

Article

On the Lanczos Method for Computing Some Matrix Functions

Ying Gu ^{1,*}, Hari Mohan Srivastava ^{2,3,4,5,6}  and Xiaolan Liu ¹

¹ School of Sciences and Arts, Suqian University, Suqian 223800, China; 23046@squ.edu.cn

² Department of Mathematics and Statistics, University of Victoria, Victoria, BC V8W 3R4, Canada; harimsri@math.uvic.ca

³ Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan

⁴ Department of Mathematics and Informatics, Azerbaijan University, AZ197 Baku, Azerbaijan

⁵ Section of Mathematics, International Telematic University Uninettuno, I-00186 Rome, Italy

⁶ Center for Converging Humanities, Kyung Hee University, 26 Kyungheedaero, Dongdaemun-gu, Seoul 02447, Republic of Korea

* Correspondence: 23095@squ.edu.cn

Abstract: The study of matrix functions is highly significant and has important applications in control theory, quantum mechanics, signal processing, and machine learning. Previous work has mainly focused on how to use the Krylov-type method to efficiently calculate matrix functions $f(A)\beta$ and $\beta^T f(A)\beta$ when A is symmetric. In this paper, we mainly illustrate the convergence using the polynomial approximation theory for the case where A is symmetric positive definite. Numerical results illustrate the effectiveness of our theoretical results.

Keywords: matrix logarithm function; Chebyshev polynomial; Legendre polynomial; Lanczos method; Krylov subspace

MSC: 65K10; 65K05; 90C20



Citation: Gu, Y.; Srivastava, H.M.; Liu, X. On the Lanczos Method for Computing Some Matrix Functions. *Axioms* **2024**, *13*, 764. <https://doi.org/10.3390/axioms13110764>

Academic Editor: Vladimir Rakocevic

Received: 7 October 2024

Revised: 30 October 2024

Accepted: 1 November 2024

Published: 4 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The study of matrix functions is highly significant and has important applications in control theory, quantum mechanics, signal processing, and machine learning. Some studies on discrete matrix functions have been investigated in [1,2]. Let $A \in \mathbb{R}^{n \times n}$ be a large sparse matrix and $\beta \in \mathbb{R}^n$ be a nonzero real vector. Many research works have focused on computing $f(A)\beta$ and $\beta^T f(A)\beta$ [3–5] using the Krylov-type method, where f is a function that makes $f(A) \in \mathbb{R}^{n \times n}$ well defined and $\beta^T f(A)\beta$ is called a *bilinear form* when A is symmetric [6]. Theoretical results and numerical experiments have shown that $f(A)\beta$ and $\beta^T f(A)\beta$ can be calculated efficiently using the Krylov-type method [7–10].

In this paper, we consider the special case that matrix A is symmetric positive definite and $f(\cdot) = \log(\cdot)$, $f(\cdot) = (\cdot)^{-\frac{1}{2}}$ or $f(\cdot) = (\cdot)^{-1}$. The logarithm of A is any matrix X that holds $\exp(X) = A$ (refer to ([11], Chapter 11), [12,13], or [14]). Let the eigen decomposition of A be

$$A = U\Lambda U^T,$$

with

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

where $U \in \mathbb{R}^{n \times n}$, $U^T U = I$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$. If an eigenvalue of A is less than 0, then $\log(A)$ is not well defined ([11], Chapter 11). Thus, if A is symmetric, it

must be positive definite. The matrix logarithm function is defined for A as follows ([11], Definition 1.2):

$$\log(A) := U \cdot \begin{pmatrix} \log(\lambda_1) & & & \\ & \log(\lambda_2) & & \\ & & \ddots & \\ & & & \log(\lambda_n) \end{pmatrix} \cdot U^T \in \mathbb{R}^{n \times n}.$$

Similarly, the matrix square root is defined for A as follows ([11], Definition 1.2):

$$A^{\frac{1}{2}} := U \cdot \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{pmatrix} \cdot U^T \in \mathbb{R}^{n \times n}.$$

The estimation of the log-determinant of matrix A is widely used in, for example, Markov random field models [15], lattice quantum chromodynamics [16], statistical learning [17], and so on. Moreover, the calculation of the bilinear form $\beta^T \log(A) \beta$ is crucial for the estimation of the log-determinant of matrix A . The computation of $A^{-\frac{1}{2}} \beta$ arises in the context of Markov function problems and domain decomposition methods [18]. The computation of $\text{trace}(A^{-1})$ has received much attention [6,19]. The utilization of modified moments [6,20], Monte Carlo methods [21], and Gaussian quadratures [6,22] to estimate $\text{trace}(A^{-1})$ for symmetric positive definite matrices has been proposed. These estimates have a wide range of applications in mathematics [6], physics [23], and statistics [24]. Moreover, the computation of $\beta^T A^{-1} \beta$ is important for estimating the trace of the matrix inverse [25,26]. The Lanczos method is a method commonly used to compute $\log(A) \beta$, $\beta^T \log(A) \beta$, $A^{-\frac{1}{2}} \beta$ and $\beta^T A^{-1} \beta$.

Using the polynomial approximation theory, we prove, in this paper, that the Lanczos method converges at rates

$$\mathcal{O}\left(k^{-1} \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right), \mathcal{O}\left(k^{-1} \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k-1}\right), \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right) \text{ and } \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k-1}\right)$$

for computing $\log(A) \beta$, $\beta^T \log(A) \beta$, $A^{-\frac{1}{2}} \beta$, and $\beta^T A^{-1} \beta$, respectively. Here, k denotes the iterations and $\kappa = \|A\| \|A^{-1}\|$, with $\|\cdot\|$ denoting the Euclidean norm of a matrix or a vector in this paper.

Throughout our paper, we let A be symmetric positive definite and organized the paper as follows. We present the Lanczos approximation to $\log(A) \beta$, $\beta^T \log(A) \beta$, $A^{-\frac{1}{2}} \beta$, and $\beta^T A^{-1} \beta$ in Section 2. In Section 3, a convergence analysis of the Lanczos method for computing $\log(A) \beta$, $\beta^T \log(A) \beta$, $A^{-\frac{1}{2}} \beta$, and $\beta^T A^{-1} \beta$ is provided. Numerical results are presented in Section 4 to illustrate the effectiveness of our bounds, and in Section 5, we present a conclusion.

2. Lanczos Approximation of $\log(A) \beta$, $\beta^T \log(A) \beta$, $A^{-\frac{1}{2}} \beta$, and $\beta^T A^{-1} \beta$

For the Krylov subspace

$$\mathcal{K}_k(A, \beta) = \text{span}\{\beta, A\beta, \dots, A^{k-1}\beta\},$$

we can usually obtain an orthonormal basis Q_k of it using the Lanczos process [6]. Moreover, it holds that

$$AQ_k = Q_k S_k + \theta_k q_{k+1} e_1^T, \quad (1)$$

where

$$Q_k = [q_1, q_2, \dots, q_k] \in \mathbb{R}^{n \times k}, q_1 = \frac{\beta}{\|\beta\|},$$

$$[Q_k \ q_{k+1}]^T [Q_k \ q_{k+1}] = I_{k+1}, e_1 = (1 \ 0 \dots 0)^T$$

and

$$S_k = Q_k^T A Q_k = \begin{pmatrix} \omega_1 & \theta_1 & & & \\ \theta_1 & \omega_2 & \theta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \theta_{k-2} & \omega_{k-1} & \theta_{k-1} \\ & & & \theta_{k-1} & \omega_k \end{pmatrix} \in \mathbb{R}^{k \times k}.$$

Thus, $\|\beta\| Q_k \log(S_k) e_1$, $\|\beta\|^2 e_1^T \log(S_k) e_1$, $\|\beta\| Q_k S_k^{-\frac{1}{2}} e_1$, and $\|\beta\|^2 e_1^T S_k^{-1} e_1$ are taken as approximations of $\log(A)\beta$, $\beta^T \log(A)\beta$, $A^{-\frac{1}{2}}\beta$, and $\beta^T A^{-1}\beta$, respectively ([11], Section 13.2).

If the Lanczos process (1) terminates at iteration $k_{\max}$ for the first time, then $\theta_{k_{\max}} = 0$ and

$$AQ_{k_{\max}} = Q_{k_{\max}} S_{k_{\max}}. \quad (2)$$

Let the eigen decomposition of $S_{k_{\max}}$ be

$$S_{k_{\max}} = W \cdot \text{diag}(d_1, d_2, \dots, d_{k_{\max}}) \cdot W^T.$$

Hence, $\{d_1, d_2, \dots, d_{k_{\max}}\} \subseteq \{\lambda_1, \lambda_2, \dots, \lambda_n\}$. By ([11], Definition 1.4), we have that there exist three polynomials q , p , and h such that $A^{-\frac{1}{2}} = q(A)$, $\log(A) = p(A)$ and $h(A) = A^{-1}$. Hence, $q(\lambda_i) = \sqrt{\lambda_i}$, $p(\lambda_i) = \log(\lambda_i)$, and $h(\lambda_i) = \lambda_i^{-1}$, $i = 1, 2, \dots, n$. Thus, it can be seen that $q(d_i) = \sqrt{d_i}$, $p(d_i) = \log(d_i)$, $h(d_i) = d_i^{-1}$, $i = 1, 2, \dots, k_{\max}$,

$$\begin{aligned} \log(A)\beta &= p(A)\beta \\ &= \|\beta\| \cdot p(A) Q_{k_{\max}} e_1 \\ &\stackrel{(2)}{=} \|\beta\| \cdot Q_{k_{\max}} p(S_{k_{\max}}) e_1 \\ &= \|\beta\| \cdot Q_{k_{\max}} W \cdot \text{diag}(p(d_1), p(d_2), \dots, p(d_{k_{\max}})) \cdot W^T e_1 \\ &= \|\beta\| \cdot Q_{k_{\max}} W \cdot \text{diag}(\log(d_1), \log(d_2), \dots, \log(d_{k_{\max}})) \cdot W^T e_1 \\ &= \|\beta\| \cdot Q_{k_{\max}} \log(S_{k_{\max}}) e_1, \end{aligned}$$

and

$$\beta^T \log(A)\beta = \|\beta\| \cdot \beta^T Q_{k_{\max}} \log(S_{k_{\max}}) e_1 = \|\beta\|^2 \cdot e_1^T \log(S_{k_{\max}}) e_1. \quad (3)$$

Similarly, it holds that

$$A^{-\frac{1}{2}}\beta = \|\beta\| \cdot S_{k_{\max}}^{-\frac{1}{2}} e_1 \quad \text{and} \quad \beta^T A^{-1}\beta = \|\beta\|^2 \cdot e_1^T S_{k_{\max}}^{-1} e_1.$$

Thus, when the Lanczos process terminates, $\log(A)\beta$, $\beta^T \log(A)\beta$, $A^{-\frac{1}{2}}\beta$, and $\beta^T A^{-1}\beta$ can be found exactly.

3. Main Achievements

The convergence of the Lanczos approximations to $\|\beta\| Q_k \log(S_k) e_1$, $\|\beta\|^2 e_1^T \log(S_k) e_1$, $\|\beta\| Q_k S_k^{-\frac{1}{2}} e_1$, and $\|\beta\|^2 e_1^T S_k^{-1} e_1$ is mainly considered in this section. We first give the following three lemmas.

Lemma 1 ([27], eq.(5.6.2)), and ([28], p. 449)). For any $|z| \leq 1$ and $x \in [-1, 1]$, we have

$$\log(1 - 2xz + z^2) = -2 \sum_{n=1}^{\infty} \frac{T_n(x)z^n}{n},$$

and

$$\frac{1}{\sqrt{1 - 2xz + z^2}} = \sum_{n=1}^{\infty} P_n(x)z^n, \quad (4)$$

where $T_n(x)$ represents the first kind of n -th degree Chebyshev polynomial

$$T_n(x) := \cos(n \arccos x), \quad x \in [-1, 1], \quad n \in \mathbb{N}_+,$$

and $P_n(x)$ represents the n -th degree Legendre polynomial

$$P_n(x) = \frac{(-1)^n}{2^n n!} \cdot \frac{d^n}{dx^n} (1 - x^2)^n, \quad x \in [-1, 1], \quad n \in \mathbb{N}_+.$$

Moreover,

$$\max_{|x| \leq 1} |P_n(x)| = 1. \quad (5)$$

Lemma 2 ([8], Lemma 2.2). If f is analytic on $[\lambda_n, \lambda_1]$, then

$$\frac{\|f(A)\beta - \|\beta\| Q_j f(S_j) e_1\|}{\|\beta\|} \leq 2 \cdot \min_{p_j \in \mathcal{P}_{j-1}} \max_{\lambda \in [\lambda_n, \lambda_1]} |f(\lambda) - p(\lambda)|, \quad (6)$$

$$\frac{|\beta^T f(A)\beta - \|\beta\|^2 e_1^T f(S_j) e_1|}{\|\beta\|^2} \leq 2 \cdot \min_{p_j \in \mathcal{P}_{2j-1}} \max_{\lambda \in [\lambda_n, \lambda_1]} |f(\lambda) - p(\lambda)|, \quad (7)$$

where $j \in \mathbb{N}_+$, and $\mathcal{P}_s$ denotes the polynomials set with degrees no higher than $s \in \mathbb{N}_+$.

Lemma 3 ([29]). For any $\omega > 1$, we have

$$\min_{\phi \in \mathcal{P}_j} \max_{x \in [-1, 1]} \left| \frac{1}{x - \omega} - \phi(x) \right| \leq \frac{(\omega + \sqrt{\omega^2 - 1})^{-j}}{\omega^2 - 1}.$$

The main achievements of this article are presented below.

Theorem 1. Let $\varkappa = \|A\| \|A^{-1}\|$. For $k \geq 1$,

$$\epsilon_1^{(k)} := \frac{\|\log(A)\beta - \|\beta\| \cdot Q_k \log(S_k) e_1\|}{\|\log(A)\beta\|} \leq \frac{2\|\beta\|}{\|\log(A)\beta\|} \cdot \frac{\sqrt{\varkappa} + 1}{k} \left(\frac{\sqrt{\varkappa} - 1}{\sqrt{\varkappa} + 1} \right)^k, \quad (8)$$

and

$$0 \leq \epsilon_2^{(k)} := \frac{\|\beta\|^2 \cdot e_1^T \log(S_k) e_1 - \beta^T \log(A) \beta}{|\beta^T \log(A) \beta|} \leq \frac{\|\beta\|^2}{|\beta^T \log(A) \beta|} \cdot \frac{\sqrt{\varkappa} + 1}{k} \left(\frac{\sqrt{\varkappa} - 1}{\sqrt{\varkappa} + 1} \right)^{2k}. \quad (9)$$

Proof. Let

$$\eta := \frac{\lambda_1 + \lambda_n}{\lambda_1 - \lambda_n} = \frac{\varkappa + 1}{\varkappa - 1}. \quad (10)$$

Then,

$$z_0 := \eta - \sqrt{\eta^2 - 1} = \frac{\sqrt{\varepsilon} - 1}{\sqrt{\varepsilon} + 1} \in (0, 1), \text{ and } 1 + z_0^2 = 2z_0\eta. \quad (11)$$

According to Lemma 1, we have

$$\log(2z_0(\eta - x)) = \log(1 - 2xz_0 + z_0^2) = -2 \sum_{n=1}^{\infty} \frac{T_n(x)z_0^n}{n}. \quad (12)$$

Consider the linear transformation

$$\lambda = \frac{\lambda_n - \lambda_1}{2}x + \frac{\lambda_1 + \lambda_n}{2}, \quad x \in [-1, 1],$$

which maps $x \in [-1, 1]$ bijectively onto $\lambda \in [\lambda_n, \lambda_1]$. Then,

$$\begin{aligned} \frac{\|\log(A)\beta\|_{\epsilon_1}^{(k)}}{\|\beta\|} &\stackrel{(6)}{\leq} 2 \min_{p \in \mathcal{P}_{k-1}} \max_{\lambda \in [\lambda_n, \lambda_1]} |\log \lambda - p(\lambda)| \\ &= 2 \min_{p \in \mathcal{P}_{k-1}} \max_{-1 \leq x \leq 1} \left| \log \left(\frac{\lambda_n - \lambda_1}{2}x + \frac{\lambda_1 + \lambda_n}{2} \right) - p \left(\frac{\lambda_n - \lambda_1}{2}x + \frac{\lambda_1 + \lambda_n}{2} \right) \right| \\ &= 2 \min_{q \in \mathcal{P}_{k-1}} \max_{-1 \leq x \leq 1} \left| \log \left[\frac{\lambda_n - \lambda_1}{2}(x - \eta) \right] - q(x) \right| \\ &= 2 \min_{q \in \mathcal{P}_{k-1}} \max_{-1 \leq x \leq 1} \left| \log \left[\frac{\lambda_1 - \lambda_n}{4z_0}(2z_0(\eta - x)) \right] - q(x) \right| \\ &= 2 \min_{q \in \mathcal{P}_{k-1}} \max_{-1 \leq x \leq 1} \left| \log(2z_0(\eta - x)) - q(x) + \log \left(\frac{\lambda_1 - \lambda_n}{4z_0} \right) \right|. \end{aligned}$$

Let

$$q(x) - \log \left(\frac{\lambda_1 - \lambda_n}{4z_0} \right) = -2 \sum_{n=1}^{k-1} \frac{T_n(x)z_0^n}{n} \in \mathcal{P}_{k-1}.$$

By (12) and $|T_n(x)| \leq 1$ for $x \in [-1, 1]$, we obtain that

$$\begin{aligned} \frac{\|\log(A)\beta\|_{\epsilon_1}^{(k)}}{\|\beta\|} &\leq 2 \max_{x \in [-1, 1]} \left| \log(2z_0(\eta - x)) + 2 \sum_{n=1}^{k-1} \frac{T_n(x)z_0^n}{n} \right| \\ &\leq 4 \sum_{n=k}^{\infty} \frac{z_0^n}{n} = 4z_0^k \sum_{n=0}^{\infty} \frac{z_0^n}{n+k} \\ &\leq \frac{4z_0^k}{k} \sum_{n=0}^{\infty} z_0^n = \frac{4z_0^k}{k(1-z_0)}, \end{aligned}$$

which yields (8).

Furthermore, a linear map $\Psi : \mathbb{R}^{n_1 \times n_1} \rightarrow \mathbb{R}^{n_2 \times n_2}$ is called a *unital positive linear map* ([30], p. 5) if the following hold:

- (i) A is symmetric positive definite $\Rightarrow \Psi(A)$ is also symmetric positive definite;
- (ii) $\Psi(I_{n_1}) = I_{n_2}$.

Define

$$\Phi(\Delta) = \Delta(1 : k, 1 : k) \in \mathbb{R}^{k \times k}, \text{ where } \Delta \in \mathbb{R}^{k_{\max} \times k_{\max}}.$$

Thus, Φ is a unital positive linear map. By ([30], Corollary 1.8),

$\log(S_k) - \Phi(\log(S_{k_{\max}})) = \log(\Phi(S_{k_{\max}})) - \Phi(\log(S_{k_{\max}}))$ is positive semidefinite.

It follows that

$$\begin{aligned} 0 &\leq e_1^T \log(S_k) e_1 - e_1^T \Phi(\log(S_{k_{\max}})) e_1 \\ &= e_1^T \log(S_k) e_1 - e_1^T \log(S_{k_{\max}}) e_1 \\ &\stackrel{(3)}{=} e_1^T \log(S_k) e_1 - \|\beta\|^{-2} \cdot \beta^T \log(A) \beta, \end{aligned}$$

which yields the first inequality in (9). In light of (7), we see that

$$\begin{aligned} &\frac{|\beta^T \log(A) \beta| \epsilon_2^{(k)}}{\|\beta\|^2} \\ &\leq 2 \min_{h \in \mathcal{P}_{2k-1}} \max_{\lambda \in [\lambda_n, \lambda_1]} |\log(\lambda) - h(\lambda)| \\ &= 2 \min_{h \in \mathcal{P}_{2k-1}} \max_{-1 \leq x \leq 1} \left| \log\left(\frac{\lambda_n - \lambda_1}{2} x + \frac{\lambda_1 + \lambda_n}{2}\right) - h\left(\frac{\lambda_n - \lambda_1}{2} x + \frac{\lambda_1 + \lambda_n}{2}\right) \right| \\ &= 2 \min_{h \in \mathcal{P}_{2k-1}} \max_{-1 \leq x \leq 1} \left| \log(2z_0(\eta - x)) - h\left(\frac{\lambda_n - \lambda_1}{2} x + \frac{\lambda_1 + \lambda_n}{2}\right) + \log\left(\frac{\lambda_1 - \lambda_n}{4z_0}\right) \right|. \end{aligned}$$

Analogously, consider

$$h\left(\frac{\lambda_n - \lambda_1}{2} x + \frac{\lambda_1 + \lambda_n}{2}\right) - \log\left(\frac{\lambda_1 - \lambda_n}{4z_0}\right) = -2 \sum_{n=1}^{2k-1} \frac{T_n(x) z_0^n}{n} \in \mathcal{P}_{2k-1}.$$

Recall that $|T_n(x)| \leq 1$ for $x \in [-1, 1]$. We then have

$$\begin{aligned} 0 &\leq \frac{|\beta^T \log(A) \beta| \epsilon_2^{(k)}}{\|\beta\|^2} \\ &\leq 2 \max_{-1 \leq x \leq 1} \left| \log(2z_0(\eta - x)) + 2 \sum_{n=1}^{2k-1} \frac{T_n(x) z_0^n}{n} \right| \\ &\stackrel{(12)}{\leq} 4 \sum_{n=2k}^{\infty} \frac{z_0^n}{n} = 4z_0^{2k} \sum_{n=0}^{\infty} \frac{z_0^n}{n+2k} \\ &\leq \frac{2z_0^{2k}}{k} \sum_{n=0}^{\infty} z_0^n = \frac{2z_0^{2k}}{k(1-z_0)}, \end{aligned}$$

which yields (9). $\square$

Theorem 2. For $k \geq 1$,

$$\epsilon_3^{(k)} := \frac{\|A^{-\frac{1}{2}} \beta - \|\beta\| Q_k S_k^{-\frac{1}{2}} e_1\|}{\|A^{-\frac{1}{2}} \beta\|} \leq \frac{2\|\beta\|(\sqrt{\kappa}+1)}{\|A^{-\frac{1}{2}} \beta\| \sqrt{\lambda_1 - \lambda_n}} \left(\frac{\sqrt{\varkappa}-1}{\sqrt{\varkappa}+1} \right)^{k+\frac{1}{2}} \quad (13)$$

Proof. By (11) and Lemma 1,

$$\frac{1}{\sqrt{2z_0(\eta-x)}} = \frac{1}{\sqrt{1-2xz_0+z_0^2}} = \sum_{n=1}^{\infty} P_n(x) z_0^n.$$

It holds that

$$\begin{aligned}
& \frac{\left\| A^{-\frac{1}{2}} \beta - \|\beta\| Q_k S_k^{-\frac{1}{2}} e_1 \right\|}{\|\beta\|} \leq 2 \cdot \min_{p_j \in \mathcal{P}_{j-1}} \max_{\lambda \in [\lambda_n, \lambda_1]} \left| \frac{1}{\sqrt{\lambda}} - p(\lambda) \right| \\
& = 2 \min_{p \in \mathcal{P}_{k-1}} \max_{x \in [-1, 1]} \left| \left(\frac{\lambda_n - \lambda_1}{2} x + \frac{\lambda_1 + \lambda_n}{2} \right)^{-\frac{1}{2}} - p \left(\frac{\lambda_n - \lambda_1}{2} x + \frac{\lambda_1 + \lambda_n}{2} \right) \right| \\
& \stackrel{(10)}{=} 2 \min_{q \in \mathcal{P}_{k-1}} \max_{x \in [-1, 1]} \left| \left[\frac{\lambda_1 - \lambda_n}{4z_0} (2z_0(\eta - x)) \right]^{-\frac{1}{2}} - q(x) \right| \\
& = \frac{2}{\sqrt{\frac{\lambda_1 - \lambda_n}{4z_0}}} \min_{\tilde{q} \in \mathcal{P}_{k-1}} \max_{x \in [-1, 1]} \left| (2z_0(\eta - x))^{-\frac{1}{2}} - \tilde{q}(x) \right| \\
& \leq 4 \sqrt{\frac{z_0}{\lambda_1 - \lambda_n}} \cdot \max_{x \in [-1, 1]} \left| (2z_0(\eta - x))^{-\frac{1}{2}} - \sum_{n=1}^{k-1} P_n(x) z_0^n \right| \\
& \stackrel{(4)}{\leq} 4 \sqrt{\frac{z_0}{\lambda_1 - \lambda_n}} \cdot \sum_{n=k}^{\infty} \left(\max_{-1 \leq x \leq 1} |P_n(x)| \right) z_0^n \stackrel{(5)}{=} 4 \sqrt{\frac{z_0}{\lambda_1 - \lambda_n}} \cdot \frac{z_0^k}{1 - z_0} \\
& = \frac{2(\sqrt{\kappa} + 1) z_0^{k+\frac{1}{2}}}{\sqrt{\lambda_1 - \lambda_n}},
\end{aligned}$$

which yields (13). $\square$

Finally, we provide a convergence analysis of $\left| \beta^T A^{-1} \beta - \|\beta\|^2 e_1^T S_k^{-1} e_1 \right|$.

Theorem 3. For $k \geq 1$,

$$0 \leq \epsilon_4^{(k)} := \frac{\beta^T A^{-1} \beta - \|\beta\|^2 e_1^T S_k^{-1} e_1}{\beta^T A^{-1} \beta} \leq \frac{\|\beta\|^2}{\lambda_n \cdot \beta^T A^{-1} \beta} \left(\frac{\sqrt{\varkappa} - 1}{\sqrt{\varkappa} + 1} \right)^{2k-1}. \quad (14)$$

Proof. On the one hand, we deduce that

$$\begin{aligned}
& \beta^T A^{-1} \beta - \|\beta\|^2 e_1^T S_k^{-1} e_1 \\
& = \|\beta\|^2 \left(e_1^T Q_{k_{\max}}^T A^{-1} Q_{k_{\max}} e_1 - e_1^T S_k^{-1} e_1 \right) \\
& = \|\beta\|^2 \left(e_1^T S_{k_{\max}}^{-1} e_1 - e_1^T S_k^{-1} e_1 \right) \geq 0.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
& \frac{\beta^T A^{-1} \beta - \|\beta\|^2 e_1^T S_k^{-1} e_1}{\|\beta\|^2} \stackrel{(6)}{\leq} 2 \min_{p_k \in \mathcal{P}_{2k-1}} \max_{\lambda \in [\lambda_n, \lambda_1]} \left| \frac{1}{\lambda} - p(\lambda) \right| \\
& = 2 \min_{p \in \mathcal{P}_{2k-1}} \max_{x \in [-1, 1]} \left| \left(\frac{(\lambda_n - \lambda_1)x}{2} + \frac{\lambda_1 + \lambda_n}{2} \right)^{-1} - p \left(\frac{(\lambda_n - \lambda_1)x}{2} + \frac{\lambda_1 + \lambda_n}{2} \right) \right| \\
& \stackrel{(10)}{=} \frac{4}{\lambda_1 - \lambda_n} \cdot \min_{\tilde{p} \in \mathcal{P}_{2k-1}} \max_{x \in [-1, 1]} \left| \frac{1}{x - \eta} - \tilde{p}(x) \right| \\
& \stackrel{\text{Lemma 3}}{=} \frac{4}{\lambda_1 - \lambda_n} \cdot \frac{(\eta + \sqrt{\eta^2 - 1})^{-(2k-1)}}{\eta^2 - 1} \\
& = \frac{4}{\lambda_n(\varkappa - 1)} \cdot \frac{(\eta + \sqrt{\eta^2 - 1})^{-(2k-1)}}{\eta^2 - 1}.
\end{aligned}$$

Recall that $\eta = \frac{\kappa+1}{\kappa-1}$; it follows that

$$\eta + \sqrt{\eta^2 - 1} = \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1} \text{ and } \eta^2 - 1 = \frac{4\kappa}{(\kappa - 1)^2} > \frac{4}{\kappa - 1},$$

which yields (14). $\square$

Remark 1. It is well known that the rate of convergence in computing $A^{-1}\beta$ using the conjugate gradient method is

$$\varsigma_k := \frac{\|\beta - A\mathbf{y}_k\|}{\|\beta\|} \leq \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right),$$

where $\mathbf{y}_k$ is the approximate solution of $A^{-1}\beta$ derived from the k -th step of the conjugate gradient method for solving linear system $A\mathbf{y} = \beta$. It follows that $\epsilon_4^{(k)} = \mathcal{O}(\varsigma_k^2)$. Thus, in practice, we can estimate the size of ϵ_4 from the size of ς_k .

Remark 2. According to Theorems 1–3, the convergence rates of $\epsilon_1^{(k)}$, $\epsilon_2^{(k)}$, $\epsilon_3^{(k)}$, and $\epsilon_4^{(k)}$ are

$$\mathcal{O}\left(k^{-1}\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right), \mathcal{O}\left(k^{-1}\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k-1}\right), \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right) \text{ and } \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k-1}\right),$$

respectively. The smaller the condition number κ of A , the faster these four quantities converge to zero. It should be noted that the convergence rate of $\epsilon_2^{(k)}$ is significantly higher than $\epsilon_1^{(k)}$ and $\epsilon_3^{(k)}$ and slightly higher than that of $\epsilon_4^{(k)}$. The numerical experiments in Section 4 illustrate these two facts.

If f is analytic on $[\lambda_n, \lambda_1]$, then

$$\left|\|\beta\|^2 \cdot \mathbf{e}_1^T f(S_k) \mathbf{e}_1 - \beta^T f(A) \beta\right| \leq \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k}\right).$$

In contrast, for $f(\cdot) = \log(\cdot)$, the convergence rate provided by (9) is superior.

4. Numerical Experiments

We want to show the effectiveness of the bounds provided by (8), (9), (13), and (14) with two examples.

Example 1. Let the matrix $B = \text{diag}(t_{in}^{[a,b]})$ [31], where

$$t_{in}^{[a,b]} = \left(\frac{b-a}{2}\right) \left(t_{in} + \left(\frac{a+b}{b-a}\right)\right) \text{ with } t_{in} = \cos \frac{(2i-1)\pi}{2n}, i = 1, 2, \dots, n.$$

Here, $n = 5000$, $a = -10$, $b = 10$, and $t_{in}^{[a,b]}$ represents the n -th translated Chebyshev zero node on $[a, b]$. Let

$$A = B + \zeta I, \quad \beta = \frac{\mathbf{v}}{\|\mathbf{v}\|},$$

with

$$\mathbf{v} = (1 \ 1 \ \dots \ 1)^T \in \mathbb{R}^{5000}, \text{ and } \zeta > 10.$$

The selection of ζ is listed in Table 1.

Table 1. Example 1: Different values of κ with different choices of ζ .

	$\zeta = 10 + 10^{-1}$	$\zeta = 10 + 10^{-2}$	$\zeta = 10 + 10^{-3}$	$\zeta = 10 + 10^{-4}$
κ	2.01×10^2	2.00×10^3	2.00×10^4	1.99×10^5

Example 2. Consider the Strakoš matrix

$$A = \text{diag}(\{\alpha_i\}_{i=1}^n) \text{ with } \alpha_i = \alpha_1 + \left(\frac{i-1}{n-1}\right)(\alpha_n - \alpha_1)\rho^{n-i}, i = 1, 2, \dots, n = 10000.$$

The distribution of the eigenvalue is controlled by the parameter ρ . We take

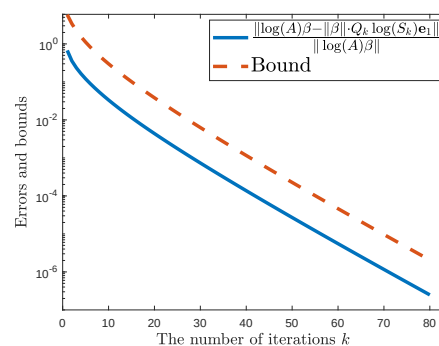
$$\alpha_1 = 1, \rho = 0.99 \text{ and } \beta = \frac{v}{\|v\|}, \text{ with } v = (1 \ 1 \ \dots \ 1)^T \in \mathbb{R}^{10000}$$

in this example. The selection of α_n is listed in Table 2.

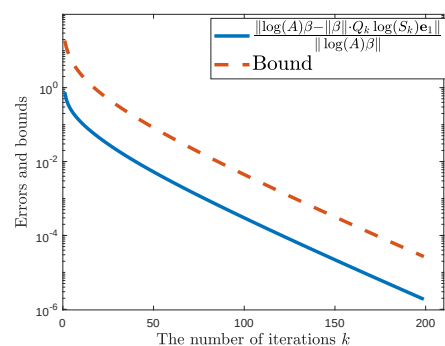
Table 2. Example 2: Different values of κ with different choices of α_n .

	$\alpha_n = 10^{-2}$	$\alpha_n = 10^{-3}$	$\alpha_n = 10^{-4}$	$\alpha_n = 10^{-5}$
κ	1×10^2	1×10^3	1×10^4	1×10^5

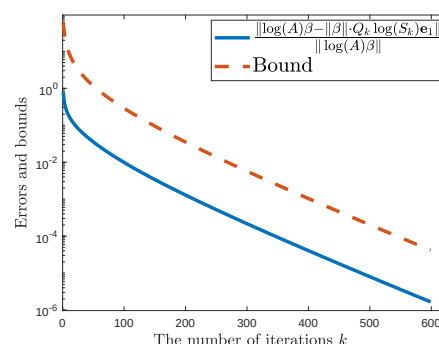
So, we obtain different values of κ with different choices of ζ and α_n ; refer to Tables 1 and 2. In Figures 1–4, we plot the curves of $\epsilon_1^{(k)}$, $\epsilon_2^{(k)}$, $\epsilon_3^{(k)}$, and $\epsilon_4^{(k)}$ with different κ , as well as the bounds in (8) and (9), respectively. It is obvious to see from Figures 1–4 that (i) the iterations increase with the increasing of κ , which is consistent with our results (see (8), (9), (13), and (14)); (ii) the convergence rate of $\epsilon_2^{(k)}$ is significantly higher than $\epsilon_1^{(k)}$ and $\epsilon_3^{(k)}$ and slightly higher than that of $\epsilon_4^{(k)}$, which is also consistent with our results; (iii) the curves of the bounds in (8), (9), (13), and (14) are almost parallel to the curves of the real values, except for $\epsilon_3^{(k)}$ in Example 2 (see Figure 4c,d). That is, our bounds can evaluate the convergence rates of $\epsilon_1^{(k)}$, $\epsilon_2^{(k)}$, $\epsilon_3^{(k)}$, and $\epsilon_4^{(k)}$ effectively, in most instances. All these show the effectiveness of the theoretical results proposed in our paper.



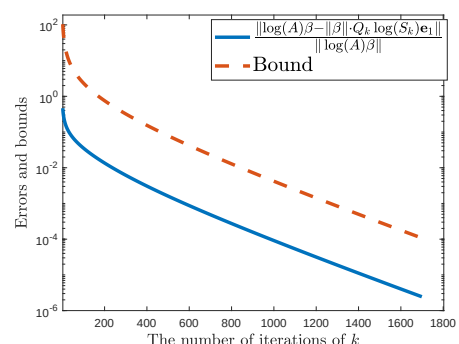
(a) $\zeta = 10 + 10^{-1}$ and $\kappa = 2.01 \times 10^2$.



(b) $\zeta = 10 + 10^{-2}$ and $\kappa = 2.00 \times 10^3$.



(c) $\zeta = 10 + 10^{-3}$ and $\kappa = 2.00 \times 10^4$.



(d) $\zeta = 10 + 10^{-4}$ and $\kappa = 1.99 \times 10^5$.

Figure 1. Cont.

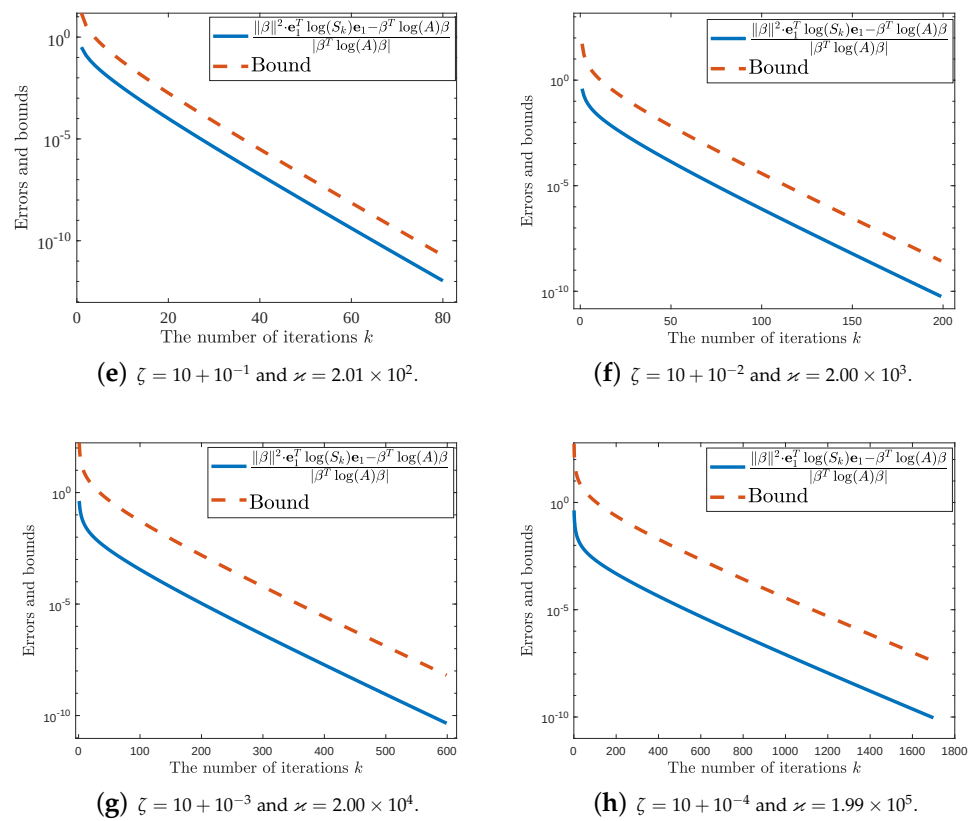


Figure 1. Example 1: Lines correspond to $\epsilon_1^{(k)}$ and $\epsilon_2^{(k)}$; refer to Theorem 1.

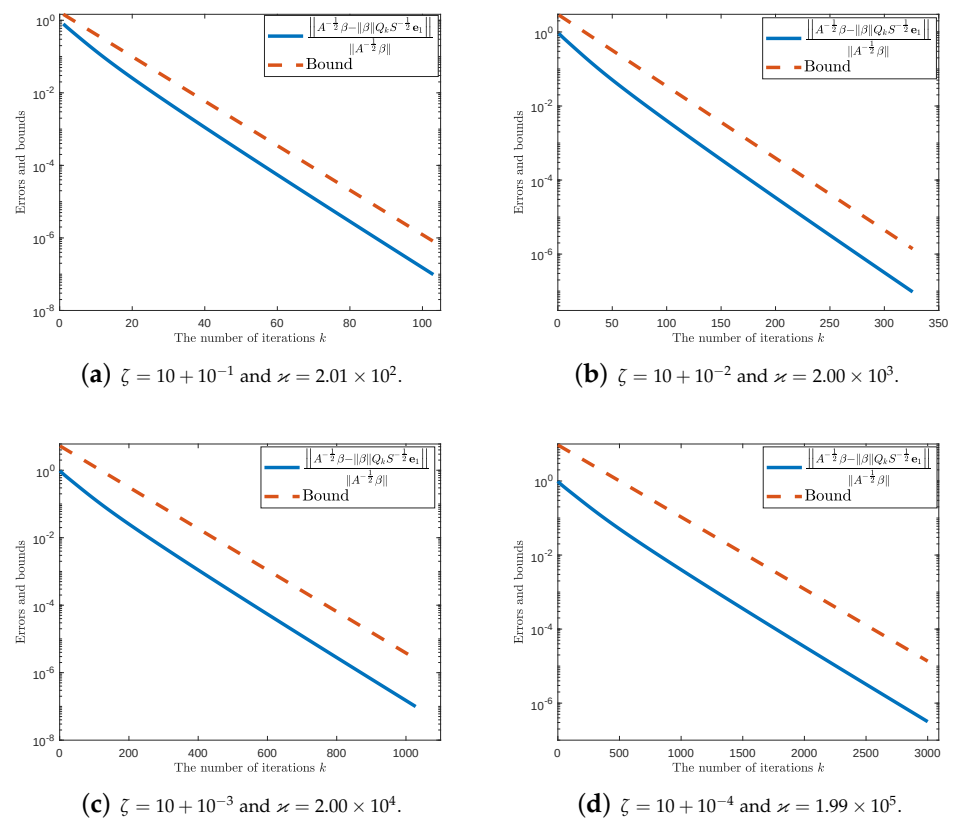


Figure 2. Cont.

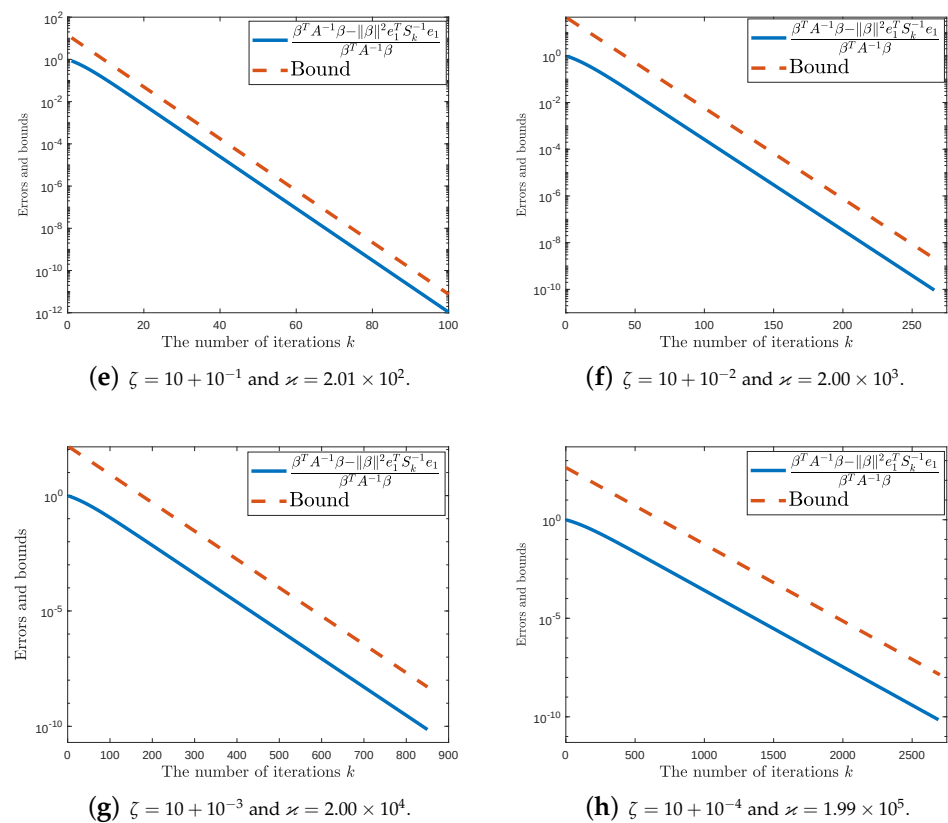


Figure 2. Example 1: Lines correspond to $\epsilon_3^{(k)}$ and $\epsilon_4^{(k)}$; refer to Theorems 2 and 3.

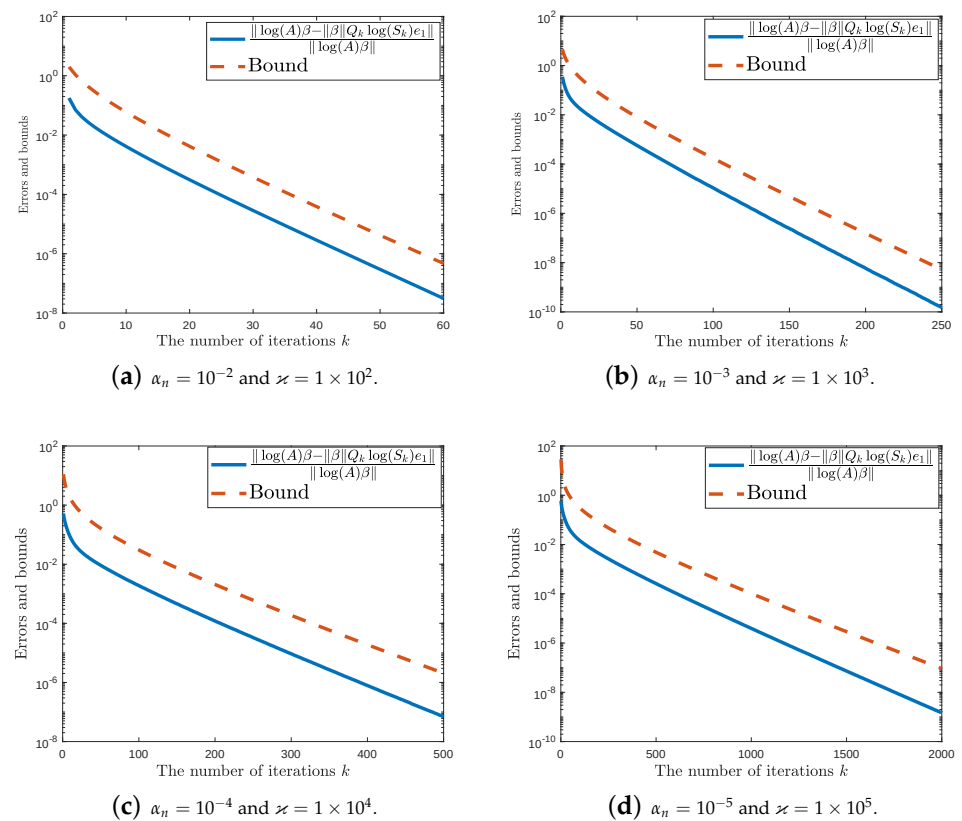


Figure 3. Cont.

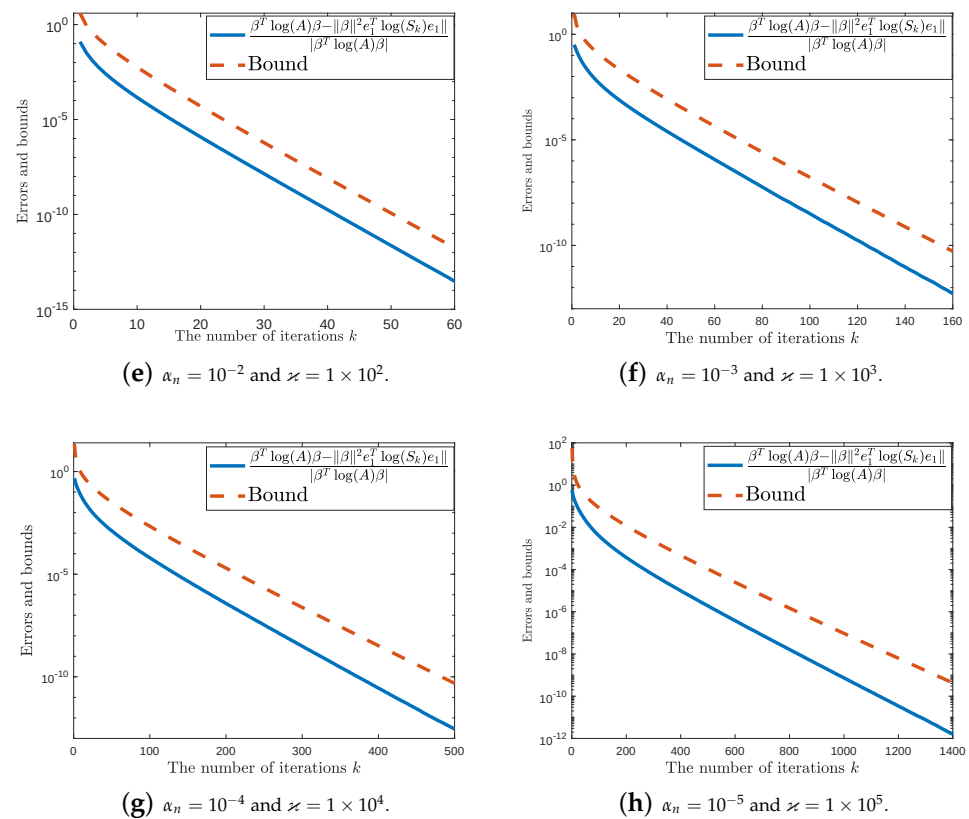


Figure 3. Example 2: Lines correspond to $\epsilon_1^{(k)}$ and $\epsilon_2^{(k)}$; refer to Theorem 1.

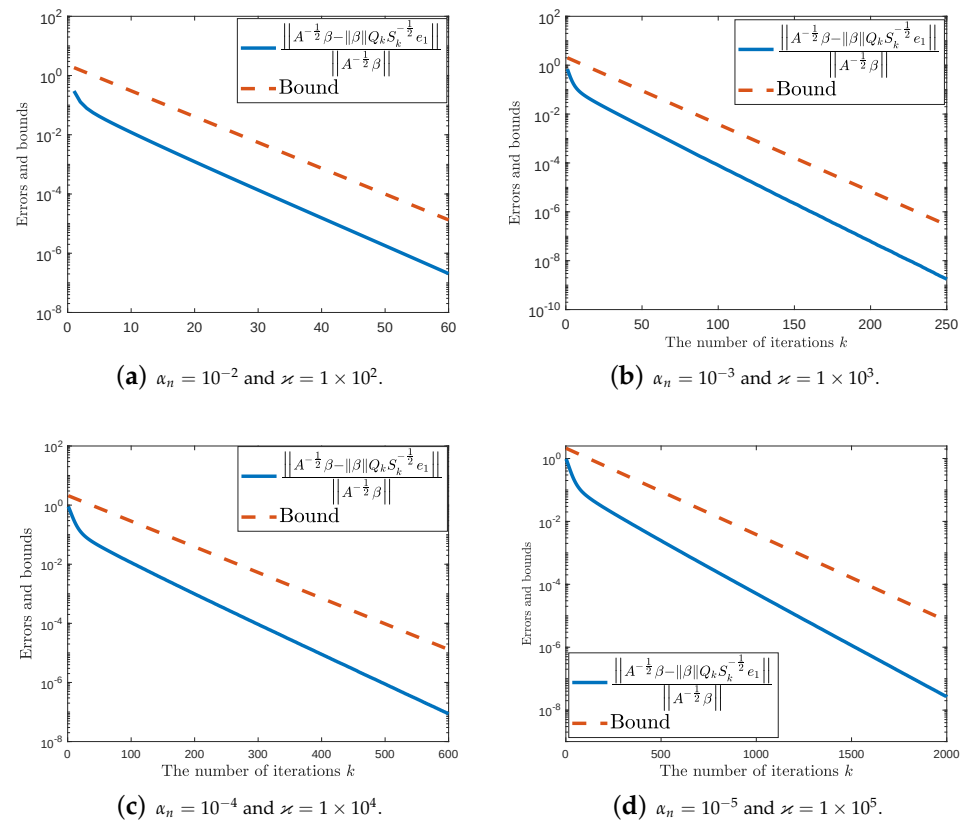


Figure 4. Cont.

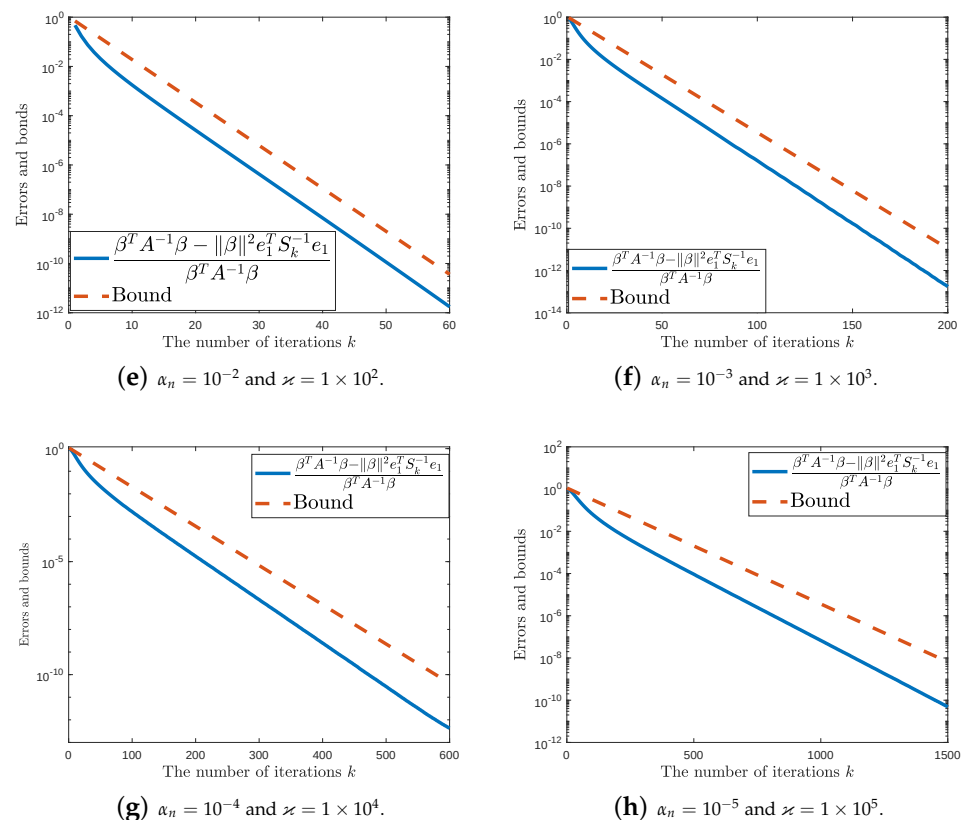


Figure 4. Example 2: Lines correspond to $\epsilon_3^{(k)}$ and $\epsilon_4^{(k)}$; refer to Theorems 2 and 3.

5. Concluding Remarks

We established the convergence rates of the Lanczos method for computing $\log(A)\beta$, $\beta^T \log(A)\beta$, $A^{-\frac{1}{2}}\beta$, and $\beta^T A^{-1}\beta$, as

$$\mathcal{O}\left(k^{-1}\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right), \mathcal{O}\left(k^{-1}\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k-1}\right), \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k\right) \text{ and } \mathcal{O}\left(\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2k-1}\right),$$

respectively. Here, A is assumed to be symmetric positive definite. It is not difficult to see that the convergence rate of $\epsilon_2^{(k)}$ is significantly higher than $\epsilon_1^{(k)}$ and $\epsilon_3^{(k)}$ and slightly higher than that of $\epsilon_4^{(k)}$. Numerical experiments illustrate the effectiveness of our theoretical results.

The Lanczos method can also be used to compute some other matrix functions, e.g., $\exp(A)\beta$, $A^p\beta$ ($0 < p < 1$), $\sin(A)\beta$, $\cos(A)\beta$, and so on. The key to convergence analysis is to find the Chebyshev expansion of these functions. Furthermore, our analysis is confined to the scenario where A is symmetric positive definite. The situation is complicated for a nonsymmetric matrix whose eigenvalues are often complex. Therefore, we must resort to the theory of polynomial approximation over complex domains for convergence analysis. Extending this analysis to cases where A is nonsymmetric warrants further investigation. Finally, it is clear from the numerical experiments that the bounds we provide are not sharp enough when the condition number is large, and it is also an interesting question to investigate whether these upper bounds can be improved further.

Author Contributions: Y.G. prepared the mathematica programs, tables, and figures; Y.G., H.M.S. and X.L. wrote the manuscript text. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Suqian University Youth Foundation (Grant No. 2023XQNA15) and Suqian Sci & Tech Program (Grant No. K202419).

Data Availability Statement: This manuscript has no associated data.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cuchta, T.; Grow, D.; Wintz, N. Discrete matrix hypergeometric functions. *J. Math. Anal. Appl.* **2023**, *518*, 126716. [\[CrossRef\]](#)
2. Cuchta, T.; Luketic, R. Discrete hypergeometric Legendre polynomials. *Mathematics* **2021**, *9*, 2546. [\[CrossRef\]](#)
3. Güttel, S. Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection. *GAMM-Mitteilungen* **2013**, *36*, 8–31. [\[CrossRef\]](#)
4. Ilic, M.D.; Turner, I.W.; Simpson, D.P. A restarted Lanczos approximation to functions of a symmetric matrix. *IMA J. Numer. Anal.* **2009**, *30*, 1044–1061. [\[CrossRef\]](#)
5. Ubaru, S.; Chen, J.; Saad, Y. Fast estimation of $\text{tr}(f(A))$ via stochastic Lanczos quadrature. *SIAM J. Matrix Anal. Appl.* **2017**, *38*, 1075–1099. [\[CrossRef\]](#)
6. Golub, G.H.; Meurant, G. *Matrices, Moments and Quadrature with Applications*; Princeton University Press: Princeton, NJ, USA, 2009; Volume 30.
7. Frommer, A.; Schweitzer, M. Error bounds and estimates for Krylov subspace approximations of Stieltjes matrix functions. *BIT Numer. Math.* **2015**, *56*, 865–892. [\[CrossRef\]](#)
8. Chen, T.; Hallman, E. Krylov-aware stochastic trace estimation. *SIAM J. Matrix Anal. Appl.* **2023**, *44*, 1218–1244. [\[CrossRef\]](#)
9. Druskin, V.L.; Knizhnerman, L.A. Error bounds in the simple Lanczos procedure for computing functions of symmetric matrices and eigenvalues. *Comput. Math. Math. Phys.* **1991**, *31*, 20–30.
10. Frommer, A.; Kahl, K.; Lippert, T.; Rittich, H. 2-norm error bounds and estimates for Lanczos approximations to linear systems and rational matrix functions. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 1046–1065. [\[CrossRef\]](#)
11. Higham, N.J. *Functions of Matrices: Theory and Computation*; SIAM: Philadelphia, PA, USA, 2008.
12. Cuchta, T.; Grow, D.; Wintz, N. A dynamic matrix exponential via a matrix cylinder transformation. *J. Math. Anal. Appl.* **2019**, *479*, 733–751. [\[CrossRef\]](#)
13. Golub, G.H.; Van Loan, C.F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, USA, 2012.
14. Saad, Y. Analysis of Some Krylov Subspace Approximations to the Matrix Exponential Operator. *SIAM J. Numer. Anal.* **1992**, *29*, 209–228. [\[CrossRef\]](#)
15. Wainwright, M.J.; Jordan, M.I. Log-determinant relaxation for approximate inference in discrete Markov random fields. *IEEE Trans. Signal Process.* **2006**, *54*, 2099–2109. [\[CrossRef\]](#)
16. Thron, C.; Dong, S.J.; Liu, K.F.; Ying, H.P. Padé- Z_2 estimator of determinants. *Phys. D-Rev. Part. Fields Gravit. Cosmol.* **1998**, *57*, 1642–1653. [\[CrossRef\]](#)
17. Affandi, H.; Fox, E.; Adams, R.; Taskar, B. Learning the parameters of determinantal point process kernels. In Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 21–26 June 2014; pp. 1224–1232.
18. Arioli, M.; Loghin, D. *Matrix Square-Root Preconditioners for the Steklov-Poincaré Operator*; Technical Report RAL-TR-2008-003; Rutherford Appleton Laboratory: Didcot, UK, 2008.
19. Tang, J.; Saad, Y. *A Probing Method for Computing the Diagonal of the Inverse of a Matrix*; Report UMSI-2010-42; Minnesota Supercomputer Institute, University of Minnesota: Minneapolis, MN, USA, 2010.
20. Meurant, G. Estimates of the trace of the inverse of a symmetric matrix using the modified Chebyshev algorithms. *Numer. Algorithms* **2009**, *51*, 309–318. [\[CrossRef\]](#)
21. Bai, Z.; Fahey, M.; Golub, G.H. Some large Scale computation problems. *J. Comput. Appl. Math.* **1996**, *74*, 71–89. [\[CrossRef\]](#)
22. Bai, Z.; Golub, G.H. Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices. *Ann. Numer. Math.* **1997**, *4*, 29–38.
23. Dong, S.J.; Liu, K.F. Stochastic estimation with Z_2 noise. *Phys. Lett. B* **1994**, *328*, 130–136. [\[CrossRef\]](#)
24. Ortner, B.; Krauter, A.R. Lower bounds for the determinant and the trace of a class of Hermitian matrices. *Linear Algebra Its Appl.* **1996**, *236*, 147–180. [\[CrossRef\]](#)
25. Brezinski, C.; Fika, P.; Mitrouli, M. Moments of a linear operator, with applications to the trace of the inverse of matrices and the solution of equations. *Numer. Linear Algebra Appl.* **2012**, *19*, 937–953. [\[CrossRef\]](#)
26. Wu, L.; Laeuchli, J.; Kalantzis, V.; Stathopoulos, A.; Gallopoulos, E. Estimating the trace of the matrix inverse by interpolating from the diagonal of an approximate inverse. *J. Comput. Phys.* **2016**, *326*, 828–844. [\[CrossRef\]](#)
27. Beals, R.; Wong, R. *Special Functions and Orthogonal Polynomials*; Cambridge University Press: Cambridge, UK, 2016.
28. Olver, F.; Lozier, D.; Boisvert, R.; Clark, C. *The NIST Handbook of Mathematical Functions*; Cambridge University Press: New York, NY, USA, 2010.
29. Bernstein, S.N. Sur l'ordre de la meilleure approximation des fonctions continues par les polynômes de degré donné. *R. Acad. Med. Belg.* **1912**, *4*, 1–104.
30. Zhan, X. *Matrix Inequalities*; Springer: Berlin/Heidelberg, Germany, 2002.
31. Zhang, L.; Shen, C.; Li, R. On the generalized Lanczos trust-region method. *SIAM J. Optim.* **2017**, *27*, 2110–2142. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.