Journal of Computational Mathematics, Vol.25, No.5, 2007, 512–521.

ON THE RAYLEIGH QUOTIENT FOR SINGULAR VALUES *

Xiaoshan Chen and Wen Li
(*School of Mathematical Sciences, South China Normal University, Guangzhou, 510631, P. R. China*
Email: cxs333@21th.com.cn   liwen@scnu.edu.cn)

Abstract

In this paper, the theoretical analysis for the Rayleigh quotient matrix is studied, some results of the Rayleigh quotient (matrix) of Hermitian matrices are extended to those for arbitrary matrix on one hand. On the other hand, some unitarily invariant norm bounds for singular values are presented for Rayleigh quotient matrices. Our results improve the existing bounds.

Mathematics subject classification: 65F10.
Key words: The singular value, Rayleigh quotient, Canonical angle matrix.

1. Introduction

Let $C^{m\times n}$ be the set of all $m\times n$ complex matrices, and let $C^m = C^{m\times 1}$. Without loss of generality we always assume that $m\geq n$ in this paper. By $\|\cdot\|$ we denote a unitarily invariant norm. Especially, by $\|\cdot\|_2$ and $\|\cdot\|_F$ we denote the spectral norm and the Frobenius norm, respectively. A^* stands for the conjugate transpose of a matrix A. Let $\sigma(A)$ be the set of the singular values of A, I_k be the identity matrix of order k. For the column vectors x and y, the angle $\theta(x,y)\in[0,\frac{\pi}{2}]$ between x and y is defined by

$$\theta(x,y) = \arccos\frac{|x^*y|}{\sqrt{x^*x\cdot y^*y}}.$$

More generally, the canonical angle matrix $\Theta(X,\widetilde{X})$ between two subspaces spanned by the columns of $X\in C^{n\times k}$ and $\widetilde{X}\in C^{n\times k}$ is defined as [1]

$$\Theta(X,\widetilde{X}) = \mathrm{diag}(\theta_1,...,\theta_k),$$

where X and $\widetilde{X}$ have orthonormal columns, $\pi/2\geq\theta_1\geq...\geq\theta_k\geq 0$ and $\{\cos\theta_i\}_{i=1}^k$ are the singular values of $X^*\widetilde{X}$.

Let $A\in C^{n\times n}$ be a Hermitian matrix. The Rayleigh quotient of A with respect to $x\in C^n$ is defined by

$$\rho(x) = \frac{x^*Ax}{x^*x},\quad 0\neq x\in C^n.$$

More generally, let $\tilde{U}_k\in C^{n\times k}$ $(1<k\leq n)$ and $\tilde{U}_k^*\tilde{U}_k = I_k$. Then the matrix

$$N = \tilde{U}_k^*A\tilde{U}_k$$

is called the Rayleigh quotient matrix of the Hermitian matrix A with respect to $\tilde{U}_k$.

The Rayleigh quotient (matrix) plays an important role in computing eigenvalues and eigenvectors. In particular, it can be applied to the eigenvector computations in Principal Component Analysis in image processing [8]. It has been studied by many authors (see, e.g., [2,4-8,14,15]).

* Received December 15, 2005; final revised September 1, 2006; accepted October 1, 2006.

Theorem 1.1. [1] *Let $A \in C^{n\times n}$ be a Hermitian matrix, and let λ be the eigenvalue of A and u be the eigenvector corresponding to λ. If a vector $\tilde{u}$ satisfies* $\sin\theta(u,\tilde{u}) = \mathcal{O}(\varepsilon)$, *then*

$$\rho(\tilde{u}) = \frac{\tilde{u}^*A\tilde{u}}{\tilde{u}^*\tilde{u}} = \lambda + \mathcal{O}(\varepsilon^2).$$

Theorem 1.1 shows that the precision of the Rayleigh quotient $\rho(\tilde{u})$ as an approximate eigenvalue of a Hermitian matrix A is higher than that of $\tilde{u}$ as its approximate eigenvector. The converse of Theorem 1.1 was considered by Li [8] who obtained the following theorem.

Theorem 1.2. [8] *Let $A \in C^{n\times n}$ be a Hermitian matrix with eigenvalues*

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \tag{1.1}$$

and corresponding orthonormal eigenvectors $u_1, u_2, \cdots, u_n$. If

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

where $\varepsilon \geq 0$, then

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

Furthermore, Li [8] extended Theorem 1.2 to a more general case.

Theorem 1.3. [8] *Let $A \in C^{n\times n}$ be a Hermitian matrix with eigenvalues* (1.1) *and corresponding eigenvectors* (1.2). *Let $U_k = (u_1, ..., u_k)$, and let $\tilde{U}_k$ be $n\times k$ and have orthonormal columns. If*

$$trace(\tilde{U}_k^*A\tilde{U}_k) \geq \lambda_1 + ... + \lambda_k - \varepsilon^2,$$

where $\varepsilon \geq 0$, then

$$||\sin\Theta(U_k,\tilde{U}_k)||_2 \leq \frac{\varepsilon}{\sqrt{\lambda_k-\lambda_{k+1}}}.$$

The following theorem provides a bound on the eigenvalues of $\tilde{U}_k^*A\tilde{U}_k$ as an approximation to those of A (see [5,11]).

Theorem 1.4. [5, 11] *Let $A \in C^{n\times n}$ be a Hermitian matrix with eigenvalues* (1.1) *and $\tilde{U}_k \in C^{n\times k}$ have orthonormal columns. Let $N = \tilde{U}_k^*A\tilde{U}_k$ and $R = A\tilde{U}_k - \tilde{U}_kN$. If the eigenvalues of N are $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_k$, then there is a permutation τ of $\{1,2,\cdots,n\}$ such that*

$$\sqrt{\sum_{i=1}^{k}(\nu_i - \lambda_{\tau(i)})^2} \leq \|R\|_F.$$

The Rayleigh quotient of Hermitian matrices for eigenvalue problems can be extended to the Rayleigh quotient (matrix) of an arbitrary matrix for singular value problems. Let

$$\mathcal{X} = \{x \mid x \in C^m, \|x\|_2 = 1\} \qquad \mathcal{Y} = \{y \mid y \in C^n, \|y\|_2 = 1\}.$$

The Rayleigh quotient of an arbitrary matrix $A \in C^{m\times n}$ for the singular value problem is defined by

$$\rho(x,y) = x^*Ay, \quad x \in \mathcal{X}, y \in \mathcal{Y}. \tag{1.2}$$

Similarly, the Rayleigh quotient matrix of an arbitrary matrix $A \in C^{m\times n}$ for the singular values is defined by

$$M = \tilde{U}_k^* A \tilde{V}_k, \tag{1.3}$$

where $\tilde{U_k} \in C^{m\times k}, \tilde{V_k} \in C^{n\times k}, \tilde{U_k}^* \tilde{U_k} = \tilde{V}_k^* \tilde{V}_k = I_k$ (see, e.g., [3]).

A natural question is whether or not Theorems 1.1-1.4 can be extended to the Rayleigh quotient of an arbitrary matrix A for the singular value problem. In this paper, we will consider this question.

The rest of this paper is organized as follows. In Section 2, we show that Theorems 1.1-1.3 are true for the Rayleigh quotient of an arbitrary matrix A for the singular value problem. In Section 3, the singular values of the Rayleigh quotient matrices will be discussed. We construct a result corresponding to Theorem 1.4, which improves the one by Liu in [3]. Moreover, some bounds are obtained in some unitarily invariant norm.

2. Approximate Subspace Variations for the Rayleigh Quotient

In this section we consider approximate subspace variations of the Rayleigh quotient for the singular value problems. Some corresponding results for the Rayleigh quotient of a Hermitian matrix are given. The following theorem corresponds to Theorem 1.1.

Theorem 2.1. *Let $A \in C^{m\times n}$, and let u_1 and v_1 be the left and right singular vectors corresponding to the singular value σ_1, respectively. If $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ satisfy*

$$\sin\theta(x, u_1) = \mathcal{O}(\varepsilon) \quad \textit{and} \quad \sin\theta(y, v_1) = \mathcal{O}(\varepsilon), \tag{2.1}$$

then

$$|\, \rho(x, y)\, | - \sigma_1 = \mathcal{O}(\varepsilon^2). \tag{2.2}$$

Proof. If $\sin\theta(x, u_1) = 0$ and $\sin\theta(y, v_1) = 0$, then it is obvious that (2.2) holds. We only consider the case that $\sin\theta(x, u_1) \neq 0$ and $\sin\theta(y, v_1) \neq 0$. The result (2.2) can be proved by the analogous approach for the case that $\sin\theta(x, u_1) \neq 0$ and $\sin\theta(y, v_1) = 0$ or $\sin\theta(x, u_1) = 0$ and $\sin\theta(y, v_1) \neq 0$.

Without loss of generality, we may assume that $\|u_1\|_2 = \|v_1\|_2 = 1$. The singular value decomposition of A is given by

$$U^* A V = \begin{pmatrix} \Sigma \\ 0 \end{pmatrix},$$

where $U = (u_1, u_2, \cdots, u_m) \in C^{m\times m}$ and $V = (v_1, v_2, \cdots, v_n) \in C^{n\times n}$ are unitary, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$. Put $w = U^* x = (\eta_1, \eta_2, \cdots, \eta_m)^T$ and $z = V^* y = (\gamma_1, \gamma_2, \cdots, \gamma_n)^T$. Then

$$x = Uw = \eta_1 u_1 + \eta_2 u_2 + \cdots + \eta_m u_m \quad \text{and} \quad y = Ve = \gamma_1 v_1 + \gamma_2 v_2 + \cdots \gamma_n v_n.$$

Similar to the proof of [1, Theorem 7.1], it follows from (2.1) that

$$\begin{array}{lll} \eta_1 = 1 + \mathcal{O}(\varepsilon), & \eta_i = \mathcal{O}(\varepsilon), & i = 2, \cdots, m \\ \gamma_1 = 1 + \mathcal{O}(\varepsilon), & \gamma_j = \mathcal{O}(\varepsilon), & j = 2, \cdots, n. \end{array} \tag{2.3}$$

Since

$$\rho(x, y) = x^* A y = w^* (U^* A V) e = \overline{\eta}_1 \gamma_1 \sigma_1 + \overline{\eta}_2 \gamma_2 \sigma_2 + \cdots + \overline{\eta}_n \gamma_n \sigma_n,$$

where $\overline{a}$ is a conjugate complex of a complex number a, it follows from (2.3) that

$$|\rho(x,y)| = |\overline{\eta}_1||\gamma_1|\sigma_1 + \mathcal{O}(\varepsilon^2) = \sqrt{1-\sum_{i=2}^{m}|\eta_i|^2}\sqrt{1-\sum_{i=2}^{n}|\gamma_i|^2}\sigma_1 + \mathcal{O}(\varepsilon^2).$$

Using $\sqrt{1-x} = 1 - \frac{1}{2}x - \frac{1}{8}x^2 + \cdots$ and the above result yields

$$|\rho(x,y)| = \sigma_1 + \mathcal{O}(\varepsilon^2).$$

This completes the proof. □

The following theorem is the converse of Theorem 2.1, which corresponds to Theorem 1.2.

Theorem 2.2. *Assume that $A \in C^{m\times n}$ with singular values*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n, \tag{2.4}$$

corresponding left and right singular orthonormal vectors

$$u_1, u_2, \cdots, u_m \quad \textit{and} \quad v_1, v_2, \cdots, v_n, \tag{2.5}$$

respectively. If

$$| x^*Ay | \geq \sigma_1 - \varepsilon^2, \quad x \in \mathcal{X}, y \in \mathcal{Y},$$

where $\varepsilon \geq 0$, then

$$\sqrt{\sin^2\theta(x,u_1) + \sin^2\theta(y,v_1)} \leq \sqrt{\frac{2}{\sigma_1 - \sigma_2}}\,\varepsilon.$$

Proof. Write $x = \alpha_1 u_1 + \beta_1 u$ and $y = \alpha_2 v_1 + \beta_2 v$, where $\|u\|_2 = 1, u \perp u_1, \|v\|_2 = 1$ and $v \perp v_1$. Then from [8] we have

$$| \alpha_1 | = \cos\theta(x,u_1), \quad | \beta_1 | = \sin\theta(x,u_1);$$

$$| \alpha_2 | = \cos\theta(y,v_1), \quad | \beta_2 | = \sin\theta(y,v_1).$$

From the fact that $u \perp u_1$ and $v \perp v_1$, it follows that $| u^*Av | \leq \sigma_2$, and thus we have

$$\begin{aligned}
& | x^*Ay | = | (\alpha_1 u_1 + \beta_1 u)^* A(\alpha_2 v_1 + \beta_2 v) | \\
= \ & | \sigma_1 \overline{\alpha}_1 \alpha_2 + \overline{\beta}_1 \beta_2 u^*Av | \\
\leq \ & \sigma_1 | \overline{\alpha}_1 \alpha_2 | + | \overline{\beta}_1 \beta_2 || u^*Av | \\
\leq \ & \sigma_1 | \overline{\alpha}_1 \alpha_2 | + \sigma_2 | \overline{\beta}_1 \beta_2 | \\
= \ & \sigma_1 \cos\theta(x,u_1)\cos\theta(y,v_1) + \sigma_2 \sin\theta(x,u_1)\sin\theta(y,v_1) \\
\leq \ & \frac{1}{2}\sigma_1[\cos^2\theta(x,u_1) + \cos^2\theta(y,v_1)] + \frac{1}{2}\sigma_2[\sin^2\theta(x,u_1) + \sin^2\theta(y,v_1)] \\
= \ & \sigma_1 - \frac{1}{2}(\sigma_1 - \sigma_2)[\sin^2\theta(x,u_1) + \sin^2\theta(y,v_1)],
\end{aligned}$$

which together with the assumption that $| x^*Ay | \geq \sigma_1 - \varepsilon^2$ gives

$$(\sigma_1 - \sigma_2)[\sin^2\theta(x,u_1) + \sin^2\theta(y,v_1)] \leq 2\varepsilon^2.$$

This implies the result. The proof is complete. □

In order to get the result corresponding to Theorem 1.3, the following lemma is useful.

Lemma 2.1. *Let $A \in C^{m\times p}$, $B \in C^{p\times q}$ and $C \in C^{q\times m}$. Let $k = \min\{m, p, q\}$ and the singular values of A, B and C are ordered by $\sigma_1(A) \geq \cdots \geq \sigma_{\min\{m,p\}}(A) \geq 0$, $\sigma_1(B) \geq \cdots \geq \sigma_{\min\{p,q\}}(B) \geq 0$ and $\sigma_1(C) \geq \cdots \geq \sigma_{\min\{q,m\}}(C) \geq 0$. Then*

$$| \, trace(ABC) \, | \leq \sum_{i=1}^{k} \sigma_i(A)\sigma_i(B)\sigma_i(C).$$

Proof. By augmenting the involved matrices with zero blocks, the rectangular matrices will be changed to the square matrices. Then the result follows immediately from (3.3.13) and (3.3.22) of [13]. The proof is complete. □

Theorem 2.3. *Assume that $A \in C^{m\times n}$ with singular values (2.4) and corresponding left and right singular orthonormal vectors (2.5). Let $U_k = (u_1, u_2, \cdots, u_k)$, $V_k = (v_1, v_2, \cdots, v_k)$, and let $\tilde{U}_k$ and $\tilde{V}_k$ be $m \times k$ and $n \times k$ matrices with their orthonormal columns, respectively. If*

$$|trace(\tilde{U}_k^* A \tilde{V}_k)| \geq \sigma_1 + \sigma_2 + \cdots + \sigma_k - \varepsilon^2, \tag{2.6}$$

where $\varepsilon \geq 0$, then

$$\sqrt{\|\sin\Theta(U_k, \tilde{U}_k)\|_2^2 + \|\sin\Theta(V_k, \tilde{V}_k)\|_2^2} \leq \sqrt{\frac{2}{\sigma_k - \sigma_{k+1}}}\,\varepsilon.$$

Proof. Let $U_{m-k} = (u_{k+1}, u_{k+2}, \cdots, u_m)$ and $V_{n-k} = (v_{k+1}, v_{k+2}, \cdots, v_n)$. Thus we have

$$U_k^* A V_k = \Sigma_k \quad \text{and} \quad U_{m-k}^* A V_{n-k} = \begin{pmatrix} \Sigma_{n-k} \\ 0 \end{pmatrix}, \tag{2.7}$$

where $\Sigma_k = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_k)$ and $\Sigma_{n-k} = \mathrm{diag}(\sigma_{k+1}, \sigma_{k+2}, \cdots, \sigma_n)$. Obviously, there are $Y_1 \in C^{k\times k}$, $Z_1 \in C^{(m-k)\times k}$, $Y_2 \in C^{k\times k}$ and $Z_2 \in C^{(n-k)\times k}$ such that

$$\tilde{U}_k = U_k Y_1 + U_{m-k} Z_1 \quad \text{and} \quad \tilde{V}_k = V_k Y_2 + V_{n-k} Z_2, \tag{2.8}$$

which, together with the fact that $\tilde{U}_k^* \tilde{U}_k = \tilde{V}_k^* \tilde{V}_k = I_k$, give

$$Y_1^* Y_1 + Z_1^* Z_1 = I_k \quad \text{and} \quad Y_2^* Y_2 + Z_2^* Z_2 = I_k. \tag{2.9}$$

Let

$$0 \leq c_1 \leq c_2 \leq \cdots \leq c_k \quad \text{and} \quad s_1 \geq s_2 \geq \cdots \geq s_k \geq 0$$

be the singular values of Y_1 and Z_1, respectively. Similarly, let

$$0 \leq \tilde{c}_1 \leq \tilde{c}_2 \leq \cdots \leq \tilde{c}_k \quad \text{and} \quad \tilde{s}_1 \geq \tilde{s}_2 \geq \cdots \geq \tilde{s}_k \geq 0$$

be the singular values of Y_2 and Z_2, respectively. By (2.9), we have

$$c_i^2 + s_i^2 = 1 \quad \text{and} \quad \tilde{c}_i^2 + \tilde{s}_i^2 = 1. \tag{2.10}$$

Since $U_{m-k}^* \tilde{U}_k = Z_1$ and $V_{n-k}^* \tilde{V}_k = Z_2$, from Lemma 7.1 of [1] we get

$$\|\sin\Theta(U_k, \tilde{U}_k)\|_2 = \|Z_1\|_2 = s_1 \quad \text{and} \quad \|\sin\Theta(V_k, \tilde{V}_k)\|_2 = \|Z_2\|_2 = \tilde{s}_1. \tag{2.11}$$

From (2.7) and (2.8) we have

$$\tilde{U}_k^* A \tilde{V}_k = Y_1^* \Sigma_k Y_2 + Z_1^* \begin{pmatrix} \Sigma_{n-k} \\ 0 \end{pmatrix} Z_2,$$

and hence

$$trace(\tilde{U}_k^* A \tilde{V}_k) = trace(Y_1^* \Sigma_k Y_2) + trace(Z_1^* \begin{pmatrix} \Sigma_{n-k} \\ 0 \end{pmatrix} Z_2).$$

By (2.10) and Lemma 3.1, we get

$$\begin{aligned}
|trace(\tilde{U}_k^* A \tilde{V}_k)| &\leq \mid trace(Y_1^* \Sigma_k Y_2) \mid + \mid trace(Z_1^* \begin{pmatrix} \Sigma_{n-k} \\ 0 \end{pmatrix} Z_2) \mid \\
&\leq \sigma_1 c_k \tilde{c}_k + \cdots + \sigma_k c_1 \tilde{c}_1 + \sigma_{k+1} s_1 \tilde{s}_1 + \cdots + \sigma_{2k} s_k \tilde{s}_k \\
&\leq \tfrac{1}{2}\sigma_1 (c_k^2 + \tilde{c}_k^2) + \cdots + \tfrac{1}{2}\sigma_k (c_1^2 + \tilde{c}_1^2) \\
&\quad + \tfrac{1}{2}\sigma_{k+1}(s_1^2 + \tilde{s}_1^2) + \cdots + \tfrac{1}{2}\sigma_{2k}(s_k^2 + \tilde{s}_k^2) \\
&= \sigma_1 + \cdots + \sigma_k - \tfrac{1}{2}\sigma_1 (s_k^2 + \tilde{s}_k^2) - \cdots - \tfrac{1}{2}\sigma_k (s_1^2 + \tilde{s}_1^2) \\
&\quad + \tfrac{1}{2}\sigma_{k+1}(s_1^2 + \tilde{s}_1^2) + \cdots + \tfrac{1}{2}\sigma_{2k}(s_k^2 + \tilde{s}_k^2) \\
&\leq \sigma_1 + \cdots + \sigma_k - \tfrac{1}{2}(\sigma_k - \sigma_{k+1})(s_1^2 + \tilde{s}_1^2),
\end{aligned}$$

which together with the assumptions (2.6) and (2.11) gives the desired estimate. The proof is complete. □

Remark 2.1. It can be noted that when the matrix A is Hermitian positive semidefinite, then Theorems 2.2, 2.3 and Lemma 2.1 reduce to Theorems 1.1, 1.2 and 1.3, respectively.

3. The singular value variations for the Rayleigh quotient

In this section, we study the singular value variations for the Rayleigh quotient matrix. Firstly, we provide a result corresponding to Theorem 1.4. The following lemma can be found in Li [9].

Lemma 3.1. ([9]) *Let $Z = (z_{ij})$ be an $n \times n$ doubly stochastic matrix and let $M = (\theta_{ij}) \in C^{n\times n}$. Then there exists a permutation τ of $\{1, 2, \cdots, n\}$ such that*

$$\sum_{i,j=1}^{n} \mid \theta_{ij} \mid z_{ij} \geq \sum_{i=1}^{n} \mid \theta_{\tau_{(i)},i} \mid .$$

Theorem 3.1. *Assume that $A \in C^{m\times n}$ and $\tilde{U}_k \in C^{m\times k}, \tilde{V}_k \in C^{n\times k}$ with $\tilde{U}_k^* \tilde{U}_k = \tilde{V}_k^* \tilde{V}_k = I_k$. Let*

$$R = A\tilde{V}_k - \tilde{U}_k T, \quad S = A^* \tilde{U}_k - \tilde{V}_k T^*, \quad T \in C^{k\times k}.$$

Suppose $\gamma_1, \gamma_2, \cdots, \gamma_k$ are the singular values of a matrix T and $\sigma_1, \sigma_2, \cdots, \sigma_n$ are the singular values of a matrix A . Then there exists a permutation τ of $\{1, 2, \cdots, m\}$ such that

$$\sqrt{\sum_{j=1}^{k} \mid \gamma_j - \sigma_{\tau_{(j)}} \mid^2} \leq \sqrt{\frac{\|R\|_F^2 + \|S\|_F^2}{2}}, \tag{3.1}$$

where $\sigma_{\tau_{(j)}} = 0$, if $\tau_{(j)} > n$.

Proof. Let

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^* \quad \text{and} \quad T = W\Gamma Q^*$$

be the singular value decompositions of A and T, respectively. Here, $U \in C^{m\times m}, V \in C^{n\times n}, W, Q \in C^{k\times k}$ are unitary, and $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n), \Gamma = \mathrm{diag}(\gamma_1, \gamma_2, \cdots, \gamma_k)$. Since the Frobenius norm is unitarily invariant, we get

$$\|R\|_F = \|U^*RQ\|_F = \| \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} (V^*\tilde{V}_kQ) - (U^*\tilde{U}_kW)\Gamma\|_F$$

and

$$\|S\|_F = \|V^*SW\|_F = \|(\Sigma, 0)(U^*\tilde{U}_kW) - (V^*\tilde{V}_kQ)\Gamma\|_F.$$

Without loss of generality, we may suppose $A = \begin{pmatrix} \Sigma \\ 0 \end{pmatrix}, T = \Gamma$. Then

$$R = \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} \tilde{V}_k - \tilde{U}_k\Gamma \quad \text{and} \quad S = (\Sigma, 0)\tilde{U}_k - \tilde{V}_k\Gamma.$$

Hence

$$\|R\|_F^2 = \sum_{j=1}^{k}\sum_{i=1}^{n} \mid \sigma_i v_{ij} - \gamma_j u_{ij} \mid^2 + \sum_{j=1}^{k}\sum_{i=n+1}^{m} \mid \gamma_j u_{ij} \mid^2 \tag{3.2}$$

and

$$\|S\|_F^2 = \sum_{j=1}^{k}\sum_{i=1}^{n} \mid \sigma_i u_{ij} - \gamma_j v_{ij} \mid^2, \tag{3.3}$$

where u_{ij} is the (i, j) element of a matrix $\tilde{U}_k$ and v_{ij} is the (i, j) element of a matrix $\tilde{V}_k$. Since

$$\begin{aligned}
& \sum_{j=1}^{k}\sum_{i=1}^{n} \mid \sigma_i v_{ij} - \gamma_j u_{ij} \mid^2 + \sum_{j=1}^{k}\sum_{i=1}^{n} \mid \sigma_i u_{ij} - \gamma_j v_{ij} \mid^2 \\
= & \sum_{j=1}^{k}\sum_{i=1}^{n} (\mid \sigma_i v_{ij} - \gamma_j u_{ij} \mid^2 + \mid \sigma_i u_{ij} - \gamma_j v_{ij} \mid^2) \\
= & \sum_{j=1}^{k}\sum_{i=1}^{n} [(\sigma_i^2 + \gamma_j^2)(\mid u_{ij} \mid^2 + \mid v_{ij} \mid^2) - 4\sigma_i\gamma_j Re(u_{ij}\overline{v}_{ij})] \\
\geq & \sum_{j=1}^{k}\sum_{i=1}^{n} [(\sigma_i^2 + \gamma_j^2)(\mid u_{ij} \mid^2 + \mid v_{ij} \mid^2) - 2\sigma_i\gamma_j(\mid u_{ij} \mid^2 + \mid v_{ij} \mid^2)] \\
= & \ 2\sum_{j=1}^{k}\sum_{i=1}^{n} (\sigma_i - \gamma_j)^2 (\tfrac{|u_{ij}|^2 + |v_{ij}|^2}{2}),
\end{aligned} \tag{3.4}$$

where $Re(a)$ stands for real part of a complex a, it follows from (3.2)-(3.4) that

$$\|R\|_F^2 + \|S\|_F^2 \geq 2\left\{\sum_{j=1}^{k}\sum_{i=1}^{n} (\sigma_i - \gamma_j)^2 (\frac{\mid u_{ij} \mid^2 + \mid v_{ij} \mid^2}{2}) + \sum_{j=1}^{k}\sum_{i=n+1}^{m} (0 - \gamma_j)^2 \frac{\mid u_{ij} \mid^2}{2}\right\}. \tag{3.5}$$

Let $\tilde{U} = (\tilde{U}_k, \tilde{U}_{m-k}) = (u_{ij}) \in C^{m\times m}, \tilde{V} = (\tilde{V}_k, \tilde{V}_{n-k}) = (v_{ij}) \in C^{n\times n}$ be unitary. Define a matrix $Z = (z_{ij}) \in C^{m\times m}$ by

$$z_{ij} = \begin{cases} \frac{|u_{ij}|^2 + |v_{ij}|^2}{2} & : \ i, j = 1, 2, \cdots, n \\ \mid u_{ij} \mid^2 & : \ otherwise. \end{cases} \tag{3.6}$$

Then it is easy to see that Z is an $m \times m$ doubly stochastic matrix. Now an $m \times m$ matrix $M = (\theta_{ij})$ is defined by

$$\theta_{ij} = \begin{cases} (\sigma_i - \gamma_j)^2 & : \quad i = 1, 2, \cdots, n, j = 1, 2, \cdots, k \\ (0 - \gamma_j)^2 & : \quad i = n+1, \cdots, m, j = 1, 2, \cdots, k \\ 0 & : \quad i = 1, 2, \cdots, m, j = k+1, \cdots, m. \end{cases} \tag{3.7}$$

In terms of (3.6) and (3.7), the inequality (3.5) can be rewritten as

$$\|R\|_F^2 + \|S\|_F^2 \geq 2 \sum_{j=1}^{m} \sum_{i=1}^{m} \theta_{ij} z_{ij}. \tag{3.8}$$

Applying Lemma 3.1 to (3.8), we know that there exists a permutation τ of $\{1, 2, \cdots, m\}$ such that

$$\|R\|_F^2 + \|S\|_F^2 \geq 2 \sum_{j=1}^{m} \theta_{\tau(j),j} = 2 \sum_{i=1}^{k} \theta_{\tau(j),j} = 2 \sum_{j=1}^{k} \mid \sigma_{\tau(j)} - \gamma_j \mid^2 .$$

This implies the inequality (3.1). The proof is complete. $\square$

Remark 3.1. Under the conditions of Theorem 3.1, Liu [4] showed that there is a permutation τ of $\{1, 2, \cdots, m\}$ such that

$$\sqrt{\sum_{j=1}^{k} \mid \gamma_j - \sigma_{\tau(j)} \mid^2} \leq \sqrt{\|R\|_F^2 + \|S\|_F^2}.$$

It is obvious that Theorem 3.1 improves Liu's bound by a factor $\frac{1}{\sqrt{2}}$.

Remark 3.2. When A is Hermitian positive semidefinite, Theorem 3.1 reduces to Theorem 1.4.

Obviously, Theorem 3.1 also holds if $T= M= \tilde{U}_k^* A \tilde{V}_k$. Next we provide some other bounds on the singular values of $M = \tilde{U}_k^* A \tilde{V}_k$ as an approximation to those of A under some unitarily invariant norms. First we introduce the definition of Q-norms. A unitarily invariant norm $|| \cdot ||$ is called a Q-norm (e.g. see [12]) if there exists another unitarily invariant norm $|| \cdot ||'$ such that $||Y|| = (||Y^*Y||')^{\frac{1}{2}}$, which is denoted by $|| \cdot ||_Q$. It is noted that the Ky-Fan p-k norm is a Q-norm for $p \geq 2$; in fact,

$$||Y||_{k;p} \equiv (\sum_{i=1}^{k} \sigma_i^p)^{1/p} = ||Y^*Y||_{k;p/2}^{\frac{1}{2}}$$

for $p \geq 2$ and $k = 1, ..., n$. It is easy to prove that both the spectral norm and the Frobenius norm are also Q-norms.

Lemma 3.2. ([10]) *Let A have the block form*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Then for a Q-norm $\| \cdot \|_Q$, we have

$$\|A\|_Q^2 \leq \|A_{11}\|_Q^2 + \|A_{12}\|_Q^2 + \|A_{21}\|_Q^2 + \|A_{22}\|_Q^2.$$

Theorem 3.2. *In Theorem* 3.1*, let* T *be replaced by* $M = \tilde{U}_k^* A \tilde{V}_k$*. Then there is a permutation* τ *of* $\{1, 2, \cdots, n\}$ *such that*

$$\|diag(\gamma_1 - \sigma_{\tau(1)}, \gamma_2 - \sigma_{\tau(2)}, \cdots, \gamma_k - \sigma_{\tau(k)})\|_Q \le \sqrt{\|R\|_Q^2 + \|S\|_Q^2}. \tag{3.9}$$

In particular

$$\sqrt{\sum_{i=1}^{k} (\gamma_i - \sigma_{\tau(i)})^2} \le \sqrt{\|R\|_F^2 + \|S\|_F^2}. \tag{3.10}$$

Proof. We take $\tilde{U}_{m-k} \in C^{m\times(m-k)}$ and $\tilde{V}_{n-k} \in C^{n\times(n-k)}$ such that $\tilde{U} = (\tilde{U}_k, \tilde{U}_{m-k})$ and $\tilde{V} = (\tilde{V}_k, \tilde{V}_{n-k})$ are unitary matrices. Then we get

$$\begin{aligned} \tilde{U}^* A \tilde{V} &= \begin{pmatrix} \tilde{U}_k^* A \tilde{V}_k & \tilde{U}_k^* A \tilde{V}_{n-k} \\ \tilde{U}_{m-k}^* A \tilde{V}_k & \tilde{U}_{m-k}^* A \tilde{V}_{n-k} \end{pmatrix} \\ &= \begin{pmatrix} M & 0 \\ 0 & \tilde{U}_{m-k}^* A \tilde{V}_{n-k} \end{pmatrix} + \begin{pmatrix} 0 & \tilde{U}_k^* A \tilde{V}_{n-k} \\ \tilde{U}_{m-k}^* A \tilde{V}_k & 0 \end{pmatrix}. \end{aligned} \tag{3.11}$$

For any unitarily invariant norm $\|\cdot\|$ we have

$$\|R\| = \|\tilde{U}^*(A\tilde{V}_k - \tilde{U}_k M)\| = \|\tilde{U}_{m-k}^* A \tilde{V}_k\| \tag{3.12}$$

and

$$\|S\| = \|\tilde{V}^*(A^*\tilde{U}_k - \tilde{V}_k M^*)\| = \|\tilde{V}_{n-k}^* A^* \tilde{U}_k\| = \|\tilde{U}_k^* A \tilde{V}_{n-k}\|. \tag{3.13}$$

From (3.11) and perturbation theory [1, Theorem 3.10] of the singular values, there is a permutation τ of $\{1, 2, \cdots, n\}$ such that for any unitarily invariant norm $\|\cdot\|$,

$$\begin{aligned} &\|\mathrm{diag}(\mu_1 - \sigma_{\tau(1)}, \mu_2 - \sigma_{\tau(2)}, \cdots, \mu_k - \sigma_{\tau(k)})\| \\ &\le \left\| \begin{pmatrix} 0 & \tilde{U}_k^* A \tilde{V}_{n-k} \\ \tilde{U}_{m-k}^* A \tilde{V}_k & 0 \end{pmatrix} \right\|. \end{aligned} \tag{3.14}$$

If the unitarily invariant norm $\|\cdot\|$ is assumed to be a Q-norm, then it follows from Lemma 3.2 and (3.14) that

$$\begin{aligned} &\|\mathrm{diag}(\mu_1 - \sigma_{\tau(1)}, \mu_2 - \sigma_{\tau(2)}, \cdots, \mu_k - \sigma_{\tau(k)})\|_Q^2 \\ &\le \|\tilde{U}_k^* A \tilde{V}_{n-k}\|_Q^2 + \|\tilde{U}_{m-k}^* A \tilde{V}_k\|_Q^2. \end{aligned} \tag{3.15}$$

Now the result follows from (3.12), (3.13) and (3.15). Since the Frobenius norm is a $Q-$norm, (3.10) holds. The proof is complete. □

From the above proof, it is easy to deduce the following bound under any unitarily invariant norm.

Corollary 3.1. *Under the assumption of Theorem* 3.2*, there is a permutation* τ *of* $\{1, 2, \cdots, n\}$ *such that for any unitarily invariant norm* $\|\cdot\|$*,*

$$\|diag(\gamma_1 - \sigma_{\tau(1)}, \gamma_2 - \sigma_{\tau(2)}, \cdots, \gamma_k - \sigma_{\tau(k)})\| \le \|R\| + \|S\|.$$

Remark 3.3. It is noted that the two permutations in (3.1) and (3.10) are different. One is for $\{1, 2, \cdots, m\}$ and the other is for $\{1, 2, \cdots, n\}$. However, the following example shows that Theorem 3.1 is not true for the permutation of $\{1, 2, \cdots, n\}$.

Example 3.1. Let

$$A = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{U}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tilde{V}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

where $\sigma_1 > 0, \sigma_2 > 0$. Taking

$$T = \tilde{U}_2^* A \tilde{V}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

then

$$R = A\tilde{V}_2 - \tilde{U}_2 T = A \quad \text{and} \quad S = A^* \tilde{U}_2 - \tilde{V}_2 T^* = T.$$

Hence for any permutation τ of $\{1,2\}$, the left-hand side of (3.1) is equal to $\sqrt{\sigma_1^2 + \sigma_2^2} = \|A\|_F$ and the right-hand side of (3.1) is $\|A\|_F/\sqrt{2}$. This example also illustrates that the bound in (3.10) is sharp.

Acknowledgments. This work was supported by NNSF of China (10671077) and NSF of Guangdong Province (031496 and 06025061).

References

[1] J.G. Sun, Matrix Perturbation Analysis, Science Press, Beijing, 2001.
[2] J.G. Sun, Eigenvalues of Rayleigh quotient matrices, *Numer. Math.*, **59** (1991), 603-614.
[3] X.G. Liu, On Rayleigh quotient theory for the eigenproblem and the singular value problem, *J. Comput. Math.*, **Supplementary issue** (1992), 216-224.
[4] X.G. Liu and Y. Xu, On the Rayleign quotient matrix, *Math. Numer. Sinica.*, **2** (1990), 208-213 (in Chinese).
[5] X.G. Liu, The perturbation of matrix eigenvalues associated with invariant subspaces, *J. Ocean Univ. Qingdao*, **19** (1989), 91-95.
[6] G.W. Stewart, Two simple residual bounds for the eigenvalues of a Hermitian matrix, *SIAM J. Matrix Anal. Appl.*, **12** (1991), 205-208.
[7] R.C. Li, On eigenvalues of a Rayleigh quotient matrix, *Linear Algebra Appl.*, **169** (1992), 249-255.
[8] R.C. Li, Accuracy of computed eigenvectors via optimizing a Rayleigh quotient, *BIT*, **44** (2004), 585-593.
[9] R.C. Li, Relative perturbation theory: I. Eigenvalue and singular value variations, *SIAM J. Matrix Anal. Appl.*, **19** (1998), 956-982.
[10] W. Li, Some new the perturbation bounds for subunitary polar factors, *Acta Math. Sinica*, **21** (2005), 1515-1520.
[11] W. Kahan, Numerical linear algebra, *Can. Math. Bull.*, **9** (1966), 757-801.
[12] R. Bhatia, Matrix Analysis, Springer Press, New York, 1997.
[13] R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.
[14] M.-X. Pang, New estimates for singular values of a matrix. *J. Comput. Math.*, **23** (2005), 199-204.
[15] W. Li and J.-X. Chen, The eigenvalue perturbation bound for arbitrary matrices. *J. Comput. Math.*, **24** (2006), 141-148.