Линейная алгебра и геометрия

Лекция 12

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

23 марта 2021 г.

Теорема

Для любого оператора $\mathcal{A}$ в пространстве V над алгебраически замкнутым полем существует жорданов базис (в котором оператор имеет жорданову нормальную форму). Жорданова нормальная форма оператора единственна с точностью до перестановки блоков (клеток).

На языке матриц теорема Жордана означает, что любая квадратная комплексная матрица $A \in \mathrm{Mat}_n(\mathbb{C})$ подобна матрице J, состоящей из жордановых клеток, т. е. $J = C^{-1}AC$ для некоторой невырожденной матрицы C.

В отличие от жордановой формы, жорданов базис оператора далеко не единствен. Например, для тождественного оператора id любой базис будет жордановым.

Теорема Жордана не имеет места над полем $\mathbb{R}$. Например, оператор, заданный матрицей $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ в $\mathbb{R}^2$, не приводится к жордановой форме (докажите).

На практике для нахождения жордановой формы и жорданова базиса оператора, заданного матрицей A, можно использовать следующий алгоритм.

Сначала находим характеристический многочлен и его корни $\lambda_1, \ldots, \lambda_k$ (собственные значения) с кратностями.

Затем находим корневые подпространства R_{λ_i}. Для этого возводим матрицу $A - \lambda_i E$ в степень до тех пор, пока не наступит стабилизация ранга: $\mathrm{rk}(A - \lambda_i E)^{m_i} = \mathrm{rk}(A - \lambda_i E)^{m_i+1}$. Тогда $R_{\lambda_i} = \mathrm{Ker}(A - \lambda_i E)^{m_i}$, а число m_i будет размером максимальной жордановой клетки, отвечающей λ_i.

Далее в каждом пространстве R_{λ_i} находим нормальный базис для нильпотентного оператора $(\mathcal{A} - \lambda_i \cdot \mathrm{id})|_{R_{\lambda_i}}$; объединение этих базисов и будет жордановым базисом для $\mathcal{A}$.

Зная жорданову форму, легко вычислить минимальный многочлен оператора.

Предложение

Минимальный аннулирующий многочлен оператора $\mathcal{A}$ над полем $\mathbb{C}$ есть $P(t)=\prod_{i=1}^{k}(t-\lambda_i)^{m_i}$, где $\lambda_1,\ldots,\lambda_k$ — собственные значения $\mathcal{A}$, а m_i — размер максимальной жордановой клетки, отвечающей λ_i.

Доказательство.

Мы имеем $R_{\lambda_i}=\mathrm{Ker}(\mathcal{A}-\lambda_i\,\mathrm{id})^{m_i}$, поэтому многочлен $(t-\lambda_i)^{m_i}$ является минимальным аннулирующим для оператора $\mathcal{A}|_{R_{\lambda_i}}$. Любой вектор $\boldsymbol{v}\in V$ представляется в виде $\boldsymbol{v}=\sum_i \boldsymbol{v}_i$, где $\boldsymbol{v}_i\in R_{\lambda_i}$. Так как $P(\mathcal{A})$ содержит множитель $(\mathcal{A}-\lambda_i\,\mathrm{id})^{m_i}$, мы имеем $P(\mathcal{A})\boldsymbol{v}_i=\mathbf{0}$, т. е. $P(\mathcal{A})\boldsymbol{v}=\mathbf{0}$ и многочлен $P(t)$ аннулирует оператор $\mathcal{A}$. С другой стороны, любой многочлен $Q(t)$, аннулирующий оператор $\mathcal{A}$, делится на минимальный аннулирующий многочлен для оператора $\mathcal{A}|_{R_{\lambda_i}}$, т. е. на $(t-\lambda_i)^{m_i}$, для каждого λ_i. Следовательно, $Q(t)$ делится на $P(t)$, и $P(t)$ — минимальный многочлен. $\square$

2.11. Многочлены и функции от матриц. Экспонента

Одним из важных применений жордановой формы является эффективное вычисление многочленов и функций от операторов (матриц).

Сначала получим формулу для многочлена от жордановой клетки.

Предложение

Пусть $f(t)$ — многочлен. Тогда его значение на жордановой клетке J_λ размера n вычисляется по формуле

$$f(J_\lambda) = f\begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \dots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & \ddots & \ddots & \vdots \\ & & f(\lambda) & \frac{f'(\lambda)}{1!} \\ 0 & & & f(\lambda) \end{pmatrix}.$$

Доказательство.

Вначале по индукции проверим формулу для многочлена $f(t) = t^m$, т. е. для m-й степени жордановой клетки. Пусть $(J_\lambda)^m = (a_{ij}^m)$, т. е. (ij)-й элемент матрицы $(J_\lambda)^m$ есть a_{ij}^m. По предположению индукции, доказываемая формула для $f(t) = t^{m-1}$ даёт

$$a_{ij}^{m-1} = \frac{f^{(j-i)}(\lambda)}{(j-i)!} = C_{m-1}^{j-i}\lambda^{m-1-j+i},$$

где $C_k^i = \frac{k!}{i!(k-i)!}$ — биномиальный коэффициент; мы считаем $C_k^i = 0$ при $i < 0$ или $i > k$. Тогда из соотношения $(J_\lambda)^m = (J_\lambda)^{m-1}J_\lambda$ по правилу умножения матриц вычисляем

$$a_{ik}^m = \sum_{j=1}^n a_{ij}^{m-1}a_{jk}^1 = \sum_{j=1}^n C_{m-1}^{j-i}\lambda^{m-1-j+i}C_1^{k-j}\lambda^{1-k+j} =$$

$$= C_{m-1}^{k-1-i}\lambda^{m+i-k} + C_{m-1}^{k-i}\lambda^{m+i-k} = C_m^{k-i}\lambda^{m-k+i},$$

что и требовалось. Осталось заметить, что доказываемая формула линейна по f, а значит она верна для любого многочлена $f(t)$. $\square$

На основе формулы для многочлена от жордановой клетки мы можем также вычислять многочлены от матриц в жордановой форме J, так как многочлен применяется к такой матрице поблочно.

Теперь если A — произвольная матрица, то мы можем привести её к жордановой форме, т. е. найти жорданову матрицу J, для которой $A = CJC^{-1}$. При возведении матрицы A в степень мы получаем

$$A^m = (CJC^{-1})^m = CJC^{-1}CJC^{-1} \dots CJC^{-1} = CJ^mC^{-1}.$$

Тогда аналогичная формула верна и для произвольного многочлена $f(t)$:

$$f(A) = Cf(J)C^{-1}.$$

При помощи этой формулы мы можем вычислить любой многочлен от матрицы A, зная её жорданову форму и жорданов базис.

Определение

Пусть $\mathcal{A}$ — оператор в комплексном пространстве с собственными значениями $\lambda_1, \ldots, \lambda_k$ кратностей $r_{\lambda_1}, \ldots, r_{\lambda_k}$ соответственно. Говорят, что многочлены $f(t)$ и $g(t)$ совпадают на спектре оператора $\mathcal{A}$, если

$$f^{(j)}(\lambda_i) = g^{(j)}(\lambda_i) \quad \text{при } i = 1, \ldots, k,\ j = 0, \ldots, r_{\lambda_i} - 1.$$

Предложение

Если многочлены $f(t)$ и $g(t)$ совпадают на спектре $\mathcal{A}$, то $f(\mathcal{A}) = g(\mathcal{A})$.

Доказательство.

При вычислении многочлена $f(A)$ по формуле $f(A) = Cf(J)C^{-1}$ и формуле для многочлена от жордановой клетки используются производные многочлена $f(t)$ в точках λ_i порядка не выше $r_{\lambda_i} - 1$. □

Это утверждение позволяет находить многочлен $f(t)$ большой степени от матрицы A (например, возводить матрицу в большую степень), вовсе не вычисляя её жордановой формы.

Предложение

Пусть $\mathcal{A}$ — оператор в n-мерном пространстве. Для любого многочлена $f(t)$ существует многочлен $g(t)$ степени меньше n, совпадающий с $f(t)$ на спектре оператора $\mathcal{A}$.

Доказательство.

Пусть $P_{\mathcal{A}}(t)$ — характеристический многочлен для $\mathcal{A}$. Тогда $P_{\mathcal{A}}(t) = \prod_{i=1}^{k}(\lambda_i - t)^{r_{\lambda_i}}$ — многочлен степени n, причём $P_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. Разделим $f(t)$ на $P_{\mathcal{A}}(t)$ с остатком:

$$f(t) = q(t)P_{\mathcal{A}}(t) + g(t),$$

где $g(t)$ — многочлен-остаток, имеющий степень $< n$. Подставив λ_i в соотношение выше и дифференцируя требуемое число раз, получим

$$f^{(j)}(\lambda_i) = g^{(j)}(\lambda_i) \quad \text{при } i = 1, \dots, k, \ j = 0, \dots, r_{\lambda_i} - 1,$$

т. е. f и g совпадают на спектре $\mathcal{A}$. Кроме того, подставляя в соотношение выше оператор $\mathcal{A}$, получим $f(\mathcal{A}) = g(\mathcal{A})$. Это показывает, что $g(t)$ — требуемый многочлен. □

На практике многочлен $g(t)$ степени не выше $n-1$, удовлетворяющий соотношениям

$$f^{(j)}(\lambda_i) = g^{(j)}(\lambda_i) \quad \text{при } i = 1, \ldots, k,\ j = 0, \ldots, r_{\lambda_i} - 1,$$

находится при помощи формул интерполяции или методом неопределённых коэффициентов. Тогда мы имеем $f(A) = g(A)$.

Если же нам известна жорданова форма матрицы A, то можно ещё снизить степень многочлена g, заменив в соотношениях выше числа r_{λ_i} на числа m_{λ_i} — размеры максимальных жордановых клеток с λ_i на диагонали (жорданов базис для этого знать не обязательно).

На самом деле формулу для вычисления многочлена от жордановой клетки и формулу $f(A) = Cf(J)C^{-1}$ можно использовать также для вычисления более общих функций f от операторов или матриц (а не только многочленов).

Если функция f гладкая и хорошо приближается многочленами (такие функции называются аналитическими) в окрестности собственных значений матрицы A, то можно определить $f(A)$ как предел последовательности многочленов, получаемых обрезанием ряда Тейлора для f. При этом, однако, необходимо обосновывать сходимость получаемых последовательностей (или рядов) из матриц. Мы этого делать не будем, а скажем лишь, что таким образом можно определить e^A, $\sin A$, $\cos A$, а также $\sqrt{A}$ и $\ln A$ для матриц с положительными собственными значениями.

Можно также использовать формулу $f(A) = Cf(J)C^{-1}$ в качестве определения $f(A)$ для функций f, которые определены на спектре A.

Мы рассмотрим экспоненту более подробно.

Определение

Экспонентой оператора $\mathcal{A}$ называется оператор

$$e^{\mathcal{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{A}^k = \mathrm{id} + \mathcal{A} + \frac{\mathcal{A}^2}{2} + \frac{\mathcal{A}^3}{6} + \ldots .$$

Предложение

Имеет место соотношение $\det e^{\mathcal{A}} = e^{\operatorname{tr}\mathcal{A}}$.

Доказательство.

Приведём матрицу оператора к жордановой форме: $A = CJC^{-1}$. Используя формулу $f(A) = Cf(J)C^{-1}$ и формулу для функции $f(t) = e^t$ от жордановой клетки, вычисляем

$$\det e^A = \det(Ce^J C^{-1}) = \det e^J = \prod_{i=1}^{k} e^{r_{\lambda_i}\lambda_i} = e^{\sum_{i=1}^k r_{\lambda_i}\lambda_i} = e^{\operatorname{tr} J} = e^{\operatorname{tr} A}.$$

$\square$

Основное свойство числовой экспоненты $e^a e^b = e^{a+b}$, вообще говоря, *нарушается* для экспоненты операторов. Однако есть важный случай, когда оно выполнено.

Предложение

Если операторы $\mathcal{A},\mathcal{B}\colon V\to V$ коммутируют, т. е. $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$, то $e^{\mathcal{A}}e^{\mathcal{B}}=e^{\mathcal{A}+\mathcal{B}}$.

Доказательство.

Мы имеем

$$e^{\mathcal{A}}e^{\mathcal{B}}=$$

$$=\Bigl(\sum_{i=0}^{\infty}\frac{1}{i!}\mathcal{A}^i\Bigr)\Bigl(\sum_{j=0}^{\infty}\frac{1}{j!}\mathcal{B}^j\Bigr)=\sum_{i,j\geqslant 0}\frac{1}{i!j!}\mathcal{A}^i\mathcal{B}^j=\sum_{k\geqslant 0}\sum_{i=0}^{k}\frac{1}{i!(k-i)!}\mathcal{A}^i\mathcal{B}^{k-i}=$$

$$=\sum_{k\geqslant 0}\frac{1}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i!)}\mathcal{A}^i\mathcal{B}^{k-i}=\sum_{k\geqslant 0}\frac{1}{k!}(\mathcal{A}+\mathcal{B})^k=e^{\mathcal{A}+\mathcal{B}}.$$

Коммутативность $\mathcal{A}$ и $\mathcal{B}$ используется в том месте, где $(\mathcal{A}+\mathcal{B})^k$ раскладывается по формуле бинома Ньютона. $\square$