§ 11. Приведение матрицы оператора к жордановой нормальной форме

Б.М.Верников

Уральский федеральный университет,
Институт естественных наук и математики,
кафедра алгебры и фундаментальной информатики

11.1. Основной результат: формулировка и начало доказательства

Основным результатом данного параграфа является следующая теорема.

Теорема 11.1

Линейный оператор $\mathcal{A}$ в векторном пространстве V над полем F приводим к жордановой нормальной форме тогда и только тогда, когда характеристический многочлен оператора $\mathcal{A}$ разложим в поле F на линейные множители.

Из этой теоремы и следствия 11.4 из курса «Основы алгебры» вытекает

Следствие 11.1

Если V — векторное пространство над полем $\mathbb{C}$, то любой линейный оператор в V приводим к жордановой нормальной форме. $\square$

Необходимость в теореме 11.1 проверяется легко.

Доказательство необходимости в теореме 11.1. Пусть $\mathcal{A}$ — линейный оператор, матрица которого в некотором базисе жорданова. В частности, эта матрица нижнетреугольна. Обозначим скаляры, стоящие на ее главной диагонали, через $\lambda_1, \lambda_2, \ldots, \lambda_n$ (некоторые из них могут совпадать). Тогда

$$\chi_{\mathcal{A}}(x) = |A - xE| = \begin{vmatrix} \lambda_1 - x & 0 & 0 & \cdots & 0 \\ a_{21} & \lambda_2 - x & 0 & \cdots & 0 \\ a_{31} & a_{32} & \lambda_3 - x & \cdots & 0 \\ \hdotsfor{5} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & \lambda_n - x \end{vmatrix} =$$

$$= (\lambda_1 - x)(\lambda_2 - x)\cdots(\lambda_n - x) =$$

$$= (-1)^n (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n).$$

Таким образом, многочлен $\chi_{\mathcal{A}}(x)$ разложим на линейные множители. $\square$

Достаточность в теореме 11.1 будет доказана позже, после того, как мы докажем все необходимые для этого промежуточные результаты.

11.2. Разложение Фиттинга

Для доказательства теоремы 11.1 нам потребуется установить некоторые свойства линейных операторов, имеющие ненулевое ядро. Зафиксируем произвольный линейный оператор $\mathcal{H}$ в векторном пространстве V над полем F. Очевидно, что для всякого натурального k выполнены включения

$$\operatorname{Ker} \mathcal{H}^k \subseteq \operatorname{Ker} \mathcal{H}^{k+1} \tag{1}$$

и

$$\operatorname{Im} \mathcal{H}^k \supseteq \operatorname{Im} \mathcal{H}^{k+1}. \tag{2}$$

Лемма 11.1

Если $\operatorname{Ker} \mathcal{H} \neq \{\mathbf{0}\}$, *то существует натуральное число* s *такое, что*

$$\begin{aligned} \{\mathbf{0}\} \ &\subset \operatorname{Ker} \mathcal{H} \subset \operatorname{Ker} \mathcal{H}^2 \subset \cdots \subset \operatorname{Ker} \mathcal{H}^s = \\ &= \operatorname{Ker} \mathcal{H}^{s+1} = \cdots = \operatorname{Ker} \mathcal{H}^{s+m} = \cdots \end{aligned} \tag{3}$$

и

$$V \supset \operatorname{Im} \mathcal{H} \supset \operatorname{Im} \mathcal{H}^2 \supset \cdots \supset \operatorname{Im} \mathcal{H}^s = \operatorname{Im} \mathcal{H}^{s+1} = \cdots = \operatorname{Im} \mathcal{H}^{s+m} = \cdots . \tag{4}$$

Доказательство. Положим $\dim V = n$. Включение $\{\mathbf{0}\} \subset \operatorname{Ker} \mathcal{H}$ выполнено по условию. Далее, для всякого натурального k выполнены включения (1) и $\operatorname{Ker} \mathcal{H}^{k+1} \subseteq V$, и потому $\dim \operatorname{Ker} \mathcal{H}^k \leqslant \dim \operatorname{Ker} \mathcal{H}^{k+1} \leqslant n$. Следовательно, существует натуральное s такое, что

$$\{\mathbf{0}\} \subset \operatorname{Ker} \mathcal{H} \subset \operatorname{Ker} \mathcal{H}^2 \subset \cdots \subset \operatorname{Ker} \mathcal{H}^s = \operatorname{Ker} \mathcal{H}^{s+1}.$$

Проверим, что $\operatorname{Ker} \mathcal{H}^s = \operatorname{Ker} \mathcal{H}^{s+m}$ для всякого натурального m.

Предположим, напротив, что

$$\operatorname{Ker} \mathcal{H}^s = \operatorname{Ker} \mathcal{H}^{s+1} = \cdots = \operatorname{Ker} \mathcal{H}^{s+m-1} \subset \operatorname{Ker} \mathcal{H}^{s+m}$$

для некоторого m. Отметим, что $m > 1$, так как

$$\operatorname{Ker} \mathcal{H}^s = \operatorname{Ker} \mathcal{H}^{s+1}. \tag{5}$$

Из включения $\operatorname{Ker} \mathcal{H}^{s+m-1} \subset \operatorname{Ker} \mathcal{H}^{s+m}$ вытекает, что существует вектор $\mathbf{x}$ такой, что $\mathbf{x} \in \operatorname{Ker} \mathcal{H}^{s+m}$, но $\mathbf{x} \notin \operatorname{Ker} \mathcal{H}^{s+m-1}$. В частности, $\mathcal{H}^{s+m}(\mathbf{x}) = \mathbf{0}$. Положим $\mathbf{y} = \mathcal{H}^{m-1}(\mathbf{x})$. Тогда $\mathcal{H}^{s+1}(\mathbf{y}) = \mathcal{H}^{s+m}(\mathbf{x}) = \mathbf{0}$. Из равенства (5) вытекает, что $\mathcal{H}^s(\mathbf{y}) = \mathbf{0}$. Но $\mathcal{H}^s(\mathbf{y}) = \mathcal{H}^{s+m-1}(\mathbf{x}) \neq \mathbf{0}$.
Полученное противоречие доказывает, что выполнены соотношения (3). Следовательно,

$$\dim\{\mathbf{0}\} < \dim \operatorname{Ker} \mathcal{H} < \dim \operatorname{Ker} \mathcal{H}^2 < \cdots < \dim \operatorname{Ker} \mathcal{H}^s =$$
$$= \dim \operatorname{Ker} \mathcal{H}^{s+1} = \cdots = \dim \operatorname{Ker} \mathcal{H}^{s+m} = \cdots$$

Поскольку, в силу теоремы о ранге и дефекте, $\dim \operatorname{Im} \mathcal{H}^k + \dim \operatorname{Ker} \mathcal{H}^k = n$ для любого натурального k, отсюда вытекает, что

$$n > \dim \operatorname{Im} \mathcal{H} > \dim \operatorname{Im} \mathcal{H}^2 > \cdots > \dim \operatorname{Im} \mathcal{H}^s =$$
$$= \dim \operatorname{Im} \mathcal{H}^{s+1} = \cdots = \dim \operatorname{Im} \mathcal{H}^{s+m} = \cdots .$$

Отсюда и из того, что для всякого k выполнено включение (2), вытекают соотношения (4). $\square$

Разложение Фиттинга (1)

Положим $N_{\mathcal{H}} = \operatorname{Ker} \mathcal{H}^s$, $U_{\mathcal{H}} = \operatorname{Im} \mathcal{H}^s$, $\mathcal{N} = \mathcal{H}|_{N_{\mathcal{H}}}$ и $\mathcal{U} = \mathcal{H}|_{U_{\mathcal{H}}}$.

Предложение 11.1

Если $\operatorname{Ker} \mathcal{H} \neq \{\mathbf{0}\}$, *то подпространства* $N_{\mathcal{H}}$ *и* $U_{\mathcal{H}}$ *инвариантны относительно* $\mathcal{H}$, $V = N_{\mathcal{H}} \oplus U_{\mathcal{H}}$, $\mathcal{N}$ *— нильпотентный оператор, а* $\mathcal{U}$ *— автоморфизм.*

Доказательство. Пусть $\mathbf{x} \in \operatorname{Im} \mathcal{H}^s$. Тогда $\mathbf{x} = \mathcal{H}^s(\mathbf{y})$ для некоторого $\mathbf{y} \in V$. Следовательно,

$$\mathcal{H}(\mathbf{x}) = \mathcal{H}\big(\mathcal{H}^s(\mathbf{y})\big) = \mathcal{H}^{s+1}(\mathbf{y}) = \mathcal{H}^s\big(\mathcal{H}(\mathbf{y})\big),$$

и потому $\mathcal{H}(\mathbf{x}) \in \operatorname{Im} \mathcal{H}^s$. Таким образом, подпространство $\operatorname{Im} \mathcal{H}^s$ инвариантно относительно $\mathcal{H}$.

Далее, пусть $\mathbf{x} \in \operatorname{Ker} \mathcal{H}^s$, т. е. $\mathcal{H}^s(\mathbf{x}) = \mathbf{0}$. Тогда

$$\mathcal{H}^s\big(\mathcal{H}(\mathbf{x})\big) = \mathcal{H}^{s+1}(\mathbf{x}) = \mathcal{H}\big(\mathcal{H}^s(\mathbf{x})\big) = \mathcal{H}(\mathbf{0}) = \mathbf{0},$$

и потому $\mathcal{H}(\mathbf{x}) \in \operatorname{Ker} \mathcal{H}^s$. Таким образом, подпространство $\operatorname{Ker} \mathcal{H}^s$ также инвариантно относительно $\mathcal{H}$.

Разложение Фиттинга (2)

Из определения подпространства $N_{\mathcal{H}}$ вытекает, что $\mathcal{N}^s = \mathcal{O}$, и потому оператор $\mathcal{N}$ нильпотентен. Далее,

$$\mathcal{H}(U_{\mathcal{H}}) = \mathcal{H}(\operatorname{Im}\mathcal{H}^s) = \operatorname{Im}\mathcal{H}^{s+1} = \operatorname{Im}\mathcal{H}^s = U_{\mathcal{H}}.$$

Поэтому из предложения 8.1 вытекает, что оператор $\mathcal{U} = \mathcal{H}|_{U_{\mathcal{H}}}$ является автоморфизмом.

Докажем, что $N_{\mathcal{H}} \cap U_{\mathcal{H}} = \{\mathbf{0}\}$. Пусть $\mathbf{x} \in N_{\mathcal{H}} \cap U_{\mathcal{H}}$. Из того, что $\mathbf{x} \in N_{\mathcal{H}}$, вытекает, что $\mathcal{H}^s(\mathbf{x}) = \mathbf{0}$, а из того, что $\mathbf{x} \in U_{\mathcal{H}}$, следует, что $\mathbf{x} = \mathcal{H}^s(\mathbf{y})$ для некоторого вектора $\mathbf{y}$. Следовательно, $\mathcal{H}^{2s}(\mathbf{y}) = \mathcal{H}^s\big(\mathcal{H}^s(\mathbf{y})\big) = \mathcal{H}^s(\mathbf{x}) = \mathbf{0}$. Таким образом, $\mathbf{y} \in \operatorname{Ker}\mathcal{H}^{2s} = \operatorname{Ker}\mathcal{H}^s$, и значит $\mathbf{x} = \mathcal{H}^s(\mathbf{y}) = \mathbf{0}$.

Чтобы доказать равенство $V = N_{\mathcal{H}} \oplus U_{\mathcal{H}}$ (и тем самым завершить доказательство предложения), осталось проверить, что $V = N_{\mathcal{H}} + U_{\mathcal{H}}$. В самом деле, учитывая, что $N_{\mathcal{H}} \cap U_{\mathcal{H}} = \{\mathbf{0}\}$, имеем

$$\dim(N_{\mathcal{H}} + U_{\mathcal{H}}) = \dim N_{\mathcal{H}} + \dim U_{\mathcal{H}} - \dim(N_{\mathcal{H}} \cap U_{\mathcal{H}}) = \dim N_{\mathcal{H}} + \dim U_{\mathcal{H}}.$$

Из теоремы о ранге и дефекте вытекает, что $\dim N_{\mathcal{H}} + \dim U_{\mathcal{H}} = \dim V$. Таким образом, $\dim(N_{\mathcal{H}} + U_{\mathcal{H}}) = \dim V$, и потому $N_{\mathcal{H}} + U_{\mathcal{H}} = V$. Предложение доказано. $\square$

Равенство $V = N_{\mathcal{H}} \oplus U_{\mathcal{H}}$ называется *разложением Фиттинга* пространства V относительно линейного оператора $\mathcal{H}$.

11.3. Корневое разложение. Доказательство основного результата

С этого места и до конца доказательства теоремы 11.1 через $\mathcal{A}$ обозначается линейный оператор в векторном пространстве V над полем F такой, что многочлен $\chi_{\mathcal{A}}(x)$ разложим в поле F на линейные множители. Будем считать, что

$$\chi_{\mathcal{A}}(x) = (-1)^n (x-\lambda_1)^{k_1}(x-\lambda_2)^{k_2}\cdots(x-\lambda_m)^{k_m}, \qquad (6)$$

где $n = \lambda_1 + \lambda_2 + \cdots + \lambda_m = \dim V$. В частности, $\lambda_1, \lambda_2, \ldots, \lambda_m$ — все попарно различные собственные значения оператора $\mathcal{A}$.

Для всякого $i = 1, 2, \ldots, m$ положим $\mathcal{A}_i = \mathcal{A} - \lambda_i \mathcal{E}$.

Замечание 11.1

Пусть $i \in \{1, 2, \ldots, m\}$. Тогда $\operatorname{Ker} \mathcal{A}_i \neq \{\mathbf{0}\}$.

Доказательство. Если $\mathbf{x}$ — собственный вектор оператора $\mathcal{A}$, относящийся к собственному значению λ_i, то $\mathbf{x} \neq \mathbf{0}$ и

$$\mathcal{A}_i(\mathbf{x}) = (\mathcal{A} - \lambda_i\mathcal{E})(\mathbf{x}) = \mathcal{A}(\mathbf{x}) - \lambda_i\mathcal{E}(\mathbf{x}) = \lambda_i\mathbf{x} - \lambda_i\mathbf{x} = \mathbf{0}.$$

Мы видим, что $\mathbf{x} \in \operatorname{Ker} \mathcal{A}_i$, и потому $\operatorname{Ker} \mathcal{A}_i \neq \{\mathbf{0}\}$. □

Из леммы 11.1 и замечания 11.1 вытекает, что для всякого $i = 1, 2, \ldots, m$ существует число s_i такое, что

$$\{\mathbf{0}\} \subset \operatorname{Ker} \mathcal{A}_i \subset \operatorname{Ker} \mathcal{A}_i^2 \subset \cdots \subset \operatorname{Ker} \mathcal{A}_i^{s_i} = \operatorname{Ker} \mathcal{A}_i^{s_i+1} = \cdots = \operatorname{Ker} \mathcal{A}_i^{s_i+m} = \cdots \quad \text{и}$$

$$V \supset \operatorname{Im} \mathcal{A}_i \supset \operatorname{Im} \mathcal{A}_i^2 \supset \cdots \supset \operatorname{Im} \mathcal{A}_i^{s_i} = \operatorname{Im} \mathcal{A}_i^{s_i+1} = \cdots = \operatorname{Im} \mathcal{A}_i^{s_i+m} = \cdots.$$

Положим $N_i = \operatorname{Ker} \mathcal{A}_i^{s_i}$ и $U_i = \operatorname{Im} \mathcal{A}_i^{s_i}$. Ограничения оператора $\mathcal{A}_i$ на подпространства N_i и U_i будем обозначать через $\mathcal{N}_i$ и $\mathcal{U}_i$ соответственно.

Определение

Подпространство N_i называется *корневым подпространством* пространства V, соответствующим собственному значению λ_i.

Свойства подпространств N_i и U_i. Теорема о корневом разложении

Из предложения 11.1 вытекает следующий факт, который далее будет использоваться, как правило, без явных ссылок.

Следствие 11.2

Для всякого $i = 1, 2, \ldots, m$ справедливы следующие утверждения:

1) *подпространства N_i и U_i инвариантны относительно оператора $\mathcal{A}_i$;*
2) $V = N_i \oplus U_i$;
3) $\mathcal{N}_i$ — *нильпотентный оператор;*
4) $\mathcal{U}_i$ — *автоморфизм.* □

Оказывается, что справедлива следующая теорема, из которой, как мы увидим ниже, уже легко вытекает достаточность в теореме 11.1.

Теорема 11.2

$V = N_1 \oplus N_2 \oplus \cdots \oplus N_m$.

Равенство $V = N_1 \oplus N_2 \oplus \cdots \oplus N_m$ называется *корневым разложением* пространства V относительно оператора $\mathcal{A}$.

Отметим, что теорема 11.2 неоднократно пригодится нам в последующих параграфах.

Для того, чтобы доказать теорему 11.2, нам понадобятся еще две леммы.

Лемма 11.2

Пусть $i, j \in \{1, 2, \ldots, m\}$ и $i \neq j$. Тогда $N_i \subseteq U_j$.

Доказательство. Заметим, что

$$\mathcal{A}_j = \mathcal{A} - \lambda_j \mathcal{E} = \mathcal{A} - \lambda_i \mathcal{E} + \lambda_i \mathcal{E} - \lambda_j \mathcal{E} = \mathcal{A}_i + (\lambda_i - \lambda_j)\mathcal{E}.$$

Если $\mathbf{x} \in N_i$, то $\mathcal{A}_i(\mathbf{x}) \in N_i$ (поскольку N_i инвариантно относительно $\mathcal{A}_i$) и $(\lambda_i - \lambda_j)\mathbf{x} \in N_i$. Следовательно, в этом случае

$$\mathcal{A}_j(\mathbf{x}) = \big(\mathcal{A}_i + (\lambda_i - \lambda_j)\mathcal{E}\big)(\mathbf{x}) = \mathcal{A}_i(\mathbf{x}) + (\lambda_i - \lambda_j)\mathbf{x} \in N_i.$$

Мы видим, что N_i инвариантно относительно $\mathcal{A}_j$. Следовательно, ограничение оператора $\mathcal{A}_j$ на подпространство N_i можно рассматривать как линейный оператор в N_i. Обозначим этот оператор через $\mathcal{A}_{ji}$.

Проверим, что $\operatorname{Ker} \mathcal{A}_{ji} = \{\mathbf{0}\}$. В самом деле, пусть $\mathbf{v} \in \operatorname{Ker} \mathcal{A}_{ji}$. Это означает, что $\mathbf{v} \in \operatorname{Ker} \mathcal{A}_j$ и $\mathbf{v} \in N_i$. Используя первое из этих двух включений, имеем

$$\mathbf{0} = \mathcal{A}_j(\mathbf{v}) = (\mathcal{A} - \lambda_j \mathcal{E})(\mathbf{v}) = \mathcal{A}(\mathbf{v}) - \lambda_j \mathbf{v},$$

откуда $\mathcal{A}(\mathbf{v}) = \lambda_j \mathbf{v}$.

N_i содержится в U_j (2)

Следовательно,

$$\mathcal{A}_i(\mathbf{v}) = (\mathcal{A} - \lambda_i\mathcal{E})(\mathbf{v}) = \mathcal{A}(\mathbf{v}) - \lambda_i\mathbf{v} = \lambda_j\mathbf{v} - \lambda_i\mathbf{v} = (\lambda_j - \lambda_i)\mathbf{v}.$$

Отсюда вытекает, что $\mathcal{A}_i^{s_i}(\mathbf{v}) = (\lambda_j - \lambda_i)^{s_i}\mathbf{v}$. Учитывая, что $\mathbf{v} \in N_i = \operatorname{Ker}\mathcal{A}_i^{s_i}$, имеем $(\lambda_j - \lambda_i)^{s_i}\mathbf{v} = \mathbf{0}$. Поскольку $\lambda_j \neq \lambda_i$, заключаем, что $(\lambda_j - \lambda_i)^{s_i} \neq 0$, и потому $\mathbf{v} = \mathbf{0}$. Мы доказали, что $\operatorname{Ker}\mathcal{A}_{ji} = \{\mathbf{0}\}$. В силу теоремы о ранге и дефекте получаем, что $\operatorname{Im}\mathcal{A}_{ji} = N_i$. Из этого равенства и инвариантности подпространства N_i относительно оператора $\mathcal{A}_j$ вытекает, что $\mathcal{A}_j(N_i) = N_i$. Но тогда

$$N_i = \mathcal{A}_j(N_i) = \mathcal{A}_j\big(\mathcal{A}_j(N_i)\big) = \mathcal{A}_j^2(N_i) = \cdots = \mathcal{A}_j^{s_j}(N_i) \subseteq \mathcal{A}_j^{s_j}(V) = \operatorname{Im}\mathcal{A}_j^{s_j} = U_j.$$

Лемма 11.2 доказана. $\square$

Напомним, что для всякого $i = 1, 2, \ldots, m$ мы обозначаем через k_i кратность скаляра λ_i как корня многочлена $\chi_{\mathcal{A}}(x)$.

Лемма 11.3

Пусть $i \in \{1, 2, \ldots, m\}$. *Тогда* $\dim N_i = k_i$.

Доказательство. Положим $d_i = \dim N_i$. Пусть A — матрица оператора $\mathcal{A}$ в некотором базисе. Тогда оператор $\mathcal{A}_i$ имеет в том же базисе матрицу $A - \lambda_i E$. Поскольку подпространства N_i и U_i инвариантны относительно $\mathcal{A}_i$ и $V = N_i \oplus U_i$, из п. 2) предложения 10.1 вытекает, что $\chi_{\mathcal{A}_i}(x) = \chi_{\mathcal{N}_i}(x) \cdot \chi_{\mathcal{U}_i}(x)$. Это равенство выполнено при любом значении x. В частности, $\chi_{\mathcal{A}_i}(x - \lambda_i) = \chi_{\mathcal{N}_i}(x - \lambda_i) \cdot \chi_{\mathcal{U}_i}(x - \lambda_i)$. В силу предложения 10.4 $\chi_{\mathcal{N}_i}(x) = (-1)^{d_i} x^{d_i}$, и потому $\chi_{\mathcal{N}_i}(x - \lambda_i) = (-1)^{d_i}(x - \lambda_i)^{d_i}$. Объединяя сказанное, получаем, что

$$\chi_{\mathcal{A}}(x) = |A - xE| = \left|(A - \lambda_i E) - (x - \lambda_i)E\right| = \chi_{\mathcal{A}_i}(x - \lambda_i) =$$

$$= \chi_{\mathcal{N}_i}(x - \lambda_i) \cdot \chi_{\mathcal{U}_i}(x - \lambda_i) = (-1)^{d_i}(x - \lambda_i)^{d_i} \cdot \chi_{\mathcal{U}_i}(x - \lambda_i).$$

Размерность пространства N_i (2)

Мы доказали, что многочлен $\chi_{\mathcal{A}}(x)$ делится на $(x-\lambda_i)^{d_i}$. Следовательно, $d_i \leqslant k_i$. Чтобы доказать, что $d_i = k_i$, осталось проверить, что $\chi_{\mathcal{A}}(x)$ не делится на $(x-\lambda_i)^{d_i+1}$. В силу сказанного выше это равносильно тому, что скаляр λ_i не является корнем многочлена $\chi_{\mathcal{U}_i}(x-\lambda_i)$, т.е. тому, что $\chi_{\mathcal{U}_i}(\lambda_i-\lambda_i) = \chi_{\mathcal{U}_i}(0) \neq 0$. Предположим, напротив, что $\chi_{\mathcal{U}_i}(0) = 0$. Обозначим матрицу оператора $\mathcal{U}_i$ в некотором базисе через M_i. Равенство $\chi_{\mathcal{U}_i}(0) = 0$ означает, что $|M_i - 0 \cdot E| = |M_i| = 0$. Но оператор $\mathcal{U}_i$ является автоморфизмом. Поэтому из предложения 8.1 вытекает, что его матрица в любом базисе невырожденна. Следовательно, $|M_i| \neq 0$. Полученное противоречие завершает доказательство. $\square$

Следствие о простом корне характеристического многочлена

Из леммы 11.3 вытекает следующий факт, который пригодится нам в дальнейшем.

Следствие 11.3

Пусть $i \in \{1, 2, \ldots, m\}$. Если λ_i — простой корень многочлена $\chi_{\mathcal{A}}(x)$, то корневое подпространство N_i одномерно, а в качестве его базиса можно взять произвольный собственный вектор оператора $\mathcal{A}$, соответствующий собственному значению λ_i.

Доказательство. Равенство $\dim N_i = 1$ непосредственно вытекает из леммы 11.3. Пусть $\mathbf{p}_i$ — собственный вектор оператора $\mathcal{A}$, соответствующий собственному значению λ_i. Тогда $\mathbf{p}_i \neq \mathbf{0}$ и, как видно из доказательства замечания 11.1, $\mathbf{p}_i \in \operatorname{Ker} \mathcal{A}_i$. Поскольку $\operatorname{Ker} \mathcal{A}_i \subseteq \operatorname{Ker} \mathcal{A}_i^{s_i} = N_i$, мы получаем, что $\mathbf{p}_i \in N_i \setminus \{\mathbf{0}\}$. Поскольку одномерное пространство порождается любым своим ненулевым вектором, получаем, что $\{\mathbf{p}_i\}$ — базис пространства N_i. $\square$

Теперь мы можем быстро завершить

Доказательство теоремы 11.2. Требуется доказать, что $V = N_1 \oplus N_2 \oplus \cdots \oplus N_m$. Пусть $1 \leqslant i \leqslant m-1$. В силу леммы 11.2 $N_{i+1}, \ldots, N_m \subseteq U_i$, и потому $N_{i+1} + \cdots + N_m \subseteq U_i$. Поскольку $N_i \cap U_i = \{\mathbf{0}\}$, это означает, что $N_i \cap (N_{i+1} + \cdots + N_m) = \{\mathbf{0}\}$, и потому $N_i + (N_{i+1} + \cdots + N_m) = N_i \oplus (N_{i+1} + \cdots + N_m)$. Следовательно,

$$N_1 + N_2 + \cdots + N_m = N_1 \oplus (N_2 + \cdots + N_m) =$$
$$= N_1 \oplus N_2 \oplus (N_3 + \cdots + N_m) = \cdots = N_1 \oplus N_2 \oplus \cdots \oplus N_m.$$

Поскольку многочлен $\chi_{\mathcal{A}}(x)$ разложим в поле F на линейные множители, выполнено равенство $k_1 + k_2 + \cdots + k_m = \deg \chi_{\mathcal{A}}(x) = \dim V$. Учитывая сказанное в предыдущем абзаце и лемму 11.3, имеем

$$\dim(N_1 + N_2 + \cdots + N_m) = \dim N_1 + \dim N_2 + \cdots + \dim N_m =$$
$$= k_1 + k_2 + \cdots + k_m = \dim V.$$

Итак, $\dim V = \dim(N_1 + N_2 + \cdots + N_m)$, и потому $V = N_1 + N_2 + \cdots + N_m = N_1 \oplus N_2 \oplus \cdots \oplus N_m$. $\square$

Теперь все готово к тому, чтобы быстро завершить доказательство теоремы 11.1. Напомним, что необходимость в этой теореме была доказана в начале параграфа.

Доказательство достаточности в теореме 11.1. Пусть $i \in \{1, 2, \dots, m\}$. В силу теоремы 10.1 в пространстве N_i существует базис P_i, в котором матрица оператора $\mathcal{N}_i$ жорданова, причем на главных диагоналях всех ее клеток Жордана которой стоит 0. Обозначим эту матрицу через J_i, а матрицу оператора $\mathcal{A}|_{N_i}$ в базисе P_i — через A_i. Поскольку $\mathcal{N}_i = \mathcal{A}|_{N_i} - \lambda_i\mathcal{E}$, имеем $J_i = A_i - \lambda_i E$, и потому $A_i = J_i + \lambda_i E$. Ясно, что A_i — клеточно-диагональная матрица, все диагональные клетки которой имеют вид

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} + \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda_i & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda_i & \cdots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \cdots & 1 & \lambda_i \end{pmatrix},$$

т. е. являются клетками Жордана. Следовательно, матрица A_i жорданова.

Завершение доказательства основного результата (2)

Пусть $\mathbf{x} \in N_i$. Учитывая, что $\mathcal{A}_i(\mathbf{x}) \in N_i$ (поскольку N_i инвариантно относительно $\mathcal{A}_i$) и $\lambda_i\mathbf{x} \in N_i$, получаем, что

$$\mathcal{A}(\mathbf{x}) = (\mathcal{A}_i + \lambda_i\mathcal{E})(\mathbf{x}) = \mathcal{A}_i(\mathbf{x}) + \lambda_i\mathbf{x} \in N_i.$$

Это означает, что подпространство N_i (при любом $i = 1, 2, \ldots, m$) инвариантно относительно оператора $\mathcal{A}$.

Положим $P = P_1 \cup P_2 \cup \cdots \cup P_m$. Из теоремы 11.2 и замечания 3.6 вытекает, что P — базис пространства V. Поскольку подпространства $N_1, N_2, \ldots, N_m$ инвариантны относительно $\mathcal{A}$, из теоремы 11.2 и п. 1) предложения 10.1 вытекает, что матрица оператора $\mathcal{A}$ в базисе P имеет вид

$$\begin{pmatrix} A_1 & O & O & \cdots & O \\ O & A_2 & O & \cdots & O \\ O & O & A_3 & \cdots & O \\ \hdotsfor{5} \\ O & O & O & \cdots & A_m \end{pmatrix}.$$

Ясно, что эта матрица жорданова.

Теорема 11.1 полностью доказана. $\square$

11.4. Алгоритм приведения матрицы оператора к жордановой нормальной форме

Перейдем к алгоритму приведения матрицы оператора к жордановой нормальной форме, найденному В. А. Чуркиным в 1991 г. Отметим, что алгоритм 10.1 является частным случаем данного алгоритма. Прежде чем приступать к подробному изложению алгоритма, которое займет шесть слайдов, приведем его краткую общую схему.

Пусть $\mathcal{A}$ — линейный оператор в пространстве V, удовлетворяющий условию теоремы 11.1, а $\lambda_1, \lambda_2, \ldots, \lambda_m$ — его всевозможные попарно различные собственные значения. Алгоритм перебирает эти собственные значения и для всякого $i = 1, 2, \ldots, m$ находит набор векторов P_i, который является жордановым базисом корневого подпространства N_i, и матрицу J_i, которая является матрицей оператора $\mathcal{A}|_{N_i}$ в базисе P_i (и, в частности, жорданова).

Алгоритм приведения матрицы оператора к жордановой нормальной форме: краткая схема (2)

Набор векторов P_i и матрица J_i ищутся одним из трех способов, соответствующих следующим трем случаям:

а) λ_i — кратный корень многочлена $\chi_{\mathcal{A}}(x)$ и $i < m$ (шаг 1 алгоритма);

б) λ_i — кратный корень многочлена $\chi_{\mathcal{A}}(x)$ и $i = m$ (шаг 2);

в) λ_i — простой корень многочлена $\chi_{\mathcal{A}}(x)$ (шаг 3).

После рассмотрения всех собственных значений на шаге 4 алгоритма легко выписывается ответ.

На следующем слайде приведена блок-схема обсуждаемого алгоритма. Подразумевается, что собственные значения оператора $\mathcal{A}$ упорядочены так, что сначала идут все кратные корни многочлена $\chi_{\mathcal{A}}(x)$, а потом — все его простые корни.

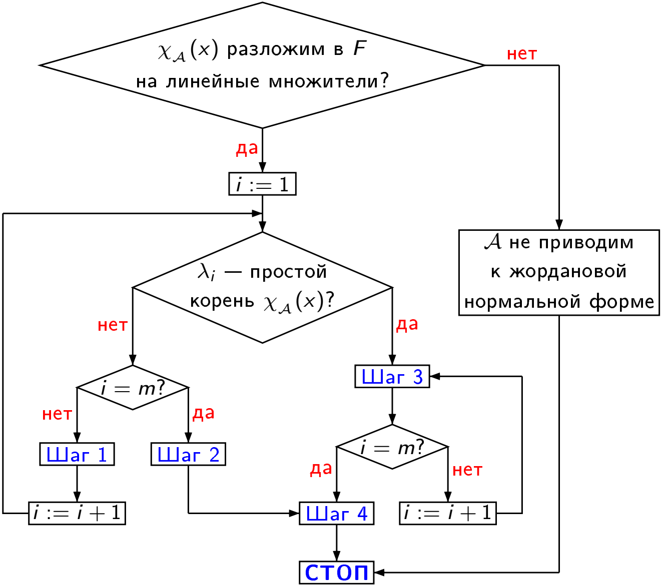

Рис. 1. Блок-схема алгоритма 11.1

Алгоритм приведения матрицы оператора к жордановой нормальной форме (1)

Перейдем к подробному изложению алгоритма.

Алгоритм 11.1 (алгоритм приведения матрицы оператора к жордановой нормальной форме), часть 1

Дан линейный оператор $\mathcal{A}$ в n-мерном векторном пространстве V над полем F. Предполагается известной матрица A этого оператора в некотором базисе.

Шаг 0. Находим многочлен $\chi_{\mathcal{A}}(x)$ и его разложение на неприводимые множители в поле F. Если не все из этих неприводимых множителей линейны, делаем вывод, что оператор не приводим к жордановой нормальной форме, и завершаем работу алгоритма. В противном случае считаем, что выполнено равенство (6). В частности, собственными значениями оператора $\mathcal{A}$ являются скаляры $\lambda_1, \lambda_2, \ldots, \lambda_m$ и только они. Пусть, как и ранее, $\mathcal{A}_i = \mathcal{A} - \lambda_i \mathcal{E}$, $A_i = A - \lambda_i E$, N_i — корневое подпространство пространства V, соответствующее собственному значению λ_i, а $\mathcal{N}_i$ — ограничение оператора $\mathcal{A}_i$ на N_i $(i = 1, 2, \ldots, m)$.

Упорядочиваем собственные значения оператора $\mathcal{A}$ так, чтобы сначала шли все кратные корни многочлена $\chi_{\mathcal{A}}(x)$, а потом — все его простые корни.

Алгоритм приведения матрицы оператора к жордановой нормальной форме (2)

Алгоритм 11.1 (алгоритм приведения матрицы оператора к жордановой нормальной форме), часть 2

Полагаем $i = 1$. Если λ_i — простой корень многочлена $\chi_{\mathcal{A}}(x)$, переходим к шагу 3. Если же λ_i — кратный корень многочлена $\chi_{\mathcal{A}}(x)$, полагаем $X_i = E$ и переходим при $i < m$ к шагу 1, а при $i = m$ — к шагу 2.

Шаг 1. Записываем матрицу $(X_i \mid X_i A_i^\top)$. С помощью элементарных преобразований всей матрицы приводим ее к виду $(B_{1,1} \mid B_{1,2})$, где $B_{1,2}$ — ступенчатая матрица[1]. Затем записываем матрицу $(B_{1,1} \mid B_{1,2} \mid B_{1,2} A_i^\top)$ и с помощью элементарных преобразований всей матрицы приводим ее к виду $(B_{2,1} \mid B_{2,2} \mid B_{2,3})$, где $B_{2,3}$ — ступенчатая матрица. Если $r(B_{1,2}) \neq r(B_{2,3})$, записываем матрицу $(B_{2,1} \mid B_{2,2} \mid B_{2,3} \mid B_{2,3} A_i^\top)$ и с помощью элементарных преобразований всей матрицы приводим ее к виду $(B_{3,1} \mid B_{3,2} \mid B_{3,3} \mid B_{3,4})$, где $B_{3,4}$ — ступенчатая матрица. Продолжаем этот процесс до тех пор, пока на каком-то шаге не окажется, что ранг крайне правого (ступенчатого) блока полученной матрицы равен рангу крайне правого (ступенчатого) блока матрицы, полученной на предыдущем шаге.

[1] Здесь и ниже в описании алгоритма вертикальные линии внутри матрицы делят ее на блоки размера $n \times n$.

Алгоритм приведения матрицы оператора к жордановой нормальной форме (3)

Алгоритм 11.1 (алгоритм приведения матрицы оператора к жордановой нормальной форме), часть 3

Предположим, что на предпоследнем шаге была построена матрица $(B_{k,1} \mid B_{k,2} \mid \cdots \mid B_{k,k+1})$, а на последнем — матрица $(B_{k+1,1} \mid B_{k+1,2} \mid \cdots \mid B_{k+1,k+2})$ и $r(B_{k,k+1}) = r(B_{k+1,k+2})$. В матрице $(B_{k,1} \mid B_{k,2} \mid \cdots \mid B_{k,k})$ рассматриваем ненулевые строки, продолжение которых в матрице $B_{k,k+1}$ является нулевой строкой. Каждая такая ненулевая строка делится на подстроки, расположенные внутри матриц $B_{k,1}$, $B_{k,2}$, ..., $B_{k,k}$. Рассматриваем эти подстроки как векторы из пространства F^n. Вычеркиваем из каждой строки нулевые векторы, а оставшиеся векторы выравниваем по правому краю.

Получилась жорданова таблица. «Сжимаем» эту жорданову таблицу до тех пор, пока оставшиеся в ней векторы не станут линейно независимыми (способ «сжатия» указан в доказательстве предложения 10.2). Обозначаем через P_i упорядоченный набор векторов из полученной жордановой таблицы (сначала перечисляются все векторы из первой строки таблицы, слева направо, потом, в том же порядке, — векторы из второй строки таблицы и т. д.).

Алгоритм приведения матрицы оператора к жордановой нормальной форме (4)

Алгоритм 11.1 (алгоритм приведения матрицы оператора к жордановой нормальной форме), часть 4

Обозначаем через J_i клеточно-диагональную матрицу, в которой все диагональные клетки являются клетками Жордана со скаляром λ_i на главной диагонали, число клеток равно числу строк в полученной жордановой таблице, а порядки клеток равны длинам этих строк, причем сначала стоит клетка, соответствующая первой строке таблицы, потом — второй строке и т. д.

Увеличиваем значение i на единицу. Если λ_i — простой корень многочлена $\chi_{\mathcal{A}}(x)$, переходим к шагу 3. Если же λ_i — кратный корень указанного многочлена, полагаем $X_i = B_{k,k+1}$ и при $i < m$ возвращаемся в начало шага 1, а при $i = m$ переходим к шагу 2.

Шаг 2. Записываем матрицу $(X_m \mid X_m A_m^\top)$ и с помощью элементарных преобразований приводим ее к виду $(B_{1,1} \mid B_{1,2})$, где $B_{1,2}$ — ступенчатая матрица. Если $B_{1,2} A_m^\top \neq O$, записываем матрицу $(B_{1,1} \mid B_{1,2} \mid B_{1,2} A_m^\top)$ и с помощью элементарных преобразований приводим ее к виду $(B_{2,1} \mid B_{2,2} \mid B_{2,3})$, где $B_{2,3}$ — ступенчатая матрица.

Алгоритм приведения матрицы оператора к жордановой нормальной форме (5)

Алгоритм 11.1 (алгоритм приведения матрицы оператора к жордановой нормальной форме), часть 5

Продолжаем описанный процесс до тех пор, пока при каком-то k не станет верным равенство $B_{k,k+1}A_m^\top = O$. Строки полученной матрицы $(B_{k,1} \mid B_{k,2} \mid \cdots \mid B_{k,k+1})$ разбиваем на подстроки, каждая из которых расположена внутри одной из матриц $B_{k,1}$, $B_{k,2}$, ..., $B_{k,k+1}$, и рассматриваем эти подстроки как векторы из пространства F^n. Вычеркиваем из каждой строки нулевые векторы и выравниваем оставшиеся векторы по правому краю.

Получилась жорданова таблица. Как и на шаге 1, «сжимаем» эту жорданову таблицу до тех пор, пока оставшиеся в ней векторы не станут линейно независимыми (способ «сжатия» указан в доказательстве предложения 10.2). Обозначаем через P_m упорядоченный набор векторов из полученной жордановой таблицы (перечисляемых в порядке, указанном на шаге 1), а через J_m — клеточно-диагональную матрицу, которая строится по тем же правилам, что и матрица J_i на шаге 1 (в частности, все ее диагональные элементы равны λ_m). Переходим к шагу 4.

Алгоритм приведения матрицы оператора к жордановой нормальной форме (6)

Алгоритм 11.1 (алгоритм приведения матрицы оператора к жордановой нормальной форме), часть 6

Шаг 3. Находим произвольный собственный вектор оператора $\mathcal{A}$, соответствующий собственному значению λ_i. Обозначаем через P_i набор векторов, состоящий из одного этого вектора, а через J_i — квадратную матрицу порядка 1, единственным элементом которой является скаляр λ_i. Если $i < m$, увеличиваем значение i на единицу и возвращаемся в начало шага 3, а если $i = m$, переходим к шагу 4.

Шаг 4. Жордановым базисом пространства V относительно оператора $\mathcal{A}$ является упорядоченный набор векторов $P = P_1 \cup P_2 \cup \dots \cup P_m$, в котором идут сначала векторы из P_1, упорядоченные указанным выше способом, потом векторы из P_2 в аналогичном порядке и т. д. Матрица оператора $\mathcal{A}$ в базисе P имеет вид

$$J = \begin{pmatrix} J_1 & O & \dots & O \\ O & J_2 & \dots & O \\ \dots & \dots & \dots & \dots \\ O & O & \dots & J_m \end{pmatrix}.$$

- Можно было бы убрать из алгоритма шаги 2 и 3 и обрабатывать все собственные значения оператора $\mathcal{A}$ так, как указано на шаге 1. Но это существенно увеличило бы время работы алгоритма.
- Действия, выполняемые на шагах 1 и 2, во многом совпадают. Есть только два различия. Первое состоит в условии выхода из реализуемого на этом шаге цикла, а второе — в том, из каких векторов состоит жорданова таблица, которую мы «сжимаем» до жорданова базиса.
- Шаг 2 почти дословно воспроизводит алгоритм 10.1. Как станет ясно из дальнейшего, это совпадение не случайно.

Алгоритм 10.1: комментарии (2)

Обозначим число кратных корней многочлена $\chi_{\mathcal{A}}(x)$ через r, а число его простых корней — через s. Очевидно, что $0 \leqslant r, s \leqslant m$ и $r + s = m$.

- Если $s = 0$, то шаг 1 выполняется $m - 1$ раз, шаг 2 — один раз, а шаг 3 не выполняется ни разу. Если при этом $m = 1$, то работа алгоритма сводится к однократному выполнению шага 2. Фактически при этом к оператору $\mathcal{A}_1 = \mathcal{A} - \lambda_1\mathcal{E}$ применяется алгоритм 10.1. Это не случайно: в рассматриваемом случае $\chi_{\mathcal{A}}(x) = (-1)^n(x - \lambda_1)^n$, и из теоремы Гамильтона–Кэли для линейных операторов вытекает, что оператор $\mathcal{A} - \lambda_1\mathcal{E}$ нильпотентен. Если, помимо всего сказанного, еще и $\lambda_1 = 0$, то оператор $\mathcal{A}$ уже сам нильпотентен, и алгоритм 11.1 «ужимается» до алгоритма 10.1.
- Если же $s \neq 0$, то шаг 1 выполняется r раз, шаг 3 — s раз, а шаг 2 не выполняется ни разу. Если при этом $r = 0$, то $s = m = n$ и работа алгоритма состоит в n-кратном выполнении шага 3. Матрица J при этом оказывается диагональной. Это и не удивительно, так как в указанном случае $\chi_{\mathcal{A}}(x) = (-1)^n(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$ и оператор $\mathcal{A}$ приводим к диагональному виду в силу следствия 10.1.

Для обоснования алгоритма 11.1 нам понадобятся еще три вспомогательных утверждения. Пусть $i, j \in \{1, 2, \ldots, m\}$ и $i \neq j$. В доказательстве леммы 11.2 показано, что пространство N_i инвариантно относительно оператора $\mathcal{A}_j$. Следующее замечание аналогично этому утверждению и по формулировке, и по доказательству.

Замечание 11.2

Пусть $i, j \in \{1, 2, \ldots, m\}$ и $i \neq j$. Пространство U_i инвариантно относительно оператора $\mathcal{A}_j$.

Доказательство. Пространство U_i инвариантно относительно $\mathcal{A}_i$. Поэтому требуемое утверждение вытекает из равенств

$$\mathcal{A}_j = \mathcal{A} - \lambda_j \mathcal{E} = \mathcal{A} - \lambda_i \mathcal{E} + \lambda_i \mathcal{E} - \lambda_j \mathcal{E} = \mathcal{A}_i + (\lambda_i - \lambda_j)\mathcal{E}$$

и замечания 10.1. □

Инвариантность некоторых подпространств относительно некоторых операторов (2)

Замечание 11.3

Для всякого $i = 1, 2, \ldots, m$, пространство N_i инвариантно относительно оператора $\mathcal{A}$.

Доказательство. Этот факт вытекает из инвариантности пространства N_i относительно оператора $\mathcal{A}_i$, равенства $\mathcal{A} = \mathcal{A}_i + \lambda_i \mathcal{E}$ и замечания 10.1. $\square$

Пересечение подпространств $U_1, U_2, \ldots, U_q$ (1)

Лемма 11.4

Для всякого $q = 1, 2, \ldots, m-1$ выполнено равенство $U_1 \cap U_2 \cap \cdots \cap U_q = N_{q+1} \oplus N_{q+2} \oplus \cdots \oplus N_m$.

Доказательство. Из леммы 11.3 вытекает, что

$$N_{q+1} \oplus N_{q+2} \oplus \cdots \oplus N_m \subseteq U_1 \cap U_2 \cap \cdots \cap U_q.$$

Поэтому достаточно показать, что

$$\dim(N_{q+1} \oplus N_{q+2} \oplus \cdots \oplus N_m) = \dim(U_1 \cap U_2 \cap \cdots \cap U_q).$$

В силу теоремы 11.2

$$\begin{aligned}\dim(N_{q+1} \oplus N_{q+2} \oplus \cdots \oplus N_m) &= \dim V - \dim(N_1 \oplus N_2 \oplus \cdots \oplus N_q) = \\ &= \dim V - \dim N_1 - \dim N_2 - \cdots - \dim N_q.\end{aligned}$$

Осталось проверить, что

$$\dim(U_1 \cap U_2 \cap \cdots \cap U_q) = \dim V - \dim N_1 - \dim N_2 - \cdots - \dim N_q.$$

Докажем это равенство индукцией по q.

База индукции. Пусть $q = 1$. В силу следствия 10.2 $V = N_1 \oplus U_1$. Следовательно, $\dim V = \dim N_1 + \dim U_1$, откуда непосредственно вытекает требуемое равенство.

Пересечение подпространств $U_1, U_2, \ldots, U_q$ (2)

Шаг индукции. Пусть $q > 1$. Из теоремы 3.1 вытекает, что

$$\begin{aligned}\dim(U_1 \cap U_2 \cap \cdots \cap U_q) &= \dim\big((U_1 \cap U_2 \cap \cdots \cap U_{q-1}) \cap U_q\big) = \\ &= \dim(U_1 \cap U_2 \cap \cdots \cap U_{q-1}) + \dim U_q - \\ &\quad - \dim(U_1 + U_2 + \cdots + U_q).\end{aligned}$$

По предположению индукции

$$\dim(U_1 \cap U_2 \cap \cdots \cap U_{q-1}) = \dim V - \dim N_1 - \dim N_2 - \cdots - \dim N_{q-1}.$$

Из следствия 10.2 вытекает, что $\dim U_q = \dim V - \dim N_q$. Наконец, из леммы 11.3 вытекает, что $U_1 \supseteq N_i$ для всякого $i = 2, 3, \ldots, m$, а $U_2 \supseteq N_1$, и потому

$$U_1 + U_2 + \cdots + U_q \supseteq U_1 + U_2 \supseteq N_i$$

для всякого $i = 1, 2, \ldots, m$. С учетом теоремы 11.2, имеем

$$U_1 + U_2 + \cdots + U_q \supseteq N_1 + N_2 + \cdots + N_m = V,$$

и потому $U_1 + U_2 + \cdots + U_q = V$. Объединяя сказанное, имеем

$$\begin{aligned}\dim(U_1 \cap U_2 \cap \cdots \cap U_q) &= (\dim V - \dim N_1 - \dim N_2 - \cdots - \dim N_{q-1}) + \\ &\quad + (\dim V - \dim N_q) - \dim V = \\ &= \dim V - \dim N_1 - \dim N_2 - \cdots - \dim N_q,\end{aligned}$$

что и требовалось доказать. $\square$

Обоснование алгоритма 11.1 (1)

Обоснование алгоритма 11.1. Если многочлен $\chi_{\mathcal{A}}(x)$ не разложим в поле F на линейные множители, то оператор $\mathcal{A}$ не приводим к жордановой нормальной форме по теореме 11.1, и потому работа алгоритма завершается, не успев начаться. Поэтому далее можно считать, что выполнено равенство (6).

Для всякого $i = 1, 2, \ldots, m$ на шагах 1–3 строятся набор векторов P_i и жорданова матрица J_i. Достаточно убедиться в том, что P_i — базис корневого подпространства N_i, а J_i — матрица оператора $\mathcal{A}|_{N_i}$ в этом базисе. В самом деле, предположим, что это так, а P и J — соответственно, набор векторов и матрица, построенные на шаге 4. Тогда из замечания 11.3, теоремы 11.2 и предложения 10.1 вытекает, что P — базис пространства V, а J — матрица оператора $\mathcal{A}$ в этом базисе. Остается заметить, что из построения матрицы J вытекает, что она жорданова.

Обоснование алгоритма 11.1 (2)

Напомним, что скаляры $\lambda_1, \lambda_2, \ldots, \lambda_m$ упорядочены так, что сначала идут все кратные корни многочлена $\chi_{\mathcal{A}}(x)$, а потом — все его простые корни. Обнаружив первый раз, что очередное собственное значение является простым корнем, алгоритм переходит к шагу 3 и выполняет указанные там действия для данного и всех последующих собственных значений, т. е. для всех простых корней многочлена $\chi_{\mathcal{A}}(x)$. Проверим, что построенные при этом наборы векторов P_i и матрицы J_i обладают нужными нам свойствами.

Если λ_i — простой корень многочлена $\chi_{\mathcal{A}}(x)$, то в качестве набора векторов P_i алгоритм выбирает множество $\{\mathbf{p}_i\}$, где $\mathbf{p}_i$ — произвольный собственный вектор оператора $\mathcal{A}$, соответствующий собственному значению λ_i. Следствие 10.3 показывает, что P_i — базис подпространства N_i. Поскольку $\mathcal{A}(\mathbf{p}_i) = \lambda_i \mathbf{p}_i$, матрицей ограничения оператора $\mathcal{A}$ на N_i в базисе P_i является квадратная матрица 1-го порядка (λ_i), являющаяся жордановой. Именно ее алгоритм выбирает в качестве матрицы J_i. Таким образом, в рассматриваемом случае набор векторов P_i и матрица J_i обладают нужными нам свойствами.

Обоснование алгоритма 11.1 (3)

Осталось понять, что алгоритм правильно обрабатывает кратные корни многочлена $\chi_{\mathcal{A}}(x)$. Работа с этими корнями происходит на шагах 1 (при $i < m$) и 2 (при $i = m$). Вкратце суть происходящего там можно охарактеризовать следующим образом. Начиная выполнение шага 1 или 2 при очередном значении i, мы получаем на вход матрицу X_i, ненулевые строки которой образуют базис пространства $N_i \oplus \cdots \oplus N_m$. Действуя на эти векторы нужное число раз оператором $\mathcal{A}_i$, мы с помощью алгоритма Чуркина получаем набор векторов, являющихся системой образующих пространства N_i, а при $i < m$ — еще и матрицу X_{i+1}, ненулевые строки которой образуют базис пространства $N_{i+1} \oplus \cdots \oplus N_m$. После этого мы с помощью алгоритма 10.1 «сжимаем» найденную систему образующих пространства N_i до базиса этого пространства и находим матрицу оператора $\mathcal{A}|_{N_i}$ в этом базисе, являющуюся жордановой. Если $i < m$, то после сказанного мы переходим к рассмотрению следующего значения i, в противном случае завершаем рассмотрение корней многочлена $\chi_{\mathcal{A}}(x)$.

Остается обосновать сказанное в предыдущем абзаце. Этому посвящены следующие пять слайдов.

Перейдем к более подробному разбору работы алгоритма на шагах 1 и 2. Пусть λ_i — кратный корень многочлена $\chi_{\mathcal{A}}(x)$. Предположим сначала, что $i = 1$ и $m > 1$. В этом случае выполняется шаг 1. Заметим, что строки матрицы $X_1 = E$ образуют (стандартный) базис пространства V, а в силу теоремы 11.2 $V = N_1 \oplus N_2 \oplus \cdots \oplus N_m$. Выполнение шага 1 начинается с рассмотрения матрицы $(X_1 \mid X_1 A_1^\top) = (E \mid A_1^\top)$. Пусть Q — тот базис пространства V, в котором оператор $\mathcal{A}$ имеет матрицу A. Мы можем считать, что в левой части матрицы $(E \mid A_1^\top)$ по строкам записаны координаты векторов базиса Q в самом этом базисе. Поскольку в правой части стоит произведение матриц $EA_1^\top$, там по строкам записаны координаты образов векторов из Q под действием оператора $\mathcal{A}_1$. Приведя правую часть к ступенчатому виду, мы найдем базис подпространства $\operatorname{Im} \mathcal{A}_1$ (ненулевые строки матрицы $B_{1,2}$). Затем, умножив $B_{1,2}$ на $A_1^\top$, приписав полученную матрицу справа к той, что была найдена ранее, и приведя новую правую часть к ступенчатому виду, мы найдем базис подпространства $\operatorname{Im} \mathcal{A}_1^2$ (ненулевые строки матрицы $B_{2,3}$). Повторяя эти действия, мы будем получать базисы подпространств $\operatorname{Im} \mathcal{A}_1^3$, $\operatorname{Im} \mathcal{A}_1^4$ и т. д., до тех пор, пока не совпадут ранги матриц $B_{k,k+1}$ и $B_{k+1,k+2}$. Последнее условие означает, что совпадают размерности подпространств $\operatorname{Im} \mathcal{A}_1^k$ и $\operatorname{Im} \mathcal{A}_1^{k+1}$, а значит, и сами эти подпространства (поскольку $\operatorname{Im} \mathcal{A}_1^{k+1} \subseteq \operatorname{Im} \mathcal{A}_1^k$). Это означает, что $\operatorname{Im} \mathcal{A}_1^k = U_1$.

Обоснование алгоритма 11.1 (5)

Отсюда и из алгоритма Чуркина вытекает, что ненулевые строки матрицы $B_{k,k+1}$ образуют базис пространства U_1. Рассмотрим ненулевые строки матрицы $(B_{k,1} \mid B_{k,2} \mid \cdots \mid B_{k,k})$, имеющие нулевые продолжения в $B_{k,k+1}$. Действуя в соответствии с алгоритмом, мы разбиваем всякую такую строку на подстроки, расположенные внутри матриц $B_{k,1}$, $B_{k,2}$, ..., $B_{k,k}$, и рассматриваем эти подстроки как векторы из пространства F^n. Затем мы вычеркиваем нулевые векторы, а оставшиеся векторы сдвигаем вправо. Ясно, что после этого векторы из одной и той же строки образуют ниль-слой, а в целом мы получаем жорданову таблицу. Оператор $\mathcal{N}_1$ нильпотентен. Поэтому «сжимая» полученную жорданову таблицу до линейно независимого набора векторов, мы находим жорданов базис пространства N_1 относительно оператора $\mathcal{N}_1$, который алгоритм и выбирает в качестве набора векторов P_1. Вид матрицы оператора $\mathcal{N}_1$ в базисе P_1 указан в алгоритме 10.1. В этой матрице на главных диагоналях всех клеток Жордана стоит число 0. Поскольку $A = A_1 + \lambda_1 E$, заменив в полученной матрице на главной диагонали нули на λ_1, мы получим матрицу оператора $\mathcal{A}|_{N_1}$, которую алгоритм и выбирает в качестве матрицы J_1. Ясно, что эта матрица жорданова. Таким образом, набор векторов P_1 и матрица J_1 обладают нужными нам свойствами.

Пусть теперь $i = 2$ и $m > 2$. В этом случае также выполняется шаг 1. При этом в качестве X_2 выступает полученная при предыдущем выполнении шага 1 матрица $B_{k,k+1}$, ненулевые строки которой образуют базис пространства U_1. В силу леммы 11.4 $U_1 = N_2 \oplus N_3 \oplus \cdots \oplus N_m$. Выполняя шаг 1 при $i = 2$, мы строим новую последовательность матриц $B_{k,1}$, $B_{k,2}$, ..., $B_{k,k+1}$, которые получаются из матрицы X_2 с помощью элементарных преобразований и умножения справа на матрицу $A_2^\top$. Эти действия не выводят строки указанных матриц за пределы пространства U_1 (для элементарных преобразований матрицы это очевидно, а для умножения на $A_2^\top$ вытекает из замечания 11.2). В частности, это означает, что ненулевые строки «новой» матрицы $B_{k,k+1}$ по-прежнему лежат в U_1. При этом мы, по сути дела, вновь применяем алгоритм Чуркина, но в модифицированном виде[2]. Следовательно, ненулевые строки матрицы $B_{k,k+1}$ лежат в U_2. Таким образом, эти строки попадают в пространство $U_1 \cap U_2$.

[2]Как и в «классическом» варианте алгоритма Чуркина, мы приписываем к координатам базисных векторов координаты их образов и приводим правую часть матрицы к ступенчатому виду. Правда, здесь координаты базисных векторов и их образов записываются не в том базисе, который эти базисные векторы образуют (как это делается в «классическом» варианте алгоритма Чуркина), а в каком-то другом. Но из обоснования алгоритма Чуркина легко усмотреть, что это ничего не меняет.

Как и при рассмотрении случая $i = 1$, проверяется, что ненулевые строки матрицы $B_{k,k+1}$ образуют базис пространства $U_1 \cap U_2$, а ненулевые строки матриц $B_{k,1}$, $B_{k,2}$, …, $B_{k,k}$, имеющие нулевые продолжения в $B_{k,k+1}$, лежат в пространстве $N_2 \cap U_1 = N_2$ (последнее равенство вытекает из леммы 11.3). Как и ранее, эти строки образуют жорданову таблицу. «Сжав» ее до жорданова базиса, мы построим набор векторов P_2 и матрицу J_2, обладающие нужными нам свойствами.

Продолжая перебирать собственные значения, являющиеся кратными корнями многочлена $\chi_{\mathcal{A}}(x)$, мы каждый раз, выполняя шаг 1 для очередного $i < m$, начинаем с найденной при рассмотрении предыдущего собственного значения матрицы X_i, ненулевые строки которой образуют базис пространства $U_1 \cap \dots \cap U_{i-1}$, которое в силу леммы 11.4 равно $N_i \oplus N_{i+1} \oplus \dots \oplus N_m$, и находим матрицу $B_{k,k+1}$, ненулевые строки образуют базис пространства $U_1 \cap \dots \cap U_i$, набор векторов P_i, являющийся базисом пространства $(U_1 \cap \dots \cap U_{i-1}) \cap N_i = N_i$ (последнее равенство вытекает из леммы 11.3), и жорданову матрицу J_i, которая является матрицей оператора $\mathcal{A}|_{N_i}$ в базисе P_i. Набор векторов P_i и матрица J_i обладают нужными нам свойствами.

Если многочлен $\chi_{\mathcal{A}}(x)$ не имеет простых корней, то, дойдя до рассмотрения собственного значения λ_m, мы получим матрицу X_m, ненулевые строки которой лежат в пространстве $U_1 \cap \dots \cap U_{m-1}$, и перейдем к шагу 2. В силу леммы 11.4 $U_1 \cap \dots \cap U_{m-1} = N_m$. Поскольку оператор $\mathcal{N}_m$ нильпотентен, жорданов базис пространства N_m относительно оператора $\mathcal{N}_m$ может быть найден так, как это описано в алгоритме 10.1. Фактически, именно это и делается на шаге 2 с той только разницей, что рассматривается не оператор $\mathcal{A}$, а оператор $\mathcal{A}_m$, и действия начинаются не с матрицы $(E \mid A^\top) = (E \mid EA^\top)$, а с матрицы $(X_m \mid X_m A_m^\top)$. Замена матрицы $(E \mid A^\top)$ на $(X_m \mid X_m A_m^\top)$ непринципиальна: она означает лишь, что мы применяем здесь модифицированный вариант алгоритма Чуркина, упомянутый на слайде, предшествующем предыдущему. А для того, чтобы в результате получилась матрица не оператора $\mathcal{A}_m$, а оператора $\mathcal{A} = \mathcal{A}_m + \lambda_m \mathcal{E}$, мы заменяем на главных диагоналях клеток Жордана полученной матрицы нули на скаляр λ_m. Таким образом, набор векторов P_m и матрица J_m, построенные на шаге 2, обладают нужными нам свойствами. Это завершает обоснование алгоритма. $\square$

11.5. Вычисление степеней матрицы

Жорданова нормальная форма матрицы имеет много различных применений. В оставшейся части параграфа мы остановимся на двух из них: к вычислению степеней матриц и нахождению минимальных многочленов линейных операторов.

Как уже отмечалось в начале данного параграфа, одним из побудительных мотивов для приведения матрицы оператора к жордановой нормальной форме, является то обстоятельство, что оно существенно облегчает возведение квадратной матрицы в произвольную степень. Покажем, как это происходит. Пусть A — матрица некоторого оператора, приводимого к жордановой нормальной форме, в каком-то базисе Q, а P — тот базис, в котором матрица этого оператора жорданова. Вычисление матрицы A^k, где k — произвольное натуральное число, легко сводится к вычислению матрицы A_P^k, поскольку $A^k = T_{QP}A_P^k(T_{QP})^{-1}$ (см. замечание 7.4), а матрица T_{QP} вычисляется легко.

Вычисление степеней матрицы (2)

Покажем, как вычисляются степени жордановых матриц. Легко понять, что вычисление степени клеточно-диагональной матрицы сводится к вычислению степеней ее диагональных клеток:

$$\text{если } A = \begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \hdotsfor{4} \\ O & O & \cdots & A_m \end{pmatrix}, \text{ то } A^k = \begin{pmatrix} A_1^k & O & \cdots & O \\ O & A_2^k & \cdots & O \\ \hdotsfor{4} \\ O & O & \cdots & A_m^k \end{pmatrix}.$$

В частности, вычисление степеней жордановой матрицы сводится к вычислению степеней ее клеток Жордана.

Приведем без доказательства результат о том, как выглядит произвольная степень клетки Жордана. Для этого нам понадобится следующий частный случай биномиальной формулы Ньютона:

$$(\lambda + 1)^k = \lambda^k + C_k^1 \lambda^{k-1} + C_k^2 \lambda^{k-2} + \cdots + C_k^{k-1} \lambda + 1. \tag{7}$$

Пусть $J_{\lambda,n}$ — клетка Жордана порядка n со скаляром λ на главной диагонали, а k — натуральное число. Тогда $J_{\lambda,n}^k = (a_{ij})$, где

$$a_{ij} = \begin{cases} 0, & \text{если либо } i < j, \text{ либо} \\ & i > j + k \text{ (в случае, когда } j + k < n); \\ C_k^{i-j} \lambda^{k-i+j}, & \text{если } j \leqslant i \leqslant \min\{j + k, n\}. \end{cases} \tag{8}$$

Вычисление степеней матрицы (3)

Иными словами, все элементы матрицы $J_{\lambda,n}^k$, расположенные выше главной диагонали, равны 0 и в каждом ее столбце, начиная с элемента на главной диагонали, сверху вниз последовательно идут слагаемые из правой части равенства (7). При этом, если в указанной части столбца «хватает места» для всех слагаемых и столбец остается не заполненным до конца, то оставшиеся места в нижней части столбца заполняются нулями. Если же места не хватает, то там стоят столько первых слагаемых из правой части равенства (7), сколько «помещается» в столбце.

Например,

$$J_{\lambda,3}^5 = \begin{pmatrix} \lambda^5 & 0 & 0 \\ 5\lambda^4 & \lambda^5 & 0 \\ 10\lambda^3 & 5\lambda^4 & \lambda^5 \end{pmatrix}, \text{ а } J_{\lambda,8}^4 = \begin{pmatrix} \lambda^4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4\lambda^3 & \lambda^4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6\lambda^2 & 4\lambda^3 & \lambda^4 & 0 & 0 & 0 & 0 & 0 \\ 4\lambda & 6\lambda^2 & 4\lambda^3 & \lambda^4 & 0 & 0 & 0 & 0 \\ 1 & 4\lambda & 6\lambda^2 & 4\lambda^3 & \lambda^4 & 0 & 0 & 0 \\ 0 & 1 & 4\lambda & 6\lambda^2 & 4\lambda^3 & \lambda^4 & 0 & 0 \\ 0 & 0 & 1 & 4\lambda & 6\lambda^2 & 4\lambda^3 & \lambda^4 & 0 \\ 0 & 0 & 0 & 1 & 4\lambda & 6\lambda^2 & 4\lambda^3 & \lambda^4 \end{pmatrix}.$$

11.6. Нахождение минимального многочлена линейного оператора

Покажем, как знание жордановой нормальной формы линейного оператора позволяет легко найти его минимальный многочлен. Пусть $\mathcal{A}$ — линейный оператор, приводимый к жордановой нормальной форме, а A — его матрица в некотором базисе, являющаяся жордановой. В силу предложения 7.5 минимальный многочлен оператора $\mathcal{A}$ совпадает с минимальным многочленом матрицы A. Поэтому достаточно понять, как выглядит минимальный многочлен жордановой матрицы.

Начнем с простейшего случая — с клетки Жордана. Пусть, как и ранее, $J_{\lambda,n}$ — клетка Жордана порядка n со скаляром λ на главной диагонали.

Лемма 11.5

Минимальный многочлен матрицы $J_{\lambda,n}$ равен $(x-\lambda)^n$.

Доказательство. Найдем характеристический многочлен матрицы $J_{\lambda,n}$:

$$|J_{\lambda,n} - xE| = \begin{vmatrix} \lambda - x & 0 & 0 & \ldots & 0 \\ 1 & \lambda - x & 0 & \ldots & 0 \\ 0 & 1 & \lambda - x & \ldots & 0 \\ \multicolumn{5}{c}{\dotfill} \\ 0 & 0 & 0 & \ldots & \lambda - x \end{vmatrix} = (\lambda - x)^n = (-1)^n (x-\lambda)^n.$$

В силу следствия 14.1 из курса «Основы алгебры» минимальный многочлен матрицы $J_{\lambda,n}$ равен $(x-\lambda)^k$ для некоторого $k \leqslant n$. При этом $(J_{\lambda,n} - \lambda E)^n = O$ в силу теоремы Гамильтона–Кэли. Осталось убедиться в том, что если $k < n$, то $(J_{\lambda,n} - \lambda E)^k \neq O$.

Положим $(J_{\lambda,n} - \lambda E)^k = (a_{ij})$. Тогда

$$a_{ij} = \begin{cases} 1, & \text{если } i - j = k; \\ 0 & \text{в противном случае.} \end{cases}$$

Это вытекает из (8) при $\lambda = 0$ (поскольку $J_{\lambda,n} - \lambda E = J_{0,n}$) и может быть легко проверено непосредственно индукцией по k. В частности, если $k < n$, то $a_{k+1\,1} = 1$, и потому $(J_{\lambda,n} - \lambda E)^k \neq O$. $\square$

Предложение 11.2

Пусть $\mathcal{A}$ — линейный оператор, приводимый к жордановой нормальной форме, A — его матрица в некотором базисе, являющаяся жордановой, а $\lambda_1, \lambda_2, \ldots, \lambda_m$ — совокупность всех попарно различных собственных значений этого оператора. Для всякого $i = 1, 2, \ldots, m$ обозначим через q_i максимальный из порядков клеток Жордана матрицы A, соответствующих собственному значению λ_i. Тогда минимальный многочлен оператора $\mathcal{A}$ равен $(x-\lambda_1)^{q_1}(x-\lambda_2)^{q_2}\cdots(x-\lambda_m)^{q_m}$.

Доказательство. Как уже отмечалось, миниальные многочлены оператора $\mathcal{A}$ и матрицы A совпадают в силу предложения 7.5. Можно считать, что

$$A = \begin{pmatrix} J_1 & O & O & \ldots & O \\ O & J_2 & O & \ldots & O \\ O & O & J_3 & \ldots & O \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ O & O & O & \ldots & J_m \end{pmatrix},$$

где J_i — жорданова матрица, во всех клетках которой на главной диагонали стоит скаляр λ_i $(i = 1, 2, \ldots, m)$. В силу предложения 14.3 из курса «Основы алгебры» и леммы 11.5 для всякого $i = 1, 2, \ldots, m$ минимальный многочлен матрицы J_i равен $(x - \lambda_i)^{q_i}$, где q_i имеет смысл, указанный в формулировке предложения. Остается еще раз сослаться на предложение 14.3 из курса «Основы алгебры». □

Отметим, что предложение 11.2 обобщает следствие 14.2 из курса «Основы алгебры».