Коллоквиум по линейной алгебре

Лектор: Чубаров И. А. • Автор: Пшеничный Никита[*], группа 109

1 курс • Весенний семестр 2024 г.

Аннотация

При подготовке данного файла я использовал курсы лекций И. А. Чубарова, Т. Е. Панова и О. В. Куликовой и книги «Курс алгебры» Э. Б. Винберга, «Сборник задач по аналитической геометрии и линейной алгебре» под редакцией Ю. М. Смирнова и «Задачи по линейной алгебре» А. А. Гайфуллина, А. В. Пенского и С. В. Смирнова.

Обо всех ошибках и опечатках пишите мне, исправлю.

Программа коллоквиума

[*] `Telegram: @pshenikita. Последняя компиляция: 7 апреля 2024 г.`

1. Векторные пространства. Линейная зависимость и независимость векторов. Базис, размерность. Примеры

Определение 1. *Линейным* (или *векторным*) *пространством* над полем $\mathbb{F}$ называется множество V с заданными на нём операциями *сложения* $+ : (u, v) \in V \times V \mapsto (u + v) \in V$ и *умножения* элементов V на элементы $\mathbb{F}$ $\cdot : (\lambda, v) \in \mathbb{F} \times V \mapsto (\lambda \cdot v) \in V$, удовлетворяющие следующим аксиомам:

$$\left.\begin{array}{l} 1.\ v + u = u + v\ \forall u, v \in V; \\ 2.\ (u + v) + w = u + (v + w)\ \forall u, v, w \in V; \\ 3.\ \exists \mathbf{0} \in V : v + \mathbf{0} = v\ \forall v \in V; \\ 4.\ \forall v \in V\ \exists (-v) \in V : v + (-v) = \mathbf{0}; \end{array}\right\} (V, +) \text{ — абелева группа}$$

$$\left.\begin{array}{l} 5.\ \lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v\ \forall u, v \in \mathbb{F},\ \forall \lambda \in \mathbb{F}; \\ 6.\ (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v\ \forall v \in V,\ \forall \lambda, \mu \in \mathbb{F}; \\ 7.\ \lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v\ \forall v \in V,\ \forall \lambda, \mu \in \mathbb{F}; \\ 8.\ 1 \cdot v = v\ \forall v \in V. \end{array}\right\} \mathbb{F} \text{ линейно действует на } V$$

Определение 2. Элементы множества V называются *векторами*, элемент $\mathbf{0}$ называется *нулевым вектором*, а элемент $(-v)$ называется *противоположным* к v. Элементы $\mathbb{F}$ называют *скалярами*.

Предложение 1.

1. $\mathbf{0} \cdot v = \lambda \cdot \mathbf{0} = \mathbf{0}\ \forall v \in V,\ \forall \lambda \in \mathbb{F}$;
2. $(-1) \cdot v = -v\ \forall v \in V$;
3. $\lambda v = \mathbf{0} \Rightarrow \lambda = 0$ или $v = \mathbf{0}$.

Доказательство.

1. $0 \cdot v + 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v \Rightarrow 0 \cdot v = \mathbf{0}$. Аналогично, $\lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0} = \lambda(\mathbf{0} + \mathbf{0}) = \lambda \cdot \mathbf{0} \Rightarrow \lambda \cdot \mathbf{0} = \mathbf{0}$;
2. $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = \mathbf{0} \Rightarrow -v = (-1) \cdot v$.
3. $\lambda \neq 0 \Rightarrow \mathbf{0} = \lambda^{-1} \lambda v = v$.

■

Пример 1.

1. Множество $\{\mathbf{0}\}$ из одного элемента является линейным пространством над любым полем.
2. Множества геометрических векторов на прямой, плоскости или пространстве являются линейными пространствами над полем $\mathbb{R}$.
3. Поле $\mathbb{F}$ является векторным пространством над самим собой.
4. Поле $\mathbb{C}$ является линейным пространством над полем $\mathbb{R}$, а поле $\mathbb{R}$ является линейным пространством над полем $\mathbb{Q}$.
5. Пусть $\mathbb{F}^n := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{F} \right\}$ — множество столбцов фиксированной длины n из элементов поля $\mathbb{F}$. Операции покоординатного сложения и умножения на скаляры

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

задают на $\mathbb{F}^n$ структуру линейного пространства над $\mathbb{F}$. Его часто называют *координатным.*

Пусть V — линейное пространство над полем $\mathbb{F}$.

Определение 3. *Линейной комбинацией* системы векторов $\{v_i : i \in I\}$ называется формальная сумма вида $\sum\limits_{i \in I} \lambda_i v_i$, в которой лишь конечное число скаляров λ_i отличны от нуля.

Примечание. Линейную комбинацию системы $\{v_i : i \in I\}$ можно также определить как функцию $i \in I \mapsto \lambda_i \in \mathbb{F}$, которая принимает ненулевое значение только на конечном числе индексов.

Определение 4. Линейная комбинация $\sum\limits_{i \in I} \lambda_i v_i$ называется *тривиальной*, если $\lambda_i = 0\ \forall i \in I$.

Определение 5. Система векторов $\{v_i : i \in I\}$ называется *линейно зависимой*, если существует нетривиальная линейная комбинация, представляющая нулевой вектор. В противном случае система называется *линейно независимой.*

Лемма 1. Если система векторов $\{v_i : i \in I\}$ линейно зависима, то в ней найдётся вектор, представленный линейной комбинацией всех остальных.

Доказательство. Пусть $\sum\limits_{i \in I} \lambda_i v_i = \mathbf{0}$, причём $\exists \lambda_j \neq 0$. Тогда $v_j = \sum\limits_{i \in I \setminus \{j\}} \dfrac{-\lambda_i}{\lambda_j} v_i$. ■

Определение 6. *Базисом* пространства V называется линейно независимая система $\{v_i : i \in I\}$, порождающая всё пространство V, т. е. такая, что каждый вектор из V представляется какой-то линейной комбинацией системы $\{v_i : i \in I\}$.

Определение 7. Линейное пространство называется *конечномерным*, если в нём существует базис, состоящий из конечного числа векторов. В противном случае пространство называется *бесконечномерным.*

Предложение 2. Представление любого вектора линейного пространства в виде линейной комбинации базисных векторов единственно.

Доказательство. Действительно, если $v = \sum\limits_{i \in I} \lambda_i v_i = \sum\limits_{i \in I} \mu_i v_i$ (где $\{v_i : i \in I\}$ — базис), то получаем $\mathbf{0} = \sum\limits_{i \in I} (\lambda_i - \mu_i) v_i$. Из линейной независимости базиса, линейная комбинация в правой части тривиальна и $\lambda_i = \mu_i\ \forall i \in I$ и два представления v совпадают. ■

Определение 8. *Линейная комбинация* системы векторов $\langle v_i : i \in I \rangle$ есть множество всевозможных линейных комбинаций $\sum\limits_{i \in I} \lambda_i v_i$.

Теорема 1. В конечномерном пространстве все базисы состоят из одного числа элементов.

Доказательство этой теоремы будет опираться на следующую лемму.

Лемма 2 (О линейной зависимости). Пусть $\{e_1, \ldots, e_m\}$ и $f_1, \ldots, f_n$ — две (конечные) линейно независимые системы, причём $\{f_1, \ldots, f_n\} \subseteq \langle e_1, \ldots, e_n \rangle \Rightarrow n \leqslant m$.

Доказательство. Пусть $f_j = a_{1j} e_1 + \ldots + a_{mj} e_m$, $a_{ij} \in \mathbb{F}$, $j = 1, \ldots, n$. Т. к. $f_1, \ldots, f_n$ — линейно независимая система векторов, то $x_1 f_1 + \ldots + x_n f_n = \mathbf{0} \Leftrightarrow x_1 = \ldots = x_n = 0$. Подставляя сюда выражения f_i через $e_1, \ldots, e_m$, получаем

$$\begin{aligned}\mathbf{0} = x_1(a_{11} e_1 + \ldots + a_{m1} e_m) + \ldots x_n(a_{1n} e_1 + \ldots + a_{mn} e_m) = \\ = (a_{11} x_1 + \ldots + a_{1n} x_n) e_1 + \ldots + (a_{m1} x_1 + \ldots + a_{mn} x_n) e_m.\end{aligned}$$

Т. к. $e_1, \dots, e_m$ — линейно независимая система, то последнее равенство равносильно

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0, \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0. \end{cases}$$

Если $n > m$, то эта система имеет ненулевое решение, что противоречит линейно независимости системы $f_1, \dots, f_n$. ■

Теперь докажем теорему 1:

Доказательство. Пусть V — конечномерное пространство. По определению, в V существует базис $e_1, \dots, e_m$. Пусть $\{f_i : i \in I\}$ — другой базис. Если этот базис бесконечен, то в нём содержится конечная линейно незавсимая система векторов $f_1, \dots, f_n$, где $n > m$. При этом, т. к. $e_1, \dots, e_m$ — базис, мы имеем $\{f_1, \dots, f_n\} \subseteq \langle e_1, \dots, e_m \rangle$, что противоречит лемме о линейной зависимости. Следовательно, базис $\{f_i : i \in I\}$ конечен, т. е. имеет вид $f_1, \dots, f_n$. Тогда $\{f_1, \dots, f_n\} \subseteq \langle e_1, \dots, e_m \rangle$ и $\{e_1, \dots, e_m\} \subseteq \langle f_1, \dots, f_m \rangle$. Отсюда $n = m$. ■

Лемма 3. В конечномерном пространстве любую линейно независимую систему можно дополнить до базиса.

Доказательство. Пусть $\{e_1, \dots, e_k\}$ — конечная подсистема в V. Тогда, если эта система максимальна по включению, то она базис. Иначе существует $e_{k+1} \in V$ такой, что система $\{e_1, \dots, e_k, e_{k+1}\}$ линейно независима. Продолжная процесс далее, за конечное число шагов получим базис (в силу конечномерности пространства V). ■

Примечание. Ниже изложен удобный алгоритм дополнения линейно независимой системы до базиса. Пусть

$$u_1 = \begin{pmatrix} u_{11} \\ \vdots \\ u_{1n} \end{pmatrix}, \quad u_2 = \begin{pmatrix} u_{21} \\ \vdots \\ u_{2n} \end{pmatrix}, \quad \dots, \quad u_m = \begin{pmatrix} u_{m1} \\ \vdots \\ u_{mn} \end{pmatrix}$$

линейно независимы. Обозначим

$$U = \begin{pmatrix} u_{11} & u_{21} & \dots & u_{m1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \dots & u_{mn} \end{pmatrix}.$$

Отметим, что $\operatorname{rk} U = m$, а $\operatorname{rk}(U \mid E) = n$. Так что нужно привести матрицу $(U \mid E)$ элементарными преобразованиями к ступенчатому виду и дополнить столбцы матрицы U единичными столбцами, вошедшими в базис матрицы $(U \mid E)$.

Лемма 4. Всякое конечномерное линейное пространство V обладает базисом. Более точно, из всякого конечного порождающего множества $S \subset V$ можно выбрать базис пространства V.

Доказательство. Если множество S линейно независимо, то по лемме 1 в нём найдётся вектор, линейно выражающийся через остальные. Выкидывая этот вектор, мы получаем порождающее множество из меньшего числа векторов. Продолжая так дальше, мы в конце концов получим линейно независимое порождающее множество, т. е. базис. ■

Примечание. Чтобы сделать это на практике, выписываем векторы в матрицу по столбцам, приводим её к ступенчатому виду и те столбцы, в которых стоят лидеры, будут базисными.

Определение 9. *Размерностью* конечномерного линейного пространства V (обозначается $\dim V$) называется число элементов в базисе V. Если V бесконечномерно, то пишут $\dim V = \infty$.

Лемма 5 (Свойство монотонности размерности)**.** Подпространство W конечномерного пространства V конечномерно, причём $\dim W \leqslant \dim V$ и равенство достигается только при $W = V$.

Доказательство. Пусть $\dim V = m$ и $e_1, \dots, e_m$ — базис пространства V. Если $\dim W > m$, то в W найдётся линейно независимая система $f_1, \dots, f_n$ с $n > m$. Причём $\{f_1, \dots, f_n\} \subseteq \langle e_1, \dots, e_m \rangle = V$, что противоречит лемме о линейной зависимости. Следовательно, $\dim W \leqslant \dim V$.

Пусть $\dim W = \dim V = m$ и пусть $f_1, \dots, f_m$ — базис в W. Тогда каждый вектор линейно выражается через $f_1, \dots, f_m$, так как иначе получили бы линейно независимую систему $f_1, \dots, f_m, e_i$ из $m+1$ векторов в V, что противоречит теореме 1. Следовательно, любой вектор из V лежит в $\langle f_1, \dots, f_m \rangle = W$, т. е. $V \subseteq W$, а обратное включение верно из условия. Итак, получаем $V = W$. ■

Примечание. В нулевом пространстве $\{\mathbf{0}\}$ базисом естественно считать пустое множество $\varnothing$. Поэтому $\dim\{\mathbf{0}\} = 0$.

2. Матрица перехода от одного базиса к другому. Изменение координат вектора при замене базиса

Определение 1. Пусть V — линейное пространство, и $e_1, \dots, e_n$ — базис в V. Любой вектор $x \in V$ единственным образом представляется в виде линейной комбинации базисных векторов: $x = \sum\limits_{i=1}^{n} x_i e_i$. Числа $x_1, \dots, x_n \in \mathbb{F}$ называются *координатами* вектора x в базисе $e_1, \dots, e_n$.

Обозначения Эйнштейна. Вместо $\sum\limits_{i=1}^{n} x_i e_i$ пишем $x^i e_i$ (суммирование производится по повторяющемуся индексу). В связи с этим обозначением, нам будет также удобно обозначать координаты вектора верхними индексами вместо нижних. Для произведения матриц: $c^i_k = a^i_j b^j_k$ (суммирование опять производится по повторяющемуся индексу). Матрица (d^j_k) является обратной к (c^i_j), если $c^i_j d^j_k = \delta^i_k$ — символ Кронекера.

Пусть в пространстве V заданы два базиса: «старый» $e_1, \dots, e_n$ и «новый» $e'_1, \dots, e'_n$. Нам будет удобно обозначать векторы нового базиса через $e_{1'}, \dots, e_{n'}$. Элементы нового базиса выражаются через элементы старого: $e_{i'} = c^i_{i'} e_i$, $i' = 1, \dots, n$. Эти формулы равносильны одному матричному равенству

$$(e_{1'}, \dots, e_{n'}) = (e_1, \dots, e_n) \cdot \begin{pmatrix} c^1_{1'} & \cdots & e^1_{n'} \\ \vdots & \ddots & \vdots \\ c^n_{1'} & \cdots & c^n_{n'} \end{pmatrix}.$$

Определение 2. Матрица $C := (c^i_{i'})$ называется *матрицей перехода* от базиса $e_1, \dots, e_n$ к базису $e_{1'}, \dots, e_{n'}$. Её столбцами являются координаты новых базисных векторов в старом базисе.

Предложение 1.

1. Матрица $C_{e' \to e} = (c^{i'}_i)$ перехода от базиса $e_{1'}, \dots, e_{n'}$ к базису $e_1, \dots, e_n$ является обратной к матрице $C_{e \to e'}$ перехода от $e_1, \dots, e_n$ к $e_{1'}, \dots, e_{n'}$, т. е. $C_{e \to e'} \cdot C_{e' \to e} = E$. В частности, матрица перехода всегда невырождена.
2. Если $e_1, \dots, e_n$, $e_{1'}, \dots, e_{n'}$, $e_{1''}, \dots, e_{n''}$ — три базиса, то для соответствующих матриц перехода выполнено $C_{e \to e'} \cdot C_{e' \to e''} = C_{e \to e''}$.

Доказательство. Первое утверждение следует из второго, если положить $e'' = e$, поэтому будем доказывать второе утверждение. Пусть $C_{e \to e'} = (c^i_{i'})$, $C_{e' \to e''} = (c^{i'}_{i''})$, $C_{e \to e''} = (c^i_{i''})$. Тогда

$$c^i_{i''} e_i = e_{i''} = c^{i'}_{i''} e_{i'} = c^{i'}_{i''} c^i_{i'} e_i = c^i_{i'} c^{i'}_{i''} e_i \Rightarrow c^i_{i''} = c^i_{i'} \cdot c^{i'}_{i''}.$$

■

Примечание. Важное практическое следствие. Заметим, что в $\mathbb{R}^n$ писать матрицу перехода от стандартного базиса к любому другому очень легко — достаточно написать базисные векторы, в которые мы хотим перейти, по столбцам матрицы. Пусть мы хотим написать матрицу перехода от базиса $a = \{a_1, a_2, \ldots, a_n\}$ к базису $b = \{b_1, b_2, \ldots, b_n\}$. Для этого можно написать матрицу A перехода от стандартного к a, потом матрицу B перехода от стандартного к b, а потом выдать ответ — $A^{-1}B$.

Примечание. Трюк от Александра Александровича (и в Винберге находил). Чтобы найти матрицу $X = A^{-1}B$ при известных A и B, не надо искать A^{-1}, а потом умножать её на B. Домножим на A слева, получим $AX = B$. Это n систем линейных уравнений с одной и той же матрицей A (решив i-ую систему, найдём i-ый столбец X). Эти системы можно решать одновременно, записав матрицу A в правой части и приписав к ней каждый столбец B по очереди. Выглядеть будет как $(A \mid B)$. Решить системы — значит привести эту матрицу $n \times 2n$ к улучшенному ступенчатому виду: $(E \mid X)$. В правой части теперь будут стоять столбцы матрицы X.

Теорема 1 (Закон изменения координат)**.** Пусть $x^1, \ldots, x^n$ — координаты вектора x в базисе $e_1, \ldots, e_n$, а $x^{1'}, \ldots, x^{n'}$ — в базисе $e_{1'}, \ldots, e_{n'}$. Тогда два набора координат связаны формулой

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = C \cdot \begin{pmatrix} x^{1'} \\ \vdots \\ x^{n'} \end{pmatrix}.$$

Доказательство. В обозначениях Эйнштейна утверждение равносильно $x^i = c^i_{i'} x^{i'}$, $i = 1, \ldots, n$. Оно верно, потому что

$$x^i e_i = x = x^{i'} e_{i'} = x^{i'} c^i_{i'} e_i \Rightarrow x^i = x^{i'} c^i_{i'} = c^i_{i'} x^{i'}.$$

■

3. Изоморфизм векторных пространств одинаковой размерности

Определение 1. Пусть V и W — линейные пространства над полем $\mathbb{F}$. Отображение $\mathcal{A} : V \to W$ называется *линейным*, если $\forall u, v \in V$, $\forall \lambda \in \mathbb{F}$ выполнено $\mathcal{A}(u + v) = \mathcal{A}u + \mathcal{A}v$, $\mathcal{A}(\lambda v) = \lambda \mathcal{A} v$.

Определение 2. Биективное линейное отображение $\mathcal{A} : V \to W$ называется *изоморфизмом*, а пространства V и W, между которыми есть изоморфизм, называются *изоморфными*.

Теорема 1. Два конечномерных пространства V и W над полем $\mathbb{F}$ изоморфны тогда и только тогда, когда они имеют одинаковые размерности.

Доказательство. Из определения изоморфизма вытекает, что свойство системы векторов быть линейно независимой и порождать всё пространство сохраняются при изоморфизмах, т. е. при изоморфизме базис переходит в базис[1]. Следовательно, если $\mathcal{A} : V \to W$ — изоморфизм, то $\dim V = \dim W$. Пусть теперь $\dim V = \dim W = n$. Выберем базисы $e_1, \ldots, e_n$ и $f_1, \ldots, f_n$ соответственно. Тогда формула $\mathcal{A}(x^i e_i) = x^i f_i$ определяет линейное отображение $\mathcal{A} : V \to W$. Оно является биекцией, т. к. формула $\mathcal{A}^{-1}(x^i f_i) = x^i e_i$ определяет обратное отображение. ■

4. Векторные подпространства, равносильность двух способов их задания. Сумма и пересечение подпространств. Формула Грассмана

Определение 1. Непустое подмножество $W \subseteq V$ линейного пространства V называется *подпространством*, если $\forall u, v \in W$, $\forall \lambda \in \mathbb{F}$ выполнено $(u + v) \in \mathbb{F}$ и $\lambda u \in W$.

[1] Это частный случай того, что «образ базиса является базисом образа» при $\operatorname{Im} \mathcal{A} = W$.

Предложение 1. Множество всех решений системы однородных линейных уравнений с n неизвестными является подпространством координатного пространства $\mathbb{F}^n$.

Доказательство. Рассмотрим произвольную систему однородных линейных уравнений:

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

Очевидно, что нулевой столбец является её решением и что произведение любого решения на число также является решением. Докажем, что сумма решений $(u_1, \ldots, u_n)^t$ и $(v_1, \ldots, v_n)^t$ является решением. Подставляя её компоненты в i-ое уравнение системы, получаем

$$a_{i1}(u_1 + v_1) + \ldots + a_{in}(u_n + v_n) = \underbrace{a_{i1}u_1 + \ldots + a_{in}u_n}_{=0} + \underbrace{a_{i1}v_1 + \ldots + a_{in}v_n}_{=0} = 0.$$

∎

Определение 2. *Фундаментальная система решений* — это базис подпространства решений однородной СЛУ.

Предложение 2. Линейная оболочка $\langle v_i : i \in I \rangle$ является линейным подпространством в V. Более того, она является наименьшим по включению линейным подпространством, содержащим все векторы системы $\{v_i : i \in I\}$.

Доказательство. Сумма векторов системы и результат умножения вектора системы на скаляр представляются линейными комбинациями и потому принадлежат линейной оболочке. Следовательно, $\langle v_i : i \in I \rangle$ — подпространство. Если W — подпространство, содержащее все векторы из системы $\{v_i : i \in I\}$, то W также содержит все векторы, представляющиеся их линейными комбинациями, а значит, $W \supseteq \langle v_i : i \in I \rangle$. ∎

Теорема 1. Способы задания подпространства с помощью однородной системы линейных уравнений и линейной оболочки равносильны.

Нам понадобится следующая лемма.

Лемма 1. Пусть матрица B состоит из столбцов фундаментальной системы решений системы $Ax = 0$ (где x — вектор). Тогда линейная система $B^t y = 0$ задаёт линейную оболочку строк матрицы A.

Доказательство. Поскольку каждый столбец матрицы B является решением системы $Ax = 0$, имеет место матричное равенство $AB = 0$, которое эквивалентно $B^t A^t = 0$. Таким образом, если матрицу B^t интерпретировать как матрицу коэффициентов некоторой линейной системы, все столбцы матрицы A^t (строки A) будут ей удовлетворять.

Допустим, что некоторый столбец, не принадлежащий линейной оболочке столбцов матрицы A^t, тоже удовлетворяет системе $B^t y = 0$. Тогда рассмотрим матрицу C^t, которая получается дописыванием к матрице A этого столбца справа; полученная матрица будет удовлетворять соотношению $B^t C^t = 0$, а следовательно, и соотношению $CB = 0$. Это означает, что столбцы матрицы B являются решениями не только линейной системы $Ax = 0$, но и системы $Cx = 0$, отличающейся от системы $Ax = 0$ одним добавленным уравнением, которое по предположению линейно не выражается через исходные уравнения. Это означает, что ранг матрицы C на единицу больше ранга матрицы A, т. е. количество свободных неизвестных у системы $Cx = 0$ на единицу меньше, чем у системы $Ax = 0$. Значит, все столбцы матрицы B не могут быть решениями системы $Cx = 0$ — противоречие. Таким образом, системе $B^t y = 0$ удовлетворяют все строки матрицы A, и притом только они. ∎

Доказательство. Пусть подпространство задано однородной системой линейных уравнений

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

Тогда задать его линейной оболочкой можно, найдя ФСР. С помощью элементарных преобразований приведём её к ступенчатому виду. Число ненулевых уравнений в этом ступенчатом виде равно $r = \operatorname{rk} A$. Поэтому общее решение будет содержать r главных неизвестных и с точностью до перенумерации неизвестных будет иметь вид

$$\begin{cases} x_1 = c_{11}x_{r+1} + c_{12}x_{r+2} + \ldots + c_{1,n-r}x_n, \\ x_2 = c_{21}x_{r+1} + c_{22}x_{r+2} + \ldots + c_{2,n-r}x_n, \\ \ldots \\ x_r = c_{r1}x_{r+1} + c_{r2}x_{r+2} + \ldots + c_{r,n-r}x_n. \end{cases}$$

Придавая поочерёдно одному из свободных неизвестных $x_{r+1}, x_{r+2}, \ldots, x_n$ значение 1, а остальным — 0, получим следующие решения системы:

$$u_1 = \begin{pmatrix} c_{11} \\ \vdots \\ c_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} c_{12} \\ \vdots \\ c_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad u_{n-r} = \begin{pmatrix} c_{1,n-r} \\ \vdots \\ c_{r,n-r} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Ранг системы векторов $\{u_1, u_2, \ldots, u_{n-r}\}$ равен рангу матрицы

$$\left(\begin{array}{cccc|cccc} c_{11} & c_{21} & \ldots & c_{r1} & 1 & 0 & \ldots & 0 \\ c_{12} & c_{22} & \ldots & c_{r2} & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{1,n-r} & c_{2,n-r} & \ldots & c_{r,n-r} & 0 & 0 & \ldots & 1 \end{array}\right)$$

Если поменять местами блоки, отделённые чертой, то получится улучшенный ступенчатый вид с количеством ступенек, равным $n - r$. Так что ранг системы векторов $\{u_1, u_2, \ldots, u_{n-r}\}$ равен количеству векторов в этой системе, поэтому она линейно независима. Эта система также порождает всё подпространство решений, т. к. любая линейная комбинация вида

$$\lambda_1 u_1 + \ldots + \lambda_{n-r} u_{n-r}$$

является решением, в котором свободные неизвестные имеют значения $\lambda_1, \ldots, \lambda_{n-r}$.

Теперь пусть подпространство задано линейной оболочкой

$$\left\langle u_1 = \begin{pmatrix} u_{11} \\ \vdots \\ u_{1n} \end{pmatrix}, \quad u_2 = \begin{pmatrix} u_{21} \\ \vdots \\ u_{2n} \end{pmatrix}, \quad \ldots, \quad u_m = \begin{pmatrix} u_{m1} \\ \vdots \\ u_{mn} \end{pmatrix} \right\rangle$$

Составим матрицу

$$U = \begin{pmatrix} u_{11} & u_{12} & \ldots & u_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m2} & \ldots & u_{mn} \end{pmatrix}$$

из строк $u_1, u_2, \ldots, u_n$. Найдём ФСР системы $Ux = 0$ (указанным выше способом), и запишем её векторы по строкам в матрицу U'. По лемме 1 пространство решений однородной системы линейных уравнений $U'y = 0$ есть линейная оболочка строк матрицы U, а это и есть данные нам векторы. ■

Предложение 3. Пересечение $V_1 \cap V_2$ подпространств в V является подпространством в V.

Доказательство. Во-первых, $\mathbf{0} \in V_1$ и $\mathbf{0} \in V_2$, поэтому $\mathbf{0} \in V_1 \cap V_2 \neq \varnothing$. Во-вторых, $\forall u, v \in V_1 \cap V_2$, $\forall \lambda \in \mathbb{F}$ сумма $u + v$ и произведение λv также лежат в V_1 и в V_2, а значит, и в $V_1 \cap V_2$. ■

Примечание. Аналогично доказывается, что для любого семейства подпространств $\{U_i : i \in I\}$ их пересечение $\bigcap\limits_{i \in I} U_i$ тоже подпространство.

Определение 3 (Сумма подпространств)**.** $V_1 + V_2 := \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$.

Предложение 4. $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$.

Доказательство. Включение $V_1 + V_2 \subseteq \langle V_1 \cup V_2 \rangle$ следует из того, что вектор $v_1 + v_2$ является линейной комбинацией векторов $v_1, v_2 \in V_1 \cup V_2$. Докажем обратное включение. Для этого рассмотрим линейную комбинацию $v = \lambda_1 u_1 + \ldots + \lambda_n u_n$ векторов $u_1, \ldots, u_n \in V_1 \cup V_2$. Можно считать, что $u_1, \ldots, u_k \in V_1$ и $u_{k+1}, \ldots, u_n \in V_2$. Тогда мы имеем $v = v_1 + v_2$, где $v_1 = \lambda_1 u_1 + \ldots + \lambda_n u_n \in V_1$ и $v_2 = \lambda_{k+1} u_{k+1} + \ldots + \lambda_n u_n \in V_2$. Следовательно, $v \in V_1 + V_2$. ■

Примечание. Само объединение $V_1 \cup V_2$ подпространств в общем случае не является подпространством. Примером служит объединение двух прямых на плоскости.

Теорема 2 (Формула Грассмана)**.** $\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$.

Доказательство. Выберем базис $e_1, \ldots, e_k$ пространства $V_1 \cap V_2$. Воспользовавшись леммой 3 из первого вопроса, можем дополнить его до базиса $e_1, \ldots, e_k, f_1, \ldots, f_\ell$ пространства V_1 и до базиса $e_1, \ldots, e_k, g_1, \ldots, g_m$ пространства V_2. Тогда мы имеем $\dim(V_1 \cap V_2) = k$, $\dim V_1 = k + \ell$, $\dim V_2 = k + m$. Докажем, что $e_1, \ldots, e_k, f_1, \ldots, f_\ell, g_1, \ldots, g_m$ — базис пространства $V_1 + V_2$. Заметим, что т. к. $V_1 + V_2 = \langle V_1 \cup V_2 \rangle$, то любой вектор из $V_1 + V_2$ выражается через эту систему векторов. Остаётся проверить, что эта система линейно независима. Пусть имеет место равенство

$$\lambda_1 e_1 + \ldots + \lambda_k e_k + \mu_1 f_1 + \ldots + \mu_\ell f_\ell + \nu_1 g_1 + \ldots + \nu_m g_m = \mathbf{0}.$$

Перепишем его в виде

$$\lambda_1 e_1 + \ldots + \lambda_k e_k + \mu_1 f_1 + \ldots + \mu_\ell f_\ell = -\nu_1 g_1 - \ldots - \nu_m g_m.$$

Вектор, стоящий в обеих частях этого равенства, лежит в $V_1 \cap V_2$ и линейно выражается через $e_1, \ldots, e_k$. Т. к. векторы $e_1, \ldots, e_k, f_1, \ldots, f_\ell$ линейно независимы по построению, получаем $\mu_1 = \ldots = \mu_\ell = 0$. Аналогично, $\nu_1 = \ldots = \nu_m = 0$. Тогда из линейной независимости $e_1, \ldots, e_k$ следует $\lambda_1 = \ldots = \lambda_k = 0$. Итак, $\dim(V_1 + V_2) = k + \ell + m$, отсюда вытекает требуемое. ■

Алгоритм вычисления базисов в $U_1 + U_2$ и $U_1 \cap U_2$ (U_1 и U_2 — подпространства конечномерного пространства V). При доказательстве формулы Грассмана мы попутно показали, что базис $U_1 + U_2$ есть объединение базисов U_1 и U_2. Отметим, что в случае, когда сумма $U_1 + U_2$ прямая, пересечение этих базисов пустое.

Чтобы найти базис пересечения, зададим оба подпространства в виде системы линейных уравнений. Система, состоящая из всех уравнений обоих систем, будет задавать пересечение этих подпространств. Для нахождения базиса ищем ФСР.

5. Прямая сумма подпространств. Внешняя прямая сумма векторных пространств

Определение 1. Сумма V_1+V_2 подпространств пространства V называется *прямой* (обозначается $V_1 \oplus V_2$), если $\forall v \in V_1+V_2$ представление $v = v_1 + v_2$, где $v_1 \in V_1$, $v_2 \in V_2$, единственно.

Теорема 1. Следующие условия эквивалентны для подпространств V_1 и V_2:

1. сумма $V_1 + V_2$ прямая;
2. $V_1 \cap V_2 = \{\mathbf{0}\}$;
3. если $\mathbf{0} = v_1 + v_2$, где $v_1 \in V_1$ и $v_2 \in V_2$, то $v_1 = v_2 = \mathbf{0}$;
4. $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$.

Доказательство. $1 \Rightarrow 2$. Пусть найдётся $v \in V_1+V_2$, $v \neq \mathbf{0}$. Тогда $\mathbf{0} = \mathbf{0}+\mathbf{0} = v+(-v)$. Получаем, что представление вектора $\mathbf{0}$ не единственно, и сумма $V_1 + V_2$ не прямая.

$2 \Rightarrow 3$. Если существует представление $\mathbf{0} = v_1 + v_2$, где $v_1 \in V_1$ и $v_2 = (-v_1) \in V_2$ и $v_1 \neq \mathbf{0}$, т. е. $v_1 \in V_1 \cap V_2 \neq \{\mathbf{0}\}$ — противоречие.

$3 \Rightarrow 1$. Пусть у вектора $v \in V$ есть два разложения $v = u_1 + u_2 = v_1 + v_2$, где $u_1, v_2 \in V_1$ и $u_2, v_2 \in V_2$. Тогда $\mathbf{0} = (u_1 - v_1) + (u_2 - v_2)$, где $u_1 - v_1 \in V_1$ и $u_2 - v_2 \in V_2$. Следовательно, $u_1 - v_1 = u_2 - v_2 = \mathbf{0}$, т. е. два разложения совпадают.

$2 \Leftrightarrow 4$. Следствие формулы Грассмана. ■

Определение 2. Сумма $V_1 + \ldots + V_n$ подпространств пространства V называется *прямой* (обозначается $V_1 \oplus \ldots \oplus V_n$), если $\forall v \in V_1 + \ldots + V_n$ представление $v = v_1 + v_2 + \ldots + v_n$, где $v_i \in V_i$, единственно.

Теорема 2. Для подпространств $V_1, \ldots, V_n$ пространства V следующие условия эквивалентны:

1. сумма $V_1 + \ldots + V_n$ прямая;
2. $V_i \cap \sum\limits_{j\neq i} V_j = \{\mathbf{0}\}$;
3. если $\mathbf{0} = v_1 + \ldots + v_n$, где $v_i \in V_i$, то $v_1 = \ldots = v_n = \mathbf{0}$;
4. $\dim\left(\sum\limits_i V_k\right) = \sum\limits_i \dim V_i$.

Доказательство. Индукция по n с помощью предыдущей теоремы. ■

Примечание. При $n \geqslant 3$ условие 2 в предыдущей теореме сильнее, чем условие $V_i \cap V_j = \{\mathbf{0}\}$ $\forall i \neq j$. Это последнее условие не гарантирует, что сумма подпространств прямая. Действительно, рассмотрим следующие три подпространства в $\mathbb{R}^2$ со стандартным базисом e_1, e_2: $V_1 = \langle e_1 \rangle$, $V_2 = \langle e_2 \rangle$ и $V_3 = \langle e_1 + e_2 \rangle$. Тогда $V_i \cap V_j = \{\mathbf{0}\}$ $\forall i \neq j$, но сумма данных подпространств не прямая, т. к.

$$e_1 + e_2 = e_1 + e_2 + \mathbf{0} = \mathbf{0} + \mathbf{0} + (e_1 + e_2).$$

Определение 3. Пусть $V_1, \ldots, V_n$ — линейные пространства над одним полем $\mathbb{F}$. Их *внейшней прямой суммой* (обозначается как $V_1 \oplus \ldots \oplus V_n$) называется линейное пространство $V_1 \times \ldots \times V_n$ с операциями, определёнными покомпонентно:

$$(u_1, \ldots, u_n) + (v_1, \ldots, v_n) = (u_1 + v_1, \ldots, u_n + v_n), \quad \lambda \cdot (v_1, \ldots, v_n) = (\lambda v_1, \ldots, \lambda v_n).$$

Предложение 1. Для любого пространства $U \subset V$ найдётся подпространство $W \subset V$ такое, что $V = U \oplus W$.

Определение 4. Такое подпространство W называется *прямым дополнением* к U.

Доказательство. Пусть $e_1, \ldots, e_k$ — базис в U. Его можно дополнить до базиса V векторами $e_{k+1}, \ldots, e_n$ (где $n = \dim V$), тогда искомое $W := \langle e_{k+1}, \ldots, e_n \rangle$. ■

Примечание. Пространство $V = V_1 + \ldots + V_m$ можно превратить в прямую сумму пространств. Рассмотрим $U_i = \{(\mathbf{0}, \ldots, \mathbf{0}, v_i, \mathbf{0}, \ldots, \mathbf{0}) : v_i \in V_i\} \subset V$. Тогда

$$(v_1, \ldots, v_m) = (v_1, \mathbf{0}, \ldots, \mathbf{0}) + (\mathbf{0}, v_2, \ldots, \mathbf{0}) + \ldots + (\mathbf{0}, \ldots, \mathbf{0}, v_m),$$

а отсюда $V = U_1 \oplus \ldots \oplus U_m$.

6. Факторпространство, его размерность. Коразмерность. Связь с решениями неоднородной системы линейных уравнений

Пусть V — линейное пространство, а $W \subseteq V$ — его подпространство.

Определение 1. *Классом смежности* вектора $v \in V$ по подпространству W называется множество $v + W := \{v + w : w \in W\}$.

Лемма 1. Равенство $v_1 + W = v_2 + W$ имеет место тогда и только тогда, когда $v_1 - v_2 \in W$.

Доказательство. Пусть $v_1 + W = v_2 + W$. Тогда $v_1 \in v_1 + W = v_2 + W$, значит $\exists w \in W : v_1 = v_2 + w$. Следовательно, $v_1 - v_2 = w \in W$. Обратно, пусть $v := v_1 - v_2 \in W$. Докажем, что $v_1 + W \subseteq v_2 + W$. Возьмём произвольный вектор $u \in v_1 + W$. Тогда $u = v_1 + w$ для некоторого $w \in W$. Мы имеем $u = v_1 + w = v_2 + (v + w)$, где $v + w \in W$. Следовательно, $u \in v_2 + W$ и $v_1 + W \subseteq v_2 + W$. Обратное включение доказывается аналогично. ■

Предложение 1. Отношение $v_1 \sim v_2 \overset{\text{def}}{\iff} u_1 - u_2 \in U$ задаёт отношение эквивалентности на V.

Доказательство. Совсем несложно проверяются все аксиомы. ■

Определение 2. *Факторпространством* линейного пространства V по подпространству W называется множество $V/W := \{v + W : v \in V\}$ с операциями сложения и умножения на скаляры:

$$(u + W) + (v + W) := (u + v) + W, \quad \lambda \cdot (v + W) := \lambda v + W.$$

Предложение 2. Приведённые выше операции определены на классах смежности корректно и задают на V/W структуру линейного пространства.

Доказательство. Проверим корректность определения операций, т. е. независимость результата операции от выбора вектора v в смежном классе $v + W$. Докажем для сложения. Если $u_1 + W = u_2 + W$ и $v_1 + W = v_2 + W$, то $u := u_1 - u_2 \in W$ и $v := v_1 - v_2 \in W$ в силу предыдущей леммы. Следовательно,

$$\begin{aligned}(u_1 + W) + (v_1 + W) = (u_1 + v_1) + W = (u_2 + v_2) + (u + v) + W = \\ = (u_2 + v_2) + W = (u_2 + W) + (v_2 + W).\end{aligned}$$

Корректность определения умножения на скаляры проверяется аналогично. Теперь докажем, что V/W — линейное пространство. Свойства 1 и 2 сразу следуют из определения. Нулём является $\mathbf{0} + W = W$, а противоположным к $v + W$ является $(-v) + W$, Проверим свойство 5:

$$\begin{aligned}\lambda \cdot ((u + W) + (v + W)) = \lambda \cdot ((u + v) + W) = (\lambda u + \lambda v) + W = \\ = (\lambda u + W) + (\lambda v + W) = \lambda(u + W) + \lambda(v + W).\end{aligned}$$

Оставшиеся свойства 6-8 проверяются аналогично. ■

Определение 3. *Коразмерностью* подпространства W линейного пространства V (обозначается через $\operatorname{codim} W$) называется $\dim V/W$.

Теорема 1. $\operatorname{codim} W = \dim V - \dim W$.

Доказательство. Пусть $\dim V = n$, $\dim W = k$ и $e_1, \ldots, e_k$ — базис в W. Дополним его до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ в V. Докажем, что классы $e_{k+1} + W, \ldots, e_n + W$ образуют базис в V/W. Вначале покажем, что они линейно независимы. Пусть $\lambda_{k+1}(e_{k+1} + W) + \ldots + \lambda_n(e_n + W) = \mathbf{0} + W$. Тогда $(\lambda_{k+1}e_{k+1} + \ldots + \lambda_n e_n) + W = \mathbf{0} + W$, т.е. $v := \lambda_1 e_1 + \ldots + \lambda_k e_k \in W$. Т.к. $e_1, \ldots, e_k$ — базис в W, то можем записать $v = \lambda_1 e_1 + \ldots + \lambda_k e_k$. Тогда получаем

$$\lambda_1 e_1 + \ldots + \lambda_k e_k - \lambda_{k+1} e_{k+1} - \ldots - \lambda_n e_n = \mathbf{0}.$$

Т.к. $e_1, \ldots, e_n$ — базис в V, то $\lambda_1 = \ldots = \lambda_{k+1} = \ldots = \lambda_n = 0$. Значит, классы $e_{k+1} + W, \ldots, e_n + W$ линейно независимы. Осталось доказать, что эти классы порождают всё пространство. Возьмём произвольный вектор $v + W \in V/W$. Разложим вектор v по базису в V: $v = \lambda_1 e_1 + \ldots + \lambda_k e_k + \lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n$. Тогда

$$\begin{aligned} v + W &= (\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n) + (\lambda_1 e_1 + \ldots + \lambda_k e_k) + W = \\ &= (\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n) + W = \lambda_{k+1}(e_{k+1} + W) + \ldots + \lambda_n(e_n + W). \end{aligned}$$

Итак, в базисе V/W ровно $n - k$ векторов. ■

Предложение 3. Совокупность всех решений произвольной совместной системы линейных уравнений есть сумма какого-либо одного её решения и подпространства решений однородной системы линейных уравнений с той же матрицей коэффициентов.

Доказательство. Пусть $u = (u_1, u_2, \ldots, u_n)^t \in \mathbb{F}^n$ — частное решение неоднородной СЛУ

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

Пусть также $v = (v_1, v_2, \ldots, v_n)^t \in \mathbb{F}^n$ — произвольное решение ассоциированной однородной системы

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

Тогда сумма $u + v$ является решением первой системы. Действительно,

$$\begin{aligned} a_{i1}(u_1 + v_1) + a_{i2}(u_2 + v_2) + \ldots + a_{in}(u_n + v_n) &= \\ = \underbrace{a_{i1}u_1 + a_{i2}u_2 + \ldots + a_{in}u_n}_{=b_i} &+ \underbrace{a_{i1}v_1 + a_{i2}v_2 + \ldots + a_{in}v_n}_{=0} = b_i. \end{aligned}$$

Обратно, если u' — произвольное решение неоднородной СЛУ, то $v = u' - u$ является решением ассоциированной однородной системы (проверяется так же). ■

Связь определения факторпространства со структурой решений неоднородной СЛУ заключается в том, что они «похожи». Я уточню это позже у Игоря Андреевича.

7. Линейные функции на векторном пространстве, их ядра. Изменение коэффициентов линейной формы при замене базиса. Сопряжённое пространство V^*, дуальный базис. Канонический изоморфизм $V \simeq V^{**}$

Сначала см. вопрос 8.

Определение 1. Линейное отображение $f : V \to \mathbb{F}$ из пространства V над полем $\mathbb{F}$ в поле $\mathbb{F}$ (одномерное векторное пространство) называется *линейной функцией* (*функционалом*).

Определение 2. $V^* := \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$ — *двойственное* (*сопряжённое*, *дуальное*) пространство к V.

Пусть $e_1, \ldots, e_n$ — базис в V. Значение линейной функции $\xi \in V^*$ на любом векторе $x = x^i e_i \in V$ определяется её значениями на базисных векторах, т. к. $\xi(x) = \xi(x^i e_i) = x^i \xi(e_i)$. Определим линейные функции $\varepsilon^1, \ldots, \varepsilon^n \in V^*$ по правилу $\varepsilon^i(e_j) = \delta^i_j$. Тогда для любого вектора $x = x^j e_j$ мы имеем

$$\varepsilon^i(x) = \varepsilon^i(x^j e_j) = x^j \varepsilon^i(e_j) = x^j \delta^i_j = x^i.$$

Определение 3. В связи с этим функции $\varepsilon^1, \ldots, \varepsilon^n$ часто называют *координатными функциями*.

Предложение 1. Линейные функции $\varepsilon^1, \ldots, \varepsilon^n$ образуют базис в V^*.

Доказательство. Линейная независимость: пусть $x_1\varepsilon^1 + \ldots + x_n\varepsilon^n = \mathbf{0}$. Это равенство означает, что линейная функция $\xi : x_i\varepsilon^i$ равна нулю на любом векторе из V. Вычислим её на векторе e_j:

$$\mathbf{0} = \xi(e_j) = x_i\varepsilon^i(e_j) = x_i\varepsilon^i(e_j) = x_i\delta^i_j = x_j.$$

Итак, $x_1 = \ldots = x_n = 0$, а значит, $\varepsilon^1, \ldots, \varepsilon^n$ линейное независимы.

Теперь проверим, что $\varepsilon^1, \ldots, \varepsilon^n$ порождают всё пространство V^*. Мы утверждаем, что любая линейная функция ξ представляется в виде линейной комбинации $\xi = \xi_i\varepsilon^i$, где $\xi_i = \xi(e_i)$. Действительно, для любого вектора $x = x^j e_j \in V$ мы имеем

$$\xi_i\varepsilon^i(x) = \xi_i x^i = \xi(e_i)x^i = \xi(x^i e_i) = \xi(x).$$

Таким образом, $\varepsilon^1, \ldots, \varepsilon^n$ — базис в V^*. ■

Определение 4. Базис $\varepsilon^1, \ldots, \varepsilon^n$ пространства V^* называется *двойственным* (*сопряжённым*, *дуальным*) базисом к $e_1, \ldots, e_n$.

Следствие 1. $\dim V = \dim V^* \Rightarrow V \simeq V^*$.

Пусть теперь $e_{1'}, \ldots, e_{n'}$ — друго базис пространства V и $C = (c^i_{i'})$ — матрица перехода от $e_{i'} = c^i_{i'} e_i$. Рассмотрим двойственные базисы $\varepsilon^1, \ldots, \varepsilon^n$ и $\varepsilon^{1'}, \ldots, \varepsilon^{n'}$.

Предложение 2. Матрица перехода от $\varepsilon^1, \ldots, \varepsilon^n$ к $\varepsilon^{1'}, \ldots, \varepsilon^{n'}$ есть $(C^{-1})^t$

Доказательство. Для любого вектора $x = x^i e_i = x^{i'} e_{i'}$ мы имеем $\varepsilon^i(x) = x^i = c^i_{i'} x^{i'} = c^i_{i'} \varepsilon^{i'}(x)$. Следовательно, $\varepsilon^i = c^i_{i'}\varepsilon^{i'}$, что эквивалентно матричному соотношению

$$\begin{pmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{pmatrix} = C \cdot \begin{pmatrix} \varepsilon^{1'} \\ \vdots \\ \varepsilon^{n'} \end{pmatrix}$$

или $(\varepsilon^{1'}, \ldots, \varepsilon^{n'}) = (\varepsilon^1, \ldots, \varepsilon^n) \cdot (C^{-1})^t$. ■

Для построения изоморфизма между пространствами V и V^* нам необходимо выбрать базис в V (и двойственный ему базис в V^*). Разные базисы дают разные изоморфизмы. Для бесконечномерных пространств ситуация иная: пространства V и V^* **никогда не изоморфны**, пространство

V^* всегда «больше». Проиллюстрируем это на примере. Обозначим через $\mathbb{F}^\infty$ пространство финитных последовательностей из элементов поля $\mathbb{F}$, а через $\widehat{F}^\infty$ — пространство всех бесконечных последовательностей элементов поля $\mathbb{F}$.

Предложение 3. $(\mathbb{F}^\infty)^* \simeq \widehat{\mathbb{F}}^\infty$.

Доказательство. Возьмём в пространстве $\mathbb{F}^\infty$ базис $\{e_i\}_{i=1}^\infty$, $e_i = (\overbrace{0,\ldots,0}^{i-1},1,0,\ldots)$. Рассмотрим отображение $\mathcal{A} : (\mathbb{F}^\infty)^* \to \widehat{\mathbb{F}}^*$, $f \mapsto (f(e_1), f(e_2),\ldots)$, которое линейной функции $f \in (\mathbb{F}^\infty)^*$ ставит в соответствие последовательность её значений на базисных векторах. Это отображение очевидно линейно. Кроме того, оно биективно: обратное отображение задаётся формулой $\mathcal{A}^{-1}(x_1, x_2,\ldots) = f \in (\mathbb{F}^\infty)^*$, где $f(e_i) = x_i$. Т. к. любой элемент $y \in \mathbb{F}^\infty$ есть (конечная) линейная комбинация элементов e_i, то линейная функция f однозначно восстанавливаются по её значениям на базисных векторах ■

Предложение 4. $\mathbb{Z}_2^\infty \not\simeq \widehat{\mathbb{Z}}_2^\infty$.

Доказательство. $\mathbb{Z}_2^\infty$ можно отождествить с множеством рациональных чисел на отрезке $[0;1]$ в двоичной записи, а $\widehat{\mathbb{Z}}_2^\infty$ — с множеством вещественных чисел на $[0;1]$ в двоичной записи. Поэтому биекции между этими множествами быть не может. ■

Следствие 2. $\mathbb{Z}_2^\infty \not\simeq (\mathbb{Z}_2^\infty)^*$.

Определение 5. $V^{**} := (V^*)^*$ — *второе сопряжённое пространство.*

Теорема 1. Пусть V — конечномерное линейное пространство. Отображение $\varphi : x \in V \mapsto \varphi_x \in V^{**}$, где $\varphi_x(\xi) := \xi(x)$ для $\xi \in V^*$, является изоморфизмом.

Доказательство. Из определения линейных функций следует, что $\varphi_{x+y} = \varphi_x + \varphi_y$ и $\varphi_{\lambda x} = \lambda\varphi_x$. Остаётся проверить, что отображение $x \mapsto \varphi_x$ биективное. Пусть $e_1,\ldots,e_n$ — базис пространства V и $\varepsilon^1,\ldots,\varepsilon^n$ — сопряжённый базис пространства V^*. Тогда

$$\omega_i(\varepsilon^j) = \varepsilon^j(e_i) = \delta^i_j,$$

так что $\omega_1,\ldots,\omega_n$ — базис пространства V^{**}, сопряжённый $\varepsilon^1,\ldots,\varepsilon^n$. Отображение $x \mapsto \varphi_x$ переводит вектор с координатами $x_1,\ldots,x_n$ в базисе $e_1,\ldots,e_n$ пространства V в вектор с такими же координатами в базисе $\omega_1,\ldots,\omega_n$ пространства V^{**}. Следовательно, оно биективно. ■

Отметим, что этот изоморфизм не зависит от выбора базисов в V и V^{**}, поэтому его часто называют *каноническим.*

Следствие 1. Всякий базис пространства V^* сопряжён некоторому базису пространства V.

Задача 1 (Из Винберга)**.** Доказать, что линейные функции $f_1,\ldots,f_n$ (где $n = \dim V$) составляют базис пространства V^* тогда и только тогда, когда не существует ненулевого вектора $x \in V$, для которого $f_1(x) = \ldots = f_n(x) = 0$.

$\triangleright$ $\Rightarrow$. Пусть $f_1,\ldots,f_n$ — базис и нашёлся такой вектор $v \in V$, $v \neq \mathbf{0}$, для которого $f_1(v) = \ldots = f_n(v) = 0$. Этот базис сопряжён некоторому базису $e_1,\ldots,e_n$ пространства V. А это значит, что вектор v ненулевой, но все координаты у него нулевые. Так не бывает.

$\Leftarrow$. Выберем базис $\varepsilon^1,\ldots,\varepsilon^n$ в V^*. Тогда $f_i = a^i_j\varepsilon^j$. Рассмотрим систему уравнений (верхними индексами обозначены координаты, а не степени)

$$\begin{cases} f_1(x) = 0, \\ f_2(x) = 0, \\ \ldots \\ f_n(x) = 0 \end{cases} \Leftrightarrow \begin{cases} a^1_1x^1 + a^1_2x^2 + \ldots + a^1_nx^n = 0, \\ a^2_1x^1 + a^2_2x^2 + \ldots + a^2_nx^n = 0, \\ \ldots \\ a^n_1x^1 + a^n_2x^2 + \ldots + a^n_nx^n = 0, \end{cases}$$

По условию не существует ненулевого вектора x такого, что $f_1(x) = f_2(x) = \ldots = f_n(x) = 0$, поэтому выписанная система не имеет решений кроме нулевого. Поэтому она определена, а значит, по правилу Крамера матрица коэффициентов невырожденна, отсюда следует линейная независимость строк. А полноту можно не доказывать, потому что количество векторов уже правильное. ◀

8. Линейные отображения и операторы. Ядро и образ, связь их размерностей. Критерий инъективности

Определение 1. Пусть V и W — линейные пространства над полем $\mathbb{F}$. Отображение $\mathcal{A} : V \to W$ называется *линейным*, если $\forall u, v \in V$, $\forall \lambda \in \mathbb{F}$ выполнено $\mathcal{A}(u + v) = \mathcal{A}u + \mathcal{A}v$ и $\mathcal{A}(\lambda v) = \lambda \mathcal{A}v$.

Определение 2. Линейное отображение $\mathcal{A} : V \to V$ из пространства V в себя называется *линейным оператором*.

Определение 3. Пусть $\mathcal{A} : V \to W$ — линейное отображение. *Ядром* $\mathcal{A}$ называется множество $\operatorname{Ker} \mathcal{A} := \{v \in V : \mathcal{A}v = \mathbf{0}\}$. *Образом* $\mathcal{A}$ называется множество $\operatorname{Im} \mathcal{A} := \{\mathcal{A}v : v \in V\}$.

Предложение 1. Пусть $\mathcal{A} : V \to W$ — линейное отображение. Тогда $\operatorname{Ker} \mathcal{A}$ — подпространство в V, а $\operatorname{Im} \mathcal{A}$ — подпространство в W.

Доказательство. Пусть $u, v \in \operatorname{Ker} \mathcal{A}$. Т. е. $\mathcal{A}u = \mathcal{A}v = \mathbf{0}$. Тогда $\mathcal{A}(u + v) = \mathcal{A}u + \mathcal{A}v = \mathbf{0}$ и $\mathcal{A}(\lambda u) = \lambda \mathcal{A}u = \mathbf{0}$. Следовательно, $u + v \in \operatorname{Ker} \mathcal{A}$ и $\lambda u \in \operatorname{Ker} \mathcal{A}$ $\forall \lambda \in \mathbb{F}$, а значит, $\operatorname{Ker} \mathcal{A}$ — подпространство в V.

Пусть теперь $x, y \in \operatorname{Im} \mathcal{A}$, т. е. $\exists u, v \in V : \mathcal{A}u = x, \mathcal{A}v = y$. Тогда $\mathcal{A}(u + v) = x + y$ и $\mathcal{A}(\lambda u) = \lambda x$. Следовательно, $x + y \in \operatorname{Im} \mathcal{A}$ и $\lambda x \in \operatorname{Im} \mathcal{A}$ $\forall \lambda \in \mathbb{F}$, а значит, $\operatorname{Im} \mathcal{A}$ — подпространство в W. ■

Лемма 1 (Критерий инъективности)**.** Линейное отображение $\mathcal{A} : V \to W$ инъективно тогда и только тогда, когда $\operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$.

Доказательство. $\Rightarrow$. Мы знаем, что $\mathcal{A}\mathbf{0} = \mathbf{0}$, а т. к. $\mathcal{A}$ инъективно, то $\mathbf{0}$ — единственный вектор из V, переходящий в $\mathbf{0}$, отсюда $\operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$.

$\Leftarrow$. Пусть $\mathcal{A}u = \mathcal{A}v \Rightarrow \mathcal{A}(u - v) = \mathbf{0}$, значит, $u - v \in \operatorname{Ker} \mathcal{A} = \{\mathbf{0}\}$, отсюда $u = v$. ■

Теорема 1. Пусть $\mathcal{A} : V \to W$ — линейное отображение. Тогда соответствие $v + \operatorname{Ker} \mathcal{A} \mapsto \mathcal{A}v$ задаёт изоморфизм между факторпространством $V/\operatorname{Ker} \mathcal{A}$ и подпространством $\operatorname{Im} \mathcal{A}$.

Доказательство. Сначала проверим, что $v + \operatorname{Ker} \mathcal{A} \mapsto \mathcal{A}v$ действительно корректно определяет отображение $\widetilde{\mathcal{A}} : V/\operatorname{Ker} \mathcal{A} \to \operatorname{Im} \mathcal{A}$. Для этого нужно проверить, что если $u + \operatorname{Ker} \mathcal{A} = v + \operatorname{Ker} \mathcal{A}$, то $\mathcal{A}u = \mathcal{A}v$. Из равенства классов смежности следует $u - v \in \operatorname{Ker} \mathcal{A}$, а отсюда $\mathcal{A}(u - v) = \mathbf{0}$, т. е. $\mathcal{A}u = \mathcal{A}v$. Итак, отображение $\widetilde{\mathcal{A}}$ определено корректно.

Линейность и сюръективность $\widetilde{\mathcal{A}}$ очевидны. Инъективность проверяется по критерию:

$$\operatorname{Ker} \widetilde{\mathcal{A}} = \{(v + \operatorname{Ker} \mathcal{A}) \in V/\operatorname{Ker} \mathcal{A} : \widetilde{\mathcal{A}}(v + \operatorname{Ker} \mathcal{A}) = \mathcal{A}v = \mathbf{0}\} = \operatorname{Ker} \mathcal{A} = \mathbf{0} + \operatorname{Ker} \mathcal{A}.$$

Итак, $\widetilde{\mathcal{A}}$ задаёт изоморфизм $V/\operatorname{Ker} \mathcal{A} \simeq \operatorname{Im} \mathcal{A}$. ■

Следствие 1. Для всякого линейного отображения $\mathcal{A} : V \to W$ мы имеем

$$\dim V = \dim \operatorname{Ker} \mathcal{A} + \dim \operatorname{Im} \mathcal{A}.$$

Предложение 2. Если в каких-то базисах пространств V и W линейное отображение $\mathcal{A} : V \to W$ имеет матрицу A, то

$$\dim \operatorname{Im} \mathcal{A} = \operatorname{rk} A.$$

Доказательство. Очевидно, что $\operatorname{Im}\mathcal{A}$ есть линейная оболочка образов базисных векторов $e_1,\dots,e_n$ пространства V и, значит, $\dim\operatorname{Im}\mathcal{A}$ есть ранг системы векторов $\mathcal{A}(e_1),\dots,\mathcal{A}(e_n)$. Но в столбцах матрицы A как раз и записаны координаты этих векторов в каком-то базисе пространства W. Следовательно, ранг этой системы векторов равен рангу матрицы A. ■

Определение 4. Множество всех линейных отображений $\mathcal{A}: V\to W$ с операциями сложения и умножения на скаляры

$$(\mathcal{A}_1+\mathcal{A}_2)(v):=\mathcal{A}_1v+\mathcal{A}_2v,\quad (\lambda\mathcal{A})(v):=\lambda(\mathcal{A}v)$$

является линейным пространством. Оно называется *пространством линейных отображений* из V в W и обозначается $\operatorname{Hom}_{\mathbb{F}}(V,W)$.

9. Задание линейных отображений (операторов) матрицами. Изменение матрицы линейного отображения при переходе к другим базисам. Нахождение ядра и образа при помощи матрицы

Пусть $\mathcal{A}: V\to W$ — линейное отображение, $e_1,\dots,e_m$ — базис в V, а $f_1,\dots,f_n$ — базис в W.

Определение 1. *Матрицей линейного отображения* $\mathcal{A}: V\to W$ по отношению к базисам $e_1,\dots,e_m$ и $f_1,\dots,f_n$ называется матрица $A=\begin{pmatrix} a_1^1 & \cdots & a_m^1\\ \vdots & \ddots & \vdots\\ a_1^n & \cdots & a_m^n\end{pmatrix}$ размера $n\times m$, в которой i-ый столбец состоит из координат вектора $\mathcal{A}(e_i)$ в базисе $f_1,\dots,f_n$: $\mathcal{A}e_i=a_i^jf_j$.

Предложение 1. Пусть $x=x^je_j$ — произвольный вектор из V, а $y=y^if_i$ — его образ в W, т.е. $y=\mathcal{A}x$. Тогда $y^i=a_j^ix^j$.

Доказательство. Действительно, $y^if_i=y=\mathcal{A}x=\mathcal{A}(x^je_j)=x^j\mathcal{A}e_j=x^ja_j^if_i$. Т.к. $\{f_i\}_{i=1}^n$ — базис, отсюда следует, что $y^i=a_j^ix^j$. ■

Предложение 2. Пусть $\dim V=m$, $\dim W=n$. Тогда $\operatorname{Hom}_{\mathbb{F}}(V,W)\simeq\operatorname{Mat}_{\mathbb{F}}(n,m)$.

Доказательство. Выберем базисы $e_1,\dots,e_m$ и $f_1,\dots,f_n$ в V и W соответственно. Определим отображение $\operatorname{Hom}_{\mathbb{F}}(V,W)\to\operatorname{Mat}_{\mathbb{F}}(n,m)$, которое сопоставляет линейному отображению его матрицу в выбранных базисах. Непосредственно проверяется, что это отображение линейно. Кроме того, оно биективно: обратное отображение сопоставляет матрице $\underset{n\times m}{A}=(a_j^i)$ линейного отображения, определяется в координатах формулой из предыдущего предложения. Следовательно, такое отображение $\operatorname{Hom}_{\mathbb{F}}(V,W)\to\operatorname{Mat}_{\mathbb{F}}(n,m)$ является изоморфизмом. ■

Теорема 1 (Закон изменения матрицы линейного отображения)**.** Имеет место соотношение $A'=D^{-1}AC$, где A — матрица линейного отображения $\mathcal{A}: V\to W$ по отношению к базисам $e_1,\dots,e_m$ и $f_1,\dots,f_n$; A' — матрица линейного отображения $\mathcal{A}$ по отношению к базисам $e_{1'},\dots,e_{m'}$ и $f_{1'},\dots,f_{n'}$; $C=C_{e\to e'}$ — матрица перехода от $e_1,\dots,e_m$ к $e_{m'},\dots,e_{m'}$; $D=D_{f\to f'}$ — матрица перехода от $f_1,\dots,f_n$ к $f_{1'},\dots,f_{n'}$.

Доказательство. Пусть $C=(c_{i'}^i)$, $A=(a_i^j)$, тогда $\mathcal{A}e_{i'}=\mathcal{A}(c_{i'}^ie_i)=c_{i'}^i\mathcal{A}e_i=c_{i'}^ia_i^jf_j$. С другой стороны, если $A'=(a_{i'}^{j'})$ и $D=(d_{j'}^j)$, то $a_{i'}^{j'}f_{j'}=a_{i'}^{j'}d_{j'}^jf_j$. Сравнивая два последних соотношения с учётом того, что ${f_j}_{j=1}^n$ — базис, получаем $a_i^jc_{i'}^i=d_{j'}^ja_{i'}^{j'}$. В матричном виде это эквивалентно $AC=DA'\Rightarrow A'=D^{-1}AC$. ■

Поиск ядра и образа линейного оператора по его матрице. Пусть имеем матрицу A оператора $\mathcal{A}$ в каком-то базисе. Приведя её к ступенчатому виду, сможем найти базис системы столбцов матрицы A (его будут составлять столбцы, в которых есть лидеры). Вспомним, что по столбцам A написаны образы базисных векторов при отображении $\mathcal{A}$, а мы нашли базис этой системы. Это значит, что найденные нами столбцы есть базис $\operatorname{Im}\mathcal{A}$.

$\operatorname{Ker}\mathcal{A}$ — это просто пространство решений СЛУ $Ax = 0$. Чтобы найти базис ядра, нам нужно просто найти её ФСР.

10. Линейные операторы. Изменение матрицы линеного оператора при переходе к новому базису. Подобные матрицы

Определение 1. Линейное отображение $\mathcal{A} : V \to V$ пространства V в себя называется *линейным оператором.*

Базис в V выбирается только один, поэтому изменение матрицы линейного оператора при переходе к новому базису выглядит так, как и в теореме 1 из 9 вопроса при $C = D$.

Определение 2. Матрицы A и B называются *подобными* тогда и только тогда, когда существует $C \in \mathrm{GL}(n)$ такая, что

$$B = C^{-1}AC.$$

Предложение 1. Если A и B подобны, то $\det A = \det B$ $\operatorname{rk} A = \operatorname{rk} B$, $\operatorname{tr} A = \operatorname{tr} B$.

Доказательство. В обозначениях из определения подобия матриц $\det B = \det(C^{-1}AC) = \det(C^{-1}C)\cdot \det A = \det A$. Второе утверждение следует из того, что ранг не меняется при домножении на невырожденную матрицу.

Непосредственным вычислением доказывается, что $\operatorname{tr}(AC) = \operatorname{tr}(CA)$ (теорема из первого семестра). Как следствие, $\operatorname{tr}(C^{-1}AC) = \operatorname{tr} A$. ■

Это значит, что определитель матрицы линейного оператора не меняется при любой замене базиса, таким образом, можно говорить об *определителе*, *следе* и *ранге линейного оператора.*

11. Векторное пространство линейных отображений. Алгебра линейных операторов. Изоморфизм алгебры матриц и алгебры линейных операторов

Утверждение, которое доказывалось в первом семестре:

Предложение 1. Пусть $U \overset{\psi}{\to} V \overset{\varphi}{\to} W$. Если φ, ψ — линейные, то $\varphi \circ \psi$ линейное.

Определение 1. Множество A с операциями сложения, умножения на скаляры и умножения элементов A, удовлетворяющее аксиомам:

1. $(A, +)$ — ассоциативное кольцо;
2. A — векторное пространство (относительно операций сложения и умножения на скаляры);
3. $\lambda(ab) = (\lambda a)b$ $\forall \lambda \in \mathbb{F}$, $\forall a, b \in A$

называеся *линейной алгеброй.*

Теорема 1. Множество всех линейных операторов на V с операциями сложения и умножения на скаляры и композиции является линейной алгеброй. Причём, это алгебра изоморфна алгебре матриц над тем же полем.

Доказательство. Первые две аксиомы уже проверены. Третья проверяется непосредственно:

$$(\lambda(ab))(v) = \lambda((ab)(v)) = \lambda a(b(v)) \Rightarrow \lambda(ab) = (\lambda a)b.$$

Вторая часть утверждения очевидна — мотивировкой именно такого определения произведения матриц было как раз то, что это будет матрица композиции линейных отображений. ■

12. Инвариантные подпространства линейного оператора. Ограничение линейного оператора на инвариантное подпространство. Вид матрицы линейного оператора при наличии инвариантных подпространств

Определение 1. Подпространство $W \subset V$ называется *инвариантным* относительно оператора $\mathcal{A} : V \to V$, если $\mathcal{A}(W) \subseteq W$.

Пример 1. $\operatorname{Ker} \mathcal{A}$ и $\operatorname{Im} \mathcal{A}$ являются инвариантными подпространствами. Пусть $W \subset V$ — инвариантное подпространство для оператора $\mathcal{A} : V \to V$. Выберем базис $e_1, \ldots, e_k$ в W и дополним его до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ в V. Пусть $A = (a^i_j)$ — матрица оператора $\mathcal{A}$ в этом базисе. Тогда $\mathcal{A}e_j = a^1_j e_j + \ldots a^k_j e_k$ при $j = 1, \ldots, k$. Это означает, что матрица A имеет вид $A = \left(\begin{array}{c|c} * & * \\ \hline 0 & * \end{array}\right)$, где в левом нижнем углу стоит матрица размера $(n-k) \times k$ из нулей. Аналогично, если имеет место разложение $V = W_1 \oplus W_2$ в прямую сумму инвариантных подпространств, $\mathcal{A}(W_1) \subset W_1$, $\mathcal{A}(W_2) \subset W_2$, то в подходящем базисе матрица оператора $\mathcal{A}$ будет иметь блочно-диагональный вид: $A = \left(\begin{array}{c|c} * & 0 \\ \hline 0 & * \end{array}\right)$.

Определение 2. Пусть $W \subset V$ — инвариантное подпространство для оператора $\mathcal{A} : V \to V$. Тогда оператор $\widehat{\mathcal{A}} : W \to W$, определённый равенством $\widehat{\mathcal{A}}w : \mathcal{A}w$ для $w \in W$, называется *ограничением* оператора $\mathcal{A}$ на подпространство W и часть обозначается как $\mathcal{A}\big|_W$.

Определение 3. Линейный оператор $\widetilde{\mathcal{A}} : V/W \to V/W$, определённый на классах смежности по правилу $\widetilde{\mathcal{A}}(v + W) = \mathcal{A}v + W$, называется *фактор-оператором.*

Определение фактор-операора корректно. Действительно, если $v + W = u + W$, то $v - u \in W$, $\mathcal{A}(v-u) \in W$, и мы имеем $\widetilde{\mathcal{A}}(v+W) = \mathcal{A}(u+v=u)+W = \mathcal{A}u+\mathcal{A}(v-u)+W = \mathcal{A}u+W = \widetilde{\mathcal{A}}(u+W)$.

Предложение 1. Пусть $W \subset V$ — инвариантное подпространство для оператора $\mathcal{A} : V \to V$. Пусть $e_1, \ldots, e_k$ — базис в W и $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ — базис в V. Тогда матрица оператора $\mathcal{A}$ в этом базисе имеет вид $A = \left(\begin{array}{c|c} \widehat{A} & * \\ \hline 0 & \widetilde{A} \end{array}\right)$, где $\widehat{A}$ — матрица ограничения $\mathcal{A}\big|_W$ в базисе $e_1, \ldots, e_k$, а $\widetilde{A}$ — матрица фактор-оператора в базисе $e_{k+1} + W, \ldots, e_n + W$ факторпространства V/W.

Доказательство. Пусть $A = \begin{pmatrix} a^1_1 & a^1_2 & \ldots & a^1_n \\ a^2_1 & a^2_2 & \ldots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n_1 & a^n_2 & \ldots & a^n_n \end{pmatrix}$. Тогда $\widehat{A} = \begin{pmatrix} a^1_1 & \ldots & a^1_k \\ \vdots & \ddots & \vdots \\ a^k_1 & \ldots & a^k_k \end{pmatrix}$ и $\widetilde{A} = \begin{pmatrix} a^{k+1}_{k+1} & \ldots & a^{k+1}_n \\ \vdots & \ddots & \vdots \\ a^n_{k+1} & \ldots & a^n_n \end{pmatrix}$.
Разложим вектор по базису: $v = v_1 e_1 + \ldots + v_k e_k + \ldots + v_{k+1} e_{k+1} + \ldots + v_n e_n$. Отметим, что в силу инвариантности подпространства W имеет место $A \cdot (v_1, \ldots, v_k, 0, \ldots, 0)^t = (\underbrace{*, \ldots, *}_{k}, 0, \ldots, 0)^t$. Итак,

$$\mathcal{A}v = A \cdot v = \begin{pmatrix} a^1_1 & a^1_2 & \ldots & a^1_n \\ a^2_1 & a^2_2 & \ldots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n_1 & a^n_2 & \ldots & a^n_n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_k \\ v_{k+1} \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a^1_1 & a^1_2 & \ldots & a^1_n \\ a^2_1 & a^2_2 & \ldots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n_1 & a^n_2 & \ldots & a^n_n \end{pmatrix} \cdot \left(\begin{pmatrix} v_1 \\ \vdots \\ v_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_{k+1} \\ \vdots \\ v_n \end{pmatrix} \right) =$$

$$= \left(\begin{array}{c|c} \widehat{A} & 0 \\ \hline 0 & 0 \end{array}\right) \cdot v + \left(\begin{array}{c|c} 0 & * \\ \hline 0 & \widetilde{A} \end{array}\right) \cdot v = \left(\begin{array}{c|c} \widehat{A} & * \\ \hline 0 & \widetilde{A} \end{array}\right) \cdot v.$$

Отсюда и следует, что в выбранном базисе матрица A имеет указанный вид. ■

Примечание. В конспектах Т. Е. Панова написано, что это утверждение примерно очевидно и доказывать мы его не будем. Поэтому изложенное выше доказательство я проделал сам, оно может содержать ошибки.

Следствие 1. Если пространство V представлено в виде прямой суммы ненулевых инвариантных относительно преобразования $\mathcal{A}$ подпространств $V = W_1 \oplus \ldots \oplus W_k$, то существует базис, в котором матрица преобразования имеет блочно-диагональный вид $A = \operatorname{diag}(A_1, \ldots, A_k) = \begin{pmatrix} A_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_k \end{pmatrix}$, где A_i — матрица ограничения $\mathcal{A}\big|_{W_i}$ преобразования $\mathcal{A}$ на подпространство W_i, $i = 1, \ldots, k$.

13. Собственные векторы и собственные значения линейного оператора. Линейная независимость собственных векторов линейного оператора, отвечающих попарно различным собственным значениям

Определение 1. Ненулевой вектор $v \in V$ называется *собственным* для оператора $\mathcal{A}$, если $\mathcal{A}v = \lambda v$ для некоторого $\lambda \in \mathbb{F}$. Число $\lambda \in \mathbb{F}$ называется *собственным значением*, если существует собственный вектор v, для которого $\mathcal{A}v = \lambda v$.

Предложение 1. Собственные векторы, отвечающие попарно различным собственным значениям $\lambda_1, \ldots, \lambda_k$ оператора $\mathcal{A}$, линейно независимы.

Доказательство. Докажем утверждение теоремы индукцией по k. При $k = 1$ доказывать нечего. Пусть $k > 1$ и

$$\mu_1 v_1 + \ldots + \mu_{k-1} v_{k-1} + \mu_k v_k = \mathbf{0} \quad (v_i \in V_{\lambda_i}, (\mu_1, \ldots, \mu_k) \neq (0, \ldots, 0))$$

Применяя оператор $\mathcal{A}$, получаем

$$\lambda_1 \mu_1 v_1 + \ldots + \lambda_{k-1} \mu_{k-1} v_{k-1} + \lambda_k \mu_k v_k = \mathbf{0}.$$

Вычитая отсюда исходное неравенство, умноженное на λ_k, получаем

$$\mu_1(\lambda_1 - \lambda_k) v_1 + \ldots + (\lambda_{k-1} - \lambda_k) v_{k-1} = \mathbf{0},$$

откуда в силу предположения индукции следует равенство нулю всех коэффициентов. Если найдётся $\mu_j \neq 0$, то $\lambda_j - \lambda_{k-1} = 0$, однако из условия собственные значения попарно различны. Противоречие. Значит, все μ_j равны нулю, и система $v_1, \ldots, v_k$ линейно независима. ■

Следствие 1. Собственные подпространства $V_{\lambda_1}, \ldots, V_{\lambda_k}$, соответствующие попарно различным собственным значениям $\lambda_1, \ldots, \lambda_k$ оператора $\mathcal{A}$, образуют прямую сумму $V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$.

14. Вычисление собственных значений и собственных векторов с помощью матрицы. Характеристический многочлен

Определение 1. Многочлен $\chi_{\mathcal{A}}(t) := \det(\mathcal{A} - t \cdot \mathrm{id})$ называется *характеристическим многочленом* оператора $\mathcal{A}$.

Так как характеристический многочлен определяется как детерминант оператора, его можно вычислять как определитель матрицы этого оператора в любом базисе:

$$\chi_{\mathcal{A}}(t) = \det(A - t \cdot E) = \det \begin{pmatrix} a_1^1 - t & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 - t & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n - t \end{pmatrix}.$$

Из этой формулы ясно, что $\deg \chi_{\mathcal{A}}(t) = \dim V$, кроме того, при $\chi_{\mathcal{A}}(t) = p_n t^n + p_{n-1}t^{n-1} + \ldots + p_0$, имеем $p_n = (-1)^n$, $p_{n-1} = (-1)^{n-1} \operatorname{tr} \mathcal{A}$, $p_0 = \det \mathcal{A}$.

Предложение 1. Собственные значения оператора $\mathcal{A}$ — это в точности корни его характеристического многочлена.

Доказательство. Если λ — собственное значение, то оператор $\mathcal{A} - \lambda \cdot \mathrm{id}$ вырожден, т. е. $\det(\mathcal{A} - \lambda \cdot \mathrm{id}) = 0$, а значит, λ — корень многочлена $\chi_{\mathcal{A}}(t)$. Обратно, если $\chi_{\mathcal{A}}(\lambda) = 0$, то $\det(\mathcal{A} - \lambda \cdot \mathrm{id}) = 0$. Поэтому $\operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id}) \neq \{\mathbf{0}\}$, а значит, λ — собственное значение. ■

Поиск собственных векторов. Сначала находим корни характеристического многочлена, т. е. собственные значения. Теперь найдём подпространство V_λ для собственного значения λ. Иными словами, хотим найти все такие векторы, что $\mathcal{A}v = \lambda v$. Это равносильно системе уравнений с матрицей $(A - \lambda E)$. Она и задаёт искомое подпространство.

15. Собственные подпространства. Неравенство между размерностью собственного подпространства и кратностью корня характеристического многочлена

Предложение 1. Все собственные векторы, отвечающие собственным значениям λ, и вектор $\mathbf{0}$ образуют подпространство, которое совпадает с $\operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$.

Доказательство. По определению ядра, равенство $\mathcal{A}v = \lambda v$ имеет место тогда и только тогда, когда $v \in \operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$. ■

Определение 1. Пусть λ — собственное значение для оператора $\mathcal{A}$. Подпространство $V_\lambda := \operatorname{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$ называется *собственным подпространством*, соответствующим λ.

Доказательство. Действительно, если $v \in V_\lambda$, то $\mathcal{A}v = \lambda v \in V_\lambda$. ■

Предложение 2. Размерность собственного подпространства V_λ не превосходит кратности λ как корня характеристического многочлена.

Доказательство. Пусть $\dim V_\lambda = k$. Выберем базис $e_1, \ldots, e_k$ в пространстве V_λ и дополним его до базиса в V. Т. к. $\mathcal{A}e_i = \lambda e_i$ при $i = 1, \ldots, k$, матрица оператора $\mathcal{A}$ в выбранном базисе имеет вид $A = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & \widetilde{A} \end{array}\right)$. Тогда $\chi_{\mathcal{A}}(t) = \det(A - t \cdot E) = (\lambda - t)^k \det(\widetilde{A} - t \cdot E) = (\lambda - t)^k \chi_{\widetilde{\mathcal{A}}}(t)$, где $\widetilde{\mathcal{A}}$ — фактор-оператор. Отсюда вытекает, что кратность корня λ не меньше $k = \dim V_\lambda$. ■

16. Диагонализируемость (матрицы) линейного оператора. Критерии диагонализируемости и достаточное условие

Определение 1. Оператор $\mathcal{A}$ называется *диагонализируемым*, если существует базис, в котором матрица этого оператора диагональна.

По определению матрицы оператора, базис, в котором матрица оператора диагональна, состоит из собственных векторов. Поэтому оператор диагонализируем тогда и только тогда, когда для него существует базис из собственных векторов.

Теорема 1 (Критерий диагонализируемости)**.** Оператор $\mathcal{A}$ в n-мерном пространстве V диагонализируем тогда и только тогда, когда его характеристический многочлен имеет в точности n корней (с учётом кратностей) и размерность каждого собственного подпространства V_λ равна кратности корня λ.

Доказательство. Предположим, что оператор $\mathcal{A}$ диагонализируем. Пусть на диагонали матрицы D оператора $\mathcal{A}$ стоят числа $\lambda_1, \ldots, \lambda_k$, причём число λ_i присутствует ровно r_i раз. Тогда мы имеем $\chi_{\mathcal{A}}(t) = \det(D - t \cdot E) = \prod\limits_{i=1}^{k} (\lambda_i - t)^{r_i}$. Следовательно, многочлен $\chi_{\mathcal{A}}(t)$ имеет $\sum\limits_{i=1}^{k} = n$ корней, и каждому корню λ_i соответствует ровно r_i линейно независимых собственных векторов, т. е. $\dim V_{\lambda_i} = r_i$.

Предположим теперь, что многочлен $\chi_{\mathcal{A}}(t)$ имеет различные корни $\lambda_1, \ldots, \lambda_k$, причём кратность корня λ_i равна r_i, $\sum\limits_{i=1}^{k} r_i = n$ и $\dim V_{\lambda_i} = r_i$. Согласно следствию 1 из вопроса 13, пространства $V_{\lambda_1}, \ldots, V_{\lambda_k}$ образуют прямую сумму, а по условию сумма их размерностей равна $n = \dim V$. Следовательно, $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$. Выбрав базис в каждом из подпространств $\{V_{\lambda_i}\}_{i=1}^{k}$ и взяв объединение этих базисов, мы получим базис пространства V, состоящий из собственных векторов. Итак, оператор $\mathcal{A}$ диагонализируем. ■

Следствие 1 (Достаточное условие диагонализируемости)**.** Если характеристический многочлен $\chi_{\mathcal{A}}(t)$ имеет $n = \dim V$ различных корней, то оператор $\mathcal{A}$ диагонализируем.

Примечание. Панов пишет, что набор собственных значений оператора $\mathcal{A}$ часто называют его *спектром* (будет в курсе функционального анализа). Если все собственные значения имеют кратность 1 как корни характеристического многочлена, то говорят о *простом спектре*. Таким образом, операторы с простым спектром диагонализируемы. Появление кратных корней является «особенностью», которая устраняется произвольно малым возмущением коэффициентов матрицы оператора. Поэтому над полем $\mathbb{C}$ «почти все операторы диагонализируемы».

К последней фразе лично я отношусь скептически, ведь (как заметил Рамиль Хакамов) любое действительное число можно произвольно малым возмущением сделать рациональным, хотя язык не повернётся сказать, что «почти все действительные числа рациональные».

UPD. Почитал, оказалось, что подмножество матриц, не являющихся диагонализируемыми над $\mathbb{C}^{n\times n}$, имеет нулевую меру Лебега. То есть, по факту это «почти никакие». Можно ещё сказать, что диагонализируемые матрицы образуют всюду плотное подмножество в топологии Зарисского (замкнутыми называются множества нулей многочленов из некоторого выделенного множества), ведь дополнение к этому подмножеству лежит в множестве, в котором дискриминант характеристического уравнения обнуляется, т. е. на гиперповерхности. Над $\mathbb{R}$ это не выполняется.

Если интересно почитать про геометрию дискриминанта, можно это сделать ЗДЕСЬ.

Пример 1.

1. Оператор, заданный матрицей $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ в стандартном базисе в $\mathbb{R}^2$ не диагонализируем, т. к. его характеристический многочлен $t^2 + 1$ не имеет вещественных корней. Однако тот же оператор над $\mathbb{C}$ диагоналазируем: в базисе $f_1 = (1, i)$, $f_2 = (1, -i)$ его матрица $\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$ диагональна.
2. Оператор, заданный матрицей $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ не диагонализируем ни над каким полем по другой причине: его характеристический многочлен $(t-1)^2$ имеет корень 1 кратности 2, но при этом размерность соответствующего собственного подпространства равна 1 (вектор $e_2 = (0, 1)$ не является собственным).

17. Аннулирующие многочлены линейного оператора (матрицы). Минимальный многочлен. Критерий диагонализируемости в терминах минимального многочлена

Пусть $\mathcal{A} : V \to V$ — оператор. Каждому многочлену $P(t) = p_o + p_1 t + \ldots + p_n t^n \in \mathbb{F}[t]$ можно сопоставить оператор $P(\mathcal{A}) := p_0 \cdot \mathrm{id} + p_1 \cdot \mathcal{A} + \ldots + p_n \mathcal{A}^n$.

Определение 1. Такой оператор называется *многочленом от оператора* $\mathcal{A}$.

Определение 2. Многочлен $P(t)$ называют *аннулирующим* оператор $\mathcal{A}$, если $P(\mathcal{A}) = 0$ (нулевой оператор).

Пример 1.

1. Многочлен $t-1$ аннулирует оператор id.
2. Многочлен $t^2 - t$ аннулирует любой проектор (потом будет написано, почему).

Предложение 1. У любого оператора существует ненулевой аннулирующий многочлен.

Доказательство. Пусть $\dim V = n$. Рассмотрим n^2+1 операторов: $\mathcal{A}^0 = \mathrm{id}, \mathcal{A}^1 = \mathcal{A}, \mathcal{A}^2, \ldots, \mathcal{A}^{n^2}$. Т. к. размерность пространства операторов равна n^2, эти операторы линейно зависимы, т. е. существуют числа $p_0, p_1, p_2, \ldots, p_{n^2}$, не все равные нулю, такие, что $p_o \cdot \mathrm{id} + p_1 \cdot \mathcal{A} + p_2 \cdot \mathcal{A}^2 + \ldots + p_{n^2} \cdot \mathcal{A}^{n^2} = 0$. Тогда ненулевой многочлен $P(t) = p_0 + p_1 t + \ldots + p_{n^2} t^{n^2}$ аннулирует $\mathcal{A}$. ■

Определение 3. Ненулевой многочлен со старшим коэффициентом 1 минимальной возможной степени среди аннулирующих многочленов оператора $\mathcal{A}$ называется *минимальным многочленом* для оператора $\mathcal{A}$.

Лемма 1. Пусть $\mathcal{A} : V \to V$ — линейный оператор конечномерного векторного пространства V. Тогда

1. минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в любом базисе;
2. минимальный многочлен является делителем любого аннулирующего многочлена;
3. минимальный многочлен единственный;
4. наборы корней у минимального многочлена и характеристического многочлена совпадают.

Доказательство.

1. Из доказательства теоремы 1 из вопроса 11 следует, что если $f(t) \in \mathbb{F}[t]$ — произвольный многочлен с коэффициентами из $\mathbb{F}$, то для матрицы A линейного оператора $\mathcal{A}$ в произвольном базисе выполняется $f(A) = 0 \Leftrightarrow f(\mathcal{A}) = 0$. Откуда и получаем первое утверждение.

2. Пусть $f(t)$ — произвольный аннулирующий многочлен для $\mathcal{A}$. Поделим многочлен $f(t)$ на минимальный многочлен $\mu_{\mathcal{A}}(t)$ линейного оператора $\mathcal{A}$ с остатком: $f(t) = \mu_{\mathcal{A}}(t) q(t) + r(t)$, при этом для остатка $r(t)$ должно выполняться $\deg r < \deg \mu_{\mathcal{A}}$. Подставляя $\mathcal{A}$ вместо t, получаем
$$r(\mathcal{A}) = f(\mathcal{A}) - \mu_{\mathcal{A}}(\mathcal{A}) q(\mathcal{A}) = 0 - 0 \cdot q(\mathcal{A}) = 0,$$
т. е. $r(t)$ аннулирует $\mathcal{A}$ и $\deg r < \deg \mu_{\mathcal{A}}$, поэтому $r(t) = 0$, т. е. деление происходит нацело.

3. Любой из двух минимальных многочленов делит другой, а их старшие коэффициенты совпадают.

4. Т. к. характеристический многочлен $\chi_{\mathcal{A}}(t)$ линейного оператора $\mathcal{A}$ является его аннулирующим многочленом (по теореме Гамильтона — Кэли, будет позднее), из п. 1 получаем, что $\mu_{\mathcal{A}}$ делит $\chi_{\mathcal{A}}$. Значит, корни минимального многочлена являются корнями характеристического многочлена. Обратно, пусть λ — корень характеристического многочлена. Тогда у $\mathcal{A}$ есть собственный вектор v с собственным значением λ. Применим к вектору v оператор $\mu_{\mathcal{A}}$. С одной стороны, расписав минимальный многочлен явно:
$$\mu_{\mathcal{A}} = t^k + b_{k-1} t^{k-1} + \ldots + b_1 t + b_0$$
и подставив линейный оператор $\mathcal{A}$ вместо переменной t, получаем
$$\begin{aligned} \mu_{\mathcal{A}}(\mathcal{A}) v &= (\mathcal{A}^k + b_{k-1}\mathcal{A}^{k-1} + \ldots + b_1 \mathcal{A} + b_0 \,\mathrm{id}) v = \\ &= \mathcal{A}^k v + b_{k-1} \mathcal{A}^{k-1} v + \ldots + b_1 \mathcal{A} v + b_0 \,\mathrm{id}\, v = \lambda^k v + b_{k-1} \lambda^{k-1} v + \ldots + b_1 \lambda v + b_0 v = \\ &= (\lambda^k + b_{k-1}\lambda^{k-1} + \ldots + b_1 \lambda + b_0) v = \mu_{\mathcal{A}}(\lambda) v. \end{aligned}$$

С другой стороны, т. к. $\mu_\mathcal{A}(\mathcal{A}) = 0$, имеем $\mu_\mathcal{A}(\mathcal{A})v = 0v = 0$. Откуда $\mu_\mathcal{A}(\lambda)v = 0$. Т. к. $v \neq 0$, то число $\mu_\mathcal{A}(\lambda)$ равно нулю, т. е. λ — корень минимального многочлена.

$\blacksquare$

Теорема 1 (Критерий диагонализируемости в терминах минимального многочлена)**.** Линейный оператор $\mathcal{A} : V \to V$ конечномерного векторного пространства V над алгебраически замкнутым полем $\mathbb{F}$ диагонализируем тогда и только тогда, когда его минимальный многочлен не имеет кратных корней.

Доказательство. Пусть сначала $e_1, \ldots, e_n$ — базис из собственных векторов для $\mathcal{A}$, причём $\mathcal{A}(e_i) = \lambda_i e_i$, т. е. матрица оператора $\mathcal{A}$ в выбранном базисе диагональна:

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Т. к. минимальный многочлен линейного оператора равен минимальному многочлену его матрицы в произвольном базисе, то $\mu_\mathcal{A}(t) = (t-\alpha_1)\ldots(t-\alpha_p)$, где $\alpha_1, \ldots, \alpha_p$ — все попарно различные числа $\lambda_1, \ldots, \lambda_n$. Это следует из того, что выписанный многочлен является аннулирующим для матрицы A, а минимальный многочлен делит любой аннулирующий многочлен и наборы корней у минимального и характеристического многочлена совпадают.

Пусть теперь минимальный многочлен не имеет кратных корней. И пусть J — жорданова нормальная форма оператора $\mathcal{A}$, которая существует в силу алгебраической замкнутости поля $\mathbb{F}$. Позже будет написано, как искать минимальный многочлен для жордановой матрицы, а сейчас воспользуемся этим результатом:

$$\mu_\mathcal{A}(t) = \mu_J(t) = (t-\alpha_1)^{m_1}\ldots(t-\alpha_p)^{m_p},$$

где $\alpha_1, \ldots, \alpha_p$ — все попарно различные собственные значения линейного оператора $\mathcal{A}$, m_i — максимальный порядок жордановой клетки с α_i на диагонали в жордановой нормальной форме J. Т. к. минимальный многочлен не имеет кратных корней, получаем, что $m_i = 1$ для всех i, т. е. в J все клетки имеют порядок 1. Это и означает, что жорданова нормальная форма диагональна. $\blacksquare$

Задача 2 (MATH.STACKEXCHANGE)**.** Пусть $\mathbb{K}$ — поле и $\mathbb{F} \supseteq \mathbb{K}$ — его расширение. Пусть также $A \in \underset{n\times n}{\mathrm{Mat}}(\mathbb{K})$ и $\mu_{A,\mathbb{K}}$ — минимальный многочлен A. Заметим, что эту матрицу A можно также рассмотреть как матрицу из $\underset{n\times n}{\mathrm{Mat}}(\mathbb{F})$, и пусть в этом случае её минимальный многочлен есть $\mu_{A,\mathbb{F}}$. Доказать, что $\mu_{A,\mathbb{K}} = \mu_{A,\mathbb{F}}$.

Нам понадобится две леммы.

Лемма 2. Если $A \in \underset{n\times n}{\mathrm{Mat}}(\mathbb{K})$, то $\deg \mu_{A,\mathbb{K}} = d$ тогда и только тогда, когда выполнены следующие два условия:

1. матрицы $E, A, A^2, \ldots, A^{d-1}$ линейно независимы над полем $\mathbb{K}$;
2. матрицы $E, A, A^2, \ldots, A^{d-1}, A^d$ линейно зависимы над полем $\mathbb{K}$;

Доказательство. $\Rightarrow$. Пусть $\mu_{A,\mathbb{K}}(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_d t^d$. Тогда $\mu_{A,\mathbb{K}}(A) = 0$, т. е.

$$a_0 E + a_1 A + a_2 A^2 + \ldots + a_d A^d = 0.$$

Причём, не все a_i равны нулю. Значит, выполнено условие 2. А условие 1 выполнено в силу минимальности $\mu_{A,\mathbb{K}}$. Действительно, если $E, A, A^2, \ldots, A^{d-1}$ линейно зависимы, то коэффициенты линейной зависимости можно выбрать в качестве коэффициентов аннулирующего многочлена степени $d-1$, а такого нет, потому что $\deg \mu_{A,\mathbb{K}}$ минимальна для аннулирующих многочленов.

$\Rightarrow$. Эти условия ровно то и означают: существует аннулирующий многочлен степени d, но не существует аннулирующего многочлена степени $d-1$. ■

Лемма 3. Есть $\mathbb{F} \supseteq \mathbb{K}$ — расширение поля и $v_1, \ldots, v_r \in \mathbb{K}^N \subseteq \mathbb{F}^N$. Тогда $v_1, \ldots, v_r$ линейно зависимы над полем $\mathbb{K}$ тогда и только тогда, когда они линейно зависимы над полем $\mathbb{F}$.

Доказательство. $\Rightarrow$. Очевидно, т. к. коэффициенты этой линейной зависимости лежат как в $\mathbb{K}$, так и в $\mathbb{F}$ в силу $\mathbb{K} \subseteq \mathbb{F}$.

$\Leftarrow$. Пусть $v_i = (v_1^i, v_1^i, \ldots, v_N^i)^t$ и пусть $x_1, \ldots, x_r \in \mathbb{F}^N$ — коэффициенты линейной зависимости $v_1, \ldots, v_r$. Тогда

$$x_1 \begin{pmatrix} v_1^1 \\ \vdots \\ v_N^1 \end{pmatrix} + x_2 \begin{pmatrix} v_1^2 \\ \vdots \\ v_N^2 \end{pmatrix} + \ldots + x_r \begin{pmatrix} v_1^r \\ \vdots \\ v_N^r \end{pmatrix} = \mathbf{0}.$$

Получили систему уравнений с матрицей коэффициентов $V = (v_j^i)_{j=1,\ldots,N}^{i=1,\ldots r}$. Система имеет ненулевое решение, значит, $\det V \neq 0$ над $\mathbb{F}$. Причём, $V \in \underset{N \times r}{\mathrm{Mat}}(\mathbb{K})$, поэтому $\det V \neq 0$ и над $\mathbb{K}$. А значит, у неё есть ненулевое решение из $\mathbb{K}$. ■

Теперь можно и задачу решить.

$\triangleright$ Очевидно, что $\mu_{A,\mathbb{F}}$ делит $\mu_{A,\mathbb{K}}$. Так как их старшие коэффициенты равны, достаточно показать, что их степени равны. А это очевидное следствие из двух предыдущих лемм. ◄

18. Теорема Гамильтона — Кэли

Теорема 1 (Гамильтон, Кэли)**.** Характеристический многочлен $\chi_{\mathcal{A}}$ оператора $\mathcal{A}: V \to V$ аннулирует этот оператор, т. е. $\chi_{\mathcal{A}}(\mathcal{A}) = 0$.

Доказательство. Для квадратно матрицы M обозначим через $\widehat{M}$ транспонированную матрицы алгебраических дополнений. Из курса алгебры, $M \cdot \widehat{M} = \det M \cdot E$. Теперь возьмём в качестве M матрицу $A - t \cdot E$, где A — матрица оператора $\mathcal{A}$ в произвольном базисе. Тогда $(A - t \cdot E)(\widehat{A - t \cdot E}) = \det(A - t \cdot E) \cdot E = \chi_{\mathcal{A}}(t) \cdot E$. По определению, элементы матрицы $\widehat{A - t \cdot E}$ являются многочленами степени не выше $n - 1$ ($n = \dim V$). Следовательно, эту матрицу можно записать в виде $\widehat{A - t \cdot E} = B_0 + t \cdot B_1 + t^2 \cdot B_2 + \ldots + t^{n-1} B_{n-1}$, где B_i — числовые матрицы. Подставив это разложение вместе с разложение $P_{\mathcal{A}}(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n$ в формулу выше, получим

$$(A - t \cdot E)(B_0 + t \cdot B_1 + t^2 \cdot B_2 + \ldots + t^{n-1} \cdot B_{n-1}) = (a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n) \cdot E.$$

Приравнивая коэффициенты при степенях t и складывая полученные уравнения, получим:

$$\left.\begin{array}{r} AB_0 = a_0 E \\ -B_0 + AB_1 = a_1 E \\ -B_1 + AB_2 = a_2 E \\ \vdots \\ -B_{n-2} + AB_{n-1} = a_{n-1} E \end{array} \right| \begin{array}{l} \cdot A \\ \cdot A^2 \\ \cdot A^3 \\ \vdots \\ \cdot A^n \end{array} \overset{+}{\Rightarrow} 0 = \underbrace{a_0 E + a_1 A + a_2 A^2 + \ldots + a_n A^n}_{\chi_{\mathcal{A}}(A) = \chi_{\mathcal{A}}(\mathcal{A})}$$

■

19. Существование одномерного или двумерного инвариантного подпространства для любого линейного оператора над полем действительных чисел

Бо́льшая часть этого вопроса — рассказ про овеществление и комплексификацию линейных пространств (взято из Панова и Винберга). При работе с линейными операторами часто бывает удобно изменить поле скаляров. Здесь рассмотрим две такие операции: переход от $\mathbb{R}$ к $\mathbb{C}$ (комплексификация) и переход от $\mathbb{C}$ к $\mathbb{R}$ (овеществление).

Определение 1. Пусть V — линейное пространство над полем $\mathbb{C}$. Рассмотрим множество $V_{\mathbb{R}}$, состоящее из тех же векторов, что и V. На $\mathbb{R}$ имеется операция сложения (та же, что и на V), а вместо операции умножения на все комплексные числа оставим лишь умножение на все вещественные числа. Тогда $V_{\mathbb{R}}$ — линейное пространство над полем $\mathbb{R}$, которое называется *овеществлением пространства* V.

Предложение 1. Пусть $e_1, \ldots, e_n$ — базис пространства V. Тогда $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ — базис пространства $V_{\mathbb{R}}$.

Доказательство. Проверим, что векторы $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ линейно независимы в $V_{\mathbb{R}}$. Пусть $\lambda_1 e_1 + \ldots + \lambda_n e_n + \mu_1 ie_1 + \ldots + \mu_n ie_n = \mathbf{0}$ в пространстве $V_{\mathbb{R}}$, где $\lambda_k, \mu_k \in \mathbb{R}$, $k = 1, \ldots, n$. Тогда в пространстве V мы имеем $(\lambda_1 + i\mu_1)e_1 + \ldots + (\lambda_n + i\mu_n)e_n = \mathbf{0}$. Т. к. векторы $e_1, \ldots, e_n$ линейно независимы в V, то $\lambda_k + i\mu_k = 0 \Rightarrow \lambda_k = \mu_k = 0$, $k = 1, \ldots, n$. Следовательно, векторы $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ линейно независимы в $V_{\mathbb{R}}$. Теперь проверим, что эти векторы порождают всё пространство $V_{\mathbb{R}}$. Возьмём $v \in V_{\mathbb{R}}$ и рассмотрим его как вектор из V. Т. к. $e_1, \ldots, e_n$ — базис в V, то $v = \alpha_1 e_1 + \ldots + \alpha_n e_n$, где $(\alpha_k = \lambda_k + i\mu_k) \in \mathbb{C}$ $(k = 1, \ldots, n)$, где $\lambda_k, \mu_k \in \mathbb{R}$. Тогда $v = \lambda_1 e_1 + \ldots + \lambda_n e_n + \mu_1 ie_1 + \ldots + \mu_n ie_n$. ■

Следствие 1. $\dim V_{\mathbb{R}} = 2 \dim V$.

Определение 2. Пусть V — комплексное пространство и $\mathcal{A} : V \to V$ — оператор. Тогда тот же оператор, рассматриваемый в пространстве $V_{\mathbb{R}}$ называется *овеществлением оператора* $\mathcal{A}$ и обозначается $\mathcal{A}_{\mathbb{R}}$.

Предложение 2. Запишем матрицу оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_n$ пространства V в виде $A+iB$, где A и B — вещественные матрицы. Тогда

1. матрица оператора $\mathcal{A}_{\mathbb{R}}$ в базисе $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ есть $\left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array}\right)$;
2. $\det \mathcal{A}_{\mathbb{R}} = |\det \mathcal{A}|^2$.

Доказательство.

1. Пусть $A = (a_k^\ell)$ и $B = (b_k^\ell)$. Тогда $\mathcal{A}_{\mathbb{R}}(e_k) = \mathcal{A}(e_k) = (a_k^\ell + ib_k^\ell)e_\ell = a_k^\ell e_\ell + ib_k^\ell e_\ell$, $\mathcal{A}_{\mathbb{R}}(ie_k) = \mathcal{A}(ie_k) = i\mathcal{A}(e_k) = i(a_k^\ell + ib_k^\ell)e_\ell = -b_k^\ell e_\ell + a_k^\ell ie_\ell$.
2. Заметим, что
$$\left(\begin{array}{c|c} A & -B \\ \hline B & A \end{array}\right) \rightsquigarrow \left(\begin{array}{c|c} A - iB & -B - iA \\ \hline B & A \end{array}\right) \rightsquigarrow$$
$$\left(\begin{array}{c|c} A - iB & -B - iA + i(A - iB) \\ \hline B & A + iB \end{array}\right) \rightsquigarrow \left(\begin{array}{c|c} A - iB & 0 \\ \hline B & A + iB \end{array}\right).$$
Отсюда получаем $\det \mathcal{A}_{\mathbb{R}} = \det(A - iB)\det(A + iB) = \overline{\det \mathcal{A}} \det \mathcal{A} = |\det \mathcal{A}|^2$.

■

Определение 3. Пусть V — вещественное пространство. *Комплексной структурой* на V называется такой оператор $\mathcal{J} : V \to V$, что $\mathcal{J}^2 = -\operatorname{id}$.

Пусть V — вещественное пространство с комплексной структурой $\mathcal{J}$. Введём на V операцию умножения на комплексные числа по правилу $(\lambda + i\mu)v = \lambda v + \mu\mathcal{J}v$. Тогда V превращается в комплексное пространство $\widetilde{V}$, для которого $\widetilde{V}_{\mathbb{R}} = V$, а овеществление оператора умножения на i есть $\mathcal{J}$. Проверка выполняется непосредственно.

Предложение 3. Пусть $\mathcal{J}$ – комплексная структура на V. Тогда:

1. размерность вещественного пространства V чётна;
2. в подходящем базисе матрица оператора $\mathcal{J}$ имеет вид $\left(\begin{array}{c|c} 0 & -E \\ \hline E & 0 \end{array}\right)$.

Доказательство. Т. к. любой базис V порождает $\widetilde{V}$, то это пространство конечномерно. А т. к. $V = \widetilde{V}_\mathbb{R}$, то $\dim V = 2\dim\widetilde{V}$ — чётно. Далее, если $e_1, \ldots, e_n$ — базис комплексного пространства $\widetilde{V}$, то $e_1, \ldots, e_n, ie_1, \ldots, ie_n$ — базис пространства $\widetilde{V}_\mathbb{R} = V$. В этом базисе оператор $\mathcal{J}$ (овеществление оператора умножения на i) имеет указанный вид. Это прямое следствия записей сверху. ■

Определение 4. Пусть V — линейное пространство над полем $\mathbb{R}$. Рассмотрим пространство $V \oplus V$, состоящее из пар (u, v), где $u, v \in V$, и введём на нём комплексную структуру следующим образом: $\mathcal{J}(u, v) := (-v, u)$. Получаемое пространство $\widetilde{V \oplus V}$ над полем $\mathbb{C}$ называется *комплексификацией* пространства V и обозначается $V_\mathbb{C}$.

Предложение 4. Пусть $e_1, \ldots, e_n$ — базис пространства V. Тогда векторы $(e_1, \mathbf{0}), \ldots, (e_n, \mathbf{0})$ образуют базис пространства $V_\mathbb{C}$.

Доказательство. Проверим линейную независимость: пусть $\alpha_1(e_1, \mathbf{0}) + \ldots + \alpha_n(e_n, \mathbf{0}) = (\mathbf{0}, \mathbf{0})$ для некоторых $\alpha_k = \lambda_k + i\mu_k \in \mathbb{C}$, $k = 1, \ldots, n$. Выкладка:

$$(\lambda_k + i\mu_k)(e_k, \mathbf{0}) = \lambda_k(e_k, \mathbf{0}) + \mu_k\mathcal{J}(e_k, \mathbf{0}) = (\lambda_k e_k, \mu_k e_k).$$

Подставляя это в линейную комбинацию выше, получаем $(\lambda_1 e_1 + \ldots + \lambda_n e_n, \mu_1 e_1, \ldots, \mu_n e_n) = \mathbf{0}, \mathbf{0}$. Из линейной независимости векторов $e_1, \ldots, e_n$ в V, получаем $\lambda_k = \mu_k = 0$ для всех k. А значит, $\alpha_1 = \ldots = \alpha_n = 0$. Итак, $(e_1, \mathbf{0}), \ldots, (e_n, \mathbf{0})$. То, что они порождают всё пространство, проверяется аналогично. ■

Следствие 2. $\dim V_\mathbb{C} = \dim V$.

Определение 5. Пусть V — пространство над $\mathbb{R}$ и $\mathcal{A} : V \to V$ — оператор. Оператор $\mathcal{A}_\mathbb{C} : V_\mathbb{C} \to V_\mathbb{C}$, заданный формулой $\mathcal{A}_\mathbb{C}(u, v) := (\mathcal{A}u, \mathcal{A}v)$, называется *комплексификацией оператора* $\mathcal{A}$.

Предложение 5. Пусть A — матрица оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_n$. Тогда оператор $\mathcal{A}_\mathbb{C}$ в базисе $(e_1, \mathbf{0}), \ldots, (e_n, \mathbf{0})$ задаётся той же матрицей A.

Доказательство. $\mathcal{A}_\mathbb{C}(e_k, \mathbf{0}) = (\mathcal{A}e_k, \mathbf{0}) = (a_k^\ell e_\ell, \mathbf{0}) = a_k^\ell(e_\ell, \mathbf{0})$. ■

Отметим, что при работе с комплексифицированным пространством $V_\mathbb{C}$ удобно записывать векторы $(u, v) \in V_\mathbb{C}$ в виде $u + iv$. Тогда действие комплексифицированного оператора $\mathcal{A}_\mathbb{C}$ записывается как $\mathcal{A}_\mathbb{C}(u + iv) = \mathcal{A}u + i\mathcal{A}v$.

Предложение 6. Пространство $(V_\mathbb{C})_\mathbb{R}$ канонически изоморфно $V \oplus V$.

Доказательство. Действительно, $V_\mathbb{C} = \widetilde{V \oplus V}$, а $(\widetilde{V \oplus V})_\mathbb{R} = V \oplus V$ (мы просто сначала добавили, а потом убрали домножение на комплексные скаляры). ■

Предложение 7. $(V_\mathbb{R})_\mathbb{C}$ канонически изоморфно $V \oplus \overline{V}$, где $\overline{V}$ — комплексно сопряжённое пространство, в котором сложение то же, что и в V, а умножение $*$ на скаляры определяется как $\lambda * v := \overline{\lambda}v$.

Доказательство. Так как $\mathbb{C}$ — алгебраически замкнутое поле, комплексная структура $\mathcal{J}$ должна иметь $n = \dim(V_\mathbb{R})_\mathbb{C}$ собственных значений, каждое из которых удовлетворяет $\lambda^2 = -1$, т. е. $\lambda = \pm 1$. Значит, можем записать $(V_\mathbb{R})_\mathbb{C} = V^+ \oplus V^-$, где V^+ и V^- — собственные подпространства для $+i$ и $-i$ соответственно. Первое из них изоморфно V, а второе — $\overline{V}$. ■

Примечание. Предложение выше Панов оставил в качестве задачи. Я привёл здесь своё решение, оно может содержать ошибки.

Оператор $\mathcal{A}: V \to V$ в нетривиальном пространстве над полем $\mathbb{C}$ имеет инвариантное подпространство размерности 1. Действительно, т. к. поле $\mathbb{C}$ алгебраически замкнуто, характеристический многочлен $\chi_{\mathcal{A}}(t)$ имеет корень λ и собственный вектор $v \in V_\lambda$. Подпространство $\langle v \rangle$ собственное размерности 1.

Теорема 1. Оператор $\mathcal{A}: V \to V$ в нетривиальном пространстве над полем $\mathbb{R}$ имеет инвариантное подпространство размерности 1 или 2.

Доказательство. Если характеристический многочлен $\chi_{\mathcal{A}}(t)$ имеет вещественный корень, то (аналогично) мы получаем одномерное инвариантное подпространство. Предположим, что $\chi_{\mathcal{A}}(t)$ не имеет вещественных корней. Пусть $\lambda + i\mu$ — комплексный корень, $\mu \neq 0$. Тогда $\lambda + i\mu$ — собственное значение комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ (напомним, что в подходящих базисах матрицы операторов $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ совпадают). Возьмём соответствующий собственный вектор $u + iv \in V_{\mathbb{C}}$. Тогда $\mathcal{A}u + i\mathcal{A}v = \mathcal{A}_{\mathbb{C}}(u + iv) = (\lambda + i\mu)(u + iv) = (\lambda u - \mu v) + i(\mu u + \lambda v)$. Следовательно, $\mathcal{A}u = \lambda u - \mu v$, а $\mathcal{A}v = \mu u + \lambda v$, и линейная оболочка $\langle u, v \rangle$ является инвариантным подпространством для $\mathcal{A}$. ■

Следствие 1. Над полем $\mathbb{R}$ любой линейный оператор приводим к блочно-диагональному виду, причём блоки имеют порядок не выше двух.

20. Корневые подпространства. Разложение пространства в прямую сумму корневых подпространств

Определение 1. Вектор $v \in V$ называется *корневым вектором* оператора $\mathcal{A}$, отвечающим числу $\lambda \in \mathbb{F}$, если существует такое m, что $(\mathcal{A} - \lambda \cdot \mathrm{id})^m v = \mathbf{0}$.

Обозначим через R_λ множество всех корневых векторов, отвечающих λ.

Предложение 1. R_λ является подпространством в V.

Доказательство. Пусть $u, v \in R_\lambda$, т. е. $(\mathcal{A} - \lambda \cdot \mathrm{id})^\ell u = (\mathcal{A} - \lambda \cdot \mathrm{id})^m v = \mathbf{0}$ для некоторых ℓ и m. Тогда $(\mathcal{A} - \lambda \cdot \mathrm{id})^\ell (\mu u) = \mathbf{0}$ $\forall \mu \in \mathbb{F}$ и $(\mathcal{A} - \lambda \cdot \mathrm{id})^{\max\{\ell, m\}}(u + v) = \mathbf{0}$. ■

Определение 2. Подпространство $R_\lambda \subseteq V$ называется *корневым подпространством* для оператора $\mathcal{A}$, отвечающим λ.

Предложение 2. Подпространство R_λ нетривиально тогда и только тогда, когда λ — собственное значение оператора $\mathcal{A}$. При этом $V_\lambda \subseteq R_\lambda$.

Доказательство. Действительно, если λ — собственное значение, то существует $v \neq \mathbf{0}$ такой, что $(\mathcal{A} - \lambda \cdot \mathrm{id})^1 v = \mathbf{0}$, т. е. $v \in R_\lambda$ и R_λ нетривиально. Отсюда же следует, что $V_\lambda \subseteq R_\lambda$.

Обратно, пусть R_λ содержит $u \neq \mathbf{0}$ такой, что $(\mathcal{A} - \lambda \cdot \mathrm{id})^m u = \mathbf{0}$, причём m минимально, т. е. $v := (\mathcal{A} - \lambda \cdot \mathrm{id})^{m-1} u \neq \mathbf{0}$. Тогда $(\mathcal{A} - \lambda \cdot \mathrm{id}) v = (\mathcal{A} - \lambda \cdot \mathrm{id})^m u = \mathbf{0}$, т. е. v — собственный вектор, отвечающий λ. ■

Далее будем рассматривать только нетривиальные корневые подпространства.

Теорема 1. Пусть $\mathcal{A}$ — оператор в пространстве V над алгебраически замкнутым полем, и пусть $\lambda_1, \ldots, \lambda_k$ — все собственные значения оператора $\mathcal{A}$. Тогда $V = R_{\lambda_1} \oplus \ldots \oplus R_{\lambda_k}$.

Доказательство будет опираться на три леммы.

Лемма 1. Подпространство R_λ инвариантно относительно любого оператора $\mathcal{A} - \mu \cdot \mathrm{id}$ (в частности, относительно $\mathcal{A}$). Ограничение $(\mathcal{A} - \mu \cdot \mathrm{id})\big|_{R_\lambda} : R_\lambda \to R_\lambda$ при $\lambda \neq \mu$ является обратимым, а при $\lambda = \mu$ нильпотентным оператором.

Доказательство. Пусть $v \in R_\lambda$, т. е. $(\mathcal{A} - \lambda \cdot \mathrm{id})v = \mathbf{0}$. Тогда

$$(\mathcal{A} - \lambda \cdot \mathrm{id})^m (\mathcal{A} - \mu \cdot \mathrm{id}) v = (\mathcal{A} - \mu \cdot \mathrm{id}) \underbrace{(\mathcal{A} - \mu \cdot \mathrm{id})^m v}_{\mathbf{0}} = \mathbf{0},$$

т. к. многочлены от оператора коммутируют. Итак, R_λ является $(\mathcal{A}-\mu\cdot\mathrm{id})$-инвариантным подпространством, и мы можем рассмотреть ограничение $(\mathcal{A}-\mu\cdot\mathrm{id})\big|_{R_\lambda}$. Пусть $v\in\mathrm{Ker}(\mathcal{A}-\mu\cdot\mathrm{id})\big|_{R_\lambda}$, т. е. $v\in R_\lambda$ и $\mathcal{A}v=\mu v$. Тогда $(\mathcal{A}-\lambda\cdot\mathrm{id})^m v=\mathbf{0}$ для некоторого m и $(\mathcal{A}-\lambda\cdot\mathrm{id})v=(\mu-\lambda)v$, а значит, $(\mu-\lambda)^m v=(\mathcal{A}-\lambda\cdot\mathrm{id})^m v=\mathbf{0}$. Следовательно, при $\mu\neq\lambda$ имеем $v=\mathbf{0}$. Тогда ядро $(\mathcal{A}-\mu\cdot\mathrm{id})\big|_{R_\lambda}$ тривиально, и этот оператор инъективен, а значит, обратим.

Наконец, если $e_1,\ldots,e_r$ — базис в R_λ и $(\mathcal{A}-\lambda\cdot\mathrm{id})^{m_i}e_i=\mathbf{0}$, то $(\mathcal{A}-\lambda\cdot\mathrm{id})^m v=\mathbf{0}$, $\forall v\in V$, где $m=\max\{m_1,\ldots,m_r\}$. ■

Лемма 2. Корневые подпространства $R_{\lambda_1},\ldots,R_{\lambda_k}$, соответствующие различным собственным значениям $\lambda_1,\ldots,\lambda_k$ образуют прямую сумму.

Доказательство. Проведём индукцию по k При $k=1$ доказывать нечего. Предположим, что утверждение доказано для $k-1$ подпространств. Докажем, что соотношение $(*)$: $v_1+\ldots+v_k=\mathbf{0}$, где $v_i\in R_{\lambda_i}$, влечёт $v_1=\ldots=v_k=\mathbf{0}$. Имеем $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p v_k=\mathbf{0}$ для некоторого p. Применив к $(*)$ оператор $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p$, получим $(\star):(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p v_1+\ldots+(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p v_{k-1}=\mathbf{0}$. Т. к. подпространства $R_{\lambda_1},\ldots,R_{\lambda_{k-1}}$ инвариантны относительно $A-\lambda_k\cdot\mathrm{id}$, мы имем $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p v_i\in R_{\lambda_i}$, $i=1,\ldots,k-1$. По предположению индукции, из $(\star)$ следует $(\mathcal{A}-\lambda_k\cdot\mathrm{id})^p v_i=\mathbf{0}$. А т. к. по предыдущей лемме оператор $\mathcal{A}-\lambda_k\cdot\mathrm{id}$ в пространства $R_{\lambda_1},\ldots,R_{\lambda_{k-1}}$ обратим, то $v_1=\ldots=v_{k-1}=\mathbf{0}$. Тогда из $(*)$ получаем $v_k=\mathbf{0}$. ■

Лемма 3. Размерность корневого подпространства R_λ равна кратности λ как корня характеристического многочлена оператора $\mathcal{A}$.

Доказательство. Обозначим через r_λ кратность корня λ. Пусть $\widehat{A}=\mathcal{A}\big|_{R_\lambda}$ — ограничение оператора $\mathcal{A}$ на R_λ и $\widetilde{\mathcal{A}}:V/R_\lambda\to V/R_\lambda$ — фактор-оператор. Тогда для характеристических многочленов мы имеем $(*):\chi_{\mathcal{A}}(t)=\chi_{\widehat{\mathcal{A}}}(t)\cdot\chi_{\widetilde{\mathcal{A}}}(t)=(\lambda-t)^{\dim R_\lambda}\chi_{\widetilde{\mathcal{A}}}(t)$, потому что $A=\left(\begin{array}{c|c}\widehat{A} & * \\ \hline 0 & \widetilde{A}\end{array}\right)$ в некотором базисе (т. к. R_λ — инвариантное подпространство). Отсюда $\dim R_\lambda\leqslant r_\lambda$.

Предположим, что $\dim R_\lambda<r_\lambda$. Тогда из $(*)$ следует, что λ является корнем многочлена $\chi_{\widetilde{\mathcal{A}}}(t)$, т. е. собственным значением оператора $\widetilde{\mathcal{A}}$. Пусть $v+R_\lambda$ — соответствующий (ненулевой) собственный вектор, т. е. $\mathcal{A}(v+R_\lambda)=\lambda(v+R_\lambda)$ или $\mathcal{A}v+R_\lambda=\lambda v+R_\lambda$. Отсюда вытекает, что $\mathcal{A}v-\lambda v=(\mathcal{A}-\lambda\cdot\mathrm{id})v\in R_\lambda$. По определению R_λ это значит, что $\mathbf{0}=(\mathcal{A}-\lambda\cdot\mathrm{id})^m(\mathcal{A}-\lambda\cdot\mathrm{id})v=(\mathcal{A}-\lambda\cdot\mathrm{id})^{m+1}v$, т. е. $v\in R_\lambda$. Но тогда $v+R_\lambda$ — нулевой вектор пространства V/R_λ. Противоречие. ■

Теперь докажем основную теорему:

Доказательство. Пусть $\dim V=n$ и r_i — кратность корня λ_i, $i=1,\ldots,k$. Тогда $\sum\limits_{i=1}^{k}r_i=n$ (здесь мы пользуемся алгебраической замкнутостью поля) и из двух предыдущих лемм вытекает, что $\dim(R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k})=\sum\limits_{i=1}^{k}r_i=\dim V$. ■

Определение 3. Разложение $V=R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}$ называется *корневым.*

21. Жордановы клетки и матрицы, их характеристические и минимальные многочлены. Жорданов базис

Определение 1. Матрица вида $J_\lambda=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}$ называется *жордановой клеткой.* Если матрица оператора $\mathcal{A}$ в некотором базисе является блочно-диагональной с блоками вида J_λ (возможно, соответствующих различным λ), то такая матрица называется *жордановой нормальной* *формой* оператора $\mathcal{A}$. Базис, в котором оператор имеет жорданову нормальную форму, называется *жордановым*.

Предложение 1. Минимальный аннулирующий многочлен оператора $\mathcal{A}$ над полем $\mathbb{C}$ есть $P(t) = \prod\limits_{i=1}^{k}(t-\lambda_i)^{m_i}$, где $\lambda_1, \ldots, \lambda_k$ — все собственные значения $\mathcal{A}$, а m_i — размер максимальной жордановой клетки, отвечающей λ_i.

Доказательство. Мы имеет $R_{\lambda_i} = \mathrm{Ker}(\mathcal{A} - \lambda_i \cdot \mathrm{id})^{m_i}$, поэтому многочлен $(t-\lambda_i)^{m_i}$ является минимальным аннулирующим для оператора $\mathcal{A}\big|_{R_{\lambda_i}}$. В силу теоремы о корневом разложении, $\forall v \in V$ представляется в виде $v = \sum\limits_i v_i$, где $v_i \in R_{\lambda_i}$. Т. к. $P(\mathcal{A})$ содержит множитель $(\mathcal{A} - \lambda_i \cdot \mathrm{id})^{m_i}$, мы имеем $P(\mathcal{A})v_i = \mathbf{0}$, т. е. $P(\mathcal{A})v = \mathbf{0}$, и многочлен $P(t)$ аннулирует оператор $\mathcal{A}$. С другой стороны, любой многочлен $Q(t)$, аннулирующий оператор $\mathcal{A}$, делится на минимальный аннулирующий многочлен для оператора $\mathcal{A}\big|_{R_{\lambda_i}}$, т. е. на $(t-\lambda_i)^{m_i}$, для каждого λ_i. Следовательно, $Q(t)$ делится на $P(t)$, и $P(t)$ — минимальный многочлен. ■

Про то, как искать жорданов базис, лучше всего НАПИСАЛ Антон Александрович Клячко. Однако там этот алгоритм описан очень неподробно, когда-нибудь он появится здесь со всеми обоснованиями и доказательствами. Также напишу, как всё-таки искать количество жордановых базисов над конечным полем. То, что Антон Александрович записывает клетки с единицами под собственными значениями, а не над, как это принято, **существенно**.

Как нетрудно заметить, у жордановой клетки минимальный многочлен совпадает с характеристическим. На самом деле, такие матрицы обладают особенным свойством.

Задача 3 (А. А. Клячко). Докажите, что матрица коммутирует только с многочленами от себя тогда и только тогда, когда её минимальный многочлен совпадает с характеристическим.

Для начала заметим, что, не ограничивая общности, можно рассматривать матрицы сразу над алгебраически замкнутым полем (см. задачу 2 в вопросе 17), тогда у них есть жордановы формы. И понятно, что вопрос о коммутировании с какой-то матрицей равносилен вопросу о коммутировании с жордановой этой матрицы (просто потому что коммутирующие матрицы коммутируют в любом базисе).

Отметим также, что совпадение минимального многочлена с характеристическим равносильно тому, что у каждой жордановой клетки своё уникальное собственное значение (следствие предложения 1).

Нам понадобятся несколько технических лемм.

Лемма 1. Пусть $A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{pmatrix}$ — жорданова матрица, причём $\chi_A(t) = \mu_A(t)$, а матрица B имеет вид

$$B = \begin{pmatrix} g_1(J_1) & & \\ & \ddots & \\ & & g_m(J_m) \end{pmatrix},$$

где g_i — многочлены. Тогда существует многочлен f такой, что $B = f(A)$.

Примечание. Это очень сильное утверждение. С помощью него мы можем «сшивать» разные многочлены, которые переводят жордановы клетки матрица A в блоки матрицы B, в единый, переводящий всю матрицу A в матрицу B. То есть, с помощью него мы сводим вопрос о коммутировании даже не к жордановым матрицам, а просто к жордановым клеткам.

Доказательство. Пусть J_i — клетка размера m_i с собственным значением λ_i на диагонали. Для каждого i положим $p_i(t) = \prod\limits_{j \neq i}(t-\lambda_j)^{m_j}$. Т. к. все λ_i различны, НОД$(p_i, (x-\lambda_i)^{m_i}) = 1$. Поэтому существуют такие многочлены a_i и b_i, что $a_i(t)p_i(t)+b_i(t)(t-\lambda_i)^{m_i}=1$. Отсюда $(a_ip_i)(J_j)$ равняется $\underset{m_i\times m_i}{E}$ при $j=i$ и 0 при $j\neq i$. Отсюда

$$(a_ip_ig_i)(A)=\begin{pmatrix}(a_ip_ig_i)(J_1)&&\\&\ddots&\\&&(a_ip_ig_i)(J_m)\end{pmatrix}=\begin{pmatrix}0&&\\&E\cdot g_i(J_i)&\\&&0\end{pmatrix}=\begin{pmatrix}0&&\\&g_i(J_i)&\\&&0\end{pmatrix}$$

где $g_i(J_i)$ стоит на месте i-го блока. Наконец, $B=\sum\limits_{i=1}^{r}(a_ip_ig_i)(A)$. ∎

Лемма 2 (Костя Зюбин)**.** Пусть у жордановой матрицы A минимальный многочлен совпадает с характеристическим. Тогда диагональ A является многочленом от неё.

К этой лемме приведём два доказательства. Первое из них предложил Костя Зюбин (он же сформулировал лемму), второе появилось в результате обсуждения этой задачи с Николаем Юрьевичем Ероховцом.

Первое доказательство. Т. к. минимальный многочлен совпал с характеристическим, то каждому уникальному собственному значению λ_i соответствует ровно одна клетка (из предложения 1), т. е. матрица выглядит вот так:

$$A=\begin{pmatrix}J_1&&\\&\ddots&\\&&J_m\end{pmatrix},$$

где J_i — жорданова клетка размером m_i с собственным значением λ_i, причём все такие λ_i попарно различны. Для каждого i рассмотрим многочлен $g_i(t):=\prod\limits_{j\neq i}(t-\lambda_j)^{m_j}$. Этот многочлен, как нетрудно заметить, будет аннулировать все клетки, кроме J_i. Это значит, что

$$g_i(A)=\begin{pmatrix}g_i(J_i)&\\&\mathbf{0}\end{pmatrix}$$

(условно будем рисовать клетку в верхнем левом углу). Пусть $\mu_{G_i}(t)$ — минимальный многочлен матрицы $G_i=g_i(J_i)$ и $\lambda_i\neq 0$ (мы потом отдельно рассмотрим случай вырожденных клеток). Тогда свободный член $\mu_{G_i}(t)$ не равен нулю. Действительно, в противном случае $\mu_{G_i}(t)=q_{G_i}(t)\cdot t\Rightarrow q_{G_i}(G_i)G_i=0\Rightarrow q_{G_i}(G_i)=0G_i^{-1}=0$, причём $\deg q_{G_i}<\deg\mu_{G_i}$ — противоречие. Теперь будем строить искомый многочлен в несколько шагов:

$$\mu_{G_i}(g_i(A))=\left(\begin{array}{c|c}0&\\\hline&\mu_{G_i}(0)E\end{array}\right)$$

$$\mu_{G_i}(g_i(A))-\mu_{G_i}(0)E=\left(\begin{array}{c|c}-\mu_{G_i}(0)E&\\\hline&0\end{array}\right)$$

$$\lambda_iE-\frac{\lambda_i}{\mu_{G_i}(0)}\mu_{G_i}(g_i(A))=\left(\begin{array}{c|c}\lambda_iE&\\\hline&0\end{array}\right).$$

И вот отсюда

$$\left(\sum_i\lambda_i\right)E-\sum_{\lambda_i\neq 0}\frac{\lambda_i}{\mu_{G_i}(0)}\cdot\mu_{G_i}(g_i(A))=\begin{pmatrix}\boxed{\lambda_1E}&&&\\&\boxed{\lambda_2E}&&\\&&\ddots&\\&&&\boxed{\lambda_mE}\end{pmatrix}.$$

А теперь заметим, что при $\lambda_j=0$ в получившейся матрице j-ый блок будет нулевой матрицей (там будет стоять пустая сумма), а ровно этого и хотелось. Так что найденный нами многочлен переводит A в её диагональ. ∎

Второе доказательство. Рассмотрим минимальный многочлен невырожденной жордановой клетки J_i (потом поговорим про вырожденные), это $(t-\lambda_i)^{m_i}$. Подставим J_i и раскроем скобки. Получим:

$$J_i^{m_i} + \ldots + (-\lambda_i)^{m_i} E = 0.$$

Заметим, что последнее слагаемое — это уже почти диагональ. Разделим обе части на $(-1)^{m_i}\lambda_i^{m_i-1}$ и перенесём в правую часть всё, кроме правого слагаемого. Получим $\lambda_i E = g_i(J_i)$. Для вырожденных клеток $g_i(t) = t^{m_i}$ (получим нулевую матрицу, а это и есть диагональ вырожденной клетки). А теперь все многочлены g_i можно «сшить» по первой лемме. ■

Лемма 3. С матрицей $D = \mathrm{diag}(\overbrace{\lambda_1, \ldots, \lambda_1}^{m_1}, \overbrace{\lambda_2, \ldots, \lambda_2}^{m_2}, \cdots, \overbrace{\lambda_k, \ldots, \lambda_k}^{m_k})$ коммутируют такие и только такие матрицы:

$$\begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix},$$

где B_i — блоки размера m_i.

Доказательство. То, что такие коммутируют, очевидно (проверяется непосредственно). Докажем, что коммутируют только такие. Пусть B коммутирует с D. Тогда умножив D на B слева, мы каждую строку B домножим на соответствующий коэффициент. А умножив D на B справа, мы каждый столбец умножим на соответствующий коэффициент. Итак, получаем:

$$\begin{pmatrix}
\lambda_1 b_{11} & \ldots & \lambda_1 b_{1m_1} & \lambda_1 b_{1,m_1+1} & \ldots & \lambda_1 b_{1,m_1+m_2} & \ldots & \lambda_1 b_{1n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\lambda_1 b_{m_1 1} & \ldots & \lambda_1 b_{m_1 m_1} & \lambda_1 b_{m_1,m_1+1} & \ldots & \lambda_1 b_{m_1,m_1+m_2} & \ldots & \lambda_1 b_{m_1 n} \\
\lambda_2 b_{m_1+1,1} & \ldots & \lambda_2 b_{m_1+1,m_1} & \lambda_2 b_{m_1+1,m_1+1} & \ldots & \lambda_2 b_{m_1+1,m_1+m_2} & \ldots & \lambda_2 b_{m_1+1,n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\lambda_2 b_{m_1+m_2,1} & \ldots & \lambda_2 b_{m_1+m_2,m_1} & \lambda_2 b_{m_1+m_2,m_1+1} & \ldots & \lambda_2 b_{m_1+m_2,m_1+m_2} & \ldots & \lambda_2 b_{m_1+m_2,n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\lambda_k b_{n1} & \ldots & \lambda_k b_{nm_1} & \lambda_k b_{n,m_1+1} & \ldots & \lambda_k b_{n,m_1+m_2} & \ldots & \lambda_k b_{nn}
\end{pmatrix} =$$

$$= \begin{pmatrix}
\lambda_1 b_{11} & \ldots & \lambda_1 b_{1m_1} & \lambda_2 b_{1,m_1+1} & \ldots & \lambda_2 b_{1,m_1+m_2} & \ldots & \lambda_k b_{1n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\lambda_1 b_{m_1 1} & \ldots & \lambda_1 b_{m_1 m_1} & \lambda_2 b_{m_1,m_1+1} & \ldots & \lambda_2 b_{m_1,m_1+m_2} & \ldots & \lambda_k b_{m_1 n} \\
\lambda_1 b_{m_1+1,1} & \ldots & \lambda_1 b_{m_1+1,m_1} & \lambda_2 b_{m_1+1,m_1+1} & \ldots & \lambda_2 b_{m_1+1,m_1+m_2} & \ldots & \lambda_k b_{m_1+1,n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\lambda_1 b_{m_1+m_2,1} & \ldots & \lambda_1 b_{m_1+m_2,m_1} & \lambda_2 b_{m_1+m_2,m_1+1} & \ldots & \lambda_2 b_{m_1+m_2,m_1+m_2} & \ldots & \lambda_k b_{m_1+m_2,n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\lambda_1 b_{n1} & \ldots & \lambda_1 b_{nm_1} & \lambda_2 b_{n,m_1+1} & \ldots & \lambda_2 b_{n,m_1+m_2} & \ldots & \lambda_k b_{nn}
\end{pmatrix}.$$

Осталось приравнять элементы на соответствующих местах. Взглянем на левые верхние подматрицы $m_1 \times m_1$. Они совпадают у обеих матриц. Затем взглянем на подматрицы $M_{j=m_1+1,\ldots,m_1+m_2}^{i=m_1+1,\ldots,m_1+m_2}$. Они тоже совпадают! И так далее. Причём, как нетрудно увидеть, совпадают в этих матрицах только эти блоки. Получили то, что хотели. ■

Лемма 4. С жордановыми клетками коммутируют только многочлены от них[2].

Доказательство. Пусть $J_\lambda = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ коммутирует с матрицей $B = (b^i_j)$. Перемножим

[2]Имеется в виду, конечно, «они и только они». Однако то, что матрица коммутирует с многочленами от себя, очевидно. Нам интересен именно тот случай, когда других коммутирующих нет.

их с обеих сторон:

$$\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}\cdot\begin{pmatrix}b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn}\end{pmatrix}=\begin{pmatrix}\lambda b_{11}+b_{21} & \lambda b_{12}+b_{22} & \dots & \lambda b_{1n}+b_{2n} \\ \lambda b_{21}+b_{31} & \lambda b_{22}+b_{32} & \dots & \lambda b_{2n}+b_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda b_{n-1,1}+b_{n1} & \lambda b_{n-1,2}+b_{n2} & \dots & \lambda b_{n-1,n}+b_{nn} \\ \lambda b_{n1} & \lambda b_{n2} & \dots & \lambda b_{nn}\end{pmatrix}$$

$$\begin{pmatrix}b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn}\end{pmatrix}\cdot\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}=\begin{pmatrix}\lambda b_{11} & b_{11}+\lambda b_{12} & \dots & b_{1,n-1}+\lambda b_{1n} \\ \lambda b_{21} & b_{21}+\lambda b_{22} & \dots & b_{2,n-1}+\lambda b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda b_{n1} & b_{n1}+\lambda b_{n2} & \dots & b_{n,n-1}+\lambda b_{nn}\end{pmatrix}.$$

Теперь приравняем соответствующие элементы получившихся матриц. Из первого столбца, $b_{i1}=0$, $i=2,\dots,n$. Теперь, из второго столбца, $b_{i2}=0$, $i=3,\dots,n$. И так далее. Получаем, что матрица B верхнетреугольная. Но это не всё, что можно про неё сказать. Приравняем клетки с координатами $(1,2)$, получим $b_{11}+\lambda b_{12}=\lambda b_{12}+b_{22}\Rightarrow b_{11}=b_{22}$. Теперь, приравняв клетки с координатами $(2,3)$, аналогично получим $b_{22}=b_{33}$. И так далее. Получаем, что $b_{11}=\ldots=b_{nn}$. Это мы приравняли диагональ на одну выше главной. Потом приравниваем диагонали ещё на одну выше, и ещё, и т. д. Получаем, что матрица B имеет вид:

$$B=\begin{pmatrix}b_0 & b_1 & \ddots & b_{n-2} & b_{n-1} \\ & b_0 & b_1 & \ddots & b_{n-2} \\ & & b_0 & \ddots & \ddots \\ & & & \ddots & b_1 \\ & & & & b_0\end{pmatrix}.$$

Или же, $B=b_0+b_1N+b_2N^2+\ldots+b_{n-1}N^{n-1}$, где $N=J_\lambda-\lambda E$ — нильпотентный оператор. А т. к. B — многочлен от $J_\lambda-\lambda E$, то это и многочлен от J_λ (подставить, раскрыть скобки). ■

Перейдём, наконец, к решению задачи.

▷ $\Leftarrow$. Пусть $\chi_A(t)=\mu_A(t)$ и $BA=AB$. По лемме 2, диагональ A — это некоторый многочлен от A, поэтому B коммутирует и с диагональю A. А значит, по лемме 3, матрица B имеет блочно-диагональный вид с блоками, равными по размерам жордановым клеткам матрицы A. Коммутирование такой матрицы с жордановой матрицей A равносильно коммутированию соответствующих блоков матрицы B с жордановыми клетками A. Значит, матрица B имеет вид

$$B=\begin{pmatrix}g_1(J_1) & & \\ & \ddots & \\ & & g_m(J_m)\end{pmatrix},$$

где g_i — многочлены (по лемме 4). По лемме 1 из этих многочленов можно «сшить» многочлен f, для которого $f(A)=B$.

$\Rightarrow$. Пусть у матрицы A характеристический многочлен не равен минимальному. Тогда у неё какому-то собственному значению λ соответствуют (хотя бы) две клетки. Понятно, что без ограничения общности можно рассматривать случай $A=\begin{pmatrix}J_1 & \\ & J_2\end{pmatrix}$, где J_1 и J_2 — жордановы клетки размера m_1 и m_2 соответственно с собственным значением λ на диагонали. По лемме 2 матрицы J и D коммутируют, т. е. $DJ=JD$. Предположим, что $D=f(A)$, где f — многочлен.

Теперь воспользуемся результатом задачи 1202 из «Сборника задач по аналитической геометрии и линейной алгебре» под ред. Смирнова. Утверждение[3] этой задачи заключается в том, что

$$f(J_i) = \begin{pmatrix} f(\lambda) & \dfrac{f'(\lambda)}{1!} & \dfrac{f''(\lambda)}{2!} & \dfrac{f'''(\lambda)}{3!} & \dots & \dfrac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & f(\lambda) & \dfrac{f'(\lambda)}{1!} & \dfrac{f''(\lambda)}{2!} & \dots & \dfrac{f^{(n-2)}(\lambda)}{(n-2)!} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & f(\lambda) \end{pmatrix}.$$

А т. к. на диагоналях J_1 и J_2 собственные значения одинаковые, то на диагонали матрицы $f(A) = \begin{pmatrix} f(J_1) & \\ & f(J_2) \end{pmatrix}$ должны стоять одинаковые числа, а для D это условие не выполняется. ◀

Примечание. Вспомним, что любая (невырожденная) матрица коммутирует со своей обратной: $AA^{-1} = A^{-1}A = E$. Однако есть такие невырожденные матрицы, которые коммутируют только с многочленами от себя (например, те же жордановы клетки). А это значит, что для матрицы A из условия задачи 3 верно, что A^{-1} — какой-то многочлен от A. На самом деле, это верно для всех невырожденных матриц, и доказывается это несложно. Эту задачу Антон Александрович предлагал нам ещё в первом семестре. Позже я напишу здесь её решение.

22. Существование жорданова базиса для нильпотентного оператора (для матрицы с единственным собственным значением — характеристическим корнем)

Определение 1. Оператор $\mathcal{A}$ называется *нильпотентным*, если $\mathcal{A}^k = 0$ для некоторого $k \in \mathbb{N}$. Минимальное число k, для которого $\mathcal{A}^k = 0$, называется *степенью нильпотентности* $\mathcal{A}$.

Пример 1. Рассмотрим оператор $\mathcal{A}$, заданный в базисе $e_1, \dots, e_n$ матрицей $\begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 \end{pmatrix}$.
Действие этого оператора на базисные векторы описывается схемой $e_n \mapsto e_{n-1} \mapsto \ldots \mapsto e_1 \mapsto \mathbf{0}$. Отсюда видно, что $\mathcal{A}^n = 0$, т. е. оператор $\mathcal{A}$ нильпотентен и имеет степень n.

Сформулируем некоторые свойства нильпотентных операторов.

Предложение 1. Пусть $\mathcal{A} : V \to V$ — нильпотентный оператор, причём $\dim V = n$. Тогда

1. единственным собственным значением $\mathcal{A}$ является 0;
2. $\mathcal{A}$ диагонализируем тогда и только тогда, когда $\mathcal{A} = 0$;
3. степень нильпотентности $\mathcal{A}$ не превосходит n.

Доказательство.

1. Пусть $\mathcal{A}^k = 0$ и $\mathcal{A}^{k-1} \neq 0$. Значит, существует $v \in V : u := \mathcal{A}^{k-1}v \neq \mathbf{0}$. Тогда $\mathcal{A}u = \mathcal{A}^k v = \mathbf{0}$, т. е. u — собственный вектор с собственным значением 0. Пусть $\lambda \neq 0$ — другое собственное значение, тогда по определению найдётся $w \neq \mathbf{0} : \mathcal{A}w = \lambda w$. Тогда $\mathbf{0} = \mathcal{A}^k w = \lambda^k w$. Отсюда $0 = \lambda^k$ и $\lambda = 0$.

2. Если $\mathcal{A}$ диагонализируем, то на диагонали его диагональной матрицы стоят собственные значения, которые все нулевые в силу п. 1. Следовательно, матрица нулевая и $\mathcal{A} = 0$.

3. Из п. 1 вытекает, что характеристический многочлен оператора $\mathcal{A}$ есть $(-t)^n$. Тогда $\mathcal{A}^k = 0$ по теореме Гамильтона — Кэли.

[3] По сути, полностью оно нам и не нужно. Нам достаточно понимать, что на диагонали $f(J_i)$ стоят $f(\lambda)$.

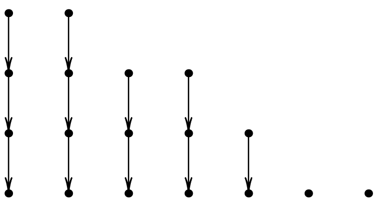
■

Теорема 1. Пусть $\mathcal{A} : V \to V$ — нильпотентный оператор. Тогда в пространстве V существует базис, в котором матрица оператора $\mathcal{A}$ имеет блочно-диагональный вид с блоками из матриц вида $\begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ $(*)$ произвольных размеров. Такой вид матрицы оператора единственный с точностью до перестановки блоков.

Определение 2. Базис, существование которого утверждается в этой теореме, называется *нормальным*, а матрица оператора в таком базисе называется *нормальным видом* (*нормальной формой*) нильпотентного оператора.

Впереди нас ждёт длинное и не очень увлекательное

Доказательство. Базис, в котором матрица оператора состоит из блоков вида $(*)$, удобно изображать в виде диаграммы.



В этой диаграмме точки изображают элементы нормального базиса, а стрелки описывают действие оператора $\mathcal{A}$. Элементы нижней строки оператор переводит в нуль, т. е. в ней стоят собственные векторы оператора (с собственным значением 0), входящие в базис. Каждый столбец соответствует одному блоку вида $(*)$, причём размер блока равен высоте столбца (количеству точек в нём). Итак, нам нужно доказать существование оператора, действие оператора $\mathcal{A}$ на элементы которого описывается диаграммой указанного вида. Проведём индукцию по размерности пространства V. Если $\dim V = 1$, то нильпотентный оператор $\mathcal{A}$ является нулевым, и любой ненулевой вектор в V образует нормальный базис. Пусть теперь $\dim V = n > 1$, и пусть для размерностей, меньших n, существование нормального базиса уже доказано. Пусть $V_0 = \operatorname{Ker} \mathcal{A}$ — подпространство собственных векторов для $\mathcal{A}$. Т. к. $\dim V > 0$, имеем $\dim V/V_0 < n$. Рассмотрим фактор-оператор $\widetilde{\mathcal{A}} : V/V_0 \to V/V_0$, $\widetilde{\mathcal{A}}(v + V_0) = \mathcal{A}v + V_0$. По индуктивному предположению $\widetilde{\mathcal{A}}$ имеет нормальный базис. Можно считать его непустым: иначе $V = V_0$ и любой базис в V_0 будет нормальным для $\mathcal{A}$. Построим диаграмму $\widetilde{D}$ для элементов нормального базиса оператора $\widetilde{\mathcal{A}}$, в каждом её столбце запишем самый верхний вектор $\widetilde{e}_i$, $i = 1, \ldots, m$ (m — количество столбцов в $\widetilde{D}$), и положим $\widetilde{e}_i =: e_i + V_0$, $e_i \in V$. Теперь построим диаграмму D из векторов пространства V следующим образом. Для $i = 1, \ldots, m$ столбец с номером i будет состоять (сверху вниз) из векторов $e_i, \mathcal{A}e_i, \ldots, \mathcal{A}^{h_i-1}e_i, \mathcal{A}^{h_i}e_i$, где h_i — высота i-го столбца в диаграмме $\widetilde{D}$. Т. к. $\widetilde{\mathcal{A}}^{h_i}\widetilde{e}_i = \mathbf{0}$, мы имеем $\mathcal{A}^{h_i}e_i \in V_0$ и $\mathcal{A}^{h_i+1}e_i = \mathbf{0}$. Выберем базис в линейной оболочке $\langle \mathcal{A}^{h_1}e_1, \ldots, \mathcal{A}^{h_m}e_m \rangle \subseteq V_0$, дополним его до базиса V_0 и поставим дополняющие векторы в качестве новых столбцов (высоты 1) в нижней строке диаграммы D; оператор $\mathcal{A}$ переводит их в нуль.

Таким образом, построенная диаграмма D из векторов пространства V имеет в точности такой вид, как требуется для нормального базиса. Нужно лишь проверить, что векторы, составляющие диаграмму, действительно образуют базис в V. Сначала покажем, что векторы из D порождают всё V. Пусть $v \in V$. Положим $\widetilde{v} := v + V_0$. По предположению $\widetilde{v} = \sum\limits_{i=1}^{m} \sum\limits_{j=0}^{h_i-1} \lambda_{ij} \widetilde{\mathcal{A}}^j \widetilde{e}_i$. Тогда $v - \sum\limits_{i=1}^{m} \sum\limits_{j=0}^{h_i-1} \lambda_{ij} \mathcal{A}^j e_i \in V_0$. Но все векторы $\mathcal{A}^j e_i$, $j \leqslant h_i - 1$, лежит в строках диаграммы D, начиная со второй снизу, а подпространство V_0 порождено векторами из нижней строки D по построению. Поэтому v можно представить в виде линейной комбинации векторов из D.

Теперь доказываем линейную независимость. Сначала докажем, что векторы нижней строки линейно независимы. Действительно, если некоторая их нетривиальная линейная комбинация равна нулю, то она должна иметь вид $\sum\limits_{i=1}^{m} \lambda_i \mathcal{A}^{h_i} e_i = \mathbf{0}$, ибо остальные элементы нижней строки дополняют базис линейной оболочки $\langle \mathcal{A}^{h_1}e_1, \ldots, \mathcal{A}^{h_m}e_m \rangle$ до базиса V_0. Но все $h_i \geqslant 1$, поэтому

$\mathcal{A}\left(\sum\limits_{i=1}^{m}\lambda_i\mathcal{A}^{h_i-1}e_i\right)=\mathbf{0}$, так что $\sum\limits_{i=1}^{m}\lambda_i\mathcal{A}^{h_i-1}e_i\in V_0$ и $\sum\limits_{i=1}^{m}\lambda_i\widetilde{\mathcal{A}}^{h_i-1}\widetilde{e}_i=\mathbf{0}$. Из последнего соотношения следует, что все $\lambda_i=0$, т. к. векторы $\widetilde{A}^{h_i-1}\widetilde{e}_i$ составляют нижнюю строку диаграммы D и являются частью базиса V/V_0.

Наконец, покажем, что если имеется любая нетривиальная линейная комбинация векторов D, равная нулю, то из неё можно получить нетривиальную линейную зависимость между векторами нижней строки D. Отметим самую верхнюю строку D, в которой имеются ненулевые коэффициенты этой линейной комбинации. Пусть номер этой строки (считая снизу) равен h. Применим к этой линейной комбинации оператор $\mathcal{A}^{h-1}$. При этом её часть, лежащая в h-ой строке, перейдёт в нетривиальную линейную комбинацию элементов нижней строки, а остальные слагаемые обратятся в нулю. Это завершает доказательсвто существования нормального базиса.

Теперь докажем единственность. Размеры блоков — это высоты столбцов диаграммы. Если расположить столбцы, как на риснуке, в порядке убывания, то их высоты однозначно определяются, если известны длины строк в диаграмме, начиная с нижней, в порядке убывания. Из предыдущего рассуждения следует, что длина нижней строки равна $\dim V_0=\dim\operatorname{Ker}\mathcal{A}$ и не зависит от выбора базиса. Длина второй снизу строки равна размерности ядра фактор-оператора $\widetilde{\mathcal{A}}$ в пространстве $V/\operatorname{Ker}\mathcal{A}$, т. е. $\dim\operatorname{Ker}\mathcal{A}^2-\dim\operatorname{Ker}\mathcal{A}$, что также не зависит от выбора базиса. Продолжая далее, мы видим, что длина k-ой снизу строки равна размерности ядра фактор-оператора в пространстве $V/\operatorname{Ker}\mathcal{A}^{k-1}$, т. е. $\dim\operatorname{Ker}\mathcal{A}^k-\dim\operatorname{Ker}\mathcal{A}^{k-1}$. Это завершает доказательство единственности. ∎

23. Существование жордановой нормальной формы матрицы над алгебраически замкнутым полем

Посмотрим, как выглядит матрица оператора $\mathcal{A}\big|_{R_\lambda}$ (ограничения оператора $\mathcal{A}$ на корневое подпространства R_λ). Т. к. $(\mathcal{A}-\lambda\cdot\mathrm{id})\big|_{R_\lambda}$ является нильпотентным оператором, в пространстве R_λ можно выбрать нормальный базис для этого оператора. Тогда матрица оператора $(\mathcal{A}-\lambda\cdot\mathrm{id})\big|_{R_\lambda}$ в этом базисе будет состоять из блоков вида $\begin{pmatrix}0&1&&\\&0&\ddots&\\&&\ddots&1\\&&&0\end{pmatrix}$. А значит, матрица оператора $\mathcal{A}\big|_{R_\lambda}$ состоит из блоков вида $\begin{pmatrix}\lambda&1&&\\&\lambda&\ddots&\\&&\ddots&1\\&&&\lambda\end{pmatrix}$, т. е. жордановых клеток.

Теорема 1. Для любого оператора $\mathcal{A}$ в пространстве V над алгебраически замкнутым полем существует жорданов базис (в котором оператор имеет жорданову нормальную форму).

Доказательство. Существование жордановой формы является прямым следствием теорем о разложении в прямую сумму корневых подпространств и существовании нормального вида для нильпотентных операторов. Действительно, пусть $\lambda_1,\ldots,\lambda_k$ — все собственные значения $\mathcal{A}$. Выберем в каждом корневом подпространстве R_{λ_i} нормальный базис для нильпотентного оператора $(\mathcal{A}-\lambda_i\cdot\mathrm{id})\big|_{R_{\lambda_i}}$. Тогда объединение этих базисов даст жорданов базис для оператора $\mathcal{A}$ в силу наличия корневого разложения $V=R_{\lambda_1}\oplus\ldots\oplus R_{\lambda_k}$ (здесь мы пользуемся алгебраической замкнутостью поля). ∎

24. Единственность жордановой нормальной формы

Теорема 1. Жорданова нормальная форма оператора единственна с точностью до перестановки клеток.

Доказательство. Надо показать, что количество жордановых клеток фиксированного размера с одним и тем же λ не зависит от выбора жорданова базиса. Выберем произвольный жорданов базис. Пусть W_{λ_i} — линейная оболочка части этого базиса, отвечающая всем клеткам с λ_i на диагонали. Тогда ограничение оператора $(\mathcal{A} - \lambda_i \cdot \mathrm{id})$ на W_{λ_i} — нильпотентный оператор, а значит, $W_{\lambda_i} \subseteq R_{\lambda_i}$. Кроме того, $W_{\lambda_1} \oplus \ldots \oplus W_{\lambda_k}$ по определению жорданова базиса и $V = R_{\lambda_1} \oplus \ldots \oplus R_{\lambda_k}$ (корневое разложение), отсюда $\dim W_{\lambda_i} = \dim R_{\lambda_i}$ и $W_{\lambda_i} = R_{\lambda_i}$. Итак, подпространства, отвечающие клеткам с собственным значением λ_i, не зависят от способа приведения к жордановой форме и равны R_{λ_i}.

Таким образом, мы свели доказательство единственности жордановой формы к случаю, когда оператор $\mathcal{A}$ имеет одно собственное значение λ. Любой жорданов базис для такого оператора будет также нормальным базисом для нильпотентного оператора $\mathcal{A} - \lambda \cdot \mathrm{id}$. Для нильпотентных операторов мы уже доказали единственость нормального вида (т. е. жордановой формы). ■

25. Билинейные функции и их матрицы. Изменение матрицы билинейной формы при замене базиса. Ранг билинейной функции. Симметрические билинейные функции

Определение 1. Пусть V — линейное пространство над полем $\mathbb{F}$. Функция $\mathcal{B} : V \times V \to \mathbb{F}$ называется *билинейной функцией*, если она линейна по каждому аргументу:

$$\mathcal{B}(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 \mathcal{B}(x_1, y) + \lambda_2 \mathcal{B}(x_2, y),$$
$$\mathcal{B}(x, \mu_1 y_1 + \mu_2 y_2) = \mu_1 \mathcal{B}(x, y_1) + \mu_2 \mathcal{B}(x, y_2)$$

для любых $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{F}$ и $x, x_1, x_2, y, y_1, y_2 \in V$.

Определение 2. *Матрицей билинейной функции* $\mathcal{B}$ в базисе $e_1, \ldots, e_n$ пространства V называется квадратная матрица $B = (b_{ij})$ размера n, где $b_{ij} = \mathcal{B}(e_i, e_j)$.

Зная матрицу $B = (b_{ij})$ билинейной функции, можно восстановить значение $\mathcal{B}(x, y)$ на любой паре векторов $x = x^i e_i$ и $y = y^j e_j$:

$$\mathcal{B}(x, y) = \mathcal{B}(x^i e_i, y^j e_j) = x^i y_j \mathcal{B}(e_i, e_j) = b_{ij} x^i y^j = x^t B y.$$

Определение 3. Выражение $B(x, y) = b_{ij} x^i y^j$ называется *билинейной формой.*

Билинейная форма представляет собой однородный многочлен степени 2 от двух наборов переменных $x^1, \ldots, x^n$ и $y^1, \ldots, y^n$, который линее по x при фиксированных y и линеен по y при фиксированных x.

Теорема 1 (Закон изменения матрицы билинейной функции)**.** Имеет место соотношение

$$B^t = C^t B C,$$

где B — матрица билинейной функции $\mathcal{B} : V \times V \to \mathbb{F}$ в базисе $e_1, \ldots, e_n$, B^t — матрица в базисе $e_{1'}, \ldots, e_{n'}$ и C — матрица перехода от базиса $e_1, \ldots, e_n$ к базису $e_{1'}, \ldots, e_{n'}$.

Доказательство. Пусть $B = (b_{ij})$, $B^t = (b'_{ij})$, $C = (c^i_{i'})$. Мы имеем

$$b'_{ij} = \mathcal{B}(e_{i'}, e_{j'}) = \mathcal{B}(c^i_{i'} e_i, c^j_{j'} e_j) = c^i_{i'} c^j_{j'} \mathcal{B}(e_i, e_j) = c^i_{i'} b_{ij} c^j_{j'},$$

что эквивалентно матричному соотношению $B^t = C^t B C$. ■

Следствие 1. Ранг матрицы билинейной функции не зависит от базиса.

Доказательство. Т. к. матрица C обратима $\operatorname{rk} B^t = \operatorname{rk}(C^t BC) = \operatorname{rk} B$. ■

Определение 4. *Рангом* билинейной функции $\mathcal{B}$ (обозначается $\operatorname{rk}\mathcal{B}$) называется ранг её матрицы в произвольном базисе. Билинейная функция $\mathcal{B}$ в пространстве V называется *невырожденной*, если $\operatorname{rk}\mathcal{B} = \dim V$.

Множество $\mathrm{B}(V)$ всех билинейных функций в пространстве V образует линейное пространство относительно операций сложения функций и умножения функций на скаляры. Сопоставление билинейной функции $\mathcal{B}$ её матрицы B в фиксированном базисе $e_1, \ldots, e_n$ устанавливает изоморфизм между пространством $\mathrm{B}(V)$ и пространством квадратных матрицы $\underset{n\times n}{\mathrm{Mat}}(\mathbb{F})$. Как и в случае пространства линейных операторов $\mathrm{Hom}(V, V)$, этот изоморфизм неканоничен, т. к. он зависит от выбора базиса.

Теорема 2. Отображение $\varphi : \mathrm{B}(V) \to \mathrm{Hom}(V, V^*)$, сопоставляющее билинейной функции $\mathcal{B}$ линейной отображение $\widetilde{\mathcal{B}} : V \to V^*$, задаваемое формулой

$$\widetilde{\mathcal{B}}(x) = \mathcal{B}(x, \cdot) \text{ для } x \in V,$$

является каноническим изоморфизмом. Здесь $\mathcal{B}(x, \cdot) \in V^*$ — линейная функция, значение которой на векторе $y \in V$ есть $\mathcal{B}(x, y)$.

Доказательство. Отображение φ линейной, т. к. билинейная функция линейна по первому аргументу x. Кроме того, отображение φ биективно: обратное отображение φ^{-1} ставит в соответствие линейному отображению $\widetilde{\mathcal{B}} : V \to V^*$ билинейную функцию $\mathcal{B}$, заданную по формуле $\mathcal{B}(x, y) = \widetilde{\mathcal{B}}(x)(y)$. Следовательно, φ — изоморфизм. Этот изоморфизм каноничен, т. к. в его конструкции не использовался базис. ■

Определение 5. Билинейная функция $\mathcal{B} : V \times V \to \mathbb{F}$ называется *симметрической*, если $\mathcal{B}(y, x) = \mathcal{B}(x, y)$, и *кососимметрической*, если $\mathcal{B}(x, x) = 0$ для любых $x, y \in V$.

Примечание. Именно такое определение кососимметрической билинейной функции правильное, потому что над полем характеристики 2 стандартное определение $\mathcal{B}(x, y) = -\mathcal{B}(y, x)$ равносильно определению симметричности. А над полем характеристики не 2 эти определения равносильны. Действительно, если $\mathcal{B}(x, x) = 0$ для любого $x \in V$, то для любых $u, v \in V$ выполняется

$$0 = \mathcal{B}(u + v, u + v) = \underbrace{\mathcal{B}(u, u)}_{0} + \mathcal{B}(u, v) + \mathcal{B}(v, u) + \underbrace{\mathcal{B}(v, v)}_{0} \Rightarrow \mathcal{B}(u, v) = -\mathcal{B}(v, u).$$

26. Квадратичные формы и их матрицы. Восстановление симметрической билинейной функции по даной квадратичной функции. Диагональный вид квадратичной формы. Алгоритм Лагранжа

Определение 1. *Квадратичной формой* над $\mathbb{F}$ называется однородный многочлен второй степени от n переменных $x = (x^1, \ldots, x^n)$, т. е. многочлен вида

$$Q(x) = Q(x^1, \ldots, x^n) = q_{ij}x^i x^j = \sum_{i=1}^{n} q_{ii}(x^i)^2 + \sum_{i<j} 2q_{ij}x^i x^j,$$

где $q_{ji} = q_{ij} \in \mathbb{F}$. Симметричная матрица $Q = (q_{ij})$ размера $n \times n$ называется *матрицей квадратичной формы*.

Если $B(x, y) = b_{ij}x^i y^j$ — симметрическая билинейная форма, то $B(x, x) = b_{ij}x^i x^j$ является квадратичной формой с матрицей B. Таким образом, квадратичная форма $B(x, x)$ полностью определяет симметрическую билинейную форму $B(x, y)$, а значит, и симметрическую билинейную функцию

$\mathcal{B}(x, y)$. Это можно увидеть и не прибегая к выбору базиса: для симметрической билинейной функции имеет место соотношение

$$\mathcal{B}(x, y) = \frac{1}{2}(\mathcal{B}(x + y, x + y) - \mathcal{B}(x, x) - \mathcal{B}(y, y)).$$

Примечание. Заметим, что здесь мы должны уметь делить на 2, поэтому формула верна только для полей с характеристикой не 2.

Определение 2. Функцию $V \to \mathbb{F}$, $x \mapsto \mathcal{B}(x, x)$ называют *квадратичной функцией.*

Теорема 1. Для симметрической билинейной функции $\mathcal{B}$ над полем характеристики, отличной от 2, существует базис, в котором её матрица диагональна. Другими словами, любую квадратичную форму $Q(x)$ линейной заменой координат $x = Cy$ можно привести к виду

$$Q(y) = r_{11}(y^1)^2 + \ldots + r_{nn}(y^n)^2.$$

Мы приведём два доказательства этого факта. В первом случае мы будем работать с квадратичными формами и координатами, а во втором — с симметрическими билинейными функциями и базисами. Каждое из доказательств будет проведено таким образом, что его можно будет использовать как алгоритм.

Первое доказательство (метод Лагранжа)**.** Пусть $Q(x) = q_{ij}x^i x^j$ — квадратичная форма. Доказательство заключается в последовательном упрощении $Q(x)$, использующем основное и два вспомогательных преобразования.

Основное преобразование производится, если в квадратичной форме $Q(x) = q_{ij}x^i x^j$ первый коэффициент q_{11} не равен нулю. Тогда имеем

$$\begin{aligned}
Q(x^1, \ldots, x^n) &= q_{11}(x^1)^2 + 2q_{12}x^1x^2 + \ldots + 2q_{1n}x^1x^n + \sum_{i,j>1} q_{ij}x^ix^j = \\
&= q_{11}\left(x^1 + \frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n\right)^2 - q_{11}q_{11}\left(\frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n\right)^2 + \sum_{i,j>1} q_{ij}x^ix^j = \\
&= q_{11}\left(x^1 + \frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n\right)^2 + Q'(x^2, \ldots, x^n),
\end{aligned}$$

где $Q'(x^2, \ldots, x^n)$ — некоторая квадратичная форма от $n-1$ переменных. Теперь сделаем замену координат

$$\begin{gathered}
u^1 = x^1 + \frac{q_{12}}{q_{11}}x^2 + \ldots + \frac{q_{1n}}{q_{11}}x^n, \\
u^2 = x^2, \quad \ldots, \quad u^n = x^n.
\end{gathered}$$

В результате $Q(x)$ преобразуется к виду

$$Q(u^1, \ldots, u^n) = q_{11}(u^1)^2 + Q'(u^2, \ldots, u^n).$$

Если в форме $Q'(u^2, \ldots, u^n)$ первый коэффициент (т. е. q'_{22}) не равен нулю, то мы снова можем применить основное преобразование, и т. д.

Первое вспомогательное преобразование производится, если $q_{11} = 0$, но существует $q_{ii} \neq 0$. В этом случае мы делаем замену $u^1 = x^i, u^i = x^1$, а остальные координаты без изменений. В результате получаем $q'_{11} \neq 0$.

Второе вспомогательное преобразование производится, если все коэффициенты q_{ii} при квадратах равны нулю, но при этом есть хотя бы один ненулевой коэффициент (в противном случае $Q(x) \equiv 0$ уже имеет нужный вид). Пусть $q_{ij} \neq 0$, где $i < j$. Произведём замену координат

$$x^i = u^i, \quad x^j = u^i + u^j, \quad x^k = u^k \text{ при } k \neq i, j.$$

В результате форма $Q(x)$ преобразуется к виду

$$Q(x) = 2q_{ij}x^i x^j + \ldots = 2q_{ij}u^i(u^i + u^j) + \ldots = 2q_{ij}(u^i)^2 + \ldots,$$

где $\ldots$ означает члены, не содержащие квадратов. Далее мы можем применить предыдущие преобразования.

Последовательно применяя основное преобразование и (если нужно) вспомогательные преобразования, мы приводим форму $Q(x)$ к диагональному виду. ■

Первое доказательство (метод поиска базиса)**.** Пусть $B = (b_{ij})$ — матрица билинейной функции $\mathcal{B}$ в исходном базисе $e_1, \ldots, e_n$.

Основное преобразование производится, если $b_{11} = \mathcal{B}(e_1, e_1) \neq 0$. Выберем новый базис следующим образом:

$$\begin{aligned}
&e_{1'} = e_1,\\
&e_{2'} = e_2 - \frac{\mathcal{B}(e_1, e_2)}{\mathcal{B}(e_1, e_1)} e_1 = e_2 - \frac{b_{12}}{b_{11}} e_1,\\
&\quad\vdots\\
&e_{n'} = e_n - \frac{\mathcal{B}(e_1, e_n)}{\mathcal{B}(e_1, e_1)} e_1 = e_n - \frac{b_{1n}}{b_{11}} e_1.
\end{aligned}$$

В результате получим $\mathcal{B}(e_{1'}, e_{1'}) = 0$ при $i > 1$. Таким образом, матрица билинейной функции $\mathcal{B}$ в новом базисе принимает вид

$$B' = \left(\begin{array}{c|ccc}
b_{11} & 0 & \ldots & 0\\
\hline
0 & & & \\
\vdots & & \widetilde{B}' & \\
0 & & &
\end{array}\right),$$

где $\widetilde{B}'$ — матрица размера $(n-1)\times(n-1)$ билинейной функции $\mathcal{B}$ на подпространстве $\langle e_{2'}, \ldots, e_{n'}\rangle$. Далее мы работаем уже с этой матрицей $\widetilde{B}'$.

Первое вспомогательное преобразование производится, если $b_{11} = 0$, но имеется $b_{ii} \neq 0$. Тогда делаем замену, меняющую местами первый и i-ый базисный векторы.

Второе вспомогательное преобразование производится, если все b_{ii} равны нулю, но при этом билинейная функция $\mathcal{B}$ не является тождественно нулевой, т. е. $b_{ij} = \mathcal{B}(e_i, e_j) \neq 0$ для некоторых $i < j$. Произведём замену базиса

$$e_{i'} = e_i + e_j, \quad e_{j'} = e_j, \quad e_{k'} = e_k \text{ при } k \neq i, j.$$

Далее можем применить предыдущие преобразования.

Последовательно применяя основное преобразование и дополняя его в необходимых случаях вспомогательными преобразованиями, мы получаем базис $f_1, \ldots, f_n$, в котором матрица билинейной функции $\mathcal{B}$ имеет диагональный вид. ■

Обратим внимание, что основное и вспомогательное преобразование в обоих доказательствах — это одно и то же преобразование, просто в первом случае оно записано через координаты, а во втором — через базисы. Так что диагональные матрицы, получаемые первым и вторым методом, совпадают, как и все промежуточные матрицы.

Примечание. Если при приведении матрицы билинейной функции (квадратичной формы) к диагональному виду использовалось лишь основное преобразование, то матрица перехода от исходного базиса к базису, в котором матрица имеет диагональный вид, является верхнетреугольной.

Задача 4. Над полем $\mathbb{Z}_2$ симметрическая билинейная функция с матрицей $\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$ не приводится к диагональному виду заменой базиса.

Приведём здесь два решения. Первое моё, второе — Антона Александровича.

$\triangleright$ Пусть можно, и матрица замены базисов имеет вид $C = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$. Произведём замену:

$$C^t \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} C = \begin{pmatrix}a & c\\ b & d\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}c & a\\ d & b\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}2ac & bc+ad\\ bc+ad & 2bd\end{pmatrix}.$$

Тут можно рассуждать по-разному:

1. Над $\mathbb{Z}_2$ имеем $2 = 0$, поэтому на диагонали у такой матрицы стоят нули. Поэтому если она диагональная, то нулевая, а нулевая матрица — она нулевая в любом базисе.
2. Вне диагонали должны стоять нули, поэтому $bc + ad = 0$. Над $\mathbb{Z}_2$ имеем $-1 = 1$, поэтому $0 = ad - bc = \det C$, но матрица перехода между базисами должна быть невырожденной.

$\blacktriangleleft$

$\triangleright$ Пусть $\mathcal{B}$ — билинейная функция с такой матрицей в базисе e_1, e_2. Тогда

$$\mathcal{B}(x,x) = \mathcal{B}(x^1e_1 + x^2e_2, x^1e_1 + x^2e_2) = x^1\underbrace{\mathcal{B}(e_1,e_1)}_{0} + x^2\underbrace{\mathcal{B}(e_2,e_2)}_{0} + x^1x^2\underbrace{(\mathcal{B}(e_1,e_2)+\mathcal{B}(e_1,e_2))}_{\text{Над } \mathbb{Z}_2 \text{ это } 0} = 0.$$

Значит, при любой замене базиса на диагонали будут стоять нули. А ещё при любой замене базиса матрица должна быть симметрической. Поэтому она либо нулевая, либо такая, как есть. Ранг должен сохраняться, поэтому она такая, как есть, т. е. не диагональная. $\blacktriangleleft$

27. Нормальный (канонический) вид квадратичной формы над полями действительных и комплексных чисел. Закон инерции

Над полем $\mathbb{R}$ квадратичную форму можно далее упростить:

Предложение 1. Для любой симметрической билинейной функции $\mathcal{B}$ в пространстве над полем $\mathbb{R}$ существует базис, в котором её матрица имеет диагональный вид с 1, -1 и 0 на диагонали.

Доказательство. Сначала с помощью теоремы 1 приведём квадратичную форму к виду

$$Q(u) = r_{11}(u^1)^2 + \ldots + r_{nn}(u^n)^2.$$

Если $r_{ii} > 0$, то замена $y^i = \sqrt{r_{ii}}u^i$ приводит слагаемое $r_{ii}(u^i)^2$ к виду $(y^i)^2$. Если же $r_{ii} < 0$, то замена $y^i = \sqrt{|r_{ii}|}u^i$ приводит слагаемое $r_{ii}(u^i)^2$ к виду $-(y^i)^2$. В результате получаем требуемый вид квадратичной формы с коэффициентами 1, -1 и 0. $\blacksquare$

Определение 1. Вид, описанный в последнем предложении, называется *нормальным видом* вещественной симметрической билинейной формы (вещественной квадратичной формы).

Над полем $\mathbb{C}$ квадратичную форму можно ещё больше упростить:

Предложение 2. Для любой симметрической билинейной функции $\mathcal{B}$ над полем $\mathbb{C}$ существует базис, в котором её матрица имеет диагональный вид с 1 и 0 на диагонали.

Доказательство. Сначала мы с помощью последнего предложения приведём квадратичную форму к виду $(y^1)^2 + \ldots + (y^p)^2 - (y^{p+1})^2 - \ldots - (y^{p+q})^2$. Затем сделаем замену координат $y^k = z^k$ при $k \leqslant p$ и $y^k = iz^k$ при $k > p$. В результате получим требуемый вид, где $r = p + q = \operatorname{rk} Q$. $\blacksquare$

Определение 2. Вид, описанный в последнем предложении, называется *нормальным видом* комплексной симметрической билинейной формы (комплексной квадратичной формы).

В случае симметрической билинейной формы над $\mathbb{C}$ нормальный вид зависит только от её ранга, и поэтому мы получаем:

Предложение 3. Две комплексные симметрические билинейные формы (комплексные квадратичные формы) получаются друг из друга линейной заменой координат только и только тогда, когда их ранги совпадают.

В случае симметрических билинейных форм над $\mathbb{R}$ ситуация сложнее: их нормальный вид не определяется одним лишь рангом, а зависит ещё от количества 1 и -1 на диагонали матрицы. Оказывается, что нормальный вид такой формы не зависит от способа приведения к нормальному виду.

Теорема 1 (Закон инерции)**.** Количество 1, -1 и 0 на диагонали нормального вида матрицы вещественной симметрической билинейной функции $\mathcal{B}$ не зависит от способа приведения к нормальному виду.

Примечание. Важно понимать, что человечество на самом деле не умеет приводить квадратичные формы к какому-то адекватному виду. Мы хоть что-то знаем только про очень узкие ситуации — симметричная (кососимметричная) матрица, только над полями $\mathbb{R}$ или $\mathbb{C}$ и т. п. Даже над полем $\mathbb{Q}$ понять, являются две квадратичные формы эквивалентными, сложно. Есть инвариант в виде ранга, есть замечание, что $\det A' = \det(C^t AC) = (\det C)^2 \det A$ (т. е. отношение определителей должно быть квадратом элемента поля). Но вот примерно на этом какие-то нормальные соображения заканчиваются. Ходят слухи, что в НМУ на «Алгебре 2» в 2024 году учили что-то понимать про поля типа $\mathbb{Z}_p$.