J. Math. Anal. Appl. 462 (2018) 613–636

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Evaluating non-analytic functions of matrices

Nir Sharon [a,*], Yoel Shkolnisky [b]

[a] The Program in Applied and Computational Mathematics, Princeton University, Princeton NJ, USA
[b] School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel

ARTICLE INFO

Article history:
Received 30 October 2017
Available online xxxx
Submitted by J.A. Ball

Keywords:
Matrix functions
Chebyshev polynomials
Matrix Chebyshev expansion
Convergence rates
Jordan blocks

ABSTRACT

The paper revisits the classical problem of evaluating $f(A)$ for a real function f and a matrix A with real spectrum. The evaluation is based on expanding f in Chebyshev polynomials, and the focus of the paper is to study the convergence rates of these expansions. In particular, we derive bounds on the convergence rates which reveal the relation between the smoothness of f and the diagonalizability of the matrix A. We present several numerical examples to illustrate our analysis.

1. Introduction

We revisit the problem of lifting a real function $f\colon \mathbb{R} \to \mathbb{R}$ to a matrix function $f\colon \mathcal{M}_k(\mathbb{R}) \to \mathcal{M}_k(\mathbb{R})$, where $\mathcal{M}_k(\mathbb{R})$ is the set of square real matrices of size $k \times k$ having real spectrum. When f is a polynomial, such lifting is straightforward since addition and powers are well-defined for square matrices. When f is not a polynomial, there are several standard methods to define the above-mentioned lifting. If f is analytic having a Taylor expansion whose convergence radius is larger than the spectral radius of A, then the Taylor expansion $f(x) = \sum_{n=0}^{\infty} \alpha_n x^n$ yields $f(A) = \sum_{n=0}^{\infty} \alpha_n A^n$. If f is not analytic, it is required that at least $f \in C^{m-1}$, where m is the size of the largest Jordan block of A. This condition allows defining $f(A)$ on each of the Jordan blocks of A. This latter approach has several equivalent definitions, see e.g. [10, Chapter 1].

Chebyshev polynomials are ubiquitous in applied mathematics and engineering.[1] These polynomials arise as solutions of a Sturm–Liouville ODE and are used in numerous approximation methods, ranging from classical PDE methods [16] to modern methods for image denoising [18]. Motivated by the favorable

* Corresponding author.
E-mail addresses: nsharon@math.princeton.edu (N. Sharon), yoelsh@post.tau.ac.il (Y. Shkolnisky).

[1] “Chebyshev polynomials are everywhere dense in numerical analysis”. This quote by Philip Davis and George Forsythe is the opening sentence of [16].

https://doi.org/10.1016/j.jmaa.2018.02.029
0022-247X/

numerical properties of Chebyshev polynomials in representing and approximating scalar functions, we rigorously study the use of Chebyshev expansions for matrix functions. The idea of evaluating a matrix function by its Chebyshev expansion is not new and has been used in applications before. For example, [26] uses Chebyshev expansion in spectral methods for solving PDEs. In this context of solving PDEs, the Chebyshev polynomials are also examined as a special case of ultraspherical polynomials [5], and Faber polynomials [25]. Expansion in the Faber polynomials (for complex spectra) appears in [19] for matrices. Chebyshev polynomials are also widely used for pseudospectral methods, see [27] and reference therein. In all of the above papers, one assumes that f is a smooth function. One of the most frequently used analytic function is the exponential function, naturally arising in solving differential equations. In [1] Chebyshev polynomials are proved to be an effective alternative to Krylov techniques for calculating $\exp(A)v$, for a given vector v. Another application of Chebyshev polynomials is in representing the best matrix 2-norm approximation for analytic functions, over the space of polynomials of a fixed degree [15]. In [3,24] matrix Chebyshev polynomials are used for slicing the spectrum of a matrix in order to extract interior eigenvalues. Another kind of spectrum filtering is presented in [18] for the construction of image denoising operators. We can also find matrix Chebyshev polynomials in calculating matrix functions of symmetric matrices [7], computing square roots of the covariance matrix of Gaussian random fields [4] and facilitating the estimation of autoregressive models [20].

1.1. Our contribution

In this paper, we generalize the study of matrix Chebyshev expansions to cases where the matrix is not necessarily diagonalizable, and the function is not necessarily analytic. In such cases, there exists a trade-off between two factors; "how much" the matrix is far from being diagonalizable, as expressed by the size of the largest Jordan block in the Jordan form of the matrix, and "how smooth is the function". Specifically, as the size of the largest Jordan block increases ("less diagonalizable") the smoothness of the function required to guarantee the convergence of the matrix Chebyshev expansion, increases as well.

In the current paper, we mainly focus on examining convergence issues and proving convergence rates. The convergence rate of the matrix Chebyshev expansion is crucial for determining how many coefficients one has to use to approximate $f(A)$ to a prescribed accuracy using a truncated Chebyshev expansion. Thus, the convergence rate directly affects the efficiency of approximation algorithms that use matrix Chebyshev expansions.

Chebyshev polynomials are naturally define on $[-1, 1]$, so without loss of generality, we assume that the spectrum of A is linearly transformed to this segment using an estimated bound on the spectral radius of A (for example, by estimating the eigenvalue of A with the largest magnitude). We henceforth assume that all eigenvalues of A lie in $[-1, 1]$. Given this assumption, we divide our analysis into two cases: a case where there is no a priori information about the distribution of the eigenvalues over $[-1, 1]$, and a case where all non-semisimple eigenvalues are concentrated inside $(-1, 1)$. As expected, the latter case implies fewer restrictions on the smoothness of the function. In particular, if we denote by m the size of the largest Jordan block of A. Then, the convergence of the matrix Chebyshev expansion is guaranteed in the first case for any $f \in C^{2m-2}$ where $f^{(2m-1)}$ is of bounded variation. In the second case, it is guaranteed for any $f \in C^{m-1}$ where $f^{(m)}$ is of bounded variation. As for convergence rates, we summarize our results for the different cases in Table 1. Each row in the table presents the requirements, from both the matrix and the function, so that an expansion of length N has an error bound of $cN^{-\ell}$ ($\ell > 1$). The constant c is independent of N and ℓ but does depend on the factors specified in the 'Remarks' column. We note two properties from the table. First, for matrices with a concentrated spectrum, the smoothness assumptions on f are weaker than in the general case. Second, one can guarantee a constant which is independent of the size of the matrix by imposing additional regularity on the function f. Further conclusions and their proofs appear in the text.