On the Arithmetic-Geometric-Harmonic-Mean Inequalities for Positive Definite Matrices

T. Ando
Division of Applied Mathematics
Research Institute of Applied Electricity
Hokkaido University
Sapporo, Japan

Dedicated to Professor A. M. Ostrowski on his ninetieth birthday

Submitted by Chandler Davis

ABSTRACT

A sharper form of the arithmetic-geometric-mean inequality for a pair of positive definite matrices is presented, which is a natural analog of the relation

\[
\sqrt{\alpha \beta} + \left( \frac{\alpha - \beta}{2} \right)^2 = \left( \frac{\alpha + \beta}{2} \right)^2
\]

for positive numbers \(\alpha, \beta\). Correspondingly a sharper form of the geometric-harmonic-mean inequality is obtained.

1. INTRODUCTION AND THEOREM

Let \(A\) and \(B\) be positive definite matrices of order \(n\). Their arithmetic mean, in symbols \(A \triangle B\), and harmonic mean, in symbols \(A \dagger B\), are naturally defined respectively by

\[
A \triangle B = \frac{A + B}{2} \quad \text{and} \quad A \dagger B = 2(A^{-1} + B^{-1})^{-1}.
\]

As the definition of the geometric mean, in symbols \(A \# B\), we adopt

\[
A \# B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}.
\]

Despite its appearance, this definition of the geometric mean is symmetric with respect to A and B, and it has strong support from the side of physics (see [4], [6]). The following arithmetic-geometric-harmonic-mean inequalities hold (see [1]):

\[ A!B \leq A \# B \leq A \triangle B. \quad (1) \]

Here the order relation \( S \leq T \) for Hermitian matrices \( S, T \) means the positive semidefiniteness of \( T - S \).

In case \( A \) and \( B \) are positive scalars, or more generally, \( A \) and \( B \) are commuting, (1) has more exact formulations: let

\[ D = \frac{A - B}{2}; \]

then

\[ (A \# B)^2 + D^2 - (A \triangle B)^2 \quad (2) \]

and

\[ (A!B)^2 + \{(A \# B)(A \triangle B)^{-1}D\}^2 \leq (A \# B)^2. \quad (3) \]

The identities (2)–(3), or even the inequality \( D^2 \leq (A \triangle B)^2 \), do not hold in general. To endow (2)–(3) with suitable meanings in the general case, note that \( S \leq T \) is characterized by the condition that \( \psi(S) \leq \psi(T) \) for all positive linear functionals \( \psi \) on \( M_n \), the space of all \( n \times n \) matrices. Here \( \psi \) is said to be positive if \( R \geq 0 \) implies \( \psi(R) \geq 0 \).

Our main results are the following versions of (2)–(3).

**Theorem.** For any positive linear functional \( \psi \) on \( M_n \), the following inequalities hold:

\[ \psi(A \# B)^2 + \psi(D)^2 \leq \psi(A \triangle B)^2 \quad (4) \]

and

\[ \psi(A!B)^2 + \psi((A \# B)(A \triangle B)^{-1}D)^2 \leq \psi(A \# B)^2. \quad (5) \]
If $D$ and $(A \# B)(A \triangle B)^{-1}D$ are replaced by their respective moduli, (4) and (5) are modified as follows:

$$
\psi(A \# B)^2 + \psi(|D|)^2 \leq \sup_{\text{unitary } U} \psi(U^*(A \triangle B)U)^2,
$$

and

$$
\psi(A \triangle B)^2 + \psi((A \# B)(A \triangle B)^{-1}D)^2 \leq \sup_{\text{unitary } U} \psi(U^*(A \# B)U)^2.
$$

The inequalities (4) and (5) can be expressed in terms of order relations among certain matrices, while (6) and (7) can be formulated by using majorization relations, explained later.

2. PROOF OF THEOREM

It is known (see [1]) that the geometric mean $A \# B$ is characterized as the maximum of all $C \geq 0$ for which the $2n$-square matrix $\begin{bmatrix} A & C \\ C & B \end{bmatrix}$ is positive semidefinite. As a consequence,

$$
S(A \# B)S = (SAS) \# (SBS) \quad \text{for all positive definite } S.
$$

Note that if $\begin{bmatrix} A & C \\ C & B \end{bmatrix}$ is positive semidefinite, then for any positive linear functional $\psi$ on $M_n$ the 2-square matrix

$$
\begin{bmatrix}
\psi(A) & \psi(C) \\
\psi(C) & \psi(B)
\end{bmatrix}
$$

is positive semidefinite, so that

$$
\psi(C)^2 \leq \psi(A) \cdot \psi(B),
$$

and in particular

$$
\psi(A \# B)^2 \leq \psi(A) \cdot \psi(B).
$$
Let
\[ C = \frac{1}{2} (A \Delta B)^{-1/2} A (A \Delta B)^{-1/2}. \]

Then application of (8) produces from the definitions of the means
\[
(A \Delta B)^{-1/2} (A \# B) (A \Delta B)^{-1/2} = 2 C^{1/2} (I - C)^{1/2},
\]
and
\[
(A \Delta B)^{-1/2} (A \mid B) (A \Delta B)^{-1/2} = 4 C (I - C),
\]

Then it follows from these relations again via (8) that
\[
A \# B = (A \Delta B) \# \left((A \Delta B) - D (A \Delta B)^{-1} D\right), \quad (11)
\]
\[
A \mid B = (A \# B) \# \left((A \# B) - D (A \Delta B)^{-1} (A \# B) (A \Delta B)^{-1} D\right), \quad (12)
\]
and that \((A \# B)(A \Delta B)^{-1} D\) is Hermitian. According to (10), the identity (11) implies
\[
\psi(A \# B)^2 \leq \psi(A \Delta B) \left(\psi(A \Delta B) - \psi\left(D (A \Delta B)^{-1} D\right)\right). \quad (13)
\]
Since
\[
\begin{bmatrix}
A \Delta B & D \\
D & D (A \Delta B)^{-1} D
\end{bmatrix}
\]
is positive semidefinite, (9) yields
\[
\psi(D)^2 \leq \psi(A \Delta B) \psi\left(D (A \Delta B)^{-1} D\right),
\]
which together with (13) proves (4).
To see (6), use a representation: \( |D| = V^* D \) for some unitary \( V \). Since
\[
\begin{bmatrix}
V^* (A \Delta B) V & |D| \\
|D| & D (A \Delta B)^{-1} D
\end{bmatrix}
\]
is positive semidefinite, (9) yields

\[ \psi(|D|)^2 \leq \psi(D (A \Delta B)^{-1} D) \sup_{\text{unitary } U} \psi(U^*(A \Delta B) U), \]

which together with (13) proves (6).

(5) and (7) can be proved analogously on the basis of identity (12).

3. VARIANTS

Since for any positive numbers \( \xi, \eta \)

\[ (\xi^2 + \eta^2)^{1/2} = \sup_{0 \leq \alpha \leq 1} (\sqrt{\alpha \xi} + \sqrt{1 - \alpha \eta}), \]

(4) and (5) admit the following variants.

**Corollary 1.** For any real \( \alpha \) with \( 0 \leq \alpha \leq 1 \) the following inequalities hold:

\[ \sqrt{\alpha} (A \# B) \pm \sqrt{1 - \alpha} D \leq A \Delta B, \quad (4') \]

and

\[ \sqrt{\alpha} (A \# B) \pm \sqrt{1 - \alpha} (A \# B)(A \Delta B)^{-1} D \leq A \# B. \quad (5') \]

It is known (see [2, p. 72]) that the order relation \( S \leq T \) between two Hermitian matrices \( S, T \) implies \( \lambda_j(S) \leq \lambda_j(T) \), \( j = 1, \ldots, n \), where \( \lambda_1(S) \geq \cdots > \lambda_n(S) \) and \( \lambda_1(T) \geq \cdots > \lambda_n(T) \) are the eigenvalues of \( S \) and \( T \), respectively, arranged in decreasing order with multiplicities counted. A Hermitian matrix \( S \) is said to be **submajorized** by another \( T \), in symbols \( S \prec T \), if

\[ \sum_{j=1}^{k} \lambda_j(S) \leq \sum_{j=1}^{k} \lambda_j(T), \quad k = 1, \ldots, n. \]

If the equality occurs at \( k = n \) in the above, \( S \) is said to be **majorized** by \( T \), in symbols \( S \prec T \). The order relation \( S \leq T \) implies submajorization \( S \prec T \), but not conversely. According to a theorem of Ky Fan (see [2, p. 77], [3, p. 51]),
for any fixed orthogonal projection \( P \) of rank \( k \) (\( \leq n \))

\[
\sum_{j=1}^{k} \lambda_j(S) = \sup_{\text{unitary } U} \text{tr}(PU^*SU).
\]

Thus \( S \preceq T \) follows from the condition that for all positive linear functionals \( \psi \) on \( M_n \)

\[
\psi(S) \leq \sup_{\text{unitary } U} \psi(U^*TU).
\] (14)

The converse is also true (cf. [5]). In fact, diagonalization of Hermitian matrices and the fundamental theorem on majorization for scalar vectors of Hardy, Littlewood, and Polya (see [3, p. 221]) show that if \( S \preceq T \), then there exist unitary matrices \( U_1, \ldots, U_N \) and positive numbers \( \alpha_1, \ldots, \alpha_N \) such that

\[
\sum_{j=1}^{N} \alpha_j = 1 \quad \text{and} \quad S \preceq \sum_{j=1}^{N} \alpha_j U_j^*TU_j,
\]

which implies (14).

Now (6) and (7) admit the following variants.

**COROLLARY 2.** For any real \( \alpha \) with \( 0 \leq \alpha \leq 1 \) the following submajorizations hold:

\[
\sqrt{\alpha} (A \# B) + \sqrt{1-\alpha} |D| \preceq A \triangle B,
\] (6')

and

\[
\sqrt{\alpha} (A \triangledown B) + \sqrt{1-\alpha} |(A \# B)(A \triangle B)^{-1}D| \preceq A \# B.
\] (7')

Submajorization between two positive semidefinite matrices can be characterized by using unitarily invariant seminorms. Recall that a seminorm \( \| \cdot \|_{\bullet} \) on \( M_n \) is said to be unitarily invariant if \( \|UXV\|_{\bullet} = \|X\|_{\bullet} \) for all unitary \( U, V \) and all \( X \in M_n \). Typical examples are the trace norm \( \|X\|_1 = \sum_{j=1}^{n} \lambda_j(\|X\|) \), the Hilbert-Schmidt norm \( \|X\|_2 = (\sum_{j=1}^{n} \lambda_j(\|X\|)^2)^{1/2} \), and the spectral norm \( \|X\|_{\infty} = \lambda_1(\|X\|) \). According to the results of von Neumann and of Ky Fan (see [3, p. 263]), \( S \preceq T \) for positive semidefinite \( S, T \) is equivalent to the condition that
\[ \|S\|_* \leq \|T\|_* \] for all unitarily invariant seminorms \( \|\cdot\|_* \). Thus (6') and (7') admit the following variants.

**Corollary 3.** For any unitarily invariant seminorm \( \|\cdot\|_* \) and any real \( \alpha \) with \( 0 < \alpha < 1 \) the following inequalities hold:

\[
\| \sqrt{\alpha} (A \# B) + \sqrt{1 - \alpha} |D| \| \|_* \leq \| A \triangle B \|_* \tag{6'}
\]

and

\[
\| \sqrt{\alpha} (A! B) + \sqrt{1 - \alpha} (A \# B)(A \triangle B)^{-1} D \| \|_* \leq \| A \# B \|_* \tag{7'}
\]

*For valuable discussions the author would like to thank Fumio Kubo.*

**References**


*Received 23 March 1982*