Записки научных
семинаров ПОМИ
Том 303, 2003 г.

Н. В. Илюшечкин

О НЕКОТОРЫХ ТОЖДЕСТВАХ НАД ЭЛЕМЕНТАМИ СИММЕТРИЧЕСКОЙ МАТРИЦЫ

Введение

В некоторых сплошных средах встречается не тензорная [1], а более общая зависимость тензора напряжений от тензора деформаций. Свойства упомянутой зависимости приводят к постановке следующей задачи.

Пусть $\mathrm{Sym}(n)$ – пространство n-мерных вещественных симметрических матриц. Требуется описать семейство бесконечно гладких отображений

$$\mathcal{F} : \mathrm{Sym}(n) \longrightarrow \mathrm{Sym}(n),$$

обладающих следующим свойством: для любой матрицы $X \in \mathrm{Sym}(n)$ и такой ортогональной матрицы C, что $C^{-1}XC$ – диагональная матрица, $C^{-1}\mathcal{F}(X)C$ тоже является диагональной матрицей.

Можно поставить также задачу об описании семейства бесконечно гладких отображений

$$\mathcal{G} : \mathrm{Sym}(n) \longrightarrow \mathrm{Sym}(n),$$

в каком-то смысле двойственного к предыдущему семейству. Именно, будем считать, что из диагональности матрицы $C^{-1}XC$ следует равенство нулю всех диагональных элементов матрицы $C^{-1}\mathcal{G}(X)C$.

Ниже будут доказаны следующие теоремы.

Теорема 1. *Если $\mathcal{F}$ – отображение из первого семейства, то существуют такие бесконечно гладкие функции*

$$a_m : \mathrm{Sym}(n) \longrightarrow R$$

для $m = 0, 1, \dots, n-1$, что

$$\mathcal{F}(X) = a_0(X)E + a_1(X)X + \dots + a_{n-1}(X)X^{n-1}, \tag{0.1}$$

то есть $\mathcal{F}(X)$ является матричным многочленом степени $n-1$ с переменными коэффициентами.

Теорема 2. *Для любого отображения $\mathcal{G}$ из второго семейства существует такая бесконечно гладкая кососимметрическая матрица $A(X)$, что*

$$\mathcal{G}(X) = A(X)X - XA(X), \tag{0.2}$$

то есть $\mathcal{G}(X)$ является матричным коммутатором X с переменной кососимметрической матрицей.

Попытки доказательства сформулированных выше теорем привели к установлению некоторых тождеств, представляющих, по мнению автора, и самостоятельный интерес.

§1. Об объемах, порожденных функциями от нормальной матрицы

Хотя наше исследование посвящено симметрическим матрицам, все же одна необходимая в дальнейшем формула тривиально обобщается на более общий случай нормальных матриц, причем с комплексными элементами.

Обозначим теперь через $\mathrm{Mat}_C(n)$ пространство квадратных комплексных матриц порядка n, которое мы будем рассматривать как унитарное пространство размерности n^2. Пусть $Z \in \mathrm{Mat}_C(n)$ – нормальная матрица, и пусть $f_1(\zeta), \ldots, f_n(\zeta)$ – комплекснозначные функции комплексного аргумента, определенные на спектре $\lambda_1, \ldots, \lambda_n$ матрицы Z. Нас будет интересовать объем параллелепипеда, натянутого в $\mathrm{Mat}_C(n)$ на векторы $f_1(Z), \ldots, f_n(Z)$.

Вычисление объема начнем с одного простого замечания. Пусть $\mathbf{w}_1, \ldots, \mathbf{w}_l$ – векторы унитарного пространства размерности $m \geq l$. Объем параллелепипеда, натянутого на эти векторы, будем обозначать $V(\mathbf{w}_1, \ldots, \mathbf{w}_l)$. Если все $\mathbf{w}_j$ мыслить как векторы-столбцы $\mathbf{w}_j = (w_{1j}, \ldots, w_{mj})^T$, то из их координат можно составить матрицу $W = (w_{ij})$, где $i = 1, \ldots, m$ и $j = 1, \ldots, l$. Обозначим через $\mathbb{M}_r(W)$ сумму квадратов абсолютных величин всех миноров порядка r матрицы W. Тогда, как известно [2],

$$V^2(\mathbf{w}_1, \ldots, \mathbf{w}_l) = \mathbb{M}_l(W). \tag{1.1}$$

Заметим, что объем не изменится, если все векторы $\mathbf{w}_j$ подвергнуть одному и тому же унитарному преобразованию Q, то есть

$$V(\mathbf{w}_1, \ldots, \mathbf{w}_l) = V(Q\mathbf{w}_1, \ldots, Q\mathbf{w}_l). \tag{1.2}$$

В самом деле, квадрат объема $V^2(\mathbf{w}_1, \ldots, \mathbf{w}_l)$ совпадает с определителем матрицы Грама, порожденной векторами $\mathbf{w}_1, \ldots, \mathbf{w}_l$, а скалярные произведения $(\mathbf{w}_j, \mathbf{w}_k) = (Q\mathbf{w}_j, Q\mathbf{w}_k)$, которые суть элементы матрицы Грама, не изменяются при унитарном преобразовании всего пространства.

Вернемся к вычислению объема $V(f_1(Z), \ldots, f_n(Z))$. Из элементов матриц $f_1(Z), \ldots, f_n(Z)$ можно образовать новую матрицу $\Phi(Z)$ размера $n^2 \times n$, в ν-м ($\nu = 1, \ldots, n$) столбце которой стоят элементы матрицы $f_\nu(Z)$, взятые в определенном, одном и том же для всех ν, порядке. Тогда в силу (1.1) будем иметь

$$V^2(f_1(Z), \ldots, f_n(Z)) = \mathbb{M}_n(\Phi(Z)). \tag{1.3}$$

Заметим, что преобразование подобия $N = U^{-1}MU$, где переменные M, N взяты из $\mathrm{Mat}_C(n)$, а U – постоянная унитарная матрица, можно рассматривать как унитарное преобразование пространства $\mathrm{Mat}_C(n)$. Выберем унитарную матрицу U так, чтобы матрица $L = U^{-1}ZU$ была диагональной:

$$L = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Тогда, разумеется, $f_\nu(L) = U^{-1}f_\nu(Z)U$ и

$$f_\nu(L) = \mathrm{diag}(f_\nu(\lambda_1), \ldots, f_\nu(\lambda_n)).$$

Из (1.2) благодаря унитарности преобразования подобия получаем

$$V(f_1(Z), \ldots, f_n(Z)) = V(f_1(L), \ldots, f_n(L)),$$

и вследствие (1.3) для определения искомого объема нам остается найти величину $\mathbb{M}_n(\Phi(L))$. Но в матрице $\Phi(L)$ имеется всего один ненулевой минор порядка n, образованный диагональными элементами матриц $f_\nu(Z)$, так что $\mathbb{M}_n(\Phi(L))$ есть квадрат абсолютной величины упомянутого минора. Извлекая квадратный корень, получим окончательную формулу

$$V(f_1(Z), \ldots, f_n(Z)) = \left\| \begin{matrix} f_1(\lambda_1) & f_2(\lambda_1) & \ldots & f_n(\lambda_1) \\ f_1(\lambda_2) & f_2(\lambda_2) & \ldots & f_n(\lambda_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\lambda_n) & f_2(\lambda_n) & \ldots & f_n(\lambda_n) \end{matrix} \right\|.$$

Например, пусть все $f_\nu(\zeta) = p_\nu(\zeta)$ – многочлены степени $n-1$:

$$p_\nu(\zeta) = b_{\nu 0} + b_{\nu 1}\zeta + \cdots + b_{\nu, n-1}\zeta^{n-1}.$$

Коэффициенты этих многочленов образуют квадратную матрицу порядка n:

$$B = \begin{pmatrix} b_{10} & b_{11} & \dots & b_{1,n-1} \\ b_{20} & b_{21} & \dots & b_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n0} & b_{n1} & \dots & b_{n,n-1} \end{pmatrix}.$$

Тогда, как видно из теоремы об умножении определителей,

$$V(p_1(Z), \dots, p_n(Z)) = |\det(B)||W(\lambda_1, \dots, \lambda_n)|,$$

где W – определитель Вандермонда, порожденный числами $\lambda_1, \dots, \lambda_n$. Возводя в квадрат, получим

$$V^2(p_1(Z), \dots, p_n(Z)) = |\det(B)|^2 |\mathrm{Dis}(\chi_Z)|,$$

где χ_Z – характеристический многочлен матрицы Z, а $\mathrm{Dis}(\chi_Z)$ – его дискриминант. В частности,

$$V^2(E, Z, \dots, Z^{n-1}) = |\mathrm{Dis}(\chi_Z)|.$$

Вследствие (1.3) алгебраический аналог последней формулы приобретает вид

$$|\mathrm{Dis}(\chi_Z)| = \mathbb{M}_n(Z_*), \tag{1.4}$$

где Z_* – матрица размера $n^2 \times n$, в ν-м ($\nu = 0, 1, \dots, n-1$) столбце которой стоят элементы матрицы Z^ν, взятые в определенном, одном и том же для всех ν, порядке.

В случае, когда нормальная матрица $Z = X \in \mathrm{Sym}(n)$ симметрична, формулы (1.3) и (1.4) могут быть несколько видоизменены. Предлагаемое изменение связано с тем, что размерность пространства $\mathrm{Sym}(n)$, которую мы всегда будем обозначать через $N = C_{n+1}^2$, меньше размерности n^2 пространства $\mathrm{Mat}_C(n)$ или числа строк в матрицах $\Phi(X)$ и X_*.

Перейдем к точным определениям. Матрицу $X = (x_{ij}) \in \mathrm{Sym}(n)$ мы впредь будем рассматривать как столбец $\widehat{X}$ размерности N, элементы которого суть x_{ii} и $\sqrt{2}x_{ij}$ при $i < j$, расположенные в порядке

$$11, 22, \dots, nn, 12, 13, \dots, 1n, 23, \dots, 2n, \dots, n-1\, n. \tag{1.5}$$

Выбор множителя $\sqrt{2}$ обусловлен тем, что преобразование подобия

$$X = C^{-1} Y C, \tag{1.6}$$

где X и Y – переменные симметрические матрицы, а C – постоянная ортогональная матрица, сохраняет квадратичную форму

$$\sum_{i=1}^{n} x_{ii}^2 + 2\sum_{i<j} x_{ij}^2.$$

Через E_{ij} будем обозначать квадратную матрицу порядка n, на пересечении i-й строки и j-го столбца которой стоит единица, а все остальные элементы равны нулю. Определим также симметрические матрицы Γ_{ii} при $i = 1, \ldots, n$ и Γ_{ij} при $i < j$ равенствами

$$\Gamma_{ii} = E_{ii}, \qquad \Gamma_{ij} = \frac{1}{\sqrt{2}}(E_{ij} + E_{ji}). \tag{1.7}$$

Тогда $\mathrm{Sym}(n)$ можно рассматривать как евклидово пространство размерности N с ортонормированным базисом (1.7), расположенным в порядке (1.5).

Рассуждениями, аналогичными приведенным выше, можно вывести формулы, подобные (1.3) и (1.4), но в размерности N. Именно, вместо матрицы $\Phi(X)$ из (1.3) можно использовать матрицу $\Phi_s(X)$ размера $N \times n$, в ν-м ($\nu = 1, \ldots, n$) столбце которой стоят элементы матрицы $f_\nu(X)$, расположенные в порядке (1.5) и умноженные, где это нужно, на $\sqrt{2}$. Тогда формула (1.3) преобразуется к виду

$$V^2(f_1(X), \ldots, f_n(X)) = \mathbb{M}_n(\Phi_s(X)),$$

а формула (1.4) – к виду

$$\mathrm{Dis}(\chi_X) = \mathbb{M}_n(X_!), \tag{1.8}$$

где матрица $X_!$ образована столбцами $\widehat{X^0}, \widehat{X}, \ldots, \widehat{X^{n-1}}$, расположенными в естественном порядке. Вообще, если $X = (x_{ij}) \in \mathrm{Sym}(n)$, то элементы матрицы X^m для $m = 0, 1, \ldots, n-1$ обозначим через $x_{ij}^{(m)}$. Таким образом, в строке $X_!$ с номером ii стоят элементы

$$1, x_{ii}, x_{ii}^{(2)}, \ldots, x_{ii}^{(n-1)},$$

а строка с номером ij при $i < j$ образована элементами

$$0, \sqrt{2}x_{ij}, \sqrt{2}x_{ij}^{(2)}, \ldots, \sqrt{2}x_{ij}^{(n-1)}.$$

Например, в случае $n = 3$ матрица $X_!$ имеет вид

$$X_! = \begin{pmatrix} 1 & x_{11} & x_{11}^{(2)} \\ 1 & x_{22} & x_{22}^{(2)} \\ 1 & x_{33} & x_{33}^{(2)} \\ 0 & \sqrt{2}x_{12} & \sqrt{2}x_{12}^{(2)} \\ 0 & \sqrt{2}x_{13} & \sqrt{2}x_{13}^{(2)} \\ 0 & \sqrt{2}x_{23} & \sqrt{2}x_{23}^{(2)} \end{pmatrix}.$$

§2. Дискриминантные многочлены

Как мы увидим ниже, максимальные миноры матрицы $X_!$ имеют прямое отношение к доказательству теоремы 1. Минор порядка n матрицы $X_!$, образованный строками с двойными индексами $i_1j_1, \ldots, i_nj_n$, обозначим через $\langle i_1j_1, \ldots, i_nj_n\rangle(X)$. Это – однородный многочлен степени C_n^2 от элементов матрицы X. Например, при $n = 3$ имеем

$$\langle 11, 22, 33\rangle(X) = \begin{vmatrix} 1 & x_{11} & x_{11}^2 + x_{12}^2 + x_{13}^2 \\ 1 & x_{22} & x_{12}^2 + x_{22}^2 + x_{23}^2 \\ 1 & x_{33} & x_{13}^2 + x_{23}^2 + x_{33}^2 \end{vmatrix} =$$

$$= (x_{22} - x_{11})(x_{33} - x_{11})(x_{33} - x_{22}) -$$

$$-(x_{22} - x_{11})x_{12}^2 + (x_{33} - x_{11})x_{13}^2 - (x_{33} - x_{22})x_{23}^2;$$

$$\langle 33, 12, 23\rangle(X) = 2\begin{vmatrix} x_{12} & x_{11}x_{12} + x_{12}x_{22} + x_{13}x_{23} \\ x_{23} & x_{12}x_{13} + x_{22}x_{23} + x_{23}x_{33} \end{vmatrix} =$$

$$= 2x_{12}x_{23}(x_{33} - x_{11}) + 2x_{13}(x_{12}^2 - x_{23}^2).$$

Как мы уже доказали, сумма квадратов всех этих миноров равна дискриминанту $\mathrm{Dis}(\chi_X)$. Поэтому миноры $\langle i_1j_1, \ldots, i_nj_n\rangle(X)$ мы в дальнейшем будем называть дискриминантными многочленами.

Скажем несколько слов о многочлене $T(X) = \langle 11, \ldots, nn\rangle(X)$, который занимает центральное в каком-то смысле положение во множестве дискриминантных многочленов. Именно, мы докажем его неприводимость.

Определим сначала действие группы подстановок $S(n)$ на множестве квадратных матриц. Пусть $A = (a_{ij})$ – матрица порядка n, а $\sigma \in S(n)$ – подстановка. Тогда по определению $(\sigma A)_{ij} = a_{\sigma(i)\sigma(j)}$. Нетрудно видеть, что σ – кольцевой гомоморфизм, то есть

$$\sigma(A + B) = \sigma(A) + \sigma(B), \qquad \sigma(AB) = \sigma(A)\sigma(B).$$

В частности, для $X \in \mathrm{Sym}(n)$ имеем

$$\sigma(X^m) = [\sigma(X)]^m. \tag{2.1}$$

Можно также определить действие группы $S(n)$ на множестве $\{P(X)\}$ многочленов, зависящих от переменных x_{ij}. Разумеется, по определению

$$\sigma P(\ldots, x_{ij}, \ldots) = P(\ldots, x_{\sigma(i)\sigma(j)}, \ldots),$$

причем σ – автоморфизм кольца многочленов.

Приступим к доказательству неприводимости многочлена $T(X)$. Так как

$$T(X) = \begin{vmatrix} 1 & x_{11} & x_{11}^{(2)} & \ldots & x_{11}^{(n-1)} \\ 1 & x_{22} & x_{22}^{(2)} & \ldots & x_{22}^{(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{nn} & x_{nn}^{(2)} & \ldots & x_{nn}^{(n-1)} \end{vmatrix}, \tag{2.2}$$

то вследствие (2.1) для любой подстановки σ имеем

$$\sigma x_{ij}^{(m)} = x_{\sigma(i)\sigma(j)}^{(m)},$$

так что

$$\sigma T(X) = \pm T(X), \tag{2.3}$$

где знак зависит от четности подстановки.

Далее, элементы x_{ii} назовем диагональными, а все остальные переменные – недиагональными. Нетрудно видеть, что элемент $x_{ii}^{(m)}$ можно представить в виде $x_{ii}^{(m)} = x_{ii}^m + \widetilde{x}_{ii}^{(m)}$, где каждый одночлен многочлена $\widetilde{x}_{ii}^{(m)}$ содержит недиагональные переменные. Значит, можно написать

$$T(X) = W(x_{11}, x_{22}, \ldots, x_{nn}) + \widetilde{T}(X), \tag{2.4}$$

где W – как и прежде, определитель Вандермонда, а каждый одночлен многочлена $\widetilde{T}(X)$ содержит недиагональные переменные. Однако из (2.2) видно, что в $\widetilde{T}(X)$ нет одночлена, содержащего только недиагональные переменные.

Теперь предположим, что $Q(X)$ – неприводимый делитель $T(X)$. В силу (2.4) многочлен $Q(X)$ имеет вид

$$Q(X) = R(x_{11}, x_{22}, \ldots, x_{nn}) + \widetilde{Q}(X), \tag{2.5}$$

где R зависит только от диагональных переменных и является делителем W, а в каждый одночлен $\widetilde{Q}(X)$ входят недиагональные переменные. Из (2.3) следует, что для любой подстановки $\sigma \in S(n)$ многочлен $\sigma(Q)$ тоже является неприводимым делителем многочлена $T(X)$, а из (2.4) и (2.5) видно, что $\sigma(R)$ – делитель многочлена W. Так как W распадается в произведение линейных форм $x_{jj} - x_{ii}$, а для любых двух различных таких форм существует подстановка, переводящая первую из них во вторую, то любой другой неприводимый делитель многочлена $T(X)$ имеет вид $\sigma(Q)$ для некоторой подстановки σ. Далее нужно рассмотреть три случая.

1. Q – форма первой степени. Тогда $T(X)$ распадается в произведение линейных форм. При этом, если $\widetilde{Q}(X) \equiv 0$, то $T(X) = W(x_{11}, \dots, x_{nn})$, что очевидным образом неверно. Если же $\widetilde{Q}(X) \not\equiv 0$, то и все $\sigma(\widetilde{Q})$ не равны тождественно нулю и не зависят от диагональных переменных. Но тогда $T(X)$ содержит одночлены, не зависящие от диагональных переменных, что также неверно.

2. Q – форма степени выше первой. Тогда R – произведение некоторых сомножителей вида $x_{jj} - x_{ii}$, причем все они различны. Легко показать, что всегда можно подобрать такую подстановку σ, что $\sigma(R) \neq R$, но у R и $\sigma(R)$ будет одинаковый делитель вида $x_{jj} - x_{ii}$ (исключение составляет случай $n = 4$, когда степень формы Q равна двум, о нем ниже). Так как формы Q и $\sigma(Q)$ различны и неприводимы, то обе они входят в разложение многочлена $T(X)$ на неприводимые множители. Поэтому W делится на произведение $R\sigma(R)$, содержащее множитель $(x_{jj} - x_{ii})^2$, что невозможно.

3. Исключительный случай $n = 4$,

$$R = (x_{jj} - x_{ii})(x_{ll} - x_{kk}),$$

где (i, j, k, l) – перестановка чисел 1,2,3,4. В этом случае $T(X)$ распадается в произведение трех квадратичных форм. Очевидно, форма $\widetilde{Q}$ должна иметь первую степень по диагональным переменным и первую – по недиагональным. Тогда любой одночлен в $T(X)$ имеет по крайней мере третью степень по диагональным переменным. Покажем, что на самом деле это не так. Придадим переменным x_{ij} следующие значения:

$$x_{11} = x, \qquad x_{12} = x_{23} = x_{34} = t,$$

а все остальные переменные положим равными нулю. Тогда, как нетрудно видеть,

$$\langle 11, 22, 33, 44\rangle(X) = x^2t^4,$$

откуда следует, что $T(X)$ обязан содержать одночлены, имеющие степень по диагональным переменным не выше второй – противоречие.

Оказывается, дискриминантные многочлены можно получить из совсем других соображений. Именно, пусть $\mathrm{so}(n)$ – пространство кососимметрических матриц порядка n. Размерность этого пространства мы будем обозначать через $M = C_n^2$. Матрицу $A = (a_{kl}) \in \mathrm{so}(n)$ можно рассматривать как столбец $\bar{A}$ размерности M, элементы которого $\sqrt{2}a_{kl}$ при $k < l$ расположены в порядке (1.5), но, разумеется, без первых n элементов. Таким образом, мы будем рассматривать $\mathrm{so}(n)$ как евклидово пространство размерности M с ортонормированным базисом

$$\Delta_{ij} = \frac{1}{\sqrt{2}}(E_{ij} - E_{ji}) \tag{2.6}$$

при $i < j$, расположенным в порядке (1.5). Пусть $X \in \mathrm{Sym}(n)$. Рассмотрим линейное отображение

$$X^{\square} : \mathrm{so}(n) \longrightarrow \mathrm{Sym}(n),$$

задаваемое формулой

$$X^{\square}(A) = AX - XA.$$

Матрицу этого отображения в базисах (1.7) и (2.6) обозначим через $X^{!}$. С учетом введенных ранее обозначений имеет место матрично-векторное равенство $\widehat{H} = X^{!}\bar{A}$, где $H = X^{\square}(A)$. Дадим описание матрицы $X^{!}$. Прежде всего, она имеет размер $N \times M$. Ее строки пронумерованы двойными индексами ij при $i \leqslant j$ и расположены в порядке (1.5). Столбцы пронумерованы двойными индексами kl при $k < l$ и расположены в том же порядке.

Очевидно, если обозначить

$$\Psi_{kl} = X^{\square}(\Delta_{kl}),$$

то $\widehat{\Psi}_{kl}$ есть kl-ый столбец матрицы $X^{!}$.

Поэтому строка матрицы $X^{!}$ с индексом ii выглядит следующим образом. Если $k < i$, то в столбце ki находится число

$-\sqrt{2}x_{ki}$, а если $k > i$, то в столбце ik находится число $\sqrt{2}x_{ik}$. Все остальные элементы строки равны нулю.

Теперь опишем строку с индексом ij при $i < j$. Если $k < i$, то в столбце ki лежит число $-x_{kj}$, а в столбце kj – число $-x_{ki}$. Если $i < k < j$, то в столбце ik находится число x_{kj}, а в столбце kj – число $-x_{ik}$. Если $j < k$, то столбец ik содержит число x_{jk}, а столбец jk – число x_{ik}. Наконец, в столбце ij расположено число $x_{jj} - x_{ii}$. Все остальные элементы строки равны нулю.

Например, пусть $n = 3$. Тогда

$$X^! = \begin{pmatrix} \sqrt{2}x_{12} & \sqrt{2}x_{13} & 0 \\ -\sqrt{2}x_{12} & 0 & \sqrt{2}x_{23} \\ 0 & -\sqrt{2}x_{13} & -\sqrt{2}x_{23} \\ x_{22} - x_{11} & x_{23} & x_{13} \\ x_{23} & x_{33} - x_{11} & -x_{12} \\ -x_{13} & -x_{12} & x_{33} - x_{22} \end{pmatrix}.$$

Нас будут интересовать максимальные миноры матрицы $X^!$. Минор, содержащий все строки, кроме $i_1j_1, \dots, i_nj_n$, обозначим через $\{i_1j_1, \dots, i_nj_n\}(X)$. Очевидно, что это – однородный многочлен степени M от элементов x_{ij} матрицы X. Например, при $n = 3$ имеем

$$\{11, 22, 33\}(X) = \begin{vmatrix} x_{22} - x_{11} & x_{23} & x_{13} \\ x_{23} & x_{33} - x_{11} & -x_{12} \\ -x_{13} & -x_{12} & x_{33} - x_{22} \end{vmatrix} = \langle 11, 22, 33\rangle(X);$$

$$\{33, 12, 23\}(X) = \begin{vmatrix} \sqrt{2}x_{12} & \sqrt{2}x_{13} & 0 \\ -\sqrt{2}x_{12} & 0 & \sqrt{2}x_{23} \\ x_{23} & x_{33} - x_{11} & -x_{12} \end{vmatrix} = -\langle 33, 12, 23\rangle(X).$$

Мы докажем, что и вообще, при любом наборе индексов справедливо равенство

$$\{i_1j_1, \dots, i_nj_n\}(X) = \pm\langle i_1j_1, \dots, i_nj_n\rangle(X). \tag{2.7}$$

§3. Об изоморфизме Ходжа над пространством симметрических операторов

Чтобы при доказательстве равенства (2.7) избежать утомительных вычислений с ассоциированными матрицами, мы будем работать в инвариантной ситуации, то есть не с матрицами, а с линейными операторами. При этом мы сохраним все предыдущие обозначения, но инвариантные объекты будем отмечать жирным шрифтом. В основу доказательства мы положим известное свойство изоморфизма Ходжа [3], гласящее, что координаты соответствующих друг другу при этом изоморфизме кососимметрических тензоров могут отличаться только знаком. Введем необходимые обозначения и вкратце напомним нужные нам свойства инвариантных объектов.

Пусть $\mathcal{L}$ – вещественное n-мерное евклидово пространство со скалярным умножением $(\ ,\ )$. Через $\mathbf{End}(n)$ обозначим пространство линейных операторов над $\mathcal{L}$. Нулевой оператор в $\mathbf{End}(n)$ будем обозначать через $\mathbf{O}$.

Если в пространстве $\mathcal{L}$ задан ортонормированный базис $\mathbf{e}_1, \ldots, \mathbf{e}_n$, то в пространстве $\mathbf{End}(n)$ тоже можно естественным образом ввести ортогональную систему координат. Для этого нам сначала придется определить конструкцию тензорного произведения $(\mathbf{a} \otimes \mathbf{b})$ для каждой упорядоченной пары векторов $\mathbf{a}, \mathbf{b} \in \mathcal{L}$. Именно, $(\mathbf{a} \otimes \mathbf{b}) \in \mathbf{End}(n)$ – линейный оператор, определяемый формулой

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{x}) = (\mathbf{b}, \mathbf{x})\mathbf{a}.$$

Всевозможные произведения $(\mathbf{e}_i \otimes \mathbf{e}_j)$ элементов базиса $\mathcal{L}$ образуют базис пространства $\mathbf{End}(n)$, который мы будем называть базисом, сопутствующим базису $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Очевидно, для любых i, j, k имеем

$$(\mathbf{e}_i \otimes \mathbf{e}_j)(\mathbf{e}_k) = \delta_{jk}\mathbf{e}_i, \tag{3.1}$$

где δ_{jk} – символ Кронекера. Оператору $(\mathbf{e}_i \otimes \mathbf{e}_j)$ в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ соответствует матрица E_{ij}. Разумеется, если $A = (a_{ij})$ – матрица оператора $\mathbf{A} \in \mathbf{End}(n)$ в том же базисе, то

$$\mathbf{A} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(\mathbf{e}_i \otimes \mathbf{e}_j).$$

В пространстве $\mathbf{End}(n)$ можно естественным образом определить скалярное умножение $\langle\ ,\ \rangle$, тесно связанное с умножением $(,)$ в $\mathcal{L}$. Для любых двух операторов $\mathbf{A}, \mathbf{B} \in \mathbf{End}(n)$ их скалярное произведение вводится формулой

$$\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{A}^*\mathbf{B}),$$

где $\mathbf{A}^*$ – оператор, сопряженный с $\mathbf{A}$, а символом tr обозначается след линейного оператора. В любом ортонормированном базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ след определяется формулой

$$\mathrm{tr}\mathbf{A} = \sum_{i=1}^{n} (\mathbf{A}\mathbf{e}_i, \mathbf{e}_i).$$

Нетрудно найти попарные скалярные произведения элементов базиса $(\mathbf{e}_i \otimes \mathbf{e}_j)$. Именно,

$$\langle (\mathbf{e}_i \otimes \mathbf{e}_j), (\mathbf{e}_k \otimes \mathbf{e}_l) \rangle = \delta_{ik}\delta_{jl}, \tag{3.2}$$

то есть сопутствующий базис ортонормирован и координаты оператора $\mathbf{A} \in \mathbf{End}(n)$ в этом базисе совпадают с элементами матрицы A этого оператора в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Оператор $\mathbf{A}$ называется симметрическим, если $\mathbf{A}^* = \mathbf{A}$, и кососимметрическим, если $\mathbf{A}^* = -\mathbf{A}$. Подпространство симметрических операторов будем обозначать символом $\mathbf{Sym}(n)$, кососимметрических – символом $\mathbf{so}(n)$.

Пусть $\mathbf{X} \in \mathbf{Sym}(n)$. Так как матрица $X = (x_{ij})$ оператора $\mathbf{X}$ в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ симметрична, то

$$\mathbf{X} = \sum_{i=1}^{n} x_{ii}(\mathbf{e}_i \otimes \mathbf{e}_i) + \sum_{i<j} x_{ij}((\mathbf{e}_i \otimes \mathbf{e}_j) + (\mathbf{e}_j \otimes \mathbf{e}_i)).$$

Рассмотрим симметрические операторы

$$\mathbf{\Gamma}_{ii} = (\mathbf{e}_i \otimes \mathbf{e}_i), \qquad \mathbf{\Gamma}_{ij} = \frac{1}{\sqrt{2}}(\mathbf{e}_i \otimes \mathbf{e}_j) + \frac{1}{\sqrt{2}}(\mathbf{e}_j \otimes \mathbf{e}_i), \tag{3.3}$$

где $i < j$. Этим операторам соответствуют введенные ранее матрицы Γ_{ii} и Γ_{ij}. Очевидно, любой оператор $\mathbf{X} \in \mathbf{Sym}(n)$ может быть разложен по системе (3.3):

$$\mathbf{X} = \sum_{i=1}^{n} x_{ii}\Gamma_{ii} + \sqrt{2}\sum_{i<j} x_{ij}\mathbf{\Gamma}_{ij}.$$

Кроме того, из (3.2) нетрудно вывести, что система (3.3) ортонормирована. Следовательно, эта система операторов образует ортонормированный базис евклидова пространства $\mathbf{Sym}(n)$ размерности N, причем координатами оператора $\mathbf{X} \in \mathbf{Sym}(n)$ в этом базисе будут диагональные элементы x_{ii} его матрицы X в базисе

$\mathbf{e}_1, \ldots, \mathbf{e}_n$, а также умноженные на $\sqrt{2}$ недиагональные элементы x_{ij} этой же матрицы при $i < j$. Элементы базиса (3.3) всегда будем располагать в порядке (1.5).

Теперь пусть $\mathbf{A} \in \mathbf{so}(n)$. Так как матрица $A = (a_{ij})$ оператора $\mathbf{A}$ в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$ кососимметрична, то

$$\mathbf{A} = \sum_{i<j} a_{ij}((\mathbf{e}_i \otimes \mathbf{e}_j) - (\mathbf{e}_j \otimes \mathbf{e}_i)),$$

так что в пространстве $\mathbf{so}(n)$ естественно возникает ортонормированный базис

$$\mathbf{\Delta}_{ij} = \frac{1}{\sqrt{2}}(\mathbf{e}_i \otimes \mathbf{e}_j) - \frac{1}{\sqrt{2}}(\mathbf{e}_j \otimes \mathbf{e}_i) \tag{3.4}$$

при $i < j$. Элементы этого базиса расположим в том же порядке (1.5). Им, очевидно, соответствуют матрицы Δ_{ij}. Координатами оператора $\mathbf{A} \in \mathbf{so}(n)$ в этом базисе будут умноженные на $\sqrt{2}$ недиагональные элементы a_{ij} (при $i < j$) его матрицы A в базисе $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

С этого момента фиксируем в пространстве $\mathcal{L}$ какую-то одну определенную ориентацию, то есть допустимыми будем считать только преобразования базисов с собственно ортогональными матрицами. Докажем, что при таких преобразованиях и сопутствующие базисы (3.3) и (3.4) также подвергаются собственно ортогональным преобразованиям. Пусть базисы $\mathbf{e}_1, \ldots, \mathbf{e}_n$ и $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ связаны формулами

$$\mathbf{e}'_i = \sum_{j=1}^{n} c_{ji}\mathbf{e}_j,$$

где матрица C – собственно ортогональная, $\det C = 1$. Тогда легко вывести формулы замены сопутствующих базисов:

$$\begin{gathered} \mathbf{\Gamma}'_{ii} = \sum_{k=1}^{n} c_{ki}^2 \mathbf{\Gamma}_{kk} + \sqrt{2}\sum_{k<l} c_{ki}c_{li}\mathbf{\Gamma}_{kl}, \\ \mathbf{\Gamma}'_{ij} = \sqrt{2}\sum_{k=1}^{n} c_{ki}c_{kj}\mathbf{\Gamma}_{kk} + \sum_{k<l}(c_{ki}c_{lj} + c_{kj}c_{li})\mathbf{\Gamma}_{kl}, \\ \mathbf{\Delta}'_{ij} = \sum_{k<l}(c_{ki}c_{lj} - c_{kj}c_{li})\mathbf{\Delta}_{kl}. \end{gathered} \tag{3.5}$$

Пусть $C^\wedge$ – матрица перехода от старого базиса к новому в пространстве $\mathbf{Sym}(n)$, $C_\wedge$ – соответствующая матрица в пространстве $\mathbf{so}(n)$. Так как все базисы, связанные формулами (3.5), ортонормированы, то обе матрицы $C^\wedge$ и $C_\wedge$ ортогональны. Например, если $n = 2$ и

$$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

то

$$C^\wedge = \begin{pmatrix} a^2 & b^2 & -\sqrt{2}ab \\ b^2 & a^2 & \sqrt{2}ab \\ \sqrt{2}ab & -\sqrt{2}ab & a^2 - b^2 \end{pmatrix},$$

где $a^2 + b^2 = 1$. При этих же условиях $C_\wedge = 1$. Докажем, что матрицы $C^\wedge$ и $C_\wedge$ собственно ортогональны, то есть их определители положительны. В самом деле, так как матрица C собственно ортогональна, то ее можно непрерывной деформацией преобразовать в единичную матрицу. Итак, существуют такие непрерывные функции $c_{ij}(t)$, определенные при $0 \leqslant t \leqslant 1$, что $c_{ij}(0) = \delta_{ij}, c_{ij}(1) = c_{ij}$, и матрица $C(t) = (c_{ij}(t))$ ортогональна при любом t. Рассмотрим семейство преобразований (3.5), в которых элементы c_{ij} заменены на $c_{ij}(t)$. Тогда мы получим непрерывное семейство ортонормированных базисов $\mathbf{\Gamma}'_{ij}(t)$ и $\mathbf{\Delta}'_{ij}(t)$, причем, как нетрудно видеть,

$$\mathbf{\Gamma}'_{ij}(0) = \mathbf{\Gamma}_{ij}, \quad \mathbf{\Gamma}'_{ij}(1) = \mathbf{\Gamma}'_{ij}, \qquad \mathbf{\Delta}'_{ij}(0) = \mathbf{\Delta}_{ij}, \quad \mathbf{\Delta}'_{ij}(1) = \mathbf{\Delta}'_{ij}.$$

Значит, матрицы $C^\wedge$ и $C_\wedge$ могут быть непрерывно продеформированы в единичные матрицы соответствующих размерностей.

Таким образом, задание ориентации в пространстве $\mathcal{L}$ автоматически влечет задание ориентаций в пространствах $\mathbf{Sym}(n)$ и $\mathbf{so}(n)$. Положительно ориентированными мы будем считать базисы (3.3) и (3.4).

Размерности пространств $\mathbf{Sym}(n)$ и $\mathbf{so}(n)$ связаны соотношением $M + n = N$. Над пространством $\mathbf{Sym}(n)$ можно построить внешнюю алгебру

$$\bigwedge(\mathbf{Sym}(n)) = \bigoplus_{p=0}^{N} \bigwedge^p(\mathbf{Sym}(n)). \tag{3.6}$$

В этой алгебре нас будут интересовать два поливектора $\mathbf{X}_!$ и $\mathbf{X}^!$ степеней n и M соответственно, связанных с произвольным оператором $\mathbf{X} \in \mathbf{Sym}(n)$. Первый из них определяется как внешнее произведение степеней $\mathbf{X}^q$ оператора $\mathbf{X}$ при $q = 0, 1, \ldots, n-1$:

$$\mathbf{X}_! = \mathbf{X}^0 \wedge \mathbf{X}^1 \wedge \cdots \wedge \mathbf{X}^{n-1}.$$

Второй поливектор определяется несколько сложнее. Рассмотрим линейное отображение

$$\mathbf{X}^\square : \mathbf{so}(n) \longrightarrow \mathbf{Sym}(n),$$

задаваемое формулой

$$\mathbf{X}^\square(\mathbf{A}) = \mathbf{AX} - \mathbf{XA}.$$

Пусть $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_M$ – произвольный положительно ориентированный ортонормированный базис пространства $\mathbf{so}(n)$. Обозначим $\mathbf{\Psi}_r = \mathbf{X}^\square(\mathbf{\Phi}_r)$. Тогда по определению

$$\mathbf{X}^! = \mathbf{\Psi}_1 \wedge \mathbf{\Psi}_2 \wedge \cdots \wedge \mathbf{\Psi}_M.$$

Докажем корректность этого определения. Пусть $\mathbf{\Phi}'_1, \ldots, \mathbf{\Phi}'_M$ – другой положительно ориентированный ортонормированный базис пространства $\mathbf{so}(n)$, связанный с первым базисом собственно ортогональной матрицей перехода G. Тогда в силу линейности оператора $\mathbf{X}^\square$ системы векторов $\mathbf{\Psi}_1, \ldots, \mathbf{\Psi}_M$ и $\mathbf{\Psi}'_1, \ldots, \mathbf{\Psi}'_M$, где $\mathbf{\Psi}'_r = \mathbf{X}^\square(\mathbf{\Phi}'_r)$, будут связаны той же матрицей перехода G. А так как $\det G = 1$, то

$$\mathbf{\Psi}'_1 \wedge \cdots \wedge \mathbf{\Psi}'_M = (\det G)\mathbf{\Psi}_1 \wedge \cdots \wedge \mathbf{\Psi}_M = \mathbf{X}^!.$$

Очевидно, рассмотренные ранее матрицы $X_!$ и $X^!$ суть просто матрицы поливекторов $\mathbf{X}_!$ и $\mathbf{X}^!$, если X – матрица оператора $\mathbf{X}$, а в качестве базисных векторов $\mathbf{\Phi}_r$ в $\mathbf{so}(n)$ взяты операторы $\mathbf{\Delta}_{ij}$. Таким образом, миноры $\langle i_1 j_1, \ldots, i_n j_n \rangle(X)$ и $\{i_1 j_1, \ldots, i_n j_n\}(X)$ представляют собой координаты поливекторов $\mathbf{X}_!$ и $\mathbf{X}^!$ соответственно.

Оказывается, между поливекторами $\mathbf{X}_!$ и $\mathbf{X}^!$ есть очень простая связь. Напомним, что на слагаемых формулы (3.6) можно определить изоморфизм Ходжа

$$\mathcal{H}_p : \bigwedge^p(\mathbf{Sym}(n)) \longrightarrow \bigwedge^{N-p}(\mathbf{Sym}(n)),$$

связывающий слагаемые, имеющие дополнительные степени p и $N - p$. В частности, таковыми будут пространства кососимметрических тензоров степеней n и M. Нетрудно видеть, что при выбранной нами ориентации пространства $\mathbf{Sym}(n)$ справедливо соотношение

$$\mathcal{H}_n(\mathbf{\Gamma}_{11} \wedge \mathbf{\Gamma}_{22} \wedge \dots \wedge \mathbf{\Gamma}_{nn}) = \\ = \mathbf{\Gamma}_{12} \wedge \dots \wedge \mathbf{\Gamma}_{1n} \wedge \mathbf{\Gamma}_{23} \wedge \dots \wedge \mathbf{\Gamma}_{2n} \wedge \dots \wedge \mathbf{\Gamma}_{n-1,n}. \qquad (3.7)$$

Докажем, что изоморфизм Ходжа $\mathcal{H}_n$ переводит поливектор $\mathbf{X}_!$ в поливектор $\mathbf{X}^!$, то есть

$$\mathcal{H}_n(\mathbf{X}_!) = \mathbf{X}^!. \qquad (3.8)$$

Прежде всего, из (3.1), (3.3) и (3.4) можно вывести, что в случае $i < j$ справедливы равенства

$$\mathbf{\Gamma}_{ii}^{\Box}(\mathbf{\Delta}_{ij}) = -\mathbf{\Gamma}_{ij}, \quad \mathbf{\Gamma}_{jj}^{\Box}(\mathbf{\Delta}_{ij}) = \mathbf{\Gamma}_{ij}, \quad \mathbf{\Gamma}_{kk}^{\Box}(\mathbf{\Delta}_{ij}) = \mathbf{O}, \qquad (3.9)$$

где $k \neq i, j$.

Далее, выберем в $\mathcal{L}$ такой базис $\mathbf{e}_1, \dots, \mathbf{e}_n$, в котором матрица оператора $\mathbf{X}$ имеет диагональный вид:

$$X = \mathrm{diag}(\lambda_1, \dots, \lambda_n).$$

Тогда для любого $q \geqslant 0$ будет

$$\mathbf{X}^q = \lambda_1^q \mathbf{\Gamma}_{11} + \lambda_2^q \mathbf{\Gamma}_{22} + \dots + \lambda_n^q \mathbf{\Gamma}_{nn},$$

и, следовательно,

$$\mathbf{X}_! = W(\lambda_1, \dots, \lambda_n) \mathbf{\Gamma}_{11} \wedge \mathbf{\Gamma}_{22} \wedge \dots \wedge \mathbf{\Gamma}_{nn}.$$

Наконец, из (3.9) вытекает, что

$$\mathbf{X}^{\Box}(\mathbf{\Delta}_{ij}) = (\lambda_j - \lambda_i) \mathbf{\Gamma}_{ij},$$

откуда

$$\mathbf{X}^! = \prod_{i<j} (\lambda_j - \lambda_i) \times \mathbf{\Gamma}_{12} \wedge \dots \wedge \mathbf{\Gamma}_{1n} \wedge \mathbf{\Gamma}_{23} \wedge \dots \wedge \mathbf{\Gamma}_{2n} \wedge \dots \wedge \mathbf{\Gamma}_{n-1,n},$$

причем коэффициент при внешнем произведении является, очевидно, тем же самым определителем Вандермонда $W(\lambda_1, \dots, \lambda_n)$. Равенство (3.8) следует теперь из (3.7).

В произвольной системе координат поливекторы $\mathbf{X}_!$ и $\mathbf{X}^!$ имеют представления

$$\mathbf{X}_! = \sum \langle i_1 j_1, \dots, i_n j_n \rangle (X) \mathbf{\Gamma}_{i_1 j_1} \wedge \dots \wedge \mathbf{\Gamma}_{i_n j_n},$$

$$\mathbf{X}^! = \sum \{i_1j_1, \ldots, i_nj_n\}(X)\mathbf{\Gamma}_{k_1l_1} \wedge \cdots \wedge \mathbf{\Gamma}_{k_Ml_M},$$

где $k_1l_1, \ldots, k_Ml_M$ – множество двойных индексов, дополняющее набор $i_1j_1, \ldots, i_nj_n$ до ряда (1.5). Так как при изоморфизме Ходжа имеем

$$\mathcal{H}_n(\mathbf{\Gamma}_{i_1j_1} \wedge \cdots \wedge \mathbf{\Gamma}_{i_nj_n}) = \pm \mathbf{\Gamma}_{k_1l_1} \wedge \cdots \wedge \mathbf{\Gamma}_{k_Ml_M},$$

то равенство (2.7) следует из (3.8). Формула (2.7) дает основания называть дискриминантными многочленами также и миноры $\{i_1j_1, \ldots, i_nj_n\}(X)$.

§4. Доказательство основных теорем

Начнем с доказательства теоремы 1. Условимся называть регулярными те матрицы $X \in \mathrm{Sym}(n)$, у которых все собственные числа попарно различны. Множество всех регулярных матриц обозначим через $\mathbb{R}$. Все остальные симметрические матрицы назовем критическими.

Пусть X – регулярная матрица. Тогда существует такая ортогональная матрица C, что

$$C^{-1}XC = L = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \tag{4.1}$$

причем все значения $\lambda_1, \lambda_2, \ldots, \lambda_n$ попарно различны. По условию задачи

$$C^{-1}\mathcal{F}(X)C = M = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_n).$$

Диагональные матрицы $E, L, \ldots, L^{n-1}$ линейно независимы, поэтому диагональная матрица M единственным образом представима в виде линейной комбинации матриц $E, L, \ldots, L^{n-1}$:

$$M = a_0E + a_1L + \cdots + a_{n-1}L^{n-1},$$

откуда

$$\mathcal{F}(X) = a_0E + a_1X + \cdots + a_{n-1}X^{n-1}.$$

Таким образом, если матрица X регулярна, то числа $a_0, a_1, \ldots, a_{n-1}$ в выражении (0.1) единственным образом определяются через элементы X.

В дальнейшем под гладкостью какой-либо функции будем понимать ее бесконечную гладкость. Докажем, что функции $a_0(X), a_1(X), \ldots, a_{n-1}(X)$ являются гладкими на множестве $\mathbb{R}$. Элементы матрицы $\mathcal{F}(X)$ обозначим через $\theta_{ij}(X)$. Очевидно, для любой матрицы $X \in \mathbb{R}$ и при произвольном наборе двойных индексов $i_1j_1, i_2j_2, \ldots, i_nj_n$ справедливо равенство

$$\begin{pmatrix} x^{(0)}_{i_1j_1} & x^{(1)}_{i_1j_1} & \ldots & x^{(n-1)}_{i_1j_1} \\ x^{(0)}_{i_2j_2} & x^{(1)}_{i_2j_2} & \ldots & x^{(n-1)}_{i_2j_2} \\ \vdots & \vdots & \ddots & \vdots \\ x^{(0)}_{i_nj_n} & x^{(1)}_{i_nj_n} & \ldots & x^{(n-1)}_{i_nj_n} \end{pmatrix} \begin{pmatrix} a_0(X) \\ a_1(X) \\ \vdots \\ a_{n-1}(X) \end{pmatrix} = \begin{pmatrix} \theta_{i_1j_1}(X) \\ \theta_{i_2j_2}(X) \\ \vdots \\ \theta_{i_nj_n}(X) \end{pmatrix},$$

правая часть которого определена и гладка на всем множестве $\mathrm{Sym}(n)$. Если $i_r < j_r$, то r-ю строку равенства можно умножить на $\sqrt{2}$. Умножая после этого обе части равенства слева на матрицу, присоединенную к матрице, которая стоит в левой части, получим для $m = 0, 1, \ldots, n-1$ соотношения

$$\langle i_1j_1, \ldots, i_nj_n\rangle(X)a_m(X) = \varphi_{i_1j_1,\ldots,i_nj_n,m}(X), \tag{4.2}$$

в которых все $\varphi_{i_1j_1,\ldots,i_nj_n,m}$ определены и гладки на всем пространстве $\mathrm{Sym}(n)$.

Пусть $X_0 \in \mathbb{R}$. В точке X_0 матрицы $E, X, \ldots, X^{n-1}$ линейно независимы, то есть ранг матрицы $X_!$ в точности равен n. Значит, у $X_!$ есть невырожденный в точке X_0 минор $\langle i_1j_1, \ldots, i_nj_n\rangle(X)$. Очевидно, существует такая окрестность $\Omega \ni X_0$, что всюду в Ω имеем

$$\langle i_1j_1, \ldots, i_nj_n\rangle(X) \neq 0.$$

Отсюда и из (4.2) следует, что $a_m(X) \in C^\infty(\Omega)$ для всех m.

Очевидно, теперь нам нужно доказать, что функции $a_m(X)$ продолжимы по непрерывности на критическое множество и гладки на нем. Как мы видели, для любого $m = 0, 1, \ldots, n-1$ и любого набора двойных индексов $k_1l_1, \ldots, k_nl_n$ функция $f(X)$, определенная на множестве $\mathbb{R}$ равенством

$$f(X) = \langle k_1l_1, \ldots, k_nl_n\rangle(X)a_m(X),$$

может быть гладко продолжена на все пространство $\mathrm{Sym}(n)$. Следовательно, если из различных миноров $\langle k_1l_1, \ldots, k_nl_n\rangle(X)$ составить какую-нибудь линейную комбинацию

$$P(X) = \sum P_{k_1l_1,\ldots,k_nl_n}\langle k_1l_1, \ldots, k_nl_n\rangle(X),$$

в которой все $P_{k_1l_1,\ldots,k_nl_n}$ – константы, то функция $g(X) = P(X)a_m(X)$ будет гладкой на $\mathrm{Sym}(n)$. Пусть $Q(X)$ – какая-то другая линейная комбинация и $h(X) = Q(X)a_m(X)$. Очевидно,

$$h(X)P(X) = g(X)Q(X) \tag{4.3}$$

во всех точках множества $\mathbb{R}$. Но так как критические матрицы составляют в $\mathrm{Sym}(n)$ множество нулевой меры, а функции $g(X)$ и $h(X)$ гладки всюду, то равенство (4.3) верно на всем множестве $\mathrm{Sym}(n)$. Теперь, если бы удалось доказать, исходя из (4.3), что $g(X)$ делится на $P(X)$ в окрестности какой-нибудь критической точки X_0 и частное от деления является гладкой функцией, то очевидно, что функцию $a_m(X)$ можно было бы продолжить в окрестность точки X_0 с сохранением гладкости. Для обоснования возможности таких делений при, разумеется, специальном подборе многочленов $P(X)$ и $Q(X)$ нам понадобится теорема Мальгранжа [4], формулировка которой приводится ниже.

Теорема Мальгранжа. *Пусть Ω – открытое множество в R^N и $f_1, \ldots, f_r$ – аналитические в Ω функции. Пусть $\varphi \in C^\infty(\Omega)$. Тогда φ можно представить в виде*

$$\varphi = \sum_{i=1}^{r} f_i \psi_i,$$

где $\psi_i \in C^\infty(\Omega)$, в том и только в том случае, если для любого $a \in \Omega$ разложение Тейлора $T_a\varphi$ принадлежит идеалу, порожденному разложениями $T_a f_i$ в кольце формальных степенных рядов в точке a.

Как видно, в теореме Мальгранжа речь идет о делении гладких функций на аналитические. У нас в роли аналитических функций будут всегда выступать многочлены. Поэтому коснемся вопроса о делении формального степенного ряда на многочлен. Как известно, любой многочлен от нескольких переменных можно единственным образом представить в виде суммы однородных многочленов. Тот из них, который имеет наименьшую степень, обычно называется начальной формой данного многочлена.

Мы будем использовать следующее достаточное условие делимости. Пусть P_1 и P_2 – многочлены от нескольких переменных, R_1 и R_2 – формальные степенные ряды от тех же переменных, причем $R_1P_2 = R_2P_1$. Предположим, что начальные формы многочленов P_1 и P_2 взаимно просты. Нетрудно доказать, что тогда и сами многочлены P_1 и P_2 взаимно просты в кольце формальных степенных рядов. Но так как упомянутое кольцо факториально, то ряд R_1 делится на P_1, а ряд R_2 – на P_2.

В связи с этим встает вопрос о степенях начальных форм дискриминантных многочленов, разложенных в конечный ряд Тейлора в окрестности какой-либо критической точки X_0. Из (1.8) видно, что для решения этого вопроса имеет смысл найти порядок нуля дискриминанта $\mathrm{Dis}(\chi_X)$ в каждой критической точке.

Для решения поставленной задачи введем понятие корневого типа симметрической матрицы. Строку $(m_1, m_2, \ldots, m_k)$ назовем корневым типом матрицы $X \in \mathrm{Sym}(n)$, если у X ровно k различных собственных чисел $\lambda_1, \lambda_2, \ldots, \lambda_k$, причем каждое значение λ_i имеет кратность m_i. Разумеется, $m_1 + m_2 + \cdots + m_k = n$. Будем говорить, что два корневых типа $(m_1, m_2, \ldots, m_k)$ и $(r_1, r_2, \ldots, r_l)$ эквивалентны, если $k = l$ и во второй из этих строк можно так переставить числа, что она совпадет с первой.

Далее, скажем, что корневой тип $(r_1, r_2, \ldots, r_l)$ получается из корневого типа $(m_1, m_2, \ldots, m_k)$ посредством элементарной операции, если $l = k - 1$ и найдутся такие целые числа α, β, что $1 \leqslant \alpha < \beta \leqslant k$ и корневой тип

$$(m_1, \ldots, m_{\alpha-1}, m_\alpha + m_\beta, m_{\alpha+1}, \ldots, m_{\beta-1}, m_{\beta+1}, \ldots, m_k)$$

эквивалентен типу $(r_1, r_2, \ldots, r_l)$. Наконец, будем говорить, что один корневой тип предшествует другому, если этот другой может быть получен из первого посредством нескольких элементарных операций.

Очевидно, что корневой тип $(1, 1, \ldots, 1)$ регулярной матрицы предшествует любому другому, а корневому типу (n) скалярной матрицы предшествует всякий другой.

Предположим, что $X_0 \in \mathrm{Sym}(n)$ – критическая матрица корневого типа $(m_1, m_2, \ldots, m_k)$. Как мы уже говорили, нас интересует поведение линейных комбинаций дискриминантных многочленов в окрестности точки X_0 и порядок нуля дискриминанта $\mathrm{Dis}(\chi_X)$ в точке X_0. Приведем X_0 к диагональному виду преобразованием подобия $Y_0 = CX_0C^{-1}$. Преобразование (1.6) осуществляет соответствие между окрестностями точек X_0 и Y_0. Применительно к $\mathrm{Sym}(n)$ как к евклидову пространству размерности N это преобразование можно записать в виде $\widehat{X} = C^{\wedge}\widehat{Y}$, где $C^{\wedge}$ – ортогональная матрица порядка N. Из (1.6) следует, что для любого q имеем $X^q = C^{-1}Y^qC$ или, в другой записи, $\widehat{X^q} = C^{\wedge}\widehat{Y^q}$. Объединяя эти равенства при $q = 0, 1, \ldots, n-1$, получим формулу преобразования

$$X_! = C^\wedge Y_!.$$

Отсюда, а также из формулы Бине–Коши, следует, что каждый многочлен $\langle i_1 j_1, \ldots, i_n j_n \rangle (X)$ является линейной комбинацией с постоянными коэффициентами многочленов $\langle i_1 j_1, \ldots, i_n j_n \rangle (Y)$ и наоборот. Что же касается дискриминанта, то он инвариантен относительно преобразований подобия. Поэтому можно считать, что X_0 – диагональная матрица вида

$$X_0 = \begin{pmatrix} L_1 & & & \\ & L_2 & & \\ & & \ddots & \\ & & & L_k \end{pmatrix}, \quad \text{где} \quad L_\alpha = \begin{pmatrix} \lambda_\alpha & & \\ & \ddots & \\ & & \lambda_\alpha \end{pmatrix} \tag{4.4}$$

является скалярной матрицей размерности m_α. Теперь уже нетрудно доказать (с использованием, например, кругов Гершгорина [2]), что искомый порядок нуля равен 2ν, где

$$\nu = C^2_{m_1} + C^2_{m_2} + \cdots + C^2_{m_k}.$$

Отсюда и из (1.8) следует, что любой дискриминантный многочлен имеет в точке X_0 нуль порядка не менее ν.

Вернемся к доказательству теоремы. Итак, мы оценили снизу степени начальных форм дискриминантных многочленов, разложенных в ряд Тейлора в окрестности какой-нибудь критической точки. Теперь можно приступить к построению многочленов $P(X)$ и $Q(X)$ в равенстве (4.3) со взаимно простыми начальными формами.

Как мы уже делали раньше, разобьем множество переменных x_{ij} на две группы: диагональные переменные x_{ii} и недиагональные переменные x_{ij} с $i < j$. Теперь сделаем одно замечание. Пусть $R(Z, W)$ – однородный многочлен степени M от двух групп переменных $Z = (z_1, z_2, \ldots, z_k)$ и $W = (w_1, w_2, \ldots, w_l)$. Тогда его можно единственным образом представить в виде

$$R(Z, W) = \sum_{s+t=M} R_{s,t}(Z, W),$$

где каждый многочлен $R_{s,t}(Z, W)$ является однородным по переменным Z со степенью однородности s и по переменным W

со степенью однородности t. У одного из ненулевых многочленов $R_{s,t}(Z,W)$ значение s окажется наибольшим. Этот многочлен назовем старшей частью многочлена $R(Z,W)$ по переменным Z. Очевидно, что если у двух однородных многочленов $R_1(Z,W)$ и $R_2(Z,W)$ взаимно просты старшие части по Z, то и сами эти многочлены взаимно просты.

Старшую часть любого однородного многочлена $K(X)$ по диагональным переменным будем обозначать через $\Delta K(X)$. Каждый элемент $x_{ij}^{(m)}$ матрицы X^m является однородным многочленом степени m. Нетрудно доказать, что

$$\Delta x_{ii}^{(m)} = x_{ii}^m, \qquad \Delta x_{ij}^{(m)} = x_{ij}\frac{x_{jj}^m - x_{ii}^m}{x_{jj} - x_{ii}}. \tag{4.6}$$

Отсюда вытекает формула

$$\Delta\langle 11, 22, \ldots, nn\rangle(X) = W(x_{11}, x_{22}, \ldots, x_{nn}), \tag{4.7}$$

где W – определитель Вандермонда. Далее, пусть $1 \leqslant i < j \leqslant n$. Обозначим через $P_{ij}(X)$ определитель, который получается из определителя $\langle 11, 22, \ldots, nn\rangle(X)$, если в нем строку jj заменить строкой ij. Вследствие (4.6) имеем

$$\Delta P_{ij}(X) = \frac{x_{ij}}{x_{jj} - x_{ii}} W(x_{11}, x_{22}, \ldots, x_{nn}). \tag{4.8}$$

Теперь пусть X_0 – критическая матрица. Докажем, что коэффициенты $a_m(X)$ могут быть продолжены в некоторую окрестность точки X_0 с сохранением гладкости. Доказательство поведем индукцией по корневому типу матрицы X_0. Базой индукции послужит корневой тип $(1, 1, \ldots, 1)$ регулярной матрицы. Будем считать, что продолжение возможно в некоторые окрестности всех точек, корневой тип которых предшествует типу матрицы X_0.

В силу сказанного выше мы вправе предположить, что X_0 – диагональная матрица вида (4.4). Около точки X_0 выберем такую окрестность Ω, что все ее точки имеют корневой тип, предшествующий или эквивалентный типу X_0, и положим $X = X_0 + H$. Начальную форму тейлоровского разложения любого многочлена $K(X)$ в точке X_0 будем обозначать через $K_{X_0}(H)$. Далее, пару индексов (i, j) назовем допустимой, если найдется такое целое t, что

$$m_0 + m_1 + \cdots + m_t + 1 \leqslant i < j \leqslant m_0 + m_1 + \cdots + m_{t+1}$$

(мы считаем, что $m_0 = 0$, $0 \leqslant t \leqslant n-1$). Возьмем в равенстве (4.3)

$$Q(X) = \langle 11, 22, \ldots, nn\rangle(X).$$

Нетрудно видеть, что $Q(X)$ имеет в точке X_0 нуль порядка ν вида (4.5), причем в силу (4.7) старшая по диагональным переменным часть начальной формы $Q_{X_0}(H)$ имеет вид

$$\Delta Q_{X_0}(H) = B \prod (h_{jj} - h_{ii}),$$

где $B \neq 0$ – некоторая константа, зависящая только от X_0, а произведение берется по всем допустимым парам.

Точно так же, если (i, j) – допустимая пара, то для многочлена $P_{ij}(X)$ вблизи X_0 благодаря (4.8) имеем

$$\Delta P_{ij, X_0}(H) = \frac{h_{ij}}{h_{jj} - h_{ii}} \Delta Q_{X_0}(H).$$

Поэтому, если в равенстве (4.3) взять

$$P(X) = \sum P_{ij}(X),$$

где суммирование распространено на все допустимые пары, то получим

$$\Delta P_{X_0}(H) = \Delta Q_{X_0}(H) \sum \frac{h_{ij}}{h_{jj} - h_{ii}},$$

откуда видно, что многочлены $\Delta P_{X_0}(H)$ и $\Delta Q_{X_0}(H)$ взаимно просты. Но тогда взаимно просты и начальные формы $P_{X_0}(H)$ и $Q_{X_0}(H)$, так что в точке X_0 формальный ряд $T_{X_0} g(X)$ делится на $T_{X_0} P(X)$.

Далее, во всех точках окрестности Ω, корневой тип которых предшествует корневому типу матрицы X_0, по предположению индукции коэффициент $a_m(X)$ является гладкой функцией. Поэтому в них формальный ряд для $g(X)$ делится на разложение $P(X)$.

Наконец, в окрестности Ω остаются точки, корневой тип которых эквивалентен типу матрицы X_0. В этих точках по доказанному ранее многочлены $P(X)$ и $Q(X)$ имеют нули порядка не менее ν. Если же взять окрестность Ω достаточно малой, то формы степени ν многочленов $P(X)$ и $Q(X)$, разложенных в любой точке множества Ω, будут сколь угодно близки к соответствующим формам этих многочленов, разложенных в точке X_0. Поэтому, во-первых, $P(X)$ и $Q(X)$ имеют в точках эквивалентного корневого типа из Ω нули порядка в точности ν, а во-вторых, начальные формы разложений многочленов $P(X)$ и $Q(X)$ в этих точках взаимно просты. Поэтому и в них формальное разложение функции $g(X)$ делится на разложение многочлена $P(X)$.

Утверждение о взаимной простоте достаточно близких форм нуждается в пояснении. Пусть $\alpha_X(H)$ и $\beta_X(H)$ – формы степени ν от переменных H с коэффициентами, непрерывно зависящими от переменных X. Если бы при каком-то значении X они не были взаимно простыми, то нашлись бы такие ненулевые формы $\gamma(H)$ и $\delta(H)$ степени в точности $\nu - 1$, что

$$\alpha_X(H)\delta(H) = \beta_X(H)\gamma(H). \tag{4.9}$$

Последнее равенство есть система линейных уравнений на коэффициенты форм $\gamma(H)$ и $\delta(H)$. При $X = X_0$ эта система имеет только нулевое решение, то есть ее ранг не меньше числа неизвестных. Но матрица системы (4.9) непрерывно зависит от X. Поэтому при X, достаточно близких к X_0, ее ранг тоже не меньше числа неизвестных. Следовательно, система (4.9) имеет только нулевое решение, и формы $\alpha_X(H), \beta_X(H)$ взаимно просты.

Итак, в любой точке X некоторой окрестности $\Omega \ni X_0$ формальный ряд $T_X g$ принадлежит идеалу, порожденному рядом $T_X P$. По теореме Мальгранжа в области Ω можно разделить $g(X)$ на $P(X)$, причем частное будет гладкой функцией. Тем самым доказана продолжаемость любого коэффициента $a_m(X)$ в окрестность точки X_0 с сохранением гладкости. Так как X_0 – любая критическая точка, то функции $a_m(X)$ могут быть продолжены на все критическое множество. Теорема 1 полностью доказана. Разумеется, эта теорема остается справедливой и в случае, когда функция $\mathcal{F}$ определена не на всем пространстве $\mathrm{Sym}(n)$, а на некоторой его области.

Используя равенство (2.7), можно тем же способом доказать теорему 2. Отличие здесь будет только в доказательстве базы индукции, которое мы сейчас и проведем. Пусть матрица X регулярна. Тогда существует такая ортогональная матрица C, что справедливо равенство (4.1). Пусть

$$C^{-1}\mathcal{G}(X)C = T = (t_{ij}),$$

причем все t_{ii} равны нулю. Тогда можно подобрать такую кососимметрическую матрицу B, что

$$T = BL - LB. \tag{4.10}$$

Именно, элементы b_{ij} матрицы B однозначно определяются из соотношения $t_{ij} = (\lambda_j - \lambda_i)b_{ij}$.

Подвергая (4.10) преобразованию подобия, получим (0.2), где $A(X) = CBC^{-1}$. Таким образом, на регулярном множестве $\mathbb{R}$ матрица $A(X)$ существует и однозначна. Так как соотношение (0.2) может быть записано в матрично-векторной форме

$$\widehat{\mathcal{G}(X)} = X^! \overline{A(X)},$$

то продолжаемость элементов $a_{ij}(X)$ матрицы $A(X)$ на критическое множество доказывается так же, как и прежде, с использованием дискриминантных многочленов $\{i_1 j_1, \ldots, i_n j_n\}(X)$.

Автор благодарен своему учителю А. С. Мищенко, а также П. П. Мосолову, А. Т. Фоменко, С. В. Дужину и В. И. Арнольду за постановку задач и внимание к работе.

Литература

1. А. И. Лурье, *Теория упругости*, Наука, М. (1970).
2. Ф. Р. Гантмахер, *Теория матриц*, Наука, М. (1988).
3. М. М. Постников, *Лекции по геометрии. Семестр II. Линейная алгебра*, Наука, М. (1986).
4. Мальгранж, *Идеалы дифференцируемых функций*, Мир, М. (1968).
5. Н. В. Илюшечкин, *Об одном классе гладких матричнозначных функций*, — УМН **40**, No. 1 (1985), 201–202.
6. Н. В. Илюшечкин, *Дискриминант характеристического многочлена нормальной матрицы*, — Матем. заметки **51**, No. 3 (1992), 16–23.
7. Н. В. Илюшечкин, *О некоторых классах гладких преобразований пространства симметрических матриц*, — Матем. сборник **184**, No. 9 (1993), 89–102.

Ilyushechkin N. V. On some identities for the elements of a symmetric matrix.

Let Sym(n) be the space of n-dimensional real symmetric matrices, and let $X \in \mathrm{Sym}(n)$. The matrices $E, X, X^2, \ldots, X^{n-1}$ can be regarded as vectors of Euclidean space of dimension n^2. Denote by $V(E, X, \ldots, X^{n-1})$ the volume of the parallelepiped built on these vectors. It is proved that

$$V^2(E, X, \ldots, X^{n-1}) = D(X),$$

where D(X) is the discriminant of the characteristic polynomial of the matrix X. Two classes of smooth maps of the space $\mathrm{Sym}(n)$ are described.

Федеральное Государственное
Унитарное Предприятие
Научно-Производственное
Объединение "Агат"

E-mail: nikita@ilyush.mccme.ru

Поступило 21 мая 2003 г.