Interpolating between the Arithmetic-Geometric Mean and Cauchy-Schwarz matrix norm inequalities

Koenraad M.R. Audenaert

Department of Mathematics, Royal Holloway University of London,
Egham TW20 0EX, United Kingdom
Department of Physics and Astronomy, Ghent University,
S9, Krijgslaan 281, B-9000 Ghent, Belgium

Abstract

We prove an inequality for unitarily invariant norms that interpolates between the Arithmetic-Geometric Mean inequality and the Cauchy-Schwarz inequality.

Key words: eigenvalue inequality, matrix norm inequality

1 Introduction

In this paper we prove the following inequality for unitarily invariant matrix norms:

**Theorem 1** Let $|||.|||$ be any unitarily invariant norm. For all $n \times n$ matrices $X$ and $Y$, and all $q \in [0,1]$,

$$||.||^2 \leq |||qX^*X + (1-q)Y^*Y||| |||1-q)X^*X + qY^*Y|||.$$ (1)

For $q = 0$ or $q = 1$, this reduces to the known Cauchy-Schwarz (CS) inequality for unitarily invariant norms ([2], inequality (IX.32))

$$||.||^2 \leq |||X^*X||| |||Y^*Y|||.$$
For $q = 1/2$ on the other hand, this yields the arithmetic-geometric mean (AGM) inequality ([2], inequality (IX.22))

$$\|\|XY^*\|\| \leq \frac{1}{2}\|\|X^*X + Y^*Y\|\|.$$  

Thus, inequality (1) interpolates between the AGM and CS inequalities for unitarily invariant norms.

In Section 2 we prove an eigenvalue inequality that may be of independent interest. The proof of Theorem 1 follows easily from this inequality, in combination with standard majorisation techniques; this proof is given in Section 3.

2 Main technical result

For any $n \times n$ matrix $A$ with real eigenvalues, we will denote these eigenvalues sorted in non-ascending order by $\lambda_k(A)$. Thus $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. Singular values will be denoted as $\sigma_k(A)$, again sorted in non-ascending order.

**Theorem 2** Let $A$ and $B$ be $n \times n$ positive semidefinite matrices. Let $q$ be a number between 0 and 1, and let $C(q) = qA + (1 - q)B$. Then, for all $k = 1, \ldots, n$,

$$\lambda_k(AB) \leq \lambda_k(C(q)C(1 - q)).$$  

**Proof.** The main tool used in the proof is the following eigenvalue monotonicity result due to Lax and Weinberger (see, e.g. Theorem XIII.4.5 in [2]). Let $X$ and $Y$ be $n \times n$ matrices with real eigenvalues such that $Y - X$ has non-negative real eigenvalues. Then $\lambda_k(X) \leq \lambda_k(Y)$, for all $k = 1, \ldots, n$. In the special case that $X$ and $Y$ are Hermitian, this reduces to the well-known Weyl monotonicity principle.

Before we can use this tool, however, we must first reduce the statement of the theorem to a special case. To do so, we use a technique due to Ando [1] that was also used in [3] (see its Section 4, which we follow quite closely). Throughout the proof, we will keep $k$ fixed. We will also assume that $A$ is non-singular. If the theorem holds for all non-singular $A$, then by continuity it must also hold for singular $A$.

If $B$ has rank less than $k$, then $\lambda_k(AB) = 0$ and there is nothing to prove. Henceforth we will therefore assume that $B$ has rank at least $k$. By scaling $A$ and $B$ we can ensure that $\lambda_k(AB) = 1$.  

2
We will now try and find a positive semidefinite matrix $B'$ of rank exactly $k$ with $B' \leq B$ and such that $AB'$ has $k$ eigenvalues equal to 1 and all others equal to 0.

By hypothesis, $AB$ has at least $k$ eigenvalues larger than or equal to 1. Since $AB$ is similar to $A^{1/2}BA^{1/2}$, the same is true for the latter. Therefore, there exists a rank-$k$ projector $P$ satisfying $P \leq A^{1/2}BA^{1/2}$. Let $B' = A^{-1/2}PA^{-1/2}$. Thus, $B' \leq B$, $B'$ is rank $k$, and $AB'$ has the requested spectrum.

Passing to an eigenbasis of $B'$, we can write $B' = \begin{pmatrix} B_{11} & 0 \\ 0 & 0 \end{pmatrix}$, where $B_{11}$ is a $k \times k$ positive definite matrix. In that same basis, let us partition $A$ conformally with $B'$ as $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix}$. Since $A^{1/2}B'A^{1/2}$ is a rank $k$ projector, so is $R := (B')^{1/2}A(B')^{1/2}$. We have

$$R = \begin{pmatrix} (B_{11})^{1/2}A_{11}(B_{11})^{1/2} & 0 \\ 0 & 0 \end{pmatrix}.$$

Since the top-left block of $R$ is a $k \times k$ matrix, and $R$ itself is a rank-$k$ projector, the block must be identical to the $k \times k$ identity matrix. Thus we have $(B_{11})^{1/2}A_{11}(B_{11})^{1/2} = I$, which implies that $B_{11} = (A_{11})^{-1}$. We therefore have, in an eigenbasis of $B'$,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix}, \quad B' = \begin{pmatrix} (A_{11})^{-1} & 0 \\ 0 & 0 \end{pmatrix} \leq B.$$

Clearly, $C'(q) := qA + (1 - q)B'$ satisfies $C'(q) \leq C(q)$, so that $\lambda_k(C'(q)C'(1 - q)) \leq \lambda_k(C(q)C(1 - q))$. On the other hand, $\lambda_k(AB') = \lambda_k(AB) = 1$. It is left to show that $\lambda_k(C'(q)C'(1 - q)) \geq 1$.

We are now in a position to apply the Lax-Weinberger monotonicity principle. We will consider the matrices $X = qAB' + (1 - q)B'A$ and $Y = C'(q)C'(1 - q)$. The matrix $Y$ clearly has real eigenvalues, as $C'(q)$ and $C'(1 - q)$ are positive semidefinite. That $X$ also has real eigenvalues is not so obvious; it is not true in general for $qAB + (1 - q)BA$, which explains why the reduction step to $B'$ was necessary.
An explicit expression for $X$ is

$$
X = \begin{pmatrix}
\mathbb{I} & (1-q)(A_{11})^{-1}A_{12} \\
qA_{12}^*(A_{11})^{-1} & 0
\end{pmatrix}.
$$

Replacing $(A_{11})^{-1}A_{12}$ by its singular value decomposition $U\Sigma V^*$ gives

$$
X = \begin{pmatrix}
\mathbb{I} & (1-q)U\Sigma V^* \\
qV\Sigma^TU^* & 0
\end{pmatrix}.
$$

Here, $\Sigma$ is a (pseudo)-diagonal $k \times (n-k)$ matrix with $m := \min(k, n-k)$ diagonal elements $\sigma_j \geq 0$. Thus, $X$ is (unitarily) similar to

$$
X' = \begin{pmatrix}
\mathbb{I} & (1-q)\Sigma \\
q\Sigma^TU^* & 0
\end{pmatrix}.
$$

This $X'$ can be written as a direct sum of $m$ matrices $X_j = \begin{pmatrix} 1 & (1-q)\sigma_j \\
q\sigma_j & 0 \end{pmatrix}$ and, depending on the sign of $n-2k$, either $[0]_{n-2k}$ or $\mathbb{I}_{2k-n}$. Since $X_j$ has trace 1 and non-positive real determinant, it has one eigenvalue larger than or equal to 1, and one non-positive eigenvalue. Hence, $X$ has $k$ eigenvalues larger than or equal to 1, and $n-k$ non-positive ones. In particular, $\lambda_k(X) \geq 1$.

Thus both $X$ and $Y$ have real eigenvalues. To see that $Y - X$ has non-negative real eigenvalues, note that

$$
Y = C'(q)C'(1-q)
= (qA + (1-q)B')(qB' + (1-q)A)
= q(1-q)(A^2 + B'^2) + q^2AB' + (1-q)^2B'A
= q(1-q)(A - B')^2 + qAB' + (1-q)B'A.
$$

Thus, $Y - X = q(1-q)(A - B')^2$, which is positive semidefinite and obviously has non-negative eigenvalues. This means that the Lax-Weinberger monotonicity principle applies, and we get $\lambda_j(X) \leq \lambda_j(Y)$ for $j = 1, \ldots, n$. Combined with the previously obtained fact $\lambda_k(X) \geq 1$, we finally get $\lambda_k(Y) \geq 1$, which proves the theorem. $\square$
Writing $A = X^*X$ and $B = Y^*Y$, for general matrices $X$ and $Y$, and noting that
\[ \lambda_k^{1/2}(AB) = \lambda_k^{1/2}(YX^*XY^*) = \sigma_k(XY^*), \]
we can write (2) as a singular value inequality:
\[ \sigma_k^2(XY^*) \leq \lambda_k((qX^*X + (1-q)Y^*Y)((1-q)X^*X + qY^*Y)). \] (3)

For $p = 1/2$, Theorem 2 gives
\[ \lambda_k^{1/2}(AB) \leq \frac{1}{2} \lambda_k(A + B). \] (4)

and (3) becomes the well-known AGM inequality for singular values ([2], inequality (IX.20))
\[ \sigma_k(XY^*) \leq \frac{1}{2} \sigma_k(X^*X + Y^*Y). \]

We suspect that inequality (2) also holds as a singular value inequality; that is:

**Conjecture 1** Let $A$ and $B$ be $n \times n$ positive semidefinite matrices. Let $q$ be a number between 0 and 1, and let $C(q) = qA + (1-q)B$. Then, for all $k = 1, \ldots, n$,
\[ \sigma_k(AB) \leq \sigma_k(C(q)C(1-q)). \] (5)

### 3 Proof of Theorem 1

Using Theorem 2 and some standard arguments, the promised norm inequality is easily proven.

For all positive semidefinite matrices $A$ and $B$, and any $r > 0$, we have the weak majorisation relation
\[ \lambda^r(AB) \preceq_w \lambda^r(A) \cdot \lambda^r(B), \]
where $\cdot$ denotes the elementwise product for vectors. This relation follows from combining Weyl’s majorant inequality ([2], inequality (II.23))
\[ |\lambda(AB)|^r \preceq_w \sigma^r(AB) \]
with the singular value majorisation relation ([2], inequality (IV.41))

\[ \sigma^r(AB) \prec_w \sigma^r(A) \cdot \sigma^r(B) \]

From (3) we immediately get, for any \( r > 0 \),

\[ \sigma^{2r}(XY^*) \prec_w \lambda^r ((qX^*X + (1-q)Y^*Y) ((1-q)X^*X + qY^*Y)). \]

Hence,

\[ \sigma^{2r}(XY^*) \prec_w \lambda^r (qX^*X + (1-q)Y^*Y) \cdot \lambda^r((1-q)X^*X + qY^*Y)). \]

If we now apply Hölder’s inequality for symmetric gauge functions \( \Phi \),

\[ \Phi(|x \cdot y|) \leq \Phi(|x|^p)^{1/p} \Phi(|y|^{p'})^{1/p'}, \]

where \( x, y \in \mathbb{C}^n \) and \( 1/p + 1/p' = 1 \), we obtain

\[ \Phi(\sigma^{2r}(XY^*)) \leq \Phi(\lambda^r (qX^*X + (1-q)Y^*Y) \cdot \lambda^r((1-q)X^*X + qY^*Y)) \]
\[ \leq \Phi(\lambda^{rp}(qX^*X + (1-q)Y^*Y))^{1/p} \Phi(\lambda^{rp'}((1-q)X^*X + qY^*Y))^{1/p'}. \]

Hence, for any unitarily invariant norm,

\[ |||XY^*|||^2 \| \leq |||(qX^*X + (1-q)Y^*Y)^{rp}|||^1/p |||(1-q)X^*X + qY^*Y)^{rp'}|||^1/p'. \]

Theorem \([\Pi]\) follows by setting \( r = 1/2 \) and \( p = p' = 2 \). \( \square \)

**Acknowledgments**

We acknowledge support by an Odysseus grant from the Flemish FWO.

**References**

[1] T. Ando, “Matrix Young inequalities”, Operator Theory: Advances and Applications 75, 33–38 (1995).
[2] R. Bhatia, *Matrix Analysis*, Springer, Heidelberg (1997).
[3] S.W. Drury, “On a question of Bhatia and Kittaneh”, Linear Algebra Appl. 437, 1955–1960 (2012).