Операторы - 2

Часть 1. Собственные и корневые векторы

Пусть A — оператор в векторном пространстве V над полем K.

Определение 1.1. Вектор $v \in V$ называется собственным вектором оператора $A : V \to V$, если $Av = \lambda v$, $\lambda \in K$. Число $\lambda \in K$, для которого существует нетривиальный собственный вектор, называется собственным значением оператора A.

Лемма 1.2. *Множество собственных векторов оператора A с собственным значением λ является векторным подпространством в V.*

Доказательство: если $Av_1 = \lambda v_1$ и $Av_2 = \lambda v_2$, то $A(v_1 + v_2) = Av_1 + Av_2 = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$. □

Лемма 1.3. *Число λ является собственным значением оператора $A \iff \chi_A(\lambda) = 0$.*

Доказательство: из определения собственного вектора следует, что λ — собственное значение $\iff$ $\operatorname{Ker}(\lambda \operatorname{Id} - A) \neq 0 \iff \det(\lambda \operatorname{Id} - A) = \chi_A(\lambda) = 0$. □

Следствие 1.4. *Если характеристический многочлен $\chi_A(t)$ имеет $n = \dim V$ различных корней, то существует базис, в котором матрица оператора A диагональна.*

Доказательство: пусть $\lambda_1, \dots, \lambda_n$ — корни характеристического многочлена, а $v_1, \dots, v_n$ — соответствующие собственные векторы. Тогда $v_1, \dots, v_n$ — линейно независимы. Действительно, если $\sum x_i v_i = 0$, то $\forall i$ имеем $0 = \prod_{j \neq i}(A - \lambda_j \operatorname{Id})(\sum_{k=1}^n x_k v_k) = \sum_{k=1}^n x_k \prod_{j \neq i}(\lambda_k - \lambda_j) v_k = x_i \prod_{j \neq i}(\lambda_i - \lambda_j) v_i$, значит $x_i = 0$. Следовательно, векторы $v_1, \dots, v_n$ образуют базис. Наконец, легко видеть, что матрица оператора A относительно этого базиса равна $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$. □

Определение 1.5. Вектор $v \in V$ называется корневым вектором для оператора A относительно $\lambda \in K$, если $(A - \lambda \operatorname{Id})^k v = 0$ для какого-либо числа $k \in \mathbb{Z}$.

Лемма 1.6. *Множество V_λ корневых векторов оператора A относительно числа $\lambda \in K$ является векторным подпространством, инвариантным относительно оператора A.*

Доказателство: если $(A - \lambda \operatorname{Id})^{k_1} v_1 = 0$ и $(A - \lambda \operatorname{Id})^{k_2} v_2 = 0$, то $(A - \lambda \operatorname{Id})^{k_1 + k_2}(v_1 + v_2) = 0$, а если $(A - \lambda \operatorname{Id})^k v = 0$, то $(A - \lambda \operatorname{Id})^k Av = A(A - \lambda \operatorname{Id})^k v = 0$. □

Часть 2. Жорданова нормальная форма

В этой части лекции мы будем предполагать, что поле K алгебраически замкнуто, то есть всякий многочлен с коэффициентами в K имеет корень в K. Примером алгебраически замкнутого поля является поле $\mathbb{C}$.

Лемма 2.1. *Векторное пространство V раскладывается в прямую сумму корневых подпространств.*

Доказательство: рассмотрим цепочки векторных подпространств $V \supset \operatorname{Im}(A - \lambda \operatorname{Id}) \supset \operatorname{Im}(A - \lambda \operatorname{Id})^2 \supset \dots$ и $0 \subset \operatorname{Ker}(A - \lambda \operatorname{Id}) \subset \operatorname{Ker}(A - \lambda \operatorname{Id})^2 \subset \dots$ (вложения очевидны). Размерности в первой цепочке уменьшаются, а во второй — увеличиваются. Кроме того, $\dim \operatorname{Ker}(A - \lambda \operatorname{Id})^p + \dim \operatorname{Im}(A - \lambda \operatorname{Id})^p = \dim V$, следовательно, существует $k \in \mathbb{Z}$, такое что размерности (а значит, и подпространства) в цепочках при $p \geqslant k$ не меняются. Иначе говоря, при $p \geqslant k$ имеем $\operatorname{Ker}(A - \lambda \operatorname{Id})^p = \operatorname{Ker}(A - \lambda \operatorname{Id})^k$ и $\operatorname{Im}(A - \lambda \operatorname{Id})^p = \operatorname{Im}(A - \lambda \operatorname{Id})^k$. Ясно, что первое из этих подпространств совпадает с корневым подпространством V_λ. Обозначим второе подпространство через $\widehat{V}_\lambda$ и покажем, что $V = V_\lambda \oplus \widehat{V}_\lambda$. Заметим, что $\dim V_\lambda + \dim \widehat{V}_\lambda = \dim V$, поэтому достаточно проверить, что $V_\lambda \cap \widehat{V}_\lambda = 0$. Действительно, мы знаем, что $(A - \lambda \operatorname{Id})\widehat{V}_\lambda = \widehat{V}_\lambda$, то есть оператор $(A - \lambda \operatorname{Id})$ на $\widehat{V}_\lambda$ сюрьективен, следовательно обратим, следовательно все его степени обратимы, следовательно не существует вектора $0 \neq v \in \widehat{V}_\lambda$, такого что $(A - \lambda \operatorname{Id})^m v = 0$.

Таким образом мы показали, что $\forall\lambda \in K$ имеем разложение в прямую сумму $V = V_\lambda \oplus \widehat{V}_\lambda$. Доакажем теперь утверждение теоремы индукцией по $\dim V$. В случае $\dim V = 0$ доказывать нечего. Пусть теперь $\dim V > 0$. Преположим, что λ — собственное значение оператора A (оно существует, так как в силу алгебраической замкнутости поля K многочлен $\chi_A(t)$ имеет корень). Тогда $\dim V_\lambda > 0$, следовательно $\dim \widehat{V}_\lambda < \dim V$. Кроме того, разложение $V = V_\lambda \oplus \widehat{V}_\lambda$ по построению инвариантно относительно оператора A. Рассмотрим ограничение A на $\widehat{V}_\lambda$. Тогда по предположению индукции $\widehat{V}_\lambda = \oplus_{i=1}^m (\widehat{V}_\lambda)_{\lambda_i}$ — прямая сумма корневых подпространств. Легко видеть, что $(\widehat{V}_\lambda)_{\lambda_i}$ является корневым подпространством оператора A в V и разложение $V = V_\lambda \oplus (\oplus_{i=1}^m (\widehat{V}_\lambda)_{\lambda_i})$ — искомое. □

Квадратная матрица размера $k \times k$, у которой на диагонали стоят числа λ, над диагональю — единицы, а в остальных местах — нули, называется жордановой клеткой и обозначается $J_k(\lambda)$.

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

Теорема 2.2. *Пусть A — оператор в V. Тогда существует базис, в котором матрица оператора A — блочно диагональная с жордановыми клетками на диагонали.*

Доказательство: разложим V в сумму корневых подпространств. Так как корневые подпространства инвариантны относительно A, то достаточно построить жорданов базис для ограничения A на любое корневое пространство V_λ. Более того, ясно, что достаточно построить жорданов базис для оператора $A - \lambda\,\mathrm{Id}$. Таким образом, общее утверждение сводится к случаю, когда $A^k = 0$ для некоторого $k \geqslant 0$.

Преположим, что $A^{k-1} \neq 0$. Рассмотрим цепочку $0 \subset \operatorname{Ker} A \subset \operatorname{Ker} A^2 \subset \dots \subset \operatorname{Ker} A^{k-1} \subset \operatorname{Ker} A^k = V$ и докажем, что существует последовательность подпространств $W_p \subset \operatorname{Ker} A^p$, такая что условие

$$(*_p) \qquad \operatorname{Ker} A^p = [W_p \oplus A(W_{p+1}) \oplus \dots \oplus A^{k-p}(W_k)] \oplus \operatorname{Ker} A^{p-1}$$

выполняется при $p = 1, \dots, k$. Действительно, воспользуемся убывающей индукцией по p. При $p = k$ утверждение очевидно. Предположим, что $1 < p < k$ и последовательность пространств $W_p, \dots, W_k$ такова, что $(*_p)$ выполняется. Заметим, что $\left[W_p \oplus A(W_{p+1}) \oplus \dots \oplus A^{k-p}(W_k)\right] \cap \operatorname{Ker} A = 0$, следовательно

$$A(W_p) \oplus A^2(W_{p+1}) \oplus \dots \oplus A^{k-p+1}(W_k) = A(W_p \oplus A(W_{p+1}) \oplus \dots \oplus A^{k-p}(W_k)) \subset A(\operatorname{Ker} A^p) \subset \operatorname{Ker} A^{p-1}.$$

Кроме того, $\left[A(W_p) \oplus A^2(W_{p+1}) \oplus \dots \oplus A^{k-p+1}(W_k)\right] \cap \operatorname{Ker} A^{p-2} = 0$, следовательно существует W_{p-1}, так что выполняется условие $(*_{p-1})$.

Пусть теперь $W_1, \dots, W_k$ — последовательность подпространств, для которой условие $(*_p)$ выполнено при всех $1 \leqslant p \leqslant k$. Подставив равенства $(*_p)$ одно в другое, получим

$$V = \operatorname{Ker} A^k = W_k \oplus [W_{k-1} \oplus A(W_k)] \oplus [W_{k-2} \oplus A(W_{k-1}) \oplus A^2(W_k)] \oplus \dots \oplus [W_1 \oplus A(W_2) \oplus \dots \oplus A^{k-1}(W_k)].$$

Перегруппировывая члены, получаем

$$V = [W_k \oplus A(W_k) \oplus \dots \oplus A^{k-1}(W_k)] \oplus [W_{k-1} \oplus A(W_{k-1}) \oplus \dots \oplus A^{k-2}(W_{k-1})] \oplus \dots \oplus [W_2 \oplus A(W_2)] \oplus W_1.$$

Выберем теперь базис $w_{p,i}$ в каждом из пространств W_p. Заметим, что по построению $A^p(W_p) = 0$, но $W_p \cap \operatorname{Ker} A^{p-1} = 0$, поэтому векторы $e_{p,i,j} = A^j w_{p,i}$ $(j = 0, \dots, p-1)$ образуют базис в V. Более того, для любых p и i подпространство в V, порожденное векторами $e_{p,i,0}, \dots, e_{p,i,p-1}$ инвариантно относительно A, а матрица A в этом базисе является жордановой клеткой $J_p(0)$. □

Замечание 2.3. Если поле K не является алгебраически замкнутым, то жорданова базиса, вообще говоря, может и не быть. Например, оператор умножения на i в $\mathbb{C} \cong \mathbb{R}^2$ не имеет жорданова базиса над $\mathbb{R}$.

Как видно из доказательства теоремы жорданова базис не является единственным. Однако, количество жордановых клеток вида $J_k(\lambda)$ не зависит от его выбора.

Лемма 2.4. *Пусть $n_{\lambda,k}$ — количество жордановых клеток вида $J_k(\lambda)$ в матрице оператора A в некоторм жордановом базисе. Тогда $\sum_{p=1}^k p n_{\lambda,p} + \sum_{p=k+1}^\infty k n_{\lambda,p} = \dim \operatorname{Ker}(A - \lambda\,\mathrm{Id})^k$ для всех λ и k.*

Доказательство: прямое вычисление.

Замечание 2.5. Жорданов базис очень удобен для вычислений с оператогм A. Например, легко видеть, что

$$\operatorname{tr} A = \sum_{\lambda,k} k n_{\lambda,k}(A)\lambda, \quad \det A = \prod_{\lambda,k} \lambda^{k n_{\lambda,k}(A)}, \quad \chi_A(t) = \prod_{\lambda,k} (t-\lambda)^{k n_{\lambda,k}(A)},$$

где $n_{\lambda,k}(A)$ — количество жордановых клеток вида $J_k(\lambda)$. Отсюда видно, что характеристический многочлен *не определяет* жорданову нормальную форму оператора.

Задание оператора A в векторном пространстве V равносильно заданию в V структуры модуля над кольцом многочленов $K[t]$ (многочлен $f(t) = a_n t^n + \dots + a_0$ действует оператором $f(A) = a_n A^n + \dots + a_0$). Легко видеть, что при таком соответствии оператору, состоящему из одной жордановой клетки $J_k(\lambda)$ соответствует модуль $K[t]/(t-\lambda)^k K[t]$. Поэтому теорему 2.2 можно переформулировать следующим способом.

Следствие 2.6. *Если K — алгебраически замкнутое поле и V — $K[t]$-модуль, такой что $\dim V < \infty$, то $V \cong \bigoplus_{\lambda,k} K[t]/(t-\lambda)^k K[t]$.*

Вот еще одна полезная переформулировка. Оператор называется полупростым, если его матрица диагональна в некотором базисе. Оператор A называется нильпотентным, если $A^m = 0$ для некоторого m.

Следствие 2.7. *Всякий оператор A представляется в виде $A = A_{ss} + A_n$, где A_{ss} — полупрост, A_n — нильпотентен и $A_{ss}A_n = A_nA_{ss}$.*

Замечание 2.8. Такое разложение называется жордановым. Можно показать, что оно единственно.

Часть 3. Теорема Гамильтона–Кэли

Теорема 3.1. *Пусть A — оператор в конечномерном векторном пространстве V над произвольным полем K (не обязательно алгебраически замкнутым). Тогда $\chi_A(A) = 0$.*

Доказательство: из теоремы 3.5 предыдущей лекции и определения характеристического многочлена следует, что

$$(\dagger) \qquad \chi_A(t)E = (\widehat{tE-A}) \cdot (tE-A) = (tE-A) \cdot (\widehat{tE-A})$$

(равенство в кольце $\mathrm{Mat}_{n\times n}(K[t]) \cong \mathrm{Mat}_{n\times n}(K)[t]$). Более того, это равенство выполняется также и в кольце $R = Z_A(\mathrm{Mat}_{n\times n}(K)[t]) = (Z_A(\mathrm{Mat}_{n\times n}))[t]$ (Z_A — централизатор оператора A). Действительно, ясно что $tE - A \in R$ и $\chi_A(t)E \in R$. Наконец, из ($\dagger$) следует, что $(\widehat{tE-A})$ коммутирует с $tE-A$, а значит и с A, следовательно $(\widehat{tE-A}) \in R$. Заметим теперь, что отображение

$$\varphi : R \to \mathrm{Mat}_{n\times n}(K), \qquad \textstyle\sum_p B_p t^p \mapsto \sum_p B_p A^p \qquad (B_p \in Z_A(\mathrm{Mat}_{n\times n}(K)))$$

является гомоморфизмом колец. Действительно,

$$\varphi(\textstyle\sum B_p t^p)\varphi(\sum C_q t^q) = (\sum B_p A^p)(\sum C_q A^q) = \sum B_p A^p C_q A^q = \sum B_p C_q A^p A^q = \sum B_p C_q A^{p+q} = \varphi(\sum B_p C_q t^{p+q}).$$

Применяя φ к ($\dagger$) получаем $\chi_A(A) = \varphi(\chi_A(t)E) = \varphi((\widehat{tE-A})) \cdot \varphi(tE-A) = \varphi((\widehat{tE-A})) \cdot (A-A) = 0.$ $\square$

Замечание 3.2. Если поле K — алгебраически замкнуто, то теорема Гамильтона–Кэли легко доказывается непосредственным вычислением оператора $\chi_A(A)$ в жордановом базисе.

Следствие 3.3. *Пусть $f_A : K[t] \to \mathrm{End}_K(A)$ — гомоморфизм, переводящий многочлен $\sum a_p t^p \in K[t]$ в оператор $\sum_p a_p A^p \in \mathrm{End}_K(A)$. Тогда $\chi_A(t) \in \mathrm{Ker}\, f_A$.*

Определение 3.4. Ненулевой многочлен минимальной степени со старшим коэффициентом 1 в идеале $\mathrm{Ker}\, f_A \subset K[t]$ называется минимальным многочленом оператора A и обозначается $m_A(t)$.

Примеры 3.5.
1. Если $m_A(t) = t - \lambda$, то $A = \lambda E$.
2. Если $m_A(t) = t^2 - t$, то A называется проектором или идемпотентным оператором.
3. Если $m_A(t) = t^n$, то A является нильпотентным оператором.
4. Если $m_A(t) = (t-1)^n$, то A называется унипотентным оператором.

Лемма 3.6. *Минимальный многочлен оператора A делит его характеристический многочлен.*

Доказательство: $\chi_A(t) = m_A(t)q(t) + r(t)$, где $\deg r(t) < \deg m_A(t)$, но $r(t) \in \mathrm{Ker}\, f_A$, поэтому $r(t) = 0$. $\square$

Следствие 3.7. *Если A — оператор в n-мерном пространстве, то $\deg m_A(t) \leqslant n$.*

Замечание 3.8. Можно доказать также, что характеристический многочлен делит некоторую степень минимального многочлена.