BIT Numerical Mathematics **44:** 585–593, 2004.

ACCURACY OF COMPUTED EIGENVECTORS VIA OPTIMIZING A RAYLEIGH QUOTIENT ⋆

REN-CANG LI[1,⋆⋆]

[1] *Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA. email: rcli@ms.uky.edu*

Abstract.

This paper establishes converses to the well-known result: for any vector $\tilde{u}$ such that the sine of the angle $\sin\theta(u,\tilde{u}) = \mathrm{O}(\epsilon)$, we have

$$\rho(\tilde{u}) \stackrel{\mathrm{def}}{=} \frac{\tilde{u}^* A\tilde{u}}{\tilde{u}^*\tilde{u}} = \lambda + \mathrm{O}(\epsilon^2),$$

where λ is an eigenvalue and u is the corresponding eigenvector of a Hermitian matrix A, and "*" denotes complex conjugate transpose. It shows that if $\rho(\tilde{u})$ is close to A's largest eigenvalue, then $\tilde{u}$ is close to the corresponding eigenvector with an error proportional to the square root of the error in $\rho(\tilde{u})$ as an approximation to the eigenvalue and inverse proportional to the square root of the gap between A's first two largest eigenvalues. A subspace version of such an converse is also established. Results as such may have interest in applications, such as eigenvector computations in Principal Component Analysis in image processing where eigenvectors may be computed by optimizing Rayleigh quotients with the Conjugate Gradient method.

AMS subject classification (2000): 15A42, 65F15.

Key words: accuracy, Rayleigh quotient, eigenvector, eigenvalue gap.

1 Introduction.

Let A be Hermitian with an eigenvalue λ and the corresponding eigenvector u. The following is well-known: For any vector $\tilde{u}$ such that the sine of the angle $\sin\theta(u,\tilde{u}) = \mathrm{O}(\epsilon)$, we have

$$\text{(1.1)} \qquad \text{Rayleigh quotient} \quad \rho(\tilde{u}) \stackrel{\mathrm{def}}{=} \frac{\tilde{u}^* A\tilde{u}}{\tilde{u}^*\tilde{u}} = \lambda + \mathrm{O}(\epsilon^2)$$

(see, e.g., [7, 8]), where "*" denotes complex conjugate transpose, and $0 \le \theta(u,\tilde{u}) \le \pi/2$ is defined as

$$\text{(1.2)} \qquad \theta(u,\tilde{u}) = \arccos \frac{|u^*\tilde{u}|}{\sqrt{u^*u \cdot \tilde{u}^*\tilde{u}}}.$$

⋆ Received April 2004. Revised June 2004. Communicated by Axel Ruhe.

⋆⋆ This work was supported in part by a NSF/CAREER award under Grant No. CCR-9875201.

In this paper, we are interested in establishing converses to this statement. Our motivation is from eigenvector computations in Principal Component Analysis in image processing [5, 9], where eigenvectors may be computed by optimizing Rayleigh quotients with the Conjugate Gradient method. This paper is based on an earlier technical report [4].

One may expect that if (1.1) holds then $\tilde{u}$ would be close to an eigenvector u of A. But this may not be true without additional assumptions. Here are two examples.

EXAMPLE 1.1. Let $\delta > 0$ (a tiny number, say 10^{-10}), and

$$A = \begin{pmatrix} 1+\delta & \\ & 1-\delta \end{pmatrix}, \qquad \tilde{u} = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}.$$

Then $\rho(\tilde{u}) = 1 + \delta/2 \equiv (1+\delta) - \epsilon^2$, where $\epsilon = \sqrt{\delta/2}$. But the angle between $\tilde{u}$ and the eigenvector $(1, 0)^*$ of A is $\pi/6$. This is caused by the small gap between the eigenvalue $1 + \delta$ and the rest eigenvalue of A. This example demonstrates that just because a Rayleigh quotient $\rho(\tilde{u})$ is very close to A's largest eigenvalue it does not mean vector $\tilde{u}$ is anywhere near any of A's eigenvectors.

EXAMPLE 1.2. Let

$$A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}, \qquad \tilde{u} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Then $\rho(\tilde{u}) = 2$ *exactly*, but $\tilde{u}$ is orthogonal to the eigenvector corresponding to the eigenvalue 2, nor near any other eigenvectors. This example indicates that a problem may arise when the Rayleigh quotient approaches an interior eigenvalue.

This paper presents two converse results. (One in fact is a special case of the other.) Examples 1.1 and 1.2 give hints as to what kinds of conditions we might need for such converses to exist, e.g., a reasonable gap assumption, $\rho(\tilde{u})$ close to A's exterior (largest or smallest) eigenvalues, and/or additional restriction on $\tilde{u}$ for interior eigenvalues. Our converses may not be immediately applicable to estimate errors for practical purposes but rather they shed lights on how accurate approximate eigenvectors computed via optimizing Rayleigh quotients may be expected.

2 Main result.

The following theorem is in fact a special case ($k = 1$) of Theorem 2.2. But because of its simplicity in conclusion and proof, we single it out.

THEOREM 2.1. *Assume that $n \times n$ matrix A is Hermitian with eigenvalues*

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \tag{2.1}$$

and corresponding orthonormal eigenvectors

$$u_1, u_2, \ldots, u_n. \tag{2.2}$$

If

$$\frac{\tilde{u}_1^* A \tilde{u}_1}{\tilde{u}_1^* \tilde{u}_1} \geq \lambda_1 - \epsilon^2,$$

where $\epsilon \geq 0$, *then*

$$\sin\theta(u_1, \tilde{u}_1) \leq \frac{\epsilon}{\sqrt{\lambda_1 - \lambda_2}}. \tag{2.3}$$

PROOF. We may assume $\tilde{u}_1^* \tilde{u}_1 = 1$. Write $\tilde{u}_1 = \alpha u_1 + \beta v$, where $v \perp u_1$ and $v^* v = 1$. This can be realized by taking $\alpha = u_1^* \tilde{u}_1$, $\hat{v} = \tilde{u}_1 - \alpha u_1$, $\beta = \sqrt{\hat{v}^* \hat{v}}$, and $v = \hat{v}/\beta$ if $\beta > 0$, or v any unit vector orthogonal to u_1 otherwise. It can be seen that

$$|\alpha| = \cos\theta(u_1, \tilde{u}_1), \qquad |\beta| = \sin\theta(u_1, \tilde{u}_1).$$

We have

$$\tilde{u}_1^* A \tilde{u}_1 = |\alpha|^2 \lambda_1 + |\beta|^2 v^* A v.$$

Since $v \perp u_1$, $v^* A v \leq \lambda_2$. Now use $\tilde{u}_1^* A \tilde{u}_1 \geq \lambda_1 - \epsilon^2$ to get

$$\lambda_1 \sin^2\theta(u_1, \tilde{u}_1) - \epsilon^2 \leq |\beta|^2 v^* A v \leq \lambda_2 \sin^2\theta(u_1, \tilde{u}_1),$$

and thus $(\lambda_1 - \lambda_2)\sin^2\theta(u_1, \tilde{u}_1) \leq \epsilon^2$, as expected. □

Notay [6], in our notation, proved the following inequality

$$\tan\theta(u_1, \tilde{u}_1) \leq \frac{\epsilon}{\sqrt{(\lambda_1 - \epsilon^2) - \lambda_2}}, \tag{2.4}$$

if $\epsilon^2 < \lambda_1 - \lambda_2$. This is not a restrictive assumption. In fact our inequality (2.3) is only useful when the assumption is true because $\sin\theta(u_1, \tilde{u}_1) \leq 1$, always. It can be shown (2.3) and (2.4) are equivalent.

Example 1.2 says that $\tilde{u}$ may be nowhere near an eigenvector even if $\rho(\tilde{u})$ is close to an interior eigenvalue. Going through the above proof, we may find that an additional condition: $\tilde{u}$ is orthogonal, say to u_1, will prevent this from happening.

COROLLARY TO THE PROOF. *Assume that* $n \times n$ *matrix* A *is Hermitian with eigenvalues* (2.1) *and corresponding eigenvectors* (2.2). *If* $\tilde{u}$ *is orthogonal to* $u_1, u_2, \ldots, u_{k-1}$, *and*

$$\frac{\tilde{u}^* A \tilde{u}}{\tilde{u}^* \tilde{u}} \geq \lambda_k - \epsilon^2,$$

then

$$\sin\theta(u_k, \tilde{u}) \leq \frac{\epsilon}{\sqrt{\lambda_k - \lambda_{k+1}}}. \tag{2.5}$$

To state the next theorem, we need to define the canonical angles between two subspaces spanned by the columns of X and of $\widetilde{X}$, respectively, where both X and $\widetilde{X}$ have orthonormal columns. The canonical angle matrix [8, p. 43] $\Theta(X,\widetilde{X})$ is defined as

$$\Theta(X,\widetilde{X}) = \operatorname{diag}(\theta_1,\theta_2,\ldots,\theta_k),$$

where $\pi/2 \geq \theta_1 \geq \theta_2 \geq \cdots \geq \theta_k \geq 0$ and $\{\cos\theta_i\}_{i=1}^k$ are the singular values of $X^*\widetilde{X}$. Thus for the *spectral norm* $\|\cdot\|$

$$\|\sin\Theta(X,\widetilde{X})\| = \sin\theta_1.$$

THEOREM 2.2. *Assume that $n\times n$ matrix A is Hermitian with eigenvalues* (2.1) *and corresponding eigenvectors* (2.2). *Let $U_k = [u_1,u_2,\ldots,u_k]$, and let $\widetilde{U}_k$ be $n\times k$ and have orthonormal columns. If*

$$\operatorname{trace}(\widetilde{U}_k^* A\widetilde{U}_k) \geq \lambda_1+\lambda_2+\cdots+\lambda_k-\epsilon^2,$$

then

$$\|\sin\Theta(U_k,\widetilde{U}_k)\| \leq \frac{\epsilon}{\sqrt{\lambda_k-\lambda_{k+1}}}. \tag{2.6}$$

PROOF. Write $\widetilde{U}_k = U_kY + VZ$, where V is $n\times\ell$ and $V^*U_k = 0$. $\widetilde{U}_k^{\mathrm{T}}\widetilde{U}_k = I_k$, the $k\times k$ identity matrix, implies that $Y^*Y+Z^*Z = I_k$. This can be realized by taking $Y = U_k^*\widetilde{U}_k$, $\widehat{V} = \widetilde{U}_k - U_kY$. It can be checked that $U_k^*\widehat{V} = 0$. Orthogonalize the columns of $\widehat{V}$ to yield V with ℓ orthonormal vectors ($\ell\leq k$) and $\widehat{V} = VZ$. Since must then $U_k^*V = 0$, $k+\ell\leq n$. So

$$\ell \leq \min\{k, n-k\}.$$

Let $V^*AV = Q\Omega Q^*$ be an eigendecomposition, where Q is $\ell\times\ell$ and unitary, and $\Omega = \operatorname{diag}(\mu_1,\mu_2,\ldots,\mu_\ell)$. Since $V^*U_k = 0$, all $\mu_i\leq\lambda_{k+1}$; so we may arrange μ_i's in descending order, i.e.,

$$\lambda_{k+1} \geq \mu_1 \geq \mu_2 \geq \cdots \geq \mu_\ell. \tag{2.7}$$

Let $\Lambda_k = \operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_k)$. We have

$$\widetilde{U}_k^*A\widetilde{U}_k = Y^*\Lambda_kY + Z^*V^*AVZ = Y^*\Lambda_kY + Z^*Q\Omega Q^*Z.$$

Thus

$$\begin{aligned}\operatorname{trace}(\widetilde{U}_k^*A\widetilde{U}_k) &= \operatorname{trace}(Y^*\Lambda_kY) + \operatorname{trace}(Z^*Q\Omega Q^*Z)\\ &= \operatorname{trace}(YY^*\Lambda_k) + \operatorname{trace}(Q^*ZZ^*Q\Omega)\end{aligned}$$

to give

$$\operatorname{trace}((I_k - YY^*)\Lambda_k) - \epsilon^2 \leq \operatorname{trace}(Q^*ZZ^*Q\Omega).$$

Let the eigenvalues of $I_k - YY^*$ be

$$\sigma_1 \geq \sigma_1 \geq \cdots \geq \sigma_k \geq 0,$$

which are also the eigenvalues of ZZ^* (and of Q^*ZZ^*Q as well), including or excluding some zeros. Therefore there are at most ℓ σ_i's are nonzero. Notice that

$$\sigma_1 = \|\sin\Theta(U_k, \widetilde{U}_k)\|^2.$$

Now since (they will be justified later).

$$\text{trace}((I_k - YY^*)\Lambda_k) \geq \sigma_1\lambda_k + \sigma_2\lambda_{k-1} + \cdots + \sigma_k\lambda_1, \tag{2.8}$$

$$\text{trace}(Q^*ZZ^*Q\Omega) \leq \sigma_1\mu_1 + \sigma_2\mu_2 + \cdots, \tag{2.9}$$

we conclude that

$$\sigma_1\lambda_k + \sigma_2\lambda_{k-1} + \cdots + \sigma_k\lambda_1 - \epsilon^2 \leq \sigma_1\mu_1 + \sigma_2\mu_2 + \cdots. \tag{2.10}$$

Since at most ℓ σ_i's are nonzero and $\ell \leq \{k, n-k\}$, we have from (2.10) and (2.7) that

$$\begin{aligned} \sigma_1\lambda_k &\leq \sigma_1\mu_1 + \sum_{i=2}^{\ell} \sigma_i(\mu_i - \lambda_{k-i+1}) + \epsilon^2 \\ &\leq \sigma_1\lambda_{k+1} + \epsilon^2, \end{aligned}$$

to get $\sigma_1(\lambda_k - \lambda_{k+1}) \leq \epsilon^2$ as expected. □

How does Theorem 2.2 do if compared to existing error bounds in terms of residual? It appears this question is not easy to answer in its generality, i.e., for $k > 1$. Nevertheless, for $k = 1$, i.e., Theorem 2.1, it is sharper than [8, p. 250]

$$\sin\theta(u_1, \tilde{u}_1) \leq \frac{\|A\tilde{u}_1 - \rho(\tilde{u}_1)\tilde{u}_1\|}{\rho(\tilde{u}_1) - \lambda_2}, \tag{2.11}$$

a single vector version of Davis–Kahan $\sin\theta$ theorem [2]. Here all notation and conditions of Theorem 2.1 and its proof are inherited. We shall now show the right-hand side (2.11) is no smaller than that of (2.3). Denote $r = A\tilde{u}_1 - \rho(\tilde{u}_1)\tilde{u}_1$, $c = |\alpha| = \cos\theta(u_1, \tilde{u}_1)$, $s = |\beta| = \sin\theta(u_1, \tilde{u}_1)$, and assume $\epsilon^2 < \lambda_1 - \lambda_2$. We have

$$\begin{aligned} \rho(\tilde{u}_1) &= c^2\lambda_1 + s^2\rho(v), \\ A\tilde{u}_1 &= \alpha\lambda_1 u_1 + \beta A v, \\ r &= \alpha s^2[\lambda_1 - \rho(v)]u_1 + \beta[A - \rho(\tilde{u}_1)]v, \\ \|r\|^2 &= c^2[s^2(\lambda_1 - \rho(v))]^2 + s^2\|[A - \rho(\tilde{u}_1)]v\|^2. \end{aligned}$$

For comparison purpose, we make ϵ as small as possible, for which $\rho(\tilde{u}_1) = \lambda_1 - \epsilon^2$. This gives

$$s^2[\lambda_1 - \rho(v)] = \epsilon^2.$$

Next expand v as a linear combination of u_i for $i \geq 2$ to get $v = \sum_{i=2}^n \xi_i u_i$. Then $\sum_{i=2}^n |\xi_i|^2 = 1$, and

$$\begin{aligned}
\rho(v) &= \sum_{i=2}^n \lambda_i |\xi_i|^2, \\
\|(A - \rho(\tilde{u}_1))v\|^2 &= \sum_{i=2}^n [\lambda_i - \rho(\tilde{u}_1)]^2 |\xi_i|^2 \\
&\geq [\rho(\tilde{u}_1) - \lambda_2] \sum_{i=2}^n [\rho(\tilde{u}_1) - \lambda_i] |\xi_i|^2 \\
&= [\rho(\tilde{u}_1) - \lambda_2][\rho(\tilde{u}_1) - \rho(v)] \\
&= [\rho(\tilde{u}_1) - \lambda_2][\lambda_1 - \epsilon^2 - \rho(v)].
\end{aligned}$$

Put all together to get

$$\|r\|^2 \geq c^2\epsilon^4 + [\rho(\tilde{u}_1) - \lambda_2](\epsilon^2 - s^2\epsilon^2) = c^2\epsilon^2(\lambda_1 - \lambda_2) \geq (\lambda_1 - \epsilon^2 - \lambda_2)\epsilon^2$$

upon using $c^2 = 1 - s^2 \geq (\lambda_1 - \epsilon^2 - \lambda_2)/(\lambda_1 - \lambda_2)$ by (2.3). Therefore

$$\frac{\|r\|}{\rho(\tilde{u}_1) - \lambda_2} \geq \frac{\epsilon}{\sqrt{\lambda_1 - \epsilon^2 - \lambda_2}} \geq \frac{\epsilon}{\sqrt{\lambda_1 - \lambda_2}},$$

as was claim.

We now justify the inequalities (2.8) and (2.9). To this end, we need a result from majorization [1, 3]. Given two sequences of real numbers $\{\alpha_i\}_{i=1}^m$ and $\{\beta_i\}_{i=1}^m$, we say that $\{\beta_i\}_{i=1}^m$ *majorizes* $\{\alpha_i\}_{i=1}^m$ if

$$\sum_{i=1}^j \alpha_i^\downarrow \leq \sum_{i=1}^j \beta_i^\downarrow, \quad \text{for } j = 1, 2, \dots, m$$

with equality holds for $j = m$, where $\{\alpha_i^\downarrow\}_{i=1}^m$ is from re-ordering $\{\alpha_i\}_{i=1}^m$ in decreasing order, i.e.,

$$\alpha_1^\downarrow \geq \alpha_2^\downarrow \geq \cdots \geq \alpha_m^\downarrow$$

(similarly for $\{\beta_i^\downarrow\}_{i=1}^m$). We also use notation $\alpha_i^\uparrow$ obtained from re-ordering $\{\alpha_i\}_{i=1}^m$ as well but in increasing order.

The following lemma, together with the fact that *the sequence of the diagonal entries of a Hermitian matrix is majorized by the sequence of its eigenvalues* [1, Exercise II.1.12, p. 35], yield both (2.8) and (2.9).

LEMMA 2.3. *Let $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_m \geq 0$. If $\{\beta_i\}_{i=1}^m$ majorizes $\{\alpha_i\}_{i=1}^m$, then*

$$\sum_{i=1}^m \gamma_i \beta_i^\uparrow \leq \sum_{i=1}^m \gamma_i \alpha_i \leq \sum_{i=1}^m \gamma_i \beta_i^\downarrow.$$

PROOF. Set

$$p_j = \sum_{i=1}^{j} \beta_i^{\uparrow}, \qquad s_j = \sum_{i=1}^{j} \alpha_i, \qquad t_j = \sum_{i=1}^{j} \beta_i^{\downarrow}, \qquad p_0 = s_0 = t_0 = 0.$$

Since $\{\beta_i\}_{i=1}^m$ majorizes $\{\alpha_i\}_{i=1}^m$, we have

$$p_j \le s_j \le t_j, \qquad p_m = s_m = t_m$$

and thus

$$\begin{aligned}
\sum_{i=1}^{m} \gamma_i \alpha_i &= \sum_{i=1}^{m} (s_i - s_{i-1})\gamma_i \\
&= \sum_{i=1}^{m} s_i \gamma_i - \sum_{i=2}^{m} s_{i-1}\gamma_i \\
&= s_m \gamma_m + \sum_{i=1}^{m-1} s_i(\gamma_i - \gamma_{i+1}) \\
&\le t_m \gamma_m + \sum_{i=1}^{m-1} t_i(\gamma_i - \gamma_{i+1}) \\
&= \sum_{i=1}^{m} \gamma_i \beta_i^{\downarrow}, \\
\sum_{i=1}^{m} \gamma_i \alpha_i &= s_m \gamma_m + \sum_{i=1}^{m-1} s_i(\gamma_i - \gamma_{i+1}) \\
&\ge p_m \gamma_m + \sum_{i=1}^{m-1} p_i(\gamma_i - \gamma_{i+1}) \\
&= \sum_{i=1}^{m} \gamma_i \beta_i^{\uparrow},
\end{aligned}$$

as required. □

3 Conclusions.

We conclude this paper with the following remarks.

(1) The results here show that the error in computed eigenvectors (eigenspaces) via optimizing a Rayleigh quotient is expected to be about the square root of the error in the Rayleigh quotient as approximations to the eigenvalues. This could be bad news. Fortunately in the applications when it is used now, computed eigenvectors (eigenspaces) with such accuracy are good enough for the practical purpose [5, 9].

(2) Example 1.2 shows that $\tilde{u}$ is in no way near an eigenvector even if $\rho(\tilde{u})$ is an *exact* interior eigenvalues. In actual computations as in [5, 9], a deflation is performed after approximations to $u_1, \ldots, u_{k-1}$ is gotten in a way equivalent to optimizing $\rho(\tilde{u})$ subject to $\tilde{u}$ orthogonal to the computed vectors. Of course this is not the same as $\tilde{u} \perp u_1, \ldots, u_{k-1}$, as required by Corollary to the Proof, but it makes sure $\tilde{u}$ has tiny component in directions of $u_1, \ldots, u_{k-1}$.

(3) Replacing A by $-A$ in Theorems 2.1 and 2.2 will yield results for the eigenvectors corresponding the smallest eigenvalues.

(4) Inequality (2.3) is asymptotically sharp as the following example shows.

$$A = \begin{pmatrix} 1 & \\ & 2 \end{pmatrix}, \qquad \tilde{u}_1 = \begin{pmatrix} \epsilon \\ 1 \end{pmatrix},$$
$$\frac{\tilde{u}_1^* A \tilde{u}_1}{\tilde{u}_1^* \tilde{u}_1} = 2 - \epsilon^2 + \mathcal{O}(\epsilon^4), \qquad \sin\theta(\tilde{u}_1, u_1) = \epsilon + \mathcal{O}(\epsilon^3).$$

Its subspace version (2.6) is asymptotically sharp, too. For example,

$$A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}, \qquad \tilde{u}_1 = \frac{1}{\sqrt{1+\epsilon^2}} \begin{pmatrix} \epsilon \\ 1 \\ 0 \end{pmatrix}, \qquad \tilde{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$
$$\operatorname{trace}(\widetilde{U}_2^* A \widetilde{U}_2) = 5 - \epsilon^2 + \mathcal{O}(\epsilon^4), \qquad \sin\Theta(\widetilde{U}_2, U_2) = \epsilon + \mathcal{O}(\epsilon^3).$$

(5) It appears to be very complicated to compare Theorem 2.2 to Davis–Kahan $\sin\theta$ theorem [8, p. 250] for $k > 1$ in general. But we did show Theorem 2.1 is sharper than the vector version Davis–Kahan $\sin\theta$ theorem.

Acknowledgment.

The author wishes to thank an anonymous referee for his many constructive comments and for his suggestion of adding a comparison of results here to the well-known error bounds in terms of the residual. This led to the discovery that Theorem 2.1 is sharper than the vector version Davis–Kahan $\sin\theta$ theorem.

REFERENCES

1. R. Bhatia, *Matrix Analysis*, Graduate Texts in Math. 169, Springer, New York, 1996.
2. C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM Journal on Numerical Analysis, 7 (1970), pp. 1–46.
3. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
4. R.-C. Li, *Accuracy of computed eigenvectors via optimizing a Rayleigh quotient*, Technical Report 2000-28, Department of Mathematics, University of Kentucky, 2000. Avaliable at `http://www.ms.uky.edu/~math/MAreport/PDF/00-28.pdf`.
5. H. Murase and M. Lindenbaum, *Partial eigenvalue decomposition of large images using the spatial temporal adaptive method*, IEEE Transactions on Image Processing, 4 (1995), pp. 620–629.
6. Y. Notay, *Convergence analysis of inexact Rayleigh quotient iteration*, SIAM Journal on Matrix Analysis and Applications, 24 (2003), pp. 627–644.
7. B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

8. G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
9. X. Yang, T. K. Sarkar, and E. Arvas, *A survey of conjugate gradient algorithms for solution of extreme eigen-problems of a symmetric matrix*, IEEE Transactions on Acoustics, Speech and Signal Processing, 37 (1989), pp. 1550–1555.