Theoretical error bounds on the convergence of the Lanczos and block Lanczos methods

Tianruo Yang
Department of Computer and Information Science
Linköping University, S-581 83 Linköping, Sweden
Phone: +46-13-281763, Fax: +46-13-282666
Email: tiaya@ida.liu.se
Tianchu Yang
Department of Mechanical Engineering
Linköping University, S-581 83 Linköping, Sweden
Phone: +46-13-214838, Fax: +46-13-214838

May 2, 1997

Abstract

In this paper the new theoretical error bounds on the convergence of the Lanczos and the block-Lanczos methods are established based on results given by Saad. Similar further inequalities are found for the eigenelements by using bounds on the acute angle between the exact eigenvectors and the Krylov subspace spanned by $x_0, Ax_0, \cdots, A^{n-1}x_0$, where x_0 is the initial starting vector of the process. The same analysis is extended to the block-Lanczos method. Several numerical experiments are presented in order to permit a comparison between the actual rates of convergence of the Lanczos method with the theoretical error bounds.

Keywords

Lanczos and block lanczos, error bound, convergence, eigenvalue

1 Convergence of the Lanczos method

1.1 Introduction

We will follow Saad's notation suggested in [1]. Let $\lambda_1 > \lambda_2 > \cdots > \lambda_N$ be the ordered eigenvalues of A and $\phi_1, \phi_2, \cdots, \phi_N$ the associated eigenvectors of norm one. Given a starting vector x_0, the method of Lanczos provides a simple way of realizing the Ritz-Galerkin projection process on the subspace E_n spanned by the Krylov vectors $x_0, Ax_0, \cdots, A^{n-1}x_0$, where $n \leq N$ [2, 3]. If we denote $\pi_n(A)$

as the orthogonal projection on the subspace E_n, then one computes the eigenvalues $\lambda_1^{(n)} > \lambda_2^{(n)} > \cdots > \lambda_n^{(n)}$ of the operator $\pi_n A_{E_n} : E_n \to E_n$ with their associated eigenvectors $\phi_1^{(n)}, \phi_2^{(n)}, \cdots, \phi_n^{(n)}$ and take $\lambda_i^{(n)}, \phi_i^{(n)}$ as approximations to λ_i, ϕ_i.

The numerical performance of the Lanczos method has been studied by several authors [4, 3, 5, 6, 7]. They also gave several variants of the Lanczos method to compute effectively a few of the extreme eigenelements of A and indicated how eigenvalue estimates can be obtained via the Lanczos method, but revealed nothing about the rate of convergence. It is quite natural to ask how rapidly would be the approximate eigenelements $\lambda_i^{(n)}, \phi_i^{(n)}$ converge to λ_i, ϕ_i in exact precision. The first result was given by Kaniel [8], and corrected by Paige [9]. Exploiting the Courant characterization property, the alternative generalization of the Kaniel-Paige's result was shown by Saad in [1]. In his paper, a number of results were also established for eigenvectors. Saad's results, are the improvement of the similar results of Kaniel and Paige's. Kaniel and Paige's error bounds incorporate the computed Ritz vectors explicitly, and Saad's error bound does not directly estimate the goodness of the corresponding Ritz vectors. Note that it does not use the computed eigenvalue approximations. Instead it is a measure of how good the entire Krylov subspace is. As expected this bound on the eigenvector approximation indicates an accuracy which is essentially the square root of the indicated accuracy for eigenvalue approximations.

1.2 Notation and some basic properties

We will use the same notation as Saad. Because the sequence $\lambda_i^{(n)}, n = i, \cdots, N$ is finite, we can not study the convergence of $\lambda_i^{(n)}$ when $n \to \infty$. So, we have to deal with a compact operator A on a Hilbert space E. Let $\lambda_1 > \lambda_2 > \cdots > \lambda_k$ be k largest positive eigenvalues numbered in a decreasing order and all the other eigenvalues λ_j of A are assumed to satisfy $\lambda_k > \lambda_j$. The almost same results can be obtained for the negative part of the spectrum with essentially the same results. It is very easy to show that the results apply to the k largest eigenvalues of a finite dimensional operator on a space E of dimension N, numbered in decreasing order, but now we must restrict ourselves in the case of $n \leq N$.The acute angle $\theta(x, E_n)$ between a vector $x \neq 0$ and the subspace E_n is defined by

$$\theta(x, E_n) = \arcsin \frac{\| (I - \pi_n(A))x \|}{\| x \|}. \tag{1}$$

Denote by λ_{inf} the infimum of the spectrum of A and P_i the eigenprojection associated with λ_i, that is the orthogonal projection on the eigenspace corresponding to λ_i.

We now recall some results in Saad's paper [1] about the behavior of the acute angle between the eigenvector ϕ_i and the subspace E_n by giving a bound for $\tan\theta(\phi_i, E_n)$. Here the subspace E_n is generated successively by an orthonormal basis $v_1, \ldots, v_n$ when the Lanczos method is performed with a starting vector x_0. The proof is given in [10, 1].

Lemma 1.1 *Let P_{n-1} denote the space of polynomials of degree not exceeding $n-1$. If P_i is the eigenprojection associated with λ_i, and if $P_i x_0 \neq 0$, we set*

$$\hat{x}_0 = \frac{(I-P_i)x_0}{\| (I-P_i)x_0 \|}, \qquad t_{i,n} = \inf_{p \in P_{n-1} \quad p(\lambda_i)=1} \| p(A)\hat{x}_0 \|, \tag{2}$$

Then

$$\tan\theta(\phi_i, E_n) = t_{i,n} \tan\theta(\phi_i, x_0).$$

Lemma 1.1, leads to the following theorem:

Theorem 1.2 *Let $\lambda_1 > \lambda_2 > \cdots > \lambda_k$ be the largest k eigenvalues of A, and P_i the eigenprojection associated with $\lambda_i (i < k)$. Assume that $P_i x_0 \neq 0$, and consider the eigenvector $\phi_i = P_i x_0 / \| P_i x_0 \|$ associated with λ_i. Set:*

$$\gamma_i = 1 + 2\frac{\lambda_i - \lambda_{i+1}}{\lambda_{i+1} - \lambda_{inf}}, \qquad K_i = \prod_{j=1}^{i-1} \frac{\lambda_j - \lambda_{inf}}{\lambda_j - \lambda_i}, \quad K_1 = 1 \tag{3}$$

Then

$$\frac{\| (I - \pi_n(A))\phi_i \|}{\| \pi_n(A)\phi_i \|} \leq \frac{K_i}{T_{n-i}(\gamma_i)} \frac{\| (I - \pi_1(A))\phi_i \|}{\| \pi_1(A)\phi_i \|}. \tag{4}$$

where $T_k(x)$ is the Chebyshev polynomial of the first kind of degree k.

Here $\pi_1(A)$ is the orthogonal projection on the subspace E_1 spanned by x_0. The inequality can also be written like this:

$$\tan\theta(\phi_i, E_n) \leq \frac{K_i}{T_{n-i}(\gamma_i)} \tan\theta(\phi_i, x_0). \tag{5}$$

From Theorem 1.2, It easily follows that the acute angle between ϕ_i and E_n decreases at least as rapidly as $K_i / T_{n-i}(\gamma_i)$. When n is large, then

$$T_{n-i}(\gamma_i) \simeq 1/2(\gamma_i + \sqrt{\gamma_i^2 - 1})^{n-i}, \tag{6}$$

is greater than 1, and depends on the gap $\lambda_i - \lambda_{inf}$ and the spread $\lambda_{i+1} - \lambda_{inf}$. We remark that the first $i-1$ eigenvalues do not interfere in the coefficient $\tau_i = \gamma_i - \sqrt{\gamma_i^2 - 1}$, which estimates the rate of convergence to zero of the bound on $\theta(\phi_i, E_n)$. From this theorem, we can also learn that there is at least one vector in E_n which converges to the eigenvector ϕ_i with a rate superior to τ_i above when n increases. This is true even when λ_i is not simple, since the only condition required is that $P_i x_0 \neq 0$. The proof of the theorem reveals that when λ_i is multiple there is only one such vector in E_n which is close to ϕ_i. This shows in particular that a multiple eigenvalue will be approximated by at most one eigenvalue of T_n. For details of the above remarks, see [1].

1.3 The error bounds of Kaniel-Paige and Saad

We review the error bounds of Kaniel-Paige and Saad in this subsection suggested in [11]. Each element s in the Krylov subspace E_n has the special form:

$$s = \sum_{i=0}^{n-1} (A^i x_0)\xi_i = \sum_{i=0}^{n-1} (\xi_i A^i) x_0 = \pi_n(A) x_0, \tag{7}$$

where $\pi_n(A)$ is a polynomial of degree $< m$.

Now we cite a basic lemma from [11] which can be used to derive bounds on $\lambda_i - \lambda_i^{(n)}$ for each $i = 1, 2, \ldots, n$.

Lemma 1.3 *Let h be the normalized projection of x_0 orthogonal to Z_i, where the subspace Z_i =span$(\phi_1, \phi_2, \ldots, \phi_i)$. For each $\pi_i(A) \in P_{n-1}$ and each $i \leq n$ the Rayleigh quotient ρ satisfies*

$$\rho(\pi_i(A)x_0; A - \lambda_i) \leq (\lambda_i - \lambda_{inf})[\tan\theta(\phi_i, x_0)\frac{\|\pi_i(A)h\|}{\pi_i(\lambda_i)}]^2. \tag{8}$$

As we said before, the error bound on $\lambda_i - \lambda_i^{(n)}$ depends on several quantities. However, as will be seen shortly, the leading role is played by the Chebyshev polynomial T_{n-i} whose steep climb outside the interval $[-1, 1]$ helps to explain the excellent approximations obtained from Krylov subspaces.

The error bounds come from choosing a polynomial π in Lemma1.3 such that, among other things

1. $|\pi_i(\lambda_i)|$ is large while $\|\pi_i(A)h\|$ is small, and
2. $\rho(s; A - \lambda_i) \geq 0$ where $s = \pi_i(A)x_0$.

The second requirement concerns the left side of the inequality in Lemma 1.3, namely $\rho(s; A - \lambda_i)$. The following facts are known:

1. $0 < \lambda_i - \lambda_i^{(n)}$.
2. $\lambda_i - \lambda_i^{(n)} \leq \rho(s; A - \lambda)$. if $s - \phi_j^{(n)}$ for all $j < i$.
3. $\lambda_i - \lambda_i^{(n)} \leq \rho(s; A - \lambda_i) + \sum_{j=1}^{i-1} (\lambda_j - \lambda_{inf})\varepsilon_j^2$, where ε_j represents the sine of the angle between ϕ_j and $\phi_j^{(n)}$. if $s - \phi_j$ for all $j < i$.

The complete explanation can be found in [11]. It is clear from the first inequality that if $\rho(s; A - \lambda_i) < 0$ then, a fortiori, $\rho(s; A - \lambda_i) < \lambda_i - \lambda_i^{(n)}$ and Lemma 1.3 can not be used to bound $\lambda_i - \lambda_i^{(n)}$. Hence we have to turn the attention to the second one, which yields Saad's bounds, or the third one which yields Kaniel-Paige's bounds.

Theorem 1.4 (Kaniel-Paige's Bounds) *The Rayleigh-Ritz approximations* $(\lambda_i^{(n)}, \phi_i^{(n)})$ *from* E_n *to* (λ_i, ϕ_i) *satisfy*

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq (\lambda_i - \lambda_{inf}) \frac{K_i^2}{T_{n-i}^2(\gamma_i)} \tan^2 \theta(\phi_i, x_0) + \sum_{j=1}^{i-1} (\lambda_j - \lambda_{inf}) \sin^2 \theta(\phi_j, \phi_j^{(n)}),$$

where $\sin^2 \theta(\phi_j, \phi_j^{(n)})$ *are bounded by*

$$\sin^2 \theta(\phi_j, \phi_j^{(n)}) \leq \frac{\lambda_j - \lambda_j^{(n)} + \sum_{k=1}^{j-1} (\lambda_k - \lambda_{j+1}) \sin^2 \theta(\phi_k, \phi_k^{(n)})}{\lambda_j - \lambda_{j+1}}, \qquad (9)$$

in which

$$\gamma_i = 1 + \frac{2(\lambda_i - \lambda_{i+1})}{(\lambda_{i+1} - \lambda_{inf})}, \qquad K_i = \sum_{j=1}^{i-1} \frac{\lambda_j - \lambda_{inf}}{\lambda_j - \lambda_i}.$$

Kaniel-Paige's error bounds incorporate the computed Ritz vectors explicitly, and the effect upon their bounds as one move into the interior of the spectrum is much more obvious. Those bounds indicate specific decreases in accuracy of the computed eigenvalues and of the corresponding Ritz vectors as we move into the interior of the spectrum.

Theorem 1.5 (Saad's Bounds) *Let* $\lambda_1^{(n)} \geq \ldots \geq \lambda_n^{(n)}$ *be the Ritz values derived from* E_n *and let* (λ_i, ϕ_i) *be the eigenpair of* A. *For each* $i = 1, \ldots, n$

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq (\lambda_i - \lambda_{inf}) [\frac{K_i^{(n)}}{T_{n-i}(\gamma_i)}]^2 \tan^2 \theta(\phi_i, x_0), \qquad (10)$$

where

$$K_i^{(n)} = \prod_{j=1}^{i-1} \frac{\lambda_j^{(n)} - \lambda_{inf}}{\lambda_j^{(n)} - \lambda_i}, \quad \textit{and} \qquad K_1^{(n)} = 1. \qquad (11)$$

These error bounds indicate that for many matrices and for relatively small n, several of the extreme eigenvalues of A, that is several of the algebraically-largest or algebraically-smallest of the eigenvalues of A, are well approximated by eigenvalues of the corresponding Lanczos matrices. In practice it is not always true that both ends of the spectrum of the given matrix are equally well-approximated. However, it is generally true that at least one end of the spectrum is approximated well. Some examples in [12] illustrate, however, that in some cases the Lanczos matrix must be as large as the original matrix before good eigenvalue approximations are obtained for any of the eigenvalues of A, including the extreme ones. In next section, we will present a more compact approximation and some related results for eigenelements.

1.4 New theoretical bounds

Before our error bounds, we need two lemmas for the proof.

Lemma 1.6 *Let P_n^* be the set of all polynomials of degree n, with leading coefficient equal to 1, let $\overline{p}(x)$ be the polynomial of P_n^* which minimizes $\|p(A)x_0\|$ over all elements of P_n^*. Then the approximate eigenvalues $\lambda_i^{(n)}$ are the roots of $\overline{p}$, and if we set $q_i(x) = \overline{p}(x)/(x - \lambda_i^{(n)})$, the vector $q_i(A)x_0$ are associated approximate eigenvectors.*

The Lemma 1.6 is a known conclusion. About the proof, see [10, 13] for detail.

Lemma 1.7 *For $j = 1, 2, \ldots, i$, let $\phi_j^{(n)}$ be the approximate eigenvectors associated with $\lambda_j^{(n)}$ and $F_i^{(n)}$ the subspace of E_n, orthogonal to $\phi_1^{(n)}, \phi_2^{(n)}, \ldots,$ $\phi_{i-1}^{(n)}$. Then $x \in F_i^{(n)}$ if and only if $x = p(A)x_0$, where p is a polynomial of degree $\leq n-1$ such that $p(\lambda_1^{(n)}) = p(\lambda_2^{(n)}) = \cdots = p(\lambda_{i-1}^{(n)}) = 0$.*

The Lemma 1.7 is a simple consequence of Lemma 1.6.

Theorem 1.8 *Let $\lambda_0 = \lambda_1 > \lambda_2 \ldots > \lambda_n$ be eigenvalues of A with $i \leq k$ with associated eigenvector ϕ_i such that $(\phi_i, x_0) \neq 0$, and assume that $\lambda_{i-1}^{(n)} > \lambda_i$. Let γ_i be defined as in the Theorem 1.2. Then*

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq \min(\Psi, \Theta, \Omega), \tag{12}$$

in which

$$\Psi = (\lambda_i - \lambda_{inf}) - (\lambda_{i-1} - \lambda_{inf}) \frac{T_{n-i}^2(\gamma_{i-1})}{T_{n-i}^2(\gamma_{i-1}) + \tan^2\theta(\phi_{i-1}, x_0)\left(K_{i-1}^{(n)}\right)^2},$$

$$\Theta = (\lambda_i - \lambda_{inf}) - (\lambda_{i+1} - \lambda_{inf}) \frac{T_{n-i}^2(\gamma_i)}{T_{n-i}^2(\gamma_i) + \tan^2\theta(\phi_i, x_0)\left(K_i^{(n)}\right)^2},$$

$$\Omega = (\lambda_i - \lambda_{inf}) - (\lambda_i - \lambda_{inf}) \frac{T_{n-i}^2(\gamma_i) - \left(K_i^{(n)}\right)^2 \tan^2\theta(\phi_i, x_0)}{T_{n-i}^2(\gamma_i)}.$$

Here Ω is Saad's bounds, and $\gamma_0 = \gamma_1$,

$$K_i^{(n)} = \prod_{j=1}^{i-1} \frac{\lambda_j^{(n)} - \lambda_{inf}}{\lambda_j^{(n)} - \lambda_i}, \quad K_1^{(n)} = K_0^{(n)} = 1, \qquad \gamma_i = 1 + \frac{2(\lambda_i - \lambda_{i+1})}{(\lambda_{i+1} - \lambda_{inf})}.$$

Proof:

1. Let us prove the first part of Theorem 1.8:

Making use of the Courant characterization of the eigenvalues of symmetric operators, we have

$$\lambda_i^{(n)} = \max_{u \in F_i^{(n)}} \frac{(Au, u)}{\|u\|^2}.$$

Let $u \in F_i^{(n)}$, $u = p(A)x_0 = \sum_{j=1}^{\infty} \alpha_j p(\lambda_j)\phi_j$, where the α_j are the expansion coefficients of x_0 in the eigenbasis ϕ_j, Then we get

$$\begin{aligned} \frac{(Au, u)}{\|u\|^2} &= \frac{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2(\lambda_j - \lambda_i + \lambda_i)}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2} \\ &= \lambda_i + \frac{\sum_{j=1}^{i-1} p^2(\lambda_j)\alpha_j^2(\lambda_j - \lambda_i) - \sum_{j=i}^{\infty} p^2(\lambda_j)\alpha_j^2(\lambda_i - \lambda_j)}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2}. \end{aligned}$$

For $1 \le j \le i-1$, $\lambda_j - \lambda_i \ge \lambda_{i-1} - \lambda_i$ and for $j \ge i$, $\lambda_i - \lambda_j \le \lambda_i - \lambda_{inf}$
Thus

$$\frac{(Au, u)}{\|u\|^2} \ge \lambda_i + \frac{\sum_{j=1}^{i-1}(\lambda_{i-1} - \lambda_i)p^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2} - \frac{\sum_{j=i}^{\infty}(\lambda_i - \lambda_{inf})p^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2}.$$

Extend $\lambda_{i-1} - \lambda_i$ of the second term as $\lambda_{i-1} - \lambda_{inf} + \lambda_{inf} - \lambda_i$, we know

$$\begin{aligned} \frac{(Au, u)}{\|u\|^2} &\ge \lambda_i - \frac{\sum_{j=1}^{\infty}(\lambda_i - \lambda_{inf})p^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2} + \frac{\sum_{j=1}^{i-1}(\lambda_{i-1} - \lambda_{inf})p^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2} \\ &\ge \lambda_{inf} + (\lambda_{i-1} - \lambda_{inf})\frac{\sum_{j=1}^{i-1} p^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2}. \end{aligned}$$

Since $\sum_{j=1}^{i-1} p^2(\lambda_j)\alpha_j^2 \ge p^2(\lambda_{i-1})\alpha_{i-1}^2$

$$\frac{(Au, u)}{\|u\|^2} \ge \lambda_{inf} + (\lambda_{i-1} - \lambda_{inf})\frac{1}{1 + \frac{\sum_{j=i}^{\infty} p^2(\lambda_j)\alpha_j^2}{p^2(\lambda_{i-1})\alpha_{i-1}^2}}. \tag{13}$$

From Lemma 1.7, $u \in F_i^{(n)}$ implies that $p(\lambda_1^{(n)}) = p(\lambda_2^{(n)}) = \cdots = p(\lambda_{i-1}^{(n)}) = 0$. In other words, this means that $p(x)$ can be written as $p(x) = (x - \lambda_1^{(n)})\cdots(x - \lambda_{i-1}^{(n)})q(x)$, with $q(x)$ having degree at most $n-i$. Hence for all $u \in F_1^{(n)}$

$$\begin{aligned} \frac{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2}{p^2(\lambda_{i-1})\alpha_{i-1}^2} &= \sum_{j=i}^{\infty} \frac{\left(\lambda_j - \lambda_1^{(n)}\right)^2 \cdots \left(\lambda_j - \lambda_{i-1}^{(n)}\right)^2 q^2(\lambda_j)\alpha_j^2}{\left(\lambda_{i-1} - \lambda_1^{(n)}\right)^2 \cdots \left(\lambda_{i-1} - \lambda_{i-1}^{(n)}\right)^2 q^2(\lambda_{i-1})\alpha_j^2}, \\ &\le \frac{\left(\lambda_1^{(n)} - \lambda_{inf}\right)^2 \cdots \left(\lambda_{i-1}^{(n)} - \lambda_{inf}\right)^2}{\left(\lambda_1^{(n)} - \lambda_{i-1}\right)^2 \cdots \left(\lambda_{i-1}^{(n)} - \lambda_{i-1}\right)^2} \sum_{j=i}^{\infty} \frac{q^2(\lambda_j)\alpha_j^2}{q^2(\lambda_{i-1})\alpha_{i-1}^2}. \end{aligned}$$

Thus

$$\max_{u\in F_i^{(n)}} \frac{\sum_{j=i}^{\infty} p^2(\lambda_j)\alpha_j^2}{p^2(\lambda_{i-1})\alpha_{i-1}^2} \le \left(K_{i-1}^{(n)}\right)^2 \max_{q\in P_{n-i}} \sum_{j=i}^{\infty} \frac{q^2(\lambda_j)\alpha_j^2}{q^2(\lambda_{i-1})\alpha_{i-1}^2}. \quad (14)$$

Applying (13) into (14), we immediately get the following inequalities

$$\begin{aligned}
\lambda_i^{(n)} &= \max_{u\in F_i^{(n)}} \frac{(Au,u)}{\|u\|^2},\\
&\ge \max_{u\in F_i^{(n)}} (\lambda_{inf} + (\lambda_{i-1}-\lambda_{inf}) \frac{1}{1+\left(K_{i-1}^{(n)}\right)^2 \max_{q\in P_{n-i}} \sum_{j=i}^{\infty} \frac{q^2(\lambda_j)\alpha_j^2}{q^2(\lambda_{i-1})\alpha_{i-1}^2}}),\\
&= \lambda_{inf} + (\lambda_{i-1}-\lambda_{inf}) \frac{1}{1+\left(K_{i-1}^{(n)}\right)^2 \min_{q\in P_{n-i}} \max_{j\ge i} \sum_{j=i}^{\infty} \frac{q^2(\lambda_j)\alpha_j^2}{q^2(\lambda_{i-1})\alpha_{i-1}^2}},\\
&\ge \lambda_{inf} + (\lambda_{i-1}-\lambda_{inf}) \frac{1}{1+\left(K_{i-1}^{(n)}\right)^2 \sum_{j\ge i} \frac{\alpha_j^2}{\alpha_{i-1}^2} \min_{q\in P_{n-i}} \max_{j\ge i} \|\frac{q(\lambda_j)}{q(\lambda_{i-1})}\|^2}.
\end{aligned}$$

Since

$$\min_{q\in P_{n-i}} \max_{j\ge i} \|\frac{q(\lambda_j)}{q(\lambda_{i-1})}\|^2 \le \frac{1}{T_{n-i}^2(\gamma_{i-1})}, \qquad \sum_{j\ge i} \frac{\alpha_j^2}{\alpha_{i-1}^2} \le \tan^2\theta(\phi_{i-1}, x_0). \quad (15)$$

From (15), it is very easy to show

$$\lambda_i^{(n)} \ge \lambda_{inf} + (\lambda_{i-1}-\lambda_{inf}) \frac{T_{n-i}^2(\gamma_{i-1})}{T_{n-i}^2(\gamma_{i-1}) + \tan^2\theta(\phi_{i-1}, x_0)(K_{i-1}^{(n)})}.$$

Hence, we get the first part of Theorem 1.8.

2. Let us finish the second part of Theorem 1.8:

Using the similar ideas, we get the following inequalities

$$\begin{aligned}
\frac{(Au,u)}{\|u\|^2} &= \lambda_i + \frac{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2(\lambda_j-\lambda_i)}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2},\\
&= \lambda_i + \frac{\sum_{j=1}^{i} p^2(\lambda_j)\alpha_j^2(\lambda_j-\lambda_i)}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2} - \frac{\sum_{j=i+1}^{\infty} p^2(\lambda_j)\alpha_j^2(\lambda_i-\lambda_j)}{\sum_{j=1}^{\infty} p^2(\lambda_j)\alpha_j^2}.
\end{aligned}$$

For $1\le j\le i$, $\lambda_j-\lambda_i \ge \lambda_{i+1}-\lambda_i$ and for $j\ge i+1$, $\lambda_i-\lambda_j \le \lambda_i-\lambda_{inf}$

$$\begin{aligned}
\frac{(Au,u)}{\|u\|^2} &\ge \lambda_{inf} + (\lambda_{i+1}-\lambda_{inf}) \frac{1}{1+\frac{\sum_{j=i+1}^{\infty} p^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{i} p^2(\lambda_j)\alpha_j^2}},\\
&\ge \lambda_{inf} + (\lambda_{i+1}-\lambda_{inf}) \frac{1}{1+\frac{\sum_{j=i+1}^{\infty} p^2(\lambda_j)\alpha_j^2}{p^2(\lambda_i)\alpha_i^2}},
\end{aligned}$$

$$\geq \quad \lambda_{inf} + (\lambda_{i+1} - \lambda_{inf}) \frac{T_{n-i}^2(\gamma_i)}{T_{n-i}^2(\gamma_i) + \tan^2\theta(\phi_i, x_0)(K_i^{(n)})^2}.$$

Hence, we get the second part of Theorem 1.8. ■

The corresponding inequalities for eigenvectors state as

Theorem 1.9 *Let $\lambda_0 = \lambda_1 > \lambda_2 \ldots > \lambda_n$ be eigenvalues of A with $i \leq k$ with associated eigenvector ϕ_i such that $\|\phi_i\| = 1$. Let $P_i^{(n)}$ denote the approximate eigenprojection associated with $\lambda_i^{(n)}$, and $d_{i,n} = \min_{j \neq i} |\lambda_i - \lambda_j^{(n)}|$, and $r_n = \|(I - \pi_n)A\pi_n\|$. Then*

$$\|(I - P_i^{(n)})\phi_i\| \leq \sqrt{1 + \frac{r_n^2}{d_{i,n}^2}} \|(I - \pi_n)\phi_i\|. \tag{16}$$

or

$$\sin\theta(\phi_i, \phi_i^{(n)}) \leq \sqrt{1 + \frac{r_n^2}{d_{i,n}^2}} \sin\theta(\phi_i, E_n) \leq \sqrt{1 + \frac{r_n^2}{d_{i,n}^2}} \frac{K_i}{T_{n-i}(\gamma_i)} \tan\theta(\phi_i, x_0). \tag{17}$$

The last bound comes from Theorem 1.2. About the detail of proof, see [1].

1.5 Refined error bounds

From Theorem 1.8, the bounds reveal that they are weak in the case where λ_i is close to λ_{i+1}, for λ_i is then close to 1 and the right side of (13), (13) and (13) can decrease too slowly to 0. It is quite natural to improve them by generalizing them so that they allow to take advantage of a particular structure of the spectrum. This idea can be achieved by choosing a more appropriate polynomial. In this section, we denote p any integer such that $0 \leq p \leq n - i$. We call L the set of the P integers $i+1, i+2, \cdots, i+p$.

Theorem 1.10 *Let $\lambda_0 = \lambda_1 > \lambda_2 \ldots > \lambda_n$ be eigenvalues of A with $i \leq k$ with associated eigenvector ϕ_i such that $(\phi_i, x_0) \neq 0$, and assume that $\lambda_{i-1}^{(n)} > \lambda_i$. Let γ_i and $K_i^{(n)}$ be defined as same as in Theorem 1.2, and*

$$x_L = \prod_{j \in L} (A - \lambda_j) x_0, \qquad y_L = (I - P_1 - P_2 - \cdots - P_{i-1}) x_L.$$

and

$$\gamma_i = 1 + \frac{2(\lambda_i - \lambda_{i+p+1})}{(\lambda_{i+p+1} - \lambda_{inf})}.$$

Then

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq \min(\overline{\Psi}, \overline{\Theta}, \overline{\Omega}), \tag{18}$$

in which

$$
\begin{aligned}
\overline{\Psi} &= (\lambda_i - \lambda_{inf}) - (\lambda_{i-1} - \lambda_{inf}) \frac{T^2_{n-i-p}(\gamma_{i-1})}{T^2_{n-i-p}(\gamma_{i-1}) + \tan^2\theta(\phi_{i-1}, y_L)\left(K^{(n)}_{i-1}\right)^2}, \\
\overline{\Theta} &= (\lambda_i - \lambda_{inf}) - (\lambda_{i+1} - \lambda_{inf}) \frac{T^2_{n-i-p}(\gamma_i)}{T^2_{n-i-p}(\gamma_i) + \tan^2\theta(\phi_i, y_L)\left(K^{(n)}_{i}\right)^2}, \\
\overline{\Omega} &= (\lambda_i - \lambda_{inf}) - (\lambda_i - \lambda_{inf}) \frac{T^2_{n-i-p}(\gamma_i) - \left(K^{(n)}_{i}\right)^2 \tan^2\theta(\phi_i, y_L)}{T^2_{n-i-p}(\gamma_i)}.
\end{aligned}
$$

Proof:

We can set $\hat{E}_n$ which contains all elements of the form $u = p(A)x_L$, where p is any polynomial of degree at most $n-p-1$, is a subspace of E_n orthogonal to $\phi_{i+1}, \phi_{i+2}, \cdots, \phi_{i+p}$. Let $\hat{F}^{(n)}_i$ be the subspace of E_n orthogonal to the subspace spanned by $\phi^{(n)}_1, \phi^{(n)}_2, \cdots, \phi^{(n)}_{i-1}$. Then $\hat{F}^{(n)}_i \subset F^{(n)}_i$ and we can repeat the proof of Theorem 1.8 with $\hat{F}^{(n)}_i$ instead of $F^{(n)}_i$, x_L instead of x_0, and P_{n-p-i} instead of P_{n-i}. ∎

By majorizing the term $\tan(\phi_i, y_L)$ we can obtain the following weakening of the bound (18).

Theorem 1.11 *Under the same assumptions as in Theorem 1.10, let*

$$
K_L = \prod_{j \in L} \frac{\lambda_j - \lambda_{inf}}{\lambda_i - \lambda_j}, \qquad K_L = 1 \quad \textit{if } p = 0.
$$

Then, we have the following inequalities

$$
0 \le \lambda_i - \lambda^{(n)}_i \le \min(\hat{\Psi}, \hat{\Theta}, \hat{\Omega}), \tag{19}
$$

in which

$$
\begin{aligned}
\hat{\Psi} &= (\lambda_i - \lambda_{inf}) - (\lambda_{i-1} - \lambda_{inf}) \frac{T^2_{n-i-p}(\gamma_{i-1})}{T^2_{n-i-p}(\gamma_{i-1}) + \tan^2\theta(\phi_{i-1}, x_0)\left(K^{(n)}_{i-1}\right)^2 K_L^2}, \\
\hat{\Theta} &= (\lambda_i - \lambda_{inf}) - (\lambda_{i+1} - \lambda_{inf}) \frac{T^2_{n-i-p}(\gamma_i)}{T^2_{n-i-p}(\gamma_i) + \tan^2\theta(\phi_i, x_0)\left(K^{(n)}_{i}\right)^2 K_L^2}, \\
\hat{\Omega} &= (\lambda_i - \lambda_{inf}) - (\lambda_i - \lambda_{inf}) \frac{T^2_{n-i-p}(\gamma_i) - \left(K^{(n)}_{i}\right)^2 \tan^2\theta(\phi_i, x_0) K_L^2}{T^2_{n-i-p}(\gamma_i)}.
\end{aligned}
$$

Proof:

Because $x_0 = \sum_{k=1}^{\infty} \alpha_k \phi_k$ and $\hat{\alpha}_k = \prod_{j \in L} (\lambda_k - \lambda_j)\alpha_k$, where $\hat{\alpha}_j$ are the expansion coefficients of x_L in the eigenbasis, the term $\sum_{j \ge p+i} \hat{\alpha}^2_j / \hat{\alpha}^2_{i-1}$ can be bounded by

$$\sum_{j\geq p+i}\frac{\hat{\alpha}_j^2}{\hat{\alpha}_{i-1}^2}\leq K_L^2\sum_{j=i+p}^{\infty}\frac{\alpha_j^2}{\alpha_{i-1}^2}\leq K_L^2\tan^2\theta(\phi_i,x_0).$$

■

Moreover, We also can select p optimally, in other words, the right side of is minimal over all possible p. This gives the optimal bounds.

1.6 Example of these bounds

Let us compare Kaniel-Paige's, Saad's and new bounds with the following example which was considered in Kaniel's original paper.
Eigenvalues of A

$$\lambda_1=1.0,\quad \lambda_2=0.95,\quad \lambda_3=0.9453,\quad 0\leq\lambda_j\leq 0.94\quad \text{for}\quad j\geq 4$$

We assume that $\tan\theta(\phi_i,x_0)=100$ for $i=1,2,3$. If Lanczos algorithm is interrupted at $n=53$ and $K_i^{(n)}$ of Saad's bound equals to the corresponding factors K_i in Kaniel-Paige's bounds to the given accuracy. There is no loss in taking all angles to be acute.

- For the 1st eigenvalue, we have

$$\gamma_0=1.105,\ \gamma_1=1.105,\ K_1^{(n)}=K_1=1,\ T_{52}(1.105)=9.109e+09.$$

Kaniel-Paige:

$$0\leq\lambda_1-\lambda_1^{(n)}\leq 1.205e-16,$$

$$\varepsilon_1^2=\sin^2(\phi_1,\phi_1^{(n)})\leq 2.410e-15.$$

Saad:

$$0\leq\lambda_1-\lambda_1^{(n)}\leq 1.205e-16.$$

New:

$$\Psi=1.205e-16,\quad \Theta=5.000e-02,\quad \Omega=1.205e-16.$$

$$0\leq\lambda_1-\lambda_1^{(n)}\leq 1.205e-16.$$

- For the 2nd eigenvalue, we have

$$\gamma_1=1.105,\quad \gamma_2=1.010,\quad K_2^{(n)}=K_2=20,\quad K_1^{(n)}=K_1=1.$$

$$T_{51}(1.010) = 6.741e+02, \qquad T_{51}(1.105) = 5.783e+09.$$

Kaniel-Paige:

$$0 \leq \lambda_2 - \lambda_2^{(n)} \leq 8.363e+00,$$

$$\varepsilon_2^2 = \sin^2(\phi_2, \phi_2^{(n)}) \leq 1.779e+03.$$

Saad:

$$0 \leq \lambda_2 - \lambda_2^{(n)} \leq 8.363e+00.$$

New:

$$\Psi = 5.000e-02, \quad \Theta = 8.536e-01, \quad \Omega = 8.363e+00.$$

$$0 \leq \lambda_2 - \lambda_2^{(n)} \leq 5.000e-02.$$

- For the 3rd eigenvalue, we have

$$\gamma_2 = 1.010, \quad \gamma_3 = 1.011, \quad K_2^{(n)} = K_2 = 20, \quad K_3^{(n)} = K_3 = 220.$$

$$T_{50}(1.011) = 9.127e+02, \qquad T_{50}(1.010) = 5.851e+02.$$

Kaniel-Paige:

$$0 \leq \lambda_3 - \lambda_3^{(n)} \leq 2.239e+03.$$

Saad:

$$0 \leq \lambda_3 - \lambda_3^{(n)} \leq 5.492e+02.$$

New:

$$\Psi = 9.437e-01, \quad \Theta = 8.704e-02, \quad \Omega = 5.492e+02.$$

$$0 \leq \lambda_1 - \lambda_1^{(n)} \leq 8.704e-02.$$

From this example, we can see that Saad's bounds are simpler and tighter than Kaniel-Paige's bounds. The difference between new and Saad's bounds depends on the spectrum of A. Combining these two bounds Ψ and Θ with Saad's bounds Ω, the new bounds can give us the better understanding on the rate of convergence of the Lanczos method.

Table 1: Lanczos algorithm of Example 1 stopped with $n = 15$,

i		1	2
Observed $\tan\theta(\phi_i, E_n)$		3.90e-06	3.51e-04
Bound for $\tan\theta(\phi_i, E_n)$		1.54e-05	1.20e-03
Observed $\lambda_i - \lambda_i^{(n)}$		2.06e-11	1.02e-07
Kaniel-Paige's bound for $\lambda_i - \lambda_i^{(n)}$		6.50e-10	3.70e-06
Saad's bound for $\lambda_i - \lambda_i^{(n)}$		6.50e-10	3.71e-06
New bounds for $\lambda_i - \lambda_i^{(n)}$	Ψ	7.17e-11	1.74e-07
	Θ	1.55e+00	1.25e+00
	Ω	6.50e-10	3.71e-06
Observed $\sin\theta(\phi_i, \phi_i^{(n)})$		4.07e-06	3.08e-04
Bound for $\sin\theta(\phi_i, \phi_i^{(n)})$		2.45e-05	1.90e-03

1.7 Numerical experiments

In this subsection, we compare the effective quantities $\theta(\phi_i, E_n)$, $\lambda_i - \lambda_i^{(n)}$ and $\theta(\phi_i, \phi_i^{(n)})$ with their theoretical bounds. All tests were performed on SUN workstation using double precision.

The first example is a diagonal matrix A of order $N = 50$, with the following distribution for the eigenvalues:

$$\lambda_1 = 1.8, \quad \lambda_2 = 0.25, \quad \lambda_k = \cos\frac{(2k-5)\pi}{2(N-2)}, \quad \text{for} \quad k = 3, \cdots, N$$

We assume that the starting vector x_0 is the vector $e = (1, 1, 1, \cdots, 1)^T$, which forms the same acute angle with each eigenvector of A. The eigenvector ϕ_i is the ith vector of the canonical basis, and therefore $\tan\theta(\phi_i, x_0) = \sqrt{N-1} = 7$. The Lanczos algorithm with full reorthogonalization was run and stopped at $n = 15$ and at $n = 18$. The special distribution of the spectrum suggests using the refined bound with $p = 1$ for the first eigenvalue and the nonrefined one $p = 0$ for the second eigenvalue. Based on these above assumption, we get the following results in Table 1 and Table 2 and when the Lanczos algorithm with full reorthogonalization was run and stopped at $n = 15$.

The second example we test is also a 50 by 50 diagonal matrix with diagonal elements

$$\lambda_1 = 1.8, \ \ \lambda_2 = 1.6, \ \ \lambda_3 = 1.4, \ \ \lambda_4 = 1.2, \ \text{and} \ \lambda_k = 1 - \frac{k-1}{N}, \ \ k = 5, \cdots, N$$

We assume that the starting vector x_0 is the vector $e = (1, 1, 1, \cdots, 1)^T$, which forms the same acute angle with each eigenvector of A. The eigenvector ϕ_i is the ith vector of the canonical basis, and therefore $\tan\theta(\phi_i, x_0) = \sqrt{N-1} = 7$. The Lanczos algorithm with full reorthogonalization was run and stopped at

Table 2: Lanczos algorithm of Example 1 stopped with $n = 18$,

i		1	2
Observed $\tan\theta(\phi_i, E_n)$		1.06e-07	2.63e-05
Bound for $\tan\theta(\phi_i, E_n)$		4.3e-07	9.15e-05
Observed $\lambda_i - \lambda_i^{(n)}$		1.97e-14	5.60e-10
Kaniel-Paige's bound for $\lambda_i - \lambda_i^{(n)}$		5.10e-13	2.01e-08
Saad's bound for $\lambda_i - \lambda_i^{(n)}$		5.10e-10	2.01e-08
New bounds for $\lambda_i - \lambda_i^{(n)}$	Ψ	9.90e-14	3.7e-09
	Θ	1.50e+00	1.25e-00
	Ω	5.10e-13	2.01e-08
Observed $\sin\theta(\phi_i, \phi_i^{(n)})$		9.3e-08	2.85e-05
Bound for $\sin\theta(\phi_i, \phi_i^{(n)})$		6.8e-07	1.4e-04

$n = 15$. Based on these above assumption, we get the following results in Table 3 when the Lanczos algorithm was run and stopped at $n = 15$.

1.8 Conclusion

Observe that the starting vector enters these bounds given in this paper through its projection on the eigenvector which is being approximated and its projection on the subspace corresponding to the first k eigenvector of A. The key component in these bound is however Chebyshev polynomial. We know that if $\frac{\gamma_j - 1}{2}(n - j) > 1$ then this polynomial grows exponentially, so that these error bounds decay as we increase n.

We can not say that a certain kind of error bound is the best. In some cases, the results are quite bad. How to improve the theoretical bounds for the rate of convergence of Lanczos method is still a open problem. Our work only have been to give more results to provide the better understanding about the convergence of the Lanczos method.

2 Convergence of the block-Lanczos method

The literature contains two basically different types of block Lanczos procedures, iterative and non-iterative. For iterative procedures see Cullum and Donath [14] and Golub and Underwood [15, 16], who replace the single vector x_0 by a system of r independent vectors $(x_1, x_2, \cdots, x_r)$. For non-iterative procedure see Lewis [17], Ruhe [18], and Scott [19], who mimic the single vector Lanczos procedures. Chain of blocks are generated, the length of the chain depends upon what one is trying to compute and upon the amount of storage available. The Lanczos blocks may or may not be reorthogonalized as they are generated. Ritz vectors may or may not be computed simultaneously with the eigenvalues. In each case

Table 3: Lanczos algorithm of Example 2 stopped with $n = 15$,

i		1	2	3
p		3	2	1
Observed $\tan\theta(\phi_i, E_n)$		4.00e-07	9.35e-06	1.17e-04
Bound for $\tan\theta(\phi_i, E_n)$		3.39e-06	7.87e-05	9.81e-04
Observed $\lambda_i - \lambda_i^{(n)}$		1.20e-13	8.64e-11	1.04e-08
Kaniel-Paige's bound for $\lambda_i - \lambda_i^{(n)}$		2.04e-11	2.01e-05	1.64e-04
Saad's bound for $\lambda_i - \lambda_i^{(n)}$		2.04e-11	9.79e-06	1.33e-06
New bounds for $\lambda_i - \lambda_i^{(n)}$	Ψ	2.40e-12	3.75e-05	6.43e-05
	Θ	6.52e-10	8.63e-09	5.44e-07
	Ω	2.04e-11	9.79e-06	1.33e-06
Observed $\sin\theta(\phi_i, \phi_i^{(n)})$		4.06e-07	9.55e-06	1.21e-04
Bound for $\sin\theta(\phi_i, \phi_i^{(n)})$		5.08e-06	1.18e-04	1.47e-03

subsets of the eigenvalues of the block tridiagonal matrices T_n generated are used as approximations to eigenvalues of A. Iterative procedures, on each iteration k, use the block recursion to generate a small projection matrix. First the relevant eigenvalues and eigenvector of these small projection matrices are computed, and then the corresponding Ritz vectors are computed and used as updated approximations to the desired eigenvectors. If on iteration k convergence has not yet been achieved, another iteration is carried out using these updated eigenvector approximations as the starting block. The iteration continue until convergence is achieved.

We have proposed the new theoretical error bounds on the rate of convergence of the Lanczos method in the previous section. Here the block generalization of Lanczos method can be treated as a system U_0 of r vectors $U_0 = (x_1, x_2, \cdots, x_r)$ instead of a single vector x_0 [14, 16, 18]. We will also follow Saad's notation suggested in [1]. Theorem 2 in [1] has shown that there is no loss of generality in assuming that the eigenvalues of A are of multiplicity not exceed r. The largest k positive eigenvalues of A under consideration will therefore be numbered in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq \lambda_{k+1}$ and $\lambda_i \leq \lambda_{k+1}$ if $j > k+1$. Each eigenvalue of A appears at most r times. The same numbering is assumed for $\phi_1, \phi_2, \cdots, \phi_N$ the associated eigenvectors of norm one. If we denote by $\pi_n(A)$ the orthogonal projection on the subspace E_n spanned by $U_0, AU_0, \cdots, A^{n-1}U_0$, then one can compute the eigenvalues $\lambda_1^{(n)} \geq \lambda_2^{(n)} \geq \cdots \geq \lambda_n^{(k)}$ of the operator $\pi_n A_{E_n} : E_n \rightarrow E_n$ with their associated eigenvectors $\phi_1^{(n)}, \phi_2^{(n)}, \cdots, \phi_k^{(n)}$ and take $\lambda_i^{(n)}, \phi_i^{(n)}$ as approximations to λ_i, ϕ_i.

It is well-known that the Lanczos and block Lanczos method have the attractive feature when n increases, the computed extreme eigenelements rapidly become good approximations to the exact ones, and are satisfactorily accurate if

n far less than N. It is very natural to ask how rapidly would the approximation eigenelements $\lambda_i^{(n)}, \phi_i^{(n)}$ converge to λ_i, ϕ_i, if exact arithmetic were performed. Golub and Underwood [15, 16] have studied the convergence of the process and obtained theoretical error bounds, generalizing Kaniel's results [8], for the s largest eigenvalues. That type of estimate they got illustrates the importance of the effective local gaps, but does not illustrate the potential positive effect of the outer loop iteration of an iterative block Lanczos procedure on reducing the overall effective spread and thereby improving the convergence rate. In this section, we will extend our new theoretical error bounds on the rate of the convergence of the Lanczos method to the block Lanczos version by using bounds on the acute angle between the exact eigenvectors and Krylov subspace spanned by $U_0, AU_0, \cdots, A^{n-1}U_0$, where $U_0 = (x_1, x_2, \cdots, x_r)$ instead of a single vector x_0. Instead of concentrating on the eigenvalues as Kaniel and Paige did, we follow the approach suggested by Saad to estimate first the angle between ϕ_i and subspace E_n. It is quite clear that $\lambda_i - \lambda_i^{(n)}$ and $\|\phi_i - \phi_i^{(n)}\|$ may be analyzed in term of $\|(I - \pi_n(A))\phi_i\|$ [20]. The number $\|(I - \pi_n(A))\phi_i\|$ is by definition the sine of the angle between ϕ_i and subspace E_n. The analysis in term of the angle between ϕ_i and subspace E_n has many advantages, as can be seen in [1]. In particular, it yields good estimates of the convergence rates for both eigenvalues and eigenvectors for block Lanczos method.

2.1 The error bounds of Golub-Underwood and Saad

We use the same notation with before. Comparing with the previous bounds, the alternative estimates of convergence can be obtained by estimating how well the eigenvalues of one of the small projection matrices T_s generated by the Lanczos tridiagonalization approximate eigenvalues of A. Saad [1] provided a variety of estimates of this type. Before recall Saad's error bounds, we restate the Golub and Underwood's results first.

Theorem 2.1 (Golub and Underwood's bounds) *Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of real symmetric matrix A. Let $\phi_1, \ldots, \phi_n$ be corresponding orthonormalized eigenvectors of A. Apply the block Lanczos recursion to A generating $s+1$ blocks and let $\lambda_1^{(n)} \geq \ldots \geq \lambda_{q(s+1)}^{(n)}$ be the eigenvalues of the small block tridiagonal matrix T_{s+1}^n generated on one iteration k. Let $Y = [Y_1^T, Y_2^T]^T = \Phi U_1^n$ be the matrix of the projections of the starting block U_1^n on the eigenvectors of A, where Y_1 is the matrix composed of the corresponding projections on the desired eigenvectors of A. Assume that σ_{min}, the smallest singular value of Y_1, is greater than 0. Then*

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq (\lambda_i - \lambda_{inf}) \frac{\tan^2 \theta}{T_s^2(\frac{1+\gamma_i}{1-\gamma_i})}, \tag{20}$$

where

$$\theta = \arccos(\sigma_{min}), \qquad \gamma_i = \frac{\lambda_i - \lambda_{q+1}}{\lambda_i - \lambda_{inf}}.$$

and the T is the Chebyshev polynomial of the first kind.

This type of estimate illustrates the importance of the effective local gaps $|\lambda_i - \lambda_{q+1}|$. but does not illustrates the potential positive effect of the outer iteration of an iterative block Lanczos procedure on reducing the overall spread and thereby improving the convergence rate.

Theorem 2.2 (Saad's Bounds) *Let λ_i be an eigenvalue of A and ϕ_i an associated eigenvector of norm one. Let assume that the vectors $\pi_1(A)\phi_j$ are linearly independent. Let vector $\hat{x}_i$ be the vector of E_1 whose orthogonal projection on the subspace spanned by $\{\phi_i, \ldots, \phi_{i+r-1}\}$ is the vector ϕ_i*

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq (\lambda_i - \lambda_{inf})\Big[\frac{K_i^{(n)}\|\phi_i - \hat{x}_i\|}{T_{n-i}(\hat{\gamma}_i)}\Big]^2. \tag{21}$$

where

$$K_i^{(n)} = \prod_{\lambda_j^{(n)} \in \sigma_i^{(n)}} \frac{\lambda_j^{(n)} - \lambda_{inf}}{\lambda_j^{(n)} - \lambda_i} \quad \textit{and} \qquad K_1^{(n)} = 1, \tag{22}$$

and

$$\hat{\gamma}_i = 1 + \frac{2(\lambda_i - \lambda_{i+r})}{(\lambda_{i+r} - \lambda_{inf})}.$$

$\sigma_i^{(n)}$ is the set of the first $i-1$ approximate eigenvalues

2.2 New theoretical bounds

In order to state the main inequality we need the following lemma

Lemma 2.3 *Let E_1 be the subspace spanned by the initial system of vectors U_0 and $\pi_1(A)$, the orthogonal projection on E_1. Let us assume that the initial system U_0 is such that vectors $\pi_1(A)\phi_i, \pi_1(A)\phi_{i+1}, \cdots, \pi_1(A)\phi_{i+r-1}$ are independent. Then there exists in E_1 a unique vector $\hat{x}_i$ such that*

$$(\hat{x}_i, \phi_j) = \delta_{ij} \qquad \textit{for} \quad j = i, i+1, \cdots, i+r-1.$$

The vector $\hat{x}_i$ defined by this lemma is the vector of E_1 whose orthogonal projection on the invariant subspace spanned by $\{\phi_i, \phi_{i+1}, \cdots, \phi_{i+r-1}\}$ is exactly ϕ_i. The rate of convergence can be studied in term of the number $\|(I - \pi_n(A))\phi_i\|/\|\pi_n(A)\phi_i\|$ which we now want to estimate. A similar approach to that of Theorem 1.2 would be quite difficult, so we have to extend the inequality of (4) to the block case. The following theorem generalizes the inequalities of Theorem 1.2 to the block Lanczos method.

Theorem 2.4 *Let λ_i be an eigenvalue of A and ϕ_i an associated eigenvector of norm one. Let us assume that the vectors $\pi_1(A)\phi_j, j = i, \cdots, i+r-1$, are linearly independent. Let $\hat{x}_i$ be the vector defined by Lemma, that is the vector of E_1 whose orthogonal projection on the subspace spanned by $\{\phi_i, \cdots, \phi_{i+r-1}\}$ is the vector ϕ_i. Let us set*

$$\hat{\gamma}_i = 1 + 2\frac{\lambda_i - \lambda_{i+r}}{\lambda_{i+r} - \lambda_{inf}}, \qquad K_i = \prod_{\lambda_j \in \sigma_i} \frac{\lambda_j - \lambda_{inf}}{\lambda_j - \lambda_i}, \quad K_1 = 1. \tag{23}$$

Then

$$\frac{\| (I - \pi_n(A))\phi_i \|}{\| \pi_n(A)\phi_i \|} \leq \frac{K_i}{T_{n-i}(\hat{\gamma}_i)} \| \phi_i - \hat{x}_i \| . \tag{24}$$

where $T_k(x)$ is the Chebyshev polynomial of the first kind of degree k and σ_i is the set of the first $i-1$ distinct eigenvalues.

About the detail of the proof, please see [1]. Here we make some remarks on this theorem. It is quite easy to know that Theorem 2.4 is a extension of Theorem 1.2. When U_0 reduces to a single vector, that is when $r = 1$, then the inequality (24) gives back its analogue (4). In a certain sense Theorem 2.4 can be viewed as an optimal extension of Theorem 1.2. Let us consider the case where $r = 2, x_1 = \sum_{j\neq 2} \alpha_j \phi_j, x_2 = \phi_2$. The block Lanczos method will provide the same approximation as with the simple Lanczos method because the second vector x_2 does not contain any more information with that contained in x_1. Any vector x in E_n can be expressed like this $x = p_1(A)x_1 + p_2(A)x_2$, where p_1 and p_2 are two polynomials of degree not exceeding $n-1$, and its angle with ϕ_1 will satisfy

$$\tan^2 \theta(\phi_1, x) = \frac{1}{\alpha_1^2 p_1^2(\lambda_1)}(p_2^2(\lambda_2) + \sum_{j\geq 3} p_1^2(\lambda_j)\alpha_j^2). \tag{25}$$

The minimum of (25) is reached when $p_2 = 0$ and when

$$\sum_{j\geq 3} \frac{\alpha_j^2}{\alpha_1^2}\frac{p_1^2(\lambda_j)}{p_1^2(\lambda_1)},$$

is minimum over all polynomials p_1 of degree less than n. This shows that in this case the process reduces to the simple Lanczos process, except that the eigenvalue λ_2 is skipped. Then the equality of (24) may be achieved by choosing a suitable sequence of λ_k and α_k. This allows us to say that when $i = 1$, the result of Theorem 2.4 is optimal in a certain sense.

Theorem 2.5 (New Block Bounds) *Let λ_i be an eigenvalue of A and ϕ_i an associated eigenvector of norm one. Let assume that the vectors $\pi_1(A)\phi_j$ are linearly independent. Let vector $\hat{x}_i$ be the vector of E_1 whose orthogonal projection on the subspace spanned by $\{\phi_i, \ldots, \phi_{i+r-1}\}$ is the vector ϕ_i*

$$0 \leq \lambda_i - \lambda_i^{(n)} \leq \min(\Psi, \Theta, \Omega), \tag{26}$$

in which

$$\Psi = (\lambda_i - \lambda_{inf}) - (\lambda_{i-1} - \lambda_{inf})\frac{T^2_{n-i}(\hat{\gamma}_{i-1})}{T^2_{n-i}(\hat{\gamma}_{i-1}) + \|\phi_{i-1} - \hat{x}_{i-1}\|^2 (K^{(n)}_{i-1})^2},$$

$$\Theta = (\lambda_i - \lambda_{inf}) - (\lambda_{i+1} - \lambda_{inf})\frac{T^2_{n-i}(\hat{\gamma}_i)}{T^2_{n-i}(\hat{\gamma}_i) + \|\phi_i - \hat{x}_i\|^2 (K^{(n)}_i)^2},$$

$$\Omega = (\lambda_i - \lambda_{inf}) - (\lambda_i - \lambda_{inf})\frac{T^2_{n-i}(\hat{\gamma}_i) - (K^{(n)}_i)^2 \|\phi_i - \hat{x}_i\|^2}{T^2_{n-i}(\hat{\gamma}_i)}.$$

Here Ω is Saad's bounds, and $\hat{\gamma}_0 = \hat{\gamma}_1$,

$$K^{(n)}_i = \prod_{\lambda^{(n)}_j \in \sigma^{(n)}_i} \frac{\lambda^{(n)}_j - \lambda_{inf}}{\lambda^{(n)}_j - \lambda_i}, \quad K^{(n)}_1 = K^{(n)}_0 = 1. \quad \hat{\gamma}_i = 1 + \frac{2(\lambda_i - \lambda_{i+r})}{(\lambda_{i+r} - \lambda_{inf})}.$$

Proof:

1. Let us prove the first part of Theorem 2.5:

Let $t_i(x)$ be the polynomial defined by

$$t_i(x) = \prod_{\lambda^{(n)}_j \in \sigma^{(n)}_i} (x - \lambda^{(n)}_j) T_{n-i}(\hat{\alpha}_i x - \hat{\beta}_i). \tag{27}$$

If $i = 1$, we take $\prod_{j=1}^{i-1}(x - \lambda^{(n)}_j) = 1$, with

$$\hat{\alpha}_i = \frac{2}{\lambda_{i+r} - \lambda_{inf}}, \qquad \hat{\beta}_i = \frac{\lambda_{i+r} + \lambda_{inf}}{\lambda_{i+r} - \lambda_{inf}}. \tag{28}$$

Then we can consider the vector $\varphi_i = t_i(A)\hat{x}_i$,

- Since degree $t_i \le n-1$ and $\hat{x}_i \in E_1$, $\varphi_i \in E_n$.
- φ_i is orthogonal to each approximate eigenvector $\phi(n)_j$ for $j \le i-1$. From (27) and (28) we can write φ_i as

  $$\varphi_i = (A - \lambda^{(n)}_j)u.$$

  where u is a vector of E_n, and

  $$(\varphi_i, \phi^{(n)}_j) = ((A - \lambda^{(n)}_j)u, \phi^{(n)}_j) = 0.$$

  We can make use of the fact that $A - \lambda^{(n)}_j I$ is self-adjoint, and that $(A - \lambda^{(n)}_j I)\phi^{(n)}_j$ is orthonormal to the subspace E_n.

- Making use of the Courant characterization of the eigenvalues of symmetric operators, we have

$$\lambda_i^{(n)} \geq \frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2}, \tag{29}$$

because

$$\lambda_i^{(n)} = \max_{u-\phi_j^{(n)}, j=1,\ldots,i-1} \frac{(Au, u)}{\|u\|^2}.$$

Let $\hat{x}_i = \sum_{j=1}^{\infty} \alpha_j \phi_j$, where the α_j are the expansion coefficients in the eigenbasis ϕ_j, Then from (29) we get

$$\begin{aligned}
\frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2} &= \frac{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2(\lambda_j - \lambda_i + \lambda_i)}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}, \\
&= \lambda_i + \frac{\sum_{j=1}^{i-1} t_i^2(\lambda_j)\alpha_j^2(\lambda_j - \lambda_i) - \sum_{j=i}^{\infty} t_i^2(\lambda_j)\alpha_j^2(\lambda_i - \lambda_j)}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}.
\end{aligned}$$

For $1 \leq j \leq i-1$, $\lambda_j - \lambda_i \geq \lambda_{i-1} - \lambda_i$ and for $j \geq i$, $\lambda_i - \lambda_j \leq \lambda_i - \lambda_{inf}$
Thus

$$\frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2} \geq \lambda_i + \frac{\sum_{j=1}^{i-1}(\lambda_{i-1} - \lambda_i)t_i^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2} - \frac{\sum_{j=i}^{\infty}(\lambda_i - \lambda_{inf})t_i^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}.$$

Extend $\lambda_{i-1} - \lambda_i$ of the second term as $\lambda_{i-1} - \lambda_{inf} + \lambda_{inf} - \lambda_i$, we know

$$\begin{aligned}
\frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2} &\geq \lambda_i - \frac{\sum_{j=1}^{\infty}(\lambda_i - \lambda_{inf})t_i^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2} + \frac{\sum_{j=1}^{i-1}(\lambda_{i-1} - \lambda_{inf})t_i^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}, \\
&\geq \lambda_{inf} + (\lambda_{i-1} - \lambda_{inf})\frac{\sum_{j=1}^{i-1} t_i^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}.
\end{aligned}$$

Since $\sum_{j=1}^{i-1} t_i^2(\lambda_j)\alpha_j^2 \geq t_i^2(\lambda_{i-1})\alpha_{i-1}^2$

$$\frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2} \geq \lambda_{inf} + (\lambda_{i-1} - \lambda_{inf})\frac{1}{1 + \frac{\sum_{j=i}^{\infty} t_i^2(\lambda_j)\alpha_j^2}{t_i^2(\lambda_{i-1})\alpha_{i-1}^2}}. \tag{30}$$

From (29), it is clear that

$$\lambda_i^{(n)} \geq \max_{\varphi_i \in E_n} \frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2}.$$

Completing the proof with the similar approach as we did in the previous section, and

$$\sum_{j \geq i} \frac{\alpha_j^2}{\alpha_{i-1}^2} \leq \|\phi_{i-1} - \hat{x}_{i-1}\|^2 = \tan^2\theta(\phi_{i-1}, \hat{x}_{i-1}).$$

We can get the following inequality

$$\lambda_i^{(n)} \geq \lambda_{inf} + (\lambda_{i-1} - \lambda_{inf}) \frac{T_{n-i}^2(\hat{\gamma}_{i-1})}{T_{n-i}^2(\hat{\gamma}_{i-1}) + \|\phi_{i-1}, \hat{x}_{i-1}\|^2 (K_{i-1}^{(n)})}.$$

Hence, we get the first part of Theorem 2.5.

2. Let us finish the second part of Theorem 2.5:

Using the similar ideas, we can get the following inequalities

$$\begin{aligned}
\frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2} &= \lambda_i + \frac{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2(\lambda_j - \lambda_i)}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}, \\
&= \lambda_i + \frac{\sum_{j=1}^{i} t_i^2(\lambda_j)\alpha_j^2(\lambda_j - \lambda_i)}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2} - \frac{\sum_{j=i+1}^{\infty} t_i^2(\lambda_j)\alpha_j^2(\lambda_i - \lambda_j)}{\sum_{j=1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}.
\end{aligned}$$

For $1 \leq j \leq i$, $\lambda_j - \lambda_i \geq \lambda_{i+1} - \lambda_i$ and for $j \geq i+1$, $\lambda_i - \lambda_j \leq \lambda_i - \lambda_{inf}$

$$\begin{aligned}
\frac{(A\varphi_i, \varphi_i)}{\|\varphi_i\|^2} &\geq \lambda_{inf} + (\lambda_{i+1} - \lambda_{inf}) \frac{1}{1 + \frac{\sum_{j=i+1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}{\sum_{j=1}^{i} t_i^2(\lambda_j)\alpha_j^2}}, \\
&\geq \lambda_{inf} + (\lambda_{i+1} - \lambda_{inf}) \frac{1}{1 + \frac{\sum_{j=i+1}^{\infty} t_i^2(\lambda_j)\alpha_j^2}{t_i^2(\lambda_i)\alpha_i^2}}.
\end{aligned}$$

To complete the proof it is sufficient to notice that

$$\forall j \geq i+1, \qquad \frac{t_i^2(\lambda_j)}{t_i^2(\lambda_i)} \leq \frac{(K_i^{(n)})^2}{T_{n-i}^2(\hat{\gamma}_i)}.$$

and that

$$\sum_{j \geq i+1} \frac{\alpha_j^2}{\alpha_{i-1}^2} \leq \|\phi_i - \hat{x}_i\|^2 = \tan^2 \theta(\phi_i, \hat{x}_i).$$

Hence, we get the second part of Theorem 2.5. ■

For the corresponding eigenvectors, When λ_i is simple, it is clear that the proof of Theorem 2.5 is valid for the block method.

Theorem 2.6 *Let λ_i be eigenvalues of A with associated eigenvector ϕ_i such that $\|\phi_i\| = 1$. Let $P_i^{(n)}$ denote the approximate eigenprojection associated with $\lambda_i^{(n)}$, and $d_{i,n} = \min_{j \neq i} |\lambda_i - \lambda_j^{(n)}|$, and $r_n = \|(I - \pi_n)A\pi_n\|$. Then*

$$\|(I - P_i^{(n)})\phi_i\| \leq \sqrt{1 + \frac{r_n^2}{d_{i,n}^2}} \|(I - \pi_n)\phi_i\|, \tag{31}$$

or

$$\sin \theta(\phi_i, \phi_i^{(n)}) \leq \|(I - P_i^{(n)})\phi_i\| \leq \sqrt{1 + \frac{r_n^2}{d_{i,n}^2}} \|(I - \pi_n)\phi_i\|. \tag{32}$$

Table 4: Lanczos algorithm of Example 1 stopped with $n = 15$,

i		1	2
Observed $\tan\theta(\phi_i, E_n)$		7.31e-08	9.44e-06
Bound for $\tan\theta(\phi_i, E_n)$		1.14e-07	2.20e-04
Observed $\lambda_i - \lambda_i^{(n)}$		1.91e-14	8.60e-11
Saad's bound for $\lambda_i - \lambda_i^{(n)}$		3.94e-14	1.27e-07
New bounds for $\lambda_i - \lambda_i^{(n)}$	Ψ	3.94e-14	1.18e-07
	Θ	5.27e-06	5.06e-06
	Ω	3.94e-14	1.27e-07

When λ_i is of multiplicity m, which does not exceed r, then that proof can be carried out with the projection $P_i^{(n)} + P_{i+1}^{(n)} + \cdots + P_{i+m+1}^{(n)}$ instead of $P_i^{(n)}$, to yield

Theorem 2.7 *Let λ_i be eigenvalues of A with associated eigenvector ϕ_i such that $\|\phi_i\| = 1$. Let $P_i^{(n)}$ denote the approximate eigenprojection associated with $\lambda_i^{(n)}$, and*
$d_{i,n} = \min_{j \neq i,\dots,i+r-1} |\lambda_i - \lambda_j^{(n)}|$, and $r_n = \|(I - \pi_n)A\pi_n\|$. Then

$$\|(I - \sum_{i}^{i+m-1} P_j^{(n)})\phi_i\| \leq \sqrt{1 + \frac{r_n^2}{d_{i,n}^2}} \|(I - \pi_n)\phi_i\|. \tag{33}$$

2.3 Numerical experiments

In this section, we compare the effective quantities $\theta(\phi_i, E_n)$, $\lambda_i - \lambda_i^{(n)}$ and with their theoretical bounds. All tests were performed on SUN workstation using double precision.

The first example is a diagonal matrix A of order $N = 70$, with the following distribution for the eigenvalues:

$$\lambda_1 = 2, \quad \lambda_2 = 1.5, \quad \lambda_k = \cos\frac{(2k-2)\pi}{2(N-2)}, \quad \text{for} \quad k = 3, \cdots, N$$

We assume that the starting system $U_0 = (x_1, x_2)$ was chosen as follows. Let

$$e = (1, 1, 1, \cdots, 1)^T, \qquad g = (1, -1, 1, -1, \cdots, 1, -1)^T,$$

Then

$$x_1 = \frac{e}{\|e\|}, \qquad x_2 = \frac{g}{\|g\|}.$$

We assume that we use two dimensional blocks or $r = 2$. The eigenvector ϕ_i is the ith vector of the canonical basis, and therefore $\tan\theta(\phi_i, x_0) = \sqrt{N-1}$. The

Table 5: Lanczos algorithm of Example 2 stopped with $n = 12$,

i		1	2	3
Observed $\tan\theta(\phi_i, E_n)$		4.66e-07	4.11e-06	2.74e-05
Bound for $\tan\theta(\phi_i, E_n)$		7.40e-07	4.11e-04	2.33e-02
Observed $\lambda_i - \lambda_i^{(n)}$		3.14e-13	1.60e-11	5.54e-10
Saad's bound for $\lambda_i - \lambda_i^{(n)}$		1.06e-12	2.62e-07	7.38e-04
New bounds for $\lambda_i - \lambda_i^{(n)}$	Ψ	1.06e-12	2.55e-07	2.33e-05
	Θ	2.45e-13	4.63e-09	6.34e-07
	Ω	1.06e-12	2.62e-07	7.38e-04

Lanczos algorithm with full reorthogonalization was run and stopped at $n = 15$. Based on these above assumption, we get the following results in Table 4 and when the Lanczos algorithm with full reorthogonalization was run and stopped at $n = 15$.

The second example we test is a three dimensional blocks. A is of order $N = 60$, with eigenvalues

$$\lambda_1 = 2, \;\; \lambda_2 = 1.6, \;\; \lambda_3 = 1.4, \;\; \lambda_4 = 1, \;\; \text{and} \;\; \lambda_k = 1 - \frac{k-3}{N}, \;\; k = 5, \cdots, N$$

We assume that the starting system $U_0 = (x_1, x_2, x_3)$. Let

$$f = (1, 0, -1, 1, 0, -1, \cdots, 1, 0, -1)^T, \quad g = (1, -2, 1, 1, -2, 1, \cdots, 1, -2, 1)^T,$$

Then

$$x_1 = \frac{e}{||e||}, \qquad x_2 = \frac{f}{||f||}, \qquad x_3 = \frac{g}{||g||}.$$

The eigenvector ϕ_i is the ith vector of the canonical basis, and therefore $\tan\theta(\phi_i, x_0) = \sqrt{N-1}$. The Lanczos algorithm with full reorthogonalization was run and stopped at $n = 12$. Based on these above assumption, we get the following results in Table 5 when the Lanczos algorithm was run and stopped at $n = 12$.

2.4 Conclusion

We have established new theoretical error bounds for block Lanczos method by using bounds on the acute angle between the exact eigenvectors and Krylov subspace spanned by $U_0, AU_0, \cdots, A^{n-1}U_0$, where $U_0 = (x_1, \cdots, x_r)$ instead of a single vector x_0. From previous sections, it is easy to show that the bounds on the rate of the block version are superior to those of the single vector processor.

As Underwood said in [21], there are many other possibilities to obtain a priori bounds by using vectors of E_n of the form $P_n(A)\hat{x}$, where P_n is a polynomial of degree at most $n-1$, and where $\hat{x}$ is a vector of E_1. We can not say that a certain kind of error bound is the best. In some cases, the results are quite bad. Our work only have been to give more results to provide the better understanding about the convergence of the Lanczos method. It is also possible to derive refined error bounds for the block Lanczos method similar to those of Lanczos method.

References

[1] Y. Saad. On the rates of convergence of the Lanczos and block-Lanczos methods. *SIAM Journal on Numerical Analysis*, 17:687–706, 1980.

[2] C. Lanczos. An iteration method for the solution of the eigenvalues problem of linear differential and integral operators. *Journal of Research of National Bureau of Standards*, 45:255–282, 1950.

[3] C. C. Paige. Computational variants of the Lanczos method for the eigenproblem. *Journal of the Institute of Mathematics and its Applications*, pages 373–381, 1972.

[4] W. Kahan and B. N. Parlett. How far should you go with the Lanczos process. In J. Bunch and D. J. Rose, editors, *Sparse Matrix Computations*, pages 131–144. Academic Press, New York, 1976.

[5] C. C. Paige. Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix. *Journal of the Institute of Mathematics and its Applications*, 18:341–349, 1976.

[6] D. S. Scott. Analysis of the symmetric Lanczos process. Technical Report M78/40, University of California, Berkeley, 1978.

[7] J. H. Wilkinson. *The Algebraic Eigenvalue Problems*. Clarendon Press, Oxford, England, 1965.

[8] S. Kaniel. Estimates for some computational techniques in linear algebra. *Mathematics of Computation*, 20:369–378, 1966.

[9] C. C. Paige. *The computation of eigenvalues and eigenvectors of very large sparse matrices*. PhD thesis, University of London, 1971.

[10] Y. Saad. Calcul de valeurs propres de grandes matrices hermitiennes par des techniques de partitionnement. Master's thesis, University of Grenoble, 1974.

[11] B. N. Parlett. *The Symmetric Eigenvalue Problem*. Prentice Hall, Englewood Cliffs, New York, 1980.

[12] J. K. Cullum and R. A. Willoughby. *Lanczos Algorithm for Large Symmetric Eigenvalue Computation*, volume 1 of *Scientific Computing*. Birkhäuser, Boston, 1985.

[13] J. S. Vandergraft. Generalized Rayleigh method with application to finding eigenvalues of large matrices. *Linear Algebra and its Applications*, 4:353–368, 1971.

[14] J. K. Cullum and W. E. Donath. A block Lanczos algorithm for computing the q algebraically largest eigenvalues and a corresponding eigenspace of large, sparse, real symmetric matrices. In *Proceedings of IEEE Conference on Decision and Control*, Phoenix, Arizona, 1974.

[15] G. H. Golub and R. Underwood. Methods for computing eigenvalues of sparse matrix equation. In Pereyra V., editor, *Actas Del Seminario Sombre Methods Numericos Modernos*, volume 1, pages 125–148. University Central de Venezuela, 1975.

[16] G. H. Golub and R. Underwood. The block Lanczos method for computing eigenvalues. In J. Rice, editor, *Mathematical Software*, volume 3, pages 364–377. Academic Press, New York, 1977.

[17] J. Lewis. Algorithm for sparse matrix eigenvalue problems. Technical Report CS-77-595, Department of Computer Science, Stanford University, 1977.

[18] A. Ruhe. Implementation aspects of band Lanczos algorithm for computation of eigenvalue of large sparse matrices. *Mathematics of Computation*, 33:680–687, 1979.

[19] D. S. Scott. Block Lanczos software for symmetric eigenvalue problems. Technical Report ORNL-CSD-48, Oak Ridge National Laboratory, Union Carbide Corporation, Oak Ridge, Tennessee, 1979.

[20] F. Chatelin and J. Lemordant. La method de Rayleigh-Ritz appliquee a des operateurs differentiels elloptiques. *Numerische Mathematik*, 23:215–222, 1975.

[21] R. R. Underwood. *An iterative block Lanczos method for the solution of large sparse symmetric eigenproblem*. CS-75-496, Department of Computer Science, Stanford University, Stanford, CA, 1975.