

Nick Higham

# What Is a Matrix Function?

If you give an array as the argument to a mathematical function in a programming language or problem solving environment you are likely to receive the result of applying the function to each component of the array. For example, here MATLAB exponentiates the integers from 0 to 3:

```
>> A = [0 1; 2 3], X = exp(A)
A =
     0     1
     2     3
X =
 1.0000e+00  2.7183e+00
 7.3891e+00  2.0086e+01
```

When the array represents a matrix this componentwise evaluation is not useful, as it does not obey the rules of matrix algebra. To see what properties we might require, consider the matrix square root. This function is notable historically because Cayley considered it in the 1858 paper in which he introduced matrix algebra.

A square root of an  $n \times n$  matrix  $A$  is a matrix  $X$  such that  $X^2 = A$ . Any square root  $X$  satisfies

$$AX = X^2X = XX^2 = XA,$$

so  $X$  commutes with  $A$ . Also,  $A^* = (X^2)^* = (X^*)^2$ , so  $X^*$  is a square root of  $A^*$  (here, the superscript  $*$  denotes the conjugate transpose).

Furthermore, for any nonsingular matrix  $Z$  we have

$$(Z^{-1}XZ)^2 = Z^{-1}X^2Z = Z^{-1}AZ.$$

If we choose  $Z$  as a matrix that takes  $A$  to its Jordan canonical form  $J$  then we have  $(Z^{-1}XZ)^2 = J$ , so that  $Z^{-1}XZ$  is a square root of  $J$ , or in other words  $X$  can be expressed as  $X = Z\sqrt{J}Z^{-1}$ .

Generalizing from these properties of the matrix square root it is reasonable to ask that a function  $f(A)$  of a square matrix  $A$  satisfies

- $f(A)$  commutes with  $A$ ,
- $f(A^*) = f(A)^*$ ,
- $f(A) = Zf(J)Z^{-1}$ , where  $A$  has the Jordan canonical form  $A = ZJZ^{-1}$ .

(These are of course not the only requirements; indeed  $f(A) \equiv I$  satisfies all three conditions.)

Many definitions of  $f(A)$  can be given that satisfy these and other natural conditions. We will describe just three definitions (a notable omission is a definition in terms of polynomial interpolation).

## Power Series

For a function  $f$  that has a power series expansion we can define  $f(A)$  by substituting  $A$  for the variable. It can be shown that the matrix series will be convergent for matrices whose eigenvalues lie within the radius of convergence of the scalar power series. For example, where  $\rho$  denotes the spectral radius. This definition is natural for functions that have a power series expansion, but it is rather limited in its applicability.

$$\begin{aligned}\cos(A) &= I - \frac{A^2}{2!} + \frac{A^4}{4!} - \frac{A^6}{6!} + \cdots, \\ \log(I + A) &= A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \cdots, \quad \rho(A) < 1,\end{aligned}$$

## Jordan Canonical Form Definition

If  $A$  has the Jordan canonical form

$$Z^{-1}AZ = J = \text{diag}(J_1, J_2, \dots, J_p),$$

where

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} = \lambda_k I + N \in \mathbb{C}^{m_k \times m_k},$$

then

$$f(A) := Zf(J)Z^{-1} = Z\text{diag}(f(J_k))Z^{-1},$$

where

$$f(J_k) := \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \cdots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}.$$

Notice that  $f(J_k)$  has the derivatives of  $f$  along its diagonals in the upper triangle. Write  $J_k = \lambda_k I + N$ , where  $N$  is zero apart from a superdiagonal of 1s. The formula for  $f(J_k)$  is just the Taylor series expansion

$$f(J_k) = f(\lambda_k I + N) = f(\lambda_k)I + f'(\lambda_k)N + \frac{f''(\lambda_k)}{2!}N^2 + \dots + \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!}N^{m_k-1}.$$

The series is finite because  $N^{m_k} = 0$ : as  $N$  is powered up the superdiagonal of 1s moves towards the right-hand corner until it eventually disappears.

An immediate consequence of this formula is that  $f(A)$  is defined only if the necessary derivatives exist: for each eigenvalue  $\lambda_k$  we need the existence of the derivatives at  $\lambda_k$  of order up to one less than the size of the largest Jordan block in which  $\lambda_k$  appears.

When  $A$  is a diagonalizable matrix the definition simplifies considerably: if  $A = ZDZ^{-1}$  with  $D = \text{diag}(\lambda_i)$  then  $f(A) = Zf(D)Z^{-1} = Z\text{diag}(f(\lambda_i))Z^{-1}$ .

## Cauchy Integral Formula

For a function  $f$  analytic on and inside a closed contour  $\Gamma$  that encloses the spectrum of  $A$  the matrix  $f(A)$  can be defined by the Cauchy integral formula

$$f(A) := \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1} dz.$$

This formula is mathematically elegant and can be used to provide short proofs of certain theoretical results.

## Equivalence of Definitions

These and several other definitions turn out to be equivalent under suitable conditions. This was recognized by Rinehart, who wrote in 1955

“There have been proposed in the literature since 1880 eight distinct definitions of a matrix function, by Weyr, Sylvester and Buchheim, Giorgi, Cartan, Fantappiè, Cipolla, Schwerdtfeger and Richter ... All of the definitions except those of Weyr and Cipolla are essentially equivalent.”

## Computing a Matrix Function in MATLAB

In MATLAB, matrix functions are distinguished from componentwise array evaluation by the postfix “m” on a function name. Thus `expm`, `logm`, and `sqrtn` are matrix function counterparts of `exp`, `log`, and `sqrt`. Compare the example at the start of this article with

```
>> A = [0 1; 2 3], X = expm(A)
A =
    0    1
    2    3
X =
 5.2892e+00  8.4033e+00
 1.6807e+01  3.0499e+01
```

The MATLAB function `funm` calculates  $f(A)$  for general matrix functions  $f$  (subject to some restrictions).

## References

This is a minimal set of references, which contain further useful references within.

- Nicholas J. Higham, **Functions of Matrices: Theory and Computation** (<https://www.maths.manchester.ac.uk/~higham/fm/index.php>), Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.
- Roger Horn and Charles Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
- R. F. Rinehart, **The Equivalence of Definitions of a Matric Function** (<https://doi.org/10.1080/00029890.1955.11988651>), Amer. Math. Monthly 62(6), 395-414, 1955.

## Related Blog Posts

- **Arthur Buchheim (1859–1888)** (<https://nhigham.com/2013/01/31/arthur-buchheim/>) (2013)
- **Update of Catalogue of Software for Matrix Functions** (<https://nhigham.com/2020/03/27/update-of-catalogue-of-software-for-matrix-functions/>) (2020)
- **What Is a Matrix Square Root?** (<https://nhigham.com/2020/05/21/what-is-a-matrix-square-root/>) (2020)
- **What Is the Matrix Exponential?** (<https://nhigham.com/2020/05/28/what-is-the-matrix-exponential/>) (2020)

This article is part of the “What Is” series, available from <https://nhigham.com/category/what-is> (<https://nhigham.com/category/what-is>) and in PDF form from the GitHub repository <https://github.com/higham/what-is> (<https://github.com/higham/what-is>).

Posted on **June 9, 2020** by **Nick Higham** Posted in **what-is** Tagged **matrix\_function**.

**Website Powered by WordPress.com.**