Mathematical
Inequalities
& Applications
Volume 18, Number 3 (2015), 1079–1086 doi:10.7153/mia-18-83

DETERMINANTAL INEQUALITIES FOR BLOCK TRIANGULAR MATRICES

MINGHUA LIN

Dedicated to Stephen Drury,
whose beautiful works on matrix analysis
inspire the author

(*Communicated by S. Puntanen*)

Abstract. Let $T = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$ be an n-square matrix, where X, Z are r-square and $(n-r)$-square, respectively. Among other determinantal inequalities, it is proved that

$$\det(I_n + T^*T) \geqslant \det(I_r + X^*X) \cdot \det(I_{n-r} + Z^*Z)$$

with equality if and only if $Y = 0$.

1. Introduction

The well known Fischer inequality [4, p. 506] states that if $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}$ is positive semidefinite, then

$$\det A \leqslant \det A_{11} \cdot \det A_{22}. \tag{1}$$

As any positive semidefinite matrix A can be factorized as $A = T^*T$ with $T = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$ being comformally partitioned as A, inequality (1) can be written as

$$\det X^*X \cdot \det Z^*Z = \det T^*T \leqslant \det X^*X \cdot \det(Y^*Y + Z^*Z). \tag{2}$$

This paper presents some results that complement (2). We believe our results are of new pattern concerning determinantal inequalities. Let us fix some notation. The matrices considered here have entries from the field of complex numbers. $X', \overline{X}, X^*$ stand for transpose, (entrywise)conjugate, conjugate transpose of X, respectively. For two n-square Hermitian matrices X, Y, we write $X \geqslant Y$ to mean $X - Y$ is positive semidefinite (so $X \geqslant 0$ means X is positive semidefinite). The n-square identity matrix is denoted by I_n. The Frobenius norm of X is denoted by $\|X\|_F$. It is known that

Mathematics subject classification (2010): 15A45.
Keywords and phrases: Determinantal inequality, block triangular matrices.

$\|X\|_F = \sqrt{\operatorname{tr} X^*X}$, where tr denotes the trace. If $X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$ is an n-square matrix with X_{11} nonsingular, then the Schur complement of X_{11} in X is defined by $X/X_{11} = X_{22} - X_{21}X_{11}^{-1}X_{12}$. A well known property of the Schur complement is $\det X = \det X_{11} \cdot \det(X/X_{11})$. Finally, for an n-square matrix, we denote by $\lambda_j(X)$ and $\sigma_j(X)$, $j = 1,\dots,n$, the eigenvalues and singular values of X, respectively, such that $|\lambda_1(X)| \geqslant \cdots \geqslant |\lambda_n(X)|$ and $\sigma_1(X) \geqslant \cdots \geqslant \sigma_n(X)$.

2. Main results

We present the following result, showing that when more matrices are summed, the identity in (2) becomes an inequality.

THEOREM 1. *Let* $T_k = \begin{bmatrix} X_k & Y_k \\ 0 & Z_k \end{bmatrix}$, $k = 1,\dots,m$, *be* n*-square conformally partitioned matrices. Then*

$$\det\left(\sum_{k=1}^m T_k^*T_k\right) \geqslant \det\left(\sum_{k=1}^m X_k^*X_k\right) \cdot \det\left(\sum_{k=1}^m Z_k^*Z_k\right). \tag{3}$$

Proof. By a standard continuity argument, we may assume $X_k^*X_k$ is nonsingular for $k = 1,\dots,m$. As

$$\begin{bmatrix} X_k^*X_k & X_k^*Y_k \\ Y_k^*X_k & Y_k^*Y_k \end{bmatrix} = \begin{bmatrix} X_k & Y_k \end{bmatrix}^* \begin{bmatrix} X_k & Y_k \end{bmatrix} \geqslant 0,$$

summing for k from 1 to m gives

$$\begin{bmatrix} \sum_{k=1}^m X_k^*X_k & \sum_{k=1}^m X_k^*Y_k \\ \sum_{k=1}^m Y_k^*X_k & \sum_{k=1}^m Y_k^*Y_k \end{bmatrix} \geqslant 0.$$

Hence,

$$\sum_{k=1}^m Y_k^*Y_k - \left(\sum_{k=1}^m Y_k^*X_k\right)\left(\sum_{k=1}^m X_k^*X_k\right)^{-1}\left(\sum_{k=1}^m X_k^*Y_k\right) \geqslant 0.$$

On the other hand, $T_k^*T_k = \begin{bmatrix} X_k^*X_k & X_k^*Y_k \\ Y_k^*X_k & Y_k^*Y_k + Z_k^*Z_k \end{bmatrix}$. Thus

$$\begin{aligned}\left(\sum_{k=1}^m T_k^*T_k\right) \Big/ \left(\sum_{k=1}^m X_k^*X_k\right) &= \sum_{k=1}^m (Y_k^*Y_k + Z_k^*Z_k) - \left(\sum_{k=1}^m Y_k^*X_k\right)\left(\sum_{k=1}^m X_k^*X_k\right)^{-1}\left(\sum_{k=1}^m X_k^*Y_k\right) \\ &\geqslant \sum_{k=1}^m Z_k^*Z_k \geqslant 0.\end{aligned}$$

Applying the determinant on both sides gives the desired inequality. $\square$

REMARK 1. By a simple induction, Theorem 1 can be extended to the $p \times p$ ($p > 2$) block upper triangular case.

A full characterization of the equality case in (3) is nasty, so we do not include it here. We extract a special case of Theorem 1 with $m = 2$ for later use. Moreover, the equality case is concise.

COROLLARY 1. *Let* $T = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$ *be an* n*-square matrix, where* X, Z *are* r*-square and* $(n-r)$*-square, respectively. Then*

$$\det(I_n + T^*T) \geqslant \det(I_r + X^*X) \cdot \det(I_{n-r} + Z^*Z). \tag{4}$$

Equality holds in (4) if and only if $Y = 0$.

Proof. It suffices to show the equality case. If $Y = 0$, clearly (4) becomes an equality. Conversely, if the equality in (4) holds, then

$$\det\left(I_{n-r} + Z^*Z + Y^*Y - Y^*X(I_r + X^*X)^{-1}X^*Y\right) = \det(I_{n-r} + Z^*Z). \tag{5}$$

As $Y^*Y - Y^*X(I_r + X^*X)^{-1}X^*Y \geqslant 0$, the equality (5) holds only if $Y^*Y - Y^*X(I_r + X^*X)^{-1}X^*Y = 0$, i.e,

$$Y^*\left(I_r - X(I_r + X^*X)^{-1}X^*\right)Y = 0.$$

Now for any $j = 1, \ldots, r$,

$$\begin{aligned} \lambda_j\left(I_r - X(I_r + X^*X)^{-1}X^*\right) &= 1 - \frac{\lambda_{r-j+1}(X^*X)}{1 + \lambda_{r-j+1}(X^*X)} \\ &= \frac{1}{1 + \lambda_{r-j+1}(X^*X)} > 0. \end{aligned}$$

So $I_r - X(I_r + X^*X)^{-1}X^*$ is positive definite, forcing $Y = 0$. □

COROLLARY 2. *Let* $T_k = \begin{bmatrix} X_k & Y_k \\ 0 & Z_k \end{bmatrix}$, $k = 1, \ldots, m$, *be* n*-square conformally partitioned matrices. If* X_k, Z_k *are all normal matrices, then*

$$\det\left(\sum_{k=1}^m T_k^*T_k\right) \geqslant \left|\det\left(\sum_{k=1}^m \overline{X}_k X_k\right)\right| \cdot \left|\det\left(\sum_{k=1}^m \overline{Z}_k Z_k\right)\right|. \tag{6}$$

Proof. In view of Theorem 1, it suffices to show

$$\det\left(\sum_{k=1}^m X_k^*X_k\right) \geqslant \left|\det\left(\sum_{k=1}^m \overline{X}_k X_k\right)\right|.$$

As

$$\begin{bmatrix} \sum_{k=1}^m \overline{X}_k X_k' & \sum_{k=1}^m \overline{X}_k X_k \\ \sum_{k=1}^m X_k^* X_k' & \sum_{k=1}^m X_k^* X_k \end{bmatrix} = \sum_{k=1}^m \begin{bmatrix} X_k' & X_k \end{bmatrix}^* \begin{bmatrix} X_k' & X_k \end{bmatrix} \geqslant 0,$$

this yields (see [4, p. 445, P25])

$$\begin{bmatrix} \det\left(\sum_{k=1}^m \overline{X}_k X_k'\right) & \det\left(\sum_{k=1}^m \overline{X}_k X_k\right) \\ \det\left(\sum_{k=1}^m X_k^* X_k'\right) & \det\left(\sum_{k=1}^m X_k^* X_k\right) \end{bmatrix} \geqslant 0,$$

and so

$$\det\left(\sum_{k=1}^m \overline{X}_k X_k'\right) \cdot \det\left(\sum_{k=1}^m X_k^* X_k\right) \geqslant \left|\det\left(\sum_{k=1}^m \overline{X}_k X_k\right)\right|^2.$$

But X_k, $k = 1, \dots, m$, are normal, and so

$$\begin{aligned} \det\left(\sum_{k=1}^m \overline{X}_k X_k'\right) &= \det\left(\sum_{k=1}^m \overline{X}_k X_k'\right)' \\ &= \det\left(\sum_{k=1}^m X_k X_k^*\right) = \det\left(\sum_{k=1}^m X_k^* X_k\right), \end{aligned}$$

as desired. □

The author does not know if there is a simple characterization for the equality case in (6). The following example shows that (6) may fail without the normality assumption.

EXAMPLE 1. Taking

$$T_1 = \left[\begin{array}{cc|cc} -9 & 10 & 5 & 12 \\ -7 & 10 & -11 & -10 \\ \hline 0 & 0 & -2 & 3 \\ 0 & 0 & 2 & 26 \end{array}\right], \qquad T_2 = \left[\begin{array}{cc|cc} 13 & -16 & 3 & 3 \\ -7 & 9 & 3 & 11 \\ \hline 0 & 0 & 3 & -16 \\ 0 & 0 & -7 & -13 \end{array}\right],$$

a simple calculation using Matlab gives

$$\begin{aligned} \det(T_1^* T_1 + T_2^* T_2) &= 1.25 \times 10^8 \\ &< \left|\det\left(X_1^2 + X_2^2\right)\right| \cdot \left|\det\left(Z_1^2 + Z_2^2\right)\right| = 9.93 \times 10^8. \end{aligned}$$

Our next result says that, to some extent, Corollary 2 can be improved.

THEOREM 2. *Let* $T = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$ *be an* n*-square matrix, where* X, Z *are* r*-square and* $(n-r)$*-square, respectively. Then*

$$\det(I_n + T^*T) \geqslant \det\left(I_r + \overline{X}X\right) \cdot \det\left(I_{n-r} + \overline{Z}Z\right). \tag{7}$$

Equality holds in (7) if and only if $Y = 0$ *and* X, Z *are symmetric.*

REMARK 2. Compared with (6), we don't use absolute value on the right hand side of (7). This is because $\det\left(I_r+\overline{X}X\right)\geqslant 0$, an observation by Djoković [1, 2]. However, it is possible that $\det\left(\sum_{k=1}^m \overline{X}_k X_k\right)<0$ in (6). For example, taking

$$X_1 = X_2' = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

a simple calculation gives

$$\det(\overline{X}_1 X_1 + \overline{X}_2 X_2) = \det\begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix} = -12 < 0.$$

We need a lemma, which plays a key role in establishing the equality case in Theorem 2.

LEMMA 1. *Let X be an n-square matrix. Then*

$$\det(I_n + X^*X) \geqslant \det\left(I_n+\overline{X}X\right). \tag{8}$$

Equality holds in (8) if and only if X is symmetric.

Proof. From the proof of Corollary 2, we have

$$\det(I_n+XX^*)\cdot\det(I_n+X^*X)\geqslant \det\left(I_n+\overline{X}X\right)^2.$$

Note that

$$\det(I_n+\overline{X}X') = \det(I_n+\overline{X}X')' = \det(I_n+XX^*) = \det(I_n+X^*X).$$

This proves (8).

If X is symmetric, then $\overline{X}=X^*$ and so $\det\left(I_n+\overline{X}X\right)=\det(I_n+X^*X)$.

We show the other way around. It is clear that

$$\det\left(I_n+\overline{X}X\right) = \prod_{j=1}^n \left(1+\lambda_j(\overline{X}X)\right) \leqslant \prod_{j=1}^n \left(1+|\lambda_j(\overline{X}X)|\right)$$

with the second inequality becoming an equality only if $\lambda_j(\overline{X}X)\geqslant 0$ for all j.

By Weyl's inequality [6, p. 317], for $k=1,\ldots,n$,

$$\begin{aligned}\prod_{j=1}^k |\lambda_j(\overline{X}X)| &\leqslant \prod_{j=1}^k \sigma_j(\overline{X}X)\\ &\leqslant \prod_{j=1}^k \sigma_j(\overline{X})\sigma_j(X) = \prod_{j=1}^k \sigma_j^2(X) = \prod_{j=1}^k \sigma_j(X^*X),\end{aligned}$$

where equality holds when $k=n$. The strict convexity of the function $f(t)=\log(1+e^t)$ ([6, p. 156]) implies that

$$\prod_{j=1}^n \left(1+|\lambda_j(\overline{X}X)|\right) \leqslant \prod_{j=1}^n \left(1+\sigma_j(X^*X)\right) = \det(I_n+X^*X)$$

with the first inequality becoming an equality only if $\overline{X}X$ is normal.

Thus, if $\det\left(I_n+\overline{X}X\right)=\det(I_n+X^*X)$, then $\overline{X}X \geqslant 0$ and $\lambda_j(\overline{X}X)=\sigma_j(X^*X)$ for all j. In particular, $\operatorname{tr}\overline{X}X=\operatorname{tr}X^*X$. We shall show the latter implies that X is symmetric. Compute

$$\begin{aligned}\|X-X'\|_F^2 &= \operatorname{tr}(X-X')^*(X-X')\\ &= \operatorname{tr}X^*X-\operatorname{tr}\overline{X}X-\operatorname{tr}(\overline{X}X)^*+\operatorname{tr}\overline{X}X'\\ &= 2\left(\operatorname{tr}X^*X-\operatorname{tr}\overline{X}X\right)=0,\end{aligned}$$

and so $X=X'$, as required. □

REMARK 3. As pointed out by a referee, Lemma 1 possesses a much shorter proof using exterior products and the Cauchy-Schwarz inequality.

Proof of Theorem 2. The inequality (7) follows from (4) and (8).
If $Y=0$ and X,Z are symmetric, then

$$\begin{aligned}\det(I_n+T^*T) &= \det(I_r+X^*X)\cdot\det(I_{n-r}+Z^*Z)\\ &= \det\left(I_r+\overline{X}X\right)\cdot\det\left(I_{n-r}+\overline{Z}Z\right).\end{aligned}$$

Conversely, if the equality holds in (7), then actually we have

$$\begin{aligned}\det(I_n+T^*T) &= \det(I_r+X^*X)\cdot\det(I_{n-r}+Z^*Z)\\ &= \det\left(I_r+\overline{X}X\right)\cdot\det\left(I_{n-r}+\overline{Z}Z\right).\end{aligned}$$

In view of Corollary 1, the first equality gives $Y=0$. By Lemma 1, the second equality implies that X,Z are symmetric. □

An immediate consequence of Theorem 2 is the following, which is due to Drury [3, Lemma 4].

COROLLARY 3. (Drury's inequality) *Let $T=[t_{ij}]$ be an n-square upper triangular matrix. Then*

$$\det(I_n+T^*T)\geqslant\prod_{j=1}^{n}(1+|t_{jj}|^2).$$

Equality holds if and only if T is diagonal.

3. More results

The absolute value of a matrix X is defined to be the matrix $|X|=(X^*X)^{1/2}$, the unique positive semidefinite square root of X^*X. The inequality (3) can be rewritten as

$$\det\left(\sum_{k=1}^{m}|T_k|^2\right)\geqslant\det\left(\sum_{k=1}^{m}|X_k|^2\right)\cdot\det\left(\sum_{k=1}^{m}|Z_k|^2\right). \tag{9}$$

The following result is an extension of Corollary 1.

THEOREM 3. *Let* $T = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$ *be an* n*-square matrix, where* X, Z *are* r*-square and* $(n-r)$*-square, respectively. Then for any* $p > 0$

$$\det(I_n + |T|^p) \geqslant \det(I_r + |X|^p) \cdot \det(I_{n-r} + |Z|^p). \tag{10}$$

Equality holds in (10) if and only if $Y = 0$.

Proof. The proof is similar to the one given in [3, Lemma 4]. Let $X = U_1|X|$, $Z = U_2|Z|$ be the polar decomposition of X, Z, respectively. Taking $U = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix}$ (so U is unitary) gives $U^*T = \begin{bmatrix} |X| & U_1^*Y \\ 0 & |Z| \end{bmatrix}$. We have by Weyl's inequality [6, p. 317]

$$\prod_{j=1}^{k} |\lambda_j(U^*T)| \leqslant \prod_{j=1}^{k} \sigma_j(U^*T) = \prod_{j=1}^{k} \sigma_j(T), \qquad k = 1, \ldots, n,$$

i.e.,

$$\prod_{j=1}^{k} \lambda_j\left(\begin{bmatrix} |X| & 0 \\ 0 & |Z| \end{bmatrix}\right) \leqslant \prod_{j=1}^{k} \sigma_j(T), \qquad k = 1, \ldots, n,$$

where equality holds when $k = n$.

By the convexity of the function $f(t) = \log(1 + e^{pt})$ for $p > 0$, we obtain from [6, p. 156] that

$$\det(I_r + |X|^p) \cdot \det(I_{n-r} + |Z|^p) = \det\left(I_n + \begin{bmatrix} |X|^p & 0 \\ 0 & |Z|^p \end{bmatrix}\right) \leqslant \det(I_n + |T|^p).$$

This proves (10).

If $Y = 0$, then clearly

$$\det(I_n + |T|^p) = \det(I_r + |X|^p) \cdot \det(I_{n-r} + |Z|^p).$$

Conversely, if

$$\det(I_n + |T|^p) = \det(I_r + |X|^p) \cdot \det(I_{n-r} + |Z|^p),$$

the strict convexity of $f(t) = \log(1 + e^{pt})$, $p > 0$, gives that

$$\prod_{j=1}^{k} \lambda_j\left(\begin{bmatrix} |X| & U_1Y \\ 0 & |Z| \end{bmatrix}\right) = \prod_{j=1}^{k} \sigma_j\left(\begin{bmatrix} |X| & U_1Y \\ 0 & |Z| \end{bmatrix}\right), \qquad k = 1, \ldots, n.$$

And so $\begin{bmatrix} |X| & U_1Y \\ 0 & |Z| \end{bmatrix}$ is normal (see, e.g., [5, p. 157, P19]]), which is the case only if $U_1Y = 0$ and therefore $Y = 0$. $\square$

Nevertheless, (9) does not have such an analogue. We show by an example that, in general, it is not true that

$$\det(|T_1|+|T_2|) \geqslant \det(|X_1|+|X_2|)\cdot\det(|Z_1|+|Z_2|), \tag{11}$$

where T_1, T_2, are as in Theorem 1.

EXAMPLE 2. Taking

$$T_1 = \left[\begin{array}{cc|cc} 2 & -3 & 9 & -1 \\ -4 & 15 & 1 & -19 \\ \hline 0 & 0 & 0 & -2 \\ 0 & 0 & -4 & 19 \end{array}\right], \qquad T_2 = \left[\begin{array}{cc|cc} 0 & 1 & 6 & 0 \\ 4 & -12 & 12 & 10 \\ \hline 0 & 0 & 14 & -2 \\ 0 & 0 & 23 & -3 \end{array}\right],$$

a simple calculation using Matlab gives

$$\det(|T_1|+|T_2|) = 5193.1 < \det(|X_1|+|X_2|)\cdot\det(|Z_1|+|Z_2|) = 20248.$$

Acknowledgements. The author is grateful to two referees for valuable suggestions. In particular, a referee communicated to the author a shorter proof of Lemma 1 and several insightful comments. The research is supported by a PIMS postdoctoral fellowship.

REFERENCES

[1] D. Z. DJOKOVIĆ, *On some representations of matrices*, Linear Multilinear Algebra **4** (1976) 33–40.
[2] D. Z. DJOKOVIĆ, O. P. LOSSERS, *Problem E2525,* $\det(I+A\overline{A}) \geqslant 0$, Amer. Math. Monthly **83** (1976) 483–484.
[3] S. W. DRURY, *The canonical correlations of a* 2×2 *block matrix with given eigenvalues*, Linear Algebra Appl. **354** (2002) 103–117.
[4] R. A. HORN, C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, 2nd ed., 2013.
[5] R. A. HORN, C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
[6] A. W. MARSHALL, I. OLKIN, B. ARNOLD, *Inequalities: Theory of Majorization and Its Applications*, Springer, New York, 2nd ed., 2011.

(Received October 21, 2014)

Minghua Lin
Department of Mathematics
Shanghai University
Shanghai, 200444, China
and
Department of Mathematics and Statistics
University of Victoria
Victoria, BC, Canada, V8W 3R4
e-mail: mlin87@ymail.com