Глава 12

Линейные отображения и преобразования

12.1. Определения и примеры

Рассмотрим два линейных пространства V, W, заданных над одним и тем же полем F. Отображение $\varphi : V \to W$ называется *линейным отображением*, или *линейным оператором*, если выполнены следующие свойства (свойства линейности):

1) $\varphi(x + y) = \varphi x + \varphi y$ для произвольных векторов $x, y \in V$.

2) $\varphi(\alpha x) = \alpha \cdot \varphi x$ для произвольных вектора $x \in V$ и числа $\alpha \in F$.

Из определения следует, что изоморфизм — это линейное отображение, являющееся биекцией. Отображение $\varphi : V \to V$ называется *преобразованием* пространства V.

Из свойств линейности отображения по индукции легко вывести следующее обобщение:

$$\varphi \left(\sum_{i=1}^{s} \alpha_i a_i \right) = \sum_{i=1}^{s} \alpha_i \varphi a_i, \tag{12.1}$$

где $a_i \in V$, $\alpha_i \in F$ $(i = 1, 2, \ldots, s)$. В свойстве 2) полагая $\alpha = 0$, получаем

$$\varphi o = o.$$

Заметим, что в левой и правой частях этого равенства символом o обозначены нулевые векторы разных (если $V \neq W$) пространств.

Обозначим $\Phi(V, W)$ множество всех линейных отображений, действующих из V в W.

Рассмотрим некоторые примеры линейных отображений, а также некоторые примеры отображений, которые не являются линейными.

Нулевое отображение. Определим отображение $\theta : V \to W$ равенством $\theta x = o$ для произвольного $x \in V$. Отображение θ, очевидно, является линейным. Оно называется *нулевым* отображением. Заметим, что отображение ψ, определяемое по формуле $\psi x = a$, где a — фиксированный ненулевой вектор пространства W, линейным не является.

Тождественное преобразование. Определим преобразование $\varepsilon : V \to V$ формулой $\varepsilon x = x$ для произвольного $x \in V$. Отображение ε, очевидно, является линейным. Оно называется *тождественным* преобразованием.

Преобразование проектирования. Пусть пространство V есть прямая сумма подпространств V_1 и V_2. Преобразование $\varphi : V \to V$, ставящее в соответствие вектору x его проекцию y на V_1 параллельно V_2, называется *преобразованием проектирования*. Докажем его линейность. Пусть $x = x_1 + x_2$, $y = y_1 + y_2$; $x_1, y_1 \in V_1$, $x_2, y_2 \in V_2$. Тогда $x+y = (x_1+y_1)+(x_2+y_2)$, $x_1 + y_1 \in V_1$, $x_2 + y_2 \in V_2$ и поэтому $\varphi(x + y) = x_1 + y_1 = \varphi x + \varphi y$. Свойство 1) выполнено. Далее, $\alpha x = \alpha x_1 + \alpha x_2$, поэтому $\varphi(\alpha x) = \alpha\varphi x$. Свойство 2) выполнено.

Преобразование вращения. Пусть преобразование $\varphi : \mathbf{V}_2 \to \mathbf{V}_2$ каждому вектору x ставит в соответствие вектор φx, получающийся из x поворотом вокруг полюса на заданный угол α. Это преобразование называется *преобразованием поворота* (*на угол* α). Нетрудно видеть, что оно является линейным. Также линейным является преобразование поворота пространства $\mathbf{V}_3$ вокруг фиксированной прямой, проходящей через начало координат.

(Нелинейное) преобразование сдвига. Отображение $\varphi : V \to V$, ставящее в соответствие вектору x вектор $\varphi x = a+x$, называется *преобразованием сдвига*, при этом фиксированный вектор a называется вектором *сдвига*. Если $a \neq o$, то преобразование сдвига линейным не является.

Пример преобразования арифметического пространства. Пусть $A = (a_{ij}) \in F^{m\times n}$. Рассмотрим отображение $\varphi :\ F^n \to F^m$, ставящее в соответствие вектору $x = (x_1, x_2, \ldots, x_n)^\top \in F^n$ вектор $\varphi x = (y_1, y_2, \ldots, y_m)^\top \in F^m$, такой, что

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n. \end{aligned}$$

Используя матричные операции, данное отображение можно записать короче:

$$\varphi x = Ax. \tag{12.2}$$

Легко проверить, что отображение φ — линейное. Далее мы увидим, что формула (12.2) задает общий вид линейного отображения F^n в F^m.

Преобразования дифференцирования и интегрирования. *Преобразование дифференцирования* D пространства многочленов $F[x]$ каждому многочлену ставит в соответствие его производную:

$$Df(x) = f'(x).$$

Преобразование дифференцирования часто обозначается $\dfrac{d}{dx}$. *Преобразование интегрирования* I пространства многочленов $F[x]$ определим как отображение, ставящее в соответствие всякому многочлену его первообразную, в точке 0 принимающую значение 0:

$$If(x) = \int_0^x f(t)dt.$$

Очевидно, каждое из этих преобразований — линейное.

12.2. Матрица линейного отображения

Пусть $\mathbf{e} = \{e_1, e_2, \dots, e_n\}$ — некоторый базис пространства V. Пусть

$$x = \sum_{j=1}^{n} \alpha_j e_j \in V$$

т. е. $[x]_{\mathbf{e}} = (\alpha_1, \alpha_2, \dots, \alpha_n)^\top$. Из (12.1) получаем:

$$\varphi x = \sum_{j=1}^{n} \alpha_j \varphi e_j. \tag{12.3}$$

Таким образом, доказано следующее

Утверждение 12.1. *Линейное отображение восстанавливается однозначно по образам базисных векторов.*

Пусть $\mathbf{f} = \{f_1, f_2, \dots, f_m\}$ — базис пространства W. Так как $\varphi e_j \in W$ $(j = 1, 2, \dots, n)$, то φe_j раскладывается по базису $\mathbf{f}$. Пусть

$$\varphi e_j = \sum_{i=1}^{m} \alpha_{ij} f_i,$$

т. е. $[\varphi e_j]_{\mathbf{f}} = (\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj})^\top$. Матрица, j-й столбец которой есть столбец координат $[\varphi e_j]_{\mathbf{f}}$ $(j = 1, 2, \dots, n)$, называется *матрицей отображения* φ, построенной в базисах $\mathbf{e}$ и $\mathbf{f}$. Обозначение: $[\varphi]_{\mathbf{f},\mathbf{e}}$ или просто $[\varphi]$, если понятно из контекста, в каких базисах построена матрица. Из определения получаем

$$[\varphi]_{\mathbf{f},\mathbf{e}} = (\alpha_{ij}) \in F^{m \times n}.$$

Пример 12.2. Найдем матрицу оператора дифференцирования D пространства многочленов степени не выше n в пространство многочленов степени не выше $n-1$, если в качестве базисов выбраны

$$\mathbf{e} = \langle 1, t, t^2, \dots, t^n \rangle \qquad \text{и} \qquad \mathbf{f} = \langle 1, t, t^2, \dots, t^{n-1} \rangle$$

соответственно. Находим

$$\begin{array}{rcl} D1 &=& 0 \quad = 0 \cdot 1 + 0 \cdot t + \ldots + 0 \cdot t^{n-1}, \\ Dt &=& 1 \quad = 1 \cdot 1 + 0 \cdot t + \ldots + 0 \cdot t^{n-1}, \\ Dt^2 &=& 2t \quad = 0 \cdot 1 + 2 \cdot t + \ldots + 0 \cdot t^{n-1}, \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ Dt^n &=& nt^{n-1} = 0 \cdot 1 + 0 \cdot t + \ldots + n \cdot t^{n-1}, \end{array}$$

откуда

$$[D]_{\mathbf{f},\mathbf{e}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \multicolumn{5}{c}{\dots\dots\dots\dots\dots} \\ 0 & 0 & 0 & \dots & n \end{pmatrix} \in F^{n \times (n+1)}.$$

Утверждение 12.3. *Для произвольного вектора x из пространства V*

$$[\varphi x]_{\mathbf{f}} = [\varphi]_{\mathbf{f},\mathbf{e}} [x]_{\mathbf{e}}. \tag{12.4}$$

Доказательство. Переходя к координатам в выражении (12.3), получаем

$$[\varphi x]_{\mathbf{f}} = \sum_{j=1}^{n} \alpha_j [\varphi e_j]_{\mathbf{f}}.$$

В правой части имеем линейную комбинацию столбцов матрицы $[\varphi]_{\mathbf{f}}$, откуда следует (12.4). ∎

По аналогии с (12.4) рассмотрим равенство

$$[\varphi x]_{\mathbf{f}} = A[x]_{\mathbf{e}}, \tag{12.5}$$

где A — произвольная матрица из $F^{m\times n}$. Приведенная формула каждому вектору $x \in V$ ставит в соответствие вектор $\varphi x \in W$, таким образом, *определяет* отображение $\varphi : V \to W$. Исследуем его.

Утверждение 12.4. *Отображение $\varphi : V \to W$, заданное формулой (12.5), является линейным, причем A — его матрица: $[\varphi]_{\mathbf{f},\mathbf{e}} = A$.*

Доказательство. Для произвольных $x, y \in V$ имеем

$$[\varphi(x+y)]_{\mathbf{f}} = A[x+y]_{\mathbf{e}} = A([x]_{\mathbf{e}} + [y]_{\mathbf{e}}) = A[x]_{\mathbf{e}} + A[y]_{\mathbf{e}} = [\varphi x]_{\mathbf{f}} + [\varphi y]_{\mathbf{f}} = [\varphi x + \varphi y]_{\mathbf{f}},$$

следовательно, $\varphi(x+y) = \varphi x + \varphi y$. Для произвольного $x \in V$ и произвольного $\alpha \in F$ имеем

$$[\varphi(\alpha x)]_{\mathbf{f}} = A[\alpha x]_{\mathbf{e}} = A(\alpha[x]_{\mathbf{e}}) = \alpha A[x]_{\mathbf{e}} = \alpha[\varphi x]_{\mathbf{f}} = [\alpha\varphi x]_{\mathbf{f}},$$

следовательно, $\varphi(\alpha x) = \alpha\varphi x$. Итак, отображение φ — линейное.

Проверим теперь, что $[\varphi]_{\mathbf{f},\mathbf{e}} = A$. Подставим e_j $(j = 1, 2, \ldots, n)$ в (12.5): $[\varphi e_j]_{\mathbf{f}} = A[e_j]_{\mathbf{e}}$. Так как j-я компонента столбца $[e_j]_{\mathbf{e}}$ равна 1, а остальные компоненты равны 0, то $A[e_j]_{\mathbf{e}}$ есть j-й столбец матрицы A, значит A, по определению, есть матрица отображения φ. ∎

Итак, формула (12.5) задает общий вид линейного отображения.

Замечание 12.5. Матрица линейного отображения пространства F^n в F^m, определяемого формулой (12.2), записанная в стандартных базисах этих пространств, есть A. Следовательно, формула (12.2) задает общий вид линейного отображения из F^n в F^m. Это позволяет отождествить линейные отображения с матрицами и перенести терминологию с отображений на матрицы.

Переформулировка результатов этого раздела приводит нас к следующему.

Следствие 12.6. *Отображение, ставящее в соответствие всякому линейному отображению его матрицу в паре фиксированных базисов, является биекцией из $\Phi(V, W)$ в $F^{m\times n}$.*

12.3. Операции с линейными отображениями

Определим некоторые операции над линейными отображениями.

Суммой двух отображений $\varphi, \psi \in \Phi(V, W)$ называется отображение χ, такое, что $\chi x = \varphi x + \psi x$ для произвольного $x \in V$. Обозначение для суммы: $\chi = \varphi + \psi$.

Произведением отображения $\varphi \in \Phi(V, W)$ на число $\alpha \in F$ называется отображение χ, такое, что $\chi x = \alpha(\varphi x)$ для произвольного $x \in V$. Обозначение для произведения отображения на число: $\chi = \alpha\varphi$.

Рассмотрим три пространства U, V, W, заданных над одним и тем же полем F. Пусть ψ, φ — линейные отображения, действующие из U в V и из V в W соответственно: $U \xrightarrow{\psi} V \xrightarrow{\varphi} W$. *Произведением*, или *композицией*, *отображений* φ и ψ называется отображение θ, такое, что $\chi x = \varphi(\psi x)$. Обозначение для произведения отображений: $\chi = \varphi\psi$.

Итак, для произвольного $x \in V$

$$(\varphi + \psi)x = \varphi x + \psi x, \qquad (\alpha\varphi)x = \alpha(\varphi x), \qquad (\varphi\psi)x = \varphi(\psi x).$$

Линейные операции над отображениями — это сложение отображений и умножение их на числа.

Утверждение 12.7 (Линейность результата операций)**.**

1) Пусть $\varphi, \psi \in \Phi(V, W)$, $\alpha \in F$, *тогда отображения* $\varphi + \psi$, $\alpha\varphi$ *— линейные, причем* $[\varphi + \psi] = [\varphi] + [\psi]$, $[\alpha\varphi] = \alpha[\varphi]$.

2) Пусть $\varphi \in \Phi(V, W)$, $\psi \in \Phi(U, V)$, *тогда отображение* $\varphi\psi$ *— линейный, причем* $[\varphi\psi] = [\varphi][\psi]$.

Доказательство. Все свойства доказываются аналогично. Приведем для примера два способа доказательства п. 2.

1 способ. Пусть

$$\begin{aligned} \mathbf{g} &= \{g_1, g_2, \ldots, g_l\} \text{ — базис пространства } U, \\ \mathbf{e} &= \{e_1, e_2, \ldots, e_n\} \text{ — базис пространства } V, \\ \mathbf{f} &= \{f_1, f_2, \ldots, f_m\} \text{ — базис пространства } W; \end{aligned} \tag{12.6}$$

$$[\varphi]_{\mathbf{f},\mathbf{e}} = (\alpha_{ij}) \in F^{m\times n}, \qquad [\psi]_{\mathbf{e},\mathbf{g}} = (\beta_{jk}) \in F^{n\times l}. \tag{12.7}$$

Теперь получаем

$$[(\varphi\psi)x]_{\mathbf{f}} = [\varphi(\psi x)]_{\mathbf{f}} = [\varphi]_{\mathbf{f},\mathbf{e}}[\psi x]_{\mathbf{e}} = [\varphi]_{\mathbf{f},\mathbf{e}}[\psi]_{\mathbf{e},\mathbf{g}}[x]_{\mathbf{g}}.$$

Обозначим $A = [\varphi]_{\mathbf{f},\mathbf{e}}[\psi]_{\mathbf{e},\mathbf{g}}$, тогда

$$[(\varphi\psi)x]_{\mathbf{f}} = A[x]_{\mathbf{g}}.$$

По лемме 12.4 отображение $\varphi\psi$ — линейное и $[\varphi\psi]_{\mathbf{f},\mathbf{g}} = [\varphi]_{\mathbf{f},\mathbf{e}}[\psi]_{\mathbf{e},\mathbf{g}}$.

2 способ. Докажем вначале линейность. Для произвольных $x, y \in U$ имеем

$$(\varphi\psi)(x + y) = \varphi(\psi(x + y)) = \varphi(\psi x + \psi y) = \varphi(\psi x) + \varphi(\psi y) = (\varphi\psi)x + (\varphi\psi)y.$$

Для произвольных $x \in U$, $\alpha \in F$ имеем

$$(\varphi\psi)(\alpha x) = \varphi(\psi(\alpha x)) = \varphi(\alpha\psi x) = \alpha\varphi(\psi x) = \alpha(\varphi\psi)x.$$

Итак, отображение $\varphi\psi$ — линейное.

Теперь вычислим $[\varphi\psi]_{\mathbf{f},\mathbf{g}} = (\gamma_{ik})$. В системе обозначений (12.6), (12.7) для $k = 1, 2, \ldots, l$ получаем

$$(\varphi\psi)g_k = \varphi(\psi g_k) = \varphi\sum_{j=1}^{n}\beta_{jk}e_j = \sum_{j=1}^{n}\beta_{jk}\varphi e_j = \sum_{j=1}^{n}\beta_{jk}\sum_{i=1}^{m}\alpha_{ij}f_i = \sum_{i=1}^{m}f_i\sum_{j=1}^{n}\alpha_{ij}\beta_{jk}.$$

По определению матрицы линейного отображения получаем, что $\gamma_{ik} = \sum\limits_{j=1}^{n}\alpha_{ij}\beta_{jk}$, или $[\varphi\psi]_{\mathbf{f},\mathbf{g}} = [\varphi]_{\mathbf{f},\mathbf{e}}[\psi]_{\mathbf{e},\mathbf{g}}$. ∎

Замечание 12.8. Сумма, произведение и произведение на число линейных отображений являются линейными отображениями в том числе и для бесконечномерного пространства. 2-й способ доказательства проходит и для этого случая.

Утверждение 12.9. *Для любых* $\varphi, \psi, \chi \in \Phi(V, W)$, $\alpha, \beta \in F$ *справедливы равенства:*

1. $\varphi + \psi = \psi + \varphi$,

2. $\varphi + (\psi + \chi) = (\varphi + \psi) + \chi$,

3. $\varphi + \theta = \varphi$, *где* θ *— нулевое отображение,*

4. $\varphi + (-1)\varphi = \theta$ *(таким образом,* $(-1)\varphi$ *— отображение* противоположное *отображению* φ*),*

5. $1\varphi = \varphi$,

6. $\alpha(\beta\varphi) = (\alpha\beta)\varphi$,

7. $(\alpha + \beta)\varphi = \alpha\varphi + \beta\varphi$,

8. $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$.

Доказательство. Доказательство приведенных выше свойств не должно вызывать затруднений. Докажем для примера двумя способами последнее свойство.
1 способ (по определению). Для произвольного $x \in V$ имеем

$$(\alpha(\varphi + \psi))x = \alpha((\varphi + \psi)x) = \alpha(\varphi x + \psi x) = \alpha(\varphi x) + \alpha(\psi x) = (\alpha\varphi)x + (\alpha\psi)x,$$

т. е. $\alpha(\varphi + \psi) = \alpha\varphi + \alpha\psi$.
2 способ (через свойства матриц). Для произвольного $x \in V$ имеем

$$[\alpha(\varphi + \psi)x]_{\mathbf{f}} = [\alpha(\varphi + \psi)]_{f,e}[x]_{\mathbf{e}} = [\alpha\varphi]_{\mathbf{f},\mathbf{e}}[x]_{\mathbf{e}} + [\alpha\psi]_{\mathbf{f},\mathbf{e}}[x]_{\mathbf{e}} = [\alpha\varphi x]_{\mathbf{f}} + [\alpha\psi x]_{\mathbf{f}} = [\alpha\varphi x + \alpha\psi x]_{\mathbf{f}},$$

откуда вытекает доказываемое. ∎

Замечание 12.10. 1-й способ доказательства проходит также и для случая бесконечномерного линейного пространства.

Следствие 12.11. *Множество* $\Phi(V, W)$ *относительно ранее введенных операций сложения отображений и умножения их на числа является линейным пространством. Отображение, ставящее в соответствие всякому отображению его матрицу в паре фиксированных базисов, является изоморфизмом* $\Phi(V, W)$ *в* $F^{m\times n}$.

Доказательство. Замкнутость операций сложения отображений и умножения их на числа вытекает из утверждения о линейности результата операций. Аксиомы линейного пространства доказаны в предыдущем утверждении. Биективность указанного в формулировке следствия вытекает из следствия 12.6. Сохранение операций следует из утверждения о линейности результата операций. ∎

Утверждение 12.12. *Пусть* U, V, W, Q *— линейные пространства, заданные над одним полем* F. *Тогда для любых линейных отображений* $\varphi, \varphi_1 \in \Phi(V, W)$, $\psi, \psi_1 \in \Phi(U, V)$, $\chi \in \Phi(Q, U)$, *и произвольных* $\alpha, \beta \in F$ *справедливы равенства:*

1. $\varphi(\psi\chi) = (\varphi\psi)\chi$,

2. $\alpha(\varphi\psi) = (\alpha\varphi)\psi = \varphi(\alpha\psi)$,

3. $(\varphi + \varphi_1)\psi = \varphi\psi + \varphi_1\psi$,

4. $\varphi(\psi + \psi_1) = \varphi\psi + \varphi\psi_1$.

Доказательство. Как и свойства линейных операций сформулированные свойства легко могут быть выведены из определения, либо из соответствующих свойств матриц. ∎

12.4. Изменение матрицы линейного отображения при замене базисов

Пусть $\varphi \in \Phi(V, W)$,

$$\begin{aligned} &\mathbf{e} = \{e_1, e_2, \ldots, e_n\}\,, \quad \mathbf{e}' = \{e'_1, e'_2, \ldots, e'_n\} \text{ — два базиса пространства } V, \\ &\mathbf{f} = \{f_1, f_2, \ldots, f_n\}\,, \quad \mathbf{f}' = \{f'_1, f'_2, \ldots, f'_n\} \text{ — два базиса пространства } W. \end{aligned}$$

Рассмотрим, как меняется матрица отображения φ при переходе от базисов $\mathbf{e}$ и $\mathbf{f}$ к базисам $\mathbf{e}'$ и $\mathbf{f}'$ соответственно.

Утверждение 12.13.

$$[\varphi]_{\mathbf{f}',\mathbf{e}'} = [\mathbf{f}']_{\mathbf{f}}^{-1}[\varphi]_{\mathbf{f},\mathbf{e}}[\mathbf{e}']_{\mathbf{e}}. \tag{12.8}$$

Доказательство. Из равенств

$$[x]_{\mathbf{e}} = [\mathbf{e}']_{\mathbf{e}}[x]_{\mathbf{e}'}, \qquad [\varphi x]_{\mathbf{f}} = [\mathbf{f}']_{\mathbf{f}}[\varphi x]_{\mathbf{f}'}, \qquad [\varphi x]_{\mathbf{f}} = [\varphi]_{\mathbf{f},\mathbf{e}}[x]_{\mathbf{e}}$$

легко получается формула

$$[\varphi x]_{\mathbf{f}'} = [\mathbf{f}']_{\mathbf{f}}^{-1}[\varphi]_{\mathbf{f},\mathbf{e}}[\mathbf{e}']_{\mathbf{e}}[x]_{\mathbf{e}'},$$

справедливая для произвольного $x \in V$. Обозначив $A = [\mathbf{f}']_{\mathbf{f}}^{-1}[\varphi]_{\mathbf{f},\mathbf{e}}[\mathbf{e}']_{\mathbf{e}}$, по утверждению 12.4 получаем требуемое. ∎

Пусть A, B — две матрицы из $F^{m\times n}$. Говорят, что матрица B *эквивалентна* матрице A, если для некоторых невырожденных матриц $R \in F^{m\times m}$, $S \in F^{n\times n}$ выполнено

$$B = RAS. \tag{12.9}$$

Легко проверить, что отношение эквивалентности обладает свойствами рефлексивности, симметричности и транзитивности.

Следствие 12.14. *Матрицы одного и того же линейного отображения в разных базисах эквивалентны. Любая матрица, эквивалентная заданной матрице линейного отображения, является матрицей того же отображения, записанной в некоторой паре базисов.*

Доказательство. Если обозначить $A = [\varphi]_{\mathbf{f},\mathbf{e}}$, $B = [\varphi]_{\mathbf{f}',\mathbf{e}'}$, $R = [\mathbf{f}']_{\mathbf{f}}^{-1}$, $S = [\mathbf{e}']_{\mathbf{e}}$, то (12.8) примет вид (12.9). ∎

12.5. Образ и ядро линейного отображения

Образом, или *множеством значений*, отображения $\varphi \in \Phi(V, W)$, называется множество векторов из W, для которых в V существует по крайней мере один прообраз. Обозначение: φV или $\operatorname{Im}\varphi$. Итак, по определению,

$$\varphi V = \{y \in W : \ \exists\, x \in V, \varphi x = y\}.$$

Ядром, или *нуль-пространством*, отображения $\varphi \in \Phi(V, W)$, называется множество векторов из V, обращающихся в $o \in W$. Обозначение: $\operatorname{Ker}\varphi$. Итак, по определению,

$$\operatorname{Ker}\varphi = \{x \in V : \varphi x = o\}.$$

Утверждение 12.15. *Образ и ядро линейного отображения $\varphi \in \Phi(V, W)$ являются подпространствами в W и V соответственно.*

Доказательство. Докажем, что $\operatorname{Im}\varphi$ — подпространство. Очевидно, что $\operatorname{Im}\varphi \neq \emptyset$. Для $y_1, y_2 \in \operatorname{Im}\varphi$ найдутся такие $x_1, x_2 \in V$, что $\varphi x_1 = y_1$, $\varphi x_2 = y_2$. Так как $y_1 + y_2 = \varphi(x_1 + x_2)$ и $x_1 + x_2 \in V$, то $y_1 + y_2 \in \operatorname{Im}\varphi$. Пусть $y = \varphi x \in V$, $\alpha \in F$, тогда $\alpha y = \alpha\varphi x = \varphi(\alpha x) \in \operatorname{Im}\varphi$.

Докажем, что $\operatorname{Ker}\varphi$ — подпространство. Так как $\varphi o = o$, то $o \in \operatorname{Ker}\varphi$, поэтому $\operatorname{Ker}\varphi \neq \emptyset$. Пусть $x_1 \in \operatorname{Ker}\varphi$, $x_2 \in \operatorname{Ker}\varphi$, тогда $\varphi(x_1 + x_2) = \varphi x_1 + \varphi x_2 = o$, следовательно, $x_1 + x_2 \in \operatorname{Ker}\varphi$. Пусть $x \in \operatorname{Ker}\varphi$, $\alpha \in F$, тогда $\varphi(\alpha x) = \alpha\varphi x = o$, следовательно, $\alpha x \in \operatorname{Ker}\varphi$. ∎

Размерность образа линейного отображения называется его *рангом* и обозначается $\operatorname{rank}\varphi$, размерность ядра линейного отображения называется *дефектом* и обозначается $\operatorname{def}\varphi$.

Теорема 12.16 (Ранг и дефект отображения)**.**

1. *Пусть $\mathbf{e} = \{e_1, e_2, \ldots, e_n\}$ и $\mathbf{f} = \{f_1, f_2, \ldots, f_m\}$ — базисы пространств V, W соответственно, тогда* $\operatorname{rank}\varphi = \operatorname{rank}[\varphi]_{\mathbf{f},\mathbf{e}}$.

2. *Имеет место равенство* $\operatorname{def}\varphi + \operatorname{rank}\varphi = \dim V$.

Доказательство. 1) Рассмотрим множество $[\varphi V]$ координатных столбцов всех векторов из $\varphi V = \{y = \varphi x : x \in V\}$. Имеем $[\varphi V] = \{y = [\varphi]_{\mathbf{f},\mathbf{e}} x : x \in F^n\}$, таким образом, $[\varphi V]$ является линейной оболочкой столбцов матрицы $[\varphi]_{\mathbf{f},\mathbf{e}}$ и имеет размерность $\operatorname{rank}[\varphi]_{\mathbf{f},\mathbf{e}}$. Очевидно, $\dim V = \dim[\varphi V]$.

2) Рассмотрим множество $[\operatorname{Ker} V]$ координатных столбцов всех векторов из $\operatorname{Ker}\varphi$. Имеем $[\operatorname{Ker}\varphi] = \{x \in F^n : [\varphi]_{\mathbf{f},\mathbf{e}} x = 0\}$, таким образом, $[\operatorname{Ker}\varphi]$ есть множество решений системы линейных уравнений, поэтому $\dim[\operatorname{Ker}\varphi] = n - \operatorname{rank}[\varphi]_{f,e}$, или $\operatorname{def}\varphi = n - \operatorname{rank}\varphi$. ∎

Пусть $A \in F^{m\times n}$. Множество решений однородной системы $Ax = 0$ называется *ядром*, или *нуль-пространством*, матрицы A.

Отношение эквивалентности разбивает множество матриц заданных размеров на классы эквивалентности. Рассмотим вопрос о «наиболее простом» представителе из каждого такого класса.

Теорема 12.17. *Пусть* $A \in F^{m\times n}$ *и* $\operatorname{rank} A = r$. *Тогда* A *эквивалентна матрице*

$$I_r = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{pmatrix}, \tag{12.10}$$

в которой на диагонали стоят r *единиц, а все непомеченные элементы равны* 0.

Доказательство. Понятно, что с помощью элементарных преобразований строк и столбцов матрицы A ее можно привести к виду (12.10): например, вначале применяем алгоритм Жордана–Гаусса к A, а затем его же к $A^\top$. Элементарные преобразования не меняют ранга матрицы, поэтому количество единиц в (12.10) равно r. Преобразования со строчками равносильны домножению на некоторую невырожденную матрицу R слева, а преобразования со столбцами — домножению на некоторую невырожденную матрицу S справа, т. е. $I_r = RAS$ и, следовательно, $I_r \sim A$. ∎

Следствие 12.18. *Для любого* $\varphi \in \Phi(V, W)$ *найдется такая пара базисов* $e_1, e_2, \ldots, e_n$ *и* $f_1, f_2, \ldots, f_m$ *пространств* V *и* W *соответственно, что*

$$\varphi e_j = f_j \quad (j = 1, 2, \ldots, r), \qquad \varphi e_j = o \quad (j = r+1, r+2, \ldots, n),$$

где $r = \operatorname{rank} \varphi$.

Доказательство. Пусть $e_1, e_2, \ldots, e_n$ и $f_1, f_2, \ldots, f_m$ — базисы, в которых матрица преобразования имеет вид (12.10). Теперь доказываемое следует из определения матрицы линейного отображения. ∎

Теорема 12.19. *Пусть* $A \in F^{m\times n}$, $B \in F^{m\times n}$. *Для того, чтобы матрицы* A *и* B *были эквивалентны, необходимо и достаточно, чтобы* $\operatorname{rank} A = \operatorname{rank} B$.

Доказательство. Необходимость. Так как $A \sim B$, то существуют такие невырожденные R и S, что $B = RAS$, откуда $\operatorname{rank} B = \operatorname{rank} A$.
Достаточность. Пусть $\operatorname{rank} A = \operatorname{rank} B = r$. По теореме 12.17 имеем $I_r \sim A$, $I_r \sim B$, следовательно, $A \sim B$. ∎

Теоремы 12.17, 12.19 отвечают на вопрос, как выглядит наиболее простая матрица из заданного класса эквивалентных матриц. Также мы ответили на вопрос, как выглядят базисы, в которых матрица линейного отображения имеет наиболее простой вид.

12.6. Линейные преобразования

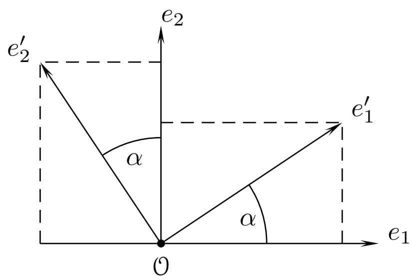
Напомним, что линейным преобразованием пространства V называется линейное отображение $\varphi : V \to V$.

- Преобразование называется *тождественным* (единичным) и обозначается ε, если $\varepsilon x = x$ для любого $x \in V$.
- Преобразование ψ_1 называется *левым обратным* для преобразования φ, если $\psi_1\varphi = \varepsilon$.
- Преобразование ψ_2 называется *правым обратным* для преобразования φ, если $\varphi\psi_2 = \varepsilon$.
- Преобразование называется *обратным* для преобразования φ и обозначается φ^{-1}, если $\varphi\varphi^{-1} = \varphi\varphi^{-1} = \varepsilon$. Преобразование φ называется *невырожденным*, если для него существует обратное преобразование.

Матрицей преобразования φ в базисе $e_1, e_2, \dots, e_n$ называется матрица, составленная из координатных столбцов $[\varphi e_j]_{\mathbf{e}}$ $(j = 1, 2, \dots, n)$. Для матрицы преобразования используется обозначение $[\varphi]_{\mathbf{e}}$. Для произвольных φ, ψ из $\Phi(V, V)$ и произвольного α из F, используя свойства матриц линейных отображений, получаем

$$[\alpha\varphi]_{\mathbf{e}} = \alpha[\varphi]_{\mathbf{e}}, \quad [\varphi + \psi]_{\mathbf{e}} = [\varphi]_{\mathbf{e}} + [\psi]_{\mathbf{e}}, \quad [\varphi\psi]_{\mathbf{e}} = [\varphi]_{\mathbf{e}}[\psi]_{\mathbf{e}}. \tag{12.11}$$

Пример 12.20. Найдем матрицу преобразования поворота плоскости $\mathbf{V}_2$ на угол α в ортонормированном базисе $\mathbf{e} = \langle e_1, e_2 \rangle$ (ср. с примером 9.67).



$$\varphi e_1 = \cos\alpha \cdot e_1 + \sin\alpha \cdot e_2, \qquad \varphi e_2 = -\sin\alpha \cdot e_1 + \cos\alpha \cdot e_2,$$

поэтому

$$[\varphi]_{\mathbf{e}} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Эта матрица называется *матрицей поворота*. При повороте вектор с координатами $(x_1, x_2)^\top$ переходит в вектор с координатами $(x_1^*, x_2^*)^\top$:

$$\begin{cases} x_1^* = x_1\cos\alpha - x_2\sin\alpha, \\ x_2^* = x_1\sin\alpha + x_2\cos\alpha. \end{cases}$$

Заметим, что обратное преобразование — поворот на угол $-\alpha$ — определяется матрицей

$$[\varphi^{-1}]_{\mathbf{e}} = [\varphi]_{\mathbf{e}}^{-1} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Пусть $\mathbf{e} = \{e_1, \ldots, e_n\}$, $\mathbf{e}' = \{e'_1, \ldots, e'_n\}$ — два базиса пространства V. Согласно утверждению 12.13 матрицы преобразования φ в этих базисах связаны формулой

$$[\varphi]_{\mathbf{e}'} = [\mathbf{e}']_{\mathbf{e}}^{-1}[\varphi]_{\mathbf{e}}[\mathbf{e}']_{\mathbf{e}}. \tag{12.12}$$

Пусть A, B — две матрицы из $F^{n\times n}$. Говорят, что матрица B *подобна* матрице A, если для некоторой невырожденной матрицы $Q \in F^{n\times n}$ выполнено

$$B = Q^{-1}AQ. \tag{12.13}$$

Легко проверить, что отношение подобия матриц обладает свойствами рефлексивности, симметричности и транзитивности.

Сравнивая формулы (12.12) и (12.13), получаем

Следствие 12.21. *Матрицы одного и того же линейного преобразования в разных базисах подобны. Любая матрица, подобная заданной матрице линейного преобразования, является матрицей того же преобразования, записанной в некотором базисе.*

В пространстве $\Phi(V, V)$ определим операцию возведения в степень:

$$\varphi^m = \underbrace{\varphi\varphi\ldots\varphi}_{m \text{ раз}} \qquad (m \in \mathbb{N}).$$

Из (12.11) по индукции получаем $[\varphi^m] = [\varphi]^m$.

Значением многочлена

$$f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \ldots + a_{m-1}\lambda + a_m \in F[\lambda] \tag{12.14}$$

от преобразования φ назовем преобразование

$$f(\varphi) = a_0\varphi^m + a_1\varphi^{m-1} + \ldots + a_{m-1}\varphi + a_m\varepsilon.$$

Значением многочлена (12.14) *от матрицы* $A \in F^{n\times n}$ назовем матрицу

$$f(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mE.$$

Из (12.11) по индукции получаем $[f(\varphi)]_{\mathbf{e}} = f([\varphi]_{\mathbf{e}})$.

Утверждение 12.22. *Пусть φ — преобразование пространства V, а $f(t)$, $g(t)$ — многочлены из $F[t]$. Тогда*

$$f(\varphi)g(\varphi) = g(\varphi)f(\varphi), \qquad (f+g)(\varphi) = f(\varphi) + g(\varphi), \qquad (fg)(\varphi) = f(\varphi)g(\varphi).$$

12.7. Собственные числа и собственные векторы преобразования

Подпространство W линейного пространства V называется *инвариантным относительно преобразования* φ, если $\varphi W \subseteq W$, т. е. $\varphi x \in W$ для любого x из W. *Сужением преобразования* (*индуцированным преобразованием*) φ на инвариантное подпространство W называется преобразование $\psi : W \to W$ такое, что $\psi x = \varphi x$ для любого $x \in W$. Обозначение: $\varphi \mid W = \psi$. Очевидно, преобразование $\varphi \mid W$ — линейное.

Упражнение 12.23. Докажите, что образ и ядро линейного преобразования являются инвариантными пространствами. Опишите соответствующие индуцированные преобразования.

Ненулевой вектор $x \in V$ называется *собственным вектором* преобразования φ, если для некоторого $\lambda \in F$ выполнены условия

$$\varphi x = \lambda x. \tag{12.15}$$

Число λ при этом называется *собственным числом*, или *собственным значением* преобразования φ. Говорят также, что собственный вектор x *принадлежит*, или *относится* к собственному числу λ. Таким образом, собственный вектор — это ненулевой вектор, который отображается в коллинеарный.

Обозначим V_λ множество всех собственных векторов, принадлежащих собственному значению λ, дополненное нулевым вектором, иными словами:

$$V_\lambda = \{x \in V : \varphi x = \lambda x\}.$$

Равенство $\varphi x = \lambda x$ можно записать эквивалентным способом как $(\varphi - \lambda\varepsilon)x = o$. Отсюда получаем

$$V_\lambda = \operatorname{Ker}(\varphi - \lambda\varepsilon).$$

Следствие 12.24. *Множество V_λ является подпространством пространства V.*

Подпространство V_λ называется *собственным подпространством* преобразования φ, относящимся к собственному числу λ.

Упражнение 12.25. Доказать что V_λ инвариантно относительно φ.

Утверждение 12.26. *Если x — собственный вектор, то для любого ненулевого $\alpha \in F$ вектор αx — тоже собственный, относящийся к тому же собственному числу.*

Доказательство.

$$\varphi(\alpha x) = \alpha(\varphi x) = \alpha(\lambda x) = \lambda(\alpha x),$$

т. е. αx — собственный вектор, относящийся к собственному числу λ. ∎

Утверждение 12.27. *Вектор является собственным тогда и только тогда, когда он является базисным вектором одномерного инвариантного подпространства.*

Доказательство. Пусть x — собственный вектор. По предыдущему утверждению αx для любого $\alpha \neq 0$ — тоже собственный, относящийся к тому же собственному числу, поэтому подпространство $L(x)$ инвариантно.

Пусть $L(x)$, где x — некоторый ненулевой вектор, инвариантно, т. е. $\varphi x \in L(x)$, или $\varphi x = \lambda x$ для некоторого $\lambda \in F$, следовательно, вектор x — собственный. ∎

Таким образом, собственный вектор определяет прямую (одномерное подпространство), относительно которой происходит гомотетия с коэффициентом λ.

Исследуем задачу нахождения собственных векторов преобразования φ. Пусть $\mathbf{e} = \{e_1, \ldots, e_n\}$ — произвольный базис пространства V. Вектор x — собственный тогда и только тогда, когда $x \neq o$ и $(\varphi - \lambda\varepsilon)x = o$ для некоторого $\lambda \in F$. Переходя к координатам, получаем эквивалентное условие:

$$\left([\varphi]_{\mathbf{e}} - \lambda E\right)[x]_{\mathbf{e}} = o. \tag{12.16}$$

Итак, вектор x является собственным тогда и только тогда, когда его координатный столбец $[x]_{\mathbf{e}}$ является нетривиальным решением квадратной системы линейных уравнений (12.16). Для существования такого x необходимо и достаточно, чтобы

$$\det\left([\varphi]_{\mathbf{e}} - \lambda E\right) = 0. \tag{12.17}$$

По аналогии с (12.17) для произвольной матрицы $A = (a_{ij}) \in F^{n\times n}$ относительно неизвестного λ рассмотрим уравнение

$$\det(A - \lambda E) = 0,$$

называемое *характеристическим уравнением матрицы A*. Правая его часть

$$\chi_A(\lambda) = \det(A - \lambda E) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} \in F[\lambda]$$

представляет собой многочлен от λ, который называется *характеристическим многочленом матрицы A*.

Лемма 12.28. *Характеристические многочлены подобных матриц совпадают.*

Доказательство. Пусть $B = Q^{-1}AQ$ для некоторой невырожденной матрицы Q. Тогда

$$\chi_B(\lambda) = \det(B - \lambda E) = \det\Big(Q^{-1}AQ - \lambda Q^{-1}EQ\Big) =$$

$$= \det\Big(Q^{-1}(A - \lambda E)Q\Big) = (\det Q)^{-1}\det(A - \lambda E)\det Q = \det(A - \lambda E) = \chi_A(\lambda).$$

Итак, $\chi_A(\lambda) = \chi_B(\lambda)$. ∎

Из леммы следует, что уравнение (12.17), записанное для одного и того же преобразования φ в разных базисах, имеет один и тот же вид и поэтому корректны следующие определения. Уравнение (12.17) называется *характеристическим уравнением преобразования φ*. Левая часть этого уравнения

$$\chi(\lambda) = \det([\varphi]_{\mathbf{e}} - \lambda E)$$

называется *характеристическим многочленом преобразования φ*.

Из всего вышесказанного следует

Теорема 12.29. *Для того, чтобы $\lambda \in F$ являлось собственным числом преобразования φ необходимо и достаточно, чтобы оно являлось корнем характеристического многочлена этого преобразования.*

Итак, все корни характеристического многочлена, принадлежащие F, и только они являются собственными числами преобразования.

12.7.1.* Сумма определителей

Из свойства линейности определителя получаем.

Теорема 12.30 (о сумме определителей)**.** *Определитель суммы двух матриц A и B порядка n равен сумме всех 2^n определителей матриц, получающихся из A заменой части столбцов A соответствующими столбцами (т. е. столбцами с теми же номерами) из B (включая сами матрицы A и B). Аналогичное утверждение справедливо и для строк матриц.*

Пример 12.31. Для матриц второго порядка:

$$\det(A+B) = \begin{vmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12}+b_{12} \\ a_{21} & a_{22}+b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & a_{12}+b_{12} \\ b_{21} & a_{22}+b_{22} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & b_{12} \\ a_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & a_{12} \\ b_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

Пример 12.32. Вычислим определитель

$$\Delta = \begin{vmatrix} a_1b_1+x_1 & a_1b_2 & a_1b_3 & \dots & a_1b_n \\ a_2b_1 & a_2b_2+x_2 & a_2b_3 & \dots & a_2b_n \\ a_3b_1 & a_3b_2 & a_3b_3+x_3 & \dots & a_3b_n \\ \dots & \dots & \dots & \dots & \dots \\ a_nb_1 & a_nb_2 & a_nb_3 & \dots & a_nb_n+x_n \end{vmatrix}.$$

Заметим, что Δ есть определитель суммы двух матриц:

$$A = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x_n \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \dots & \dots & \dots & \dots \\ a_nb_1 & a_nb_2 & \dots & a_nb_n \end{pmatrix}$$

Применим к Δ теорему 12.30. Определитель матрицы, получающейся из A заменой более 1 столбца столбцами матрицы B, равен нулю, так как содержит по крайней мере два пропорциональных столбца. Поэтому Δ равен определителю матрицы A плюс сумма n определителей вида

$$\begin{vmatrix} x_1 & 0 & 0 & a_1b_j & 0 & 0 \\ 0 & x_2 & 0 & a_2b_j & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & x_{j-1} & a_{j-1}b_j & 0 & 0 \\ 0 & 0 & 0 & a_jb_j & 0 & 0 \\ 0 & 0 & 0 & a_{j+1}b_j & x_{j+1} & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & a_nb_j & 0 & x_n \end{vmatrix} = x_1x_2\dots x_{j-1}a_jb_jx_{j+1}\dots x_n$$

(раскрыли по первым $j-1$ столбцам). Таким образом,

$$\Delta = x_1x_2\dots x_n + \sum_{j=1}^{n} x_1x_2\dots x_{j-1}a_jb_jx_{j+1}\dots x_n.$$

12.7.2. Выражение коэффициентов характеристического многочлена через главные миноры матрицы

Миноры матрицы $A \in F^{n\times n}$ вида $A\binom{j_1,\dots,j_k}{j_1,\dots,j_k}$ называются *главными минорами* матрицы A. В частности,

$$A\begin{pmatrix}1,2,\dots,n\\1,2,\dots,n\end{pmatrix} = \det A.$$

Сумма диагональных элементов (сумма главных миноров первого порядка) матрицы A называется ее *следом* и обозначается $\operatorname{tr} A$.

Теорема 12.33. *Коэффициент s_k многочлена*

$$\det(A - \lambda E) = (-\lambda)^n + s_1(-\lambda)^{n-1} + \ldots + s_{n-1}(-\lambda) + s_n \tag{12.18}$$

равен сумме главных миноров порядка k. В частности, $s_1 = \operatorname{tr} A$, $s_n = \det A$.

Доказательство. Положим $B = -\lambda E$. По теореме 12.30 определитель $\det(A + B)$ можно представить в виде суммы 2^n определителей вида

- столбцы с номерами $j_1, \dots, j_k$, где $1 \le j_1 < \ldots < j_k \le n$, совпадают с соответствующими столбцами матрицы A (столбцы первой группы),
- в остальных столбцах на диагонали стоит $-\lambda$, на остальных местах 0 (столбцы второй группы).

Раскладывая каждый определитель по столбцам второй группы, мы получим, что он равен

$$(-\lambda)^{n-k} A\begin{pmatrix}j_1,\dots,j_k\\j_1,\dots,j_k\end{pmatrix}.$$

Для окончания доказательства осталось собрать слагаемые с одинаковым множителем λ^{n-k}. ∎

Пример 12.34. Распишем формулу (12.18) для $n = 2, 3$:

$$\begin{vmatrix} a_{11}-\lambda & a_{12}\\ a_{21} & a_{22}-\lambda\end{vmatrix} = \lambda^2 - (a_{11}+a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21});$$

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13}\\ a_{21} & a_{22}-\lambda & a_{23}\\ a_{31} & a_{32} & a_{33}-\lambda\end{vmatrix} =$$

$$= -\lambda^3 + (a_{11}+a_{22}+a_{33})\lambda^2 - \left(\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix} + \begin{vmatrix} a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix} + \begin{vmatrix} a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}\right)\lambda + \begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}.$$

12.7.3. Матрица Фробениуса

Матрица вида

$$F(a_1,\dots,a_n) = \begin{pmatrix} -a_1 & -a_2 & \dots & -a_{n-1} & -a_n\\ 1 & 0 & & 0 & 0\\ 0 & 1 & & 0 & 0\\ & & \ddots & & \\ 0 & 0 & & 1 & 0\end{pmatrix}$$

называется *матрицей Фробениуса*.

Утверждение 12.35. *Характеристический многочлен матрицы Фробениуса $F(a_1, \ldots, a_n)$ равен*

$$\Phi(a_1, \ldots, a_n; \lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n.$$

Доказательство. Доказательство проведем индукцией по n. Основание индукции легко проверяется. Найдем теперь $\det(\lambda E - F(a_1, \ldots, a_n))$. Раскладывая этот определитель по последнему столбцу и пользуясь предположением индукции, получаем

$$a_n(-1)^{1+n}(-1)^{n-1} + (-\lambda)\big((-\lambda)^{n-1} + a_1(-\lambda)^{n-2} + \ldots + a_{n-1}\big) = \Phi(a_1, \ldots, a_n; \lambda).$$

Утверждение доказано. ∎

Матрица Фробениуса $F(a_1, \ldots, a_n)$ называется также *сопровождающей матрицей* многочлена $\Phi(a_1, \ldots, a_n; \lambda)$.

Следствие 12.36. *Всякий многочлен степени n со старшим коэффициентом 1 может быть характеристическим многочленом некоторой квадратной матрицы порядка n.*

12.8. Диагонализируемость линейного преобразования

Преобразование называется *диагонализируемым*, если существует базис, в котором матрица преобразования имеет диагональный вид.

Теорема 12.37 (Очень простое, но очень важное утверждение)**.** *Преобразование диагонализируемо тогда и только тогда, когда существует базис из собственных векторов преобразования.*

Доказательство. Необходимость Пусть в базисе $e_1, \ldots, e_n$

$$[\varphi]_{\mathbf{e}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Так как j-й столбец матрицы $[\varphi]_{\mathbf{e}}$ есть

$$[\varphi e_j]_{\mathbf{e}} = (0, \ldots, 0, \lambda_j, 0, \ldots, 0)^\top,$$

то $\varphi e_j = \lambda_j e_j$, т. е. e_j — собственный вектор ($j = 1, \ldots, n$).
Достаточность Пусть базис $e_1, \ldots, e_n$ состоит из собственных векторов, тогда $\varphi e_j = \lambda_j e_j$ ($j = 1, \ldots, n$), откуда

$$[\varphi]_{\mathbf{e}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Теорема доказана. ∎

Пусть характеристический многочлен преобразования φ линейного пространства V, заданного над полем F, имеет вид

$$\chi(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s} q(\lambda),$$
$$\text{где } \lambda_i \in F, \lambda_i \neq \lambda_j \text{ при } i \neq j,$$
$$q(\lambda) \text{ корней в } F \text{ не имеет.}$$

Назовем *алгебраической кратностью* собственного значения λ_i величину k_i (т. е. его кратность, как корня характеристического многочлена).

Назовем *геометрической кратностью* собственного значения λ_i размерность собственного подпространства, принадлежащего собственному значению λ_i:

$$d_i = \dim V_{\lambda_i},$$

иными словами, величину $\operatorname{def}(\varphi - \lambda_i \varepsilon)$, или, что то же, максимальное число линейно независимых решений системы

$$([\varphi]_{\mathbf{e}} - \lambda_i E)x = 0.$$

Лемма 12.38. *Геометрическая кратность собственного числа не превосходит алгебраической.*

Доказательство. Пусть геометрическая кратность собственного значения λ_0 равна d_0. Следовательно, найдется система линейно независимых собственных векторов $e_1, \ldots, e_{d_0}$, относящихся к собственному значению λ_0. Дополним ее до базиса и в нем построим матрицу преобразования:

$$[\varphi]_{\mathbf{e}} = \left(\begin{array}{ccc|c} \lambda_0 & & & \\ & \ddots & & B \\ & & \lambda_0 & \\ \hline & 0 & & C \end{array}\right).$$

Очевидно, что

$$\chi_\varphi = \det(\lambda E - [\varphi]_{\mathbf{e}}) = (\lambda - \lambda_0)^{d_0} p(\lambda), \tag{12.19}$$

где $p(\lambda) \in F[\lambda]$. Из (12.19) получаем, что алгебраическая кратность собственного значения λ_0 не меньше d_0. ∎

Лемма 12.39. *Собственные векторы, относящиеся к разным собственным числам, линейно независимы.*

Доказательство. Пусть $\lambda_1, \ldots, \lambda_s$ — некоторые попарно различные собственные числа преобразования φ. В i-й строке следующей таблицы выписана произвольная линейно независимая система собственных векторов, относящихся к собственному значению λ_i:

$$\begin{array}{lcl} e_{11}, e_{12}, \ldots, e_{1d_1} & \to & \lambda_1; \\ e_{21}, e_{22}, \ldots, e_{2d_2} & \to & \lambda_2; \\ \ldots & & \\ e_{s1}, e_{s2}, \ldots, e_{sd_s} & \to & \lambda_s. \end{array} \tag{12.20}$$

Покажем, что объединенная система собственных векторов (12.20) линейно независима.

Доказательство проведем индукцией по s. При $s = 1$ система линейно независима по предположению. Предположим теперь, что векторы, стоящие в первых $s - 1$ строках таблицы (12.20), линейно независимы и покажем, что все векторы в (12.20) линейно независимы. Рассмотрим произвольную равную нулю линейную комбинацию этих векторов:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} e_{ij} = 0. \tag{12.21}$$

Применим к обеим частям (12.21) преобразование φ:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} \underbrace{\varphi e_{ij}}_{\lambda_i e_{ij}} = 0, \tag{12.22}$$

откуда

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} \lambda_i e_{ij} = 0. \tag{12.23}$$

Теперь домножим обе части (12.21) на λ_s:

$$\sum_{i=1}^{s} \sum_{j=1}^{d_i} \alpha_{ij} \lambda_s e_{ij} = 0. \tag{12.24}$$

Вычитая (12.24) из (12.23), получаем

$$\sum_{i=1}^{\cancel{s}\, s-1} \sum_{j=1}^{d_i} \alpha_{ij} (\lambda_i - \lambda_s) e_{ij} = 0.$$

В последнем равенстве имеем нулевую линейную комбинацию векторов, по предположению индукции линейно независимых. Поэтому все коэффициенты этой линейной комбинации равны нулю:

$$\alpha_{ij} \underbrace{(\lambda_i - \lambda_s)}_{\neq 0} = 0 \quad (i = 1, \ldots, s-1),$$

Откуда

$$\alpha_{ij} = 0 \quad (i = 1, \ldots, s-1;\ j = 1, \ldots, d_i).$$

Теперь в правой части (12.21) остается лишь одна сумма:

$$\sum_{j=1}^{d_s} \alpha_{sj} e_{sj} = 0.$$

Так как векторы $e_{s1}, \ldots, e_{sd_s}$ линейно независимы, то

$$\alpha_{sj} = 0 \quad (j = 1, \ldots, d_s).$$

Итак, все коэффициенты в произвольной линейной комбинации (12.21) равны нулю, поэтому система (12.20) линейно независима. ∎

Из двух предыдущих лемм и утверждения получаем следующий результат.

Теорема 12.40 (Критерий диагонализируемости)**.** *Преобразование диагонализируемо тогда и только тогда, когда сумма геометрических кратностей всех собственных чисел совпадает с размерностью пространства.*

12.9.* Быстрое преобразование Фурье

Преобразование $\varphi : \mathbb{C}^n \to \mathbb{C}^n$, определяемое формулой $\varphi x = F_n x$, где

$$F_n = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \omega^{(n-1)^2} \end{pmatrix}$$

и ω — какой-либо первообразный корень n-й степени из 1, называется (*дискретным*) *преобразованием Фурье*. Обычно в качестве ω берут

$$\omega_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n} \quad \text{или} \quad \omega_{n-1} = \cos\frac{2\pi}{n} - i\sin\frac{2\pi}{n}.$$

Матрица F_n называется *матрицей Фурье*. Обратим внимание, что эта матрица есть матрица Вандермонда, составленная для всех значений корня n-й степени из 1:

$$F_n = W(1, \omega, \omega^2, \dots, \omega^{n-1}).$$

Пусть x — некоторый вектор их $\mathbb{C}^n$. Обычное вычисление матрично-векторного произведения $F_n x$ требует $O(n^2)$ арифметических операций (предполагая, что все значения ω^j вычислены заранее). Пусть $n = 2^k$, где $k \in \mathbb{Z}$. Покажем, как в этом случае можно учесть специальный вид матрицы F_n и произведение $F_n x$ вычислить с помощью только $O(n \log n)$ арифметических операций.

Будем нумеровать строки и столбцы матрицы F_n, а также компоненты вектора x индексами от 0 до $n-1$. Переставим в F_n столбцы так, что в начале идут столбцы с четными индексами, а затем — с нечетными. Получим матрицу $F_n P_n$, где P_n — соответствующая матрица-перестановка. Напомним, что $P^{-1} = P^\top$. Определим

$$D_n = \begin{pmatrix} 1 & & & & \\ & \omega & & & \\ & & \omega^2 & & \\ & & & \ddots & \\ & & & & \omega^{n-1} \end{pmatrix}.$$

Пусть $n = 2^k$, $k \geq 1$. Обозначим $m = 2^{k-1} = n/2$. Можно проверить, что

$$F_n P = \begin{pmatrix} F_m & D_m F_m \\ F_m & -D_m F_m \end{pmatrix} = \begin{pmatrix} E_m & D_m \\ E_m & -D_m \end{pmatrix} \cdot \begin{pmatrix} F_m & O \\ O & F_m \end{pmatrix}.$$

Тогда

$$F_n x = F_n P P^\top x = \begin{pmatrix} E_m & D_m \\ E_m & -D_m \end{pmatrix} \cdot \begin{pmatrix} F_m x^{(0)} \\ F_m x^{(1)} \end{pmatrix}, \tag{12.25}$$

где $x^{(0)}$ — вектор, составленный из компонент вектора x с четными индексами, а $x^{(1)}$ — с нечетными.

Согласно (12.25) для вычисления преобразования Фурье $F_n x$ достаточно два раза рекурсивно вычислить это преобразование для векторов длины $m = n/2$: $F_m x^{(0)}$ и $F_m x^{(1)}$, и, кроме того, потратить еще m умножений (на диагональные элементы матрицы D_m), m сложений и m вычитаний.

Обозначим $t(n)$ количество арифметических операций, используемых для вычисления преобразования Фурье $F_n x$, где $n = 2^k$. Имеем

$$t(n) = 3 \cdot \frac{n}{2} + 2 \cdot t\Big(\frac{n}{2}\Big) \qquad (n \geq 2).$$

Так как для вычисления преобразования Фурье при $n = 1$ вообще не требуется арифметических операций, то $t(1) = 0$. Теперь по индукции легко доказать, что

$$t(n) = \frac{3}{2} n \log_2 n.$$

Полученный алгоритм называется алгоритмом *быстрого преобразования Фурье.*

Упражнение 12.41. Доказать, что

$$F_n^{-1} = \frac{1}{n} F_n,$$

откуда, в частности, получить, что при $n = 2^k$ обратное преобразование Фурье также можно вычислить за время $O(n \log n)$.

12.10. Аннулирующий многочлен

Говорят, что многочлен $f(\lambda)$ *аннулирует* (обращает в ноль) преобразование φ на подпространстве $W \subseteq V$, если $W \subseteq \operatorname{Ker} f(\varphi)$, иными словами, $f(\varphi)x = 0$ для произвольного вектора $x \in W$. В данном случае $f(\lambda)$ называется также *аннулирующим многочленом* преобразования φ на подпространстве (относительно подпространства) W.

Аннулирующий многочлен минимальной степени со старшим коэффициентом 1 называется *минимальным аннулирующим*, или просто *минимальным.*

Пример 12.42.

1) Для нулевого преобразования θ минимальным аннулирующим многочленом является константа 1.
2) Минимальным многочленом, аннулирующим тождественное преобразование ε на произвольном подпространстве, является, очевидно, $f(\lambda) = \lambda - 1$.
3) Пусть

$$[\varphi] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

В данном случае второй базисный вектор e_2 собственный, поэтому относительно подпространства $L(e_2)$ аннулирующим (и минимальным) многочленом, является $\lambda - 1$. Относительно всего пространства, легко проверить, аннулирующим является многочлен $f(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$. Действительно,

$$[f(\varphi)] = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^2 - 2\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Заметим, что многочлен $f(\lambda)$ совпадает с характеристическим.

Утверждение 12.43. *Минимальный многочлен существует и единственен.*

Доказательство. Существование Вначале мы докажем существование аннулирующего многочлена. В линейном пространстве $\Phi(V, V)$ рассмотрим систему векторов $\varepsilon, \varphi, \varphi^2, \ldots, \varphi^{n^2}$. Так как $\dim \Phi(V, V) = n^2$, то рассматриваемая система линейно зависима, т. е. найдутся коэффициенты $\alpha_0, \alpha_1, \ldots, \alpha_{n^2}$, не все равные нулю, такие, что

$$\alpha_0\varepsilon + \alpha_1\varphi + \ldots \alpha_{n^2}\varphi^{n^2} = 0.$$

Очевидно, что ненулевой многочлен $f(\lambda) = \alpha_0 + \alpha_1\lambda + \ldots \alpha_{n^2}\lambda^{n^2}$ аннулирует преобразование φ на любом подпространстве $W \subseteq V$. Отсюда следует существование минимального аннулирующего многочлена.

Единственность Пусть $f_1(\lambda), f_2(\lambda)$ – минимальные многочлены преобразования φ на подпространстве W. Докажем, что многочлен $f_1(\lambda) - f_2(\lambda)$ является аннулирующим. Действительно,

$$(f_1(\varphi) - f_2(\varphi))x = f_1(\varphi)x - f_2(\varphi)x = 0$$

для произвольного вектора $x \in W$. У разности $f_1(\lambda) - f_2(\lambda)$ старшие члены взаимно уничтожаются, следовательно, либо $f_1(\lambda) - f_2(\lambda) = 0$, либо степень $f_1(\lambda) - f_2(\lambda)$ меньше степени многочленов $f_1(\lambda)$, $f_2(\lambda)$. Последнее однако невозможно, так как $f_1(\lambda), f_2(\lambda)$ – минимальные, следовательно, $f_1(\lambda) - f_2(\lambda) = 0$, и поэтому $f_1(\lambda) = f_2(\lambda)$. ∎

Утверждение 12.44. *Пусть $f(\lambda)$ – минимальный многочлен преобразования φ относительно подпространства W. Тогда множество всех многочленов, аннулирующих φ на W, есть множество $\{f(\lambda)q(\lambda) : \; q(\lambda) \in F[\lambda]\}$.*

Доказательство. Так как $q(\varphi)f(\varphi)x = q(\varphi)o = o$ для произвольного $x \in W$, то $f(\lambda)q(\lambda)$ — аннулирующий.

Покажем, что других аннулирующих многочленов, кроме многочленов вида $f(\lambda)q(\lambda)$, нет. Для этого поделим произвольный аннулирующий многочлен $g(\lambda)$ с остатком на $f(\lambda)$:

$$g(\lambda) = f(\lambda)q(\lambda) + r(\lambda), \tag{12.26}$$

причем либо $r(\lambda) = 0$, либо $\deg r(\lambda) < \deg f(\lambda)$. Так как $r(\varphi)x = g(\varphi)x - q(\varphi)f(\varphi)x = 0$ для произвольного $x \in W$, то $r(\lambda)$ – также аннулирующий, поэтому $r(\lambda) = 0$. Теперь из (12.26) получаем $g(\lambda) = f(\lambda)q(\lambda)$. ∎

Утверждение 12.45. *Пусть многочлены $f_1(\lambda)$, $f_2(\lambda)$ являются минимальными аннулирующими многочленами для преобразования φ на подпространствах W_1, W_2 соответственно, причем $W_1 \subseteq W_2$. Тогда $f_2 \vdots f_1$.*

Доказательство. Так как $W_1 \subseteq W_2$, то $f_2(\lambda)$ аннулирует φ на W_1. Доказываемое теперь следует из предыдущего утверждения. ∎

Утверждение 12.46. *Пусть многочлены $f_1(\lambda)$, $f_2(\lambda)$ являются минимальными аннулирующими многочленами для преобразования φ на подпространствах W_1, W_2 соответственно. Тогда* $\text{НОК}(f_1(\lambda), f_2(\lambda))$ *является минимальным аннулирующим многочленом для преобразования φ на подпространстве $W_1 + W_2$.*

Доказательство. Сначала докажем, что $f(\lambda) = \text{НОК}(f_1(\lambda), f_2(\lambda))$ аннулирует преобразование φ на подпространстве $W_1 + W_2$. Пусть

$$f(\lambda) = q_1(\lambda)f_1(\lambda) = q_2(\lambda)f_2(\lambda).$$

Для произвольного вектора $x = x_1 + x_2 \in W$, где $x_1 \in W_1$, $x_2 \in W_2$, имеем $f(\varphi)x = f(\varphi)(x_1 + x_2) = f(\varphi)x_1 + f(\varphi)x_2 = q_1(\varphi)f_1(\varphi)x_1 + q_2(\varphi)f_2(\varphi)x_2 = 0$. Итак, многочлен $f(\lambda)$ – аннулирующий на $W_1 + W_2$.

Докажем теперь, что $f(\lambda)$ – минимальный многочлен, т.е. из всех аннулирующих многочленов многочлен $f(\lambda)$ имеет минимальную степень. Так как $W_1 \subseteq W_1 + W_2$ и $W_2 \subseteq W_1 + W_2$, то по предыдущему утверждению произвольный аннулирующий на подпространстве $W_1 + W_2$ многочлен должен делиться без остатка как на многочлен $f_1(\lambda)$, так и на многочлен $f_2(\lambda)$. Из всех многочленов, удовлетворяющих этим свойствам, минимальную степень имеет $f(\lambda)$. Следовательно, этот многочлен минимальный аннулирующий. ∎

12.10.1. Метод Крылова построения минимального многочлена

В данном разделе мы опишем метод Крылова[1] нахождения минимального многочлена, аннулирующего преобразование φ на всем пространстве V.

Пусть $e_1, e_2, \ldots, e_n$ – базис пространства V. Для произвольного $i \in \{1, 2, \ldots, n\}$ в силу конечномерности пространства найдется натуральное s (зависящее от i), такое, что система

$$e_i, \varphi e_i, \varphi^2 e_i, \ldots, \varphi^{s-1} e_i$$

линейно независима, а система

$$e_i, \varphi e_i, \varphi^2 e_i, \ldots, \varphi^{s} e_i$$

[1] академик А. Н. Крылов (1863–1945)

линейно зависима. Тогда найдутся числа $\alpha_0, \alpha_1, \ldots, \alpha_{s-1}$, такие, что

$$\alpha_0 e_i + \alpha_1 \varphi e_i + \ldots + \alpha_{s-1} \varphi^{s-1} e_i + \varphi^s e_i = o.$$

Легко видеть, что многочлен

$$f_i(\lambda) = \alpha_0 + \alpha_1 \lambda + \ldots + \alpha_{s-1} \lambda^{s-1} + \lambda^s$$

является минимальным на подпространстве $L(e_i)$.

По утверждению 12.46 минимальным многочленом, аннулирующим φ на всем пространстве V, является НОК$(f_1(\lambda), \ldots, f_n(\lambda))$.

Пример 12.47. Построить минимальный аннулирующий многочлен преобразования φ, заданного в некотором базисе матрицей

$$[\varphi] = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

Для базиса e_1, e_2, e_3, в котором задана матрица преобразования, имеем

$$[e_1] = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad [\varphi e_1] = \begin{pmatrix} 0 \\ -4 \\ -2 \end{pmatrix}, \quad [\varphi^2 e_1] = \begin{pmatrix} -4 \\ -16 \\ -8 \end{pmatrix};$$

$$[e_2] = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad [\varphi e_2] = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, \quad [\varphi^2 e_2] = \begin{pmatrix} 4 \\ 12 \\ 4 \end{pmatrix};$$

$$[e_3] = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad [\varphi e_3] = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

В каждой строке приведенной таблицы векторы линейно зависимые. Найдем их нулевые нетривиальные комбинации:

$$\begin{array}{lll} 4e_1 - 4\varphi e_1 + \varphi^2 e_1 = 0 & \to & f_1(\lambda) = 4 - 4\lambda + \lambda^2 = (\lambda - 2)^2; \\ 4e_2 - 4\varphi e_2 + \varphi^2 e_2 = 0 & \to & f_2(\lambda) = 4 - 4\lambda + \lambda^2 = (\lambda - 2)^2; \\ 2e_3 - \varphi e_3 = 0 & \to & f_3(\lambda) = \lambda - 2. \end{array}$$

Минимальный многочлен имеет вид $f(\lambda) = \text{НОК}(f_1(\lambda), f_2(\lambda), f_3(\lambda)) = \lambda - 2$. Как и в примере 12.42 в данном случае характеристический многочлен $\chi(\lambda) = (\lambda - 2)^3$ является аннулирующим. В следующей теореме утверждается, что это справедливо для любого преобразования.

Теорема 12.48 (Теорема Гамильтон–Кэли)**.** *Любая квадратная матрица является корнем своего характеристического многочлена. Любое линейное преобразование является корнем своего характеристического многочлена.*

Доказательство. Мы покажем, что произвольная матрица A является корнем своего характеристического многочлена $\chi_A(\lambda)$, т. е. $\chi_A(A) = O$.

Рассмотрим квадратную матрицу $\mathrm{adj}(\lambda E - A) = (\alpha_{ij})$, в которой $\alpha_{ij} = (-1)^{i+j} A\binom{j}{i}$. Имеем

$$\mathrm{adj}(\lambda E - A) \cdot (\lambda E - A) = \det(\lambda E - A) \cdot E.$$

В полученное равенство подставим $\lambda = A$, тогда в его правой части получаем $\chi_A(A)$, а в левой O. ∎

Замечание 12.49. Доказанная теорема равносильна утверждению, что характеристический многочлен преобразования является аннулирующим многочленом того же преобразования на всем пространстве.

Замечание 12.50. Приведенное доказательство не основывалось на результатах данного раздела и является другим доказательством существования аннулирующего многочлена.

Теорема 12.51 (Теорема о корнях аннулирующего многочлена)**.** *Пусть многочлен*

$$\chi(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s}$$
$$(\lambda_i \neq \lambda_j \text{ } \textit{при} \text{ } i \neq j)$$

является характеристическим многочленом преобразования φ*. Тогда минимальным многочленом, аннулирующим* φ *на всем пространстве, является многочлен*

$$\mu(\lambda) = (\lambda - \lambda_1)^{l_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{l_s}, \tag{12.27}$$

где $l_1, \ldots, l_s$ *– некоторые натуральные числа, такие, что* $1 \leq l_i \leq k_i$ $(i = 1, 2, \ldots, s)$.

Доказательство. По теореме Гамильтона–Кэли (теорема 12.48) характеристический многочлен является аннулирующим. Минимальный многочлен является его делителем и, поэтому, имеет вид (12.27), причем $l_i \leq k_i$ $(i = 1, 2, \ldots, s)$.

Теперь докажем, что $l_i \geq 1$. Для этого рассмотрим собственный вектор x. Имеем $\varphi x = \lambda_i x$. Легко видеть, что многочлен $f_i(\lambda) = \lambda - \lambda_i$ является минимальным аннулирующим преобразование φ на подпространстве $L(x)$. Необходимое неравенство следует теперь из утверждения 12.46. ∎

Замечание 12.52. Из доказанной теоремы следует, что множество корней минимального многочлена без учета их кратностей есть в точности множество характеристических чисел. Кратности, с которыми они встречаются в минимальном многочлене не превосходят алгебраических кратностей характеристических чисел.

12.11. Жорданова форма линейного преобразования

Жордановой клеткой называется квадратная матрица порядка n, следующего вида:

$$J_n(\lambda_0) = \begin{pmatrix} \lambda_0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & \lambda_0 & \ddots & 0 & 0 \\ \ldots & \ldots & \ldots & \ddots & \ddots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_0 \end{pmatrix}.$$

Жордановой матрицей J называется блочно-диагональная матрица, составленная из жордановых клеток:

$$J = \operatorname{diag}(J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \ldots, J_{k_s}(\lambda_s)).$$

Базис пространства V называется *жордановым*, если в этом базисе матрица преобразования является жордановой (имеет *жорданову форму*). В данном случае базис также называют каноническим базисом преобразования, а саму матрицу преобразования в этом базисе – *каноническим видом*, или *жордановой формой* преобразования.

Теорема 12.53 (Теорема Жордана). *Для любого преобразования φ комплексного линейного n-мерного пространства существует жорданов базис. Жорданова форма определена единственным с точностью до перестановки жордановых клеток образом.*

Теорема 12.54 (Терема Жордана (матричная формулировка)). *Для любой матрицы $A \in \mathbb{C}^{n\times n}$ найдется такая невырожденная матрица Q, что $J = Q^{-1}AQ$ есть жорданова матрица. Матрица J определяется единственным образом с точностью до перестановки жордановых клеток.*

Следствие 12.55 (Критерий подобия матриц на полем C). *Для того, чтобы матрицы A и B из $A \in \mathbb{C}^{n\times n}$ были подобны над полем $\mathbb{C}$ необходимо и достаточно совпадение их жордановых форм.*

Доказательство теоремы Жордана приводится в трех следующих разделах.

12.12.* Разложение пространства в прямую сумму корневых подпространств

Рассмотрим минимальный многочлен

$$f(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}\dots(\lambda - \lambda_s)^{k_s},$$

аннулирующий преобразование φ на всем пространстве V.

Множество $P_i = \operatorname{Ker}(\varphi - \lambda_i\varepsilon)^{k_i}$ называется *корневым подпространством*, принадлежащим собственному числу (корню) λ_i $(i = 1, 2, \dots, s)$.

Теорема 12.56. *Пространство V есть прямая сумма корневых подпространств P_i:*

$$V = P_1 \dotplus P_2 \dotplus \dots \dotplus P_s.$$

Доказательство. Легко видеть, что многочлены

$$f_i(\lambda) = \frac{f(\lambda)}{(\lambda - \lambda_i)^{k_i}}, \quad (i = 1, 2, \dots, s)$$

взаимно просты, поэтому найдутся такие $u_1(\lambda), u_2(\lambda), \dots, u_s(\lambda)$, что

$$1 = \sum_{i=1}^{s} f_i(\lambda)u_i(\lambda),$$

откуда

$$\varepsilon = \sum_{i=1}^{s} f_i(\varphi)u_i(\varphi),$$

следовательно, для любого $x \in V$

$$x = \sum_{i=1}^{s} \underbrace{f_i(\varphi)u_i(\varphi)x}_{x_i}.$$

Обозначим $x_i = f_i(\varphi)u_i(\varphi)x$, тогда
$$x = \sum_{i=1}^{s} x_i. \tag{12.28}$$
Докажем, что формула (12.28) задает разложение произвольного вектора x по прямой сумме $P_1 \dot{+} \ldots \dot{+} P_s$.

Сперва покажем, что $x_i \in P_i$. Действительно,
$$(\varphi - \lambda_i\varepsilon)^{k_i} x_i = \underbrace{(\varphi - \lambda_i\varepsilon)^{k_i} f_i(\varphi)}_{f(\varphi)=\theta} u_i(\varphi)x = o.$$

Теперь докажем, что рассматриваемая сумма прямая. Для этого достаточно установить единственность разложения по пространствам P_i $(i = 1, 2, \ldots, s)$ для нулевого вектора. Пусть
$$o = \sum_{j=1}^{s} x_j, \text{ где } x_j \in P_j \qquad (12.29).$$
Применим к обеим частям равенства (12.29) преобразование $f_i(\varphi)$, тогда
$$o = \sum_{j=1}^{s} f_i(\varphi)x_j = f_i(\varphi)x_i,$$
последнее равенство справедливо в силу
$$f_i(\varphi)x_j = \prod_{\substack{\nu=1\\ \nu\neq i}}^{s} (\varphi - \lambda_\nu\varepsilon)^{k_\nu} x_j = o.$$
Так как многочлены $f_i(\lambda)$ и $(\lambda - \lambda_i)^{k_i}$ взаимно простые, то существуют такие $u(\lambda)$, $v(\lambda)$, что
$$1 = u(\lambda)f_i(\lambda) + v(\lambda)(\lambda - \lambda_i)^{k_i},$$
поэтому
$$x_i = u(\varphi)\underbrace{f_i(\varphi)x_i}_{o} + v(\varphi)\underbrace{(\varphi - \lambda_i\varepsilon)^{k_i} x_i}_{o,\ \text{т. к. } x_i \in P_i} = o.$$
Итак, компонента разложения $x_i = o$ $(i = 1, 2 \ldots, s)$ определяется однозначно, поэтому рассматриваемая сумма прямая. ∎

Замечание 12.57. Вместо $f(\lambda)$ в теореме можно рассматривать любой аннулирующий многочлен преобразования φ. Формулировка теоремы и доказательство при этом никак не изменятся.

Замечание 12.58. Корневые пространства P_i $(i = 1, 2, \ldots, s)$ инвариантны относительно преобразования φ.

Замечание 12.59. Сужение преобразования φ на подпространство P_i (индуцированное преобразование) имеет одно собственное значение λ_i.

12.13.* Построение жорданова базиса

В силу замечаний 12.58, 12.59 мы можем на время ограничиться рассмотрением линейного преобразования φ пространства V с одним собственным числом λ_0. В этом случае минимальный многочлен имеет вид $f(\lambda) = (\lambda - \lambda_0)^k$. Обозначим $\psi = \varphi - \lambda_0 \varepsilon$, тогда

$$\psi^k = \theta. \tag{12.30}$$

Назовем *высотой* вектора $x \in V$ такое $h \geq 0$, что $\psi^{h-1}x \neq 0$, но $\psi^h x = 0$. В нашем случае высота произвольного вектора не превосходит k. Существует единственный вектор высоты 0 — нулевой вектор. Векторы высоты 1 это в точности собственные векторы преобразования.

Обозначим $L_h = \operatorname{Ker} \psi^h$. Иными словами, L_h — это множество векторов высоты $\leq h$. Имеем

$$\{0\} = L_0 \subseteq \underbrace{L_1}_{\substack{\text{собственное}\\ \text{подпространство}}} \subseteq \ldots \subseteq L_k = L_{k+1} = \ldots = V.$$

Следующая лемма показывает, что $L_h \neq L_{h+1}$ при $h \leq k-1$, и поэтому

$$\{0\} = L_0 \subset L_1 \subset \ldots \subset L_k = L_{k+1} = \ldots = V.$$

Лемма 12.60. *Если $L_h = L_{h+1}$, то $L_h = L_l$ для любого $l \geq h$.*

Доказательство. От противного. Пусть

$$L_h = L_{h+1} \neq L_{h+2},$$

тогда найдется такой $x \in V$, что

$$x \notin L_h = L_{h+1}, \quad x \in L_{h+2}.$$

Используя определение пространств L_{h+2} и L_{h+1}, соответственно получаем

$$\begin{aligned} \psi^{h+2}x = 0 \quad &\Rightarrow \quad \psi^{h+1}\psi x = 0 \quad \Rightarrow \quad \psi x \in L_{h+1}, \\ \psi^{h+1}x \neq 0 \quad &\Rightarrow \quad \psi^{h}\psi x \neq 0 \quad \Rightarrow \quad \psi x \notin L_{h}. \end{aligned}$$

Итак, $\psi x \in L_{h+1}$, однако $\psi x \notin L_h$, что невозможно, так как $L_h = L_{h+1}$. ∎

Опишем алгоритм построения жорданова базиса.

- на предварительном шаге необходимо построить базисы пространств $L_1, L_2, \ldots, L_k$;
- далее найти такую линейно независимую систему $a_{11}, \ldots, a_{1t_1}$, для которой справедливо соотношение
$$L_k = L_{k-1} \dotplus L(a_{11}, \ldots, a_{1t_1}); \tag{12.31}$$
- для каждого $h = 2, 3, \ldots, k$ необходимо найти линейно независимую систему $a_{h1}, \ldots, a_{ht_h}$ (возможно, $t_h = 0$), такую, что
$$\begin{aligned} L_{k-h+1} = L_{k-h} \dotplus \\ \dotplus L(\psi^{h-1}a_{11}, \ldots, \psi^{h-1}a_{1t_1}, \ldots, \psi a_{h-1,1}, \ldots, \psi a_{h-1,t_{h-1}}, a_{h1}, \ldots, a_{ht_h}), \end{aligned} \tag{12.32}$$
где $L_0 = \{o\}$ — нулевое подпространство.

Итак, мы имеем

$$
\begin{aligned}
L_k &= L_{k-1} \dotplus L(a_{11}, \ \ldots, a_{1t_1}),\\
L_{k-1} &= L_{k-2} \dotplus L(\psi a_{11}, \ \ldots, \psi a_{1t_1}, \ a_{21}, \ \ldots, a_{2t_2}),\\
L_{k-2} &= L_{k-3} \dotplus L(\psi^2 a_{11}, \ldots, \psi^2 a_{1t_1}, \psi a_{21}, \ldots, \psi a_{2t_2}, a_{31}, \ldots, a_{3t_3}),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
L_{k-h+2} &= L_{k-h+1} \dotplus L(\psi^{h-2} a_{11}, \ldots, \psi^{h-2} a_{1t_1}, \ldots, \ a_{h-1,1}, \ \ldots, a_{h-1,t_{h-1}}),\\
L_{k-h+1} &= L_{k-h} \quad \dotplus L(\psi^{h-1} a_{11}, \ldots, \psi^{h-1} a_{1t_1}, \ldots, \psi a_{h-1,1}, \ldots, \psi a_{h-1,t_{h-1}}, a_{h1}, \ldots, a_{ht_h}),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
L_1 &= L_0 \dotplus L(\psi^{k-1} a_{11}, \ldots, \psi^{k-1} a_{1t_1}, \psi^{k-2} a_{21}, \ldots, \psi^{k-2} a_{2t_2}, \ldots, \psi a_{k-1,1}, \ldots, \psi a_{k-1,t_k}, a_{k1}, \ldots, a_{kt_k}).
\end{aligned}
$$

В настоящем пункте мы докажем, что построенная система векторов

$$
\begin{array}{ccccccc}
a_{11}, \ldots, & a_{1t_1}, & & & & & \\
\psi a_{11}, \ldots, & \psi a_{1t_1}, & a_{21}, \ldots, & a_{2t_2}, & & & \\
\vdots & \vdots & \vdots & \vdots & & & \\
\psi^{k-1} a_{11}, \ldots, & \psi^{k-1} a_{1t_1}, & \psi^{k-2} a_{21}, \ldots, & \psi^{k-2} a_{2t_2}, & \ldots & a_{k1}, \ldots & a_{kt_k}
\end{array}
\tag{12.33}
$$

образует жорданов базис пространства V.

Сперва докажем возможность построения системы (12.33).

Лемма 12.61. *Системы векторов*

$$a_{h1}, \ldots, a_{ht_h} \quad (h = 1, 2, \ldots, k),$$

удовлетворяющие условиям (12.31)–(12.32)*, существуют.*

Доказательство. Лемма доказывается индукцией по h.
Основание индукции $h = 1$ Система $a_{11}, \ldots a_{1t_1}$ является дополнением базиса пространства L_{k-1} до базиса пространства L_k. Так как $L_{k-1} \subset L_k$, то такая система существует.
Индуктивный переход $h - 1 \to h$ Пусть $u_1, \ldots, u_p$ — базис подпространства L_{k-h}, а $v_1, \ldots, v_r$ — базис подпространства L_{k-h+1}. Обозначим векторы

$$\psi^{h-2} a_{11}, \ldots, \psi^{h-2} a_{1t_1}, \ldots, a_{h-1,1}, \ldots, a_{h-1,t_{h-1}}$$

через $w_1, \ldots, w_q$ соответственно. По предположению индукции, векторы

$$v_1, \ldots, v_r, w_1, \ldots, w_q$$

составляют базис подпространства L_{k-h+2}. Итак, имеем

$$
\begin{aligned}
L_{k-h+2} &= L_{k-h+1} \dotplus L(w_1, \ldots, w_q),\\
L_{k-h+1} &= L(v_1, \ldots, v_r),\\
L_{k-h} &= L(u_1, \ldots, u_p).
\end{aligned}
$$

Докажем, что векторы

$$u_1, \ldots, u_p, \psi w_1, \ldots, \psi w_q$$

образуют линейно независимую систему в L_{k-h+1} и, следовательно, могут быть дополнены (векторами $a_{h1}, \dots, a_{ht_h}$) до базиса пространства L_{k-h+1}. Для этого рассмотрим линейную комбинацию

$$\sum_{j=1}^{p} \alpha_j u_j + \sum_{i=1}^{q} \beta_i \psi w_i = 0. \tag{12.34}$$

Применим преобразование ψ^{k-h} к обеим частям (12.34):

$$\sum_{j=1}^{p} \alpha_j \underbrace{\psi^{k-h} u_j}_{0} + \psi^{k-h+1} \sum_{i=1}^{q} \beta_i w_i = 0.$$

Получим

$$\psi^{k-h+1} \sum_{i=1}^{q} \beta_i w_i = 0 \quad \Rightarrow \quad \sum_{i=1}^{q} \beta_i w_i \in L_{k-h+1}.$$

Раскладывая эту сумму по базису пространства L_{k-h+1}, получаем

$$\sum_{i=1}^{q} \beta_i w_i = \sum_{j=1}^{r} \gamma_j v_j \tag{12.35}$$

для некоторых $\gamma_1, \dots, \gamma_r$. Так как по предположению индукции векторы $v_1, \dots, v_r, w_1, \dots, w_q$ линейно независимы, то в (12.35)

$$\beta_i = 0 \quad (i = 1, \dots, q); \qquad \gamma_j = 0 \quad (j = 1, \dots, r).$$

Теперь из (12.34) получаем

$$\sum_{j=1}^{p} \alpha_j u_j = 0.$$

Так как векторы $u_1, \dots, u_p$ линейно независимы, то $\alpha_j = 0$ $(j = 1, \dots, r)$. Итак, в линейной комбинации (12.34) коэффициенты равны нулю и, следовательно, векторы $v_1, \dots, v_r, \psi w_1, \dots, \psi w_q$ линейно независимы. ∎

Лемма 12.62. *Система (12.33) образует базис пространства V.*

Доказательство. По построению L_1 есть линейная оболочка векторов, стоящих в нижней строке таблицы (12.33). Подпространство L_2 есть линейная оболочка векторов, стоящих в двух последних строках. Поднимаясь так далее снизу вверх по таблице (12.33) мы получаем, что пространство $L_k = V$ есть линейная оболочка векторов системы (12.33). ∎

Вертикальный ряд векторов в таблице (12.33) назовем *цепочкой*. Векторы в одной цепочке будем рассматривать в последовательности от собственного вектора к вектору максимальной высоты (снизу вверх). Покажем, что каждой цепочке соответствует одна жорданова клетка. Таким образом, вся совокупность векторов (12.33) является жордановым базисом.

Лемма 12.63. *Каждой цепочке векторов $a_1, \dots a_r$ таблицы (12.33) в матрице преобразования соответствует жорданова клетка.*

Доказательство. Для доказательства достаточно заметить, что вектор a_1 — собственный, поэтому $\varphi a_1 = \lambda_0 a_1$, и $a_j = \psi a_{j+1}$ $(j = 2, 3, \dots, r)$, поэтому $\varphi a_{j+1} = \lambda_0 a_{j+1} + a_j$. ∎

В заключение пункта отметим, что в случае нескольких собственных значений преобразования φ необходимо проделать те же вычисления с каждым собственным числом.

12.14.* Единственнось жордановой формы

Лемма 12.64. *Для произвольного натурального* i

$$\operatorname{rank}(J_n(\lambda_0)-\lambda E))^i = \begin{cases} n-i, & \text{если} \quad \lambda=\lambda_0, \quad n-i>0, \\ 0, & \text{если} \quad \lambda=\lambda_0, \quad n-i\le 0, \\ n, & \text{если} \quad \lambda\ne\lambda_0. \end{cases}$$

Обозначим через $l_j = l_j(\lambda_0)$ число жордановых клеток вида $J_j(\lambda_0)$ в жорданой форме J преобразования φ. Пусть $r_j = r_j(\lambda_0) = \operatorname{rank}(\varphi-\lambda_0\varepsilon)^j$. Следующая лемма показывает, что жорданова форма преобразования определена однозначно с точностью до перестановки диагональных клеток.

Теорема 12.65. *Для числа жордановых клеток порядка* j *в жордановой форме* J *справедливо соотношение*

$$l_j(\lambda_0) = r_{j-1}(\lambda_0) - 2r_j(\lambda_0) + r_{j+1}(\lambda_0).$$

Доказательство. Выразим величины

$$r_j(\lambda_0) = \operatorname{rank}(\varphi-\lambda_0\varepsilon)^j = \operatorname{rank}(J-\lambda_0 E)^j$$

через $l_j(\lambda_0)$. При возведении блочно-диагональной матрицы в степень каждый блок возводится в степень независимо и матрица сохраняет блочную структуру. Таким образом, величина $r_j(\lambda_0)$ равна сумме рангов каждого блока. Обозначим через $\bar r$ сумму рангов блоков, соответствующих собственным числам, отличным от λ_0. Пользуясь леммой, получаем

$$r_j(\lambda_0) = l_{j+1} + 2l_{j+2} + \ldots + (p-j)l_p + \bar r = \sum_{i=j+1}^{p}(i-j)l_i + \bar r,$$

где p — максимальный порядок жордановой клетки, откуда

$$r_{j-1} = \sum_{i=j}^{p}(i-j+1)l_i = l_j + 2l_{j+1} + \sum_{i=j+2}^{p}(i-j+1)l_i = l_j + 2l_{j+1} + \sum_{i=j+2}^{p}(i-j)l_i + \sum_{i=j+2}^{p} l_i,$$

$$r_j = \sum_{i=j+1}^{p}(i-j)l_i = l_{j+1} + \sum_{i=j+2}^{p}(i-j)l_i, \qquad r_{j+1} = \sum_{i=j+2}^{p}(i-j-1)l_i = \sum_{i=j+2}^{p}(i-j)l_i - \sum_{i=j+2}^{p} l_i.$$

Легко видеть, что $r_{j-1} - 2r_j + r_{j+1} = l_j$. ∎

Итак, число жордановых клеток заданного порядка с заданным числом на диагонали полностью определяется формулами из предыдущей теоремы и зависит только от преобразования, но никак не от базиса. Поэтому *жорданова форма (но не базис) линейного преобразования определены однозначно с точностью до перестановок диагональных клеток.*

12.15.* Матричные функции

В настоящем разделе введем понятие знаяения функции от матрицы.

12.15.1. Значение многочлена от матрицы

Понятие значения многочлена от матрицы было введено естественным образом. Покажем, как во многих случаях удается упростить вычисления.

Пример 12.66. Для матрицы

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$

найдем A^{2000}.

Минимальный многочлен матрицы A совпадает с характеристическим и равен $f(\lambda) = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2$. Разделим λ^{2000} на $f(\lambda)$ с остатком:

$$\lambda^{2000} = q(\lambda)(\lambda-2)^2 + a_0\lambda + a_1. \tag{12.36}$$

Нам нет необходимости находить явный вид многочлена $q(\lambda)$. Для нахождения коэффициентов a_0, a_1 подставим в (12.36) вместо λ значение 2. Тогда получим:

$$2^{2000} = 2a_0 + a_1. \tag{12.37}$$

Теперь продифференцируем (12.36):

$$2000\lambda^{1999} = 2q(\lambda)(\lambda-2) + a_0,$$

и снова подставим сюда вместо λ значение 2. Получим новое соотношение:

$$2000 \cdot 2^{1999} = a_0. \tag{12.38}$$

Из (12.37), (12.38) находим $a_0 = 1000 \cdot 2^{2000}$, $a_1 = -1999 \cdot 2^{2000}$. Теперь подставим в (12.36) матрицу A. Так как $(\lambda-2)^2$ — ее минимальный многочлен, то первое слагаемое в правой части обратится в ноль. Получаем

$$A^{2000} = a_0A + a_1E = 1000 \cdot 2^{2000} \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix} - 1999 \cdot 2^{2000} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2^{2000} \begin{pmatrix} -999 & 1000 \\ -1000 & 1001 \end{pmatrix}.$$

Значение многочлена от матрицы удобно находить, зная ее жорданову форму. Рассмотрим этот способ.

Утверждение 12.67. *Если матрицы A, B подобны, причем $B = Q^{-1}AQ$, тогда для любого многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ выполнено соотношение*

$$g(B) = Q^{-1}g(A)Q. \tag{12.39}$$

Утверждение 12.68. *Для произвольного многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ и произвольной блочно-диагональной матрицы* $\operatorname{diag}(A_1, \ldots, A_p)$

$$g(\operatorname{diag}(A_1, \ldots, A_p)) = \operatorname{diag}(g(A_1), \ldots, g(A_p)). \tag{12.40}$$

Утверждение 12.69.

$$(J_n(\lambda_0))^m = \begin{pmatrix} \lambda_0^m & \binom{m}{1}\lambda_0^{m-1} & \binom{m}{2}\lambda_0^{m-2} & \ldots & \binom{m}{n-1}\lambda_0^{m-n+1} \\ 0 & \lambda_0^m & \binom{m}{1}\lambda_0^{m-1} & \ldots & \binom{m}{n-2}\lambda_0^{m-n+2} \\ \ldots & & & & \\ 0 & 0 & 0 & \ldots & \lambda_0^m \end{pmatrix}. \tag{12.41}$$

Доказательство. Утверждение легко доказывается индукцией по m. ∎

Утверждение 12.70. *Для произвольного многочлена $g(\lambda) \in \mathbb{C}[\lambda]$ выполняется соотношение*

$$g(J_n(\lambda_0)) = \begin{pmatrix} g(\lambda_0) & \dfrac{g'(\lambda_0)}{1!} & \dfrac{g''(\lambda_0)}{2!} & \dots & \dfrac{g^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & g(\lambda_0) & \dfrac{g'(\lambda_0)}{1!} & \dots & \dfrac{g^{(n-2)}(\lambda_0)}{(n-2)!} \\ \dots & & & & \\ 0 & 0 & 0 & \dots & g(\lambda_0) \end{pmatrix}. \tag{12.42}$$

Доказательство. Если $g(\lambda) = \lambda^m$, то формула 12.41 превращается в 12.42. Далее доказательство не должно вызывать затруднений. ∎

Пример 12.71. Решим задачу из примера 12.66 другим способом. Жорданова форма $J = Q^{-1}AQ$ матрицы A и соответствующая трансформирующая матрица Q равны

$$J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \qquad Q = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix},$$

По утверждению 12.69 имеем

$$J^{2000} = \begin{pmatrix} 2^{2000} & 2000 \cdot 2^{1999} \\ 0 & 2^{2000} \end{pmatrix} = \begin{pmatrix} 2^{2000} & 1000 \cdot 2^{2000} \\ 0 & 2^{2000} \end{pmatrix} = 2^{2000} \begin{pmatrix} 1 & 1000 \\ 0 & 1 \end{pmatrix}.$$

По утверждению 12.67 получаем

$$A^{2000} = QJ^{2000}Q^{-1} = 2^{2000} \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1000 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} = 2^{2000} \begin{pmatrix} -999 & 1000 \\ -1000 & 1001 \end{pmatrix}.$$

12.15.2. Функции от матриц

Пусть

$$f(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s}$$

— минимальный многочлен матрицы A. Будем говорить, что произвольная функция $g(\lambda) : \mathbb{C} \to \mathbb{C}$ *задана на спектре матрицы* A, если существует набор значений

$$\begin{aligned} & g(\lambda_1), g'(\lambda_1), \ldots, g^{k_1-1}(\lambda_1); \\ & g(\lambda_2), g'(\lambda_2), \ldots, g^{k_2-1}(\lambda_1); \\ & \dots \\ & g(\lambda_s), g'(\lambda_s), \ldots, g^{k_s-1}(\lambda_1). \end{aligned} \tag{12.43}$$

В данном случае говорят также, что *функция $g(\lambda)$ на спектре матрицы A принимает значения* (12.43). Обозначим через $g_A(\lambda)$ многочлен минимальной степени, совпадающий на спектре матрицы A с функцией $g(\lambda)$ (так называемый интерполяционный многочлен Лагранжа–Эрмита, заданный на спектре матрицы A). Значение $g(A)$ функции $g(\lambda)$ определим как значение многочлена $g_A(\lambda)$ от матрицы A:

$$g(A) = g_A(A).$$

Сразу отметим, что, по предыдущей теореме, *степень интерполяционного многочлена Лагранжа–Эрмита, заданного на спектре матрицы A, меньше степени минимального многочлена этой матрицы.*

Пример 12.72. Для матрицы
$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$
найти e^A.

Минимальный многочлен $f(\lambda) = (\lambda - 2)^2$ имеет вторую степень, поэтому ищем интерполяционный многочлен в виде
$$g_A(\lambda) = f_1\lambda + f_2.$$
На спектре матрицы A функция $g(\lambda) = e^\lambda$ принимает следующие значения: $g(2) = e^2$, $g'(2) = e^2$, поэтому
$$\begin{aligned} g_A(2) &= 2f_1 + f_2 = e^2, \\ g'_A(2) &= f_1 = e^2; \end{aligned}$$
откуда получаем $f_1 = e^2$, $f_2 = -e^2$, следовательно, $g_A(\lambda) = e^2\lambda - e^2$, поэтому
$$e^A = g_A(A) = e^2A - e^2E = \begin{pmatrix} 0 & e^2 \\ -e^2 & 2e^2 \end{pmatrix}.$$

Следующее утверждение показывает корректность введенного определения для того случая, когда функция $g(\lambda)$ сама является многочленом.

Теорема 12.73. *Пусть $g(\lambda)$, $h(\lambda)$ — многочлены из $\mathbb{C}[\lambda]$, A — матрица из $\mathbb{C}^{n\times n}$. Для того, чтобы $g(A) = h(A)$ необходимо и достаточно, чтобы значения многочленов $g(\lambda)$, $h(\lambda)$ совпадали на спектре матрицы A.*

Доказательство. Пусть
$$f(\lambda) = (\lambda - \lambda_1)^{k_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{k_s}$$
— минимальный многочлен матрицы A, тогда
$$\begin{aligned} & g(A) = h(A) \\ \Leftrightarrow \quad & g(A) - h(A) = 0 \\ \Leftrightarrow \quad & g(\lambda) - h(\lambda) \text{ — аннулирующий многочлен матрицы } A \\ \Leftrightarrow \quad & g(\lambda) - h(\lambda) = q(\lambda)f(\lambda) \text{ для некоторого } q(\lambda) \in \mathbb{C}[\lambda] \\ \Leftrightarrow \quad & \lambda_i \text{ — корень } g(\lambda) - h(\lambda) \text{ кратности } \geq k_i \ (i = 1, 2, \ldots, s) \\ \Leftrightarrow \quad & g^{(j)}(\lambda_i) - h^{(j)}(\lambda_i) = 0 \ (i = 1, \ldots, s; j = 0, \ldots, k_i - 1). \end{aligned}$$
∎

Пример 12.74. Для матрицы
$$A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$$
найти A^{2000}.

Этот пример уже рассматривался (см. примеры 12.66, 12.71). Приведем еще один способ решения. Для функции $g(\lambda) = \lambda^{2000}$ найдем интерполяционный многочлен
$$g_A(\lambda) = f_1\lambda + f_2,$$
заданный на спектре матрицы A. Имеем
$$\begin{aligned} g_A(2) &= 2f_1 + f_2 = 2^{2000}, \\ g'_A(2) &= f_1 = 2000 \cdot 2^{1999}; \end{aligned}$$

откуда $f_1 = 2000 \cdot 2^{1999}$, $f_2 = -1999 \cdot 2^{2000}$, следовательно,

$$g_A(\lambda) = 2000 \cdot 2^{1999}\lambda - 1999 \cdot 2^{2000} = 2^{2000}(1000\lambda - 1999).$$

Подставляя $\lambda = A$ в $g_A(\lambda)$, получаем

$$A^{2000} = 2^{2000} \cdot \begin{pmatrix} -999 & 1000 \\ -1000 & 1001 \end{pmatrix}.$$

Следующие утверждения являются аналогами утверждений 12.67, 12.68, 12.70.

Утверждение 12.75. *Если матрицы A, B подобны, причем $B = Q^{-1}AQ$, тогда для любой функции $g(\lambda)$, определенной на спектре матрицы A, $g(B)$ существует, причем*

$$g(B) = Q^{-1}g(A)Q.$$

Доказательство. Так как A и B подобны, то у них совпадают минимальные многочлены и, следовательно, наборы значений функции $g(\lambda)$ на их спектре, поэтому $g(B)$ существует и $g_A(\lambda) = g_B(\lambda)$. Теперь по утверждению 12.67 получаем

$$g(B) = g_B(B) = g_A(B) = Q^{-1}g_A(A)Q = Q^{-1}g(A)Q.$$

∎

Утверждение 12.76. *Для произвольной функции $g(\lambda)$, определенной на спектре блочно-диагональной матрицы $\mathrm{diag}(A_1, \ldots, A_p)$ функция $g(A_i)$ существует для каждого $i = 1, 2 \ldots, p$ и*

$$g(\mathrm{diag}(A_1, \ldots, A_p)) = \mathrm{diag}(g(A_1), \ldots, g(A_p)).$$

Доказательство. Пусть $f_i(\lambda)$ — минимальный многочлен матрицы A_i ($i = 1, 2, \ldots, p$), а $f(\lambda)$ — минимальный многочлен матрицы A. Так как $f(\lambda) = \text{НОК}(f_1(\lambda), \ldots, f_p(\lambda))$, то все корни многочлена $f_i(\lambda)$ являются корнями многочлена $f(\lambda)$, причем в последнем каждый из корней имеет не меньшую кратность (спектр матрицы A_i содержится в спектре матрицы A). Следовательно, $g(A_i)$ существует, $g_A(\lambda) = g_{A_i}(\lambda)$ и поэтому, используя утверждение 12.68, получаем

$$\begin{aligned} &g(\mathrm{diag}(A_1, \ldots, A_p)) \\ &\quad = g_A(\mathrm{diag}(A_1, \ldots, A_p)) \\ &\quad = \mathrm{diag}(g_A(A_1), \ldots, g_A(A_p)) \\ &\quad = \mathrm{diag}(g_{A_1}(A_1), \ldots, g_{A_p}(A_p)) \\ &\quad = \mathrm{diag}(g(A_1), \ldots, g(A_p)). \end{aligned}$$

∎

Утверждение 12.77. *Пусть функция $g(\lambda)$ определена на спектре матрицы $J_n(\lambda_0)$, тогда справедлива формула* (12.42).

Доказательство. Для краткости обозначим $A = J_n(\lambda_0)$. Минимальный многочлен матрицы A равен $f(\lambda) = (\lambda - \lambda_0)^n$, поэтому интерполяционный многочлен $g_A(\lambda)$ удовлетворяет n условиям вида

$$g_A^{(j)}(\lambda_0) = g^{(j)}(\lambda_0) \quad (j = 0, 1, \ldots, n-1),$$

откуда по утверждению 12.70 получаем доказываемое. ∎