On best approximation by polynomials of matrices

Vance Faber, Jörg Liesen, Petr Tichý

Abstract

An important part of approximation theory is concerned with the approximation of a given function f on some (compact) set Ω in the complex plane by polynomials. Classical results in this area deal with the best approximation problem

$$\min_{p \in \mathcal{P}_m} \|f - p\|_\Omega \qquad \text{where} \qquad \|g\|_\Omega = \max_{z \in \Omega} |g(z)|, \tag{1}$$

and $\mathcal{P}_m$ is the set of polynomials of degree at most m.

Scalar approximation problems of the form (1) have been studied since the mid 1850s. Accordingly, numerous results on existence and uniqueness of the solution as well as estimates for the value of (1) are known. Here we consider a problem that at first sight looks similar, but apparently is much less understood: Let f be a function that is analytic in a neighbourhood of the spectrum of a given matrix $A \in \mathbb{C}^{n \times n}$, so that $f(A)$ is well defined, and let $\|\cdot\|$ be a given matrix norm. Consider the matrix approximation problem

$$\min_{p \in \mathcal{P}_m} \|f(A) - p(A)\|. \tag{2}$$

We ask two basic questions: *Does this problem have a unique solution? Can we understand some properties of polynomials that solve this kind of best approximation problems?*

An answer to the first question depends of course on the norm used in (2). If the norm is known to be *strictly convex*, as for example the Frobenius norm, then (2) is guaranteed to have a uniquely defined solution as long as the value of (2) is positive. A useful matrix norm that is met in many applications is the matrix 2-norm (spectral norm), which for a given matrix A is equal to the largest singular value of A. This norm is not strictly convex, and thus the general result on uniqueness of best approximation in linear spaces with a strictly convex norm does not apply. In this presentation we will consider matrix approximation problems in the matrix 2-norm.

It is well known that when the function f is analytic in an open neighborhood of the spectrum of the matrix $A \in \mathbb{C}^{n \times n}$, then $f(A)$ is a well-defined complex $n \times n$ matrix. In fact, $f(A) = p_f(A)$, where p_f is a polynomial that depends on the values and possibly the derivatives of f on the spectrum of A. Therefore, f in (2) can be thought to be a polynomial of degree, say, $m + \ell + 1$ ($m \geq 0$, $\ell \geq 0$). Then, as we show in [4], the problem (2) can equivalently be written in the form

$$\min_{h \in \mathcal{P}_m} \|A^{m+1} g(A) - h(A)\|, \tag{3}$$

where g is a given polynomial of degree at most ℓ. We can also consider a related problem

$$\min_{g \in \mathcal{P}_\ell} \|A^{m+1} g(A) - h(A)\|, \tag{4}$$

where h is a given polynomial of degree at most m, and the best $g \in \mathcal{P}_\ell$ is sought.

Special cases of problems (3) and (4) have been considered by Greenbaum and Trefethen [2] in the context of convergence of Krylov subspace methods for solving linear systems and eigenvalue problems. In particular, they considered the approximation problems

$$\min_{h \in \mathcal{P}_m} \|A^{m+1} - h(A)\| \quad \text{and} \quad \min_{g \in \mathcal{P}_\ell} \|A g(A) - I\|, \tag{5}$$

the first one called the *ideal Arnoldi approximation problem*, and the second one called *ideal GMRES approximation problem*. Greenbaum and Trefethen seem to be the first who studied existence and uniqueness of polynomials that solve the problems (5). In particular, in [2] it is shown that if the minima in (5) are nonzero, and if A is nonsingular in the case of ideal GMRES, then both problems (5) have a unique minimizer. In the first part of this talk we present the results of our paper [4] and generalize results of Greenbaum and Trefethen to problems of the form (3) and (4):

Provided that the minimum in (3) is nonzero, the problem (3) has a unique minimizer. Provided that the minimum in (4) is nonzero and A is nonsingular, the problem (4) has a unique minimizer.

Note that the assumption of nonsingularity in the second case is in general necessary.

In a later paper, Toh and Trefethen [5] have called the polynomial that solves the problem

$$\min_{h \in \mathcal{P}_m} \|A^{m+1} - h(A)\| = \min_{p \in \mathcal{M}_{m+1}} \|p(A)\| \;\equiv\; \varphi_{m+1}(A) \tag{6}$$

the $(m+1)$st *Chebyshev polynomial of A*. Here $\mathcal{M}_{m+1}$ denotes the class of monic polynomials of degree $m+1$. The reason for this terminology is the following: When the matrix A is normal, i.e. unitarily diagonalizable, problem (6) for the matrix 2-norm becomes a scalar approximation problem of the form (1), with Ω being the set of eigenvalues of A, and the resulting monic polynomial is the $(m+1)$st Chebyshev polynomial on the (discrete) set of eigenvalues of A. Apart from the work of Greenbaum, Toh and Trefethen, and some further numerical examples in the recent book [3], very little has been published on Chebyshev polynomials of matrices, let alone the more general problem (2).

In second part of the talk we will present results of our recent paper [1] that studies general properties of Chebyshev polynomials of matrices. In some cases, these properties turn out to be generalizations of well known properties of Chebyshev polynomials of compact sets in the complex plane. For example, it is well known that Chebyshev polynomials for compact sets are characterized by alternation properties. A similar property can also be shown for Chebyshev polynomials of block diagonal matrices: We show that the minimal norm $\varphi_{m+1}(A)$ is attained on several blocks simultaneously. We also present explicit formulas of the Chebyshev polynomials of certain classes of matrices, including Jordan blocks, perturbed Jordan blocks and special classes of bidiagonal matrices, and explore the relation between Chebyshev polynomials of these classes of matrices and Chebyshev polynomials of sets in the complex plane.

References

[1] V. Faber, J. Liesen and P. Tichý, *On Chebyshev polynomials of matrices*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2205-2221.

[2] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.

[3] N. J. Higham, *Functions of Matrices: Theory and Computation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[4] J. Liesen and P. Tichý, *On best approximations of polynomials in matrices in the matrix 2-norm*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 853–863.

[5] K.-C. Toh and L. N. Trefethen, *The Chebyshev polynomials of a matrix*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 400–419.