

Тема II. Линейные операторы

§ 9. Корневые подпространства

М.В.Волков

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

2021/2022 учебный год

Пусть V – линейное пространство над произвольным полем F , $\dim V < \infty$, $\mathcal{A}: V \rightarrow V$ – линейный оператор.

Мы хотим найти базис пространства V , в котором матрица оператора \mathcal{A} устроена как можно проще и действие оператора \mathcal{A} наиболее понятно. Если поле F произвольно, то скалярного произведения на V , вообще говоря, нет, и построения из предыдущих лекций, увы, неприменимы. Вспомним идеи, связанные с инвариантными подпространствами.

Определение

Пусть V – векторное пространство, $\mathcal{A}: V \rightarrow V$ – линейный оператор. Подпространство $S \subseteq V$ называется *инвариантным относительно \mathcal{A}* или *\mathcal{A} -инвариантным*, если $\mathcal{A}x \in S$ для любого $x \in S$.

Мы знаем, что если пространство V является *прямой суммой* ненулевых \mathcal{A} -инвариантных подпространств S_1, \dots, S_t , то в базисе пространства V , полученном объединением базисов подпространств S_1, \dots, S_t , матрица оператора \mathcal{A} будет блочно-диагональной, причем i -й диагональный блок будет матрицей ограничения \mathcal{A} на подпространство S_i в выбранном в этом подпространстве базисе. Но где взять такие подпространства?

Для любого линейного оператора $\mathcal{A}: V \rightarrow V$ его ядро $\text{Ker } \mathcal{A}$ и образ $\text{Im } \mathcal{A}$ будут \mathcal{A} -инвариантными подпространствами. Однако для «типичного» оператора \mathcal{A} эти подпространства имеют ненулевое пересечение, и потому разложить пространство V в их прямую сумму не получится.

Так как $\mathcal{A}^{k+1}\mathbf{x} = \mathcal{A}(\mathcal{A}^k\mathbf{x})$ для любого $\mathbf{x} \in V$, имеем $\text{Ker } \mathcal{A}^k \subseteq \text{Ker } \mathcal{A}^{k+1}$ для любого натурального k . Кроме того, ясно, что $\text{Im } \mathcal{A}^k \supseteq \text{Im } \mathcal{A}^{k+1}$. Получаем две цепочки подпространств – растущую и убывающую:

$$\begin{aligned} \{0\} &\subseteq \text{Ker } \mathcal{A} \subseteq \text{Ker } \mathcal{A}^2 \subseteq \dots \subseteq \text{Ker } \mathcal{A}^s \subseteq \dots \\ V &\supseteq \text{Im } \mathcal{A} \supseteq \text{Im } \mathcal{A}^2 \supseteq \dots \supseteq \text{Im } \mathcal{A}^s \supseteq \dots \end{aligned}$$

Так как $\dim V < \infty$, в обеих цепочках возникнут равные подпространства. Пусть s таково, что $\text{Im } \mathcal{A}^s = \text{Im } \mathcal{A}^{s+1}$. Индукцией по k покажем, что тогда имеет место $\text{Im } \mathcal{A}^{s+k} = \text{Im } \mathcal{A}^s$ для любого натурального k . База очевидна. Шаг: $\text{Im } \mathcal{A}^{s+k} = \mathcal{A}(\text{Im } \mathcal{A}^{s+k-1}) = \mathcal{A}(\text{Im } \mathcal{A}^s) = \text{Im } \mathcal{A}^{s+1} = \text{Im } \mathcal{A}^s$.

По теореме о сумме ранга и дефекта $\dim \text{Im } \mathcal{A}^s + \dim \text{Ker } \mathcal{A}^s = \dim V$. Поэтому обе цепочки стабилизируются одновременно.

Заметим еще, что все подпространства в обеих цепочках \mathcal{A} -инвариантны.

Положим $U_{\mathcal{A}} := \text{Im } \mathcal{A}^s$, $N_{\mathcal{A}} := \text{Ker } \mathcal{A}^s$, где s – наименьшее со свойством $\text{Im } \mathcal{A}^s = \text{Im } \mathcal{A}^{s+1} = \dots$. Как отмечено, тогда и $\text{Ker } \mathcal{A}^s = \text{Ker } \mathcal{A}^{s+1} = \dots$. Инвариантное подпространство $U_{\mathcal{A}}$ называется **1-компонентой** оператора \mathcal{A} , а инвариантное подпространство $N_{\mathcal{A}}$ – его **0-компонентой**. Оператор, некоторая степень которого – нулевой оператор, называют **нильпотентным**.

Предложение (лемма Фиттинга)

$V = U_{\mathcal{A}} \oplus N_{\mathcal{A}}$ для любого линейного оператора $\mathcal{A}: V \rightarrow V$. При этом ограничение \mathcal{A} на 1-компоненту $U_{\mathcal{A}}$ – невырожденный оператор, а ограничение \mathcal{A} на 0-компоненту $N_{\mathcal{A}}$ – нильпотентный оператор.

Доказательство. Поскольку $\dim \text{Im } \mathcal{A}^s + \dim \text{Ker } \mathcal{A}^s = \dim V$, равенство $V = U_{\mathcal{A}} \oplus N_{\mathcal{A}}$ равносильно равенству $U_{\mathcal{A}} \cap N_{\mathcal{A}} = \{\mathbf{0}\}$. Возьмем вектор $\mathbf{x} \in U_{\mathcal{A}} \cap N_{\mathcal{A}}$. Раз $\mathbf{x} \in U_{\mathcal{A}} = \text{Im } \mathcal{A}^s$, имеем $\mathbf{x} = \mathcal{A}^s \mathbf{y}$ для некоторого \mathbf{y} . Раз $\mathbf{x} \in N_{\mathcal{A}} = \text{Ker } \mathcal{A}^s$, имеем $\mathbf{0} = \mathcal{A}^s \mathbf{x} = \mathcal{A}^{2s} \mathbf{y}$. Отсюда $\mathbf{y} \in \text{Ker } \mathcal{A}^{2s} = \text{Ker } \mathcal{A}^s$, а значит, $\mathbf{x} = \mathcal{A}^s \mathbf{y} = \mathbf{0}$.

По определению 0-компоненты ограничение \mathcal{A}^s на нее – нулевой оператор, т.е. ограничение \mathcal{A} на нее – нильпотентный оператор. Из того, что $U_{\mathcal{A}} \cap \text{Ker } \mathcal{A}^s = \{\mathbf{0}\}$, следует, что ограничение \mathcal{A}^s на 1-компоненту невырождено, но тогда невырождено и ограничение \mathcal{A} на нее. □

Пусть V – линейное пространство над произвольным полем F , $\dim V < \infty$, $\mathcal{A}: V \rightarrow V$ – линейный оператор. Предполагаем, что все собственные значения оператора \mathcal{A} лежат в F , – этого можно добиться, расширив F до поля разложения характеристического многочлена оператора \mathcal{A} . Возьмем собственное значение α оператора \mathcal{A} и положим $\mathcal{A}_\alpha := \mathcal{A} - \alpha\mathcal{E}$. Отметим, что у \mathcal{A} и \mathcal{A}_α одинаковые инвариантные подпространства.

Определение

Корневым подпространством линейного оператора \mathcal{A} , соответствующим собственному значению α , называется 0-компонента $N_{\mathcal{A}_\alpha}$ оператора \mathcal{A}_α .

Более подробно, корневое подпространство, отвечающее собственному значению α , – это $\text{Ker}(\mathcal{A} - \alpha\mathcal{E})^s$, где s – наименьшее число со свойством $\text{Ker}(\mathcal{A} - \alpha\mathcal{E})^s = \text{Ker}(\mathcal{A} - \alpha\mathcal{E})^{s+1} = \dots$. Число s называют *высотой* корневого подпространства $\text{Ker}(\mathcal{A} - \alpha\mathcal{E})^s$, а его ненулевые вектора – *корневыми векторами, принадлежащими собственному значению* α .

Отметим, что собственные вектора оператора \mathcal{A} , принадлежащие α , являются корневыми – ведь собственные вектора лежат в $\text{Ker}(\mathcal{A} - \alpha\mathcal{E})$. Итак, корневые вектора – обобщение собственных.

Зачем нужны корневые вектора? Вспомним, чем хороши собственные вектора: матрица оператора в базисе из собственных векторов диагональна, причем по диагонали стоят собственные значения. Однако базис из собственных векторов есть далеко не у всех операторов.

Пример. Рассмотрим оператор дифференцирования \mathcal{D} на пространстве квадратных трехчленов над \mathbb{R} . Матрица оператора \mathcal{D} в стандартном

базисе $1, x, x^2$ равна $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. Ее характеристический многочлен

$\begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3$ имеет корень 0 кратности 3. Собственные вектора

оператора \mathcal{D} , принадлежащие 0, суть ненулевые константы, поэтому у \mathcal{D} нет базиса из собственных векторов. Отметим, что *базис из корневых векторов* у \mathcal{D} *есть*: многочлены $1, x, x^2$ – корневые вектора.

Мы докажем, что базис из корневых векторов есть у любого линейного оператора, собственные значения которого лежат в поле скаляров. После этого надо будет еще разобраться, как устроена матрица оператора в базисе из корневых векторов, но это тема следующей лекции.

Спектр оператора \mathcal{A} – это множество всех его собственных значений. Спектр оператора \mathcal{A} обозначается через $\text{Spec } \mathcal{A}$.

Теорема (корневое разложение)

Если V – линейное пространство над полем F , а $\mathcal{A}: V \rightarrow V$ – такой линейный оператор, что $\text{Spec } \mathcal{A} \subseteq F$, то V – прямая сумма всех корневых подпространств оператора \mathcal{A} . Размерность корневого подпространства, отвечающего $\alpha \in \text{Spec } \mathcal{A}$, равна кратности корня α в характеристическом многочлене оператора \mathcal{A} .

Если корневое подпространство, отвечающее α , обозначить через V_α , корневое разложение запишется так:

$$V = \bigoplus_{\alpha \in \text{Spec } \mathcal{A}} V_\alpha.$$

Доказательство. Индукция по $\dim V$. База $\dim V = 1$ очевидна.

Пусть $\dim V > 1$. Возьмем $\alpha \in \text{Spec } \mathcal{A}$ и запишем для оператора $\mathcal{A}_\alpha = \mathcal{A} - \alpha \mathcal{E}$ разложение Фиттинга: $V = U_{\mathcal{A}_\alpha} \oplus N_{\mathcal{A}_\alpha} = U_{\mathcal{A}_\alpha} \oplus V_\alpha$.

Составим базис пространства V из базисов прямых слагаемых U_{A_α} и V_α . Поскольку подпространства U_{A_α} и V_α инвариантны относительно \mathcal{A} , матрица оператора \mathcal{A} в этом базисе имеет вид $A = \begin{pmatrix} B & O \\ O & C \end{pmatrix}$, где B и C – матрицы ограничений \mathcal{A} на U_{A_α} и V_α в базисах этих подпространств, а через O обозначены нулевые блоки соответствующих размеров. Тогда характеристический многочлен $f_{\mathcal{A}}(\lambda)$ оператора \mathcal{A} равен произведению характеристических многочленов матриц B и C :

$$f_{\mathcal{A}}(\lambda) = |A - \lambda E| = \begin{vmatrix} B - \lambda E & O \\ O & C - \lambda E \end{vmatrix} = |B - \lambda E| \cdot |C - \lambda E|.$$

Пусть β – корень многочлена $|C - \lambda E|$. Возьмем в подпространстве V_α собственный вектор \mathbf{x} , принадлежащий β . Применим к \mathbf{x} оператор \mathcal{A}_α :

$$\mathcal{A}_\alpha \mathbf{x} = (\mathcal{A} - \alpha \mathcal{E}) \mathbf{x} = \mathcal{A} \mathbf{x} - \alpha \mathbf{x} = \beta \mathbf{x} - \alpha \mathbf{x} = (\beta - \alpha) \mathbf{x}.$$

Отсюда $\mathcal{A}_\alpha^2 \mathbf{x} = (\beta - \alpha)^2 \mathbf{x}$, $\mathcal{A}_\alpha^3 \mathbf{x} = (\beta - \alpha)^3 \mathbf{x}$, и т.д. Если s – высота корневого подпространства V_α , то $\mathcal{A}_\alpha^s \mathbf{x} = \mathbf{0}$, так как $V_\alpha = \text{Ker } \mathcal{A}_\alpha^s$. Отсюда $(\beta - \alpha)^s \mathbf{x} = \mathbf{0}$. Поскольку $\mathbf{x} \neq \mathbf{0}$, заключаем, что $(\beta - \alpha)^s = 0$, т.е. $\beta = \alpha$. Раз все корни многочлена $|C - \lambda E|$ равны α , то $|C - \lambda E| = (-1)^k (\lambda - \alpha)^k$, где k – размер матрицы C , т.е. $\dim V_\alpha$.

Множитель $|B - \lambda E|$ в разложении $f_A(\lambda) = |B - \lambda E| \cdot |C - \lambda E|$ – это характеристический многочлен ограничения оператора A на подпространство U_{A_α} , т.е. на 1-компоненту оператора $A_\alpha = A - \alpha E$. Ограничение A_α на свою 1-компоненту невырождено, т.е. $A_\alpha x \neq 0$ для ненулевых $x \in U_{A_\alpha}$. Это значит, что $Ax \neq \alpha x$, если $x \in U_{A_\alpha}$ и $x \neq 0$, т.е. подпространство U_{A_α} не содержит собственных векторов оператора A , принадлежащих α . Поэтому α не является корнем многочлена $|B - \lambda E|$. Из равенства $f_A(\lambda) = |B - \lambda E| \cdot |C - \lambda E| = |B - \lambda E| \cdot (-1)^k (\lambda - \alpha)^k$ заключаем, что $k = \dim V_\alpha$ есть кратность корня α в характеристическом многочлене оператора A .

Для завершения доказательства остается применить предположение индукции к ограничению оператора A на подпространство U_{A_α} , размерность которого меньше размерности V . □

Отметим, что теорема о корневом разложении сводит задачу об устройстве «простейшей» матрицы произвольного линейного оператора к случаю нильпотентных операторов.

Следствие (Камилл Жордан, 1870)

Если $A: V \rightarrow V$ – такой линейный оператор, что $\text{Spec } A$ содержится в поле скаляров, то в V можно выбрать базис из его корневых векторов. В этом базисе матрица оператора A блочно-диагональна.

Число диагональных блоков равно $|\text{Spec } A|$, размер блока, отвечающего $\alpha \in \text{Spec } A$, равен кратности k корня α в характеристическом многочлене оператора A , а блок равен $\alpha E_k + A_\alpha$, где A_α есть матрица ограничения оператора $A_\alpha = A - \alpha \mathcal{E}$ на его 0-компоненту.

Если $\text{Spec } A = \{\alpha_1, \alpha_2, \dots, \alpha_t\}$ и

$$f_A(\lambda) = \pm(\lambda - \alpha_1)^{k_1}(\lambda - \alpha_2)^{k_2} \dots (\lambda - \alpha_t)^{k_t},$$

то

$$A = \begin{pmatrix} \alpha_1 E_{k_1} + A_{\alpha_1} & O & \dots & O \\ O & \alpha_2 E_{k_2} + A_{\alpha_2} & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & \alpha_t E_{k_t} + A_{\alpha_t} \end{pmatrix}.$$

Пусть $f = a_p x^p + a_{p-1} x^{p-1} + \dots + a_1 x + a_0$ – многочлен над полем F , а \mathcal{A} – линейный оператор на пространстве V над F . *Значением многочлена f от \mathcal{A}* называется оператор

$$f(\mathcal{A}) = a_p \mathcal{A}^p + a_{p-1} \mathcal{A}^{p-1} + \dots + a_1 \mathcal{A} + a_0 \mathcal{E},$$

где \mathcal{E} – единичный оператор. Многочлен f *аннулирует оператор \mathcal{A}* , если $f(\mathcal{A}) = \mathcal{O}$, где \mathcal{O} – нулевой оператор.

Предложение (существование аннулирующих многочленов)

Для любого линейного оператора на конечномерном пространстве существует ненулевой аннулирующий многочлен.

Доказательство. Если $\dim V = n$, то размерность пространства всех линейных операторов на пространстве V равна n^2 . Поэтому для любого оператора \mathcal{A} система $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \dots, \mathcal{A}^{n^2}$ линейно зависима. Значит, есть такие скаляры $a_0, a_1, a_2, \dots, a_{n^2}$, не все равные 0, что

$$a_0 \mathcal{E} + a_1 \mathcal{A} + a_2 \mathcal{A}^2 + \dots + a_{n^2} \mathcal{A}^{n^2} = \mathcal{O}.$$

Если $f := a_0 + a_1 x + a_2 x^2 + \dots + a_{n^2} x^{n^2}$, то $f \neq 0$ и $f(\mathcal{A}) = \mathcal{O}$. □

Минимальным многочленом линейного оператора \mathcal{A} называется ненулевой многочлен наименьшей степени, аннулирующий \mathcal{A} .

Предложение (свойство минимального многочлена)

Минимальный многочлен линейного оператора делит любой его аннулирующий многочлен.

Доказательство. Пусть f – произвольный аннулирующий многочлен оператора \mathcal{A} , а m – минимальный многочлен. Поделим с остатком: $f = qm + r$, где $\deg r < \deg m$. Подставив в это равенство оператор \mathcal{A} , получим $f(\mathcal{A}) = q(\mathcal{A})m(\mathcal{A}) + r(\mathcal{A})$, откуда $r(\mathcal{A}) = \mathcal{O}$. Но m – ненулевой многочлен наименьшей степени, аннулирующий \mathcal{A} , следовательно, $r = 0$ и m делит f . □

Из предложения следует, что минимальный многочлен оператора единствен с точностью до ассоциированности.

Теорема (о минимальном многочлене)

Пусть $\mathcal{A}: V \rightarrow V$ – линейный оператор, $\text{Spec } \mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_t\}$ содержится в поле скаляров и s_i – высота корневого подпространства V_{α_i} . Минимальный многочлен оператора \mathcal{A} есть

$$m(\lambda) := (\lambda - \alpha_1)^{s_1} (\lambda - \alpha_2)^{s_2} \dots (\lambda - \alpha_t)^{s_t}.$$

Доказательство. Сперва поймем, что многочлен $m(\lambda)$ аннулирует \mathcal{A} . Пользуясь корневым разложением, представим произвольный вектор $\mathbf{x} \in V$ как $\mathbf{x} = \sum_{i=1}^t \mathbf{x}_i$, где $\mathbf{x}_i \in V_{\alpha_i}$. По определению корневого подпространства имеем $(\mathcal{A} - \alpha_i \mathcal{E})^{s_i} \mathbf{x}_i = \mathbf{0}$. Поэтому

$$m(\mathcal{A})\mathbf{x} = \sum_{i=1}^t m(\mathcal{A})\mathbf{x}_i = \sum_{i=1}^t \prod_{j \neq i} (\mathcal{A} - \alpha_j \mathcal{E})^{s_j} \cdot (\mathcal{A} - \alpha_i \mathcal{E})^{s_i} \mathbf{x}_i = \mathbf{0}.$$

Здесь использовано, что многочлены от оператора перестановочны.

По доказанному выше минимальный многочлен должен делить $m(\lambda)$. Проверим, что никакой собственный делитель этого многочлена не аннулирует оператор \mathcal{A} .

Любой собственный делитель многочлена $m(\lambda)$ имеет вид

$$n(\lambda) := (\lambda - \alpha_1)^{q_1} (\lambda - \alpha_2)^{q_2} \cdots (\lambda - \alpha_t)^{q_t},$$

где $q_i \leq s_i$ для всех $i = 1, \dots, t$ и $q_j < s_j$ для какого-то j . Возьмем вектор $\mathbf{y} \in \text{Ker}(\mathcal{A} - \alpha_j \mathcal{E})^{s_j} \setminus \text{Ker}(\mathcal{A} - \alpha_j \mathcal{E})^{s_j - 1}$ – он существует по определению высоты корневого подпространства. Тогда $(\mathcal{A} - \alpha_j \mathcal{E})^{q_j} \mathbf{y} \neq \mathbf{0}$, откуда

$$n(\mathcal{A})\mathbf{y} = \prod_{i \neq j} (\mathcal{A} - \alpha_i \mathcal{E})^{q_i} \cdot (\mathcal{A} - \alpha_j \mathcal{E})^{q_j} \mathbf{y} \neq \mathbf{0},$$

ибо при $i \neq j$ оператор $\mathcal{A} - \alpha_i \mathcal{E}$ невырожден на подпространстве V_{α_i} . \square

Ясно, что высота корневого подпространства не больше его размерности. Поэтому характеристический многочлен $(\lambda - \alpha_1)^{k_1} (\lambda - \alpha_2)^{k_2} \cdots (\lambda - \alpha_t)^{k_t}$ делится на минимальный многочлен $(\lambda - \alpha_1)^{s_1} (\lambda - \alpha_2)^{s_2} \cdots (\lambda - \alpha_t)^{s_t}$. Из этого наблюдения вытекает красивое следствие:

Теорема Гамильтона–Кэли (Фердинанд Георг Фробениус, 1878)

Характеристический многочлен оператора аннулирует оператор.

Это – мощное тождество, неочевидное даже для 3×3 -матриц.

Другим важным следствием теоремы о минимальном многочлене является критерий диагоналируемости линейного оператора. Напомним, что *диагоналируемыми* или *приводимыми к диагональному виду* называют операторы, допускающие базис из собственных векторов.

Теорема (критерий диагоналируемости)

Оператор, спектр которого лежит в поле скаляров, диагоналируем, если и только если его минимальный многочлен не имеет кратных корней.

Доказательство. Если минимальный многочлен не имеет кратных корней, то высота каждого корневого подпространства равна 1, а тогда все корневые вектора будут собственными. Поэтому у такого оператора есть базис из собственных векторов.

Обратно, если оператор \mathcal{A} диагоналируем и $\text{Spec } \mathcal{A} = \{\alpha_1, \alpha_2, \dots, \alpha_t\}$, то используя диагональную матрицу оператора \mathcal{A} , легко подсчитать, что многочлен $(\lambda - \alpha_1)(\lambda - \alpha_2) \cdots (\lambda - \alpha_t)$ аннулирует \mathcal{A} . \square

Отмечавшийся ранее факт, что операторы, у которых характеристический многочлен не имеет кратных корней, диагоналируемы, является простейшим частным случаем доказанного сейчас критерия.

Как использовать критерий диагоналируемости? Вычисление минимального многочлена – непростая задача, удобной явной формулы (как для характеристического многочлена) в этом случае нет.

Оказывается, над полем нулевой характеристики можно проверить, имеются ли у минимального многочлена кратные корни, не вычисляя его! Применим технику отделения кратных множителей из §1.6.

Пусть f – характеристический многочлен линейного оператора \mathcal{A} , определенного на векторном пространстве над полем характеристики 0.

Вычислим многочлен $g := \frac{f}{\text{НОД}(f, f')}$. Если $f = (\lambda - \alpha_1)^{k_1} \cdots (\lambda - \alpha_t)^{k_t}$,

то $g = (\lambda - \alpha_1) \cdots (\lambda - \alpha_t)$ – многочлен без кратных корней с теми же корнями, что у f . Теперь подсчитаем значение $g(\mathcal{A})$. Если $g(\mathcal{A}) = \mathcal{O}$, то g – минимальный многочлен оператора \mathcal{A} . В этом случае у минимального многочлена нет кратных корней. Если же $g(\mathcal{A}) \neq \mathcal{O}$, то у минимального многочлена есть кратные корни.

Отметим еще, что иногда удастся вывести отсутствие у минимального многочлена оператора \mathcal{A} кратных корней из свойств самого оператора. Например, если $\mathcal{A}^2 = \mathcal{A}$ (*идемпотентный* оператор), то \mathcal{A} аннулируется многочленом $\lambda^2 - \lambda = \lambda(\lambda - 1)$ без кратных корней. Поэтому минимальный многочлен идемпотентного оператора не имеет кратных корней.