J. of Inequal. & Appl., 1999, Vol. 3, pp. 137–142
Reprints available directly from the publisher
Photocopying permitted by license only

© 1999 OPA (Overseas Publishers Association) N.V.
Published by license under
the Gordon and Breach Science
Publishers imprint.
Printed in Malaysia.

Inequalities on the Singular Values of an Off-Diagonal Block of a Hermitian Matrix*

CHI-KWONG LI[†] and ROY MATHIAS

Department of Mathematics, College of William & Mary, Williamsburg, VA 23187, USA

(Received 26 November 1997; Revised 4 February 1998)

A majorization relating the singular values of an off-diagonal block of a Hermitian matrix and its eigenvalues is obtained. This basic majorization inequality implies various new and existing results.

Keywords: Hermitian matrix; Eigenvalue; Singular value; Majorization

AMS Subject Classification: 15A42; 15A60

1 INTRODUCTION

Let $\lambda_1(H) \geq \cdots \geq \lambda_n(H)$ denote the eigenvalues of an $n \times n$ Hermitian matrix H. For an $m \times n$ complex matrix X, let $\sigma_i(X) = \sqrt{\lambda_i(X^*X)}$ denote the ith singular value for $i = 1, \ldots, k$, where $k = \min\{m, n\}$, and let $\sigma(X) = (\sigma_1(A), \ldots, \sigma_k(A))$ be the vector of singular values of X. In [2], the following result was obtained as a generalization of a result in [6].

THEOREM 1 *Suppose H is an $n \times n$ positive definite matrix. Then for any $n \times k$ matrix X such that $X^*X = I_k$,*

$$\operatorname{tr}(X^*H^2X - (X^*HX)^2) \leq \frac{1}{4}\sum_{j=1}^{k}(\lambda_j(H) - \lambda_{n-j+1}(H))^2.$$

* This research was supported by grants from the NSF.

[†] Corresponding author. E-mail: ckli@math.wm.edu.

In fact, this result was conjectured by J. Durbin, and proved in [1] and [3], independently. This result is important in the context of studying the relative performance of the least squares estimator and the best linear unbiased estimator in a linear model [1]. Observe that if U is a unitary matrix of the form $[X \mid Y]$, then X^*HY is a $k \times n$ matrix, and

$$\operatorname{tr}(X^*H^2X - (X^*HX)^2) = \sum_{j=1}^{m} \sigma_j(X^*HY)^2,$$

where $m = \min\{k, n-k\}$.

In the following, we obtain a majorization result that will allow one to deduce a whole family of inequalities including Theorem 1. In Refs. [1–3], the proof of Theorem 1 was done using partial differentiation to locate the optimal matrix that yields the upper bound of $\operatorname{tr}(X^*H^2X - (X^*HX)^2)$. In our case, we use different approaches to give two proofs for our result – Theorem 2 – that connect our problem to other subjects.

We need some more definitions to state our result. Given two real (row or column) vectors $x, y \in \mathbf{R}^n$, we say that x is *weakly majorized* by y, denoted by $x \prec_w y$, if the sum of the k largest entries of x is not larger than that of y for each $k = 1, \ldots, n$. If in addition the sum of the entries of each of the vectors is the same then we say that x is *majorized* by y.

THEOREM 2 *Let H be an $n \times n$ Hermitian matrix. Then for any unitary matrix U of the form $[X \mid Y]$, where X is $n \times k$ matrix, we have*

$$\sigma(X^*HY) \prec_w \tfrac{1}{2}(\lambda_1(H) - \lambda_n(H), \ldots, \lambda_m(H) - \lambda_{n-m+1}(H)),$$

where $m = \min\{k, n-k\}$. Consequently, for any Schur convex increasing function $f: \mathbf{R}^m \rightarrow \mathbf{R}$, we have

$$f(\sigma(X^*HY)) \leq f\left(\tfrac{1}{2}(\lambda_1(H) - \lambda_n(H), \ldots, \lambda_m(H) - \lambda_{n-m+1}(H))\right).$$

Note that if we take $f(x) = \sum_{j=1}^{m} x_i^2$ in Theorem 2, we obtain Theorem 1. In fact, there are many other interesting Schur convex functions (see [5, Chapter 1] for details). For instance, $f(x) = \sum_{j=1}^{p} |x_j|^p$ with $p \geq 1$ and the kth elementary symmetric function $E_k(x_1, \ldots, x_m)$ with $1 \leq k \leq m$ are such examples.

2 PROOFS

We first give a proof of Theorem 2 using the theory of majorization (see [5] for the general background) and a reduction of the problem to the 2×2 case.

First proof of Theorem 2 Assume, without loss of generality, that $U = I$ and that $k \leq (n-k)$. Write

$$H = \begin{pmatrix} H_{11} & B \\ B^* & H_{22} \end{pmatrix}$$

where H_{11} is $k \times k$ and H_{22} is $(n-k) \times (n-k)$. Let

$$B = W\Sigma V$$

be a singular value decomposition of B, where V and W are unitary. Then the eigenvalues of the matrix

$$\tilde{H} = \begin{pmatrix} W^* & 0 \\ 0 & V \end{pmatrix} H \begin{pmatrix} W & 0 \\ 0 & V^* \end{pmatrix}$$

are the same as those of H. The 2×2 principal submatrix of $\tilde{H}$ lying in rows and columns i and $k+i$ is

$$\tilde{H}[i, k+i] = \begin{pmatrix} \tilde{h}_{ii} & \sigma_i(B) \\ \sigma_i(B) & \tilde{h}_{k+i,k+i} \end{pmatrix}, \quad i = 1, \ldots, k.$$

One easily checks that Theorem 2 is true when $n=2$ and $k=1$. As a result, if

$$\tilde{H}[i, k+i] = R_i^* \begin{pmatrix} \mu_i & 0 \\ 0 & \eta_i \end{pmatrix} R_i,$$

where $\mu_i \geq \eta_i$ and $R_i^* R_i = I_2$, then

$$\sigma_i(B) \leq (\mu_i - \eta_i)/2.$$

Let R be the $n \times n$ unitary matrix obtained from I_n by replacing $I_n[i, k+i]$ by R_i for all $i = 1, \ldots, k$. Then $(R^* \tilde{H} R)_{ii} = \mu_i$ and

$(R^*\tilde{H}R)_{k+i,k+i} = \eta_i$. Since the vector of diagonal entries of $R^*\tilde{H}R$ is majorized by the vector of eigenvalues of $R^*\tilde{H}R$ (e.g., see [5, Chapter 9, B.1]), for any $t = 1, \ldots, k$, we have

$$\sum_{i=1}^{t} \mu_i = \sum_{i=1}^{t} (R^*\tilde{H}R)_{ii} \leq \sum_{i=1}^{t} \lambda_i(H)$$

and

$$\sum_{i=1}^{t} \eta_i = \sum_{i=1}^{t} (R^*\tilde{H}R)_{k+i,k+i} \geq \sum_{i=1}^{t} \lambda_{n-i+1}(H).$$

It follows that

$$\sum_{i=1}^{t} \sigma_i(B) \leq \sum_{i=1}^{t} (\mu_i - \eta_i)/2 \leq \sum_{i=1}^{t} (\lambda_i(H) - \lambda_{n-i+1}(H))/2.$$

Next, we give a proof of Theorem 2 using the theory of the C-numerical range (e.g., see [4] and its references for the general background).

Second proof of Theorem 2 By the singular value decomposition, one can find unitary matrices V and W of appropriate sizes such that

$$(VX^*HYW)_{jj} = \sigma_j(X^*HY), \quad j = 1, \ldots, m.$$

Thus for any positive integer t with $1 \leq t \leq m$, if we let $C_t = \sum_{j=1}^{t} E_{k+j,j}$, where $\{E_{11}, E_{12}, \ldots, E_{nn}\}$ denotes the standard basis for $n \times n$ matrices, then

$$\begin{aligned}
\sum_{j=1}^{t} \sigma_j(X^*HY) &\leq \max\left\{ \left| \sum_{j=1}^{t} (RX^*HYS)_{jj} \right| : \ R,\ S \text{ unitary} \right\} \\
&\leq \max\left\{ \left| \sum_{j=1}^{t} (Z^*HZ)_{k+j,j} \right| : \ Z \text{ unitary} \right\} \\
&= \max\{ |\mathrm{tr}(Z^*HZC_t)| : \ Z \text{ unitary} \} \\
&= \max\{ |\mathrm{tr}(HZC_tZ^*)| : \ Z \text{ unitary} \},
\end{aligned}$$

which can be viewed as the H-numerical radius $r_H(C_t)$ of C_t (e.g., see [4] for the general background). Moreover, since

$$C_t = \begin{pmatrix} 0 & 0 \\ * & 0 \end{pmatrix}$$

is in the so-called shift block form and $(C_t + C_t^*)/2$ has eigenvalues

$$\overbrace{1/2, \ldots, 1/2}^{t}, 0, \ldots, 0, \overbrace{-1/2, \ldots, -1/2}^{t},$$

we have (e.g., see [4, (5.1) and (5.2)])

$$\begin{aligned} r_H(C_t) &= \max\{|\mathrm{tr}(HZ(C_t + C_t^*)Z^*)/2|: \ Z \text{ unitary}\} \\ &= \sum_{j=1}^{n} \lambda_j((C_t + C_t^*)/2)\lambda_j(H) \\ &= \sum_{j=1}^{t} (\lambda_j(H) - \lambda_{n-j+1}(H))/2. \end{aligned}$$

Remarks Suppose that $\lambda_1 \geq \cdots \geq \lambda_n$ are given real numbers and that k is a positive integer such that $1 < k < n$. Let $m = \min\{k, n-k\}$. One can construct 2×2 matrices H_i with eigenvalues $\lambda_i, \lambda_{n-m+i}$, and off-diagonal entries equal to $(\lambda_i - \lambda_{n-m+i})/2$. Applying a suitable permutation similarity to the matrix

$$H = H_1 \oplus \cdots \oplus H_m \oplus \mathrm{diag}(\lambda_{m+1}, \ldots, \lambda_{n-m})$$

will yield a matrix

$$\tilde{H} = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix},$$

where B is $k \times (n-k)$ such that

$$B_{ij} = \begin{cases} (\lambda_i - \lambda_{n-i+1})/2 & \text{if } 1 \leq i = j \leq k, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $\tilde{H}$ has eigenvalues $\lambda_1, \ldots, \lambda_n$. Thus, we see that our result in Theorem 2 is best possible.

In the context of statistics one is interested in real symmetric matrices. Since Theorem 2 is true for Hermitian matrices it is *a fortiori* true for real symmetric matrices. It cannot be improved in the case of real symmetric matrices either because the matrix constructed in the example above is a real symmetric matrix.

Acknowledgement

We thank Professor G. Styan for drawing our attention to [1,2], and Professor Z. Jia for sending us a copy of the paper.

References

[1] P. Bloomfield and G.S. Watson, The inefficiency of least squares, *Biometrika* **62** (1975), 121–128.
[2] Z. Jia, An extension of Styan's inequality (Chinese), *Gongcheng Shuxue Xuebao* (*J. Engineering Mathematics*) **13** (1996), 122–126.
[3] M. Knott, On the minimum efficiency of least squares, *Biometrika* **62** (1975), 129–132.
[4] C.K. Li, *C*-numerical ranges and *C*-numerical radii, *Linear and Multilinear Algebra* **37** (1994), 51–82.
[5] A.W. Marshall and I. Olkin, *Inequalities: The Theory of Majorization and its Applications*, Academic Press, 1979.
[6] G.P.H. Styan, On some inequalities associated with ordinary least squares and the Kantorovich inequality, *Acta Univ. Tampere, Ser. A.* **153** (1983), 158–166.

Note added in proof

X. Zhan has another proof of our main theorem, which is also obtained independently by R. Bhatia, F.C. Silva, P. Assouad and J.A. Dias da Silva.