Алгебра

Дмитрий Абрамов, Антон Ермилов, Наташа Мурашкина, Дмитрий Саютин

По лекциям Афанасьевой Софьи

17 июня 2017 г.

Содержание

1. Векторное пространство, базис

1.1. Линейная алгебра

Замечание. K - поле

***Определение* 1.1.** Линейное уравнение: $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$

решение линейного уравнения - вектор

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

***Определение* 1.2.** Система линейных уравненений

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_n + \ldots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_n + \ldots + a_{2,n}x_n = b_2 \\ \ldots \\ a_{m,1}x_1 + a_{m,2}x_n + \ldots + a_{m,n}x_n = b_m \end{cases}$$

Матрица — возможное представление системы линейных уравнений

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Расширенная матрица системы уравнений — дописать к справа b

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} & b_m \end{pmatrix}$$

***Определение* 1.3.** Система совместна — имеет хотя бы одно решение. Несовместна — иначе.

***Определение* 1.4.** Системы эквивалентны, если множества их решений совпадают.

Преобразования системы, не меняющие мн-ва решений:

1. умножение строки (уравнения) на $x \in K \neq 0$
2. Прибавление к одной строке (уравнению) другой домноженной на любое $x \in K$
3. Перемена строк (уравнений) местами

Замечание. $\alpha \in K$

$$\alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

$$\begin{pmatrix} \alpha_{1,1} \\ \alpha_{2,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix} x_1 + \begin{pmatrix} \alpha_{1,2} \\ \alpha_{2,2} \\ \vdots \\ \alpha_{m,2} \end{pmatrix} x_2 + \ldots + \begin{pmatrix} \alpha_{1,n} \\ \alpha_{2,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Замечание. (K — поле)

***Определение* 1.5.** Векторное (линейное) пространство — мн-во V со следующими операциями:
Поэлементное сложение $+ : V \times V \to V$
Поэлементное умножение на скаляр $* : K \times V \to V$

Эти операции обладают следующими свойствами:

1. V — абелева группа, относительно $+$
2. $\lambda(u + v) = \lambda u + \lambda v$
3. $(\lambda + \mu)u = \lambda u + \mu u$
4. $(\lambda\mu)v = \lambda(\mu v)$
5. $1 \in K \quad 1 \cdot v = v, \forall v \in V$

Замечание. Договоренности:

- Элементы поля K обозначаем греческими буквами — $\alpha, \beta, \gamma \in K$
- Элементы V латинскими — $u, v, x, y \in V$
- Элементы векторного пространства — векторы
- Элементы K — скаляры

Примеры:

1. Пространство столбцов (длины n) K^n
   Упражнение: 1 — векторное пространство
2. Пространство строк (длины n)
3. Пространство матриц — $K^{n \times m}$ или $Mat_{n \times m}(K)$ таблица $n \times m$
4. (a) $\mathbb{C}$ (комплексные) — векторное пространство над полем $\mathbb{R}$
   (b) $\mathbb{R}$ — векторное пространство над полем $\mathbb{Q}$

(c) K, L — поля, и $K \subset L$ L — векторное пространство над полем K

5. Многочлены с коэффициентами из поля K
6. K^∞ — строки бесконечной длины $(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2 \ldots)$ $\alpha(a_1, a_2, \ldots) = (\alpha a_1, \alpha a_2, \ldots)$
7. Пространство непрерывных функций $[0, 1] \to R$

Замечание. Пусть V — векторное пространство над полем K

***Определение* 1.6.** Подмножество $U \subseteq V$ — векторное подпространство, если U — векторное пространство относительно тех же операций, что и на V

Лемма. $U \subseteq V$ Тогда U — подпространство $\Leftrightarrow$

1. $\forall u \in U, \alpha \in K: \ \alpha u \in U$
2. $\forall u, v \in U: \ u + v =\in U$

***Определение* 1.7.** Пусть $u_1 \ldots u_n$ — векторы.
Линейная комбинация с коэффециентами $\alpha_1 \ldots \alpha_n \in K$ — это $\sum\limits_{i=1}^{n} \alpha_i u_i$

***Определение* 1.8.** .

1. Подпространство порожденное подмножеством (линейная оболочка множества X, $X \subset V$) $\langle X \rangle = \bigcap\limits_{X \subset U} U$ наименьшее подпространство, содержащее X
2. X — система образующих пространства V, если его линейная оболочка совпадает с V ($\langle X \rangle = V$)
3. Пространство V — конечномерное, если $\exists X : |X| < \infty, \ \langle X \rangle = V$

Лемма. $\langle X \rangle = \{\sum\limits_{i_1}^{n} \alpha_i x_i \mid \alpha_i \in K, x_i \in X\}$ — множество линейных комбинаций элементов X

Доказательство. Докажем $\subset$:
Правая часть $\supset X$, правая часть — векторное подпространство V
$\alpha \in K$ $\alpha \sum\limits_{i_1}^{n} \alpha_i x_i = \sum\limits_{i_1}^{n} \alpha\alpha_i x_i$
$\sum\limits_{i_1}^{n} \alpha_i x_i + \sum\limits_{j_1}^{m} \alpha_i y_j, \ x_i, y_i \in X$ — тоже линейная комбинация

Докажем $\supset$
очевидно, что если $x \in X$ то все линейные комбинации $\in \langle X \rangle$

□

Лемма. если $v \in \langle X \rangle$, то $\langle X \rangle = \langle X \cup \{v\} \rangle$

Доказательство. $\subset$
$A \subset B \subset V \Rightarrow \langle A \rangle \subseteq \langle B \rangle$
$\supset$
По лемме: $\langle X \cup \{v\} \rangle$, $v = \sum\limits_{j_1}^{m} \beta_j x_j$ $\sum\limits_{i_1}^{n} \alpha_i x_i + \alpha v = \sum\limits_{i_1}^{n} \alpha_i x_i + \sum\limits_{j_1}^{m} \alpha\beta_j x_j$ □

1. $\mathbb{C}$ над $\mathbb{R}$
   $a + bi = a \cdot 1 + bi = \alpha(1 - i) + \beta(1 + i)$
2. K^n

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \end{pmatrix} = x_i \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

***Определение* 1.9.** Векторы $u_1 \ldots u_n$ — линейно зависимы, если $\exists \alpha_1 \ldots \alpha_n \in K, \exists \alpha_i \neq 0$, такие, что $\sum\limits_{i_1}^{n} \alpha_i x_i = 0,\ 0 \in V$ 0 — это вектор!

Иначе, линейно независимы. ($\forall \alpha_1 \ldots \alpha_n \in K$ что $\sum\limits_{i_1}^{n} \alpha_i x_i = 0 \Rightarrow \forall \alpha_i :\ \alpha_i = 0$])

***Определение* 1.10.** Базис — линейно независимая система образующих.

Базис точки (0) — $\varnothing$

Пусть $u_1 \ldots u_n \in V$

Теорема 1.1. Следующие условия попарно эквивалентны:

1. $u_1 \ldots u_n$ — базис
2. $u_1 \ldots u_n$ — максимальная (по включению) линейно независимая система векторов
3. $u_1 \ldots u_n$ — минимальная (по включению) система образующих
4. $\forall v \in V :\ \exists(!) \alpha_1 \ldots \alpha_n \in K,\ v = \sum\limits_{i=1}^{n} \alpha_i u_i$

Доказательство.

Теорема 1.2 $(1 \Rightarrow 2)$**.**

Доказательство. $u_1 \ldots u_n$ — Базис $\Rightarrow$ (по определению) это линейно независимая система
Пусть $U \supset \{u_1 \ldots u_n\}$, пусть $v \in U$ $\{u_1 \ldots u_n\}$. Т.к. $u_1 \ldots u_n$ - система образующих, то $\exists \alpha_1 \ldots \alpha_n \in K,\ v = \sum\limits_{i=1}^{n} \alpha_i u_i \Rightarrow$
$\Rightarrow v - \sum\limits_{i=1}^{n} \alpha_i u_i = 0$ но это линейная комбинация v и $\{u_1 \ldots u_n\} \Rightarrow U$ — линейно зависима. $\square$

Теорема 1.3 $(2 \Rightarrow 3)$**.**

Доказательство. а) покажем, что $\{u_1 \ldots u_n\}$ — система образующих
$v \in V$ либо $v = u_i$ тогда v является линейной комбинацией
Либо $v \notin \{u_1 \ldots u_n\}$
Рассмотрим $\{u_1 \ldots u_n\} \cup \{v\}$, т.к. $\{u_1 \ldots u_n\}$ — максимальная линейно независимая система $\Rightarrow$ $\{u_1 \ldots u_n\} \cup \{v\}$ — линейно зависимая $\Rightarrow \exists \alpha_1 \ldots \alpha_n \in K, \exists \alpha_i \neq 0, \sum\limits_{i=1}^{n} \alpha_i u_i + \alpha v = 0$
Если $\alpha = 0 \Rightarrow \sum\limits_{i=1}^{n} \alpha_i u_i = 0$ противоречие — $\{u_1 \ldots u_n\}$ линейно независима

Значит $\alpha \neq 0 \Rightarrow v = -\sum\limits_{i=1}^{n} \frac{\alpha_i}{\alpha} u_i$

б) минимальная
$Z \subset \{u_1 \dots u_n\}$ докажем, что $\langle Z \rangle \neq V$
Пусть $u_n \neq z$, пусть $\langle Z \rangle = V$, то $u_n \in Z$, $u_n - \sum\limits_{i=1}^{n-1} \alpha_i u_i = 0$, линейно зависимая, противоречие □

Теорема 1.4 $(3 \Rightarrow 4)$**.**

Доказательство. $\{u_1 \dots u_n\}$ — система образующих $\Rightarrow \forall v \in V \exists \alpha_1 \dots \alpha_n,\ v = \sum\limits_{i=1}^{n} \alpha_i u_i$

От противного: Пусть также $\exists \beta_1 \dots \beta_n :\ v = \sum\limits_{i=1}^{n} \beta_i u_i$

Рассмотрим $\sum\limits_{i=1}^{n} (\alpha_i - \beta_i) u_i = 0$. Пусть $\alpha_n \neq \beta_n$, тогда $u_n = -\sum\limits_{i=1}^{n-1} \frac{\alpha_i - \beta_i}{\alpha_n - \beta_n} u_i$. $u_n \in \langle u_1 \dots u_{n-1} \rangle$
По лемме выше $\langle u_1 \dots u_{n-1} \rangle = \langle u_1 \dots u_n \rangle = V$ — противоречие с минимальностью. □

Теорема 1.5 $(4 \Rightarrow 1)$**.**

Доказательство. а) $\{u_1 \dots u_n\}$ — система образующих — из 4 очевидно
б) минимальная независимая
$0 \in V\ \exists (!) \alpha_1 \dots \alpha_n,\ \sum\limits_{i=1}^{n-1} \alpha_i = 0,\ \alpha_i = 0 \forall i$. То есть, если $\sum\limits_{i=1}^{n} \beta_i u_i = 0 \Rightarrow \beta_i = 0 \forall i$ — линейная независимость. □

□

Теорема 1.6 (О существовании базиса)**.** $X \subseteq Y \subseteq V$. X — линейно независимо, Y — система образующих , тогда ∃ базис, содержащий X и содежащийся в Y.

Доказательство. $A = \{B \mid B-$ линейно независимо, $X \subseteq B \subseteq Y\}$

1. $A \neq \varnothing$ т.к. $X \in A$
2. $\subseteq$ — частичный порядок на множестве A
3. Пусть $Z \subset A$ и Z линейно упорядочено (по отношению $\subseteq$)
   возьмем $\bigcup\limits_{B \in Z} B \supseteq B,\ \forall B \in Z$ — верхняя грань.
   а) $X \subseteq \bigcup\limits_{B \in Z} \subseteq Y$
   б) Докажем от противного, что $\bigcup\limits_{B \in Z}$ линейно независимо
   Пусть $\sum\limits_{j=1}^{m} \alpha_j b_j = 0\ \forall j \exists B_j \in Z,\ b_j \in B \Rightarrow$ Z — минимальное упорядоченное. $\Rightarrow \exists N :$ $b_1 \dots b_m \in B_N \Rightarrow$ Но B_N линейно независимо $\Rightarrow \alpha_1 = \dots = \alpha_m = 0$

   Выполняются условия леммы Цорна $\Rightarrow$ в A ∃ максимальный элемент B_{max}, $X \subseteq B_{max} \subseteq Y$
   Докажем, что B_{max} — система образующих.

   $v \in Y \notin B_{max} \Rightarrow B_{max} \cup \{v\}$ — линейно зависима (так как не лежит в A).
   Значит ∃ нетривиальная $\sum\limits_{i_1}^{n} \alpha_i b_i + \alpha v = 0$

$b_i \in B_{max}$, $\alpha \neq 0$, $v = \sum\limits_{i_1}^{n} \frac{\alpha_i}{\alpha} b_i \in \langle B_{max} \rangle$

$Y \subset \langle B_{max} \rangle \Rightarrow \langle Y \rangle \subset \langle B_m \rangle \Rightarrow V = \langle B_{max} \rangle$ □

2. Равномощность базисов, линейные отображения

Лемма. (О замене)

B – базис векторного пространства V над полем K. (То есть $\langle B\rangle = V$ и B линейно независимо.)

Пусть вектор $u \in B$, вектор $v \in V$, но $v \notin \langle B \setminus \{u\}\rangle$.

Тогда $B \setminus \{u\} \cup \{v\}$ – базис V.

Доказательство. Докажем по определению два пункта.

1. Множество векторов $B \setminus \{u\} \cup \{v\}$ является системой образующих пространства V. Чтобы доказать это, достаточно показать, что можно выразить вектор u через вектора $B \setminus \{u\} \cup \{v\}$.

   $v = \sum \alpha_i b_i + \alpha u$, $b_i \in B \setminus \{u\}$, $\alpha, \alpha_i \in K$.

   $\alpha \neq 0$, так как $v \notin \langle B \setminus \{u\}\rangle$.

   $u = \frac{1}{\alpha} v - \sum \frac{\alpha_i}{\alpha} b_i \in B \setminus \{u\} \cup \{v\}$.

   $V = \langle B\rangle \subset B \setminus \langle \{u\} \cup \{v\}\rangle$.

2. $B \setminus \{u\} \cup \{v\}$ линейно независимо.

   $\beta v + \sum \beta_i v_i = 0$, $v_i \in B \setminus \{u\}$

   Подставим значение v: $\beta \sum \alpha_i b_i + \beta\alpha u + \sum \beta_i v_i = 0$

   Коэффициент $\beta\alpha$ перед u должен равняться нулю, так как в двух суммах нет векторов, которые могли бы сократиться с u. Но $\alpha \neq 0 \Rightarrow \beta = 0$. $\Rightarrow \sum \beta_i v_i = 0 \Rightarrow \beta_i = 0 \ \forall i$. Значит, действительно, для получения нуля подходит только тривиальная комбинация. □

Теорема 2.1. Любые два базиса равномощны.

Доказательство. Пусть V – конечномерное пространство. B, C – базисы пространства V, $|B| > |C| = n$.

B – минимальная система образующих $\Rightarrow B \setminus \{b_1\}$ – не является системой образующих.

Если $C \subset \langle B \setminus \{b_1\}\rangle$, значит $V = \langle C\rangle \subset \langle B \setminus \{b_1\}\rangle \neq V$. Противоречие.

Поясним. $C \subset \langle B \setminus \{b_1\}\rangle$, значит в $\langle B \setminus \{b_1\}\rangle$ также лежат все комбинации векторов из C, C – базис $\Rightarrow \langle B \setminus \{b_1\}\rangle = V$, то есть $B \setminus \{b_1\}$ – базис, но тогда B не базис (так как не минимально). Противоречие.

Значит, $C \not\subset \langle B \setminus \{b_1\}\rangle \Rightarrow \exists c_1 \notin \langle B \setminus \{b_1\}\rangle \Rightarrow$ по лемме о замене $B \setminus \{b_1\} \cup \{c_1\}$ – базис. $|B \setminus \{b_1\} \cup \{c_1\}| = |B| > |C|$.

Продолжим выкидывать и добавлять векторы: $B \setminus \{b_1\} \cup \{c_1\} \setminus \{b_2\}$ и т.д.

$B_1 = B \setminus \{b_1\} \cup \{c_1\} ... \setminus \{b_n\} \cup \{c_n\}$. $|B_1| = |B|$, $C \subset B_1$. Однако раз C – базис и B_1 – базис, то в B_1 не должны содержаться никакие другие лишние вектора, помимо векторов базиса C (так как базис – минимальная система образующих), которые там обязательно присутствуют по построению. Значит, $|B_1| = |C| \Rightarrow |B| = |C|$.

Замечание. Векторы c_i не совпадают, поэтому после каждой итерации мощность B не меняется.

□

Замечание. Теорема верна и в бесконечномерном случае (без доказательства).

***Определение* 2.1.** $\dim_K V$ – размерность пространства V над полем K, то есть мощность базиса. Иногда индекс K опускается.

Пример. 1. K^n – пространство столбцов длины n. $\dim K^n = n$.

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

Назовём столбцы $\{e_1, ..., e_n\}$ стандартным базисом K^n (В столбце e_i на i-том месте стоит 1, на остальных местах — 0). Это действительно базис:

$$\sum_{i=1}^{n} \alpha_i e_i = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

2. ${}^nK$ – пространство строчек длины n, $\dim {}^nK = n$.
3. Пространство $\mathbb{C}$ над полем $\mathbb{R}$. $\dim_{\mathbb{R}} \mathbb{C} = 2$. Пример базиса: $\{1, i\}$.
4. Пространство $\mathbb{R}$ над полем $\mathbb{Q}$. $\dim_{\mathbb{Q}} \mathbb{R} = \infty$.
5. $K[X]$. Базис: $1, x, x^2, x^3$ и т.д. $\dim K[X] = \infty$
6. $C[0,1]$ – функции, непрерывные на отрезке $[0,1]$. $\dim C[0,1] = \infty$.

***Определение* 2.2.** Линейное отображение, или линейный оператор, или гомоморфизм пространств. V, U – векторные поля над K.

$\varphi : V \to U$ называется линейным отображением, если выполняются следующие два условия:

1. $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)\ \forall v_1, v_2 \in V$.
2. $\varphi(\alpha v) = \alpha\varphi(v)\ \forall v \in V, \alpha \in K$.

Свойства.

1. $\varphi(0) = 0$.
2. $\varphi(-v) = -\varphi(v)$.
3. $\varphi(u - v) = \varphi(u) - \varphi(v)$.

Замечание. Линейное отображение аналогично гомоморфизму колец, доказательство свойств аналогично.

Пример. Линейные отображения.

1. Поворот плоскости на угол α, $R^\alpha : \mathbb{R}^2 \to \mathbb{R}^2$.

2. Проекция точки на ось OX, $f_x : \mathbb{R}^2 \to \mathbb{R}$, $(a, b) \mapsto a$.

3. Производная, $D : K[x] \to K[x]$.

***Определение* 2.3.** $\varphi, \psi : V \to V$ – эндоморфизмы (линейные отображения из себя в себя).

Замечание. Не путать эндоморфизм с автоморфизмом. Авторморфизм – это биективный эндоморфизм.

Свойства.

1. $\varphi \circ \psi$ – линейное отображение.

2. $(\varphi + \psi)(v) = \varphi(v) + \psi(v)$.

3. $(\alpha\psi)(v) = \alpha\psi(v), \alpha \in K$.

***Определение* 2.4.** V – векторное пространство над K, V – кольцо относительно той же операции сложения.

$\alpha(ab) = (\alpha a)b \ \ \forall \alpha \in K, a, b \in V$.

Тогда V – алгебра над K.

***Определение* 2.5.** $\operatorname{End} V$ – алгебра эндоморфизмов пространства V.

***Определение* 2.6.** V – векторное пространство над K, $U \leqslant V$.

Фактор-пространство $V/_U$ – векторное пространство, множество элементов фактор-пространства $V/_U$ совпадает с множеством элементов фактор-группы $V/_U$ абелевых групп $(V, +)$ и $(U, +)$.

Элементы фактор-пространства имеют вид $\{x + U\}$. Сложение уже определено в фактор-группе $((a + U) + (b + U) = (a + b) + U)$, определим умножение: $\alpha(x + U) = \alpha x + U, \alpha \in K, x \in V$.

***Определение* 2.7.** $\varphi : U \to V$ – линейное отображение.

$\operatorname{Ker} \varphi = \{u \in U \mid \varphi(u) = 0\}$

$\operatorname{Im} \varphi = \{v \in V \mid \varphi(u) = v, u \in U\}$

(Определения такие же, как в теориях групп и колец.)

Упражнение: Найти ядро и образ в отображениях из примеров.

1. Поворот плоскости на угол α, $R^\alpha : \mathbb{R}^2 \to \mathbb{R}^2$.
   $\operatorname{Ker} R^\alpha = \{0\}$, $\operatorname{Im} R^\alpha = \mathbb{R}^2$.

2. Проекция точки на ось OX, $f_x : \mathbb{R}^2 \to \mathbb{R}$, $(a, b) \mapsto a$.
   $\operatorname{Ker} f_x = \{(0, b)\}$, $\operatorname{Im} f_x = \mathbb{R}$ – вся прямая OX.

3. Производная, $D : K[x] \to K[x]$.
   $\operatorname{Ker} D = K$ – все элементы поля K, то есть константы. $\operatorname{Im} D = K[X]$.

Замечание. $\varphi : U \to V$ – линейное отображение.

1. φ – инъективно $\Leftrightarrow \operatorname{Ker} \varphi = \{0\}$.

2. $\forall b \in V$ множество решений уравнения $\varphi(x) = b$, то есть $\varphi^{-1}(b)$, имеет вид $a + \operatorname{Ker} \varphi$, где $\varphi(a) = b$.

Теорема 2.2. (О гомоморфизме)

U, V – векторное пространство, $\varphi : U \to V$ – линейное отображение.

Тогда $U/_{\operatorname{Ker}\varphi} \cong \operatorname{Im}\varphi$.

Замечание. Пусть $U, W \leqslant V$. Тогда

$U \cap W \leqslant V$.

$U + W = \{u + w \mid u \in U, w \in W\} \leqslant V$, причём иногда $\neq V$.

(Например, $\mathbb{R} + \mathbb{R} \leqslant \mathbb{R}^3$, $\{(a, 0)\} + \{(0, b)\}$ порождает плоскость, подпространство в R^3.)

$\{U \cup W\} = U + W$.

Замечание. Наши следующие вопросы:

$\dim V/_U$ — ?

$\dim(U + W)$ — ?

3. Размерности, суммы, кольцо матриц

3.1. Леммы о размерностях

Замечание. Пусть U и V — векторные пространство, и $\varphi\colon U \to V$ — линейное отображение между ними.

$\varphi^{-1}(b) = \{\varphi(x) = b\} = \operatorname{Ker}\varphi + a$, где a — произвольный прообраз.

$\operatorname{Ker}\varphi + a$ является аффинным подпространством.

***Определение* 3.1.** Непустое множество V называется *аффинным подпространством*, если для любых $v, u \in V$ и чисел $\alpha, \beta \in K$, таких что $\alpha + \beta = 1$, верно $\alpha v + \beta u \in V$

Замечание. В определении выше можно рассматривать любую конечную сумму $\sum \alpha_i v_i$, для $\sum \alpha_i = 1$. Несложно показать, что это одно и то же.

Замечание. Определение аффинного подпространство весьма схоже с определением подпространства. Если убрать ограничение $\alpha + \beta = 1$, то получится в точности необходимое и достаточное условие на подпространство.

Замечание. Афинное подпространство **не** является подпространством в общем случае

Теорема 3.1. Пусть U, V — векторные пространства над K, $\varphi\colon U \to V$ — линейное отображение. Тогда $\operatorname{Im}\varphi \cong U/_{\operatorname{Ker}\varphi}$

Доказательство. Построим отображение слева направо: $f\colon \operatorname{Im}\varphi \to U/_{\operatorname{Ker}\varphi}$

$f(u) := \varphi^{-1}(u) = v + \operatorname{Ker}\varphi$, где $\varphi(v) = u$.

Несложно увидеть, что отображение получилось линейным, сюръекция очевидна, ядро $\operatorname{Ker} f$ тривиально (только $\operatorname{Ker}\varphi$ отображается в $0 + \operatorname{Ker}\varphi$). □

Лемма. Пусть U — подпространство V, U и V имеют конечную размерность.

Тогда $\dim V/_U = \dim V - \dim U$

Доказательство. Пусть $\dim V = n$, $\dim U = m$.

Пусть $u_1, \ldots, u_m$ — базис U, дополним его до базиса V, пусть $u_1, \ldots, u_m, u_{m+1}, \ldots, u_n$ — базис V.

Пусть для $x \in U$, $\overline{x} := x + U$ – класс элемента x в факторе $V/_U$.

Покажем, что $\overline{u_{m+1}}, \overline{u_{m+2}}, \ldots, \overline{u_n}$ — базис $V/_U$.

Предположим, что этот набор является линейно независимым: $\sum_{i=m+1}^{n} \alpha_i \overline{u_i} = 0$

Тогда $\sum_{i=m+1}^{n} \alpha_i u_i = t \in U$, потому что в факторе это выражение равно 0.

Выразим t в базисе U: $t = \sum_{i=1}^{m} \beta_i u_i$

Итого имеем $0 = t - t = \sum_{i=1}^{m} \beta_i u_i - \sum_{i=m+1}^{n} \alpha_i u_i$

Сумма нетривиальна, если хотя бы одно α_i было не равно 0.

Противоречие с линейной независимостью $u_1, \ldots, u_n$.

Докажем, что любой $\overline{v} \in V/_U$ выражается.

Разложим v в V: $v = \sum_{i=1}^{n} \alpha_i u_i$

Разложение в $V/_U$ отличается не сильно: $\overline{v} = \sum_{i=1}^{n} \alpha_i \overline{u_i} = \sum_{i=m+1}^{n} \alpha_i \overline{u_i}$.

Пользуемся тем, что фактор — линейное отображение, выкидываем младшие $\overline{u_i}$, так как они равны 0 в факторе. □

Теорема 3.2. $\dim U = \dim \operatorname{Ker} \varphi + \dim \operatorname{Im} \varphi$.

Доказательство. Применим лемму о размерности фактора и аналог теоремы о гомоморфизме. □

***Определение* 3.2.** Определим ранг линейного отображения как размерность его образа.

$\operatorname{rank} \varphi := \dim \operatorname{Im} \varphi$

3.2. Суммы подпространств, прямые и не очень

***Определение* 3.3.** Пусть $U, W \leqslant V$. Тогда *суммой подпространств* (или *суммой Минковского*) называется множество всех попарных сумм;

$U + W := \{u + w \mid u \in U, w \in W\}$

Упражнение. $U + W \leqslant V$, $U \cap W \leqslant V$

Замечание. Введённое определение, а также выводы из него соответствуют аналогичному определению в теории идеалов. Там тоже множество всех попарных сумм идеалом являлось и пересечение идеалом являлось, а объединение — нет.

Замечание. Сумма подпространств является коммутативной операцией: $V + U = U + V$

Лемма. Пусть $U, W, U_1, \ldots, U_n \leqslant V$, тогда:

1. $\langle \cup U_i \rangle = U_1 + U_2 + \ldots + U_n$
2. $U + W = U \iff W \leqslant U$.

Доказательство.

1. Более-менее аналогично предыдущим соответствующим теоремам. Можно доказать, показав два включения.
2. Пусть $W \leqslant U$, тогда $U + W \subseteq U$, при чём равенство достигается, например второе слагаемое можно сделать 0.

   Пусть $U + W = U$, тогда $\forall w \in W \; w = 0 + w \in U + W = U$. □

***Определение* 3.4.** Подпространство V является *внутренней прямой суммой* подпространств $V_1, \ldots, V_m$ ($V = \oplus_i V_i$), если любой вектор $v \in V$ раскладывается единственным образом в сумму $v = v_1 + \ldots + v_m$, где $v_i \in V_i$.

***Определение* 3.5.** *Внешней прямой суммой* пространств $V_1, \ldots, V_m$ ($V = \oplus_i V_i$) называется множество кортежей $(v_1, \ldots, v_m)$ с операциями $+$ и $\cdot$ определёнными покомпонентно.

Замечание. Хотя заданные операции отличаются по смыслу (например первое больше похоже на прямую сумму в теории групп, а второе на прямую сумму в теории колец (прямое произведение в группах), обозначим их одинаково.

Лемма. Следующие условия эквивалентны:

1. Сумма подпространств $U_1 + \ldots + U_n$ является прямой

2. $0 = v_1 + \ldots + v_n \implies \forall i\colon v_i = 0$, где $v_i \in U_i$
3. $\forall i\colon U_i \cap (U_1 + \ldots + U_{i-1} + U_{i+1} + \ldots + U_n) = \{0\}$

Следствие. Сумма $U + V$ является прямой $\iff U \cap V = \{0\}$.

Доказательство. ($1 \implies 2$) В прямой сумме любой элемент представляется единственным образом.

В том числе и ноль, при чём одно из представлений нуля мы знаем, значит других нет.

($2 \implies 1$) Ноль представляется единственным образом, но пусть у x есть два представления:

$x = x_1 + \ldots + x_n = y_1 + \ldots + y_n$

Тогда $(x_1 - y_1) + \ldots + (x_n - y_n) = 0$, и значит каждое слагаемое равно нулю.

($1 \implies 3$) Пусть $W_i := U_1 + \ldots + U_{i-1} + U_{i+1} + \ldots + U_n$

Если $x \in U_i \cap W_i$ и $x \neq 0$, то x можно выразить двумя способами, через U_i и через сумму составляющих W_i.

Значит пересечение тривиально.

($3 \implies 2$) Пусть $0 = u_1 + \ldots + u_n$, где $u_i \in U_i$

Пусть не все u_i равны 0, тогда $\exists k\colon u_k \neq 0$

Тогда $u_k = -\sum\limits_{i \neq k} u_i$

Тогда $u_k \in U_k \cap W_k$, при чём $u_k \neq 0$, противоречие. □

В следующей теореме мы докажем, что внешняя прямая сумма, в некотором смысле, совпадает с внутренней.

В следующих утверждениях обозначим внутреннюю прямую сумму через $\oplus$, а внешнюю через $\oplus^e$.

Утверждение 3.3. 1. Пусть $V_1, V_2, \ldots, V_n$ — векторные пространства над K.

Тогда отображения $\mu_i\colon V_i \to V_1 \oplus^e \ldots \oplus^e V_n$ $(v_i \mapsto (0, \ldots, 0, v_i, 0, \ldots, 0))$ являются мономорфизмами.

2. Если $V = V_1 \oplus^e \ldots \oplus^e V_n$, то прямая сумма $\mu_1(V_1) \oplus \ldots \oplus \mu_n(V_n)$ изоморфна внешней (то есть V).
3. Если $V = V_1 \oplus \ldots \oplus V_n$, тогда объединение базисов $V_1, \ldots, V_n$ является базисом V.
4. $\dim(V_1 \oplus \ldots \oplus V_n) = \dim V_1 + \ldots + \dim V_n$.

Замечание. После второго пункта внешняя прямая сумма это то же самое, что внутренняя сумма внутри пространства внешней, поэтому пункты 3 и 4 применимы как для внутренней суммы, так и для внешней.

Доказательство. 1. Очевидно (есть гомоморфизм, и он инъекция).

2. Рассмотрим отображение $f\colon V_1 \oplus^e \ldots \oplus^e V_n \to \mu_1(V_1) \oplus \ldots \oplus \mu_n(V_n)$

$f(v_1, \ldots, v_n) := \mu_1(v_1) + \ldots + \mu_n(v_n) = (v_1, \ldots, v_n)$

Явно видно, что отображение оказалось тождественным.

3. Пусть $\dim V_i = n_i$, и V_i имеет базис $v_{i,1}, \ldots, v_{i,n_i}$.

   Проверим линейную независимость:

   $\sum_i (\sum_{j=1}^{n_i} \alpha_{i,j} v_{i,j}) = 0$, но $(\sum_{j=1}^{n_i} \alpha_{i,j} v_{i,j}) \in V_i$.

   Так как V — это прямая сумма, то это эквивалентно $\sum_{j=1}^{n_i} \alpha_{i,j} v_{i,j} = 0$ (для всех i), но это базис подпространства, поэтому $\alpha_{i,j} = 0$ для всех i, j.

   Кроме того нужно заметить, что каждый элемент выражается, для этого достаточно разложить этот элемент в сумму элементов из V_i, а каждый из полученных элементов выразить через соответствующий базис.

4. Следует из предыдущего пункта (мы даже явно построили базис). □

Теорема 3.4 (Формула Грассмана)**.** Пусть $U, W \leqslant V$, тогда $\dim(U + W) = \dim U + \dim W - \dim U \cap W$.

Замечание. Равенство выше в некотором смысле напоминает формулу включений-исключений.

Доказательство. Рассмотрим линейное отображение $\varphi\colon U \oplus^e W \to V$ $((u, w) \mapsto u + w)$.

Заметим, что φ — линейное отображение (убедитесь сами).

$\operatorname{Im}\varphi = U + W$, $\operatorname{Ker}\varphi = \varphi^{-1}(0) = \{(u, -u) \mid u \in U, -u \in W\} = \{(u, -u) \mid u \in U \cap W\} \cong U \cap W$.

Изоморфизм почти очевиден: $(u, -u) \mapsto u$

$\dim U \oplus^e W = \dim \operatorname{Ker}\varphi + \dim \operatorname{Im}\varphi = \dim(U \cap W) + \dim(U + W)$

Но $\dim U \oplus^e W = \dim U + \dim W$, подставим выше и теорема доказана. □

Замечание. Обратите внимание, что мы пользуемся внешней прямой суммой, а не внутренней.

Так как U и W произвольные пространства, то их внутренней суммы возможно не существует (если они, например, нетривиально пересекаются).

3.3. Кольцо матриц

***Определение* 3.6.** Матрица размером $n \times m$ — это таблица из n строк и m столбцов, состоящая из элементов поля K.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,m} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,m} \end{pmatrix}$$

***Определение* 3.7.** Обозначается множество матриц как $K^{n \times m}$ или $\operatorname{Mat}_{n \times m}(K) = \operatorname{Mat}(n, m, K)$.

В последней записи если $n = m$, то второй параметр можно опустить.

Замечание. Обратите внимание на порядок: сначала идёт число строк, затем число столбцов.

***Определение* 3.8.** Матрицы одинаковых размеров можно складывать и домножать на скаляр:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,m} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,m} \end{pmatrix} + \begin{pmatrix} b_{1,1} & b_{1,2} & \ldots & b_{1,m} \\ b_{2,1} & b_{2,2} & \ldots & b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \ldots & b_{n,m} \end{pmatrix} := \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \ldots & a_{1,m} + b_{1,m} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \ldots & a_{2,m} + b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} + b_{n,1} & a_{n,2} + b_{n,2} & \ldots & a_{n,m} + b_{n,m} \end{pmatrix}$$

$$\alpha \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,m} \\ b_{2,1} & b_{2,2} & \dots & b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \dots & b_{n,m} \end{pmatrix} := \begin{pmatrix} \alpha b_{1,1} & \alpha b_{1,2} & \dots & \alpha b_{1,m} \\ \alpha b_{2,1} & \alpha b_{2,2} & \dots & \alpha b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha b_{n,1} & \alpha b_{n,2} & \dots & \alpha b_{n,m} \end{pmatrix}$$

Лемма. Множество матриц с размером $n \times m$ является векторным пространством размерности nm.

Проверка всех условий векторного пространства остаётся в качестве упражнения.

***Определение* 3.9.** *Стандартным* базисом назовём множество матриц, в которых ровно один ненулевой элемент — единица в некотором месте.

$A = \sum_{i,j} a_{i,j} e_{i,j}$

***Определение* 3.10.** Пусть $A \in K^{n \times m}$, $B \in K^{m \times k}$

$AB := C$, где $C \in K^{n \times k}$, а $c_{i,j} = \sum_{l=1}^{m} a_{i,l} b_{l,j}$

Замечание. Такое определение умножения матриц может показаться странным на первый взгляд.

На самом деле матрица A размером $n \times m$ задаёт линейное отображение из K^m в K^n:

$f(x) := Ax$.

Можно убедиться (расписав формулы), что матрице композиции отображений соответствует матрица, получающаяся именно таким умножением, иначе говоря $(BA)x = B(Ax)$ для любого вектора x.

Замечание. Если записать вторую матрицу в виде столбцов $B = (b_1 \mid b_2 \mid \dots \mid b_k)$, $(b_i \in K^m)$, то $AB = (Ab_1 \mid Ab_2 \mid \dots \mid Ab_k)$.

Замечание. Умножение матриц ассоциативно, так как операция композиции отображений является ассоциативной.

Более формально, скомбинируем два предыдущих замечания:

Мы знаем, что $(BA)x = B(Ax)$, где x — любой вектор.

Покажем, что $(BA)C = B(AC)$ для любой матрицы C.

Представим $C = (c_1 \mid \dots \mid c_k)$ в столбцовой записи и подставим:

$(BA)(c_1 \mid \dots \mid c_k) = ((BA)c_1 \mid \dots \mid (BA)c_k) = (B(Ac_1) \mid \dots \mid B(Ac_k)) = B(Ac_1 \mid Ac_2 \mid \dots \mid Ac_k) = B(AC)$

Лемма. $K^{n \times n}$ — кольцо с единицей.

Доказательство. 1. $E_4 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Подставьте в определение и убедитесь, аналогично для всех остальных n.

2. Ассоциативность доказана в замечании.

3. $A(B + C) = AB + AC$, так как матрица — линейное отображение.

4. $(A + B)C = AC + BC$, несложно убедиться из определения. □

Замечание. В этом кольце (для $n \geqslant 2$) есть нильпотенты (и тем более делители нуля):

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Замечание. Умножение матриц не коммутативно (композиция отображений не является коммутативной)

Более того, если матрицы не квадратные, то при существовании произведения AB произведение BA вообще может не существовать.

Лемма. Пусть V — векторное пространство над K, $e = (e_1, e_2, \ldots, e_n)$ — базис V.

Введём $\varphi_e \colon V \to K^n$, это отображение переводит элемент V в столбец его коэффициентов при разложении по базису.

Тогда φ_e — линейное отображение и вообще изоморфизм.

Доказательство. Линейность очевидна, сюръективность тоже (по набору коэффициентов можно построить нужный вектор), также ядро тривиально (так как e — базис, то только тривиальная линейная комбинация равна нулю). □

Следствие. Пусть V — конечномерное пространство, тогда $V \cong K^{\dim V}$.

В частности, все V с одинаковой конечным $\dim V$ изоморфны друг другу (так как они все изоморфны $K^{\dim V}$).

Замечание. Построенное отображение, вообще-то говоря, зависит от выбора базиса e.

Более того, выбирая разные базисы и отображения получатся разные (смотрите далее).

4. Матрица оператора, матрица перехода, решение систем линейных уравнений

4.1. Связь линейных операторов с матрицами

$V \simeq K^n, \;\; n = \dim V$

Уже знаем, что такое гомоморфизм в векторном пространстве.

$K^n \simeq U \to V \simeq K^m$

Соответственно, каждому гомоморфизму в векторном пространстве будет соответствовать гомоморфизм в пространстве столбцов.

Обозначение.

$e = (e_1, e_2, \ldots, e_n)$ — базис V.

$$\forall\, x \in V : x = \sum_{i=1}^{n} \alpha_i e_i = (e_1, e_2, \ldots, e_n) \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}$$

$$x^e = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n \end{pmatrix}; \quad x = ex^e$$

Лемма.

Линейное отображение однозначно определяется образом базисных векторов.

Переформулировка:

$\begin{cases} e = (e_1, \ldots, e_n) \text{ — базис пространства } U \\ f = (f_1, \ldots, f_n), \;\; f_i \in V \end{cases}$

Тогда $\exists$ единственное отображение $\phi : U \to V$, такое что $\phi(e_i) = f_i, \;\; 1 \leqslant i \leqslant n$

Доказательство.

Единственность.

Если $\phi : U \to V$ — линейное, то $\phi(u) = \phi\left(\sum_{i=1}^{n} \alpha_i e_i\right) = \sum_{i=1}^{n} \alpha_i \phi(e_i) = \sum_{i=1}^{n} \alpha_i f_i$.

Отображение $U \to V, \quad \sum_{i=1}^{n} \alpha_i e_i \mapsto \sum_{i=1}^{n} \alpha_i f_i$ является линейным. □

Лемма.

e — базис U.

$\phi : U \to V$ — линейное отображение.

ϕ — инъективно $\iff$ $\phi(e_1), \ldots, \phi(e_n)$ — линейно независимы.

ϕ — сюръективно $\iff$ $\phi(e_1), \ldots, \phi(e_n)$ — система образующих.

ϕ — биекция $\iff$ $\phi(e_1), \ldots, \phi(e_n)$ — базис.

Доказательство.

1) $\phi(e_1), \ldots, \phi(e_n)$ — линейно зависимы $\iff$ $\exists\, \alpha_1, \ldots, \alpha_n \in K$, не все равные 0.

$\sum_{i=1}^{n} \alpha_i \phi(e_i) = 0 \iff \phi\left(\sum_{i=1}^{n} \alpha_i e_i\right) = 0$ (в силу линейности ϕ).

Следовательно, ϕ не инъективно, поскольку ядро нетривиально.

$\phi(e_1), \ldots, \phi(e_n)$ — линейно независимы $\implies \left(\sum_{i=1}^{n} \alpha_i \phi(e_i) = 0 \iff \forall\, \alpha_i = 0\right)$.

Следовательно, ϕ — инъективно.

2) $\operatorname{Im}\phi = \{\phi(u) : u \in U\} = \{\phi\left(\sum_{i=1}^{n} \alpha_i e_i\right) : \alpha_i \in K\} = \{\sum_{i=1}^{n} \alpha_i \phi(e_i) : \alpha_i \in K\} = \langle \phi(e_1), \ldots, \phi(e_n)\rangle$

3) 1 + 2 □

Следствие.

Два конечномерных пространства изоморфны $\iff$ они имеют одинаковые размерности.

Обозначение.

$e = (e_1, \ldots, e_n)$ — базис U.

$f = (f_1, \ldots, f_n)$ — базис V.

$\phi : U \to V$ — линейное отображение.

ϕ однозначно определяется матрицей.

$\phi_e^f = \phi_{e \to f} = \left(\phi(e_1)^f, \ldots, \phi(e_n)^f\right)$

Обозначение.

$u_1, \ldots, u_n \in U, \quad \phi : U \to V$ — линейное отображение

$u = (u_1, \ldots, u_n)$

$\phi(u) = (\phi(u_1), \ldots, \phi(u_n))$

Утверждение 4.1.

1) $\exists! A \in Mat_{m\times n}(K) : \forall\, u \in U \quad (\phi(u))^f = Au^e$

2) $\forall\, A \in Mat_{m\times n}(K) \quad \exists! \phi : (\phi(u))^f = Au^e$

То есть, линейный оператор ϕ однозначно определяется некоторой матрицей.

Доказательство.

$(\phi(u))^f = (\phi(eu^e))^f = (\phi(e)u^e)^f = \phi_e^f \cdot u^e$ □

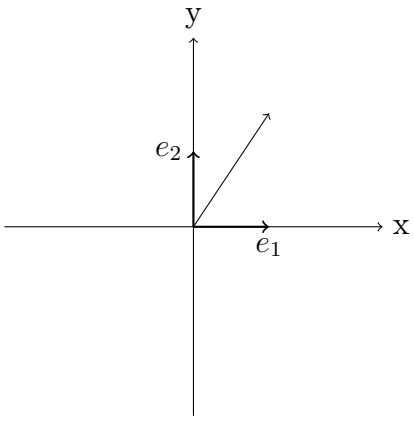
***Определение* 4.1.**

ϕ_e^f — матрица линейного отображения ϕ в базисах e и f.

$\phi_e := \phi_e^e$

Пример.

$\phi : \mathbb{R}^2 \to \mathbb{R}^2$ — поворот на угол α



$(\phi(e_1)^e,\ \phi(e_2)^e) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ — матрица линейного оператора.

***Определение* 4.2.**

$V = U \oplus W$

$e_1, \ldots, e_m$ — базис U.

$f_1, \ldots, f_k$ — базис W.

$e_1, \ldots, e_m, f_1, \ldots, f_k$ — базис V.

$\phi : V \to V, \quad (u, v) \mapsto (u, 0)$ — проекция.

$$\phi_{e,f} = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Заметим, что $\phi^2 = \phi$.

Пример.

$K[x]$ — кольцо многочленов.

$K_3[x] = \{f(x) \in K[x] : \deg f \leqslant 3\}$

$e = (1, x, x^2, x^3)$ — базис $K_3[x]$

$\phi = \frac{d}{dx}$

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

Утверждение 4.2.

1) U, V, W — конечномерные векторные пространства.

e, f, g — соответствующие базисы.

$\phi : U \to V,\ \psi : V \to W$ — линейные операторы.

$U \xrightarrow{\phi} V \xrightarrow{\psi} W, \quad \psi \circ \phi : U \to W$

Тогда $(\psi \circ \phi)^g_e = \psi^g_f \cdot \phi^f_e$

2) Если $U = V = W, \quad e = f = g$, то $(\psi \circ \phi)_e = \psi_e \cdot \phi_e$.

Доказательство.

$((\psi \circ \phi)^g_e)_j = (\psi \circ \phi(e_j))^g = \big(\psi(\phi(e_j))\big)^g = \psi^g_f \cdot (\phi(e_j))^f = \psi^g_f \cdot (\phi^f_e)_j = (\psi^g_f \cdot \phi^f_e)_j$ □

Теорема 4.3.

U, V — векторные пространства над K.

e — базис U.

$e = (e_1, \ldots, e_n), \quad n = \dim U$

$f = (f_1, \ldots, f_n)$ — базис V, $m = \dim V$

Тогда:

1) Имеется следующий изоморфизм векторных пространств:

$(U \to V) \to Mat_{m\times n}(K)$ (из множества линейных отображений в множество матриц)

$\phi \mapsto \phi_e^f$

2) Имеется изоморфизм алгебр

$\operatorname{End} U \to Mat_n(K)$

4.2. Замена базиса. Матрица перехода.

***Определение* 4.3.**

V — векторное пространство над K.

e, f — базисы V.

$f = f_1, \ldots, f_n$

$f_i = \sum_{j=1}^n C_i^j e_j$

$\begin{pmatrix} c_i^1 \\ \cdots \\ c_i^n \end{pmatrix} = f_i^e$

$(f_1, \ldots, f_n) = (e_1, \ldots, e_n)C$

$C_f^e = (f_1^e, f_2^e, \ldots, f_n^e)$ — матрица перехода от e к f.

Замечание. $C_e^e = E$

Утверждение 4.4.

1) $\phi : V \to V$

$\phi(e_i) = f_i$

$\phi_e^f = C_f^e$

2) $id : V_f \to V_e$

$(id)_f^e = C_f^e$

Лемма.

$C_e^f = (C_f^e)^{-1}$

Доказательство.

$C_e^f \cdot C_f^e = (id)_e^f \cdot (id)_f^e = (id)_f^f = C_f^f = E$ □

Утверждение 4.5.

$f = (f_1, \ldots, f_n)$ — линейно независимые $\iff \begin{pmatrix} \forall\, a, b \in K^m \\ fa = fb \iff a = b \end{pmatrix} \iff \begin{pmatrix} \forall\, A, B \in Mat_{m\times n}(K) \\ fA = fB \iff A = B \end{pmatrix}$

Утверждение 4.6.

$f = (f_1, \ldots, f_n), \quad e = (e_1, \ldots, e_n)$ — базисы V.

$v \in V$

$v^f = C_e^f v^e$

Доказательство.

$v = fv^f = ev^e = fC_e^f v^e$

$v^f = C_e^f v^e$ □

4.3. Изменение матрицы опеатора при замене базиса

Утверждение 4.7.

$\phi : U \to V$ — линейный оператор.

e, e' — базисы U.

f, f' — базисы V.

$\phi^{f'}_{e'} = C^{f'}_{f} \cdot \phi^{f}_{e} \cdot C^{e}_{e'}$

Доказательство.

$\phi^{f'}_{e'} = (id_v \circ \phi \circ id_u)^{f'}_{e'} = (id_v \circ \phi)^{f'}_{e} \cdot id^{e}_{e'} = (id)^{f'}_{f} \cdot \phi^{f}_{e} \cdot (id)^{e}_{e'} = C^{f'}_{f} \cdot \phi^{f}_{e} \cdot C^{e}_{e'}$ □

4.4. Решение системы линейных уравнений

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,n}x_n = b_1 \\ \vdots \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n = b_m \end{cases}$$

$$A = \begin{pmatrix} a_{1,1} & \ldots & \ldots \\ \ldots & \ddots & \ldots \\ \ldots & \ldots & a_{m,n} \end{pmatrix} \in Mat_{m\times n}(K)$$

$$b = \begin{pmatrix} b_1 \\ \ldots \\ b_m \end{pmatrix} \in K^m, \quad x = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} \in K^n$$

$$A \cdot \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} = b, \quad Ax = b$$

$\phi : K^n \to K^m, \quad x \mapsto Ax$

Решения $Ax = b$.

$\phi^{-1}(b) = x_0 + \operatorname{Ker}\phi$

$x_0 \in \phi^{-1}(b)$, x_0 — частное решение $Ax = b$.

$\operatorname{Ker}\phi$ — решения $Ax = 0$ (однородной системы).

$\dim \operatorname{Ker}\phi = n - \operatorname{rank}\phi$

***Определение* 4.4.**

Набор базисных векторов пространства $\operatorname{Ker}\phi$ — фундаментальная система решений однородной системы.

Замечание.

Легко решить:

$$\begin{pmatrix} a_{1,1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Заметим также, что систему легко решить, если выше главной диагонали стоят произвольные коэффициенты.

***Определение* 4.5.**

$(A|b)$ — расширенная матрица системы.

***Определение* 4.6.**

Элементарные преобразования:

1) Прибавление к i-ой строки j-ой, умножение на $\alpha \in K$.

2) Умножение i-ой строки на $\alpha \neq 0,\ \alpha \in K$.

3)Перестановка строк.

***Определение* 4.7.**

Элементарные матрицы — матрицы вида:

$E + \alpha e_i^j, \quad \alpha \in K,\ 1 \leqslant i = j \leqslant n$

***Определение* 4.8.**

Ступенчатая матрица — матрица, у которой все нулевые строки внизу, а номера первых ненулевых коэффициентов в очередной строке строго возрастают.

Теорема 4.8.

1) Всякую матрицу можно привести к ступенчатому виду с помощью элементарных преобразований.

2) $\forall\, A \in Mat_{m\times n}(K)\ \ \exists\, l \geqslant 0\ \ \exists\, B_1, \ldots, B_l \in Mat_m(K) : B_1 \cdot \ldots \cdot B_l \cdot A$ — ступенчатая матрица.

$B_1, \ldots, B_l$ — элементарные матрицы.

5. Ранг, полилинейные отображения, форма объёма, определитель

5.1. Транспонирование матриц

***Определение* 5.1.** Пусть $A \in K^{m\times n}$ и $(A)_ij = (a_{ij})$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$. Транспонированная матрица $A^T \in K^{n\times m}$ — это такая матрица, что $(A^T)_{ji} = a_{ij}$.

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \end{pmatrix}^T = \begin{pmatrix} a_{11} & \vdots \\ \vdots & \vdots \\ a_{1n} & \vdots \end{pmatrix} = \begin{pmatrix} a_{11} & \vdots \\ \vdots & \vdots \\ a_{1n} & \vdots \end{pmatrix}$$

***Определение* 5.2.** Обозначим группу обратимых матриц размера $n \times n$ над K как $\mathrm{GL}(n, K) = \mathrm{GL}_n(K) = (K^{n\times n})^*$. (GL от general linear group — полная линейная группа.)

Лемма. 1. $A, B \in K^{m\times n}$. Тогда $(A + B)^T = A^T + B^T$.

2. $A \in K^{m\times n}$, $B \in K^{n\times l}$. Тогда $(AB)^T = B^T A^T$.

3. $A \in K^{n\times n}$ и $A \in \mathrm{GL}_n(K)$. Тогда $A \in \mathrm{GL}_n(K)$ и $(A^{-1})^T = (A^T)^{-1}$.

***Определение* 5.3.** Пусть $A \in K^{m\times n}$.

1. Ранг системы векторов — размерность её линейной оболочки.
2. Пусть $a : U \to V$ — линейное отображение. Ранг этого отображения $\operatorname{rank} a = \dim \operatorname{Im} a$.
3. Ранг матрицы по строкам, $\operatorname{rank}_r A$ — ранг системы строк.
4. Ранг матрицы по столбцам, $\operatorname{rank}_c A$ — ранг системы столбцов.

Замечание. Понятно, что $\operatorname{rank}_r A = \operatorname{rank}_c A^T$.
Вскоре вообще поймём, что ранг матрицы по строкам совпадает с рангом матрицы по столбцам.

Замечание. Пусть $a : U \to V$ — линейное отображение, e и f — базисы U и V соответственно. Тогда $\operatorname{rank} a = \operatorname{rank}_c a_e^f$ — ранг по столбцам соответствующей матрицы перехода от базиса e к f.

Доказательство. Знаем, что $\operatorname{Im} a = \langle \varphi(e_1), \dots, \varphi(e_n) \rangle$. Тогда $\operatorname{rank} a = \dim \operatorname{Im} a = \operatorname{rank}_c a_e^f$. □

Лемма. **TODO:** Утверждения 1 и 2 совпадают. В пункте 2 нужно домножать с другой стороны. Вечером в конспекте Сони появится правильная формулировка и доказательство (дыра обнаружилась на консультации).

1. Пространство, порождённое строками (столбцами) матрицы, не изменяется при элементарных преобразованиях строк (столбцов) матрицы.
2. (a) Пусть $A \in K^{m\times n}$, $B \in \mathrm{GL}_n(K)$. Тогда $\operatorname{rank}_c(AB) = \operatorname{rank}_c A$.
   (b) Пусть $A \in K^{m\times n}$, $C \in \mathrm{GL}_m(K)$. Тогда $\operatorname{rank}_r(CA) = \operatorname{rank}_r A$.
3. $\operatorname{rank}_c A$ ($\operatorname{rank}_r A$) не изменяется при элементарных преобразованиях строк (столбцов).

Доказательство. 1. Нужно показать, что ни одно из трёх элементарных преобразований не изменяет пространство. Домножение строки на скаляр и перемена строк местами (1 и 3 элементарные преобразования) не меняют пространство. Распишем прибавление к одной строки другой, домноженной на скаляр (2 элементарное преобразование).

Пусть $A = \begin{pmatrix} A^1 \\ \vdots \\ A^n \end{pmatrix}$, где A^i — i-тая строка матрицы A.

Пусть $U := \langle A^1, \ldots, A^n \rangle$ и $V := \langle A^1, \ldots, A_i + \alpha A_j, \ldots, A^n \rangle, \alpha \in K$ — пространства, порождённые строками до и после преобразований соответственно.

Знаем, что $A^i + \alpha A^j \in U$. С другой стороны, $A_i = (A^i + \alpha A^j) - \alpha A^j \in V$ Значит $U = V$.

2. (a) Пусть $B \in \mathrm{GL}_n(K)$, $A \in K^{m \times n}$. Рассмотрим следующие отображения: $\underset{g}{K^n} \xrightarrow{id} \underset{e}{K^n} \xrightarrow{a} \underset{f}{K^m}$, где g — какой-то базис, e, f — стандартные базисы, а a — отображение такое, что $a : x \mapsto Ax$. Тогда $A = a_e^f$. Пусть базис $g := (B_1, \ldots, B_n)$, где B_i — столбцы матрицы B, то есть $B = id_g^e$.

   Распишем произведение матриц: $AB = a_e^f \cdot id_g^e = (a \circ id)_g^f = a_g^f$. Тогда $\mathrm{rank}_c(AB) = \mathrm{rank}_c\, a_g^f = \mathrm{rank}\, a = \mathrm{rank}_c\, A = \mathrm{rank}_r\, A$.

   (b) Распишем через транспонирование и воспользуемся пунктом a.

   $\mathrm{rank}_r\, CA = \mathrm{rank}_c (CA)^T = \mathrm{rank}_c (A^T C^T) = \mathrm{rank}_c\, A^T$.

3. Элементарные преобразования строк соответствуют домножению исходной матрицы A на какую-то матрицу ε элементарного преобразования (мы вывели такие матрицы ранее), то есть $A \rightsquigarrow \varepsilon A$. Матрица ε обратима, значит по пункту 3a $\mathrm{rank}_c(\varepsilon A) = \mathrm{rank}_c\, A$, то есть ранг по столбцам не изменяется.

   Элементарные преобразования столбцов соответствуют домножению транспонированной матрицы A^T на элементарные. Аналогично, по пункту 3b получаем требуемое. □

Следствие. Ранг матрицы по строкам (столбцам) равен рангу ступенчатой матрицы, к которой приводится данная.

Теорема 5.1. Пусть $A \in K^{m \times n}$. Тогда $\mathrm{rank}_c\, A = \mathrm{rank}_r\, A$.

Доказательство.

$$A \rightsquigarrow \begin{pmatrix} * & * & \ldots & * & * \\ 0 & * & \ldots & * & * \\ 0 & 0 & \ddots & * & * \\ 0 & 0 & \ldots & * & * \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} * & 0 & \ldots & 0 & 0 \\ 0 & * & \ldots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ldots & * & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} := \overline{A}.$$

Элементарными преобразованиями строк приведём матрицу к ступенчатому виду. Элементарными преобразованиями столбцов добьёмся того, чтобы ненулевые элементы стояли только на главной диагонали. Домножим ненулевые строки на обратные к единственному ненулевому элементу в строке, чтобы получить единицы на диагонали. Получим почти единичную матрицу, возможно с нулями в конце. Обозначим количество ненулевых строк (или столбцов, что то же самое) за l. Заключаем, что $\mathrm{rank}_c\, A = \mathrm{rank}_c\, \overline{A} = l$ и $\mathrm{rank}_r = \mathrm{rank}_r\, \overline{A} = l$. □

Замечание. 1. $\operatorname{rank} A = \operatorname{rank} A^T$.

2. Ранг матрицы $\operatorname{rank} A$ — количество ненулевых строк в её ступенчатом виде.

Теорема 5.2. Кронкера-Капелли

Пусть $A \in K^{m\times n}$, $b \in K^m$. Система линейных уравнений $Ax = b$ совместна (имеет решения) $\Leftrightarrow \operatorname{rank} A = \operatorname{rank}(A \mid b)$.

Другими словами, система линейных уравнений совместна тогда и только тогда, когда ранг матрицы её коэффициентов равен рангу расширенной матрицы.

Доказательство. Докажем стрелочку ($Ax = b$ совместна) $\Rightarrow$ ($\operatorname{rank} A = \operatorname{rank}(A \mid B)$). Пусть система совместна. Тогда b является линейной комбинацией столбцов матрицы A. Тогда ранг системы столбцов не изменится при добавлении b.

Пусть система несовместна. Тогда b не является линейной комбинацией столбцов матрицы A. Тогда при добавлении b к системе, её ранг изменится (увеличится).

Стрелочка в другую сторону доказывается аналогично, путём прочтения доказательства снизу вверх. □

Замечание. Пусть $A \in K^{n\times n}$. Тогда $A \in GL_n(k) \Leftrightarrow \operatorname{rank} A = n$.

Доказательство. Пусть a — соответствующее матрице отображение: $a : x \mapsto Ax$.

$\operatorname{rank} A = n \Rightarrow \operatorname{rank} a = \dim \operatorname{Im} a = n$. Тогда ядро a тривиально. Значит, a — инъекция. Но a также является сюръекцией по построению, значит a — биекция, и a обратимо. □

***Определение* 5.4.** $\mathrm{GL}(V)$ — группа обратимых линейных операторов, действующих $V \to V$. $\mathrm{GL}(V) = (\operatorname{End} V)^*$.

5.2. Определитель

***Определение* 5.5.** Данное определение вводит интуицию относительно объёма. Его нет в билетах. Аналог ориентированного объёма — форма объёма — введена ниже в конспекте с необходимой формальностью.

Пусть $(A_1, \ldots, A_n) \in \mathbb{R}^n$ — столбцы квадратной матрицы A. Комбинацию столбцов $\prod(A_1, \ldots, A_n) = \{x_1A_1 + \ldots + x_nA_n \mid x_i \in [0, 1]\}$ будем называть параллелепипедом, чьи рёбра — столбцы матрицы (натянули параллелепипед на столбцы матрицы).

Неформально введём ориентированный объём по индукции. $\mathrm{Vol}^{(n)}(A_1, \ldots, A_n) = \mathrm{Vol}^{(n-1)} \cdot h_{A_n}$, где h_{A_n} — условно «высота» вектора A_n (не вводим понятие, а даём интуицию).

Знак ориентированного объёма зависит от того, правильно ли направленны векторы относительно каких-то фиксированных направлений.

Ориентированный объём должен обладать следующими свойствами:

1. Антисимметричность: $V(A_1, A_2) = -V(A_2, A_1)$.
2. Линейность по первому аргументу: $V(A_1 + A_3, A_2) = V(A_1, A_2) + V(A_3, A_2)$; $V(\alpha A_1, A_2) = \alpha V(A_1, A_2)$. (Линейность по второму аргументу следует из 1).
3. Объём единичного гиперкуба (многомерного куба) равен единице: $V(\prod(E)) = 1$.

***Определение* 5.6.** Пусть V — векторное пространство над K. $f : \underbrace{V \times \ldots \times V}_{k \text{ раз}} \to K$ — полилинейное отображение с валентностью k (k-форма), если оно линейно по каждому из аргументов, то есть $\forall\ u, v \in V$, $\alpha \in K$ выполняется:

1. $f(\ldots, u + v, \ldots) = f(\ldots, u, \ldots) + f(\ldots, v, \ldots)$.
2. $f(\ldots, \alpha u, \ldots) = \alpha f(\ldots, u, \ldots)$.

Полилинейные формы с валентностью 2 называются билинейными.

Лемма. $e = (e_1, \ldots, e_n)$ — базис V. $f : \underbrace{V \times \ldots \times V}_{k \text{ раз}} \to K$ — полилинейное отображение. Пусть $A = (a_{ij}) = (v_1^e, \ldots, v_k^e)$. Тогда

$$f(v_1, \ldots, v_k) = \sum_{i_1, \ldots, i_k = 1}^{n} f(e_{i_1}, ..., e_{i_n}) \cdot a_{i_1 1} \cdot \ldots \cdot a_{i_k k}.$$

Доказательство. Распишем по определению линейности по каждому аргументу постепенно, получим требуемое.

$$f(v_1, \ldots, v_k) = f(\sum_{i_1=1}^{n} e_{i_1} a_{i_1 1}, \ldots, \sum) = \sum_{i_1=1}^{n} f(e_{i_1}, ..., \sum) \cdot a_{i_1 1} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} f(e_{i_1}, e_{i_2}, ..., \sum) \cdot a_{i_1 1} \cdot a_{i_2 2}.$$

□

Пример. 1. Скалярное произведение.

2. Пусть $V = K^n$, $f : K^n \times K^n \to K$.

$$f(\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}, \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}) = \sum_{1 \leqslant i,j \leqslant n} \alpha_{ij} x_i y_i.$$

***Определение* 5.7.** Пусть ω — полилинейная k-форма $\underbrace{V \times \ldots \times V}_{k \text{ раз}} \to K$. Форма ω симметрическая, если $\forall\ 1 \leqslant i, j \leqslant k$ выполняется: $w(\ldots, v_i, v_j, \ldots) = w(\ldots, v_j, v_i, \ldots)$.

Замечание. Обозначим $\mathrm{Multi}_k V$ — пространство k-форм, $\mathrm{SMulti}_k V$ — пространство симметрических k-форм.

Симметрические полилинейные отображения — подпространство в пространстве всех полилинейных отображений (той же валентности), то есть $\mathrm{SMulti}_k V \leqslant \mathrm{Multi}_k V$.

Определено действие группы перестановок на пространство форм: $S_k \times \mathrm{Multi}_k V \to \mathrm{Multi}_k V$, $(\sigma\omega)(v_1, \ldots, v_k) = \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$. Это действительно действие, то есть выполняется $(\sigma\tau)\omega = \sigma(\tau\omega)$ — слева сначала перемешали перемешку, а потом перемешали аргументы, справа дважды перемешали аргументы, равенство сохранится. Также выполняется $\mathrm{id}\,\omega = \omega$.

Заметим, что для пространства симметрических форм выполняется: $\mathrm{SMulti}_k V = \{\omega \in \mathrm{Multi}_k V \mid \forall\ \sigma \in S_k : (\sigma\omega)(v_1, \ldots, v_k) = \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\}$

5.3. Антисимметрические формы

***Определение* 5.8.** $w : \underbrace{V \times \ldots \times V}_{k \text{ раз}} \to K$ — полилинейная k-форма является антисимметрической, если $\forall\ u_i\ \omega(\ldots, u_i, \ldots, u_i, \ldots) = 0$.

$\mathrm{AMulti}_k V$ — пространство полилинейных антисимметрических k-форм.

Лемма. 1. $\forall\ \omega \in \mathrm{AMulti}_k V\ \omega(\ldots, a, \ldots, b, \ldots) = -\omega(\ldots, b, \ldots, a, \ldots)$.

2. Если $\mathrm{Char}\, K \neq 2$, то условие (1) равносильно антисимметричности.

3. $\forall\ \omega \in \mathrm{AMulti}_k V, \forall\ \sigma \in S_k$ условие (1) равносильно $\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \mathrm{sign}\,\sigma\omega(v_1, \ldots, v_k)$.

Доказательство. 1. $0 = \omega(\ldots, a+b, \ldots, a+b, \ldots) = \omega(\ldots, a, \ldots, a+b, \ldots) + \omega(\ldots, b, \ldots, a+b, \ldots) = \omega(\ldots, a, \ldots, a, \ldots) + \omega(\ldots, a, \ldots, b, \ldots) + \omega(\ldots, b, \ldots, a, \ldots) + \omega(\ldots, b, \ldots, b, \ldots) = \omega(\ldots, a, \ldots, b, \ldots) + \omega(\ldots, b, \ldots, a, \ldots)$.

2. $\omega(\ldots, a, \ldots, a, \ldots) = -\omega(\ldots, a, \ldots, a, \ldots)$, то есть $2\omega(\ldots, a, \ldots, a, \ldots) = 0$.

□

Лемма. Пусть $\omega \in \mathrm{AMulti}_k V$. Тогда $\omega(\ldots, v_i, \ldots, v_j, \ldots) = \omega(\ldots, v_i + \alpha v_j, \ldots, v_j, \ldots)$.

Доказательство. $= \omega(\ldots, v_i, \ldots, v_j, \ldots) + \alpha \underbrace{\omega(\ldots, v_j, \ldots, v_j, \ldots)}_{=0}$. □

5.4. Форма объёма

Пусть V — векторное пространство над K, $\dim V = n$, $e = (e_1, \ldots, e_n)$ — базис V. Пусть A — матрица, $A = (a)_{ij} = (v_1^e, \ldots, v_n^e)$.

***Определение* 5.9.** Форма объёма на V — антисимметрическая полилинейная n-форма.

Лемма. Пусть ω — форма объёма, $\omega\colon \underbrace{V \times \ldots \times V}_{n \text{ раз}} \to K$. Тогда выполняется:

$$\omega(v_1, \ldots, v_n) = \omega(e_1, \ldots, e_n) \sum_{\sigma \in S_n} \mathrm{sign}\,\sigma \cdot (v_1^e)^{\sigma(1)} \cdot \ldots \cdot (v_n^e)^{\sigma(n)} =$$
$$= \omega(e_1, \ldots, e_n) \sum_{\sigma \in S_n} \mathrm{sign}\,\sigma \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}.$$

Доказательство. По лемме о полилинейных отображениях:

$$\omega(v_1, \ldots, v_n) = \sum_{i_1, \ldots, i_n = 1}^{n} \omega(e_{i_1}, \ldots, e_{i_n}) \cdot a_{i_1 1} \cdot \ldots \cdot a_{i_n n} =$$
$$= \sum_{\sigma \in S_n} \omega(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} =$$
$$= \omega(e_1, \ldots, e_n) \sum_{\sigma \in S_n} \mathrm{sign}\,\sigma \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}.$$

Индексы $i_1, \ldots, i_n$ пробегают все значения от 1 до n (то есть если выписать их, получатся все строки длины n над алфавитом из n букв). Однако если два каких-то индекса совпадают

$i_p = i_q$, то ω от этого набора аргументов обнулится. Значит, останутся только те наборы индексов, которые соответствуют перестановкам. Но значение ω на перестановках равны по модулю, и знак зависит от чётности перестановки. Тогда вынесем значение ω на любой перестановке (например, тождественной), оставив под суммой только вычисление соответствующего знака перестановки. □

***Определение* 5.10.** Определитель матрицы A — это $\det A = \sum_{\sigma \in S_n} \operatorname{sign} \sigma \cdot a_{\sigma(1)1} \dots a_{\sigma(n)n}$.

Замечание. По определению, $\det E = 1$.

Замечание. Использовав определение определителя, предыдущую лемму можно записать в следующем виде: $\omega(v_1, \dots, v_n) = \omega(e_1, \dots, e_n) \det(v_1^e, \dots, v_n^e)$

6. Форма объёма, определитель и его свойства, минор

Лемма. Определитель матрицы — полилинейная антисимметричная форма её столбцов

Доказательство. Покажем линейность по i-ому столбцу.

Пусть $B = (b_1 \mid \ldots \mid b_n)$, причём $b_i = \alpha v + \beta u$, тогда:

$$\det(B) = \det(b_1, \ldots, b_{i-1}, \alpha v + \beta u, b_{i+1}, \ldots, b_n) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^{n} b_{\sigma(j),j} =$$

$$= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\ b_{\sigma(i),i} \prod_{j \neq i} b_{\sigma(j),j} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)(\alpha v_{\sigma(i)} + \beta u_{\sigma(i)}) \prod_{j \neq i} b_{\sigma(j),j} =$$

$$= \alpha \det(b_1, \ldots, b_{j-1}, v, b_{j+1}, \ldots, b_n) + \beta \det(b_1, \ldots, b_{j-1}, u, b_{j+1}, \ldots, b_n)$$

Покажем антисимметричность, иначе говоря, $b_k = b_l$ для некоторого $k \neq l$, покажем, что $\det B = 0$.

В S_n есть нормальная подгруппа $A_n \trianglelefteq S_n$ чётных перестановок, с индексом $|S_n : A_n| = 2$

Пусть $\tau = (k\, l)$ — транспозиция k и l ($\operatorname{sign}(\tau) = -1$), тогда $S_n = A_n \sqcup A_n\tau$

$$\det B = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} b_{\sigma(i),i} = \sum_{\sigma \in A_n} \prod_{i=1}^{n} b_{\sigma(i),i} - \sum_{\sigma \in A_n} \prod_{i=1}^{n} b_{\sigma(\tau(i)),i} =$$

$$= \sum_{\sigma \in A_n} b_{\sigma(l),l}\ b_{\sigma(k),k} \prod_{i \notin \{k,l\}} b_{\sigma(i),i} - \sum_{\sigma \in A_n} b_{\sigma(\tau(k)),k}\ b_{\sigma(\tau(l)),l} \prod_{i \notin \{k,l\}} b_{\sigma(\tau(i)),i} =$$

$$= \sum_{\sigma \in A_n} b_{\sigma(l),l}\ b_{\sigma(k),k} \prod_{i \notin \{k,l\}} b_{\sigma(i),i} - \sum_{\sigma \in A_n} b_{\sigma(l),k}\ b_{\sigma(k),l} \prod_{i \notin \{k,l\}} b_{\sigma(i),i} = 0$$

В последнем равенстве пользуемся тем, что $b_k = b_l$. □

***Определение* 6.1.** Пусть $e = (e_1, \ldots, e_n)$ — базис.

Тогда $\operatorname{Vol}^e(v_1, \ldots, v_n) := \det(v_1^e, \ldots, v_n^e)$

Замечание. Соответственно Vol^e для любого базиса e является формой объёма.

Теорема 6.1. Пусть V — векторное пространство над K размерности n, $e = (e_1, \ldots, e_n)$ — базис.

Тогда

1. $\operatorname{Vol}^e(e_1, \ldots, e_n) = 1$
2. $\forall$ формы объёма w, $w(v_1, \ldots, v_n) = w(e_1, \ldots, e_n) \operatorname{Vol}^e(v_1, \ldots, v_n)$
   Если $f = (f_1, \ldots, f_n)$ — базис, то $\operatorname{Vol}^f = \det C_e^f \operatorname{Vol}^e$, где C_e^f матрица перехода из e в f.
3. Пространство форм объёма одномерное.
4. Пусть w — ненулевая форма объёма, тогда $v_1, \ldots, v_n$ — базис $V \iff w(v_1, \ldots, v_n) \neq 0$.

5. Для $A \in K^{n\times n}$: $\det A \neq 0 \iff$ столбцы A линейно независимы.
6. A — обратима $\iff \det A \neq 0$
   $\mathrm{GL}(K^{n\times n}) = \{A \in K^{n\times n} \mid \det(A) \neq 0\}$

Доказательство. 1. $\mathrm{Vol}^e(e_1, \dots, e_n) = \det(e_1^e, \dots, e_n^e) = \det(E) = 1$
   Коэффициенты разложения базиса на самого себя образуют единичную матрицу.
2. Посмотрите на это замечание о формах объёма, и на определение Vol.

   Возьмём форму объёма Vol^f: $\mathrm{Vol}^f(v_1, \dots, v_n) = \mathrm{Vol}^f(e_1, \dots, e_n) * \mathrm{Vol}^e(v_1, \dots, v_n)$
   $\mathrm{Vol}^f(e_1, \dots, e_n) = \det\left(e_1^f, \dots, e_n^f\right) = \det\left(C_e^f\right)$
3. Рассмотрим произвольный базис $e_1, \dots, e_n$. Любая форма объёма по предущему пункту однозначно определеяется константой $w(e_1, \dots, e_n)$, которая может принимать любое значение.
4. ($w \neq 0$, v_i — базис $\implies$ $w(v_1, \dots, v_n) \neq 0$)
   Так как $w \neq 0$, то $\exists x_1, \dots, x_n \in V$: $w(x_1, \dots, x_n) \neq 0$
   $w(x_1, \dots, x_n) = w(v_1, \dots, v_n)\,\mathrm{Vol}^v(x_1, \dots, x_n)$
   Так как $w(x_1, \dots, x_n) \neq 0$, то и $w(v_1, \dots, v_n) \neq 0$

   ($w \neq 0$, $w(v_1, \dots, v_n) \neq 0 \implies$ v_i — базис)
   Пусть $v_1, \dots, v_n$ — не базис, тоесть некоторый вектор v_i выражается через остальные (без потери общности пусть выражается n-ый).
   Тогда $v_n = \sum_{i=1}^{n-1} \alpha_i v_i$
   $w(v_1, \dots, v_n) = \sum \alpha_i (v_1, \dots, v_{n-1}, v_i) = 0$
5. Следует из предыдущего пункта, рассмотрим пространство векторов и форму объёма в виде определителя.
6. A обратима $\iff$ обратим соответствующий линейный оператор, а он обратим $\iff$ образы $e_1, \dots, e_n$ — линейно независимы, а они линейно независимы $\iff$ столбцы A линейно независимы $\iff \det(A) \neq 0$. □

Замечание. Определитель det — единственная форма объёма, такая что $\det E = 1$

Маркер вычитки

Теорема 6.2. Пусть V — векторное пространство размерности n, $a \in \mathrm{End}(v)$, w — форма объёма.

1. Пусть $w_a \colon V^n \to K$, $w_a(v_1, \dots, v_n) := w(a(v_1), \dots, a(v_n))$.
   Тогда w_a — форма объёма.
2. Значение выражения $w(a(e_1), \dots, a(e_n))/w(e_1, \dots, e_n)$ не зависит ни от w, ни от базиса e (для ненулевых форм объёма w)

3. Пусть A — матрица оператора a в базисе e ($A = a_e^e$), тогда.
   Тогда $\det(A) = \mathrm{Vol}^e(a(e_1), \dots, a(e_n)) = w(a(e_1), \dots, a(e_n))/w(e_1, \dots, e_n)$
   Более того, $\det(A) = \det(a_e^e)$ не зависит от выбора базиса e

Доказательство. 1. Пусть A — матрица оператора a в e.

$$w(a(v_1), \dots, a(v_n)) = w(e_1, \dots, e_n) \det(a(v_1)^e, \dots, a(v_n)^e) =$$
$$= w(e_1, \dots, e_n) \det(A * (v_1^e \mid \dots \mid v_n^e)) = (w(e_1, \dots, e_n) \det(A)) \det(v_1^e, \dots, v_n^e)$$

   TODO: !!!!

2. $w \neq 0 \implies w(e_1, \dots, e_n) \neq 0$.
   Покажем независимость от формы w. Пусть $\hat{w}$ — ненулевая форма объёма. Так как пространство форм объёма одномерно, то $\hat{w} = cw$ для $c \in K^*$.

$$\frac{\hat{w}(a(e_1), \dots, a(e_n))}{\hat{w}(e_1, \dots, e_n)} = \frac{cw(a(e_1), \dots, a(e_n))}{cw(e_1, \dots, e_n)} = \frac{w(a(e_1), \dots, a(e_n))}{w(e_1, \dots, e_n)}$$

   Покажем независимость от базиса. Пусть $e_1, \dots, e_n$ — базис V.
   Так как w_a — форма объёма, то $w_a = dw$, для $d \in K$

$$\frac{w(a(e_1), \dots, a(e_n))}{w(e_1, \dots, e_n)} = \frac{dw(e_1, \dots, e_n)}{w(e_1, \dots, e_n)} = d$$

3. Так как Vol^e — тоже форма объёма, то по предыдущему пункту

$$\frac{w(a(e_1), \dots, a(e_n))}{w(e_1, \dots, e_n)} = \frac{\mathrm{Vol}^e(a(e_1), \dots, a(e_n))}{\mathrm{Vol}^e(e_1, \dots, e_n)} = \frac{\mathrm{Vol}^e(a(e_1), \dots, a(e_n))}{1} = \det(A)$$

   Так $\det(A) = \mathrm{Vol}^e(a(e_1), \dots, a(e_n)) = w(a(e_1), \dots, a(e_n))/w(e_1, \dots, e_n)$, и последнее не зависит от w и e, значит $\det(a_e^e)$ независит от e □

$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)}$

***Определение* 6.2.** Для $a \in \mathrm{End}(V)$, $\det a := w(a(e_1), \dots, a(e_n))/w(e_1, \dots, e_n))$

Или, $\det(a) := \det(a_e^e)$, по предыдущей лемме это эквивалентные определения, более того, ни то, ни другое не зависит от e (и в первом w).

Лемма. Пусть $A, B \in K^{n \times n}$, $a, b \in \mathrm{End}(v)$, тогда:

1. $\det(a \circ b) = \det a \det b$, $\det(AB) = \det A \det B$
2. $\mathrm{GL}(K^{n \times n}) = \{A \mid \det A \neq 0\}$
   $\mathrm{GL}(V) = \{a \in \mathrm{End}(v) \mid \det a \neq 0\}$
3. $\det\colon \mathrm{GL}(V) \to K^*$ (det — гомоморфизм групп)
   $\det\colon \mathrm{GL}(K^{n \times n}) \to K^*$ (det — гомоморфизм групп)

Доказательство. 1. Пусть $e = (e_1, \dots, e_n)$ — некоторый базис V.

Пусть b — обратим, тогда $b(e_1), \dots, b(e_n)$ — базис V.

$\det(a \circ b) = \frac{w(a(b(e_1)),\dots,a(b(e_n)))}{w(e_1,\dots,e_n)} = \frac{w(a(b(e_1)),\dots,a(b(e_n)))}{w(b(e_1),\dots,b(e_n))} \frac{w(b(e_1),\dots,b(e_n))}{w(e_1,\dots,e_n)}$

$\det(a \circ b) = \det a \det b$

Если b необратим, то $\det b = 0$ ($b(e_1), \dots, b(e_n)$ линейно зависимы).

Тогда $ab(e_1), \dots, ab(e_n)$ — тем более линейно зависимы.

Итого имеем $\det(a \circ b) = 0, \det b = 0$, формула верна.

2. Первая строчка уже была 6-ом пункте этой теоремы.

Хотим показать, что $\mathrm{GL}(V) = \{a \in \mathrm{End}(v) \mid \det(a) \neq 0\} = \{a \in \mathrm{End}(v) \mid \det(a_e^e) \neq 0\}$

В частности нужно показать, что $\det(a) = \det(a_e^e) \neq 0 \iff$ оператор a — обратим.

Но это правда, оператор обратим, тогда и только тогда, когда его матрица обратима.

Это следует из какой-то(TODO) леммы о том, что оператор $\iff$ образ базиса — базис.

3. TODO □

Несколько полезных свойств, рекомендуется попробовать доказать их самостоятельно.

Доказательства написаны ниже.

Лемма (Свойства определителя)**.** 1. $\det A = \det A^T$

2. Если в матрице A есть нулевой столбец или строка, то $\det A = 0$
3. $\det(A_1, \dots, A_n) = \det(A_1, \dots, A_i + \alpha A_j, \dots, A_n)$
4. $\det(\dots, A_i, \dots, \alpha A_i, \dots) = 0$
5. $\det(A_1, \dots, \alpha A_i, \dots, A_n) = \alpha \det(A_1, \dots, A_i, \dots, A_n)$

Лемма. 1. $\det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix} = \det A \det B$ (A и B — квадратные матрицы)

2. $$\det \begin{pmatrix} A_1 & * & * & \dots & * \\ 0 & A_2 & * & \dots & * \\ 0 & 0 & A_3 & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A_n \end{pmatrix} = \det A_1 \dots \det A_n.$$

Доказательство.

1. $$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_i a_{i,\sigma(i)} = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_i a_{i,\sigma^{-1}(i)} = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_i a_{\sigma(i),i} = \det A^T$$

Небольшое пояснение:

- Суммировать по всем перестановкам это то же самое, что суммировать по всем обратным перестановкам (каждое слагаемое и там, и там реализуется).

- $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$ (при такой замене знак у слагаемых сохранится).
- Внутри произведения можно поменять порядок умножения (каждый множитель и так, и так реализуется).

2. В сумме внутри определителя все произведения нулевые.

3. det — форма объёма, применим антисимметричность и полилинейность:
$\det(A_1, \dots, A_i + \alpha A_j, \dots, A_n) = \det(A_1, \dots, A_n) + \alpha \det(A_1, \dots, A_{i-1}, A_j, A_{i+1}, \dots, A_n)$
$= \det(A_1, \dots, A_n) + 0 = \det(A_1, \dots, A_n)$ □

Доказательство. 1. $\det \begin{pmatrix} A & * \\ 0 & B \end{pmatrix}$ **TODO:**

2. Применить индукцию и предыдущий пункт. □

***Определение* 6.3.** Пусть $B \in K^{n \times n}$, $1 \leqslant i, j \leqslant n$

Минором в позиции i, j ($M_{i,j}$) называется определитель матрицы, полученной вычёркиванием i-го столбца и j-ой строки.

Лемма. $\det A_{i,j} = \sum$

7. Обратные Матрицы, формулы Крамера

7.1. Обратные матрицы, формулы Крамера

***Определение* 7.1.** $B \in Mat(n, K)$
$Adj(B) = (\text{алгебраические дополнения})^T$ называется присоеденненной матрицей к B

Теорема 7.1. $A \in Mat(n, K)$
Тогда $A \cdot Adj(A) = Adj(A) \cdot A = (\det A) \cdot E$.
В частности, если матрица $A \in \mathrm{GL}(n, K)$, то $A^{-1} = \frac{1}{\det A} Adj(A)$.

Доказательство. $C = A \cdot Adj(A)$
$C_{ii} = \sum\limits_{j=1}^{n} A_{ij} Adj(A)_{ji} = \det A$ (разложение по строке).
$C_{ij} = 0$ $(i \neq j)$ □

Теорема 7.2. Формула Крамера $A \in Mat(n, K),\ b \in K^n,\ A_{ij} = a_{ij}$

Система линейных уравнений $Ax = b$ имеет единственное решение $\Leftrightarrow \Delta = \det A \neq 0$

решение: $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $x_i = \frac{\Delta_i}{\Delta}$, где $\Delta_i = \det \begin{pmatrix} a_{1,1} & \dots & b_{1,i} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \dots & b_{n,i} & \dots & a_{n,n} \end{pmatrix}$

Доказательство. A - матрица отображения $K^n \to K^n$
$a : x \to Ax$
$a^{-1}(b) = x_0 + \operatorname{Ker} a$

Решаем систему $Ax = b$, зная, что $\Delta = \det A \neq 0$.

Тогда $x = A^{-1}b = \frac{Adj(A) \cdot b}{\det A} = \frac{Adj(A) \cdot b}{\Delta} = \begin{pmatrix} \Delta_1/\Delta \\ \vdots \\ \Delta_n/\Delta \end{pmatrix}$ □

Теорема 7.3. Ранг матрицы равен размеру наибольшей квадратной подматрицы (пересечения r выбранных строк и r выбранных столбцов), определитель которой не равен 0 (невырожденный).

Доказательство. $A \in Mat(n, m)$, $r = \operatorname{rank} A$
k - размер наибольшей квадратной подматрицы с ненулевым определителем.
Есть подматрица размера $k \times k$, строки которой линейно независимы, а значит соответствующие строки матрицы A тоже. Значит $\operatorname{rank} A \geqslant k$. Теперь $r = \operatorname{rank} A$, то найдутся r линейно независимых строк. Они образуют подматрицу ранга r, в ней есть r линейно независимых столбцов. Рассмотрим эти столбцы, получим подматрицу размера $r \times r$, она обратима, так как ее ранг равен размеру, значит $\operatorname{rank} A \leqslant k$. □

7.2. Двойственное пространство

***Определение* 7.2.** Линейная функция на векторном пространстве V называется 1 – форма. То есть, линейное отображение : $V \to K$.

Замечание.

1. Линейное отображение однозначно определяется образами базисных векторов.
2. $e = (e_1, \ldots, e_n)$ — базис V, $a : V \to K$ $a_e \in \text{Multi}(1, n)$

Примеры:

1. $\forall u \in Mat(1, n)$ задает линейную функцию
   $K^n \to K$
   $x \rightarrowtail ux$
2. $V = \{f : X \to K\}$, X — множество, $x_0 \in X$
   $V \to K$
   $f \rightarrowtail f(x_0)$
3. $e = (e_1, \ldots, e_n)$ - базис V
   Тогда $1 \leqslant i \leqslant n$
   $e^i : V \to K$
   $v \rightarrowtail (v^e)^i$

Замечание. e^i — это отображение, сопоставляющее v i-ую координату вектора v в базисе e.

***Определение* 7.3.** Пространство линейных функций на V называется двойственным (сопряженным, дуальным) к V и обозначается V^*, элементы двойственного пространства называются ко-векторами.

Теорема 7.4. $\dim V < \infty$ Тогда:

1. $\dim V = \dim V^*$ (а значит $V \simeq V^*$)
2. $e = (e_1, \ldots, e_n)$ - базис пространства V. Тогда $e^1, \ldots, e^n$ (функции) - базис V

Доказательство. 1. Линейная независимость: $\sum\limits_{i=1}^{n} \alpha_i e^i = 0 \Rightarrow \sum\limits_{i=1}^{n} \alpha_i e^i(e_j) = 0 \ \forall 1 \leqslant i \leqslant n$, $\alpha_j = 0$

2. Система образующих: $f \in V^*$, $f(v) = \sum\limits_{i=1}^{n} e^i(v) \cdot f(e_i)$, $f = \sum\limits_{i=1}^{n} e^i \cdot f(e_i)$

□

***Определение* 7.4.** Базис $e^1, \ldots, e^n$ называется двойственным к базису $e_1, \ldots, e_n$

Замечание. V - конечномерное.
$V^{**} \simeq V^* \simeq V$ (двойственное к двойственному). Изоморфизмы не канонические (зависят от базиса), а вот $V^{**} \simeq V$ — канонический.

Теорема 7.5. $\dim V < \infty$
$V \to V^{**}$ - канонический изоморфизм
$v \rightarrowtail f_v$, где $f_v(\phi) = \phi(v)$ $(\phi \in V^*)$

Доказательство. $f_{v+w} = f_v + f_w,\ f_{\lambda v} = \lambda f_v,\ \lambda \in K$
$\dim V = \dim V^* = \dim V^{**}$
e — Базис V, $e = e_1, \ldots, e_n$. $f_{e_i}(e^j) = e^j(e_i) = 0$, если $i \neq j$ и $= 1$, если $i = j \Rightarrow f_{e_1}, \ldots, f_{e_n}$ - двойственное к $e^1, \ldots, e^n$ □

***Определение* 7.5.** Двойственное отображение $\phi : U \to V$ — линейное отображение
$\phi^* : V^* \to U^*$
$\phi^*(f) = f \circ \phi$
то есть ϕ^* — это $U \xrightarrow{\phi} V \xrightarrow{f} K$

Лемма. e — базис U, f — базис V, $\tilde{e}$ — двойственный к e базис U, $\tilde{f}$ — двойственный к f базис V, $a : U \to V$ — линейное отображение. Тогда $(a_{e\to f})^T = (a^*)_{\tilde{f}\to\tilde{e}}$

Доказательство. $((a^*)_{\tilde{f}\to\tilde{e}})_i = (a^*(f^i))^{\tilde{e}} = (f^i \circ a)^{\tilde{e}} = \begin{pmatrix} f^i(a(e_1)) \\ \vdots \\ f^i(a(e_n)) \end{pmatrix}$ (это i строка матрицы $a_{e\to f}$)

□

Следствие. $a^{**} = a$ (отождествляются)

7.3. Инвариантные подпространства

Замечание. $a \in \operatorname{End} V$
V - конечномерное векторное пространство над K
$U \leqslant V$
$a(U) = aU \leqslant V$

***Определение* 7.6.** Подпространство $U \leqslant V$ инвариантно относительно оператора a, если $aU \leqslant U$

Примеры:

1. $\operatorname{Ker} a,\ V,\ \operatorname{Im}\{a\}$
2. $\frac{d}{dx} : K[x] \to K[x]$, здесь $K_n[x] = \{f \in K[x] | \deg f \leqslant n\}$ — инвариантное подпространство.

Замечание.
1. U — инвариантное, $a \in \operatorname{End} V$
   $a|_U : U \to U$, то есть $a|_U \in \operatorname{End} U$
2. U — инвариантное относительно a
   $e_1, \ldots, e_m$ - базис U, дополним до $e_1, \ldots, e_m, \ldots e_n$ — базис V.
   $a_e = \begin{pmatrix} A_1 & * \\ 0 & * \end{pmatrix}$
   $A_1 = (a|_U)_e$

Теорема 7.6. $a \in \operatorname{End} V$, тогда V раскладывается в прямую сумму двух инвариантных подпространств (т.е. $\exists U, W \leqslant V$, U, W — инвариантные и $V = U \oplus W$) $\Leftrightarrow \exists e$ — базис V, такой, что $a_e = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$

Доказательство. $\Rightarrow$
$V = U \oplus W$

$e_1, \ldots, e_m$ — базис U, $e_{m+1}, \ldots, e_n$ — базис W, тогда $e_1, \ldots, e_n$ — базис V
$a_e = \begin{pmatrix} (a|_U)_{e_1,\ldots,e_m} & 0 \\ 0 & (a|_W)_{e_{m+1},\ldots,e_n} \end{pmatrix}$
$\Leftarrow$
$e_1, \ldots, e_n$ — базис V
$a_e = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$
$U = \langle e_1, \ldots, e_m \rangle$, $W = \langle e_{m+1}, \ldots, e_n \rangle$
$V = U \oplus W$, U, V — инвариантны, $u \in U$, $u = \sum_{i=1}^{m} \alpha_i e_i$, $a(u) = \sum_{i=1}^{m} \alpha_i a(e_i) \in U$ □

Теорема 7.7. Предыдущая теорема справедлива для прямой суммы из n слагаемых.

Доказательство. Индукция. □

7.4. Немного алгебры

***Определение* 7.7.** V — векторное пространство над полем K, V — кольцо (с той же операцией $+$)
$\alpha(ab) = (\alpha a)b = a(\alpha b)$ $\forall a, b \in V$, $\alpha \in K$
Тогда V — алгебра над полем K

Примеры: $\operatorname{End} V$, $Matn, K$, $K[x]$

***Определение* 7.8.** $f(x) = f_0 + f_1(x) + \ldots + f_n(x^n)$
Положим $f(a) = f_0 \cdot id + f_1 a + \ldots + f_n a^n$, $f(a) \in \operatorname{End}(V)$

Замечание. $\phi : K[x] \to \operatorname{End} V$
$\phi : f \rightarrowtail f(a)$
это гомоморфизм алгебр

$I = \operatorname{Ker} \phi$ — идеал $K[x]$, I — главный, $I = \mu_a(x) \cdot K[x]$

7.5. Минимальный многочлен (Minimal polynomial)

***Определение* 7.9.** Многочлен $\mu_a \in K[x]$ называется минимальным многочленом оператора a, если $\deg \mu = min\{\deg f | f \neq 0, f(a) = 0\}$ и старший коэффициент $= 1$

Замечание. Минимальный многочлен существует и единственный.

8. Многочлены операторов, спектр, собственные числа

8.1. Многочлены эндоморфизмов

Замечание. Если $f(x)$ — некоторый многочлен над полем K, а a – некоторый оператор над векторным пространством V, то $f(a)$ тоже будет оператором над V.

$f(x) \in K[x]$, $a \in \mathrm{End}(V) \implies f(a) \in \mathrm{End}(V)$.

Лемма. 1. Пусть $f(x) \in K[x], a \in \mathrm{End}(V)$, тогда:

$a^k(\mathrm{Ker}\, f(a)) \subseteq \mathrm{Ker}\, f(a)$

2. Пусть $f, g \in K[x]$, $g \vdots f$, $a \in \mathrm{End}(V)$, тогда:

$\mathrm{Ker}\, f(a) \subseteq \mathrm{Ker}\, g(a)$

Доказательство. 1. Пусть $f(a)(v) = 0$. Покажем, что $f(a)(a^k(v)) = 0$.

$f(a)(a^k(v)) = (\sum_t c_t a^t)(a^k(v)) = (\sum_t c_t\, a^{t+k})v = a^k f(a)(v) = 0$

2. $g \vdots f \implies g = hf$ для некоторого h.

$g(a) = h(a)f(a)$

Если $f(a)(v) = 0$, то $g(a)(v) = 0$. □

Утверждение 8.1. Пусть $f_1, \ldots, f_k$ — попарно взаимно просты.

Тогда $\mathrm{Ker}(f_1 * \ldots * f_k)(a) = \bigoplus_{i=1}^{k} \mathrm{Ker}\, f_i(a)$

Доказательство. Докажем для $k = 2$ (если $k = 1$, то утверждение очевидно, а $k > 2$ можно по индукции доказать через $k = 2$)

Так как f_1, f_2 — взаимно простые, то существуют $g_1, g_2 \in K[x]$, такие что $1 = g_1 f_1 + g_2 f_2$

$\mathrm{id} = g_1(a)f_1(a) + g_2(a)f_2(a)$

Тогда для любого v: $v = g_1(a)(f_1(a)(v)) + g_2(a)(f_2(a)(v))$

Пусть $v \in \mathrm{Ker}(f_1(a)f_2(a)) = \mathrm{Ker}(f_2(a)f_1(a))$

Тогда $(f_2(a)f_1(a))(v) = 0$, то есть $f_1(a)(v) \in \mathrm{Ker}\, f_2(a)$

Пусть $g_1(x) = \beta_0 + \beta_1 x + \ldots + \beta_l x^l$

Воспользуемся леммой выше и подставим $f_1(a)(v)$.

$g_1(a)(f_1(a)(v)) = \beta_0 f_1(a)(v) + \beta_1 a(f_1(a)(v)) + \ldots + \beta_l a^l(f_1(a)(v)) \in \mathrm{Ker}\, f_2(a)$

Каждое слагаемое $\in \mathrm{Ker}\, f_2(a)$, а значит и сумма тоже.

Аналогично получается, что $g_2(a)(f_2(a)(v)) \in \mathrm{Ker}\, f_1(a)$, а значит мы разложили $v \in \mathrm{Ker}(f_1(a)f_2(a))$ в сумму.

Осталось показать, что сумма получилась прямая.

Пусть $v \in \mathrm{Ker}\, f_1(a) \cap \mathrm{Ker}\, f_2(a)$, тогда

$v = g_1(a)(f_1(a)(v)) + g_2(a)(f_2(a)(v)) = g_1(a)(0) + g_2(a)(0) = 0$. □

8.2. Проекторы, след

***Определение* 8.1.** Если $V = U \oplus W$, $a\colon V \to V$, $a(u+w) := u$, то оператор a называется проектором.

Замечание. В частности, для любого проектора верно, что:

$a^2 = a$ (иначе говоря, $a^2 - a = 0$)

Замечание. На самом деле, любой оператор, обладающий таким свойством, является проектором (а само пространство в таком случае раскладывается в прямую сумму).

Доказательство. $a^2 - a = 0$

То есть $V = \operatorname{Ker}(a^2 - a) = \operatorname{Ker}(a) \oplus \operatorname{Ker}(a - \mathrm{id}_V)$

Действительно, многочлен x и многочлен $x - 1$ — взаимно просты над $K[x]$.

$1 = 1 * x + (-1) * (x - 1)$ □

***Определение* 8.2.** Оператор a, такой что $a^2 = \mathrm{id}$, называется отражением.

Замечание. Если a — отражение над V и $\operatorname{Char} K \neq 2$, то $V = \operatorname{Ker}(a - \mathrm{id}) \oplus \operatorname{Ker}(a + \mathrm{id})$

Доказательство. $a^2 - \mathrm{id} = (a - \mathrm{id})(a + \mathrm{id})$

Доказательство взаимной простоты (заметьте, что требуется, что $2 \neq 0$):

$1 = \frac{1}{2} * (x + 1) - \frac{1}{2} * (x - 1)$ □

***Определение* 8.3.** Пусть $A \in K^{n \times n}$

Тогда следом матрицы A называется $\operatorname{tr} A := \sum_{i=1}^{n} A_{i,i}$

Лемма. Для $A, B \in K^{n \times n}$ $\operatorname{tr} AB = \operatorname{tr} BA$

Доказательство. Нужно увидеть, что суммируются просто одинаковые элементы.

$\operatorname{tr} AB = \sum_k (AB)_{k,k} = \sum_k \sum_i a_{k,i} b_{i,k} = \sum_i \sum_k b_{i,k} a_{k,i} = \sum_i (BA)_{i,i} = \operatorname{tr} BA$ □

***Определение* 8.4.** Следом оператора называется след его матрицы в каком-то базисе.

Лемма (о корректности определения)**.** След матрицы оператора не зависит от выбора базиса.

Доказательство. $\operatorname{tr} C^{-1}AC = \operatorname{tr} A$ для обратимой матрицы C.

Действительно, $\operatorname{tr} C^{-1}AC = \operatorname{tr}(C^{-1})(AC) = \operatorname{tr}(AC)(C^{-1}) = \operatorname{tr} A$. □

Утверждение 8.2. 1. Функция tr линейна: $\operatorname{tr}(\alpha A + \beta B) = \alpha \operatorname{tr}(A) + \beta \operatorname{tr}(B)$

2. $\operatorname{tr}\left(A^T\right) = \operatorname{tr}(A)$

Доказательство. Очевидно из определения tr. □

8.3. Спектр и характеристический многочлен оператора

Пусть $a \in \operatorname{End}(V)$

***Определение* 8.5.** Ненулевой $v \in V$ называется собственным вектором оператора a, если $a(v) = \lambda v$ для $\lambda \in K$.

***Определение* 8.6.** Если есть $v \in V$, $v \neq 0$, такой что $a(v) = \lambda v$, то λ наывается собственным числом оператора a.

Утверждение 8.3. $a \in \operatorname{End} V$. λ — собственное число оператора $a \iff \det(a - \lambda \operatorname{id}) = 0$

Доказательство. $\exists v \neq 0 \colon a(v) = \lambda v \iff \exists v \neq 0 \colon v \in \operatorname{Ker}(a - \lambda \operatorname{id}) \iff$

$\iff \operatorname{Ker}(a - \lambda \operatorname{id}) \neq \{0\} \iff \det(a - \lambda \operatorname{id}) = 0$ □

***Определение* 8.7.** Спектром оператора называется множество всех собственных чисел этого оператора.

$\operatorname{Spec}(a) := \{\lambda \in K \mid (a - \lambda \operatorname{id}) \notin \operatorname{GL}(V)\}$

***Определение* 8.8.** Характеристическим многочленом матрицы $A \in K^{n \times n}$ называется

$\chi_A(t) := \det(tE - A)$

Замечание. $\chi_A(t) \in K[t]$, $\deg \chi_A(t) = n$

***Определение* 8.9.** Пусть $a \in \operatorname{End} V$, тогда $\chi_a(t) := \chi_{a_e}(t)$

Замечание (о корректности). $\chi_{a_e}(t) = \det(t \operatorname{id} - a_e^e) = \det\left(t \operatorname{id} - a_f^f\right) = \chi_{a_f}$

Левое и правое равенство берётся по определению. Обсудим среднее. Под определителем и слева и справа записан одинаковый линейный оператор: $v \mapsto t * v - a(v)$.

И, так как мы знаем, что определитель не зависит от базиса, то равенство верно.

Утверждение 8.4.

1. Множество собственных чисел оператора a совпадает с множеством корней χ_a.
2. $\chi_{a_e} = \chi_{a_f}$

   $\operatorname{Spec} a = \operatorname{Spec} a_e = \operatorname{Spec} a_f$
3. $\chi_a(t) = t^n - (\operatorname{tr} a)t^{n-1} + \ldots + (-1)^n \det a$.

Их доказательство почти очевидным образом следует из предыдущих утверждений.

Следствие. Если $K = \mathbb{C}$ (или другое алгебраически замкнутое поле), то любой $a \in \operatorname{End}(V)$ (для конечномерного V) имеет собственный вектор.

Теорема 8.5 (Кэли-Гамильтона)**.** Пусть V — конечномерное, $a \in \operatorname{End} V$.

Тогда $\chi_a(a) = 0$

Теорема 8.6 (Кэли-Гамильтона для матриц)**.** Если $A \in K^{n \times n}$, то $\chi_A(A) = 0$

Замечание. Эти две формулировки эквивалентны.

Действительно, матрицу можно воспринимать как оператор над $V = K^n$, а любой оператор можно записать в каком-нибудь базисе (так как χ оператора не зависит от базиса).

Доказательство. Будем доказывать формулировку для матриц.

Пусть $\chi_A(t) = t^n + c_{n-1}t^{n-1} + \ldots + c_0$

Пусть $B := \operatorname{Adj}(tE - A)$, то есть $B \in K[t]^{n \times n}$

Тогда $(tE - A)B = \det(tE - A)E = \chi_A(t)E$

Разложим B: $B = \sum_{i=0}^{n-1} t^i B_i$, $B \in K^{n \times n}[t]$.

$\chi_A(t)E = (tE - A)B = (tE - A)\sum_{i=0}^{n-1} t^i B_i = \sum_{i=0}^{n-1} t^{i+1} B_i - \sum_{i=0}^{n-1} t^i AB_i =$
$= t^n B_{n-1} + \sum_{i=1}^{n-1} t^i (B_{i-1} - AB_i) - AB_0$

Тогда $\chi_A(A)E = \chi_A(A) = A^n B_{n-1} + \sum_{i=1}^{n-1}(A^i B_{i-1} - A^{i+1} B_i) - AB_0$

Обратите внимание, что в сумме соседние слагаемые "убивают" друг-друга, так как в них записаны одни и те же мономы. Остающиеся два слагаемых как раз совпадают с двумя слагаемыми вне суммы.

То есть $\chi_A(A) = 0$. □

9. Собственные числа и жорданова форма оператора

9.1. Связь характерестического многочлена с минимальным

Факт.

χ_a — характеристический многочлен, $\chi_a(a) = 0$.

μ_a — минимальный многочлен, $\mu_a(a) = 0$.

Следствие.

1. $\mu_a \mid \chi_a$
2. Собственные числа и только они являются корнями μ_a

Доказательство.

Первый пункт следует из того, что μ_a — минимальный многочлен, зануляющий оператор a.

Докажем второй пункт.

Поскольку $\mu_a | \chi_a$, то $\{\text{корни } \mu_a\} \in \operatorname{Spec} a$.

Покажем теперь, что $\forall\, \lambda \in \operatorname{Spec} a : \mu_a(\lambda) = 0$.

Действительно, $\lambda \in \operatorname{Spec} a \implies \exists\, v \neq 0 : a(v) = \lambda v$. И, в частности, $a^k(v) = \lambda^k v$.

Тогда $0 = \mu_a(a)(v) = \sum\limits_{i=0}^{m} c_i a^i(v) = \sum\limits_{i=0}^{m} c_i \lambda^i v = \mu_a(\lambda) v \implies \mu_a(\lambda) = 0$. □

9.2. Кратности собственных чисел

***Определение* 9.1.**

1. $V_\lambda(a) := \operatorname{Ker}(a - \lambda \cdot \mathrm{id})$ — собственное подпространство, соответствующее собственному числу $\lambda \in K$.
2. Пусть $\lambda \in \operatorname{Spec} a$. Тогда:

   Алгебраическая кратность $\alpha(\lambda, a)$ — кратность λ как корня характеристического многочлена $\chi_a(x)$.

   Геометрическая кратность $\gamma(\lambda, a) = \dim \operatorname{Ker}(a - \lambda \cdot \mathrm{id})$.

Лемма.

Геометрическая кратность не превосходит алгебраической кратности.

То есть $\gamma(\lambda, a) \leqslant \alpha(\lambda, a)$.

Доказательство.

Рассмотрим подпространство $\operatorname{Ker}(a - \lambda \cdot \mathrm{id})$.

Выберем в нём базис. Т.к. все выбранные векторы в выбранном базисе будут являться собственными векторами оператора a, то его матрица в таком базисе будет являться диагональной (с собственными числами λ на главной диагонали).

Дополнив теперь этот базис до базиса пространства V, получим матрицу вида:

$$a_e = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & * \end{array}\right)$$

Здесь размер клетки с собственным числом λ равен γ.

Вспомним теперь, что значение характеристического многочлена не зависит от выбора базиса.

Тогда, вспомнив также, что $\chi_a(\lambda) = \det(a - \lambda \cdot \mathrm{id})$, увидим, что кратность корня λ в χ_a будет хотя бы γ.

Следовательно, $(x - \lambda)^\gamma \mid \chi_a$. □

9.3. Собственные значения и корневые пространства

Теорема 9.1.

1. Сумма собственных подпространств является прямой.
2. Собственные векторы, соответствующие различным собственным числам, линейно независимы.

Доказательство.

1. Пусть $\lambda_1, \lambda_2, \ldots, \lambda_k$ — собственные числа.

   Если все собственные числа различны, то $(x - \lambda_i)$ взаимно просты.

   Тогда $\operatorname{Ker}\left(\prod_{i=1}^{k}(x - \lambda_i)\right)(a) = \bigoplus_{i=1}^{k} \operatorname{Ker}(a - \lambda_i \cdot \mathrm{id})$ (по доказанной ранее теоремы).

2. $v_1, v_2, \ldots, v_k$ — собственные векторы, соответствующие различным собственным числам.

   $a(v_i) = \lambda_i v_i,\ v_i \neq 0 \implies v_i \in \operatorname{Ker}(a - \lambda_i \cdot \mathrm{id})$

   $\operatorname{Ker}(a - \lambda_i \cdot \mathrm{id})$ — собственное подпространство, соответствующее собственному числу λ.

   Вспомнив теперь, что сумма собственных подпространств — прямая, поймём, что ноль представим в виде линейной комбинации v_i единственным образом. А именно, когда все коэффициенты при v_i равны 0.

   Следовательно, v_i линейно независимы.

□

***Определение* 9.2.**

Оператор a диагонализуем, если существует базис e, в котором его матрица a_e диагональна.

Теорема 9.2.

Следующие условия эквивалентны.

1. a — диагонализуем.
2. $\mu_a = \prod\limits_{\lambda \in \operatorname{Spec} a} (x - \lambda)$.
3. $V = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} \operatorname{Ker}(a - \lambda \cdot \mathrm{id})$.
4. $\dim V = \sum\limits_{\lambda \in \operatorname{Spec} a} \gamma(\lambda, a)$.

Доказательство.

"1 $\implies$ 2"

По определению диагонализуемости, существует такой базис $e = (e_1, e_2, \ldots, e_n)$, что оператор a_e — диагональный.

Знаем, что все собственные числа являются корнями μ_a.

Следовательно, по теореме Безу $\prod\limits_{\lambda \in \operatorname{Spec} a} (x - \lambda) \Big| \mu_a$.

Осталось показать, что $\mu_a \Big| \prod\limits_{\lambda \in \operatorname{Spec} a} (x - \lambda)$. То есть, $\prod\limits_{\lambda \in \operatorname{Spec} a} (x - \lambda)(a) = 0$.

Поскольку образ оператора однозначно задаётся образом базисных векторов, покажем, что $\forall\, e_i : \prod\limits_{\lambda \in \operatorname{Spec} a} (x - \lambda)(a)(e_i) = 0$.

Действительно. Поскольку оператор a является диагональным в выбранном базисе, то e_i является собственным вектором.

И, соответственно $(x - \lambda_i)(a)(e_i) = (a - \lambda_i \cdot \mathrm{id})(e_i) = 0$.

Поскольку это верно для любого базисного элемента, то получаем, что $\mu_a = \prod\limits_{\lambda \in \operatorname{Spec} a} (x - \lambda)$.

"2 $\implies$ 3"

$$V = \operatorname{Ker} \mu_a(a) = \operatorname{Ker} \left(\prod_{\lambda \in \operatorname{Spec} a} (x - \lambda) \right) (a) = \bigoplus_{\lambda \in \operatorname{Spec} a} \operatorname{Ker}(a - \lambda \cdot \mathrm{id})$$

"3 $\implies$ 4"

$$\dim V = \sum_{\lambda \in \operatorname{Spec} a} \dim \operatorname{Ker}(a - \lambda \cdot \mathrm{id}) = \sum_{\lambda \in \operatorname{Spec} a} \gamma(\lambda, a)$$

"4 $\implies$ 1"

Уже знаем, что сумма собственных подпространств — прямая.

Также знаем, что $\dim V = \sum\limits_{\lambda \in \operatorname{Spec} a} \gamma(\lambda, a) = \sum\limits_{\lambda \in \operatorname{Spec} a} \dim \operatorname{Ker}(a - \lambda \cdot \mathrm{id})$.

Следовательно, получаем, что $V = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} \operatorname{Ker}(a - \lambda \cdot \mathrm{id})$.

Выберем базис в каждом собственном подпространстве $\operatorname{Ker}(a - \lambda \cdot \mathrm{id})$. Объединение этих базисов будет в точности являться базисом пространства V, а матрица оператора в этом базисе будет диагональной. □

***Определение* 9.3.**

V — векторное пространство, $v \in V$.

v — корневой вектор оператора a, соответствующий собственному числу λ, если существует такое число m, что $(a - \lambda \cdot \mathrm{id})^m v = 0$.

Наименьшее такое m называется высотой вектора v.

Замечание.

Собственный вектор является корневым вектором с высотой 1.

Замечание.

Корневые векторы образуют корневое подпространство $V^\lambda(a) := \bigcup\limits_{j \in \mathbb{N}} \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^j$.

$v \in \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^m,\ u \in \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^n \implies (\alpha v + \beta u) \in \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{\max\{m,n\}}$

Лемма.

1. $V^\lambda(a)$ — инвариантно относительно оператора a.
2. $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^j \subseteq \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1}$
3. Если два ядра совпали в цепочке вложенных ядер, то и далее все ядра будут совпадать.
   То есть, если $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^j = \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1}$, то $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1} = \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+2}$.

Доказательство.

1. $v \in V^\lambda(a)$
   $a(v) = (a - \lambda \cdot \mathrm{id})v + \lambda v,\ \lambda v \in V^\lambda(a)$
   $v \in V^\lambda(a) \implies \exists\, m : (a - \lambda \cdot \mathrm{id})^m v = 0 \implies (a - \lambda \cdot \mathrm{id})^{m-1}\left((a - \lambda \cdot \mathrm{id})v\right) = 0$
   Получили, что $(a - \lambda \cdot \mathrm{id})v \in V^\lambda(a)$.
   Следовательно, $(a - \lambda \cdot \mathrm{id})v + \lambda v \in V^\lambda(a) \implies a(v) \in V^\lambda(a)$.
2. Достаточно заметить, что если $(a - \lambda \cdot \mathrm{id})^j v = 0$, то $(a - \lambda \cdot \mathrm{id})^{j+1} \cdot v = (a - \lambda \cdot \mathrm{id}) \cdot 0 = 0$.
   Следовательно, $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^j \subseteq \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1}$.
3. Пусть $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^j = \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1}$, но $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1} \neq \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+2}$.
   Значит, существует $v \in \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+2} \setminus \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1}$.
   $(a - \lambda \cdot \mathrm{id})v \in \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{j+1} \setminus \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^j$.
   Однако мы предположили, что эти два подпространства совпадают. Следовательно, их разность — пустое множество.
   Получили противоречие.

□

Следствие.

$V^\lambda(a) = \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^m$, где m — такое минимальное число, что $\mathrm{Ker}(a - \lambda \cdot \mathrm{id})^m = \mathrm{Ker}(a - \lambda \cdot \mathrm{id})^{m+1}$.

Теорема 9.3.

Пусть χ_a раскладывается в произведение многочленов степени 1 в $K[x]$.

Тогда $V = \bigoplus\limits_{\lambda \in \mathrm{Spec}\, a} V^\lambda(a)$.

Доказательство.

$V = \operatorname{Ker} \chi_a(a) = \operatorname{Ker} \prod\limits_{\lambda \in \operatorname{Spec} a} (x-\lambda)^{\alpha(\lambda,a)}(a) = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} \operatorname{Ker}(a - \lambda \cdot \mathrm{id})^{\alpha(\lambda,a)} = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} V^{\lambda}(a)$ □

Замечание.

Если K — алгебраически замкнуто, то χ_a раскладывается в произведение многочленов степени 1.

9.4. Жорданов базис. Случай нильпотентного оператора

***Определение* 9.4.**

Пусть $b \in \operatorname{End}(V)$.

b — нильпотентен, если $b^m = 0$ для некоторого m.

***Определение* 9.5.**

$\operatorname{ht} v = \min\{k | b^k v = 0\}$ — высота вектора v.

Лемма.

Если $h = \operatorname{ht} v$, то $v, bv, b^2v, \ldots, b^{h-1}v$ — линейно независимы.

Доказательство.

Рассмотрим линейную комбинацию данных векторов:

$\alpha_0 v + \alpha_1 bv + \ldots + \alpha_{h-1} b^{h-1} v = 0$

Домножим все элементы данного выражения на b^{h-1}. Заметим, что все элементы, кроме $\alpha_0 v$ — занулятся.

Следовательно, $\alpha_0 = 0$.

Домножим все элементы данного выражения на b^{h-2}. Уже знаем, что $\alpha_0 = 0$. Значит, не занулится только элемент $\alpha_1 bv$.

Следовательно, $\alpha_1 = 0$.

Продолжив так далее, получим, что все $\alpha_i = 0$. □

***Определение* 9.6.**

$\langle v, bv, \ldots, b^{h-1}v \rangle$ — циклическое (инвариантное) подпространство нильпотентного оператора b, порождённое v.

На самом деле, оно является минимальным инвариантным подпространством, содержащим v.

Лемма.

1. Циклическое подпространство инвариантно относительно b.
2. Матрица $b\Big|_{\langle v, bv, \ldots, b^{h-1}v \rangle}$ в базисе $b^{h-1}v, \ldots, bv, v$ имеет следующий вид:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

3. U — инвариантное циклическое подпространство, $\dim U = h$

   $e \in U \setminus bU \implies U = \langle e, be, \ldots, b^{h-1}e \rangle$

4. U — инвариантное циклическое подпространство, $U = \langle v, bv, \ldots, b^{h-1}v\rangle$

   Тогда $\operatorname{Ker} b\Big|_U = \langle b^{h-1}v\rangle$.

Доказательство.

1. $U = \langle v, bv, \ldots, b^{h-1}v\rangle$

   $u \in U,\ u = \alpha_0 v + \alpha_1 bv + \ldots + \alpha_{h-1}b^{h-1}v$

   $bu = \alpha_0 bv + \alpha_1 b^2 v + \ldots + \alpha_{h-2}b^{h-1}v \in U$.

2. Рассмотрим данную матрицу. Заметим, что она попросту циклически сдвигает элементы вектора, зануляя последний: $(x_1, x_2, \ldots, x_{h-1}, x_h) \mapsto (x_2, x_3, \ldots, x_h, 0)$

   Ровно таким образом действует наш оператор на заданном базисе.

3. $U = \langle v, bv, \ldots, b^{h-1}v\rangle$

   $e = \alpha_0 v + \alpha_1 bv + \ldots + \alpha_{h-1}b^{h-1}v,\ \alpha_0 \neq 0$

   Достаточно показать, что $\operatorname{ht} e = h$.

   Действительно, $b^{h-1}e = \alpha_0 b^{h-1}v \neq 0 \implies U = \langle e, be, \ldots, b^{h-1}e\rangle$.

4. $b(\alpha_0 v + \alpha_1 bv + \ldots + \alpha_{h-1}b^{h-1}v) = 0 \implies \alpha_0 bv + \alpha_1 b^2 v + \ldots + \alpha_{h-2}b^{h-1}v = 0$

   Получили, что $\alpha_0 = \alpha_1 = \ldots = \alpha_{h-2} = 0$ и, соответственно, $\alpha_{h-1}b^{h-1}v \in \langle b^{h-1}v\rangle$

□

9.5. Относительный базис

***Определение* 9.7.**

$U \leqslant V$

Векторы $v_1, v_2, \ldots, v_k$ называются линейно независимыми относительно подпространства U, если $\alpha_1 v_1 + \ldots + \alpha_k v_k \in U \implies \forall\, \alpha_i = 0$

***Определение* 9.8.**

$v_1, v_2, \ldots, v_k$ — базис V относительно U, если v_i линейно независимы относительно U и $\forall v \in V : v = \alpha_1 v_1 + \ldots + \alpha_k v_k + u,\ u \in U$

Замечание.

Любое линейно независимое множество относительно U можно дополнить до относительного базиса.

Лемма.

Пусть $v_1, v_2, \ldots, v_s \in \operatorname{Ker} b^{j+1}$ и линейно независимы относительно $\operatorname{Ker} b^j$.

Тогда $bv_1, bv_2, \ldots, bv_s \in \operatorname{Ker} b^j$ и линейно независимы относительно $\operatorname{Ker} b^{j-1}$.

Доказательство.

Пусть $\alpha_1 bv_1 + \alpha_2 bv_2 + \ldots + \alpha_s bv_s \in \operatorname{Ker} b^{j-1}$.

Тогда $b^{j-1}(\alpha_1 bv_1 + \alpha_2 bv_2 + \ldots + \alpha_s bv_s) = 0 \iff b^j(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_s v_s) = 0$.

Следовательно, $\alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_s v_s \in \operatorname{Ker} b^j \implies \alpha_1, \alpha_2, \ldots, \alpha_s = 0$. □

Теорема 9.4.

1. V раскладывается в прямую сумму циклических (инвариантных) подпространств нильпотентного оператора b.
2. Количество слагаемых равно $\dim \operatorname{Ker} b$.

Доказательство.

$b^m = 0,\ V = \operatorname{Ker} b^m$

$\operatorname{Ker} b \subseteq \operatorname{Ker} b^2 \subseteq \ldots \subseteq \operatorname{Ker} b^m = V$

Рассмотрим базис $v_1, v_2, \ldots, v_s$ пространства $V = \operatorname{Ker} b^m$ относительно $\operatorname{Ker} b^{m-1}$.

Выберем теперь базис пространства $\operatorname{Ker} b^{m-1}$ относительно $\operatorname{Ker} b^{m-2}$. Для этого заметим, что векторы $bv_1, bv_2, \ldots, bv_s$ линейно независимы. Значит, их можно дополнить до базиса векторами $u_1, u_2, \ldots, u_t$.

Таким образом, у нас уже есть базис пространства $\operatorname{Ker} b^m$ относительно $\operatorname{Ker} b^{m-2}$ — это объединение выбранных базисов.

Действуя так и далее, выберем базис всего пространства V. Для удобства, покажем описанные действия на следующей схеме:

$V = \operatorname{Ker} b^m$ относительно $\operatorname{Ker} b^{m-1}$	$v_1, \ldots, v_s$			
$\operatorname{Ker} b^{m-1}$ относительно $\operatorname{Ker} b^{m-2}$	$bv_1, \ldots, bv_s$	$u_1, \ldots, u_t$		
$\operatorname{Ker} b^{m-2}$ относительно $\operatorname{Ker} b^{m-3}$	$b^2v_1, \ldots, b^2v_s$	$bu_1, \ldots, bu_t$	$w_1, \ldots, w_q$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
$\operatorname{Ker} b$	$b^{m-1}v_1, \ldots, b^{m-1}v_s$	$b^{m-2}u_1, \ldots, b^{m-2}u_t$	$b^{m-3}w_1, \ldots, b^{m-3}w_q$	$\cdots$

Заметим, что, по сути, мы выбирали базисы подпространств

$$\operatorname{Ker} b^m / \operatorname{Ker} b^{m-1},\ \operatorname{Ker} b^{m-1} / \operatorname{Ker} b^{m-2}, \ldots, \operatorname{Ker} b.$$

А поскольку $\operatorname{Ker} b \subseteq \operatorname{Ker} b^2 \subseteq \ldots \subseteq \operatorname{Ker} b^m$, то выбранные векторы являются линейно независимыми, и $\sum\limits_{i=0}^{m-1} \dim \operatorname{Ker} b^{i+1} / \operatorname{Ker} b^i = \dim \operatorname{Ker} b^m = \dim V$.

Следовательно, выбранные векторы действительно образуют базис пространства V.

Перейдём теперь к доказательству теоремы:

1. Мы научились строить базис пространства V. Заметим, что векторы вида $v, bv, b^2v, \ldots$ образуют циклическое подпространство.

   Таким образом, каждый "столбик" в полученной нами диаграмме соответствует базису некоторого циклического пространства.

   Отсюда следует разложение V в прямую сумму циклический подпространств.

2. $V = \bigoplus\limits_{i=1}^{k} V_i,\ V_i$ — циклическое подпространство.

   $\operatorname{Ker} b = \bigoplus\limits_{i=1}^{k} \operatorname{Ker} b\Big|_{V_i} \implies \dim \operatorname{Ker} b = \sum\limits_{i=1}^{k} \dim \operatorname{Ker} b\Big|_{V_i} = k$

□

Замечание.

Пусть a — произвольный оператор.

$V = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} V^{\lambda}(a)$

$b = (a - \lambda \cdot \mathrm{id})\Big|_{V^{\lambda}(a)}$ — нильпотентный оператор.

$a\Big|_{V^{\lambda}(a)} = b + \lambda \cdot \mathrm{id}$

Отсюда получаем, что существует базис, в котором матрица оператора a имеет вид

$$a_e = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}, \text{ где } J_i = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

***Определение* 9.9.**

Базис, в котором матрица a_e имеет указанный блочно-диагональный вид, называется жордановым, а сама форма матрицы — жордановой.

Теорема 9.5.

1. Если χ_a раскладывается в произведение многочленов степени 1, то существует жорданов базис.

2. Пусть a_e — матрица оператора в базисе e над алгебраически замкнутым полем K.

   Тогда $\exists\, C \in \mathrm{GL}_n(K) : C^{-1}\, a_e\, C$ — жорданова форма оператора.

Доказательство.

Докажем первый пункт (второй следует из первого).

Пусть $\chi_a(t) = \prod\limits_{\lambda \in \operatorname{Spec} a} (t - \lambda)^{\alpha(\lambda, a)}$.

Тогда знаем, что $V = \operatorname{Ker} \chi(a) = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} \operatorname{Ker}(a - \lambda \cdot \mathrm{id})^{\alpha(\lambda, a)}$

Заметим, что $\operatorname{Ker}(a - \lambda \cdot \mathrm{id})^{\alpha(\lambda, a)} = V^{\lambda}(a)$. Следовательно, $V = \bigoplus_{\lambda \in \operatorname{Spec} a} V^{\lambda}(a)$. □

10. Функции от операторов

Теорема 10.1.

1. Количество жордановых клеток не зависит от выбора базиса.
2. Сумма порядков жордановых клеток соответствующего λ равна алгебраической кратности $\alpha(\lambda, a)$.

Доказательство.

1. Знаем, что $V = \bigoplus\limits_{\lambda \in \operatorname{Spec} a} V^\lambda(a)$.

   Мы знаем, что количество жордановых клеток в матрице оператора $a\Big|_{V^\lambda(a)}$ не зависит от выбора базиса.

   Следовательно, их количество не зависит от выбора базиса и в матрице оператора a.

2. Заметим, что сумма размеров жордановых клеток с собственным числом λ — это в точности $\dim V^\lambda(a)$.

   А мы уже знаем, что $\dim V^\lambda(a) = \alpha(\lambda, a)$.

□

Замечание.

Минимальный многочлен $J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ равен $\mu_J = (t-\lambda)^m$.

Пояснение:

Очевидно, что $(t-\lambda)^m$ является характеристическим многочленом оператора $J_m(\lambda)$.

Также можно заметить, что оператор $J_m(\lambda) - \lambda E$ является нильпотентным и соответствует оператору, который переводит вектор $(x_1, x_2, \ldots, x_{m-1}, x_m)$ в вектор $(x_2, x_3, \ldots, x_m, 0)$. В частности, потому m — минимальная степень, обнуляющая $J_m(\lambda) - \lambda E$.

10.1. Многочлен от матрицы

Пусть у нас есть оператор $a \in \operatorname{End}(V)$ и некоторый многочлен $f \in K[x]$. Сведём вычисление значения многочлена от матрицы к вычислению многочлена степени, меньшей $n = \deg \chi_a$.

А именно, положим $f = q\chi_a + r$, $\deg r < n$. Тогда:

1. $f(a) = r(a)$
2. Если же имеет место разложение $\chi_a = \prod (x - \lambda_i)^{m_i}$, то многочлен r однозначно задаётся условиями:
   $f^{(k)}(\lambda_i) = r^{(k)}(\lambda_i) \ \ \forall\, 0 \leqslant k < m_i$

Далее везде рассматриваем алгебраически замкнутые поля (если не оговорено обратное).

Факт.

$A \in K^{n\times n}$

$A = C^{-1}JC,\ \ C \in \mathrm{GL}_n(K)$

Тогда:

$A^k = (C^{-1}JC)^k = \underbrace{(C^{-1}JC)(C^{-1}JC)\dots(C^{-1}JC)}_{k\ \ times} = C^{-1}J^kC$

В частности, если имеется многочлен $p \in K[x]$, то $p(A) = C^{-1}p(J)C$.

Факт.

Положим $A = J_n(\lambda)$. Тогда

$$A^k = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & (\lambda^k)' & \frac{(\lambda^k)''}{2!} & & \frac{(\lambda^k)^{(n-1)}}{(n-1)!} \\ & \lambda^k & (\lambda^k)' & & \\ & & \ddots & \ddots & \\ & & & \lambda^k & (\lambda^k)' \\ & & & & \lambda^k \end{pmatrix}$$

Для доказательства сего факта достаточно заметить, что $\frac{(\lambda^n)^{(k)}}{k!} = \binom{n}{k}\lambda^{n-k}$, а далее воспользоваться методом математической индукции, показав, как изменяется форма матрицы J_n^k при домножении на J_n.

В частности, пусть $p \in K[x]$. Тогда:

$$p(A) = \begin{pmatrix} p(\lambda) & p'(\lambda) & \frac{p''(\lambda)}{2!} & & \frac{p^{(n-1)}(\lambda)}{(n-1)!} \\ & p(\lambda) & p'(\lambda) & & \\ & & \ddots & \ddots & \\ & & & p(\lambda) & p'(\lambda) \\ & & & & p(\lambda) \end{pmatrix}$$

Замечание.

$p(J_n(\lambda)) = \sum_{k=0}^{n-1} \frac{p^{(k)}(\lambda)}{k!} J_n^k(0)$

10.2. Норма оператора

***Определение* 10.1.**

Вспомним, что такое норма:

$\|\cdot\| : V \to K$

1. $\|x\| > 0,\ x \neq 0$
   $\|0\| = 0$
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$
3. $\|x + y\| \leqslant \|x\| + \|y\|$

***Определение* 10.2.**

$\rho(u, v) = \|u - v\|$ — метрика.

Свойства сходимости.

1. В $\mathbb{R}^n$ сходимость по норме $\iff$ покоординатная сходимость.
2. В $\mathbb{R}^n$ все нормы эквивалентны.

***Определение* 10.3.**

$U,\ V$ — нормированные векторные пространства.

$a : U \to V$ — линейное отображение.

$\|a\| := \sup\limits_{\|v\|=1} \|a(v)\| = \sup\limits_{v\neq 0} \frac{\|a(v)\|}{\|v\|}$

Замечание.

$\dim U,\ \dim V < \infty$

$\|a\| = \max\limits_{\|v\|=1} \|a(v)\|$

В частности, далее будем рассматривать только конечномерные пространства.

Утверждение 10.2.

V — пространство с $\|\cdot\|$.

$a, b \in \mathrm{End}(V),\ \ \|a \cdot b\| \leqslant \|a\| \cdot \|b\|$

Доказательство.

$$\|ab\| = \max_{v\neq 0} \frac{\|ab(v)\|}{\|v\|} = \max_{b(v)\neq 0} \frac{\|ab(v)\|}{\|b(v)\|} \cdot \frac{\|b(v)\|}{\|v\|} \leqslant$$
$$\leqslant \max_{b(v)\neq 0} \frac{\|ab(v)\|}{\|b(v)\|} \cdot \max_{b(v)\neq 0} \frac{\|b(v)\|}{\|v\|} \leqslant \max_{b(v)\neq 0} \frac{\|ab(v)\|}{\|b(v)\|} \cdot \max_{v\neq 0} \frac{\|b(v)\|}{\|v\|} = \|a\| \cdot \|b\| \qquad \square$$

***Определение* 10.4.**

Пусть дан набор чисел $c_i \in K$. Радиусом сходимости ряда $\sum c_i x^i$ будем называть (каким бы ни было поле K) вещественное число $R \in \mathbb{R}$, такое что $\frac{1}{R} = \overline{\lim} \sqrt[n]{|c_n|}$. Разумеется, иногда R бывает бесконечным.

Факт.

Пусть ряд $\sum c_i x^i$ имеет радиус сходимости R. Тогда для всех $x \in K$, таких что $|x| < R$, ряд $\sum c_i x^i$ сходится (в том числе, абсолютно), а при всех $|x| > R$ — расходится.

Утверждение 10.3.

V — пространство с $\|\cdot\|$

$a \in \mathrm{End}(V)$ — линейное отображение.

Пусть $f(x) = \sum\limits_{n=0}^{+\infty} c_n x^n$ — ряд с радиусом сходимости R.

Тогда если $\|a\| < R$, то $f(a) = \sum\limits_{n=0}^{+\infty} c_n a^n$ — сходится.

Доказательство.

$\left\| \sum\limits_{n=0}^{+\infty} c_n a^n \right\| \leqslant \sum\limits_{n=0}^{+\infty} \|c_n a^n\| \leqslant \sum\limits_{n=0}^{+\infty} |c_n| \cdot \|a\|^n$

Но $\|a\| < R$, следовательно, этот ряд сходится.

А поскольку сходится $\|f(a)\|$, то сходится и сам $f(a)$. $\square$

10.3. Экспонента от оператора

***Определение* 10.5.**

V — линейное пространство.

$a \in \mathrm{End}(V)$

$e^a = \sum\limits_{k=0}^{\infty} \frac{a^k}{k!}$

Теорема 10.4.

1. V — нормированное пространство.

   $a, b \in \mathrm{End}(V)$

   Если $a \circ b = b \circ a$, то $e^{a+b} = e^a \circ e^b$.

2. $A \in \mathbb{C}^{n\times n}$

   $\det e^A = e^{\mathrm{tr}\, A}$

   $e^{A^T} = (e^A)^T$ (следует из покоординатной сходимости)

Доказательство.

1. $e^{a+b} = \sum\limits_{k=0}^{+\infty} \frac{1}{k!}(a+b)^k = \sum\limits_{k=0}^{+\infty} \frac{1}{k!} \sum\limits_{i=0}^{k} \binom{k}{i} a^i b^{k-i}$

   $e^a \circ e^b = \left(\sum\limits_{k=0}^{+\infty} \frac{a^k}{k!}\right)\left(\sum\limits_{l=0}^{+\infty} \frac{b^l}{l!}\right) = \sum\limits_{k,\, l\geqslant 0} \frac{1}{k!\, l!} a^k b^l = \sum\limits_{k,\, l\geqslant 0} \frac{1}{(k+l)!}\binom{k+l}{l} a^k b^l = \sum\limits_{n=0}^{+\infty} \sum\limits_{k,\, l\geqslant 0}^{k+l=n} \frac{1}{(k+l)!}\binom{k+l}{l} a^k b^l$

   Очевидно, что полученные записи эквивалентны.

   Что же касается коммутативности, то ею мы пользуемся при раскрытии бинома.

2. $e^A = \sum\limits_{n=0}^{+\infty} \frac{A^n}{n!} = \sum\limits_{n=0}^{+\infty} C^{-1} \frac{J^n}{n!} C = C^{-1} \left(\sum\limits_{n=0}^{+\infty} \frac{J^n}{n!}\right) C = C^{-1} e^J C$

   Таким образом, получаем, что $\det e^A = \det e^J$.

   Вспомнив теперь, как выглядит матрица многочлена от жордановой клетки, поймём, что $\det e^J = e^{\mathrm{tr}\, J} = e^{\mathrm{tr}\, A}$.

□

Теорема 10.5.

Рассмотрим оператор $a \in \mathrm{End}(V)$ и некоторый ряд $f(x) = \sum\limits_{n=0}^{+\infty} c_n x^n$ над полем $\mathbb{C}$.

Положим m_i — алгебраическая кратность собственного числа λ_i оператора a.

Пусть для многочлена $r \in \mathbb{C}[x]$, такого что $\deg r < n = \deg \chi_a$ выполняется условие

$$f^{(k)}(\lambda_i) = r^{(k)}(\lambda_i), \quad 0 \leqslant k < m_i$$

Тогда $f(a) = r(a)$.

Доказательство.

$$f_m(x) := \sum_{i=0}^{m} c_i x^i, \quad f_m = q_m \chi_a + r_m$$

$$\begin{array}{ccc|c} f_m(a) & = & r_m(a) & f_m^{(k)}(\lambda_i) = r_m^{(k)}(\lambda_i) \\ \downarrow & {\scriptstyle m\to\infty} & \downarrow & \\ f(a) & & r(a) & f^{(k)}(\lambda_i) = r^{(k)}(\lambda_i) \end{array}$$

□

11. Многочлены над конечными полями

11.1. Многочлены над кольцами

Утверждение 11.1. R область целостности $\implies$ $R[x]$ — область целостности.

Доказательство. Если у одного многочлена старший коэффициент a, а у другого старший коэффициент b, то у произведения будет ab.

Если $a \neq 0$ и $b \neq 0$, то и $ab \neq 0$. □

Замечание. По тем же причинам $\deg(fg) = \deg(f) + \deg(g)$.

Замечание (напоминание). f — неприводим, если $f = gh \implies f \sim g$ или $f \sim h$.

Если R — факториально, то $R[x]$ — то же.

Далее в этой главе R — факториальное, ассоциативное и коммутативное кольцо.

***Определение* 11.1.** Пусть $f(x) \in R[x]$, $f(x) = a_0 + a_1x + \ldots + a_nx^n$.

Тогда *содержанием* многочлена f называется $d(f) := \gcd(a_i)$.

Замечание. Вспомним, что gcd определён с точностью до ассоциированности, то есть домножения на обратимый элемент.

Замечание. Если $c \in R$, $f \in R[x]$, то $d(cf) = c\,d(f)$.

Лемма (Гаусса)**.** Пусть R — факториально, $f, g \in R[x]$.

Тогда $d(fg) = d(f)d(g)$.

Доказательство. 1. Рассмотрим случай $d(f) = d(g) = 1$.

Пусть $f(x) = a_0 + a_1x + \ldots + a_nx^n$, $g(x) = b_0 + b_1x + \ldots + b_mx^m$

Пусть $d(fg) \neq 1$, тогда $\exists$ простой p, такой что $d(fg) \vdots p$.

Пусть s, t — наименьшие индексы, для которых $a_s \not\vdots p$, $b_t \not\vdots p$.

Рассмотрим коэффициент fg при x^{s+t}

$0 = \sum\limits_{i+j=s+t} a_ib_j = a_sb_t \pmod p$

А значит $a_sb_t \vdots p$, а так как p — простой, то $a_s \vdots p$ или $b_t \vdots p$ (противоречие).

2. Общий случай.

Так как $d(\ldots)$ — gcd всех коэффициентов многочлена, то на него можно делить.

Пусть $f = d(f)f_0$, $g = d(g)g_0$, в частности $d(f_0) = d(g_0) = 1$.

Тогда $d(fg) = d(d(f)f_0\,d(g)g_0) = d(f)d(g)d(f_0g_0) = d(f)d(g)$ □

Следствие. Если $f \in \mathbb{Z}[x]$, f неприводим в $\mathbb{Z}[x]$, то f неприводим и в $\mathbb{Q}[x]$.

Доказательство. Пусть $f = gh$, $g, h \in \mathbb{Q}[x]$.

Вынесем из g и h дробь, чтобы они были в $\mathbb{Z}[x]$, затем вынесем из этого содержание, итого:

$f = \frac{p}{q} g_0 h_0$, где $p, q \in \mathbb{Z}$, $g_0, h_0 \in \mathbb{Z}[x]$, $d(g_0) = d(h_0) = 1$.

$qf = pg_0h_0$

$d(qf) = d(pg_0h_0)$, то есть $qd(f) = p$

$qf = qd(f)g_0h_0$

$f = d(f)g_0h_0 = (d(f)g_0)\, h_0$. □

Следствие. Если $f \in R[x]$, f неприводим в $R[x]$, то f неприводим и в $(\operatorname{Quot} R)[x]$.

Доказательство совпадает с доказательством частного случая $R = \mathbb{Z}[x]$.

11.2. Возведение в p-ую степень, извлечение p-ого корня

Лемма. Пусть p — простое, V — векторное пространство над $\mathbb{F}_p$, $a_i \in V$.

Тогда $(a_1 + \ldots + a_n)^p = a_1^p + a_2^p + \ldots + a_n^p$

Доказательство.

- $n = 1$, говорить не о чем.
- $n = 2$.

  $(a_1 + a_2)^p = \sum_i C_p^i a_1^i a_2^{p-i}$

  Для $1 \leqslant i \leqslant p-1$, $C_p^i = \frac{p!}{i!(p-i)!} \vdots p$, так как числитель на p делится, а знаменатель нет.
- $n \geqslant 3$.

  $(a_1 + a_2 + \ldots + a_n)^p = (a_1 + (a_2 + \ldots + a_n))^p = a_1^p + (a_2 + \ldots + a_n)^p = a_1^p + a_2^p + \ldots + a_n^p$

  Пользуемся индукцией и случаем $n = 2$. □

Следствие. Если в векторном пространстве над полем $\mathbb{F}_p$ рассмотреть функцию $f\colon x \mapsto x^p$, то окажется, что она эндоморфизм.

- $f(xy) = (xy)^p = x^p y^p = f(x)f(y)$
- $f(x + y) = (x + y)^p = x^p + y^p = f(x) + f(y)$
- $f(cx) = (cx)^p = c^p x^p = cx^p = cf(x)$

Замечание (алгоритм извлечения корня). Пусть мы ищем корень $b_0 + b_1 x + \ldots + b_k x^k$,

то есть такой $a_0 + a_1 x + \ldots + a_n x^n$, что $(a_0 + a_1 x + \ldots + a_n x^n)^p = b_0 + b_1 x + \ldots + b_k x^k$

$(a_0 + a_1 x + \ldots + a_n x^n)^p = a_0^p + a_1^p x^p + \ldots + a_n^p x^{np} = a_0 + a_1 x^p + \ldots + a_n x^{np}$

Видно, что не каждый многочлен $\sum_i b_i x^i$ являются степенью p, а только тот, у которого все коэффициенты при степенях x, кроме делящихся на p, равны 0.

Соответственно чтобы извлечь корень нужно превратить все мономы x^{kp} в x^k

11.3. Критерий Эйзенштейна

Теорема 11.2. Пусть $f \in \mathbb{Z}[x]$, $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$, p — простое.

$a_0, a_1, \ldots, a_{n-1} \vdots p$, $a_n \nmid p$, $a_0 \nmid p^2$

Тогда f неприводим над $\mathbb{Z}[x]$ (а значит, по предыдущему следствию, и над $\mathbb{Q}[x]$)

Доказательство. Предположим, что $f(x) = (\sum_{k=0}^{m} b_k x^k)(\sum_{k=0}^{s} c_l x^l)$, где $m, s > 0$.

$a_0 \vdots p$, $a_0 \nmid p^2$, $a_0 = b_0 c_0$, значит либо $b_0 \vdots p$, либо $c_0 \vdots p$, но не вместе.

Рассмотрим случай, когда $b_0 \vdots p$, $c_0 \nmid p$ (другой абсолютно аналогичен).

Рассмотрим множество $\{i \mid b_i \nmid p\}$, оно непусто так как m там лежит, а значит имеет минимальный элемент i_0.

$a_{i_0} = \sum_{j=0}^{i_0} b_j c_{i_0 - j} = b_{i_0} c_0 \pmod p$.

$c_0 \nmid p$, $b_{i_0} \nmid p$, $a_{i_0} \vdots p$ (заметим, что $i_0 \leqslant m < n$).

Противоречие. □

Пример. Покажем, что многочлен $f(x) = 1 + x + \ldots + x^{p-1}$ неприводим для любого простого p.

Сразу заметим, что $f(x) = \frac{x^p - 1}{x - 1}$.

Во-вторых заметим, что раскладывать многочлен $f(x)$ это то же самое, что раскладывать многочлен $f(x+1)$ (разложение одного несложно переделать в разложение другого).

$f(x+1) = \frac{(x+1)^p - 1}{x} = \frac{x^p + \sum_{k=1}^{p-1} C_p^k x^k}{x} = x^{p-1} + \sum_{k=1}^{p-1} C_p^k x^{k-1}$.

Так как $C_p^k = \frac{p!}{k!(p-k)!} \vdots p$, то по критерию Эйзенштейна получаем требуемое.

11.4. Конечные поля

Замечание. Число многочленов над $\mathbb{F}_p[x]$ степени $n \geqslant 1$ равно $(p-1)p^n$.

Действительно, первый коэффициент можно выбрать $p - 1$ способом, а остальные n как угодно (многочлен степени n определяется $n + 1$ коэффициентом).

Утверждение 11.3. Пусть F — конечное поле, тогда

1. $|F| = p^m$ для некоторого простого p и некоторого целого m, при чём $p = \operatorname{Char} F$.
2. Пусть $q = |F|$, тогда $\forall a \in F^* \colon a^{q-1} = 1$.

Доказательство. • Рассмотрим канонический гомомрфизм $\varphi \colon \mathbb{Z} \to F$ ($1 \mapsto 1$, остальное определяется из этого).

Так как поле конечно, то у φ есть нетривиальное ядро $p\mathbb{Z}$, для некоторого $p > 0$.

В частности $\operatorname{Char} F$ по определению и равна p.

Заметим, что p — простое, если нет и $p = ab$, тогда элементы a, b — являются делителями 0, чего не бывает в полях.

Тогда существует вложение из $\mathbb{F}_p$ в F, в частности можно воспринимать F как расширение поля $\mathbb{F}_p$ ($\mathbb{F}_p \subset F$).

Таким образом, F — это векторное пространство над $\mathbb{F}_p$ (можно проверить, что все аксиомы соблюдаются).

$\dim_{\mathbb{F}_p} F = m < \infty$, так как F — конечно.

Тем самым, $|F| = p^m$.

- F^* — группа порядка $q-1$, Поэтому $a^{q-1} = 1$ по теореме Лагранжа. □

Замечание. Последний результат можно сформулировать также в виде $\forall x \in \mathbb{F}_p \colon x^p = x$

11.5. Приводимость многочленов, алгоритм Берлекэмпа

Замечание. Существуют многочлены сколь угодно большой степени, являющиеся приводимыми по модулю любого простого, но не являющиеся приводимым в целых числах.

Примером такого многочлена может служить x^4+1 (доказательство предлагается в качестве упражнения, разбор есть в конце секции).

Теорема 11.4. Пусть $f(x) \in \mathbb{F}_p[x]$

1. Если $h(x) \in \mathbb{F}_p[x]$, $h^p = h \pmod f$, то $f(x) = \prod_{a \in \mathbb{F}_p} \gcd(f(x), h(x) - a)$
2. Если $f(x) = f_1(x) \dots f_k(x)$ (f_i — попарно различные, неприводимые), то $h^p(x) = h(x) \pmod f \iff (\forall i \colon h(x) = a_i \pmod{f_i}$, где $a_i \in \mathbb{F}_p)$

   При чём каждому набору $(a_1, \dots, a_k)$ соответствует ровно один многочлен h ($\deg h < \deg f$) такого свойства.

Доказательство. 1. $F(x) := \prod_{a \in \mathbb{F}_p} \gcd(f, h(x) - a)$

$h(x) - a$ попарно взаимно просты для различных a, а значит имеет имеет место равенство: $F(x) = \gcd(f, \prod_{a \in \mathbb{F}_p} h(x) - a)$, в частности $f(x) \vdots F(x)$ даже без $h^p = h \pmod f$.

В поле $\mathbb{F}_p$, $y^p - y = \prod_{a \in \mathbb{F}_p}(y - a)$

Действительно, мы знаем, что все элементы поля являются корнями, а корней у многочлена степени p не более p.

Этот результат можно обобщить до $h^p(x) - h(x) = \prod_{a \in \mathbb{F}_p}(h(x) - a)$, это следует из того, что мы доказали разложение многочлена выше и сделали замену переменной $y \to h(x)$.

Тем самым $\prod_{a \in \mathbb{F}_p}(h(x) - a) = h^p(x) - h(x) \vdots f(x)$, а значит $F(x) = f(x)$.

2. Последняя часть является прямым следствием Китайской Теоремы об остатках (напомним, что теорема утверждает, что если $I_1, \dots, I_n$ — попарно взаимно простые идеалы, то $R/_{I_1 I_2 \dots I_n} \cong R/_{I_1} \oplus \dots \oplus R/_{I_n}$).

   Заметим, что $\gcd(f_i, f_j) = 1$ для $i \neq j$, потому что если это не так, тогда хотя бы один из них не является неприводимым, а так как $\gcd = 1$, то есть линейная комбинация дающая 1, то есть f_i попарно взаимно простые.

   Пусть $h^p(x) = h(x) \pmod f$, покажем, что $h(x) = a_i \pmod{f_i}$

   $f \mid h(x)^p - h(x) = \prod_{a \in \mathbb{F}_p}(h(x) - a)$

   В частности, $\forall i \colon f_i \mid h^p(x) - h(x) = \prod_{a \in \mathbb{F}_p}(h(x) - a)$

   Так как f_i является неприводимым, то $h(x) - a_i \vdots f_i$

   Пусть $\forall i \colon h(x) = a_i \pmod{f_i}$, покажем, что $h^p = h \pmod f$

Так как $f(x) = f_1(x) \dots f_k(x)$, где все f_i — различны и неприводимы, то достаточно показать, что $h^p(x) = h(x) \pmod{f_i}$ для всех i.

$h(x) = a_i \pmod{f_i}$

$h^p(x) = a_i^p \pmod{f_i} = a_i \pmod{f_i}$ □

Замечание. Если перефразировать теорему, то первая её часть утверждает, что если найти многочлен h, такой что $h^p = h \pmod{f}$, то можно разложить f, вторая часть предлагает некоторый инструмент для доказательства алгоритма берлекэмпа в случае, если у f нет кратных множителей.

***Определение* 11.2.** Если f не имеет кратных множителей, то многочлен $h(x) \in \mathbb{F}_p[x]$, такой что $h^p = h \pmod{f}$ и $1 \leqslant \deg h(x) < \deg f(x)$ называется f-разлагающим многочленом (потому что задаёт какое-то частичное разложение f).

Замечание. Почему не 0? Любая константа действительно задаёт разложение, но оно имеет вид $f = \gcd(f, 0)$. Почему $< \deg f$? Потому что по модулю f всё большее не интересно.

Начнём с того, что научимся избавляться от кратных множителей.

Лемма. Пусть f неприводим, тогда $\gcd(f, f') = 1$.

Доказательство. Предлагается в качестве упражнения и есть в конце главы. □

Лемма. Пусть $f = f_1^{n_1} \dots f_k^{n_k}$, где f_i — различные неприводимые многочлены над полем $\mathbb{F}_p$

Тогда $\gcd(f, f') = \prod\limits_{i, n_i \vdots p} f_i^{n_i} \prod\limits_{i, n_i \not\vdots p} f_i^{n_i - 1}$.

Доказательство. $f' = \sum_i n_i f_i' (f_1^{n_1} \dots f_i^{n_i - 1} \dots f_k^{n_k}) = \sum_i n_i \frac{f f_i'}{f_i}$

$$\gcd(f, f') = \gcd(f_1^{n_1} \dots f_k^{n_k}, \textstyle\sum_i n_i \frac{f f_i'}{f_i}) = f_1^{n_1 - 1} \dots f_k^{n_k - 1} \gcd(f_1 \dots f_k, \sum_i n_i f_i' (f_1 \dots f_{i-1} f_{i+1} f_k))$$

Если $n_i \vdots p$, то соответствующее слагаемое в сумме убивается, и у всех остальных слагаемых суммы тогда находится общий множитель f_i, который можно вынести, преобразуем. Для упрощения индексов перенумеруем f_i, так чтобы все f_i с $n_i \vdots p$ занимали места $f_1, \dots, f_t$, а все остальные $f_{t+1}, \dots, f_k$.

$$\gcd(f, f') = f_1^{n_1} \dots f_t^{n_t} f_{t+1}^{n_{t+1} - 1} \dots f_k^{n_k - 1} \gcd(f_{t+1} \dots f_k, \textstyle\sum_{i=t+1}^{k} n_i f_{t+1} \dots f_{i-1} f_i' f_{i+1} f_k)$$

Для завершения доказательства покажем, что сумма $\sum_{i=t+1}^{k} n_i f_{t+1} \dots f_{i-1} f_i' f_{i+1} f_k$ не делится на f_j для $j \in [t+1; k]$.

$$\textstyle\sum_{i=t+1}^{k} (n_i f_{t+1} \dots f_{i-1} f_i' f_{i+1} f_k) = n_j f_{t+1} \dots f_{j-1} f_j' f_{j+1} \dots f_k \pmod{f_j}$$

Так как f_j не приводим и не совпадает с другим f_i, а также $f_j' \not\vdots f_j$, то последнее не 0. □

Лемма. Пусть $f = f_1^{n_1} \dots f_k^{n_k}$, где f_i — различные неприводимые над $\mathbb{Z}$, $\mathbb{Q}$ или $\mathbb{R}$.

Тогда $\gcd(f, f') = f_1^{n_1 - 1} f_2^{n_2 - 1} \dots f_k^{n_k - 1}$

Доказательство. Доказательство абсолютно аналогично предыдущему пункту и потому для краткости опущено.

Единственная разница в том, что так как Char $= 0$, то не возникает особого случая при $n_i \vdots p$. □

Замечание (набросок алгоритма). Если $f = f_1^{n_1} \dots f_k^{n_k}$, то по лемме выше

$$f = \gcd(f, f') \prod_{i, n_i \nvdots p} f_i$$

В частности, вычислив $\gcd(f, f')$ мы сводим задачу разложения f к разложению, собственно gcd, а также второй части, назовём её $g(x)$.

Так как $g(x)$ не имеет общих множителей, то для него применим общий алгоритм, о котором пойдёт речь ниже, пока лишь предположим, что мы нашли разложение $g(x)$.

Часть множителей, полученных из $g(x)$ лежит и в gcd, выделим их оттуда и вместо gcd у нас останется $\prod\limits_{i, n_i \vdots p} f_i^{n_i}$, так как $n_i \vdots p$, то gcd равен некоторому многочлену в степени p.

Извлечём корень, и вызовемся от него рекурсивно.

Замечание (основной случай). Научимся факторизовать многочлен в том случае, если он не имеет кратных множителей.

Как нам показал первый пункт теоремы выше, для этого достаточно найти какой-то f-разлагающий многочлен $h(x)$. Заметим, что условие задаёт $h^p = h \pmod f$ какие-то линейные уравнения на h, изучим их.

Пусть $n = \deg f$, $h(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1}$.

$h^p(x) = a_0 + a_1 x^p + \dots + a_{n-1} x^{p(n-1)}$

Пусть для $i \in [0; n-1]$, $x^{iq} = \sum_{j=0}^{n-1} b_{j,i} x^j \pmod f$, то есть разложим x^{iq} по модулю f.

$h^p(x) = \sum_{i=0}^{n-1} a_i \sum_{j=0}^{n-1} b_{j,i} x^j$, $h(x) = \sum_{j=0}^{n-1} a_j x^j$

Тогда по-сути задана система уравнений на a_i: $\sum_{i=0}^{n-1} a_i b_{j,i} = a_j$ для всех j.

Пусть B — матрица из $b_{i,j}$, а $\vec{a}$ — вектор из a_i, тогда надо решить систему $(B - E)\vec{a} = 0$.

Тем самым пространство решений $h^p = h \pmod f$ изоморфно $\mathbb{F}_p^k$ для некоторого k.

Выберем базис $h_1, \dots, h_k$ пространства решений $(B - E)\vec{a} = 0$.

Так как $(1, 0, \dots, 0)$ всегда является решением, то можно выбрать базис, в котором $h_1 = (1, 0, \dots, 0)$

Если $k = 1$, то в качестве h можно взять только константу, это означает, что многочлен неприводим. Если ставить себе цель проверить на неприводимость, то на этом можно закончить.

Оказывается, что верно даже более сильное утверждение: найденное $k = \dim \operatorname{Ker}(B - E)$ совпадает с числом неприводимых множителей у f (пусть k').

- По второму пункту теоремы 11.4 мы знаем, что число многочленов h, таких что $h^p = h \pmod f$ равно $p^{k'}$, ибо между множеством допустимых h и кортежами $(a_1, \dots, a_{k'})$ по теореме существует взаимо однозначное соответствие.
- Мы только что нашли все решения $h^p = h \pmod f$ из $(B-E)\vec{a} = 0$, тем самым заявив, что их p^k.
- Итог: $k' = k$

Пусть $k \geqslant 2$. Можно было бы взять какой-то произвольный нетривиальный элемент, например h_2, объявить его как h, частично факторизовать по первому пункту теоремы 11.4 и решать задачу рекурсивно, но можно сделать лучше.

Чатично разложим f по h_2: $f(x) = \prod_{a \in \mathbb{F}_p} \gcd(f(x), h_2(x) - a)$.

Возможно нам повезло и мы уже сразу обрели все k множителей, но скорее всего нет, поэтому с каждым из (нетривиальных, то есть $\neq 1$) полученных множителей выше можно проделать разложение по h_3, и так далее.

Почему это корректно?

- Если $h^p - h$ делился на f, то $h^p - h$ делится и на делитель f (необходимое условие применения первого пункта теоремы 11.4).
- Все множители мы найдём. Об этом как раз второй пункт теоремы 11.4.

  Чтобы найти все множители нужно для каждой пары i, j $(i \neq j)$ разложить по h, такому что $h(x) \pmod{f_i} \neq h(x) \pmod{f_j}$. То есть каждую пару нужно "разъединить".

  Ну и второй пункт утверждает, что такой h есть, так как по любому набору a_i можно построить h.

  Почему такой h найдётся именно среди перебираемых $h_1, h_2, \ldots, h_n$ (давайте с теоретической точки зрения считать, что мы перебрали h_1, просто он не вносит никакого разбиения)?

  Любой h является линейной комбинацией h_t, то есть если бы $h_t(x) \pmod{f_i} = h_t(x) \pmod{f_j}$ для всех t, то и не было бы h, который их бы разделил. Противоречие.

Замечание. С практической точки зрения для больших p считать матрицу B можно с помощью быстрого возведения в степень

Правда ускорения это добавит только если мы проверяли многочлен на неприводимость, иначе всё равно придётся перебирать все элементы поля (это происходит в тот момент, когда мы пишем равенство $f(x) = \prod_a \gcd(f(x), h(x) - a)$).

11.6. Лемма Гензеля

***Определение* 11.3.** Если $f = f_1 f_2 \pmod{p^m}$, где $f, f_1, f_2 \in \mathbb{Z}[x]$, $\deg(f) = \deg(f_1) + \deg(f_2)$, старший коэффициент f_1 равен 1, старший коэффициент f взаимно прост с p, f_1 и f_2 взаимно просты по модулю p (то есть существуют многочлены u, v, что $uf_1 + vf_2 = 1 \mod p$), то назовём такое разложение *разложением по Гензелю*

NB: определение выше не является общепринятым и введено для краткости.

Замечание. Если u' и v' другие такие многочлены $(u'f_1 + v'f_2 = 1 \pmod{p})$, то $u' = u + wf_2$, $v' = v - wf_2$, то есть многочлены u и v однозначно заданы по модулю p, если потребовать, что $\deg u < \deg f_2$ и $\deg v < \deg f_1$.

Возможно не очевидно, почему двух неравенств $\deg u < \deg f_2$ и $\deg v < \deg f_1$ можно достичь одновременно. Давайте выберем u, такой что $\deg u < \deg f_2$, тогда $\deg(uf_1) < \deg(f_1 f_2)$, $\deg(1) = 0$, а значит $\deg(vf_2) < \deg(f_1 f_2)$ и $\deg v < \deg f_1$.

***Определение* 11.4.** *Продолжением Гензеля* для разложения по Гензелю $f = f_1 f_2 \pmod{p^m}$ назовём разложение по Гензелю $f = \bar{f}_1 \bar{f}_2 \pmod{p^{m+1}}$, где дополнительно $\bar{f}_i = f_i \pmod{p^m}$, и $\deg \bar{f}_i = \deg f_i$.

Лемма (Гензеля)**.** Пусть $m \geqslant 1$ и $f = f_1 f_2 \pmod{p^m}$ — разложение по Гензелю. Тогда существует продолжение Гензеля $f = \bar{f}_1 \bar{f}_2 \pmod{p^{m+1}}$. При этом многочлены $\bar{f}_i$ определены однозначны по модулю p^{m+1}.

Доказательство. Условия $\bar{f}_i = f_i \pmod{p^m}$, $\deg \bar{f}_i = \deg f_i$ означают, что

$\bar{f}_i = f_i + p^m g_i$, где $g_i \in \mathbb{Z}[x]$. Так как многочлен f_1 и $\bar{f}_1$ имеют старший коэффициент 1 и $\deg \bar{f}_1 = \deg f_1$, то $\deg g_1 < \deg f_1$.

Изучим условие $f = \bar{f}_1 \bar{f}_2 \pmod{p^m}$.

$\bar{f}_1 \bar{f}_2 = f_1 f_2 + p^m(g_1 f_2 + g_2 f_1) + p^{2m} g_1 g_2 = f \pmod{p^{m+1}}$

Так как $p^{2m} \vdots p^{m+1}$, то имеем $p^m(g_1 f_2 + g_2 f_1) = f - f_1 f_2 \pmod{p^{m+1}}$

Так как $f = f_1 f_2 \pmod{p^m}$, то $f - f_1 f_2 \vdots p^m$ и можно поделить.

$g_1 f_2 + g_2 f_1 = d \pmod{p}$, где $d = (f - f_1 f_2)/p^m$, $d \in \mathbb{Z}[x]$.

Как мы помним, $u f_1 + v f_2 = 1 \pmod{p}$, то есть $g_1 = d * v + w f_1 \pmod{p}$, $g_2 = d * u - w f_2 \pmod{p}$ для $w \in \mathbb{Z}[x]$.

Так как $\deg(g_1) < \deg(f_1)$, то g_1 определён однозначно (по модулю p), а это значит, что однозначно определён g_2.

Нужно лишь убедиться, что $\deg(g_2) < \deg(f_2)$.

Это следует из равенства $g_1 f_2 + g_2 f_1 = d \pmod{p}$, $d = (f - f_1 f_2)/p^m$.

Так как $\deg f = \deg f_1 + \deg f_2$, то $\deg d < \deg f$. Также $\deg(g_1) < \deg(f_1)$, то есть $\deg(g_1 f_2) < \deg f$. А значит и $\deg(g_2 f_1) < \deg f$ и $\deg g_2 < \deg f_2$. □

11.7. Разбор упражнений

Лемма. Пусть f неприводим, тогда $\gcd(f, f') = 1$.

Доказательство. Заметим, что $f' \neq 0$.

Действительно, если $f = a_0 + a_1 x + \ldots a_n x^n$, то $f' = a_1 + 2a_2 x + \ldots n a_n x^{n-1}$, что равно нулю, если все $a_i = 0$, кроме $i \vdots p$.

Но в таком случае (см алгоритм извлечения p-ого корня) f является p-ой степенью какого-то многочлена, что противоречит его неприводимости.

Так как f неприводим, то $\gcd(f, f') = 1$ или $\gcd(f, f') = f$, но так как $f' \neq 0$, то $1 \leqslant \deg f' < \deg f$, а значит последнее невозможно. □

Утверждение 11.5. $x^4 + 1$ неприводим в $\mathbb{Q}[x]$, $\mathbb{Z}[x]$.

Доказательство. $x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$ над $\mathbb{R}[x]$.

Так как любое разложение над $\mathbb{Q}[x]$ и $\mathbb{Z}[x]$ является разложением и над $\mathbb{R}[x]$, то имея разложения над $\mathbb{R}[x]$ можно утверждать, что разложение над $\mathbb{Q}[x]$ и над $\mathbb{Z}[x]$ может только объединить некоторые скобки этого разложения.

Соответственно исходное разложение не существует в $\mathbb{Q}[x]$ и в $\mathbb{Z}[x]$, а объединяя скобки получаем только $x^4 + 1$ □

Утверждение 11.6. $x^4 + 1$ приводим в любом $\mathbb{F}_p$ для простого p.

Доказательство. Отдельно разберём случай $p = 2$ (так как для него не сработает теория о символах Лежандра), тогда $x^4 + 1 = (x + 1)^4$.

Предположим, что $x^4 + 1 = (x^2 + ax + b)(x^2 + cx + d)$ для некоторых констант a, b, c, d.

Возможно имеет место более сильное разложение (например $1+1+2$ или $1+1+1+1$), но из этого уже будет следовать приводимость.

$$x^4+1=(x^2+ax+b)(x^2+cx+d) \iff \begin{cases} a+c=0 \\ b+d+ac=0 \\ ad+bc=0 \\ bd=1 \end{cases} \iff \begin{cases} a=-c \\ b+d-a^2=0 \\ ad-ba=0 \\ bd=1 \end{cases}$$

Верно одно из двух, либо $a=0$, либо $b=d$, разберём оба случая.

- Случай $a=0$

  $$\begin{cases} a=c=0 \\ b+d=0 \\ bd=1 \end{cases} \iff \begin{cases} a=c=0 \\ b=-d \\ b^2=-1 \end{cases}$$

  В частности, если уравнение $x^2=-1$ разрешимо, то мы нашли разложение.

- Случай $b=d$

  $$\begin{cases} b=d \\ a=-c \\ a^2=2b \\ b^2=1 \end{cases}$$

  Уравнение $b^2=1$ всегда разрешимо и имеет решения 1 и -1 $(p-1)$.

  Проверим оба.

- $b=d=1$

  $$\begin{cases} b=d=1 \\ a=-c \\ a^2=2 \end{cases}$$

  Мы нашли разложение если $x^2=2$ разрешимо.

- $b=d=-1$

  $$\begin{cases} b=d=-1 \\ a=-c \\ a^2=-2 \end{cases}$$

  Мы нашли разложение если $x^2=-2$ разрешимо.

Заметим, что хотя бы одно из трёх уравнений разрешимо, действительно разрешимость первого задаёт символ лежандра $\left(\frac{-1}{p}\right)$, второго $\left(\frac{2}{p}\right)$, и третьего $\left(\frac{-2}{p}\right)$.

Поэтому если даже первые два уравнения не разрешимы (символы равны -1), то символ последнего по мультипликативности получается $+1$ и уравнение разрешимо. □

Замечание. Соответственно частичное разложение x^4+1 для $p \geqslant 3$ выглядит одним из следующих образов:

- $x^4+1=(x^2+\sqrt{-1})(x^2-\sqrt{-1})$
- $x^4+1=(x^2+\sqrt{2}x+1)(x^2-\sqrt{2}x+1)$
- $x^4+1=(x^2+\sqrt{-2}x-1)(x^2-\sqrt{-2}x-1)$

12. Факториальные кольца и многочлены

12.1. Факториальность кольца многочленов над факториальным кольцом

***Определение* 12.1.**

$f = a_0 + a_1x + \ldots + a_nx^n \in R[x]$

Если $d(f) = \gcd(a_0, a_1, \ldots, a_n) = 1$, то f называется примитивным.

Замечание.

Вспомним определение ассоциированности.

А именно, пусть $a,\ b \in R$, где R — область целостности.

Тогда элементы a и b ассоциированны (или $a \sim b$), если $a = \varepsilon b,\ \varepsilon \in R^*$.

Замечание.

В евклидовых кольцах gcd определён с точностью до делителей единицы.

Лемма.

$f,\ g \in R[x]$

$$\left.\begin{array}{l}\alpha f = \beta g \\ d(f) = d(g) = 1 \\ \alpha,\ \beta \in R \setminus \{0\}\end{array}\right\} \implies \begin{array}{l}\alpha \sim \beta \text{ в } R \\ f \sim g \text{ в } R[x]\end{array}$$

Доказательство.

Покажем, что $\alpha \sim \beta$:

$f = a_0 + a_1x + \ldots + a_nx^n,\ \ g = b_0 + b_1x + \ldots + b_mx^m$

$$\alpha f = \beta g,\ \ \alpha,\ \beta \in R \setminus \{0\} \implies \begin{cases} n = m \\ \alpha a_i = \beta b_i \quad \forall\, i = 0\,..\,n \end{cases}$$

$\gcd(a_0, a_1, \ldots, a_n) = 1,\ \ \gcd(b_0, b_1, \ldots, b_n) = 1$

$\gcd(\alpha a_0, \alpha a_1, \ldots, \alpha a_n) = \alpha \sim \beta = \gcd(\beta b_0, \beta b_1, \ldots, \beta b_n)$

Покажем теперь, что $f \sim g$:

$\alpha = \varepsilon\beta,\ \varepsilon \in R^*$

$$\left.\begin{array}{l}\alpha f = \varepsilon\beta f,\ \ \alpha f = \beta g \\ \beta \neq 0\end{array}\right\} \implies \varepsilon\beta f - \beta g = \beta(\varepsilon f - g) = 0 \implies \varepsilon f = g, \text{ т.е. } f \sim g \qquad \square$$

Теорема 12.1.

R — факториально $\implies$ $R[x]$ — факториально.

Доказательство.

$f \in R[x],\ f \neq 0$

Покажем существование разложения на неприводимые. Для этого воспользуемся индукцией по $\deg f$.

База: $\deg f = 0,\ \ f(x) = c \in R \setminus \{0\} \implies$ существует разложение на неприводимые.

Переход: $\deg f > 0$.

$f = d(f) \cdot f_1$, где $d(f) \in R \setminus \{0\}$

Если f_1 неприводим, то уже получили искомое разложение.

Иначе же получаем, что существует представление $f_1 = g_1 \cdot h_1$, где g_1 и h_1 — раскладываются на неприводимые по предположению индукции.

Покажем теперь единственность разложения.

$f = p_1 p_2 \dots p_k \cdot q_1 q_2 \dots q_m = p'_1 p'_2 \dots p'_l \cdot q'_1 q'_2 \dots q'_n$, где $p_i, p'_i \in R$ — неприводимые коэффициенты; $q_i, q'_i \in R[x]$ — неприводимые многочлены

По доказанной выше лемме, $p_1 p_2 \dots p_k \sim p'_1 p'_2 \dots p'_l$, $q_1 q_2 \dots q_m \sim q'_1 q'_2 \dots q'_n$.

Уже знаем, что разложение над факториальным кольцом на неприводимые множители единственно. То есть, $k = l$, и существует такая перестановка $\tau_1 \in S_k$, что $p_i \sim p'_{\tau_1(i)}$.

Ещё нам известно, что q_i и q'_i — неприводимы в $R[x]$. Следовательно, по следствию из леммы Гаусса, они также неприводимы и в $(\operatorname{Quot} R)[x]$ (которое факториально) $\implies n = m$ и $q_i \sim q'_{\tau_2(i)}$ для некоторой перестановки $\tau_2 \in S_n$.

Таким образом, показали, что $R[x]$ — факториально. □

12.2. Многочлены от многих переменных

***Определение* 12.2.**

Многочленом от многих переменных называется конечная сумма мономов вида

$f(x_1, x_2, \dots, x_n) = \sum a_{k_1,k_2,\dots,k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$

***Определение* 12.3.**

$R[x_1, x_2] := R[x_1][x_2]$

$R[x_1, x_2, \dots, x_n] := R[x_1, x_2, \dots, x_{n-1}][x_n]$

В частности, если R — факториально, то и $R[x_1, x_2, \dots, x_n]$ также будет факториально.

***Определение* 12.4.**

Назовём многочлен однородным степени d, если в нём нет мономов степени, отличной от d.

Замечание.

1. Пусть R — поле.

   Однородные многочлены степени d образуют подпространство $R^{(d)}[x_1, x_2, \dots, x_n]$ размерности $\binom{n+d-1}{d}$.

2. $R[x_1, x_2, \dots, x_n] = \bigoplus\limits_{d \geqslant 0} R^{(d)}[x_1, x_2, \dots, x_n]$.

Обозначение.

R — область целостности.

$f,\ g \in R[x]$ — многочлены.

$f = g$, если они совпадают покоэффициентно (формальное равенство).

$f \equiv g$, если $\forall\, \lambda \in R : f(\lambda) = g(\lambda)$ (функциональное равенство).

Теорема 12.2.

R — поле, $|R| = \infty$.

$f,\ g \in R[x]$ — многочлены.

Тогда $f \equiv g \iff f = g$.

Доказательство.

Положим $f(x) = a_0 + a_1x + \ldots + a_nx^n,\ g(x) = b_0 + b_1x + \ldots + b_mx^m$.

"$\Longleftarrow$" Пусть $f = g$.

Тогда $n = m$ и $\forall\, k = 0..n : a_k = b_k$.

Очевидно, что в таком случае $\forall\, \lambda \in R : f(\lambda) = g(\lambda)$, то есть $f \equiv g$.

"$\Longrightarrow$" Пусть $f \equiv g$.

Рассмотрим многочлен $h = f - g \in R[x]$.

Положим $h = c_0 + c_1x + \ldots + c_lx^l$, где $l = \max\{n, m\}$.

Мы знаем, что $\forall\, \lambda \in R : f(\lambda) = g(\lambda) \implies \forall\, \lambda \in R : h(\lambda) = f(\lambda) - g(\lambda) = 0$.

Вспомним, что любой отличный от нуля многочлен степени l не может иметь более l корней.

Но многочлен h имеет бесконечное количество корней. Следовательно, он тождественно равен нулю.

Таким образом, $h = 0 \implies f - g = 0 \implies f = g$. □

Следствие.

R — поле, $|R| = \infty$.

$f,\ g \in R[x_1, x_2, \ldots, x_n]$ — многочлены.

Тогда $f \equiv g \iff f = g$.

Доказательство.

Положим

$$f(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{s} a_k\, x_1^{p_{k,1}} x_2^{p_{k,2}} \ldots x_n^{p_{k,n}}$$

$$g(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{t} b_k\, x_1^{q_{k,1}} x_2^{q_{k,2}} \ldots x_n^{q_{k,n}}$$

Если $f = g$, то функциональное равенство очевидно. Потому покажем следствие в другую сторону.

Воспользуемся индукцией по числу переменных.

Для одной переменной всё знаем. Рассмотрим случай n переменных, где $n > 1$.

Рассмотрим переменную x_n в многочленах f и g. Разобьём мономы f и g на группы в зависимости от того, в какой степени входит x_n в тот или иной моном.

Пусть m — максимальная из степеней x_n в многочленах f и g. Таким образом, можем получить следующее представление данных многочленов:

$$f(x_1, x_2, \ldots, x_n) = \sum_{k=0}^{m} f_k(x_1, x_2, \ldots, x_{n-1})\, x_n^k$$

$$g(x_1, x_2, \ldots, x_n) = \sum_{k=0}^{m} g_k(x_1, x_2, \ldots, x_{n-1})\, x_n^k$$

Мы знаем, что такие многочлены функционально равны.

Следовательно, $\forall\, x_1, x_2, \ldots x_{n-1} \in R \quad \forall k = 0..m : f_k(x_1, x_2, \ldots x_{n-1}) = g_k(x_1, x_2, \ldots, x_{n-1})$.

Однако, по предположению индукции, $f_k \equiv g_k \implies f_k = g_k$.

Значит, $f \equiv g \implies f = g$. □

12.3. Теорема Гильберта о базисе

Теорема 12.3.

Пусть в R любой идеал конечнопорождён.

Тогда и в $R[x_1, x_2, \ldots, x_n]$ — тоже.

Доказательство.

Достаточно доказать для $R[x]$.

Пусть I — произвольный идеал, $I \subseteq R[x]$.

Рассмотрим множество старших коэффициентов многочленов из I. Обозначим его за J.

Заметим, что J — идеал в R. Действительно:

$f(x) = ax^n + \ldots \in I,\ g(x) = bx^m + \ldots \in I \implies$

$\implies x^m f(x) + x^n g(x) = (a+b)x^{n+m} + \ldots \in I \implies a + b \in J$

$f(x) = ax^n + \ldots \in I \implies cf(x) = cax^n + \ldots \in I \implies ca \in J$

По предположению идеал J конечнопорождён в R. Значит, существует некоторый набор его образующих $a_1, a_2, \ldots, a_s$.

Рассмотрим произвольный набор $f_1, f_2, \ldots, f_s \in I$, такой что $f_i(x) = a_i x^{m_i} + \ldots$

Положим $n = \max\limits_{i=1..s} \deg f_i$.

Если $n = 0$, то заметим, что $a_1, a_2, \ldots, a_s$ также являются образующими и идеала I. А потому будем рассматривать случай $n > 0$.

Покажем, что $\forall\, f \in I\ \ \exists\, \lambda_1, \lambda_2, \ldots, \lambda_s \in R[x],\ g \in I,\ \deg g < n : f(x) = g(x) + \sum\limits_{i=1..s} \lambda_i f_i$.

Воспользуемся индукцией по $\deg f$.

Если $\deg f < n$, то можем выбрать $g = f$ и $\lambda_i = 0$.

Пусть $\deg f \geqslant n,\ f(x) = b_N x^N + \ldots$

$b_N \in J \implies b_N = \sum\limits_{i=1..s} \alpha_i a_i,\ \alpha_i \in R$

Тогда выберем $g(x) = f(x) - \sum\limits_{i=1..s} \alpha_i x^{N-m_i} f_i(x)$ и $\lambda_i = \alpha_i x^{N-m_i} \implies \deg g < N$. И воспользуемся предположением индукции.

Осталось показать, что $\exists$ конечный набор многочленов $h_i \in I$, что $\forall\, g \in I$, $\deg g < n$ представим в виде линейной комбинации h_i.

Положим J_{n-1} — идеал, состоящий из старших членов коэффициентов многочленов из I степени $n-1$. Любой такой идеал конечнопорождён $\implies J_{n-1} = \langle b_1, b_2, \ldots, b_t \rangle$.

Следовательно, можем выбрать такой набор $g_1, g_2, \ldots, g_t \in I$, что $g_i(x) = b_i x^{n-1} + \ldots$

Осталось заметить, что $\forall\, g \in I,\ \deg g = n-1$ существует представление в виде

$g(x) = \widetilde{g}(x) + \sum\limits_{i=1..t} \mu_i g_i(x)$, где $\mu_i \in R$ и $\deg \widetilde{g}(x) < n-1$

Получили сведение задачи от степени $n-1$ к степени $n-2$.

Построив теперь аналогичную конструкцию для многочленов степени $n-2, n-3, \ldots, 0$, найдём оставшиеся образующие.

□

13. Базис Грёбнера и симметрические многочлены

13.1. Базис Грёбнера

Мотивация.

Определим задачу вхождения:

Пусть идеал $I \lhd K[x_1, x_2, \ldots, x_n]$ задан своим базисом $I = \langle f_1, f_2, \ldots, f_k \rangle$.

Требуется найти алгоритм, позволяющий за конечное число шагов выяснить, принадлежит ли заданный многочлен h идеалу I.

В редких случаях сделать это бывает просто (к примеру, если $I = \langle f \rangle$, где $f \in K[x]$). В общем же случае эта задача не так тривиальна.

Это и является предпосылкой для введения базиса Грёбнера.

Замечание.

Для нахождения базиса Грёбнера идеала $I \lhd K[x_1, x_2, \ldots, x_n]$ нам потребуется определить для многочлена $f(x_1, x_2, \ldots, x_n)$ понятие старшего члена.

Существует несколько способов однозначного определения старшего члена многочлена. Далее будет приведён только один такой способ, называемый *лексикографическим*.

***Определение* 13.1.**

Многочлен $f = ax_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}, a \in K$ называется *одночленом* или *мономом*.

Любому такому моному можно сопоставить набор $(k_1, k_2, \ldots, k_n)$ целых неотрицательных чисел, который называют *набором степеней*.

***Определение* 13.2.**

Будем говорит, что набор $(k_1, k_2, \ldots, k_n)$ больше набора $(l_1, l_2, \ldots, l_n)$, если существует такой индекс i, что $k_1 = l_1,\ k_2 = l_2,\ \ldots,\ k_{i-1} = l_{i-1},\ k_i > l_i$.

Заметим, что любые два набора степеней одинаковой длины сравнимы подобным образом.

***Определение* 13.3.**

Положим многочлен $f = \sum_i a_{i_1, i_2, \ldots, i_n} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$.

Определим старший член $\operatorname{lt} f$ как наибольший моном многочлена f.

Лемма (о старшем члене).

Старший член произведения многочленов есть произведение их старших членов.

То есть, $\operatorname{lt} f \cdot g = \operatorname{lt} f \cdot \operatorname{lt} g$.

Доказательство.

Пусть старшим членам $\operatorname{lt} f$ и $\operatorname{lt} g$ соответствуют наборы степеней $(k_1, k_2, \ldots, k_n)$ и $(l_1, l_2, \ldots, l_n)$.

Положим s и t — некоторые мономы f и g, не равные $\operatorname{lt} f$ и $\operatorname{lt} g$ одновременно. Пусть им соответствуют наборы степеней $(k'_1, k'_2, \ldots, k'_n)$ и $(l'_1, l'_2, \ldots, l'_n)$.

Для определённости, положим, что $\operatorname{lt} f > s$ и $\operatorname{lt} g \geqslant t$.

Заметим, что моному $\operatorname{lt} f \cdot \operatorname{lt} g$ будет соответствовать набор степеней $(k_1 + l_1, k_2 + l_2, \ldots, k_n + l_n)$, а моному $s \cdot t$ — набор степеней $(k'_1 + l'_1, k'_2 + l'_2, \ldots, k'_n + l'_n)$.

Но так как мы положили $(k_1, k_2, \ldots, k_n) > (k'_1, k'_2, \ldots, k'_n)$ и $(l_1, l_2, \ldots, l_n) \geqslant (l'_1, l'_2, \ldots, l'_n)$, то набор, соответствующий $\operatorname{lt} f \cdot \operatorname{lt} g$, будет строго больше набора, соответствующего $s \cdot t$.

Следовательно, $\operatorname{lt} fg = \operatorname{lt} f \cdot \operatorname{lt} g$. □

***Определение* 13.4.**

Пусть $I = \langle f_1, f_2, \ldots f_k \rangle \lhd K[x_1, x_2, \ldots, x_n]$.

Определим для многочлена $h \in K[x_1, x_2, \ldots, x_n]$ операцию редукции.

Пусть для некоторого многочлена f_i верно, что $\operatorname{lt} h \vdots \operatorname{lt} f_i$. То есть, $\operatorname{lt} h = q \cdot \operatorname{lt} f_i$, где q — некоторый моном.

Тогда положим $h_r = h - q \cdot f_i$, при этом $\operatorname{lt} h_r < \operatorname{lt} h$.

Лемма.

$I = \langle f_1, f_2, \ldots, f_k \rangle \lhd K[x_1, x_2, \ldots, x_n],\ h \in K[x_1, x_2, \ldots, x_n]$

Тогда $h \in I \iff h_r \in I$, где h_r — многочлен, полученный из h при помощью редукции.

Доказательство.

Достаточно показать, что если $h \in I$, то $h - h_r \in I$.

Однако заметим, что $h - h_r = q \cdot f_i$ для некоторого i, а $q \cdot f_i \in I$ (f_i является одним из порождающих элементов идеала I). □

Следствие.

Пусть $I = \langle f_1, f_2, \ldots, f_k \rangle$.

Тогда если многочлен h редуцируется к нулю при помощи $f_1, f_2, \ldots, f_n$, то $h \in I$.

***Определение* 13.5.**

Набор порождающих $f_1, f_2, \ldots, f_k$ идеала $I = \langle f_1, f_2, \ldots, f_k \rangle$ называется *базисом Грёбнера*, если любой многочлен $h \in I$ редуцируется к нулю при помощи $f_1, f_2, \ldots, f_k$.

Теорема 13.1.

Пусть $I = \langle f_1, f_2, \ldots, f_k \rangle$.

Тогда следующие условия эквивалентны:

1. $\forall\, h \in I \quad \exists\, i :\ \operatorname{lt} h \vdots \operatorname{lt} f_i$
2. Любой многочлен $h \in I$ редуцируется к нулю.
3. $h \in I \iff \begin{cases} h = \sum\limits_{i=1}^{k} f_i g_i \\ \operatorname{lt} h = \max\limits_{i=1}^{k} \operatorname{lt} f_i \cdot \operatorname{lt} g_i \end{cases}$
4. Пусть L — идеал, порождённый старшими членами многочленов из I.

   Тогда $L = \langle \operatorname{lt} f_1, \operatorname{lt} f_2, \ldots, \operatorname{lt} f_k \rangle$.

Доказательство.

“1 $\implies$ 2”

Заметим, что, т.к. $\operatorname{lt} h \vdots \operatorname{lt} f_i$ для некоторого i, то мы можем редуцировать наш многочлен h, избавившись от его старшего члена и перейдя к многочлену h_r, такому что $\operatorname{lt} h_r < \operatorname{lt} h$.

Для редуцированного многочлена h_r так же будет верно, что $\operatorname{lt} h_r \vdots \operatorname{lt} f_i$ для некоторого i. То есть, сможем повторить редукцию.

Осталось лишь показать, что можно редуцировать многочлен h к нулю за конечное число действий.

Для этого рассмотрим следующую индукцию по числу переменных:

Рассмотрим кольцо многочленов от одной переменной и некоторый многочлен $h \in I$, его старшему члену $\operatorname{lt} h$ соответствует набор степеней (p_1).

Поскольку при редукции мы обязательно перейдём к меньшему набору, то за p_1 действие мы точно редуцируем его к нулю.

Рассмотрим теперь кольцо многочленов от $n > 1$ переменных. Пусть старшему члену $\operatorname{lt} h$ некоторого многочлена $h \in I$ соответствует набор переменных $(p_1, p_2, \ldots, p_n)$.

Так как набор степеней $(p_2, \ldots, p_n)$ за конечное количество действий можно свести к нулю (по предположению индукции), то мы за конечное число действий сможем уменьшить p_1 на единичку и перейти к набору $(p_1 - 1, p'_2, \ldots, p'_n)$.

Следовательно, мы за конечное количество действий сможем редуцировать к нулю степень p_1, а затем и все оставшиеся степени.

"2 $\implies$ 3"

Будем строить сумму в процессе редукции: когда редуцируем h с помощью f_i, в соответствующем ему многочлене g_i появляется моном вида $\frac{\operatorname{lt} h}{\operatorname{lt} f_i}$.

Заметим, что на первом же шаге редукции мы избавимся от старшего члена h. А значит, этот старший член в полученной нами сумме встретится всего один раз и будет наибольшим.

Значит, в силу леммы о старшем члене, получаем, что $\operatorname{lt} h = \max\limits_{i=1}^{k} \operatorname{lt} f_i \cdot \operatorname{lt} g_i$

"3 $\implies$ 4"

Пусть $h = \sum\limits_{i=1}^{k} f_i g_i$ и $\operatorname{lt} h = \max\limits_{i=1}^{k} \operatorname{lt} f_i \cdot \operatorname{lt} g_i$.

Так как L — идеал, порождённый старшими членами многочленов из I, то он также содержит и старшие члены f_i.

Однако $\operatorname{lt} h = \operatorname{lt} f_i \cdot \operatorname{lt} g_i$ для некоторого i, и это верно для любого $h \in I$.

Следовательно, $\operatorname{lt} h \in \langle \operatorname{lt} f_1, \operatorname{lt} f_2, \ldots, \operatorname{lt} f_k \rangle$.

Получили, что $L = \langle \operatorname{lt} f_1, \operatorname{lt} f_2, \ldots, \operatorname{lt} f_k \rangle$.

"4 $\implies$ 1"

Так как $\operatorname{lt} h \in L = \langle \operatorname{lt} f_1, \operatorname{lt} f_2, \ldots, \operatorname{lt} f_k \rangle$, то старший член h представим в виде линейной комбинации старших членов $f_1, f_2, \ldots, f_k$. А именно, $\operatorname{lt} h = \sum_i \alpha_i x^{k_i} \operatorname{lt} f_i$

Осталось лишь заметить, что степени всех мономов в данной сумме совпадают и равны степени $\operatorname{lt} h$.

Откуда можно сделать вывод, что для произвольного $\operatorname{lt} f_i$ из данной суммы будет верно, что $\operatorname{lt} h \vdots \operatorname{lt} f_i$. □

Теорема 13.2 (о существовании базиса Грёбнера).
В любом идеале $I \lhd K[x_1, x_2, \ldots, x_n]$ существует базис Грёбнера.

Доказательство.
Рассмотрим идеал $L = \langle \operatorname{lt} f | f \in I \rangle \lhd K[x_1, x_2, \ldots, x_n]$.

Заметим, что по теореме Гильберта о базисе, в L существует конечный набор порождающих. Обозначим их за $g_1, g_2, \ldots, g_k$.

Пусть $g_{i,j}$ — j-ый моном многочлена g_i. Здесь считаем, что мономы отсортированы в обратном

лексикографическом порядке (то есть, $g_{i,1} = \operatorname{lt} g_i,\ g_{i,2} = \operatorname{lt}(g_i - \operatorname{lt} g_i)$ и т.д.).

Покажем, что любой такой моном лежит в L.

Для этого покажем, что если g_i лежит в L, то $g_i - \operatorname{lt} g_i$ также лежит в L.

Действительно. Поскольку $g_i \in L$, то он представим в виде линейной комбинации старших членов многочленов из I. В частности, получаем, что $\operatorname{lt} g_i \in L$. Следовательно, их разность лежит в L.

Следовательно, в L лежат все многочлены вида $g_i,\ g_i - g_{i,1},\ g_i - g_{i,1} - g_{i,2},\ \ldots$

Обозначим их за $h_{i,1},\ h_{i,2},\ h_{i,3},\ \ldots$

Осталось лишь заметить, что т.к. любой многочлен $h_{i,j} \in L$, то $g_{i,j} = h_{i,j} - h_{i,j+1} \in L$.

Таким образом, поскольку мономы $g_{i,j}$ лежат в L (и, более того, порождают его), то в I найдутся многочлены $f_1, \ldots, f_s$, такие что $\operatorname{lt} f_i = g_{j,k}$ для некоторых $j,\ k$.

Эти многочлены и будут образовывать базис Грёбнера (по 4 определению). □

13.2. Алгоритм Бухбергера

Замечание.

***Определение* 13.6.**
Пусть $I \lhd K[x_1, x_2, \ldots, x_n]$ — идеал, и $f_1, f_2, \ldots, f_k$ — его базис.

Говорят, что многочлены f_i и f_j имеют *зацепление*, если их старшие члены $\operatorname{lt} f_i$ и $\operatorname{lt} f_j$ делятся одновременно на какой-то многочлен w, отличный от константы.

***Определение* 13.7.**
Пусть многочлены f_i и f_j имеют зацепление. При этом $\operatorname{lt} f_i = a \cdot x^\alpha,\ \operatorname{lt} f_j = b \cdot x^\beta$.

Определим S-многочлен пары f_i и f_j следующим образом:

Тогда $S(f, g) := \frac{x^\gamma}{ax^\alpha} \cdot f_i - \frac{x^\gamma}{bx^\beta} \cdot f_j$, где $x^\gamma = \operatorname{lcm}(x^\alpha, x^\beta)$.

***Определение* 13.8.**
Говорят, что зацепление f_i и f_j *разрешимо*, если многочлен $S(f_i, f_j)$ редуцируется к нулю с помощью $f_1, f_2, \ldots, f_k$.

Теорема 13.3.
Базис $f_1, f_2, \ldots, f_k$ идеала $I = \langle f_1, f_2, \ldots, f_k \rangle$ является базисом Грёбнера $\iff \forall\, i, j : S(f_i, f_j)$ редуцируется к нулю с помощью $f_1, f_2, \ldots, f_k$.

Для доказательства теоремы нам потребуется следующая лемма:

Лемма.
Пусть $f_1, f_2, \ldots, f_s \in K[x_1, x_2, \ldots, x_n],\ \operatorname{lt} f_i = a_i x^\alpha$.

Положим $h = \sum\limits_{i=1}^{s} \lambda_i f_i,\ \lambda_i \in K$.

Тогда если $\operatorname{lt} h < x^\alpha$, то имеет место равенство $h = \sum\limits_{i<j} \mu_{i,j} S(f_i, f_j),\ \mu_{i,j} \in K$.

Доказательство.
Заметим, что, поскольку степени старших членов всех f_i совпадают, то $S(f_i, f_j) = \frac{f_i}{a_i} - \frac{f_j}{a_j}$.

Запишем теперь нашу сумму следующим образом:

$$h = \sum_{i=1}^{s} \lambda_i f_i = \lambda_1 a_1 \left(\frac{f_1}{a_1} - \frac{f_2}{a_2}\right) + (\lambda_1 a_1 + \lambda_2 a_2)\left(\frac{f_2}{a_2} - \frac{f_3}{a_3}\right) + \ldots + (\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_s a_s)\left(\frac{f_s}{a_s}\right)$$

Поймём, что $\lambda_1 a_1 + \lambda_2 a_2 + \dots \lambda_s a_s = 0$.

Действительно. Поскольку $\operatorname{lt}\left(\frac{f_i}{a_i} - \frac{f_{i+1}}{a_{i+1}}\right) < x^\alpha$, а мы предположили, что старший член всей суммы меньше x^α, то коэффициент при $\frac{f_s}{a_s}$ должен быть равен нулю.

Таким образом, получили, что $h = \sum\limits_{i=1}^{s-1}(\lambda_1 a_1 + \dots + \lambda_i a_i) \cdot S(f_i, f_{i+1})$. □

Доказательство. (**теоремы**)

“$\Longrightarrow$”

Покажем, что если $f_1, f_2, \dots, f_k$ являются базисом Грёбнера идеала I, то $\forall\, i, j : S(f_i, f_j)$ редуцируется к нулю.

Действительно, поскольку $f_i, f_j \in I$, то любая их линейная комбинация также лежит в I. В частности, $S(f_i, f_j) \in I$.

Но поскольку $f_1, f_2, \dots, f_k$ — базис Грёбнера, то любой элемент из идеала редуцируется к нулю.

“$\Longleftarrow$”

Вспомним следующее определение базиса Грёбнера:

$$f_1, f_2, \dots, f_k \text{ — базис Грёбнера, если } h \in I \iff \begin{cases} h = \sum\limits_{i=1}^{k} f_k g_k \\ \operatorname{lt} h = \max\limits_{i=1}^{k} \operatorname{lt} f_i \cdot \operatorname{lt} g_i \end{cases}$$

Пусть $\forall\, i, j : S(f_i, f_j)$ редуцируется к нулю, и некоторый многочлен $h = \sum\limits_{i=1}^{k} f_i g_i$.

Положим $\operatorname{lt} h = a x^\alpha,\ \operatorname{lt} f_i = b_i x^{\beta_i},\ \operatorname{lt} g_i = c_i x^{\gamma_i}$.

Определим тогда $x^\delta := \max\limits_{i=1}^{k}\ x^{\beta_i} \cdot x^{\gamma_i}$, и среди всех разложений h выберем такое, в котором x^δ минимально.

Заметим, что при этом всегда верно, что $x^\delta \geqslant x^\alpha$. И если $x^\delta = x^\alpha$, то заданный набор является базисом Грёбнера.

Предположим, что $x^\delta > x^\alpha$.

Перегруппируем слагаемые в нашей сумме таким образом, что $x^{\beta_i} \cdot x^{\gamma_i} = x^\delta$ при $i = 1 \,..\, m$ и $x^{\beta_i} \cdot x^{\gamma_i} < x^\delta$ при $i > m$.

Теперь наш многочлен h представим следующим образом:

$$h = \sum_{i=1}^{m} c_i x^{\gamma_i} f_i + \sum_{i=1}^{m} (g_i - c_i x^{\gamma_i}) f_i + \sum_{i=m+1}^{k} f_i g_i.$$

Положим $f = \sum\limits_{i=1}^{m} c_i x^{\gamma_i} f_i$.

Заметим, что $\operatorname{lt} f < x^\delta$, поскольку старший член всей нашей исходной суммы равен $x^\alpha < x^\delta$. Значит, наш многочлен f попадает под условие доказанной выше леммы. Получаем:

$$f = \sum_{i=1}^{m} c_i x^{\gamma_i} f_i = \sum_{1 \leqslant i < j \leqslant m} \mu_{i,j} S(x^{\gamma_i} f_i, x^{\gamma_j} f_j)$$

Теперь заметим, что $S(x^{\gamma_i} f_i, x^{\gamma_j} f_j) \vdots S(f_i, f_j)$, а $S(f_i, f_j)$ редуцируется к нулю:

$$S(x^{\gamma_i} f_i, x^{\gamma_j} f_j) = \frac{x^\delta}{b_i x^{\beta_i + \gamma_i}} x^{\gamma_i} f_i - \frac{x^\delta}{b_j x^{\beta_j + \gamma_j}} x^{\gamma_j} f_j = \frac{x^\delta}{b_i x^{\beta_i}} f_i - \frac{x^\delta}{b_j x^{\beta_j}} f_j = \frac{x^\delta}{\operatorname{lcm}(x^{\beta_i}, x^{\beta_j})} \cdot S(f_i, f_j)$$

Следовательно, $S(x^{\gamma_i} f_i, x^{\gamma_j} f_j)$ также редуцируется к нулю. А поскольку его старший член

меньше x^δ, то он представляется в виде линейной комбинации f_i, каждый старший член которой меньше x^δ.

Поскольку эту верно для любого $S(f_i, f_j)$, то и весь многочлен g можно представить в виде линейной комбинации f_i, каждый старший член которой меньше x^δ.

Значит, получили такое разложение h, в котором $\max\limits_{i=1}^{k} \operatorname{lt} f_i \cdot \operatorname{lt} g_i < x^\delta$.

Противоречие. □

Теорема 13.4.
Если многочлены f и g не зацеплены, то $S(f, g)$ редуцируется к нулю.

Доказательство.
Рассмотрим идеал $I = \langle f, g \rangle$. Покажем, что f и g задают базис Грёбнера данного идеала.

То есть, $\forall\, h = af + bg : \operatorname{lt} h \vdots \operatorname{lt} f \cup \operatorname{lt} h \vdots g$.

Для определённости рассмотрим такое разложение h, что степень многочлена a в нём минимальная возможная.

Пусть старшие члены многочленов af и bg в представлении h не сокращаются. Тогда старший член h делится на старший член одного из многочленов f или g.

Пусть же это не так. То есть, $\operatorname{lt} af + \operatorname{lt} bg = 0 \implies \operatorname{lt} a \cdot \operatorname{lt} f = -\operatorname{lt} b \cdot \operatorname{lt} g$.

Так как $\operatorname{lt} f$ и $\operatorname{lt} g$ взаимнопросты, $\operatorname{lt} a = \operatorname{lt} g \cdot w$ и $\operatorname{lt} b = -\operatorname{lt} f \cdot \operatorname{lt} w$ для некоторого монома w. Тогда получаем:

$h = af + bg = f(gw + a_1) + g(-fw + b_1) = fa_1 - gb_1 \implies \operatorname{lt} a_1 < \operatorname{lt} a$

Получили противоречие. □

Алгоритм.
Пусть $f_1, f_2, \ldots, f_k$ — набор многочленов, являющихся базисом идеала I.

Опишем теперь сам алгоритм нахождения базиса Грёбнера:

1. Проверим, есть ли в наборе зацепления.

   Если нет, то набор является базисом Грёбнера. Иначе переходим к пункту 2.

2. По найденному зацеплению многочленов f_i и f_j составим многочлен $S(f_i, f_j)$ и при помощи последовательности редукций сведём его к нередуцируемому многочлену $\widetilde{f}$.

   Если $\widetilde{f} \equiv 0$, то рассмотрим другое, не рассмотренное ранее зацепление. Иначе добавим $\widetilde{f}$ к базису и повторим алгоритм.

3. Если все имеющиеся зацепления ранее рассматривались, то наш текущий набор — базис Грёбнера.

Замечание.
У нас есть алгоритм. Осталось показать, что он конечен.

Для этого докажем следующую теорему:

Теорема 13.5.
Для каждого набора $f_1, f_2, \ldots, f_k \in K[x_1, x_2, \ldots, x_n]$ после редуцирования конечного числа зацеплений мы получим набор $f_1, \ldots, f_k, f_{k+1}, \ldots f_m$, в котором каждое зацепление разрешимо.

Доказательство.

Будем рассуждать от противного.

Пусть при редуцировании многочленов $S(f_i, f_j)$ возникает бесконечно много нередуцируемых многочленов.

Рассмотрим идеал, порождённый их старшими членами. По теореме Гильберта о базисе в нём существует конечный базис из числа образующих.

Тогда у любого последующего многочлена старший член будет делиться на старший член одного из "базисных" многочленов. И, следовательно, этот многочлен можно редуцировать.

Значит, построенный базис будет конечен. □

13.3. Приведённый базис Грёбнера

Замечание.

Заметим, что в общем случае полученный с помощью описанного алгоритма базис Грёбнера можно упростить, избавившись от каких-то его элементов.

А именно, рассмотрим первое упрощение:

Пусть f_1 и f_2 — элементы базиса Грёбнера, такие что $\operatorname{lt} f_1 \vdots \operatorname{lt} f_2$. Тогда можно избавиться от элемента f_1 в нашем базисе.

Можно заметить, что оставшиеся элементы по-прежнему будут задавать базис.

***Определение* 13.9.**

Базис Грёбнера $f_1, f_2, \ldots, f_k$ называется минимальным, если $\operatorname{lt} f_i$ не делится на $\operatorname{lt} f_j$ при $i \neq j$.

Теорема 13.6.

Минимальный базис Грёбнера идеала I существует и определён однозначно (с точностью до равенства старших членов).

Доказательство.

Существование минимального базиса следует из алгоритма его получения из произвольного базиса Грёбнера.

Покажем единственность.

Пусть $f_1, f_2, \ldots, f_n$ и $g_1, g_2, \ldots, g_m$ — минимальные базисы Грёбнера.

Рассмотрим элемент f_i. Заметим, что $\operatorname{lt} f_i \vdots \operatorname{lt} g_j$ для некоторого элемента j (в силу того, что многочлены g образуют базис Грёбнера).

В то же время, $\operatorname{lt} g_j \vdots \operatorname{lt} f_k$ для некоторого k.

Значит, $f_i \vdots f_k \implies i = k$ в силу минимальности базиса f.

Из всего этого можем сделать вывод, что $n = m$ и $\forall\, i \;\; \exists\, j : \operatorname{lt} f_i = \operatorname{lt} g_j$. □

Замечание.

Введём второе упрощение.

А именно, пусть f_1 и f_2 — многочлены базиса Грёбнера, и некоторый моном f_1 делится на старший член f_2. Тогда мы можем редуцировать этот моном с помощью f_2.

***Определение* 13.10.**

Базис $f_1, f_2, \ldots, f_k$ называется приведённым, ни один член многочлена f_i не делится на старший член многочлена f_j для всех $i \neq j$.

Иными словами, f_i равен остатку от деления f_i на многочлены $f_1, \ldots, f_{i-1}, f_{i+1}, \ldots f_k$.

Замечание.

Приведённый базис Грёбнера является минимальным.

Теорема 13.7.

Приведённый базис Грёбнера идеала I существует и определён однозначно.

Доказательство.

Пусть $f_1, f_2, \ldots, f_k$ — минимальный базис Грёбнера.

Покажем существование приведённого базиса. Для этого положим:

h_1 — редукция f_1 по модулю $f_2, f_3, \ldots, f_k$

h_2 — редукция f_2 по модулю $h_1, f_3, \ldots, f_k$

И так далее...

В силу редукции ни один из мономов многочлена h_i не будет делиться на старший член любого из многочленов $h_1, \ldots h_{i-1}, h_{i+1}, \ldots h_k$.

Докажем единственность.

Пусть $f_1, f_2, \ldots, f_k$ и $g_1, g_2, \ldots, g_k$ — приведённые базисы Грёбнера. При этом $\operatorname{lt} f_1 = \operatorname{lt} g_1,\ \operatorname{lt} f_2 = \operatorname{lt} g_2, \ldots, \operatorname{lt} f_k = \operatorname{lt} g_k$.

Предположим, что многочлен $h_i := f_i - g_i \neq 0$.

Поскольку $h_i \in I$, то $\operatorname{lt} h_i$ делится на некоторый старший член $\operatorname{lt} g_j$ и $\operatorname{lt} f_k$.

С другой стороны, $\operatorname{lt} h_i$ является одним из нестарших членов многочленов f_i или g_i. А значит, его можно было бы редуцировать.

Получили противоречие $\square$

13.4. Симметрические многочлены

***Определение* 13.11.** Многочлен $f \in K[x_1, \dots, x_n]$ называется симметрическим, если $\forall \sigma \in S_n$ $f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$.

Важный пример: $\sigma_0 = 1$, $\sigma_1 = x_1 + \ldots + x_n$, $\sigma_2 = x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_3 + \ldots + x_{n-1}x_n$, $\sigma_k = \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} x_{i_1} \dots x_{i_k}$ — элементарные симметрические многочлены.

Пример не элементарного для $n = 2$: $x_1^5 + x_2^5 + x_1x_2 + 1$.

Замечание. Множество всех симметрических многочленов образует кольцо $R \leqslant K[x_1, \dots, x_n]$.

Например, проверим замкнутость умножения. Пусть f_1, f_2 — симметрические. Тогда действительно $f(x_1, \dots, x_n) = f_1(x_1, \dots, x_n)f_2(x_1, \dots, x_n) = f_1(x_{\sigma(1)}, \dots, x_{\sigma(n)})f_2(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$.

Замечание. Стандартный способ получать симметрические многочлены — это взять любой многочлен $f \in K[Y_1, \dots, Y_n]$ и подставить в него $\sigma_1, \dots, \sigma_n$.

Теорема 13.8. (О выражении симметрического многочлена через элементарные.)

Пусть f — симметрический многочлен. Существует единственный многочлен $g \in K[Y_1, \dots, Y_n]$, такой что:

$f(x_1, \dots, x_n) = g(\sigma_1(x_1, \dots, x_n), \dots, \sigma_n(x_1, \dots, x_n))$.

Доказательство. Существование. Первый моном выглядит так:

$f = ax_1^{\lambda_1}x_2^{\lambda_2} \dots x_n^{\lambda_n} + \ldots$, причём обязательно выполняется $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ (из-за симметричности).

Рассмотрим многочлен $f_1 := f - a\sigma_1^{\lambda_1 - \lambda_2}\sigma_2^{\lambda_2 - \lambda_3} \dots \sigma_n^{\lambda_n}$.

Старшими членами (lt) элементарных симметрических многочленов $\sigma_1, \dots, \sigma_n$ являются $x_1,\ x_1x_2, \dots, x_1x_2 \dots x_n$.

Распишем $\operatorname{lt}(\sigma_1^{\lambda_1 - \lambda_2} \dots \sigma_n^{\lambda_n}) = x_1^{\lambda_1 - \lambda_2}(x_1x_2)^{\lambda_2 - \lambda_3}(x_1x_2x_3)^{\lambda_3 - \lambda_4} \dots (x_1 \dots x_n)^{\lambda_n} = x_1^{\lambda_1}x_2^{\lambda_2} \dots x_n^{\lambda_n}$.

Первый моном сократился, то есть $\operatorname{lt} f_1 < \operatorname{lt} f$. f_1 является симметрическим многочленом (так как получен операциями над элементами кольца симметрических многочленов), поэтому можем продолжить алгоритм. В конечном итоге придём к многочлену со старшим членом 0, тогда искомое разложение будет получено.

Итак, мы получили последовательность: $f_1 = f - a_1g_1(\sigma_1, \dots, \sigma_n)$, $f_2 = f_1 - a_2g_2$ и так далее. То есть $f = f_1 + a_1g_1 = f_2 + a_2g_2 + a_1g_1 = \sum a_ig_i$. То есть мы разложили f в линейную комбинацию многочленов g_i с параметрами $\sigma_1, \dots, \sigma_n$. Приведём подобные слагаемые в этой линейной комбинации, получим искомый многочлен $g = \sum a_ig_i$.

Единственность.

Пусть многочлен f имеет два различных представления: $f = g_1(\sigma_1, \dots, \sigma_n) = g_2(\sigma_1, \dots, \sigma_n)$, $g_1 \neq g_2$. Рассмотрим $g(Y_1, \dots, Y_n) = g_1(Y_1, \dots, Y_n) - g_2(Y_1, \dots, Y_n) \neq 0$, причём $g(\sigma_1, \dots, \sigma_n) = 0$.

Произвольный моном многочлена g имеет вид $aY_1^{k_1} \dots Y_n^{k_n}$. Рассмотрим старшие члены, получающиеся от различных мономов, при подстановке $\sigma_1, \dots, \sigma_n$.

$\operatorname{lt}(aY_1^{k_1} \dots Y_n^{k_n} \mid_{\sigma_1, \dots, \sigma_n}) = ax_1^{k_1 + \ldots + k_n}x_2^{k_2 + \ldots + k_n} \dots x_n^{k_n}$.

(Так как x_1 встречается в старшем мономе каждого элементарного симметрического многочлена, x_2 встречается в каждом, кроме как в σ_1 и т.д.)

Значит, различным мономам (то есть при разных наборах $k_1, \dots, k_n$) соответствуют различные старшие члены. Среди них есть самый старший, который ни с чем не сможет сократиться,

значит $\operatorname{lt} g(\sigma_1, \ldots, \sigma_n) \neq 0$ вопреки предположению. □

Теорема 13.9. (Виета.) $f = (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n) = \sum_{k=0}^{n} c_k x^k$.

$$c_k = (-1)^{n-k} \sum_{i_1 < \ldots < i_{n-k}} \alpha_{i_1} \ldots \alpha_{i_{n-k}} = (-1)^{n-k} \sigma_{n-k}(\alpha_1, \ldots, \alpha_n).$$

Чтобы получить коэффициент c_k, выберем k скобок, откуда берём x, а из оставшихся скобок получим $n - k$ коэффициентов α.