Глава 12. Линейные преобразования линейных пространств

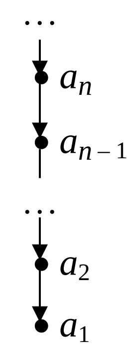
§ 5. Жорданова форма матрицы линейного преобразования

Из теоремы о корневом разложении следует, что теперь надо изучить, как может быть устроена матрица линейного преобразования на корневом подпространстве. Поэтому в первой части этого параграфа будем считать, что $L = L_f(\alpha)$ для некоторого собственного числа α преобразования f.

Лемма 2, п. 4 из § 3 утверждает, что $(f - \alpha\varepsilon)\,(L_n(\alpha)) \subseteq L_{n-1}(\alpha)$. Это наблюдение делает естественным следующее понятие.

Определение 1. Последовательность векторов a_1, a_2, …, a_n, … называется ***α-серией***, если $a_1 \neq 0$, $(f - \alpha\varepsilon)(a_1) = 0$ и $(f - \alpha\varepsilon)(a_n) = a_{n-1}$ для $n > 1$.

Мы будем изображать α-серию так, как показано справа. Стрелочки показывают нам, куда переходит вектор после применения к нему преобразования $f - \alpha\varepsilon$.

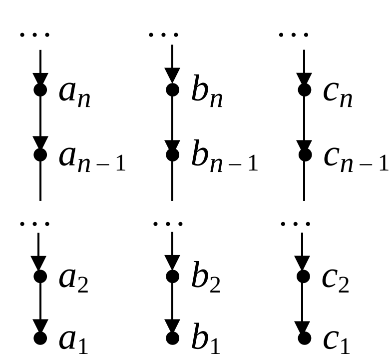


Ясно, $a_1 \in L_1(\alpha)$, $a_2 \in L_2(\alpha)$, …, $a_n \in L_n(\alpha)$, … .

Мы знаем, что a_1 – собственный вектор преобразования f. И, вообще говоря, с каждого собственного вектора начинается некоторая α-серия. Я не буду применять двойную индексацию, а просто писать, что другая серия начинается с b_1, третья – с c_1 и т.д. Впрочем, если мы возьмём произвольный вектор из пространства L и будем последовательно применять преобразование $f - \alpha\varepsilon$, то на каком-то шаге получим 0, а на предыдущем собственный вектор. У нас и в этом случае появится α-серия. Иными словами, каждый вектор пространства L может быть включён в некоторую α-серию.

Определение 2. ***Диаграммой*** называется объединение некоторого набора α-серий, у которых собственные векторы линейно независимы.

Диаграмму будем обозначать буквой ***D*** и изображать так, как показано справа.



Ясно, что количество α-серий в диаграмме конечно, поскольку конечна размерность пространства L. На самом деле конечна вся диаграмма, что вытекает из следующей теоремы.

Теорема 1. Любая диаграмма является линейно независимой системой векторов.

Доказательство. Положим $D_i = \boldsymbol{D} \cap L_i(\alpha)$. Видно, что $D_1 = \{a_1, b_1, c_1, \ldots\}$, $D_2 = \{a_1, b_1, c_1, \ldots, a_2, b_2, c_2, \ldots\}$ и т.д.

Индукцией по i докажем, что система векторов D_i линейно независима.

Б.И. выполнена по определению диаграммы.

Ш.И. Пусть уже доказано, что D_{i-1} линейно независима. Запишем линейную комбинацию векторов из D_i:

$$\rho_i a_i + \sigma_i b_i + \tau_i c_i + \ldots + \rho_{i-1} a_{i-1} + \sigma_{i-1} b_{i-1} + \tau_{i-1} c_{i-1} + \ldots + \ldots + \rho_1 a_1 + \sigma_1 b_1 + \tau_1 c_1 + \ldots = 0.$$

Применим к обеим частям равенства преобразование $f - \alpha\varepsilon$. Получим

$$\rho_i a_{i-1} + \sigma_i b_{i-1} + \tau_i c_{i-1} + \ldots + \rho_2 a_1 + \sigma_2 b_1 + \tau_2 c_1 + \ldots = 0.$$

Но это линейная комбинация векторов из D_{i-1}. Ввиду линейной независимости имеем $\rho_i = \sigma_i = \tau_i = \ldots = \rho_2 = \sigma_2 = \tau_2 = \ldots = 0$. Тогда исходная линейная комбинация превращается в $\rho_1 a_1 + \sigma_1 b_1 + \tau_1 c_1 + \ldots = 0$, откуда $\rho_1 = \sigma_1 = \tau_1 = \ldots = 0$. □

Следствие. Каждая α-серия конечна и порождает инвариантное относительно f. подпространство.

Доказательство очевидно.

Определение 3. Диаграмма называется **максимальной**, если она содержит наибольшее число элементов.

Теорема 2. Максимальная диаграмма является базисом пространства L.

Пусть $\boldsymbol{D}$ – некоторая максимальная диаграмма. Ввиду теоремы 1 нам требуется доказать лишь то, что $\boldsymbol{D}$ – это система образующих для пространства L. Для доказательства нам потребуется ещё одно рабочее понятие. Поскольку L – α-корневое пространство, для каждого $x \in L$ найдётся натуральное n, для которого $(f - \alpha\varepsilon)^n(x) = 0$. Наименьшее n, для которого это выполняется, будем называть ***высотой вектора*** x и обозначать $h(x)$.

Докажем индукцией по $h(x)$ следующее утверждение: если $h(x) = k$, то $x \vdash D_k$.

Б.И. $k = 1$. Если x не выражается линейно через D_1, добавим его в диаграмму $\boldsymbol{D}$. Это снова будет диаграмма, что противоречит максимальности $\boldsymbol{D}$. Фактически мы показали, что D_1 – базис пространства $L_1(\alpha)$.

Ш.И. Пусть утверждение доказано для всех векторов, высота которых не больше $k - 1$. Пусть $h(x) = k$. Положим $y = (f - \alpha\varepsilon)(x)$. Заметим, что $h(y) = k - 1$. По предположению индукции $y \vdash D_{k-1}$. Запишем его выражение через векторы из D_{k-1}:

$$y = \rho_{k-1}a_{k-1} + \sigma_{k-1}b_{k-1} + \tau_{k-1}c_{k-1} + \ldots + \rho_{k-2}a_{k-2} + \sigma_{k-2}b_{k-2} + \tau_{k-2}c_{k-2} + \ldots .$$

Отметим, что хотя бы один из коэффициентов ρ_{k-1}, σ_{k-1}, τ_{k-1} … отличен от 0, ибо в противном случае $h(y) < k - 1$, а тогда $h(x) < k$. Будем считать, что в записи элемента y присутствуют только ненулевые слагаемые.

Рассмотрим далее два случая.

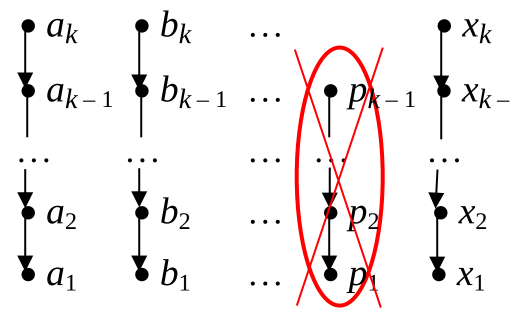
1) Каждый из векторов a_{k-1}, b_{k-1}, c_{k-1}, …, p_{k-1}, фигурирующих в записи вектора y, не является последним в своей α-серии. Пусть $z = x - (\rho_k a_k + \sigma_k b_k + \tau_k c_k + \ldots + \xi_k p_k)$. Тогда

$$(f - \alpha\varepsilon)(z) = (f - \alpha\varepsilon)(x) - (f - \alpha\varepsilon)(\rho_k a_k) - (f - \alpha\varepsilon)(\sigma_k b_k) - (f - \alpha\varepsilon)(\tau_k c_k) - \ldots - (f - \alpha\varepsilon)(\xi_k p_k) =$$

$$= y - \rho_{k-1}a_{k-1} - \sigma_{k-1}b_{k-1} - \tau_{k-1}c_{k-1} - \ldots. - \xi_{k-1}p_{k-1} = \rho_{k-2}a_{k-2} + \sigma_{k-2}b_{k-2} + \tau_{k-2}c_{k-2} + \ldots$$

Следовательно, $h((f - \alpha\varepsilon)(z)) \le k - 2$. Но тогда $h(z) \le k - 1$. По предположению индукции $z \vdash D_{k-1}$. А тогда $x \vdash D_k$, поскольку $a_k, b_k, c_k, \ldots, p_k \in D_k$.

2) Один из векторов a_{k-1}, b_{k-1}, c_{k-1}, …, p_{k-1}, фигурирующих в записи вектора y, является последним в своей α-серии. Поскольку нет разницы, в каком порядке располагаются векторы в изображении диаграммы $\boldsymbol{D}$, будем считать, что это вектор p_{k-1}. Дальше поступим так.



Пусть $x = x_k$ и положим $x_{k-1} = (f - \alpha\varepsilon)(x_k)$, $x_{k-2} = (f - \alpha\varepsilon)(x_{k-1})$, …, $x_1 = (f - \alpha\varepsilon)(x_2)$. Вырежем из диаграммы α-серию p и вклеим вместо неё α-серию x. Докажем, что у нас снова получится диаграмма. Для этого надо проверить, что векторы, расположенные в нижнем ряду линейно независимы. Это множество векторов можно записать так: $(D_1 \setminus \{p_1\}) \cup \{x_1\}$.

С одной стороны, D_1 – базис $L_1(\alpha)$. Поэтому система $(D_1 \setminus \{p_1\}) \cup \{x_1\} \vdash D_1$. С другой стороны,

$$y = \rho_{k-1}a_{k-1} + \sigma_{k-1}b_{k-1} + \tau_{k-1}c_{k-1} + \ldots + \xi_{k-1}p_{k-1} + \rho_{k-2}a_{k-2} + \sigma_{k-2}b_{k-2} + \tau_{k-2}c_{k-2} + \ldots,$$

Поэтому $x_1 = (f - \alpha\varepsilon)^{k-2}(y) = \rho_{k-1}a_1 + \sigma_{k-1}b_1 + \tau_{k-1}c_{k1} + \ldots + \xi_{k-1}p_1$. Отсюда

$$p_1 = \frac{1}{\xi_{k-1}}(x_1 - \rho_{k-1}a_1 - \sigma_{k-1}b_1 - \tau_{k-1}c_{k1} - \ldots).$$

Следовательно, $D_1 \vdash (D_1 \setminus \{p_1\}) \cup \{x_1\}$. Значит, системы D_1 и $(D_1 \setminus \{p_1\}) \cup \{x_1\}$ эквивалентны, содержат одинаковое число векторов, поэтому $(D_1 \setminus \{p_1\}) \cup \{x_1\}$ –

тоже базис и, значит, линейно независимая система. Но тогда получается диаграмма, содержащая больше элементов, чем в $\boldsymbol{D}$ – противоречие. □

Теперь приступим к конструированию матрицы преобразования f. Благодаря следствию из Теоремы 1 и Теореме 2, α-корневое подпространство является прямой суммой инвариантных подпространств, порождённых всеми α-сериями, образующими максимальную диаграмму. Следовательно, матрица преобразования f будет квазидиагональной, если в качестве базиса взять выписанные последовательно α-серии максимальной диаграммы:

$$a_1, a_2, \dots, a_{n_1}, b_1, b_2, \dots, b_{n_2}, \dots, h_1, h_2, \dots, h_{n_m}.$$

Осталось записать, как выглядит матрица преобразования f, когда базисом является α-серия. Возьмём для примера α-серию $a_1, a_2, \dots, a_{n_1}$.

$(f - \alpha\varepsilon)(a_1) = 0 \quad \Rightarrow \quad f(a_1) = \alpha a_1;$

$(f - \alpha\varepsilon)(a_2) = a_1 \quad \Rightarrow \quad f(a_2) = \alpha a_2 + a_1;$

$(f - \alpha\varepsilon)(a_3) = a_2 \quad \Rightarrow \quad f(a_3) = \alpha a_3 + a_2;$

…

$(f - \alpha\varepsilon)(a_{n_1}) = a_{n_1-1} \quad \Rightarrow \quad f(a_{n_1}) = \alpha a_{n_1} + a_{n_1-1}.$

Значит, матрица преобразования в этом базисе выглядит так:

$$\begin{pmatrix} \alpha\ 1\ 0 \dots 0\ 0 \\ 0\ \alpha\ 1 \dots 0\ 0 \\ 0\ 0\ \alpha \dots 0\ 0 \\ \dots \dots \dots \dots \dots \\ 0\ 0\ 0 \dots\ \alpha\ 1 \\ 0\ 0\ 0 \dots 0\ \alpha \end{pmatrix}.$$

Матрица такого вида называется ***жордановой клеткой***.

Вернёмся к общей задаче. Пусть пространство L разложилось в прямую сумму корневых подпространств. Если теперь для каждого корневого подпространства в качестве базиса взять последовательность соответствующих корневых серий из максимальной диаграммы и всё вместе рассмотрим как базис всего L, то матрица преобразования f в этом базисе будет квазидиагональной и на её диагонали стоят жордановы клетки. Матрица такого вида называется ***жордановой формой*** матрицы преобразования f. Фактически доказана следующая

Теорема 3 (существования жордановой формы). Если характеристический многочлен линейного преобразования f линейного пространства L разлагается на линейные множители, то существует такой базис пространства L, в котором матрица преобразования f имеет жорданову форму.

Сформулируем также матричную версию этой теоремы, вспоминая, что переход к другому базису – это замена матрицы на ей подобную.

Теорема 3′. Если характеристический многочлен квадратной матрицы разлагается на линейные множители, то она подобна матрице в жордановой форме.

Из этой теоремы мы получим одно удивительное следствие.

Теорема 4 (Гамильтона – Кэли). Каждая квадратная матрица является корнем своего характеристического многочлена.

Доказательство. Прежде всего, заметим, что если для некоторого многочлена $g(x)$ и некоторой матрицы A выполнено равенство $g(A) = 0$, то и для любой матрицы B, подобной матрице A, выполняется равенство $g(B) = 0$. Действительно, поскольку $B = T^{-1}AT$ для некоторой обратимой матрицы T, $g(B) = g(T^{-1}AT) = T^{-1}g(A)T = 0$.

Пусть теперь A – произвольная квадратная матрица размеров $n \times n$, $g(x)$ – её характеристический многочлен. Рассмотрим сначала случай, когда $g(x)$ разлагается на линейные множители:

$$g(x) = (-1)^n(x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\dots(x-\alpha_m)^{k_m},$$

где $k_1 + k_2 + \dots + k_m = n$. Вместо матрицы A мы можем подставить в $g(x)$ её жорданову форму, получаемую благодаря теореме 3′. Обозначим её B. Вот её общий вид:

$$\left(\begin{array}{c|c|c|c|c}
\begin{matrix} \alpha_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & \alpha_1 & 1 & \dots & 0 & 0 \\ 0 & 0 & \alpha_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha_1 & 1 \\ 0 & 0 & 0 & \dots & 0 & \alpha_1 \end{matrix} & & & & \\ \hline
 & \begin{matrix} \alpha_1 & 1 & 0 & \dots & 0 \\ 0 & \alpha_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha_1 \end{matrix} & & & \\ \hline
 & & \dots & & \\ \hline
 & & & \begin{matrix} \alpha_2 & 1 & 0 & \dots & 0 \\ 0 & \alpha_2 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha_2 \end{matrix} & \\ \hline
 & & & & \dots
\end{array}\right).$$

? Каковы размеры матрицы, записанной в левом верхнем углу и ограниченной красными линиями (т.е. до начала клеток с числами α_2 на диагонали матрицы)?

Вычислим значение $(x - \alpha_1)^{k_1}$ после подстановки вместо x матрицы B:

$$\left(\begin{array}{c|c} \begin{matrix} 0\ 1\ 0 \dots 0\ 0 \\ 0\ 0\ 1 \dots 0\ 0 \\ 0\ 0\ 0 \dots 0\ 0 \\ \dots \dots \dots \dots \dots \\ 0\ 0\ 0 \dots\ 0\ 1 \\ 0\ 0\ 0 \dots\ 0\ 0 \end{matrix} & \\ \qquad\qquad \begin{matrix} 0\ 1\ 0 \dots 0 \\ 0\ 0\ 1 \dots 0 \\ \dots \dots \dots \dots \\ 0\ 0\ 0 \dots\ 0 \end{matrix} & \\ \qquad\qquad\qquad\qquad \dots & \\ \hline & \begin{matrix} \alpha_2 - \alpha_1 & 1 & 0 \dots 0 \\ 0 & \alpha_2 - \alpha_1 & 1 \dots 0 \\ \dots & \dots \dots & \dots \dots \\ 0 & 0 & 0 \dots \alpha_2 - \alpha_1 \end{matrix} \\ & \qquad\qquad\qquad \dots \end{array}\right)^{k_1}$$

? Как возвести квазидиагональную матрицу в степень?

! Проверьте, что при возведении матрицы $\begin{pmatrix} 0\ 1\ 0 \dots 0\ 0 \\ 0\ 0\ 1 \dots 0\ 0 \\ 0\ 0\ 0 \dots 0\ 0 \\ \dots \dots \dots \dots \dots \\ 0\ 0\ 0 \dots\ 0\ 1 \\ 0\ 0\ 0 \dots 0\ 0 \end{pmatrix}$ в степень, не меньшую, чем размер матрицы, получается нулевая матрица.

Значит, после возведения матрицы в степень k_1 в верхнем левом красном углу получится нулевая матрица.

Аналогично при вычислении $(B - \alpha_2 E)^{k_2}$ возникнет нулевая матрица в том квадрате, где у матрицы B стоит число α_2. И т.д. Чтобы вычислить $g(B)$, надо перемножить все получившиеся квазидиагональные матрицы. Ясно, что получится нулевая матрица.

Чтобы доказать теорему и в том случае, когда характеристический многочлен не разлагается на линейные множители, мы используем однажды уже применявшийся приём. Мы рассмотрим алгебраически замкнутое поле $\bar{F}$, содержащее поле F, и уже над $\bar{F}$ заменим матрицу A на подобную ей матрицу B в жордановой форме. Характеристический многочлен у матриц A и B один и тот же, и, значит, его значение от матрицы A равно 0 по уже доказанному. $\square$