Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

ELA

MATRIX INEQUALITIES BY MEANS OF EMBEDDING*

TIAN-GANG LEI†, CHING-WAH WOO‡, AND FUZHEN ZHANG§

Abstract. In this expository study some basic matrix inequalities obtained by embedding bilinear forms $\langle Ax, x\rangle$ and $\langle Ax, y\rangle$ into 2×2 matrices are investigated. Many classical inequalities are reproved or refined by the proposed unified approach. Some inequalities involving the matrix absolute value $|A|$ are given. A new proof of Ky Fan's singular value majorization theorem is presented.

Key words. Eigenvalue, Majorization, Matrix absolute value, Matrix inequality, Matrix norm, Normal matrix, Positive semidefinite matrix, Singular value, Spread, Wielandt inequality.

AMS subject classifications. 15A42, 15A63, 47B15.

1. Introduction. It has been evident that 2×2, ordinary or partitioned, matrices play an important role in various matrix problems. For example, the well-known Toeplitz-Hausdorff Theorem on the convexity of numerical range can be proved by a reduction to the 2×2 case; see, e.g., [6, p. 18]. While the sets of values x^*Ax and y^*Ax with some constraints on vectors x and y have been extensively studied as numerical ranges or fields of values, we shall inspect a number of matrix equalities and inequalities that involve the quadratic terms x^*Ax and x^*Ay through the standpoint of *embedding*, where $*$ denotes conjugate ($^-$) transpose (^{T}). Namely, we will embed x^*Ax and x^*Ay in 2×2 matrices of the forms $\begin{pmatrix} x^*Ax & \star \\ \star & \star \end{pmatrix}$ and $\begin{pmatrix} x^*Ay & \star \\ \star & \star \end{pmatrix}$, respectively, where $\star$ stands for some entries irrelevant to our discussions, so that the results on 2×2 matrices can be utilized to derive equalities or inequalities of x^*Ax and x^*Ay. This idea is further used to "couple" matrices A and X in the form $\begin{pmatrix} \star & \langle A, X\rangle \\ \star & \star \end{pmatrix}$ when a trace inequality involving $\operatorname{tr}(AX^*) = \langle A, X\rangle$ is to be studied.

As usual, we write $A \geq 0$ if A is a positive semidefinite matrix, i.e., $x^*Ax \geq 0$ for all vectors x of appropriate size. The notation $A \leq B$ or $B \geq A$ means that $B - A \geq 0$ for Hermitian A and B of the same size. For any matrix A, $|A|$ is the *matrix absolute value* of A, defined to be $(A^*A)^{1/2}$. Denote by $\langle u, v\rangle$ the inner product of vectors u and v in a vector space. In particular, for matrices A and B in the unitary space (symbolized by $\mathbb{M}_{m,n}$ or simply $\mathbb{M}_n$ if $m = n$) of all $m \times n$ complex matrices, $\langle A, B\rangle = \operatorname{tr}(B^*A)$, and for $x, y \in \mathbb{C}^n$, $\langle x, y\rangle = y^*x$.

We shall examine a variety of important matrix inequalities by a unified approach and obtain some new inequalities as well. We will then extend our studies to the matrix absolute values and Key Fan singular value majorization theorem. We must

* Received by the editors on 27 October 2003. Accepted for publication on 1 April 2004. Handling Editor: Miroslav Fiedler.

†Department of Mathematical and Physical Sciences, National Natural Science Foundation of China, Beijing 100085, China (leitg@mail.nsfc.gov.cn).

‡Department of Mathematics, City University of Hong Kong, Hong Kong, China (mawoo@cityu.edu.hk).

§Division of Math, Science, and Technology, Nova Southeastern University, Fort Lauderdale, Florida 33314, USA (zhang@nova.edu) and Shenyang Normal University, Shenyang, China.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

ELA

point out that many inequalities in this paper are classical and they have been proved in a number of different ways.

2. Embedding approach. We begin with a lemma in which the inequalities may have appeared in a fragmentary literature. For example, (2.1) is the Corollary of Theorem 1 in [10].

LEMMA 2.1. *Let* $A = \left(\begin{smallmatrix} a & b \\ \bar{b} & c \end{smallmatrix}\right)$ *be a* 2×2 *Hermitian matrix and let* α *and* β *be the (necessarily real) eigenvalues of* A *with* $\alpha \geq \beta$. *Then*

$$2|b| \leq \alpha - \beta. \tag{2.1}$$

Further, if A *is positive definite, that is, if* $\alpha \geq \beta > 0$, *then*

$$\frac{|b|}{\sqrt{ac}} \leq \frac{\alpha - \beta}{\alpha + \beta}, \tag{2.2}$$

$$\frac{|b|}{a} \leq \frac{\alpha - \beta}{2\sqrt{\alpha\beta}}, \tag{2.3}$$

and

$$\frac{|b|}{\sqrt{c}} \leq \sqrt{\alpha} - \sqrt{\beta}. \tag{2.4}$$

Proof. The inequalities (2.1) and (2.2) follow from the observation that

$$\alpha,\ \beta = \frac{(a+c) \pm \sqrt{(a-c)^2 + 4|b|^2}}{2}.$$

To show (2.3), notice that for any real parameter t

$$(\alpha + \beta)t - \alpha\beta \leq \frac{(\alpha+\beta)^2}{4\alpha\beta} t^2.$$

Put $t = a$. Replace $\alpha + \beta$ with $a + c$, $\alpha\beta$ with $ac - |b|^2$ on the left hand side.

To prove (2.4), use, for all $t > 0$,

$$(\alpha + \beta) - \alpha\beta t \leq t^{-1} + (\sqrt{\alpha} - \sqrt{\beta})^2.$$

Set $t = c^{-1}$ and $\alpha + \beta = a + c$, $\alpha\beta = ac - |b|^2$ on the left hand side. ∎

REMARK 1. We shall frequently use an equivalent form of (2.2):

$$|b|^2 \leq \left(\frac{\alpha - \beta}{\alpha + \beta}\right)^2 ac. \tag{2.5}$$

In addition, (2.3) holds for c in place of a. This reveals the inequality

$$4|b| \leq \frac{\alpha^2 - \beta^2}{\sqrt{\alpha\beta}}. \tag{2.6}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

In the same spirit, from (2.4), one obtains

$$\sqrt{2}|b| \le (\sqrt{\alpha} - \sqrt{\beta})\sqrt{\alpha + \beta}. \tag{2.7}$$

It is worth noticing that, as $\alpha - \beta$ is considered to be the spread of the Hermitian matrix A, the above inequalities (2.1)–(2.7) give lower bounds for the spreads of the matrices A, A^2, $A^{1/2}$ in terms of the entries of A.

We proceed to inspect, using embedding techniques, some matrix equalities and inequalities that have frequently made their appearance; that is, we formulate a matrix inequality in terms of a sesquilinear form involving $\langle Ax, x\rangle$ or $\langle Ax, y\rangle$ as an inequality involving the entries of a matrix or a submatrix of the original matrix.

The Cauchy-Schwarz inequality. The classical Cauchy-Schwarz inequality (see, e.g., [5, p. 261]) states that for any n-column complex vectors x, $y \in \mathbb{C}^n$,

$$|\langle x, y\rangle|^2 \le \langle x, x\rangle\langle y, y\rangle. \tag{2.8}$$

Proof. This well-known inequality is traditionally proved by examining the discriminant of the quadratic function $\langle x + ty, x + ty\rangle$ in t. We now notice that

$$(x, y)^*(x, y) = \begin{pmatrix} x^* \\ y^* \end{pmatrix} (x, y) = \begin{pmatrix} x^*x & x^*y \\ y^*x & y^*y \end{pmatrix} \ge 0.$$

The inequality follows at once by taking the determinant of the 2×2 matrix. Equality in (2.8) occurs if and only if the rank of (x, y) is 1; that is, x and y are linearly dependent. □

The Wielandt inequality. Let A be a nonzero n-square positive semidefinite matrix having eigenvalues $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$. The Wielandt inequality (see, e.g., [5, p. 443]) asserts that for all orthogonal x, $y \in \mathbb{C}^n$,

$$|x^*Ay|^2 \le \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 (x^*Ax)(y^*Ay). \tag{2.9}$$

Proof. Inequality (2.9) involves the quadratic terms x^*Ax, x^*Ay, and y^*Ay. It is natural for us to think of the 2×2 matrix

$$\begin{pmatrix} x^*Ax & x^*Ay \\ y^*Ax & y^*Ay \end{pmatrix}.$$

If $\lambda_n = 0$, (2.9) follows immediately by taking the determinant. Let $\lambda_n > 0$ and M be the 2×2 matrix. Then $M = (x, y)^*A(x, y)$, which is bounded from below by $\lambda_n(x, y)^*(x, y)$ and from above by $\lambda_1(x, y)^*(x, y)$. We may assume that x and y are orthonormal by scaling both sides of (2.9). Then $\lambda_n I_2 \le M \le \lambda_1 I_2$ and thus the eigenvalues λ and μ of M with $\lambda \ge \mu$ are contained in $[\lambda_n, \lambda_1]$. Therefore $\frac{\lambda-\mu}{\lambda+\mu} \le \frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}$ since $\frac{t-1}{t+1}$ is monotone in t.

An application of (2.5) to M results in

$$|x^*Ay|^2 \le \left(\frac{\lambda - \mu}{\lambda + \mu}\right)^2 (x^*Ax)(y^*Ay) \le \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^2 (x^*Ax)(y^*Ay). \qquad \square$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

ELA

In a similar manner, by (2.3), we have

$$|x^*Ay| \le \frac{\lambda-\mu}{2\sqrt{\lambda\mu}}(x^*Ax), \quad |x^*Ay| \le \frac{\lambda-\mu}{2\sqrt{\lambda\mu}}(y^*Ay).$$

Note that $\frac{\lambda-\mu}{2\sqrt{\lambda\mu}} = \frac{1}{2}\left(\sqrt{\frac{\lambda}{\mu}} - \sqrt{\frac{\mu}{\lambda}}\right)$. Since $t - \frac{1}{t}$ is an increasing function in t, we have

$$|x^*Ay| \le \frac{\lambda_1-\lambda_n}{2\sqrt{\lambda_1\lambda_n}} \min\{x^*Ax,\ y^*Ay\} \tag{2.10}$$

and, likewise, by (2.4),

$$|x^*Ay|^2 \le (\sqrt{\lambda_1} - \sqrt{\lambda_n})^2 \min\{x^*Ax,\ y^*Ay\}. \tag{2.11}$$

REMARK 2. The first part of the above proof is essentially the same as in the literature (see, e.g., [5, pp. 441–442]). We include it for completeness. The proof also shows that the spectrum of the matrix M is contained in the interval $[\lambda_n, \lambda_1]$. In fact, for any n-square complex matrix A, with M defined as above for orthonormal x and y, the numerical range of M is contained in that of A. This is because, for $z = (a, b)^T \in \mathbb{C}^2$,

$$z^*Mz = z^*(x, y)^*A(x, y)z = (ax + by)^*A(ax + by).$$

Note that if z is a unit vector in $\mathbb{C}^2$, then so is $ax+by$ in $\mathbb{C}^n$ as x and y are orthonormal.

Refinement of an inequality. Consider the 2×2 partitioned matrix

$$M = \begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \ge 0,$$

where $A \in \mathbb{M}_m$, $C \in \mathbb{M}_n$ are Hermitian, $B \in \mathbb{M}_{m,n}$. For $x \in \mathbb{C}^m$, $y \in \mathbb{C}^n$, let

$$N = \begin{pmatrix} x^* & 0 \\ 0 & y^* \end{pmatrix} M \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} = \begin{pmatrix} x^*Ax & x^*By \\ y^*B^*x & y^*Cy \end{pmatrix}.$$

If x and y are unit, let U be an $(m+n) \times (m+n)$ unitary matrix with the 1st column $\binom{x}{0}$ and the 2nd column $\binom{0}{y}$. It follows that

$$U^*MU = \begin{pmatrix} N & \star \\ \star & \star \end{pmatrix} = \begin{pmatrix} x^*Ax & x^*By & \star \\ y^*B^*x & y^*Cy & \star \\ & \star & \star \end{pmatrix}.$$

If the eigenvalues of N are α and β, then, by the interlacing eigenvalue theorem (see, e.g., [13, p. 222]), α and β are contained in $[\mu, \lambda]$, where λ and μ are the largest and smallest eigenvalues of M, respectively. We may assume $\mu > 0$. By (2.5),

$$|x^*By|^2 \le \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2 (x^*Ax)(y^*Cy) \le \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2 (x^*Ax)(y^*Cy). \tag{2.12}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

Inequality (2.12) is stronger than the one in Theorem 7.7.7(a) of [5, p. 473] or Theorem 6.26 [13, p. 203].

Example. Let

$$M = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 3 & 1 \\ 0 & 1 & 1 & 3 \end{pmatrix}.$$

Then M is a positive definite matrix having eigenvalues

$$\mu = \frac{3-\sqrt{5}}{2}, \quad \frac{3+\sqrt{5}}{2}, \quad \frac{5-\sqrt{13}}{2}, \quad \frac{5+\sqrt{13}}{2} = \lambda,$$

and so

$$\left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2 = 0.70045...$$

By setting

$$A = B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix},$$

we have

$$|x^*By|^2 \le 0.71(x^*Ax)(y^*Cy), \text{ for all } x, y \in \mathbb{C}^2.$$

Remark 3. There is no fixed universal scalar less than 1 that fits in (2.12) in the place of $\left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2$ (or $\left(\frac{\lambda_1-\lambda_n}{\lambda_1+\lambda_n}\right)^2$ in (2.9)), for $\left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2$ can be 1.

A theorem of Mirsky. The spreads of Hermitian matrices have been studied by many authors (see, e.g., [11]). Recall that the spread of a Hermitian matrix A is defined to be $s(A) = \lambda_{\max} - \lambda_{\min}$, where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of A, respectively. It is shown in [9] and [10] that

$$s(A) = 2 \sup_{u,v} |u^*Av|, \tag{2.13}$$

where the upper bound is taken with respect to all pairs of orthonormal u, v.

Proof. To show (2.13), place the term u^*Av in a 2×2 matrix as follows: Let

$$M = (u,v)^*A(u,v) = \begin{pmatrix} u^*Au & u^*Av \\ v^*Au & v^*Av \end{pmatrix}.$$

Let U be a unitary matrix with u as the first column and v the second. Then

$$U^*AU = \begin{pmatrix} M & \star \\ \star & \star \end{pmatrix}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

ELA

By the interlacing eigenvalue theorem, $|\lambda_{\max} - \lambda_{\min}| \geq |\lambda - \mu|$, where λ and μ are the eigenvalues of M. On the other hand $|\lambda - \mu| \geq 2|u^*Av|$ by (2.1). It follows that $s(A) \geq 2|u^*Av|$. For the other direction of the inequality, that $s(A) \leq 2\sup_{u,v}|u^*Av|$ is seen, as in [10], by taking $u = \frac{1}{\sqrt{2}}(x+iy)$ and $v = \frac{1}{\sqrt{2}}(x-iy)$, where x and y are the orthonormal eigenvectors belonging to $\lambda_{\max}$ and $\lambda_{\min}$, respectively. ∎

REMARK 4. Equation (2.13) is proved in two separate papers. The inequality "$\geq$" follows from a discussion on the results of normal matrices in [9] (see Eq. (6)), while "$\leq$" is shown in [10].

A theorem of Bellman. Let $A > 0$, i.e., A is positive definite. Then as shown in [3],

$$\langle A^{-1}x, x\rangle = \max_y \left[2\,\mathrm{Re}(x,y) - (Ay,y)\right]. \tag{2.14}$$

Proof. The proof suggested in [3] is by integration and limit process. Now

$$\begin{pmatrix} A^{-1} & I \\ I & A \end{pmatrix} \geq 0 \quad \Rightarrow \quad \begin{pmatrix} x^*A^{-1}x & x^*y \\ y^*x & y^*Ay \end{pmatrix} \geq 0.$$

By pre- and post-multiplication of $(1,-1)$ and $(1,-1)^T$, respectively, we have

$$(1,-1)\begin{pmatrix} x^*A^{-1}x & x^*y \\ y^*x & y^*Ay \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = x^*A^{-1}x + y^*Ay - 2\,\mathrm{Re}(x^*y) \geq 0$$

or

$$x^*A^{-1}x \geq 2\,\mathrm{Re}(x^*y) - y^*Ay.$$

Thus

$$x^*A^{-1}x \geq \max_y \left[2\,\mathrm{Re}(x^*y) - y^*Ay\right].$$

Equality follows by taking $y = A^{-1}x$. ∎

An interesting application of the representation (2.14) is to deduce the well-known matrix inequality

$$A \geq B > 0 \Rightarrow B^{-1} \geq A^{-1}.$$

Partitioned matrices. Let $M = \left(\begin{smallmatrix} A & B \\ B^* & C \end{smallmatrix}\right) \geq 0$ and let λ and μ be the largest and smallest eigenvalues of M, respectively. If A is nonsingular, then (see, e.g., [14])

$$B^*A^{-1}B \leq \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2 C. \tag{2.15}$$

If A, B and C are all n-square, taking the determinants of both sides reveals

$$|\det B|^2 \leq \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^{2n} \det A\ \det C,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

which yields

$$|\det B|^2 \le \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2 \det A \ \det C. \tag{2.16}$$

In contrast, if we let

$$N = \begin{pmatrix} \det A & \det B \\ \det B^* & \det C \end{pmatrix}$$

(note that $M \ge 0 \Rightarrow N \ge 0$) and let α and β be the eigenvalues of N, then by (2.5),

$$|\det B|^2 \le \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2 \det A \ \det C. \tag{2.17}$$

Inequalities (2.16) and (2.17) seem unrelated, and both can be compared to the well-known but weaker determinantal inequality $|\det B|^2 \le \det A \ \det C$.

REMARK 5. Inequalities (2.16) and (2.17) are not comparable. Take

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & 0 & 4 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} \det A & \det B \\ \det B^* & \det C \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 8 \end{pmatrix}$$

Then the 4×4 matrix is singular, whereas the 2×2 is nonsingular. Thus

$$\left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2 < 1 = \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2.$$

Take

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 100 & 0 \\ 1 & 1 & 0 & 100 \end{pmatrix}, \quad \begin{pmatrix} \det A & \det B \\ \det B^* & \det C \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 10000 \end{pmatrix}.$$

The eigenvalues of the 4×4 matrix are $\lambda = 100$, 55.477..., 45.523..., $\mu = 1$, while the eigenvalues of the 2×2 matrix are $\alpha = 10000, \beta = 1$. Upon computation,

$$0.9996... = \left(\frac{\alpha-\beta}{\alpha+\beta}\right)^2 > \left(\frac{\lambda-\mu}{\lambda+\mu}\right)^2 = 0.9607....$$

Positivity and inner product. Let $X \in \mathbb{M}_n$. It is well-known that $X \ge 0 \Leftrightarrow \langle X, Y\rangle \ge 0$ for all n-square $Y \ge 0$. It would be tempting to generalize the statement for block matrices. Theorem 7.7.7 (a) and (d) in [5, p. 473] has shed a little light on this. In fact, by writing $x^*Ay = \operatorname{tr}(yx^*A) = \langle A, xy^*\rangle$ and with slight modification of the form of Theorem 7.7.7(a), we get

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \ge 0 \quad \Leftrightarrow \quad \begin{pmatrix} \langle A, xx^*\rangle & \langle B, xy^*\rangle \\ \langle B^*, yx^*\rangle & \langle C, yy^*\rangle \end{pmatrix} \ge 0, \tag{2.18}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

for all x and y of appropriate sizes. This generalizes to the following result: Let $A \in \mathbb{M}_m$, $B \in \mathbb{M}_{m,n}$, and $C \in \mathbb{M}_n$. Then

$$\begin{pmatrix} A & B \\ B^* & C \end{pmatrix} \geq 0 \quad \Leftrightarrow \quad \begin{pmatrix} \langle A, P\rangle & \langle B, Q\rangle \\ \langle B^*, Q^*\rangle & \langle C, R\rangle \end{pmatrix} \geq 0, \tag{2.19}$$

whenever the conformally partitioned matrix $\left(\begin{smallmatrix} P & Q \\ Q^* & R \end{smallmatrix}\right) \geq 0$.

Proof. (2.18) ensures "$\Leftarrow$" in (2.19) by taking $P = xx^*$, $Q = xy^*$, and $R = yy^*$. For the other direction, let W be an $(m+n) \times (m+n)$ matrix such that

$$\begin{pmatrix} P & Q \\ Q^* & R \end{pmatrix} = WW^*.$$

Let U and V be the matrices consisting of the first m rows of W and the remaining rows of W, respectively, and denote the ith column of U by u_i and the ith column of V by v_i for each i. Then

$$P = \sum_{i=1}^{m+n} u_i u_i^*, \quad Q = \sum_{i=1}^{m+n} u_i v_i^*, \quad R = \sum_{i=1}^{m+n} v_i v_i^*.$$

By using (2.18) again, the block matrix with the inner products in (2.19), when written as a sum of $(m+n)$ positive semidefinite 2×2 matrices, is positive semidefinite. $\square$

The equivalence statement (2.19) is seen in [2, p. 20] with a different proof by the cone property of self-duality.

Note that the positive semidefiniteness of the block matrix with inner products as entries in (2.19) implies the trace inequality

$$|\operatorname{tr}(BQ^*)|^2 \leq \operatorname{tr}(AP)\operatorname{tr}(CR).$$

This trace inequality will be extensively used later. As an application, for any positive definite $X \in \mathbb{M}_m$ and for any A, $B \in \mathbb{M}_{m,n}$, since block matrices

$$\begin{pmatrix} X & I \\ I & X^{-1} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} AA^* & AB^* \\ BA^* & BB^* \end{pmatrix}$$

are both positive semidefinite, we obtain the existing trace inequality

$$|\operatorname{tr}(A^*B)|^2 \leq \operatorname{tr}(A^*XA)\operatorname{tr}(B^*X^{-1}B).$$

REMARK 6. The above idea of coupling matrices in the form $\langle X, Y\rangle$ in block matrices may be extended and rephrased in terms of completely positive *bilinear* maps. It has been evident that completely positive linear maps (see, e.g., [1]) are a useful tool for deriving matrix or operator inequalities. In view of the matrix Kronecker product and Hadamard product as bilinear matrix forms, we consider a bilinear map f from $\mathbb{M}_n \times \mathbb{M}_n$ to a matrix space (or a number field). We say that f is positive if $f(X,Y) \geq 0$ whenever $X, Y \geq 0$, and further call f completely positive if

$$\begin{pmatrix} f(A,P) & f(B,Q) \\ f(B^*,Q^*) & f(C,R) \end{pmatrix} \geq 0 \quad \text{whenever} \quad \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}, \begin{pmatrix} P & Q \\ Q^* & R \end{pmatrix} \geq 0.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

It is immediate that for each fixed $X \geq 0$ in $\mathbb{M}_n$, $f_X(Y) = f(X,Y)$ is a completely positive linear map. Similarly $f_Y(X)$ is a completely positive linear map for fixed $Y \geq 0$. This definition may generalize to block matrices of higher dimensions.

Since the Kronecker product ($\otimes$) and the Hadamard product ($\circ$) of positive semidefinite matrices are positive semidefinite and because a principal submatrix of a positive semidefinite matrix is positive semidefinite, we see that $f(X,Y) = X \otimes Y$ and $f(X,Y) = X \circ Y$ are completely positive bilinear maps. We saw earlier in (2.19) that $f(X,Y) = \langle X, Y\rangle$ is the same kind.

3. The matrix absolute value $|A|$**.** For any matrix $A \in \mathbb{M}_{m,n}$, the matrix absolute value (or modulus) of A is defined to be $|A| = (A^*A)^{1/2}$. The matrix absolute value has many interesting properties. For instance, $|A^{-1}| = |A^*|^{-1}$ if A is square and invertible, and $|A \otimes B| = |A| \otimes |B|$ for A and B of any sizes. If $A = UDV$ is a singular value decomposition of a square A, where U and V are unitary, then $|A| = V^*DV$ and $|A^*| = UDU^*$. It is easy to check that A is normal, i.e., $A^*A = AA^*$, if and only if $|A| = |A^*|$. In addition, if A is normal, then for all $\alpha,\ \beta > 0$

$$|\alpha A + \beta A^*| \leq \alpha|A| + \beta|A^*|. \tag{3.1}$$

In particular, by taking $\alpha = \beta = \frac{1}{2}$,

$$|A + A^*| \leq |A| + |A^*|, \tag{3.2}$$

which does not hold in general for non-normal matrices [12].

In what follows we make use of (2.19) with the block matrix

$$\begin{pmatrix} |A^*| & A \\ A^* & |A| \end{pmatrix}, \tag{3.3}$$

which is easily seen to be positive semidefinite by singular value decomposition.

Notice that (2.19), applied to (3.3), gives the trace inequality

$$|\operatorname{tr}(AQ^*)|^2 \leq \operatorname{tr}(|A^*|P)\ \operatorname{tr}(|A|R) \tag{3.4}$$

whenever

$$\begin{pmatrix} P & Q \\ Q^* & R \end{pmatrix} \geq 0.$$

Following are immediate consequences of (3.4) for a square A:

a) Setting $P = Q = R = I$ yields $|\operatorname{tr} A| \leq \operatorname{tr}|A|$ (see, e.g., [13, p. 260]).

b) Putting $P = Q = R = J$, the matrix all whose entries are 1, we have

$$|\Sigma(A)|^2 \leq \Sigma(|A^*|)\ \Sigma(|A|), \tag{3.5}$$

where $\Sigma(X) = \sum_{ij} x_{ij}$, the sum of all entries of $X = (x_{ij})$. Note that (3.5) implies

$$|\Sigma(A)| \leq \Sigma(|A|), \quad \text{if } A \text{ is normal.} \tag{3.6}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

c) Letting $P = |A|$, $Q = A^*$, and $R = |A^*|$, we obtain the trace inequality

$$|\operatorname{tr} A^2| \leq \operatorname{tr}(|A|\,|A^*|). \tag{3.7}$$

Note: $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ shows $\operatorname{tr}(AA^*) \not\leq \operatorname{tr}(|A|\,|A^*|)$, while $|\operatorname{tr} A^2| \leq \operatorname{tr}(AA^*)$ is valid.

d) Replacing P, Q, and R with yy^*, yx^*, and xx^*, respectively, we get

$$|\langle Ax, y\rangle|^2 \leq \langle |A|x, x\rangle\langle |A^*|y, y\rangle. \tag{3.8}$$

Remark 7. It is interesting to notice that the converse of (3.6) is invalid; that is, the inequality $|\Sigma(A)| \leq \Sigma(|A|)$ in (3.6) does not imply the normality of A. Take the non-normal matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Then

$$\Sigma(|A^*|) = \sqrt{2} < \Sigma(A^*) = \Sigma(A) = 2 < 2\sqrt{2} = \Sigma(|A|).$$

This example also indicates that $|\Sigma(X)|$ can be greater than $\Sigma(|X|)$ for some matrices X. In addition, note that (3.5) is a special case of (3.8) (see, e.g., [7]).

We now turn our attention to matrix normality. Let $A \in \mathbb{M}_n$. It is known that

$$|\langle Ax, x\rangle| \leq \langle |A|x, x\rangle \text{ for all } x \in \mathbb{C}^n \quad \Leftrightarrow \quad A \text{ is normal}$$

or

$$|\langle Ax, x\rangle| \leq \langle |A^*|x, x\rangle \text{ for all } x \in \mathbb{C}^n \quad \Leftrightarrow \quad A \text{ is normal.}$$

Here we present a proof different than the ones in [7] and [4] by using (3.3).

Proof. If A is normal, then $|A| = |A^*|$. An application of (3.4) with $P = Q = R = xx^*$ immediately yields that for all $x \in \mathbb{C}^n$,

$$|\langle Ax, x\rangle| \leq \langle |A|x, x\rangle.$$

The necessity is done in the same way as in [7] or [4] by an induction on n, by assuming A to be upper triangular, and by taking $x = (1, 0, \ldots, 0)^T$. ◻

Besides, one may prove for any square matrix A

$$|\langle A_a x, x\rangle| \leq \langle A_{|a|}x, x\rangle,$$

where

$$A_a = \frac{1}{2}(A + A^*) \quad \text{and} \quad A_{|a|} = \frac{1}{2}(|A| + |A^*|).$$

Note that

$$|\langle A_a x, x\rangle| \leq \langle |A|x, x\rangle$$

holds for normal matrices A but not for general A by taking $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

4. The Ky Fan singular value majorization. There exist a great number of fascinating inequalities on matrix singular values. We study some well-known and fundamentally important inequalities due to Ky Fan. To state Fan's result, let A be any $m\times n$ matrix and denote the singular values of A by $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A)$ (which are the same as the eigenvalues of $|A|$). The Fan singular value majorization theorem states that

$$\sum_{i=1}^{k} \sigma_i(A+B) \leq \sum_{i=1}^{k} \sigma_i(A) + \sum_{i=1}^{k} \sigma_i(B), \tag{4.1}$$

for $k = 1, 2, \ldots, q = \min\{m, n\}$; or written in majorization form [8, p. 243],

$$\sigma(A+B) \prec_w \sigma(A) + \sigma(B).$$

With this, the matrix norm defined for each k by $N_k(A) = \sum_{i=1}^{k} \sigma_i(A)$, known as the Ky Fan k-norm, is subadditive; i.e., $N_k(A+B) \leq N_k(A) + N_k(B)$.

Fan's theorem follows at once from the representation (see, e.g., [6, p. 195])

$$\sum_{i=1}^{k} \sigma_i(A) = \max\{|\operatorname{tr}(X^*AY)| : \ X \in \mathbb{M}_{m,k}, Y \in \mathbb{M}_{n,k}, \ X^*X = I_k = Y^*Y\}. \tag{4.2}$$

Proof of (4.2). The term $\operatorname{tr}(X^*AY) = \operatorname{tr}(YX^*A) = \langle A, XY^*\rangle$ suggests that we couple the matrices A and XY^*. So we take the positive semidefinite matrices

$$\begin{pmatrix} |A^*| & A \\ A^* & |A| \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} XX^* & XY^* \\ YX^* & YY^* \end{pmatrix}.$$

It is immediate by (3.4) that

$$|\operatorname{tr}(X^*AY)|^2 \leq \operatorname{tr}(X^*|A^*|X) \operatorname{tr}(Y^*|A|Y).$$

Note that if $P \in \mathbb{M}_n$ is positive semidefinite having eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, and $U \in \mathbb{M}_{n,k}$ is such that $U^*U = I_k$, then $\operatorname{tr}(U^*PU) \leq \lambda_1 + \cdots + \lambda_k$. Therefore, noticing that $|A^*|$ and $|A|$ have the same eigenvalues, we have

$$\operatorname{tr}(X^*|A^*|X) \ \text{ and } \ \operatorname{tr}(Y^*|A|Y) \leq \sigma_1(A) + \cdots + \sigma_k(A).$$

Thus

$$|\operatorname{tr}(X^*AY)| \leq \sigma_1(A) + \cdots + \sigma_k(A).$$

For the other direction, let $A = VDW$ be a singular value decomposition of A with the ith largest singular value of A in the (i, i)-position of D for each i, where $V \in \mathbb{M}_m$ and $W \in \mathbb{M}_n$ are unitary matrices. Then the extremal value is attained by taking X and Y to be the first k columns of V and W^*, respectively. $\square$

The representation (4.2) is traditionally and commonly proved using stochastic matrix theory [8, p. 511] or by eigenvalue and singular value inequalities for matrix products [6, p. 196]. The following two special cases are of interest in their own right.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 11, pp. 66-77, April 2004

ELA

For the case $k = 1$ (see, e.g, [13, p. 91]), the largest singular value or the spectral norm of A is given by

$$\sigma_1(A) = \max_{\|x\|=\|y\|=1} |\langle Ax, y\rangle|$$

and this can be shown directly via the pair of positive semidefinite matrices

$$\begin{pmatrix} \sigma_1(A)I & A \\ A^* & \sigma_1(A)I \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} yy^* & yx^* \\ xy^* & xx^* \end{pmatrix}.$$

For the case $k = m = n$ (see, e.g., [5, p. 430]), the Ky Fan n-norm of A or the trace of $|A|$ has the representation

$$\sigma_1(A) + \sigma_2(A) + \cdots + \sigma_n(A) = \max_{\text{unitary } U \in \mathbb{M}_n} |\operatorname{tr}(UA)| \tag{4.3}$$

and this is proved by simply taking the pair of positive semidefinite matrices

$$\begin{pmatrix} |A^*| & A \\ A^* & |A| \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} I & U^* \\ U & I \end{pmatrix}.$$

REFERENCES

[1] T. Ando. Operator-Theoretic Methods for Matrix Inequalities. Manuscript, March 1998.
[2] T. Ando. Norms and Cones in Tensor Products of Matrices. Manuscript, March 2001.
[3] R. Bellman. Some inequalities for the square Root of a Positive Definite Matrix. *Linear Algebra Appl.,* 1:321–324, 1968.
[4] L. Elsner and Kh.D. Ikramov. Normal matrices: an update. *Linear Algebra Appl.,* 285:291–303, 1998.
[5] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, New York, 1985.
[6] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, New York, 1991.
[7] C. R. Johnson and F.-Z. Zhang. An operator inequality and matrix normality. *Linear Algebra Appl.,* 240:105–110, 1996.
[8] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York, 1979.
[9] L. Mirsky. The spread of a matrix. *Mathematika,* 3:127–130, 1956.
[10] L. Mirsky. Inequalities for normal and Hermitian matrices. *Duke Math. J.,* 24:591–599, 1957.
[11] R. C. Thompson. Root spreads for polynomials and Hermitian matrix pencils. *Linear Algebra Appl.,* 220:419–433, 1995.
[12] G. Visick. Private communication, 2003.
[13] F.-Z. Zhang. *Matrix Theory: Basic Results and Techniques.* Springer, New York, 1999.
[14] F.-Z. Zhang. Equivalence of the Wielandt inequality and the Kantorovich inequality. *Linear and Multilinear Algebra,* 48:275–279, 2001.