8th International Conference on
Combinatorics, Cryptography, Computer Science and Computation
November. 15-16, 2023

Approximating Entries of Functions of Band Matrices

Majed Hamadi[1], Nezam Mahdavi-Amiri[2]
Department of Mathematical Sciences, Sharif University of Technology, Tehran, Iran

[1]e-mail: majed.hamadi77@sharif.edu
[2]e-mail: nezamm@sharif.edu

Abstract

We analyze and approximate the elements of a band matrix function and present methods for computing the matrix function and its trace with a complexity of $\mathcal{O}(n)$. We also present an $\mathcal{O}(1)$ algorithm for the function of band-Toeplitz matrices.

Keywords: matrix function, banded matrix, numerical range, divide-and-conquer algorithm
AMS Mathematical Subject Classification [2010]: 15A15, 65F60

1 Introduction

Recently, there has been interest in computing matrix functions, such as the matrix exponential or the matrix square root [8]. Function of a matrix apear in sevral problems, including the numerical solution of partial differential equations [6], electronic structure calculations [1], and social network analysis [7]. A matrix function, as an analytic function f of a square matrix A, can be represented through a contour integral,

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

where f is analytic on and inside a closed contour Γ that encloses $\Lambda(A)$[8]. Furthermore, if A is diagonalizable, that is,

$$A = P\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}P^{-1},$$

[1]Speaker, majed.hamadi77@sharif.edu
[2]nezamm@sharif.edu

then to compute $f(A)$, we will have:

$$f(A) = P \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} P^{-1}.$$

Cortinovis et al. [4] presented a divide-and-conquer algorithm, Algorithm 1, for functions of band matrices. Here, we investigate Algorithm 1 in Section 3 and derive a new error bound for approximating the trace of a band matrix function.

In Section 3, we first establish approximations for the elements of the band matrix function (Theorem 3.1). Then, we propose new approximations and error bounds for the trace of the band matrix function (Corollary 3.3). Additionally, we offer an approximation for a submatrix of the band matrix function (Theorem 3.4). In Section 4, we introduce an algorithm based on the work of [4], employing a new definition for a submatrix. In Section 5, we introduce a new method with a computational complexity of $\mathcal{O}(1)$ for computing the function of band-Toeplitz and band-Hankel matrices.

2 Preliminaries

Before delving into the approximation, in this section we present some notations and definitions. For an arbitrary matrix $A \in \mathbb{C}^{n\times n}$, and a non-negative integer r, we define the set α_p as follows:

$$\alpha_p(r) := \Big\{ t \in \{1,\dots,n\} \;\Big|\; |t-p| \le r \Big\}.$$

We also define the submatrix $\mathcal{B}_r([A]_{ij})$ of A as follows:

$$\mathcal{B}_r([A]_{ij}) := A[\alpha_i(r), \alpha_j(r)]. \tag{1}$$

Based on this definition, the submatrix $\mathcal{B}_r([A]_{ij})$ contains elements from matrix A being around the element $[A]_{ij}$. This submatrix has a size of $s_j \times s_i$, where $s_p = \max(\alpha_p) - \min(\alpha_p) + 1$. Also, we define the function $\varphi_x : \alpha_x(r) \to \{1,\dots,n\}$ as follows:

$$\varphi_x(p) = p - \min(\alpha_x(r)) + 1. \tag{2}$$

Now, the *field of values* or *numerical range* for matrix A is defined as a subset of $\mathbb{C}$:

$$\mathcal{W}(A) := \{\mathbf{x}^H A \mathbf{x} \mid \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\|_2 = 1\}.$$

It is worth noting that numerical range of a matrix is a convex, bounded, and compact set that contains the eigenvalues of the matrix. Also, if A_k is a principal submatrix of A, then $\mathcal{W}(A_k) \subseteq \mathcal{W}(A)$.

Now, we can state the result, due to Crouzeix and Palencia [5], that if $A \in \mathbb{C}^{n\times n}$ and the function f is an analytic function over $\mathcal{W}(A)$ and bounded on its boundary, then

$$\|f(A)\|_2 \leq \mathcal{Q}\|f\|_{\mathcal{W}(A)}, \tag{3}$$

where $2 \leq \mathcal{Q} \leq 1+\sqrt{2}$, and

$$\|f\|_{\mathcal{W}(A)} := \sup_{z\in\mathcal{W}(A)} |f(z)|.$$

The result given as (3) is helpful for finding an error bound for the approximation considered in the following sections. Now, considering Π_k as the set of polynomials of degree at most k, with $p_k \in \Pi_k$, for the entry (i,j), we have:

$$\left|\left[f(A)-p_k(A)\right]_{ij}\right| \leq \|(f-p_k)(A)\|_2 \leq \mathcal{Q}\|f-p_k\|_{\mathcal{W}(A)}. \tag{4}$$

3 Band Matrix Function Approximation

A matrix with the non-zero elements being located in a band around the main diagonal is called a band matrix. An m-band matrix A is defined as follows:

$$[A]_{ij} = 0, \quad |i-j| > m,$$

where m is a non-negative integer number. Next, we will derive an approximation for an element of a function of an m-band matrix.

Theorem 3.1. *Let $A \in \mathbb{C}^{n\times n}$ be an m-band matrix, and k be a non-negative integer. For an entry (i,j) of the matrix $f(A)$, with $x := \left\lfloor \frac{i+j}{2} \right\rfloor$, we have:*

$$\left|\left[f(A)\right]_{ij} - \left[f\left(\mathcal{B}_k\right)\right]_{\varphi_x(i),\varphi_x(j)}\right| \leq 2\mathcal{Q} \min_{p_k\in\Pi_k} \|f-p_k\|_{\mathcal{W}(A)},$$

where $\mathcal{B}_k := \mathcal{B}_{\lceil mk/2\rceil}\left([A]_{xx}\right)$ as defined in (1), φ_x is the function defined in (2), and $\mathcal{Q} = 1+\sqrt{2}$.

Proof. Assume A is an m-band matrix. We can write $[A^k]_{ij}$, for $k \geq 0$, as

$$\left[A^k\right]_{ij} = \sum_{|i-t_1|\leq m} \sum_{|t_1-t_2|\leq m} \cdots \sum_{|t_{k-1}-j|\leq m} [A]_{it_1}[A]_{t_1t_2}\cdots[A]_{t_{k-1}j}. \tag{5}$$

For $1 \leq p \leq k-1$, we have $|t_p - t_{p+1}| \leq m$. Thus, for any i and j, using the triangle inequality and applying inequalities between indices appropriately, we have $|i-t_p| \leq pm$, and $|j-t_p| \leq m(k-p)$. Now, using the inequalities obtained from the indices, we can deduce that all indices t_p satisfy the following inequality:

$$\left|\left\lfloor \frac{i+j}{2} \right\rfloor - t_p\right| \leq \left\lceil \frac{mk}{2} \right\rceil. \tag{6}$$

Now, consider the smallest submatrix that includes all the elements in A as in relation (5) based on the inequality (6). This submatrix is $\mathcal{B}_k := \mathcal{B}_{\lceil mk/2\rceil}([A]_{xx})$. Since $\mathcal{B}_k$ is a principal submatrix, then it is an m-band matrix. Therefore, with respect to the relation between the entries of $\mathcal{B}_k$ and A, that is, $[A]_{ij} = [\mathcal{B}_k]_{\varphi_x(i),\varphi_x(j)}$, for all $0 \leq s \leq k$, we have:

$$[A^s]_{ij} = [\mathcal{B}_k^s]_{\varphi_x(i),\varphi_x(j)}.$$

Hence, for a polynomial $p_k \in \Pi_k$, we have $[p_k(A)]_{ij} = [p_k(\mathcal{B}_k)]_{\varphi_x(i),\varphi_x(j)}$. For $[f(A)]_{ij}$, using (4) we get

$$\left|\left[f(A) - p_k(A)\right]_{ij}\right| \leq \mathcal{Q}\|f - p_k\|_{\mathcal{W}(A)}.$$

Similarly, we can apply the same for $\mathcal{B}_k$ and get

$$\left|\left[f(\mathcal{B}_k) - p_k(\mathcal{B}_k)\right]_{\varphi_x(i),\varphi_x(j)}\right| \leq \mathcal{Q}\|f - p_k\|_{\mathcal{W}(\mathcal{B}_k)}.$$

Since $\mathcal{B}_k$ is a principal submatrix of A, we have $\mathcal{W}(\mathcal{B}_k) \subseteq \mathcal{W}(A)$, which implies $\|f-p_k\|_{\mathcal{W}(\mathcal{B}_k)} \leq \|f-p_k\|_{\mathcal{W}(A)}$. Combining the two recent inequalities for A and $\mathcal{B}_k$, we arrive at

$$\left|\left[f(A)\right]_{ij} - \left[f(\mathcal{B}_k)\right]_{\varphi_x(i),\varphi_x(j)}\right| \leq 2\mathcal{Q}\|f - p_k\|_{\mathcal{W}(A)}.$$

And this completes the proof. □

We now have the next corollary for the elements of diagonal of $p_k(A)$.

Corollary 3.2. *Let $A \in \mathbb{C}^{n\times n}$ be an m-band matrix, and k be a non-negative integer. For $p_k \in \Pi_k$, we have:*

$$\left[p_k(A)\right]_{ii} = \left[p_k(\mathcal{B}_{k_i})\right]_{\varphi_i(i),\varphi_i(i)},$$

where $\mathcal{B}_{k_i} := \mathcal{B}_{\lceil mk/2\rceil}([A]_{ii})$ is as defined in (1), φ_i is the function as in (2), and $\mathcal{Q} = 1 + \sqrt{2}$.

Proof. The proof can easily be obtained from Theorem 3.1. □

Next, let us define $\mathcal{T}_f^{(k)}(A)$ as follows:

$$\mathcal{T}_f^{(k)}(A) := \sum_{i=1}^{n} \Big(f\big(\mathcal{B}_{\lceil mk/2\rceil}([A]_{ii})\big)\Big)_{\varphi_i(i),\varphi_i(i)}. \tag{7}$$

Then, we have the next result for approximating trace of $f(A)$.

Corollary 3.3. *Let $A \in \mathbb{C}^{n\times n}$ be an m-band matrix, and k be a non-negative integer. Then, we have:*

$$\left|\mathrm{tr}\big(f(A)\big) - \mathcal{T}_f^{(k)}(A)\right| \leq 2n\mathcal{Q} \min_{p_k\in\Pi_k} \|f - p_k\|_{\mathcal{W}(A)},$$

where $\mathcal{Q} = 1 + \sqrt{2}$ and $\mathcal{T}_f^{(k)}(A)$ as defined in (7).

Proof. The proof can easily be obtained from Theorem 3.1. □

Theorem 3.1 was suitable for approximating an entry of the matrix $f(A)$. However, if we are interested in approximating $f(A)$ entirely, then use of Theorem 3.1 incurs a high computational cost. Next, we derive an approximation for principal submatrices of the matrix $f(A)$ when A is an m-band matrix.

Theorem 3.4. *Assume $A \in \mathbb{C}^{n\times n}$ is an m-band matrix and k is a non-negative integer. Then we have:*

$$\left\|\mathcal{B}_\eta([f(A)]_{ii}) - \mathcal{B}_\eta\left(\left[f\left(\mathcal{B}_{2\eta}\right)\right]_{\eta\eta}\right)\right\|_2 \leq 2\mathcal{Q} \min_{p\in\Pi_k} \|f-p\|_{\mathcal{W}(A)},$$

where $\mathcal{B}_{2\eta} := \mathcal{B}_{2\eta}\left([A]_{ii}\right)$, $\eta := \left\lceil \frac{mk}{2} \right\rceil$, and $\mathcal{Q} = 1+\sqrt{2}$.

Proof. Let (i', j') be indices so that $|i-i'| \leq \eta$ and $|i-j'| \leq \eta$. Then, for $x = \lfloor \frac{i'+j'}{2} \rfloor$, we have $|i-x| \leq \eta$, implying that $\mathcal{B}_\eta([A]_{xx})$ is the principal submatrix of $\mathcal{B}_{2\eta}([A]_{ii})$. According to Theorem 3.1, we have:

$$\mathcal{B}_\eta([p_k(A)]_{ii}) = \mathcal{B}_\eta\left(\left[p_k\left(\mathcal{B}_{2\eta}\right)\right]_{\eta\eta}\right).$$

Now, using inequality (4) for the matrix A, we get

$$\left\|\mathcal{B}_\eta\left([f(A) - p_k(A)]_{ii}\right)\right\|_2 \leq \mathcal{Q}\|f-p_k\|_{\mathcal{W}(A)},$$

and, since $\mathcal{B}_{2\eta}$ is the principal submatrix of A, we have:

$$\left\|\mathcal{B}_\eta\left([f(\mathcal{B}_{2\eta}) - p_k(\mathcal{B}_{2\eta})]_{\eta\eta}\right)\right\|_2 \leq \mathcal{Q}\|f-p_k\|_{\mathcal{W}(A)}.$$

Now, by summing up the two previous inequalities, the result is at hand. □

Theorem 3.4 provides us with the possibility that by a slight increase in computational cost, we can approximate a larger number of entries of $f(A)$. Now, considering the fact that entries far from the diagonal become exponentially small [2], we can safely ignore them. This shows that the band matrix function is approximately a banded matrix.

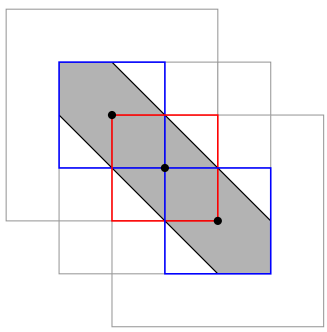

Figure 1: A part of the band of a band matrix.

Let us assume that, by removing small entries, the function of an m-band matrix is transformed into an m'-band matrix. On the main diagonal of this m'-band matrix, we can select three points, each at a distance of $2m'$ apart, as shown in Figure 1. Next, we draw submatrices around these points that cover the entries of this band matrix (Theorem 3.4). In the following section, we will present a method for computing $f(A)$, where A is an m-band matrix.

4 Algorithm 1: Divide and Conquer Method for Band Matrix Function

Here, we present a method for computing the function of an m-band matrix. For that, we first need to define several submatrices of the m-band matrix $A \in \mathbb{C}^{n\times n}$. Assume that a number $s \ll n$ has been chosen such that $n = ks$, and also $b := \frac{s}{2m}$, where m is the bandwidth of the matrix A. Now, we define the following matrices:

- Matrix $D := \mathrm{blkdiag}(D_1, \dots, D_k)$,where $D_i := \mathcal{B}_{s/2}\big([A]_{\eta_i\eta_i}\big)$ and $\eta_i := \frac{(2i-1)s}{2}$.
- Matrix $\tilde{B} := \mathrm{blkdiag}(B_1, \dots, B_{k-1})$, where $B_i := \mathcal{B}_{s/2}\big([A]_{\beta_i\beta_i}\big)$ and $\beta_i := is$ by starting from index $\frac{s}{2} + 1$.
- Matrix $\tilde{C} := \mathrm{blkdiag}(C_1^{(1)}, C_1^{(2)}, \dots, C_{k-1}^{(1)}, C_{k-1}^{(2)})$, where the C_i are diagonal block submatrices of size $\frac{s}{2} \times \frac{s}{2}$ from A starting from the entry $\frac{s}{2} + 1$.

Now, we define $B := \mathrm{blkdiag}(Z, \tilde{B}, Z)$ and $C := \mathrm{blkdiag}(Z, \tilde{C}, Z)$, where $Z := \mathtt{zeros}(\frac{s}{2})$. Using the decomposition introduced earlier, as $A = D + B - C$, we can now describe Algorithm 1, as presented in [4].

Algorithm 1 Approximation Algorithm for a Banded Matrix Function.

Input: Banded matrix $A \in \mathbb{C}^{n\times n}$ with bandwidth m, value s for block size, and function f.
Output: Approximation $\mathrm{approx}_f^{(s)}(A)$ for $f(A)$.

1: Compute matrices $f(D)$, $f(\tilde{B})$, and $f(\tilde{C})$ by applying f to only the blocks D_i, B_i, and C_i.
2: Define $f_B := \mathrm{blkdiag}(Z, f(\tilde{B}), Z)$ and $f_C := \mathrm{blkdiag}(Z, f(\tilde{C}), Z)$, where $Z := \mathtt{zeros}(\frac{s}{2})$.
3: $\mathrm{approx}_f^{(s)}(A) \leftarrow f(D) + f_B - f_C$.

In [4], the following result was established for estimating the error of Algorithm 1.

Theorem 4.1 ([4])**.** *Let A be an m-band matrix. The output* $\mathrm{approx}_f(A)$ *of Algorithm 1 satisfies:*

$$\|f(A) - \mathrm{approx}_f(A)\|_2 \le 4(1+\sqrt{2}) \min_{p\in\Pi_b} \|f - p\|_{\mathcal{W}(A)}.$$

Theorem 4.1 ensures that Algorithm 1 converges to the solution and provides an estimate for the error. This estimate is particularly accurate when f can be approximated with a polynomial of degree b in the numerical range of A. Also, for Algorithm 1, the computational complexity is $\mathcal{O}(nm^2)$ for an m-band matrix in $\mathbb{C}^{n\times n}$.

5 Algorithm 2: Divide and Conquer Method for Band-Toeplitz Matrix Function

Here, we are to find an algorithm for functions of band-Toeplitz matrices. We represent Toeplitz matrices using the complex function a in L^1 on the unit circle as follows [3]:

$$a(x) = \sum_{k=-\infty}^{\infty} a_k x^k; \quad x = e^{i\theta}, \ \theta \in [0, 2\pi]. \tag{8}$$

The coefficients are given by

$$a_k = \frac{1}{2\pi} \int_0^{2\pi} a(e^{ix}) e^{-ikx} \, \mathrm{d}x. \tag{9}$$

We define a Toeplitz matrix, denoted as $T_n(a)$, as follows:

$$T_n(a) = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \dots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \dots & a_{-(n-2)} \\ a_2 & a_1 & a_0 & \dots & a_{-(n-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_0 \end{pmatrix}.$$

In other words, $[T_n(a)]_{ij} = a_{i-j}$, where a_k is defined as in (9). Now, consider a non-negative integer, denoted by m, such that for (8), we have:

$$a(x) = \sum_{k=-\infty}^{\infty} a_k x^k = \sum_{k=-m}^{m} a_k x^k.$$

It follows that the operator $T_n(a)$ represents a Toeplitz matrix and is an m-band matrix. Assume that a number $s \ll n$ has been chosen such that $n = ks$, and $b := \frac{s}{2m}$.

Algorithm 2 Approximation Algorithm for a Band-Toeplitz Matrix Function.

Input: Band-Toeplitz matrix $T_n(a) \in \mathbb{C}^{n\times n}$ with bandwidth m, value s for block size, and function f.
Output: Approximation $\mathrm{Tapprox}_f^{(s)}(T_n(a))$ for $f(T_n(a))$.

1: Compute $f_1 := f(T_s(a))$, and $f_{\frac{1}{2}} := f(T_{\frac{s}{2}}(a))$.
2: Set

$$f_T^{(1)} := \begin{pmatrix} f_1 & & & \\ & f_1 & & \\ & & \ddots & \\ & & & f_1 \end{pmatrix}, \quad f_T^{(2)} := \begin{pmatrix} Z & & & & \\ & f_1 & & & \\ & & \ddots & & \\ & & & f_1 & \\ & & & & Z \end{pmatrix}, \quad f_T^{(3)} := \begin{pmatrix} Z & & & & \\ & f_{\frac{1}{2}} & & & \\ & & \ddots & & \\ & & & f_{\frac{1}{2}} & \\ & & & & Z \end{pmatrix},$$

where $Z := \mathsf{zeros}(\frac{s}{2})$.
3: $\mathrm{Tapprox}_f^{(s)}(A) \leftarrow f_T^{(1)} + f_T^{(2)} - f_T^{(3)}$.

Theorem 4.1 ensures the convergence of Algorithm 2 to $f(T_n(a))$. Assuming that the computational complexities for computing $f(T_s(a))$ and $f(T_{\frac{s}{2}}(a))$ are C_s and $C_{\frac{s}{2}}$, respectively, the computational complexity of Algorithm 2 can be expressed as $\mathcal{O}\big(C_s + C_{\frac{s}{2}}\big)$. Since $s \ll n$ is independent of n, the complexity of approximation is of $\mathcal{O}(1)$.

5.1 Divide and Conquer Method for Persymmetric Band-Hankel Matrix Function

Let us define $J_n \in \mathbb{R}^{n\times n}$ as a backward identity matrix. This means that for elements with $i + j = n + 1$, we have $[J_n]_{ij} = 1$, while for all other elements, $[J_n]_{ij} = 0$. For example, in the case of $n = 3$, we have:

$$J_3 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

We call H_n a persymmetric band-Hankel matrix if there exists a symmetric band-Toeplitz matrix $T_n(a)$ such that [3]

$$H_n = J_n T_n(a).$$

In fact, we also have $H_n = J_n T_n(a) = T_n(a) J_n$. Now, let f be an analytic function as

$$f(z) = \sum_{k=0}^{\infty} \alpha_k z^k.$$

To compute $f(H_n)$, we can write

$$f(H_n) = J_n \frac{f(T_n(a)) - f(-T_n(a))}{2} + \frac{f(T_n(a)) + f(-T_n(a))}{2}. \tag{10}$$

Now, by computing $f(T_n(a))$ and $f(-T_n(a))$, we can compute $f(H_n)$. For that, we simply utilize Algorithm 2 for both $f(T_n(a))$ and $f(-T_n(a))$, and subsequently apply the equation (10).

6 Conclusion

We have proposed new methods for computing functions of band matrices, band-Toeplitz matrices, and band-Hankel matrices. We have also presented approximations for the entries and the trace of functions of band matrices.

References

[1] C. Bekas, E. Kokiopoulou, and Y. Saad, An estimator for the diagonal of a matrix, Appl. Numer. Math., 57 (2007), 1214–1229.

[2] M. Benzi and N. Razouk, Decay bounds and $\mathcal{O}(n)$ algorithms for approximating functions of sparse matrices, Electron. Trans. Numer. Anal., 28 (2007), 16–39.

[3] A. Böttcher, S.M. Grudsky, Spectral Properties of Banded Toeplitz Matrices, SIAM, Philadelphia, 2005.

[4] A. Cortinovis, D. Kressner, and S. Massei. Divide-and-conquer methods for functions of matrices with banded or hierarchical low-rank structure. SIAM J. Matrix Anal. Appl., 43 (2022), 151–177.

[5] M. Crouzeix and C. Palencia, The numerical range is a $(1+\sqrt{2})$-spectral set, SIAM J. Matrix Anal. Appl., 38 (2017), 649–655.

[6] V. Druskin, S. Güttel, and L. Knizhnerman, Near-optimal perfectly matched layers forindefinite Helmholtz problems, SIAM Rev., 58 (2016), 90–116.

[7] E. Estrada and D. J. Higham, Network properties revealed through matrix functions, SIAM Rev., 52 (2010), 696–714.

[8] N. J. Higham, Functions of Matrices: Theory and Computation, SIAM, Philadelphia, PA, 2008.