On Eigenvalues of a Rayleigh Quotient Matrix

Ren-cang Li
Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago
Box 4348
Chicago, Illinois 60680

Submitted by Richard A. Brualdi

ABSTRACT

This note deals with the following problem: Let A be an $n \times n$ Hermitian matrix, and Q and $\tilde{Q}$ be two $n \times m$ $(n > m \geqslant 1)$ matrices both with orthonormal column vectors. How do the eigenvalues of the $m \times m$ Hermitian matrix $\tilde{Q}^H A \tilde{Q}$ differ from those of the $m \times m$ Hermitian matrix $Q^H A Q$? We give a positive answer to one of the unsolved problems raised recently by Sun.

In what follows, we will consider the following interesting problem concerning the spectral variation of a Rayleigh quotient matrix:

Let A be an $n \times n$ Hermitian matrix, and Q and $\tilde{Q}$ be two $n \times m$ $(n > m \geqslant 1)$ matrices both with orthonormal column vectors. How do the eigenvalues of the $m \times m$ Hermitian matrix $\tilde{Q}^H A \tilde{Q}$ differ from those of the $m \times m$ Hermitian matrix $Q^H A Q$?

Here the superscript H denotes the conjugate transpose of a matrix. This problem arises often in computation methods for symmetric matrix eigenvalues such as the power method, the QR method, the simultaneous method, and several other techniques now available. Thus it is of great importance from the point of view not only of theoretical analysis of some iterative algorithms but also of practical applications.

Much work has been done for this problem in various aspects so far, e.g., [2], [5], [9], [11], and [12]. In [5], Liu and Xu showed the following:

Let $\lambda_1, \dots, \lambda_m$ and $\tilde{\lambda}_1, \dots, \tilde{\lambda}_m$ be the eigenvalues (arranged in ascending order) of $Q^H A Q$ and of $\tilde{Q}^H A \tilde{Q}$, respectively. If

$$AQ = Q\Lambda, \qquad \textit{where} \quad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_m), \tag{1}$$

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

then

$$\max_{1 \leqslant j \leqslant m} |\lambda_j - \tilde{\lambda}_j| \leqslant 4\|A\|_2 \left\|\sin\Theta(Q,\tilde{Q})\right\|_2^2, \tag{2}$$

$$\sqrt{\sum_{j=1}^{m} |\lambda_j - \tilde{\lambda}_j|^2} \leqslant 4\|A\|_2 \left\|\sin\Theta(Q,\tilde{Q})\right\|_F^2, \tag{3}$$

where $\|\cdot\|_2$ *and* $\|\cdot\|_F$ *denote the spectral norm and Frobenius norm, respectively, and* $\Theta(Q,\tilde{Q})$ *is the angle between* Q *and* $\tilde{Q}$ *defined by*

$$\Theta(Q,\tilde{Q}) = \arcsin\left(I^{(m)} - \tilde{Q}^H Q Q^H \tilde{Q}\right)^{1/2}. \tag{4}$$

Here $I^{(m)}$ is the $m \times m$ identity matrix. A very interesting and important fact illustrated by (2) and (3) is that if Q differs from $\tilde{Q}$ by $O(\epsilon)$ in the sense of the angle between them, then there exists a one-one pairing between the eigenvalues of $\tilde{Q}^H A \tilde{Q}$ and those of $Q^H A Q$ such that every two corresponding eigenvalues differ from each other by $O(\epsilon^2)$. This coincides with the well-known fact for the case of $m = 1$ that if A is Hermitian with eigenvalue λ and corresponding eigenvector x and if $\sin\theta(x,\tilde{x}) = O(\epsilon)$, then

$$\frac{\tilde{x}^H A \tilde{x}}{\tilde{x}^H \tilde{x}} = \lambda + O(\epsilon^2)$$

(see, e.g., [10, Chapter 3, §7.1]).

The purpose of this note is to extend the inequalities (2) and (3) to cover *unitarily invariant norms* with or without the assumption (1). One of our results provides also a positive answer to one of the unsolved problems raised by Sun [12]. In the following, we use $\mathbb{C}^{p\times n}$ for the set of $p \times n$ complex matrices, and $\mathcal{U}_n \subset \mathbb{C}^{n\times n}$ for the set of $n \times n$ unitary matrices. To say that the norm $|||\cdot|||$ is *unitarily invariant* on $\mathbb{C}^{p\times n}$ means it satisfies, besides the usual properties of any norm, also (see, e.g., [10, Chapter 2, §3])

(1) $|||UAV||| = |||A|||$ for any $U \in \mathcal{U}_p$ and $V \in \mathcal{U}_n$;
(2) $|||A||| = \|A\|_2$ for any $A \in \mathbb{C}^{p\times n}$ with rank $A = 1$.

Two unitarily invariant norms used frequently are $\|\cdot\|_2$ and $\|\cdot\|_F$.

It is well known that any unitarily invariant norm $|||\cdot|||$ on $\mathbb{C}^{p\times n}$ corresponds to a symmetric gauge function $\Phi(\xi_1,\dots,\xi_N)$, where $N = \min\{p,n\}$, and vice versa. By extension according to this property, we can define a unitarily invariant norm $|||\cdot|||$ on $\mathbb{C}^{p_1\times n_1}$ $(p_1 \leqslant n,\ n_1 \leqslant n)$ consistent with the original one as $|||A||| = \Phi(\sigma_1,\dots,\sigma_{N_1},0,\dots,0)$ if $A \in \mathbb{C}^{p_1\times n_1}$ with singular values $\sigma_1,\dots,\sigma_{N_1}$ $(N_1 = \min\{p_1, n_1\})$. In this note, very often matrices with different dimensions enter our arguments together, so we make an agreement: assume we first have a matrix space with sufficiently large dimension $P\times N$ and with a unitarily invariant norm $|||\cdot|||$ on it; then by the extension mentioned, on every matrix space with dimension smaller there exists the extended unitarily invariant norm, denoted also by $|||\cdot|||$. In this way, we have (see [10, Chapter 5, §1])

$$|||CD||| \leqslant \begin{cases} \|C\|_2\,|||D||| \\ |||C|||\,\|D\|_2 \end{cases} \quad \text{for any} \quad C\in\mathbb{C}^{p\times n},\quad D\in\mathbb{C}^{n\times l}. \tag{5}$$

We begin our study by stating the following powerful lemma which can be proved by using the CS decomposition of a unitary matrix (see [1, 8, 10]).

Lemma 1. *Let $Q, \tilde{Q}\in\mathbb{C}^{n\times m}$ $(1\leqslant m\leqslant n-1)$ with $Q^HQ = \tilde{Q}^H\tilde{Q} = I^{(m)}$. Then there exist $U\in\mathscr{U}_n$ and $V, \tilde{V}\in\mathscr{U}_m$ such that:*

(1) *For* $2m \leqslant n$

$$UQV = \begin{matrix} m \\ m \\ n-2m \end{matrix}\overset{m}{\begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix}}, \qquad U\tilde{Q}\tilde{V} = \begin{matrix} m \\ m \\ n-2m \end{matrix}\overset{m}{\begin{pmatrix} \Gamma \\ \Sigma \\ 0 \end{pmatrix}}, \tag{6a}$$

where

$$\Gamma = \operatorname{diag}(\gamma_1,\dots,\gamma_m), \qquad \Sigma = \operatorname{diag}(\sigma_1,\dots,\sigma_m),$$

$$\gamma_i = \cos\theta_i,\ \sigma_i = \sin\theta_i,\quad 0\leqslant\theta_i\leqslant\frac{\pi}{2}, \qquad \text{for}\quad i = 1,2,\dots,m. \tag{6b}$$

(2) *For* $2m \geqslant n$

$$UQV = \begin{matrix} \\ n-m \\ 2m-n \\ n-m \end{matrix}\begin{matrix} n-m & 2m-n \\ \begin{pmatrix} I & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix} \end{matrix}, \qquad U\tilde{Q}\tilde{V} = \begin{matrix} \\ n-m \\ 2m-n \\ n-m \end{matrix}\begin{matrix} n-m & 2m-n \\ \begin{pmatrix} \Gamma & 0 \\ 0 & I \\ \Sigma & 0 \end{pmatrix} \end{matrix}, \tag{7a}$$

where

$$\Gamma = \operatorname{diag}(\gamma_1, \dots, \gamma_{n-m}), \qquad \Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_{n-m}),$$

$$\gamma_i = \cos\theta_i, \;\; \sigma_i = \sin\theta_i, \;\; 0 \leqslant \theta_i \leqslant \frac{\pi}{2}, \qquad \textit{for} \quad i = 1, 2, \dots, n-m. \tag{7b}$$

Assume for the moment that $2m \leqslant n$. Thus (6a) and (6b) hold. As V and $\tilde{V}$ are both unitary, the eigenvalues of $Q^H A Q$ and those of $V^H Q^H A Q V$ are exactly the same. So are those of $\tilde{Q}^H A \tilde{Q}$ and of $\tilde{V}^H \tilde{Q}^H A \tilde{Q} \tilde{V}$. By Wielandt's theorem [13] (cf. [3, 4, 14]), we have for any unitarily invariant norm $|||\cdot|||$

$$\left|\left|\left| \operatorname{diag}\left(\lambda_1 - \tilde{\lambda}_1, \dots, \lambda_m - \tilde{\lambda}_m\right) \right|\right|\right| \leqslant \left|\left|\left| V^H Q^H A Q V - \tilde{V}^H \tilde{Q}^H A \tilde{Q} \tilde{V} \right|\right|\right|. \tag{8}$$

Now, we are in a position to bound the right hand side of (8) by the angle between Q and $\tilde{Q}$. Denote by $B = UAU^H$. Then

$$\begin{aligned} V^H Q^H A Q V - \tilde{V}^H \tilde{Q}^H A \tilde{Q} \tilde{V} &= (I, 0, 0) B \begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix} - (\Gamma, \Sigma, 0) B \begin{pmatrix} \Gamma \\ \Sigma \\ 0 \end{pmatrix} \\ &= -(\Gamma - I, \Sigma, 0) B \begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix} - (I, 0, 0) B \begin{pmatrix} \Gamma - I \\ \Sigma \\ 0 \end{pmatrix} \\ &\quad -(\Gamma - I, \Sigma, 0) B \begin{pmatrix} \Gamma - I \\ \Sigma \\ 0 \end{pmatrix}. \end{aligned} \tag{9}$$

The m singular values of $(\Gamma - I, \Sigma, 0)$ as well as those of $(\Gamma - I, \Sigma, 0)^H$ are easily given by

$$\sqrt{(1-\gamma_i)^2 + \sigma_i^2} = \sqrt{2(1 - \cos\theta_i)} = 2 \sin\frac{\theta_i}{2} \leqslant \sqrt{2}\,\sin\theta_i \qquad \text{for} \quad i = 1, 2, \dots, m.$$

So it follows from (4) [whence $\sin\Theta(Q, \tilde{Q}) = \tilde{V} \operatorname{diag}(\sin\theta_1, \dots, \sin\theta_m) \tilde{V}^H$] that

$$|||(\Gamma - I, \Sigma, 0)||| = \left|\left|\left| \begin{pmatrix} \Gamma - I \\ \Sigma \\ 0 \end{pmatrix} \right|\right|\right| = 2 \left|\left|\left| \sin\tfrac{1}{2}\Theta(Q, \tilde{Q}) \right|\right|\right| \leqslant \sqrt{2} \left|\left|\left| \sin\Theta(Q, \tilde{Q}) \right|\right|\right|, \tag{10}$$

which, together with (9), (5), and $\|B\|_2 = \|A\|_2$, says

$$\left|\left|\left| V^H Q^H A Q V - \tilde{V}^H \tilde{Q}^H A \tilde{Q} \tilde{V} \right|\right|\right|$$

$$\leqslant 4\|A\|_2 \left|\left|\left| \sin \tfrac{1}{2}\Theta(Q,\tilde{Q}) \right|\right|\right| \left(1 + \left\| \sin \tfrac{1}{2}\Theta(Q,\tilde{Q}) \right\|_2\right) \tag{11a}$$

$$\leqslant 2\|A\|_2 \left|\left|\left| \sin \Theta(Q,\tilde{Q}) \right|\right|\right| \left(\sqrt{2} + \left\| \sin \Theta(Q,\bar{Q}) \right\|_2\right). \tag{11b}$$

We claim that (11a) and (11b) also hold for the case of $2m \geqslant n$, for which the proof based upon (7) is quite similar and is omitted here.

A consequence of (8) and (11) is

THEOREM 2. *Let A be an $n \times n$ Hermitian matrix, and let Q and $\tilde{Q}$ be two $n \times m$ $(n > m \geqslant 1)$ matrices with $Q^H Q = \tilde{Q}^H \tilde{Q} = I$. If the eigenvalues $\lambda_1, \dots, \lambda_m$ of $Q^H A Q$ and the eigenvalues $\tilde{\lambda}_1, \dots, \tilde{\lambda}_m$ of $\tilde{Q}^H A \tilde{Q}$ are arranged in ascending order, then for any unitarily invariant norm $|||\cdot|||$ we have*

$$\left|\left|\left| \operatorname{diag}\left(\lambda_1 - \tilde{\lambda}_1, \dots, \lambda_m - \tilde{\lambda}_m\right) \right|\right|\right|$$

$$\leqslant 4\|A\|_2 \left|\left|\left| \sin \tfrac{1}{2}\Theta(Q,\tilde{Q}) \right|\right|\right| \left(1 + \left\| \sin \tfrac{1}{2}\Theta(Q,\tilde{Q}) \right\|_2\right) \tag{12a}$$

$$\leqslant 2\|A\|_2 \left|\left|\left| \sin \Theta(Q,\tilde{Q}) \right|\right|\right| \left(\sqrt{2} + \left\| \sin \Theta(Q,\tilde{Q}) \right\|_2\right). \tag{12b}$$

Theorem 2 is derived without the assumption (1). But what will happen if (1) is indeed satisfies? Again we consider the case of $2m \leqslant n$, and a similar consideration for the case of $2m \geqslant n$ is omitted here.

We see that B has the form

$$B = \begin{pmatrix} \Lambda & 0 \\ 0 & B_{22} \end{pmatrix} \qquad \text{with} \quad B_{22} \in \mathbb{C}^{(n-m)\times(n-m)},$$

from which it follows that

$$\begin{aligned} V^H Q^H A Q V - \tilde{V}^H \tilde{Q}^H A \tilde{Q} \tilde{V} &= \Lambda - \Gamma^2 \Lambda - (\Sigma, 0) B_{22} \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} \\ &= \Sigma^2 \Lambda - (\Sigma, 0) B_{22} \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} \\ &= \Sigma \left[\Sigma \Lambda - (I, 0) B_{22} \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} \right]. \end{aligned} \tag{13}$$

Imitating the proof of Theorem 4.2 in Sun [12], we have

$$\left|\!\left|\!\left| \Sigma\Lambda - (I,0)B_{22}\begin{pmatrix}\Sigma\\0\end{pmatrix} \right|\!\right|\!\right| \leqslant \max_{\substack{1\leqslant i\leqslant m\\ m+1\leqslant j\leqslant n}} |\lambda_i - \lambda_j|\, |||\Sigma|||,$$

as clearly

$$\Sigma\Lambda - (I,0)B_{22}\begin{pmatrix}\Sigma\\0\end{pmatrix} = \Sigma(\Lambda - \alpha I) - (I,0)(B_{22} - \alpha I)\begin{pmatrix}\Sigma\\0\end{pmatrix} \qquad \text{for any } \alpha,$$

where we denote $\lambda(B_{22}) = \{\lambda_j \text{ for } j = m+1,\dots,n\}$. Therefore from (13)

$$\begin{aligned} &\left|\!\left|\!\left| V^H Q^H A Q V - \tilde{V}^H \tilde{Q}^H A \tilde{Q} \tilde{V} \right|\!\right|\!\right| \\ &\leqslant \max_{\substack{1\leqslant i\leqslant m\\ m+1\leqslant j\leqslant n}} |\lambda_i - \lambda_j| \cdot \|\Sigma\|_2\, |||\Sigma||| \\ &\leqslant \max_{\substack{1\leqslant i\leqslant m\\ m+1\leqslant j\leqslant n}} |\lambda_i - \lambda_j| \left|\!\left|\!\left| \sin\Theta(Q,\tilde{Q}) \right|\!\right|\!\right| \cdot \left\| \sin\Theta(Q,\tilde{Q}) \right\|_2. \end{aligned}$$

So we also have

THEOREM 3. *To the hypotheses of Theorem 2 add these: the equation (1) holds, and $\lambda(A) = \{\lambda_i$ for $i = 1,\dots,n\}$. Then for any unitarily invariant norm $|||\cdot|||$ we have*

$$\begin{aligned} &\left|\!\left|\!\left| \operatorname{diag}\left(\lambda_1 - \tilde{\lambda}_1, \dots, \lambda_m - \tilde{\lambda}_m\right) \right|\!\right|\!\right| \\ &\leqslant \max_{\substack{1\leqslant i\leqslant m\\ m+1\leqslant j\leqslant n}} |\lambda_i - \lambda_j| \left|\!\left|\!\left| \sin\Theta(Q,\tilde{Q}) \right|\!\right|\!\right| \cdot \left\| \sin\Theta(Q,\tilde{Q}) \right\|_2 \qquad (14a) \\ &\leqslant 2\|A\|_2 \left|\!\left|\!\left| \sin\Theta(Q,\tilde{Q}) \right|\!\right|\!\right| \cdot \left\| \sin\Theta(Q,\tilde{Q}) \right\|_2. \qquad (14b) \end{aligned}$$

The inequality (14b) for $|||\cdot||| = \|\cdot\|_2$ improves (2) due to Liu and Xu [5]. The inequality (14a) is, as a matter of fact, the unsolved Problem 5.3 of Sun [12]. The differences among the inequalities (12) and (14) show that the assumption (1) plays an essential role. As can be easily seen, approximately, if $\Theta(Q,\tilde{Q}) = O(\epsilon)$, then the right hand sides of (12) are still $O(\epsilon)$, while those of (14) are $O(\epsilon^2)$.

The author wishes to thank the anonymous referee for drawing his attention to [2], [9], [11], and [12], which led him to rewrite and improve Theorem 3 to its present form.

REFERENCES

1 G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins U.P., 1983.
2 S.-P. Han, Relative eigenvalues of Hermitian matrices, *Linear Algebra Appl.* 81:75–88 (1986).
3 A. J. Hoffman, and H. W. Wielandt, The variation of the spectrum of a normal matrix, *Duke Math. J.* 20:37–39 (1953).
4 V. B. Lidskii, The proper values of the sum and product of symmetric matrices, *Dokl. Akad. Nauk SSSR* 75:796–772 (1950).
5 Xin-guo Liu and Ya-ge Xu, On the Rayleigh quotient matrix (in Chinese), *Math. Numer. Sinica* 12(2):208–213 (1990).
6 L. Mirsky, Symmetric gauge functions and unitarily invariant norms, *Quart. J. Math. Oxford* 11:50–59 (1960).
7 B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
8 G. W. Stewart, On the perturbation of pseudoinverses, projections and linear square problem, *SIAM Rev.* 19:634–664 (1977).
9 G. W. Stewart, Two Simple Residual Bounds for the Eigenvalues of Hermitian Matrices, Technical Report CS-TR 2364, Dept. of Computer Science, Univ. of Maryland.
10 G. W. Stewart and Ji-guang Sun, *Matrix Perturbation Analysis*, Academic, New York, 1990.
11 Ji-guang Sun, Relative eigenvalue of a definite matrix pairs, *Linear Algebra Appl.* 139:253–267 (1990).
12 Ji-guang Sun, Eigenvalues of Rayleigh quotient matrices, *Numer. Math.*, submitted for publication.
13 H. Wielandt, An extremum property of sums of eigenvalues, *Proc. Amer. Math. Soc.* 6:106–110 (1955).
14 H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linear partielles Differentialgleichungen, *Math. Ann.* 71:441–479 (1912).

Received 5 February 1991; final manuscript accepted 25 April 1991