V. D. Shmatkov, On diagonalization of matrices in an arbitrary field,
Fundam. Prikl. Mat., 2016, Volume 21, Issue 3, 233–239

https://www.mathnet.ru/eng/fpm1744

О диагонализации матриц в произвольном поле

В. Д. ШМАТКОВ
*Рязанский государственный
радиотехнический университет*
e-mail: shmatkov-vadim@yandex.ru

УДК 512.64

Ключевые слова: подобие матриц, функции от матриц.

Аннотация

В данной работе приведён простой способ диагонализации матриц в произвольном поле, с помощью которого можно вычислять функции от матриц.

Abstract

V. D. Shmatkov, On diagonalization of matrices in an arbitrary field, Fundamentalnaya i prikladnaya matematika, vol. 21 (2016), no. 3, pp. 233—239.

The paper presents a simple way for diagonalization of matrices in an arbitrary field, with which one can calculate functions of matrices.

Пусть P — поле. Обозначим через $\mathrm{M}_n(P)$ множество всех матриц над полем P. Пусть $A \in \mathrm{M}_n(P)$.

Рассмотрим характеристический многочлен матрицы A и для него поле разложения F. Будем считать, что $A \in \mathrm{M}_n(F)$. По аналогу теоремы Шура существует обратимая матрица $T \in \mathrm{M}_n(F)$, такая что $T^{-1}AT = B$ — верхнетреугольная матрица. Пусть $f_i(x)$, $i = 1, \ldots, n$, — множество многочленов над полем F, каждый из которых имеет корень 0. Рассмотрим поле рациональных дробей $F(X)$ (поле частных многочленов). Будем считать, что $B \in \mathrm{M}_n\bigl(F(X)\bigr)$.

Рассмотрим диагональную матрицу

$$d\bigl(f_i(x)\bigr) = \begin{pmatrix} f_1(x) & & 0 \\ & \ddots & \\ 0 & & f_n(x) \end{pmatrix}.$$

Пусть $B\bigl(f_i(x)\bigr) = B + d\bigl(f_i(x)\bigr)$. Ясно, что $B\bigl(f_i(x)\bigr) \in \mathrm{M}_n\bigl(F(X)\bigr)$.

Рассмотрим такие $f_i(x)$, что все элементы $B(i,i)+f_i(x)$, $i = 1, \ldots, n$, попарно различны. Тогда у матрицы $B\bigl(f_i(x)\bigr)$ все собственные значение попарно различны и её можно диагонализировать, т. е. существует обратимая $S \in \mathrm{M}_n\bigl(F(X)\bigr)$, такая что

$$S^{-1}B\bigl(f_i(x)\bigr)S = \begin{pmatrix} B(1,1) + f_1(x) & & 0 \\ & \ddots & \\ 0 & & B(n,n) + f_n(x) \end{pmatrix}. \tag{1}$$

Фундаментальная и прикладная математика, 2016, том 21, № 3, с. 233—239.

Если $x = 0$, то правая часть равенства (1) равна

$$\begin{pmatrix} B(1,1) & & 0 \\ & \ddots & \\ 0 & & B(n,n) \end{pmatrix}.$$

Следовательно, после сопряжения в левой части и подстановки $x = 0$ левая часть равенства (1) станет равна

$$\begin{pmatrix} B(1,1) & & 0 \\ & \ddots & \\ 0 & & B(n,n) \end{pmatrix}.$$

Ясно, что

$$B\big(f_i(x)\big) = S \begin{pmatrix} B(1,1) + f_1(x) & & 0 \\ & \ddots & \\ 0 & & B(n,n) + f_n(x) \end{pmatrix} S^{-1}. \tag{2}$$

Возведём матрицу в левой части равенства (2) в степень $k \in \mathbb{N}$. Тогда

$$B^k\big(f_i(x)\big) = S \begin{pmatrix} \big(B(1,1) + f_1(x)\big)^k & & 0 \\ & \ddots & \\ 0 & & \big(B(n,n) + f_n(x)\big)^k \end{pmatrix} S^{-1}. \tag{3}$$

Далее,

$$\begin{aligned} B^k = B^k\big(f_i(x)\big)\big|_{x=0} = \\ = S \begin{pmatrix} \big(B(1,1) + f_1(x)\big)^k & & 0 \\ & \ddots & \\ 0 & & \big(B(n,n) + f_n(x)\big)^k \end{pmatrix} S^{-1} \Bigg|_{x=0}. \end{aligned} \tag{4}$$

Обозначение $|_{x=0}$ показывает, что после вычисления выражения мы подставляем значение $x = 0$.

Пусть

$$f(B) = a_0 + a_1 B + \ldots + a_k B^k \,—$$

матричный многочлен. Тогда из (4) следует, что

$$f(B) = S \begin{pmatrix} f\big(B(1,1) + f_1(x)\big) & & 0 \\ & \ddots & \\ 0 & & f\big(B(n,n) + f_n(x)\big) \end{pmatrix} S^{-1} \Bigg|_{x=0}. \tag{5}$$

Здесь

$$f\big(B(i,i)+f_i(x)\big)=\sum_{j=0}^{k} a_j\big(B(i,i)+f_i(x)\big)^j\,-$$

соответствующий многочлен для $B(i,i)+f_i(x)$.

Если ряд

$$f(B)=\sum_{i=0}^{\infty} a_i B^i$$

сходится по какой-то матричной норме, то

$$f(B)=S\begin{pmatrix} f\big(B(1,1)+f_1(x)\big) & & 0 \\ & \ddots & \\ 0 & & f\big(B(n,n)+f_n(x)\big)\end{pmatrix}S^{-1}\Bigg|_{x=0}. \tag{6}$$

Вспоминая, что $A=TBT^{-1}$, и считая, что $T_1=TS$, имеем

$$f(A)=T_1\begin{pmatrix} f\big(B(1,1)+f_1(x)\big) & & 0 \\ & \ddots & \\ 0 & & f\big(B(n,n)+f_n(x)\big)\end{pmatrix}T_1^{-1}\Bigg|_{x=0}. \tag{7}$$

Если F — поле действительных или комплексных чисел и переменная x принимает значения в F, то

$$f(A)=\lim_{x\to 0} T_1\begin{pmatrix} f\big(B(1,1)+f_1(x)\big) & & 0 \\ & \ddots & \\ 0 & & f\big(B(n,n)+f_n(x)\big)\end{pmatrix}T_1^{-1}. \tag{8}$$

Формулы (7), (8) позволяют вычислить функцию от матрицы, если функция представляется степенным матричным рядом или многочленом.

Хорошо известно, что ряды

$$e^A=\sum_{p=0}^{\infty}\frac{A^p}{p!},\quad \cos A=\sum_{p=0}^{\infty}\frac{(-1)^pA^{2p}}{(2p)!},\quad \sin A=\sum_{p=0}^{\infty}\frac{(-1)^pA^{2p+1}}{(2p+1)!}$$

сходятся для всех A. Поэтому формулы (7), (8) можно, в частности, применить для вычислений e^A, $\sin A$, $\cos A$.

Обозначим через

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

минимальный многочлен матрицы A. (Здесь $\lambda_1,\lambda_2,\ldots,\lambda_s$ — все различные характеристические числа матрицы A.) Степень этого многочлена равна

$$m=\sum_{k=1}^{s} m_k.$$

Существует единственный многочлен $r(\lambda)$ степени m, принимающий те же значения на спектре A, что и f, т. е.

$$r(\lambda_k) = f(\lambda_k),\ r'(\lambda_k) = f'(\lambda_k),\ \ldots,\ r^{m_{k-1}}(\lambda_k) = f^{m_{k-1}}(\lambda_k) \quad (k = 1, 2, \ldots, s).$$

Этот многочлен называется интерполяционным многочленом Лагранжа—Сильвестра для функции $f(\lambda)$ на спектре A (см. [1, гл. 5, § 1, 2]). Как показано в [1, гл. 5, § 4], это определение f совпадает с определением, данным нами для степенных матричных рядов и многочленов.

Будем считать, что функция f вместе с производными до $f^{m_{k-1}}$ определена в окрестности λ_k. Тогда, используя результаты [1, с. 102, утверждения 2, 3], имеем

$$f\left(T_1 \begin{pmatrix} B(1,1)+f_1(x) & & 0 \\ & \ddots & \\ 0 & & B(n,n)+f_n(x) \end{pmatrix} T_1^{-1}\right) =$$
$$= T_1 \begin{pmatrix} f\bigl(B(1,1)+f_1(x)\bigr) & & 0 \\ & \ddots & \\ 0 & & f\bigl(B(n,n)+f_n(x)\bigr) \end{pmatrix} T_1^{-1}.$$

Следовательно,

$$f(A) = \lim_{x\to 0} T_1 \begin{pmatrix} f\bigl(B(1,1)+f_1(x)\bigr) & & 0 \\ & \ddots & \\ 0 & & f\bigl(B(n,n)+f_n(x)\bigr) \end{pmatrix} T_1^{-1}. \qquad (9)$$

Как показано в [1, гл. 5, § 5], при таком определении f для обратимой матрицы

$$\left(\sqrt{A}\right)^2 = A, \quad e^{\ln A} = A, \quad AA^{-1} = E.$$

Пример 1. Пусть дана действительная матрица

$$B = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Диагонализируем

$$B\bigl(f_i(x)\bigr) = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1+x & 2 \\ 0 & 0 & 1-x \end{pmatrix}$$

с помощью матрицы

$$S = \begin{pmatrix} 1 & \dfrac{1}{x} & \dfrac{1-2x}{x^2} \\ 0 & 1 & -\dfrac{1}{x} \\ 0 & 0 & 1 \end{pmatrix},$$

столбцы которой являются собственными векторами матрицы $B\big(f_i(x)\big)$. Имеем

$$f(B)=\lim_{x\to 0}\begin{pmatrix} f(1) & \dfrac{f(1+x)-f(1)}{x} & \dfrac{2f(1)(x-1)+f(1+x)+f(1-x)(1-2x)}{x^2} \\ 0 & f(1+x) & \dfrac{f(1+x)-f(1-x)}{x} \\ 0 & 0 & f(1-x) \end{pmatrix}.$$

Если f дифференцируема, то

$$f(B)=\begin{pmatrix} f(1) & f'(1) & f''(1)+2f'(1) \\ 0 & f(1) & 2f'(1) \\ 0 & 0 & f(1) \end{pmatrix}.$$

Пример 2. Пусть дана действительная матрица

$$A=\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

Тогда

$$T^{-1}AT=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}=B,$$

где

$$T=\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Диагонализируем

$$B\big(f_i(x)\big)=\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2+x & 0 \\ 0 & 0 & 2-x \end{pmatrix},$$

с помощью матрицы

$$S=\begin{pmatrix} 1 & \dfrac{1}{x} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

столбцы которой являются собственными векторами матрицы $B\big(f_i(x)\big)$. По формуле (6) имеем

$$f(B)=\begin{pmatrix} f(2) & \dfrac{f(2+x)-f(2)}{x} & 0\\ 0 & f(2+x) & 0\\ 0 & 0 & f(2-x)\end{pmatrix}\Bigg|_{x=0} =$$

$$=\begin{pmatrix} f(2) & \lim\limits_{x\to 0}\dfrac{f(2+x)-f(2)}{x} & 0\\ 0 & f(2) & 0\\ 0 & 0 & f(2)\end{pmatrix}.$$

Если f дифференцируема, то

$$f(B)=\begin{pmatrix} f(2) & f'(2) & 0\\ 0 & f(2) & 0\\ 0 & 0 & f(2)\end{pmatrix},\quad f(A)=\begin{pmatrix} f(2)-2f'(2) & f'(2) & 0\\ -4f'(2) & 2f'(2)+f(2) & 0\\ -2f'(2) & f'(2) & f(2)\end{pmatrix}.$$

Хорошо известно, что ряд Тейлора функции $f=e^{At}$ сходится для всех At. Поэтому

$$e^{At}=\begin{pmatrix} e^{2t}-2te^{2t} & -2te^{2t} & 0\\ -4te^{2t} & 2te^{2t}+e^{2t} & 0\\ -2te^{2t} & te^{2t} & e^{2t}\end{pmatrix}.$$

Пример 3. Пусть

$$B=\begin{pmatrix} 1 & 1 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

$B\in \mathrm{M}_n(F)$, $F=\mathrm{GF}(2)$. Диагонализируя матрицу

$$B\big(f_i(x)\big)=\begin{pmatrix} 1 & 1 & 1\\ 0 & 1+x & 0\\ 0 & 0 & 1+x^2\end{pmatrix}$$

с помощью матрицы

$$S=\begin{pmatrix} 1 & \dfrac{1}{x} & \dfrac{1}{x^2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

имеем

$$f(B)=\begin{pmatrix} f(1) & \dfrac{f(1+x)-f(1)}{x} & \dfrac{f(1)-f(1+x^2)}{x^2}\\ 0 & f(1+x) & 0\\ 0 & 0 & f(1+x^2)\end{pmatrix}\Bigg|_{x=0}.$$

Например,

$$B^{100}=\begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Литература

[1] Гантмахер Ф. Р. Теория матриц. — М.: Наука, 1967.
[2] Ланкастер П. Теория матриц. — М.: Наука, 1970.
[3] Хорн Р., Джонсон Ч. Матричный анализ. — М.: Мир, 1989.