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
1991 MSC: 15A60

1 Introduction

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

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

$$|||XY^*|||^2 \le |||qX^*X + (1-q)Y^*Y||| \quad |||(1-q)X^*X + qY^*Y|||. \tag{1}$$

For $q = 0$ or $q = 1$, this reduces to the known Cauchy-Schwarz (CS) inequality for unitarily invariant norms [2, (IX.32)], [3], [6]

$$|||XY^*|||^2 \le |||X^*X||| \quad |||Y^*Y|||.$$

Email address: koenraad.audenaert@rhul.ac.uk (Koenraad M.R. Audenaert).

For $q = 1/2$ on the other hand, this yields the arithmetic-geometric mean (AGM) inequality [2, (IX.22)], [4]

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

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 $\lambda_1(A) \ge \cdots \ge \lambda_n(A)$. Singular values will be denoted as $\sigma_k(A)$, again sorted in non-ascending order.

Our main technical tool in proving Theorem 1 is the following eigenvalue inequality, which may be of independent interest:

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 $k = 1, \ldots, n$,*

$$\lambda_k(AB) \le \lambda_k(C(q)C(1-q)). \tag{2}$$

Putting $A = X^*X$ and $B = Y^*Y$, for $n \times n$ 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:

$$\sigma_k^2(XY^*) \le \lambda_k((qX^*X + (1-q)Y^*Y)((1-q)X^*X + qY^*Y)). \tag{3}$$

For $p = 1/2$, Theorem 2 gives

$$\lambda_k^{1/2}(AB) \le \frac{1}{2}\,\lambda_k(A+B) \tag{4}$$

and (3) becomes the well-known AGM inequality for singular values [2, inequality (IX.20)]

$$\sigma_k(XY^*) \le \frac{1}{2}\,\sigma_k(X^*X + Y^*Y).$$

The following modification of inequality (2), proven by Drury for the case $q = 1/2$ [5], does not hold for all $q \in [0,1]$:

$$\sigma_k(AB) \le \sigma_k(C(q)C(1-q)).$$

We are grateful to Swapan Rana for informing us about counterexamples.

Proof of Theorem 2. We first reduce the statement of the theorem to a special case using a technique that is due to Ando [1] and that was also used in [5, Section 4].

Throughout the proof, we will keep k fixed. If either A or B has rank less than k, then $\lambda_k(AB) = 0$ and (2) holds trivially. We will therefore assume that A and B have rank at least k. By scaling A and B we can ensure that $\lambda_k(AB) = 1$.

We will now try and find a positive semidefinite matrix B' of rank exactly k with $B' \le B$ and such that AB' has k eigenvalues equal to 1 and all others equal to 0. By hypothesis, AB and hence $A^{1/2}BA^{1/2}$ have at least k eigenvalues larger than or equal to 1. Therefore, there exists a rank-k projector P satisfying $P \le A^{1/2}BA^{1/2}$. Let B' be a rank-k matrix such that $A^{1/2}B'A^{1/2} = P$. If A is invertible, we simply have $B' = A^{-1/2}PA^{-1/2}$; otherwise the generalised inverse of A is required. Then $B' \le B$ and AB' has the requested spectrum.

Passing to an eigenbasis of B', we can decompose B' as the direct sum $B' = B_{11} \oplus [0]_{n-k}$, where B_{11} is a $k \times k$ positive definite block. In that same basis, we partition A conformally with B' as $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{pmatrix}$. Since $A^{1/2}B'A^{1/2} = P$ is a rank k projector, so is

$$R := (B')^{1/2}A(B')^{1/2} = (B_{11})^{1/2}A_{11}(B_{11})^{1/2} \oplus [0]_{n-k}.$$

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

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

Clearly, $C'(q) := qA + (1-q)B'$ satisfies $C'(q) \leq C(q)$, so that

$$\lambda_k(C'(q)C'(1-q)) \leq \lambda_k(C(q)C(1-q)),$$

while still $\lambda_k(AB') = \lambda_k(AB) = 1$. It is now left to show that $\lambda_k(C'(q)C'(1-q)) \geq 1$.

A further reduction is possible. Let

$$A' = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{12}^*(A_{11})^{-1}A_{12} \end{pmatrix},$$

which has rank k and satisfies $0 \leq A' \leq A$. Let also $C''(q) := qA' + (1-q)B'$, for which $0 \leq C''(q) \leq C'(q)$. Then $\lambda_k(C''(q)C''(1-q)) \leq \lambda_k(C'(q)C'(1-q))$.

Introducing $F := A_{11} > 0$, $G := A_{12}A_{12}^* \geq 0$ and $s := (1-q)/q > 0$, we have

$$\begin{aligned} C''(q) &= q \begin{pmatrix} F & A_{12} \\ A_{12}^* & A_{12}^*F^{-1}A_{12} \end{pmatrix} + (1-q) \begin{pmatrix} F^{-1} & \\ & 0 \end{pmatrix} \\ &= q \begin{pmatrix} I & \\ & A_{12}^* \end{pmatrix} \begin{pmatrix} F + sF^{-1} & I \\ I & F^{-1} \end{pmatrix} \begin{pmatrix} I & \\ & A_{12} \end{pmatrix} \end{aligned}$$

so that

$$\begin{aligned} &\lambda_k(C''(q)C''(1-q)) \\ &= q(1-q)\lambda_k \left(\begin{pmatrix} I & \\ & G \end{pmatrix} \begin{pmatrix} F + sF^{-1} & I \\ I & F^{-1} \end{pmatrix} \begin{pmatrix} I & \\ & G \end{pmatrix} \begin{pmatrix} F + s^{-1}F^{-1} & I \\ I & F^{-1} \end{pmatrix} \right), \end{aligned}$$

where each factor is a $2k \times 2k$ matrix. Noting that

$$\begin{pmatrix} F + sF^{-1} & I \\ I & F^{-1} \end{pmatrix} = \begin{pmatrix} s^{1/2}F^{-1/2} & F^{1/2} \\ 0 & F^{-1/2} \end{pmatrix} \begin{pmatrix} s^{1/2}F^{-1/2} & 0 \\ F^{1/2} & F^{-1/2} \end{pmatrix},$$

we then have $\lambda_k(C''(q)C''(1-q)) = q(1-q)\lambda_k(Z^*Z) = q(1-q)\sigma_k^2(Z)$, where

$$Z = \begin{pmatrix} s^{1/2}F^{-1/2} & 0 \\ F^{1/2} & F^{-1/2} \end{pmatrix} \begin{pmatrix} I & \\ & G \end{pmatrix} \begin{pmatrix} s^{-1/2}F^{-1/2} & F^{1/2} \\ 0 & F^{-1/2} \end{pmatrix} = \begin{pmatrix} F^{-1} & s^{1/2} \\ s^{-1/2} & F+H \end{pmatrix},$$

and $H := F^{-1/2}GF^{-1/2} \geq 0$. The singular values of Z are the same as those of

$$X := \begin{pmatrix} s^{1/2} & F^{-1} \\ F+H & s^{-1/2} \end{pmatrix}.$$

By the Fan-Hoffman theorem [2, Proposition III.5.1], the singular values of X are bounded below by the ordered eigenvalues of the Hermitian part of X: $\sigma_j(X) \geq \lambda_j((X+X^*)/2)$ for $j = 1, \ldots, 2k$. Thus,

$$\lambda_k(C''(q)C''(1-q)) \geq q(1-q)\lambda_k^2(Y),$$

$$\text{with } Y := \begin{pmatrix} s^{1/2} & K \\ K & s^{-1/2} \end{pmatrix} \text{ and } K := (F+H+F^{-1})/2.$$

Clearly, $K \geq (F+F^{-1})/2 \geq I$. It is easily checked that the k largest eigenvalues of Y are given by

$$\lambda_j(Y) = \frac{1}{2}\left(s^{1/2} + s^{-1/2} + \sqrt{(s^{1/2}+s^{-1/2})^2 - 4 + 4\lambda_j^2(K)}\right), \quad j = 1, \ldots, k.$$

As this expression is a monotonously increasing function of $\lambda_j(K)$, and $\lambda_j(K) \geq 1$, we obtain the lower bound $\lambda_k(Y) \geq s^{1/2} + s^{-1/2}$. Then, finally,

$$\begin{aligned} \lambda_k(C''(q)C''(1-q)) &\geq q(1-q)\ (s^{1/2}+s^{-1/2})^2 \\ &= q(1-q)\left(\left(\frac{1-q}{q}\right)^{1/2} + \left(\frac{q}{1-q}\right)^{1/2}\right)^2 \\ &= (1-q+q)^2 = 1, \end{aligned}$$

from which it follows that $\lambda_k(C'(q)C'(1-q)) \geq 1$. $\square$

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

$$\lambda^r(AB) \prec_w \lambda^r(A) \cdot \lambda^r(B),$$

where '$\cdot$' denotes the elementwise product for vectors. This relation follows from combining the fact that AB has non-negative eigenvalues with Weyl's majorant inequality [2, (II.23)],

$$|\lambda(AB)|^r \prec_w \sigma^r(AB)$$

and 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 $r > 0$,

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

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 Φ,

$$\Phi(|x \cdot y|) \le \Phi(|x|^p)^{1/p}\, \Phi(|y|^{p'})^{1/p'},$$

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

$$\begin{aligned}\Phi(\sigma^{2r}(XY^*)) &\le \Phi(\lambda^r(qX^*X + (1-q)Y^*Y) \cdot \lambda^r((1-q)X^*X + qY^*Y))) \\ &\le \Phi(\lambda^{rp}(qX^*X + (1-q)Y^*Y))^{1/p}\, \Phi(\lambda^{rp'}((1-q)X^*X + qY^*Y)))^{1/p'}.\end{aligned}$$

Hence, for any unitarily invariant norm,

$$|||\, |XY^*|^{2r}\, ||| \le |||(qX^*X + (1-q)Y^*Y)^{rp}|||^{1/p}\, |||((1-q)X^*X + qY^*Y)^{rp'}|||^{1/p'}.$$

Theorem 1 now follows by setting $r = 1/2$ and $p = p' = 2$. $\square$

Acknowledgments

We acknowledge support by an Odysseus grant from the Flemish FWO. We are grateful to Professor Bhatia for pointing out a serious mistake in an earlier circulated version of this paper.

References

[1] T. Ando, "Matrix Young inequalities", Operator Theory: Advances and Applications **75**, 33–38 (1995).

[2] R. Bhatia, "Perturbation inequalities for the absolute value map in norm ideals of operators", Journal of Operator Theory **19**, 129–136 (1988).

[3] R. Bhatia, *Matrix Analysis*, Springer, Heidelberg (1997).

[4] R. Bhatia and F. Kittaneh, "On the singular values of a product of operators", SIAM J. Matrix Analysis **11**, 272–277 (1990).

[5] S.W. Drury, "On a question of Bhatia and Kittaneh", Linear Algebra Appl. **437**, 1955–1960 (2012).

[6] R. Horn and R. Matthias, "Cauchy-Schwarz inequalities associated with positive semidefinite matrices", Linear Algebra Appl. **142**, 63–82 (1990).