Г л а в а 4

ЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ ВЕКТОРНЫХ ПРОСТРАНСТВ

§ 30. Линейные преобразования

30.1. Определения и примеры. Пусть V – векторное пространство над полем P. Отображение $\varphi : V \longrightarrow V$ называется *линейным преобразованием* пространства V, если

$$(\alpha u + \beta v)\varphi = \alpha(u\varphi) + \beta(v\varphi)$$

для всех $\alpha, \beta \in P$ и $u, v \in V$.

П р и м е р ы линейных преобразований.

1) Преобразование 0, переводящее всякий вектор в нулевой вектор.

2) Тождественное преобразование ε, переводящее всякий вектор $u \in V$ в себя.

3) Пусть $P[x]$ – пространство многочленов. Преобразование φ, являющееся дифференцированием, линейно, что следует из свойства производных: $(\alpha f + \beta g)' = \alpha f' + \beta g'$ для любых $\alpha, \beta \in P$ и $f, g \in P[x]$.

4) Для всякой подстановки π из S_n можно определить преобразование φ_π на пространстве n–к P^n, полагая

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)\varphi_\pi = (\alpha_{1\pi}, \alpha_{2\pi}, \ldots, \alpha_{n\pi}).$$

В частности, если $\pi = (12)$, то

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)\varphi_{(12)} = (\alpha_2, \alpha_1, \alpha_3, \ldots, \alpha_n).$$

5) Функция $f : V \longrightarrow P$ называется *линейной*, если

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$$

для всех $\alpha, \beta \in P$ и $u, v \in V$. Пусть $e_1, e_2, \ldots, e_n$ – база пространства V, $f_1, f_2, \ldots, f_n$ – линейные функции из V в P. Определим преобразование $\varphi : V \longrightarrow V$, положив

$$u\varphi = f_1(u)e_1 + f_2(u)e_2 + \ldots + f_n(u)e_n$$

для всякого вектора u из V. Это преобразование линейно. В некотором смысле это общий пример линейного преобразования.

30.2 Матрица линейного преобразования. Пусть φ – линейное преобразование векторного пространства V,

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

– база V. Действуя на каждый базисный вектор e_i пространства V преобразованием φ, получим

$$\begin{cases} e_1\varphi = \alpha_{11}e_1 + \alpha_{12}e_2 + \ldots + \alpha_{1n}e_n, \\ e_2\varphi = \alpha_{21}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{2n}e_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n\varphi = \alpha_{n1}e_1 + \alpha_{n2}e_2 + \ldots + \alpha_{nn}e_n, \end{cases} \quad (1)$$

где $\alpha_{ij} \in P$.

Матрицу

$$[\varphi] = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}$$

будем называть *матрицей линейного преобразования* φ *в базе* e. Если необходимо подчеркнуть, что эта матрица в базе e, то будем писать $[\varphi]_e$. Кратко система (1) записывается в виде

$$e\varphi = [\varphi]e.$$

П р и м е р. В пространстве многочленов $P[x]$ выберем базу $1, x, x^2, \ldots$. Чтобы найти матрицу преобразования дифференцирования φ, действуем на каждый вектор преобразованием φ. Получим

$$\begin{cases} 1\varphi = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + \ldots, \\ x\varphi = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + \ldots, \\ x^2\varphi = 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Тогда матрица дифференцирования

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

Если рассмотреть пространство $P_{\leq n}[x]$ – многочлены степени не превосходящей n для некоторого фиксированного n, то это конечномерное пространство с базой $1, x, x^2, \dots, x^n$ и матрица диффернцирования имеет вид

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

30.3. Координаты образа вектора. Пусть u – произвольный вектор из V, разлагая его по базе e, получим

$$u = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = [u]e, \tag{2}$$

где $[u] = (\alpha_1, \alpha_2, \dots, \alpha_n)$ – координаты вектора u в базе e. Подействуем теперь на вектор u преобразованием φ, получим

$$u\varphi = [u\varphi]e, \tag{3}$$

где $[u\varphi]$ – координаты вектора $u\varphi$. Хотим найти связь между координатами $[u]$ и $[u\varphi]$.

Вспоминая определение матрицы преобразования φ, имеем

$$e\varphi = [\varphi]e.$$

Ввиду линейности φ

$$u\varphi = (\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n)\varphi = \alpha_1(e_1\varphi) + \alpha_2(e_2\varphi) + \dots + \alpha_n(e_n\varphi) =$$

$$= [u](e\varphi) = [u]([\varphi]e) = ([u][\varphi])e.$$

Напомним следующее утверждение, которое было доказано в прошлом семестре.

Л е м м а. *Пусть $v_1, v_2, \dots, v_n$ – линейно независимая система векторов, X и Y – матрицы размера $m \times n$. Если положить $v = (v_1, v_2, \dots, v_n)'$, где штрих означает транспонирование, то из равенства $Xv = Yv$ следует равенство матриц: $X = Y$.*

Используя эту лемму, получим искомое равенство

$$[u][\varphi] = [u\varphi],$$

т. е. координаты вектора $u\varphi$ есть произведение координат вектора u на матрицу преобразования φ.

30.4. Связь между матрицами линейного преобразования в разных базах. Пусть

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}, \quad e' = \begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_n \end{pmatrix},$$

– две базы векторного пространства V. Мы можем найти матрицы линейного преобразования φ в этих базах:

$$e\varphi = [\varphi]e, \tag{4}$$

$$e'\varphi = [\varphi]'e', \tag{5}$$

получим соответственно матрицы $[\varphi]$ и $[\varphi]'$. Возникает естественный вопрос: как связаны эти матрицы?

Мы знаем, что базы e и e' связаны матрицей перехода, т. е.

$$e' = Te, \tag{6}$$

для некоторой невырожденной матрицы $T = (t_{ij}) \in M_n(P)$. Тогда, используя равенства (4)–(5) и (6), получим

$$(T[\varphi])e = T([\varphi]e) = T(e\varphi) = (Te)\varphi = e'\varphi = [\varphi]'e' = [\varphi]'(Te) = ([\varphi]'T)e.$$

Воспользовавшись сформулированной выше леммой, получим

$$T[\varphi] = [\varphi]'T.$$

Домножив обе части справа на T^{-1}, получим окончательную формулу

$$[\varphi]' = T[\varphi]T^{-1}.$$

30.5. Алгебра линейных преобразований. Символом L обозначим множество линейных преобразований векторного пространства V. Определим на L операции сложения, умножения и умножение на скаляр из P.

Пусть φ и ψ – два линейных преобразования из L. Их *суммой* $\varphi + \psi$ назовем преобразование, которое действует по правилу

$$u(\varphi + \psi) = u\varphi + u\psi, \quad u \in V.$$

Их *произведением* $\varphi \cdot \psi$ назовем преобразование, которое действует по правилу

$$u(\varphi \cdot \psi) = (u\varphi)\psi,$$

и, наконец, для произвольного α из P определим преобразование $\alpha\varphi$, полученное умножением на скаляр α формулой

$$u(\alpha\varphi) = (\alpha u)\varphi.$$

Легко убедиться, что L относительно этих операций является алгебраической системой:

$$\langle\, L;\, +,\, \cdot,\, f_\alpha\, (\alpha \in P)\, \rangle.$$

О п р е д е л е н и е. Алгебраическая система

$$\langle\, A;\, +,\, \cdot,\, f_\alpha\, (\alpha \in P)\, \rangle$$

называется *линейной алгеброй* над полем P, если алгебраическая система $\langle\, A;\, +,\, f_\alpha\, (\alpha \in P)\, \rangle$ является векторным пространством над полем P, а $\langle\, A;\, +,\, \cdot\, \rangle$ является кольцом и выполняется следующая аксиома:

$$\alpha(\varphi\psi) = (\alpha\varphi)\psi = \varphi(\alpha\psi), \text{ для любых } \varphi, \psi \in A, \quad \alpha \in P.$$

П р и м е р ы линейных алгебр. 1) Множество квадратных матриц $M_n(P)$ над полем P с операциями сложения, умножения и умножение на скаляр образует линейную алгебру над полем P.

2) Множество многочленов $P[x_1, x_2, \ldots, x_n]$ над полем P с операциями сложения, умножения и умножение на скаляр образует линейную алгебру над полем P.

Покажем, что множество линейных преобразований векторного пространства является линейной алгеброй, которая изоморфна алгебре матриц. Более точно, справедлива

Т е о р е м а. *Множество линейных преобразований векторного пространства V размерности n над полем P является линейной алгеброй, изоморфной алгебре матриц $M_n(P)$.*

Д о к а з а т е л ь с т в о. Установим изоморфизм. Пусть e – некоторая база пространства V. Сопоставим каждому линейному преобразованию φ векторного пространства V его матрицу $[\varphi]$ в базе e. Покажем, что это отображение

$$\omega : L \longrightarrow M_n(P) \text{ по правилу } \varphi\omega = [\varphi]$$

и есть искомый изоморфизм.

Так как каждому линейному преобразованию соответствует его матрица, которая определяется однозначно при фиксированной базе, то наше отображение однозначно.

Предположим, что для двух линейных преобразований $\varphi, \psi \in L$ их матрицы равны: $[\varphi] = [\psi]$. Покажем, что тогда и $\varphi = \psi$. Для этого достаточно показать, что для произвольного вектора $v \in V$ справедливо равенство $v\varphi = v\psi$. Так как $v = [v]e$, то $v\varphi = [v][\varphi]e$. С другой стороны, $v\psi = [v][\psi]e$, а так как матрицы равны, то отсюда следует, что $v\varphi = v\psi$.

Проверим, что отображение $\omega : L \longrightarrow M_n(P)$ является отображением *на*. Пусть $A \in M_n(P)$, определим отображение φ по правилу $v\varphi = ([v]A)e \in V$. Проверим, что φ линейно. Выберем $v_1, v_2 \in V$, $\alpha_1, \alpha_2 \in P$, тогда

$$(\alpha_1 v_1 + \alpha_2 v_2)\varphi = [\alpha_1 v_1 + \alpha_2 v_2]Ae = (\alpha_1[v_1] + \alpha_2[v_2])Ae.$$

Воспользовавшись дистрибутивностью, получим

$$(\alpha_1[v_1] + \alpha_2[v_2])Ae = (\alpha_1[v_1]A + \alpha_2[v_2]A)e = \alpha_1[v_1]Ae + \alpha_2[v_2]Ae = \alpha_1(v_1\varphi) + \alpha_2(v_2\varphi).$$

Проверим, что матрица $[\varphi]$ в базе e равна A. Возьмем вектор e_i из базы e и подействуем на него преобразованием φ:

$$e_i\varphi = [0, \ldots, 0, 1, 0, \ldots, 0]Ae = a_{i1}e_1 + a_{i2}e_2 + \ldots + a_{in}e_n,$$

Следовательно, $[\varphi] = A$, т. е. для каждой матрицы A построили линейное преобразование матрица которого равна A.

Проверим, что построенное отображение ω сохраняет операции. Возьмем преобразования $\varphi, \psi \in L$ и подействуем на базу их произведением:

$$e(\varphi \cdot \psi) = (e\varphi)\psi = ([\varphi]e)\psi = [\varphi](e\psi) = [\varphi]([\psi]e) = ([\varphi][\psi])e.$$

С другой стороны

$$e(\varphi\psi) = [\varphi\psi]e.$$

Отсюда, ввиду леммы $[\varphi\psi] = [\varphi][\psi]$. Таким образом операция умножения сохраняется.

Рассмотрим теперь суму двух линейных преобразований. По определению матрицы линейного преобразования, $e(\varphi+\psi) = [\varphi+\psi]e$. С другой стороны, по определению суммы линейных преобразований:

$$e(\varphi + \psi) = e\varphi + e\psi = [\varphi]e + [\psi]e = ([\varphi] + [\psi])e$$

Опять ввиду леммы, приходим к равенству $[\varphi + \psi] = [\varphi] + [\psi]$. Следовательно, операция сложения сохраняется.

Пусть теперь $\alpha \in P$. Тогда

$$e(\alpha\varphi) = (\alpha e)\varphi = (\alpha[\varphi])e.$$

С другой стороны, $e(\alpha\varphi) = [\alpha\varphi]e$, а потому $[\alpha\varphi] = \alpha[\varphi]$. Изоморфизм установлен. Теорема доказана.

Из этой теоремы, в частности, следует, что линейные преобразования удовлетворяют аксиоме ассоциативности умножения, дистрибутивности и т. д.

§ 31. Образ и ядро линейного преобразования

31.1. Образ и ядро – подпространства векторного пространства. С каждым линейным преобразованием φ можно связать его *образ*

$$\operatorname{Im}\varphi = \{v\varphi \mid v \in V\} = V\varphi,$$

и *ядро*

$$\operatorname{Ker}\varphi = \{v \in V \mid v\varphi = 0\}.$$

Справедлива

Т е о р е м а 1. *Для всякого линейного преобразования φ его образ $\operatorname{Im}\varphi$ и ядро $\operatorname{Ker}\varphi$ являются подпространствами пространства V.*

Д о к а з а т е л ь с т в о. Чтобы проверить, что $\operatorname{Im}\varphi$ является подпространством, возьмем $u_1, u_2 \in \operatorname{Im}\varphi$ и α_1, $\alpha_2 \in P$, и покажем, что их линейная комбинация $\alpha_1 u_1 + \alpha_2 u_2$ лежит в образе $\operatorname{Im}\varphi$. Так как u_1 и u_2 лежат в образе, то для некоторых векторов v_1, v_2 имеем равенства $v_1\varphi = u_1$, $v_2\varphi = u_2$. Следовательно,

$$\alpha_1 u_1 + \alpha_2 u_2 = \alpha_1(v_1\varphi) + \alpha_2(v_2\varphi) = (\alpha_1 v_1 + \alpha_2 v_2)\varphi \in \operatorname{Im}\varphi.$$

Выберем теперь векторы u_1 и u_2 из $\operatorname{Ker}\varphi$. Действуя на линейную комбинацию $\alpha_1 u_1 + \alpha_2 u_2$ преобразованием φ, получим

$$(\alpha_1 u_1 + \alpha_2 u_2)\varphi = \alpha_1(u_1\varphi) + \alpha_2(u_2\varphi) = \alpha_1 0 + \alpha_2 0 = 0,$$

т. е. линейная комбинация $\alpha_1 u_1 + \alpha_2 u_2$ лежит в ядре $\operatorname{Ker}\varphi$. Теорема доказана.

Следующая теорема устанавливает связь между размерностью ядра и образа линейного преобразования.

Т е о р е м а 2. *Справедлива следующая формула*

$$\dim(\operatorname{Im}\varphi) + \dim(\operatorname{Ker}\varphi) = \dim V.$$

Д о к а з а т е л ь с т в о. Пусть $u_1, u_2, \ldots, u_s$ – база ядра $\operatorname{Ker}\varphi$, а $v_1, v_2, \ldots, v_r$ – база образа $\operatorname{Im}\varphi$. Для каждого v_i символом w_i обозначим его прообраз при преобразовании φ, т. е. $w_i\varphi = v_i$, $i = 1, 2, \ldots, r$.

Покажем, что система

$$u_1, u_2, \ldots, u_s, w_1, w_2, \ldots, w_r,$$

образует базу пространства V. Проверим вначале, что эти векторы линейно независимы. Рассмотрим их линейную комбинацию

$$\sum_{i=1}^{s} \alpha_i u_i + \sum_{j=1}^{r} \beta_j w_j = 0,$$

и, применяя преобразование φ, получим

$$0 = \sum_{i=1}^{s} \alpha_i (u_i\varphi) + \sum_{j=1}^{r} \beta_j (w_j\varphi) = \sum_{j=1}^{r} \beta_j v_j.$$

Так как векторы $v_1, v_2, \ldots, v_r$ линейно независимы, то все $\beta_j = 0$, а тогда из равенства

$$\sum_{i=1}^{s} \alpha_i u_i = 0,$$

ввиду линейной независимости системы $u_1, u_2, \ldots, u_s$ заключаем, что и $\alpha_i = 0$. Следовательно, система векторов линейно независима.

Покажем, что система векторов является максимальной. Рассмотрим произвольный вектор $v \in V$. Тогда вектор $v\varphi$ лежит в $\operatorname{Im}\varphi$, а потому разлагается по базе, т. е.

$$v\varphi = \sum_{j=1}^{r} \gamma_j v_j$$

для некоторых $\gamma_j \in P$. Рассмотрим вектор

$$v - (\sum_{j=1}^{r} \gamma_j w_j)$$

и покажем, что он лежит в ядре $\operatorname{Ker}\varphi$. Действительно,

$$\left(v - \sum_{j=1}^{r} \gamma_j w_j\right)\varphi = v\varphi - \sum_{j=1}^{r} \gamma_j v_j = 0.$$

Следовательно,

$$v - \sum_{j=1}^{r} \gamma_j w_j = \sum_{i=1}^{s} \delta_i u_i$$

для некоторых δ_i из P, а потому мы имеем искомое разложение:

$$v = \sum_{j=1}^{r} \gamma_j w_j + \sum_{i=1}^{s} \delta_i u_i.$$

Теорема доказана.

31.2. Невырожденные линейные преобразования.

Теорема 3. *Для линейного преобразования φ векторного пространства V следующие утверждения эквивалентны*:

1) $\operatorname{Im}\varphi = V$;
2) $\operatorname{Ker}\varphi = 0$;
3) φ – *взаимно однозначно и на*;
4) *обратное преобразование φ^{-1} существует и линейно*;
5) *матрица $[\varphi]$ невырождена в любой базе*;
6) *матрица $[\varphi]$ невырождена в некоторой базе.*

Если выполнено любое из этих условий, то преобразование φ называется *невырожденным.*

Доказательство. 1) $\Leftrightarrow$ 2). Равенство $\operatorname{Im}\varphi = V$, по предыдущей теореме, равносильно тому, что $\dim(\operatorname{Im}\varphi) = \dim V$ и $\dim(\operatorname{Ker}\varphi) = 0$, а это равносильно тому, что, $\operatorname{Ker}\varphi = 0$.

2) $\Rightarrow$ 3). Предположим, что $u\varphi = v\varphi$, тогда $(u-v)\varphi = 0$, т. е. $u - v \in \operatorname{Ker}\varphi$, а так как $\operatorname{Ker}\varphi = 0$, то $u - v = 0$ и $u = v$. Следовательно, φ унивалентно. Так как из 2) следует 1), то $\operatorname{Im}\varphi = V$, а это и означает, что φ – отображение *на*.

3) $\Rightarrow$ 4). Так как для любого вектора $v \in V$ существует единственный вектор $u \in V$ такой, что $u\varphi = v$, то в качестве φ^{-1} возьмем отображение, которое переводит вектор v в вектор u, т. е. $v\varphi^{-1} = u$. Проверим, что справедливы равенства $\varphi\varphi^{-1} = \varphi^{-1}\varphi = \varepsilon$, где ε – тождественное отображение. Действительно,

$$u(\varphi\varphi^{-1}) = (u\varphi)\varphi^{-1} = v\varphi^{-1} = u$$

и аналогично

$$v\varphi^{-1}\varphi = u\varphi = v.$$

Следовательно, φ^{-1} действительно является обратным отображением. Чтобы показать, что оно линейно, возьмем v_1 и v_2 из V такие, что $u_1\varphi = v_1$ и $u_2\varphi = v_2$. Так как

$$(\alpha_1 u_1 + \alpha_2 u_2)\varphi = \alpha_1 v_1 + \alpha_2 v_2,$$

то это и показывает, что φ^{-1} линейно.

4) $\Rightarrow$ 5). Из равенства $\varphi\varphi^{-1} = \varepsilon$ следует равенство матриц $[\varphi\varphi^{-1}] = [\varepsilon] = E$. В свою очередь, $[\varphi\varphi^{-1}] = [\varphi][\varphi^{-1}]$, а потому $[\varphi]$ – невырождено и $\det[\varphi] \cdot \det[\varphi^{-1}] \neq 0$.

5) $\Rightarrow$ 6). Очевидно.

6) $\Rightarrow$ 1). Пусть $e_1, e_2, \ldots, e_n$ – такая база пространства V, в которой матрица $[\varphi]$ невырождена. Возьмем вектор $v \in V$ и найдем его координаты $[v]$ в базе e. Рассмотрим систему $X[\varphi] = [v]$, относительно неизвестных $X = (x_1, x_2, \ldots, x_n)$. Эта система имеет единственное решение X^0. Положим $u = X^0 e$, тогда

$$u\varphi = X^0(e\varphi) = X^0([\varphi]e) = [v]e = v.$$

Таким образом, мы показали, что для любого вектора v найдется вектор $u \in V$ такой, что $u\varphi = v$, но это и означает, что $\mathrm{Im}\varphi = V$. Теорема доказана.

§ 32. Инвариантные подпространства

32.1. Инвариантные подпространства и неприводимость. Пусть U – подпространство векторного пространства V, а φ – линейное преобразование пространства V.

О п р е д е л е н и е. Говорим, что подпространство U *инвариантно относительно* φ, если $U\varphi \subseteq U$.

П р и м е р. Пусть $V = P[x]$ – пространство всех многочленов над полем P, $U = P_3[x]$ – подпространство многочленов степени не выше 3-й, φ – дифференцирование. Тогда $U\varphi \subseteq U$, т. е. U инвариантно относительно φ.

О п р е д е л е н и е. Пространство V называется *неприводимым относительно* φ, если V не содержит собственных (отличных от нулевого и всего V) инвариантных относительно φ подпространств.

О п р е д е л е н и е. Матрица следующего вида

$$\left(\begin{array}{c|c} A & * \\ \hline \mathbf{0} & B \end{array}\right), \quad A \in M_r(P), \quad B \in M_s(P),$$

называется *полураспавшейся.*

Т е о р е м а 1. *Для линейного преобразования φ векторного пространства V равносильны следующие утверждения*:

1) *V содержит собственные, инвариантные относительно φ подпространства*;

2) *в подходящей базе матрица $[\varphi]$ полураспавшаяся.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Пусть U – собственное инвариантное относительно φ подпространство, т. е. $0 < U < V$ и $U\varphi \subseteq U$. Тогда найдется такая база

$$e_1, e_2, \ldots, e_r, e_{r+1}, e_{r+2}, \ldots, e_n$$

пространства V, что векторы $e_{r+1}, e_{r+2}, \ldots, e_n$ образуют базу подпространства U. Легко видеть, что в этой базе

$$[\varphi] = \left(\begin{array}{c|c} * & * \\ \hline \mathbf{0} & * \end{array}\right)$$

т. е. является полураспавшейся.

2) $\Rightarrow$ 1). Пусть в некоторой базе $e_1, e_2, \ldots, e_n$ матрица $[\varphi]$ является полураспавшейся:

$$[\varphi] = \left(\begin{array}{c|c} * & * \\ \hline \mathbf{0} & * \end{array}\right)$$

Возьмем в качестве U подпространство, натянутое на векторы $e_{r+1}, e_{r+2}, \ldots, e_n$. Ясно, что U является собственным подпространством и $U\varphi \subseteq U$. Теорема доказана.

32.2. Индуцированные преобразования. Пусть U – инвариантное подпространство относительно преобразования φ векторного пространства V. Определим преобразование

$$\varphi_U : U \longrightarrow U$$

по правилу

$$\varphi_U : u \longmapsto u\varphi$$

и будем называть *индуцированным преобразование подпространства* U.

Рассмотрим фактор-пространство V/U пространства V по U и определим преобразование

$$\overline{\varphi}_U : V/U \longrightarrow V/U$$

по правилу

$$\overline{\varphi}_U : v + U \longmapsto v\varphi + U, \quad v \in V.$$

Покажем, что это определение не зависит от случайного выбора представителей смежных классов. Напомним, что $v_1 \equiv v_2 (\operatorname{mod} U)$ если $v_1 - v_2 \in U$. Следовательно, $v + U = v + u + U$ для некоторого вектора $u \in U$. Тогда $(v + u)\varphi + U = v\varphi + U$, т. е. определение $\overline{\varphi}_U$ не зависит от выбора представителей.

Преобразование $\overline{\varphi}_U$ называется преобразованием, *индуцированным* φ в фактор-пространстве. Легко проверить, что преобразования φ_U и $\overline{\varphi}_U$ являются линейными. Связь трех преобразований φ, φ_U и $\overline{\varphi}_U$ дает

Т е о р е м а 2. *В согласованных базах матрица преобразования φ имеет такой вид*

$$[\varphi] = \left(\begin{array}{c|c} [\overline{\varphi}_U] & * \\ \hline \mathbf{0} & [\varphi_U] \end{array}\right)$$

Д о к а з а т е л ь с т в о. Мы знаем, что справедлива формула

$$\dim V/U = \dim V - \dim U.$$

Укажем согласованные базы, в которых матрица имеет нужный вид. Пусть

$$e_1, e_2, \dots, e_r, e_{r+1}, e_{r+2}, \dots, e_n$$

– такая база векторного пространства V, что векторы $e_{r+1}, e_{r+2}, \dots, e_n$ образуют базу подпространства U. Как мы знаем, смежные классы

$$e_1 + U, e_2 + U, \dots, e_r + U$$

образуют базу фактор-пространства V/U. Действуя, на базисные векторы пространства V преобразованием φ, получим

$$\begin{cases} e_1\varphi & = & \alpha_{11}e_1 + \ldots + \alpha_{1r}e_r + \alpha_{1,r+1}e_{r+1} + \ldots + \alpha_{1n}e_n, \\ e_2\varphi & = & \alpha_{21}e_1 + \ldots + \alpha_{2r}e_r + \alpha_{2,r+1}e_{r+1} + \ldots + \alpha_{2n}e_n, \\ \ldots.. & & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_r\varphi & = & \alpha_{r1}e_1 + \ldots + \alpha_{rr}e_r + \alpha_{r,r+1}e_{r+1} + \ldots + \alpha_{rn}e_n, \\ e_{r+1}\varphi & = & 0 + \ldots + 0 + \alpha_{r+1,r+1}e_{r+1} + \ldots + \alpha_{r+1,n}e_n, \\ \ldots.. & & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n\varphi & = & 0 + \ldots + 0 + \alpha_{n,r+1}e_{r+1} + \ldots + \alpha_{nn}e_n. \end{cases}$$

По определению,

$$(e_i + U)\overline{\varphi}_U = e_i\varphi + U, \quad i = 1, 2, \dots, r,$$

а потому

$$(e_i + U)\varphi = \alpha_{i1}(e_1 + U) + \alpha_{i2}(e_2 + U) + \ldots \alpha_{ir}(e_r + U).$$

Таким образом,

$$\begin{array}{l} (e_1 + U)\overline{\varphi}_U = e_1\varphi + U = \alpha_{11}(e_1 + U) + \alpha_{12}(e_2 + U) + \ldots \alpha_{1r}(e_r + U), \\ (e_2 + U)\overline{\varphi}_U = e_2\varphi + U = \alpha_{21}(e_1 + U) + \alpha_{22}(e_2 + U) + \ldots \alpha_{2r}(e_r + U), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (e_r + U)\overline{\varphi}_U = e_r\varphi + U = \alpha_{r1}(e_1 + U) + \alpha_{r2}(e_2 + U) + \ldots \alpha_{rr}(e_r + U). \end{array}$$

Следовательно, матрица $[\varphi]$ действительно имеет указанный вид. Теорема доказана.

32.3. Одномерные инвариантные подпространства, собственные векторы и собственные значения. Предположим, что подпространство U, инвариантное относительно φ, имеет размерность 1. Тогда U порождено любым ненулевым вектором $u \in U$ и

$$u\varphi = \alpha u, \quad \alpha \in P.$$

Поиск одномерных инвариантных подпространств сводится к отысканию таких u и α. При этом u называется *собственным вектором преобразования* φ, а α – его *собственным значением.*

Пусть $e_1, e_2, \dots, e_n$ – база V, $[u]$ – координаты вектора u, а $[\varphi]$ – матрица преобразования φ в этой базе. Тогда

$$[u\varphi] = [\alpha u],$$

или

$$[u][\varphi] = \alpha[u].$$

Вынося координатную строку $[u]$, получим

$$[u]([\varphi] - \alpha E) = 0.$$

Чтобы существовал ненулевой вектор u, удовлетворяющий этому равенству необходимо и достаточно, чтобы определитель

$$\det([\varphi] - \alpha E)$$

был равен нулю. Рассмотрим определитель $\chi(\lambda) = \det([\varphi] - \lambda E)$, который является многочленом степени n относительно λ и собственные значения должны быть его корнями. Этот многочлен называется *характеристическим многочленом преобразования* φ. Справедлива

Т е о р е м а 3. а) *Характеристический многочлен не зависит от выбора базы, по которой он построен.* б) *Всякое собственное значение преобразования φ является корнем характеристического многочлена.* в) *Всякий корень характеристического многочлена, лежащий в основном поле, является собственным значением.*

Д о к а з а т е л ь с т в о. а) Пусть $e'_1, e'_2, \dots, e'_n$ – другая база V. Тогда $e' = Te$, где T – матрица перехода от базы e к базе e'. Обозначим $[\varphi]'$ матрицу преобразования φ в базе e'. Как мы знаем,

$$[\varphi]' = T[\varphi]T^{-1}.$$

Поэтому

$$\det([\varphi]' - \lambda E) = \det(T[\varphi]T^{-1} - \lambda E) = \det(T\{[\varphi] - \lambda E\}T^{-1}) =$$

$$= \det T \cdot \det([\varphi] - \lambda E) \cdot \det T^{-1} = \det([\varphi] - \lambda E).$$

Следовательно, характеристический многочлен зависит от преобразования φ, а не от вида матрицы, соответствующей φ.

б) Это доказательство проведено выше. Каждое собственное значение обязано быть корнем характеристического многочлена.

в) Пусть $\chi(\alpha) = 0$, т. е. α – корень характеристического многочлена, лежащий в P. Надо доказать, что α – собственное значение. Возьмем какую-нибудь базу $e_1, e_2, \dots, e_n$ пространства V. Рассмотрим систему линейных уравнений

$$X([\varphi] - \alpha E) = 0, \quad X = (x_1, x_2, \dots, x_n)$$

так как ранг этой системы меньше числа неизвестных, существует ненулевое решение $X^0 = (x_1^0, x_2^0, \ldots, x_n^0)$. Рассмотрим вектор

$$u = x_1^0 e_1 + x_2^0 e_2 + \ldots + x_n^0 e_n.$$

Для него

$$[u]([\varphi] - \alpha E) = 0,$$

а потому $u(\varphi - \alpha\varepsilon) = 0$ или $u\varphi = \alpha u$, т. е нашли собственный вектор $u \neq 0$, а потому α – собственное значение. Теорема доказана.

Упражнение. Где нужно, чтобы α лежало в основном поле?

32.4. Теорема Гамильтона-Кэли.

Теорема 4. *Всякое линейное преобразование является корнем своего характеристического многочлена, т. е. если $\chi(\lambda)$ – характеристический многочлен преобразования φ, то $\chi(\varphi) = 0$.*

Доказательство. Достаточно доказать эту теорему для матриц, так как алгебра линейных преобразований изоморфна алгебре матриц. Пусть $A \in M_n(P)$ и

$$\chi(\lambda) = \det(A - \lambda E)$$

– ее характеристический многочлен. Надо доказать, что $\chi(A) = 0$. Пусть

$$\chi(\lambda) = a_0 + a_1\lambda + \ldots + a_n\lambda^n.$$

Символом $B(\lambda)$ обозначим матрицу, присоединенную к матрице $A - \lambda E$ (напомним, что присоединенной к матрице $C = (c_{ij})$ называется матрица $(C_{ij})'$, составленная из алгебраических дополнений C_{ij} к элементам c_{ij} и транспонированная). Тогда

$$(A - \lambda E)B(\lambda) = \chi(\lambda)E.$$

Пример. Если дана матрица

$$\begin{pmatrix} 7 - 3\lambda^2 + \lambda^3 & 1 - \lambda \\ -8 + \lambda & -3 + \lambda^2 \end{pmatrix},$$

то ее можно представить в таком виде:

$$\begin{pmatrix} 7 - 3\lambda^2 + \lambda^3 & 1 - \lambda \\ -8 + \lambda & \lambda^2 - 3 \end{pmatrix} = \begin{pmatrix} 7 & 1 \\ -8 & -3 \end{pmatrix} + \lambda \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \lambda^2 \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix} + \lambda^3 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Раскладывая матрицу $B(\lambda)$ по степеням λ, получим

$$B(\lambda) = B_0 + B_1\lambda + B_2\lambda^2 + \ldots + B_{n-1}\lambda^{n-1}, \quad B_k \in M_n(K).$$

Из равенства $(A-\lambda E)B(\lambda) = \chi(\lambda)E$, приравнивая матрицы при одинаковых степенях λ, получаем следующую систему равенств

$$\begin{array}{ll} AB_0 & = a_0E, \\ AB_1 - B_0 & = a_1E, \\ AB_2 - B_1 & = a_2E, \\ \quad\vdots & \quad\vdots \\ AB_{n-1} - B_{n-2} & = a_{n-1}E, \\ \qquad\quad -B_{n-1} & = a_nE. \end{array}$$

Умножим первое из этих равенств на E, второе – на A, третье на A^2 и т. д. Наконец, последнее равенство умножим на A^n. Складывая полученные равенства, получим в левой часть 0, а правой

$$a_0E + a_1A + a_2A^2 + \ldots + a_nA^n = \chi(A).$$

Следовательно,

$$0 = \chi(A).$$

Теорема доказана.

Упражнение. Можно предложить такое доказательство теоремы Гамильтона–Кэли. В равенство

$$\chi(\lambda) = \det(A - \lambda E)$$

вместо λ подставим матрицу A, получим

$$\chi(A) = \det(A - AE) = 0.$$

Найдите ошибку в этом рассуждении.

§ 33. Нильпотентные и полупростые преобразования

33.1. Нильпотентные преобразования. Следующая теорема является определением и устанавливает некоторые свойства нильпотентных преобразований.

Теорема 1. *Для линейного преобразования φ следующие условия равносильны*:

1) $\varphi^N = 0$ *при подходящем* $N = N(\varphi)$;

2) *в подходящей базе*

$$[\varphi] = \begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}.$$

При любом из этих условий преобразование φ называется *нильпотентным.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Индукция по $n = \dim V$. При $n = 1$ очевидно, что $\varphi = 0$. Пусть далее $n > 1$. Если $\varphi = 0$, то утверждение очевидно. Пусть $\varphi \neq 0$, выберем N такое, что $\varphi^N = 0$, но $\varphi^{N-1} \neq 0$. Значит, существует вектор $u \in V$ такой, что $u\varphi^{N-1} \neq 0$. Обозначим $v = u\varphi^{N-1}$. Понятно, что $v \neq 0$ и $v\varphi = 0$. Пусть e_1, e_2, $\ldots$, e_{n-1}, v – база V, W – подпространство, порожденное вектором v. Тогда в этой базе

$$[\varphi] = \left(\begin{array}{c|c} & * \\ [\overline{\varphi}_W] & \vdots \\ & * \\ \hline 0 \ldots 0 & 0 \end{array}\right).$$

По определению индуцированного преобразования:

$$\overline{\varphi}_W : a + W \longmapsto a\varphi + W.$$

Применить это преобразование N раз это все равно, что к a применить преобразование φ, т. е.

$$(a + W)\overline{\varphi}_W^N = a\varphi^N + W = W.$$

Следовательно, $\overline{\varphi}_W$ удовлетворяет условию 1), т. е. любой класс переходит под действием $\overline{\varphi}_W^N$ в 0.

Учитывая, что $\dim V/W = n - 1$, видим, что в подходящей базе индуцированное преобразование имеет матрицу

$$[\overline{\varphi}_W]' = \begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}.$$

Возьмем в смежных классах

$$e'_1 + W, e'_2 + W, \ldots, e'_{n-1} + W$$

этой базы по представителю: e'_1, e'_2, $\ldots$, e'_{n-1}. Тогда в базе

$$e'_1, e'_2, \ldots, e'_{n-1}, v$$

имеем матрицу

$$[\varphi]' = \left(\begin{array}{ccc|c} 0 & & * & * \\ & \ddots & & \vdots \\ \mathbf{0} & & 0 & * \\ \hline 0 & \ldots & 0 & 0 \end{array}\right).$$

2) $\Rightarrow$ 1). Пусть в некоторой базе матрица преобразования φ имеет вид

$$A = \begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix}.$$

Надо доказать, что в некоторой степени эта матрица равна нулевой матрице. Покажем, что

$$A^i = \begin{pmatrix} 0 & \dots & 0 & & * \\ & \ddots & & \ddots & \\ & & \ddots & & 0 \\ & & & \ddots & \vdots \\ \mathbf{0} & & & & 0 \end{pmatrix}.$$

т. е. имеет i нулевых диагоналей, начиная с главной. Индукция по i. Для $i = 1$ утверждение справедливо, так как по условию, матрица A имеет нулевую главную диагональ.

Предположим, что равенство справедливо для $i-1$ и установим его для i. Обозначим символом E_{rs} *матричную единицу*, т. е матрицу из $M_n(P)$, у которой на месте (r, s) стоит 1, а на всех остальных местах – нули. Тогда матрица A имеет вид

$$A = \sum_{q-p \geq 1} a_{pq} E_{pq}.$$

Заметим, что для матричных единиц справедливо следующее правило умножения

$$E_{kl} \cdot E_{pq} = \begin{cases} E_{kq} & \text{при} \quad l = p, \\ 0 & \text{при} \quad l \neq p. \end{cases}$$

Воспользовавшись предположением индукции, получим

$$A^i = A^{i-1} \cdot A = \sum_{l-k \geq i-1} b_{kl} E_{kl} \cdot \sum_{q-p \geq 1} a_{pq} E_{pq} = \sum_{q-k \geq i} b_{kp}\, a_{pq} E_{kq}.$$

Эта матрица имеет i нулевых диагоналей.

Учитывая установленное, равенство, если положить $N \geq n$, то $A^N = 0$. Теорема доказана.

Непосредственно из доказательства получается такое

С л е д с т в и е. *В качестве $N(\varphi)$ можно взять $n = \dim V$.*

33.2. Полупростые преобразования.

Т е о р е м а 2. *Для линейного преобразования φ следующие условия равносильны*:

1) *все пространство V распадается в прямую сумму $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$ инвариантных и неприводимых относительно φ подпространств*;

2) *в подходящей базе*

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

где никакая матрица A_i не подобна полураспавшейся матрице, т. е. ни в какой базе не станет полураспавшейся. При любом из этих условий преобразование φ *называется полупростым.*

Доказательство. 1) $\Rightarrow$ 2). Пусть $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$ – прямая сумма инвариантных и неприводимых относительно φ подпространств. Берем в каждом V_i по базе и объединяем их. В полученной базе

$$e_1, e_2, \ldots, e_n$$

матрица преобразования φ имеет вид

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix}.$$

При этом $V_i\varphi \subseteq V_i$. Нужно понять, что каждая клетка A_i не подобна полураспавшейся матрице, а это равносильно неприводимости подпространства относительно φ.

2) $\Rightarrow$ 1). Пусть в базе

$$e_1, e_2, \ldots, e_n$$

матрица преобразования

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

где каждая клетка A_i не подобна полураспавшейся матрице. Разобьем базу пространства V на отрезки, стоящие против соответствующих клеток. Пусть V_i – подпространство, натянутое на i-й отрезок. Ясно, что $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$, $V_i\varphi \subseteq V_i$, и нужно понять, что каждое подпространство V_i неприводимо. Если бы оно было приводимым, то нашлось бы собственное подпространство U такое, что

$$0 < U < V_i, \quad U\varphi \subseteq U,$$

и в этом случае матрица A_i была бы подобна полураспавшейся. Противоречие. Теорема доказана.

С л е д с т в и е. *Если φ – нильпотентно и полупросто, то $\varphi = 0$.*

Д о к а з а т е л ь с т в о. Если φ полупросто, то, по теореме 2, в некоторой базе

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

а так как φ – нильпотентно, то каждая матрица A_i нильпотентна, а потому, по теореме 1, подобна матрице такого вида

$$\begin{pmatrix} 0 & & * \\ & \ddots & \\ \mathbf{0} & & 0 \end{pmatrix},$$

но из полупростоты следует, что A_i не подобна полураспавшейся. Следовательно, матрица A_i является квадратной степени 1 нулевой матрицей, но это и означает, что φ – нулевое преобразование.

33.3. Полупростота над полем, содержащим характеристические корни преобразования. В дальнейшем будем называть корни характеристического многочлена *характеристическими корнями.* Справедлива

Т е о р е м а 3. *Пусть φ – линейное преобразование векторного пространства V над полем P, причем, P содержит все характеристические корни преобразования φ. Тогда равносильны*:

1) *φ – полупросто*;

2) *в подходящей базе $[\varphi]$ – диагональна*;

3) *существует многочлен $f \in P[\lambda]$, не имеющий кратных корней и такой, что $f(\varphi) = 0$.*

Д о к а з а т е л ь с т в о. 1) $\Rightarrow$ 2). Пусть φ – полупросто, т. е. в подходящей базе

$$[\varphi] = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_s \end{pmatrix},$$

где никакая матрица A_i не подобна полураспавшейся матрице. Докажем, что все эти клетки одномерные. Допустим, что какая-то клетка имеет размерность ≥ 2, т. е. $\dim V_i \geq 2$ для соответствующего подпространства V_i. Выберем в V_i вектор v, являющийся собственным относительно φ, т. е $v\varphi = \alpha v$. Очевидно, характеристический корень преобразования $\varphi|_{V_i}$ является характеристическим корнем преобразования φ, а потому α лежит в P. Пусть

$$e_1, e_2, \ldots, e_m, v$$

– база подпространства V_i. В этой базе

$$[\varphi|_{V_i}] = \left(\begin{array}{ccc|c} & & & * \\ & * & & \vdots \\ & & & * \\ \hline 0 & \ldots & 0 & \alpha \end{array}\right),$$

т. е. матрица A_i подобна полураспавшейся. Противоречие.

У п р а ж н е н и е. Привести пример полупростой матрицы не представимой в диагональном виде.

2) $\Rightarrow$ 3). Пусть в некоторой базе матрица $[\varphi]$ – диагональна и многочлен $F(\lambda) \in P[\lambda]$ аннулирует преобразование φ, т. е. $F(\varphi) = 0$. В частности, в качестве $F(\lambda)$ можно взять характеристический многочлен преобразования φ. Разложим его на линейные множители:

$$F(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \ldots (\lambda - \lambda_s)^{r_s},$$

где все корни λ_j различны. Возьмем многочлен

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_s)$$

и число

$$r = \max\{r_1, r_2, \ldots, r_s\}.$$

Тогда $f(\lambda)^r = F(\lambda)d(\lambda)$ для некоторого многочлена $d(\lambda)$ и матрица $f([\varphi])$ диагональна. Рассмотрим матрицу $[f(\varphi)] = f([\varphi])$. Если ее возвести в степень r, то получим нулевую матрицу, но это значит, что на диагонали стояли нули, т. е. $f(\varphi) = 0$.

3) $\Rightarrow$ 1). Пусть $f(\varphi) = 0$ и

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_s),$$

где все λ_j различны. Надо доказать, что φ полупросто. Рассмотрим многочлены

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

которые в совокупности взаимно просты. Поэтому существуют многочлены $g_i \in P[\lambda]$ такие, что

$$g_1 f_1 + g_2 f_2 + \ldots + g_s f_s = 1.$$

Обозначим

$$V_i = V f_i(\varphi), \quad i = 1, 2, \ldots, s,$$

и докажем, что V является прямой суммой подпространств V_i.

а) Докажем вначале, что $V = V_1 + V_2 + \ldots + V_s$. Действительно, пусть $v \in V$, тогда из равенства

$$g_1(\varphi) f_1(\varphi) + g_2(\varphi) f_2(\varphi) + \ldots + g_s(\varphi) f_s(\varphi) = \varepsilon$$

получим

$$v = v\varepsilon = v g_1(\varphi) f_1(\varphi) + v g_2(\varphi) f_2(\varphi) + \ldots + v g_s(\varphi) f_s(\varphi).$$

Заметим, что первое слагаемое лежит в V_1, второе – в V_2 и т. д., и, наконец, последнее лежит в V_s. Следовательно, $V = V_1 + V_2 + \ldots + V_s$.

б) Докажем, что сумма $V = V_1 + V_2 + \ldots + V_s$ является прямой. Для этого достаточно доказать, что нулевой вектор имеет единственное разложение. Пусть

$$0 = u_1 f_1(\varphi) + u_2 f_2(\varphi) + \ldots + u_s f_s(\varphi).$$

Покажем, что все $u_i f_i(\varphi) = 0$. Для этого к обеим частям равенства применим преобразование $f_i(\varphi)$. Заметим, что если $k \neq i$, то

$$u_k f_k(\varphi) f_i(\varphi) = u_k f(\varphi) h(\varphi) = 0 \ \text{ для некоторых } \ h_{ki}(\lambda) \in P[\lambda],$$

и мы имеем равенство

$$0 = u_i f_i^2(\varphi).$$

Так как $f_i(\lambda)$ и $\lambda - \lambda_i$ взаимно просты, то существуют многочлены p_i и q_i из $P[\lambda]$ такие, что

$$f_i(\lambda) p_i(\lambda) + (\lambda - \lambda_i) q_i(\lambda) = 1.$$

Отсюда,

$$f_i(\varphi) p_i(\varphi) + (\varphi - \lambda_i \varepsilon) q_i(\varphi) = \varepsilon,$$

и

$$u_i f_i(\varphi) = u_i f_i(\varphi)\varepsilon = u_i f_i^2(\varphi) p_i(\varphi) + u_i f(\varphi) q_i(\varphi) = 0,$$

так как $u_i f_i(\varphi)^2 = 0$ и $u_i f(\varphi) = 0$, т. е. эта сумма $V = V_1 + V_2 + \ldots + V_s$ является прямой.

в) Чтобы показать, что каждое подпространство V_i инвариантно относительно φ, заметим, что $V_i(\varphi - \lambda_i \varepsilon) = 0$. Действительно, произвольный вектор из V_i имеет вид $u f_i(\varphi)$ для некоторого $u \in V$. Действуя преобразованием $\varphi - \lambda_i \varepsilon$, получим

$$(u f_i(\varphi))(\varphi - \lambda_i \varepsilon) = u f(\varphi) = u0 = 0.$$

Следовательно, если $u_i \in V_i$, то $u_i \varphi = \lambda_i u_i$, а потому V_i инвариантно относительно φ.

Ввиду того, что преобразование φ_{V_i} является растяжением в λ_i раз, то выбирая базу в каждом V_i и объединяя их, получим базу пространства V, в которой матрица преобразования φ имеет вид

$$[\varphi] = \left(\begin{array}{ccc|c|ccc} \lambda_1 & & \mathbf{0} & & & & \\ & \ddots & & & & \mathbf{0} & \\ \mathbf{0} & & \lambda_1 & & & & \\ \hline & \mathbf{0} & & \ddots & & \mathbf{0} & \\ \hline & & & & \lambda_s & & \mathbf{0} \\ & \mathbf{0} & & & & \ddots & \\ & & & & \mathbf{0} & & \lambda_s \end{array}\right)$$

т. е. является диагональной матрицей, а потому φ полупросто. Теорема доказана.

Заметим, что диагональная матрица всегда полупроста, но обратное верно не всегда.

33.4. Разложение преобразования на полупростую и нильпотентную компоненты.

Т е о р е м а 4. *Пусть φ – линейное преобразование вектороного пространства V над полем P и P содержит все характеристические корни φ. Тогда*

а) *существует разложение*

$$\varphi = \psi + \theta,$$

где ψ – полупростое преобразование, а θ – нильпотентное, и при этом $\psi\theta = \theta\psi$;

б) *такое разложение единственно*;

в) *ψ и θ являются многочленами от φ над полем P.*

При этом ψ называется *полупростой компонентой*, а θ – *нильпотентной.*

Д о к а з а т е л ь с т в о. Докажем существование. Покажем, что найдется многочлен $f(\lambda)$ из $P[\lambda]$ без кратных корней и $m \in \mathbb{N}$ такие, что $f^m(\varphi) = 0$. Пусть $F(\varphi) = 0$ для некоторого многочлена $F(\lambda) \in P[\lambda]$ (по теореме Гамильтона-Кэли такой многочлен существует). Разложим его на линейные множители:

$$F(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \ldots (\lambda - \lambda_s)^{r_s}, \quad \lambda_i \neq \lambda_j \text{ при } i \neq j.$$

Возьмем

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_s).$$

Тогда $f^m(\varphi) = 0$ при $m = \max\{r_1, r_2, \ldots, r_s\}$. Так как f без кратных корней, то $(f, f') = 1$, а потому существуют $u, v \in P[\lambda]$ такие, что

$$f\,u + f'\,v = 1.$$

Рассмотрим рекуррентную схему, позволяющую по каждому многочлену $g(\lambda)$ из $P[\lambda]$ построить последовательность многочленов $g^{(0)}(\lambda), g^{(1)}(\lambda), \ldots$ из $P[\lambda]$:

$$\begin{cases} g^{(0)}(\lambda) = g(\lambda), \\ g^{(i+1)}(\lambda) = g^{(i)}(\lambda - f(\lambda)\, v(\lambda)), \quad i = 0, 1, \ldots, \end{cases}$$

и применим ее к многочлену $e(\lambda) = \lambda$. Получим последовательность

$$e^{(0)}(\lambda), e^{(1)}(\lambda), \ldots, e^{(k)}(\lambda),$$

где k – наименьшее натуральное число, удовлетворяющее неравенству $2^k \geq m$. Положим $\psi = e^{(k)}(\varphi)$ и $\theta = \varphi - \psi$. Покажем, что эти преобразования и будут искомыми.

По теореме 3 преобразование полупросто, если оно является корнем многочлена без кратных корней. Рассмотрим

$$f(\psi) = f(e^{(k)}(\varphi)).$$

Если мы докажем что $f(e^{(k)}(\lambda)) = f^{(k)}(\lambda))$ и то, что $f^{(k)}(\lambda)$ делится на $f^{2^k}(\lambda)$, то, учитывая, что $f^{2^k}(\varphi) = 0$, получим, что и $f^{(k)}(\varphi) = 0$. Таким образом, полупростота преобразования ψ будет доказана.

Чтобы доказать, что θ нильпотентно, надо доказать, что $\lambda - f^{(k)}(\lambda)$ делится на $f(\lambda)$.

Л е м м а 1. *Справедливо равенство*

$$f(e^{(i)}(\lambda)) = f^{(i)}(\lambda).$$

Д о к а з а т е л ь с т в о проведем индукцией по i. При $i = 0$ имеем

$$f(e^{(0)}(\lambda)) = f(\lambda) = f^{(0)}(\lambda).$$

Предположим, что утверждение справедливо для i и докажем его для $i+1$. Имеем

$$f(e^{(i+1)}(\lambda)) = f(e^{(i)}(\lambda - f(\lambda)\, v(\lambda))) = f^{(i)}(\lambda - f(\lambda)\, v(\lambda)) = f^{(i+1)}(\lambda).$$

Лемма доказана.

Л е м м а 2. $f^{(i)}(\lambda)$ *кратно* $f^{2^i}(\lambda)$.

Д о к а з а т е л ь с т в о проведем индукцией по i. При $i = 0$ видим, что $f^{(0)}(\lambda)$ кратно $f^{2^0}(\lambda) = f(\lambda)$. При $i = 1$ имеем

$$f^{(1)}(\lambda) = f(\lambda - f(\lambda)\, v(\lambda)).$$

Воспользовавшись формулой Тейлора:

$$f(\lambda - a) = f(\lambda) - \frac{f'(\lambda)}{1!} a + \frac{f''(\lambda)}{2!} a^2 - \ldots,$$

получим

$$f(\lambda - f(\lambda)\,v(\lambda)) = f(\lambda) - \frac{f'(\lambda)}{1!} f(\lambda)\,v(\lambda) + \frac{f''(\lambda)}{2!} f^2(\lambda) v^2(\lambda) - \ldots =$$

$$= f(\lambda)\,[1 - f'(\lambda)\,v(\lambda) + f(\lambda)\,[\ldots]] = f(\lambda)\,[f(\lambda) u(\lambda) + f(\lambda)\,[\ldots]] = f^2(\lambda)\,[\ldots]\,.$$

Предположим, что формула верна для i и докажем ее для $i+1$. Имеем

$$f^{(i+1)}(\lambda) = f^{(i)}(\lambda - f(\lambda)\,v(\lambda)) = f^{2^i}(\lambda - f(\lambda)\,v(\lambda)) \cdot [\ldots] = [f^{(1)}(\lambda)]^{2^i} [\ldots] = [f(\lambda)]^{2^{i+1}} [\ldots],$$

где мы использовали установленный выше факт, что $f^{(1)}(\lambda)$ делится на $f^2(\lambda)$. Лемма доказана.

Л е м м а 3. *Многочлен $\lambda - e^{(i)}(\lambda)$ делится на многочлен $f(\lambda)$.*

Д о к а з а т е л ь с т в о проведем индукцией по i. При $i = 0$ видим, что $\lambda - e^{(0)}(\lambda) = 0$, а 0 кратен $f(\lambda)$, так как $0 = f(\lambda) \cdot 0$.

Предположим, что лемма справедлива для i. Рассмотрим

$$\lambda - e^{(i+1)}(\lambda) = \lambda - e^{(i)}(\lambda - f(\lambda)\,v(\lambda)) - f(\lambda)\,v(\lambda) + f(\lambda)\,v(\lambda).$$

так как $\mu - e^{(i)}(\mu)$ кратно $f(\mu)$, где $\mu = \lambda - f(\lambda)\,v(\lambda)$ то

$$\lambda - e^{(i+1)}(\lambda) = f(\lambda - f(\lambda)\,v(\lambda)) \cdot [\ldots] + f(\lambda)\,v(\lambda)$$

В предыдущей лемме было установлено, что $f(\lambda - f(\lambda)\,v(\lambda))$ делится на $f(\lambda)$. Следовательно, $\lambda - e^{(i+1)}(\lambda)$ делится на $f(\lambda)$. Лемма доказана.

Мы установили, что преобразование ψ полупросто, а θ – нильпотентно. Как следует из построения, ψ и θ являются многочленами от φ, а потому $\psi\theta = \theta\psi$. Следовательно, утверждения а) и в) теоремы установлены.

Установим пункт б). Допустим, что найдется еще одна пара ψ', θ' преобразований таких, что $\varphi = \psi' + \theta'$ и при этом ψ' полупросто, а θ' нильпотентно. Рассматривая разность двух равенств

$$\varphi = \psi + \theta, \quad \varphi = \psi' + \theta',$$

получим

$$\psi - \psi' = \theta' - \theta.$$

У п р а ж н е н и е. Привести пример двух полупростых преобразований, разность которых не является полупростым. Привести пример двух нильпотентных преобразований, разность которых не является нильпотентным.

Л е м м а 4. *Преобразования ψ и ψ' перестановочны, преобразования θ и θ' перестановочны.*

Д о к а з а т е л ь с т в о. Очевидно, что ψ' перестановочно с ψ'. По условию ψ' перестановочно с θ'. Следовательно, ψ' перестановочно с $\psi' + \theta' = \varphi$, т. е. ψ' перестановочно с φ, а потому перестановочно с любым многочленом от φ, но среди таких многочленов есть многочлен, равный ψ. Следовательно, ψ' перестановочно с ψ. Аналогично проверяется, что θ' перестановочно с θ.

Л е м м а 5. *Если преобразования ψ_1 и ψ_2 полупросты и перестановочны, то $\psi_1 \pm \psi_2$ – полупросто.*

Д о к а з а т е л ь с т в о. По теореме 3 в некоторой базе ψ_1 диагональна и ψ_2 диагональна в некоторой базе, но эти базы различны. Возьмем базу в которой ψ_1 диагональна

$$[\psi_1] = \left(\begin{array}{ccc|ccc|c|ccc}
\alpha & & & & & & & & & \\
& \ddots & & & & & & & & \\
& & \alpha & & & & & & & \\
\hline
& & & \beta & & & & & & \\
& & & & \ddots & & & & & \\
& & & & & \beta & & & & \\
\hline
& & & & & & \ddots & & & \\
\hline
& & & & & & & \omega & & \\
& & & & & & & & \ddots & \\
& & & & & & & & & \omega
\end{array}\right),$$

и разобьем матрицу на клетки так, что в каждой клетке стоят одинаковые диагональные элементы. Пусть в этой же базе

$$[\psi_2] = \left(\begin{array}{c|c|c|c}
X_{11} & X_{12} & \cdots & X_{1s} \\
\hline
X_{21} & X_{22} & \cdots & X_{2s} \\
\hline
\cdots & \cdots & \cdots & \cdots \\
\hline
X_{s1} & X_{s2} & \cdots & X_{ss}
\end{array}\right),$$

где матрица разбита на клетки тех же размеров, что и $[\psi_1]$. Так как преобразования перестановочны, то

$$[\psi_1][\psi_2] = \left(\begin{array}{c|c|c|c}
\alpha X_{11} & \alpha X_{12} & \cdots & \alpha X_{1s} \\
\hline
\beta X_{21} & \beta X_{22} & \cdots & \beta X_{2s} \\
\hline
\cdots & \cdots & \cdots & \cdots \\
\hline
\omega X_{s1} & \omega X_{s2} & \cdots & \omega X_{ss}
\end{array}\right), \quad [\varphi_2][\varphi_1] = \left(\begin{array}{c|c|c|c}
X_{11}\alpha & X_{12}\beta & \cdots & X_{1s}\omega \\
\hline
X_{21}\alpha & X_{22}\beta & \cdots & X_{2s}\omega \\
\hline
\cdots & \cdots & \cdots & \cdots \\
\hline
X_{s1}\alpha & X_{s2}\beta & \cdots & X_{ss}\omega
\end{array}\right).$$

Откуда имеем систему:

$$\begin{cases} \alpha X_{11} = X_{11}\alpha, & \alpha X_{12} = X_{12}\beta, \quad \cdots \quad \alpha X_{1s} = X_{1s}\omega, \\ \beta X_{21} = X_{21}\alpha, & \beta X_{22} = X_{22}\beta, \quad \cdots \quad \beta X_{2s} = X_{2s}\omega, \\ \qquad \cdots & \qquad \cdots \qquad\quad \cdots \qquad\quad \cdots \\ \omega X_{s1} = X_{s1}\alpha, & \omega X_{32} = X_{32}\beta, \quad \cdots \quad \omega X_{ss} = X_{ss}\omega. \end{cases}$$

Из этой системы видим, что все клетки вне главной диагонали нулевые, т. е.

$$[\psi_2] = \left(\begin{array}{c|c|c|c} X_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & X_{22} & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & X_{ss} \end{array}\right)$$

и каждая из этих клеток полупроста. Если матрица удовлетворяет многочлену без кратных корней, то и каждая клетка удовлетворяет этому многочлену. Поэтому найдется база в которой каждая клетка матрицы $[\psi_2]'$ имеет диагональный вид:

$$[\psi_2]' = \left(\begin{array}{ccc|ccc|c|ccc} * & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & * & & & & & & & \\ \hline & & & * & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & * & & & & \\ \hline & & & & & & \ddots & & & \\ \hline & & & & & & & * & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & * \end{array}\right),$$

но первая матрица при этом не изменится так как каждой клетке соответствует подпространство, состоящее из собственных векторов преобразования $[\psi_1]$, а потому

$$[\psi_1]' = \left(\begin{array}{ccc|ccc|c|ccc}
\alpha & & & & & & & & & \\
& \ddots & & & & & & & & \\
& & \alpha & & & & & & & \\
\hline
& & & \beta & & & & & & \\
& & & & \ddots & & & & & \\
& & & & & \beta & & & & \\
\hline
& & & & & & \ddots & & & \\
\hline
& & & & & & & \omega & & \\
& & & & & & & & \ddots & \\
& & & & & & & & & \omega
\end{array}\right).$$

Следовательно, $[\psi_1 \pm \psi_2]$ диагональна. Лемма доказана.

Л е м м а 6. *Если преобразования θ_1 и θ_2 нильпотентны и перестановочны, то $\theta_1 \pm \theta_2$ – нильпотентно.*

Д о к а з а т е л ь с т в о. Пусть $\theta_1^{n_1} = 0$, $\theta_2^{n_2} = 0$, тогда, воспользовавшись перестановочностью θ_1 и θ_2, получим

$$(\theta_1 + \theta_2)^N = \sum_{k=0}^{N} C_N^k \theta_1^k \theta_2^{N-k} = 0 \text{ при } N = n_1 + n_2.$$

Теорема доказана.

33.5. Полупростота над полем действительных чисел. Над полем комплексных чисел полупростота равносильна диагонализируемости матрицы. Над полем действительных чисел это уже неверно.

Т е о р е м а 5. *Над полем действительных чисел линейное преобразование φ полупросто тогда и только тогда, когда в некоторой базе его матрица клеточно-диагональна и каждая диагональная клетка либо одномерна, либо двумерна и не подобна полураспавшейся.*

Д о к а з а т е л ь с т в о. По теореме о полупростом преобразовании, в некоторой базе его матрица клеточно-диагональна. Покажем, что каждая клетка имеет порядок ≤ 2. Для этого достаточно доказать, что всякая $A \in M_n(\mathbb{R})$ при $n \geq 3$ подобна полураспавшейся. Для этого надо найти инвариантные подпространства.

Возьмем характеристический многочлен матрицы A и найдем его характеристический корень. Если корень вещественный, то собственный вектор порождает инвариантное подпространство, а потому матрица A подобна полураспавшейся. Пусть характеристический корень комплексный: $\alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$. Рассмотрим $\mathbb{C}^n$ – пространство строк над полем комплексных чисел. Для него характеристический корень есть собственное значение над $\mathbb{C}$. Следовательно, существует w – собственный вектор, отвечающий собственному значению $\alpha + i\beta$ т. е.

$$wA = (\alpha + i\beta)w, \quad w \neq 0.$$

У вектора w можно отделить действительную и мнимую части:

$$w = u + iv, \quad u, v \in \mathbb{R}^n,$$

тогда

$$(u + iv)A = (\alpha + i\beta)(u + iv).$$

Рассмотрим подпространство U, натянутое на векторы u и v. Учитывая равенства

$$\begin{cases} uA = \alpha u - \beta v, \\ vA = \beta u + \alpha v, \end{cases}$$

заключаем, что $UA \subseteq U$, т. е. нашли подпространство U, которое инвариантно относительно A и U – не более чем двумерно. Следовательно, исходная матрица подобна полураспавшейся. Теорема доказана.

Г л а в а 5

ЖОРДАНОВА ФОРМА МАТРИЦЫ

§ 34. Теорема Жордана

34.1. Задача о подобии матриц. Как мы знаем, матрицы одного преобразования в разных базах связаны равенством

$$[\varphi]' = T[\varphi]T^{-1},$$

где T – невырожденная матрица. Возникает задача: распознать, когда две матрицы являются матрицами одного преобразования, записанного в разных базах.

О п р е д е л е н и е. Говорим, что матрица $A \in M_n(P)$ *подобна* матрице $B \in M_n(P)$, если существует матрица $X \in GL_n(P)$ такая, что $A = X^{-1}BX$.

Заметим вначале, что отношение подобия является отношением эквивалентности.

1) A подобна A, что следует из равенства $A = E^{-1}AE$.

2) Если A подобна B, то B подобна A. Действительно, если $A = X^{-1}BX$, то $B = XAX^{-1}$.

3) Если A подобна B, а B подобна C, то A подобна C. Действительно, если $A = X^{-1}BX$ и $B = Y^{-1}CY$, то $A = (XY)^{-1}C(YX)$.

Следовательно, все множество матриц $M_n(P)$ распадается на смежные классы относительно отношения подобия. Мы хотим научиться выбирать в каждом смежном классе по представителю, а все другие матрицы из этого класса приводить к выбранному представителю.

О п р е д е л е н и е. 1) *Жордановой клеткой* степени r называется матрица

следующего вида
$$\begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix} \in M_r(P).$$

2) *Жорданова матрица* – матрица, которая распадается на жордановы клетки:
$$\left(\begin{array}{c|c|c|c} J_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & J_2 & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & J_s \end{array}\right).$$
где J_i – жорданова клетка.

П р и м е р. При $n = 4$ имеем следующие жордановы матрицы:
$$\begin{pmatrix} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 1 \\ 0 & 0 & 0 & \alpha \end{pmatrix}, \quad \left(\begin{array}{ccc|c} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & \alpha & 0 \\ \hline 0 & 0 & 0 & \beta \end{array}\right), \quad \left(\begin{array}{cc|cc} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ \hline 0 & 0 & \beta & 1 \\ 0 & 0 & 0 & \beta \end{array}\right),$$
$$\left(\begin{array}{cc|c|c} \alpha & 1 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ \hline 0 & 0 & \beta & 0 \\ \hline 0 & 0 & 0 & \gamma \end{array}\right), \quad \left(\begin{array}{c|c|c|c} \alpha & 0 & 0 & 0 \\ \hline 0 & \beta & 0 & 0 \\ \hline 0 & 0 & \gamma & 0 \\ \hline 0 & 0 & 0 & \delta \end{array}\right).$$

В каждом смежном классе существует несколько жордановых матриц, но они отличаются лишь порядком клеток.

Целью настоящего параграфа является доказательство следующей теоремы.

Т е о р е м а Ж о р д а н а. *Всякая квадратная матрица над полем, содержащим все ее характеристические корни, подобна некоторой жордановой матрице. Эта жорданова матрица определяется однозначно с точностью до порядка жордановых клеток.*

Для доказательства теоремы нам надо найти базу, в которой матрица имеет жорданову форму. Построение такой базы разбивается на несколько этапов. Вначале найдем базу, в которой матрица распадается на клетки с одним и тем же характеристическим корнем α, а потом уже каждую из этих клеток разобьем на жордановы клетки.

34.2. Корневое разложение.

О п р е д е л е н и е. Пусть V – векторное пространство над полем P, φ – линейное преобразование V, α – его собственное значение. Вектор $v \in V$ называется *корневым относительно* φ, если $v(\varphi - \alpha\varepsilon)^\rho = 0$ при подходящем $\rho \in \{0, 1, 2, \ldots\}$. Наименьшее такое ρ называется *высотой* v относительно φ.

Если $\rho = 1$, то $v\varphi = \alpha v$, т. е. v – собственный вектор. Следовательно, собственный вектор – корневой вектор высоты 1.

Т е о р е м а 1. 1) *Множество всех корневых векторов относительно преобразования* φ*, отвечающих одному и тому же собственному значению* α *преобразования* φ*, является подпространством пространства* V*.* 2) *Это подпространство инвариантно относительно* φ*.* Оно называется *корневым подпространством.*

Д о к а з а т е л ь с т в о. 1) Пусть V_α – множество корневых векторов относительно φ, отвечающие собственному значению α. Пусть

$$v_1(\varphi - \alpha\varepsilon)^{\rho_1} = 0, \quad v_2(\varphi - \alpha\varepsilon)^{\rho_2} = 0,$$

и $\beta_1, \beta_2 \in P$. Тогда

$$(\beta_1 v_1 + \beta_2 v_2)(\varphi - \alpha\varepsilon)^\rho = 0, \quad \text{где } \rho = \max\{\rho_1, \rho_2\}.$$

Следовательно, V_α является подпространством.

2) Докажем, что $V_\alpha\varphi \subseteq V_\alpha$. Надо доказать, что если $v_1 \in V_\alpha$, то вектор $v_1\varphi$ является корневым. Рассмотрим

$$(v_1\varphi)(\varphi - \alpha\varepsilon)^{\rho_1}$$

и, учитывая, что всякий многочлен от φ перестановочен с самим φ, получаем

$$(v_1\varphi)(\varphi - \alpha\varepsilon)^{\rho_1} = v_1(\varphi - \alpha\varepsilon)^{\rho_1}\varphi = 0\varphi = 0.$$

Теорема доказана.

У п р а ж н е н и е. Высоты векторов из V_α не превосходят кратности корня α в характеристическом многочлене преобразования φ, т. е. если

$$\chi(\lambda) = (\lambda - \alpha)^k \chi_1(\lambda), \quad \chi_1(\alpha) \neq 0,$$

то для любого v из V_α имеем

$$v(\varphi - \alpha\varepsilon)^k = 0.$$

Т е о р е м а 2. *Пусть поле* P *содержит все характеристические корни:* α_1, α_2, $\ldots$, α_s *преобразования* φ*. Пусть* V_1, V_2, $\ldots$, V_s *– соответствующие корневые подпространства, отвечающие* α_1, α_2, $\ldots$, α_s*. Тогда* $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$. Это разложение называется *корневым разложением.*

Д о к а з а т е л ь с т в о. Пусть характеристический многочлен преобразования φ имеет вид

$$\chi(\lambda) = (\lambda - \alpha_1)^{r_1}(\lambda - \alpha_2)^{r_2} \dots (\lambda - \alpha_s)^{r_s}, \quad \alpha_i \neq \alpha_j \text{ при } i \neq j.$$

Рассмотрим многочлены

$$\chi_i(\lambda) = \frac{\chi(\lambda)}{(\lambda - \alpha_i)^{r_i}}, \quad i = 1, 2, \dots, s.$$

а) Покажем, что $V_i = V\chi_i(\varphi)$. Установим включение $\supseteq$. Рассмотрим

$$v\chi_i(\varphi)(\varphi - \alpha_i\varepsilon)^{r_i} = v\chi(\varphi).$$

По теореме Гамильтона–Кэли $\chi_i(\varphi) = 0$, а потому $v\chi_i(\varphi)$ – корневой вектор. Следовательно, включение $\supseteq$ установлено.

Обратно. Пусть $w \in V_i$ и $w(\varphi - \alpha_i\varepsilon)^\rho = 0$. Надо доказать, что $w \in V\chi_i(\varphi)$. Так как $(\lambda - \alpha_i)^\rho$ и $\chi_i(\lambda)$ взаимно просты, то существуют многочлены $p_i(\lambda)$, $q_i(\lambda)$ из $P[\lambda]$ такие, что

$$p_i(\lambda)(\lambda - \alpha_i)^\rho + q_i(\lambda)\chi_i(\lambda) = 1.$$

Тогда

$$p_i(\varphi)(\varphi - \alpha_i\varepsilon)^\rho + q_i(\varphi)\chi_i(\varphi) = \varepsilon,$$

и отсюда

$$w = w\varepsilon = w[p_i(\varphi)(\varphi - \alpha_i\varepsilon)^\rho + q_i(\varphi)\chi_i(\varphi)] = [wq_i(\varphi)]\chi_i(\varphi),$$

т. е. равенство $V_i = V\chi_i(\varphi)$ доказано.

б) Покажем, что $V = V_1 + V_2 + \ldots + V_s$. Пусть $v \in V$. Так как многочлены $\chi_1(\lambda)$, $\chi_2(\lambda)$, …, $\chi_s(\lambda)$ взаимно просты, то существуют многочлены $h_1(\lambda)$, $h_2(\lambda)$, …, $h_s(\lambda)$ такие, что

$$h_1(\lambda)\chi_1(\lambda) + h_2(\lambda)\chi_2(\lambda) + \ldots + h_s(\lambda)\chi_s(\lambda) = 1.$$

Тогда

$$h_1(\varphi)\chi_1(\varphi) + h_2(\varphi)\chi_2(\varphi) + \ldots + h_s(\varphi)\chi_s(\varphi) = \varepsilon.$$

Отсюда

$$v = v\varepsilon = (vh_1(\varphi))\chi_1(\varphi) + (vh_2(\varphi))\chi_2(\varphi) + \ldots + (vh_s(\varphi))\chi_s(\varphi).$$

Заметим, что первое слагаемое лежит в V_1, второе в V_2 и т. д. Следовательно, мы получили нужное разложение.

в) Покажем, что $V = V_1 \oplus V_2 \oplus \ldots \oplus V_s$. Для этого достаточно показать, что нулевой вектор имеет единственную запись. Пусть

$$0 = u_1\chi_1(\varphi) + u_2\chi_2(\varphi) + \ldots + u_s\chi_s(\varphi).$$

Надо доказать, что все $u_i\chi_i(\varphi)=0$. Применяя к нашему равенству преобразование $\chi_i(\varphi)$, получим
$$0=u_i\chi_i(\varphi)^2,$$
так как слагаемое $u_k\chi_k(\varphi)\chi_i(\varphi)$ при $k\neq i$ обращается в нуль. Далее, учитывая, что многочлены $\chi_i(\lambda)$ и $(\lambda-\alpha_i)^{r_i}$ взаимно просты, найдем многочлены $f_i(\lambda)$, $g_i(\lambda)$ такие, что
$$\chi_i(\lambda)f_i(\lambda)+(\lambda-\alpha_i)^{m_i}g_i(\lambda)=1.$$
Тогда
$$\chi_i(\varphi)f_i(\varphi)+(\varphi-\alpha_i\varepsilon)^{r_i}g_i(\varphi)=\varepsilon$$
и
$$u_i\chi_i(\varphi)=u_i\chi_i(\varphi)\varepsilon=u_i\chi_i(\varphi)[\chi_i(\varphi)f_i(\varphi)+(\varphi-\alpha_i\varepsilon)^{r_i}g_i(\varphi)]=$$
$$=(u_i\chi_i^2(\varphi))f_i(\varphi)+u_i(\chi_i(\varphi)(\varphi-\alpha_i\varepsilon)^{r_i})g_i(\varphi)=u_i\chi(\varphi)g_i(\varphi)=0,$$
т. е. сумма действительно прямая. Теорема доказана.

Таким образом, мы можем разложить V в сумму инвариантных подпространств. Выбрав в каждом из подпространств базу, и объединив их, найдем базу V. В этой базе матрица преобразования φ имеет вид
$$[\varphi]=\left(\begin{array}{c|c|c|c} * & \mathbf{0} & \cdots & \mathbf{0}\\ \hline \mathbf{0} & * & \cdots & \mathbf{0}\\ \hline \cdots & \cdots & \cdots & \cdots\\ \hline \mathbf{0} & \mathbf{0} & \cdots & * \end{array}\right),$$
где каждой диагональной клетке соответствует единственное собственное значение.

34.3. Канонический вид нильпотентного преобразования. Ограничимся отдельным корневым подпространством. Всякий вектор из этого подпространства удовлетворяет равенству
$$v(\varphi-\alpha_i\varepsilon)^{r_i}=0.$$
Следовательно, преобразование $\psi=\varphi-\alpha_i\varepsilon$ является нильпотентным преобразованием корневого подпространства V_i. Для этого преобразования мы должны научиться строить базу, в которой матрица распадается на жордановы клетки.

О п р е д е л е н и е. Пусть ψ – нильпотентное преобразование векторного пространства V. Назовем подпространство U *циклическим относительно* ψ, если оно обладает базой e_1, e_2, …, e_k такой, что
$$e_1\psi=e_2, e_2\psi=e_3,\dots,e_k\psi=0.$$

Сама эта база называется *циклической.* Заметим, что циклическое подпространство U инвариантно относительно ψ и в циклической базе матрица преобразования имеет вид

$$[\psi_U] = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & 0 \end{pmatrix} \in M_k(P). \tag{1}$$

Покажем, что для всякого нильпотентного преобразования существует база, в которой матрица преобразования является клеточно-диагональной, а каждая диагональная клетка имеет вид (1).

Т е о р е м а 3. *Пусть* ψ *– нильпотентное преобразование пространства* V*. Тогда* 1) *существует разложение* V *в прямую сумму циклических относительно* ψ *подпространств*; 2) *если* z_i *– количество* i*-мерных циклических подпространств в этом разложении, то оно выражается формулой*

$$z_i = r_{i-1} - 2r_i + r_{i+1},$$

где $r_i = \dim \psi^i = \dim (V\psi^i)$.

Заметим, что из второй части этой теоремы вытекает, в частности, что какое бы разложение на циклические слагаемые не взяли, количество слагаемых будет одно и то же.

Д о к а з а т е л ь с т в о. 1) Пусть $\psi^p = 0$, $\psi^{p-1} \neq 0$. Надо построить базу, которая состоит из циклических отрезков. Рассмотрим

$$W_i = \operatorname{Ker} \psi^i, \quad i = 0, 1, 2, \dots, p.$$

Понятно, что все ядра вложены друг в друга:

$$0 = W_0 \subseteq W_1 \subseteq \dots \subseteq W_{p-1} \subseteq W_p = V.$$

Действительно, если $v\psi^{i-1} = 0$, то $v\psi^i = 0$. Будем двигаться по цепочке справа налево. Рассмотрим фактор-пространство W_i/W_{i-1}. Векторами в нем являются смежные классы. При этом классы

$$w_1 + W_{i-1}, \quad w_2 + W_{i-1}, \quad \dots, \quad w_l + W_{i-1}$$

линейно независимы, если $w_1, w_2, \dots, w_l$ линейно независимы по модулю W_{i-1}.

В дальнейшем всякое утверждение о смежных классах будем формулировать как такое же утверждение об их представителях, добавляя «по модулю ...».

Строим базу для цепочки

$$0 = W_0 \subseteq W_1 \subseteq \ldots \subseteq W_{p-1} \subseteq W_p = V.$$

Чтобы найти базу $e_1, e_2, \ldots, e_q$ по $\bmod W_{p-1}$, надо взять базу W_{p-1}, дополнить ее до базы V и дополняющие векторы годятся в качестве $e_1, e_2, \ldots, e_q$. Иными словами,

$$e_1, e_2, \ldots, e_q - \text{ база } W_p \text{ по } \bmod W_{p-1}$$

это значит, что

$$e_1 + W_{p-1}, e_2 + W_{p-1}, \ldots, e_q + W_{p-1}$$

– база W_p/W_{p-1}. Применим преобразование ψ, получим векторы

$$e_1\psi, e_2\psi, \ldots, e_q\psi,$$

которые лежат в W_{p-1}.

Так как $\psi^p = 0$, а $\psi^{p-1} \neq 0$, то построенные векторы линейно независимы по $\bmod W_{p-2}$. Действительно, если

$$\alpha_1(e_1\psi) + \alpha_2(e_2\psi) + \ldots + \alpha_q(e_q\psi) \equiv 0 \ (\bmod W_{p-2}), \quad \alpha_i \in P,$$

то это равносильно тому, что

$$\alpha_1(e_1\psi + W_{p-2}) + \alpha_2(e_2\psi + W_{p-2}) + \ldots + \alpha_q(e_q\psi + W_{p-2}) = W_{p-2},$$

т. е.

$$\alpha_1(e_1\psi) + \alpha_2(e_2\psi) + \ldots + \alpha_q(e_q\psi) \in W_{p-2}$$

или

$$(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q)\psi \in W_{p-2}.$$

Следовательно,

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_q e_q \in W_{p-1},$$

но мы выбирали векторы $e_1, e_2, \ldots, e_q$ как дополняющие базу W_{p-1} до базы W_p, а потому $\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$, т. е. $e_1\psi$, $e_2\psi$, $\ldots$, $e_q\psi$, линейно независимы по $\bmod W_{p-2}$.

Всякую линейно независимую систему можно дополнить до базы:

$$e_1\psi, e_2\psi, \ldots, e_q\psi, f_1, f_2, \ldots, f_s - \text{ база } W_{p-1} \text{ по } \bmod W_{p-2}.$$

Подействовав ψ, получим векторы

$$e_1\psi^2, e_2\psi^2, \ldots, e_q\psi^2, f_1\psi, f_2\psi, \ldots, f_s\psi, \in;$$

которые лежат в W_{p-2}.

Покажем что все они линейно независимы по $\operatorname{mod} W_{p-3}$.

Действительно, пусть

$$\alpha'_1(e_1\psi^2)+\alpha'_2(e_2\psi^2)+\ldots+\alpha'_q(e_q\psi^2)+\beta_1(f_1\psi)+\beta_2(f_2\psi)+\ldots+\beta_s(f_s\psi)\equiv 0\,(\operatorname{mod} W_{p-3}),$$

для некоторых $\alpha'_i, \beta_j \in P$. Тогда

$$[\alpha'_1(e_1\psi)+\alpha'_2(e_2\psi)+\ldots+\alpha'_q(e_q\psi)+\beta_1 f_1+\beta_2 f_2+\ldots+\beta_s f_s]\psi \in W_{p-3},$$

и

$$[\alpha'_1(e_1\psi)+\alpha'_2(e_2\psi)+\ldots+\alpha'_q(e_q\psi)+\beta_1 f_1+\beta_2 f_2+\ldots+\beta_s f_s] \in W_{p-2},$$

но векторы $e_1\psi$, $e_2\psi$, $\ldots$, $e_q\psi$, f_1, f_2, $\ldots$, f_s были линейно независимы по $\operatorname{mod} W_{p-2}$. Следовательно, все $\alpha'_i = \beta_j = 0$.

Дополним до базы:

$$e_1\psi^2, e_2\psi^2, \ldots, e_q\psi^2, f_1\psi, f_2\psi, \ldots, f_s\psi, g_1, g_2, \ldots, g_t - \text{ база } W_{p-2} \text{ по } \operatorname{mod} W_{p-3}.$$

Продолжая этот процесс, получим

$$e_1\psi^{p-1}, e_2\psi^{p-1}, \ldots, e_q\psi^{p-1}, f_1\psi^{p-2}, f_2\psi^{p-2}, \ldots, f_s\psi^{p-2}, g_1\psi^{p-3}, g_2\psi^{p-3}, \ldots, g_t\psi^{p-3}, \ldots,$$

$$h_1, h_2, \ldots, h_l - \text{ база } W_1 \text{ по } \operatorname{mod} W_0.$$

Покажем, что если объединить все построенные системы векторов, то получим базу V. Докажем вначале, что она линейно независима. Пусть

$$\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_q e_q+\alpha'_1(e_1\psi)+\alpha'_2(e_2\psi)+\ldots+\alpha'_q(e_q\psi)+\beta_1(f_1\psi)+\beta_2(f_2\psi)+\ldots+\beta_s(f_s\psi)++\ldots+$$

$$+\gamma_1 h_1+\gamma_2 h_2+\ldots+\gamma_t h_t = 0.$$

Заметим, что вектор из левой части лежит в W_{p-1}. Следовательно, все $\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$. Оставшийся вектор лежит в W_{p-1}. Следовательно, все $\alpha'_i = \beta_j = 0$, и т. д. Таким образом, система векторов линейно независима.

Докажем, что построенная система векторов максимальна. Пусть $v \in V$. Тогда

$$v \equiv \alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_q e_q\,(\operatorname{mod} W_{p-1}),$$

при подходящих α_i. Рассмотрим вектор

$$v-(\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_q e_q) \in W_{p-1}.$$

Для него

$$v-(\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_q e_q) \equiv +\alpha'_1(e_1\psi)+\alpha'_2(e_2\psi)+\ldots+\alpha'_q(e_q\psi)+$$

$$\beta_1 f_1 + \beta_2 f_2 + \ldots + \beta_s f_s \pmod{W_{p-2}}.$$

Далее рассмотрим

$$v-(\alpha_1 e_1+\alpha_2 e_2+\ldots+\alpha_q e_q)-(\alpha_1'(e_1\psi)+\alpha_2'(e_2\psi)+\ldots+\alpha_q'(e_q\psi)+\beta_1 f_1+\beta_2 f_2+\ldots+\beta_s f_s+) \in W_{p-2},$$

и т. д. Продолжая этот процесс, получим, что вектор v есть линейная комбинация векторов построенной системы.

Весь процесс построения удобно заносить в следующую таблицу

$$\begin{array}{rlllll}
V = W_p & & & & & \\
\cup & e_1, \ldots, e_q & & & & - \text{база } W_p \pmod{W_{p-1}}; \\
W_{p-1} & & & & & \\
\cup & e_1\psi, \ldots, e_q\psi & f_1, \ldots, f_s, & & & - \text{база } W_{p-1} \pmod{W_{p-2}}; \\
\cup & & & & & \\
 & \ldots & \ldots & \ldots & & \\
\cup & & & & & \\
W_1 & & & & & \\
\cup & e_1\psi^{p-1}, \ldots, e_q\psi^{p-1}, & f_1\psi^{p-2}, \ldots, f_s\psi^{p-2}, & \ldots, & h_1, h_2, \ldots, h_l, & \\
 & & & & - \text{база } W_1 \pmod{W_0}. & \\
0 = W_0 & & & & &
\end{array}$$

Видим, что каждый столбец этой системы образует циклическую базу.

2) Пусть z_i – число i-мерных циклических подпространств. Обозначим $d_i = \dim W_i$. Из первой строки таблицы видим, что число z_p циклических подпространств размерности p равно q и равно $\dim W_p - \dim W_{p-1}$. Из второй строки таблицы видим, что число z_{p-1} циклических подпространств размерности $p-1$ равно r и сумма $z_p + z_{p-1}$ равна $\dim W_{p-1} - \dim W_{p-2}$ и. д.

В общем случае, из i-й строки построенной базы, получаем равенство

$$z_{i+1} + z_{i+2} + \ldots + z_p = d_{i+1} - d_i,$$

из $(i+1)$-й – равенство

$$z_i + z_{i+1} + \ldots + z_p = d_i - d_{i-1}.$$

Вычитая из второго равенства первое, получим

$$z_i = -d_{i-1} + 2d_i - d_{i+1}.$$

Учитывая, что $r_i + d_i = n$ (размерность образа плюс размерность ядра равна размерности пространства), приходим к равенству

$$z_i = -d_{i-1} + 2d_i - d_{i+1} = r_{i-1} - 2r_i + r_{i+1}.$$

Теорема доказана.

34.4. Доказательство теоремы Жордана. Докажем вначале существование.

Пусть дано линейное преобразование φ векторного пространства V. Надо построить такую базу пространства V, в которой матрица преобразования φ имеет жорданову форму. Представим V в виде прямой суммы корневых подпространств

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_s,$$

где каждое подпространство V_i соответствует собственному значению α_i. Рассмотрим ограничение φ на V_i, получим индуцированное преобразование φ_{V_i}. Это линейное преобразование подпространства V_i. Если рассмотреть ограничение $(\varphi - \alpha_i\varepsilon)|_{V_i}$, то это тоже линейное преобразование подпространства V_i. Как мы знаем, справедливы равенства

$$\begin{gathered} V_1(\varphi - \alpha_1\varepsilon)^{\rho_1} = 0, \\ V_2(\varphi - \alpha_2\varepsilon)^{\rho_2} = 0, \\ \vdots \\ V_s(\varphi - \alpha_s\varepsilon)^{\rho_s} = 0. \end{gathered}$$

Выбирая $\rho = \max\{\rho_1, \rho_2, \ldots, \rho_s\}$, видим, что $(\varphi - \alpha_i\varepsilon)|_{V_i} = (\varphi - \alpha_i\varepsilon)_{V_i}$ – нильпотентно ступени не выше ρ. Допустим, что

$$V_i = V_{i_1} \oplus V_{i_2} \oplus \ldots \oplus V_{i_k}$$

– прямая сумма циклических подпространств. Выбирая в каждом V_{i_j} циклическую базу, объединив их, получим, что в этой базе

$$[(\varphi - \alpha_i\varepsilon)_{V_i}] = \left(\begin{array}{c|c|c|c} J_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & J_2 & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & J_k \end{array}\right),$$

где каждая клетка имеет такой вид

$$J_l = \begin{pmatrix} 0 & 1 & & & \mathbf{0} \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & 0 \end{pmatrix}.$$

Если рассмотреть в этой базе матрицу $[\varphi_{V_i}]$, то она будет иметь вид

$$[\varphi_{V_i}] = \left(\begin{array}{c|c|c|c} J_1(\alpha_i) & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & J_2(\alpha_i) & \cdots & \mathbf{0} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & J_k(\alpha_i) \end{array}\right),$$

где каждая клетка является клеткой жордана

$$J_l(\alpha_i) = \begin{pmatrix} \alpha_i & 1 & & & \mathbf{0} \\ & \alpha_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha_i \end{pmatrix}.$$

Объединяя все такие базы подпространств V_i, получим базу, в которой матрица преобразования φ имеет жорданову форму.

Докажем единственность. Пусть некоторое линейное преобразование имеет две жордановы матрицы. Как мы знаем, z_i – количество i-мерных клеток вычисляется по формуле:

$$z_i = r_{i-1} - 2r_i + r_{i+1}.$$

Следовательно, количество клеток в обеих жордановых матрицах одинаково. Может меняться только порядок клеток. Теорема доказана.

§ 35. Полиномиальные матрицы

35.1. Задача об эквивалентности λ-матриц. *Полиномиальной* называют матрицу, элементами которой являются многочлены. Рассмотрим кольцо многочленов $P[\lambda]$ над полем P от переменной λ. Тогда полиномиальная матрица

$$F = \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{pmatrix}, \quad f_{ij} \in P[\lambda].$$

Часто полиномиальные матрицы называют «λ-матрицами». Частный случай λ-матриц – характеристические матрицы.

О п р е д е л е н и е. *Элементарными преобразованиями* λ-матриц называются следующие преобразования: 1) умножение строки на ненулевой скаляр; 2) прибавление к одной строке другой, умноженной на ненулевой скаляр; $1)'$ умножение столбца на ненулевой скаляр; $2)'$ прибавление к одному столбцу другого, умноженного на ненулевой скаляр.

О п р е д е л е н и е. Две λ-матрицы F_1 и F_2 называются *эквивалентными*, если от F_1 можно перейти к F_2 цепочкой элементарных преобразований:

$$F_1 = G_1 \longrightarrow G_2 \longrightarrow \dots \longrightarrow G_s = F_2.$$

П р и м е р. Если в λ-матрице переставить две строки, то полученная матрица ей эквивалентна. Действительно, пусть F – некоторая λ-матрица. Выполним следующие элементарные преобразования:

1) умножим j-ю строку на 1 и прибавим к i-й;
2) умножим i-ю строку на -1 и прибавим к j-й;
3) умножим j-ю строку на 1 и прибавим к i-й;
4) j-ю строку умножим на -1.

Видим, что полученная матрица получается из F перестановкой i-й и j-й строк.

Нетрудно убедиться, что введенное отношение действительно является отношением эквивалентности. Относительно этого отношения эквивалентности множество λ-матриц распадается на смежные классы.

О п р е д е л е н и е. *Канонической* называется диагональная λ-матрица:

$$\operatorname{diag}(d_1(\lambda), d_2(\lambda), \ldots, d_n(\lambda))$$

такая, что: 1) $d_i(\lambda)|d_{i+1}(\lambda)$, $i = 1, 2, \ldots, n-1$; 2) если многочлен $d_i(\lambda)$ отличен от нуля, то его старший коэффициент равен единице.

Из этого определения следует, что каноническая матрица имеет вид:

$$\operatorname{diag}(1, 1, \ldots, 1, f_1, f_2, \ldots f_s, 0, 0, \ldots, 0),$$

т. е. единицы стоят вверху, а нули – внизу.

О п р е д е л е н и е. Элементы $d_i(\lambda)$ называются *инвариантными множителями* канонической матрицы.

Докажем, что в каждом смежном классе лежит по одной канонической матрице.

35.2. Приведение λ-матрицы к каноническому виду.

Т е о р е м а 1. *Всякая λ-матрица с помощью элементарных преобразований может быть приведена к каноническому виду.*

Д о к а з а т е л ь с т в о проведем индукцией по n – размерности матрицы. При $n = 1$ матрица является многочленом. Если его старший коэффициент равен единице, то мы имеем каноническую матрицу, в противном случае, поделив на старший коэффициент, приходим к канонической матрице.

Предположим, что теорема справедлива для матриц размера $n-1$ и рассмотрим λ-матрицу F размера n. Если $F = 0$, то доказывать нечего. Пусть $F \neq 0$. Среди всех матриц, эквивалентных F, выберем матрицу, у которой на месте $(1,1)$ стоит многочлен наименьшей степени со старшим коэффициентом 1, т. е.

$$F \text{ эквивалентна матрице } \begin{pmatrix} f_{11} & f_{12} & \ldots & f_{1n} \\ f_{21} & f_{22} & \ldots & f_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ f_{n1} & f_{n2} & \ldots & f_{nn} \end{pmatrix}$$

Утверждается, что $f_{11}|f_{1i}$ и $f_{11}|f_{i1}$ при $i = 2, 3, \ldots, n$. Действительно, если $f_{1i} = f_{11}q_{1i} + r_{1i}$ и $r_{1i} \neq 0$, то $\deg r_{1i} < \deg f_{1i}$, но тогда, используя элементарные преобразования, получим матрицу, эквивалентную F, у которой на месте $(1, 1)$ стоит r_{1i}. Противоречие. Аналогично проверяется, что и $f_{11}|f_{i1}$. Используя элементарные преобразования, перейдем от матрицы F к матрице

$$\begin{pmatrix} f_{11} & 0 & \ldots & 0 \\ 0 & g_{22} & \ldots & g_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & g_{n2} & \ldots & g_{nn} \end{pmatrix}, \quad g_{ij} \in P[\lambda].$$

Заметим, что в ней $f_{11}|g_{ij}$. Действительно, прибавляя i-ю строку к первой, также как и выше устанавливаем, что $f_{11}|g_{ij}$. По предположению индукции существует цепочка элементарных преобразований:

$$\begin{pmatrix} g_{22} & \ldots & g_{2n} \\ \ldots & \ldots & \ldots \\ g_{n2} & \ldots & g_{nn} \end{pmatrix} \longrightarrow \ldots \longrightarrow \begin{pmatrix} h_2 & \ldots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \ldots & h_n \end{pmatrix},$$

где последняя матрица диагональна, $h_i(\lambda)|h_{i+i}(\lambda)$ и если $h_i(\lambda) \neq 0$, то старший коэффициент равен единице. Распространяя эту цепочку элементарных преобразований с матрицы (g_{ij}), до нашей матрице, получим

$$\left(\begin{array}{c|ccc} f_{11} & 0 & \ldots & 0 \\ \hline 0 & g_{22} & \ldots & g_{2n} \\ \vdots & \ldots & \ldots & \ldots \\ 0 & g_{n2} & \ldots & g_{nn} \end{array}\right) \longrightarrow \ldots \longrightarrow \left(\begin{array}{c|ccc} f_{11} & 0 & \ldots & 0 \\ \hline 0 & h_2 & \ldots & \mathbf{0} \\ \vdots & & \ddots & \\ 0 & \mathbf{0} & \ldots & h_n \end{array}\right).$$

Как легко заметить, последняя матрица является канонической. Теорема доказана.

35.3. Единственность канонического вида λ-матриц. Пусть F – λ-матрица, для каждого $k = 1, 2, \ldots, n$ обозначим $D_{k,F}(\lambda)$ (или просто $D_k(\lambda)$) приведенный наибольший общий делитель миноров k-го порядка матрицы F, если они не все равны нулю и положим $D_{k,F}(\lambda) = 0$ в противном случае. Таким образом, каждой λ-матрице мы сопоставим n многочленов $D_1(\lambda)$, $D_2(\lambda)$, …, $D_n(\lambda)$. Покажем, что они остаются инвариантными при элементарных преобразованиях.

Т е о р е м а 2. *Если матрицы F и G эквивалентны, то $D_{k,F}(\lambda) = D_{k,G}(\lambda)$, $k = 1, 2, \ldots, n$.*

Д о к а з а т е л ь с т в о. Так как F эквивалентна G, то существует цепочка элементарных преобразований, переводящая F в G. Если мы покажем, что $D_{k,F}(\lambda)$ не меняется при выполнении одного элементарного преобразования, то и любая цепочка не меняет $D_{k,F}(\lambda)$.

1) Предположим, что G получается из F при помощи умножения некоторой строки на ненулевой скаляр. Если эта строка входит в минор, то после применения элементарного преобразования, он умножается на скаляр, но когда начнем вычислять приведенный общий делитель, то эти скаляры никак не влияют. Аналогично проверяется, что и при умножении столбца на ненулевой скаляр $D_{k,F}(\lambda)$ не меняются.

2) Пусть G получается из F после умножения i-й строки на $f \in P$, $f \neq 0$, и прибавления к j-й строке. В зависимости от того, как расположен рассматриваемый минор, возможны четыре случая:

а) i-я и j-я строки проходят через минор;

б) i-я строка проходит через минор, а j-я – нет;

в) ни i-я ни j-я строки не проходят через минор;

г) j-я строка проходит через минор, а i-я – нет.

В случаях а)–в) очевидно, что после элементарного преобразования 2) минор не меняется.

Докажем, что и в случае г) утверждение теоремы справедливо. Пусть соответствующий минор матрицы G имеет вид

$$M_G = \begin{vmatrix} \dots & \dots & \dots & \dots \\ f_{j,\nu_1} + f_{i,\nu_1} & f_{j,\nu_2} + f_{i,\nu_2} & \dots & f_{j,\nu_k} + f_{i,\nu_k} \\ \dots & \dots & \dots & \dots \end{vmatrix},$$

где выделена j-я строка. Представим его в виде суммы двух определителей:

$$M_G = \begin{vmatrix} \dots & \dots & \dots & \dots \\ f_{j,\nu_1} & f_{j,\nu_2} & \dots & f_{j,\nu_k} \\ \dots & \dots & \dots & \dots \end{vmatrix} + f \begin{vmatrix} \dots & \dots & \dots & \dots \\ f_{i,\nu_1} & f_{i,\nu_2} & \dots & f_{i,\nu_k} \\ \dots & \dots & \dots & \dots \end{vmatrix} = M_F + fN_F,$$

где символом N_F обозначили второй минор, который отличается от соответствующего минора матрицы F тем, что у него вместо j-й строки стоит соответствующая часть i-й строки матрицы F. Рассмотрим $D_{k,F}$. По условию,

$$D_{k,F}|M_F, \quad D_{k,F}|N_F$$

Следовательно, $D_{k,F}|M_G$, а потому $D_{k,F}|D_{k,G}$. Аналогично проверяется, что $D_{k,G}|D_{k,F}$. Следовательно, $D_{k,F} = D_{k,G}$.

Случай, когда G получается из F при помощи элементарного преобразования $2)'$ разбирается аналогично. Теорема доказана.

Т е о р е м а 3. *Для всякой λ-матрицы F: 1) $d_1 d_2 \dots d_k = D_{k,F}$; 2) канонический вид λ-матрицы определяется ею однозначно.*

Д о к а з а т е л ь с т в о. 1) Как мы знаем, λ-матрица F эквивалентна диагональной матрице

$$\mathrm{diag}(d_1, d_2, \dots, d_n).$$

Рассмотрим миноры порядка k полученной диагональной матрицы. Очевидно, если некоторая строка проходит через минор, а одноименный столбец – нет, то этот минор равен нулю. Следовательно, всякий ненулевой минор порядка k равен $d_{i_1}d_{i_2}\dots d_{i_k}$. Среди таких миноров есть и минор $d_1d_2\dots d_k$, а это и будет $D_{k,F}$, так как $d_1|d_2$, $d_2|d_3$, и т. д., то

$$d_1d_2\dots d_k|d_{i_1}d_{i_2}\dots d_{i_k},$$

т. е. $d_1d_2\dots d_k$ делит любой минор порядка k, а так как увеличить его нельзя, то он и будет наибольшим общим делителем миноров порядка k.

2) По теореме 2 набор $D_{1,F}, D_{2,F}, \dots, D_{n,F}$ определяется единственным образом. Так как $d_k = D_{k,F}/D_{k-1,F}$, то и d_k определяется единственным образом. Следовательно, каноническая матрица определяется единственным образом. Теорема доказана.

Таким образом, мы решили исходную задачу, показав, что всякая λ-матрица эквивалентна единственной канонической матрице.

35.4. Эквивалентность λ-матриц унимодулярным матрицам.

О п р е д е л е н и е. *Унимодулярная матрица* – это λ-матрица, определитель которой лежит в самом поле P и не равен нулю.

Т е о р е м а 4. *Матрицы F и G эквивалентны тогда и только тогда, когда существуют унимодулярные матрицы X и Y такие, что $F = XQY$.*

Д о к а з а т е л ь с т в о. Установим импликацию $\Rightarrow$. Пусть F эквивалентна G, т. е. существует цепочка элементарных преобразований

$$F \longrightarrow \dots \longrightarrow G. \tag{1}$$

Надо построить унимодулярные матрицы X и Y, удовлетворяющие равенству $F = XQY$.

Заметим, что каждому элементарному преобразованию соответствует умножение на некоторую матрицу. Например, элементарному преобразованию 1) (умножение i-й строки матрицы на ненулевой скаляр α) соответствует умножению слева на диагональную матрицу

$$\mathrm{Diag}(1, \dots, 1, \alpha, 1, \dots, 1),$$

где α стоит на i-м месте. Элементарному преобразованию 2) (умножение i-й строки на ненулевой элемент f и прибавление к j-й строке) соответствует умножению слева на трансвекцию $T_{ij}(f)$. Элементарным преобразованиям $1)'$ и $2)'$ соответствуют умножение справа на диагональную матрицу и, соответственно, на трансвекцию.

Таким образом, цепочке элементарных преобразований (1) соответствует произведение матриц:

$$X_1X_2\dots X_sFX_{s+1}X_{s+2}\dots X_t = G,$$

где каждая X_i либо диагональная матрица, либо трансвекция. Учитывая, что

$$\det \mathrm{Diag}(1, \dots, 1, \alpha, 1, \dots, 1) = \alpha, \quad \det T_{ij}(f) = 1,$$

заключаем, что матрицы

$$X_1X_2\dots X_s = X^{-1},\quad X_{s+1}X_{s+2}\dots X_t = Y^{-1}$$

являются унимодулярными, а так как обратные к ним также унимодулярные, то построенные матрицы и будут искомыми.

$\Leftarrow$. Пусть $F = XGY$, где X и Y – унимодулярные матрицы. Сама матрица X зависит от λ, а ее определитель – не зависит. Очевидно, $D_{n,X}(\lambda) = 1$ так как является приведенным многочленом и с точностью до умножения на скаляр совпадает с $\det X$. Все остальные $D_{k,X}(\lambda)$ при $1 \le k < n$ равны 1, что следует из того, что $d_1d_2\dots d_n = 1$ и ни один из d_i не содержит λ. Следовательно, все $D_{k,X}(\lambda) = 1$, а потому по теореме 3 и все $d_i = 1$. Таким образом, матрица X эквивалентна единичной матрице. Аналогично проверяется, что любая унимодулярная матрица эквивалентна единичной матрице. Следовательно, существуют цепочки элементарных преобразований:

$$X \longrightarrow \dots \longrightarrow E,$$

$$Y \longrightarrow \dots \longrightarrow E.$$

или на матричном языке:

$$X = X_1X_2\dots X_sEX_{s+1}X_{s+2}\dots X_t,$$

$$Y = Y_1Y_2\dots Y_pEY_{p+1}Y_{p+2}\dots Y_q,$$

где X_i, Y_j – матрицы, соответствующие элементарным преобразованиям. Отсюда

$$F = XGY = X_1X_2\dots X_tGY_1Y_2\dots Y_q,$$

а значит, существует цепочка элементарных преобразований:

$$G \longrightarrow \dots \longrightarrow F.$$

Следовательно, G эквивалентна F. Теорема доказана.

35.5. Деление λ-матриц. Если

$$A(\lambda) = \left(\sum_k a_{ij}^k\lambda^k\right)_{i,j=1}^n$$

– некоторая λ-матрица, то ее можно представить в виде

$$A(\lambda) = A_0 + A_1\lambda + \dots + A_s\lambda^s,$$

где каждая матрица A_i является матрицей из $M_n(P)$. Если в этом разложении $A_s \neq 0$, то назовем s *степенью* матрицы $A(\lambda)$ и будем обозначать $s = \deg A(\lambda)$.

П р и м е р.

$$\begin{pmatrix} \lambda^2+1 & \lambda \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\lambda + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\lambda^2.$$

О п р е д е л е н и е. λ-матрица $A(\lambda)$ называется *регулярной*, если $\deg A_s \neq 0$.

Л е м м а 1. *Для любых λ-матриц $A(\lambda), B(\lambda)$ справедливы неравенства:*

$$\deg\left(A(\lambda) + B(\lambda)\right) \leq \max\left\{\deg A(\lambda), \deg B(\lambda)\right\},$$

$$\deg\left(A(\lambda) \cdot B(\lambda)\right) \leq \deg A(\lambda) + \deg B(\lambda).$$

Если хотя бы одна из матриц $A(\lambda)$ или $B(\lambda)$ регулярна, то последнее неравенство превращается в равенство.

Л е м м а 2. *Пусть $A(\lambda)$ и $B(\lambda)$ – две λ-матрицы одной размерности, причем $A(\lambda)$ регулярна. Тогда существуют однозначно определенные матрицы $Q_l(\lambda)$, $R_l(\lambda)$, $Q_r(\lambda)$, $R_r(\lambda)$, такие, что*

$$B(\lambda) = A(\lambda)Q_l(\lambda) + R_l(\lambda), \quad B(\lambda) = Q_r(\lambda)A(\lambda) + R_r(\lambda),$$

где $R_l(\lambda)$ либо равна нулю, либо ее степень меньше $\deg(A(\lambda))$*, тоже для $R_r(\lambda)$.*

Доказательство лемм 1 и 2 такое же как для обычных многочленов.

35.6. Скалярная эквивалентность λ-матриц. Матрицы $A(\lambda)$ и $B(\lambda)$ называются *скалярно эквивалентными*, если существуют невырожденные скалярные матрицы P и Q (их элементы не зависят от λ) такие, что

$$A(\lambda) = PB(\lambda)Q.$$

Т е о р е м а 5 (критерий эквивалентности). *Регулярные λ-матрицы $A\lambda + B$ и $C\lambda + D$ первой степени эквивалентны тогда и только тогда, когда они скалярно эквивалентны.*

Д о к а з а т е л ь с т в о. $\Leftarrow$. Следует из того, что условие скалярной эквивалентности более сильное, чем условие унимодулярной эквивалентности.

$\Rightarrow$. Пусть

$$A\lambda + B = X(C\lambda + D)Y, \tag{2}$$

где X, Y – унимодулярные матрицы. Разделим X слева на $A\lambda + B$, а Y – справа:

$$X = (A\lambda + B)X_1 + X_2, \quad Y = Y_1(A\lambda + B) + Y_2. \tag{3}$$

Степени X_2, Y_2 нулевые или они равны нулю. В любом случае, они не зависят от λ. Следовательно, из равенства (2) имеем

$$X^{-1}(A\lambda + B) = (C\lambda + D)Y_1(A\lambda + B) + (C\lambda + D)Y_2,$$

или

$$[X^{-1} - (C\lambda + D)Y_1](A\lambda + B) = (C\lambda + D)Y_2.$$

Как мы знаем, если X – унимодулярна, то X^{-1} тоже унимодулярна. Положим

$$T = X^{-1} - (C\lambda + D)Y_1. \tag{4}$$

Сравнивая степени левой и правой частей, видим, что T – скалярная матрица, т. е.

$$T(A\lambda + B) = (C\lambda + D)Y_2,$$

где T и Y_2 – скалярные матрицы. Надо показать, что T – невырожденная матрица. Преобразуем равенство (4):

$$XT = E - X(C\lambda + D)Y_1,$$

или

$$E = XT + X(C\lambda + D)Y_1.$$

Первое вхождение X заменим на выражение из (3), а произведение $X(C\lambda + D)$ – на выражение из (2), получим

$$E = (A\lambda + B)X_1T + X_2T + (A\lambda + B)Y^{-1}Y_1,$$

т. е.

$$E - X_2T = (A\lambda + B)(X_1T + Y^{-1}Y_1).$$

Видим, что в левой части стоит скалярная матрица, а матрица $A\lambda + B$ содержит λ. Следовательно, $X_1T + Y^{-1}Y_1 = 0$, т. е.

$$E - X_2T = 0 \Leftrightarrow X_2T = E,$$

а потому $\det T \neq 0$. Из равенства

$$A\lambda + B = T^{-1}(C\lambda + D)Y_2,$$

приравнивая матрицы при λ, получаем

$$A = T^{-1}CY_2$$

и вычисляя определитель

$$\det A = \det T^{-1} \cdot \det C \cdot \det Y_2,$$

видим, что так как $\det A \neq 0$ то и $\det Y_2 \neq 0$. Теорема доказана.

35.7. Критерий подобия скалярных матриц.

Т е о р е м а 6. *Две скалярные матрицы A и B подобны тогда и только тогда, когда их характеристические матрицы эквивалентны.*

Д о к а з а т е л ь с т в о. Пусть A и B подобны, т. е. $A = T^{-1}BT$ для некоторой невырожденной матрицы T. Тогда

$$T^{-1}(B - \lambda E)T = T^{-1}BT - \lambda T^{-1}ET = A - \lambda E,$$

т. е.

$$A - \lambda E \sim B - \lambda E.$$

Обратно. Пусть

$$A - \lambda E \sim B - \lambda E$$

– это регулярные матрицы первой степени. По доказанной теореме $A - \lambda E$ и $B - \lambda E$ скалярно эквивалентны, т. е. существуют невырожденные матрицы P и Q такие, что

$$A - \lambda E = P(B - \lambda E)Q.$$

Приравнивая матрицы при одинаковых степенях λ, получим

$$A = PBQ, \quad -E = -PEQ.$$

Из последнего равенства

$$E = PQ,$$

т. е.

$$P = Q^{-1}.$$

Следовательно

$$A = Q^{-1}BQ.$$

Теорема доказана.

35.8. Элементарные делители λ-матриц. Любая λ-матрица определяется набором линейных множителей в нормальной форме.

О п р е д е л е н и е. Пусть F – нормальная диагональная форма некоторой λ-матрицы A:

$$F = \mathrm{Diag}(f_1, f_2, \ldots, f_n), \quad \text{старший коэффициент } f_1 \text{ равен 1}.$$

Выбросим из набора $f_1, f_2, \ldots, f_n$ те множители, которые равны 0 или 1. Остальные разложим над основным полем P на неразложимые множители:

$$f_i = \varepsilon_{i_1}^{\alpha_{i_1}} \varepsilon_{i_2}^{\alpha_{i_2}} \ldots \varepsilon_{i_{s_i}}^{\alpha_{i_{s_i}}}, \quad \varepsilon_{i_j} \neq \varepsilon_{i_t} \text{ при } j \neq t,$$

где ε_{i_j} – неразложим над P. Тогда множество

$$\{\varepsilon_{i_j}^{\alpha_{i_j}}\}$$

называется *системой элементарных делителей* матрицы A.

П р и м е р. Пусть

$$A \sim F = \mathrm{Diag}(1, 1, \lambda, \lambda, \lambda(1+\lambda), \lambda^2(1+\lambda), 0, 0).$$

Тогда

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

– система элементарных делителей матрицы A.

Т е о р е м а 7. *Система элементарных делителей матрицы A, ее ранг и размерность определяют A с точностью до эквивалентности.*

П р и м е р. Пусть размерность матрицы A равна 7, ранг равен 5, а серия элементарных делителей:

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

Найдем нормальную диагональную форму матрицы A. Так как $7 - 5 = 2$, то в конце стоит два нуля и

$$A \sim \mathrm{Diag}(1, \lambda, \lambda, \lambda^2(1+\lambda), f_5, 0, 0).$$

При этом $f_5 = \lambda^i(1+\lambda)^j$ для некоторых i и j. Учитывая, что все f_i делят f_5, заключаем, что $f_5 = \lambda^2(1+\lambda)^2$.

Т е о р е м а 8. *Система элементарных делителей клеточно-диагональной матрицы*

$$G = \mathrm{Diag}(G_1, G_2, \ldots, G_t)$$

совпадает с объединением систем элементарных делителей ее диагональных клеток G_1, G_2, $\ldots$, G_t.

Д о к а з а т е л ь с т в о. 1) Пусть G – диагональная матрица:

$$G = \mathrm{diag}(g_1, g_2, \ldots, g_n)$$

и ее нормальная диагональная форма имеет вид:

$$F = \mathrm{diag}(f_1, f_2, \ldots, f_n).$$

Хотим показать, что элементарные множители для $f_1, f_2, \ldots, f_n$ совпадают с элементарными множителями для $g_1, g_2, \ldots, g_n$.

Можно считать, что $g_i(x) \neq 0$, $i = 1, 2, \ldots, n$ и

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s$$

– все элементарные делители $g_1, g_2, \ldots, g_n$, т. е.

$$g_i = a_i \varepsilon_1^{m_{i,1}} \varepsilon_2^{m_{i,2}} \ldots \varepsilon_s^{m_{i,s}}.$$

Покажем, что ε_i входят в f_i с теми же степенями. Так как

$$\det G = g_1 g_2 \ldots g_n = a f_1 f_2 \ldots f_n = a \prod_{i=1}^{n} f_i = a \prod_{i,j} \varepsilon_{ij}^{\alpha_{ij}}$$

– разложение по серии элементарных делителей F. Учитывая, что разложение на неразложимые множители однозначно, видим, что каждый элементарный делитель матрицы F входит в множество элементарных делителей матрицы G.

Как выглядит элементарный делитель G, соответствующий ε_1? Вычислим определители миноров k-го порядка матрицы G. Не нарушая общности, можно считать, что

$$m_{11} \leq m_{21} \leq \ldots \leq m_{n1}.$$

Вычислим наибольший общий делитель миноров k-го порядка матрицы G и обозначим его $(D_k(G))$

$$\{g_{i_1} g_{i_2} \ldots g_{i_k} \mid i_1 \leq i_2 \leq \ldots \leq i_k\}.$$

Видим, что

$$g_{i_1} g_{i_2} \ldots g_{i_k} = \varepsilon_1^{m_{i_1,1} + m_{i_2,1} + \ldots + m_{i_k,1}} h(\lambda).$$

Наименьшая степень

$$\varepsilon_1^{m_{1,1} + m_{2,1} + \ldots + m_{k,1}} F_k(\lambda),$$

$$f_k = d_k(G) = \frac{D_k(G)}{D_{k-1}(G)} = \varepsilon^{m_{k1}} F_k^*(\lambda).$$

Так как $F_k^*(\lambda)$ не делится на $F_k(\varepsilon_1)$. Набор элементарных делителей ε_1 совпадает.

2) Пусть

$$F = \operatorname{diag}(F_1, F_2, \ldots, F_s)$$

– клеточно-диагональная матрица. Цепочкой элементарных преобразований приведем каждую клетку к диагональному виду:

$$F_1 \text{ эквивалентна } G_1 - \text{ каноническая диагональная,}$$

$$F_2 \text{ эквивалентна } G_2 - \text{ каноническая диагональная,}$$

....

$$F_s \text{ эквивалентна } G_s - \text{ каноническая диагональная.}$$

Тогда F эквивалентна диагональной матрице

$$G = \mathrm{diag}(G_1, G_2, \ldots, G_s).$$

По рассмотренному выше случаю

$$\{\text{набор элементарных делителей } F\} = \{\text{набор элементарных делителей } G\} =$$

$$= \bigcup_{i=1}^{s} \{ \text{ элементарные делители } G_i\} = \bigcup_{i=1}^{s} \{ \text{ элементарные делители } F_i\}.$$

Теорема доказана.

35.9. Другое доказательство теоремы Жордана. Докажем предварительно такое утверждение.

Л е м м а 3. *Если*

$$A = \begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}$$

– жорданова клетка степени m, то характеристическая матрица $\lambda E - A$ имеет единственный элементарный делитель $(\lambda - \alpha)^m$.

Д о к а з а т е л ь с т в о. Рассмотрим характеристическую матрицу

$$\lambda E - A = \begin{pmatrix} \lambda - \alpha & -1 & & & \mathbf{0} \\ & \lambda - \alpha & -1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & -1 \\ \mathbf{0} & & & & \lambda - \alpha \end{pmatrix}.$$

Чтобы найти инвариантные множители, найдем наибольшие общие делители миноров:

$$D_m(\lambda) = (\lambda - \alpha)^m, \quad D_{m-1}(\lambda) = 1, \quad d_1 d_2 \ldots d_k = D_k.$$

Если вычеркнуть первый столбец и последнюю строку, то этот минор равен ± 1. Значит

$$d_1(\lambda) = d_2(\lambda) = \ldots = d_{m-1}(\lambda) = 1, \quad d_m(\lambda) = (\lambda - \alpha)^m,$$

и наша матрица эквивалентна матрице

$$\begin{pmatrix} 1 & & & & \mathbf{0} \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & (\lambda-\alpha)^m \end{pmatrix}.$$

По определению элементарного делителя $(\lambda-\alpha)^m$ – единственный элементарный делитель матрицы $\lambda E - A$. Лемма доказана.

Переходим непосредственно к доказательству теоремы Жордана. Докажем вначале существование.

Пусть $A \in M_n(P)$ и P содержит все характеристические корни матрицы A. Рассмотрим характеристическую матрицу $\lambda E - A$ и найдем все ее элементарные делители:

$$(\lambda-\alpha_i)^{n_{ij}}.$$

По каждому из этих многочленов напишем клетку жордана:

$$J_{ij} = \begin{pmatrix} \alpha_i & 1 & & & \mathbf{0} \\ & \alpha_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha_i \end{pmatrix}$$

– клетка жордана степени n_{ij}. Из этих клеток составим матрицу жордана J. Утверждается, что эта матрица искомая. Действительно, по предыдущей теореме и лемме множество элементарных делителей характеристической матрицы $\lambda E - A$ совпадает с множеством элементарных делителей матрицы $\lambda E - J$. Матрица

$$\begin{pmatrix} \lambda E - J_{11} & & & & \mathbf{0} \\ & \lambda E - J_{12} & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & \lambda E - J_{rs} \end{pmatrix}$$

является клеточно-диагональной и каждая клетка дает по одному элементарному делителю $(\lambda-\alpha_i)^{n_{ij}}$. Поэтому канонический вид матрицы $\lambda E - A$ совпадает с каноническим видом матрицы $\lambda E - J$, т. е. $\lambda E - A$ эквивалентна $\lambda E - J$. Как мы знаем, если матрицы унимодулярны и эквивалентны, то они скалярно эквивалентны, т. е.

$$\lambda E - A = U(\lambda E - J)V,$$

где U и V не зависят от λ. Сравнивая матрицы при одинаковых степенях λ в левой и правой частях, получаем

$$E = UV, \quad -A = -UJV.$$

Следовательно,

$$A = V^{-1}JV.$$

и мы доказали, что матрица A подобна жордановой матрице.

Докажем единственность. Надо доказать, что если жорданова матрица J_1 подобна жордановой матрице J_2, то они совпадают с точностью до порядка жордановых клеток. Из подобия жордановых матриц следует, что

$$\lambda E - J_1 \text{ эквивалентна } \lambda E - J_2.$$

Отсюда

$$\{\text{элементарные делители } \lambda E - J_1\} = \{\text{элементарные делители } \lambda E - J_2\},$$

а потому

$$\{\text{клетки жордана в } J_1\} = \{\text{клетки жордана в } J_2\}.$$

Теорема доказана.

С л е д с т в и е. *Матрица, у которой все характеристические корни лежат в основном поле, подобна диагональной тогда и только тогда, когда все элементарные делители ее характеристической матрицы линейны, т. е. имеют вид* $\lambda - \alpha_i$.

§ 36. Функции от матриц

36.1. Многочлены от матриц. В этом параграфе будем рассматривать случай, когда поле является полем комплексных чисел. Как мы знаем, в этом случае всякая матрица приводится к жордановой форме.

Т е о р е м а 1. *Если* $A \in M_n(\mathbb{C})$, $f \in \mathbb{C}[\lambda]$ *и* A *подобна матрице жордана:* $A = T^{-1}JT$, *где*

$$J = J_1 \oplus J_2 \oplus \ldots \oplus J_s = \begin{pmatrix} J_1 & & & & \mathbf{0} \\ & J_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & J_s \end{pmatrix}$$

то

1) $f(A) = T^{-1}f(J)T$;

2) $f(J) = f(J_1) \oplus f(J_2) \oplus \ldots \oplus f(J_s)$;

3) значение f от жордановой клетки степени m определяется равенством

$$f(\begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}) = \begin{pmatrix} f(\alpha) & \frac{f'(\alpha)}{1!} & \ldots & \frac{f^{(m-1)}(\alpha)}{(m-1)!} \\ 0 & f(\alpha) & \ldots & \frac{f^{(m-2)}(\alpha)}{(m-2)!} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & f(\alpha) \end{pmatrix}.$$

Д о к а з а т е л ь с т в о. 1) Если

$$f(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \ldots + a_m,$$

то

$$f(A) = f(T^{-1}JT) = a_0(T^{-1}JT)^m + a_1(T^{-1}JT)^{m-1} + \ldots + a_mE =$$
$$= a_0(T^{-1}J^mT) + a_1(T^{-1}J^{m-1}T) + \ldots + a_mE = T^{-1}f(J)T.$$

2) Если $B = B_1 \oplus B_2 \oplus \ldots \oplus B_s$ и $C = C_1 \oplus C_2 \oplus \ldots \oplus C_s$ – две клеточно-диагональные матрицы и размеры клетки B_i совпадают с размерами клетки C_i для всех $i = 1, 2, \ldots, s$, то

$$B + C = (B_1 + C_1) \oplus (B_2 + C_2) \oplus \ldots \oplus (B_s + C_s),$$

$$BC = B_1C_1 \oplus B_2C_2 \oplus \ldots \oplus B_sC_s,$$

$$\alpha B = \alpha B_1 \oplus \alpha B_2 \oplus \ldots \oplus \alpha B_s.$$

Отсюда

$$f(J) = f(J_1) \oplus f(J_2) \oplus \ldots \oplus f(J_s).$$

3) По формуле Тейлора:

$$f(\lambda) = \sum_{k=0}^{m} \frac{f^{(k)}(\alpha)}{k!}(\lambda - \alpha)^k.$$

Рассмотрим клетку жордана

$$B = \begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}$$

Для нее
$$f(B)=\sum_{k=0}^{m}\frac{f^{(k)}(\alpha)}{k!}(B-\alpha E)^k.$$
Обозначим
$$C=B-\alpha E=\begin{pmatrix} 0 & 1 & & & \mathbf{0}\\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ \mathbf{0} & & & & 0\end{pmatrix}$$
Остается доказать, что
$$C^k=\begin{pmatrix} 0 & \dots & 1 & & & \mathbf{0}\\ & 0 & \dots & 1 & & \\ & & \ddots & & \ddots & \\ & & & \ddots & & 1\\ & & & & \ddots & \vdots\\ \mathbf{0} & & & & & 0\end{pmatrix}$$
где единицы стоят на k-й диагонали от главной.

Индукция по k. Для $k=1$ утверждение тривиально. Предположим, что формула справедлива для k и рассмотрим $k+1$. Имеем
$$C^k=\sum_{i=1}^{m-k}E_{i,i+k},$$
где E_{ij} – матрица, у которой на месте (i,j) стоит единица, а на всех остальных местах – нули. Для таких матриц справедлива следующая формула умножения:
$$E_{kl}E_{rs}=\begin{cases} E_{ks} & \text{при}\quad l=r,\\ 0 & \text{при}\quad l\neq r.\end{cases}$$
Тогда
$$C^{k+1}=C^kC=\sum_{i=1}^{m-k}E_{i,i+k}\sum_{j=1}^{m-1}E_{j,j+1}=\sum_{i=1}^{m-k-1}E_{i,i+k+1}.$$
Теорема доказана.

36.2. Функции от матриц. Многочлен Лагранжа-Сильвестера. Пусть $f:\mathbb{C}\longrightarrow\mathbb{C}$ – некоторая функция, $A\in M_n(\mathbb{C})$ и $(\lambda-\alpha_i)^{n_{ij}}$ – все элементарные делители характеристической матрицы $\lambda E - A$. Хотим определить $f(A)$, т. е. значение функции f от матрицы A.

Предположим, что выполняется следующее условие:

$$\textit{существуют все производные } f^{(k)}(\alpha_i), \quad k = 0, 1, \dots, n_{ij} - 1, \quad i = 1, 2, \dots, s. \tag{1}$$

Найдем жорданову форму матрицы A:

$$J = J_1 \oplus J_2 \oplus \dots \oplus J_l,$$

где J_i – жордановы клетки, $A = T^{-1}JT$.

О п р е д е л е н и е. Положим

1) $f(A) = T^{-1}f(J)T$,

2) $f(J) = f(J_1) \oplus f(J_2) \oplus \dots \oplus f(J_l)$,

3) $$f\left(\begin{pmatrix} \alpha_i & 1 & & & \mathbf{0} \\ & \alpha_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha_i \end{pmatrix}\right) = \begin{pmatrix} f(\alpha_i) & \frac{f'(\alpha_i)}{1!} & \dots & \frac{f^{(m-1)}(\alpha_i)}{(m-1)!} \\ 0 & f(\alpha_i) & \dots & \frac{f^{(m-2)}(\alpha_i)}{(m-2)!} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & f(\alpha_i) \end{pmatrix},$$

где m – размер рассматриваемой жордановой клетки.

Надо доказать корректность этого определения, т. е. показать, что оно не зависит от T и J.

Т е о р е м а 2. *Пусть для функции* $f : \mathbb{C} \longrightarrow \mathbb{C}$ *и матрицы* $A \in M_n(\mathbb{C})$ *выполнено условие* (1). *Тогда*: 1) *существует многочлен* $p(\lambda) \in \mathbb{C}[\lambda]$ *такой, что*

$$p^{(k)}(\alpha_i) = f^{(k)}(\alpha_i), \quad k = 0, 1, \dots, n_{ij} - 1, \quad i = 1, 2, \dots, s; \tag{2}$$

2) $p(A) = f(A)$ *и поэтому* $f(A)$ *от случайного выбора* J *и* T *не зависит.*

Д о к а з а т е л ь с т в о. Если доказана первая часть теоремы, то из нее следует и вторая часть. Докажем первую часть.

По значениям многочлена мы можем восстановить многочлен, используя формулу Лагранжа, но у нас есть и значения производных. Будем искать многочлен $p(\lambda)$ в виде

$$p(\lambda) = \sum_{i=1}^{s} \left[\beta_{i0} + \beta_{i1}(\lambda - \alpha_i) + \dots + \beta_{i,n_i-1}(\lambda - \alpha_i)^{n_i-1}\right] \cdot$$
$$\cdot(\lambda - \alpha_1)^{n_1} \dots (\lambda - \alpha_{i-1})^{n_{i-1}}(\lambda - \alpha_{i+1})^{n_{i+1}} \dots (\lambda - \alpha_s)^{n_s},$$

где $n_i = \max_j n_{ij}$. Покажем, что коэффициенты β_{ij} можно подобрать так, чтобы выполнялось условие (2). Подставив в нашу формулу значение α_i, получим

$$p(\alpha_i) = \beta_{i0} A_{i0}^{(0)}, \text{ где } A_{i0}^{(0)} = (\alpha_i - \alpha_1)^{n_1} \dots (\alpha_i - \alpha_{i-1})^{n_{i-1}} (\alpha_i - \alpha_{i+1})^{n_{i+1}} \dots (\alpha_i - \alpha_s)^{n_s} \neq 0,$$

т. е. от всей суммы остается только слагаемое с номером i, от выражения в квадратных скобках остается только коэффициент β_{i0}.

Найдем значение производной в точке α_i. Для этого обозначим

$$u(\lambda) = \left[\beta_{i0} + \beta_{i1}(\lambda - \alpha_i) + \ldots + \beta_{i,n_i-1}(\lambda - \alpha_i)^{n_i-1}\right],$$

$$v(\lambda) = (\lambda - \alpha_1)^{n_1} \dots (\lambda - \alpha_{i-1})^{n_{i-1}} (\lambda - \alpha_{i+1})^{n_{i+1}} \dots (\lambda - \alpha_s)^{n_s}.$$

Нетрудно убедиться, что

$$p'(\alpha_i) = u'(\alpha_i)\, v(\alpha_i) + u(\alpha_i)\, v'(\alpha_i) = \beta_{i0} A_{i0}^{(1)} + \beta_{i1} A_{i1}^{(1)}, \text{ где } A_{i1}^{(1)} = A_{i0}^{(0)}.$$

Далее находим вторую производную:

$$(u'v + uv')' = uv'' + 2u'v' + u''v,$$

и, подставляя значение α_i, получим

$$p''(\alpha_i) = u(\alpha_i)v''(\alpha_i) + 2u'(\alpha_i)v'(\alpha_i) + u''(\alpha_i)v(\alpha_i).$$

Аналогичным образом приходим к формулам

$$\begin{array}{ll}
p''(\alpha_i) = \beta_{i0} A_{i0}^{(2)} + \beta_{i1} A_{i1}^{(2)} + 2\beta_{i2} A_{i2}^{(2)}, & \text{где } A_{i2}^{(2)} = A_{i0}^{(0)}, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots \\
p^{(k)}(\alpha_i) = \beta_{i0} A_{i0}^{(k)} + \beta_{i1} A_{i1}^{(k)} + \ldots + k!\beta_{ik} A_{ik}^{(k)}, & \text{где } A_{ik}^{(k)} = A_{i0}^{(0)}, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots \\
p^{(r_i-1)}(\alpha_i) = \beta_{i0} A_{i0}^{(r_i-1)} + \beta_{i1} A_{i1}^{(r_i-1)} + \ldots + (r_i-1)!\beta_{i,r_i-1} A_{i,r_i-1}^{(r_i-1)}, & \text{где } A_{i,r_i-1}^{(r_i-1)} = A_{i0}^{(0)}.
\end{array}$$

Это треугольная система линейных уравнений. Решая ее, найдем β_{ij}. Теорема доказана.

В этой теореме мы интерполировали функцию по значениям многочлена и значениям производной. Построенный многочлен $p(\lambda)$ называется *многочленом Лагранжа-Сильвестера.*

36.3. Ряды от матриц. Рассмотрим ряд

$$a_0 + a_1\lambda + a_2\lambda^2 + \ldots, \quad a_i \in \mathbb{R}.$$

Пусть $A \in M_n(\mathbb{R})$. Можем определить ряд

$$a_0E + a_1A + a_2A^2 + \ldots$$

Его частичной суммой называется матрица

$$f_k(A) = a_0E + a_1A + a_2A^2 + \ldots + a_kA^k.$$

Если существует предел:

$$\lim_{k\to\infty} (f_k(A))_{ij} = b_{ij}, \quad 1 \le i, j \le n,$$

то матрица $B = (b_{ij})$ называется *суммой ряда.*

Если A – жорданова клетка:

$$A = \begin{pmatrix} \alpha & 1 & & & \mathbf{0} \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ \mathbf{0} & & & & \alpha \end{pmatrix}$$

то

$$f_k(A) = \begin{pmatrix} f_k(\alpha) & \frac{f'_k(\alpha)}{1!} & \ldots & \frac{f_k^{(m-1)}(\alpha)}{(m-1)!} \\ 0 & f_k(\alpha) & \ldots & \frac{f_k^{(m-2)}(\alpha)}{(m-2)!} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & f_k(\alpha) \end{pmatrix}.$$

Тогда из теоремы 2 получается

Т е о р е м а 3. *Ряд*

$$\sum_{k=0}^{\infty} a_kA^k$$

сходится тогда и только тогда, когда сходится ряд

$$\sum_{k=0}^{\infty} a_k\lambda^k$$

и все их производные до (n_i-1)-го порядка в точке α_i, где $(\lambda-\alpha_i)^{n_{ij}}$ – элементарные делители матрицы A, $n_i = \max_j n_{ij}$.

П р и м е р. Мы знаем, что

$$\mathrm{e}^{\lambda} = 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots$$

Для матрицы A положим

$$e^A = E + \frac{1}{1!}A + \frac{1}{2!}A^2 + \dots$$

По теореме 3 этот ряд сходится.

У п р а ж н е н и е. Если $AB = BA$, то $e^A \cdot e^B = e^{A+B}$.

Рассмотрим множество матриц

$$\{At \mid t \in \mathbb{R}\}.$$

По упражнению

$$e^{At} \cdot e^{As} = e^{A(t+s)},$$

т. е. отображение $t \longmapsto e^{At}$ переводит сумму в произведение. В частности, множество

$$\{e^{At} \mid t \in \mathbb{R}\}$$

образует подгруппу по умножению. Она называется *однопараметрической подгруппой.* Группу $GL_n(\mathbb{R})$ можно получить как конечное произведение однопараметрических подгрупп (группа Ли).

36.4. Приложение функции от матриц к решению систем линейных дифференциальных уравнений. Рассмотрим простейшее дифференциальное уравнение

$$y' = ay, \quad a \in \mathbb{R},$$

где $y = y(x)$ – неизвестная функция. Решим его методом разделения переменных. Из уравнения

$$\frac{dy}{dx} = ay$$

получим

$$\frac{dy}{y} = a dx.$$

Интегрируя это равенство, найдем общее решение:

$$y = c e^{ax},$$

где c – произвольная постоянная.

Рассмотрим теперь систему обыкновенных дифференциальных уравнений с постоянными коэффициентами:

$$Y' = AY,$$

где
$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad Y' = \begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{pmatrix}, \quad A \in M_n(\mathbb{R}).$$
Тогда легко проверить, что ее решение имеет вид
$$Y = \mathrm{e}^{Ax}C,$$
где e^{Ax} – функция от матрицы Ax, а C – вектор-столбец постоянных:
$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$