Линейная алгебра и геометрия

Лекция 9

Тарас Евгеньевич Панов

Механико-математический факультет МГУ, 1 курс, весна 2021 г.

Текст лекций и другие учебные материалы доступны на странице:
http://higeom.math.msu.su/people/taras/teaching

12 марта 2021 г.

Теорема

а) *Оператор* $\mathcal{A}\colon V \to V$ *в нетривиальном пространстве над полем* $\mathbb{C}$ *имеет инвариантное подпространство размерности 1.*

б) *Оператор* $\mathcal{A}\colon V \to V$ *в нетривиальном пространстве над полем* $\mathbb{R}$ *имеет инвариантное подпространство размерности 1 или 2.*

Доказательство

Для доказательства а) заметим, что так как поле $\mathbb{C}$ алгебраически замкнуто, характеристический многочлен $P_{\mathcal{A}}(t)$ имеет корень λ. Значит оператор $\mathcal{A}$ имеет собственный вектор $\boldsymbol{v}$, т. е. $\mathcal{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ и $\langle \boldsymbol{v} \rangle$ — одномерное инвариантное подпространство.

Доказательство (продолжение).

Докажем б). Если характеристический многочлен $P_{\mathcal{A}}(t)$ имеет вещественный корень, то мы получаем одномерное инвариантное подпространство. Предположим, что $P_{\mathcal{A}}(t)$ не имеет вещественных корней. Пусть $\lambda + i\mu$ — комплексный корень, $\mu \neq 0$. Тогда $\lambda + i\mu$ — собственное значение комплексифицированного оператора $\mathcal{A}_{\mathbb{C}}$ (напомним, что в подходящих базисах матрицы операторов $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ совпадают). Возьмём соответствующий собственный вектор $\boldsymbol{u} + i\boldsymbol{v} \in V_{\mathbb{C}}$. Тогда

$$\mathcal{A}\boldsymbol{u} + i\mathcal{A}\boldsymbol{v} = \mathcal{A}_{\mathbb{C}}(\boldsymbol{u} + i\boldsymbol{v}) = (\lambda + i\mu)(\boldsymbol{u} + i\boldsymbol{v}) = (\lambda\boldsymbol{u} - \mu\boldsymbol{v}) + i(\mu\boldsymbol{u} + \lambda\boldsymbol{v}).$$

Следовательно, $\mathcal{A}\boldsymbol{u} = \lambda\boldsymbol{u} - \mu\boldsymbol{v}$ и $\mathcal{A}\boldsymbol{v} = \mu\boldsymbol{u} + \lambda\boldsymbol{v}$, и линейная оболочка $\langle\boldsymbol{u}, \boldsymbol{v}\rangle \subset V$ является инвариантным подпространством для $\mathcal{A}$. $\square$

Предложение

Размерность собственного подпространства $V_\lambda = \mathrm{Ker}(\mathcal{A} - \lambda \cdot \mathrm{id})$ *не превосходит кратности* λ *как корня характеристического многочлена.*

Доказательство.

Пусть $\dim V_\lambda = k$. Выберем базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ в пространстве V_λ и дополним его до базиса в V. Так как $\mathcal{A}\boldsymbol{e}_i = \lambda \boldsymbol{e}_i$ при $i = 1, \ldots, k$, матрица оператора $\mathcal{A}$ в выбранном базисе имеет вид

$$A = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & \widetilde{A} \end{array}\right).$$

Тогда $P_\mathcal{A}(t) = \det(A - tE) = (\lambda - t)^k \det(\widetilde{A} - tE) = (\lambda - t)^k P_{\widetilde{\mathcal{A}}}(t)$, где $\widetilde{\mathcal{A}}$ — фактор-оператор. Отсюда вытекает, что кратность корня λ не меньше $k = \dim V_\lambda$. □

Теорема (Гамильтона–Кэли)

Характеристический многочлен $P_{\mathcal{A}}(t)$ *оператора* $\mathcal{A}\colon V \to V$ *аннулирует этот оператор, т. е.* $P_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$*.*

Мы приведём два доказательства этого фундаментального факта. Первое доказательство более элементарное, но использует специальный трюк. Второе доказательство идейно проще, но использует понятие факторпространства.

Первое доказательство

Для квадратной матрицы M обозначим через $\widehat{M}$ матрицу того же размера, на ij-м месте которой стоит алгебраическое дополнение ji-го элемента матрицы M. Как известно из курса алгебры, имеет место тождество

$$M\widehat{M} = \det M \cdot E.$$

Первое доказательство (продолжение)

Теперь возьмём в качестве M матрицу $A - tE$, где A — матрица оператора $\mathcal{A}$ в произвольном базисе. Тогда

$$(A - tE)(\widehat{A - tE}) = \det(A - tE) \cdot E = P_{\mathcal{A}}(t)E. \qquad (1)$$

По определению элементы матрицы $\widehat{A - tE}$ являются многочленами от t степени не выше $n - 1$, где $n = \dim V$. Следовательно, эту матрицу можно записать в виде

$$\widehat{A - tE} = B_0 + tB_1 + t^2 B_2 + \ldots + t^{n-1} B_{n-1},$$

где B_i — числовые матрицы. Подставив это разложение вместе с разложением характеристического многочлена $P_{\mathcal{A}}(t) = a_0 + a_1 t + \ldots + a_n t^n$ в формулу (1), получим

$$(A - tE)(B_0 + tB_1 + t^2 B_2 + \ldots + t^{n-1} B_{n-1}) = (a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n)E.$$

Доказательство (окончание).

$$(A-tE)(B_0+tB_1+t^2B_2+\ldots+t^{n-1}B_{n-1})=(a_0+a_1t+a_2t^2+\ldots+a_nt^n)E.$$

Приравнивая коэффициенты при различных степенях t, получим

$$\begin{aligned}
AB_0&=a_0E,\\
-B_0+AB_1&=a_1E,\\
-B_1+AB_2&=a_2E,\\
&\ldots\\
-B_{n-2}+AB_{n-1}&=a_{n-1}E\\
-B_{n-1}&=a_nE
\end{aligned}$$

Умножив слева обе части второго равенства на A, третьего — на A^2, и т. д., и сложив все полученные равенства, получим

$$0=a_0E+a_1A+a_2A^2+\ldots+a_nA^n.$$

В правой части стоит результат подстановки A в характеристический многочлен. □

Второе доказательство

Вначале докажем теорему над алгебраически замкнутым полем, например, над полем $\mathbb{C}$. Проведём индукцию по размерности V. Если V одномерно, то $\mathcal{A} = \lambda \cdot \mathrm{id}$ — умножение на скаляр λ. Тогда $P_{\mathcal{A}}(t) = \lambda - t$, а значит $P_{\mathcal{A}}(\mathcal{A}) = \lambda \cdot \mathrm{id} - \mathcal{A} = \mathcal{O}$.

Пусть теперь $\dim V = n$, и предположим, что теорема доказана для пространств размерности $n-1$. Выберем инвариантное одномерное подпространство U для $\mathcal{A}$; это подпространство порождено собственным вектором $\boldsymbol{u}$ с собственным значением λ. Дополним вектор $\boldsymbol{u}$ до базиса. В этом базисе матрица оператора $\mathcal{A}$ имеет вид

$$A = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & \widetilde{A} \end{array}\right),$$

где $\widetilde{A}$ — матрица фактор-оператора $\widetilde{\mathcal{A}}$, действующего в фактор-пространстве V/U.

Второе доказательство (продолжение)

$$A = \left(\begin{array}{c|c} \lambda & * \\ \hline 0 & \widetilde{A} \end{array}\right),$$

где $\widetilde{A}$ — матрица фактор-оператора $\widetilde{\mathcal{A}}$, в фактор-пространстве V/U. Отсюда вытекает, что

$$P_{\mathcal{A}}(t) = \det(A - tE) = (\lambda - t)\det(\widetilde{A} - tE) = (\lambda - t)P_{\widetilde{\mathcal{A}}}(t),$$

где $P_{\widetilde{\mathcal{A}}}(t)$ — характеристический многочлен фактор-оператора. Так как $\dim V/U = n - 1$, по предположению индукции $P_{\widetilde{\mathcal{A}}}(\widetilde{\mathcal{A}}) = \mathcal{O}$. По определению фактор-оператора это означает, что для любого вектора $\boldsymbol{v} \in V$ имеем $P_{\widetilde{\mathcal{A}}}(\mathcal{A})\boldsymbol{v} \in U$, т. е. $P_{\widetilde{\mathcal{A}}}(\mathcal{A})\boldsymbol{v} = \mu\boldsymbol{u}$ для некоторого μ. Следовательно,

$$P_{\mathcal{A}}(\mathcal{A})\boldsymbol{v} = (\lambda \cdot \mathrm{id} - \mathcal{A})P_{\widetilde{\mathcal{A}}}(\mathcal{A})\boldsymbol{v} = (\lambda \cdot \mathrm{id} - \mathcal{A})(\mu\boldsymbol{u}) = \mathbf{0},$$

так как $\mathcal{A}\boldsymbol{u} = \lambda\boldsymbol{u}$.

Итак, $P_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$ для операторов в комплексных пространствах.

Второе доказательство (окончание).

Для доказательства теоремы над полем $\mathbb{R}$ воспользуемся комплексификацией $\mathcal{A}_{\mathbb{C}}$ оператора $\mathcal{A}$. Так как в соответствующих базисах пространств V и $V_{\mathbb{C}}$ матрицы операторов $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ совпадают, мы имеем $P_{\mathcal{A}}(t) = P_{\mathcal{A}_{\mathbb{C}}}(t)$ и $P_{\mathcal{A}}(A) = P_{\mathcal{A}_{\mathbb{C}}}(A) = 0$. □

2.7. Диагонализируемые операторы

Определение

Оператор $\mathcal{A}$ называется **диагонализируемым**, если существует базис, в котором матрица этого оператора диагональна.

По определению матрицы оператора, базис, в котором матрица оператора диагональна, состоит из собственных векторов. Поэтому оператор диагонализируем тогда и только тогда, когда для него существует базис из собственных векторов.

Теорема (критерий диагонализируемости)

Оператор $\mathcal{A}$ в n-мерном пространстве V диагонализируем тогда и только тогда, когда его характеристический многочлен имеет в точности n корней (с учётом кратностей), и размерность каждого собственного подпространства V_λ равна кратности корня λ.

Для доказательства теоремы нам понадобится лемма.

Лемма

Собственные подпространства $V_{\lambda_1}, \ldots, V_{\lambda_k}$, соответствующие попарно различным собственным значениям $\lambda_1, \ldots, \lambda_k$ оператора $\mathcal{A}$, образуют прямую сумму $V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$.

Доказательство леммы.

Проведём индукцию по k. При $k = 1$ утверждение очевидно.

По определению прямой суммы мы должны проверить, что соотношение

$$\boldsymbol{v}_1 + \ldots + \boldsymbol{v}_k = \mathbf{0}, \tag{2}$$

где $\boldsymbol{v}_i \in V_{\lambda_i}$, влечёт $\boldsymbol{v}_1 = \ldots = \boldsymbol{v}_k = \mathbf{0}$. Применив к (2) оператор $\mathcal{A}$, получим

$$\lambda_1 \boldsymbol{v}_1 + \ldots + \lambda_k \boldsymbol{v}_k = \mathbf{0}$$

Умножим (2) на λ_k и вычтем из предыдущего соотношения:

$$(\lambda_1 - \lambda_k)\boldsymbol{v}_1 + \ldots + (\lambda_{k-1} - \lambda_k)\boldsymbol{v}_{k-1} = \mathbf{0}.$$

По предположению индукции получаем $(\lambda_1 - \lambda_k)\boldsymbol{v}_1 = \ldots = (\lambda_{k-1} - \lambda_k)\boldsymbol{v}_{k-1} = \mathbf{0}$. Так как по условию $\lambda_1 - \lambda_k \neq 0, \ldots, \lambda_{k-1} - \lambda_k \neq 0$, получаем $\boldsymbol{v}_1 = \ldots = \boldsymbol{v}_{k-1} = \mathbf{0}$. Тогда из (2) следует, что и $\boldsymbol{v}_k = \mathbf{0}$. □

Сформулируем одно полезное следствие этой леммы.

Следствие

Собственные векторы оператора, отвечающие различным собственным значениям, линейно независимы.

Доказательство теоремы.

Предположим, что оператор $\mathcal{A}$ диагонализируем. Пусть на диагонали матрицы D оператора $\mathcal{A}$ стоят числа $\lambda_1, \ldots, \lambda_k$, причём число λ_i присутствует r_i раз. Тогда мы имеем $P_{\mathcal{A}}(t) = \det(D - tE) = \prod_{i=1}^{k}(\lambda_i - t)^{r_i}$. Следовательно, многочлен $P_{\mathcal{A}}(t)$ имеет $\sum_{i=1}^{k} r_i = n$ корней, и каждому корню λ_i соответствует r_i линейно независимых собственных векторов, т. е. $\dim V_{\lambda_i} = r_i$.

Предположим теперь, что многочлен $P_{\mathcal{A}}(t)$ имеет различные корни $\lambda_1, \ldots, \lambda_k$, причём кратность корня λ_i равна r_i, $\sum_{i=1}^{k} r_i = n$ и $\dim V_{\lambda_i} = r_i$. Согласно лемме, пространства $V_{\lambda_1}, \ldots, V_{\lambda_k}$ образуют прямую сумму, а по условию сумма их размерностей равна $n = \dim V$. Следовательно, $V = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_k}$. Выбрав базис в каждом из подпространств V_{λ_i} и взяв объединение этих базисов, мы получим базис пространства V, состоящий из собственных векторов. Итак, оператор $\mathcal{A}$ диагонализируем. $\square$

Следствие

Пусть характеристический многочлен $P_{\mathcal{A}}(t)$ имеет $n = \dim V$ различных корней. Тогда оператор $\mathcal{A}$ диагонализируем.

Набор собственных значений оператора $\mathcal{A}$ часто называют его спектром (эта терминология будет прояснена в курсе функционального анализа, когда будут рассматриваться операторы с непрерывным спектром в бесконечномерных пространствах). Если все собственные значения имеют кратность 1 как корни характеристического многочлена, то говорят о простом спектре. Таким образом, операторы с простым спектром диагонализируемы. Появление кратных корней является «особенностью», которая устраняется произвольно малым возмущением коэффициентов матрицы оператора. Поэтому над полем $\mathbb{C}$ «почти все» операторы диагонализируемы.

Пример

1. Оператор, заданный матрицей $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ в стандартном базисе $\mathbb{R}^2$, не диагонализируем, так как его характеристический многочлен t^2+1 не имеет вещественных корней.
Однако тот же оператор в $\mathbb{C}^2$ диагонализируем: в базисе $\boldsymbol{f}_1 = (1\; i)$, $\boldsymbol{f}_2 = (1\; -i)$ его матрица $\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$ диагональна.

2. Оператор, заданный матрицей $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, не диагонализируем ни над каким полем по другой причине: его характеристический многочлен $(t-1)^2$ имеет корень 1 кратности 2, но при этом размерность соответствующего собственного подпространства равна 1 (вектор $\boldsymbol{e}_2 = (0\,1)$ не является собственным).