Project for Sections 8.6, 8.9, 10.5
Matrix Functions Methods

Vladimir Dobrushkin, Ph.D.
Brown University

In many applications, such as differential equations, digital image processing (image recognition), and quantum mechanics, we see the need for evaluation of a function whose argument is a square matrix. The purpose of this exposition is to demonstrate the versatility of the method of matrix functions. The definition of function of a square matrix and the required background for working with functions of matrices will be presented. There are several approaches to define a function of matrix. The method presented here is credited to J. Sylvester[1] and requires just skills through calculus.

Matrix functions will help find the inverse of a square matrix $\mathbf{A}$. As one can see from section 8.6, actual calculation of $\mathbf{A}^{-1}$ is quite demanding. Indeed, the inverse matrices are used throughout Chapter 8. However, with matrix functions, one can just evaluate the function $f(\lambda) = \lambda^{-1}$ where $\lambda = \mathbf{A}$.

The powers of a square matrix $\mathbf{A}$ were calculated in section 8.9. Powers are important in many applications, especially in the probability theory and quantum mechanics. This problem was reduced for solving a system of algebraic equations and the reader was referred back to section 8.6. However, to find the m-th power one can just consider the function $f(\lambda) = \lambda^m$ and substitute the given matrix $\mathbf{A}$ for λ.

The construction of orthogonal matrices from section 8.10 involves the operation $\sqrt{\mathbf{A}\,\mathbf{B}}$ for the product of the matrices. This requires determining a square root of the matrix. Instead, this can be done with the aid of the function $f(\lambda) = \lambda^{1/2}$ (assuming that the only one branch of the analytic function is considered).

Chapter 10 is dedicated to solving first-order vector differential equations $\dfrac{\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{f}(t)$, for some given square matrix $\mathbf{A}$ and unknown vector function $\mathbf{x}(t)$. Its solution is expressed through the matrix exponential $e^{\mathbf{A}t}$, which can be found by calculating $e^{\lambda t}$ for $\lambda = \mathbf{A}$. Second-order equations require definition more complicated functions of matrices such as $\cos(\sqrt{\lambda}t)$ and $\sin(\sqrt{\lambda}t)/\sqrt{\lambda}$.

Recall that the **characteristic polynomial** $\Delta(\lambda)$ of a square matrix $\mathbf{A}$ is the determinant of the matrix $\lambda\mathbf{I} - \mathbf{A}$, that is,

$$\Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}).$$

Obviously $\Delta(\lambda)$ has leading term λ^n. Any solution of the characteristic equation $\Delta(\lambda) = 0$ is said to be **eigenvalue**. The set of all eigenvalues is called the **spectrum** of the matrix $\mathbf{A}$, denoted $\sigma(\mathbf{A})$.

A scalar polynomial $q(\lambda) = q_n\lambda^n + q_{n-1}\lambda^{n-1} + \cdots + q_1\lambda + q_0$ is called an **annihilating polynomial** (or *annulled polynomial*) of the square matrix $\mathbf{A}$, if $q(\mathbf{A}) = 0$, that is,

$$q(\mathbf{A}) \equiv q_n\mathbf{A}^n + q_{n-1}\mathbf{A}^{n-1} + \cdots + q_1\mathbf{A} + q_0\mathbf{I} = 0,$$

where $\mathbf{I}$ is the identity matrix (see page 353, §8.1). An annihilating polynomial $\psi(\lambda)$ of least degree with highest coefficient 1 is called a **minimal polynomial** of $\mathbf{A}$. It is known that the degree of the minimal polynomial of an n-by-n matrix $\mathbf{A}$ is less than or equal to n; moreover, every matrix $\mathbf{A}$ is annulled by its characteristic polynomial (Hamilton - Cayley theorem 8.26, see §8.9).

Let $f(\lambda)$ be a function defined on the spectrum of the matrix $\mathbf{A}$. Assuming that the minimal polynomial of $\mathbf{A}$ has only simple roots, we define $f(\mathbf{A})$ as

$$f(\mathbf{A}) = \sum_{k=1}^{s} f(\lambda_k)\, Z_k(\mathbf{A}), \tag{1}$$

[1]James Joseph Sylvester (1814 – 1897) was a man of many talents who took music lessons from Gounoud and was prouder of his "high C" than his matrix achievements. Florence Nightingale was one of his students.

where

$$Z_k(\mathbf{A}) = \frac{(\mathbf{A}-\lambda_1)\cdots(\mathbf{A}-\lambda_{k-1})(\mathbf{A}-\lambda_{k+1})\cdots(\mathbf{A}-\lambda_s)}{(\lambda_k-\lambda_1)\cdots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\cdots(\lambda_k-\lambda_s)}, \quad k = 1, 2, \dots s.$$

are so called Sylvester's auxiliary matrices.

Example 1. Let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{bmatrix} -3 & 2 & 2 \\ -6 & 5 & 2 \\ -7 & 4 & 4 \end{bmatrix}.$$

Then the minimal polynomial of this matrix is its characteristic polynomial, that is,

$$\psi(\lambda) = \det \begin{bmatrix} \lambda+3 & -2 & -2 \\ 6 & \lambda-5 & -2 \\ 7 & -4 & \lambda-4 \end{bmatrix} = (\lambda-1)(\lambda-2)(\lambda-3).$$

The spectrum of the given matrix $\mathbf{A}$ consists of three distinct values, $\{1, 2, 3\}$. Hence the auxiliary matrices are

$$Z_1(\mathbf{A}) = \frac{(\mathbf{A}-2\mathbf{I})(\mathbf{A}-3\mathbf{I})}{(1-2)(1-3)} = \frac{1}{2} \begin{bmatrix} -5 & 2 & 2 \\ -6 & 3 & 2 \\ -7 & 4 & 2 \end{bmatrix} \cdot \begin{bmatrix} -6 & 2 & 2 \\ -6 & 2 & 2 \\ -7 & 4 & 1 \end{bmatrix},$$

$$Z_2(\mathbf{A}) = \frac{(\mathbf{A}-\mathbf{I})(\mathbf{A}-3\mathbf{I})}{(2-1)(2-3)} = - \begin{bmatrix} -4 & 2 & 2 \\ -6 & 4 & 2 \\ -7 & 4 & 3 \end{bmatrix} \cdot \begin{bmatrix} -6 & 2 & 2 \\ -6 & 2 & 2 \\ -7 & 4 & 1 \end{bmatrix},$$

$$Z_3(\mathbf{A}) = \frac{(\mathbf{A}-\mathbf{I})(\mathbf{A}-2\mathbf{I})}{(3-1)(3-2)} = \frac{1}{2} \begin{bmatrix} -4 & 2 & 2 \\ -6 & 4 & 2 \\ -7 & 4 & 3 \end{bmatrix} \cdot \begin{bmatrix} -5 & 2 & 2 \\ -6 & 3 & 2 \\ -7 & 4 & 2 \end{bmatrix},$$

Therefore,

$$Z_1(\mathbf{A}) = \begin{bmatrix} 2 & 1 & -2 \\ 2 & 1 & -2 \\ 2 & 1 & -2 \end{bmatrix}, \quad Z_2(\mathbf{A}) = \begin{bmatrix} 2 & -4 & 2 \\ 2 & -4 & 2 \\ 3 & -6 & 3 \end{bmatrix}, \quad Z_3(\mathbf{A}) = \begin{bmatrix} -3 & 3 & 0 \\ -4 & 4 & 0 \\ -5 & 5 & 0 \end{bmatrix}.$$

It is easy to see that the minimal polynomial $\psi(\lambda)$ annihilates the given matrix because

$$\psi(\mathbf{A}) = (\mathbf{A}-\mathbf{I})(\mathbf{A}-2\mathbf{I})(\mathbf{A}-3\mathbf{I}) = \begin{bmatrix} -4 & 2 & 2 \\ -6 & 4 & 2 \\ -7 & 4 & 3 \end{bmatrix} \begin{bmatrix} -5 & 2 & 2 \\ -6 & 3 & 2 \\ -7 & 4 & 2 \end{bmatrix} \begin{bmatrix} -6 & 2 & 2 \\ -6 & 2 & 2 \\ -7 & 4 & 1 \end{bmatrix} = \mathbf{0}.$$

For functions $f(\lambda) = (\lambda+1)/(\lambda+2)$ and $g(\lambda) = e^{\lambda t}$ we have

$$\begin{aligned} f(\mathbf{A}) &= \frac{\mathbf{A}+\mathbf{I}}{\mathbf{A}+2\mathbf{I}} = \frac{1+1}{1+2} Z_1(\mathbf{A}) + \frac{2+1}{2+2} Z_2(\mathbf{A}) + \frac{3+1}{3+2} Z_3(\mathbf{A}) \\ &= \frac{2}{3} \begin{bmatrix} 2 & 1 & -2 \\ 2 & 1 & -2 \\ 2 & 1 & -2 \end{bmatrix} + \frac{3}{4} \begin{bmatrix} 2 & -4 & 2 \\ 2 & -4 & 2 \\ 3 & -6 & 3 \end{bmatrix} + \frac{4}{5} \begin{bmatrix} -3 & 3 & 0 \\ -4 & 4 & 0 \\ -5 & 5 & 0 \end{bmatrix}. \end{aligned}$$

Similarly,

$$\begin{aligned} g(\mathbf{A}) &= e^{\mathbf{A}t} = e^t Z_1(\mathbf{A}) + e^{2t} Z_2(\mathbf{A}) + e^{3t} Z_3(\mathbf{A}) \\ &= e^t \begin{bmatrix} 2 & 1 & -2 \\ 2 & 1 & -2 \\ 2 & 1 & -2 \end{bmatrix} + e^{2t} \begin{bmatrix} 2 & -4 & 2 \\ 2 & -4 & 2 \\ 3 & -6 & 3 \end{bmatrix} + e^{3t} \begin{bmatrix} -3 & 3 & 0 \\ -4 & 4 & 0 \\ -5 & 5 & 0 \end{bmatrix}. \end{aligned}$$

To find the inverse matrix use the function $h(\lambda) = \lambda^{-1}$ to obtain

$$\mathbf{A}^{-1} = Z_1(\mathbf{A}) + \frac{1}{2}\, Z_2(\mathbf{A}) + \frac{1}{3}\, Z_3(\mathbf{A}) = \begin{bmatrix} 2 & 0 & -1 \\ \frac{5}{3} & \frac{1}{3} & -1 \\ \frac{11}{6} & -\frac{1}{3} & -\frac{1}{2} \end{bmatrix}.$$

Example 2. Let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

While its characteristic polynomial is $\det(\lambda\mathbf{I} - \mathbf{A}) = (\lambda+2)^2(\lambda-1)$, the corresponding minimal polynomial, $\psi(\lambda) = (\lambda+2)(\lambda-1) = \lambda^2 + \lambda - 2$, has degree 2. Indeed,

$$\mathbf{A}^2 + \mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 1 & -3 & -3 \\ 3 & 7 & 3 \\ -3 & -3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \mathbf{0}.$$

Then the auxiliary matrices are

$$Z_1(\mathbf{A}) = \frac{\mathbf{A} + 2\mathbf{I}}{1+2} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad Z_{-2}(\mathbf{A}) = \frac{\mathbf{A} - \mathbf{I}}{-2-1} = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

For the function $f(\lambda) = e^{\lambda t}$, we have

$$e^{\mathbf{A}t} = e^{-2t} \begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix} + e^t \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Since we know the auxiliary matrices, the inverse of the matrix $\mathbf{A}$, which corresponds to the function $f(\lambda) = \lambda^{-1}$, can be obtain almost without efforts:

$$\mathbf{A}^{-1} = Z_1(\mathbf{A}) - \frac{1}{2}\, Z_{-2}(\mathbf{A}) = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -1 & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & \frac{3}{2} & \frac{3}{2} \\ -\frac{3}{2} & -2 & -\frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} & 1 \end{bmatrix}$$

Problems. Using this new technique, functions of matrices, rework the following problems from the textbook *Advanced Engineering Mathematics*, 3/e, by Dennis Zill and Michael Cullen.

1. Page 393, Problems 15, 16, 17, 18
2. Page 394, Problems 43
3. Page 407, Problems 3 – 5
4. Page 591, Problems 1, 2
5. Page 603, Problems 11, 12, 15, 16, 18