Matrix functions

Nathan Rousselot

Mathematical Engineering Department

KU Leuven

July 12, 2023

1 Introduction

In this document we introduce the notion of matrix functions. Say we have a function $f : \mathbb{C} \to \mathbb{C}$, then we can define the function f on a matrix $\mathbf{A}$ as follows: $f : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$. In the following, we will first provide some theretical background on matrix function. The, we will describe the numerical issues coming with manipulating matrix functions. Finally, we will provide algorithm for efficient computation of matrix fuctions.

2 Theoretical background

2.1 Natural Definition

Polynomial functions

Let $p : \mathbb{C} \to \mathbb{C}$ be a polynomial function of degree d:

$$p(t) = \sum_{k=0}^{d} c_k t^k \tag{1}$$

Then, considering a matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$, and posing $\mathbf{A}^0 = I_n$, we can define the polynomial function $p : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ on a matrix $\mathbf{A}$ as follows:

$$p(\mathbf{A}) = \sum_{k=0}^{d} c_k \mathbf{A}^k \tag{2}$$

Rational functions

Let $f : \mathbb{C} \to \mathbb{C}$ be a rational function of the form:

$$f(t) := \frac{p(t)}{q(t)} \tag{3}$$

It is not immediate how one would approach this function with a matrix. We want to define

$$f(\mathbf{A}) := q(\mathbf{A})^{-1}p(\mathbf{A}) \tag{4}$$

However, this is not well defined if $q(\mathbf{A})$ is singular. In other words, we need to make sure that $q(\mathbf{A})$ is invertible. This is the case if and only if $q(\lambda) \neq 0$ for all eigenvalues λ of $\mathbf{A}$. This is a very strong condition, and it is not always possible to find a rational function f such that $q(\lambda) \neq 0$ for all eigenvalues λ of $\mathbf{A}$. Furthermore, there is an ambiguity in the definition of $f(\mathbf{A})$ as one could want to express $f(\mathbf{A}) := p(\mathbf{A})q(\mathbf{A})^{-1}$ instead. Indeed, in matrix computations, multiplication is not (necessarly) commutative. Thus, in some cases, $q(\mathbf{A})^{-1}p(\mathbf{A}) \neq p(\mathbf{A})q(\mathbf{A})^{-1}$.

Theorem 1. *Let $p : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ and $q : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ be two polynomial functions of degree d. Then, if $q(\mathbf{A})$ is non-singular, and if $p(\mathbf{A})$ and $q(\mathbf{A})$ commute, then $f(\mathbf{A}) := q(\mathbf{A})^{-1}p(\mathbf{A}) = p(\mathbf{A})q(\mathbf{A})^{-1}$.*

Proof. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a matrix such that $q(\mathbf{A})$ is non-singular, and such that $p(\mathbf{A})$ and $q(\mathbf{A})$ commute. Then, we have:

$$\begin{aligned}
q(\mathbf{A})^{-1}p(\mathbf{A}) &= q(\mathbf{A})^{-1}p(\mathbf{A})q(\mathbf{A})q(\mathbf{A})^{-1} \\
&= q(\mathbf{A})^{-1}q(\mathbf{A})p(\mathbf{A})q(\mathbf{A})^{-1} \\
&= p(\mathbf{A})q(\mathbf{A})^{-1}
\end{aligned}$$

$\square$

Power Series

Let $f : \mathbb{C} \to \mathbb{C}$ be a function that can be expressed as a power series:

$$f(t) = \sum_{k=0}^{\infty} c_k t^k \tag{5}$$

Then, considering a matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$, and posing $\mathbf{A}^0 = I_n$, we can define the power series function $f : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ on a matrix $\mathbf{A}$ as follows:

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^k \tag{6}$$

In the scalar case, we know that the power series converges if $|t| < r$, where r is the radius of convergence. Obviously this translates in the matrix power series

Theorem 2. *Let $f : \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$ be a matrix power series. Then, the series converges if and only if $\rho(\mathbf{A}) < r$, where $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, and r is the radius of convergence of the scalar power series.*

Proof is provided in Frommer and Simoncini 2008. In the case of a finite-order Laurent Series, *i.e*:

$$f(t) = \sum_{k=-d}^{d} c_k t^k \tag{7}$$

For the matrix case, we need to ensure convergence (similarly to power series), but also ensure existance of the inverse, as Laurent series do have negative powers. If both of those conditions are satisfied, we can write the Laurent series as a matrix function:

$$f(\mathbf{A}) = \sum_{k=-d}^{d} c_k \mathbf{A}^k \tag{8}$$

2.2 Spectrum-Based Definition

Diagonalizable Matrices

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a diagonalizable matrix. That means that there exists a matrix $\mathbf{V} \in \mathbb{C}^{n \times n}$ such that $\mathbf{V}$ is invertible, and $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}$, where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{A}$:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{9}$$

Theorem 3. *Let* $\mathbf{A} \in \mathbb{C}^{n \times n}$ *be a diagonalizable matrix. Then we can define the function* $f(\mathbf{A})$ *as:*

$$f(\mathbf{A}) := \mathbf{V} f(\mathbf{\Lambda}) \mathbf{V}^{-1} \tag{10}$$

with

$$f(\mathbf{\Lambda}) = \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{bmatrix} \tag{11}$$

This property is very handy, as it allows us to compute matrix functions by simply applying the function to the eigenvalues of the matrix. Computationally, this avoids inverting matrices, and lots of matrix products. However, this property is only valid for diagonalizable matrices. This property puts constraints on $f(\mathbf{A})$, as its eigenvectors must form a basis $\mathbf{F}^n$. Another more practical constraint, that is sufficient but not necessary, is if $\mathbf{A}$ is a full rank matrix, then it is diagonalizable.

Defective Matrices

In some cases, the matrix $\mathbf{A}$ is not diagonalizable, that means the sum of the dimensions of the eigenspaces is less than n, we call that a *Defective Matrix*. In that case, we can generalize the principle of diagonalization using the Jordan canonical form of $\mathbf{A}$:

$$\mathbf{A} = \mathbf{V} \mathbf{J} \mathbf{V}^{-1} \tag{12}$$

where $\mathbf{J}$ is a Jordan matrix, and $\mathbf{V}$ is a matrix containing the generalized eigenvectors of $\mathbf{A}$. The Jordan matrix is a block diagonal matrix, where each block is a Jordan block. A Jordan block is a matrix of the form:

$$\mathbf{J}_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \tag{13}$$

Theorem 4. *Let* $\mathbf{A} \in \mathbb{C}^{n \times n}$ *be a defective matrix. Then we can define the function* $f(\mathbf{A})$ *as:*

$$f(\mathbf{A}) := \mathbf{V} f(\mathbf{J}) \mathbf{V}^{-1} \tag{14}$$

4

with

$$
f(\mathbf{J}) = \begin{bmatrix} f(\mathbf{J}_1(\lambda_1)) & 0 & \cdots & 0 \\ 0 & f(\mathbf{J}_2(\lambda_2)) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\mathbf{J}_k(\lambda_k)) \end{bmatrix} \tag{15}
$$

and

$$
f(\mathbf{J}_i(\lambda_i)) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{f''(\lambda_i)}{2!} & \cdots & \frac{f^{(k-1)}(\lambda_i)}{(k-1)!} \\ 0 & f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{f^{(k-2)}(\lambda_i)}{(k-2)!} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_i) & f'(\lambda_i) \\ 0 & 0 & \cdots & 0 & f(\lambda_i) \end{bmatrix} \tag{16}
$$

Obviously, both definitions of matrix functions, based on diagonalization and Jordan canonical form, presuppose that the spectral radius of $\mathbf{A}$, $\rho(\mathbf{A})$, is less than r, the radius of convergence.

References

Frommer, Andreas and Valeria Simoncini (2008). "Matrix functions". In: *Model order reduction: theory, research aspects and applications*. Springer, pp. 275–303.