Б. Б. Комраков

МАТРИЧНЫЙ АНАЛИЗ КУРС ЛЕКЦИЙ

Минск
БГУ
2005

ПРЕДИСЛОВИЕ

В настоящее время теория матриц находит широкое применение в любой области современной математики и во многих приложениях.

Учебное пособие составлено на основе курса лекций по матричному анализу, который автор читал несколько лет студентам специальности "Прикладная математика"факультета прикладной математики и информатики Белорусского государственного университета.

Пособие позволит студентам подготовиться к сдаче экзамена по матричному анализу и снимет необходимость использования большого количества дополнительной литературы.

Следует отметить, что кроме теоретического материала пособие содержит также и большое количество примеров решения типовых задач и может использоваться для подготовки к практическим занятиям.

Для успешного усвоения материала пособия необходимо знание линейной алгебры в объеме программы первого курса ФПМИ БГУ.

Пособие может быть полезно также студентам любых естественнонаучных специальностей, требующих навыков в решении практических задач матричного анализа.

При написании пособия были использованы некоторые материалы из книг [1], [2], [3].

1. ПСЕВДООБРАТНАЯ МАТРИЦА

1.1. Скелетное разложение матрицы

Лемма 1. *Для любой матрицы* $A \in P_{m,n}$ ($\mathbb{C}_{m,n}$ *или* $\mathbb{R}_{m,n}$) *верно*

$$\operatorname{rank} A^*A = \operatorname{rank} A,$$
$$\operatorname{rank} AA^* = \operatorname{rank} A,$$

где $A^* = \bar{A}^T$.

Доказательство. Напомним, что матрицей Грама системы векторов $\{g_1, \ldots, g_n\}$ называется следующая матрица:

$$A_G = \begin{bmatrix} (g_1, g_1) & \ldots & (g_1, g_n) \\ \multicolumn{3}{c}{\dotfill} \\ (g_n, g_1) & \ldots & (g_n, g_n) \end{bmatrix},$$

причем $\operatorname{rank} A_G = \operatorname{rank} G$.

Рассмотрим $(\bar{A}_1, \ldots, \bar{A}_n)$ — систему столбцов матрицы $\bar{A}$. Пусть $M = A^*A$, тогда $m_{ij} = \bar{A}_i^T \cdot A_j = (\bar{A}_i, \bar{A}_j)$, то есть A^*A — матрица Грама системы векторов $(\bar{A}_1, \ldots, \bar{A}_n)$, следовательно,

$$\operatorname{rank} A^*A = \operatorname{rank}(\bar{A}_1, \ldots, \bar{A}_n) = \operatorname{rank} \bar{A} = \operatorname{rank} A.$$

Аналогично доказывается второе утверждение. □

Теорема 1. *(О скелетном разложении матрицы.) Пусть* $A \in P_{m,n}$, $\operatorname{rank} A = r > 0$, *тогда существуют* $B \in P_{m,r}$, $C \in P_{r,n}$, *такие, что* $A = B \cdot C$ *и* $\operatorname{rank} B = \operatorname{rank} C = r$.

Доказательство. Выберем произвольные r линейно независимых столбцов матрицы A и объявим их столбцами матрицы

$$B = (B_1, \ldots, B_r),$$

$B_1, \ldots, B_r$ линейно независимы и образуют базис системы столбцов матрицы $A = (A_1, \ldots, A_n)$.

Тогда

$$A_i = c_{1i}B_1 + \dots + c_{ri}B_r, \qquad i = \overline{1,n}.$$

Значит,

$$(A_1, \dots, A_n) = (B_1, \dots, B_r) \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \vdots & \vdots & \dots & \vdots \\ c_{r1} & c_{r2} & \dots & c_{rn} \end{bmatrix}.$$

Таким образом, мы получили $A = B \cdot C$, где C — матрица перехода от базиса $(B_1, \dots, B_r)$ к системе векторов $(A_1, \dots, A_n)$.

Значит,

$$\operatorname{rank} C = \operatorname{rank}(A_1, \dots, A_n) = \operatorname{rank} A = r.$$

□

Замечание. Скелетное разложение матрицы определено в общем случае неоднозначно.

Замечание. Если матрица $A \in P_{m,n}$ имеет ранг, равный n (или m), то в качестве матрицы B удобно взять саму матрицу A (матрицу E_m), а в качестве матрицы C — матрицу E_n (матрицу A).

1.2. Определение псевдообратной матрицы

Пусть $A \in P_{m,n}$, тогда псевдообратной для матрицы A называется матрица $A^+ \in P_{n,m}$, такая, что будут выполняться следующие равенства:

$$AA^+A = A, \quad A^+AA^+ = A^+,$$
$$(A^+A)^* = A^+A, \quad (AA^+)^* = AA^+.$$

Теорема 2. *Для любой матрицы $A \in P_{m,n}$ существует псевдообратная матрица A^+.*

Доказательство. Рассмотрим произвольное скелетное разложение матрицы A: $A = BC$.

Заметим, что

$$B^*B \in P_{r,r}, \ \operatorname{rank} B^*B = \operatorname{rank} B = r (\text{согласно лемме 1}) \Rightarrow \exists (B^*B)^{-1},$$
$$C \cdot C^* \in P_{r,r}, \ \operatorname{rank} CC^* = \operatorname{rank} C = r \Rightarrow \exists (CC^*)^{-1}.$$

Покажем, что матрица $A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$ является псевдообратной матрицей для матрицы A. Для этого проверим условия псевдообратной матрицы:

$$\begin{aligned}
AA^+A &= BCC^*(CC^*)^{-1}(B^*B)^{-1}B^*BC = BC = A,\\
A^+AA^+ &= C^*(CC^*)^{-1}(B^*B)^{-1}B^*BCC^*(CC^*)^{-1}(B^*B)^{-1}B^* =\\
&= C^*(CC^*)^{-1}(B^*B)^{-1}B^* = A^+,\\
(A^+A)^* &= (C^*(CC^*)^{-1}(B^*B)^{-1}B^*BC)^* =\\
&= C^*((CC^*)^{-1})^*C = C^*(CC^*)^{-1}C =\\
&= C^*(CC^*)^{-1}(B^*B)^{-1}B^*BC = A^+A.
\end{aligned}$$

Аналогично $(AA^+)^* = AA^+$. □

Следствие. *Если столбцы матрицы $A \in P_{m,n}$ линейно независимы, то $A^+ = (A^*A)^{-1}A^*$; если строки матрицы $A \in P_{m,n}$ линейно независимы, то $A^+ = A^*(AA^*)^{-1}$.*

Доказательство. Пусть столбцы матрицы A линейно независимы, значит, $\operatorname{rank} A = n \Rightarrow$

$$B = A,\ C = E_n \text{ (см. замечание 2 к теореме 1)}, \quad A = B \cdot E_n = A \cdot E_n.$$

Тогда

$$A^+ = E_n^*(E_nE_n^*)^{-1}(A^*A)^{-1}A^* = (A^*A)^{-1}A^*.$$

Утверждение про строки доказывается аналогично. □

Заметим, что теорема 2 дает нам способ вычисления псевдообратной матрицы A^+ для любой ненулевой матрицы A, но не отвечает на вопрос о единственности A^+ (напомним, что скелетное разложение матрицы определено неоднозначно).

Теорема 3. *Для любой матрицы $A \in P_{m,n}$ псевдообратная матрица A^+ определена однозначно.*

Доказательство. Пусть для некоторой матрицы A существуют две псевдообратные матрицы: A_1^+ и A_2^+.

Тогда верны следующие равенства:

$$\begin{aligned}
&AA_1^+A = AA_2^+A = A,\\
&\quad A_1^+ = A_1^+AA_1^+ = A_1^+(AA_1^+)^* = A_1^+(A_1^+)^*A^*,\\
&\quad A_2^+ = A_2^+(A_2^+)^*A^*,
\end{aligned}$$

$$A_1^+ = A_1^+ A A_1^+ = (A_1^+ A)^* A_1^+ = A^*(A_1^+)^* A_1^+,$$
$$A_2^+ = A^*(A_2^+)^* A_2^+.$$

Введем новые матрицы:

$$B = A_2^+ - A_1^+,$$
$$C = A_2^+(A_2^+)^* - A_1^+(A_1^+)^*,$$
$$D = (A_2^+)^* A_2^+ - (A_1^+)^* A_1^+.$$

Заметим, что

$$ABA = AA_2^+A - AA_1^+A = A - A = 0,$$
$$B = CA^* = A^*D.$$

Тогда

$$(BA)^*(BA) = A^*B^*BA = A^*(A^*D)^*BA = A^*D^*ABA = 0.$$

Согласно лемме 1

$$\operatorname{rank}(BA)^*(BA) = 0 \Rightarrow \operatorname{rank}(BA) = 0 \Rightarrow BA = 0.$$

Рассмотрим

$$BB^* = B(CA^*) = BAC^* = 0 \Rightarrow \operatorname{rank} BB^* = \operatorname{rank} B = 0 \Rightarrow B = 0 \Rightarrow$$
$$A_2^+ - A_1^+ = 0 \Rightarrow A_2^+ = A_1^+,$$

следовательно, псевдообратная матрица определена однозначно. □

Следствие. *Если матрица A квадратная и невырожденная, то псевдообратная матрица A^+ совпадает с обратной матрицей A^{-1}.*

Доказательство. Пусть $A \in P_{n,n}$ и $\det A \neq 0$. Тогда

$$A^+ = (A^*A)^{-1}A^* = A^{-1}(A^*)^{-1}A^* = A^{-1},$$

причем A^+ определена однозначно. □

Укажем некоторые свойства псевдообратной матрицы:

1) $(A^+)^* = (A^*)^+$;
2) $(A^+)^+ = A$;
3) $(A^+A)^2 = A^+A$;
4) $(AA^+)^2 = AA^+$.

Пример. Найдем псевдообратную матрицу A^+ для матрицы

$$A = \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & -2 & 0 \\ 2 & 2 & -4 & 2 \end{bmatrix}.$$

Найдем $\operatorname{rank} A$:

$$A \sim \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & -2 & 0 \end{bmatrix} \quad \Rightarrow \quad \operatorname{rank} A = 2.$$

В качестве столбцов матрицы B возьмем первые два столбца матрицы A:

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 2 \end{bmatrix}.$$

Найдем матрицу C. Для этого решим уравнение $BC = A$:

$$\left[\begin{array}{cc|cccc} 1 & 0 & 1 & 0 & -1 & 1 \\ 0 & 2 & 0 & 2 & -2 & 0 \\ 2 & 2 & 2 & 2 & -4 & 2 \end{array}\right] \sim \left[\begin{array}{cc|cccc} 1 & 0 & 1 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 0 \end{array}\right] \Rightarrow$$

$$\Rightarrow C = \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix}, \quad C^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \\ 1 & 0 \end{bmatrix}.$$

Тогда

$$CC^* = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad (CC^*)^{-1} = \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}.$$

$$C^*(CC^*)^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} = \frac{1}{5}\begin{pmatrix} 2 & -1 \\ -1 & 3 \\ -1 & -2 \\ 2 & -1 \end{pmatrix}.$$

$$B^*B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 8 \end{bmatrix},$$

$$(B^*B)^{-1} = \frac{1}{24}\begin{bmatrix} 8 & -4 \\ -4 & 5 \end{bmatrix}.$$

$$(B^*B)^{-1}B^* = \frac{1}{24}\begin{pmatrix} 8 & -4 \\ -4 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 2 \end{pmatrix} = \frac{1}{12}\begin{pmatrix} 4 & -4 & 4 \\ -2 & 5 & 1 \end{pmatrix}.$$

$$A^+ = \frac{1}{60}\begin{bmatrix} 2 & -1 \\ -1 & 3 \\ -1 & -2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 4 & -4 & 4 \\ -2 & 5 & 1 \end{bmatrix} = \frac{1}{60}\begin{bmatrix} 10 & -13 & 7 \\ -10 & 19 & -1 \\ 0 & -6 & -6 \\ 10 & -13 & 7 \end{bmatrix}.$$

1.3. Нормальное псевдорешение системы линейных уравнений

Пусть есть система m линейных уравнений с n неизвестными, которой соответствует матричное уравнение

$$AX = B, \quad \text{где} \quad A \in P_{m,n}, \quad X \in P_{n,1} \quad B \in P_{m,1}.$$

Столбец $Y = B - AX$ называется невязкой столбца X. Если X — решение системы, невязка равна нулю, если система несовместна, попытаемся найти столбец X, длина невязки которого минимальна. Такой столбец называется псевдорешением системы $AX = B$. Псевдорешение минимальной длины называется нормальным псевдорешением системы $AX = B$.

Теорема 4. *Нормальное псевдорешение системы $AX = B$ всегда существует, единственно и вычисляется по формуле $X_0 = A^+B$.*

Доказательство. Нужно показать, что $|B - AX_0| \leq |B - AX|$, $\forall X \in P_{n,1}$, тогда X_0 — псевдорешение, кроме того, $|X_0| \leq |X|$, $\forall X \in P_{n,1}$, такого, что $|B - AX_0| = |B - AX|$.

Рассмотрим произвольный столбец X.

$$B - AX = B - AX_0 + AX_0 - AX =$$
$$= B - AX_0 + A(X_0 - X) = T + S,$$

где

$$T = B - AX_0, \qquad S = A(X - X_0).$$

Рассмотрим

$$|B - AX|^2 = (T + S)^* \cdot (T + S) =$$
$$= T^*T + T^*S + S^*T + S^*S = |T|^2 + T^*S + S^*T + |S|^2.$$

$$T^*S = (B - AX_0)^*A(X_0 - X) = (B - AA^+B)^*A(X_0 - X) =$$

$$
\begin{aligned}
&= ((E_m - AA^+)B)^* A(X_0 - X) = \\
&= B^*(E_m - AA^+)^* A(X_0 - X) = \\
&= B^*(E_m - AA^+) A(X_0 - X) = \\
&= B^*(A - AA^+A)(X_0 - X) = \\
&= B^*(A - A)(X_0 - X) = 0.
\end{aligned}
$$

Доказали, что $T^*S = 0$. Но тогда $S^*T = ((S^*T)^*)^* = (T^*S)^* = 0^* = 0$. Таким образом, $S^*T = T^*\ S = 0$.

Получили, что $|B - AX|^2 = |T|^2 + |S|^2 = |B - AX_0|^2 + |A(X_0 - X)|^2$.

Значит, $|B - AX_0| \leq |B - AX|$, $\forall X \in P_{n,1}$. Мы показали, что X_0— псевдорешение системы $AX = B$.

Докажем, что псевдорешение X_0 — нормальное.

Пусть X — столбец, такой, что $|B - AX_0| = |B - AX|$. Тогда $A(X_0 - X) = 0$.

Пусть $Z = X - X_0$.

Тогда

$$X = X_0 + Z.$$

Заметим, что $AZ = 0$.

Тогда

$$
\begin{aligned}
|X|^2 &= (X_0 + Z)^*(X_0 + Z) = X_0^*X_0 + X_0^*Z + Z^*X_0 + Z^*Z = \\
&= |X_0|^2 + X_0^*Z + Z^*X_0 + |Z|^2. \\
X_0^*Z &= (A^+B)^*Z = (A^+AA^+B)^*Z = (A^+B)^*(A^+A)^*Z = \\
&= (A^+B)^*A^+AZ = 0 \ \ (\text{так как } \ AZ = 0) \Rightarrow X_0^*Z = 0.
\end{aligned}
$$

Аналогично: $Z^*X_0 = ((Z^*X_0)^*)^* = (X_0^*Z)^* = 0$.

Получили, что $|X|^2 = |X_0|^2 + |Z|^2 \Rightarrow |X_0| \leq |X| \Rightarrow X_0$ — нормальное псевдорешение системы $AX = B$. $\square$

Пример. Найдем нормальное псевдорешение системы линейных уравнений:

$$
\begin{cases}
x_1 - 2x_2 + x_3 = 1 \\
-2x_1 + 4x_2 - 2x_3 = 1 \\
x_1 - x_3 = 1.
\end{cases}
$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 1\\ -2 & 4 & -2 & 1\\ 1 & 0 & -1 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -2 & 1 & 1\\ 0 & 0 & 0 & 3\\ 1 & 0 & -1 & 1 \end{array}\right] \Rightarrow$$

система несовместна, $\operatorname{rank} A = 2$.

Здесь

$$A = \begin{pmatrix} 1 & -2 & 1\\ -2 & 4 & -2\\ 1 & 0 & -1 \end{pmatrix}, \qquad S = \begin{pmatrix} 1\\ 1\\ 1 \end{pmatrix}.$$

Найдем A^+. Для этого найдем скелетное разложение $A = B^*C$:

$$B = \begin{pmatrix} 1 & -2\\ -2 & 4\\ 1 & 0 \end{pmatrix}.$$

Найдем C:

$$\left[\begin{array}{cc|ccc} 1 & -2 & 1 & -2 & 1\\ -2 & 4 & -2 & 4 & -2\\ 1 & 0 & 1 & 0 & -1 \end{array}\right] \sim \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 0 & -1\\ 0 & 2 & 0 & 2 & -2 \end{array}\right] \sim \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 0 & -1\\ 0 & 1 & 0 & 1 & -1 \end{array}\right].$$

Отсюда $C = \begin{bmatrix} 1 & 0 & -1\\ 0 & 1 & -1 \end{bmatrix}$.

$$A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*.$$

$$CC^* = \begin{bmatrix} 1 & 0 & -1\\ 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0\\ 0 & 1\\ -1 & -1 \end{bmatrix} = \begin{pmatrix} 2 & 1\\ 1 & 2 \end{pmatrix}.$$

$$(CC^*)^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1\\ -1 & 2 \end{bmatrix}.$$

$$C^*(CC^*)^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 0\\ 0 & 1\\ -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1\\ -1 & 2 \end{bmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1\\ -1 & 2\\ -1 & -1 \end{pmatrix}.$$

$$B^*B = \begin{pmatrix} 1 & -2 & 1\\ -2 & 4 & 0 \end{pmatrix} \cdot \begin{bmatrix} 1 & -2\\ -2 & 4\\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & -10\\ -10 & 20 \end{bmatrix},$$

$$(B^*B)^{-1} = \frac{1}{20}\begin{bmatrix} 20 & 10\\ 10 & 6 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 10 & 5\\ 5 & 3 \end{bmatrix}.$$

$$(B^*B)^{-1}B^* = \frac{1}{10}\begin{bmatrix}10 & 5\\ 5 & 3\end{bmatrix}\cdot\begin{bmatrix}1 & -2 & 1\\ -2 & 4 & 0\end{bmatrix} = \frac{1}{10}\begin{bmatrix}0 & 0 & 10\\ -1 & 2 & 5\end{bmatrix}.$$

Получаем, что

$$A^+ = \frac{1}{30}\begin{bmatrix}1 & -2 & 15\\ -2 & 4 & 0\\ 1 & -2 & -15\end{bmatrix}.$$

Тогда нормальное псевдорешение

$$X_0 = A^+S = \frac{1}{30}\begin{bmatrix}1 & -2 & 15\\ -2 & 4 & 0\\ 1 & -2 & -15\end{bmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{1}{30}\begin{pmatrix}14\\2\\-16\end{pmatrix} = \frac{1}{15}\begin{pmatrix}7\\1\\-8\end{pmatrix}.$$

$$X_0 = \frac{1}{15}(7,1,-8).$$

Найдем длину невязки:

$$Y = S - AX_0 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1}{15}\begin{pmatrix}1 & -2 & 1\\ -2 & 4 & -2\\ 1 & 0 & -1\end{pmatrix}\begin{pmatrix}7\\1\\-8\end{pmatrix} =$$

$$= \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1}{15}\begin{pmatrix}-3\\6\\15\end{pmatrix} =$$

$$= \frac{1}{5}\begin{pmatrix}5\\5\\5\end{pmatrix} - \frac{1}{5}\begin{pmatrix}-1\\2\\5\end{pmatrix} = \frac{1}{5}\begin{pmatrix}6\\3\\0\end{pmatrix}.$$

$$Y = \frac{1}{5}(6,3,0) = \frac{3}{5}(2,1,0).$$

$$|Y| = \frac{3\sqrt{5}}{5}.$$

2. ФУНКЦИИ ОТ МАТРИЦ

2.1. Функции от матричного аргумента

Пусть имеется матрица $A \in C_{n,n}$ и функция $f : \mathbb{C} \to \mathbb{C}$. Требуется определить $f(A)$.

В частном случае, когда $f(x) = \alpha_m x^m + \alpha_{m-1}x^{m-1} + \dots + \alpha_1 x + \alpha_0$ — многочлен относительно x, $f(A)$ мы находим непосредственной подстановкой:

$$f(A) = \alpha_m A^m + \alpha_{m-1}A^{m-1} + \dots + \alpha_1 A + \alpha_0 E_n.$$

Пусть $\psi(\lambda)$— минимальный многочлен матрицы A.

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}, \;\; m_1 + m_2 + \dots + m_s = m,$$

где $m = \deg \psi$, $\lambda_1, \dots, \lambda_s$ — все различные собственные значения матрицы A.

Пусть имеются два многочлена $g(\lambda)$ и $h(\lambda)$, такие, что $g(A) = h(A)$, тогда

$$d(\lambda) = g(\lambda) - h(\lambda)$$

— аннулирующий многочлен матрицы A (так как $d(A) = g(A) - h(A) = 0$) и, значит, делится на $\psi(\lambda)$ без остатка.

Отсюда следует, что

$$d(\lambda_k) = 0,\; d'(\lambda_k) = 0, \dots, d^{(m_k-1)}(\lambda_k) = 0, \;\; k = \overline{1, s}.$$

Значит,

$$g(\lambda_k) = h(\lambda_k),\; g'(\lambda_k) = h'(\lambda_k), \dots, g^{(m_k-1)}(\lambda_k) = g^{(m_k-1)}(\lambda_k), \;\; k = \overline{1, s}.$$

Множество, состоящее из m чисел

$$f(\Lambda_A) = \{f(\lambda_k), f'(\lambda_k), \dots f^{(m_k-1)}(\lambda_k) \mid k = \overline{1, s}\},$$

будем называть значениями функции f на спектре матрицы A. Если все они существуют, то будем говорить, что функция f определена на спектре матрицы A. Если функция f не определена на спектре матрицы A, то не определено и $f(A)$.

Мы показали, что если $g(A) = h(A)$, то $g(\Lambda_A) = h(\Lambda_A)$. Нетрудно заметить, что верно и обратное, то есть если $g(\Lambda_A) = h(\Lambda_A)$, то $g(A) = h(A)$.

Отсюда следует, что значения многочлена g на спектре матрицы A однозначно определяют матрицу $g(A)$, то есть все многочлены, значения которых на спектре матрицы A совпадают с $g(\Lambda_A)$, будут иметь матричное значение $g(A)$.

Таким же образом определим и произвольную функцию матричного аргумента: значения функции $f(\lambda)$ на спектре матрицы A должны полностью определять $f(A)$, то есть все функции, значения которых на спектре матрицы A совпадают, должны иметь одно и то же матричное значение $f(A)$. Но тогда, для того чтобы найти $f(A)$, достаточно найти любой многочлен $g(\lambda)$, который на спектре матрицы A принимал бы те же значения, что и функция f. Тогда $f(A) = g(A)$.

Определение 1. Пусть функция f определена на спектре матрицы A, тогда $f(A) = g(A)$, где $g(\lambda)$ — любой многочлен, такой, что $f(\Lambda_A) = g(\Lambda_A)$.

Но всегда можно добиться того, что степень многочлена, с помощью которого мы определяем $f(A)$, будет меньше m. Действительно, разделим $g(\lambda)$ на $\psi(\lambda)$:

$$g(\lambda) = \psi(\lambda)q(\lambda) + r(\lambda), \text{ где } \deg r < m.$$

Тогда

$$g(A) = \psi(A)q(A) + r(A), \text{ но } \psi(A) = 0,$$

следовательно, $g(A) = r(A)$, значит, $f(A) = r(A)$.

Существует ровно один многочлен, такой, что $f(\Lambda_A) = r(\Lambda_A)$ и $\deg r < m$, который определяется интерполяционными условиями:

$$f(\lambda_k) = r(\lambda_k),\ f'(\lambda_k) = r'(\lambda_k), \ldots,\ f^{(m_k-1)}(\lambda_k) = r^{(m_k-1)}(\lambda_k),\ k = \overline{1, s}.$$

Этот многочлен называется *интерполяционный многочлен Лагранжа – Сильвестра.* Тогда мы можем дать новое определение $f(A)$.

Определение 2. Пусть функция f определена на спектре матрицы A, тогда $f(A) = r(A)$, где $r(\lambda)$ — интерполяционный многочлен Лагранжа – Сильвестра.

Пример. Рассмотрим матрицу

$$H_n = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \quad H_n \in \mathbb{C}_{n,n}.$$

Несложно проверить, что

$$H_n^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}, \dots, \; H_n^{n-1} = \begin{bmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 \end{bmatrix}, \; H_n^n = 0.$$

$\psi(\lambda) = \lambda^n$ — минимальный многочлен матрицы H_n.

Используя интерполяционные условия:

$$r(0) = f(0), \; r'(0) = f'(0), \dots, r^{(n-1)}(0) = f^{(n-1)}(0),$$

находим многочлен Лагранжа – Сильвестра:

$$r(\lambda) = f(0) + \frac{f'(0)}{1!}\lambda + \frac{f''(0)}{2!}\lambda^2 + \dots + \frac{f^{(n-1)}(0)}{(n-1)!}\lambda^{n-1}.$$

Тогда

$$f(H_n) = r(H_n) = f(0)E_n + \frac{f'(0)}{1!}H_n +$$

$$+ \frac{f''(0)}{2!}H_n^2 + \dots + \frac{f^{(n-1)}(0)}{(n-1)!}H_n^{n-1} =$$

$$= \begin{bmatrix} f(0) & \frac{f'(0)}{1!} & \frac{f''(0)}{2!} & \dots & \frac{f^{(n-2)}(0)}{(n-2)!} & \frac{f^{(n-1)}(0)}{(n-1)!} \\ 0 & f(0) & \frac{f'(0)}{1!} & \dots & \frac{f^{(n-3)}(0)}{(n-3)!} & \frac{f^{(n-2)}(0)}{(n-2)!} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & f(0) & \frac{f'(0)}{1!} \\ 0 & 0 & 0 & \dots & 0 & f(0) \end{bmatrix}.$$

В частном случае, при $n = 4$:

$$f(H_4) = \begin{bmatrix} f(0) & f'(0) & \frac{1}{2}f''(0) & \frac{1}{6}f'''(0) \\ 0 & f(0) & f'(0) & \frac{1}{2}f''(0) \\ 0 & 0 & f(0) & f'(0) \\ 0 & 0 & 0 & f(0) \end{bmatrix}.$$

Найдем $\cos H$:

$$\begin{aligned} f(\lambda) &= \cos\lambda, \\ f'(\lambda) &= -\sin\lambda, \\ f''(\lambda) &= -\cos\lambda, \\ f'''(\lambda) &= \sin\lambda, \\ \cos H &= \begin{bmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Найдем e^H:

$$e^H = \begin{bmatrix} 1 & 1 & \frac{1}{2} & \frac{1}{6} \\ 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Свойства функции от матрицы

(1) Пусть $\lambda_1, \ldots, \lambda_n$ — все собственные значения матрицы $A \in \mathbb{C}_{n,n}$, тогда
$f(\lambda_1), \ldots, f(\lambda_n)$ — собственные значения $f(A)$.

Доказательство. Сначала докажем это свойство для многочлена. Запишем характеристический многочлен матрицы A:

$$\Delta(\lambda) = \det(A - \lambda E_n) = (\lambda - \lambda_1)\ldots(\lambda - \lambda_n).$$

Пусть дан многочлен $g(\mu)$. Разложим его на неприводимые множители над $\mathbb{C}$:

$$g(\mu) = \alpha_l(\mu - \mu_1)\ldots(\mu - \mu_l).$$

Подставим в g матрицу A:

$$g(A) = \alpha_l(A - \mu_1 E_n)\ldots(A - \mu_l E_n).$$

Для удобства вычислений будем считать, что многочлены определены с точностью до знака, то есть $g(\mu) = -g(\mu)$ (на их корни это не влияет). Тогда

$$\det g(A) = \alpha_l^n \det(A - \mu_1 E_n), \ldots, \det(A - \mu_l E_n) = \alpha_l^n \Delta(\mu_1) \ldots \Delta(\mu_l) =$$

$$= \alpha_l^n \prod_{i=1}^{l} \prod_{k=1}^{n} (\mu_i - \lambda_k) = \alpha_l^n \prod_{k=1}^{n} \prod_{i=1}^{l} (\mu_i - \lambda_k) = g(\lambda_1) g(\lambda_2) \ldots g(\lambda_n).$$

Мы показали, что

$$\det g(A) = g(\lambda_1) g(\lambda_2) \ldots g(\lambda_n).$$

Заменив в этом равенстве $g(\mu)$ на $g(\mu) - \lambda$, получаем, что

$$\det(g(A) - \lambda E_n) = (g(\lambda_1) - \lambda)(g(\lambda_2) - \lambda) \ldots (g(\lambda_n) - \lambda).$$

Отсюда следует, что если $\lambda_1, \ldots, \lambda_n$ — собственные значения A, то $g(\lambda_1), \ldots, g(\lambda_n)$ — собственные значения $g(A)$.

Пусть теперь f — произвольная функция, определенная на Λ_A.

Тогда

$$f(A) = r(A), \quad f(\lambda_i) = r(\lambda_i), \quad i = \overline{1, n},$$

где $r(\lambda)$ — многочлен Лагранжа–Сильвестра.

Если $\lambda_1, \ldots, \lambda_n$ — собственные значения A, то $r(\lambda_1), \ldots, r(\lambda_n)$ — собственные значения $r(A)$, по $r(\lambda_1) = f(\lambda_1), \ldots, r(\lambda_n) = f(\lambda_n)$.

Следовательно, $f(\lambda_1), \ldots, f(\lambda_n)$ — собственные значения матрицы $f(A)$. □

(2) Пусть матрицы A и B подобны, причем $B = S^{-1}AS$, тогда *$f(B) = S^{-1}f(A)S$.*

Доказательство. Подобные матрицы имеют одинаковые минимальные многочлены, следовательно, $f(\Lambda_A) = f(\Lambda_B)$. Значит, существует $r(\lambda)$, такой, что $f(A) = r(A)$, $f(B) = r(B)$.

Заметим, что если $B = S^{-1}AS$, то $B^k = S^{-1}A^kS$. Отсюда следует, что для любого многочлена g: $g(B) = S^{-1}g(A)S$. Значит, $r(B) = S^{-1}r(A)S$, следовательно, $f(B) = S^{-1}f(A)S$. □

(3) Если $A = \mathrm{diag}\{A_1, \ldots, A_k\}$, то $f(A) = \mathrm{diag}\{f(A_1), \ldots, f(A_k)\}$.

Доказательство. Пусть $f(A) = r(A)$, где $r(\lambda)$ — интерполяционный многочлен Лагранжа – Сильвестра для функции f на спектре матрицы A. Легко заметить, что

$$f(A) = r(A) = \mathrm{diag}\{r(A_1), \ldots, r(A_k)\}.$$

Так как $f(\Lambda_A) = r(\Lambda_A)$, а $\Lambda_{A_i} \subset \Lambda_A$, $i = \overline{1, k}$, то

$$f(A_i) = r(A_i), \quad i = \overline{1, k}.$$

□

2.2. Интерполяционный многочлен Лагранжа – Сильвестра

(1) Матрица $A \in C_{n,n}$ не имеет кратных собственных значений, то есть $\lambda_i \neq \lambda_j$, $i, j = \overline{1, n}$, $(i \neq j)$. Тогда

$$\psi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n)$$

и интерполяционные условия имеют вид:

$$f(\lambda_i) = r(\lambda_i), \quad i = \overline{1, n}.$$

В этом случае в качестве $r(\lambda)$ можно взять обычный многочлен Лагранжа:

$$r(\lambda) = \sum_{k=1}^{n} \frac{(\lambda - \lambda_1) \ldots (\lambda - \lambda_{k-1})(\lambda - \lambda_{k+1}) \ldots (\lambda - \lambda_n)}{(\lambda_k - \lambda_1) \ldots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \ldots (\lambda_k - \lambda_n)} f(\lambda_k).$$

Тогда

$$f(A) = r(A) =$$

$$= \sum_{k=1}^{n} \frac{(A - \lambda_1 E_n) \ldots (A - \lambda_{k-1} E_n)(A - \lambda_{k+1} E_n) \ldots (A - \lambda_n E_n)}{(\lambda_k - \lambda_1) \ldots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \ldots (\lambda_k - \lambda_n)} f(\lambda_k).$$

(2) Пусть характеристический многочлен имеет кратные корни, а минимальный не имеет:

$$\psi(\lambda) = (\lambda - \lambda_1) \ldots (\lambda - \lambda_m).$$

Этот случай аналогичен предыдущему:

$$r(\lambda) = \sum_{k=1}^{m} \frac{(\lambda - \lambda_1) \ldots (\lambda - \lambda_{k-1})(\lambda - \lambda_{k+1}) \ldots (\lambda - \lambda_m)}{(\lambda_k - \lambda_1) \ldots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \ldots (\lambda_k - \lambda_m)} f(\lambda_k).$$

$$f(A) = r(A) =$$

$$= \sum_{k=1}^{m} \frac{(A-\lambda_1 E_n)\dots(A-\lambda_{k-1}E_n)(A-\lambda_{k+1}E_n)\dots(A-\lambda_m E_n)}{(\lambda_k-\lambda_1)\dots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\dots(\lambda_k-\lambda_m)} f(\lambda_k).$$

(3) Минимальный многочлен имеет кратные корни:

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\dots(\lambda-\lambda_s)^{m_s}, \quad m_1+\dots+m_s=m.$$

Представим выражение $\frac{r(\lambda)}{\psi(\lambda)}$ в виде суммы простых дробей. Это возможно, так как $\deg r < \deg \psi$.

$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s} \left(\frac{\alpha_{k_1}}{(\lambda-\lambda_k)^{m_k}} + \frac{\alpha_{k_2}}{(\lambda-\lambda_k)^{m_k-1}} + \dots + \frac{\alpha_{k_{m_k}}}{(\lambda-\lambda_k)} \right), \qquad (1)$$

где $\alpha_{ki} \in \mathbb{C}, \quad k=\overline{1,s}, \quad i=\overline{1,m_k}$.

Умножим (1) слева и справа на $(\lambda-\lambda_k)^{m_k}$ и введем новое обозначение:

$$\psi_k(\lambda) = \frac{\psi(\lambda)}{(\lambda-\lambda_k)} m_k.$$

Получаем:

$$\frac{r(\lambda)}{\psi_k(\lambda)} = \alpha_{k_1} + \alpha_{k_2}(\lambda-\lambda_k) + \dots + \alpha_{k_{m_k}}(\lambda-\lambda_k)^{m_k-1} + \\ + (\lambda-\lambda_k)^{m_k} h(\lambda),$$

где $h(\lambda_k) \neq \infty, \ k=\overline{1,s}$.

Отсюда следует, что

$$\alpha_{k_1} = \left(\frac{r(\lambda)}{\psi_k(\lambda)} \right) \Bigg|_{\lambda=\lambda_k} = \frac{r(\lambda_k)}{\psi_k(\lambda_k)},$$

$$\alpha_{k_2} = \left(\frac{r(\lambda)}{\psi_k(\lambda)} \right)' \Bigg|_{\lambda=\lambda_k}.$$

Таким образом мы можем найти все α_{k_i}:

$$\alpha_{k_i} = \frac{1}{(i-1)!} \left(\frac{r(\lambda)}{\psi_k(\lambda)} \right)^{(i-1)} \Bigg|_{\lambda=\lambda_{ki}}, \ i=\overline{1,m_k} \quad k=\overline{1,s}.$$

Отсюда следует, что все коэффициенты α_{ki} выражаются через значения многочлена $r(\lambda)$ на спектре матрицы A, которые нам известны:

$$f(\lambda_k) = r(\lambda_k), \quad f'(\lambda_k) = r'(\lambda_k), \dots, f^{(m_k-1)}(\lambda_k) = r^{(m_k-1)}(\lambda_k), \ k=\overline{1,s}.$$

Используя это, мы можем заменить значения многочлена $r(\lambda)$ и его производных на соответствующие значения функции $f(\lambda)$ и ее производных в точках $\lambda_k,\ k=\overline{1,s}$.

Тогда

$$\alpha_{k_i}=\frac{1}{(i-1)!}\left(\frac{f(\lambda)}{\psi_k(\lambda)}\right)^{(i-1)}\Bigg|_{\lambda=\lambda_k},\quad i=\overline{1,m_k}\quad k=\overline{1,s}.$$

Умножив обе части равенства (1) на $\psi(\lambda)$, получаем формулу для вычисления интерполяционного многочлена Лагранжа – Сильвестра для функции $f(\lambda)$ на спектре матрицы A.

$$r(\lambda)=\sum_{k=1}^{s}(\alpha_{k_1}+\alpha_{k_2}(\lambda-\lambda_k)+\cdots+\alpha_{k_{m_k}}(\lambda-\lambda_k)^{m_k-1})\psi_k(\lambda),$$

где

$$\alpha_{k_i}=\frac{1}{(i-1)!}\left(\frac{f(\lambda)}{\psi_k(\lambda)}\right)^{(i-1)}\Bigg|_{\lambda=\lambda_k},\quad i=\overline{1,m_k}\quad k=\overline{1,s}.$$

Пример. Пусть некоторая матрица A имеет минимальный многочлен

$$\psi(\lambda)=(\lambda-\lambda_1)^2(\lambda-\lambda_2)^3,$$

функция $f(\lambda)$ определена на спектре матрицы A. Необходимо записать интерполяционный многочлен Лагранжа – Сильвестра для вычисления $f(A)$.

$$\psi_1(\lambda)=(\lambda-\lambda_2)^3,$$
$$\psi_2(\lambda)=(\lambda-\lambda_1)^2.$$

Значит,

$$r(\lambda)=(\alpha+\beta(\lambda-\lambda_1))(\lambda-\lambda_2)^3+$$
$$+(\gamma+\sigma(\lambda-\lambda_2)+\varepsilon(\lambda-\lambda_2)^2)(\lambda-\lambda_1)^2.$$

Найдем коэффициенты:

$$\alpha=\frac{f(\lambda_1)}{(\lambda_1-\lambda_2)^3}.$$

Чтобы найти β, необходимо найти производную функции $\left(\frac{f(\lambda)}{(\lambda-\lambda_2)^3}\right)'$:

$$\left(\frac{f(\lambda)}{(\lambda-\lambda_2)^3}\right)'\Bigg|_{\lambda=\lambda_1}=\frac{f'(\lambda)(\lambda-\lambda_2)^3-3f(\lambda)(\lambda-\lambda_2)^2}{(\lambda-\lambda_2)^6}\Bigg|_{\lambda=\lambda_1}.$$

Следовательно,

$$\beta = \frac{f'(\lambda_1)(\lambda_1-\lambda_2) - 3f(\lambda_1)}{(\lambda_1-\lambda_2)^4}.$$

Аналогично

$$\gamma = \frac{f(\lambda_2)}{(\lambda_1-\lambda_2)^2},$$

$$\sigma = \frac{f'(\lambda_2)(\lambda_2-\lambda_1) - 2f(\lambda_2)}{(\lambda_2-\lambda_1)^3}, \quad \varepsilon = \frac{1}{2}\left(\frac{f(\lambda)}{(\lambda-\lambda_1)^2}\right)''\Bigg|_{\lambda=\lambda_2}.$$

Пример. Найдем $\sqrt{A}$ для матрицы

$$A = \begin{bmatrix} 5 & -1 & 2 \\ 1 & 3 & 2 \\ 0 & 0 & 4 \end{bmatrix}.$$

Характеристический многочлен матрицы A:

$$\det(A-\lambda E_3) = (\lambda-4)^3.$$

Так как $\operatorname{rank}(A-4E_3)=1$, минимальный многочлен матрицы A:

$$\psi(\lambda) = (\lambda-4)^2.$$

Тогда интерполяционный многочлен Лагранжа – Сильвестра имеет вид:

$$r(\lambda) = \alpha + \beta(\lambda-4).$$

Найдем α и β:

$$\alpha = \sqrt{4} = 2; \quad \beta = (\sqrt{\lambda})'\Big|_{\lambda=4} = \frac{1}{4}.$$

Тогда

$$r(\lambda) = 2 + \frac{1}{4}(\lambda-4) = \frac{1}{4}(\lambda+4).$$

Отсюда мы получаем, что

$$\sqrt{A} = r(A) = \frac{1}{4}(A+4E_3).$$

Значит,

$$\sqrt{A} = \frac{1}{4}\begin{bmatrix} 9 & -1 & 2 \\ 1 & 7 & 2 \\ 0 & 0 & 8 \end{bmatrix}.$$

3. МАТРИЧНЫЕ УРАВНЕНИЯ

3.1. Решение уравнения $AX = XB$

Рассмотрим матричное уравнение $AX = XB$,

$$\text{где } \ B \in \mathbb{C}_{n,n} \ \ A \in \mathbb{C}_{m,m}, \ \ X \in C_{m,n}.$$

Рассмотрим элементарные делители матрицы $A: \ (\lambda - \lambda_1)^{p_1}$, $(\lambda - \lambda_2)^{p_2}, \ \ldots, \ (\lambda - \lambda_u)^{p_u}$, где $p_1 + \cdots + p_n = m$.

Рассмотрим элементарные делители матрицы $B: \ (\lambda - \mu_1)^{q_1}$, $(\lambda - \mu_2)^{q_2}, \ \ldots, \ (\lambda - \mu_v)^{q_v}$, где $q_1 + \cdots + q_v = n$.

Через $\widetilde{A}$ и $\widetilde{B}$ обозначим жордановы нормальные формы матриц A и B, а через U и V соответствующие матрицы перехода:

$$\widetilde{A} = U^{-1}AU, \qquad \widetilde{B} = V^{-1}BV.$$

Тогда

$$\widetilde{A} = \mathrm{diag}\{\lambda_1 Ep_1 + Hp_1, \ \lambda_2 Ep_2 + Hp_2, \ldots, \lambda_u Ep_u + Hp_u\},$$

$$\widetilde{B} = \mathrm{diag}\{\mu_1 Eq_1 + Hq_1, \ldots, \mu_v Eq_v + Hq_v\}.$$

Здесь

$$\lambda_i Ep_i + Hp_i = \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix} \in \mathbb{C}_{p_i,p_i}$$

— жордановы клетки матрицы $A, \ i = \overline{1, u}$,

$$\mu_i Eq_i + Hq_i = \begin{pmatrix} \mu_i & 1 & 0 & \ldots & 0 & 0 \\ 0 & \mu_i & 1 & \ldots & 0 & 0 \\ \multicolumn{6}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & \mu_i & 1 \\ 0 & 0 & 0 & \ldots & 0 & \mu_i \end{pmatrix} \in \mathbb{C}_{q_i,q_i}$$

— жордановы клетки матрицы $B, \ i = \overline{1, v}$.

Подставив в уравнение $AX = XB$ матрицы $\widetilde{A}$ и $\widetilde{B}$, получаем:

$$U\widetilde{A}U^{-1}X = XV\widetilde{B}V^{-1}.$$

Домножив слева на U^{-1}, а справа на V, получаем:

$$\widetilde{A}U^{-1}XV = U^{-1}XV\widetilde{B}.$$

Вводя новые обозначения $\widetilde{X} = U^{-1}XV$, получаем уравнение $\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{B}$. Если нам удастся решить уравнение $\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{B}$, то мы легко сможем найти решение и уравнения $AX = XB$, используя формулу $X = U\widetilde{X}V^{-1}$.

Матрица $\widetilde{X}$ разбивается на блоки: $\widetilde{X} = (X_{\alpha\beta})$, где $X_{\alpha\beta} \in \mathbb{C}_{p_\alpha,q_\beta}$, $\alpha = \overline{1,u}$, $\beta = \overline{1,v}$ (в соответствии с блочно-диагональным видом матриц $\widetilde{A}$ и $\widetilde{B}$). Тогда уравнение $\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{B}$ распадается на $u \cdot v$ матричных уравнений:

$$(\lambda_\alpha E_{p_\alpha} + H_{p_\alpha})X_{\alpha\beta} = X_{\alpha\beta}(\mu_\beta E_{q_\beta} + H_{q_\beta}), \;\; \alpha = \overline{1,u}, \; \beta = \overline{1,v}$$

или

$$(\mu_\beta - \lambda_\alpha)X_{\alpha\beta} = H_{p_\alpha}X_{\alpha\beta} - X_{\alpha\beta}H_{q_\beta}, \;\; \alpha = \overline{1,u}, \; \beta = \overline{1,v}. \qquad (1)$$

Возможны два случая:

(1) $\mu_\beta \neq \lambda_\alpha$.

Умножая обе части равенства на $\mu_\beta - \lambda_\alpha$ и заменяя $(\mu_\beta - \lambda_\alpha)X_{\alpha\beta}$ на $H_{p_\alpha}X_{\alpha\beta} - X_{\alpha\beta}H_{q_\beta}$, получаем:

$$\begin{aligned}(\mu_\beta - \lambda_\alpha)^2 X_{\alpha\beta} &= H_{p_\alpha}(\mu_\beta - \lambda_\alpha)X_{\alpha\beta} - (\mu_\beta - \lambda_\alpha)X_{\alpha\beta}H_{q_\beta} = \\ &= H_{p_\alpha}(H_{p_\alpha}X_{\alpha\beta} - X_{\alpha\beta}H_{q_\beta}) - (H_{p_\alpha}X_{\alpha\beta} - X_{\alpha\beta}H_{q_\beta})H_{q_\beta} = \\ &= H_{p_\alpha}^2 X_{\alpha\beta} - 2H_{p_\alpha}X_{\alpha\beta}H_{q_\beta} + X_{\alpha\beta}H_{q_\beta}^2 = \\ &= \sum_{i+j=2} (-1)^j k_{ij} H_{p_\alpha}^i X_{\alpha\beta} H_{q_\beta}^j, \text{ где } k_{ij} \in \mathbb{C}.\end{aligned}$$

Проделывая эту операцию r раз, получаем:

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = \sum_{i+j=r} (-1)^j k_{ij} H_{p_\alpha}^i X_{\alpha\beta} H_{q_\beta}^j.$$

Напомним, что

$$(H_{p_\alpha})^{p_\alpha} = 0, \;\; (H_{q_\beta})^{q_\beta} = 0.$$

Если $r \geq p_\alpha + q_\beta$, то выполняется хотя бы одно из условий $i \geq p_\alpha, \; j \geq q_\beta$. Отсюда $(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = 0$, значит, $X_{\alpha\beta} = 0$.

(2) $\mu_\beta = \lambda_\alpha$.
Тогда
$$H_{p_\alpha} X_{\alpha\beta} = X_{\alpha\beta} H_{q_\beta}. \tag{2}$$
Мы опять получили уравнение вида $AX = XB$, но с матрицами специального вида. Напомним, что в матрицах H_{p_α} и H_{q_β} элементы первой наддиагонали равны единице, а все остальные нулю. Учитывая это, в зависимости от значений p_α и q_β, получаем:

2.1. $p_\alpha = q_\beta$.
Тогда
$$X_{\alpha\beta} = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{p_\alpha-1} & a_{p_\alpha} \\ 0 & a_1 & a_2 & \dots & a_{p_\alpha-2} & a_{p_\alpha-1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_1 & a_2 \\ 0 & 0 & 0 & \dots & 0 & a_1 \end{bmatrix} = T_{p_\alpha}, \tag{3}$$
то есть решением уравнения является квадратная матрица, в которой все элементы ниже главной диагонали равны нулю, элементы главной диагонали — некоторому произвольному параметру a_1, элементы первой поддиагонали — параметру a_2 и т. д.

2.2. $p_\alpha < q_\beta$.
Тогда
$$X_{\alpha\beta} = \begin{pmatrix} 0 & T_{p_\alpha} \end{pmatrix}, \tag{4}$$
где нулевой блок имеет размер $p_\alpha \times (q_\beta - p_\alpha)$.

2.3. $p_\alpha > q_\beta$.
Тогда
$$X_{\alpha\beta} = \begin{pmatrix} T_{q_\beta} \\ 0 \end{pmatrix}, \tag{5}$$
где нулевой блок имеет размер $(p_\alpha - q_\beta) \times q_\beta$.

Про матрицы (3), (4), (5) говорят, что они имеют *правильную верхнюю треугольную форму*. Количество произвольных параметров в них равно $\min\{p_\alpha, q_\beta\}$.

Пример.

(1) $p_\alpha = q_\beta = 4$.

$$X_{\alpha\beta} = \begin{bmatrix} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{bmatrix}.$$

(2) $p_\alpha = 3,\ q_\beta = 5$.

$$X_{\alpha\beta} = \begin{bmatrix} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{bmatrix}.$$

(3) $p_\alpha = 5,\ q_\beta = 3$.

$$X_{\alpha\beta} = \begin{bmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Итак, уравнение $H_{p_\alpha} X_{\alpha\beta} = X_{\alpha\beta} H_{q_\beta}$ в качестве решения имеет произвольную правильную верхнюю треугольную матрицу.

Введем новые обозначения:

$$d_{\alpha\beta}(\lambda) = \text{НОД}\ \left((\lambda - \lambda_\alpha)^{p_\alpha},\ (\lambda - \mu_\beta)^{q_\beta}\right),$$
$$\sigma_{\alpha\beta} = \deg d_{\alpha\beta}(\lambda).$$

Заметим, что

$$\sigma_{\alpha\beta} = \begin{cases} 0,\ \lambda_\alpha \neq \mu_\beta, \\ \min\{p_\alpha, q_\beta\}, \quad \lambda_\alpha = \mu_\beta. \end{cases}$$

Тогда число произвольных параметров в $\widetilde{X}$ (а значит, и в X) равно

$$N = \sum_{\alpha=1}^{u} \sum_{\beta=1}^{v} \sigma_{\alpha\beta}.$$

Обозначив через $X_{\widetilde{A}\widetilde{B}}$ решение уравнения $\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{B}$, сформулируем полученные результаты в виде следующей теоремы:

Теорема 1. *Общее решение уравнения $AX = XB$, где*

$A \in \mathbb{C}_{m,m}, \quad B \in \mathbb{C}_{n,n}$,

$A = U\widetilde{A}U^{-1} = U \operatorname{diag}(\lambda_1 E p_1 + H_{p_1}, \ldots, \lambda_u E p_u + H_{p_u}) U^{-1}$,

$$p_1 + \cdots + p_u = m,$$
$$B = V\widetilde{B}V^{-1} = V\,\mathrm{diag}(\mu_1 Eq_1 + H_{q_1}, \ldots, \mu_v Eq_v + H_{q_v})V^{-1},$$
$$q_1 + \cdots + q_v = n,$$

может быть найдено по формуле $X = UX_{\widetilde{A}\widetilde{B}}V^{-1}$.

Здесь $X_{\widetilde{A}\widetilde{B}}$ *— общее решение уравнения*

$$\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{B},\ X_{\widetilde{A}\widetilde{B}} = (X_{\alpha\beta}),\ \ \alpha = \overline{1,u},\ \beta = \overline{1,v},\ X_{\alpha,\beta} \in \mathbb{C}_{p_\alpha,q_\beta}.$$

Если $\lambda_\alpha \neq \mu_\beta$, *то* $X_{\alpha\beta} = 0$, *если* $\lambda_\alpha = \mu_\beta$, *то* $X_{\alpha\beta}$ *— произвольная правильная верхняя треугольная матрица.*

Матрица X *зависит от* N *произвольных параметров:*

$$X = \sum_{i=1}^{N} c_i X_i,\ \textit{где}\ N = \sum_{\alpha=1}^{u}\sum_{\beta=1}^{v} \sigma_{\alpha\beta},$$

где

$$\sigma_{\alpha\beta} = \deg \textit{НОД}\,\{(\lambda-\lambda_\alpha)^{p_\alpha},\ \ (\lambda-\mu_\beta)^{q_\beta}\},\ \ \alpha = \overline{1,u},\ \beta = \overline{1,v}.$$

Замечание. Частные решения $X_1, \ldots, X_N$ (X_i получается из X, если параметру c_i присвоить значение единица, а остальным — нуль) линейно независимы и образуют фундаментальную систему решений. Действительно, если это не так, тогда существует нетривиальная линейная комбинация $\sum\limits_{i=1}^{N} c_i X_i = 0$, то есть при ненулевых значениях некоторых параметров c_i матрица X, а значит, и $X_{\widetilde{A}\widetilde{B}}$ равны нулю, что невозможно.

Следствие. *Если матрицы* A *и* B *не имеют одинаковых собственных значений, то уравнение* $AX = XB$ *имеет только нулевое решение, то есть* $X = 0$.

Пример. Найти общее решение уравнения:

$$\begin{bmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{bmatrix} \cdot X = X \cdot \begin{bmatrix} -3 & 2 \\ 1 & -2 \end{bmatrix}.$$

Найдем жорданову нормальную форму для матрицы A и матрицу перехода U. Характеристическое уравнение имеет вид

$$(\lambda+1)^3 = 0.$$

Значит, единственное собственное значение $\lambda = -1$. Подставляем собственное значение в характеристическую матрицу:

$$A + E = \begin{bmatrix} 3 & 6 & -15 \\ 1 & 2 & -5 \\ 1 & 2 & -5 \end{bmatrix} \sim (12 - 5).$$

Пространство собственных векторов, соответствующих $\lambda = 1$, имеет вид:

$$\{5\beta - 2\alpha,\ \alpha,\ \beta \mid\ \alpha, \beta \in \mathbb{C}\}.$$

Ищем условие, при котором у собственного вектора существует присоединенный:

$$\left[\begin{array}{ccc|c} 3 & 6 & -15 & 5\beta - 2\alpha \\ 1 & 2 & -5 & \alpha \\ 1 & 2 & -5 & \beta \end{array}\right],$$

Следовательно, если $\alpha = \beta$, то система будет иметь решение.

Возьмем $\alpha = 1$, тогда собственный вектор $e_1 = (3, 1, 1)$, в качестве присоединенного к нему можно взять вектор $e_2 = (1, 0, 0)$, в качестве второго собственного вектора возьмем $e_3 = (-2, 1, 0)$. Получаем жорданов базис:

$$\begin{aligned} e_1 &= (3, 1, 1), \\ e_2 &= (1, 0, 0), \\ e_3 &= (-2, 1, 0). \end{aligned}$$

Отсюда следует, что

$$\widetilde{A} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad U = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Система элементарных делителей матрицы A:

$$(\lambda + 1)^2,\ (\lambda + 1).$$

Находим $\widetilde{B}$ и V:

$$\det(B - \lambda E_2) = \begin{vmatrix} -3 - \lambda & 2 \\ 1 & -2 - \lambda \end{vmatrix} =$$

$$= (3 + \lambda)(2 + \lambda) - 2 = \lambda^2 + 5\lambda + 4 = 0, \Rightarrow\ \lambda_1 = -1, \quad \lambda_2 = -4.$$

Отсюда

$$\widetilde{B} = \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}.$$

В качестве жорданова базиса можно взять собственные векторы $e_1 = (1,1),\ e_2 = (-2,1)$.

Система элементарных делителей матрицы B:

$$(\lambda+1),\ (\lambda+4).$$

$$V = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}, \quad V^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}.$$

Найдем $X_{\widetilde{A}\widetilde{B}}$:

$$A : (\lambda+1)^2,\ (\lambda+1), \quad B : (\lambda+1),\ (\lambda+4), \ \Rightarrow$$

$$X_{\widetilde{A}\widetilde{B}} = \left[\begin{array}{c|c} a & 0 \\ 0 & 0 \\ \hline b & 0 \end{array}\right], \quad a, b \in \mathbb{C}.$$

Найдем X:

$$X = \frac{1}{3}\begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{bmatrix} a & 0 \\ 0 & 0 \\ b & 0 \end{bmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} =$$

$$= \frac{1}{3}\begin{pmatrix} 3a-2b & 0 \\ a+b & 0 \\ a & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} =$$

$$= \frac{1}{3}\begin{pmatrix} 3a-2b & 6a-4b \\ a+b & 2a+2b \\ a & 2a \end{pmatrix}, \text{ где } a, b \in \mathbb{C}.$$

В качестве базиса пространства решений можно взять матрицы:

$$X_1 = \begin{pmatrix} 3 & 6 \\ 1 & 2 \\ 1 & 2 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} -2 & -4 \\ 1 & 2 \\ 0 & 0 \end{pmatrix}.$$

3.2. Решение уравнения $AX = XA$

Пусть дана матрица $A \in \mathbb{C}_{m,m}$. Будем решать следующую задачу: найти все матрицы $X \in \mathbb{C}_{m,m}$, перестановочные с матрицей A. Для этого нам необходимо найти общее решение уравнения $AX = XA$. Так как уравнение $AX = XA$ является частным случаем уравнения $AX = XB$, то для его решения воспользуемся теоремой 1 и сформулируем новую теорему:

Теорема 2. *Общее решение уравнения* $AX = XA$, $A \in \mathbb{C}_{m,m}$,

$$A = U\widetilde{A}U^{-1} = U\,\mathrm{diag}\{\lambda_1 Ep_1 + H_{p_1}, \ldots, \lambda_u Ep_u + H_{p_u}\}U^{-1},$$
$$p_1 + \cdots + p_u = m,$$

может быть найдено по формуле $X = UX_{\widetilde{A}}U^{-1}$.

Здесь $X_{\widetilde{A}}$ *— общее решение уравнения*

$$\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{A}, \quad X_{\widetilde{A}} = (X_{\alpha\beta}),\ \alpha, \beta = \overline{1,u},\ X_{\alpha\beta} \in \mathbb{C}_{p_\alpha, p_\beta}.$$

Если $\lambda_\alpha \neq \lambda_\beta$, *то* $X_{\alpha\beta} = 0$, *если* $\lambda_\alpha = \lambda_\beta$, *то* $X_{\alpha\beta}$ *— произвольная правильная верхняя треугольная матрица. Матрица* X *зависит от* N *произвольных параметров:*

$$N = \sum_{\alpha,\beta=1}^{u} \sigma_{\alpha\beta},$$

где

$$\sigma_{\alpha\beta} = \deg \textit{НОД}((\lambda - \lambda_\alpha)^{p_\alpha}(\lambda - \lambda_\beta)^{p_\beta}), \quad \alpha, \beta = \overline{1,u}$$

Пример.

$$A : (\lambda - \lambda_1)^4, (\lambda - \lambda_2)^3, (\lambda - \lambda_1)^2, (\lambda - \lambda_2).$$

Тогда общее решение уравнения $\widetilde{A}\widetilde{X} = \widetilde{X}\widetilde{A}$ имеет следующий вид:

$$X_{\widetilde{A}} = \left[\begin{array}{cccc|ccc|cc|c}
a & b & c & d & 0 & 0 & 0 & e & f & 0 \\
0 & a & b & c & 0 & 0 & 0 & 0 & e & 0 \\
0 & 0 & a & b & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & g & h & i & 0 & 0 & j \\
0 & 0 & 0 & 0 & 0 & g & h & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 \\
\hline
0 & 0 & k & l & 0 & 0 & 0 & m & n & 0 \\
0 & 0 & 0 & k & 0 & 0 & 0 & 0 & m & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & p & 0 & 0 & q
\end{array}\right].$$

Рассмотрим инвариантные множители матрицы A:

$$i_1(\lambda), i_2(\lambda), \ldots, i_t(\lambda),\ i_{t+1} = \cdots = i_m = 1.$$

Пусть

$$n_j = \deg i_j(\lambda), \text{ тогда } n_1 \geq n_2 \geq \cdots \geq n_t > n_{t+1} = \cdots = n_m = 0.$$

Так как каждый нетривиальный инвариантный множитель является произведением нескольких попарно взаимно простых элементарных делителей, то количество параметров в решении уравнения $AX = XA$ может быть найдено по следующей формуле:

$$N = \sum_{j,k=1}^{t} \varepsilon_{jk},$$

где

$$\varepsilon_{jk} = \deg \text{НОД}(i_j(\lambda),\ i_k(\lambda)) = \min\{n_j, n_k\}.$$

Отсюда мы получаем, что

$$N = n_1 + 3n_2 + 5n_3 + \cdots + (2t-1)n_t.$$

Таким образом, доказана следующая теорема.

Теорема 3. *Число линейно независимых матриц, перестановочных с матрицей $A \in \mathbb{C}_{n,n}$, определяется формулой*

$$N = n_1 + 3n_2 + \cdots + (2t-1)n_t,$$

где $i_1(\lambda), \ldots, i_t(\lambda)$ — непостоянные инвариантные множители матрицы A, $n_j = \deg i_j(\lambda)$, $j = \overline{1, t}$.

Замечание. Ясно, что $m = n_1 + n_2 + \cdots + n_t$. Отсюда следует, что $N \geq n$, причем $N = n \Leftrightarrow t = 1$, то есть все элементарные делители матрицы A попарно взаимно просты.

Пример. Пусть матрица имеет следующие элементарные делители:

$$(\lambda - \lambda_1)^5,\ (\lambda - \lambda_2)^3,\ (\lambda - \lambda_1)^3,\ (\lambda - \lambda_2)^2,\ (\lambda - \lambda_1).$$

Значит, нетривиальные инвариантные множители матрицы A имеют вид:

$$i_1(\lambda) = (\lambda - \lambda_1)^5(\lambda - \lambda_2)^3, \quad n_1 = 8;$$
$$i_2(\lambda) = (\lambda - \lambda_1)^3(\lambda - \lambda_2)^2, \quad n_2 = 5;$$
$$i_3(\lambda) = \lambda - \lambda_1, \quad n_3 = 1.$$

Тогда, согласно теореме 3, количество линейно независимых матриц, перестановочных с матрицей A, равно

$$N = 8 + 3 \cdot 5 + 5 \cdot 1 = 28.$$

Теорема 4. *Пусть*

$$AB = BA, \text{ где } A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

$$A_1 \in \mathbb{C}_{n,n}, \ A_2 \in \mathbb{C}_{m,m},$$

причем матрицы A_1 и A_2 не имеют одинаковых собственных значений. Тогда

$$B = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix},$$

$$B_1 \in \mathbb{C}_{n,n}, \quad B_2 \in \mathbb{C}_{m,m}.$$

Доказательство. Разобьем матрицу B на блоки в соответствии с блоками матрицы A:

$$B = \begin{pmatrix} B_1 & X \\ Y & B_2 \end{pmatrix},$$

где

$$B_1 \in \mathbb{C}_{n,n}, \ B_2 \in \mathbb{C}_{m,m}, \ X \in \mathbb{C}_{n,m}, \ Y \in \mathbb{C}_{m,n}.$$

Проверяя условие $AB = BA$, получаем, что

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \cdot \begin{bmatrix} B_1 & X \\ Y & B_2 \end{bmatrix} = \begin{bmatrix} B_1 & X \\ Y & B_2 \end{bmatrix} \cdot \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

$$\begin{bmatrix} A_1B_1 & A_1X \\ A_2Y & A_2B_2 \end{bmatrix} = \begin{bmatrix} B_1A_1 & XA_2 \\ YA_1 & B_2A_2 \end{bmatrix}.$$

Отсюда следует, что

$$A_2Y = YA_1 \Rightarrow Y = 0, \quad A_1X = XA_2 \Rightarrow X = 0,$$

так как A_1 и A_2 не имеют одинаковых собственных значений (см. следствие к теореме 1). $\square$

3.3. Решение уравнения $AX - XB = C$

Пусть дано уравнение $AX - XB = \mathbb{C}$, где

$$A \in \mathbb{C}_{m,m}, \ B \in \mathbb{C}_{n,n}, \ C \in \mathbb{C}_{m,n}, \ X \in \mathbb{C}_{m,n}.$$

Это матричное уравнение эквивалентно системе $m \cdot n$ линейных уравнений относительно элементов матрицы X.

Рассмотрим соответствующее однородное уравнение

$$AX - XB = 0.$$

Если матрицы A и B не имеют одинаковых собственных значений, то уравнение $AX - XB = C$ имеет единственное решение; если же матрицы A и B имеют одинаковые собственные значения, то в зависимости от C возможны два варианта:

(1) Уравнение не имеет решения.

(2) $X = X_0 + X_{\widetilde{A}\widetilde{B}}$, где X_0 — произвольное частное решение уравнения $AX - XB = C$, а $X_{\widetilde{A}\widetilde{B}}$ — общее решение уравнения $AX - XB = 0$.

4. СОПРЯЖЕННОЕ ОТОБРАЖЕНИЕ

4.1. Сопряженное пространство

Пусть V_n — векторное пространство над полем $\mathbb{R}$, $n = \dim V_n$, $f : V_n \to \mathbb{R}$ — вещественная функция, определенная над $\mathbb{R}$. Напомним, что функция f называется линейной, если выполняются следующие условия:

1) $f(u + v) = f(u) + f(v),\ \forall u, v \in V_n$;

2) $f(\alpha u) = \alpha f(u),\ \forall u \in V_n,\ \alpha \in \mathbb{R}$.

Пусть

$$E = \{e_1, \ldots, e_n\} \text{ — базис } V_n,\ u \in V_n.$$

Тогда

$$\begin{aligned} u &= \alpha_1 e_1 + \cdots + \alpha_n e_n, \\ f(u) &= f(\alpha_1 e_1 + \cdots + \alpha_n e_n) = \alpha_1 f(e_1) + \cdots + \alpha_n f(e_n) = \\ &= a_1 \alpha_1 + \cdots + a_n \alpha_n, \end{aligned}$$

где $a_i = f(e_i),\ i = \overline{1, n}$.

Отсюда видно, что при фиксированном базисе E любому набору из n чисел $a_1, \ldots, a_n$ соответствует линейная функция f, причем только одна.

Обозначим через

$$\tau_E(u) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha$$

координатный столбец вектора u в базисе E. Тогда $f(u) = a\alpha$, где $a = (a_1, \ldots, a_n)$. На множестве линейных функций мы можем определить операции сложения функций и умножения функции на число:

$$(f + g)(u) = f(u) + g(u), \quad \forall u \in V_n.$$

$$(\alpha f)(u) = \alpha f(u), \quad \forall u \in V_n,\ \forall \alpha \in \mathbb{R}.$$

Если функция f в базисе E определяется числами $(a_1, \ldots, a_n)$, а функция g в базисе E определяется числами $(b_1, \ldots, b_n)$, то функция $f + g$ определяется числами $(a_1 + b_1, \ldots, a_n + b_n)$, а функция αf — числами $(\alpha a_1, \ldots, \alpha a_n)$ в том же базисе E.

Нетрудно проверить, что множество линейных функций образует векторное пространство.

Определение. Пространство линейных функций, определенных на векторном пространстве V_n, называется *сопряженным пространством к* V_n и обозначается V_n^*.

Понятно, что V_n^* изоморфно n-мерному арифметическому пространству, значит,

$$\dim V_n^* = n, V_n^{**} = V_n.$$

Действие функции на вектор — это отображение из $V_n^* \times V_n$ в $\mathbb{R}$, $(f, u) \mapsto f(u)$.

В дальнейшем вместо $f(u)$ мы будем использовать обозначение $\langle f, u \rangle$, чтобы показать “равноправие"векторов из V_n и V_n^*, причем

$$\langle f, u \rangle = \langle u, f \rangle.$$

Получаем, что каждой паре $f \in V_n^*$ и $u \in V_n^*$ ставится в соответствие число $\langle f, u \rangle$, причем выполняются следующие условия:

1) $\langle f_1 + f_2, u \rangle = \langle f_1, u \rangle + \langle f_2, u \rangle$;
2) $\langle \alpha f, u \rangle = \alpha \langle f, u \rangle$;
3) $\langle f, u_1 + u_2 \rangle = \langle f, u_1 \rangle + \langle f, u_2 \rangle$;
4) $\langle f, \alpha u \rangle = \alpha \langle f, u \rangle$;

$\forall f, f_1, f_2 \in V_n^*,\ \forall u, u_1, u_2 \in V_n,\ \forall \alpha \in \mathbb{R}$.

Свойства 3 и 4 — это определение линейной функции, а 1 и 2 — переформулированные определения суммы линейных функций и произведения линейной функции на число.

Функцию $f \in V_n^*$ и вектор $u \in V_n$ назовем *ортогональными*, если $\langle f, u \rangle = 0$.

Пусть

$$E = \{e_1, \ldots, e_n\} \text{— базис } V_n,$$
$$F = \{f_1, \ldots, f_n\} \text{— базис } V_n^*.$$

Базисы E и F называются *биортогональными* (обозначаются $E \perp F$), если

$$\langle f_i, e_k \rangle = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}, \qquad i, k = \overline{1, n}.$$

Теорема 1. *Для любого базиса E пространства V_n существует, и притом только один, базис F пространства V_n^*, такой, что $E \perp F$.*

Доказательство. Условие биортогональности базисов E и F означает, что

$$\begin{gathered} \langle f_1, e_1 \rangle = 1,\ \langle f_1, e_2 \rangle = 0, \ldots, \langle f_1, e_n \rangle = 0; \\ \langle f_2, e_1 \rangle = 0,\ \langle f_2, e_2 \rangle = 1, \ldots, \langle f_2, e_n \rangle = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \langle f_n, e_1 \rangle = 0,\ \langle f_n, e_2 \rangle = 0, \ldots, \langle f_n, e_n \rangle = 1. \end{gathered}$$

По любому набору из n чисел, которые определяют действия функции на векторы базиса, можно построить функцию, причем только одну. Таким образом, набор $(1, 0, \ldots 0)$ однозначно определяет функцию f_1, набор $(0, 1, \ldots 0)$ — функцию $f_2, \ldots$, $(0, 0, \ldots 1)$ — функцию f_n. Функции $f_1, \ldots, f_n$ линейно независимы, так как линейно независимы соответствующие им наборы из n чисел, следовательно, образуют базис пространства V_n^*. □

Пример.

Рассмотрим пространство $\mathbb{R}_2$.

Пусть

$$E = \{(2, 3), (1, 5)\},\ (2, 3) = e_1,\ (1, 5) = e_2.$$

Найдем базис

$$F = \{f_1, f_2\}, \quad \text{такой, что} \quad F \perp E.$$

Произвольная линейная функция f действует на вектор (x, y):

$$f(x, y) = ax + by.$$

Функция f_1 определяется следующими условиями:

$$\begin{cases} \langle f_1, e_1 \rangle & = 1 \\ \langle f_1, e_2 \rangle & = 0, \end{cases} \quad \Rightarrow \quad \begin{cases} 2a + 3b & = 1 \\ a + 5b & = 0. \end{cases}$$

Аналогично определяем f_2:

$$\begin{cases} \langle f_2, e_1\rangle & = 0 \\ \langle f_2, e_2\rangle & = 1, \end{cases} \quad \Rightarrow \quad \begin{cases} 2a + 3b & = 0 \\ a + 5b & = 1. \end{cases}$$

Решаем системы линейных уравнений:

$$\left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 1 & 5 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 5 & 0 & 1 \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & \frac{5}{7} & -\frac{3}{7} \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} \end{array}\right].$$

Записываем ответ:

$$f_1(x,y) = \frac{5}{7}x - \frac{1}{7}y,$$
$$f_2(x,y) = -\frac{3}{7}x + \frac{2}{7}y.$$

Пусть

$$E = \{e_1, \dots, e_n\} \text{ — базис } V_n,$$
$$F = \{f_1, \dots, f_n\} \text{ — базис } V_n^*, \quad \text{причем} \quad E \perp F.$$

Рассмотрим вектор $u \in V_n$.

$$u = \alpha_1 e_1 + \dots \alpha_n e_n.$$

Тогда

$$\langle f_i, u\rangle = \langle f_i, \alpha_1 e_1 + \dots \alpha_n e_n\rangle = \alpha_i \langle f_i, e_i\rangle = \alpha_i, \quad i = \overline{1,n}.$$

Значит, $\alpha_i = \langle f_i, u\rangle$, $i = \overline{1,n}$.

Пусть $f \in V_n^*$, тогда $f = \beta_1 f_1 + \dots \beta_n f_n$.

$$\langle f, e_i\rangle = \langle \beta_1 f_1 + \dots + \beta_n f_n, e_i\rangle = \beta_i \langle f_i, e_i\rangle = \beta_i, \quad i = \overline{1,n}.$$

Отсюда следует, что $\langle f, e_i\rangle = \beta_i$.

Выразим $\langle f, u\rangle$ через координаты векторов f в базисе F и u в базисе E:

$$\langle f, u\rangle = \langle \beta_1 f_1 + \dots + \beta_n f_n,\, \alpha_1 e_1 + \dots + \alpha_n e_n\rangle =$$
$$= \sum_{i,k=1}^{n} \alpha_i \beta_k \langle f_k, e_i\rangle = \sum_{i=1}^{n} \alpha_i \beta_i.$$

Если базисы биортогональны, то

$$\langle f, u\rangle = \sum_{i=1}^{n} \alpha_i \beta_i.$$

Если базисы не являются ортогональными, то

$$\langle f, u\rangle = \sum_{i,k=1}^{n} \alpha_i \beta_k \langle f_k e_i\rangle.$$

Видно, что в биортогональных базисах значение $\langle f, u\rangle$ вычисляется намного проще.

Вернемся к предыдущему примеру. Найдем координаты вектора $u = (7, 10)$ в базисе E, используя базис F:

$$\begin{aligned}
&u = \alpha e_1 + \beta e_2,\\
&\alpha = \langle f_1, u\rangle = \frac{5}{7}\cdot 7 - \frac{1}{7}\cdot 10 = \frac{25}{7},\\
&\beta = \langle f_2, u\rangle = -\frac{3}{7}\cdot 7 + \frac{2}{7}\cdot 10 = -\frac{1}{7}.
\end{aligned}$$

Получаем, что $\tau_E(u) = \frac{1}{7}\begin{pmatrix} 25 \\ -1 \end{pmatrix}$.

Найдем координаты функции f в базисе $\{f_1, f_2\}$, используя базис E:

$$\begin{aligned}
&f(x, y) = x - 2y,\\
&f = \alpha f_1 + \beta f_2.\\
&\alpha = \langle f, e_1\rangle = 2 - 6 = -4;\\
&\beta = \langle f, e_2\rangle = 1 - 10 = -9.
\end{aligned}$$

Получаем, что $\tau_F(f) = \begin{pmatrix} -4 \\ -9 \end{pmatrix}$.

4.2. Ортогональное дополнение

Определение. Пусть V_k — подпространство векторного пространства V_n. Множество векторов из V_n^*, которые ортогональны всем векторам из V_k, называется *ортогональным дополнением* к пространству V_k и обозначается $V_k^{\perp}$.

Таким образом,

$$V_k^{\perp} = \{f \in V_n^* \mid \langle f, u\rangle = 0,\ \forall u \in V_k\}.$$

Теорема 2. *Ортогональное дополнение подпространства V_k является подпространством пространства V_n,* $\dim V_k^{\perp} = n - k$.

Доказательство. Пусть $E = \{e_1 \ldots, e_k\}$ — базис V_k, тогда $u = \alpha_1 e_1 + \cdots + \alpha_k e_k$. Покажем, что ортогональность функции f всем векторам из V_k равносильна ортогональности f векторам базиса E пространства V_k. Заметим, что

$$\langle f, u\rangle = \alpha_1\langle f, e_1\rangle + \cdots + \alpha_k\langle f, e_k\rangle.$$

Отсюда следует, что

$$\langle f, u\rangle = 0,\ \forall u \in V_k \ \Leftrightarrow\ \langle f, e_i\rangle = 0,\ i = \overline{1, k}.$$

Пусть

$$\begin{aligned} e_1 &= (a_{11}, a_{12}, \ldots a_{1n}),\\ &\vdots\\ e_k &= (a_{k1}, a_{k2}, \ldots a_{kn}). \end{aligned}$$

Тогда

$$\langle f, e_i\rangle = x_1 a_{i1} + x_2 a_{i2} + \cdots + x_n a_{in},\ \ i = \overline{1, k}.$$

Числа $x_1, \ldots, x_n$, определяющие функцию f, являются решениями следующей системы уравнений:

$$\begin{cases} x_1 a_{11} + \cdots + x_n a_{1n} & = 0\\ \cdots & \\ x_1 a_{k1} + \cdots + x_n a_{kn} & = 0. \end{cases}$$

Общее решение однородной системы образует подпространство в пространстве всех векторов $(x_1, \ldots, x_n)$.

$$\operatorname{rank}\begin{pmatrix} a_{11} & \ldots & a_{1n}\\ \ldots & \ldots & \ldots\\ a_{k1} & \ldots & a_{kn} \end{pmatrix} = k,$$

так как векторы $e_1, \ldots, e_k$ линейно независимы (образуют базис V_k). Отсюда следует, что общее решение этой системы образует подпространство в пространстве всех функций: $V_k^{\perp} \subset V_n^*$, и

$$\dim V_k^{\perp} = n - \operatorname{rank} A = n - k.$$

$\square$

Следствие. $(V_k^{\perp})^{\perp} = V_k$.

Доказательство. По определению

$$(V_k^\perp)^\perp = \{u \in V_n \mid \langle u, f\rangle = 0,\ \forall f \in V_k^\perp\}.$$

Следовательно, если $u \in V_k$, то $u \in (V_k^\perp)^\perp$. Получаем, что

$$V_k \subset (V_k^\perp)^\perp,\ \dim V_k = k,\ \dim (V_k^\perp)^\perp = n - (n-k) = k.$$

Значит, $V_k = (V_k^\perp)^\perp$.

□

4.3. Определение сопряженного отображения

Пусть даны два векторных пространства V_n и V_m и линейное отображение $\varphi : V_n \to V_m$. Пусть $f \in V_m^*$, $\ f : V_m \to \mathbb{R}$ — линейная функция.

Тогда можно определить новую функцию $f \circ \varphi : V_n \to \mathbb{R}$, которая является линейной, как композиция линейных отображений. Получим новое отображение, которое каждой линейной функции $f \in V_m^*$ ставит в соответствие линейную функцию $f \circ \varphi \in V_n^*$. Обозначим его φ^*.

Итак, $\varphi^*(f) = f \circ \varphi$.

Подействовав функцией на вектор $u \in V_n$, получаем

$$\langle \varphi^*(f), u\rangle = \langle f, \varphi(u)\rangle,\ \forall\ u \in V_n.$$

Докажем, что отображение φ^* — линейно. Проверим следующие условия:

1) $\varphi^*(f_1 + f_2) = \varphi^*(f_1) + \varphi^*(f_2), \forall f_1, f_2 \in V_m^*$,

2) $\varphi^*(\alpha f) = \alpha\varphi^*(f),\ \forall f \in V_m^*,\ \forall \alpha \in \mathbb{R}$.

Доказательство.

$$\begin{aligned} 1)\quad & \langle \varphi^*(f_1 + f_2), u\rangle = \langle f_1 + f_2, \varphi(u)\rangle = \\ =& \langle f_1, \varphi(u)\rangle + \langle f_2, \varphi(u)\rangle = \langle \varphi^*(f_1), u\rangle + \langle \varphi^*(f_2), u\rangle = \\ =& \langle \varphi^*(f_1) + \varphi^*(f_2), u\rangle,\ \forall u \in V_n. \end{aligned}$$

Показали, что

$$\varphi^*(f_1 + f_2) = \varphi^*(f_1) + \varphi^*(f_2).$$

$$\begin{aligned} 2)\quad & \langle \varphi^*(\alpha f), u\rangle = \langle \alpha f, \varphi(u)\rangle = \alpha\langle f, \varphi(u)\rangle = \\ & = \alpha\langle \varphi^*(f), u\rangle = \langle \alpha\varphi^*(f), u\rangle,\ \forall u \in V_n,\ \forall \alpha \in \mathbb{R}. \end{aligned}$$

Отсюда следует, что

$$\varphi^*(\alpha f) = \alpha\varphi^*(\alpha).$$

$\square$

Определение. Линейное отображение $\varphi^* : V_m^* \to V_n^*$, определяемое формулой

$$\varphi^*(f) = f \circ \varphi,$$

называется *сопряженным* линейному отображению $\varphi : V_m \to V_n$.

Рассмотрим отображение

$$\begin{aligned} &\varphi^{**} : V_n^{**} \to V_m^{**}, \text{ или} \\ &\varphi^{**} : V_n \to V_m. \end{aligned}$$

По определению сопряженного отображения:

$$\begin{aligned} &\langle \varphi^{**}(u), f \rangle = \langle u, \varphi^*(f) \rangle, \ \forall f \in V_m^* \\ &\langle f, \varphi^{**}(u) \rangle = \langle \varphi^*(f), u \rangle, \\ &\langle f, \varphi^{**}(u) \rangle = \langle f, \varphi(u) \rangle, \Rightarrow \\ &\varphi^{**}(u) = \varphi(u), \quad \forall u \in V_n \Rightarrow \\ &\varphi^{**} = \varphi. \end{aligned}$$

Пусть $E = \{e_1, \ldots, e_n\}$ — базис пространства V_n, $H = \{h_1, \ldots, h_m\}$ — базис пространства V_m. Отображение φ в базисах E и H имеет матрицу

$$\tau_E^H(\varphi) = A.$$

Рассмотрим

$$u \in V_n, \ u = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

$$\tau_E(u) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha.$$

$$\tau_H(\varphi(u)) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \beta.$$

Координатные столбцы векторов u и $\varphi(u)$ связаны формулой $\beta = A\alpha$. Рассмотрим функцию $f \in V_m^*$. Если зафиксировать базис H, то функцию заменим на соответствующую строку из чисел:

$$f \overset{H}{\sim} (b_1, \ldots, b_m) = b.$$

Тогда

$$\langle f, \varphi(u) \rangle = b\beta.$$

Аналогично

$$\varphi^*(f) \in V_n^*, \quad \varphi^*(f) \overset{E}{\sim} (a_1, \ldots, a_n) = a,$$
$$\langle \varphi^*(f), u \rangle = a\alpha.$$

Сопряженное отображение определяется формулой

$$\langle \varphi^*(f), u \rangle = \langle f, \varphi(u) \rangle, \quad \forall u \in V_n.$$

Переходя к координатной записи, получаем

$$a\alpha = b\beta,$$

или

$$a\alpha = bA\alpha, \quad \forall \alpha \in \mathbb{R}_n,$$

значит,

$$a = bA,$$

или

$$a^T = A^T b^T.$$

Напомним, что $b_i = \langle f, e_i \rangle, \quad i = \overline{1, m}$, то есть $\{b_1, \ldots, b_m\}$ — это координаты функции f в базисе $H^\perp$ пространства V_m^*, который биортогонален базису H.

Получаем, что

$$b^T = \tau_{H^\perp}(f).$$

Аналогично получаем

$$a^T = \tau_{E^\perp}(\varphi^*(f)),$$

где $E^\perp$ — базис пространства V_m^* биортогональный базису E, следовательно,

$$A^T = \tau_{H^\perp}^{E^\perp}(\varphi^*).$$

Таким образом, мы доказали следующую теорему.

Теорема 3. *Если линейное отображение $\varphi : V_n \to V_m$ в базисах E пространства V_n и H пространства V_m имеет матрицу*

$$\tau_E^H(\varphi) = A, \text{ то } \tau_{H^\perp}^{E^\perp}(\varphi^*) = A^T.$$

Следствие.

$$\operatorname{rank} \varphi = \operatorname{rank} \varphi^*.$$

Доказательство. Это следует из того, что

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

□

Укажем некоторые свойства сопряженного отображения:

1) $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$;
2) $(\varphi^{-1})^* = (\varphi^*)^{-1}$;
3) $(\varphi + \psi)^* = \varphi^* + \psi^*$;
4) $(\alpha\varphi)^* = \alpha\varphi^*,\ \forall \alpha \in \mathbb{R}$.

Доказательство. Все свойства доказываются при помощи координатной записи отображений φ и φ^*. Для примера докажем первое свойство.

Пусть $\tau(\varphi) = A,\ \tau(\psi) = B,$ тогда $\tau(\varphi \circ \psi) = A \cdot B$,

$$\tau((\varphi \circ \psi)^*) = (AB)^\top = B^\top A^\top = \tau(\psi^*)\tau(\varphi^*).$$

Отсюда следует, что

$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*.$$

□

4.4. Сопряженное преобразование

Пусть $\varphi : V_n \to V_n$ — линейное преобразование пространства V_n, тогда $\varphi^* : V_n^* \to V_n^*$ — преобразование пространства V_n^*, определяемое формулой

$$\varphi^*(f) = f \circ \varphi, \quad \forall f \in V_n^*,$$

называется *сопряженным* преобразованию φ.

Теорема 4. *Если преобразование $\varphi : V_n \to V_n$ в базисе E пространства V_n имеет матрицу $\tau_E(\varphi) = A$, то сопряженное преобразование φ^* в базисе $E^\perp$ пространства V_n^*, биортогональном базису E, имеет матрицу $\tau_E(\varphi^*) = A^\top$.*

Эта теорема является частным случаем теоремы 3.

Теорема 5. *Собственные значения преобразований φ и φ^* совпадают, при этом равные собственные значения имеют одинаковую кратность. Если собственному значению λ преобразования φ соответствует k линейно независимых собственных векторов, то k линейно независимых собственных векторов будет соответствовать этому собственному значению и для φ^*.*

Доказательство. Пусть E — базис V_n, тогда $\tau_E(\varphi) = A$, а $\tau_{E^\perp}(\varphi^*) = A^\top$.

$$\det(A - \lambda E) = \det(A - \lambda E)^\top = \det(A^\top - \lambda E).$$

Так как характеристические уравнения для φ и φ^* совпадают, то, значит, совпадают собственные значения и их кратности.

Так как $\mathrm{rank}(A - \lambda E) = \mathrm{rank}(A^\top - \lambda E)$, то количество линейно независимых векторов, соответствующих одному и тому же собственному значению для φ и φ^*, будет совпадать. $\square$

Теорема 6. *Если подпространство $V_k \subset V_n$ инвариантно относительно преобразования φ, то его ортогональное дополнение $V_k^\perp$ инвариантно относительно φ^*: $\varphi^*(V_k^\perp) \subset V_k^\perp$.*

Доказательство. Так как V_k инвариантно относительно φ, $\varphi(u) \in V_k,\ \forall u \in V_k$. По определению ортогонального дополнения

$$\langle f, \varphi(u)\rangle = 0,\ \ \forall f \in V_k^\perp,$$
$$\langle \varphi^*(f), u\rangle = \langle f, \varphi(u)\rangle = 0,\ \ \forall u \in V_k \Rightarrow \varphi^*(f) \in V_k^\perp,$$
$$\Rightarrow \varphi^*(V_k^\perp) \subset V_k^\perp,$$

то есть $V_k^\perp$ инвариантно относительно φ^*. $\square$

Теорема 7. *Пусть $\varphi(u) = \lambda u, \varphi(f) = \mu(f)$. Тогда, если $\lambda \neq \mu$,то $\langle f, u\rangle = 0$.*

Доказательство. Воспользуемся формулой

$$\langle \varphi^*(f), u\rangle = \langle f, \varphi(u)\rangle, \quad \forall f \in V_n^*,\ \forall u \in V_n.$$

Тогда

$$\begin{aligned}
&\langle \mu f, u\rangle = \langle f, \lambda u\rangle,\\
&\mu\langle f, u\rangle = \lambda\langle f, u\rangle,\\
&(\lambda - \mu)\langle\langle f, u\rangle = 0.
\end{aligned}$$

Получаем, что $\langle f, u\rangle = 0$, так как $\lambda \neq \mu$. $\square$

Теорема 8. *Если базис E пространства V_n состоит из собственных векторов преобразования $\varphi : V_n \to V_n$, то его биортогональный базис $E^\perp$ состоит из собственных векторов преобразования φ^*.*

Доказательство. Матрица преобразования φ в базисе E, состоящем из собственных векторов, диагональна:

$$\tau_E(\varphi) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{pmatrix} = A,$$

где $\lambda_1, \dots, \lambda_n$ — собственные значения.

Матрица сопряженного преобразования φ^* в базисе $E^\perp$ $\tau_{E^\perp}(\varphi^*) = A^\top = A$ тоже диагональна, значит, базис $E^\perp$ состоит из собственных векторов преобразования φ^*. $\square$

4.5. Сопряженное отображение евклидовых пространств

Пусть $\mathcal{E}_n$ и $\mathcal{E}_m$ — евклидовы пространства, $\varphi : \mathcal{E}_n \to \mathcal{E}_m$ — линейное отображение. Пусть $E = \{e_1, \dots, e_n\}$— базис $\mathcal{E}_n$.

Тогда скалярное произведение (u, v) векторов $u \in \mathcal{E}_n$ может быть определено следующим образом: $(u, v) = \alpha^\top G_E \beta$, где

$$\tau_E(u) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha, \quad \tau_E(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \beta,$$

$$G_E = \begin{pmatrix} (e_1, e_1) & \dots & (e_1, e_n) \\ \dots & \dots & \dots \\ (e_n, e_1) & \dots & (e_n, e_n) \end{pmatrix} \text{ — матрица Грама базиса } E.$$

Теорема 9. *Евклидово пространство изоморфно своему сопряженному пространству. То есть существует изоморфизм* $is : \mathcal{E}_n^* \to \mathcal{E}_n$, *который каждой функции* $f \in \mathcal{E}_n^*$ *ставит в соответствие вектор* $v \in \mathcal{E}_n$ *и такой, что* $(v, u) = \langle f, u \rangle$, $\forall u \in \mathcal{E}_n$.

Доказательство. Выберем какой–нибудь базис $E = \{e_1, \dots, e_n\}$ пространства $\mathcal{E}_n$. Тогда функции $f \in \mathcal{E}_n^*$ соответствует единственный набор из n чисел

$$f \overset{E}{\sim} (a_1, \dots, a_n) = a.$$

Равенство $(v, u) = \langle f, u \rangle$ можно заменить на его координатную запись:

$$a\alpha = \beta^\top G_E \alpha,$$

где G_E — матрица Грама базиса E,

$$\tau_E(u) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha, \quad \tau_E(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \beta.$$

Отсюда следует, что $a = \beta^\tau G_E$.

Значит, $\beta = G_E^{-1} a^\top$ (матрица Грама невырождена, так как E — базис $\mathcal{E}_n$).

Напомним, что $a^\tau = \tau_{E^\perp}(f)$ — координатный столбец функции f в базисе $E^\perp$, биортогональном базису E. Отсюда следует, что пространства $\mathcal{E}_n$ и $\mathcal{E}_n^*$ — изоморфны, причем $G_E^{-1} = \tau_{E^\perp}^E(is)$ — матрица изоморфизма is в базисах $E^\perp$ и E. $\square$

Замечание. Можно ввести в $\mathcal{E}_n^*$ скалярное произведение следующим образом:
$$(f,g) = (is(f), is(g)), \quad \forall f,g \in \mathcal{E}_n^*,$$
тогда is — изоморфизм евклидовых пространств $\mathcal{E}_n$ и $\mathcal{E}_n^*$.

Итак, евклидово пространство изоморфно своему сопряженному, следовательно, в дальнейшем мы можем отождествить евклидово пространство $\mathcal{E}_n$ с сопряженным ему пространством $\mathcal{E}_n^*$.

Тогда определение сопряженного отображения будет выглядеть следующим образом:

Определение. Отображение $\varphi^* : \mathcal{E}_n \to \mathcal{E}_m$, определяемое равенством

$$(\varphi^*(v), u) = (v, \varphi(u)),\ \forall u \in \mathcal{E}_n,\ \forall v \in \mathcal{E}_m, \tag{1}$$

называется *сопряженным* отображению $\varphi : \mathcal{E}_n \to \mathcal{E}_m$.

Заметим, что в левой части равенства стоит скалярное произведение в пространстве $\mathcal{E}_n$, а в правой — в пространстве $\mathcal{E}_m$.

Рассмотрим координатную запись формулы (1).

Пусть $E = \{e_1, \ldots, e_n\}$ — базис $\mathcal{E}_n$, $H = \{h_1, \ldots, h_m\}$ — базис $\mathcal{E}_m$.
$$\tau_E(u) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha,$$
$$\tau_H(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \beta.$$
Отображения φ и φ^* имеют в базисах E и H следующие матрицы:
$$\tau_E^H(\varphi) = A, \quad \tau_H^E(\varphi^*) = B.$$
Тогда
$$\tau_E(\varphi^*(v)) = B\beta, \quad \tau_H(\varphi(u)) = A\alpha.$$
Получаем координатную запись формулы (1):
$$(B\beta)^\top G_E \alpha = \beta^\top G_H (A\alpha),$$
где G_E и G_H — матрицы Грама базисов E и H соответственно.
$$\beta^\top B^\top G_E \alpha = \beta^\top G_H A \alpha, \quad \forall \alpha \in \mathbb{R}_n,\ \forall \beta \in \mathbb{R}_m.$$

Отсюда следует, что

$$B^\top G_E = G_H A.$$

$$B = G_E^{-1} A^\top G_H.$$

Таким образом, мы получили формулу, связывающую матрицы отображений φ и φ^*. Она значительно сложнее, чем раньше (если $\tau(\varphi) = A$, то $\tau(\varphi^*) = A^\top$), так как, отождествив пространства $\mathcal{E}_n$ и $\mathcal{E}_n^*$, мы выбираем один и тот же базис, а не два биортогональных друг другу. Чтобы сохранить формулу $B = A^\top$, матрицы G_E и G_H должны быть единичными, то есть базисы E и H — ортонормированными. Полученный результат сформулируем в виде следующей теоремы.

Теорема 10. *Если отображение $\varphi : \mathcal{E}_n \to \mathcal{E}_m$ в ортонормированных базисах пространств $\mathcal{E}_n$ и $\mathcal{E}_m$ имеет матрицу A, то его сопряженное отображение φ^* в тех же базисах имеет матрицу $A^\top$.*

Пример. Пусть $E = \{e_1, e_2\}$ и $H = \{h_1, h_2\}$ — базисы пространства $\mathbb{R}_2$, причем базис E — ортонормированный. Базисы E и H связаны соотношением

$$\begin{cases} e_1 &= h_1 - 2h_2, \\ e_2 &= h_1 + h_2. \end{cases}$$

Пусть преобразование $\varphi : \mathbb{R}_2 \to \mathbb{R}_2$ в базисе H имеет матрицу

$$\tau_H(\varphi) = \begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix} = A.$$

Найдем $\tau_H(\varphi^*)$.

Сначала найдем матрицу преобразования φ в базисе E:

$$\tau_E(\varphi) = S^{-1} A S,$$

где $S = \tau_H(E)$ — матрица перехода от базиса H к базису E.

$$S = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}, \quad \text{тогда} \quad S^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Находим $\tau_E(\varphi)$.

$$\tau_E(\varphi) = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} =$$

$$= \frac{1}{3}\begin{pmatrix} 1 & -5 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 11 & -4 \\ 4 & 1 \end{pmatrix}.$$

Так как базис E ортонормированный,

$$\tau_E(\varphi^*) = (\tau_E(\varphi))^\top = \frac{1}{3}\begin{pmatrix} 11 & 4 \\ -4 & 1 \end{pmatrix}.$$

Находим $\tau_H(\varphi^*)$.

$$\tau_H(\varphi^*) = S\tau_E(\varphi^*)S^{-1} = \frac{1}{9}\begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 11 & 4 \\ -4 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} =$$

$$= \frac{1}{9}\begin{pmatrix} 7 & 5 \\ -26 & -7 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 17 & -2 \\ -40 & 19 \end{pmatrix}.$$

Пользуясь тем, что мы отождествили евклидовы пространства $\mathcal{E}_n$ и $\mathcal{E}_m$ с их сопряженными $\mathcal{E}_n^*$ и $\mathcal{E}_m^*$, можно определить два новых линейных преобразования: $\varphi \circ \varphi^* : \mathcal{E}_m \to \mathcal{E}_m$ и $\varphi^* \circ \varphi : \mathcal{E}_n \to \mathcal{E}_n$.

Свойства этих преобразований похожи. Действительно,

$$\varphi \circ \varphi^* = (\varphi^*)^* \circ \varphi^* = [\psi = \varphi^*] = \psi^* \circ \psi.$$

Изучим некоторые свойства этих преобразований на примере $\varphi^* \circ \varphi$. Заметим, что справедливо следующее соотношение:

$$(\varphi^* \circ \varphi(u), u) = (\varphi(u), \varphi(u)) \geq 0,$$

причем

$$(\varphi^* \circ \varphi(u), u) = 0 \Leftrightarrow \varphi(u) = 0.$$

Теорема 11. *Преобразование $\varphi^* \circ \varphi$ — самосопряженное, его собственные значения неотрицательны.*

Доказательство.

$$(\varphi^* \circ \varphi)^* = \varphi^* \circ (\varphi^*)^* = \varphi^* \circ \varphi,$$

следовательно, $\varphi^* \circ \varphi$ — самосопряженное.

Пусть λ — собственное значение, а u— собственный вектор преобразования $\varphi^* \circ \varphi$:

$$(\varphi^* \circ \varphi)(u) = \lambda u.$$

Тогда

$$(\varphi^* \circ \varphi(u), u) = \lambda(u, u) \geq 0 \Rightarrow \lambda \geq 0.$$

□

Теорема 12. $\operatorname{Ker} \varphi^* \circ \varphi = \operatorname{Ker} \varphi,\ \operatorname{Im} \varphi^* \circ \varphi = \operatorname{Im} \varphi^*$.

Доказательство. Пусть

$$u \in \operatorname{Ker}\varphi \Rightarrow \varphi(u) = 0 \Rightarrow \varphi^* \circ \varphi(u) = 0 \Rightarrow$$
$$\Rightarrow u \in \operatorname{Ker}\varphi^* \circ \varphi \Rightarrow \operatorname{Ker}\varphi \subset \operatorname{Ker}\varphi^* \circ \varphi.$$

Пусть

$$u \in \operatorname{Ker}\varphi^* \circ \varphi \Rightarrow \varphi^* \circ \varphi(u) = 0 \Rightarrow$$
$$\Rightarrow (\varphi^* \circ \varphi(u), u) = 0 \Rightarrow \varphi(u) = 0 \Rightarrow$$
$$\Rightarrow u \in \operatorname{Ker}\varphi \Rightarrow \operatorname{Ker}\varphi^* \circ \varphi \subset \operatorname{Ker}\varphi \Rightarrow$$
$$\Rightarrow \operatorname{Ker}\varphi = \operatorname{Ker}\varphi^* \circ \varphi.$$

Докажем второе утверждение теоремы. Напомним, что для отображения $\varphi : \mathcal{E}_n \to \mathcal{E}_m$ справедлива следующая формула:

$$n = \operatorname{rank}\varphi + \dim\operatorname{Ker}\varphi,$$

причем $\operatorname{rank}\varphi = \dim\operatorname{Im}\varphi$.

Заметим, что $\operatorname{rank}\varphi = \operatorname{rank}\varphi^* \circ \varphi$, так как $\operatorname{Ker}\varphi = \operatorname{Ker}\varphi^* \circ \varphi$, а отображаемое пространство $\mathcal{E}_n$ одно и то же. Так как $\varphi(\mathcal{E}_n) \subset \mathcal{E}_m$, значит, $\varphi^* \circ \varphi(\mathcal{E}_n) \subset \varphi^*(\mathcal{E}_m)$, то есть $\operatorname{Im}\varphi^* \circ \varphi \subset \operatorname{Im}\varphi^*$, но $\dim\operatorname{Im}\varphi^* \circ \varphi = \operatorname{rank}\varphi^* \circ \varphi = \operatorname{rank}\varphi = \operatorname{rank}\varphi^* = \dim\operatorname{Im}\varphi^*$, значит, $\operatorname{Im}\varphi^* \circ \varphi = \operatorname{Im}\varphi^*$. □

Следствие. $\operatorname{rank}\varphi^* \circ \varphi = \operatorname{rank}\varphi \circ \varphi^*$.

Доказательство. Так как $\operatorname{rank}\varphi^* \circ \varphi = \operatorname{rank}\varphi$, то $\operatorname{rank}\varphi \circ \varphi^* = \operatorname{rank}\varphi^*$, значит, $\operatorname{rank}\varphi^* \circ \varphi = \operatorname{rank}\varphi \circ \varphi^*$, так как $\operatorname{rank}\varphi = \operatorname{rank}\varphi^*$. □

Теорема 13. *Пусть u — собственный вектор преобразования $\varphi^* \circ \varphi$, соответствующий собственному значению $\lambda \neq 0$, тогда $\varphi(u)$— собственный вектор преобразования $\varphi \circ \varphi^*$, соответствующий тому же собственному значению λ. При этом линейно независимым собственным векторам $u_1, \ldots, u_k$ преобразования $\varphi^* \circ \varphi$ с собственными значениями $\lambda_1, \ldots, \lambda_k$, $\lambda_i \neq 0$, $i = \overline{1, k}$ будут соответствовать линейно независимые собственные векторы $\varphi(u_1), \ldots, \varphi(u_k)$ преобразования $\varphi \circ \varphi^*$.*

Доказательство. Пусть $(\varphi^* \circ \varphi)(u) = \lambda u$, где $\lambda \neq 0$. Тогда $(\varphi^* \circ \varphi)(u) \neq 0$, значит, $\varphi(u) \neq 0$. Подействуем отображением φ:

$$\varphi((\varphi^* \circ \varphi)(u)) = \lambda\varphi(u),$$
$$(\varphi \circ \varphi^*)(\varphi(u)) = \lambda\varphi(u),$$

$\Rightarrow \varphi(u)$ — собственный вектор преобразования $\varphi^* \circ \varphi$, соответствующий собственному значению λ.

Пусть $u_1, \dots, u_k$ — линейно независимые векторы:

$$(\varphi^* \circ \varphi)(u_i) = \lambda_i u_i,\ i = \overline{1, k}, \lambda_i \neq 0.$$

Предположим, что векторы $\varphi(u_1), \dots, \varphi(u_k)$ линейно зависимы. Запишем их нетривиальную линейную комбинацию, равную нулю:

$$\alpha_1 \varphi(u_1) + \dots + \alpha_k \varphi(u_k) = 0.$$

Подействуем отображением φ^*:

$$\alpha_1 \varphi^* \circ \varphi(u_1) + \dots + \alpha_k \varphi^* \circ \varphi(u_k) = 0.$$

Получаем, что

$$\alpha_1 \lambda_1 u_1 + \dots + \alpha_k \lambda_k u_k = 0,$$

но $u_1, \dots, u_k$ — линейно независимы, а так как $\lambda_i \neq 0,\ i = \overline{1, k}$, то $\alpha_i = 0,\ i = \overline{1, k}$. Отсюда следует, что $\varphi(u_1), \dots, \varphi(u_k)$ — линейно независимы. $\square$

4.6. Сингулярные базисы отображения

Пусть дано отображение $\varphi : \mathcal{E}_n \to \mathcal{E}_m$, тогда $\varphi^* \circ \varphi : \mathcal{E}_n \to \mathcal{E}_n$ — преобразование пространства $\mathcal{E}_n$. Согласно теореме 11 преобразование $\varphi^* \circ \varphi$ — самосопряженное. Напомним, что любое самосопряженное преобразование обладает ортонормированным базисом, состоящим из собственных векторов.

Определение. *Первым сингулярным базисом* отображения φ называется ортонормированный базис пространства $\mathcal{E}_n$, состоящий из собственных векторов преобразования $\varphi^* \circ \varphi$, упорядоченных таким образом, что соответствующие собственные значения не возрастают.

Напомним, что собственные значения преобразования $\varphi^* \circ \varphi$ неотрицательны.

Если $r = \operatorname{rank} \varphi$, то $\lambda_i > 0$ при $i \leq r$ и $\lambda_i = 0$ при $i > r$.

Пусть $E = \{e_1, \dots, e_n\}$ — первый сингулярный базис отображения φ.

Тогда

$$(\varphi(e_i), \varphi(e_j)) = (\varphi^* \circ \varphi(e_i), e_j) = (\lambda_i e_i, e_j) = \lambda_i (e_i, e_j) = \begin{cases} \lambda_i, & i = j \\ 0, & i \neq j. \end{cases}$$

Отсюда следует, что векторы $\varphi(e_1), \ldots, \varphi(e_n)$ попарно ортогональны.

Кроме того,
$$(\varphi(e_i), \varphi(e_i)) = \lambda_i = \mid \varphi(e_i) \mid^2,$$
значит, $\mid \varphi(e_i) \mid = \sqrt{\lambda_i}$ — длина вектора $\varphi(e_i)$. Получаем, что $\varphi(e_i) \neq 0$ при $i \leq r$, $\varphi(e_i) = 0$ при $i > r$.

Определение. Числа $\alpha_i = \sqrt{\lambda_i}$, где λ_i — собственные значения преобразования $\varphi^* \circ \varphi$, $i = \overline{1, n}$, называются *сингулярными числами* отображения φ.

Тогда
$$|\varphi(e_i)| = \begin{cases} \alpha_i, & i \leq r \\ 0, & i > r. \end{cases}$$
Заметим, что $\frac{1}{\alpha_1}\varphi(e_1), \ldots, \frac{1}{\alpha_r}\varphi(e_r)$— ортонормированная система векторов. $\blacksquare$

Определение. *Вторым сингулярным базисом* отображения φ называется ортонормированный базис $H = \{h_1, \ldots, h_m\}$ пространства $\mathcal{E}_m$, первые m векторов которого имеют вид $h_i = \frac{1}{\alpha_i}\varphi(e_i)$, $i = \overline{1, r}$, где r — $\operatorname{rank}\varphi$, $e_1, \ldots, e_r$ — первые r векторов первого сингулярного базиса, $\alpha_1, \ldots, \alpha_r$ — ненулевые сингулярные числа φ.

Из определения следует, что сингулярные базисы определяются неоднозначно.

Теорема 14. *Пусть дано отображение*
$$\begin{aligned} &\varphi : \mathcal{E}_n \to \mathcal{E}_m, \\ &E = \{e_1, \ldots, e_n\} \text{ — первый сингулярный базис } \varphi, \\ &H = \{h_1, \ldots, h_m\} \text{ — второй сингулярный базис } \varphi. \end{aligned}$$
Матрица отображения φ в паре сингулярных базисов имеет вид:
$$\tau_E^H(\varphi) = \begin{pmatrix} \mathcal{D}_r & 0 \\ 0 & 0 \end{pmatrix} = \mathcal{D},$$
где
$$\mathcal{D}_r = \begin{pmatrix} \alpha_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \alpha_r \end{pmatrix}, \ r = \operatorname{rank}\varphi,$$
$\alpha_1, \ldots, \alpha_r$ — сингулярные числа отображения φ.

Доказательство. Находим $\tau_E^H(\varphi)$. Для этого векторы $\varphi(e_i),\ i = \overline{1,n}$ выражаем через базис H, $\varphi(e_1)$ выражаем через H и записываем в первый столбец, и так далее.

Получаем, что

$$\begin{aligned}
&\varphi(e_i) = \alpha_i\Big(\frac{1}{\alpha_i}\varphi(e_i)\Big) = \alpha_i h_i,\ i = \overline{1,r},\\
&\varphi(e_i) = 0,\ i > r.
\end{aligned}$$

Отсюда следует утверждение теоремы:

$$\tau_E^H(\varphi) = \begin{bmatrix} \alpha_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \alpha_r & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

□

Пусть имеется некоторая матрица $A \in \mathbb{R}_{m,n}$. Мы можем считать, что $A = \tau_k(\varphi)$ — это матрица некоторого отображения $\varphi : \mathcal{E}_n \to \mathcal{E}_m$ в канонических базисах пространств $\mathcal{E}_n$ и $\mathcal{E}_m$.

Согласно теореме 14 $\tau_E^H(\varphi) = \mathcal{D}$ — матрица отображения φ в сингулярных базисах, тогда $\mathcal{D} = T^{-1}AS$, где $S = \tau_K(E)$, $T = \tau_K(H)$ — соответствующие матрицы перехода. Заметим, что матрицы T и S ортогональны, так как и канонические, и сингулярные базисы ортонормированные.

Приведем матричную формулировку теоремы 14.

Теорема 15. *Для любой матрицы $A \in \mathbb{R}_{m,n}$ существуют ортогональные матрицы T и S, такие, что $\mathcal{D} = TAS$.*

Кроме того, мы показали, что любая матрица $A \in \mathbb{R}_{m,n}$ может быть представлена в следующем виде: $A = Q\mathcal{D}P$, где Q и P — ортогональные матрицы.

Это разложение называется *сингулярным разложением* матрицы A.

Теорема 16. *Пусть $A = Q\mathcal{D}P$, причем $QQ^\top = E_m, PP^\top = E_n$,*

$$\mathcal{D} = \begin{pmatrix} \mathcal{D}_r & 0 \\ 0 & 0 \end{pmatrix}, \quad \textit{где} \quad \mathcal{D}_r = \begin{bmatrix} \alpha_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \alpha_r \end{bmatrix},$$

причем $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_r > 0,\ r = \operatorname{rank}\varphi$.

Тогда:

1) строки матрицы P образуют первый сингулярный базис;
2) столбцы матрицы Q — второй сингулярный базис;
3) первые r сингулярных чисел матрицы $A:\ \alpha_1, \ldots, \alpha_r$, остальные сингулярные числа равны нулю.

Доказательство. Пусть p_i — i-й столбец матрицы $P^\top$.

Тогда

$$A^*Ap_i = (Q\mathcal{D}P)^\top (Q\mathcal{D}P)p_i =$$

$$= P^\top \mathcal{D}^\top \mathcal{D} P p_i = \left[\varepsilon_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} i,\ \varepsilon_i \in \mathbb{R}_{n,1} \right] = P^\top \mathcal{D}^\top \mathcal{D} \varepsilon_i =$$

$$= \left[\mathcal{D}^\top \mathcal{D} = \begin{bmatrix} \alpha_1^2 & \ldots & 0 & 0 \\ 0 & \ldots & \alpha_r^2 & 0 \\ \ldots & \ldots & \ldots & \\ 0 & \ldots & 0 & 0 \end{bmatrix} \right] = \begin{cases} 0, & i > r \\ P\alpha_i^2 \varepsilon_i, & i \le r \end{cases} =$$

$$= \begin{cases} 0, & i > r \\ \alpha_i^2 p_i, & i \le r. \end{cases}$$

Показали, что

$$A^*Ap_i = \begin{cases} \alpha_i^2 p_i, & i \le r \\ 0, & i > r. \end{cases}$$

Отсюда следует, что векторы $p_1, \ldots, p_n$ полностью удовлетворяют определению первого сингулярного базиса, а $\alpha_1, \ldots, \alpha_n$ — сингулярные числа.

Пусть q_i — i-й столбец матрицы Q.

Тогда

$$Ap_i = Q\mathcal{D}Pp_i = Q\mathcal{D}\varepsilon_i =$$

$$= \left[\widetilde{\varepsilon}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix} i,\ \widetilde{\varepsilon}_i \in \mathbb{R}_{m,1} \right] = \begin{cases} 0, & i > r \\ \alpha_i Q \widetilde{\varepsilon}_i, & i \le r. \end{cases} = \begin{cases} 0, & i > r \\ \alpha_i q_i, & i \le r. \end{cases}$$

Значит,

$$Ap_i = \begin{cases} 0, & i > r \\ \alpha_i q_i, & i \le r \end{cases}.$$

Таким образом, $q_i = \frac{1}{\alpha_i} Ap_i,\ i \le r$. Отсюда следует, что $\{q_1, \ldots, q_n\}$ — второй сингулярный базис. □

Теорема 17. *Пусть E — первый сингулярный базис отображения φ, H — второй сингулярный базис отображения φ. Тогда H — первый сингулярный базис отображения φ^*, E — второй сингулярный базис отображения φ^*. Сингулярные числа φ и φ^* совпадают.*

Доказательство. Пусть $\tau_K(\varphi) = A$ и $A = Q\mathcal{D}P$ — сингулярное разложение матрицы A, тогда $\tau_K(\varphi^*) = A^\top$, $A^\top = P^\top \mathcal{D}^\top Q^\top$.

Отсюда и следует утверждение теоремы. □

Теорема 18. *Пусть для отображения φ существует φ^{-1}, E — первый сингулярный базис отображения φ, H — второй сингулярный базис отображения φ, $\alpha_1, \ldots, \alpha_n$ — сингулярные числа отображения φ, тогда первый сингулярный базис отображения φ^{-1} отличен от H не более чем порядком векторов, а второй сингулярный базис отличается от E не более чем порядком векторов. Сингулярные числа φ^{-1} будут $\frac{1}{\alpha_1} \ldots \frac{1}{\alpha_n}$.*

Доказательство. Пусть $\tau_K(\varphi) = A$ и $A = Q\mathcal{D}P$ — сингулярное разложение матриц A, тогда $A^{-1} = \tau_K(\varphi^{-1})$ и $A^{-1} = (Q\mathcal{D}P)^{-1} = P^{-1}\mathcal{D}^{-1}Q^{-1} = P^\top \mathcal{D}^{-1} Q^\top$.

$$\mathcal{D} = \begin{bmatrix} \alpha_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \alpha_n \end{bmatrix} \Rightarrow \mathcal{D}^{-1} = \begin{bmatrix} \frac{1}{\alpha_1} & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \frac{1}{\alpha_n} \end{bmatrix}.$$

Отсюда и следует утверждение теоремы. □

Пример. Найти сингулярные числа и сингулярные базисы отображения $\varphi : \mathbb{R}_2 \to \mathbb{R}_3$ такого, что

$$\varphi(x, y) = (x + 2y, -y, x).$$

Пусть A — матрица отображения φ в канонических базисах пространств $\mathbb{R}_2$ и $\mathbb{R}_3$.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Сингулярные числа отображения φ — это квадратные корни из собственных значений матрицы A.

$$A^\top \cdot A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}.$$

$$\det(A^\top A - \lambda E_2) = \lambda^2 - 7\lambda + 6 = (\lambda - 6)(\lambda - 1).$$

Собственные значения матрицы $A^\top A$: $\lambda_1 = 6$ и $\lambda_2 = 1$.

Значит, сингулярные числа матрицы A (отображения φ):

$$\alpha_1 = \sqrt{6}, \qquad \alpha_2 = 1.$$

Первый сингулярный базис — это ортонормированный базис пространства $\mathbb{R}_2$, состоящий из собственных векторов матрицы $A^\top A$.

Пусть $\lambda_1 = 6$, тогда

$$A^\top A - 6E_2 = \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} \sim (2 \;\; -1).$$

Пусть $\lambda_2 = 1$, тогда

$$A^\top A - E_2 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \sim (1 \quad 2).$$

В качестве первого сингулярного базиса отображения φ можно взять векторы e_1 и e_2:

$$\begin{cases} e_1 & = \frac{1}{\sqrt{5}}(1, 2), \\ e_2 & = \frac{1}{\sqrt{5}}(2, -1). \end{cases}$$

Так как $\operatorname{rank}\varphi = 2$, первые два вектора второго сингулярного базиса находятся с помощью векторов первого сингулярного базиса и сингулярных чисел:

$$h_i = \frac{1}{\alpha_i} A e_i, \quad i = 1, 2.$$

$$h_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{30}} \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix},$$

$$h_2 = \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Дополняем систему векторов $\{h_1, h_2\}$ до ортонормированного базиса пространства $\mathbb{R}_3$. В качестве третьего вектора базиса можно взять, например, вектор $\frac{1}{\sqrt{6}}(1, 2, -1)$. Итак, мы нашли второй сингулярный базис отображения φ:

$$
\begin{aligned}
h_1 &= \frac{1}{\sqrt{30}}(5, -2, 1),\\
h_2 &= \frac{1}{\sqrt{5}}(0, 1, 2),\\
h_3 &= \frac{1}{\sqrt{6}}(1, 2, -1).
\end{aligned}
$$

Также укажем сингулярное разложение матрицы $A:\ A = Q\mathcal{D}P$.

$$
\begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix}.
$$

4.7. Сопряженное отображение комплексных пространств

В случае комплексного векторного пространства мы можем определить сопряженное пространство и сопряженное отображение так же, как и в действительном случае, все свойства при этом сохраняются. Но связь сопряженного отображения со скалярным произведением, которая имела место для евклидова пространства, не обобщается на унитарные пространства, то есть мы не сможем отождествить унитарное пространство со своим сопряженным. В связи с этим изменим определение сопряженного пространства для комплексного векторного пространства V_n.

Функция $f\colon V_n \to \mathbb{C}$ называется *полулинейной (эрмитово линейной)*, если выполняются следующие условия:

1) $f(u + v) = f(u) + f(v),\ \forall u, v \in V_n$;
2) $f(\alpha u) = \overline{\alpha} f(u),\ \forall u \in V_n,\ \forall \alpha \in \mathbb{C}$.

Пусть $E = \{e_1, \ldots, e_n\}$ — некоторый базис V_n, тогда любой вектор u можно выразить через базис E:

$$u = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Значит,
$$f(u) = f(\alpha_1 e_1 + \dots + \alpha_n e_n) = \overline{\alpha}_1 f(e_1) + \dots + \overline{\alpha}_n f(e_n).$$
Здесь
$$\tau_E(u) = \alpha = \begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_n \end{pmatrix}.$$
Если зафиксировать базис E, функцию f можно заменить строкой a:
$$f \overset{E}{\sim} (a_1, \dots, a_n) = a, \quad \text{где} \quad a_i = f(e_i),\ i = \overline{1, n}.$$
Тогда $f(u) = a\overline{\alpha}$.

Мы можем определить сумму полулинейных функций и умножение полулинейной функции на число:
$$(f_1 + f_2)(u) = f_1(u) + f_2(u),\ \forall u, v \in V_n;$$
$$(\alpha f)(u) = \alpha f(u),\ \forall u \in V_m,\ \forall \alpha \in \mathbb{C}.$$

Эти операции удовлетворяют аксиомам векторного пространства.

Определение. Пространство полулинейных функций, определенных на векторном пространстве V_n, называется *сопряженным пространством* векторному пространству V_n и обозначается V_n^*.

По-прежнему $\langle f, u \rangle = f(u)$.

Базис $F = \{f_1, \dots, f_n\}$ пространства V_n^* называется *биортогональным* базису $E = \{e_1, \dots, e_n\}$ пространства V_n, если
$$\langle f_i, e_k \rangle = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}.$$

Пусть есть линейное отображение $\varphi\colon V_n \to V_m$.

Определение. Линейное отображение $\varphi^*\colon V_m^* \to V_n^*$ называется *сопряженным* отображению φ, если
$$\langle \varphi^*(f), u \rangle = \langle f, \varphi(u) \rangle,\ \forall\, u \in V_n,\ \forall\, f \in V_m^*. \qquad (1)$$

Отсюда видим, что определение сопряженного отображения комплексных векторных пространств не отличается от определения сопряженного отображения действительных пространств. Все рассмотренные нами свойства сопряженного отображения сохраняются с небольшими изменениями, вызванными заменой линейных функций на полулинейные.

Пусть $\varphi : V_n \to V_m$,

$$E = \{e_1, \ldots, e_n\} \text{ — базис } V_n,$$
$$H = \{h_1, \ldots, h_m\} \text{ — базис } V_m.$$

Пусть вектор $u \in V_n$. Тогда

$$\tau_E(u) = \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix} = \alpha, \quad \tau_H(f(u)) = \begin{pmatrix} \beta_1 \\ \ldots \\ \beta_m \end{pmatrix} = \beta,$$

$\tau_E^H(\varphi) = A$ — матрица отображения φ в базисах E и H.

$$f \overset{H}{\sim} (b_1, \ldots, b_m) = b, \quad \varphi^*(f) \overset{E}{\sim} (a_1, \ldots, a_n) = a.$$

Тогда

$$\langle \varphi^*(f), u \rangle = a\overline{\alpha},$$
$$\langle f, \varphi(u) \rangle = b\overline{\beta}.$$

Перепишем определение (1) сопряженного отображения в координатной форме:

$$a\overline{\alpha} = b\overline{\beta} = b\overline{A\alpha}.$$

Значит, $a\overline{\alpha} = b\overline{A}\overline{\alpha}$. Получаем, что $a = b\overline{A}$. Отсюда следует, что

$$a^\top = \overline{A}^\top b^\top, \text{ или } a^\top = A^* b^\top.$$

Напомним, что

$$b^\top = \tau_{H^\perp}(f),$$
$$a^\top = \tau_{E^\perp}(\varphi^*(f)),$$

где $E^\perp$, $H^\perp$ — это базисы пространств V_n^* и V_m^*, биортогональные базисам E и H соответственно.

Показано, что $A^* = \tau_{H^\perp}^{E^\perp}(\varphi^*)$.

Таким образом, мы доказали следующую теорему.

Теорема 19. *Пусть линейное отображение $\varphi : V_n \to V_m$ в базисах E и H имеет матрицу $\tau_E^H(\varphi) = A$, тогда сопряженное отображение в биортогональных базисах имеет матрицу $\tau_{H^\perp}^{E^\perp}(\varphi^*) = A^*$.*

Свойства сопряженного отображения комплексных пространств те же, что и для действительных, за исключением того, что $(\alpha\varphi)^* = \overline{\alpha}\varphi^*$.

Пусть $\varphi\colon U_n \to U_m$ — линейное отображение унитарных пространств.

Рассмотрим $E = \{e_1, \ldots, e_n\}$ — базис U_n, $u, v \in U_n$. Напомним, что скалярное произведение (u, v) векторов u, v пространства U_n определяется формулой

$$(u, v) = \alpha^\top G_E \overline{\beta},$$

где

$$\tau_E(u) = \alpha, \quad \tau_E(v) = \beta,$$

G_E — матрица Грама в базисе E. Аналогично действительному случаю мы можем показать, что унитарное векторное пространство изоморфно своему сопряженному: $U_n \cong U_n^*$.

Тогда определение сопряженного отображения будет иметь следующий вид: отображение $\varphi^*\colon U_m \to U_n$ называется *сопряженным* отображению $\varphi\colon U_n \to U_m$, если

$$(\varphi^*(v), u) = (v, \varphi(u)),\ \forall u \in U_n,\ \forall v \in U_m. \tag{2}$$

Пусть $E = \{e_1, \ldots, e_n\}$ — базис U_n, $H = \{h_1, \ldots, h_m\}$ — базис U_m,

$$\tau_E(u) = \alpha, \quad \tau_H(v) = \beta,$$
$$\tau_E^H(\varphi) = A, \quad \tau_H^E(\varphi^*) = B,$$

следовательно,

$$(\varphi^*(v), u) = (B\beta)^\top G_E \overline{\alpha},$$
$$(v, \varphi(u)) = \beta^\top G_H \overline{A\alpha}.$$

Запишем условие (2) в координатной форме:

$$\beta^\top B^\top G_E \overline{\alpha} = \beta^\top G_H \overline{A}\overline{\alpha}.$$

Отсюда следует, что

$$B^\top G_E = G_H \overline{A}.$$

Тогда

$$B^\top = G_H \overline{A} (G_E)^{-1}.$$

Получили, что

$$B = (G_E^{-1})^\top \overline{A}^\top G_H^\top = (G_E^{-1})^\top A^* G_H^\top.$$

Теорема 20. *Пусть отображение $\varphi\colon U_n \to U_m$ в ортонормированных базисах E и H пространств U_n и U_m соответственно имеет матрицу $\tau_E^H(\varphi) = A$. Тогда $\tau_H^E(\varphi^*) = A^*$.*

Доказательство. Если базисы E и H ортонормированы, то $G_E = E_n$, $G_H = E_m$.

$$\tau_H^E(\varphi^*) = (G_E^{-1})^\top A^* G_H^\top = A^*.$$

$\square$

4.8. Экстремальные свойства собственных значений

Пусть $\varphi\colon \mathcal{E}_n \to \mathcal{E}_n$ — самосопряженное преобразование евклидового пространства $\mathcal{E}_n$.

Функция $\rho\colon E_n \backslash \{0\} \to \mathbb{R}$, определяемая формулой

$$\rho(u) = \frac{(\varphi(u), u)}{(u, u)},$$

называется *отношением Релея* для преобразования φ.

Определим единичную сферу S_n в пространстве $\mathcal{E}_n$.

Единичная сфера

$$S_n = \{u \in \mathcal{E}_n \mid |u| = 1\}.$$

Теорема 21. *Множество значений отношения Релея совпадает с множеством значений квадратичной формы* $(\varphi(u), u)$ *на единичной сфере.*

Доказательство. Пусть

$$u \in S_n \Rightarrow |u| = 1,\ |u| = \sqrt{(u, u)} \Rightarrow (u, u) = 1,$$

тогда $\rho(u) = (\varphi(u), u)$.

Пусть $u \in V_n$. Тогда

$$\rho(u) = \frac{(\varphi(u), u)}{|u|^2} = \left(\frac{1}{|u|}\varphi(u), \frac{1}{|u|}u\right) =$$

$$= \left(\varphi(\frac{1}{|u|}u), \frac{1}{|u|}u\right) =$$

$$= \left[v = \frac{1}{|u|}u,\ |v| = 1\right] = (\varphi(v), v).$$

То есть для любого вектора $u \in V_n$ найдется вектор $v \in S_n$, такой, что $\rho(u) = (\varphi(v), v)$. $\square$

Попробуем найти максимальное и минимальное значения отношения Релея в пространстве $\mathcal{E}_n$. Согласно теореме 21 для этого достаточно найти максимальное и минимальное значения квадратичной формы $(\varphi(u), u)$ на единичной сфере.

Так как φ — самосопряженное преобразование, существует ортонормированный базис $E = \{e_1, \dots, e_n\}$ пространства V_n, состоящий из собственных векторов преобразования φ: $\varphi(e_i) = \lambda_i e_i$, $i = \overline{1,n}$.

Упорядочим собственные значения так, что $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Пусть $u \in S_n$, $u = \alpha_1 e_1 + \cdots + \alpha_n e_n$, тогда $\rho(u) = (\varphi(u), u)$, так как $|u|^2 = (u, u) = 1$.

$$\begin{aligned}\rho(u) &= (\varphi(\alpha_1 e_1 + \cdots + \alpha_n e_n),\ \alpha_1 e_1 + \cdots + \alpha_n e_n) = \\ &= (\alpha_1 \varphi(e_1) + \cdots + \alpha_n \varphi(e_n),\ \alpha_1 e_1 + \cdots + \alpha_n e_n) = \\ &= (\alpha_1 \lambda_1 e_1 + \cdots + \alpha_n \lambda_n e_n,\ \alpha_1 e_1 + \cdots + \alpha_n e_n) = \\ &= \lambda_1 \alpha_1^2 + \lambda_2 \alpha_2^2 + \cdots + \lambda_n \alpha_n^2.\end{aligned}$$

Получили, что

$$\rho(u) = (\varphi(u), u) = \sum_{i=1}^{n} \lambda_i \alpha_i^2.$$

Заметим, что

$$|u| = 1 \Rightarrow \sum_{i=1}^{n} \alpha_i^2 = 1.$$

Тогда

$$\rho(u) = \sum_{i=1}^{n} \lambda_i \alpha_i^2 \leq \lambda_1 \sum_{i=1}^{n} \alpha_i^2 = \lambda_1,$$

то есть $\rho(u) \leq \lambda_1$.

Если $u = e_1$, тогда $\rho(e_1) = (\varphi(e_1), e_1) = \lambda_1$.

Кроме того,

$$\rho(u) = \sum_{i=1}^{n} \lambda_i \alpha_i^2 \geq \lambda_n \sum_{i=1}^{n} \alpha_i^2 = \lambda_n,$$

то есть $\rho(u) \geq \lambda_n$.

Если $u = e_n$, тогда

$$\rho(e_n) = (\varphi(e_n), e_n) = (\lambda_n e_n, e_n) = \lambda_n.$$

Таким образом, мы доказали следующую теорему.

Теорема 22. *Максимальное значение отношения Релея равно наибольшему собственному значению преобразования φ. Минимальное значение отношения Релея равно наименьшему собственному значению преобразования φ. Достигаются они на соответствующих собственных векторах.*

4.9. Полярное разложение линейного преобразования

Теорема 23. *Для любого преобразования $\varphi\colon \mathcal{E}_n \to \mathcal{E}_n$ существуют ортогональное преобразование $p\colon \mathcal{E}_n \to \mathcal{E}_n$ и самосопряженное преобразование $s\colon \mathcal{E}_n \to \mathcal{E}_n$ с неотрицательными собственными значениями, такие, что $\varphi = p \circ s$. Разложение $\varphi = p \circ s$ называется полярным разложением преобразования φ.*

Доказательство. Пусть $E = \{e_1, \ldots, e_n\}$ — первый сингулярный базис φ, $H = \{h_1, \ldots, h_n\}$ — второй сингулярный базис φ. Базисы E и H — ортонормированы, причем $\varphi(e_i) = \alpha_i h_i,\ i = \overline{1,n}$, где α_i — сингулярные числа преобразования φ. Если $r = \operatorname{rank}\varphi$, то при $i \leq r$ — это определение векторов $\{h_1, \ldots, h_r\}$, если $i > r$, то $\alpha_i = 0$, следовательно, $\varphi(e_i) = 0$.

Рассмотрим преобразование $p\colon \mathcal{E}_n \to \mathcal{E}_n$, которое переводит первый сингулярный базис во второй, то есть $p(e_i) = h_i,\ i = \overline{1,n}$. Оно ортогонально. Действительно, его матрица в базисе E — это матрица перехода от базиса E к базису H будет ортогональной, как матрица перехода от ортонормированного базиса к ортонормированному, а значит, p — ортогональное преобразование.

Пусть $s = p^{-1} \circ \varphi$. Покажем, что s — самосопряженное преобразование с неотрицательными собственными значениями:

$$\begin{aligned} s(e_i) &= (p^{-1} \circ \varphi)(e_i) = p^{-1}(\varphi(e_i)) = \\ p^{-1}(\alpha_i h_i) &= \alpha_i p^{-1}(h_i) = \\ &= \alpha_i e_i,\ i = \overline{1,n}. \end{aligned}$$

Получили, что $s(e_i) = \alpha_i e_i,\ i = \overline{1,n}$, то есть векторы первого сингулярного базиса — это собственные векторы, а сингулярные числа — собственные значения преобразования s.

Тогда

$$\tau_E(s) = \begin{pmatrix} \alpha_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \alpha_n \end{pmatrix} = S,$$

следовательно, $S = S^\top$.

Получили, что в ортонормированном базисе преобразование s имеет симметричную матрицу. Откуда следует: s — самосопряженное преобразование, причем его собственные значения $\alpha_i \geq 0,\ i = \overline{1, n}$ (как сингулярные числа). $\square$

Теорема 23 может быть сформулирована в матричном виде.

Теорема 24. *Любую матрицу $A \in \mathbb{R}_{n,n}$ можно представить в виде произведения матриц $P \in \mathbb{R}_{n,n}$ и $S \in \mathbb{R}_{n,n}$ $A = PS$, где P — ортогональная матрица, а S — симметрическая с неотрицательными собственными значениями ($PP^\top = E_n,\ S = S^\top$).*

Следствие. *Для любого преобразования $\varphi\colon \mathcal{E}_n \to \mathcal{E}_n$ существует ортогональное преобразование $\rho_1\colon \mathcal{E}_n \to \mathcal{E}_n$ и самосопряженное преобразование $s_1\colon \mathcal{E}_n \to \mathcal{E}_n$ с неотрицательными собственными значениями, такие, что $\varphi = s_1 \circ \rho_1$.*

Доказательство. Запишем полярное разложение преобразования φ^*:

$$\varphi^* = p \circ s, \quad \text{тогда} \quad \varphi = (p \circ s)^* = s^* \circ p^* = s \circ p^*.$$

Для ортогонального преобразования $p^* = p^{-1}$. Действительно, пусть E — ортонормированный базис пространства $\mathcal{E}_n$, пусть $\tau_E(p) = P$, тогда $\tau_E(p^*) = P^\top$, но так как p — ортонормированный, $PP^\top = E_n$, следовательно, $P^\top = P^{-1} \Rightarrow \tau_E(p^{-1}) = P^\top \Rightarrow p^* = p^{-1}$. А матрица, обратная к ортогональной, — ортогональна. Значит, p^{-1} — ортогональное преобразование. Тогда, положив $s_1 = s,\ p_1 = p^{-1}$, получим, что

$$\varphi = s \circ p^* = s_1 \circ p_1.$$

$\square$

Пример. Найдем полярное разложение матрицы

$$A = \begin{pmatrix} 1 & -7 \\ -1 & 7 \end{pmatrix}.$$

Сначала найдем сингулярные числа и сингулярные базисы матрицы A.

$$A^\top A = \begin{pmatrix} 1 & -1 \\ -7 & 7 \end{pmatrix} \begin{pmatrix} 1 & -7 \\ -1 & 7 \end{pmatrix} = \begin{pmatrix} 2 & -14 \\ -14 & 98 \end{pmatrix}.$$

$$\det(A^\top A - \lambda E_2) = \lambda^2 - 100\lambda = \lambda(\lambda - 100).$$

Отсюда следует, что собственные значения матрицы $A^\top A$ — это числа $\lambda_1 = 100$ и $\lambda_2 = 0$, значит, сингулярные числа матрицы A: $\alpha_1 = 10$ и $\alpha_2 = 0$.

Находим первый сингулярный базис.

$$A^\top A - 100E_2 = \begin{pmatrix} -98 & -14 \\ -14 & -2 \end{pmatrix} \sim (7 \quad 1).$$

$$A^\top A = \begin{pmatrix} 2 & -14 \\ -14 & 98 \end{pmatrix} \sim (1 \ -7).$$

В качестве векторов первого сингулярного базиса можно взять векторы

$$e_1 = \frac{1}{5\sqrt{2}}(-1, 7) \quad \text{и} \quad e_1 = \frac{1}{5\sqrt{2}}(7, 1).$$

Так как $\operatorname{rank} A = 1$, вектор

$$h_1 = \frac{1}{50\sqrt{2}} \begin{pmatrix} 1 & -7 \\ -1 & 7 \end{pmatrix} \begin{pmatrix} -1 \\ 7 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Тогда второй сингулярный базис матрицы A имеет вид:

$$h_1 = \frac{1}{\sqrt{2}}(-1, 1), \qquad h_2 = \frac{1}{\sqrt{2}}(1, 1).$$

Найдем матрицу P. Она удовлетворяет следующим условиям:

$$Pe_i = h_i, \quad i = \overline{1, 2}.$$

Рассмотрим матрицы $E = (e_1, e_2)$ и $H = (h_1, h_2)$, тогда $PE = H$. Значит, $E^\top P^\top = H^\top$. Решая это матричное уравнение, находим $P^\top$:

$$\left(\begin{array}{cc|cc} -1 & 7 & -5 & 5 \\ 7 & 1 & 5 & 5 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & -7 & 5 & -5 \\ 0 & 50 & -30 & 40 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & 1 & -\frac{3}{5} & \frac{4}{5} \end{array}\right).$$

Отсюда следует, что

$$P = \frac{1}{5} \begin{pmatrix} 4 & -3 \\ 3 & 4 \end{pmatrix}.$$

Находим матрицу S.

$$S = P^{-1} A = P^\top A$$

(так как матрица P ортогональна).

$$S = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -7 \\ -1 & 7 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & -7 \\ -7 & 49 \end{pmatrix}.$$

В результате получаем, что полярное разложение $A = PS$ матрицы A имеет вид:

$$\begin{pmatrix} 1 & -7 \\ -1 & 7 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 4 & -3 \\ 3 & 4 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} 1 & -7 \\ -7 & 49 \end{pmatrix}.$$

4.10. Единственность полярного разложения

Пусть $\varphi = p \circ s$ — полярное разложение преобразования $\varphi : \mathcal{E}_n \to \mathcal{E}_n$. Тогда

$$\varphi^* \circ \varphi = (p \circ s)^* \circ (p \circ s) = s^* \circ p^* \circ p \circ s = s \circ p^{-1} \circ p \circ s = s^2,$$

то есть $\varphi^* \circ \varphi = s^2$, значит, преобразование s^2 определено однозначно.

Теорема 25. *Пусть $s : \mathcal{E}_n \to \mathcal{E}_n$. Собственные векторы преобразования s будут собственными векторами преобразования s^2. Если s — самосопряженное преобразование с неотрицательными собственными значениями, то верно и обратное, то есть собственные векторы s^2 будут собственными и для s. При этом если $s^2(u) = \lambda u$, то $s(u) = \alpha u$, где $\alpha = \sqrt{\lambda}$.*

Доказательство. Пусть вектор u — собственный вектор для s, то есть $s(u) = \alpha u$, тогда $s^2(u) = s(s(u)) = s(\alpha u) = \alpha s(u) = \alpha^2 u$.

Получили, что $s^2(u) = \alpha^2 u$. Значит, собственные векторы преобразования s будут собственными векторами преобразования s^2.

Пусть $s^2(u) = \lambda u$, причем s — самосопряженное преобразование пространства $\mathcal{E}_n$ с неотрицательными собственными значениями. Берем ортонормированный базис $E = \{e_1, \dots, e_n\}$ пространства $\mathcal{E}_n$, состоящий из собственных векторов преобразования s:

$$s(e_i) = \alpha_i e_i, \quad \alpha_i \geq 0, \ i = \overline{1, n}.$$

Тогда

$$u = x_1 e_1 + \dots + x_n e_n.$$

$$s^2(u) = \lambda u = \lambda(x_1 e_1 + \dots + x_n e_n) = \sum_{i=1}^{n} \lambda x_i e_i.$$

С другой стороны,

$$s^2(u) = s^2(x_1 e_1 + \dots + x_n e_n) = x_1 s^2(e_1) + \dots + x_n s^2(e_n) =$$

$$= x_1\alpha_1^2 e_1 + \dots + x_n\alpha_n^2 e_n = \sum_{i=1}^{n} x_i\alpha_i^2 e_i.$$

Получим, что

$$s^2(u) = \sum_{i=1}^{n} \lambda x_i e_i \quad \text{и} \quad s^2(u) = \sum_{i=1}^{n} x_i\alpha_i^2 e_i.$$

Тогда

$$\sum_{i=1}^{n} (\lambda - \alpha_i^2) x_i e_i = 0,$$

следовательно,

$$(\lambda - \alpha_i^2) x_i = 0, \ i = \overline{1,n},$$

где $x_1, \dots, x_n$ — координаты вектора u в базисе E_n. Получаем, что $x_i \neq 0$ только в случае, когда $\lambda = \alpha_i^2 \Leftrightarrow \alpha_i = \sqrt{\lambda}$.

Это означает, что вектор u раскладывается только по векторам базиса E, которые соответствуют одному собственному значению $\alpha_i = \sqrt{\lambda}$ преобразования s. Значит, u — собственный вектор преобразования s. $\square$

Теорема 26. *Самосопряженное преобразование однозначно определяется своими собственными векторами и собственными значениями.*

Доказательство. Рассмотрим самосопряженное преобразование

$$s : \mathcal{E}_n \to \mathcal{E}_n.$$

Пусть $\lambda_1, \dots, \lambda_p$, $\lambda_i \neq \lambda_j$, $i \neq j$, $i, j = \overline{1,p}$, — все различные собственные значения преобразования s, $V_1, \dots, V_p$ — соответствующие собственные подпространства. Тогда любой вектор $u \in \mathcal{E}_n$ можно представить в виде $u = u_1 + \dots + u_p$, $u_i \in V_i$, причем такое представление единственно.

Разложение такого вида мы можем получить, объединяя слагаемые, соответствующие одному и тому же собственному значению в разложении вектора и по любому базису пространства $\mathcal{E}_n$ из собственных векторов преобразования s.

Докажем единственность. Пусть для некоторого вектора $u \in \mathcal{E}_n$ существует второе представление:

$$u = v_1 + \dots + v_p, \ v_i \in V_i.$$

Тогда

$$(u_1 - v_1) + (u_2 - v_2) + \dots + (u_p - v_p) = 0.$$

Следовательно, система векторов $\{u_1 - v_1, \dots, u_p - v_p\}$ линейно зависима, но это невозможно, так как все векторы соответствуют разным собственным значениям, следовательно,

$$u_i - v_i = 0, \ i = \overline{1, p}.$$

Тогда

$$s(u) = s(u_1 + \dots + u_p) = s(u_1) + \dots + s(u_p) = \lambda u_1 + \dots + \lambda_p u_p.$$

□

Преобразование s^2 определено однозначно, а так как собственные значения s неотрицательны, но согласно теореме 25 собственные векторы и собственные значения s^2 однозначно определяют собственные векторы и собственные значения s, которые, в свою очередь, однозначно определяют само преобразование s.

Ортогональное преобразование p в полярном разложении $\varphi = p \circ s$ определено неоднозначно. Однако если преобразование φ невырождено, то есть $\operatorname{rank} \varphi = n$, тогда невырождено и преобразование s. В этом случае преобразование p однозначно определяется как $p = \varphi \circ s^{-1}$.

4.11. Сингулярные числа и сингулярные базисы преобразования

Определим сингулярные числа и сингулярные базисы для отображений евклидовых пространств.

Рассмотрим некоторые свойства сингулярных чисел и сингулярных базисов преобразований.

Теорема 27. *Для любого линейного преобразования $\varphi\colon \mathcal{E}_n \to \mathcal{E}_n$*

$$\max_{u \in \mathcal{E}_n \setminus \{0\}} \frac{|\varphi(u)|}{|u|} = \alpha_1,$$

$$\min_{u \in \mathcal{E}_n \setminus \{0\}} \frac{|\varphi(u)|}{|u|} = \alpha_n,$$

где α_1 — максимальное сингулярное число преобразования φ, α_n — минимальное сингулярное число преобразования φ.

Доказательство.

$$\frac{|\varphi(u)|^2}{|u|^2} = \frac{(\varphi(u), \varphi(u))}{(u, u)} = \frac{(\varphi^* \circ \varphi(u), u)}{(u, u)} = \rho(u).$$

Таким образом, мы видим, что квадрат отношения $\frac{|\varphi(u)|}{|u|}$ равен отношению Релея для преобразования $\varphi^* \circ \varphi$. Согласно теореме 22 $\lambda_n \leq \rho(u) \leq \lambda_1$, где λ_1 и λ_n — соответственно максимальное и минимальное собственные значения преобразования $\varphi^* \circ \varphi$. Учитывая, что $\alpha_i = \sqrt{\lambda_i}$, $i = \overline{1, n}$, получаем:

$$\alpha_n \leq \frac{|\varphi(u)|}{|u|} \leq \alpha_1.$$

$\square$

Пример. Пусть φ — преобразование пространства $\mathbb{R}_2$.

Рассмотрим единичную окружность

$$S_2 = \{(x, y) \in \mathbb{R}_2 \mid x^2 + y^2 = 1\}.$$

Линейное преобразование φ переводит окружность в некоторый эллипс. Согласно теореме 27

$$\max_{u \in \mathbb{R}_2 \setminus \{0\}} \frac{|\varphi(u)|}{|u|} = \alpha_1,$$

$$\min_{u \in \mathbb{R}_2 \setminus \{0\}} \frac{|\varphi(u)|}{|u|} = \alpha_2,$$

где α_1 и α_2 — сингулярные числа преобразования φ, причем $\alpha_1 > \alpha_2$. Это значит, что α_1 и α_2 — полуоси эллипса. То есть

$$x^2 + y^2 = 1 \xrightarrow{\varphi} \frac{x^2}{\alpha_1^2} + \frac{y^2}{\alpha_2^2} = 1.$$

Теорема 28. *Для любого линейного преобразования $\varphi : \mathcal{E}_n \to \mathcal{E}_n$*

$$|\det \varphi| = \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_n,$$

где $\alpha_1, \ldots, \alpha_n$ — сингулярные числа преобразования φ.

Доказательство. Рассмотрим полярное разложение преобразования φ: $\varphi = p \circ s$.

Пусть $E = \{e_1, \ldots, e_n\}$ — первый сингулярный базис преобразования φ.

Тогда

$$\tau_E(\varphi) = A, \quad \tau_E(p) = P, \quad \tau_E(s) = S = \begin{pmatrix} \alpha_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \alpha_n \end{pmatrix},$$

где $\alpha_1, \ldots, \alpha_n$ — сингулярные числа преобразования φ.

Используя, что

$$A = PS, \quad \text{а} \quad |\det P| = 1,$$

получаем:

$$|\det\varphi| = |\det A| = |\det(P\cdot S)| = |\det P|\cdot|\det S| = \det S = \alpha_1,\ldots,\alpha_n.$$

□

Теорема 29. *Для любого линейного преобразования $\varphi : \mathcal{E}_n \to \mathcal{E}_n$*

$$|\operatorname{tr}\varphi| \leq \alpha_1 + \cdots + \alpha_n,$$

где $\alpha_1,\ldots,\alpha_n$ — сингулярные числа преобразования φ.

Доказательство. Пусть $E = \{e_1,\ldots,e_n\}$ — первый сингулярный базис преобразования φ. Тогда

$$\tau_E(\varphi) = A, \quad \tau_E(p) = P, \quad \tau_E(s) = S = \begin{pmatrix} \alpha_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \alpha_n \end{pmatrix},$$

где $\alpha_1,\ldots,\alpha_n$ — сингулярные числа преобразования φ.

Используя, что $A = PS$, а $|p_{ij}| \leq 1,\ i,j = \overline{1,n}$, получаем:

$$|\operatorname{tr}\varphi| = |\operatorname{tr}A| = |\operatorname{tr}(PS)| =$$

$$= \left|\sum_{i=1}^{n} p_{ii}\alpha_i\right| \leq \sum_{i=1}^{n}|p_{ii}||\alpha_i| \leq \sum_{i=1}^{n}\alpha_i.$$

□

Теорема 30. *Образ любого вектора при линейном преобразовании $\varphi\colon \mathcal{E}_n \to \mathcal{E}_n$ может быть найден, если известны сингулярные числа и сингулярные базисы этого преобразования.*

Доказательство. Пусть $E = \{e_1,\ldots,e_n\}$ — первый сингулярный базис преобразования φ, $H = \{h_1,\ldots,h_n\}$ — второй сингулярный базис преобразования φ, $\{\alpha_1,\ldots,\alpha_n\}$ — сингулярные числа преобразования φ. Рассмотрим вектор $u \in \mathcal{E}_n$. Пусть $u = x_1e_1 + \cdots + x_ne_n$. Заметим, что

$$(u,e_i) = x_1(e_1,e_i) + \cdots + x_n(e_n,e_i) = x_i, \quad i = \overline{1,n},$$

так как базис E ортонормированный.

Таким образом,

$$u = \sum_{i=1}^{n}(u,e_i)e_i.$$

Тогда

$$\varphi(u) = (p \circ s)(u) = p(s(u)) = p\left(s\left(\sum_{i=1}^{n}(u, e_i)e_i\right)\right) =$$
$$= [s(e_i) = \alpha_i e_i,\ i = \overline{1, n}] \ = p\left(\sum_{i=1}^{n}\alpha_i(u, e_i)e_i\right) =$$
$$= [p(e_i) = h_i,\ \overline{i = 1, n}] = \sum_{i=1}^{n}\alpha_i(u, e_i)h_i.$$

Отсюда следует, что

$$\varphi(u) = \sum_{i=1}^{n}\alpha_i(u, e_i)h_i.$$

□

4.12. Приведение матрицы линейного преобразования к треугольному виду

Пусть $\varphi : V_n \to V_n$, $\psi : V_n \to V_n$ — преобразования векторного пространства V_n. Преобразования φ и ψ называются *перестановочными* (или коммутативными), если

$$\varphi \circ \psi = \psi \circ \varphi.$$

Теорема 31. *Подпространства* $\operatorname{Ker}\psi, \operatorname{Im}\ \psi$ *инвариантны относительно* φ*, если преобразования* φ *и* ψ *перестановочны.*

Доказательство. Пусть $u \in \operatorname{Im}\psi \Rightarrow \exists v \in V_n$, такой, что $u = \psi(v)$, тогда $\varphi(u) = \varphi(\psi(v)) = \psi(\varphi(v))$, так как φ и ψ — перестановочны $\Rightarrow \varphi(u) \in \operatorname{Im}\psi$.

Аналогично с ядром:

$$u \in \operatorname{Ker}\psi \Rightarrow \psi(u) = 0 \Rightarrow \varphi(\psi(u)) = 0 \Rightarrow$$
$$\psi(\varphi(u)) = 0 \ \Rightarrow \ \varphi(u) \in \operatorname{Ker}\psi.$$

□

Множество собственных векторов преобразования φ, соответствующих собственному значению λ, образует собственное подпространство

$$V_\lambda = \operatorname{Ker}(\varphi - \lambda\,\mathrm{id}),$$

где $\mathrm{id} : V_n \to V_n$ — тождественное преобразование, то есть $\mathrm{id}(u) = u$, $\forall u \in V_n$.

Теорема 32. *Если $\varphi \circ \psi = \psi \circ \varphi$, то* $\mathrm{Ker}(\psi - \lambda\,\mathrm{id})$ *и* $\mathrm{Im}(\psi - \lambda\,\mathrm{id})$ *инвариантны относительно φ для любого числа λ.*

Доказательство. Полностью следует из теоремы 31. Действительно, если

$$\varphi \circ \psi = \psi \circ \varphi,$$

то

$$\varphi \circ (\psi - \lambda\,\mathrm{id}) = \varphi \circ \psi - \lambda\varphi, \quad (\psi - \lambda\,\mathrm{id}) \circ \varphi = \psi \circ \varphi - \lambda\varphi,$$

то есть преобразования $\psi - \lambda\,\mathrm{id}$ и φ — перестановочны. □

Следствие. $\mathrm{Ker}(\varphi - \lambda\,\mathrm{id})$ *и* $\mathrm{Im}(\varphi - \lambda\,\mathrm{id})$ *инвариантны относительно φ.*

В дальнейшем будем считать, что V_n — комплексное векторное пространство.

Теорема 33. *Любое линейное преобразование комплексного векторного пространства V_n обладает собственным вектором. В любом инвариантном подпространстве содержится собственный вектор.*

Доказательство. Собственные значения φ — это корни характеристического уравнения, но над $\mathbb{C}$ характеристическое уравнение имеет хотя бы один корень, а раз есть собственное значение, найдутся и собственные векторы.

Если $V_k \subset V_n$ — инвариантное подпространство, то есть $\varphi(V_k) \subset V_k$, мы можем рассматривать $\varphi|V_k$. Применяем первое утверждение к ограничению $\varphi|V_k$. □

Теорема 34. *Если $\varphi \circ \psi = \psi \circ \varphi$, то преобразования φ и ψ имеют общий собственный вектор.*

Доказательство. Пусть λ — собственное значение ψ, тогда множество собственных векторов $V_\lambda = \mathrm{Ker}(\psi - \lambda\,\mathrm{id})$ инвариантно относительно φ по утверждению теоремы 32, то есть $\varphi(\mathrm{Ker}(\psi - \lambda\,\mathrm{id})) \subset \mathrm{Ker}(\psi - \lambda\,\mathrm{id})$, следовательно, содержит собственный вектор φ, но все векторы $\mathrm{Ker}(\psi - \lambda\,\mathrm{id})$ — собственные для ψ. Значит, есть общий собственный вектор. □

Теорема 35. *В любом k-мерном инвариантном подпространстве линейного преобразования φ комплексного векторного пространства*

V_n можно найти $(k-1)$-мерное инвариантное подпространство, то есть существует $V_{k-1} \subset V_k$, такое, что $\varphi(V_{k-1}) \subset V_{k-1}$.

Доказательство. Сначала докажем, что для преобразования φ пространства V_n существует $(n-1)$-мерное инвариантное подпространство V_{n-1}.

Пусть λ — собственное значение преобразования φ. Тогда собственное подпространство $\operatorname{Ker}(\varphi - \lambda\,\mathrm{id})$ инвариантно относительно φ и

$$\dim \operatorname{Ker}(\varphi - \lambda\,\mathrm{id}) \geq 1,$$

значит,

$$\dim \operatorname{Im}(\varphi - \lambda\,\mathrm{id}) \leq n-1,$$

причем $\operatorname{Im}(\varphi - \lambda\,\mathrm{id})$ инвариантно относительно φ. Если

$$\dim \operatorname{Im}(\varphi - \lambda\,\mathrm{id}) = n-1,$$

то утверждение доказано; если

$$\dim \operatorname{Im}(\varphi - \lambda\,\mathrm{id}) < n-1,$$

то покажем, что любое подпространство V_{n-1}, такое, что

$$\operatorname{Im}(\varphi - \lambda\,\mathrm{id}) \subset V_{n-1},$$

будет инвариантно относительно φ.

Берем вектор $u \in V_{n-1}$. Нужно доказать, что $\varphi(u) \in V_{n-1}$. Заметим, что

$$(\varphi - \lambda\,\mathrm{id})(u) \in V_{n-1},$$

(так как $\operatorname{Im}(\varphi - \lambda\,\mathrm{id})$ инвариантно относительно φ), $\lambda u \in V_{n-1}$.

Тогда

$$(\varphi - \lambda\,\mathrm{id})(u) = \varphi(u) - \lambda u,$$
$$\varphi(u) = (\varphi - \lambda\,\mathrm{id})(u) + \lambda u.$$

Отсюда следует, что $\varphi(u) \in V_{n-1}$. Значит, мы в любом пространстве можем найти $(n-1)$-мерное инвариантное подпространство. Чтобы доказать теорему для любых k, достаточно применить данное утверждение к $\varphi \mid V_k$, где V_k — любое инвариантное относительно φ подпространство.

$\square$

Пользуясь теоремой 35, можно получить цепочку вложенных друг в друга инвариантных подпространств:

$$V_1 \subset V_2 \subset \ldots V_{n-1} \subset V_n, \quad \varphi(V_i) \subset V_i,\ i = \overline{1,n}. \tag{3}$$

Теорема 36. *Для любого преобразования $\varphi : V_n \to V_n$ существует базис $E = \{e_1, \dots, e_n\}$, в котором его матрица является верхней треугольной:*

$$\tau_E(\varphi) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ 0 & 0 & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Доказательство. Пусть имеется цепочка подпространств (3). Базис E выбираем следующим образом: $e_1 \in V_1$, тогда $\varphi(e_1) \in V_1$, значит, $\varphi(e_1) = a_{11}e_1$. Пусть $e_2 \in V_2$, $\varphi(e_2) \in V_2 \Rightarrow \varphi(e_2)$ выражаем через базис e_1 и e_2; $\varphi(e_2) = a_{12}e_1 + a_{22}e_2$ и так далее. Тогда на k-м шаге получаем $\{e_1, \dots, e_k\}$ — базис V_k,

$$\varphi(e_k) \in V_k,$$

причем

$$\varphi(e_k) = a_{1k}e_1 + a_{2k}e_2 + \dots + a_{kk}e_k.$$

Отсюда следует, что матрица $\tau_E(\varphi)$ — треугольная. □

Теорема 37. *Для любого преобразования φ унитарного пространства U_n существует ортонормированный базис, в котором его матрица является верхней треугольной.*

Доказательство. Берем базис E из предыдущей теоремы и применим к нему процесс ортогонализации Грама – Шмидта. Так как к любому вектору e_k мы прибавляем слагаемое из линейной оболочки векторов $e_1, \dots, e_{k-1}$, то мы не выходим за пределы V_k. □

4.13. Нормальные преобразования

Преобразование $\varphi : V_n \to V_n$ называется *нормальным*, если оно перестановочно со своим сопряженным, то есть $\varphi^* \circ \varphi = \varphi \circ \varphi^*$.

Теорема 38. *В унитарном пространстве нормальные преобразования и только они обладают ортонормированным базисом из собственных векторов.*

Доказательство. Согласно теореме 12

$$\mathrm{Ker}(\varphi^* \circ \varphi) = \mathrm{Ker}\,\varphi, \quad \mathrm{Im}(\varphi^* \circ \varphi) = \mathrm{Im}\,\varphi^*;$$

$$\operatorname{Ker}(\varphi\circ\varphi^*) = \operatorname{Ker}\varphi^*, \quad \operatorname{Im}(\varphi\circ\varphi^*) = \operatorname{Im}\varphi.$$

Отсюда следует, что для нормального преобразования

$$\operatorname{Ker}\varphi = \operatorname{Ker}\varphi^*, \ \ \operatorname{Im}\varphi = \operatorname{Im}\varphi^*.$$

□

Докажем вспомогательное утверждение.

Лемма. *Для любого преобразования* $\varphi : V_n \to V_n$ $(\operatorname{Im}\varphi)^\perp = \operatorname{Ker}\varphi^*$.

Доказательство. $\operatorname{Ker}\varphi^* = \{f \in V_n^* \mid \varphi^*(f) = 0\}$. Возьмем любую функцию $f \in \operatorname{Ker}\varphi$, получаем, что $< \varphi^*(f), u >= 0$, при любом $u \in V_n$. Используя

$$\langle\varphi^*(f), u\rangle = \langle f, \varphi(u)\rangle,$$

получаем

$$\langle f, \varphi(u)\rangle = 0, \ \forall u \in V_n \ \Rightarrow f \in (\operatorname{Im}\varphi)^\perp.$$

Отсюда следует, что

$$\operatorname{Ker}\varphi^* \subset (\operatorname{Im}\varphi)^\perp.$$

Утверждение $(\operatorname{Im}\varphi)^\perp \subset \operatorname{Ker}\varphi^*$ доказывается аналогично. □

В нашем случае $(\operatorname{Im}\varphi)^\perp = \operatorname{Ker}\varphi$. Это означает, что $U_n = \operatorname{Ker}\varphi \bigoplus \operatorname{Im}\varphi$.

Утверждение теоремы докажем по индукции.

При $n = 1$ существование ортонормированного базиса пространства U_1 очевидно.

Рассмотрим нормальное преобразование $\varphi : U_n \to U_n$. Пусть λ — собственное значение φ. Преобразование φ — нормальное, значит, преобразование $\varphi - \lambda\,\mathrm{id}$ — тоже нормальное.

Действительно,

$$\begin{aligned}(\varphi - \lambda\,\mathrm{id})^* \circ (\varphi - \lambda\,\mathrm{id}) &= (\varphi^* - \overline{\lambda}\,\mathrm{id}) \circ (\varphi - \lambda\,\mathrm{id}) = \\ &= \varphi^*\circ\varphi - \lambda\varphi^* - \overline{\lambda}\varphi + \lambda\overline{\lambda}\,\mathrm{id}\,. \\ (\varphi - \lambda\,\mathrm{id}) \circ (\varphi^* - \lambda\,\mathrm{id})^* &= (\varphi - \overline{\lambda}\,\mathrm{id}) \circ (\varphi^* - \overline{\lambda}\,\mathrm{id}) = \\ &= \varphi\circ\varphi^* - \lambda\varphi^* - \overline{\lambda}\varphi + \lambda\overline{\lambda}\,\mathrm{id}\,.\end{aligned}$$

Тогда

$$U_n = \operatorname{Ker}(\varphi - \lambda\,\mathrm{id}) \bigoplus \operatorname{Im}(\varphi - \lambda\,\mathrm{id}).$$

Ядро состоит из собственных векторов преобразования φ, причем

$$\dim\operatorname{Ker}(\varphi - \lambda\,\mathrm{id}) \geq 1 \Rightarrow \dim\operatorname{Im}(\varphi - \lambda\,\mathrm{id}) \leq n - 1.$$

К пространству $\operatorname{Im}(\varphi - \lambda\,\mathrm{id})$ применим предположение индукции, согласно которому в нем существует ортонормированный базис, состоящий из собственных векторов преобразования φ.

Следовательно, в U_n можно выбрать ортонормированный базис из собственных векторов преобразования φ.

Осталось показать, что если преобразование $\varphi : U_n \to U_n$ обладает ортонормированным базисом E из собственных векторов, то оно нормально. Рассмотрим

$$\tau_E(\varphi) = A = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \lambda_n \end{bmatrix},$$

где $\lambda_1, \ldots, \lambda_n$ — собственные значения φ.

Ясно, что $A^*A = AA^*$, тогда $\varphi^* \circ \varphi = \varphi \circ \varphi^*$, следовательно, φ — нормально.

4.14. Свойства нормальных преобразований

Важными примерами нормальных преобразований являются самосопряженные и унитарные. Напомним, что собственные значения самосопряженного преобразования унитарного пространства действительны. С другой стороны, любое нормальное преобразование с действительными собственными значениями является самосопряженным.

В самом деле, пусть $\varphi(e_i) = \lambda_i e_i,\ i = \overline{1,n},\ \lambda_i \in \mathbb{R}$. Тогда

$$\tau_E(\varphi) = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \lambda_n \end{bmatrix} = A, \quad \tau_E(\varphi^*) = A^* = A.$$

Следовательно, самосопряженное преобразование есть в точности нормальное преобразование с действительными собственными значениями.

Рассмотрим унитарное преобразование $\varphi : U_n \to U_n$. Пусть $\varphi(u) = \lambda u$. Тогда $(u, u) = (\varphi(u), \varphi(u)) = (\lambda u, \lambda u) = \lambda\overline{\lambda}(u, u) \Rightarrow \lambda\overline{\lambda} = 1 \Rightarrow |\lambda| = 1$.

Рассмотрим преобразование $\varphi : U_n \to U_n$, такое, что

$$\varphi^* \circ \varphi = \varphi \circ \varphi^*,$$
$$\varphi(e_i) = \lambda_i e_i,\ |\lambda_i| = 1,\ i = \overline{1,n},$$

где $E = \{e_1, \ldots, e_n\}$ — ортонормированный базис U_n, состоящий из собственных векторов φ.

Пусть

$$\tau_E(\varphi) = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{bmatrix} = A,$$

$$\tau_E(\varphi^*) = \begin{bmatrix} \overline{\lambda_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \overline{\lambda_n} \end{bmatrix} = A^*.$$

Тогда

$$AA^* = A^*A = \begin{bmatrix} \lambda_1\overline{\lambda_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_1\overline{\lambda_n} \end{bmatrix} = E_n,$$

то есть φ — унитарное преобразование.

Значит, унитарное преобразование есть в точности нормальное преобразование с собственными значениями, по модулю равными единице.

Пусть $\varphi : U_n \to U_n$, причем $\varphi^* \circ \varphi = \varphi \circ \varphi^*$. По теореме 38 отсюда следует, что существует ортонормированный базис $E = \{e_1, \dots, e_n\}$ пространства U_n, такой, что $\varphi(e_i) = \lambda_i e_i$, $i = \overline{1,n}$. Представим собственные значения φ в экспоненциальной форме: $\lambda_k = r_k e^{i\varphi_k}$, $k = \overline{1,n}$. Тогда

$$\tau_E(\varphi) = A = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{bmatrix} = \begin{bmatrix} r_1 e^{i\varphi_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & r_n e^{i\varphi_n} \end{bmatrix} =$$

$$= \begin{bmatrix} r_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & r_n \end{bmatrix} \cdot \begin{bmatrix} e^{i\varphi_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & e^{i\varphi_n} \end{bmatrix} = SP = PS, \tag{4}$$

где

$$S = \begin{bmatrix} r_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & r_n \end{bmatrix}, \quad P = \begin{bmatrix} e^{i\varphi_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & e^{i\varphi_n} \end{bmatrix}.$$

Заметим, что

$$P^*P = \begin{bmatrix} e^{-i\varphi_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & e^{-i\varphi_n} \end{bmatrix} \begin{bmatrix} e^{i\varphi_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & e^{i\varphi_n} \end{bmatrix} = E_n,$$

$$S = S^*, \quad r_i \geq 0, \quad i = \overline{1, n}.$$

Пусть $S = \tau_E(s)$, $P = \tau_E(p)$. Мы показали, что любое нормальное преобразование φ можно представить в виде композиции $\varphi = p \circ s = s \circ p$, где p — унитарное преобразование U_n, а s — самосопряженное преобразование U_n с неотрицательными собственными значениями. Значит, (4) — полярное разложение φ.

Теорема 39. *Для того чтобы преобразование $\varphi : U_n \to U_n$ было нормальным, необходимо и достаточно, чтобы сомножители в некотором полярном разложении φ были перестановочны, то есть*

$$\varphi^* \circ \varphi = \varphi \circ \varphi^* \Leftrightarrow \varphi = p \circ s = s \circ p.$$

Доказательство. Необходимость доказана выше. Проверим достаточность. Пусть $\varphi = p \circ s = s \circ p$. Заметим, что $p^* = p^{-1}$ (так как p — унитарное преобразование), $\varphi^* = (p \circ s)^* = (s \circ p)^*$. Значит, $\varphi^* = s \circ p^{-1} = p^{-1} \circ s$.

Тогда

$$\varphi \circ \varphi^* = (s \circ p) \circ (p^{-1} \circ s) = s \circ p \circ p^{-1} \circ s = s \circ p^{-1} \circ p \circ s = \varphi^* \circ \varphi.$$

Следовательно, φ — нормальное преобразование. □

Пусть $\varphi : U_n \to U_n$ — преобразование пространства U_n. Рассмотрим преобразования $\psi = \frac{1}{2}(\varphi + \varphi^*)$ и $\eta = \frac{1}{2i}(\varphi - \varphi^*)$ пространства U_n. Заметим, что $\varphi = \psi + i\eta$. Преобразование ψ назовем *действительной частью* φ, преобразование η — *мнимой частью* φ.

Заметим, что

$$\psi^* = (\frac{1}{2}(\varphi^* + \varphi))^* = \frac{1}{2}(\varphi^* + \varphi) = \psi,$$

$$\eta^* = \left(\frac{1}{2i}(\varphi - \varphi^*)\right)^* = -\frac{1}{2i}(\varphi^* - \varphi) = \eta,$$

то есть преобразования ψ и η — самосопряженные.

С другой стороны, пусть имеются любые самосопряженные преобразования ψ и η пространства U_n. Тогда они являются соответственно действительной и мнимой частью преобразования $\varphi = \psi + i\eta$.

Действительно, $\varphi^* = \psi^* - i\eta^* = \psi - i\eta$. Отсюда следует, что
$$\psi = \frac{1}{2}(\varphi + \varphi^*), \quad \eta = \frac{1}{2i}(\varphi - \varphi^*).$$

Теорема 40. *Пусть*
$$\psi = \psi^*, \quad \eta = \eta^*,$$
где $\psi, \eta : U_n \to U_n$ *— преобразования пространства* U_n. *Следующие утверждения эквивалентны:*

1) $\varphi = \psi + i\eta$*— нормальное преобразование* U_n*;*
2) *преобразования* ψ *и* η *обладают общим ортонормированным базисом, состоящим из собственных векторов;*
3) $\psi \circ \eta = \eta \circ \psi$.

Доказательство. Покажем, что 1) $\Leftrightarrow$ 3).
$$\psi \circ \eta - \eta \circ \psi = \left(\frac{1}{2}(\varphi^* + \varphi)\right) \circ \left(\frac{1}{2i}(\varphi^* - \varphi)\right) -$$
$$- \left(\frac{1}{2i}(\varphi^* - \varphi)\right) \circ \left(\frac{1}{2}(\varphi^* + \varphi)\right) =$$
$$= \frac{1}{4i}(\varphi^* \circ \varphi^* + \varphi \circ \varphi^* - \varphi \circ \varphi - \varphi^* \circ \varphi +$$
$$+ \varphi \circ \varphi + \varphi \circ \varphi^* - \varphi^* \circ \varphi - \varphi^* \circ \varphi^*) =$$
$$= \frac{1}{2i}(\varphi \circ \varphi^* - \varphi^* \circ \varphi).$$

Докажем, что 1) $\Rightarrow$ 2). Пусть преобразование φ нормально. Рассмотрим $E = \{e_1, \ldots, e_2\}$ — ортонормированный базис пространства U_n, состоящий из собственных векторов преобразования φ, $\varphi(e_i) = \lambda_i e_i$, $\overline{1, n}$. Пусть
$$\tau_E(\varphi) = A = \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & \lambda_n \end{bmatrix}.$$
Тогда
$$\tau_E(\psi) = \frac{1}{2}(A + A^*), \quad \tau_E(\eta) = \frac{1}{2i}(A - A^*).$$
В базисе E матрицы преобразований η и ψ — диагональные. Следовательно, векторы E являются собственными и для η, и для ψ. Значит, η и ψ обладают общим ортонормированным базисом, состоящим из собственных векторов.

Докажем, что 2) $\Rightarrow$ 3). Пусть $E = \{e_1, \dots, e_2\}$ — общий ортонормированный базис, состоящий из собственных векторов преобразований η и ψ. Тогда

$$\tau_E(\psi) = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \hdotsfor{3} \\ 0 & \dots & \lambda_n \end{bmatrix} = B,$$

$$\tau_E(\eta) = \begin{bmatrix} \mu_1 & \dots & 0 \\ \hdotsfor{3} \\ 0 & \dots & \mu_n \end{bmatrix} = C.$$

Видим, что $BC = CB$. Откуда следует, что η и ψ — перестановочные преобразования, то есть $\psi \circ \eta = \eta \circ \psi$.

$\square$

5. НОРМЫ ВЕКТОРОВ И МАТРИЦ

5.1. Определение нормированного пространства

Пусть есть векторное пространство V_n (действительное или комплексное). Функция $\|\cdot\| : V_n \to \mathbb{R}$, удовлетворяющая свойствам

1) $\|u\| \geq 0,\ \|u\| = 0 \Leftrightarrow u = 0$;
2) $\|\alpha u\| = |\alpha| \|u\|, \forall \alpha \in P, \quad \forall u \in V_n, \quad P \in \{\mathbb{R}, \mathbb{C}\}$;
3) $\|u + v\| \leq \|u\| + \|v\|,\ \forall u, v \in V_n$,

называется нормой.

Пара $(V_n, \|\cdot\|)$ называется нормированным пространством. Отметим еще одно свойство нормы:

$$\|u - v\| \geq |\|u\| - \|v\||,\ \forall u, v \in V_n.$$

Доказательство.

$$\|u\| - \|v\| = \|u - v + v\| - \|v\| \leq \|u - v\| + \|v\| - \|v\| = \|u - v\|.$$

Аналогично доказывается, что $\|v\| - \|u\| \leq \|u - v\|$. $\square$

В нормированном пространстве мы можем определить расстояние между векторами u и v: $\rho(u, v) = \|u - v\|$.

Свойства расстояния:

1) $\rho(u, v) \geq 0$; $\rho(u, v) = 0 \Leftrightarrow u = v$;
2) $\rho(u, v) = \rho(v, u)$;
3) $\rho(u, v) \leq \rho(u, w) + \rho(w, v)$ — неравенство треугольника.

Действительно,

$$\rho(u, v) = \|u - v\| = \|u - w + w - v\| \leq \|u - w\| + \|w - v\| = \rho(u, w) + \rho(w, v).$$

Для любого вектора u можно определить его ε-окрестность:

$$O(u, \varepsilon) = \{v \in V_n | \rho(u, v) < \varepsilon\}.$$

Пусть есть последовательность векторов (u_n). Вектор u называется пределом последовательности

$$(u_n) : \lim_{n\to\infty} u_n = u,$$

если $\forall\varepsilon > 0\ \exists N_\varepsilon,\ \forall n > N_\varepsilon,\ u_n \in O(u,\varepsilon)\ (\|u_n - u\| < \varepsilon)$.

Таким образом, определив предел последовательных векторов, мы можем перенести на нормированные пространства все понятия элементарного математического анализа (но в нашу задачу это не входит).

Рассмотрим наиболее употребительные нормы векторов. Пусть $u = (x_1, \ldots, x_n)$, $u \in P_n$.

(1) Октаэдрическая норма:

$$\|u\|_0 = \sum_{i=1}^{n} |x_i|.$$

(2) Евклидова норма (в комплексном пространстве — унитарная):

$$\|u\|_E = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

(3) Норма Гельдера:

$$\|u\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p},\ p > 1.$$

(4) Кубическая норма:

$$\|u\|_k = \max_{1\le i\le n} |x_i|.$$

Заметим, что в евклидовом (унитарном) векторном пространстве длина вектора совпадает с его евклидовой (унитарной) нормой.

$$|u| = \sqrt{(u,u)} = \sqrt{\sum_{i=1}^{n} x_i\overline{x_i}} = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \|u\|_E.$$

Несложно проверить, что октаэдрическая и кубическая нормы удовлетворяют всем аксиомам нормы. Нормой Гельдера мы не будем пользоваться, поэтому и аксиом проверять не будем.

Пример. $u \in \mathbb{C}_3,\ u = (2i,\ 1-i,\ -5)$. Тогда

$$\|u\|_0 = 2 + \sqrt{2} + 5 = 7 + \sqrt{2},$$
$$\|u\|_E = \sqrt{4+2+25} = \sqrt{31},$$
$$\|u\|_k = 5.$$

Определим единичную сферу в нормированном пространстве V_n :

$$S_n = \{u \in V_n \mid \|u\| = 1\}.$$

Теорема 1. *Норма любого вектора может быть вычислена, если известна единичная сфера.*

Доказательство. Рассмотрим V_1 — одномерное подпространство V_n. Пусть $u \in V_1$. Покажем, что любое одномерное подпространство пересекается с единичной сферой S_n. Вектор $u_0 = \frac{1}{\|u\|}u \in V_1$, но $\|u_0\| = \|\frac{1}{\|u\|},\ u\| = \frac{1}{\|u\|}\|u\| = 1$, значит, $u_0 \in S_n$. Мы показали, что $u_0 \in V_1 \bigcap S_n$.

Берем любой вектор $v \in V_1$. В пространстве V_1 есть вектор $v_0 \in V_1 \bigcap S$, а так как пространство одномерно, то $v = \alpha v_0$. Отсюда следует, что

$$\|v\| = \|\alpha v_0\| = |\alpha|\|v_0\| = |\alpha|.$$

□

Пример. Рассмотрим плоскость $\mathbb{R}_2$. На плоскости единичную сферу можно назвать единичной окружностью.

Рассмотрим единичную окружность для основных векторных норм:

(1) Евклидова норма: $S_2 = \{(x, y) \in \mathbb{R}_2 \mid x^2 + y^2 = 1\}$.
(2) Октаэдрическая норма: $S_2 = \{(x, y) \in \mathbb{R}_2 \mid |x| + |y| = 1\}$.
(3) Кубическая норма: $S_2 = \{(x, y) \in \mathbb{R}_2 \mid \max\{|x|, |y|\} = 1$.

Заметим, что октаэдрическая и кубическая векторные нормы называются так потому, что при $n = 3$ соответствующие единичные сферы — это октаэдр и куб.

Пусть векторы $u_1, u_2 \in V_n$, тогда множество векторов

$$u = \alpha u_1 + (1-\alpha)u_2,\ \alpha \in [0, 1]$$

называется отрезком с концами u_1 и u_2.

Теорема 2. *Если концы отрезка принадлежат единичному шару, то ему принадлежит и весь отрезок (единичный шар — это множество векторов*

$$\{u \in V_n \mid \|u\| \le 1\}).$$

Доказательство. Пусть $\|u_1\| \le 1,\ \|u_2\| \le 1$. Тогда

$$\|u\| = \|\alpha u_1 + (1-\alpha)u_2\| \le \|\alpha u_1\| + \|(1-\alpha)u_2\| =$$
$$= \alpha\|u_1\| + (1-\alpha)\|u_2\| \le \alpha + (1-\alpha) = 1.$$

Следовательно,

$$\|u\| \le 1,\ \forall u \in [u_1, u_2].$$

$\square$

5.2. Эквивалентность норм

Пусть $\|\cdot\|$ — норма в векторном пространстве V_n, $\alpha > 0$. Тогда несложно проверить, что функция $\|\cdot\|_1 = \alpha\|\cdot\|$, такая, что $\|u\|_1 = \alpha\|u\|$, для любого $u \in V_n$ также является нормой.

Пусть есть две нормы и для любого вектора $u \in V_n : \|u\|_1 \le \|u\|_2$. Тогда говорят, что норма $\|\cdot\|_2$ мажорирует норму $\|\cdot\|_1$, и записывают:

$$\|\cdot\|_1 \le \|\cdot\|_2.$$

Пусть $O_1(u,\varepsilon)$ есть ε-окрестность вектора $u \in V_n$ относительно нормы $\|\cdot\|_1$, то есть $v \in O_1(u,\varepsilon)$, если $\|u - v\|_1 < \varepsilon$.

Тогда

$$O_2(u,\varepsilon) \subset O_1(u,\varepsilon).$$

Действительно, если вектор $v \in O_2(u,\varepsilon)$, то $\|u-v\|_2 < \varepsilon$. Но

$$\|u-v\|_1 \le \|u-v\|_2 < \varepsilon,\ \text{следовательно},\ v \in O_1(u,\varepsilon).$$

Две нормы $\|\cdot\|_1$ и $\|\cdot\|_2$ называются *эквивалентными*, если существуют такие $\alpha_1 > 0$ и $\alpha_2 > 0$, что

$$\alpha_1\|\cdot\|_2 \le \|\cdot\|_1 \le \alpha_2\|\cdot\|_2.$$

Покажем, что отношение эквивалентности норм действительно является отношением эквивалентности.

(1) Рефлексивность:

$$\|\cdot\|_1 \cong |\cdot\|_1.$$

Достаточно взять $\alpha_1 = \alpha_2 = 1$.

(2) Симметричность:

$$\|\cdot\|_1 \cong \|\cdot\|_2 \Rightarrow \|\cdot\|_2 \cong |\cdot\|_1.$$

Это верно, так как

$$\frac{1}{\alpha_2}\|\cdot\|_1 \leq |\cdot\|_2 \leq \frac{1}{\alpha_1}\|\cdot\|_1.$$

(3) Транзитивность:

$$\|\cdot\|_1 \cong \|\cdot\|_2,\ \|\cdot\|_2 \cong \|\cdot\|_3 \Rightarrow \|\cdot\|_1 \cong \|\cdot\|_3.$$

Пусть

$$\alpha_1\|\cdot\|_1 \leq \|\cdot\|_2 \leq \alpha_2\|\cdot\|_1,\ \beta_1\|\cdot\|_2 \leq \|\cdot\|_3 \leq \beta_2\|\cdot\|_2 \Rightarrow$$
$$\Rightarrow \alpha_1\beta_1\|\cdot\|_1 \leq \|\cdot\|_3 \leq \alpha_2\beta_2\|\cdot\|_1,$$

то есть $\|\cdot\|_1 \cong \|\cdot\|_3$.

Теорема 3. *Пусть нормы $\|\cdot\|_1$ и $\|\cdot\|_2$ — эквивалентны, то есть существуют числа $\alpha_1 > 0,\ \alpha_2 > 0$, такие, что*

$$\alpha_1\|\cdot\|_2 \leq \|\cdot\|_1 \leq \alpha_2\|\cdot\|_2.$$

Тогда

$$O_2\left(u, \frac{\varepsilon}{\alpha_2}\right) \subseteq O_1(u,\varepsilon) \subseteq O_2\left(u, \frac{\varepsilon}{\alpha_1}\right),$$

для любого вектора $u \in V_n$.

Доказательство. Пусть $v \in O_1(u,\varepsilon) \Rightarrow \|u-v\|_1 < \varepsilon$.

Тогда

$$\alpha_1\|u-v\|_2 \leq \|u-v\|_1 < \varepsilon.$$

Отсюда следует, что

$$\|u-v\|_2 < \frac{\varepsilon}{\alpha_1} \Rightarrow v \in O_2\left(u, \frac{\varepsilon}{\alpha_1}\right),$$

значит,

$$O_1(u,\varepsilon) \subseteq O_2\left(u, \frac{\varepsilon}{\alpha_1}\right).$$

Пусть

$$v \in O_2\left(u, \frac{\varepsilon}{\alpha_2}\right) \Rightarrow \|u-v\|_2 < \frac{\varepsilon}{\alpha_2}.$$

Тогда

$$\|u-v\|_1 \leq \alpha_2\|u-v\|_2 < \alpha_2\frac{\varepsilon}{\alpha_2} = \varepsilon \Rightarrow$$

$$\Rightarrow \|u - v\|_1 < \varepsilon \Rightarrow v \in O_1(u, \varepsilon).$$

Отсюда следует, что

$$O_2\left(u, \frac{\varepsilon}{\alpha_2}\right) \subseteq O_1(u, \varepsilon).$$

$\square$

Теорема 4. *Таким образом, последовательность (u_n) сходится к вектору u по норме $\|\cdot\|_1$ тогда и только тогда, когда она сходится к вектору u по любой эквивалентной ей норме $\|\cdot\|_2$.*

Доказательство. Пусть (u_n) сходится к u по норме $\|\cdot\|_1$, то есть $\forall\varepsilon > 0,\ \exists N_\varepsilon,\ \forall u > N_\varepsilon,\ \|u_n - u\|_1 < \varepsilon, \Leftrightarrow u_n \in O_1(u, \varepsilon) \Rightarrow u_n \in O_2(u, \frac{\varepsilon}{\alpha_1})$ (по теореме 3).

Это означает, что

$$\forall\varepsilon > 0,\ \exists N_\varepsilon,\ \forall u > N_\varepsilon\ \|u_n - u\|_2 < \frac{\varepsilon}{\alpha_1}.$$

Таким образом, последовательность (u_n) сходится к вектору u по норме $\|\cdot\|_2$. Обратное утверждение следует из симметричности отношения эквивалентности норм. $\square$

Теорема 5. *Октаэдрическая, кубическая и евклидова нормы эквивалентны в арифметическом пространстве V_n.*

Доказательство. Пусть есть произвольный вектор $u = (x_1, \ldots, x_n) \in V_n$. Пусть $E = \{e_1, \ldots, e_n\}$ — канонический базис V_n. Выразим вектор u через базис E:

$$u = x_1(1, 0, \ldots, 0) + \cdots + x_n(0, 0, \ldots, 1) = x_1 e_1 + \cdots + x_n e_n.$$

Для любой нормы пространства V_n выполняется следующее:

$$\|u\| = \|\sum_{i=1}^{n} x_i e_i\| \leq \sum_{i=1}^{n} |x_i| \|e_i\|.$$

Заменим все $\|e_i\|$ на

$$\max_{1 \leq i \leq n} \|e_i\| = \alpha.$$

Тогда

$$\|u\| \leq \alpha \sum_{i=1}^{n} |x_i| = \alpha \|u\|_0. \tag{1}$$

С другой стороны,

$$\|u\| \le \max_{1\le i\le n} |x_i| \sum_{i=1}^{n} \|e_i\| = \beta \max_{1\le i\le n} |x_i| = \beta \|u\|_k, \tag{2}$$

где

$$\beta = \sum_{i=1}^{n} \|e_i\|.$$

Применив (1) к кубической норме, а (2) к октаэдрической норме, получаем, что:

$$\frac{1}{\alpha}\|u\|_k \le \|u\|_0 \le \beta\|u\|_k, \quad \text{следовательно,} \quad \|\cdot\|_k \cong \|\cdot\|_0.$$

Заметим, что

$$\max_{1\le i\le n} |x_i| \le \sqrt{\sum_{i=1}^{n} |x_i|^2} \le \sqrt{n} \max_{1\le i\le n} |x_i|.$$

Это означает, что

$$\|u\|_k \le \|u\|_E \le \sqrt{n}\|u\|_k, \quad \text{следовательно,} \quad \|\cdot\|_E \cong \|\cdot\|_k.$$

□

Примем следующую теорему без доказательства.

Теорема 6. *В конечномерном пространстве все нормы эквивалентны.*

5.3. Нормы матриц

Пусть матрица $A \in \mathbb{R}_{m,n}$. Будем говорить, что матричная норма в пространстве $\mathbb{R}_{m,n}$ согласована с векторными нормами в пространствах $\mathbb{R}_n$ и $\mathbb{R}_m$, если

$$\|AX\| \le \|A\|\|X\|, \ \forall X \in \mathbb{R}_n, \ \forall A \in \mathbb{R}_{m,n}.$$

Определим матричную функцию:

$$\|A\| = \sup \frac{\|AX\|}{\|X\|}, \ X \in \mathbb{R}_n, \ X \ne 0.$$

Заметим, что

$$\frac{\|AX\|}{\|X\|} = \left\| \frac{AX}{\|X\|} \right\| = \|AX_0\|, \ \text{где } \|X_0\| = 1. \tag{3}$$

Это означает, что функция (3) может быть записана в следующем виде:

$$\|A\| = \sup_{\|X\|=1} \|AX\|. \tag{4}$$

Теорема 7. *Функция*

$$\|A\| = \sup \frac{\|AX\|}{\|X\|} \quad (\textit{или} \quad \|A\| = \sup_{\|X\|=1} \|AX\|)$$

определена и является согласованной нормой в $\mathbb{R}_{m,n}$ для любых норм в $\mathbb{R}_m$ и $\mathbb{R}_n$.

Доказательство. Проверим ограниченность $\frac{\|AX\|}{\|X\|}$. Этим мы докажем существование точной верхней грани.

$$\|AX\| \leq \alpha\|AX\|_0 = \alpha \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j \right| \leq$$

$$\leq \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}||x_j| \leq \max_{1\leq j\leq n} |x_j| \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}| =$$

$$= \beta \max_{i\leq j\leq n} |x_j| = \beta\|X\|_k \leq \gamma\|X\| \Rightarrow$$

$$\Rightarrow \|AX\| \leq \gamma\|X\|, \ \ \forall x \in \mathbb{R}_n.$$

Откуда следует, что $\frac{\|AX\|}{\|X\|} \leq \gamma$, значит, отношение $\frac{\|AX\|}{\|X\|}$ ограничено. Покажем, что функция (3) определяет норму. Проверим выполнение аксиом нормы:

(1) $\|A\| \geq 0 \Leftrightarrow \frac{\|AX\|}{\|X\|} \geq 0$

$\|A\| = 0 \Leftrightarrow \sup \frac{\|AX\|}{\|X\|} = 0 \Leftrightarrow$

$\Leftrightarrow \|AX\| = 0, \ \ \forall X \in \mathbb{R}_n \Leftrightarrow AX = 0 \Leftrightarrow A = 0.$

(2)

$$\|\alpha A\| = \sup_{\|X\|=1} \|\alpha AX\| = \sup_{\|X\|=1} (|\alpha|\|AX\|) = |\alpha| \sup_{\|X\|=1} \|AX\| = |\alpha| \cdot \|A\|.$$

(3)

$$\|A + B\| = \sup_{\|X\|=1} \|(A + B)X\| = \sup_{\|X\|=1} \|AX + BX\| \leq \sup_{\|X\|=1} (\|AX\| +$$

$$+\|BX\|) \le \sup_{\|X\|=1} \|AX\| + \sup_{\|X\|=1} \|BX\| = \|A\| + \|B\|.$$

Докажем согласованность. По определению точной верхней грани

$$\frac{\|AX\|}{\|X\|} \le \sup \frac{\|AX\|}{\|X\|} = \|A\|,$$

следовательно,

$$\|AX\| \le \|A\| \cdot \|X\|, \ \forall X \in \mathbb{R}_n,$$

значит, норма (3) является согласованной матричной нормой. $\square$

Определение. Матричная норма в $\mathbb{R}_{m,n}$, определяемая формулой (3), называется *индуцированной* векторными нормами в пространствах $\mathbb{R}_n$ и $\mathbb{R}_m$.

Теорема 8. *Любая согласованная норма мажорирует индуцированную.* $\blacksquare$

Доказательство. Заметим, что любая согласованная норма матрицы A является верхней границей отношения $\frac{\|AX\|}{\|X\|}$, но верхняя грань — это наименьшая из верхних границ, то есть

$$\frac{\|AX\|}{\|X\|} \le \sup \frac{\|AX\|}{\|X\|} \le \|A\|.$$

Отсюда и следует утверждение теоремы. $\square$

Пусть $A \in \mathbb{R}_{n,n}$. Говорят, что норма сохраняет единицу, если $\|E_n\| = 1$.

Определение. Норма называется *кольцевой* (или *матричной*), если

$$\|A \cdot B\| \le \|A\| \cdot \|B\|,$$

для любых матриц $A, B \in \mathbb{R}_{n,n}$.

Мы будем придерживаться термина *кольцевая норма.* В противном случае может возникнуть ситуация, когда норма матрицы не является матричной нормой.

Рассмотрим некоторые свойства кольцевых норм:

$$\|A \cdot E_n\| \le \|A\| \cdot \|E_n\|.$$

Отсюда

$$\|A\| \le \|A\| \cdot \|E_n\|, \quad \text{следовательно,} \quad \|E_n\| \ge 1.$$

Ясно, что для кольцевой нормы $\|A^n\| \le \|A\|^n$. Кроме того,

$$\|A \cdot A^{-1}\| \le \|A\| \cdot \|A^{-1}\|,$$

значит,

$$\|E_n\| \le \|A\| \cdot \|A^{-1}\|,$$

тогда

$$\frac{\|E_n\|}{\|A\|} \le \|A^{-1}\|.$$

Отсюда следует, что $\|A^{-1}\| \ge \|A\|^{-1}$.

Теорема 9. *Любая индуцированная норма сохраняет единицу и является кольцевой.*

Доказательство. Заметим, что

$$\|E_n\|_I = \sup_{\|X\|=1} \frac{\|E_n X\|}{\|X\|} = \sup_{\|X\|=1} \frac{\|X\|}{\|X\|} = 1.$$

Докажем, что индуцированная норма является кольцевой.

$$\|A \cdot B\|_I = \sup_{\|X\|=1} \|(AB)X\| = \sup_{\|X\|=1} \|A(BX)\| \le$$

[так как $\|AY\| \le \|A\|_I \|Y\|$]

$$\le \sup_{\|X\|=1} (\|A\|_I \cdot \|BX\|) = \|A\|_I \sup_{\|X\|=1} \|BX\| = \|A\|_I \|B\|_I, \ \forall A, B \in \mathbb{R}_{n,n}$$

□

Пусть $A \in \mathbb{R}_{m,n}, \ B \in \mathbb{R}_{n,k}$. Тогда $AB \in \mathbb{R}_{m,k}$.

Можно обобщить кольцевое свойство норм и на этот случай. Оно имеет следующий вид:

$$\|A \cdot B\|_3 \le \|A\|_2 \cdot \|B\|_1, \quad \forall A \in \mathbb{R}_{m,n}, \ \forall B \in \mathbb{R}_{n,k},$$

где $\|\cdot\|_1$ — норма в $\mathbb{R}_{n,k}$, $\|\cdot\|_2$ — норма в $\mathbb{R}_{m,n}$, $\|\cdot\|_3$ — норма в $\mathbb{R}_{m,k}$.

5.4. Наиболее употребительные нормы матриц

Будем рассматривать два типа норм матриц: векторные (когда матрицы размеров $m \times n$ отождествляются со строкой (столбцом) длины $m \cdot n$) и индуцированные в пространстве матриц наиболее употребительными векторными нормами (октаэдрической, кубической и евклидовой).

Рассмотрим матричную норму, которая индуцируется евклидовой (унитарной) векторной нормой.

Такая норма называется *спектральной* и обозначается $\|A\|_S$.

$$\|A\|_S = \sup \frac{\|AX\|_E}{\|X\|_E}.$$

Теорема 10. *Спектральная норма матрицы не меняется при умножении этой матрицы на ортогональную (унитарную) матрицу, то есть*

$$\|A\|_S = \|A \cdot P\|_S = \|PA\|_S,$$

где P — ортогональная матрица.

Доказательство. Заметим, что умножение столбца $X \in P_n$ на ортогональную матрицу P не меняет его длины (евклидовой нормы):

$$\|PX\|_E = \|X\|_E, \ \forall X \in P_n,$$

тогда

$$\|A \cdot P\|_S \leq \|A\|_S \|P\|_S = \|A\|_S,$$

так как любая индуцированная норма является кольцевой. С другой стороны,

$$\|A\|_S = \|AP \cdot P^\top\|_S \leq \|AP\|_S \|P^\top\|_S = \|AP\|_S,$$

так как

$$\|P^\top\|_S = \sup \frac{\|P^\top X\|_E}{\|X\|_E} = 1.$$

Отсюда следует:

$$\|A\|_S = \|A \cdot P\|_S.$$

Утверждение $\|PA\|_S = \|A\|_S$ доказывается аналогично. □

Воспользуемся сингулярным разложением матрицы $A \in P_{m,n}$:

$$A = Q\mathcal{D}P, \quad \text{где} \quad QQ^\top = E_m, \ PP^\top = E_n.$$

$$\mathcal{D} = \begin{pmatrix} \mathcal{D}_r & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{где} \quad \mathcal{D}_r = \begin{pmatrix} \alpha_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \alpha_r \end{pmatrix}, \quad r = \operatorname{rank} A.$$

Тогда

$$\|A\|_S = \|Q\mathcal{D}P\|_S = \|D\|_S = \sup \frac{\|\mathcal{D}X\|_E}{\|X\|_E} = \frac{\sqrt{\sum\limits_{i=1}^{n} |\alpha_i x_i|^2}}{\sqrt{\sum\limits_{i=1}^{n} |x_i|^2}}.$$

Заменим все α_i на α_1.

Получаем:

$$\|A\|_S \leq \alpha_1 \frac{\sqrt{\sum\limits_{i=1}^{n} |x_i|^2}}{\sqrt{\sum\limits_{i=1}^{n} |x_i|^2}} = \alpha_1,$$

значит, $\|A\|_S \leq \alpha_1$.

Если мы рассмотрим вектор $X_0 = (1, 0, \dots, 0)$, то

$$\frac{\|\mathcal{D}X_0\|_E}{\|X_0\|_E} = \alpha_1,$$

отсюда следует, что $\|A\|_S = \alpha_1$.

Таким образом, мы доказали следующую теорему.

Теорема 11. *Спектральная норма матрицы равна ее наибольшему сингулярному числу.*

Следствие. *Если* $A \in \mathbb{R}_{n,n}$, $\det A \neq 0$, *то* $\|A^{-1}\| = \frac{1}{\alpha_n}$, *где* α_n *— наименьшее сингулярное число матрицы* A *(для невырожденной матрицы оно не равно нулю).*

Рассмотрим матричную норму, которая индуцируется октаэдрической векторной нормой.

Теорема 12. *Для любой матрицы $A \in P_{m,n}$ значение матричной нормы, индуцируемой октаэдрической векторной нормой, вычисляется по формуле*

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Доказательство. По определению значение $\|A\|_1$, индуцированное октаэдрической векторной нормой, вычисляется следующим образом:

$$\|A\|_1 = \sup_{\|X\|_0 = 1} \|AX\|_0.$$

Пусть $X \in P_n$, $\|X_0\| = 1$, то есть $\sum\limits_{j=1}^{n} |x_j| = 1$.

Тогда

$$\|AX\|_0 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le$$

$$\le \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} |x_j| \left(\sum_{i=1}^{m} |a_{ij}| \right) \le$$

$$\le \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| \left(\sum_{j=1}^{n} |x_j| \right) = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Отсюда следует, что

$$\|A\|_1 \le \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Пусть

$$\sum_{i=1}^{m} |a_{ik}| = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Рассмотрим

$$X_0 = (x_1, \ldots, x_n)^{\top}, \quad \text{где} \quad x_j = \begin{cases} 1, & j = k, \\ 0 & j \ne k. \end{cases}$$

Тогда

$$\|A\|_1 \ge \|AX_0\|_0 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right| = \sum_{i=1}^{m} |a_{ik}| = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

$\square$

Аналогично доказывается следующая теорема.

Теорема 13. *Для любой матрицы $A \in P_{m,n}$ значение матричной нормы, индуцированной кубической векторной нормой, вычисляется по формуле*

$$\|A\|_2 = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|.$$

Отождествим матрицу $A \in P_{m,n}$ со строкой (столбцом) длины $m \cdot n$ и запишем формулы для вычисления векторных норм матрицы A:

$$\|A\|_E = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2};$$

$$\|A\|_0 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|;$$

$$\|A\|_k = \max_{\substack{1\le i\le m\\ 1\le j\le n}} |a_{ij}|.$$

Заметим, что норма $\|A\|_k$ не является кольцевой, поэтому вместо нее используется обобщенная кубическая норма:

$$\|A\|_{k'} = \sqrt{mn} \max_{\substack{1\le i\le m\\ 1\le j\le n}} |a_{ij}|.$$

Пример. Запишем значения наиболее употребительных норм для матрицы

$$A = \begin{pmatrix} 1 & -2 \\ 0 & 3 \\ 1 & 5 \end{pmatrix}:$$

$$\|A\|_E = \sqrt{1+1+4+9+25} = \sqrt{40};$$

$$\|A\|_0 = 12;$$

$$\|A\|_{k'} = 5\sqrt{6};$$

$$\|A\|_1 = 10;$$

$$\|A\|_2 = 6.$$

Для вычисления спектральной нормы найдем α_1 — наибольшее сингулярное число матрицы A. Ищем собственные значения матрицы $A^\top A$.

$$A^\top A = \begin{pmatrix} 2 & -7 \\ -7 & 38 \end{pmatrix}.$$

$$\det(A^\top A - \lambda E_2) = \lambda^2 - 40\lambda + 27 = 0.$$

$$\lambda_{1,2} = 20 \pm \sqrt{373}.$$

Отсюда следует, что

$$\|A\|_S = \sqrt{20 + \sqrt{373}}.$$

6. ЛОКАЛИЗАЦИЯ СОБСТВЕННЫХ ЗНАЧЕНИЙ

6.1. Оценка модулей собственных значений

Определение. Пусть $\lambda_i,\ i = \overline{1,n}$ — собственные значения матрицы $A \in \mathbb{C}_{n,n}$. Число

$$\lambda_A = \max_{1\le i\le n} |\lambda_i|$$

называется *спектральным радиусом* матрицы A.

Теорема 1. *Если матричная норма кольцевая, то $\lambda_A \le \|A\|$ для любой матрицы $A \in \mathbb{C}_{n,n}$.*

Доказательство. Пусть $\lambda_A = |\lambda_1|,\ AX = \lambda_1 X,\ X \in \mathbb{C}_{n,1}$.

Рассмотрим матрицу $B \in \mathbb{C}_{n,n}$, у которой первый столбец равен X, а остальные — нулевые. Тогда $AB = \lambda_1 B$. Отсюда следует, что

$$|\lambda_1|\|B\| = \|\lambda_1 B\| = \|AB\| \le \|A\|\|B\|.$$

Значит, $\lambda_A \le \|A\|$, так как $\|B\| > 0$. □

Применяя теорему к различным матричным нормам, получаем

Следствие. *Пусть $\lambda_1, \dots, \lambda_n$ — собственные значения матрицы $A \in \mathbb{C}_{n,n}$, тогда для любого $k = \overline{1,n}$:*

1) $|\lambda_k| \le \sqrt{\sum\limits_{i,j=1}^{n} |a_{ij}|^2}$;

2) $|\lambda_k| \le \sum\limits_{i,j=1}^{n} |a_{ij}|$;

3) $|\lambda_k| \le n \max\limits_{1\le i,j\le n} |a_{ij}|$;

4) $|\lambda_k| \le \max\limits_{1\le j\le n} \sum\limits_{i=1}^{n} |a_{ij}|$;

5) $|\lambda_k| \le \max\limits_{1\le i\le n} \sum\limits_{j=1}^{n} |a_{ij}|$;

6) $|\lambda_k| \le \alpha_1$,

где α_1 — наибольшее сингулярное число матрицы A.

Теорема 2. *Для любой матрицы $A \in \mathbb{C}_{n,n}$ $\alpha_n \le |\lambda_k| \le \alpha_1$, где $\lambda_k,\ k = \overline{1,n}$— собственные значения матрицы A, α_1 и α_n — соответственно наибольшее и наименьшее сингулярные числа матрицы A.*

Доказательство. Пусть

$$AX = \lambda_k X,\ X \in \mathbb{C}_{n,1},\ |X| = 1.$$

Тогда

$$|AX|^2 = |\lambda_k X|^2 = |\lambda_k|^2 |X|^2 = |\lambda_k|^2.$$

Получаем:

$$\begin{aligned} |\lambda_k|^2 = |AX|^2 = (AX)^* AX = \\ = X^*(A^*A)X = (A^*AX, X) = \rho(X), \end{aligned}$$

где $\rho(X)$ — отношение Релея матрицы A.

Согласно теореме 22, глава IV, $\alpha_n^2 \le \rho(x) \le \alpha_1^2$, так как собственные значения матрицы A^*A равны квадратам сингулярных чисел матрицы A. Значит, $\alpha_n^2 \le |\lambda_k|^2 \le \alpha_1^2,\ \forall k = \overline{1,n}$. Откуда следует, что $\alpha_n \le |\lambda_k| \le \alpha_1,\ \forall k = \overline{1,n}$.

□

6.2. Оценка действительных и мнимых частей собственных значений

Пусть матрицы $A \in C_{n,n}$, $\lambda_1, \ldots, \lambda_n$ — собственные значения матрицы A.

Рассмотрим матрицы $S = \frac{1}{2}(A + A^*)$ и $T = \frac{1}{2i}(A - A^*)$, тогда $A = S + iT$. Заметим, что $S^* = S,\ T^* = T$.

Теорема 3. *Пусть матрица S имеет собственные значения $\alpha_1, \ldots, \alpha_n$, а матрица T — собственные значения $\beta_1, \ldots, \beta_n$, упорядоченные таким образом, что*

$$\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_n, \quad \beta_1 \ge \beta_2 \ge \cdots \ge \beta_n.$$

Тогда

$$\alpha_n \le \operatorname{Re}\lambda_i \le \alpha_1, \qquad \beta_n \le \operatorname{Im}\lambda_i \le \beta_1.$$

Доказательство. Пусть

$$AX = \lambda_k X,\ X \in \mathbb{C}_{n,1},\ |X| = 1.$$

Тогда

$$X^* A^* = \overline{\lambda_k} X^*.$$

Возьмем отношение Релея для матрицы S:

$$\rho(x) = (SX, X) = X^* S X = X^* \frac{1}{2}(A + A^*)X = \frac{1}{2}(X^* AX + X^* A^* X) =$$
$$= \frac{1}{2}(\lambda_k + \overline{\lambda_k})X^* X = \frac{1}{2}(\lambda_k + \overline{\lambda_k}) = \frac{1}{2} \cdot 2 \operatorname{Re} \lambda_k = \operatorname{Re} \lambda_k.$$

Отсюда следует: $\alpha_k \le \operatorname{Re} \lambda_k \le \alpha_1,\ \forall k = \overline{1, n}$.

Аналогично доказывается, что $\beta_k \le \operatorname{Im} \lambda_k \le \beta_1,\ \forall k = \overline{1, n}$. □

Пусть матрица $A \in \mathbb{C}_{n,n}$. Рассмотрим унитарную матрицу $U \in \mathbb{C}_{n,n}$, такую, что матрица $B = U^{-1}AU = U^* AU$ — верхнетреугольная. Так как унитарная норма матрицы не меняется при умножении на унитарную матрицу, то

$$\|A\|_u^2 = \|B\|_u^2 = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i \le j} |b_{ij}|^2 \ge \sum_{i=1}^{n} |\lambda_i|^2,$$

где $\lambda_i = b_{ii},\ i = \overline{1, n}$ — собственные значения матрицы A.

Мы показали, что

$$\sum_{i=1}^{n} |\lambda_i|^2 \le \|A\|_u^2.$$

Равенство возможно тогда и только тогда, когда матрица B — диагональная, следовательно, матрица A — нормальная.

Пусть $B = U^* AU$, тогда $B^* = U^* A^* U$ и для матрицы S из теоремы 3

$$U^* SU = U^* \frac{1}{2}(A + A^*)U = \frac{1}{2}U^* AU + \frac{1}{2}U^* A^* U = \frac{1}{2}(B + B^*).$$

Значит, диагональные элементы матрицы $\frac{1}{2}(B + B^*)$ равны $\operatorname{Re} \lambda_i,\ i = \overline{1, n}$. Отсюда следует:

$$\|S\|_u^2 = \|U^* SU\|_u^2 = \|\frac{1}{2}(B + B^*)\|_u^2 =$$
$$= \sum_{i=1}^{n} (\operatorname{Re} \lambda_i)^2 + \sum_{i,j=1,(i \ne j)}^{n} \left| \frac{1}{2}(b_{ij} + \overline{b}_{ji}) \right|^2 \ge \sum_{i=1}^{n} (\operatorname{Re} \lambda_i)^2.$$

Мы показали, что
$$\sum_{i=1}^{n}(\operatorname{Re}\lambda_i)^2 \le \|S\|_u^2.$$
Аналогично доказывается для матрицы T из теоремы 3:
$$\sum_{i=1}^{n}|\operatorname{Im}\lambda_i|^2 \le \|T\|_u^2$$

Тогда
$$|\operatorname{Re}\lambda_i| \le \sqrt{\sum_{i=1}^{n}(\operatorname{Re}\lambda_i)^2} \le \|S\|_v \le n \max_{i\le j, j\le n}|s_{ij}| = n \max_{i\le j, j\le n}\frac{|a_{ij}+\overline{a_{ij}}|}{2}.$$
В результате получаем:
$$|\operatorname{Re}\lambda_i| \le n \max_{1\le i,j\le n}\left|\frac{a_{ij}+\overline{a}_{ji}}{2}\right|.$$
Аналогично:
$$|\operatorname{Im}\lambda_i| \le n \max_{1\le i,j\le n}\left|\frac{a_{ij}-\overline{a}_{ji}}{2}\right|.$$

6.3. Локализационные круги

Пусть матрица $A \in C_{n,n}$. Будем говорить, что матрица A имеет доминирующую главную диагональ, если
$$|a_{ii}| > \sum_{\substack{k=1\\(k\ne i)}}^{n}|a_{ik}|, \text{ для всех } i=\overline{1,n},$$
то есть каждый диагональный элемент матрицы A по модулю больше суммы модулей остальных элементов этой строки.

Теорема 4. *Если матрица $A \in \mathbb{C}_{n,n}$ имеет доминирующую главную диагональ, то она невырождена.*

Доказательство. Пусть $\det A = 0$. Тогда существует ненулевое решение $X \in \mathbb{C}_{n,1}$ системы линейных уравнений $AX = 0$.

Пусть
$$X = (x_1, \dots, x_n)^\top,$$
$$|x_i| = \max_{1\le k\le n}|x_k|.$$

Рассматривая i-е уравнение системы $AX = 0$, получаем:

$$a_{ii}x_i = -\sum_{\substack{k=1\\(k\neq i)}}^{n} a_{ik}x_k.$$

Отсюда следует:

$$|a_{ii}||x_i| = |\sum_{\substack{k=1\\(k\neq i)}}^{n} a_{ik}x_k| \leq |x_i| \sum_{\substack{k=1\\k\neq i)}}^{n} |a_{ik}|.$$

Получаем, что

$$|a_{ij}||x_i| \leq |x_i| \sum_{\substack{k=1\\(k\neq i)}}^{n} |a_{ik}|.$$

Так как $x_i \neq 0$,

$$|a_{ii}| \leq \sum_{\substack{k=1\\(k\neq i)}}^{n} |a_{ik}|.$$

Значит, матрица A не имеет доминирующую главную диагональ. Получили противоречие. Система $AX = 0$ имеет только нулевое решение, следовательно, $\det A \neq 0$. □

Применяя утверждение теоремы 4 к матрице $A^\top$, найдем еще одно условие невырожденности матрицы A.

Если

$$|a_{ij}| > \sum_{\substack{k=1\\(k\neq i)}}^{n} |a_{ki}|, \text{ для всех } i = \overline{1,n},$$

то матрица A невырождена.

Теорема 5. *Все собственные значения матрицы $A \in \mathbb{C}_{n,n}$ находятся в объединении кругов*

$$|\lambda - a_{ii}| \leq \sum_{\substack{k=1\\(k\neq i)}}^{n} |a_{ik}|, \textit{ для всех } \quad i = \overline{1,n}.$$

Доказательство. Предположим, что λ — собственное значение матрицы A, в эти круги не входит.

Отсюда следует, что

$$|\lambda - a_{ii}| > \sum_{\substack{k=1 \\ (k \neq i)}}^{n} |a_{ik}|, \text{ для всех } \quad i = \overline{1, n}.$$

Это означает, что характеристическая матрица $A - \lambda E_n$ имеет доминирующую главную диагональ.

Тогда $\det(A - \lambda E_n) \neq 0$, значит, λ не может быть собственным значением, следовательно, наше предположение было неверно. $\square$

Следующую теорему примем без доказательства.

Теорема 6. *Множество из r локализационных кругов, не пересекающихся с остальными $n - r$ кругами, содержит ровно r собственных значений.*

Следствие. *Некратный локализационный круг действительной матрицы, не пересекающийся с остальными кругами, содержит действительное собственное значение.*

Доказательство. Так как некратный локализационный круг не пересекается с остальными, согласно теореме 6 в нем лежит ровно одно собственное значение.

Поскольку матрица действительная, центр круга находится на действительной оси, поэтому если бы собственное значение λ матрицы A, лежащее в этом круге, было комплексным, то кругу принадлежало бы и собственно значение $\overline{\lambda}$. Значит, $\lambda = \overline{\lambda}$, следовательно, $\lambda \in \mathbb{R}$. $\square$

Замечание. Так как собственные значения матриц A и $A^{\top}$ совпадают, мы можем записать второе множество локализационных кругов:

$$|\lambda - a_{ii}| \leq \sum_{\substack{k=1 \\ (k \neq i)}}^{n} |a_{ki}|, \text{ для всех } i = \overline{1, n}.$$

Содержание

Литература

1. *Д. В. Беклемишев.* Дополнительные главы линейной алгебры М. : Наука, 1983.
2. *Ф. Р. Гантмахер.* Теория матриц М. : Мир, 1967.
3. *Р. Хорн, Ч. Джонсон.* Матричный анализ М. : Мир, 1989.
4. *П. Ланкастер.* Теория матриц М. : Наука, 1978.
5. *Р. Беллман.* Введение в теорию матриц М. : Наука, 1969.
6. *М. Маркус, Х. Минк.* Обзор по теории матриц и матричных неравенств М. : Наука, 1972.
7. *В. В. Воеводин, Ю. А. Кузнецов.* Матрицы и вычисления М. : Наука, 1984.
8. *И. М. Гельфанд.* Лекции по линейной алгебре М. : Наука, 1971.
9. *С. А. Мазаник, Г. П. Размыслович, В. М. Ширяев.* Функции от матриц Мн. : БГУ, 2002.
10. *Г. П. Размыслович, М. М.Феденя, В. М. Ширяев.*Геометрия и алгебра Мн.: Университетское, 1987.
11. *Г. П. Размыслович, М. М. Феденя, В. М. Ширяев.* Сборник задач по геометрии и алгебре Мн.: Университетское, 1999.
12. *М. В. Милованов, Р. И. Тышкевич, А. С. Феденко.* Линейная алгебра и аналитическая геометрия Ч. 2. Мн.: Выш. школа, 1987.
13. *А. К. Деменчук, Б. Б. Комраков, Г. П. Размыслович, В. М. Ширяев.* Задачи по матричному анализу Мн. : БГУ, 2004.
14. *Х. Д. Икрамов.* Задачи по линейной алгебре М. : Наука, 1975.
15. *И. В. Проскуряков.* Сборник задач по линейной алгебре М. : Наука, 1984.