

NOTES ON AN ANDERSON-TAYLOR TYPE INEQUALITY*

MINGHUA LIN[†]

Abstract. As a complement to Olkin's extension of Anderson-Taylor's trace inequality, the following inequality is proved:

$$\sum_{j=1}^n \left(\sum_{i=1}^j A_i \right) A_j^{-1} \left(\sum_{i=1}^j A_i \right) \succeq \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i,$$

where the inequality is in the sense of Loewner partial order and $A_i, i = 1, \dots, n$, are positive definite matrices. Some related results for M-matrices are also discussed.

Key words. Loewner order, Trace inequality, Positive definite matrix, M-matrix.

AMS subject classifications. 15A60, 15A18, 15A42, 15A45.

1. Introduction. The notation in this article is standard. Capital letters are used to denote $k \times k$ matrices over real or complex fields. For two Hermitian matrices A and B , $A \succ B$ ($A \succeq B$) means $A - B$ is positive (semi)definite. Thus, $A \succ 0$ ($A \succeq 0$) naturally means A is positive (semi)definite. Of course, we do not distinguish $B \prec A$ ($B \preceq A$) from $A \succ B$ ($A \succeq B$). Comparison of Hermitian matrices in this way is the so called Loewner partial order. If $A \succeq 0$, then it has a unique square root $A^{1/2} \succeq 0$. The trace of A is denoted by $\text{tr } A$.

Let $\{x_j\}$ be a sequence of real vectors in $\mathbb{R}^p$ such that for some $n \geq p$, $\sum_{i=1}^n x_i x_i^T$ is nonsingular. Here x^T denotes the transpose of the vector x . Motivated by applications in probability theory, Anderson and Taylor [1, Proposition 1] proved the following trace inequality:

PROPOSITION 1.1. For $m > q \geq n$,

$$\sum_{j=q+1}^m x_j^T \left(\sum_{i=1}^j x_i x_i^T \right)^{-2} x_j \leq \text{tr} \left(\sum_{i=1}^q x_i x_i^T \right)^{-1}. \quad (1.1)$$

*Received by the editors on November 23, 2012. Accepted for publication on January 21, 2013.
 Handling Editor: Bit-Shun Tam.

[†]Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada (mlin87@ymail.com).

Zhan [12, Theorem 2] obtained the following generalization of (1.1):

PROPOSITION 1.2. *Let $A_1 \succ 0$, and $A_i \succeq 0$ for $i = 2, \dots, n$. Then,*

$$2 \operatorname{tr} A_1^{-1} > \operatorname{tr} \sum_{j=1}^n A_j \left(\sum_{i=1}^j A_i \right)^{-2}. \quad (1.2)$$

Indeed, a stronger version than (1.2) had been observed earlier by Olkin in [10], who obtained:

PROPOSITION 1.3. *Let $A_1 \succ 0$, and $A_i \succeq 0$ for $i = 2, \dots, n$. Then,*

$$2A_1^{-1} \succ \sum_{j=1}^n \left(\sum_{i=1}^j A_i \right)^{-1} A_j \left(\sum_{i=1}^j A_i \right)^{-1}. \quad (1.3)$$

In this article, we obtain an inequality complementary to (1.3). It is a matrix extension of the scalar inequality due to Heinig and Sinnamon [7, 8]. At the end, we discuss some possible extensions, where positive definite matrices are replaced by M-matrices and Loewner partial order is replaced by componentwise inequality.

2. Main results. Recall that the geometric mean of two positive definite matrices A and B , denoted by $A \sharp B$, is the positive definite solution of the Riccati equation $XB^{-1}X = A$ and it has the explicit expression

$$A \sharp B = B^{1/2} (B^{-1/2} A B^{-1/2})^{1/2} B^{1/2}.$$

From here, we find that $A \sharp B = B \sharp A$ and the monotonicity property: $A \sharp B \succeq A \sharp C$ whenever $B \succeq C \succ 0$ and $A \succ 0$. One of the motivations for geometric mean is of course the following arithmetic mean-geometric mean inequality:

$$\frac{A+B}{2} \succeq A \sharp B.$$

A remarkable property of the geometric mean is a maximal characterization by Pusz-Woronovicz [11]:

THEOREM 2.1. *Let $A, B \succ 0$. Then,*

$$A \sharp B = \max \left\{ X \mid \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \succeq 0, X = X^* \right\}. \quad (2.1)$$

The “maximum” here is in the sense of Loewner partial order. In some literature, this is also called Ando’s variational formula for the geometric mean; see e.g., [3, p. 93].

Applying this maximal characterization to the summation of positive semidefinite matrices $\begin{bmatrix} A_i & A_i \sharp B_i \\ A_i \sharp B_i & B_i \end{bmatrix}$, $i = 1, \dots, n$, we get the following matrix version of Cauchy-Schwarz inequality

$$\left(\sum_{i=1}^n A_i \right) \sharp \left(\sum_{i=1}^n B_i \right) \succeq \sum_{i=1}^n A_i \sharp B_i. \quad (2.2)$$

For more properties of the matrix geometric mean, we refer to [2, p. 101].

We need two lemmas for proving our main result.

LEMMA 2.2. *Let $A \succ 0$ and any Hermitian B . Then,*

$$A \sharp (BA^{-1}B) \succeq B. \quad (2.3)$$

Proof. We may assume B is nonsingular, a general case follows from a continuity argument. Indeed, with Theorem 2.1, the notion of geometric mean can be extended to cover the case of positive semidefinite matrices. Using the technique of Schur complements (e.g., [14, p. 92]), it is easy to see

$$\begin{bmatrix} A & B \\ B & BA^{-1}B \end{bmatrix} \succeq 0.$$

Now by (2.1), the desired inequality follows. $\square$

REMARK 2.3. Inequality (2.3) is of course a refinement of the following inequality: for $A \succ 0$ and any Hermitian B ,

$$A + BA^{-1}B \succeq 2B. \quad (2.4)$$

The inequality (2.4) was first proved in [13, Lemma 3.2] and it has been used to prove the convergence of some iterative methods for certain matrix equations in [13] and [6].

The next lemma can be found in [5, Theorem 4.2], for completeness, we include a simple proof.

LEMMA 2.4. *Let $A, B \succ 0$. Then $A \sharp B \succeq B$ if and only if $A \succeq B$.*

Proof. The “if” part is by the monotonicity property, $A \sharp B \succeq B \sharp B = B$. To show the converse, we use the explicit expression for geometric mean. $A \sharp B \succeq B$ is the same as $B^{1/2}(B^{-1/2}AB^{-1/2})^{1/2}B^{1/2} \succeq B$, or equivalently, $(B^{-1/2}AB^{-1/2})^{1/2} \succeq I$, i.e., $B^{-1/2}AB^{-1/2} \succeq I$, implying $A \succeq B$. $\square$

Now we are at the position to state our main result.

THEOREM 2.5. *Let $A_i \succ 0$ for $i = 1, \dots, n$. Then,*

$$\sum_{j=1}^n \left(\sum_{i=1}^j A_i \right) A_j^{-1} \left(\sum_{i=1}^j A_i \right) \succeq \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i. \quad (2.5)$$

Moreover, the constant $1/2$ is best possible.

Proof. Interchanging the order of summation gives

$$\begin{aligned} \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i &= \sum_{j=1}^n \sum_{k=j}^n \left(\sum_{i=1}^j A_i \right) \\ &= \sum_{j=1}^n (n-j+1) \sum_{i=1}^j A_i \\ &= \sum_{i=1}^n A_i \sum_{j=i}^n (n-j+1) \\ &= \sum_{i=1}^n \binom{n-i+2}{2} A_i \succeq \frac{1}{2} \sum_{i=1}^n (n-i+1)^2 A_i, \end{aligned}$$

i.e.,

$$2 \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i \succeq \sum_{i=1}^n (n-i+1)^2 A_i. \quad (2.6)$$

On the other hand,

$$\begin{aligned} \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i &= \sum_{j=1}^n (n-j+1) \left(\sum_{i=1}^j A_i \right) \\ &\preceq \sum_{j=1}^n ((n-j+1)^2 A_j) \# \left\{ \left(\sum_{i=1}^j A_i \right) A_j^{-1} \left(\sum_{i=1}^j A_i \right) \right\} \\ &\preceq \left\{ \sum_{j=1}^n (n-j+1)^2 A_j \right\} \# \left\{ \sum_{j=1}^n \left(\sum_{i=1}^j A_i \right) A_j^{-1} \left(\sum_{i=1}^j A_i \right) \right\} \\ &\preceq \left\{ 2 \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i \right\} \# \left\{ \sum_{j=1}^n \left(\sum_{i=1}^j A_i \right) A_j^{-1} \left(\sum_{i=1}^j A_i \right) \right\} \\ &= \left\{ \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i \right\} \# \left\{ 2 \sum_{j=1}^n \left(\sum_{i=1}^j A_i \right) A_j^{-1} \left(\sum_{i=1}^j A_i \right) \right\}, \end{aligned}$$

in which the first inequality is by Lemma 2.2, the second one is by (2.2) and the third one is by (2.6). Now (2.5) follows from using Lemma 2.4, this completes the proof of (2.5). The proof that $1/2$ is best possible is given in the appendix. $\square$

The following corollary is readily seen:

COROLLARY 2.6. *Let $A_i \succ 0$ for $i = 1, \dots, n$. Then,*

$$\operatorname{tr} \sum_{j=1}^n A_j^{-1} \left(\sum_{i=1}^j A_i \right)^2 \geq \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j \operatorname{tr} A_i. \quad (2.7)$$

We remark that the scalar version of (2.5) has appeared in [7, 8]. In the scalar case, it had been an open problem whether $1/2$ on the right hand side of (2.5) was best possible. This was first confirmed by Chao [4].

3. M-matrix analogue. A real nonsingular matrix A is an M-matrix if all its off-diagonal entries are nonpositive and $A^{-1} \geq 0$, i.e., A^{-1} is componentwise nonnegative (e.g., [9, p. 113]). Due to the resemblance between positive definite matrices and M-matrices, in this section, we explore some analogous results for M-matrices.

LEMMA 3.1. *Let $A, A+B$ be two M-matrices, with $B \geq 0$. Then,*

$$A^{-1} - (A+B)^{-1} \geq (A+B)^{-1} B (A+B)^{-1}. \quad (3.1)$$

Proof. The reverse property [9, p. 117] tells us that $A^{-1} \geq (A+B)^{-1} \geq 0$. Then

$$\begin{aligned} A^{-1} - (A+B)^{-1} &= A^{-1} B (A+B)^{-1} \\ &\geq (A+B)^{-1} B (A+B)^{-1}. \quad \square \end{aligned}$$

PROPOSITION 3.2. *Let A_1 and $A_1 + \sum_{i=2}^n A_i$ be M-matrices, with $A_i \geq 0$ for $i = 2, \dots, n$. Then,*

$$2A_1^{-1} \geq \sum_{j=1}^n \left(\sum_{i=1}^j A_i \right)^{-1} A_j \left(\sum_{i=1}^j A_i \right)^{-1}. \quad (3.2)$$

Proof. It is clear that $A_1 + \sum_{i=2}^j A_i$ is an M-matrix for $j = 2, \dots, n$ (e.g., [9, p. 117]). Note that (3.2) is the same as

$$A_1^{-1} \geq \sum_{j=2}^n \left(\sum_{i=1}^j A_i \right)^{-1} A_j \left(\sum_{i=1}^j A_i \right)^{-1}. \quad (3.3)$$

(3.3) will follow from (3.4) below by summing up for j from 2 to n .

$$\left(\sum_{i=1}^{j-1} A_i\right)^{-1} - \left(\sum_{i=1}^j A_i\right)^{-1} \geq \left(\sum_{i=1}^j A_i\right)^{-1} A_j \left(\sum_{i=1}^j A_i\right)^{-1}, \quad (3.4)$$

for $2 \leq j \leq n$. Let $A = \sum_{i=1}^{j-1} A_i$, $B = A_j$, then by Lemma 3.1, (3.4) and hence (3.3) follows. $\square$

REMARK 3.3. Note that “ $\geq$ ” in (3.2) cannot be replaced by “ $>$ ”. Consider A_i ($i = 1, \dots, n$) to be all diagonal M-matrices, then the off-diagonal entries on both sides of (3.2) are all equal.

Taking the trace in (3.2), we immediately have:

COROLLARY 3.4. Let A_1 and $A_1 + \sum_{i=2}^n A_i$ be M-matrices with $A_i \geq 0$ for $i = 2, \dots, n$. Then

$$2 \operatorname{tr} A_1^{-1} \geq \operatorname{tr} \sum_{j=1}^n A_j \left(\sum_{i=1}^j A_i\right)^{-2}. \quad (3.5)$$

It is natural to ask whether the following analogue of (2.5) holds or not. Under the same condition of Proposition 3.2, is it true

$$\sum_{j=1}^n \left(\sum_{i=1}^j A_i\right) A_j^{-1} \left(\sum_{i=1}^j A_i\right) \geq \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j A_i? \quad (3.6)$$

However, this is refuted by the following example.

EXAMPLE 3.5. Let $n = 2$, $A_1 = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. The conditions of Proposition 3.2 are satisfied. Simple calculation shows that

$$\sum_{j=1}^2 \left(\sum_{i=1}^j A_i\right) A_j^{-1} \left(\sum_{i=1}^j A_i\right) = A_1 + (A_1 + A_2) A_2^{-1} (A_1 + A_2) = \begin{bmatrix} -4 & 11 \\ 5 & -6 \end{bmatrix}$$

and

$$\frac{1}{2} \sum_{k=1}^2 \sum_{j=1}^k \sum_{i=1}^j A_i = \frac{1}{2} (3A_1 + A_2) = \begin{bmatrix} 3.5 & -4 \\ -1 & 3 \end{bmatrix}.$$

4. Appendix.

PROPOSITION 4.1. *The best constant c in the inequality,*

$$\sum_{j=1}^n \left(\sum_{i=1}^j x_i \right)^2 x_j^{-1} \geq c \sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j x_i, \quad (4.1)$$

where $x_i > 0$ for $i = 1, \dots, n$, is $1/2$.

The author is grateful to Gord Sinnamon for sending him the following simple proof, which is different from that of Chao [4].

Proof. To show that $c = 1/2$ is best possible, define $x_1 = n^{-2}$ and $x_i = (n - i + 1)^{-2} - (n - i + 2)^{-2}$ for $i = 2, \dots, n$. Observe that

$$\sum_{i=1}^j x_i = (n - j + 1)^{-2}, \quad j = 1, \dots, n.$$

We have

$$\sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j x_i = \sum_{j=1}^n (n - j + 1) \left(\sum_{i=1}^j x_i \right) = \sum_{j=1}^n (n - j + 1)^{-1} = \sum_{k=1}^n \frac{1}{k}.$$

Also,

$$\begin{aligned} 2 \sum_{j=1}^n \left(\sum_{i=1}^j x_i \right)^2 x_j^{-1} &= 2n^{-2} + 2 \sum_{j=2}^n (n - j + 1)^{-4} ((n - j + 1)^{-2} - (n - j + 2)^{-2})^{-1} \\ &= 2n^{-2} + 2 \sum_{k=1}^{n-1} k^{-4} (k^{-2} - (k + 1)^{-2})^{-1} \\ &= 2n^{-2} + \sum_{k=1}^{n-1} \frac{1}{k} \frac{2(k + 1)^2}{(2k + 1)k} \\ &= \left(2n^{-2} - n^{-1} + \sum_{k=1}^{n-1} \frac{3k + 2}{2k^3 + k^2} \right) + \sum_{k=1}^n \frac{1}{k}. \end{aligned}$$

As $n \rightarrow \infty$ the first term converges and the second diverges. It follows that

$$\lim_{n \rightarrow \infty} \frac{2 \sum_{j=1}^n \left(\sum_{i=1}^j x_i \right)^2 x_j^{-1}}{\sum_{k=1}^n \sum_{j=1}^k \sum_{i=1}^j x_i} = \lim_{n \rightarrow \infty} \frac{\left(2n^{-2} - n^{-1} + \sum_{k=1}^{n-1} \frac{3k+2}{2k^3+k^2} \right) + \sum_{k=1}^n \frac{1}{k}}{\sum_{k=1}^n \frac{1}{k}} = 1.$$

We conclude that the inequality fails for any constant $c > 1/2$. $\square$

Acknowledgment. The author thanks Gord Sinnamon for bringing [4] to his attention. Helpful comments from the referee are also acknowledged.

REFERENCES

- [1] T.W. Anderson and J.B. Taylor. An inequality for a sum of quadratic forms with applications to probability theory. *Linear Algebra Appl.*, 30:93–99, 1980.
- [2] R. Bhatia. *Positive Definite Matrices*. Princeton University Press, Princeton, 2007.
- [3] E. Carlen. *Trace Inequalities and Quantum Entropy: An Introductory Course*. In “Entropy and the Quantum: Arizona School of Analysis with Applications”, University of Arizona, March 16–20, 2009. *Contemporary Mathematics*, American Mathematical Society, Providence, RI, 2010.
- [4] L. Chao. Solution to Problem 2557. *Cruz Math.*, 35:330–332, 2009.
- [5] M. Fiedler and V. Pták. A new positive definite geometric mean of two positive definite matrices. *Linear Algebra Appl.*, 251:1–20, 1997.
- [6] C.-H. Guo and P. Lancaster. Iterative solution of two matrix equations. *Math. Comp.*, 68:1589–1603, 1999.
- [7] H. Heinig and G. Sinnamon. Problem 2557. *Cruz Math.*, 26:304, 2000.
- [8] H. Heinig and G. Sinnamon. Solution to Problem 2557. *Cruz Math.*, 27:465, 2001.
- [9] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [10] I. Olkin. An inequality for a sum of forms. *Linear Algebra Appl.*, 52:529–532, 1983.
- [11] W. Pusz and S.L. Woronowicz. Functional calculus for sesquilinear forms and the purification map. *Rep. Math. Phys.*, 8:159–170, 1975.
- [12] X. Zhan. On some matrix inequalities. *Linear Algebra Appl.*, 376:299–303, 2004.
- [13] X. Zhan. Computing the extremal positive definite solutions of a matrix equation. *SIAM J. Sci. Comput.*, 17:1167–1174, 1996.
- [14] F. Zhang. *The Schur Complement and its Applications*. Numerical Methods and Algorithms, Springer-Verlag, New York, 2005.