Записки научных
семинаров ПОМИ
Том 534, 2024 г.

А. Юрков

ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ, СОХРАНЯЮЩИЕ МАТРИЧНЫЙ ДИСКРИМИНАНТ

1. Введение

Интерес к описанию линейных отображений, сохраняющих те или иные матричные инварианты, возник ещё в XIX веке, и первый результат в этой области был получен Фробениусом в 1987 году в [5] для определителя квадратной матрицы. В XX веке был достигнут значительный прогресс в решении задач такого рода, и в обзоре [12] можно ознакомиться с достижениями в этой области на момент его публикации.

Матричный дискриминант является одним из полиномиальных матричных инвариантов. Он определяется как дискриминант характеристического многочлена матрицы, и, таким образом, показывает наличие или отсутствие у матрицы кратных собственных значений в алгебраическом замыкании основного поля. Первое известное автору упоминание матричного дискриминанта в математической литературе восходит к [9], где Куммер решает задачу о представлении дискриминанта характеристического многочлена симметричной 3×3 матрицы в виде суммы квадратов многочленов от её коэффициентов.

Первый результат, характеризующий линейные отображения, сохраняющие матричный дискриминант, был получен Пирсом в [13]. Пирс использует метод, основанный на рассмотрении определённых множеств пространств матриц, которые инвариантны относительно отображений, сохраняющих матричный дискриминант, и последующем применении основной теоремы проективной геометрии к некоторым из этих множеств (см., например [1, с. 122, теорема 2.26]). Этот подход довольно широко применяется для решения такого рода задач, например, в статье Дьёдонне [7], где доказывается теорема о линейных

Ключевые слова: линейные отображения, матричный дискриминант, квадратные матрицы.

Работа была выполнена при поддержке гранта ISF No. 1994/20 и стипендии Центра интеграции в науке Министерства алии и интеграции государства Израиль.

отображениях, сохраняющих множество матриц, определитель которых равен нулю.

При этом статья Пирса содержит несколько мест, которые, по нашему мнению, нуждаются в уточнении и дополнении. Перечислим их (в скобках указаны ссылки на соответствующие утверждения в приложении A, где они обсуждаются более подробно).

(1) Утверждение A.1 доказано только для $n \geqslant 6$, где n – порядок матриц. При этом в примечании 2 в [13] сказано, что, тем не менее, утверждение верно и при $n = 3, 4, 5$, но требует большого количества специальных вычислений.
(2) Неявно подразумевается, что характеристика $\mathrm{char}(\mathbb{F})$ поля $\mathbb{F}$ не равна 2 и 3 (утверждение A.3).
(3) Доказательство утверждения 17 в [13] содержит ошибочное утверждение о размерности пространств, которые в нём рассматриваются (утверждение A.7).
(4) Неявно предполагается, что $\mathrm{char}(\mathbb{F})$ не является делителем n (утверждение A.8).
(5) В конце статьи считается очевидным, что $c^n = 1$, тогда как на самом деле $c^{n(n-1)} = 1$ (утверждение A.1). Это приводит к тому, что формулировка основной теоремы не является верной.

В нашей статье подход Дьёдонне сочетается с другим, не менее распространённым при решении подобных задач. Он основывается на следующем наблюдении: если P – некоторый многочлен от матричных коэффициентов, то для данных матриц A, B выражение $P(A + \lambda B)$ является многочленом от λ; причём для некоторых P возможно связать $\max_A(\deg_\lambda(A + \lambda B))$ с различными свойствами B. Этот подход используется, например, в работе Маркуса и Пёрвса [10] для описания линейных отображений, сохраняющих матричные элементарные симметрические функции.

Таким образом, в предположении, что P равен матричному дискриминанту, для заданной матрицы B устанавливается связь между значением $\max_A(\deg_\lambda P(A + \lambda B))$ и свойством "быть линейной комбинацией скалярной матрицы и нильпотентной матрицы ранга 1".

Данное наблюдение также позволяет доказать теорему 3.6 и в "модулярном" случае, то есть тогда, когда $\mathrm{char}(\mathbb{F})$ делит n. Соответствующая статья готовится автором к публикации. Заметим при этом, что, хотя теоретически есть возможность провести рассуждения без разделения на "регулярный" и "модулярный" случаи, для этого потребуется значительное количество специальных вычислений для $n \leqslant 6$, а имеющееся доказательство для случая, когда $\operatorname{char}(\mathbb{F})$ делит n, не содержащее значительного объёма вычислений, практически независимо по содержанию от доказательства, приведённого в этой статье (хотя и похоже идейно).

В отличие от статьи Пирса, в нашей статье явно используется то, что $\operatorname{char}(\mathbb{F})$ не делит n, и это позволяет сразу свести задачу к отображениям, которые сохраняют одновременно след матрицы и матричный дискриминант, и доказать, что такие отображения сохраняют множество нильпотентных матриц ранга 1 (лемма 2.11).

Дальнейшее доказательство при $n = 2$ проводится прямыми вычислениями, а при $n \geqslant 3$ предварительно рассматривается действие линейного отображения, сохраняющего след матрицы и матричный дискриминант, на множестве пространств специального вида, каждое из которых порождено нильпотентными матрицами ранга 1. Таким образом удаётся найти обратимое преобразование этого отображения, приводящее его к требуемому виду, и доказать основную теорему.

Также стоит отметить, что из основной теоремы следует, в частности, что в "регулярном" случае множество отображений, сохраняющих матричный дискриминант при $n = 2$, имеет ту же структуру, что и при бóльших n. Это опровергает предположение, выдвинутое Пирсом в [12, глава 4, второе примечание после теоремы 4.1.11].

Для того, чтобы сформулировать основной результат, приведём используемые при этом обозначения, определения, а также свойства определяемых объектов.

Мы фиксируем, начиная с этого момента, натуральное число n, предполагая, что $n \geqslant 2$.

Всюду в этой статье через $\mathbb{F}$ обозначается **бесконечное** поле, **характеристика которого не делит** n.

Далее, $\mathcal{M}_m = \mathcal{M}_m(\mathbb{F})$ обозначает множество всех матриц размеров $m \times m$ с элементами из $\mathbb{F}$, через $I_m \in \mathcal{M}_m$ обозначается единичная матрица, через O_m – нулевая матрица, а через $E_{i\,j} \in \mathcal{M}_m$ – матричная единица, то есть матрица, все элементы которой равны нулю, за исключением элемента, лежащего на пересечении её i-ой строки и j-го столбца – он равен 1. Через $\operatorname{diag}(a_1, \ldots, a_m) \in \mathcal{M}_m(\mathbb{F})$ обозначается *диагональная* матрица, элементы диагонали которой задаются вектором $(a_1, \ldots, a_m)$, а все остальные элементы равны нулю.

Через $X_{i\,j}$ обозначается элемент матрицы $X \in \mathcal{M}_m$, лежащий на пересечении её i-ой строки и j-го столбца.

Матрица $X \in \mathcal{M}_m$ называется *скалярной*, если $X = cI_m$ для некоторого $c \in \mathbb{F}$.

$X^t \in \mathcal{M}_m$ обозначает матрицу, *транспонированную* к X, то есть $X^t_{i\,j} = X_{j\,i}$ для всех $1 \leqslant i, j \leqslant m$.

Если $X \in \mathcal{M}_k$, $Y \in \mathcal{M}_m$, то через $X \oplus Y \in \mathcal{M}_{k+m}$ обозначается *прямая сумма матриц* X и Y, то есть

$$X \oplus Y = \left[\begin{array}{c|c} X & 0 \\ \hline 0 & Y \end{array}\right].$$

Для всякой матрицы $A \in \mathcal{M}_m$ через $p_A(x) \in \mathbb{F}[x]$ мы обозначаем её характеристический многочлен:

$$p_A(x) = \det(xI_m - A).$$

Через $\langle S \rangle$ обозначается векторное пространство, порождённое подмножеством S данного векторного пространства V, то есть множество всевозможных линейных комбинаций элементов из S. Если $U, V \subset W$ – векторные подпространства пространства W, то через $U + V$ обозначается *сумма пространств* U и V: $U + V = \langle U \cup V \rangle$. Аналогичным образом определяется и обозначается сумма трёх и более подпространств одного и того же пространства. Через $e^{(1)}, \ldots, e^{(m)}$ обозначаются векторы *стандартного базиса* $\mathbb{F}^m$, то есть

$$e^{(1)} = (1, 0, \ldots, 0), \quad \ldots, \quad e^{(m)} = (0, \ldots, 0, 1).$$

Если $v, w \in \mathbb{F}^m$, то через (v, w) обозначается их *скалярное произведение*: $(v, w) = v_1 w_1 + \ldots + v_m w_m$.

Определение 1.1. Пусть $f, g \in \mathbb{F}[x]$ – многочлены, причём

$$f(x) = a_0 + \ldots + a_m x^m,$$

$$g(x) = b_0 + \ldots + b_k x^k.$$

Результантом $\operatorname{Res}(f,g) = \operatorname{Res}(a_0,\dots,a_m;b_0,\dots,b_k)$ называется многочлен от коэффициентов f и g, определяемый равенством

$$\operatorname{Res}(f,g) = \begin{vmatrix} a_m & \cdots & a_1 & a_0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_m & a_{m-1} & \cdots & a_0 \\ b_k & \cdots & b_1 & b_0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & b_k & b_{k-1} & \cdots & b_0 \end{vmatrix}. \tag{1.1}$$

Замечание 1.2. Таким образом, определитель в равенстве (1.1) имеет порядок $m+n$.

Напомним основные свойства результанта (в квадратных скобках приведены ссылки на соответствующие равенства из [8, глава 12]).

R1.

$$\operatorname{Res}\left(\prod_{i=1}^{m}(x-x_i), \prod_{j=1}^{k}(x-y_j)\right) = \prod_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant k}}(x_i-y_j)$$

для всех $x_1,\dots,x_m,y_1,\dots,y_k \in \mathbb{F}$. [1.3]

R2. $\operatorname{Res}\left(\prod_{i=1}^{m}(x-x_i), \prod_{j=1}^{k}(x-y_j)\right) = 0 \iff \exists 1\leqslant i\leqslant m$, $1\leqslant j\leqslant k\colon x_i = y_j$.

R3. $\operatorname{Res}(f_1\cdot f_2, g) = \operatorname{Res}(f_1,g)\cdot\operatorname{Res}(f_2,g)$ для всех $f_1,f_2,g \in \mathbb{F}[x]$. [1.4]

R4. $\operatorname{Res}(f(x+c),g(x+c)) = \operatorname{Res}(f(x),g(x))$ для всех $f,g\in\mathbb{F}[x], c\in\mathbb{F}$. [1.8]

R5. $\operatorname{Res}(\lambda^0 a_0,\dots,\lambda^m a_m;\lambda^0 b_0,\dots,\lambda^k b_k) = \lambda^{mk}\operatorname{Res}(a_0,\dots,a_m; b_0,\dots,b_k)$ для всех $f=a_0+\ldots+a_n x^m$, $g=b_0+\ldots+b_k x^k, \lambda\in\mathbb{F}$. [1.6]

Следствие 1.3 (из свойства R3)**.**

$$\operatorname{Res}\left(\prod_{i=1}^{m}(x-x_i), g\right) = \prod_{i=1}^{m} g(x_i)$$

для всех $x_1,\dots,x_m\in\mathbb{F}$, $g\in\mathbb{F}[x]$.

Определение 1.4. Для многочлена $f \in \mathbb{F}[x]$, где $f(x) = a_0 + \ldots + a_m x^m$ и $a_m \neq 0$, через $\mathrm{D}(f) = \mathrm{D}(a_0, \ldots, a_m)$ мы обозначаем *дискриминант* f. Он определяется равенством

$$\mathrm{D}(f) = \frac{(-1)^{n(n-1)/2}}{a_m} \operatorname{Res}(f, f'),$$

где f' – производная f.

Напомним основные свойства дискриминанта (в квадратных скобках приведены ссылки на соответствующие равенства из [8, глава 12]).

D1. $\mathrm{D}\left(\prod_{i=1}^m (x - x_i)\right) = \prod_{1 \leqslant i < j \leqslant m} (x_i - x_j)^2$ для всех $x_1, \ldots, x_m \in \mathbb{F}$.

D2. $\mathrm{D}\left(\prod_{i=1}^m (x - x_i)\right) = 0 \Longleftrightarrow \exists 1 \leqslant i < j \leqslant m \colon x_i = x_j$.

D3. $\mathrm{D}(f \cdot g) = \mathrm{D}(f) \cdot \mathrm{D}(g) \cdot (\operatorname{Res}(f, g))^2$ для всех $f, g \in \mathbb{F}[x]$. [1.32]

D4. $\mathrm{D}(f(x + c)) = \mathrm{D}(f(x))$ для всех $f \in \mathbb{F}[x]$, $c \in \mathbb{F}$. [1.26]

D5. $\mathrm{D}(\lambda^0 a_0, \lambda^1 a_1, \ldots, \lambda^m a_m) = \lambda^{m(m-1)} \mathrm{D}(a_0, \ldots, a_m)$ для всех $\lambda \in \mathbb{F}$. [1.24]

В [8, глава 12] также приведены явные формулы для дискриминанта многочлена второй, третьей, четвёртой и пятой степеней (формулы [1.33–1.36]).

Замечание 1.5. Некоторые из перечисленных выше свойств результанта и дискриминанта обсуждаются также в [3] и [4].

Определение 1.6. Для $X \in \mathcal{M}_m$ через $\mathrm{D}(X)$ мы обозначаем дискриминант характеристического многочлена матрицы X. Он называется *дискриминантом* X. Другими словами,

$$\mathrm{D}(X) = \mathrm{D}(p_X(x)) = \mathrm{D}(\det(xI_m - X)).$$

Очевидно, что матричный дискриминант является многочленом от матричных коэффициентов. Из свойств дискриминанта многочлена вытекают следующие свойства матричного дискриминанта.

MD1. Если $X = \operatorname{diag}(a_1, \ldots, a_m)$, то

$$\mathrm{D}(X) = \prod_{1 \leqslant i < j \leqslant m} (a_i - a_j)^2.$$

MD2. Если $X \in \mathcal{M}_m, Y \in \mathcal{M}_k$, то

$$\mathrm{D}(X \oplus Y) = \mathrm{D}\left(p_X(x) p_Y(x)\right) = \mathrm{D}(X) \cdot \mathrm{D}(Y) \cdot \left(\operatorname{Res}(p_X(x), p_Y(x))\right)^2.$$

MD3. Если $c \in \mathbb{F}$, то $\mathrm{D}(X + cI_m) = \mathrm{D}(X)$.

MD4. $\mathrm{D}(M^{-1}XM) = \mathrm{D}(X)$ для любой обратимой матрицы $M \in \mathcal{M}_m$.
MD5. $\mathrm{D}(\lambda X) = \lambda^{m(m-1)}\,\mathrm{D}(X)$ для всех $\lambda \in \mathbb{F}$.

Из последнего свойства следует, что матричный дискриминант является однородным многочленом степени $m(m-1)$ от матричных коэффициентов.

Приведём здесь явную формулу для дискриминанта 2×2 матрицы, которая нам понадобится в дальнейшем.

Если $X \in \mathcal{M}_2$, то

$$\mathrm{D}(X) = X_{1\,1}^2 - 4X_{1\,2}X_{2\,1} - 2X_{1\,1}X_{2\,2} + X_{2\,2}^2. \tag{1.2}$$

Замечание 1.7. Формула, выражающая дискриминант общей матрицы порядка 3 через её элементы над полем нулевой характеристики, содержит 144 члена, а аналогичная формула для общей матрицы порядка 4 содержит 72124 члена.

Замечание 1.8. Мы представили дискриминант $\mathrm{D}(X)$ матрицы $X \in \mathcal{M}_m$ в виде результанта характеристического многочлена и его производной. Помимо этого представления, существуют и другие: в [11] дано представление $\mathrm{D}(X)$ через степени X и X^t с помощью формулы Бине-Коши, а в [2] дано представление $\mathrm{D}(X)$ в виде суммы квадратов $n!$ многочленов, если $\mathbb{F} = \mathbb{C}$, а X является нормальной матрицей.

Мы определили все понятия, необходимые для формулировки основного результата нашей статьи. Приведём его здесь.

Теорема 1.9 (основная теорема, см. теорему 3.6)**.** *Линейное отображение T пространства матриц $\mathcal{M}_n$ удовлетворяет условию $\mathrm{D}(T(X)) = \mathrm{D}(X)$ для всех $X \in \mathcal{M}_n$ тогда и только тогда, когда существуют обратимая матрица $M \in \mathcal{M}_n$, линейное отображение $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ и $c \in \mathbb{F}$ такие, что $c^{n(n-1)} = 1$ и либо*

$$T(X) = cM^{-1}XM + \alpha(X)I_n \quad \textit{для всех } X \in \mathcal{M}_n,$$

либо

$$T(X) = cM^{-1}X^tM + \alpha(X)I_n \quad \textit{для всех } X \in \mathcal{M}_n.$$

Дальнейший текст организован следующим образом. В §2 мы доказываем вспомогательные утверждения, в том числе связываем свойство “быть нильпотентной матрицей ранга 1” со степенью матричного дискриминанта как многочлена от некоторой переменной матрицы, а также исследуем свойства линейных отображений, сохраняющих матричный дискриминант, и специальных пространств, образованных нильпотентными матрицами ранга 1. В §3 мы доказываем основную теорему, предварительно сформулировав и обосновав утверждения о преобразованиях, которым подвергаются при этом рассматриваемые отображения. Приложение A посвящено обсуждению утверждений из [13], которые, по нашему мнению, нуждаются в уточнении, дополнении или исправлении.

2. Вспомогательные утверждения и определения

В следующих трёх леммах мы дадим степенную характеризацию нильпотентных матриц ранга 1 (с точностью до сложения со скалярной матрицей).

Лемма 2.1. *Пусть $X \in \mathcal{M}_n$. Предположим, что X является линейной комбинацией скалярной матрицы и нильпотентной матрицы ранга* 1, *то есть*

$$X = cI_n + dN$$

для некоторых $c, d \in \mathbb{F}$ и некоторой нильпотентной матрицы $N \in \mathcal{M}_n$ ранга 1. *Тогда*

$$\deg_\lambda(\mathrm{D}(A + \lambda X)) \leqslant 2n - 3$$

для всех $A \in \mathcal{M}_n$.

Доказательство. Предположим, что $X = cI_n + dN$ для некоторых $c, d \in \mathbb{F}$ и некоторой нильпотентной матрицы $N \in \mathcal{M}_n$ ранга 1. Если $d = 0$, то $\mathrm{D}(A + \lambda X) = \mathrm{D}(A)$ для всех $A \in \mathcal{M}_n$, потому что операция сложения со скалярной матрицей сохраняет дискриминант. Значит, в этом случае

$$\deg_\lambda(\mathrm{D}(A + \lambda X)) = 0 \leqslant 2n - 3.$$

Предположим теперь, что $d \neq 0$. В этом случае, матрица dN также является нильпотентной матрицей ранга 1. Прибавление скалярной матрицы сохраняет дискриминант, поэтому

$$\mathrm{D}(A + \lambda X) = \mathrm{D}(A + \lambda(cI_n + dN)) = \mathrm{D}(A + \lambda(dN))$$

для всех $A \in \mathcal{M}_n, \lambda \in \mathbb{F}$, откуда

$$\deg_\lambda(\mathrm{D}(A+\lambda X)) = \deg_\lambda(\mathrm{D}(A+\lambda(cI_n+dN))) = \deg_\lambda(\mathrm{D}(A+\lambda(dN)))$$

для всех $A \in \mathcal{M}_n$. Таким образом, мы можем предположить, что X является нильпотентной матрицей ранга 1. Далее, всякая нильпотентная матрица ранга 1 подобна матричной единице $E_{1\,2}$, поэтому существует обратимая матрица $M \in \mathcal{M}_n$ такая, что $X = M^{-1}E_{1\,2}M$. Тогда

$$\begin{aligned}\mathrm{D}(A+\lambda X) &= \mathrm{D}(A+\lambda M^{-1}E_{1\,2}M)\\ &= \mathrm{D}\left(M^{-1}(MAM^{-1}+\lambda E_{1\,2})M\right) = \mathrm{D}(MAM^{-1}+\lambda E_{1\,2})\end{aligned}$$

для всех $A \in \mathcal{M}_n$, $\lambda \in \mathbb{F}$. Преобразование $A \mapsto MAM^{-1}$ является обратимым линейным преобразованием множества всех матриц, следовательно, утверждения

$$\deg_\lambda(\mathrm{D}(A+\lambda X)) \leqslant 2n-3, \quad A \in \mathcal{M}_n,$$

и

$$\deg_\lambda(\mathrm{D}(A+\lambda E_{1\,2}) \leqslant 2n-3, \quad A \in \mathcal{M}_n,$$

равносильны. По этой причине, не нарушая общности, мы предполагаем, что $X = E_{1\,2}$.

Зафиксируем $A \in \mathcal{M}_n$ и рассмотрим $p_{A+\lambda E_{1\,2}}(x)$:

$$p_{A+\lambda E_{1\,2}}(x) = \begin{vmatrix} x-A_{1\,1} & -A_{1\,2}-\lambda & \dots & -A_{1\,n}\\ -A_{2\,1} & x-A_{2\,2} & \dots & -A_{2\,n}\\ \vdots & \vdots & \ddots & \vdots\\ -A_{n\,1} & -A_{n\,2} & \dots & x-A_{n\,n}\end{vmatrix}.$$

Раскрывая этот определитель по первой строке, получаем, что

$$p_{A+\lambda E_{1\,2}}(x) = P + (A_{1\,2}+\lambda)Q$$

для некоторых многочленов $P, Q \in \mathbb{F}[x]$, причём $\deg_x(Q) = n-2$. Отсюда следует, что

$$p_{A+\lambda E_{1\,2}}(x) = x^n + b_{n-1}x^{n-1} + (a_{n-2}\lambda + b_{n-2})x^{n-2} + \ldots + (a_0\lambda + b_0)$$

для некоторых $a_0, \ldots, a_{n-2}, b_0, \ldots, b_{n-1} \in \mathbb{F}$.

Рассмотрим теперь многочлен $\mathrm{D}(A+\lambda E_{1\,2})$, записанный как результант

$$\mathrm{D}(A+\lambda E_{1\,2}) = \mathrm{Res}(p_{A+\lambda E_{1\,2}}(x), p'_{A+\lambda E_{1\,2}}(x)) = \det(R),$$

где

$$R = \begin{pmatrix} 1 & \cdots & a_1\lambda + b_1 & a_0\lambda + b_0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & b_{n-1} & \cdots & a_0\lambda + b_0 \\ n & \cdots & 2a_2\lambda + 2b_2 & a_1\lambda + b_1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & n & (n-1)b_{n-1} & \cdots & a_1\lambda + b_1 \end{pmatrix}.$$

Заметим, что $R \in \mathcal{M}_{2n-1}$ и первые два столбца R не содержат λ, а всякий элемент в остальных $2n-3$ столбцах содержит λ максимум в первой степени. Поэтому

$$\deg_\lambda \left(\prod_{1 \leqslant i \leqslant n} R_{i\,\sigma(i)} \right) \leqslant 2n - 1 - 2 = 2n - 3$$

для всякой перестановки σ, откуда $\deg_\lambda(\det(R)) \leqslant 2n-3$. Так как $\det(R) = \mathrm{D}(A + \lambda E_{1\,2})$, то мы заключаем, что $\deg_\lambda \mathrm{D}(A + \lambda E_{1\,2}) \leqslant 2n-3$. □

Лемма 2.2. *Предположим, что поле $\mathbb{F}$ алгебраически замкнуто и матрица $X \in \mathcal{M}_n$ не является линейной комбинацией скалярной матрицы и нильпотентной матрицы ранга 1, то есть*

$$X \neq cI_n + dN$$

для всех $c, d \in \mathbb{F}$ и нильпотентных матриц $N \in \mathcal{M}_n$ ранга 1. Тогда существует матрица $A \in \mathcal{M}_n$ такая, что

$$\deg_\lambda(\mathrm{D}(A + \lambda X)) \geqslant 2n - 2.$$

Доказательство. Пусть X_J – жорданова матрица, подобная X, то есть X_J жорданова и

$$X = M^{-1} X_J M$$

для некоторой обратимой матрицы $M \in \mathcal{M}_n$ (такая матрица X_J существует, так как мы предположили, что $\mathbb{F}$ алгебраически замкнуто). Тогда

$$\begin{aligned} \mathrm{D}(A + \lambda X) &= \mathrm{D}(A + \lambda M^{-1} X_J M) \\ &= \mathrm{D}\left(M^{-1}(MAM^{-1} + \lambda X_J)M\right) = \mathrm{D}(MAM^{-1} + \lambda X_J) \end{aligned}$$

для всех $A \in \mathcal{M}_n$, $\lambda \in \mathbb{F}$.

Преобразование $A \mapsto MAM^{-1}$ является обратимым линейным преобразованием множества всех матриц, поэтому

$$\max_A (\deg_\lambda(\mathrm{D}(A+\lambda X))) = \max_A (\deg_\lambda(\mathrm{D}(A+\lambda X_J)) .$$

Также, равенство

$$X_J = c_J I_n + d_J N_J, \tag{2.1}$$

в котором $c_J, d_J \in \mathbb{F}$, а $N_J \in \mathcal{M}_n$ является нильпотентной матрицей ранга 1, равносильно равенству

$$X = cI_n + dN$$

для некоторых $c, d \in \mathbb{F}$ и нильпотентной матрицы $N \in \mathcal{M}_n$ ранга 1. Действительно, достаточно умножить обе части равенства (2.1) на M^{-1} слева и на M справа.

По этим причинам, не ограничивая общности, мы предполагаем, что X жорданова и $X_{(i,i)} = \xi_i$, $\xi_i \in \mathbb{F}$. Возможны следующие случаи:

(1) X имеет как минимум два различных собственных значения;
(2) все собственные значения X равны между собой:
  (a) X содержит жорданов блок, размер которого не меньше 3;
  (b) все жордановы блоки X имеют размер не больше 2.

Рассмотрим их по-отдельности.

Случай 1: X имеет как минимум два различных собственных значения. В этом случае возьмём матрицу $A = \mathrm{diag}(a_1, \dots a_n)$, такую что все $a_1, \dots, a_n$ различны (напомним, что $\mathbb{F}$ бесконечно). Тогда

$$\begin{aligned}\mathrm{D}(A+\lambda X) &= \mathrm{D}(\mathrm{diag}(a_1+\lambda\xi_1, \dots, a_n+\lambda\xi_n)) \\ &= \prod_{i<j}(a_i - a_j + \lambda(\xi_i-\xi_j))^2 = (-1)^{\frac{n(n-1)}{2}} \prod_{1\leqslant i\neq j\leqslant n}(a_i-a_j+\lambda(\xi_i-\xi_j)),\end{aligned}$$

откуда $\deg_\lambda(\mathrm{D}(A+\lambda X)) = \#\{(i,j) \mid \xi_i \neq \xi_j\}$.

Пусть $m = \#\{j \mid \xi_j = \xi_1\}$. Тогда $1 \leqslant m \leqslant n-1$ и

$$\#\{(i,j) \mid \xi_i \neq \xi_j\} \geqslant \#\left(\{(1,j) \mid \xi_1 \neq \xi_j\} \cup \{(i,1) \mid \xi_i \neq \xi_1\}\right) = 2m(n-m).$$

Также

$$\begin{aligned}2m(n-m) - (2n-2) &= 2(n(m-1) - (m^2-1)) \\ &= 2((n-m+1)(m-1)) \geqslant 0.\end{aligned}$$

Следовательно,

$$2m(n-m) \geqslant (2n-2).$$

Поэтому

$$\deg_\lambda(\mathrm{D}(A+\lambda X)) = \#\{(i,j) \mid \xi_i \neq \xi_j\} \geqslant 2m(n-m) \geqslant 2n-2.$$

Случай 2: Все собственные значения X равны между собой. В таком случае мы предполагаем, что $\xi_i = 0$ для всех $1 \leqslant i \leqslant n$, так как сложение со скалярной матрицей не меняет матричный дискриминант, а, значит, и степень многочлена $\mathrm{D}(A+\lambda X)$, и не влияет на возможность представления X в виде линейной комбинации скалярной матрицы и нильпотентной матрицы ранга 1. Теперь рассмотрим два подслучая.

Подслучай 2(а): X содержит жорданов блок, размер которого не меньше 3. Заметим, что если $\mathrm{char}(\mathbb{F}) = 3$, то $n > 3$, так как, по нашему соглашению, $\mathrm{char}(\mathbb{F}) \nmid n$. Итак, в этом случае $X = \left[\begin{array}{c|c} X_0 & * \\ \hline 0 & X' \end{array}\right]$, где $X_0 = \left(\begin{smallmatrix} 0&1&0\\0&0&1\\0&0&0 \end{smallmatrix}\right)$, а $X' \in \mathcal{M}_{n-3}$. Рассмотрим $A = A_0 \oplus A'$, где $A_0 = \left(\begin{smallmatrix} 0&0&0\\1&0&0\\1&0&0 \end{smallmatrix}\right)$, $A' = \mathrm{diag}(a_4,\dots,a_n)$ и $a_4,\dots a_n$ – фиксированные попарно различные элементы $\mathbb{F}$. Заметим, что $*$ в X не влияет на $p_{A+\lambda X}$, а значит, и на $\mathrm{D}(A+\lambda X)$.

Рассмотрим цепочку равенств

$$\begin{aligned}
&\mathrm{D}(A+\lambda X) = \mathrm{D}((A_0 \oplus A')+\lambda(X_0 \oplus X')) = \mathrm{D}((A_0+\lambda X_0) \oplus (A'+\lambda X')) \\
&= \mathrm{D}(A_0+\lambda X_0)\,(\mathrm{Res}(p_{A_0+\lambda X_0}(x), p_{A'+\lambda X'}(x)))^2\,\mathrm{D}(A'+\lambda X') \\
&= \mathrm{D}(A_0+\lambda X_0)\,\big(\mathrm{Res}(p_{A_0+\lambda X_0}(x), p_{\mathrm{diag}(a_4,\dots a_n)}(x))\big)^2\,\mathrm{D}(\mathrm{diag}(a_4,\dots,a_n)) \\
&= (4\lambda^3 - 27\lambda^4)\left(\mathrm{Res}(-x^3+\lambda x+\lambda^2, \prod_{4\leqslant i\leqslant n}(x-a_i))\right)^2 \prod_{4\leqslant i<j\leqslant n}(a_i-a_j)^2.
\end{aligned}$$

Далее,

$$\mathrm{Res}(-x^3+\lambda x+\lambda^2, \prod_{4\leqslant i\leqslant n}(x-a_i)) = \prod_{4\leqslant i\leqslant n}(-a_i^3+\lambda a_i+\lambda^2).$$

Значит, $\deg_\lambda(\mathrm{Res}(-x^3+\lambda x+\lambda^2, \prod_{4\leqslant i\leqslant n}(x-a_i))) = 2(n-3)$.

Следовательно,

$$\deg_\lambda(\mathrm{D}(A+\lambda X))$$
$$= \deg_\lambda \left((4\lambda^3 - 27\lambda^4) \left(\mathrm{Res}(-x^3+\lambda x+\lambda^2, \prod_{4\leqslant i\leqslant n}(x-a_i)) \right)^2 \right.$$
$$\left. \times \prod_{4\leqslant i<j\leqslant n}(a_i-a_j)^2 \right) = \begin{cases} 4+2(2(n-3)) = 4n-8, & \mathrm{char}(\mathbb{F}) \neq 3; \\ 3+2(2(n-3)) = 4n-9, & \mathrm{char}(\mathbb{F}) = 3. \end{cases}$$

Заметим, что $4n-8 \geqslant 2n-2$ при $n \geqslant 3$ и $4n-9 \geqslant 2n-2$ при $n \geqslant 4$. Поэтому $\deg_\lambda(\mathrm{D}(A+\lambda X)) \geqslant 2n-2$ для данной матрицы A.

Подслучай 2(b): Каждый жорданов блок матрицы X имеет размер не больше 2. Так как X по предположению не является нильпотентной матрицей ранга 1, то X имеет как минимум два таких жордановых блока, и поэтому $n \geqslant 4$. Более того, так как n не делится на $\mathrm{char}(\mathbb{F})$, в случае $\mathrm{char}(\mathbb{F}) = 2$ мы получаем, что $n \geqslant 5$.

Таким образом, $X = X_0 \oplus X'$, где $X_0 = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \oplus \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, а $X' \in \mathcal{M}_{n-4}$. Рассмотрим $A = A_0 \oplus A'$, где $A_0 = \left(\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{smallmatrix}\right)$, $A' = \mathrm{diag}(a_5, \ldots, a_n)$, а $a_5, \ldots a_n$ – фиксированные попарно различные элементы $\mathbb{F}$. Тогда мы имеем:

$$\mathrm{D}(A+\lambda X) = \mathrm{D}((A_0 \oplus A') + \lambda(X_0 \oplus X')) = \mathrm{D}((A_0+\lambda X_0) \oplus (A'+\lambda X'))$$
$$= \mathrm{D}(A_0+\lambda X_0)\left(\mathrm{Res}(p_{A_0+\lambda X_0}(x), p_{A'+\lambda X'}(x))\right)^2 \mathrm{D}(A'+\lambda X')$$
$$= \mathrm{D}(A_0+\lambda X_0)\left(\mathrm{Res}(p_{A_0+\lambda X_0}(x), p_{\mathrm{diag}(a_4,\ldots a_n)}(x))\right)^2 \mathrm{D}(\mathrm{diag}(a_5,\ldots,a_n))$$
$$= \lambda^4(27+256\lambda^2)\left(\mathrm{Res}(x^4-x^3-\lambda^2, \prod_{5\leqslant i\leqslant n}(x-a_i))\right)^2 \prod_{5\leqslant i<j\leqslant n}(a_i-a_j)^2.$$

Рассмотрим $\mathrm{Res}(x^4-x^3-\lambda^2, \prod_{5\leqslant i\leqslant n}(x-a_i))$:

$$\mathrm{Res}(x^4-x^3-\lambda^2, \prod_{5\leqslant i\leqslant n}(x-a_i)) = \prod_{5\leqslant i\leqslant n}(a_i^4-a_i^3-\lambda^2).$$

Следовательно, $\deg_\lambda(\mathrm{Res}(x^4-x^3-\lambda^2, \prod_{5\leqslant i\leqslant n}(x-a_i))) = 2(n-4)$. Тогда

$$\begin{aligned}&\deg_\lambda(\mathrm{D}(A+\lambda X))\\ &\quad = \deg_\lambda\left(\lambda^4(27+256\lambda^2)\left(\mathrm{Res}(x^4-\lambda^2, \prod_{4\leqslant i\leqslant n}(x-a_i))^2\right)^2\right.\\ &\quad\left.\times \prod_{5\leqslant i<j\leqslant n}(a_i-a_j)^2\right) = \begin{cases}6+2(2(n-4)) = 4n-10, & \mathrm{char}(\mathbb{F})\neq 2;\\ 4+2(2(n-4)) = 4n-12, & \mathrm{char}(\mathbb{F}) = 2.\end{cases}\end{aligned}$$

Заметим, что $4n-10 \geqslant 2n-2$ при $n\geqslant 4$ и $4n-12\geqslant 2n-2$ при $n\geqslant 5$. Поэтому $\deg_\lambda(\mathrm{D}(A+\lambda X))\geqslant 2n-2$ для заданной матрицы A. □

ndremark

Следствие 2.3. *Предположим, что поле $\mathbb{F}$ алгебраически замкнуто и матрица $X\in\mathcal{M}_n$ такова, что*

$$\deg_\lambda(\mathrm{D}(A+\lambda(X)))\leqslant 2n-3$$

для всех $A\in\mathcal{M}_n$. Тогда X является линейной комбинацией скалярной матрицы и нильпотентной матрицы ранга 1, *то есть*

$$X = cI_n + dN$$

для некоторых $c, d\in\mathbb{F}$ и некоторой нильпотентной матрицы $N\in\mathcal{M}_n$ ранга 1.

Лемма 2.4. *Утверждение следствия* 2.3 *верно без предположения об алгебраической замкнутости поля $\mathbb{F}$.*

Доказательство. Пусть $X\in\mathbb{F}$ и

$$\deg_\lambda(\mathrm{D}(A+\lambda X))\leqslant 2n-3 \tag{2.2}$$

для всех $A\in\mathcal{M}_n(\mathbb{F})$. Если матрица X фиксирована, то выражение $\mathrm{D}(A+\lambda X)$ является многочленом от λ и коэффициентов матрицы A, т.е. $\mathrm{D}(A+\lambda X)\in\mathbb{F}[\lambda, A_{1\,1}, A_{1\,2},\ldots, A_{n\,n}]$. Разложим его по степеням λ :

$$\begin{aligned}\mathrm{D}(A+\lambda X) &= D_0(A_{1\,1}, A_{1\,2},\ldots, A_{n\,n})\\ &\quad + D_1(A_{1\,1}, A_{1\,2},\ldots, A_{n\,n})\lambda + \ldots + D_k(A_{1\,1}, A_{1\,2},\ldots, A_{n\,n})\lambda^k\end{aligned}$$

для некоторого целого неотрицательного k и

$$D_0,\ldots, D_k\in\mathbb{F}[A_{1\,1}, A_{1\,2},\ldots, A_{n\,n}].$$

Из неравенства (2.2) следует, что $k \leqslant 2n-3$.

Теперь заметим, что для алгебраического замыкания $\overline{\mathbb{F}} \supset \mathbb{F}$ мы будем иметь то же разложение для $\mathrm{D}(A+\lambda X)$. Поэтому неравенство (2.2) выполнено также для всех $A \in \mathcal{M}_n(\overline{\mathbb{F}})$. Значит,

$$X = c' I_n + d' N' \tag{2.3}$$

для некоторых $c', d' \in \overline{\mathbb{F}}$ и нильпотентной матрицы $N \in \mathcal{M}_n(\overline{\mathbb{F}})$ ранга 1. Докажем, что $c' \in \mathbb{F}$, а $d'N' \in \mathcal{M}_n(\mathbb{F})$.

Если мы возьмём след от левой и правой частей равенства (2.3), то получим равенство

$$\mathbb{F} \ni \operatorname{tr}(X) = \operatorname{tr}(c' I_n + d' N') = c' \operatorname{tr}(I_n) + d' \operatorname{tr}(N) = c' n,$$

так как след каждой нильпотентной матрицы равен нулю. Число n обратимо в $\mathbb{F}$, так как $\operatorname{char}(\mathbb{F}) \nmid n$, из чего мы получаем, что $c \in \mathbb{F}$. Значит, $dN = X - cI_n$ принадлежит $\mathcal{M}_n(\mathbb{F})$.

Теперь, если $N' \in \mathcal{M}_n(\mathbb{F})$, то мы можем положить $N = N'$, $d = d'$, $c = c'$ в требуемом представлении X в поле $\mathbb{F}$. Предположим, что $N' \notin \mathcal{M}_n(\mathbb{F})$. Если $d' = 0$, то в этом представлении мы можем положить N равной любой нильпотентной матрице ранга 1 из $\mathcal{M}_n(\mathbb{F})$, например. $N = E_{1\,2}$ и, далее, $d = d' = 0$, $c = c'$. Если же $d' \neq 0$, то положим $N = d'N'$, $d = 1$, $c = c'$. В любом случае, получено требуемое представление для X с помощью элементов из $\mathbb{F}$ и $\mathcal{M}_n(\mathbb{F})$ соответственно. □

Заметим, что, согласно свойству MD3 матричного дискриминанта из §1, выполнено следующее: если X – скалярная матрица, то $\mathrm{D}(A+\lambda X) = \mathrm{D}(A)$ для любых $A \in \mathcal{M}_n$ и $\lambda \in \mathbb{F}$. Докажем обратное утверждение.

Лемма 2.5. *Предположим, что матрица $X \in \mathcal{M}_n$ такова, что*

$$\mathrm{D}(A+\lambda X) = \mathrm{D}(A) \tag{2.4}$$

для любых $\lambda \in \mathbb{F}$ и $A \in \mathcal{M}_n$. Тогда $X = cI_n$ для некоторого $c \in \mathbb{F}$.

Доказательство. Из условия леммы следует, что $\deg_\lambda(\mathrm{D}(A+\lambda X))=0$ для любой матрицы $A \in \mathcal{M}_n$. По лемме 2.4, в таком случае $X = cI_n + dN$ для некоторых $c, d \in \mathbb{F}$ и нильпотентной матрицы N ранга 1. Докажем от противного, что $d = 0$.

Заметим сначала, что, так как

$$\mathrm{D}(A+\lambda dN) = \mathrm{D}(A+\lambda(X-cI_n)) = \mathrm{D}(A+\lambda X)$$

для всех $A \in \mathcal{M}_n$ и $\lambda \in \mathcal{M}_n$, то

$$\deg_\lambda(\mathrm{D}(A + \lambda dN)) = \deg_\lambda(\mathrm{D}(A + \lambda X)) = 0$$

для всех $A \in \mathcal{M}_n$. Матрица N – нильпотентная матрица ранга 1, так же как и dN. Не нарушая общности, предположим, что $dN = E_{1\,2}$. Рассмотрим два случая: $n = 2$ и $n \geqslant 3$.

Случай 1: $n = 2$**.** Так как $\mathrm{char}(\mathbb{F})$ не делит n, то $\mathrm{char}(\mathbb{F}) \neq 2$. Рассмотрим матрицу $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. Тогда

$$\mathrm{D}(A + \lambda dN) = \mathrm{D}(A + \lambda E_{1\,2}) = 4\lambda,$$

что не равно нулю, так как $\mathrm{char}(\mathbb{F}) \neq 2$. Значит,

$$\deg_\lambda(A + \lambda dN) = 1 > 0.$$

Случай 2: $n \geqslant 3$**.** Разложим $E_{1\,2}$ в прямую сумму двух матриц:

$$E_{1\,2} = E_0 \oplus O_{n-3},$$

где $E_0 = \left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$.

Рассмотрим матрицу $A = A_0 \oplus A'$, где $A_0 = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{smallmatrix}\right)$, а $A' = \mathrm{diag}(a_4, \ldots, a_n)$, и $a_4, \ldots, a_n$ – произвольные фиксированные попарно различные элементы $\mathbb{F}$. Тогда

$$\mathrm{D}(A + \lambda dN) = \mathrm{D}((A_0 \oplus A') + \lambda E_{1\,2}) = \mathrm{D}((A_0 + \lambda E_0) \oplus (A' + \lambda O_{n-3}))$$

$$= \mathrm{D}(A_0 + \lambda E_0)\left(\mathrm{Res}(p_{A_0+\lambda E_0}(x), p_{A'}(x))\right)^2 \mathrm{D}(A')$$

$$= \mathrm{D}(A_0 + \lambda E_0)\left(\mathrm{Res}(p_{A_0+\lambda E_0}(x), p_{\mathrm{diag}(a_4,\ldots a_n)}(x))\right)^2 \mathrm{D}(\mathrm{diag}(a_4, \ldots, a_n))$$

$$= (\lambda^2(-27+4\lambda))\left(\mathrm{Res}(-x^3+\lambda x+\lambda, \prod_{4\leqslant i\leqslant n}(x-a_i))\right)^2 \prod_{4\leqslant i<j\leqslant n}(a_i-a_j)^2.$$

Рассмотрим $\mathrm{Res}(-x^3 + \lambda x + \lambda, \prod_{4\leqslant i}(x - a_i))$:

$$\mathrm{Res}(-x^3 + \lambda x + \lambda, \prod_{4\leqslant i\leqslant n}(x - a_i)) = \prod_{4\leqslant i\leqslant n}(a_i^3 - \lambda a_i + \lambda).$$

Поэтому $\mathrm{Res}(-x^3 + \lambda x + \lambda, \prod_{4\leqslant i\leqslant n}(x - a_i)) \neq 0$, откуда

$$\deg_\lambda(\mathrm{D}(A + \lambda dN)) \geqslant \deg_\lambda(\mathrm{D}(A_0+\lambda E_{1\,2})) = \deg_\lambda(\lambda^2(-27+4\lambda)) \geqslant 2 > 0.$$

В обоих случаях, мы нашли матрицу A такую, что $\deg_\lambda(A+\lambda dN)>0$. Это противоречит тому, что $\deg_\lambda(A+\lambda dN)=0$, как мы установили вначале. Значит, наше предположение неверно, и $d=0$. Следовательно, $X=cI_n$ для некоторого $c\in\mathbb{F}$. □

Это свойство скалярных матриц позволяет доказать, что множество скалярных матриц является инвариантным пространством относительно действия обратимого линейного отображения, сохраняющего матричный дискриминант.

Лемма 2.6. *Предположим, что линейное обратимое отображение T пространства матриц $\mathcal{M}_n$ сохраняет матричный дискриминант. Тогда $T(I_n)=cI_n$ для некоторого $c\in\mathbb{F}$.*

Доказательство. Пусть $T(I_n)=X\in\mathcal{M}_n$. Согласно свойству MD3 матричного дискриминанта из §1, $\mathrm{D}(A+\lambda I_n)=\mathrm{D}(A)$ для любых $\lambda\in\mathbb{F}$ и $A\in\mathcal{M}_n$. Поэтому

$$\begin{aligned}\mathrm{D}(A+\lambda X)&=\mathrm{D}(A+\lambda T(I_n))=\mathrm{D}(T(T^{-1}(A))+\lambda T(I_n))\\&=\mathrm{D}(T(T^{-1}(A)+\lambda I_n))=\mathrm{D}(T^{-1}(A)+\lambda I_n)\\&=\mathrm{D}(T^{-1}(A))=\mathrm{D}(T(T^{-1}(A)))=\mathrm{D}(A)\end{aligned}$$

для любых $\lambda\in\mathbb{F}$ и $A\in\mathcal{M}_n$. Значит, по лемме 2.5, $X=cI_n$ для некоторого $c\in\mathbb{F}$. □

Следствие 2.7. *Предположим, что линейное отображение T пространства матриц $\mathcal{M}_n$ сохраняет след матрицы и матричный дискриминант. Тогда $T(I_n)=I_n$.*

Лемма 2.8. *Предположим, что линейное отображение T пространства матриц $\mathcal{M}_n$ сохраняет след матрицы и матричный дискриминант. Тогда T является биективным линейным отображением.*

Доказательство. Пусть $T(X)=0$ для некоторого $0\neq X\in M_n$. Тогда, так как T сохраняет дискриминант, то

$$\mathrm{D}(A+\lambda X)=\mathrm{D}(T(A+\lambda X))=\mathrm{D}(T(A))=\mathrm{D}(A)$$

для любых $\lambda\in\mathbb{F}$ и $A\in M_n$. Тогда, по лемме 2.5, $X=cI_n$ для некоторого $c\in\mathbb{F}$.

Далее, так как T сохраняет след, то

$$cn=\mathrm{tr}(X)=\mathrm{tr}(T(X))=\mathrm{tr}(0)=0.$$

Поскольку n обратимо в поле $\mathbb{F}$, отсюда следует, что $c = 0$. Значит, $X = 0$, откуда $\ker(T) = 0$, и поэтому линейное отображение T конечномерного пространства $\mathcal{M}_n$ является биективным. $\square$

В доказательстве основной теоремы при $n \geqslant 3$ нам понадобится характеризация линейных отображений пространства матриц $\mathcal{M}_n$, имеющих вид $X \mapsto QXR$ для некоторых $Q, R \in \mathcal{M}_n$, сохраняющих след матрицы.

Лемма 2.9. *Предположим, что матрицы $Q, R \in \mathcal{M}_n$ таковы, что $\operatorname{tr}(QXR) = \operatorname{tr}(X)$ для всех $X \in \mathcal{M}_n$. Тогда $Q = R^{-1}$.*

Доказательство. Во-первых,

$$\operatorname{tr}(QE_{i\,j}R) = \operatorname{tr}(E_{i\,j}) = 0$$

при $i \neq j$, откуда

$$\sum_{k=1}^{n} Q_{i\,k}R_{k\,j} = \operatorname{tr}(QE_{i\,j}R) = 0 \tag{2.5}$$

для всех $1 \leqslant i \neq j \leqslant n$. Во-вторых,

$$\operatorname{tr}(QE_{i\,i}R) = \operatorname{tr}(E_{i\,i}) = 1$$

для всех $1 \leqslant i \leqslant n$, и поэтому

$$\sum_{k=1}^{n} Q_{i\,k}R_{k\,i} = \operatorname{tr}(QE_{i\,i}R) = 1 \tag{2.6}$$

для всех $1 \leqslant i \leqslant n$.

Очевидно, что набор из n^2 равенств, которые могут быть получены из (2.5) и (2.6), совпадает с набором равенств, достаточных для того, чтобы выполнялось равенство $QR = I_n$, откуда $Q = R^{-1}$. $\square$

Как показывает следующая лемма, при $n = 2$ условие сохранения следа матрицы и матричного дискриминанта является достаточным для того, чтобы получить исчерпывающее описание отображений, ему удовлетворяющих.

Лемма 2.10. *Предположим, что $n = 2$, линейное отображение T пространства матриц $\mathcal{M}_2$ сохраняет след матрицы и матричный дискриминант и выполнены равенства $T(E_{1\,2}) = E_{1\,2}$, $T(E_{2\,1}) = E_{2\,1}$. Тогда либо*

$$T(X) = X$$

для всех $X \in \mathcal{M}_2$, *либо*

$$T(X) = P^{-1}X^tP$$

для всех $X \in \mathcal{M}_2$, *где*

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

Доказательство. Так как матричные единицы $E_{1\,1}, E_{1\,2}, E_{2\,1}, E_{2\,2}$ образуют базис векторного пространства $\mathcal{M}_2$ и образы матриц $E_{1\,2}$ и $E_{2\,1}$ нам уже известны, то для того, чтобы полностью определить отображение T, нам достаточно определить образы матриц $E_{1\,1}$ и $E_{2\,2}$.

Пусть $T(E_{1\,1}) = A$ и $T(E_{2\,2}) = B$. Так как T сохраняет матричный дискриминант, то

$$\mathrm{D}(T(E_{1\,1} + \lambda E_{1\,2})) = \mathrm{D}(E_{1\,1} + \lambda E_{1\,2}) = 1$$

для всех $\lambda \in \mathbb{F}$. С другой стороны,

$$\mathrm{D}(T(E_{1\,1} + \lambda E_{1\,2})) = \mathrm{D}(T(E_{1\,1}) + \lambda T(E_{1\,2})) = \mathrm{D}(A + \lambda E_{1\,2}).$$

Далее, применяя формулу для дискриминанта 2×2 матрицы (формула (1.2), §1, после свойств матричного дискриминанта) к $\mathrm{D}(A+\lambda E_{1\,2})$, мы получаем, что

$$\mathrm{D}(T(E_{1\,1}+\lambda E_{1\,2})) = \mathrm{D}(A+\lambda E_{1\,2}) = A_{1\,1}^2+4(A_{1\,2}+\lambda)A_{2\,1}-2A_{1\,1}A_{2\,2}+A_{2\,2}^2$$

для всех $\lambda \in \mathbb{F}$. Поэтому

$$1 = A_{1\,1}^2 + 4(A_{1\,2} + \lambda)A_{2\,1} - 2A_{1\,1}A_{2\,2} + A_{2\,2}^2$$

для всех $\lambda \in \mathbb{F}$, откуда $A_{2\,1} = 0$, так как мы предполагаем, что $\mathrm{char}(\mathbb{F})$ не делит $n = 2$.

Аналогично, рассматривая $\mathrm{D}(A + \lambda E_{2\,1})$, получаем, что $A_{1\,2} = 0$. Из тех же соображений следует, что $B_{2\,1} = B_{1\,2} = 0$. Таким образом, матрицы A и B диагональны.

Далее, так как T сохраняет след матрицы, то

$$1 = \mathrm{tr}(E_{1\,1}) = \mathrm{tr}(A) = A_{1\,1} + A_{2\,2} \tag{2.7}$$

и

$$1 = \mathrm{tr}(E_{2\,2}) = \mathrm{tr}(B) = B_{1\,1} + B_{2\,2}. \tag{2.8}$$

Также, так как T сохраняет матричный дискриминат, то

$$1 = \mathrm{D}(E_{1\,1}) = \mathrm{D}(A) = A_{1\,1}^2 - 2A_{1\,1}A_{2\,2} + A_{2\,2}^2 = (A_{1\,1} - A_{1\,2})^2 \tag{2.9}$$

и

$$1 = \mathrm{D}(E_{2\,2}) = \mathrm{D}(B) = (B_{1\,1} - B_{1\,2})^2. \tag{2.10}$$

Система, составленная из уравнений (2.7) и (2.9), имеет только два решения относительно $(A_{1\,1}, A_{2\,2})$: $(1, 0)$ и $(0, 1)$. Точно такие же решения относительно $(B_{1\,1}, B_{2\,2})$ имеет система, составленная из уравнений (2.8) и (2.10). Так как отображение T биективно согласно лемме 2.8, то $A \neq B$, откуда мы получаем, что возможны два случая: $A = E_{1\,1}, B = E_{2\,2}$ и $A = E_{2\,2}, B = E_{1\,1}$.

Если $A = E_{1\,1}$ и $B = E_{2\,2}$, то отображение T является тождественным отображением, так как оно тождественно на базисных векторах $E_{1\,1}, E_{1\,2}, E_{2\,1}, E_{2\,1}$ пространства $\mathcal{M}_2$.

Если $A = E_{2\,2}$ и $B = E_{1\,1}$, то непосредственные вычисления показывают, что на базисных векторах $E_{1\,1}, E_{1\,2}, E_{2\,1}, E_{2\,1}$ пространства $\mathcal{M}_2$ T совпадает с отображением $X \mapsto P^{-1}X^tP$, где $P = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, а значит, совпадает с ним на всём пространстве $\mathcal{M}_2$.

Таким образом, в любом случае отображение T совпадает с одним из двух отображений, указанных в формулировке леммы. □

Используя вышеперечисленные факты, мы можем доказать основное свойство линейных отображений, сохраняющих след матрицы и матричный дискриминант, которое нам нужно для доказательства теоремы 3.6.

Лемма 2.11. *Предположим, что линейное отображение T пространства матриц $\mathcal{M}_n$ сохраняет след матрицы и матричный дискриминант. Тогда образ $T(L)$ всякой нильпотентной матрицы L ранга* 1 *также является нильпотентной матрицей ранга* 1.

Доказательство. Согласно лемме 2.8, всякое такое отображение T является биективным. Далее, пусть $L \in \mathcal{M}_n$ – нильпотентная матрица единичного ранга. Заметим, что при этом $\mathrm{tr}(L) = \mathrm{tr}(T(L)) = 0$, так как T сохраняет след, а след всякой нильпотентной матрицы равен нулю. Кроме того, L удовлетворяет условиям леммы 2.1, и поэтому $\deg_\lambda(\mathrm{D}(A + \lambda(L))) \leqslant 2n - 3$ для всех $A \in \mathcal{M}_n(\mathbb{F})$.

Рассмотрим $T(L)$. Для всякой матрицы $A \in \mathcal{M}_n$ мы имеем

$$\begin{aligned} \mathrm{D}(A + \lambda T(L)) &= \mathrm{D}(T(T^{-1}(A)) + \lambda T(L)) \\ &= \mathrm{D}(T(T^{-1}(A) + \lambda L)) = \mathrm{D}(T^{-1}(A) + \lambda L) \end{aligned} \tag{2.11}$$

при любом $\lambda \in \mathbb{F}$. Следовательно,

$$\deg_\lambda(\mathrm{D}(A + \lambda T(L))) \leqslant 2n - 3 \tag{2.12}$$

для всех $A \in \mathcal{M}_n(\mathbb{F})$. Из леммы 2.4 следует, что в таком случае

$$T(L) = cI_n + dN \tag{2.13}$$

для некоторых $c, d \in \mathbb{F}$ и некоторой нильпотентной матрицы $N \in \mathcal{M}_n$ ранга 1. Заметим, что $d \neq 0$, так как в противном случае мы бы имели $\dim(T^{-1}(\langle I \rangle) > 1$, что противоречит биективности T.

Так как матрица N нильпотентна, то $\mathrm{tr}(N) = 0$, и поэтому, взяв след от обеих частей равенства (2.13), мы получаем:

$$\begin{aligned} \mathrm{tr}(T(L)) &= \mathrm{tr}(cI_n + dN), \\ \mathrm{tr}(L) &= \mathrm{tr}(cI_n) + \mathrm{tr}(dN), \\ 0 &= c\,\mathrm{tr}(I_n) + d\,\mathrm{tr}(N), \\ 0 &= cn + 0, \end{aligned}$$

откуда $c = 0$, потому что $\mathrm{char}(\mathbb{F})$ не делит n, а значит, n обратимо в $\mathbb{F}$.

Мы получили, что $T(L) = dN$, где $d \neq 0$. В таком случае, dN, так же как и N, является нильпотентной матрицей ранга 1. □

Как отмечено выше, лемма 2.11 находит своё применение при доказательстве основной теоремы. Однако при $n \geqslant 3$ только этого утверждения недостаточно. Прежде мы должны определить специальные пространства, порождённые нильпотентными матрицами ранга 1, а также установить их свойства.

Определение 2.12. Пусть $0 \neq v \in \mathbb{F}^n$. Через $M_0(v)$ обозначается векторное пространство, порождённое всеми нильпотентными матрицами, образ которых равен $\langle v \rangle$:

$$M_0(v) = \langle \{X \in \mathcal{M}_n \colon \mathrm{Im}(X) = \langle v \rangle, \exists k \in \mathbb{N} \colon X^k = 0\} \rangle.$$

Через $M_0^t(v)$ обозначается пространство, состоящее из всех матриц вида X^t, где $X \in M_0(v)$:

$$M_0^t(v) = \{X^t \colon X \in M_0(v)\}.$$

Замечание 2.13. Мы используем обозначение $M_0(v)$ для этих пространств, потому что в [13] пространства похожего вида обозначаются через $M(v)$ (см. определение A.5 из приложения A).

Далее в тексте мы также будем использовать следующее обозначение для матриц ранга 1.

Определение 2.14. Пусть $a, b \in \mathbb{F}^n$. Через $J(a,b) \in \mathcal{M}_n$ обозначается матрица, определяемая равенством

$$J(a,b) = \begin{pmatrix} a_1 \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} & \cdots & a_n \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \end{pmatrix}.$$

Таким образом, $(J(a,b))_{i\,j} = b_i a_j$ для всех $1 \leqslant i, j \leqslant n$.

Замечание 2.15. В некоторых работах (например, в [6]) рассуждения о матрицах ранга 1 проводятся с использованием языка тензорной алгебры, но это требует указания явного соответствия между тензорами и матрицами.

Перечислим основные свойства матриц, представимых в виде $J(a,b)$ для некоторых $a, b \in \mathbb{F}^n$:

J1. $J(\alpha a, b) = \alpha J(a,b)$ и $J(a, \beta b) = \beta J(a,b)$ для всех $a, b \in \mathbb{F}^n$ и $\alpha, \beta \in \mathbb{F}$;

J2. $J(a, b+c) = J(a,b) + J(a,c)$ и $J(a+b, c) = J(a,c) + J(b,c)$ для всех $a, b, c \in \mathbb{F}^n$.

Матрицы такого вида обладают и менее тривиальными свойствами. Докажем их.

Лемма 2.16. *Если $J(a,b) = J(a',b')$ для некоторых $a, b, a', b' \in \mathbb{F}^n$ отличных от нуля, то $a' = \alpha a$ и $b' = \frac{1}{\alpha} b$ для некоторого ненулевого $\alpha \in \mathbb{F}$.*

Доказательство. Так как a и b не равны нулю, то матрица $J(a,b)$ содержит ненулевой элемент, то есть $(J(a,b))_{i\,j} \neq 0$ для некоторых $1 \leqslant i, j \leqslant n$. Тогда

$$(J(a,b))_{i\,j} = a_j b_i = a'_j b'_i = (J(a',b'))_{i\,j}, \tag{2.14}$$

и потому все a_j, b_i, a'_j, b'_i не равны нулю. Пусть $\alpha = \frac{a'_j}{a_j}$. Тогда $\alpha \neq 0$ и $a'_j = \alpha a_j$. Из равенства (2.14) следует, что $b'_i = \frac{1}{\alpha} b_i$.

Далее, пусть $k \neq i$. Тогда из равенства

$$(J(a,b))_{k\,j} = a_j b_k = a'_j b'_k = (J(a',b'))_{k\,j}$$

следует, что $b'_k = \frac{1}{\alpha} b_k$, так как $a'_j = \alpha a_j$ и $\alpha, a_j \neq 0$. Поэтому $b'_i = \frac{1}{\alpha} b_i$ для всех $1 \leqslant i \leqslant n$, откуда $b' = \frac{1}{\alpha} b$.

Наконец, пусть $k \neq j$. Тогда из равенства

$$(J(a,b))_{i\,k} = a_k b_i = a'_k b'_i = (J(a',b'))_{i\,k}$$

по тем же соображениям следует, что $a'_k = \alpha a_k$. Поэтому $a'_i = \alpha a_i$ для всех $1 \leqslant i \leqslant n$, откуда $a' = \alpha a$ □

Лемма 2.17. *Пусть $X \in \mathcal{M}_n$ и $0 \neq v \in \mathbb{F}^n$. Тогда $\mathrm{Im}(X) = \langle v \rangle$ в том и только том случае, когда $X = J(a,v)$ для некоторого ненулевого $a \in \mathbb{F}^n$.*

Доказательство. Очевидно, что при ненулевом $a \in \mathbb{F}^n$ образ матрицы $J(a,v)$ равен $\langle v \rangle$.

Пусть теперь $X \in \mathcal{M}_n$ и $\mathrm{Im}(X) = \langle v \rangle$. Это означает, что всякий столбец X является вектором, коллинеарным v, то есть i-ый столбец X равен $a_i v$ для некоторого $a_i \in \mathbb{F}$ при всех $1 \leqslant i \leqslant n$. Так как $X \neq 0$, то $a \neq 0$. Тогда, положив $a = (a_1, \ldots, a_n)$, мы получаем, что $X = J(a,v)$. □

Лемма 2.18. *Пусть $a, b \in \mathbb{F}^n$. Тогда матрица $J(a,b)$ нильпотентна в том и только том случае, когда $(a,b) = 0$.*

Доказательство. Пусть $a, b \in \mathbb{F}^n$ и $(a,b) = 0$. Тогда

$$\begin{aligned}(J(a,b) \cdot J(a,b))_{i\,j} &= \sum_{k=1}^{n} (J(a,b))_{i\,k} \cdot (J(a,b))_{k\,j} \\ &= \sum_{k=1}^{n} a_k b_i \cdot a_j b_k = a_j b_i \cdot \sum_{k=1}^{n} a_k b_k = a_i b_j \cdot (a,b) = a_i b_j \cdot 0 = 0\end{aligned}$$

для всех $1 \leqslant i, j \leqslant n$, откуда $(J(a,b))^2 = 0$. Это означает, что $J(a,b)$ является нильпотентной матрицей.

Пусть теперь $a, b \in \mathbb{F}^n$ и матрица $J(a,b)$ нильпотентна. Так как след всякой нильпотентной матрицы равен нулю, то $\mathrm{tr}(J(a,b)) = 0$. Так как, с другой стороны,

$$\mathrm{tr}(J(a,b)) = \sum_{i=1}^{n} (J(a,b))_{i\,i} = a_i b_i = (a,b),$$

то $(a,b) = 0$. □

Лемма 2.19. *Пусть $0 \neq v \in \mathbb{F}^n$ и $X \in \mathcal{M}_n$. Тогда $X \in M_0(v)$ в том и только том случае, когда $X = J(a,v)$ для некоторого $a \in \mathbb{F}^n$ такого, что $(a,v) = 0$.*

Доказательство. Из лемм 2.17 и 2.18 следует, что все ненулевые матрицы вида $J(a, v)$, где $(a, v) = 0$, являются нильпотентными матрицами, образ которых равен $\langle v \rangle$, и потому они принадлежат $M_0(v)$.

Пусть $X \in M_0(v)$. Если $X = 0$, то $X = J(0, v)$. Если это не так, то включение $X \in M_0(v)$ означает, что X является суммой нескольких матриц ранга 1, образ которых равен v, то есть

$$X = A_1 + \ldots + A_k$$

для некоторого натурального $1 \leqslant k$ и нильпотентных матриц $A_i \in \mathcal{M}_n$, образ каждой из которых равен $\langle v \rangle$.

Из леммы 2.17 следует, что существуют $a^{(1)}, \ldots, a^{(k)} \in \mathbb{F}^n$ такие, что $A_i = J(a^{(i)}, v)$ для всех $1 \leqslant i \leqslant k$. Поэтому

$$X = A_1 + \ldots + A_k = J(a^{(1)}, v) + \ldots + J(a^{(k)}, v) = J(a^{(1)} + \ldots + a^{(k)}, v). \tag{2.15}$$

Далее, так как матрицы $A_1, \ldots, A_k$ нильпотентны, из леммы 2.18 следует, что $(a^{(i)}, v) = 0$ для всех $1 \leqslant i \leqslant k$. Поскольку

$$(a^{(1)} + \ldots + a^{(k)}, v) = (a^{(1)}, v) + \ldots + (a^{(k)}, v) = 0 + \ldots + 0 = 0,$$

то (2.15) является требуемым представлением для X. □

Следствие 2.20.

$$\dim(M_0(v)) = \dim(M_0^t(v)) = n - 1$$

для всех $0 \neq v \in \mathbb{F}^n$.

Следствие 2.21. *Для любых ненулевых векторов* $v^{(1)}, \ldots, v^{(k)} \in \mathbb{F}^n$ *выполнено неравенство*

$$\dim\left(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\right) \leqslant k(n-1).$$

Докажем лемму, характеризующую пространства вида $M_0(v)$ и $M_0^t(v)$ как максимальные пространства, все ненулевые элементы которых являются нильпотентными матрицами ранга 1.

Лемма 2.22. *Пусть все ненулевые элементы векторного пространства* $W \subset \mathcal{M}_n$ *являются нильпотентными матрицами ранга* 1. *Тогда найдётся* $0 \neq v \in \mathbb{F}^n$ *такой, что* $W \subset M_0(v)$ *или* $W \subset M_0^t(v)$.

Доказательство. Если $W = \{0\}$, то утверждение леммы выполняется очевидным образом. Пусть теперь $W \neq \{0\}$ и $0 \neq N \in W$. Фиксируем N. Заметим, что если матрица $P \in \mathcal{M}_n$ обратима и $0 \neq v \in \mathbb{F}^n$, то

$P^{-1}M_0(v)P = M_0(P^{-1}v)$, а $P^{-1}LP$ является нильпотентной матрицей ранга 1, если и только если L является нильпотентной матрицей ранга 1. Так как всякая нильпотентная матрица ранга 1 подобна матрице $E_{1\,2}$, это замечание позволяет нам, не нарушая общности, предположить, что $N = E_{1\,2}$.

Далее заметим, что наше утверждение равносильно тому, что выполнено следующее: либо $\operatorname{Im}(E_{1\,2}) = \operatorname{Im}(M)$ для всех $0 \neq M \in W$, либо $\operatorname{Im}(E_{1\,2}^t) = \operatorname{Im}(M^t)$ для всех $0 \neq M \in W$. Докажем это утверждение от противного. Предположим, что $M, K \in W$ – две ненулевые матрицы такие, что $\operatorname{Im}(E_{1\,2}) \neq \operatorname{Im}(M)$ и $\operatorname{Im}(E_{1\,2}^t) \neq \operatorname{Im}(K^t)$. Отсюда сразу следует, что матрицы M и K не пропорциональны $E_{1\,2}$.

Рассмотрим матрицы $E_{1\,2}$ и M. Так как их образы различны, найдётся $2 \leqslant i \leqslant n$ такое, что i-ая строка M не равна нулю. Предположим, что $1 \leqslant j \leqslant n$ таково, что $j \neq 2$ и $M_{i\,j} \neq 0$. Так как $\operatorname{rk}(M) = 1$, то существует ненулевой вектор (α_0, β_0) такой, что выполнены равенства

$$\alpha_0 M_{1\,2} + \beta_0 M_{1\,j} = 0 \tag{2.16}$$

и

$$\alpha_0 M_{i\,2} + \beta_0 M_{i\,j} = 0. \tag{2.17}$$

Далее, так как $0 \neq (E_{1\,2} + M) \in W$, то $\operatorname{rk}(E_{1\,2} + M) = 1$, и поэтому существует ненулевой вектор (α_1, β_1) такой, что выполнены равенства

$$\alpha_1(M_{1\,2} + 1) + \beta_1(M_{1\,j} + 0) = 0 \tag{2.18}$$

и

$$\alpha_1(M_{i\,2} + 0) + \beta_1(M_{i\,j} + 0) = 0. \tag{2.19}$$

Равенства (2.17) и (2.19) означают, что векторы (α_0, β_0) и (α_1, β_1) являются решениями уравнения

$$\alpha M_{i\,2} + \beta M_{i\,j} = 0.$$

Следовательно, векторы (α_0, β_0) и (α_1, β_1) пропорциональны, так как $(M_{i\,2}, M_{i\,j}) \neq 0$. Поэтому, не нарушая общности, мы можем предположить, что $(\alpha_1, \beta_1) = (\alpha_0, \beta_0)$.

Далее, вычитая (2.16) из (2.18), получаем, что

$$\alpha_0(M_{1\,2} + 1) + \beta_0(M_{1\,j} + 0) - (\alpha_0 M_{1\,2} + \beta_0 M_{1\,j}) = 0,$$
$$\alpha_0 = 0. \tag{2.20}$$

Так как $(\alpha_0, \beta_0) \neq 0$, из (2.20) получаем, что $\beta_0 \neq 0$. Тогда из (2.19) следует, что $M_{i\,j} = 0$, потому что $\beta_1 = \beta_0$. Противоречие. Значит, все элементы i-ой строки, кроме, быть может, второго, равны нулю. Так как $\mathrm{rk}(M) = 1$, то это верно и для остальных строк матрицы M. Значит, образ M^t равен образу $E_{1\,2}^t$. Отсюда следует, что все ненулевые элементы M находятся во втором столбце.

Повторяя аналогичные рассуждения для $E_{1\,2}^t$ и K^t, получаем, что образ K равен образу $E_{1\,2}$. Отсюда следует, что все ненулевые элементы K находятся в первой строке.

Рассмотрим теперь матрицу $K + M$. Все её ненулевые элементы содержатся в первой строке или во втором столбце.

Так как M не пропорциональна $E_{1\,2}$, то матрица $K + M$ содержит ненулевой элемент, который находится на пересечении второго столбца и строки, отличной от первой.

Аналогично, так как матрица K не пропорциональна $E_{1\,2}$, то матрица $K+M$ содержит ненулевой элемент, который находится на пересечении первой строки и столбца, отличного от второго.

Отсюда следует, что $\mathrm{rk}(K + M) \geqslant 2$. Так как $K + M \in W$, это противоречит условиям, которые накладываются на W. Значит, наше предположение неверно, и либо $\mathrm{Im}(E_{1\,2}) = \mathrm{Im}(M)$ для всех $0 \neq M \in W$, либо $\mathrm{Im}(E_{1\,2}^t) = \mathrm{Im}(M^t)$ для всех $0 \neq M \in W$. □

Следующие леммы дают нам информацию о том, как свойства пространств вида $M_0(v)$ и $M_0^t(v)$ связаны со свойствами задающих их векторов.

Лемма 2.23. *Если ненулевые векторы u и v коллинеарны, то*

$$\dim\left(M_0(u) \cap M_0(v)\right) = n - 1;$$

в противном случае,

$$\dim\left(M_0(u) \cap M_0(v)\right) = 0.$$

Доказательство. Если u и v коллинеарны, то $\langle u\rangle = \langle u\rangle$, откуда $M_0(u) = M_0(v)$, что приводит к требуемому равенству.

Если же u и v неколлинеарны, то $\langle u\rangle \neq \langle u\rangle$, а значит, всякая ненулевая матрица, принадлежащая $M_0(u)$, не принадлежит $M_0(v)$, откуда $M_0(u) \cap M_0(v) = \{0\}$. Последнее равенство означает, что

$$\dim\left(M_0(u) \cap M_0(v)\right) = 0. \qquad \square$$

Следствие 2.24. *Если $M_0(u) = M_0(v)$ для некоторых ненулевых $u, v \in \mathbb{F}^n$, то векторы u и v коллинеарны.*

Лемма 2.25. *Предположим, что ненулевая матрица $X \in \mathcal{M}_n$ и ненулевые векторы $u, v \in \mathbb{F}^n$ таковы, что $X \in M_0(u) \cap M_0^t(v)$. Тогда $(u, v) = 0$ и $X = \alpha J(v, u)$ для некоторого $\alpha \in \mathbb{F}$.*

Доказательство. Пусть $0 \neq X \in M_0(u) \cap M_0^t(v)$. Из леммы 2.19 следует, что $X = J(a_u, u)$ для некоторого $a_u \in \mathbb{F}^n$ такого, что $(a_u, u) = 0$. С другой стороны, из леммы 2.19, применённой к X^t, следует, что $X = J(v, a_v)$ для некоторого $a_v \in \mathbb{F}^n$ такого, что $(a_v, v) = 0$.

Далее, рассмотрим равенство

$$J(a_u, u) = X = J(v, a_v).$$

Из леммы 2.16 следует, что $a_u = \alpha v$ и $u = \frac{1}{\alpha} a_v$ для некоторого ненулевого $\alpha \in \mathbb{F}$, откуда

$$X = J(\alpha v, u) = \alpha J(v, u).$$

Также, так как $(a_v, v) = 0$, то

$$(u, v) = (\frac{1}{\alpha} a_v, v) = \frac{1}{\alpha}(a_v, v) = 0. \qquad \square$$

Следствие 2.26. *Если* $\dim(M_0(u) \cap M_0^t(v)) > 0$, *то* $(u, v) = 0$ *и* $\dim(M_0(u) \cap M_0^t(v)) = 1$.

Лемма 2.27. *Пусть* $0 \neq u, v \in \mathbb{F}^n$. *Тогда* $\dim(M_0(u) \cap M_0^t(v)) \leqslant 1$, *причём равенство достигается в том и только том случае, когда* $(u, v) = 0$.

Доказательство. Докажем неравенство $\dim(M_0(u) \cap M_0^t(v)) \leqslant 1$. Если $\dim(M_0(u) \cap M_0^t(v)) = 0$, то оно выполнено. Если

$$\dim\big(M_0(u) \cap M_0^t(v)\big) > 0,$$

то, согласно следствию 2.26, $\dim(M_0(u) \cap M_0^t(v)) = 1 \leqslant 1$.

Предположим теперь, что $\dim(M_0(u) \cap M_0^t(v)) = 1$. Тогда мы можем опять воспользоваться следствием 2.26 и получить, что $(u, v) = 0$.

Если же $(u, v) = 0$, то, $0 \neq J(v, u) \in M_0(u) \cap M_0^t(v)$. Это означает, что $\dim(M_0(u) \cap M_0^t(v)) > 0$, поэтому, в третий раз воспользовавшись следствием 2.26, мы заключаем, что $\dim(M_0(u) \cap M_0^t(v)) = 1$. $\square$

Следствие 2.28. *При $n \geqslant 3$ равенство $M_0(u) = M_0^t(v)$ для ненулевых векторов $u, v \in \mathbb{F}^n$ невозможно.*

Лемма 2.29. *Предположим, что для некоторого* $1 \leqslant k \leqslant n$ *ненулевые векторы* $v^{(1)}, \dots, v^{(k)} \in \mathbb{F}^n$ *линейно независимы. Тогда*

$$\dim\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\Big) = k(n-1). \tag{2.21}$$

Доказательство. Докажем это утверждение по индукции. При $k = 1$ утверждение очевидно. Предполагая теперь, что утверждение верно для некоторого k, докажем его для $k+1$. Для этого достаточно доказать, что

$$M_0(v^{(k+1)}) \cap \Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\Big) = \{0\}. \tag{2.22}$$

Пусть

$$A \in M_0(v^{(k+1)}) \cap \Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\Big).$$

Так как $A \in M_0(v^{(k+1)})$, то $\operatorname{Im}(A) \subset \langle v^{(k+1)} \rangle$.

С другой стороны, так как $A \in \big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\big)$, то существуют матрицы $A^{(1)}, \dots, A^{(k)}$ такие, что $A^{(i)} \in M_0(v^{(i)})$ для всех $1 \leqslant i \leqslant k$, и

$$A = A^{(1)} + \ldots + A^{(k)}.$$

Отсюда следует, что

$$\begin{aligned}\operatorname{Im}(A) = \operatorname{Im}(A^{(1)} + \ldots + A^{(k)}) \subset \operatorname{Im}(A^{(1)}) + \ldots + \operatorname{Im}(A^{(k)}) \\ \subset \langle v^{(1)} \rangle + \ldots \langle v^{(k)} \rangle = \langle v^{(1)}, \dots, v^{(k)} \rangle.\end{aligned}$$

Поэтому

$$\operatorname{Im}(A) \subset \langle v^{(k+1)} \rangle \cap \langle v^{(1)}, \dots, v^{(k)} \rangle.$$

Но $\langle v^{(k+1)} \rangle \cap \langle v^{(1)}, \dots, v^{(k)} \rangle = \{0\}$, так как $v^{(1)}, \dots, v^{(k+1)}$ линейно независимы. Поэтому $\operatorname{Im}(A) = \{0\}$ и $A = 0$, откуда следует равенство (2.22). □

Лемма 2.30. *Предположим, что для некоторого* $1 \leqslant k \leqslant n$ *ненулевые векторы* $v^{(1)}, \dots, v^{(k)} \in \mathbb{F}^n$ *линейно зависимы. Тогда*

$$\dim\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\Big) < k(n-1). \tag{2.23}$$

Доказательство. Сначала заметим, что, так как каждый из векторов $v^{(1)}, \dots, v^{(k)}$ отличен от нуля, то из условия леммы следует, что $k \geqslant 2$. Поэтому, не нарушая общности, мы можем предположить, что

$v^{(k)} = \alpha_1 v^{(1)} + \ldots + \alpha_{k-1} v^{(k-1)}$ для некоторых $\alpha_1, \ldots, \alpha_{k-1} \in \mathbb{F}$. Докажем сначала, что

$$\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k-1)})\Big) \cap M_0(v^{(k)}) \neq \{0\}. \tag{2.24}$$

Так как $\dim(\langle v^{(1)}, \ldots, v^{(k-1)} \rangle) < n$, а стандартное скалярное произведение на $\mathbb{F}^n$ является невырожденной билинейной формой, то существует ненулевой вектор $w \in \mathbb{F}^n$ такой, что $(v^{(i)}, w) = 0$ для всех $1 \leqslant i \leqslant k-1$. Так как $v^{(k)} \in \langle v^{(1)}, \ldots, v^{(k-1)} \rangle$, то также $(v^{(k)}, w) = 0$. Заметим теперь, что $J(w, v^{(i)}) \in M_0(v^{(i)})$ для $1 \leqslant i \leqslant k$ и

$$J(w, v^{(k)}) = \alpha_1 J(w, v^{(1)}) + \ldots + \alpha_{k-1} J(w, v^{(k-1)}).$$

При этом $J(w, v^{(k)}) \neq 0$, так как векторы $v^{(k)}$ и w не равны нулю. Таким образом, условие (2.24) выполнено, и потому

$$\dim\Big(\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k-1)})\Big) \cap M_0(v^{(k)})\Big) \geqslant 1.$$

Теперь докажем утверждение леммы. Согласно следствию 2.21,

$$\dim\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k-1)})\Big) \leqslant (k-1)(n-1).$$

Поэтому

$$\begin{aligned}
&\dim\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\Big) \\
&\quad = \dim\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k-1)})\Big) + \dim\Big(M_0(v^{(k)})\Big) \\
&\quad\quad - \dim\Big(\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k-1)})\Big) \cap M_0(v^{(k)})\Big) \\
&\quad\quad\quad \leqslant (k-1)(n-1) + (n-1) - 1 = k(n-1) - 1 < k(n-1).
\end{aligned}$$

Следовательно, неравенство (2.23) верно. □

Следствие 2.31. *Ненулевые векторы $v^{(1)}, \ldots, v^{(k)} \in \mathbb{F}^n$ линейно независимы тогда и только тогда, когда*

$$\dim\Big(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\Big) = k(n-1).$$

Покажем теперь, как действуют на множество пространств вида $M_0(v)$ и $M_0^t(v)$ биективные линейные отображения, сохраняющие множество нильпотентных матриц ранга 1.

Лемма 2.32. *Предположим, что биективное линейное отображение T сохраняет множество нильпотентных матриц ранга 1. Тогда для всякого $0 \neq v \in \mathbb{F}^n$ найдётся вектор $0 \neq w \in \mathbb{F}^n$ такой, что $T(M_0(v)) = T(M_0(w))$ или $T(M_0(v)) = T(M_0^t(w))$.*

Доказательство. По лемме 2.20, $\dim(M_0(v)) = n - 1$. Так как T – биективное линейное отображение, то также $\dim(T(M_0(v))) = n - 1$.

Из условия леммы следует, что всякий ненулевой элемент $T(M(v))$ является нильпотентной матрицей ранга 1. Значит, $T(M_0(v))$ удовлетворяет условиям леммы 2.22, и поэтому найдётся ненулевой вектор $w \in \mathbb{F}^n$ такой, что $T(M_0(v)) \subset M_0(w)$ или $T(M_0(v)) \subset M_0^t(w)$. Согласно замечанию в начале доказательства, размерности всех этих трёх пространств равны $n - 1$, и поэтому эти включения являются равенствами, то есть или $T(M_0(v)) = M_0(w)$, или $T(M_0(v)) = M_0^t(w)$ для этого w. □

Лемма 2.33. *Предположим, что $n \geqslant 3$ и биективное линейное отображение T сохраняет множество нильпотентных матриц ранга 1. Тогда выполнено одно из двух утверждений:*

(1) *для всякого $0 \neq v \in \mathbb{F}^n$ найдётся $0 \neq w \in \mathbb{F}^n$ такой, что*

$$T(M_0(v)) = M_0(w); \tag{2.25}$$

(2) *для всякого $0 \neq v \in \mathbb{F}^n$ найдётся $0 \neq w \in \mathbb{F}^n$ такой, что*

$$T(M_0(v)) = M_0^t(w). \tag{2.26}$$

Доказательство. Допустим противное. Это означает, что найдётся ненулевой вектор $u \in \mathbb{F}^n$ такой, что $T(M_0(u)) \neq T(M_0(u'))$ для всех $0 \neq u' \in \mathbb{F}^n$, и ненулевой вектор $u \in \mathbb{F}^n$ такой, что $T(M_0(v)) \neq T(M_0^t(v'))$ для всех $0 \neq u' \in \mathbb{F}^n$.

Из леммы 2.32 следует, что тогда найдутся ненулевые векторы $u'', v'' \in \mathbb{F}^n$ такие, что $T(M_0(u)) = M_0^t(u'')$ и $T(M_0(v)) = M_0(v'')$. Так как $n \geqslant 3$, существует ненулевой вектор $w \in \mathbb{F}^n$ такой, что

$$(u, w) = (v, w) = 0.$$

В таком случае, согласно лемме 2.27,

$$\dim\left(M_0(u) \cap M_0^t(w)\right) = \dim\left(M_0(v) \cap M_0^t(w)\right) = 1.$$

Поэтому, так как T является биективным линейным отображением, то

$$\dim\left(T(M_0(u)) \cap T(M_0^t(w))\right) = \dim\left(T(M_0(v)) \cap T(M_0^t(w))\right) = 1. \tag{2.27}$$

Пользуясь леммой 2.32, предположим, что $T(M_0^t(w)) = M_0(w'')$ для некоторого $w'' \in \mathbb{F}^n$. Тогда, по лемме 2.23,

$$\dim\left(T(M_0(v)) \cap T(M_0^t(w))\right) = \dim\left(M_0(v'') \cap M_0(w'')\right) = 0 \text{ или } n-1,$$

что противоречит равенству (2.27), так как $n \geqslant 3$.

Аналогично, если $T(M_0^t(w)) = M_0^t(w'')$ для некоторого $w' \in \mathbb{F}^n$, то, по лемме 2.23,

$$\dim\left(T(M_0(u)) \cap T(M_0^t(w))\right) = \dim\left(M_0^t(u'') \cap M_0^t(w')\right) = 0 \text{ или } n-1,$$

что также противоречит равенству (2.27), так как $n \geqslant 3$.

В любом случае, мы получили противоречие. Значит, наше допущение неверно, и выполнено одно из двух утверждений леммы. □

Замечание 2.34. Согласно следствиям 2.24 и 2.28, вектор w в равенствах (2.25) и (2.26) определён с точностью до пропорциональности.

Лемма 2.35. *Предположим, что биективное линейное отображение T пространства матриц $\mathcal{M}_n$ переводит пространства вида $M_0(v)$ в пространства такого же вида, $1 \leqslant k \leqslant n$ и ненулевые векторы $u^{(1)}, \ldots, u^{(k)}, v^{(1)}, \ldots, v^{(k)} \in \mathbb{F}^n$ таковы, что*

$$T(M_0(u^{(i)})) = M_0(v^{(i)})$$

для всех $1 \leqslant i \leqslant k$. Тогда векторы $u^{(1)}, \ldots, u^{(k)}$ линейно независимы в том и только том случае, когда векторы $v^{(1)}, \ldots, v^{(k)}$ линейно независимы.

Доказательство. Предположим, что векторы $u^{(1)}, \ldots, u^{(k)}$ линейно независимы. Тогда из леммы 2.29 следует, что

$$\dim\left(M_0(u^{(1)}) + \ldots + M_0(u^{(k)})\right) = k(n-1).$$

Так как линейное отображение T биективно и $T(M_0(u^{(i)})) = M(u^{(i)})$ для всех $1 \leqslant i \leqslant k$, то

$$\begin{aligned}\dim\left(M_0(v^{(1)}) + \ldots + M_0(v^{(k)})\right)& \\ = \dim\left(T\left(M_0(u^{(1)}) + \ldots + M_0(u^{(k)})\right)\right)& \\ = \dim\left(M_0(u^{(1)}) + \ldots + M_0(u^{(k)})\right) &= k(n-1).\end{aligned}$$

Из следствия 2.31 следует, что в таком случае векторы $v^{(1)}, \ldots, v^{(k)}$ линейно независимы.

Аналогичные рассуждения доказывают утверждение леммы в обратную сторону. $\square$

Следствие 2.36. *Предположим, что биективное линейное отображение T пространства матриц $\mathcal{M}_n$ удовлетворяет условиям леммы 2.35 и ненулевые векторы $v^{(1)}, \ldots, v^{(k+1)}, u^{(1)}, \ldots, u^{(k+1)}$, где $1 \leqslant k < n$, таковы, что*

$$T(M_0(u^{(i)})) = M(u^{(i)})$$

для всех $1 \leqslant i \leqslant k+1$ и

$$u^{(k+1)} \in \langle u^{(1)}, \ldots, u^{(k)} \rangle.$$

Тогда

$$v^{(k+1)} \in \langle v^{(1)}, \ldots, v^{(k)} \rangle.$$

3. Доказательство основной теоремы

Доказательство основной теоремы будет проводиться отдельно для $n \geqslant 3$ и $n = 2$. В обоих случаях, мы будем поэтапно приводить исходное линейное отображение к более простому виду. Этапы этого процесса сформулированы в виде лемм, приведённых ниже.

Лемма 3.1. *Предположим, что линейное отображение T пространства матриц $\mathcal{M}_n$ сохраняет матричный дискриминант. Тогда существует линейная функция $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ такая, что линейное отображение $S\colon \mathcal{M}_n \to \mathcal{M}_n$, определяемое равенством*

$$S(X) = T(X) + \alpha(X) I_n \tag{3.1}$$

для всех $X \in \mathcal{M}_n$, сохраняет след матрицы и матричный дискриминант.

Доказательство. Действительно, так как $\mathrm{char}(\mathbb{F})$ не делит n, то n обратимо в поле $\mathbb{F}$. Поэтому мы можем определить $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ равенством

$$\alpha(X) = \frac{1}{n}\left(\mathrm{tr}(X) - \mathrm{tr}(T(X))\right)$$

для всех $X \in \mathcal{M}_n$.

Определим далее $S\colon \mathcal{M}_n \to \mathcal{M}_n$ равенством (3.1). Отображение S сохраняет матричный дискриминант, так как разность $S(X) - T(X)$

является скалярной матрицей для всех $X \in \mathcal{M}_n$. Также

$$\begin{aligned}\operatorname{tr}(S(X)) = \operatorname{tr}\left(T(X) + \frac{1}{n}\left(\operatorname{tr}(X) - \operatorname{tr}(T(X))\right) I_n\right)\\ = \operatorname{tr}(T(X)) + \operatorname{tr}(X) - \operatorname{tr}(T(X)) = \operatorname{tr}(X)\end{aligned}$$

для всех $X \in \mathcal{M}_n$, то есть S сохраняет след, а значит, удовлетворяет всем требуемым условиям. $\square$

Лемма 3.2. *Предположим, что $n \geqslant 3$ и биективное линейное отображение T пространства матриц сохраняет след матрицы, матричный дискриминант и переводит пространства вида $M_0(v)$ в пространства такого же вида. Тогда существуют обратимая матрица $P \in \mathcal{M}_n$ и $b \in \mathbb{F}$ такие, что $b^{n(n-1)} = 1$ и отображение S, определяемое равенством*

$$S(X) = P^{-1}T(X)P \tag{3.2}$$

для всех $X \in \mathcal{M}_n$, обладает следующими свойствами:

Л3.2(1) *$S(E_{i\,j})$ пропорционально $E_{i\,j}$ для всех $1 \leqslant i \neq j \leqslant n$;*

Л3.2(2) *$S(E_{i_i}) = bE_{i\,i} + \frac{1-b}{n} I_n$ для всех $1 \leqslant i \leqslant n$.*

Доказательство. Разделим наше рассуждение на пять частей:

I нахождение матрицы P и определение отображения S;

II доказательство свойства Л3.2(1) для выбранного отображения S;

III нахождение b;

IV доказательство свойства Л3.2(2) для выбранного отображения S;

V доказательство равенства $b^{n(n-1)} = 1$.

Часть I. Нахождение матрицы P и определение отображения S. Пусть $e^{(1)}, \dots, e^{(n)}$ – стандартный базис $\mathbb{F}^n$. По условию леммы, существуют ненулевые векторы $f^{(1)}, \dots, f^{(n)}$ такие, что

$$T(M_0(e^{(i)})) = M_0(f^{(i)})$$

для всех $1 \leqslant i \leqslant n$. Из леммы 2.35 следует, что $f^{(1)}, \dots, f^{(n)}$ являются линейно независимыми, и поэтому они образуют базис пространства $\mathbb{F}^n$. Значит, существует матрица P такая, что $P^{-1}(f^{(i)}) = e^{(i)}$ для всех $1 \leqslant i \leqslant n$.

Определим теперь линейное отображение S пространства матриц равенством (3.2). Сразу отметим, что S биективно, сохраняет след матрицы и матричный дискриминант, потому что преобразование подобия биективно и сохраняет характеристический многочлен матрицы.

Заметим также, что
$$P^{-1}(M_0(f^{(i)}))P = M_0(e^{(i)})$$
для всех $1 \leqslant i \leqslant n$, откуда
$$S(M_0(e^{(i)})) = M_0(e^{(i)})$$
для всех $1 \leqslant i \leqslant n$.

Часть II. Доказательство свойства Л3.2(1) для выбранного отображения S**.** Докажем сначала, что
$$S(M_0^t(e^{(i)})) = M_0^t(e^{(i)}) \tag{3.3}$$
для всех $1 \leqslant i \leqslant n$. Действительно, из леммы 2.33 следует, что существуют $g^{(1)}, \ldots, g^{(n)}$ такие, что
$$S(M_0^t(e^{(i)})) = M_0^t(g^{(i)})$$
для всех $1 \leqslant i \leqslant n$. Пусть $1 \leqslant i \neq j \leqslant n$. Тогда, так как $(e^{(i)}, e^{(j)}) = 0$, то из леммы 2.27 следует, что
$$\dim\Big(M_0(e^{(i)}) \cap M_0^t(e^{(j)})\Big) = 1,$$
откуда
$$\dim\Big(S\Big(M_0(e^{(i)}) \cap M_0^t(e^{(j)})\Big)\Big) = 1,$$
и поэтому
$$\dim\Big(M_0(e^{(i)}) \cap M_0^t(g^{(j)})\Big) = 1.$$
Значит, по лемме 2.27, $(e^{(i)}, g^{(j)}) = 0$. Из этого следует, что, при фиксированном j, все координаты $g^{(j)}$, кроме j-ой, равны нулю. Таким образом, вектор $g^{(j)}$ коллинеарен $e^{(j)}$ для всех $1 \leqslant j \leqslant n$, откуда мы получаем, что равенство (3.3) выполнено при всех $1 \leqslant i \leqslant j$.

Далее, так как при $1 \leqslant i \neq j \leqslant n$ выполнено равенство
$$M_0(e^{(i)}) \cap M_0^t(e^{(j)}) = \langle E_{i\,j}\rangle,$$
то
$$\begin{aligned} S(E_{i\,j}) \in S(\langle E_{i\,j}\rangle) &= S\Big(M_0(e^{(i)}) \cap M_0^t(e^{(j)})\Big) \\ &= M_0(e^{(i)}) \cap M_0^t(e^{(j)}) = \langle E_{i\,j}\rangle, \end{aligned}$$

и поэтому $S(E_{i\,j})$ пропорционально $E_{i\,j}$.

Часть III. Нахождение b. Докажем, что существует b такое, что

$$S(E_{i\,i} - E_{j\,j}) = b(E_{i\,i} - E_{j\,j})$$

для всех $1 \leqslant i \neq j \leqslant n$.

Возьмём $1 \leqslant i \neq j \leqslant n$ и рассмотрим матрицу

$$B^{(i\,j)} = J(e^{(i)} - e^{(j)}, e^{(i)} + e^{(j)}) = E_{i\,i} - E_{i\,j} + E_{j\,i} - E_{j\,j}.$$

Из леммы 2.19 следует, что

$$B^{(i\,j)} \in M_0(e^{(i)} + e^{(j)}) \cap M_0^t(e^{(i)} - e^{(j)}).$$

Заметим, что векторы $e^{(i)} - e^{(j)}$ и $e^{(i)} + e^{(j)}$ принадлежат $\langle e^{(i)}, e^{(j)} \rangle$. Тогда, так как S является биективным линейным отображением, переводящим пространства вида $M_0(v)$ в пространства такого же вида, из леммы 2.32 получаем, что

$$S(B^{(i\,j)}) \in S\Big(M_0(e^{(i)} + e^{(j)}) \cap M_0^t(e^{(i)} - e^{(j)})\Big) = M_0(u^{(i\,j)}) \cap M_0^t(v^{(i\,j)})$$

для некоторых ненулевых векторов $u^{(i\,j)}, v^{(i\,j)} \in \mathbb{F}$. Так как $S(M_0(e^k)) = M_0(e^k)$ и $S(M_0^t(e^k)) = M_0^t(e^k)$ для всех $1 \leqslant k \leqslant n$, то, согласно следствию 2.36, векторы $u^{(i\,j)}$ и $v^{(i\,j)}$ принадлежат $\langle e^{(i)}, e^{(j)} \rangle$. Поэтому $S(B^{(i\,j)})$ является линейной комбинацией матриц $E_{i\,i}, E_{i\,j}, E_{j\,i}$ и $E_{j\,j}$. Так как это верно также для $S(E_{i\,j})$ и $S(E_{j\,i})$ и

$$E_{i\,i} - E_{j\,j} = B^{(i\,j)} + E_{i\,j} - E_{j\,i},$$

то $E_{i\,i} - E_{j\,j}$ также является линейной комбинацией матриц $E_{i\,i}$, $E_{i\,j}$, $E_{j\,i}$ и $E_{j\,j}$, то есть

$$S(E_{i\,i} - E_{j\,j}) = b_{i\,i}^{(i\,j)} E_{i\,i} + b_{i\,j}^{(i\,j)} E_{i\,j} + b_{j\,i}^{(i\,j)} E_{j\,i} + b_{j\,j}^{(i\,j)} E_{j\,j}$$

для некоторых $b_{i\,i}^{(i\,j)}, b_{i\,j}^{(i\,j)}, b_{j\,i}^{(i\,j)}, b_{j\,j}^{(i\,j)} \in \mathbb{F}$.

Возьмём теперь k, отличное от i и j (это возможно, так как мы предположили, что $n \geqslant 3$). Тогда, с одной стороны,

$$\begin{aligned}
S(E_{i\,i}-E_{k\,k})&=S((E_{i\,i}-E_{j\,j})+(E_{j\,j}-E_{k\,k}))=S(E_{i\,i}-E_{j\,j})+S(E_{j\,j}-E_{k\,k})\\
&= b^{(i\,j)}_{i\,i}E_{i\,i}+b^{(i\,j)}_{i\,j}E_{i\,j}+b^{(i\,j)}_{j\,i}E_{j\,i}+b^{(i\,j)}_{j\,j}E_{j\,j}\\
&\quad+b^{(j\,k)}_{j\,j}E_{j\,j}+b^{(j\,k)}_{j\,k}E_{j\,k}+b^{(j\,k)}_{k\,j}E_{k\,j}+b^{(j\,k)}_{k\,k}E_{k\,k}\\
&= b^{(i\,j)}_{i\,i}E_{i\,i}+b^{(i\,j)}_{i\,j}E_{i\,j}+b^{(i\,j)}_{j\,i}E_{j\,i}+(b^{(i\,j)}_{j\,j}+b^{(j\,k)}_{j\,j})E_{j\,j}\\
&\quad+b^{(j\,k)}_{j\,k}E_{j\,k}+b^{(j\,k)}_{k\,j}E_{k\,j}+b^{(j\,k)}_{k\,k}E_{k\,k}.
\end{aligned}\tag{3.4}$$

С другой стороны,

$$S(E_{i\,i}-E_{k\,k})=b^{(i\,k)}_{i\,i}E_{i\,i}+b^{(i\,k)}_{i\,k}E_{i\,k}+b^{(i\,k)}_{k\,i}E_{k\,i}+b^{(i\,k)}_{k\,k}E_{k\,k}.\tag{3.5}$$

Сравнивая коэффициенты при $E_{i\,j}$ и $E_{j\,i}$ в этих двух равенствах, получаем, что $b^{(i\,j)}_{i\,j}=b^{(i\,j)}_{j\,i}=0$.

Таким образом,

$$S(E_{i\,i}-E_{j\,j})=b^{(i\,j)}_{i\,i}E_{i\,i}+b^{(i\,j)}_{j\,j}E_{j\,j}$$

для всех $1\leqslant i\neq j\leqslant n$. Из того, что S сохраняет след, следует, что

$$b^{(i\,j)}_{i\,i}+b^{(i\,j)}_{j\,j}=\operatorname{tr}(S(E_{i\,i}-E_{j\,j}))=\operatorname{tr}(E_{i\,i}-E_{j\,j})=0,$$

откуда $b^{(i\,j)}_{i\,i}=-b^{(i\,j)}_{j\,j}$. То есть, если положить $b^{(i\,j)}=b^{(i\,j)}_{i\,i}$, то

$$S(E_{i\,i}-E_{j\,j})=b^{(i\,j)}\,(E_{i\,i}-E_{j\,j})$$

для всех $1\leqslant i\neq j\leqslant n$. Сравнивая эти равенства для $E_{i\,i}-E_{j\,j}$ и $E_{j\,j}-E_{i\,i}$, получаем, что $b^{(j\,i)}=b^{(i\,j)}$ для всех $1\leqslant i\neq j\leqslant n$.

Снова рассмотрим равенства (3.4) и (3.5). Они теперь выглядят следующим образом:

$$\begin{aligned}
S(E_{i\,i}-E_{k\,k})&=S((E_{i\,i}-E_{j\,j})+(E_{j\,j}-E_{k\,k}))\\
&=S(E_{i\,i}-E_{j\,j})+S(E_{j\,j}-E_{k\,k})\\
&=b^{(i\,j)}\,(E_{i\,i}-E_{j\,j})+b^{(j\,k)}\,(E_{j\,j}-E_{k\,k})
\end{aligned}\tag{3.6}$$

и

$$S(E_{i\,i}-E_{k\,k})=b^{(i\,k)}\,(E_{i\,i}-E_{k\,k})\,.\tag{3.7}$$

Сравнивая коэффициенты при $E_{i\,i}$, получаем, что $b^{(i\,j)}=b^{(i\,k)}$ для всех попарно различных $1\leqslant i,j,k\leqslant n$. Так как $n\geqslant 3$, отсюда следует, что существует $b \in \mathbb{F}$ такое, что $b^{(i\,j)} = b$ для всех $1 \leqslant i, j \leqslant n$. Таким образом, для этого b при всех $1 \leqslant i \neq j \leqslant n$ выполнено равенство

$$S(E_{i\,i} - E_{j\,j}) = b(E_{i\,i} - E_{j\,j}).$$

Часть IV. Доказательство свойства Л3.2(2) для выбранного отображения S. Докажем, что

$$S(E_{i\,i}) = bE_{i\,i} + \frac{1-b}{n} I_n$$

для всех $1 \leqslant i \leqslant n$.

Для этого заметим, что, так как n обратимо в $\mathbb{F}$,

$$E_{i\,i} = \frac{1}{n} \sum_{j \neq i} (E_{i\,i} - E_{j\,j}) + \frac{1}{n} I_n.$$

Согласно следствию 2.7, $S(I_n) = I_n$. Тогда

$$\begin{aligned} S(E_{i\,i}) &= S\left(\frac{1}{n} \sum_{j \neq i} (E_{i\,i} - E_{j\,j}) + \frac{1}{n} I_n\right) \\ &= \frac{1}{n} \sum_{j \neq i} S(E_{i\,i} - E_{j\,j}) + \frac{1}{n} S(I_n) = \frac{1}{n} \sum_{j \neq i} b(E_{i\,i} - E_{j\,j}) + \frac{1}{n} I_n \\ &= \frac{1}{n} \sum_{j \neq i} b(E_{i\,i} - E_{j\,j}) + \frac{1}{n} I_n + \frac{1}{n} bI_n - \frac{1}{n} bI_n \\ &= b\left(\frac{1}{n} \sum_{j \neq i} (E_{i\,i} - E_{j\,j}) + \frac{1}{n} I_n\right) + \frac{1}{n} I_n - \frac{1}{n} bI_n \\ &= bE_{i\,i} + \frac{1-b}{n} I_n. \end{aligned}$$

Часть V. Доказательство равенства $b^{n(n-1)} = 1$. Рассмотрим $A = \operatorname{diag}(a_1, \ldots, a_n)$, где $a_1, \ldots, a_n$ – попарно различные элементы $\mathbb{F}$. Тогда, с одной стороны,

$$\mathrm{D}(A) = \prod_{1 \leqslant i < j \leqslant n} (a_i - a_j)^2 \neq 0.$$

С другой стороны, так как

$$S(A) = b \operatorname{diag}(a_1, \ldots, a_n) + \frac{1-b}{n} (a_1 + \ldots + a_n) I_n$$

и $\mathrm{D}(S(A)) = \mathrm{D}(A)$, то

$$\prod_{1\leqslant i<j\leqslant n}(a_i - a_j)^2 = \prod_{1\leqslant i<j\leqslant n}(ba_i - ba_j)^2 = b^{n(n-1)}\prod_{1\leqslant i<j\leqslant n}(a_i - a_j)^2,$$

откуда $b^{n(n-1)} = 1$, потому что $\prod_{1\leqslant i<j\leqslant n}(a_i - a_j)^2 \neq 0$. □

Лемма 3.3. *Предположим, что $n \geqslant 3$ и биективное линейное отображение T пространства матриц $\mathcal{M}_n$ сохраняет след матрицы и матричный дискриминант, а также обладает следующими свойствами:*

(1) *$T(E_{i\,j})$ пропорционально $E_{i\,j}$ для всех $1 \leqslant i \neq j \leqslant n$;*
(2) *существует $b \in \mathbb{F}$ такое, что $b^{n(n-1)} = 1$ и $T(E_{i\,i}) = bE_{i\,i} + \frac{1-b}{n}I_n$ для всех $1 \leqslant i \leqslant n$.*

Тогда существует обратимая матрица $R \in \mathcal{M}_n$ такая, что

$$T(X) = bR^{-1}XR + \frac{1-b}{n}\operatorname{tr}(X)I_n$$

для всех $X \in \mathcal{M}_n$.

Доказательство. Определим линейное отображение S пространства матриц равенством

$$S(X) = \frac{1}{b}T(X) - \frac{1-b}{bn}\operatorname{tr}(X)I_n \tag{3.8}$$

для всех $X \in \mathcal{M}_n$. Отображение S сохраняет матричный дискриминант, так как получено из T композицией с линейными отображениями, сохраняющими матричный дискриминант.

Заметим также, что $S(E_{i\,j})$ пропорционально $E_{i\,j}$ для всех $1 \leqslant i \neq j \leqslant n$, так как этим свойством обладает отображение T, а линейные отображения, композицией с которыми получено отображение S, это свойство сохраняют.

Далее,

$$\begin{aligned} S(E_{i\,i}) &= \frac{1}{b}T(E_{i\,i}) - \frac{1-b}{bn}\operatorname{tr}(E_{i\,i})I_n \\ &= \frac{1}{b}(bE_{i\,i} + \frac{1-b}{n}I_n) - \frac{1-b}{bn}\operatorname{tr}(E_{i\,i})I_n \\ &= E_{i\,i} + \frac{1-b}{bn}I_n - \frac{1-b}{bn}I_n = E_{i\,i} \end{aligned}$$

для всех $1 \leqslant i \leqslant n$. Поэтому S является тождественным отображением на множестве диагональных матриц, а значит, сохраняет матричный след.

Рассмотрим теперь матрицу $C \in \mathcal{M}_n$, элементы которой являются коэффициентами пропорциональности при действии S на $E_{i\,j}$, то есть $C_{i\,j}$ определяется из равенства

$$S(E_{i\,j}) = C_{i\,j} E_{i\,j}$$

для всех $1 \leqslant i, j \leqslant n$. Докажем, что эта матрица имеет единичный ранг.

Её ранг не равен нулю, потому что её диагональные элементы равны единице. Докажем теперь, что любые два её столбца линейно зависимы, то есть что столбцы j_1 и j_2 матрицы C линейно зависимы для всех $1 \leqslant j_1 < j_2 \leqslant n$. Рассмотрим матрицу

$$D^{(j_1\,j_2)} = J(e^{(j_1)} - e^{(j_2)}, (1, \ldots, 1)) = \sum_{1 \leqslant i \leqslant n} (E_{i\,j_1} - E_{i\,j_2}).$$

Так как $(e^{(j_1)} - e^{(j_2)}, (1, \ldots, 1)) = 0$, она является нильпотентной матрицей ранга 1. Из леммы 2.11 следует, что $S(D^{(j_1\,j_2)})$ также является нильпотентной матрицей ранга 1. Значит, все её столбцы линейно зависимы, в том числе её столбцы j_1 и j_2.

При этом, так как $\big(S(D^{(j_1\,j_2)})\big)_{i\,j_1} = C_{i\,j_1}$ и $\big(S(D^{(j_1\,j_2)})\big)_{i\,j_2} = C_{i\,j_2}$ для всех $1 \leqslant i \leqslant n$, то столбцы j_1 и j_2 матрицы $D^{(j_1\,j_2)}$ равны столбцам j_1 и j_2 матрицы C соответственно. Отсюда следует, что столбцы j_1 и j_2 матрицы C также линейно зависимы.

Значит, C имеет ранг 1. Пусть $q \in \mathbb{F}^n$ –вектор, задающий один из её ненулевых столбцов. Тогда $\operatorname{Im}(C) = \langle q \rangle$, и поэтому из леммы 2.17 следует, что найдётся $0 \neq r \in \mathbb{F}^n$ такой, что $C = J(r, q)$, откуда

$$S(E_{i\,j}) = q_i r_j E_{i\,j}$$

для всех $1 \leqslant i, j \leqslant n$. Если мы теперь положим $Q = \operatorname{diag}(q_1, \ldots, q_n)$, $R = \operatorname{diag}(r_1, \ldots, r_n)$, то получим, что равенства

$$S(E_{i\,j}) = Q E_{i\,j} R \tag{3.9}$$

выполнены для всех $1 \leqslant i, j \leqslant n$. Так как матричные единицы образуют базис пространства $\mathcal{M}_n$, а отображение S и двухстороннее матричное умножение являются линейными отображениями, то равенство (3.9) выполнено для всех $X \in \mathcal{M}_n$.

Наконец, так как S сохраняет след матрицы, то $\operatorname{tr}(QXR) = \operatorname{tr}(X)$ для всех $X \in \mathcal{M}_n$. Из леммы 2.9 следует, что тогда $Q = R^{-1}$. Возвращаясь к равенству (3.9), получаем, что

$$R^{-1}XR = QXR = S(X) = \frac{1}{b}T(X) - \frac{1-b}{bn}\operatorname{tr}(X)I_n$$

для всех $X \in \mathcal{M}_n$, откуда

$$T(X) = bR^{-1}XR + \frac{1-b}{n}\operatorname{tr}(X)I_n$$

для всех $X \in \mathcal{M}_n$. □

Следствие 3.4. *Предположим, что $n \geqslant 3$ и биективное линейное отображение T пространства матриц сохраняет след матрицы и матричный дискриминант и переводит пространства вида $M_0(v)$ в пространства такого же вида. Тогда существуют обратимая матрица $Q \in \mathcal{M}_n$, линейное отображение $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ и $c \in \mathbb{F}$ такие, что $c^{n(n-1)} = 1$ и*

$$T(X) = cQ^{-1}XQ + \alpha(X)I_n$$

для всех $X \in \mathcal{M}_n$.

Доказательство. Действительно, так как такое отображение T удовлетворяет всем условиям леммы 3.2, то существует обратимая матрица $P \in \mathcal{M}_n$ и $b \in \mathbb{F}$ такие, что $b^{n(n-1)} = 1$, и отображение S пространства матриц $\mathcal{M}_n$, определяемое равенством

$$S(X) = P^{-1}(T(X))P \tag{3.10}$$

для всех $X \in \mathcal{M}_n$, обладает свойствами

(1) $S(E_{i\,j})$ пропорционально $E_{i\,j}$ для всех $1 \leqslant i \neq j \leqslant n$;
(2) $S(E_{i_i}) = bE_{i\,i} + \frac{1-b}{n}I_n$ для всех $1 \leqslant i \leqslant n$.

Это означает, что S удовлетворяет условиям леммы 3.3. Применяя её к S, мы получаем, что существуют обратимая матрица $R \in \mathcal{M}_n$ такая, что

$$S(X) = bR^{-1}XR + \frac{1-b}{n}\operatorname{tr}(X)I_n \tag{3.11}$$

для всех $X \in \mathcal{M}_n$. Из равенств (3.10) и (3.11) следует, что

$$P^{-1}(T(X))P = bR^{-1}XR + \frac{1-b}{n}\operatorname{tr}(X)I_n$$

для всех $X \in \mathcal{M}_n$, откуда

$$T(X) = b(RP^{-1})^{-1}X(RP^{-1}) + \frac{1-b}{n}\,\mathrm{tr}(X)I_n,$$

поэтому в формулировке леммы мы можем положить $c = b, Q = RP^{-1}$ и $\alpha(X) = \frac{1-b}{n}\,\mathrm{tr}(X)$ для всех $X \in \mathcal{M}_n$. □

Теперь докажем лемму о приведении к более простому виду отображения, сохраняющего след матрицы и матричный дискриминант, в случае $n = 2$.

Лемма 3.5. *Предположим, что $n = 2$ и биективное линейное отображение T пространства матриц сохраняет след матрицы и матричный дискриминант. Тогда существует такая обратимая матрица $P \in \mathcal{M}_2$, что для линейного отображения S, определяемого равенством*

$$S(X) = P^{-1}T(X)P$$

для всех $X \in \mathcal{M}_2$, выполнены равенства $S(E_{1\,2}) = E_{1\,2}$ и $S(E_{2\,1}) = E_{2\,1}$.

Доказательство. Согласно лемме 2.11, отображение T переводит нильпотентные матрицы ранга 1 в нильпотентные матрицы ранга 1, поэтому $T(E_{1\,2})$ является нильпотентной матрицей ранга 1. Все нильпотентные матрицы ранга 1 подобны друг другу, поэтому существует обратимая матрица $P_1 \in \mathcal{M}_2$ такая, что $P_1^{-1}T(E_{1\,2})P_1 = E_{1\,2}$.

Рассмотрим линейное отображение T_1 пространства матриц $\mathcal{M}_2$, определённое равенством

$$T_1(X) = P_1^{-1}T(X)P_1$$

для всех $X \in \mathcal{M}_2$. Это преобразование обратимо, сохраняет след матрицы и матричный дискриминант, а также переводит $E_{1\,2}$ в $E_{1\,2}$.

Пусть теперь $N = T_1(E_{2\,1})$. Согласно лемме 2.11, N является нильпотентной матрицей ранга 1.

Далее, из формулы для дискриминанта 2×2 матрицы (§1, формула (1.2) после свойств матричного дискриминанта) следует, что

$$\mathrm{D}(E_{2\,1} + \lambda E_{1\,2}) = 4\lambda, \tag{3.12}$$

поэтому

$$\mathrm{D}(N + \lambda E_{1\,2}) = \mathrm{D}(T_1(E_{2\,1} + \lambda E_{1\,2})) = \mathrm{D}(E_{2\,1} + \lambda E_{1\,2}) = 4\lambda$$

для всех $\lambda \in \mathbb{F}$.

С другой стороны, применяя формулу (1.2) к $\mathrm{D}(N+\lambda E_{1\,2})$, мы получаем, что

$$\mathrm{D}(N+\lambda E_{1\,2}) = N_{1\,1}^2 + 4(N_{1\,2}+\lambda)N_{2\,1} - 2N_{1\,1}N_{2\,2} + N_{2\,2}$$

для всех $\lambda \in \mathbb{F}$, откуда, с учётом равенства (3.12), заключаем, что

$$N_{1\,1}^2 + 4(N_{1\,2}+\lambda)N_{2\,1} - 2N_{1\,1}N_{2\,2} + N_{2\,2} = 4\lambda$$

для всех $\lambda \in \mathbb{F}$. Из этого следует, что $N_{2\,1} = 1$, так как $\operatorname{char}(\mathbb{F})$ не делит $n = 2$. Таким образом, $N = J(a, b)$, где $(a, b) = 0$, причём, так как $N_{2\,1} = 1$, мы можем предположить, что $a_1 = b_2 = 1$. Тогда $b_1 = -a_2 = a$ для некоторого $a \in \mathbb{F}$ и значит,

$$N = \begin{pmatrix} a & -a^2 \\ 1 & -a \end{pmatrix}$$

для этого a.

Положим теперь

$$P_2 = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

Очевидно, что эта матрица обратима. Непосредственные вычисления показывают, что

$$P_2^{-1}E_{1\,2}P_2 = E_{1\,2}$$

и

$$P_2^{-1}NP_2 = E_{2\,1}.$$

Отсюда мы получаем, что если определить линейное отображение S равенством

$$S(X) = (P_1P_2)^{-1}T(X)(P_1P_2)$$

для всех $X \in \mathcal{M}_n$, то $S(E_{1\,2}) = E_{1\,2}$ и $S(E_{2\,1}) = E_{2\,1}$. Поэтому матрица $P = P_1P_2$ удовлетворяет требуемым условиям. □

Теорема 3.6 (основная теорема). *Линейное отображение T пространства матриц $\mathcal{M}_n$ удовлетворяет условию $\mathrm{D}(T(X)) = \mathrm{D}(X)$ для всех $X \in \mathcal{M}_n$ тогда и только тогда, когда существуют обратимая матрица $M \in \mathcal{M}_n$, линейное отображение $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ и $c \in \mathbb{F}$ такие, что $c^{n(n-1)} = 1$ и либо*

$$T(X) = cM^{-1}XM + \alpha(X)I_n, \quad X \in \mathcal{M}_n,$$

либо

$$T(X) = cM^{-1}X^tM + \alpha(X)I_n, \quad X \in \mathcal{M}_n.$$

Доказательство. Достаточность следует из свойств MD3, MD4 и MD5 матричного дискриминанта. Докажем необходимость.

Пусть линейное отображение T пространства матриц $\mathcal{M}_n$ таково, что $\mathrm{D}(T(X)) = \mathrm{D}(X)$ для всех $X \in \mathcal{M}_n$. По лемме 3.1, существует линейное отображение $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ такое, что отображение T_1, определённое по формуле

$$T_1(X) = T(X) + \alpha(X)I_n$$

для всех $X \in \mathcal{M}_n$, сохраняет след матрицы и матричный дискриминант. Из леммы 2.8 следует, что T_1 является биекцией.

Рассмотрим теперь два случая: $n \geqslant 3$ и $n = 2$.

Случай I: $n \geqslant 3$**.** Согласно лемме 2.33, либо T_1 переводит пространства вида $M_0(v)$ в пространства такого же вида, либо T_1 переводит пространства вида $M_0(v)$ в пространства вида $M_0^t(v)$. Так как транспонирование матрицы не влияет на её характеристический многочлен, а значит, её след и дискриминант, не нарушая общности, предположим, что T_1 переводит пространства вида $M_0(v)$ в пространства такого же вида.

Далее, так как T_1 также является биективным линейным отображением, то оно удовлетворяет всем условиям следствия 3.4, воспользовавшись которым, мы получаем, что существуют обратимая матрица $P \in \mathcal{M}_n$, линейное отображение $\alpha'\colon \mathcal{M}_n \to \mathbb{F}$ и $c \in \mathbb{F}$ такие, что $c^{n(n-1)} = 1$ и

$$T_1(X) = cP^{-1}XP + \alpha'(X)I_n$$

для всех $X \in \mathcal{M}_n$.

Таким образом, выражая T через c, Q, α и α' из приведённых выше равенств, мы получаем, что

$$\begin{aligned} T(X) = T_1(X) - \alpha(X)I_n &= cQ^{-1}XQ + \alpha'(X)I_n - \alpha(X)I_n \\ &= cQ^{-1}XQ + (\alpha'(X) - \alpha(X))I_n \end{aligned}$$

для всех $X \in \mathcal{M}_n$, то есть T имеет вид, описанный в условиях теоремы.

Случай II: $n = 2$**.** Используя лемму 3.5, получаем, что существует обратимая матрица $P \in \mathcal{M}_2$ такая, что линейное отображение T_2 пространства матриц, определённое равенством

$$T_2(X) = P^{-1}T_1(X)P$$

для всех $X \in \mathcal{M}_2$, переводит $E_{1\,2}$ в $E_{1\,2}$, а $E_{2\,1}$ – в $E_{2\,1}$. Тогда, согласно лемме 2.10, либо

$$T_2(X) = X$$

для всех $X \in \mathcal{M}_2$, либо

$$T_2(X) = Q^{-1}X^tQ$$

для всех $X \in \mathcal{M}_2$, где $Q = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Выполняя обратную замену, получаем, что в первом случае

$$T(X) = T_1(X) - \alpha(X)I_n = PT_2(X)P^{-1} - \alpha(X)I_n = PXP^{-1} - \alpha(X)I_n$$

для всех $X \in \mathcal{M}_2$, а во втором случае

$$\begin{aligned} T(X) &= T_1(X) - \alpha(X)I_n = PT_2(X)P^{-1} - \alpha(X)I_n \\ &= PQ^{-1}X^tQP^{-1} - \alpha(X)I_n = (QP^{-1})^{-1}X^t(QP^{-1}) - \alpha(X)I_n \end{aligned}$$

для всех $X \in \mathcal{M}_2$, то есть и при $n = 2$ отображение T имеет вид, описанный в условиях теоремы. $\square$

Приложение A. Развёрнутые комментарии к некоторым утверждениям из статьи Пирса

Ниже приведены утверждения из статьи Пирса [13], которые, по нашему мнению, нуждаются в уточнении, дополнении или исправлении. Заметим, что наши рассуждения данные утверждения не используют.

Утверждение A.1 (см. [13, основная теорема]). *Предположим, что линейное отображение T пространства матриц $\mathcal{M}_n$ таково, что $\mathrm{D}(T(X)) = \mathrm{D}(X)$ для всех $X \in \mathcal{M}_n$. Тогда существуют обратимая матрица $M \in \mathcal{M}_n$, линейное отображение $\alpha\colon \mathcal{M}_n \to \mathbb{F}$ и $c \in \mathbb{F}$ такие, что $c^n = 1$ и*

$$T(X) = cM^{-1}XM + \alpha(X)I_n, \quad X \in \mathcal{M}_n,$$

или

$$T(X) = cM^{-1}X^tM + \alpha(X)I_n, \quad X \in \mathcal{M}_n.$$

Опровержение. Неверно, что $c^n = 1$. Из свойства MD5 матричного дискриминанта следует, что он является однородным многочленом степени $n(n-1)$. Если c является корнем степени $n(n-1)$ из единицы, но не является корнем степени n из единицы, то умножение на такое c будет являться линейным отображением, сохраняющим матричный дискриминант, однако равенство $c^n = 1$ выполняться не будет. ■

Для обсуждения следующего утверждения нам потребуется определение матричной элементарной симметрической функции.

Определение A.2. Пусть $A \in \mathcal{M}_n$. Тогда последовательность *матричных элементарных симметрических функций* $E_1(A), \ldots, E_n(A)$ определяется как знакопеременная последовательность коэффициентов многочлена $p_A(x)$ в разложении по степеням x, а именно, функции $E_1(A), \ldots, E_n(A)$ таковы, что

$$p_A(x) = x^n + (-1)^1 E_1(A)x^{n-1} + \ldots + (-1)^n E_n(A)$$

для всех $A \in \mathcal{M}_n$.

В частности, $\mathrm{tr}(A) = E_1(A)$, а $\det(A) = E_n(A)$.

Утверждение A.3 (см. [13, утверждение 9])**.** *Пусть W – подпространство $\mathcal{M}_3$ размерности 6. Тогда W содержит матрицу, дискриминант которой не равен нулю.*

Опровержение. Сначала заметим, что если $A \in \mathcal{M}_3$, то из определения матричных элементарных симметрических функций следует, что

$$p_A(x) = x^3 - E_1(A)x^2 + E_2(A)x - E_3(A).$$

Применяя к этому выражению формулу для дискриминанта многочлена третьей степени, получаем, что

$$\begin{aligned}\mathrm{D}(A) = E_1^2(A)E_2^2(A) - 4E_2^3(A) - 4E_1(A)^3E_3(A)\\ + 18E_1(A)E_2(A)E_3(A) - 27E_3^2(A).\end{aligned} \quad \text{(A.1)}$$

Поэтому, если $\mathrm{char}(\mathbb{F}) = 2$, то

$$\mathrm{D}(A) = E_1^2(A)E_2^2(A) + E_3^2(A).$$

Теперь рассмотрим пространство $W \subset \mathcal{M}_3$, определённое следующим образом:

$$W = \left\{ \begin{pmatrix} g & b & c \\ 0 & e & f \\ 0 & h & e \end{pmatrix} : b, c, d, e, f, g \in \mathbb{F} \right\}.$$

Очевидно, что $\dim(W) = 6$.

При этом, используя формулы $E_3(A) = \det(A)$, $E_1(A) = \operatorname{tr}(A)$ и

$$E_2(A) = \begin{vmatrix} A_{1\,1} & A_{1\,2} \\ A_{2\,1} & A_{2\,2} \end{vmatrix} + \begin{vmatrix} A_{1\,1} & A_{1\,3} \\ A_{3\,1} & A_{3\,3} \end{vmatrix} + \begin{vmatrix} A_{2\,2} & A_{2\,3} \\ A_{3\,2} & A_{3\,3} \end{vmatrix},$$

мы получаем, что если $A \in W$, то

$$\begin{aligned} \mathrm{D}(A) &= E_1^2(A)E_2^2(A) + E_3^2(A) \\ &= \operatorname{tr}^2(A)\left(\begin{vmatrix} A_{1\,1} & A_{1\,2} \\ A_{2\,1} & A_{2\,2} \end{vmatrix} + \begin{vmatrix} A_{1\,1} & A_{1\,3} \\ A_{3\,1} & A_{3\,3} \end{vmatrix} + \begin{vmatrix} A_{2\,2} & A_{2\,3} \\ A_{3\,2} & A_{3\,3} \end{vmatrix}\right)^2 + \det{}_3^2(A) \\ &= (g+e+e)^2(ge+ge+e^2-fh)^2 + \bigl(g(e^2-fh)\bigr)^2 \\ &= g^2(e^2-fh)^2 + g^2(e^2-fh)^2 = 0. \end{aligned}$$

Таким образом, дискриминант любого элемента W равен 0, что противоречит нашему утверждению. ■

Замечание A.4. Рассуждения в доказательстве утверждения A.3 также используют утверждение, что коэффициент при $E_3^2(A)$ в выражении (A.1) не равен нулю. Но это неверно, если $\operatorname{char}(\mathbb{F}) = 3$.

Дальнейшее утверждение использует понятие пространства $M(v)$. Сформулируем его так, как это сделано в [13].

Определение A.5 (см. [13, определение после утверждения 15])**.**

Пусть $0 \neq v \in \mathbb{F}^n$. Через $M(v)$ обозначим векторное пространство, порождённое I_n и всеми нильпотентными матрицами, образ которых равен $\langle v \rangle$. Через $M(v)'$ обозначим векторное пространство, образованное матрицами, транспонированными к элементам $M(v)$.

Утверждение A.6 (см. [13, утверждение 17])**.** *Пусть $T\colon M_n \to M_n$ – биективное линейное отображение, сохраняющее дискриминант, и пусть v_1, $v_2 \in \mathbb{F}^n$ линейно независимы. Тогда одновременное выполнение двух равенств $T(M(v_1)) = M(w_1)$ и $T(M(v_2) = M(w_2)'$ невозможно.*

Доказательство этого утверждения использует, в частности, следующее ошибочное утверждение.

Утверждение A.7. *Если $\langle w_1 \rangle \neq \langle w_2 \rangle$, то* $\dim(\langle M(w_1), M(w_2)' \rangle) = 2(n-1)$.

Опровержение. Действительно, это утверждение равносильно тому, что если $\langle w_1 \rangle \neq \langle w_2 \rangle$, то $\dim(M(w_1) \cap M(w_2)') = 2$. Но если $w_1 = (1,0,0)$, а $w_2 = (1,1,0)$, то, с одной стороны, $\langle w_1 \rangle \neq \langle w_2 \rangle$, а с другой стороны, прямые вычисления показывают, что $M(w_1) \cap M(w_2)' = \langle I_n \rangle$. ∎

Утверждение A.8 (см. [13, последний абзац]). *Утверждение* 20 *устанавливает, что если $A \in \mathcal{M}_n$ и $A = A_1 + dI_n$ <...> Это завершает доказательство основной теоремы.*

Опровержение. Таким образом, подразумевается, что всякая матрица представима в виде линейной комбинации скалярной матрицы и матрицы с нулевым следом. Это неверно, если $\operatorname{char}(\mathbb{F})$ делит n.

Действительно, пусть, например, $A = E_{1\,1}$. Тогда, с одной стороны, $\operatorname{tr}(A) = 1$. С другой стороны, если предположить, что

$$A = A_1 + dI_n \tag{A.2}$$

для некоторых $d \in \mathbb{F}$, $A_1 \in \mathcal{M}_n$ и $\operatorname{tr}(A_1) = 0$, то

$$\operatorname{tr}(A) = \operatorname{tr}(A_1 + dI_n) = \operatorname{tr}(A_1) + d\operatorname{tr}(I_n) = 0 + d \cdot n = 0 + 0 = 0,$$

так как мы предположили в начале, что $\operatorname{char}(\mathbb{F}) \mid n$, и поэтому $n = 0$ в $\mathbb{F}$. Значит, для данной матрицы A представление (A.2) невозможно. ∎

Замечание A.9. Если предположить, что $\operatorname{char}(\mathbb{F})$ не делит n, то, действительно, всякую матрицу можно представить в виде линейной комбинации скалярной матрицы и матрицы с нулевым следом, и обсуждаемое рассуждение будет верным.

Автор выражает благодарность своему научному руководителю профессору А. Э. Гутерману за постановку задачи, внимание к работе и ценные обсуждения.

Список литературы

1. Э. Артин, Геометрическая алгебра. Москва, Наука, 1969.
2. Н. В. Илюшечкин, *Дискриминант характеристического многочлена нормальной матрицы.* — Матем. заметки **51**, No. 3 (1992), 16–23.
3. С. Ленг, Алгебра. Москва, Наука, 1968.
4. В. В. Прасолов. Многочлены. Москва, МЦНМО, 2003.
5. Ф. Г. Фробениус, *О представлении конечных групп через линейные подстановки.* В кн.: Теория характеров и представлений групп. Харьков, ОНТИ, 1938, 106—127.

6. P. Botta, S. Pierce, W. Watkins, *Linear transformations that preserve the nilpotent matrices.* —- Pacific J. Math. **104**, No. 1 (1983), 39–46.
7. J. Dieudonné, *Sur une généralisation du groupe orthogonal à quatre variables.* — Arch. Math. **1**, No. 4 (1948), 282–287.
8. I. M. Gelfand, M. M. Kapranov, A. V. Zelevinsky, Discriminants, Resultants, and Multidimensional Determinants. Birkhauser, Boston, 1994.
9. E. E. Kummer, *Bemerkungen über die cubische Gleichung, durch welche die Haupt-Axen der Flächen zweiten Grades bestimmt werden.* — J. Reine Angew. Math. **26** (1843), 268–272.
10. M. Marcus, R. Purves, *Linear transformations on algebras of matrices: the invariance of the elementary symmetric functions.* — Can. J. Math. **11** (1959), 383–396.
11. B. N. Parlett, *The (matrix) discriminant as a determinant.* — Linear Algebra Appl. **355**, Nos. 1–3 (2002), 85–101.
12. S. Pierce et.al., *A survey of linear preserver problems.* — Linear Multilinear Algebra **33**, Nos. 1–2 (1992).
13. S. Pierce, *Discriminant preserving linear maps.* — Linear Multilinear Algebra **8**, No. 2 (1979), 101–114.

Yurkov A. Linear maps preserving matrix discriminant.

Let $\mathbb{F}$ be an infinite field, $\mathcal{M}_n(\mathbb{F})$ be the space of all $n \times n$ matrices over $\mathbb{F}$. The paper provides a characterization of all the linear maps on $\mathcal{M}_n$ that preserve matrix discriminant. The Pierce result on these maps is improved.

Университет имени Бар-Илана
5290002 Рамат-Ган, Израиль
E-mail: andrey.yurkov@biu.ac.il

Поступило 16 октября 2024 г.