции ортогонального оператора и самосопряженного оператора называется* ***полярным разложением*** *линейного оператора.*

14.9 Унитарное пространство.

• *Пусть V — векторное пространство над полем $\mathbb{C}$. Отображение $V \times V \to \mathbb{C}$, ставящее в соответствие упорядоченной паре векторов (a, b) число $ab \in \mathbb{C}$, называется* ***скалярным произведением****, если $\forall \alpha \in \mathbb{C}, a, b, c \in V$*

1. $ab = \overline{ba}$
2. $(\alpha a)b = \alpha(ab)$
3. $a(b + c) = ab + ac$
4. $\forall a \neq 0_v, a \cdot a \in R$ и $a \cdot a > 0$

• *Комплексное векторное пространство с определенным на нем скалярным произведением называется* ***унитарным пространством****.*

Простейшие свойства скалярного произведения:

1. $a(\alpha b) = \overline{\alpha}(ab), \quad \forall a, b \in V, \alpha \in \mathbb{C}.$

   ⧫ $a(\alpha b) = \overline{(\alpha b)a} = \overline{\alpha(ba)} = \overline{\alpha}(\overline{ba}) = \overline{\alpha}(ab).$ ⊠

2. $a(b+c) = ab + ac, \quad \forall a, b, c \in V$.

3. $a \cdot 0_v = 0_v \cdot a = 0_{\mathbb{C}}, \quad \forall a \in V$.

Из определения скалярного произведения следует, что для элементов матрицы Грама $G_B(\gamma_{ij})$ произвольной системы векторов B справедливо равенство $\gamma_{ij} = b_i b_j = \overline{b_j b_i} = \overline{\gamma_{ji}}$.

• *Матрица $B(\beta_{ij})$ называется* ***эрмитово сопряженной*** *для матрицы $A(\alpha_{ij})$, если $\beta_{ij} = \overline{\alpha_{ji}}$ (обозначение A^*), то есть $A^* = \overline{A}^T$.*

• *Матрица A называется* ***эрмитовой****, если $A = A^*$.*

• *Матрица A называется* ***унитарной****, если $A^* \cdot A = A \cdot A^* = E$.*

Таким образом матрица Грама любой системы векторов и, следовательно, матрица скалярного произведения пространства V является эрмитовой.

Теорема. *Если G_B — матрица скалярного произведения пространства V в базисе B, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ — координатные столбцы векторов $x, y \in V$ в базисе B, то $xy = X^T G_B \overline{Y}$*

♦ $xy = (\sum\limits_{i=1}^{n} x_i b_i)(\sum\limits_{j=1}^{n} y_j b_j) = \sum\limits_{i,j=1}^{n} (x_i b_i)(y_j b_j) = \sum\limits_{i,j=1}^{n} (x_i \overline{y_j})(b_i b_j) = X^T G_B \overline{Y}$ ⊠

Следствие. *Если G_A, G_B — матрицы скалярного произведения в базисах A, B соответственно, то $G_B = S^T G_A \overline{S}$, где $S = S_{A \to B}$.*

♦ Для любых векторов $x, y \in V \quad xy = X_A^T G_A \overline{A} = (SX_B)^T G_A S Y_B = X_A^T (S^T G_A S) Y_B$. Так как равенство выполняется для всех векторов из пространства V, оно выполняется и для всех векторов из базиса $B \Rightarrow G_B = S^T G_A \overline{S}$. ⊠

Следствие. *Определитель матрицы скалярного произведения равен действительному положительному числу.*

♦ Пусть система векторов $A(a_1, \ldots, a_n)$ является ортогональным базисом пространства V. Тогда матрица G_A диагональная и состоит из действительных чисел, при этом все диагональные элементы положительны. Следовательно, определитель матрицы G_A также положительное действительное число. Так как для любого базиса B выполняется $G_B = S^T G_A \overline{S}$, то $detG_B = det(S^T) \cdot det(G_A) \cdot det(\overline{S}) = det(S) \cdot det(G_A) \cdot |det(S)|^2 \cdot det(G_A) > 0$ ⊠

• ***Длиной*** *вектора $a \in V$ такого, что $a \cdot a \in \mathbb{R}$, называется дейтсвительное число $|a| = \sqrt{a \cdot a} \geqslant 0$.*

Если каждый вектор b_i ортогонального базиса B умножить на $|\bar{b}_i|$, то получим ортонормированный базис пространства V. Следовательно, в любом унитарном постранстве существует ортонормированный базис.

Теорема. *Матрица перехода от ортонормированного базиса к ортонормированному базису является унитарной, то есть $S^* \cdot S = E$.*

♦ Если системы векторов A и B — ортонормированные базисы пространства V, то $G_A = G_B = E$. При этом, $G_B = S^T G_A \overline{S}$, где $S = S_{A\to B}$. Следовательно, $E = S^T \cdot E \cdot \overline{S}$, $E = E^T = (S^T\overline{S})^T = \overline{S}^T S = S^* \cdot S$. ⊠

• *Линейный оператор унитарного пространства V называется* ***унитарным (изометрическим)****, если $f(x)f(y) = xy$ для всех векторов $x, y \in V$.*

• *Линеный оператор называется* ***самосопряженным****, если $f(x)y = xf(y)$ для всех векторов $x, y \in V$.*

Свойства унитарного и самосопряженного оператора:

1. *В ортонормированном базисе унитарный оператор имеет унитарную матрицу, самосопряженный – эрмитову.*

   ♦ Пусть $B(b_1, \ldots, b_n)$ — ортонормированный базис пространства V, $A(\alpha_{ij})$ — матрица оператора f в базисе B. Тогда
   $f(b_k)b_s = (\alpha_{1k}b_1 + \ldots + \alpha_{nk}b_n)b_s = \alpha_{1k}(b_1b_s) + \ldots + \alpha_{nk}(b_nb_s) = \alpha_{sk}$.
   $b_kf(b_s) = b_k(\alpha_{1s}b_1 + \ldots + \alpha_{ns}b_n) = \overline{\alpha_{1s}}(b_kb_1) + \ldots + \overline{\alpha_{ns}}(b_kb_n) = \overline{\alpha_{ks}}$.
   Если f — самосопряженный оператор, то $\alpha_{sk} = \overline{\alpha_{ks}}$, следовательно, матрица A эрмитова.
   Пусть f — унитарный оператор. Тогда
   $f(b_k)f(b_s) = (\sum\limits_{i=1}^{n} \alpha_{ik}b_i)(\sum\limits_{j=1}^{n} \alpha_{js}b_j) = \sum\limits_{i,j=1}^{n} (\alpha_{ik}b_i)(\alpha_{js}b_j) = \sum\limits_{i,j=1}^{n} (\alpha_{ik}\overline{\alpha_{js}})(b_ib_j) = \sum\limits_{i,j=1}^{n} (\alpha_{ik}\overline{\alpha_{js}})$.
   Получаем, что $\alpha_{ik}\overline{\alpha_{js}}$ — элемент матрицы $A^T\overline{A}$ в k-ой строке и s-ом столбце. Но
   $b_kb_s = \begin{cases} 1, \text{ если } k = s, \\ 0, \text{ если } k \neq s. \end{cases}$
   Следовательно, $A^T\overline{A} = E \Rightarrow \overline{A}^T A = E^T = E \Rightarrow A^* \cdot A = E$. ⊠

2. *Модуль собственного значения унитарного оператора равен 1. Все собственные значения самосопряженного оператора — действительные числа.*

   ♦ Пусть a — собственный вектор оператора f, соответствующий собственному значению λ, то есть $f(a) = \lambda a$.

   Если f — унитарный оператор, то $|a|^2 = aa = f(a)f(a) = (\lambda a)(\lambda a) = \lambda\lambda(aa) = |\lambda|^2|a|^2$. Так как вектор a ненулевой, то $|\lambda|^2 = 1 \Rightarrow |\lambda| = 1$.

   Если f — самосопряженный оператор, то $f(a)a = \lambda aa = \lambda|a|^2$; $af(a) = a\lambda a = \overline{\lambda}|a|^2$. Так как f — самосопряженный оператор, то $\lambda|a|^2 = \overline{\lambda}|a|^2 \Rightarrow \lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}$. ⊠

• *Любой линейный оператор унитарного пространства можно представить в виде композиции унитарного и самосопряженного операторов. Любая матрица представима в виде $A = U \cdot H$, где U — унитарная матрица, а H — эрмитова. Такое представление называется* ***полярным разложением*** *линейного оператора/матрицы.*

14.10 Векторные и матричные нормы.

Пусть V — векторное пространство над полем P ($\mathbb{R}$ или $\mathbb{C}$).

• *Отображение $\|\cdot\| : V \to \mathbb{R}$, ставящее в соответствие каждому вектору $a \in V$ действительное число $\|a\|$, называется* ***векторной нормой****, если*

1. $\|a\| \geqslant 0\ \forall a \in V\backslash\{0\}$, и $\|a\| = 0 \Leftrightarrow a = 0_v$
2. $\|\alpha a\| = |\alpha| \cdot \|a\|$
3. $\|a + b\| \leqslant \|a\| + \|b\|$

• *Векторное пространство с определенной на нем векторной нормой называется* ***нормированным****.*

• *Вектор $b \in V$ называется* ***пределом последовательности*** *векторов* $\{a_n\}_{n=1}^{\infty}$, если

$$\forall \varepsilon > 0\ \exists N_\varepsilon : \forall n > N_\varepsilon \Rightarrow \|a_n - b\| < \varepsilon.$$

• *Норма $\|\cdot\|_1$ называется* ***эквивалентной*** *норме $\|\cdot\|_2$, если $\exists \alpha_1, \alpha_2 \in \mathbb{R} : \forall a \in V,\ \alpha_1 \|a\|_2 \leqslant \|a\|_1 \leqslant \alpha_2 \|a\|_2$.*

Если вектор b — предел последовательности $\{a_n\}_{n=1}^{\infty}$ по норме $\|\cdot\|_1$, то он является пределом и по норме $\|\cdot\|_2$ (по любой эквивалентой норме).

В конечномерном векторном пространстве все векторные нормы эквивалентны.

• *Отображенрие $\|\cdot\| : P_{n,n} \to \mathbb{R}$, ставящее в соответствие каждой матрице $A \in P_{n,n}$ действительное число $\|A\|$ называется* ***матричной нормой****, если*

1. $\|A\| \geqslant 0\ \forall A \neq 0$, и $\|A\| = 0 \Leftrightarrow A = 0$
2. $\|\alpha A\| = |\alpha| \cdot \|A\|$
3. $\|A + B\| \leqslant \|A\| + \|B\|$
4. $\|AB\| \leqslant \|A\| \cdot \|B\|$

• *Матричная норма $\|\cdot\|_M$ называется* ***согласованной*** *с векторной нормой $\|\cdot\|_B$, если*

$$\forall A \in P_{n,n}, \quad \forall X \in P_{n,1} \quad \|AX\|_B \leqslant \|A\|_M \|X\|_B.$$

Теорема. *Если матричная норма $\|\cdot\|_M \in P_{n,n}$ является согласованной с векторной $\|\cdot\|_B$ и λ_0 — собственное значение матрицы A, то $|\lambda| \leqslant \|A\|_M$*

⧫ Пусть λ_0 — собственное значение матрицы A. Тогда $\exists x_0$ — ненулевой собственный вектор матрицы A, соответствующий значению λ_0. Следовательно
$\begin{cases} \|Ax_0\|_B = \|\lambda_0 x_0\|_B = |\lambda_0| \, \|x_0\|_B, \\ \|Ax_0\|_B \leqslant \|A\|_M \|x_0\|_B. \end{cases}$ $\Rightarrow$ так как $x_0 \neq 0$, то $\|x_0\|_B \neq 0 \Rightarrow$ поделив обе части неравенства на $\|x_0\|_B$, получим $|\lambda_0| \leqslant \|A\|_M$. ⊠

• *Если векторная норма $\|\cdot\|_B \in P_{n,1}$, то отображение, ставящее каждой матрице $A \in P_{n,m}$ в соответствие число $\|A\|_i = \sup\limits_{\|X\|_B = 1} \|AX\|_B$ является матричной нормой, согласованной с векторной $\|\cdot\|_B$ называется нормой,* ***индуцированной*** *векторной нормой $\|\cdot\|_B$.*

Теорема. *Любая матричная норма $\|\cdot\|_M$, согласованная с некоторой векторной нормой $\|\cdot\|_B$, мажорирует норму $\|\cdot\|_i$, индуцированную этой векторной нормой, то есть $\forall A \in P_{n,n} \quad \|A\|_i \leqslant \|A\|_M$.*

⧫ $\|A\|_i = \underset{\|X\|_B=1}{sup} \|AX\|_B \leqslant \sup\limits_{\|X\|_B=1} \|AX\|_M \|X\|_B = \|A\|_M$. ⊠

14.11 Псевдорешение линейной системы.

Рассмотрим линейную систему

$$AX = B,\ A \in \mathbb{R}_{m,n}, B \in \mathbb{R}_{m,1}, X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \qquad (14.11.1)$$

Обозначим столбцы матрицы A через $A_i \in \mathbb{R}_{m,1}$. Тогда система (14.11.1) может быть записана в виде

$$A_1x_1 + \ldots + A_nx_n = B.$$

Столбцы A_i и B являются элементами пространства $\mathbb{R}_{m,1}$. Определим в этом пространстве скалярное произведение следующим образом: если $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$, то $(X, Y) = x_1y_1 + \ldots + x_my_m = X^TY$.

Лемма. *Пусть* $L = L(A_1, \ldots, A_n)$. *Тогда* $\forall C \in L^\perp,\ A^TC = 0$

♦ i-й эл-т $A^TC = (A_i^T, C) = 0$ ⊠

• *Пусть* $y = (y_1, \ldots, y_n) \in \mathbb{R}$ *— произвольный элемент пространства* $\mathbb{R}^n$. *Столбец* $d(y) = A_1y_1 + \ldots + A_ny_n - B$ *называется* ***невязкой*** *последовательности* y *для системы* (14.11.1).

Если последовательность $y \in \mathbb{R}^n$ является решением системы (14.11.1), то ее невязка равна нулевому столбцу.

Пусть система (14.11.1) несовместна.

• ***Псевдорешением*** *несовместной системы* (14.11.1) *называется последовательность* $y \in \mathbb{R}^n$ *имеющая невязку наименьшей длины.*

Теорема. *Пусть* $L = L(A_1, \ldots, A_n)$, B_1 *— ортогональная проекция столбца* B *на подпространство* L. *Тогда множество псевдорешений системы* (14.11.1) *совпадает с множеством решений системы*

$$A_1x_1 + \ldots + A_nx_n = B_1. \qquad (14.11.2)$$

♦ Так как B_1 — ортогональная проекция столбца B на подпространство L, то $B_1 \in L \Rightarrow$ столбец B_1 представим в виде $B_1 = y_1A_1 + \ldots + y_nA_n$. Следовательно, последовательность, состоящая из коэффициентов этой линейной комбинации, является решением системы (14.11.2), значит, система (14.11.2) совместна.

С другой стороны, так как B_1 — ортогональная проекция B на L, то столбец B представим в виде $B = \underset{\in L}{B_1} + \underset{\in L^\perp}{B_2} \Rightarrow$ для любой последовательности $y = (y_1, \ldots, y_n)$ найдем длину: $|d(y)|^2 = |\underbrace{A_1y_1 + \ldots + A_ny_n - B_1}_{\in L} - \underbrace{B_2}_{\in L^\perp}|^2 =$ [по теореме Пифагора] $= |\underbrace{A_1y_1 + \ldots + A_ny_n - B_1}_{\geqslant 0}|^2 + |B_2|^2 \geqslant |B_2|^2$, причем равенство возможно, если $A_1y_1 + \ldots + A_ny_n - B_1 = 0$. То есть последовательность y является решением системы (14.11.2). Следовательно, множество псевдорешений системы (14.11.1) совпадает с множеством решений системы (14.11.2). ⊠

Теорема. *Множество псевдорешений системы* (14.11.1) *совпадает с множеством решений системы*

$$A^TAX = A^TB. \tag{14.11.3}$$

♦ Покажем, что системы (14.11.2) и (14.11.3) равносильны.

1. Пусть $x = (x_1, \ldots, x_n)$ — решение системы (14.11.2). Тогда $A_1x_1 + \ldots + A_nx_n = B_1$ домножим на A^T и получим $A^T(A_1x_1 + \ldots + A_nx_n) = A^TB_1$. Так как B_1 — проекция B на L, то $B = \underset{\in L}{B_1} + \underset{\in L^\perp}{B_2} \Rightarrow A^TB = A^T(B_1 + B_2) = A^TB_1 + \underset{=0}{A^TB_2} \Rightarrow x$ — решение системы (14.11.3).

2. Пусть $x = (x_1, \ldots, x_n)$ — решение системы (14.11.3) $\Rightarrow A^T(A_1x_1+\ldots+A_nx_n) = A^TB = A^T(B_1+B_2) = A^TB_1 + \underset{=0}{A^TB_2} \Rightarrow A^T(A_1x_1+\ldots+A_nx_n - B_1) = 0 \Rightarrow i$-й элемент столбца слева из последнего равенства имеет вид $\underbrace{A_i^T(A_1x_1 + \ldots + A_nx_n - B_1) = 0}_{\text{скалярное произведение}}$. Следовательно, столбец $(A_1x_1+\ldots+A_nx_n-B_1)$ диагонален $\forall A_i$, а значит ортогонален любому вектору из $L \Rightarrow (A_1x_1 + \ldots + A_nx_n - B_1) \in L^\perp$.

   С другой стороны, так как этот столбец является линейной комбинацией векторов из L, то он принадлежит подпространству L. Отсюда следует, что $(A_1x_1 + \ldots + A_nx_n - B_1) \in L \cap L^\perp = \{0\} \Rightarrow A_1x_1 + \ldots + A_nx_n - B_1 = 0 \Rightarrow x$ — решение системы (14.11.2)

⊠

Следствие. *Если система $AX = B$ совместна, то она равносильна системе $A^TAX = A^TB$.*

♦ По теореме решения системы (14.11.3) имеют наименьшую невязку относительно системы (14.11.1). Следовательно, если система (14.11.1) совместна, то наименьшую невязку для системы (14.11.1) имеют решения этой системы. ⊠

Следствие. *Системы $AX = 0$ и $A^TAX = 0$ равносильны.*

14.12 Нормальное псевдорешение линейной системы.

Рассмотрим несовместную линейную систему

$$A_1x_1 + \ldots + A_nx_n = B,\ A_i, \in \mathbb{R}_{m,n}, B \in \mathbb{R}_{m,1} \tag{14.12.1}$$

Обозначим $L_0 \subseteq \mathbb{R}^n$ — пространство решений приведенной системы $AX = 0$. Определим в этом пространстве скалярное произведение следующим образом: если $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$, то $xy = x_1y_1 + \ldots + x_ny_n$.

• ***Нормальным псевдорешением системы*** *системы* (14.12.1) *называется её псевдорешение, имеющее наименьшую длину.*

Теорема. *Нормальное псевдорешение системы* (14.12.1) *определяется однозначно и равно ортогональной составляющей произвольного псевдорешения системы* (14.12.1) *относительно подпространства L_0.*

♦ Покажем, что ортогональные составляющие псевдорешений системы (14.12.1) относительно L_0 равны. Пусть L_p — множество псевдорешений системы (14.12.1), тогда L_p — множество решений системы $A_1x_1 + \ldots + A_nx_n = B_1$ (2), где столбец B_1 — ортогнальная проекция столбца B на $L = L(A_1, \ldots, A_n) \Rightarrow L_p = \{x_0\} + L_0$, где x_0 — частное решение системы (2), которые является псевдорешением системы (14.12.1). Следовательно, произвольное псевдорешение системы (14.12.1) $x = x_0 + y$, где $y \in L_0$.

Обозначим через x_1 ортогональную проекцию, а через x_2 — ортогональную составляющую последовательности x на множестве $L_0 \subset \mathbb{R}^n$. Тогда x представимо в виде: $x = \underset{\in L_0}{x_1} + \underset{\in L_0}{x_2} \Rightarrow$ $x = x_1 + x_2 + y = \underbrace{(x_1 + y)}_{\in L_0} + \underset{\in L_0}{x_2} \Rightarrow x_2$ — ортогональная составляющая x относительно $L_0 \Rightarrow$ все псевдорешения имеют ортогональную составляющую равную x_2.

Покажем, что x_2 — нормальное псевдорешение системы (14.12.1). Так как $x_1 \in L_0 \subset \mathbb{R}^n$, то $-x_1 \in L_0$. Следовательно, среди псевдорешений системы (14.12.1) есть последовательность $x = x_0 + y = x_0 - x_1 = (x_1 + x_2) - x_1 = x_2 \Rightarrow x_2$ — псевдорешение системы (14.12.1). При этом любое другое псевдорешение системы (14.12.1) представимо в виде: $x' = x_0 + y'$, где $y' \neq -x_1$. Тогда $|x'|^2 = |x_0 + y'|^2 = |(x_1 + x_2) + y'|^2 = |(x_1 + y) + x_2'|^2 =$ [по теореме Пифагора] $= |x_1 + y'|^2 + |x_2|^2 > |x_2|^2 \Rightarrow x_2$ — нормальное псевдорешние системы (14.12.1), причем единственное. ⊠

Следствие. *Нормальное псевдорешение системы* (14.12.1) *является единственным псевдорешением, которое принадлежит пространству* $L_0^{\perp}$.

• *Матрица, столбцы которой являются нормальными псевдорешениями систем* $AX = E_i$, *где* E_i *— i-ый столбец единичной матрицы, называется* ***псевдообратной матрицей*** *для матрицы* A. (Обозначение: A^+)

Из определения следует, что если $A \in \mathbb{R}_{m,n}$, то $A^+ \in \mathbb{R}_{n,m}$ (Если A невырожденная, то $A^+ = A^{-1}$)

Теорема. *Нормальное псевдорешение системы* (14.12.1) *равно* A^+B.

♦ Пусть $B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \Rightarrow AX = \beta_1 E_1 + \ldots + \beta_m E_m$.

Обозначим через $C_1, \ldots, C_m$ столбцы псевдообратной матрицы A^+. Тогда C_i — нормальное псевдорешение системы $AX = E_i$. Следовательно, C_i — решение системы $A^TAX = A^TE_i \Rightarrow$ $A^TAC_i = A^TE_i \Rightarrow A^TA(\beta_1C_1 + \ldots \beta_mC_m) = A^{\perp}(\beta_1E_1 + \ldots + \beta_mE_m) \Rightarrow A^+B$ — псевдорешение системы (14.12.1). При этом, так как C_i — нормальное псевдорешение, то $C_i \in L_0^{\perp} \Rightarrow A^+B = \beta_1C_1 + \ldots \beta_mC_m \in L_0^{\perp} \Rightarrow A^+B$ — нормальное псевдорешение системы (14.12.1). ⊠