INEQUALITIES ON SINGULAR VALUES OF BLOCK TRIANGULAR MATRICES

CHI-KWONG LI AND ROY MATHIAS *

October 25, 2001

Abstract

We prove inequalities on singular values for 2×2 block triangular matrices. Using the results, we answer the three questions of Ando on Bloomfield-Watson type inequalities on eigenvalues and generalize the Kantorovich inequality.

Keywords: Eigenvalue, Kantorovich Inequality, Bloomfield-Watson Inequality, positive semidefinite matrix, triangular block matrix.
AMS(MOS) subject classification: 15A18, 15A42

1 Introduction

Let X be a $p \times q$ matrix and let $k = \min\{p, q\}$. Denote by $s(X) = (s_1(X), \ldots, s_k(X))$ the vector of decreasingly ordered singular values of X, i.e., $s_1(X) \geq \cdots \geq s_k(X)$ are the nonnegative square roots of the k largest eigenvalues of XX^*. For an $n \times n$ Hermitian matrix X let $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X))$ denote the vector of decreasingly ordered eigenvalues. Given two real vectors $x = (x_1, \ldots, x_k)$ and $y = (y_1, \ldots, y_k)$, we say that x is weakly majorized by y, denoted by $x \prec_w y$ if the sum of the m largest entries of x is not larger than that of y for $m = 1, \ldots, k$; for general background of the theory on majorization see [6]. The algebra of $n \times n$ complex matrices will be denoted by M_n.

In this note, we prove inequalities on singular values for 2×2 block triangular matrices. Using the results, we answer Ando's questions on Bloomfield-Watson type inequalities on eigenvalues, and generalize the Kantorovich inequality and some results of Demmel.

*Department of Mathematics, The College of William and Mary, Williamsburg, Virginia 23187, USA. E-mail: ckli@math.wm.edu, mathias@math.wm.edu. Both authors were partially supported by the NSF grants DMS-9704534 and DMS-0071994.

2 Main Theorem

Theorem 1 *Let* $A = \begin{pmatrix} R & 0 \\ S & T \end{pmatrix} \in M_n$ *be a block triangular matrix with singular values* $a_1 \geq \cdots \geq a_n$, *where* $R \in M_p$. *Let* $k = \min\{p, n-p\}$. *Then*

$$s(S) \prec_w (a_1 - a_n, \cdots, a_k - a_{n-k+1}). \tag{1}$$

If A *is invertible then*

$$s(T^{-1}SR^{-1}) \prec_w \left(a_n^{-1} - a_1^{-1}, \cdots, a_{n-k+1}^{-1} - a_n^{-1}\right), \tag{2}$$

$$s(SR^{-1}) \prec_w \frac{1}{2}\left(\frac{a_1}{a_n} - \frac{a_n}{a_1}, \cdots, \frac{a_k}{a_{n-k+1}} - \frac{a_{n-k+1}}{a_k}\right), \tag{3}$$

and

$$s(T^{-1}S) \prec_w \frac{1}{2}\left(\frac{a_1}{a_n} - \frac{a_n}{a_1}, \cdots, \frac{a_k}{a_{n-k+1}} - \frac{a_{n-k+1}}{a_k}\right). \tag{4}$$

Our proof of (1) relies on an elegant result of Thompson and Therianos [8]: Let

$$B = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix}$$

be an $n \times n$ Hermitian matrix with X being $q \times q$. Then for any indices $1 \leq i_1 < \cdots < i_m \leq q$ and $1 \leq j_1 < \cdots < j_m \leq n - q$

$$\sum_{l=1}^{m} \lambda_{i_l + j_l - l}(B) + \sum_{l=1}^{m} \lambda_{n-m+l}(B) \leq \sum_{l=1}^{m} \lambda_{i_l}(X) + \sum_{l=1}^{m} \lambda_{j_l}(Z).$$

Proof. Let S have singular values $s_1 \geq \cdots \geq s_k$. Note that the matrix $\tilde{A} = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$ has eigenvalues $a_1, \ldots, a_n, -a_n, \ldots, -a_1$, and is permutationally similar to

$$\begin{pmatrix} 0_p & S^* & R^* & 0 \\ S & 0_{n-p} & 0 & T \\ R & 0 & 0 & 0 \\ 0 & T^* & 0 & 0 \end{pmatrix} = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix},$$

where $Z = 0_n$ and $X \in M_n$ has the n eigenvalues $s_1, \ldots, s_k, 0, \ldots, 0, -s_k, \ldots, -s_1$. By the result of Thompson and Therianos, we have

$$\begin{aligned} -\sum_{j=1}^{m} s_j &= \sum_{j=1}^{m} \lambda_{n-m+j}(X) \\ &= \sum_{j=1}^{m} \lambda_{n-m+j}(X) + \sum_{j=1}^{m} \lambda_j(Z) \\ &\geq \sum_{j=1}^{m} \lambda_{n-m+j}(\tilde{A}) + \sum_{j=1}^{m} \lambda_{2n-j+1}(\tilde{A}) \\ &= \sum_{j=1}^{m} a_{n-m+j} - \sum_{j=1}^{m} a_j. \end{aligned}$$

Multiplying both sides by -1, we get (1).

If A is invertible, then

$$\hat{A} = (I_p \oplus -I_{n-p})A^{-1}(I_p \oplus -I_{n-p}) = \begin{pmatrix} R^{-1} & 0 \\ T^{-1}SR^{-1} & T^{-1} \end{pmatrix}$$

has singular values $a_n^{-1}, \ldots, a_1^{-1}$. Applying the inequalities (1) to $\hat{A}$, we get (2).

Next, note that

$$A\hat{A} = \begin{pmatrix} I_p & 0 \\ 2SR^{-1} & I_{n-p} \end{pmatrix}.$$

Suppose U and V are unitary matrices such that $U^*SR^{-1}V$ has r_j as the (j,j) entry for $j = 1, \ldots, k$, and all other entries zero, where $s(SR^{-1}) = (r_1, \ldots, r_k)$. Then $A\hat{A}$ has the same singular values as the matrix

$$(V \oplus U)^* A\hat{A}(V \oplus U) = \begin{pmatrix} I_p & 0 \\ 2U^*SR^{-1}V & I_{n-p} \end{pmatrix},$$

which is permutationally similar to a direct sum of I_{n-2k} and 2×2 matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 2r_j & 1 \end{pmatrix}, \qquad j = 1, \ldots, k. \tag{5}$$

Matrices of the form (5) have singular values

$$r_j + \sqrt{r_j^2 + 1} \quad \text{and} \quad \left(r_j + \sqrt{r_j^2+1}\right)^{-1} = \sqrt{r_j^2+1} - r_j.$$

Thus,

$$s(A\hat{A}) = (r_1 + \sqrt{r_1^2+1}, \ldots, r_k + \sqrt{r_k^2+1}, \underbrace{1, \ldots, 1}_{n-2k}, \sqrt{r_k^2+1} - r_k, \ldots, \sqrt{r_1^2+1} - r_1).$$

A well known result of Alfred Horn ([4], [5, Theorem 3.3.4], or [6, Chapter 9 H1]) gives

$$\prod_{j=1}^m s_j(A\hat{A}) \le \prod_{j=1}^m s_j(A)s_j(\hat{A}) = \prod_{j=1}^m (a_j/a_{n-j+1}), \qquad m = 1, \ldots, k,$$

i.e.,

$$\left(\ln s_1(A\hat{A}), \ldots, \ln s_k(A\hat{A})\right) \prec_w (\ln(a_1/a_n), \ldots, \ln(a_k/a_{n-k+1})).$$

Consider the function $f(t) = e^t - e^{-t}$ for $t > 0$. Then $f(\ln(r_j + \sqrt{r_j^2+1})) = 2r_j$ for $j = 1, \ldots, k$. Since f is increasing and convex on $(0, \infty)$ it preserves weak majorization ([6, Chapter 3, A.8 and C.1]), and so we have

$$\begin{aligned} 2(r_1, \ldots, r_k) &= \left(f(\ln s_1(A\hat{A})), \ldots, f(\ln s_k(A\hat{A}))\right) \\ &\prec_w (f(\ln(a_1/a_n)), \ldots, f(\ln(a_k/a_{n-k+1}))) \\ &= \left(\frac{a_1}{a_n} - \frac{a_n}{a_1}, \ldots, \frac{a_k}{a_{n-k+1}} - \frac{a_{n-k+1}}{a_k}\right), \end{aligned}$$

which is (3). Applying a similar argument to $\hat{A}A$, we get (4). $\square$

3 Questions of Ando

In [1], Ando raised several problems in connection with Bloomfield-Watson type inequalities for eigenvalues that arise in statistics (see also [3, Problem 7.3]). The following theorem answers his questions in the affirmative and extends scalar inequalities of Demmel [2, (62), (63), (65), (66)] to majorizations.

Theorem 2 *Let* $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ *be positive definite such that* $A_{11} \in M_p$. *Suppose* $k = \min\{p, n-p\}$ *and* A *has eigenvalues* $\lambda_1 \geq \cdots \geq \lambda_n$. *Then*

$$s(A_{22}^{-1/2} A_{21}) \prec_w \left(\sqrt{\lambda_1} - \sqrt{\lambda_n}, \ldots, \sqrt{\lambda_k} - \sqrt{\lambda_{n-k+1}} \right), \tag{6}$$

and

$$s(A_{21} A_{11}^{-1}) \prec_w \frac{1}{2} \left(\sqrt{\frac{\lambda_1}{\lambda_n}} - \sqrt{\frac{\lambda_n}{\lambda_1}}, \ldots, \sqrt{\frac{\lambda_k}{\lambda_{n-k+1}}} - \sqrt{\frac{\lambda_{n-k+1}}{\lambda_k}} \right). \tag{7}$$

Proof. Let $a_j = \sqrt{\lambda_j}$ for $j = 1, \ldots, n$. Let

$$B = \begin{pmatrix} R & 0 \\ S & T \end{pmatrix}$$

with $T = A_{22}$, $S = A_{22}^{-1/2} A_{21}$ and $R = (A_{11} - A_{12} A_{22}^{-1} A_{21})^{1/2}$. Then $A = B^* B$, and B has singular values $a_1, \ldots, a_n$. Applying Theorem 1 to the block triangular matrix B, we see that (6) is just (1).

Next, let

$$C = \begin{pmatrix} R & 0 \\ S & T \end{pmatrix}$$

with $R = A_{11}^{1/2}$, $S = A_{21} A_{11}^{-1/2}$, and $T = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{1/2}$. Then $A = CC^*$, and C has singular values $a_1, \ldots, a_n$. Applying Theorem 1 to the block triangular matrix C, we see that (7) is just (3). □

Suppose P is an $n \times k$ matrix such that $P^* P = I_k$. Then there exists a unitary U such that P is the first k columns of U. For any positive definite matrix A, we can apply Theorem 2 to the block matrix $U^* A U = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$. The results will take the more general form involving the eigenvalues and singular values of the matrices $P^*AP, P^*A^{-1}P, P^*A^2P$, etc. Many results in [1, 3] are stated in these forms, and they can be deduced from our results. We give a few examples in the following discussion. For easy reference and comparison, we state the next corollary in this manner.

Corollary 3 *Let* A *be a positive definite matrix with* $\lambda_1 \geq \cdots \geq \lambda_n$. *For any* $n \times k$ *matrix* P *such that* $P^*P = I_k$, *where* $2k \leq n$, *we have*

$$s(P^*AP - (P^*A^{-1}P)^{-1}) \prec_w \left(\left(\sqrt{\lambda_1} - \sqrt{\lambda_n} \right)^2, \ldots, \left(\sqrt{\lambda_k} - \sqrt{\lambda_{n-k+1}} \right)^2 \right) \tag{8}$$

and

$$s\left((P^*AP)^{-1}(P^*A^2P)(P^*AP)^{-1}\right) \prec_w \left(\frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n},\ldots,\frac{(\lambda_k+\lambda_{n-k+1})^2}{4\lambda_k\lambda_{n-k+1}}\right). \tag{9}$$

Proof. Let $a_1 \geq \cdots \geq a_n > 0$ so that $a_j^2 = \lambda_j$ for $j = 1, \ldots, n$.

To prove (8), we may assume that P is the first k columns of a unitary matrix U, and

$$U^*AU = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{10}$$

Then

$$P^*AP - (P^*A^{-1}P)^{-1} = A_{11} - (A_{11} - A_{21}^*A_{22}^{-1}A_{21}) = A_{21}^*A_{22}^{-1}A_{21} = (A_{22}^{-1/2}A_{21})^*(A_{22}^{-1/2}A_{21}).$$

By the majorization (6), and the fact that squaring preserves majorization (see [6, Chapter A.8 and C.1]), we have

$$\sum_{j=1}^m s_j(P^*AP - (P^*A^{-1}P)^{-1}) = \sum_{j=1}^m s_j^2(A_{22}^{-1/2}A_{21}) \leq \sum_{j=1}^m (a_j - a_{n-j+1})^2, \quad m = 1, \ldots, k.$$

Thus, (8) holds.

To prove (9), we may again assume that P is the first k columns of a unitary matrix U such that (10) holds. Then the left side of (9) is just

$$s(I_k + A_{11}^{-1}A_{21}A_{12}A_{11}^{-1}) = (1, \ldots, 1) + s(A_{21}A_{11}^{-2}A_{12}).$$

Using the square of (7), we have

$$(1,\ldots,1) + s\left(A_{21}A_{11}^{-2}A_{12}\right) \prec_w (1,\ldots,1) + \frac{1}{4}\left(\left(\frac{a_1}{a_n} - \frac{a_n}{a_1}\right)^2, \ldots, \left(\frac{a_k}{a_{n-k+1}} - \frac{a_{n-k+1}}{a_k}\right)^2\right),$$

which is the right side of (9). $\square$

We proved (8) and (9) by squaring (6) and (7). Ando proved (9) by another method in [1]. One may wonder whether it is possible to deduce (6) and (7) from (8) and (9) by taking square roots. It is not possible, since taking square roots does not preserve majorization.

Our bound (8) includes the inequality of Rao [7]:

$$\operatorname{tr}\left(P^*AP - (P^*A^{-1}P)^{-1}\right) \leq \sum_{j=1}^k \left(\sqrt{\lambda_j} - \sqrt{\lambda_{n-j+1}}\right)^2.$$

The inequality (9) includes the Kantorovich inequality. To see this, given a unit vector x, take $k = 1$ and take $P = A^{-1/2}x/(x^*A^{-1}x)^{1/2}$. Then we have the Kantorovich inequality:

$$(x^*Ax)(x^*A^{-1}x) = s_1((P^*AP)^{-1}(P^*A^2P)(P^*AP)^{-1}) \leq \frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n}.$$

In [1], Ando also asked whether the following is true for a positive definite matrix

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

with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, where $A_{11} \in M_{n-k}$ with $2k \leq n$:

$$\sum_{j=1}^{m} s_j(A_{22}^{-1/2} A_{21} A_{11}^{-1/2}) \leq \sum_{j=1}^{m} \frac{\lambda_j - \lambda_{n-j+1}}{\lambda_j + \lambda_{n-j+1}}, \qquad m = 1, \ldots, k.$$

The result is indeed true for $m = 1$ ([1], [2, Theorem 1]), but not in general:

Example 4 *Let*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \textit{with} \quad A_{11} = A_{22} = \begin{pmatrix} 6 & 0 \\ 0 & 3 \end{pmatrix} \quad \textit{and} \quad A_{12} = A_{21} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then A has eigenvalues $8, 4, 4, 2$ and $A_{22}^{-1/2} A_{21} A_{11}^{-1/2}$ has singular values $1/3, 1/3$. However

$$1/3 + 1/3 \not\leq 3/5 = (8-2)/10 + (4-4)/8.$$

References

[1] T. Ando, Bloomfield-Watson type inequalities for eigenvalues, notes of the lecture presented at the International Conference on Mathematical Analysis and its Applications, National Sun Yat-Sen University, Kaoshiung, Taiwan, January 2000.

[2] J. Demmel, The condition number of equivalence transformations that block diagonalize matrix pencils, SIAM J. Numer. Anal., 20 (1983), 599–610.

[3] S.W. Drury, S. Liu, C.-Y. Lu, S. Puntanen, and G.P.H. Styan, Some comments on several matrix inequalities with applications to canonical correlations, to appear in the Special Issue of Sankhya associated with "An International Conference in Honor of Professor C.R. Rao on the occasion of is 80th Birthday, Statistics: Reflections on the Past and Visions for the Future, The University of Texas at San Antonio, March 2000".

[4] A. Horn, On the singular values of a product of completely continuous operators, Proc. Nat. Acad. Sci. USA 36 (1950), 374-375.

[5] R. A. Horn and C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, New York, 1991.

[6] A.W. Marshall and I. Olkin, Inequalities: The Theory of Majorization and its Applications, Academic Press, New York, 1979.

[7] R.C. Rao, The inefficiency of least squares; extensions of the Kantorovich inequality, Linear Algebra Appl. 70 (1985), 249-255.

[8] R.C. Thompson and S. Therianos, Inequalities connecting the eigenvalues of a Hermitian matrices with the eigenvalues of complementary principal submatrices, Bull. Austral. Math. Soc. 6 (1972), 117-132.