Journal of
Linear and Topological Algebra
Vol. 01, No. 02, 2012, 91- 104

A New Inexact Inverse Subspace Iteration for Generalized Eigenvalue Problems

M. Amirfakhrian[a] and F. Mohammad[a,*]

[a]*Department of Mathematics, Islamic Azad University, Central Tehran Branch, PO. Code 14168-94351, Iran.*

Abstract. In this paper, we represent an inexact inverse subspace iteration method for computing a few eigenpairs of the generalized eigenvalue problem $Ax = \lambda Bx$[Q. Ye and P. Zhang, Inexact inverse subspace iteration for generalized eigenvalue problems, Linear Algebra and its Application, 434 (2011) 1697-1715]. In particular, the linear convergence property of the inverse subspace iteration is preserved.

Keywords: Eigenvalue problem; inexact inverse iteration; subspace iteration; inner-outer iteration; approximation; convergence.

1. Introduction

We want to compute a few eigenpairs of the generalized eigenvalue problem $Ax = \lambda Bx$. (1) The eigenvalues sought may be those in the extreme part of the spectrum or in the interior of the spectrum near certain given point. These types of problems arise in many engineering and scientific applications. In such applications, the matrices involved are often large and sparse. A brief description of some developed method is given in

[2] . If the extreme eigenvalues are not well-separated or if the eigenvalues sought are in the interior of the spectrum, a shift-and invert (or inverse) transformation is combined with one of the eigenproblem solvers to speed up the convergence. The use of the shift-and-invert transformation requires solving a linear system at each iterative step. Solving the linear systems by a direct method such as the QR factorization can be impractical or expensive if the dimension of the matrices is large or the matrices are not explicitly available. Alternatively, iterative method can be employed, which will be called the inner iteration while the original iterative algorithm will be called the outer iteration. The inner iteration produces, a situation that may arise in other applications as well. The question then is how the accuracy in the inner iterations affects the convergence behavior (or convergence speed) of the outer iteration as compared with the exact case. Related to this, a challenging problem in implementations is how to efficiently choose an appropriate stopping threshold for the inner iteration so that the convergence characteristic of the outer iteration can be preserved. This is a problem that has been discussed for several methods, such as inexact Krylov subspace method [6, 10, 17] inexact

*Corresponding author. Email: f.mohammad456@yahoo.com

http://jlta.iauctb.ac.ir

inverse iteration, [1, 4, 8, 12, 13], the rational Arnoldi algorithm [9, 11, 16], inexact Rayleigh Quotient-Type methods [3, 7, 14] and the Jacobi-Davidson method [15]. While these works demonstrate that the innerouter iteration technique may be an effective way for implementing some of these methods for solving the large scale eigenvalue problem, the more efficient Krylov subspace projection methods tend to require quite accurate matrix-vector products to preserve their convergence characteristic.

2. Inexact inverse subspace iteration

We consider computing the p smallest eigenvalues of the generalized eigenvalue problem $Ax = \lambda Bx$ by applying the standard subspace iteration to $A^{-1}B$, called inverse subspace iteration. we state the standard algorithm as follows.

Algorithm 1 . Inverse subspace iteration for $Ax = \lambda Bx$

(1) Input: $X_0 \in C^{n\times p}$ with $X_0^* X_0 = I$;
(2) For $k = 0, 1, \cdots$ until convergence
(3) $Y_{k+1} = A^{-1}BX_k$;
(4) $Y_{k+1} = X_{k+1}R_{k+1}$ (QR-factorization).
(5) *End.*

We try to use this algorithm when a problems where direct solution of A^{-1} is impractical or inefficient because A is too large or is not explicitly available. In these cases, an iterative method can be used to solve the linear systems $AY_{k+1} = BX_k$, called the inner iterations, while the subspace iteration itself is called the outer iteration. At each step of outer iteration, to solve Y_{k+1} for $AY_{k+1} = BX_k$, the previous iterate Y_k can be used as an initial approximation. Then we solve

$$AD_k = BX_k - AY_k \tag{1}$$

approximately, *i.e.* in this case we find D_k such that

$$\| E_k \|_2 = \| (BX_k - AY_k) - AD_k \|_2 < \epsilon_k, \tag{2}$$

where $E_k := (BX_k - AY_k) - AD_k$ and k is some given threshold. From D_k, then $Y_{k+1} = Y_k + D_k$.

Now we try to analyze the convergence characteristic of the subspace iteration under inexact solves.

Obviously, the amount of work required to solve (2) is proportional to $\| BX_k - AY_k \| /\epsilon_k$. Our analysis later leads to the use of linearly decreasing k, let $\epsilon_k = ar^k$ for some positive a and $r < 1$. However, ϵ_k is decreasing, the amount of work does not increase as $BX_k - AY_k$ will be decreasing at the same rate as well. Thus, the stopping threshold required for inner iterations is effectively a constant.

Algorithm 2. Inexact inverse subspace iteration for Ax = Bx

(1) Input: $X_0 \in C^{n\times p}$ with $X_0^* X_0 = I$; threshold parameter ϵ_k; set $Y_0 = 0$;
(2) For $k = 0, 1, ...$ until convergence
(3) $Z_k = BX_k - AY_k$;
(4) Solve $AD_k = Z_k$ such that $E_k = Z_k - AD_k$ satisfies (3);
(5) $Y_{k+1} = Y_k + D_k$;
(6) $Y_{k+1} = \bar{X}_{k+1}\bar{R}_{k+1}$ (QR-factorization);
(7) For $j = 1, ..., p$
(8) $[y_{\max}, i_{\max}] = \max(|\bar{X}_{k+1}(:, j)|)$;

(9) $X_{k+1}(:,j) = sign(\bar{X}_{k+1}(i_{\max},j)) * X_{k+1}(:,j);$
(10) $R_{k+1}(j,:) = sign(\bar{X}_{k+1}(i_{\max},j)) * R_{k+1}(j,:).$
(11) End.
(12) End.

For more details please refer to [4]

3. Convergence analysis

We discuss convergence of the subspace spanned by X_k for the inexact inverse subspace algorithm. Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the eigenvalues of $B^{-1}A$ ordered such that

$$0 <\mid \lambda_1 \mid \leqslant \cdots \leqslant \mid \lambda_p \mid < \mid \lambda_{p+1} \mid \leqslant \cdots \leqslant \mid \lambda_q \mid < \mid \lambda_{q+1} \mid \leqslant \cdots \mid \lambda_n \mid$$

and $v_1, ..., v_n$ be the corresponding eigenvectors. Suppose that we want to compute the p smallest eigenpairs in absolute value, $i.e.$,$\lambda_1, \lambda_2, \cdots, \lambda_p$, $\lambda_{p+1}, \cdots, \lambda_q$ .We are following eigenvalues of $p+1$ to q(with assumption ordered) so we have made the assumptions that $\rho :=\mid \lambda_p \mid / \mid \lambda_{p+1} \mid < 1$, $\acute{\rho} :=\mid \lambda_q \mid / \mid \lambda_{q+1} \mid < 1$.

Throughout this work, we assume that $B^{-1}A$ is diagonalizable. Let $V = [v_1, ..., v_n]$, $U = (BV)^{-H}$,then

$U^H A = \Lambda U^H B$, $AV = BV\Lambda$ where $\Lambda = \begin{pmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{pmatrix}$

$\Lambda_1 = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_p \end{pmatrix}$, $\Lambda_2 = \begin{pmatrix} \lambda_{p+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_q \end{pmatrix}$, $\Lambda_1 = \begin{pmatrix} \lambda_{q+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$

let $U = (U_1, U_2, U_3)$, $V = (V_1, V_2, V_3)$, where $U_1 \in {}^{n\times p}$, $U_2 \in {}^{n\times(p+q)}$, $U_3 \in {}^{n\times(n-(p+q)}$, $V_1 \in {}^{n\times p}$, $V_2 \in {}^{n\times(p+q)}$ $V_3 \in {}^{n\times(n-(p+q)}$, then $U_i^H A = \Lambda_i U_i^H B$, $AV_i = BV_i\Lambda_i$, Consider Algorithm 2 now. Define $X_k^{(i)} = U_i^H BX_k$. Since $U_i^H BV_j = \delta_{ij} I$ and $U_i^H AV_j = \delta_{ij}\Lambda_i$ where δ_{ij} is the Kronecker symbol, then

$$X_k = \sum_{i=1}^{3} V_i X_k^{(i)}$$

If $X_k^{(1)}$ is invertible, we define $t_k :=\parallel X_k^{(2)}(X_k^{(1)})^{-1} \parallel_2$ and $t_{k'} :=\parallel X_k^{(3)}(X_k^{(1)})^{-1} \parallel_2$ Clearly, t_k and $t_{k'}$ is a measures of the approximation of the column space of X_k to the column space of V_1. Indeed, the following proposition relates t_k and $t_{k'}$ to other measures of subspace approximation.

PROPOSITION 3.1 *Assume that $X_k^{(1)}$ is invertible and t_k and $t_{k'}$ is defined as above. Then*

$$\frac{\begin{pmatrix} t_k \\ t_{k'} \end{pmatrix}}{\parallel V^{-1} \parallel_2} \leqslant \parallel X_k(X_k^{(1)})^{-1} - V_1 \parallel_2 \leqslant \parallel V \parallel_2 (t_k + t_{k'}) \tag{3}$$

and

$$\sin \angle(\chi_k, \nu_1) \leqslant \parallel V \parallel_2 \parallel R^{-1} \parallel_2 (t_k + t_{k'})$$

where $\angle(X_k, V_1)$ *is the largest canonical angle between* $X_k = R(X_k)$ *and* $V_1 = R(V_1)$*, and* R *is defined by the QR-factorization* $V_1 = WR$ *of* V_1*.*

Proof From $X_k = V_1X_k^{(1)} + V_2X_k^{(2)} + V_3X^{(3)}$, we have

$$X_k(X_k^{(1)})^{-1} = V_1 + V_2X_k^{(2)}(X_k^{(1)})^{-1} + V_2X_k^{(3)}(X_k^{(1)})^{-1}$$

$$\| X_k(X_k^{(1)})^{-1} - V_1 \|_2 = \| V_2X_k^{(2)}(X_k^{(1)})^{-1} + V_3X_k^{(3)}(X_k^{(1)})^{-1} \|_2$$

$$\leqslant \| V_2 \|_2 \| X_k^{(2)}(X_k^{(1)})^{-1} + \| V_3 \|_2 \| X_k^{(3)}(X_k^{(1)})^{-1} \|_2 \leqslant \| V_2 \|_2 \ (t_k + t_{k'})$$

and

$$\| V_2X_k^{(2)}(X_k^{(1)})^{-1} + V_3X_k^{(3)}(X_k^{(1)})^{-1} \|_2 = \| V \begin{pmatrix} X_k^{(2)}(X_k^{(1)})^{-1} \\ X_k^{(3)}(X_k^{(1)})^{-1} \end{pmatrix} \|_2 \geqslant \frac{\begin{pmatrix} t_k \\ t_{k'} \end{pmatrix}}{\| V^{-1} \|_2}$$

(4) is proved.

Let $X_k^{\perp}$ be such that $(X_k, X_k^{\perp})$ is an $n \times n$ orthogonal matrix. Then the sine of the largest canonical angle between $\chi_k = R(X_k)$and $\nu_1 = R(V_1)$ is (see [5] for the definition)

$$\sin\angle(\chi_k, \nu_1) = \| (X_k^{\perp})^H w \|_2$$

$$= \| (X_k^{\perp})^H (w - X_k(X_k^{(1)})^{-1}R^{-1}) \|_2$$

$$= \| (X_k^{\perp})^H (V_2X_k^{(2)}(X_k^{(1)})^{-1}R^{-1} + V_3X_k^{(3)}(X_k^{(1)})^{-1}R^{-1}) \|_2$$

$$\leqslant \| V \|_2 \| R^{-1} \|_2 \ (t_k + t_{k'}).$$

■

It is clear that t_k and $t_{k'}$ are measures of the approximation of the column space of X_k. We shall next discuss the convergence of t_k and $t_{k'}$.

Lemma 3.2 *For Algorithm 2* $\| E_k \|_2 < \| B^{-1} \|_2^{-1}$ *then* Y_{k+1} *has full column rank.*

Proof From the algorithm, we have $AY_{k+1} = BX_k + E_k$. Therefore

$$X_k^H B^{-1}AY_{k+1} = I + X_k^H B^{-1}E_k$$

Since

$$\| X_k^H \|_2 = 1$$

$$\| X_k^H B^{-1}E_k \|_2 \leqslant \| B^{-1} \|_2 \| E_k \|_2 < 1$$

$X_k^H B^{-1}AY_{k+1}$ is invertible. Thus Y_{k+1} has full column rank. ■

From now on, we shall assume that $\epsilon_k \leqslant \| B^{-1} \|_2^{-1}$, so that all Y_k will have full column rank and Algorithm 2 will be well defined.

Lemma 3.3 *For Algorithm 2, if X_k is invertible, then*

$$\| (X_k^{(1)})^{-1} \|_2 \leqslant \| V \|_2 \, (1 + t_k + t_{k'})$$

Proof Proof. Since X_k has orthonormal columns, we have

$$X_k = V_1 X_k^{(1)} + V_2 X_k^{(2)} + V_3 X_k^{(3)}$$

from which it follows that $(X_k^{(1)})^{-1}$

$$X_k (X_k^{(1)})^{-1} = V_1 + V_2 X_k^{(2)} (X_k^{(1)})^{-1} + V_3 X_k^{(3)} (X_k^{(1)})^{-1}$$

since X_k have orthonormal columns, we have from property of orthonormal matrix (if Q is a orthonormal matrix then for every x we have $\| Qx \|_2 = \| x \|_2$) let

$$\| (X_k^{(1)})^{-1} \|_2 = \| X_k (X_k^{(1)})^{-1} \|_2$$

we have the following

$$\| (X_k^{(1)})^{-1} \|_2 = \| V_1 + V_2 X_k (X_k^{(1)})^{-1} + V_3 X_k^{(3)} (X_k^{(1)})^{-1} \|_2 \leqslant \| V_1 \|_2 + t_k \| V_2 \|_2 + t_{k'} \| V_3 \|_2$$

$$\leqslant (1 + t_k + t_{k'}) \| V \|_2$$

■

Lemma 3.4
Let rho $= | \lambda_p | / | \lambda_{p+1} | < 1$ *and assume that* $X_k^{(1)}$ *and* $X_{k+1}^{(1)}$ *are nonsingular. If* $\| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'}) \epsilon_k < 1$, *then*

$$t_{k+1} \leqslant \rho (t_k + t_{k'}) + \frac{\rho \| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'})^2 \epsilon_k}{1 - \| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'}) \epsilon_k}$$

Proof From the algorithm we know that $AY_{k+1} = BX_k + E_k$ and $Y_{k+1} = X_{k+1} R_{k+1}$. Since Y_{k+1} has full column rank, R_{k+1} is invertible. Then

$$AX_{k+1} R_{k+1} R_{k+1}^{(-1)} = BX_k R_{k+1}^{(-1)} + E_k R_{k+1}^{(-1)}$$

Multiplying U_i^H on the equation above, we have

$$\begin{cases} X_k^{(i)} = U_i^H B X_k \\ U_i^H A = \Delta_i U_i^H B \end{cases}$$

$$U_i^H A X_{k+1} = U_i^H B X_k R_{k+1}^{(-1)} + U_i^H E_k R_{k+1}^{(-1)}$$

$$\Lambda_i U_i^H B X_{k+1} = X_k^{(i)} R_{k+1}^{(-1)} + U_i^H E_i R_{k+1}^{(-1)}$$

where

$$U_i^H B X_{k+1} = \Lambda_i^{(-1)} X_k^{(i)} R_{k+1}^{(-1)} + \Lambda_i^{(-1)} U_i^H E_k R_{k+1}^{(-1)}$$

$$X_{k+1}^{(i)} = \Lambda_i^{(-1)} X_k^{(i)} R_{k+1}^{(-1)} + \Lambda_i^{(-1)} \Delta_k^{(i)} \tag{4}$$

$\Delta_k^{(i)} = U_i^H E_k R_{k+1}^{(-1)}$. let $i = 2$

$$\begin{aligned}
X_{k+1}^{(2)} &= \Lambda_2^{(-1)} X_k^{(2)} R_{k+1}^{(-1)} + \Lambda_2^{(-1)} \Delta_k^{(2)} \\
X_{k+1}^{(2)} (X_{k+1}^{(1)})^{(-1)} &= (\Lambda_2^{(-1)} X_k^{(2)} R_{k+1}^{(-1)} + \Lambda_2^{(-1)} \Delta_k^{(2)})(X_{k+1})^{(-1)} \\
&= (\Lambda_2^{(-1)} X_k^{(2)} R_{k+1}^{(-1)})(X_{k+1})^{(-1)} + \Lambda_2^{(-1)} \Delta_k^{(2)} (X_{k+1})^{(-1)} \\
&= \Lambda_2^{(-1)} X_k^{(2)} (X_k^{(1)})^{(-1)} \Lambda_1)(((X_k)^{(1)}) \Lambda_1)^{-1} R_{k+1}^{(-1)} (X_{k+1})^{(-1)} \\
&+ \Lambda_2^{(-1)} \Delta_k^{(2)} (X_{k+1})^{(-1)} \\
&= \Lambda_2^{(-1)} X_k^{(2)} (X_k^{(1)})^{(-1)} \Lambda_1) \Lambda_1^{-1} ((X_k)^{(1)})^{-1} R_{k+1}^{(-1)} (X_{k+1})^{(-1)} \\
&+ \Lambda_2^{(-1)} \Delta_k^{(2)} (X_{k+1})^{(-1)}
\end{aligned}$$

and now in the (4) let $i = 1$

$$X_{k+1}^{(1)} = \Lambda_1^{(-1)} X_k^{(1)} R_{k+1}^{(-1)} + \Lambda_1^{(-1)} \Delta_k^{(1)} \tag{5}$$

$$\Lambda_1^{(-1)} X_k^{(1)} R_{k+1}^{(-1)} = X_{k+1}^{(1)} - \Lambda_1^{(-1)} \Delta_k^{(1)}$$

$$X_{k+1}^{(2)}(X_{k+1}^{(1)})^{(-1)} = \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Lambda_1(X_{k+1}^{(1)} - \Lambda_1^{(-1)}\Delta_k^{(1)})(X_{k+1})^{(-1)}$$

$$+ \Lambda_2^{(-1)}\Delta_k^{(2)}(X_{k+1})^{(-1)}$$

$$= \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Lambda_1 X_{k+1}^{(1)}(X_{k+1})^{(-1)}$$

$$- \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Lambda_1\Lambda_1^{(-1)}\Delta_k^{(1)}(X_{k+1})^{(-1)} + \Lambda_2^{(-1)}\Delta_k^{(2)}(X_{k+1})^{(-1)}$$

$$= \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Lambda_1 - \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Delta_k^{(1)})(X_{k+1})^{(-1)}$$

$$+ \Lambda_2^{(-1)}\Delta_k^{(2)}(X_{k+1})^{(-1)}$$

$$= \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Lambda_1$$

$$- (X_{k+1})^{(-1)}(\Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Delta_k^{(1)} - \Lambda_2^{(-1)}\Delta_k^{(2)})$$

and now let in the (5)

$$= \Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Lambda_1 - (\Lambda_2^{(-1)}X_k^{(2)}((X_k)^{(1)})^{(-1)}\Delta_k^{(1)}$$

$$- \Lambda_2^{(-1)}\Delta_k^{(2)})(\Lambda_1^{(-1)}X_k^{(1)}R_{k+1}^{(-1)} + \Lambda_1^{(-1)}\Delta_k^{(1)})^{-1}$$

since $\Delta_k^{(i)} = U_i^H E_k R_{k+1}^{(-1)}$ and

$$(X_{k+1}^{(1)})^{-1} = (\Lambda_1^{(-1)}X_k^{(1)}R_{k+1}^{(-1)} + \Lambda_1^{(-1)}\Delta_k^{(1)})^{-1} = (\Lambda_1^{(-1)}(X_k^{(1)}R_{k+1}^{(-1)} + \Delta_k^{(1)}))^{-1}$$

$$= (X_k^{(1)}R_{k+1}^{(-1)}(I + \Delta_k^{(1)}R_{k+1}(X_k^{(1)})^{-1})^{-1}\Lambda_1$$

$$= R_{k+1}(X_k^{(1)})^{-1}(I + \Delta_k^{(1)}R_{k+1}(X_k^{(1)})^{-1})^{-1}\Lambda_1$$

Then we further simplify the expression $X_{k+1}^{(2)}(X_{k+1}^{(1)})^{-1}$ to

$$X_{k+1}^{(2)}(X_{k+1}^{(1)})^{-1} = \Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}\Lambda_1$$

$$- \Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}\Delta_k^{(1)}(X_{k+1}^{(1)})^{-1} + \Lambda_2^{-1}\Delta_k^{(2)}(X_{k+1}^{(1)})^{-1}$$

$$= \Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}\Lambda_1 - \Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}U_1^H E_k R_{k+1}^{-1}(X_{k+1}^{(1)})^{-1}$$

$$+ \Lambda_2^{-1}U_2^H E_k R_{k+1}^{-1}(X_{k+1}^{(1)})^{-1}$$

$$= \Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}\Lambda_1 - (\Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}U_1^H - \Lambda_2^{-1}U_2^H)E_k R_{k+1}^{-1}(X_{k+1}^{(1)})^{-1}$$

$$(X_{k+1}^{(1)})^{-1} = (\Lambda_1^{-1}X_k^{(1)}R_{k+1}^{-1} + \Lambda_1^{-1}\Delta_k^{(1)})^{-1} = (\Lambda_1^{-1}X_k^{(1)}R_{k+1}^{-1} + \Lambda_1^{-1}U_1^H E_k R_{k+1}^{-1})^{-1}$$

$$= (\Lambda_1^{-1}(X_k^{(1)})R_{k+1}^{-1})^{-1}(I + U_1^H E_k(X_k^{(1)})^{-1})^{-1}$$

$$= R_{k+1}(X_k^{(1)})^{-1}(I + U_1^H E_k(X_k^{(1)})^{-1})^{-1}\Lambda_1$$

$$X_{k+1}^{(2)}(X_{k+1}^{(1)})^{(-1)} = \Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}\Lambda_1$$

$$- (\Lambda_2^{-1}X_k^{(2)}(X_k^{(1)})^{-1}U_1^H - \Lambda_2^{-1}U_2^H)E_k(X_k^{(1)})^{-1}(I + U_1^H E_k(X_k^{(1)})^{-1})^{-1}\Lambda_1$$

Taking 2-norm of the above equation at both sides and using the condition

$$\| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'})\epsilon_k < 1, \tag{6}$$

we obtain the following upper bound of t_{k+1}:

$$\| X_{k+1}^{(2)}(X_{k+1}^{(1)})^{(-1)} \| \leqslant \| \Lambda_2^{-1} \|_2 \| X_k^{(2)}(X_k^{(1)})^{-1} \|_2 \| \Lambda_1 \|_2$$

$$- (\| \Lambda_2^{-1} \|_2 \| X_k^{(2)}(X_k^{(1)})^{-1} \|_2 \| U_1^H \|_2$$

$$- \| \Lambda_2^{-1} \|_2 \| U_2^H \|_2) \| E_k \|_2 \| (X_k^{(1)})^{-1} \|_2 \| (I + U_1^H E_k(X_k^{(1)})^{-1})^{-1} \|_2 \| \Lambda_1 \|_2$$

$$= \| \Lambda_2^{-1} \|_2 \| \Lambda_1 \|_2 \, t_k$$
$$+ (\| \Lambda_2^{-1} \|_2 \, t_k \, \| U_1^H \|_2$$
$$+ \| \Lambda_2^{-1} \|_2 \| U_2^H \|_2) \| E_k \|_2 \| (X_k^{(1)})^{-1} \|_2 \| (I + U_1^H E_k(X_k^{(1)})^{-1})^{-1} \|_2 \| \Lambda_1 \|_2$$

$$X_{k+1}^{(3)}(X_{k+1}^{(1)})^{(-1)} = \| \Lambda_2^{-1} \|_2 \| \Lambda_1 \|_2 \, t_{k'} + (\| \Lambda_2^{-1} \|_2 \, t_{k'} \| U_1^H \|_2 + \| \Lambda_2^{-1} \|_2 \| U_2^H \|_2) \| E_k \|_2 \| (X_k^{(1)})^{-1} \|_2 \| (I + U_1^H E_k (X_k^{(1)})^{-1})^{-1} \|_2 \| \Lambda_1 \|_2$$

$$t_{k+1} \leqslant \rho(t_k + t_{k'}) + ((\| \Lambda_2^{-1} \|_2 \, t_k + \| \Lambda_2^{-1} \|_2 + \| \Lambda_2^{-1} \|_2 \, t_{k'} + \| \Lambda_2^{-1} \|_2)$$

$$\times \| U \|_2 \| E_k \|_2 \| (X_k^{(1)})^{-1} \|_2 \| \Lambda_1 \|_2 \| (I + U_1^H E_k (X_k^{(1)})^{-1})^{-1} \|_2$$

$$\leqslant \rho(t_k + t_{k'}) + (1 + t_k + t_{k'}) \| \Lambda_1^{-1} \|_2 \frac{\| U \|_2 \| E_k \|_2 \| (X_k^{(1)})^{-1} \|_2 \| \Lambda_1 \|_2}{1 - \| E_k \|_2 \| (X_k^{(1)})^{-1} \|_2 \| U \|_2}$$

Using lemma (3.3), we know that $X_k^{(1)} \leqslant \| v \|_2 \, (1 + t_k + t_{k'})$. From this and (3) , the final bound for t_{k+1} is derived. ■

Lemma 3.5
Assume that X_0 is such that $X_0^{(1)}$ is invertible. If

$$\epsilon_k \leqslant \epsilon := \frac{(1-\rho)2t_0}{\| V \|_2 \| U \|_2 \, (1+2t_0)(\rho + 2t_0)}$$

for all k, then $t_k \leqslant t_0$.

Proof We prove $t_k \leqslant t_0$, $t_{k'} \leqslant t_0$ by induction. Supposing $X_k^{(1)}$ is nonsingular and $t_k \leqslant t_0$, $t_{k'} \leqslant t_0$ is true for some k, k', we show that $X_{k+1}^{(1)}$ is nonsingular and $t_{k+1} \leqslant 2t_0$. First note that from $\epsilon_k \leqslant \epsilon$, we have

$$\| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'})\epsilon_k \leqslant \| V \|_2 \| U \|_2 \, (1 + t_0 + t_{k'})\epsilon = \frac{(1+\rho)2t_0}{\rho + 2t_0} < 1$$

We discuss in two cases:

- Case I

$X_{k+1}^{(1)}$ is nonsingular. Then by lemma (3.4), we have

$$t_{k+1} \leqslant \rho(t_k + t_{k'}) + \frac{\rho \| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'})^2 \epsilon_k}{1 - \| V \|_2 \| U \|_2 \, (1 + t_k + t_{k'})\epsilon_k}$$

$$\leqslant \rho 2t_0 + \frac{\rho \| V \|_2 \| U \|_2 \, (1 + 2t_0)^2 \epsilon}{1 - \| V \|_2 \| U \|_2 \, (1 + 2t_0)\epsilon}$$

$$\leqslant \rho 2t_0 + \frac{\rho(1+2t_0)\frac{(1-\rho)2t_0}{\rho+2t_0}}{1 - \frac{(1-\rho)2t_0}{\rho+2t_0}} \leqslant \rho 2t_0 + \frac{\frac{\rho(1+2t_0)(1-\rho)2t_0}{\rho+2t_0}}{\frac{\rho+2t_0-2t_0+\rho 2t_0}{\rho+2t_0}}$$

$$\leqslant \rho 2t_0 + \frac{\rho(1+2t_0)(1-\rho)2t_0}{\rho(1+2t_0)} = 2t_0,$$

- Case II

$X^{(1)}_{k+1}$ is singular. Then let

$$\tilde{y}_{k+1} = y_{k+1} + \delta V_1 R_{k+1} + \mu V_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} R_{k+1} + \theta V_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} R_{k+1}$$

where $Y_{k+1} = X_{k+1} R_{k+1}$ and $\delta,\mu > 0$ are two parameters. Then we have

$$\begin{aligned}
A\tilde{y_{k+1}} &= Ay_{k+1} + \delta AV_1 R_{k+1} + \mu AV_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} R_{k+1} + \theta V_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} R_{k+1} \\
A\tilde{y_{k+1}} &= BX_k + E_k + \delta AV_1 R_{k+1} + \mu AV_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} R_{k+1} + \theta V_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} R_{k+1} \\
E_k + \delta AV_1 R_{k+1} &+ \mu AV_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} R_{k+1} + \theta V_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} R_{k+1} = \tilde{E}_k \\
A\tilde{y_{k+1}} &= BX_k + \tilde{E}_k
\end{aligned}$$

Since $\| E_k \|_2 \leqslant \epsilon_k$, we have $\| \tilde{E}_k \|_2 \leqslant \epsilon_k$ for sufficiently small δ and μ. Let $\tilde{Y}_{k+1} = \tilde{X}_{k+1}\tilde{R}_{k+1}$ be the QR-factorization and let $\tilde{X}_{k+1} = V_1\tilde{X}^{(1)}_{(k+1)} + V_2\tilde{X}^{(2)}_{k+1} + V_3\tilde{X}^{(3)}_{k+1}$. Then $\tilde{X}_{k+1}$ satisfies the same condition that X_{k+1} does and the bound on t_{k+1} applies to $\tilde{t}_{k+1} :=\| \tilde{X}^{(2)}_{k+1}(\tilde{X}^{(1)}_{k+1})^{-1} \|_2$ as well. It follows from

$$\tilde{Y}_{k+1} = V_1 X^{(1)}_{k+1}\tilde{R}_{k+1} + V_2\tilde{X}^{(2)}_{k+1}\tilde{R}_{k+1} + V_3\tilde{X}^{(3)}_{k+1}\tilde{R}_{k+1} \tag{7}$$

$$\begin{aligned}
\tilde{Y}_{k+1} &= Y_{k+1} + \delta V_1 R_{k+1} + \mu V_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} R_{k+1} + \theta V_3 R_{k+1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= (Y_{k+1}R^{-1}_{k+1} + \delta V_1 + \mu V_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \theta V_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}) R_{k+1} \\
&= (V_1 X^{(1)}_{k+1} + V_2 X^{(2)}_{k+1} + \delta V_1 + \mu V_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \theta V_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}) R_{k+1}
\end{aligned}$$

$$= [V_1(X^{(1)}_{k+1} + \delta I) + V_2(X^{(2)}_{k+1} + \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix}) + V_3(X^{(3)}_{k+1} + \theta \begin{pmatrix} 0 \\ 1 \end{pmatrix}))] R_{k+1} \tag{8}$$

now let this relation (7) is equal to (8) that

$$\begin{aligned}
&X^{(1)}_{k+1} = (X^{(1)}_{k+1} + \delta I) R_{k+1}\tilde{R}^{(-1)}_{k+1} \\
&X^{(2)}_{k+1} = (X^{(2)}_{k+1} + \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix}) R_{k+1}\tilde{R}^{(-1)}_{k+1} \qquad X^{(3)}_{k+1} = (X^{(3)}_{k+1} + \theta \begin{pmatrix} 0 \\ 1 \end{pmatrix}) R_{k+1}\tilde{R}^{(-1)}_{k+1}
\end{aligned}$$

So $X^{(1)}_{k+1}$ is nonsingular for sufficiently small $\delta > 0$. Then, by case I, we have

$$\tilde{t}_{k+1} =\| \tilde{X}^{(2)}_{k+1}(\tilde{X}^{(1)}_{k+1})^{(-1)} \|_2 \leqslant t_0$$

$$\tilde{t}_{k+1} = \parallel \tilde{X}^{(3)}_{k+1}(\tilde{X}^{(1)}_{k+1})^{(-1)} \parallel_2 \leqslant t_0$$

for all sufficiently small $\delta > 0$ and $\mu \leqslant 0$ and $\theta \leqslant 0$. However

$$\begin{aligned}
\tilde{X}^{(2)}_{k+1}(\tilde{X}^{(1)}_{k+1})^{(-1)} &= (X^{(2)}_{k+1} + \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix}) R_{k+1}\tilde{R}^{(-1)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{(-1)}\tilde{R}_{k+1}R^{(-1)}_{k+1} \\
&= (X^{(2)}_{k+1} + \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix})(X^{(1)}_{k+1} + \delta I)^{(-1)} \\
&= X^{(2)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{(-1)} + \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix}((X^{(1)}_{k+1} + \delta I)^{(-1)})
\end{aligned}$$

is unbounded as $\delta \to 0$, because if $X^{(2)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{-1}$ is unbounded, then $\tilde{t}_{k+1}$ is unbounded by setting $\mu = 0$; and if $X^{(2)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{-1}$ is bounded, then $\tilde{t}_{k+1}$ is unbounded by setting $\mu > 0$. Therefore $X^{(1)}_{k+1}$ is nonsingular and hence $t_{k+1} \leqslant 2t_0$

$$\begin{aligned}
\tilde{X}^{(3)}_{k+1}(\tilde{X}^{(1)}_{k+1})^{(-1)} &= (X^{(3)}_{k+1} + \theta \begin{pmatrix} 1 \\ 0 \end{pmatrix}) R_{k+1}\tilde{R}^{(-1)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{(-1)}\tilde{R}_{k+1}R^{(-1)}_{k+1} \\
&= (X^{(3)}_{k+1} + \theta \begin{pmatrix} 0 \\ 1 \end{pmatrix})(X^{(1)}_{k+1} + \delta I)^{(-1)} \\
&= X^{(3)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{(-1)} + \theta \begin{pmatrix} 0 \\ 1 \end{pmatrix}((X^{(1)}_{k+1} + \delta I)^{(-1)})
\end{aligned}$$

is unbounded as $\delta \to 0$, because if $X^{(3)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{-1}$ is unbounded, then $\tilde{t}_{k+1}$ is unbounded by setting $\theta = 0$; and if $X^{(3)}_{k+1}(X^{(1)}_{k+1} + \delta I)^{-1}$ is bounded, then $\tilde{t}_{k+1}$ is unbounded by setting $\theta > 0$. Therefore $X^{(1)}_{k+1}$ is nonsingular and hence $t_{k+1} \leqslant 2t_0$ proof is completed. ■

We now prove our main result on convergence of t_k and $t_{k'}$. We are interested in the case that ϵ_k is a linearly decreasing sequence.

THEOREM 3.6 *Assume that X_0 is such that X_0^{-1} is invertible. Let $\epsilon_k = a\gamma^k$ with $\gamma < 1$ and*

$$a \leqslant \frac{(1-\rho)2t_0}{\parallel V \parallel_2 \parallel U \parallel_2 (1+2t_0)(\rho+2t_0)}$$

Then we have

$$t_k \leqslant \begin{cases} 2\rho^k t_0 + ac\frac{\gamma^k - \rho^k}{\gamma - \rho} & \gamma \neq \rho \\ 2\rho^k t_0 + ack\rho^{k-1} & \gamma = \rho \end{cases}$$

where

$$c = \parallel V \parallel_2 \parallel U \parallel_2 (1+2t_0)(\rho+2t_0)$$

Proof Since

$$\epsilon_k \leqslant \frac{(1-\rho)2t_0}{\parallel V \parallel_2 \parallel U \parallel_2 (1+2t_0)(\rho+2t_0)}$$

we have $t_k \leqslant t_0$ by Lemma (3.5). Then,

$$\epsilon_k \leqslant \frac{(1-\rho)2t_0}{\| V \|_2 \| U \|_2 \,(1+2t_0)(\rho+2t_0)}$$

so we have $t_k \leqslant t_0$

$$\| V \|_2 \| U \|_2 \,(1+t_k+t_{k'})\epsilon_k \leqslant \| V \|_2 \| U \|_2 \,(1+2t_0)\epsilon \leqslant \tfrac{(1-\rho)2t_0}{(\rho+2t_0)} < 1$$

It follows from Lemma (3.4) that $t_{k+1} \leqslant \rho(t_k+t_{k'}) + c_k\epsilon_k$

$$t_{k+1} \leqslant \rho(t_k+t_{k'}) + \frac{\rho \| V \|_2 \| U \|_2 \,(1+t_k+t_{k'})^2\epsilon_k}{1- \| V \|_2 \| U \|_2 \,(1+t_k+t_{k'})\epsilon_k}$$

$$t_{k+1} \leqslant \rho(t_k+t_{k'}) + ac_k\gamma^k$$

$$t_k \leqslant \rho(t_k+t_{k'}) + ac_{k-1}\gamma^{k-1}$$

$$\vdots$$

$$t_k \leqslant \rho^k 2t_0 + ac\frac{\gamma^k-\rho^k}{\gamma-\rho}$$

now if $\rho = \gamma$ we have $t_k \leqslant \rho^k 2t_0 + ack\rho^{k-1}$.

$$t_{k+1} \leqslant \rho(t_k+t_{k'}) + c_k\epsilon_k$$

$$c_k = \frac{\rho \| V \|_2 \| U \|_2 \,(t_k+t_{\acute{k}})^2}{1- \| V \|_2 \| U \|_2 \,(t_k+t_{k'})\epsilon_k} \leqslant \frac{\rho \| V \|_2 \| U \|_2 \,(1+2t_0)^2}{1-\frac{(1-\rho)2t_0}{(\rho+2t_0)}}$$

$$\leqslant \frac{\rho \| V \|_2 \| U \|_2 \,(1+2t_0)^2}{\frac{\rho+t_0-t_0+\rho t_0}{(\rho+2t_0)}} \leqslant \frac{\rho \| V \|_2 \| U \|_2 \,(1+2t_0)^2(\rho+2t_0)}{\rho(1+2t_0)^2}$$

$$\leqslant \| V \|_2 \| U \|_2 \,(1+2t_0)(\rho+2t_0) = c$$

Therefore, $t_{k+1} \leqslant \rho(t_k+t_{k'}) + ac\gamma^k$. Solving this inequality, we obtain the bound for t_k ■

The conclusion of the above theorem is that the subspace spanned by X_k, $R(X_k)$, converges to the spectral subspace $R(V_1)$ linearly at the rate of $\max\{\rho,\gamma\}$. The condition on a is to ensure convergence and is clearly not a necessary condition. An interesting fact is that there is no gain in convergence rate if we choose $\gamma < \rho$, some shall focus on the case $\gamma > \rho$. The following corollary gives a more precise bound for the constant C and hence for t_k and $t_{k'}$ at the convergence stage.

Corollary 3.7
Let $1 > \gamma > \rho$ and $\epsilon_k = a^k$. Suppose that a is chosen such that $t_k \longrightarrow 0$ and

$t_{k'} \longrightarrow 0$. *Then*

$$\limsup \frac{t_k}{a\gamma^k} \leqslant \rho(\gamma-\rho)^{-1} \parallel V \parallel_2 \parallel U \parallel_2$$

and

$$\limsup \frac{t_{k'_0}}{a\gamma^k} \leqslant \rho(\gamma-\rho)^{-1} \parallel V \parallel_2 \parallel U \parallel_2$$

Proof Apply the main theorem to t_k starting from $k = k_0$, we have

$$t_k \leqslant 2\rho^{k-k_0} t_{k_0} + \frac{\gamma^{k-k_0} - \rho^{k-k_0}}{\gamma - \rho} a\gamma^{k_0} c_{k_0}$$

where

$$c_{k_0} = \frac{\rho \parallel V \parallel_2 \parallel U \parallel_2 (1 + t_{k_0} + t_{k'_0})^2}{1 - \rho \parallel V \parallel_2 \parallel U \parallel_2 (1 + t_{k_0} + t_{k'_0})\epsilon_{k_0}}$$

$$c_{k_0} = \frac{\rho \parallel V \parallel_2 \parallel U \parallel_2 (1 + t_{k_0} + t_{k'_0})^2}{1 - \rho \parallel V \parallel_2 \parallel U \parallel_2 (1 + t_{k_0} + t_{k'_0})\epsilon_{k_0}} \sim \rho \parallel V \parallel_2 \parallel U \parallel_2$$

$$\parallel V \parallel_2 \parallel U \parallel_2 (1 + 2t_{k_0})(\rho + 2t_{k_0}) \leqslant \frac{2(1 - \rho_{k_0})t_{k_0}}{\epsilon_{k_0}}$$

Dividing $a\gamma^k$ and taking $k \longrightarrow \infty$ first and then $k_0 \longrightarrow \infty$ in the inequality, we obtain the bound.

$$\frac{t_k}{a\gamma^k} \leqslant \frac{\rho^{k-k_0} t_{k_0}}{a\gamma^k} + \frac{\gamma^{k-k_0} - \rho^{k-k_0}}{(\gamma-\rho)a\gamma^{k-k_0}} a\gamma^{k_0} c_{k_0}$$

$$\frac{t_k}{a\gamma^k} \leqslant \frac{\rho^{k-k_0} t_{k_0}}{a\gamma^k} + \frac{\gamma^{k-k_0} - \rho^{k-k_0}}{(\gamma-\rho)} \gamma^{-(k-k_0)} c_{k_0}$$

and so $\rho < \gamma < 1$

$$\lim_{k\to\infty} (\frac{\rho}{\gamma})^k (\frac{\rho^{-k_0} t_{k_0}}{a}) = 0$$

and

$$\lim_{k\to\infty} (\frac{\rho}{\gamma})(\frac{\rho^{-k_0} t_{k_0}}{a}) + \lim_{k\to\infty} \frac{\gamma^{k-k_0}(1 - (\frac{\rho}{\gamma})^{k-k_0})}{(\gamma-\rho)} \gamma^{-(k-k_0)} c_{k_0}$$

and

$$= 0 + \lim_{k\to\infty} \frac{1-0}{\gamma-\rho} (\rho \parallel V \parallel_2 \parallel U \parallel_2) = \rho(\gamma-\rho)^{-1} \parallel V \parallel_2 \parallel U \parallel_2$$

$$\limsup \frac{t_k}{a\gamma^k} \leqslant \rho(\gamma-\rho)^{-1} \parallel V \parallel_2 \parallel U \parallel_2$$

Apply the main theorem to $t_{k'_0}$ starting from $k' = k_0$, we have

$$\limsup \frac{t_{k'}}{a\gamma^k} \leqslant \rho(\gamma - \rho)^{-1} \parallel V \parallel_2 \parallel U \parallel_2$$

■

4. Conclusions

We have presented an inexact inverse subspace iteration for computing a few smallest eigenpairs of the generalized eigenvalue problem Ax = Bx. By properly scaling the block vectors, we ensure convergence of columns in the iterative blocks, which allows using approximation from one step as an initial approximation for the next step.we analyzed convergence of the subspace to the spectral space sought.

References

[1] J. Berns-Muler, I. G. Graham and A. Spence, Inexact inverse iteration for symmtric matrices, Linear Algebra Appl, 416 (2006), 389-413.

[2] J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, Templates for the solution of algebraic eigenvalue problems: a practical guide, Philadelphia, PA, USA, 2000.

[3] M. A. Freitag, A. Spence, Convergence rates for inexact inverse iteration with application to preconditioned iterative solves, BIT, 47 (2007), 27-44.

[4] G. H. Golub and Q. Ye, Inexact inverse iterations for the generalized eigenvalue problems, BIT, 40 (1999), 672-684.

[5] G. H. Golub and C. F. Van loan, Matrix computation, Baltimore, MD, USA, 1989.

[6] G. H. Golub, Z. Zhang and H. Zha, Large sparse symmetric eigenvalue prob- lems with homogeneous linear constraints:the lanczos process with inner-outer iteration, Linear Algebra And Its Applications, 309 (2000), 289-306.

[7] Z. Jia, On convergence of the inexact rayleigh quotient iteration without and with minres, 2009.

[8] Y. Lai, K. Lin, W. Lin . An inexact inverse iteration for large sparse eigenvalue problems, Numerical Linear Algebra With Application, (1997), 425-437.

[9] R. B. Lehoucq and Karl Meerbergen, Using generalized cayley transformations within an inexact rational krylov sequence method, SIAM J. Matrix Anal. Appl., 20, 131-148.

[10] R. B. Morgan and D. S. Scott, Preconditioning the lanczos algorithm for sparse symmetric eigenvalue problems, SIAM J. Sci. Comput., 14 (1993), no. 3, 585-593.

[11] A. Ruhe, Rational krylov: A practical algorithm for large sparse nonsymmetric matrix pencils, SIAM J. Sci. Comput., 19 (1998), no. 5, 1535-1551.

[12] A. Ruhe And T. Wiberg, The method of conjugate gradients used in inverse iter- ation, BIT, 12 (1972), 543-554.

[13] P. Smit and M. H. C. Paardekooper, The effects of inexact solvers in algorithms for sym- metric eigenvalue problems, Linear Algebra and its Applications, 287 (1999), 337-357.

[14] V. Simoncini and L. Eldn, Inexact rayleigh quotient-type methods for eigenvalue compu- tations, BIT, 42 (2002), 159-182.

[15] G. Sleijpen and H. Van Der Vorst, A jacobi-davidson iteration method for linear eigenvalue problems, SIAM J. Matrix Anal. Appl., 17 (2000), 401-425.

[16] D. C. Sorensen and C. Yang, A truncated rq iteration for large scale eigenvalue calculations, SIAM J. Matrix Anal. Appl., 19 (1998), no. 4, 1045-1073.

[17] A. Stathopoulos, Y. Saad and F. Fischer, Robust preconditioning of large sparse symmetric eigenvalue problems, Journal of Computational and Applied Mathematics, 64 (1994), 197-215.