МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра компьютерной алгебры и теории чисел

Спектральное разложение функции от матриц

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 421 группы
направление 02.03.01 — Математика и компьютерные науки

механико-математического факультета

Яковлевой Анастасии Александровны

Научный руководитель
доцент, к.ф.-м.н. О.А. Королева

Зав. кафедрой
зав. каф., к.ф.-м.н., доцент А.М. Водолазов

Саратов 2023

Введение. Дифференциальные уравнения широко используются для моделирования различных процессов в физике, инженерии и других областях науки. Решение этих уравнений является трудной задачей, особенно для сложных нелинейных систем. Существуют различные методы решения дифференциальных уравнений, одним из которых является спектральное разложение функции от матриц.

Спектральное разложение позволяет представить функцию от матрицы в виде суммы собственных значений и собственных векторов матрицы. Это применение теории линейной алгебры к дифференциальным уравнениям позволяет упростить решение и улучшить его точность.

Существует много приложений спектрального разложения функций от матриц, включая решение дифференциальных уравнений. В данной работе исследуется применение этой техники для численного решения систем дифференциальных уравнений. Для реализации метода использовался Wolfram Mathematica.

Основная часть данной работы включает в себя 2 главы. Первая глава посвящена рассмотрению минимального многочлена и нормальных форм, например жардановых форм. А во второй главе рассмотрены функции от матриц, интерполяционный многочлен Лагранжа-Сильвестра , спектральное разложение и приложения к решению дифференциальных уравнений.

Основное содержание работы. Рассмотрим основные определения для спектрального разложения матриц.

Рассмотрим $n \times n$ - матрицу $A(\lambda)$, элементы которой – многочлены от λ. Предпологается, что значения для λ и коэффициенты многочленов берутся из некоторого поля F, так что если элементы матрицы A вычисляются для частного значения λ, скажем $\lambda = \lambda_0$, то $A(\lambda_0) \in F_{n\times n}$. Если имеется некоторый элемент в $A(\lambda)$, который является многочленом от λ степени l, и нет элементов в $A(\lambda)$ степени, большей l, то будем говорить, что l – степень матрицы $A(\lambda)$. Матрица $A(\lambda)$ известна как матрица многочленов, или λ - матрица. Если нужно будет выделить основное поле, то будем называть $A(\lambda)$ λ - матрицей на полем F.

Если $A(\lambda)$ имеет степень l, то (i, j) - элемент в $A(\lambda)$ может быть записан в виде

$$a_{ij}(\lambda) = a_{ij}^{(0)}\lambda^l + a_{ij}^{(1)}\lambda^{l-1} + ... + a_{ij}^{(l-1)}\lambda + a_{ij}^{(l)}$$

и имеется по крайней мере один элемент, для которого $a_{ij}^0 \neq 0$. Если определить матрицу A_r $(r = 0, 1, ..., l)$ как такую, у которой (i, j) - элемент равен $a_{ij}^{(r)}$, то будем иметь $A_0 \neq 0$ и

$$A(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + ... + A_{l-1}\lambda + A_l.$$

λ - матрица $A(\lambda)$ называется регулярной, если $det(A_0) \neq 0$.

При рассмотрении скалярного многочлена p можно записать

$$p(\lambda) = a_0\lambda^l + a_1\lambda^{l-1} + \ldots + a_l = \lambda^l a_0 + \lambda^{l-1}a_1 + \ldots + a_l.$$

Для λ-матрицы с матричным аргументом в общем случае это невозможно. Если $A(\lambda) - \lambda$-матрица над F и $B \in F_{n\times n}$, определим правое значение $A(B)$ матрицы $A(\lambda)$ в B как

$$A(B) = A_0B^l + A_1B^{l-1} + \ldots + A_l$$

и левое значение $\hat{A}(B)$ матрицы $A(\lambda)$ в B как

$$\hat{A}(B) = B^lA_0 + B^{l-1}A_1 + \ldots + A_l.$$

Теорема 1.2 Если A – квадратная матрица с характеристическим многочленом c, то $c(A) = 0$.

Эта теорема имеет такое важное следствие: если f - любой скалярный многочлен с коэффициентами в поле F и $A \in F_{n\times n}$, то существует многочлен p (зависящий от A) степени, меньшей n, для которого $f(A) = p(A)$. Чтобы убедиться в этом, находим многочлены q, r, которые будут частным и остатком при делении f на c. Тогда

$$f(\lambda) = q(\lambda)c(\lambda) + r(\lambda),$$

где r – либо нулевой многочлен, либо имеет степень меньше n. Таким образом,

$$f(A) = q(A)c(A) + r(A).$$

Ввиду теоремы 1.2, получаем $f(A) = r(A)$.

Скалярный многочлен f над F называется анулирующим многочленом для квадратной матрицы $A \in F_{n\times n}$, если $f(A) = 0$. В теореме 1.2 утверждается, что $c(A) = 0$, а поэтому всегда существует по крайней мере один аннулирующий многочлен степени n, порядка матрицы A. Конечно, можно построить аннулирующие многочлены степени, большей n, но интересны аннулирующие многочлены наименьшей возможной степени. Чтобы избежать ненужных усложнений, будем рассматривать многочлены, у которых старший коэффициент (коэффициент при наивысшей степени переменной) равен 1. Такой многочлен называется приведенным. Определим теперь приведенный многочлен ψ, являющийся аннулирующий многочленом для A наименьшей степени, как минимальный многочлен для A.

Теорема 1.3. Каждый аннулирующий многочлен матрицы A делится без остатка на минимальный многочлен для A .

Теорема 1.4. Минимальный многочлен для данной матрицы единствен.

Теперь можно исследовать связь между характеристическим и минимальным многочленами. Пока не было сказано ничего, что препятствовало бы этим многочленам быть совпадающими.

Лемма 1.1. Если $A(\lambda)$ – λ - матрица над F порядка n, то $A(\lambda)$ эквивалентна диагональной матрице $diag\{a_1(\lambda), ..., a_s(\lambda), 0, ..., 0\}$, где $s \leq n$ и $a_1(\lambda), ..., a_s(\lambda)$ – приведенные многочлены над F и $a_{i-1}(\lambda)$, $f = 2, 3, ..., s$.

Теорема 1.7. $\lambda-$матрица $A(\lambda)$ ранга r эквивалентна матрице $\operatorname{diag}\{i_1(\lambda), \ldots, i_r(\lambda), 0, \ldots, 0\}$, где $i_1(\lambda), \ldots, i_r(\lambda)$ – инвариантные многочлены для $A(\lambda)$.

Следствие. Две λ – матрицы эквивалентны $\leftrightarrow$ они имеют одинаковые инвариантные многочлены.

Для получения следствия заметим, что уже доказано, что эквивалентные λ – матрицы имеют одинаковые инвариантные многочлены. Наоборот, если две λ – матрицы имеют одинаковые инвариантные многочлены, то теорема означает, что они имеют одинаковые канонические формы Смита. Свойство транзитивности отношений эквивалентности означает тогда, что λ – матрицы эквивалентны.

Наконец, следует заметить, что если в утверждении теоремы элементы матрицы $A(\lambda)$ – это многочлены над F, то инвариантные многочлены имеют коэффициенты в том же самом поле F.

Определение 2.1. Матрица $A \in M_(C)$ называется простой, если алгебраическая кратность каждого ее собственного значения совпадает с его геометрической кратностью.

Из линейной алгебры следует, что матрица $A \in M_(C)$ простая тогда и только тогда, когда $A \sim U = diag(\lambda_i, ..., \lambda_n)$.

Если матрица A простая, тогда существует n линейно независимых собственных векторов $x_1, x_2, ..., x_n$ таких, что $Ax_i = \lambda_i x_i$, для $i = \overline{1, n}$. Запишем это равенство в матричном виде:

$$x = [x_1, x_2, ..., x_n]$$
$$AX = XU$$
$$A = XUX^{-1} = (X^{-1})^{-1}UX^{-1},$$

Предположим, что матрица $A \in B_{n\times n}$ имеет различные собственные значения $\lambda_1, \lambda_2, ..., \lambda_s$ с индексами $m_1, m_2, ..., m_s$, так что минимальный многочлен для A равен

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_s)^{m_s}.$$

Пусть g, h – многочлены над B, для которых $g(A) = h(A)$, и положим $d = g - h$. Тогда $d(A) = 0$, так что d – аннулирующий многочлен и поэтому (теорема 1.3) делится на ψ. Таким образом, существует такой многочлен p, что

$$g(\lambda) - h(\lambda) = \psi(\lambda)p(\lambda).$$

Но, очевидно, для $k = 1, 2, ..., s$

$$\psi(\lambda_k) = \psi^{(1)}(\lambda_k) = ... = \psi^{(m_k-1)}(\lambda_k) = 0,$$

откуда получаем, что

$$g(\lambda_k) = h(\lambda_k), \quad g^{(1)}(\lambda_k) = h^{(1)}(\lambda_k), ..., g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\lambda_k).$$

В общем случае m чисел

$$f(\lambda_k), \quad f^{(1)}(\lambda_k), ..., f^{(m_k-1)}(\lambda_k), \qquad k = 1, 2, ..., s,$$

называются значениями функции f на спектре матрицы A. Про любую функцию f, для которой эти числа существуют, будем говорить, что она определена на спектре матрицы A. Очевидно, каждый многочлен определен на спектре любой матрицы из $B_{n\times n}$ и, в частности, показали, что многочлены g и h имеют одинаковые значения на спектре матрицы A, если $g(A) = h(A)$.

Наоборот, если g и h – любые два многочлена с одинаковыми значениями на спектре матрицы A, то $d = g - h$ имеет корень кратности m_k в λ_k для каждого k. Следовательно, d должен делиться на ψ, и поэтому $d(A) = 0$, т.е. $g(A) = h(A)$. Доказана.

Лемма 2.1. Если g и h – многочлены на B и $A \in B_{n\times n}$, то
$g(A) = h(A) \Leftrightarrow g$ и h имеют одинаковые значения на спектре матрицы A.

Именно это свойство многочленов с матричными аргументами используется для определения $f(A)$ с более общими функциями f. Таким образом, будем требовать, чтобы все фукнции, которые определены на спектре матрицы A и принимают там одинаковые значения, приводили к одной и той же матрице $f(A)$.

Теорема 2.1. Если f – некоторая фукнция, определенная на спектре матрицы A, и g – интерполяционный многочлен минимальной степени, определенный значениями функции f на спектре A, то $f(A) = g(A)$.

Если A – простая матрица, то многочлен (2.2) принимает особенно простой вид. В этом случае имеем (следствие 2 теоремы 1.12) $m_1 = m_2 = ... = m_k = 1$, так что

$$g(\lambda) = \sum_{k=1}^{s} f(\lambda_k) l_k(\lambda),$$

где l_k – базисный многочлен лагранжева типа степени $s - 1$.

Определение 2.2. Интерполяционным многочленом Лагранжа-Сильвестра для функции $f(\lambda)$ на спектре матрицы A называется многочлен $r(\lambda)$, однозначно определяющийся интерполяционными условиями:

$$r(\lambda_k) = f(\lambda_k),\ r'(\lambda_k) = f'(\lambda_k)...r^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k) \quad (k = \overline{1, s}). \qquad (1)$$

Определение 2.3. Если функция $f(\lambda)$ определена на спектре матрицы A, то $r(\lambda)$ есть соответствующий интерполяционный многочлен Лагранжа-Сильвестра, и тогда:

$$f(A) = r(A)$$

Определение 2.4. Определим формулу интерполяционного многочлена Лагранжа-Сильвестра для случая, когда характеристический многочлен не имеет кратных корней, или же, при наличии кратных корней, минимальный многочлен, являющийся делителем характеристического, имеет только простые корни:

$$r(\lambda) = \sum_{k=1}^{n} \frac{(\lambda-\lambda_1)...(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})...(\lambda-\lambda_n)}{(\lambda_k-\lambda_1)...(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})...(\lambda_k-\lambda_n)} f(\lambda_k), \qquad (2)$$

$$f(A) = r(A) = \sum_{k=1}^{n} \frac{(A-\lambda_1 E)...(A-\lambda_{k-1}E)(A-\lambda_{k+1}E)...(A-\lambda_n E)}{(\lambda_k-\lambda_1)...(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})...(\lambda_k-\lambda_n)} f(\lambda_k).$$

Необходимо также рассмотреть общий случай. Пусть

$$\psi(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}...(\lambda-\lambda_s)^{m_s} \quad (m_1 + m_2 + ... + m_s = m).$$

Представим функцию $\frac{r(\lambda)}{\psi(\lambda)}$в виде суммы простых дробей:

$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^{s} \Big(\frac{a_{k_1}}{(\lambda-\lambda_k)^{m_k}} + \frac{a_{k_2}}{(\lambda-\lambda_k)^{m_k-1}} + ... + \frac{a_{k_{m_k}}}{\lambda-\lambda_k} \Big), \qquad (3)$$

где a_{k_j} $(j = \overline{1, m_k};\ k = \overline{1, s})$ - некоторые числа.

Умножим обе части равенства (2.5) на $(\lambda-\lambda_k)^{m_k}$ и обозначим многочлен $\psi^k(\lambda) = \frac{\psi(\lambda)}{(\lambda-\lambda_k)^{m_k}}$.

Получим:

$$\frac{r(\lambda)}{\psi^k(\lambda)} = a_{k_1} + a_{k_2}(\lambda - \lambda_k) + ... + a_{k_{m_k}}(\lambda - \lambda_k)^{m_k - 1} + (\lambda - \lambda_k)^{m_k}\rho(\lambda)\ (k = \overline{1, s}), \quad (4)$$

где $\rho(\lambda)$ - регулярная при $\lambda = \lambda_k$ рациональная функция.

Следовательно

$$\left.\begin{aligned} a_{k1} &= \left(\frac{r(\lambda)}{\psi^k(\lambda)}\right)_{\lambda=\lambda_k}, \\ a_{k2} = \left(\frac{r(\lambda)}{\psi^k(\lambda)}\right)'_{\lambda=\lambda_k} &= r(\lambda_k)\left(\frac{1}{\psi^k(\lambda)}\right)'_{\lambda=\lambda_k} + r'(\lambda_k)\frac{1}{\psi^k(\lambda_k)}, \\ \dots\dots&\dots\dots\dots\dots \end{aligned}\right\} \quad (5)$$

Из (2.7) видно, что числители a_{k_j} в правой части равенства (2.5) выражаются через значения многочлена $r(\lambda)$ на спектре матрицы A, а эти значения равны соответствующим значениям функции $f(\lambda)$ и ее производных. Следовательно можем переписать их так:

$$a_{kj} = \frac{1}{(j-1)!}\left(\frac{f(\lambda)}{\psi^k(\lambda)}\right)^{(j-1)}_{\lambda=\lambda_k} \qquad (j = \overline{1, m_k}; k = \overline{1, s}). \quad (6)$$

После нахождения всех a_{k_j} определяем $r(\lambda)$ из следующей формулы:

$$r(\lambda) = \sum_{k=1}^{s}\left(a_{k_1} + a_{k_2}(\lambda - \lambda_k) + ... + a_{k_{m_k}}(\lambda - \lambda_k)^{m_k - 1}\right)\psi^k(\lambda). \quad (7)$$

Следующее (спектральное) разложение для $f(A)$ имеет большое значение и, очевидно, тесно примыкает к уже полученной спектральной теореме для простых матриц.

Теорема 2.8. Если $A \in B_{n\times n}$, f определена на спектре A, $f_k^{(i)}$ – значение j-й производной от f в собственном значении λ_k $\quad (k = 1, 2, ..., s,$ $j = 0, 1, ..., m_k - 1)$ и m_k – индекс λ_k, то существует такие независимые от f матрицы Z_{kj}, что

$$f(A) = \sum_{k=1}^{s}\sum_{j=1}^{m_k} f_k^{(j-1)} Z_{kj}. \quad (8)$$

Кроме того, матрицы Z_{kj} линейно независимы как элементы из $B_{n\times n}$ и коммутирует с A.

Доказательство. Напомним, что f и многочлен g, определяемый равенством (2.2), принимают одинаковые значения на спектре матрицы A. Следовательно, $f(A) = g(A)$ и равенство (2.10) будет выполняться, если положить

$$Z_{kj} = \varphi_{kj}(A).$$

Так как φ_{kj} определяется свойствами минимального многочлена для A, то Z_{kj} не зависит от f. Немедленно также получаем, что каждое Z_{kj} коммутирует с A.

Остается лишь доказать, что эти матрицы линейно независимы.

Предположим, что

$$\sum_{k=1}^{s}\sum_{j=1}^{m_k} c_{kj} Z_{kj} = 0$$

и определим $h(\lambda) = \sum_{k,j} c_{k,j}\varphi_{kj}(\lambda)$ - многочлен пространства P_{m-1}. Тогда равенство $Z_{kj} = \varphi_{kj}(A)$ означает, что $h(A) = 0$. Но m - степень минимального многочлена матрицы A и, так как h аннулирует A, из теоремы 1.3 следует, что h - нулевой многочлен и, следовательно (так как φ_{kj} независимы), что $c_{kj} = 0$ для всех k и j. Таким образом, матрицы Z_{kj} линейно независимы.

Будем называть матрицы Z_{kj} компонентами матрицы A. Заметим, что, так как они линейно независимы, ни одна из них не может быть нулевой.

Рассмотрим более подробно случай простых матриц. В (2.10) имеем теперь $m_1 = m_2 = \ldots = m_s = 1$, так что

$$f(A) = \sum_{k=1}^{s} f_k Z_{k1}$$

для простой матрицы A. Сравним это равенство с результатом теоремы (2.7).

Применяя оба результата к базисным лагранжевым многочленам $l_k(\lambda)$ получаем, что Z_{k_1} равна сумме сопровождающих матриц, соответствующих собственному значению λ_k. Как тогда нетрудно видеть, Z_{k1} - идемпотентная магрица и образ Z_{k1} - это правое собственное подпространство, соответствующее λ_k (т. е. нулевое подпространство матрицы $I\lambda_k - A$). Таким образом, теорема 2.7 оказывается частным случаем теоремы 2.8. Из следствия теоремы 2.6 получаем, что

$$\sum_{k=1}^{s} Z_{k1} = I \quad \text{и} \quad Z_{k1}Z_{j1} = \delta_{kj}Z_{k1}, \tag{9}$$

хотя первый из этих результатов, как легко видеть, верен для всех матриц $A \in B_{n\times n}$, следует лишь положить $f(\lambda) = 1$ в равенстве (2.10).

Заметим, наконец, что в рассматриваемом случае $\varphi_{k1} = l_k$, $k = 1, 2, \ldots, s$, так что

$$Z_{k1} = \varphi_{k1}(A) = l_k(A) = \prod_{\substack{j=1\\ j\neq k}}^{s} (A - \lambda_i I) \Big/ \prod_{\substack{j=1\\ j\neq k}}^{s} (\lambda_k - \lambda_j)\,. \tag{10}$$

Пусть $A_1, A_2, \ldots$ - некоторая последовательность матриц, принадлежащих $C_{m\times n}$, и пусть $a_{ij}^{(p)}$ – (i,j)-элемент матрицы $A_p, p = 1, 2, \ldots$ Последовательность $A_1, A_2, \ldots$ называется сходящейся к матрице $A \in C_{m\times n}$, если существуют такие числа a_{ij} (элементы матрицы A), что $a_{ij}^{(p)} \to a_{ij}$ при $p \to \infty$ для каждой пары индексов i, j. Последовательность, которая не сходится, называется расходящейся. Таким образом, сходимость последовательностей матриц определяется как поэлементная сходимость. Введенное определение включает также в качестве частных случаев сходимость вектор-столбцов и вектор-строк.

Теорема 2.9. Пусть функции $f_1, f_2, \ldots$ определены на спектре матрицы $A \in B_{n\times n}$ $\cup A_p = f_p(A), p = 1, 2, \ldots$ Тогда последовательность $A_1, A_2, \ldots$ сходится при $p \to \infty \iff m$ скалярных последовательностей $f_1^{(j-1)}(\lambda_k), f_2^{(j-1)}(\lambda_k), \ldots$ (определенных для $k = 1, 2, \ldots$, s и $= 1, 2, \ldots, m_k$) сходятся $p \to \infty$.

Кроме того, если $\lim_{p\to\infty} f_p^{(j-1)}(\lambda_k) = f^{(j-1)}(\lambda_k)$ для всех i и k и некоторой функции f, то $\lim_{p\to\infty} A_p = f(A)$ и, наоборот, если $\lim_{p\to\infty} A_p$ существует, то существует такая определенная на спектре A функция f, что
$\lim_{p\to\infty} A_p = f(A)$.

Теорема 2.10. Если функции $u_1, u_2, \ldots$ определены на спектре $A \in B_{n\times n}$, то $\sum\limits_{p=1}^{\infty} u_p(A)$ сходится $\iff$ ряды $\sum\limits_{p=1}^{\infty} u_p^{(j-1)}(\lambda_k)$ сходятся, где $k = 1, 2, \ldots, s$ и $j = 1, 2, \ldots, m_k$.

Кроме того, если

$$\sum_{p=1}^{\infty} u_p^{(j-1)}\left(\lambda_k\right)=f(j-1)\left(\lambda_k\right)$$

для всех j и k, то

$$\sum_{p=1}^{\infty} u_p(A)=f(A)$$

и наоборот.

Можно теперь применить эту теорему к получению одного важного результата, который даст возможность оправдать наше изложение теории функций от матриц. Следующая теорема показывает общность данного в (2.1) определения, которое при первом знакомстве могло показаться несколько произвольным. Действительно, что описанные в следующей теореме свойства часто используются при определении функции от матрицы. Данный подход допускает определение $f(A)$ для более широкого класса функций f, чем описанный в теореме (2.8).

Теорема 2.11. Пусть матрица $A \in B_{n \times n}$ имеет собственные значения $\lambda_1, \lambda_2, \ldots, \lambda_n$. Если функция f имеет ряд Тейлора в точке λ_0 :

$$f(\lambda)=\sum_{p=0}^{\infty} \alpha_p\left(\lambda-\lambda_0\right)^p$$

с кругом сходимости $\left|\lambda-\lambda_0\right|=r$ и если $\left|\lambda_j-\lambda_0\right|<r, j=1,2, \ldots, n$, то $f(A)$ определена и

$$f(A)=\sum_{p=0}^{\infty} \alpha_p\left(A-\lambda_0 I\right)^p .$$

Заключение. В данной работе было исследовано применение спектрального разложения для численного решения систем дифференциальных уравнений. Были рассмотрены минимальный многочлен и нормальные формы матриц, а также функции от матриц и интерполяционный многочлен Лагранжа-Сильвестра.

Работа показала, что спектральное разложение является мощной техникой в решении систем дифференциальных уравнений, благодаря использованию теории линейной алгебры. Его приложения распространены в различных областях науки, где требуется моделирование сложных процессов.

Таким образом, исследование спектрального разложения позволяет улучшить точность численного решения дифференциальных уравнений и может оказаться полезным в различных областях науки и инженерии.