An Analog of the Cauchy-Schwarz Inequality for Hadamard Products and Unitarily Invariant Norms

Roger A. Horn and Roy Mathias
The Johns Hopkins University, Baltimore, Maryland 21218

SIAM J. Matrix Analysis and Applications, Vol. 11, No. 4, 481-498, 1990.

Abstract

We show that for any unitarily invariant norm $\|\cdot\|$ on M_n (the space of n-by-n complex matrices)

$$\|A^*B\|^2 \quad \leq \quad \|A^*A\|\,\|B^*B\| \text{ for all } A, B \in M_{m,n} \tag{1}$$

and

$$\|A \circ B\|^2 \quad \leq \quad \|A^*A\|\,\|B^*B\| \text{ for all } A, B \in M_n$$

where $\circ$ denotes the Hadamard (entrywise) product. These results are a consequence of an inequality for absolute norms on C^n

$$\|x \circ y\|^2 \leq \|x \circ \bar{x}\|\,\|y \circ \bar{y}\| \text{ for all } x, y \in C^n. \tag{2}$$

We also characterize the norms on C^n that satisfy (2), characterize the unitary similarity invariant norms on M_n that satisfy (1), and obtain related results on norms on C^n and unitary similarity invariant norms on M_n that are of independent interest.

1 Introduction and Notation

Let $M_{m,n}$ denote the space of m-by-n complex matrices and write $M_n \equiv M_{n,n}$; let $A^* \equiv \bar{A}^*$ denote the conjugate transpose of a matrix in $M_{m,n}$. Recently, Harald Wimmer [20, p. 315] conjectured that an analog of the Cauchy-Schwarz inequality holds for any unitarily invariant norm $\|\cdot\|$ on $M_{m,n}$:

$$\|A^*B\|^2 \leq \|A^*A\|\,\|B^*B\| \text{ for all } A, B \in M_{m,n} \tag{1.1}$$

For three special choices of norm $\|\cdot\|$ (the trace norm, the Frobenius norm, and the spectral norm), Wimmer proved (1.1) and identified the cases of equality.

In Section 3 we give a proof of (1.1) and a similar inequality for Hadamard products; both results follow from a simple norm inequality (Theorem 2.3) for the Hadamard product of vectors. We identify the cases of equality in this inequality and in (1.1). In Sections 3 and 4 we prove some results of independent interest on unitary similarity invariant norms. In Section 4 we provide a variety of examples and show that the set of norms satisfying (1.1) is a convex set that strictly contains the unitarily invariant norms.

We use $A \succeq 0$ to mean that A is positive semidefinite. If $A \succeq 0$ then $A^{1/2}$ denotes the unique positive semidefinite square root of A. Given $A \in M_{m,n}$ we define $|A| \equiv (A^*A)^{1/2}$. The real vector space of n-by-n Hermitian matrices is denoted by H_n. If $A, B \in H_n$, we write $A \succeq B$ if $A - B \succeq 0$. Recall that the *Hadamard product* of two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ of the same size is $A \circ B \equiv [a_{ij}b_{ij}]$. We denote the ordered singular values of any $A \in M_{m,n}$ by $\sigma_1(A) \geq \ldots \geq \sigma_q(A) \geq 0$ (where $q = \min\{m,n\}$) and define $\sigma(A) \equiv [\sigma_1(A), \ldots, \sigma_q(A)]^T \in R_+^q$; for $A \in H_n$ we denote the ordered eigenvalues of A by $\lambda_1(A) \geq \ldots \geq \lambda_n(A)$ and define $\lambda(A) \equiv [\lambda_1(A), \ldots, \lambda_n(A)]^T \in R^n$. The eigenvalues and singular values of a positive semidefinite matrix are identical. A complex matrix is a *partial isometry* if each of its singular values is 0 or 1. The *trace* of a square matrix A (the sum of its main diagonal entries, or, equivalently, the sum of its eigenvalues) is denoted by tr A.

Given $x \in C^n$ and an index set $\mathcal{I} \subset \{1, \ldots, n\}$ we define $x(\mathcal{I}) \in C^n$ by

$$x(\mathcal{I})_i = \begin{cases} x_i & \text{if } i \in \mathcal{I} \\ 0 & \text{if } i \notin \mathcal{I} \end{cases}$$

and we define $|x| \equiv [|x_i|]_{i=1}^n$. Given a vector $x \in C^n$ we define $\mathrm{diag}(x) \in M_n$ to be the diagonal matrix with i,i entry x_i. Given vectors $x, y \in R^n$ we use $x \leq y$ to mean that $x_i \leq y_i$ for $i = 1, \ldots, n$. A norm $\|\cdot\|$ on C^n is *absolute* if $\|x\| = \|\,|x|\,\|$ for all $x \in C^n$, and is *monotone* if $|x| \geq |y|$ implies $\|x\| \geq \|y\|$. These two notions were introduced by Bauer, Stoer and Witzgall in [2], where they arose naturally in the study of induced norms on M_n. It is a fact that a norm is absolute if and only if it is monotone [2, Theorem 2] or [7, Theorem 5.5.10].

If vectors $x, y \in R^n$ are given, and if τ and π are permutations of $\{1,\ 2,\ \ldots,\ n\}$ such that $x_{\tau(1)} \geq x_{\tau(2)} \geq \cdots \geq x_{\tau(n)}$ and $y_{\pi(1)} \geq y_{\pi(2)} \geq \cdots \geq y_{\pi(n)}$, we say that x is *weakly majorized* by y if

$$\sum_{i=1}^k x_{\tau(i)} \leq \sum_{i=1}^k y_{\pi(i)} \quad \text{for all } k = 1, \ldots, n.$$

If, in addition, equality holds when $k = n$, then we say that x is *majorized* by y. A function $g(\cdot) : C^n \to R_+$ is called a *symmetric gauge function* if it is a permutation invariant absolute norm on C^n. We will make frequent use of the fact that for $x, y \in C^n$ and any symmetric gauge function $g(\cdot)$

$$|x| \text{ is weakly majorized by } |y| \quad \text{implies} \quad g(x) \leq g(y). \tag{1.2}$$

Given a norm $\|\cdot\|$ on C^m we define its *dual* (with respect to the Euclidean inner product) by

$$\|x\|^D \quad \equiv \quad \max\{|y^*x| : y \in C^m, \|y\| \leq 1\}. \tag{1.3}$$

Given a norm $\|\cdot\|$ on $M_{m,n}$ we define its *dual* (with respect to the Frobenius inner product $< A, B > \equiv \mathrm{tr}\ B^*A$) by

$$\|A\|^D \equiv \max\{|\mathrm{tr}\ (B^*A)| : B \in M_{m,n}, \|B\| \leq 1\}.$$

If we take $n = 1$, then this definition specializes to (1.3). The duality theorem for norms [7, Theorem 5.5.14] states that $\|\cdot\| = (\|\cdot\|^D)^D$ for any norm $\|\cdot\|$. A norm $\|\cdot\|$ on $M_{m,n}$ is *unitarily invariant* if $\|A\| = \|UAV\|$ for all $A \in M_{m,n}$ and all unitary $U \in M_m$ and $V \in M_n$. A theorem of von Neumann [19] (or [7, Theorem 7.4.24], or [16, Theorem V.5]) states that a norm $\|\cdot\|$ on $M_{m,n}$ is unitarily invariant if and only if there is a symmetric gauge function g such that $\|X\| = g(\sigma(X))$ for all $X \in M_{m,n}$. A norm $\|\cdot\|$ on M_n is *unitary similarity invariant* if $\|A\| = \|UAU^*\|$ for all $A, U \in M_n$ with U unitary.

See [7] for further information on Hadamard products, norms, dual norms, unitarily invariant norms, symmetric gauge functions, singular values, and other concepts discussed in this paper. See [13, p. 263] for a general discussion of the connection between majorization and unitarily invariant norms.

2 An Inequality for Absolute Norms

In this section we are interested in an inequality for Hadamard products of vectors that leads directly to a proof of the matrix inequality (1.1). To obtain Theorem 2.3, the main result in this section, it is helpful to know two lemmata, whose proofs we omit. The first result is Theorem 1 in [2]; the second can be proved by an argument very similar to the proof of Lemma 3.7.

Lemma 2.1 *A norm on C^n is absolute if and only if its dual norm is absolute.*

Lemma 2.2 *Let $\|\cdot\|$ be an absolute norm on C^n and let $x \in C^n$ be given. Then*

$$\begin{aligned} \|x\| &= \max\{|y^*x| : y \in C^n \text{ and } \|y\|^D \le 1\} \\ &= \max\{z^T|x| : z \in R^n_+ \text{ and } \|z\|^D \le 1\}. \end{aligned}$$

We use the following notation in Theorem 2.3. Given $x \in C^n$ and an index set $\mathcal{I} \subset \{1, \ldots, n\}$, define $x(\mathcal{I}) \in C^n$ by

$$x(\mathcal{I})_i = \begin{cases} x_i & \text{if } i \in \mathcal{I} \\ 0 & \text{if } i \notin \mathcal{I} \end{cases}.$$

Theorem 2.3 *Let $\|\cdot\|$ be an absolute norm on C^n. Then*

$$\|x \circ y\|^2 \le \|x \circ \bar{x}\| \; \|y \circ \bar{y}\| \quad \text{for all } x, y \in C^n. \tag{2.1}$$

If $x, y \neq 0$ then equality holds in (2.1) if and only if there is a positive constant c and an index set $\mathcal{I} \subset \{1, \ldots, n\}$ such that

$$|x(\mathcal{I})| = c|y(\mathcal{I})|, \; \|x \circ \bar{x}(\mathcal{I})\| = \|x\|, \;\; \text{and} \;\; \|y \circ \bar{y}(\mathcal{I})\| = \|y\|.$$

Proof: Use Lemma 2.2 to compute

$$\|x \circ y\|^2 = [\max\{z^T(|x| \circ |y|) : z \in R^n_+ \text{ and } \|z\|^D \le 1\}]^2 \tag{2.2}$$

$$
\begin{aligned}
&= \max\{[(z^{1/2}\circ|x|)^T(z^{1/2}\circ|y|)]^2 : z\in R^n_+ \text{ and } \|z\|^D\le 1\} \\
&\le \max\{[(z^{1/2}\circ|x|)^T(z^{1/2}\circ|x|)][(z^{1/2}\circ|y|)^T(z^{1/2}\circ|y|)] : \\
&\qquad\qquad z\in R^n_+ \text{ and } \|z\|^D\le 1\} \qquad (2.3)\\
&\le \max\{(z^T|x\circ x|)] : z\in R^n_+ \text{ and } \|z\|^D\le 1\} \\
&\qquad \cdot\max\{(z^T|y\circ y|)] : z\in R^n_+ \text{ and } \|z\|^D\le 1\} \qquad (2.4)\\
&= \|\,|x\circ x|\,\|\,\|\,|y\circ y|\,\| \\
&= \|x\circ\bar{x}\|\,\|y\circ\bar{y}\|.
\end{aligned}
$$

For $z=[z_i]\in R^n_+$, we have written $z^{1/2}\equiv[z_i^{1/2}]\in R^n_+$ for the Hadamard (entrywise) nonnegative square root of z.

That the the stated conditions are sufficient for equality is clear from the monotonicity of $\|\cdot\|$:

$$
\begin{aligned}
\|x\circ y\|^2 &\ge \|x\circ y(\mathcal{I})\|^2 \\
&= \|x\circ(1/c)x(\mathcal{I})\|\,\|cy\circ y(\mathcal{I})\| \\
&= \|x\circ\bar{x}(\mathcal{I})\|\,\|y\circ\bar{y}(\mathcal{I})\| \\
&= \|x\circ\bar{x}\|\,\|y\circ\bar{y}\|
\end{aligned}
$$

and hence equality holds in (2.1).

Conversely, suppose that equality holds in (2.1) for given non-zero vectors x, y. Then any $z\in R^n_+$ that attains the maximum in (2.2) must also attain the maximum in (2.3) and both maxima in (2.4). Let $\tilde{z}$ be such a vector and define the index set $\mathcal{I}\equiv\{i:\tilde{z}_i>0\}$. Equality in (2.3) implies that there is a positive scalar c such that

$$\tilde{z}_i^{1/2}|x|_i = c\tilde{z}_i^{1/2}|y|_i \quad \text{for } \ i=1,\ldots,n$$

and hence, by the definition of $\mathcal{I}$, it follows that $|x(\mathcal{I})| = c|y(\mathcal{I})|$. Because $\tilde{z}$ attains the first maximum in (2.4), we have $\|x\circ\bar{x}\| = \tilde{z}^T|x\circ\bar{x}|$, and we can use Lemma 2.2 again to compute

$$
\begin{aligned}
\|x\circ\bar{x}(\mathcal{I})\| &= \max\{z^T|x\circ\bar{x}(\mathcal{I})| : z\in R^n_+ \text{ and } \|z\|^D\le 1\} \\
&\ge \tilde{z}^T(x\circ\bar{x}) \\
&= \|x\circ\bar{x}\|.
\end{aligned}
$$

But $\|x\circ\bar{x}\|\ge\|x\circ\bar{x}(\mathcal{I})\|$ by the monotonicity of $\|\cdot\|$, so $\|x\circ\bar{x}\| = \|x\circ\bar{x}(\mathcal{I})\|$. The same argument shows that $\|y\circ\bar{y}\| = \|y\circ\bar{y}(\mathcal{I})\|$. ■

Notice that the inequality (2.1) with the l_1 norm is the heart of the classical Cauchy-Schwarz inequality:

$$|\sum_{i=1}^n x_iy_i|^2 \le \left\{\sum_{i=1}^n |x_iy_i|\right\}^2 = \|x\circ y\|_1^2 \le \|x\circ\bar{x}\|_1\,\|y\circ\bar{y}\|_1 = \sum_{i=1}^n |x_i|^2\sum_{i=1}^n |y_i|^2.$$

It is of interest to characterize the norms on C^n that satisfy the conclusion of Theorem 2.3. We discuss the converse of the following preliminary lemma in Theorem 4.8.

Lemma 2.4 *Let $\|\cdot\|$ be a norm on C^n such that $\|x\| \leq \|\,|x|\,\|$ for all $x = [x_i] \in C^n$, where $|x| \equiv [|x_i|]$. Then the function $\nu(x) \equiv \|\,|x|\,\|$ is a norm on C^n.*

Proof: Since the function $\nu(x) \equiv \|\,|x|\,\|$ is positive definite and homogeneous on C^n, we need only show that it obeys the triangle inequality. We claim that it suffices to prove that

$$\|u\| \leq \|v\| \text{ whenever } u, v \in R_+^n \text{ and } u \leq v. \tag{2.5}$$

Since $|x+y| \leq |x| + |y|$ for all $x, y \in C^n$, (2.5) and the triangle inequality for $\|\cdot\|$ give the desired result:

$$\nu(x+y) = \|\,|x+y|\,\| \leq \|\,|x| + |y|\,\| \leq \|\,|x|\,\| + \|\,|y|\,\| = \nu(x) + \nu(y).$$

To prove (2.5), let $u, v \in R_+^n$ be given with $u \leq v$. If $u = v$, there is nothing to prove, so assume that $u \neq v$. Some corresponding entries of u and v may be equal but at least one entry of u must be strictly less than the corresponding entry of v. We shall construct a vector $w \in R_+^n$ such that $u \leq w \leq v$, $\|w\| \leq \|v\|$, and w has one more entry than v that is equal to the corresponding entry of u. A finite induction then leads to the conclusion that $\|u\| \leq \|v\|$. Define $k = \min\{i : u_i < v_i, i = 1, \ldots, n\}$ and define $v', v'' \in R^n$ by

$$v_i' = \begin{cases} v_i & \text{for } i \neq k \\ -v_i & \text{for } i = k \end{cases}, \qquad v_i'' = \begin{cases} v_i & \text{for } i \neq k \\ 0 & \text{for } i = k \end{cases}.$$

Notice that $v'' = \frac{1}{2}(v' + v)$ and $|v'| = v$. Using the hypothesis on $\|\cdot\|$, we have

$$\|v'\| \leq \|\,|v'|\,\| \leq \|v\|,$$

and hence

$$\|v''\| = \tfrac{1}{2}\|v + v'\| \leq \tfrac{1}{2}(\|v\| + \|v'\|) \leq \tfrac{1}{2}(\|v\| + \|v\|) = \|v\|. \tag{2.6}$$

Now define $\alpha \equiv u_k/v_k$, so $0 \leq \alpha < 1$. Define $w \equiv \alpha v + (1-\alpha)v''$ and notice that $w_i = v_i$ if $i \neq k$ and that $w_k = u_k$. Thus, w has one more entry than v that is equal to the corresponding entry of u. Using (2.6) we obtain

$$\|w\| = \|\alpha v + (1-\alpha)v''\| \leq \alpha\|v\| + (1-\alpha)\|v''\| \leq \alpha\|v\| + (1-\alpha)\|v\| = \|v\|,$$

as desired. ■

We can now characterize the norms that satisfy the inequality (2.1).

Theorem 2.5 *Let $\|\cdot\|$ be a norm on C^n. Then*

$$\|x \circ y\|^2 \leq \|x \circ \bar{x}\|\,\|y \circ \bar{y}\| \;\; \textit{for all} \;\; x, y \in C^n \tag{2.7}$$

if and only if

$$\|z\| \leq \|\,|z|\,\| \;\; \textit{for all} \;\; z = [z_i] \in C^n, \tag{2.8}$$

where $|z| \equiv [|z_i|]$.

Proof: Suppose $\|z\| \leq \| \, |z| \, \|$ for all $z \in C^n$. Lemma 2.4 guarantees that $\nu(x) \equiv \| \, |x| \, \|$ is an absolute norm on C^n, so we may apply Theorem 2.3 to $\nu(\cdot)$ and obtain

$$\|x \circ y\|^2 \leq \| \, |x \circ y| \, \|^2 = \nu^2(x \circ y) \leq \nu(x \circ \bar{x})\nu(y \circ \bar{y}) = \|x \circ \bar{x}\| \, \|y \circ \bar{y}\|.$$

Conversely, suppose (2.7) holds and let $z \in C^n$ be given. Define $x, y \in C^n$ by

$$x_i \equiv \begin{cases} z_i/|z_i|^{1/2} & \text{if } z_i \neq 0 \\ 0 & \text{if } z_i = 0 \end{cases}, \quad y_i \equiv |z_i|^{1/2} \quad i = 1, \ldots, n.$$

Then

$$\|z\|^2 = \|x \circ y\|^2 \leq \|x \circ \bar{x}\| \, \|y \circ \bar{y}\| = \| \, |z| \, \| \, \| \, |z| \, \| = \| \, |z| \, \|^2,$$

which is (2.7). ■

An example of a norm on C^2 that is not absolute but nevertheless satisfies the condition (2.8), and hence (2.7) as well, is $\|x\| \equiv \max\{|x_1 + x_2|, |x_1|, |x_2|\}$.

Although we have characterized the norms for which the inequality (2.7) holds in terms of the natural condition (2.8), it is not always easy to determine whether a particular norm has this property. For example, it is not known which unitarily invariant norms on $M_{m,n}$ satisfy (2.8).

3 Inequalities for Matrices

We are now ready to prove (1.1), as well as an analogous inequality for the Hadamard product of matrices, and to discuss the cases of equality.

Theorem 3.1 *Let $\| \cdot \|$ be a unitarily invariant norm on M_n. Then*

$$\|A^*B\|^2 \leq \|A^*A\| \, \|B^*B\| \quad \textit{for all } A, B \in M_{m,n} \tag{3.1}$$

and

$$\|A \circ B\|^2 \leq \|A^*A\| \, \|B^*B\| \quad \textit{for all } A, B \in M_{m,n}. \tag{3.2}$$

The inequality (3.2) has also been obtained by Okubo [14, Theorem 4.3], while (3.1) can be derived by using an argument similar to that used by Bhatia to prove Proposition 5 (another Cauchy-Schwarz type inequality) in [4]. Both of these inequalities can also be derived as corollaries of Theorem 2.3 in [10].

Proof: Let g be the symmetric gauge function associated with the unitarily invariant norm $\| \cdot \|$. A theorem of A. Horn [6, Theorem 3] gives the weak majorization relation

$$\sum_{i=1}^{k} \sigma_i(A^*B) \leq \sum_{i=1}^{k} \sigma_i(A)\sigma_i(B) \quad k = 1, \ldots, n \tag{3.3}$$

between the singular values of the product A^*B and those of A and B. Compute

$$\begin{aligned}
\|A^*B\|^2 &= g^2(\sigma_1(A^*B), \ldots, \sigma_n(A^*B)) \\
&\leq g^2(\sigma_1(A)\sigma_1(B), \ldots, \sigma_n(A)\sigma_n(B)) \\
&\leq g(\sigma_1^2(A), \ldots, \sigma_n^2(A))\; g(\sigma_1^2(B), \ldots, \sigma_n^2(B)) \\
&= g(\sigma_1(A^*A), \ldots, \sigma_n(A^*A))\; g(\sigma_1(B^*B), \ldots, \sigma_n(B^*B)) \\
&= \|A^*A\|\; \|B^*B\|.
\end{aligned}$$

The first inequality comes from combining (3.3) and (1.2), the second is an application of Theorem 2.3 to the monotone norm $g(\cdot)$, and the penultimate equality is because $\sigma_i^2(A) = \sigma_i(A^*A)$.

To prove (3.2) we use the weak majorization relation [8, Lemma 1]

$$\sum_{i=1}^{k} \sigma_i(A \circ B) \leq \sum_{i=1}^{k} \sigma_i(A)\sigma_i(B) \qquad k = 1, \ldots, n \tag{3.4}$$

for the Hadamard product and apply exactly the same argument. ■

The inequality (3.2) is a generalization to all unitarily invariant norms of a classical inequality of Schur for the spectral norm [17, Satz III, pg 8]: if $\|\cdot\|$ is chosen to be the spectral norm $\|X\|_2 \equiv \sigma_1(X)$, then (3.2) is Schur's inequality $\sigma_1(A \circ B) \leq \sigma_1(A)\sigma_1(B)$.

Theorem 3.1 allows us to make the following generalization of Theorem 2.3 in [20].

Corollary 3.2 *Let $\|\cdot\|$ be a given unitarily invariant norm on M_n. Then for all $A \in M_n$,*

$$\|A\| = \min\{\|B^*B\|^{1/2}\|C^*C\|^{1/2} : B, C \in M_n \text{ and } B^*C = A\}. \tag{3.5}$$

Proof: For any $A \in M_n$, Theorem 3.1 gives

$$\|A\| \leq \inf\{\|B^*B\|^{1/2}\|C^*C\|^{1/2} : B, C \in M_n \text{ and } B^*C = A\}. \tag{3.6}$$

That the infimum is attained and is equal to $\|A\|$ follows by setting $B = P^{1/2}$ and $C = P^{1/2}U$, where $A = PU$ is a polar decomposition of A, i.e., $P, U \in M_n, P \succeq 0$, and U is unitary. ■

In Example 4.4 we give a non-unitarily invariant norm that satisfies 3.5. It is possible to characterize the unitary similarity invariant norms that satisfy (1.1); to do so, we require the following analog of Lemma 2.4.

Lemma 3.3 *Let $\|\cdot\|$ be a unitary similarity invariant norm on M_n such that*

$$\|A\| \leq \|\,|A|\,\| \quad \text{for all } A \in M_n, \tag{3.7}$$

where $|A| \equiv (A^*A)^{1/2}$. Then $N(A) \equiv \|\,|A|\,\|$ is a unitarily invariant norm on M_n.

Proof: The unitary invariance, positivity and homogeneity of the function $N(\cdot)$ are clear, so it suffices to prove that $N(\cdot)$ obeys the triangle inequality. We have the weak majorization (see [5], or [13, p. 243], or [9, Chapter3]):

$$\sum_{i=1}^{k} \sigma_i(A+B) \le \sum_{i=1}^{k} \sigma_i(A) + \sum_{i=1}^{k} \sigma_i(B) \quad k = 1, \ldots, n,$$

which expresses the subadditivity of the Ky Fan k-norms, i.e., the vector $\sigma(A+B)$ is weakly majorized by the vector $\sigma(A)+\sigma(B)$. Define a norm $\|\cdot\|'$ on C^n by $\|x\|' \equiv \|\mathrm{diag}(x)\|$. Condition (3.7) guarantees that $\|x\|' \le \|\,|x|\,\|'$ for all $x \in C^n$, so Lemma 2.4 ensures that the function $\|x\|'' \equiv \|\,|x|\,\|'$ is a monotone norm on C^n. Since the given norm $\|\cdot\|$ is unitary similarity invariant, the norm $\|\cdot\|''$ is permutation invariant. Thus, the norm $\|\cdot\|''$ is actually a symmetric gauge function.

If $C \in M_n$ is a given positive semidefinite matrix, then there is a unitary $U \in M_n$ such that $C = U\Lambda U^*$, where $\Lambda = \mathrm{diag}(\lambda(C))$. Because the norm $\|\cdot\|$ is unitary similarity invariant,

$$\|C\| = \|U\Lambda U^*\| = \|\mathrm{diag}(\lambda(C))\| = \|\lambda(C)\|' = \|\sigma(C)\|''.$$

Now use this identity with $C \equiv |A+B|$, noting that the eigenvalues and singular values of a positive semidefinite matrix are identical, to write

$$\begin{aligned} N(A+B) &= \|\,|A+B|\,\| \\ &= \|\sigma(A+B)\|'' \\ &\le \|\sigma(A)+\sigma(B)\|'' \\ &\le \|\sigma(A)\|'' + \|\sigma(B)\|'' \\ &= \|\,|A|\,\| + \|\,|B|\,\| \\ &= N(A)+N(B). \end{aligned}$$

The first inequality uses weak majorization and the fact that $\|\cdot\|''$ is a symmetric gauge function, while the second is just the triangle inequality for $\|\cdot\|''$. ■

The condition (3.7) is sufficient for a unitary similarity invariant norm on M_n to satisfy the conclusion of Lemma 3.3, but it is not necessary. See Theorem 4.9 for a stronger version of Lemma 3.3, which provides a necessary and sufficient condition.

Another way to prove the triangle inequality for $N(A) \equiv \|\,|A|\,\|''$ is to use the matrix-valued triangle inequality [18, Therem 2]. We demonstrate this technique in the proof of Theorem 4.9.

The following characterization is a matrix analog of Theorem 2.5 for the usual matrix product.

Theorem 3.4 *Let $\|\cdot\|$ be a unitary similarity invariant norm on M_n. Then*

$$\|A^*B\|^2 \le \|A^*A\|\,\|B^*B\| \quad \textit{for all } A, B \in M_n \tag{3.8}$$

if and only if

$$\|A\| \le \|\,|A|\,\| \quad \textit{for all } A \in M_n, \tag{3.9}$$

where $|A| \equiv (A^*A)^{1/2}$. *If either (3.8) or (3.9) holds, then*

$$\|A^*B\|^2 \le \|\,|A^*B|\,\|^2 \le \|A^*A\|\,\|B^*B\| \quad \text{for all } A, B \in M_n. \tag{3.10}$$

Proof: Let $\|\cdot\|$ be a given unitary similarity invariant norm on M_n. If condition (3.9) holds, then the function $\|\,|\cdot|\,\|$ is a unitarily invariant norm on M_n by Lemma 3.3. It now follows from (3.1) that for any $A, B \in M_n$

$$\|A^*B\|^2 \le \|\,|A^*B|\,\|^2 \le \|\,|A^*A|\,\|\,\|\,|B^*B|\,\| = \|A^*A\|\,\|B^*B\|.$$

Thus (3.9) implies both (3.8) and (3.10).

Conversely, suppose that (3.8) holds and let $A \in M_n$ be given. Let $A = UP$ be a polar decomposition of A. Using the condition (3.8) and the hypothesis of unitary similarity invariance, we obtain the desired inequality:

$$\begin{aligned}
\|A\|^2 &= \|(P^{1/2}U^*)^*P^{1/2}\|^2 \\
&\le \|(P^{1/2}U^*)^*(P^{1/2}U^*)\|\,\|(P^{1/2})^*(P^{1/2})\| \\
&= \|UPU^*\|\,\|P\| \\
&= \|P\|\,\|P\| = \|\,|A|\,\|^2.
\end{aligned}$$

■

The hypothesis that the norm $\|\cdot\|$ be unitary similarity invariant is essential, as we show in Example 4.2. In Example 4.13 we exhibit a unitary similarity invariant norm that does not satisfy the condition (3.9).

We now determine the case of equality in (3.1), and to do so we require two preliminary results that are analogs of Lemmata 2.1 and 2.2. There is an analogy between absolute norms on C^n and unitarily invariant norms on M_n. A unitarily invariant norm $\|\cdot\|$ on M_n is a function only of the singular values, and hence $\|A\| = \|\,|A|\,\|$ for all $A \in M_n$ because A and $|A|$ have the same singular values.

Lemma 3.5 *A norm on* $M_{m,n}$ *is unitarily invariant if and only if its dual norm is unitarily invariant.*

Proof: Let $\|\cdot\|$ be a given unitarily invariant norm and let $A \in M_{m,n}, U \in M_m$, and $V \in M_n$ be given with U and V unitary. Then

$$\begin{aligned}
\|UAV\|^D &= \max\{|\text{tr } B^*UAV| : B \in M_{m,n},\ \|B\| \le 1\} \\
&= \max\{|\text{tr } (U^*BV^*)^*A| : B \in M_{m,n},\ \|B\| \le 1\} \\
&= \max\{|\text{tr } C^*A| : C \in M_{m,n},\ \|C\| \le 1\} \\
&= \|A\|^D,
\end{aligned}$$

which shows that $\|\cdot\|^D$ is unitarily invariant. The hypothesis that $\|\cdot\|$ is unitarily invariant is used to obtain the penultimate equality in this series of identities. The converse now follows from the duality theorem for norms. ■

Before proceeding it is convenient to isolate some simple but useful facts about the Frobenius inner product on M_n.

Lemma 3.6 *Let $A, B \in M_n$ be given.*

(a) If A and B are positive semidefinite then $\mathrm{tr}\ AB \geq 0$.

(b) If A and B are Hermitian then $\mathrm{tr}\ AB$ *is real.*

(c) Let A be positive semidefinite and let $B = H + iK$, where $H, K \in H_n$. Then $\mathrm{Re}\ \ \mathrm{tr}\ B^*A = \mathrm{tr}\ HA \leq \mathrm{tr}\ |H|A$.

Proof: If A and B are positive semidefinite then so is $A^{1/2}BA^{1/2}$, and hence $\mathrm{tr}\ \ AB = \mathrm{tr}\ \ A^{1/2}BA^{1/2} \geq 0$, which verifies (a). The assertion in (b) follows from applying (a) to the positive semidefinite matrices $A + \|A\|_2 I$ and $B + \|B\|_2 I$ and noting that the trace of a Hermitian matrix is real. The assertion in (c) that $\mathrm{Re}\ \mathrm{tr}\ \ B^*A = \mathrm{Re}\ (\mathrm{tr}\ HA - i\mathrm{tr}\ KA) = \mathrm{tr}\ HA$ follows from (b), while the inequality $\mathrm{tr}\ HA \leq \mathrm{tr}\ |H|A$ follows from the observation that $(|H| - H)$ is positive semidefinite and hence $\mathrm{tr}\ (|H| - H)A \geq 0$. ■

The following is a matrix analog of Lemma 2.2.

Lemma 3.7 *Let $\|\cdot\|$ be a given unitarily invariant norm on M_n and let $A \in M_n$ be given. Then*

$$\begin{aligned} \|A\| &= \max\{|\mathrm{tr}\ (C^*A)| : C \in M_n \ \textit{and}\ \|C\|^D \leq 1\} \\ &= \max\{\mathrm{tr}\ C|A| : C \in H_n, C \succeq 0,\ \textit{and}\ \|C\|^D \leq 1\}. \end{aligned}$$

Proof: The first identity is the duality theorem for norms. To show the second, compute

$$\begin{aligned} \|A\| &= \|\ |A|\ \| \\ &= \max\{|\mathrm{tr}\ \ C^*|A|\ | : C \in M_n \ \text{ and }\ \|C\|^D \leq 1\} \\ &= \max\{\mathrm{Re}\ \mathrm{tr}\ \ C^*|A| : C \in M_n \ \text{ and }\ \|C\|^D \leq 1\} \\ &= \mathrm{tr}\ \ C_0^*|A| \end{aligned}$$

for some $C_0 \in M_n$ with $\|C_0\|^D \leq 1$. Then $\mathrm{tr}\ C_0^*|A| \leq \mathrm{tr}\ |C_0^*||A|$ by Lemma 3.6 (c), and hence, using Lemma 3.5 for the second inequality,

$$\begin{aligned} \|A\| &\leq \mathrm{tr}\ \ |C_0^*||A| \\ &\leq \max\{\mathrm{tr}\ \ C|A| : C \in H_n, C \succeq 0,\ \text{ and }\ \|C\|^D \leq 1\} \\ &\leq \max\{\mathrm{Re}\ \ \mathrm{tr}\ \ C^*|A| : C \in M_n \ \text{ and }\ \|C\|^D \leq 1\} \\ &= \|\ |A|\ \| = \|A\|. \end{aligned}$$

Thus, all the inequalities must be equalities and the asserted identity follows. ■

We also need a well-known result expressing the monotonicity of a unitarily invariant norm with respect to multiplication by a partial isometry [3, Prop. 7.7.3]. We provide a proof that uses only the existence of the singular value decomposition.

Lemma 3.8 *Let $\|\cdot\|$ be a unitarily invariant norm on $M_{m,n}$, and let $A \in M_{m,n}$, $P_1 \in M_m$, and $P_2 \in M_n$ be given. Then*

$$\|P_1AP_2\| \leq \sigma_1(P_1)\sigma_1(P_2)\|A\|.$$

In particular, if P_1 and P_2 are partial isometries then

$$\|P_1AP_2\| \leq \|A\|.$$

Proof: Let $A \in M_{m,n}$, $P_1 \in M_m$, and $P_2 \in M_n$ and let $\|\cdot\|$ be a unitarily invariant norm on $M_{m,n}$. Assume without loss of generality that $\sigma_1(P_1) = \sigma_1(P_2) = 1$. We will show that

$$\|P_1A\| \leq \|A\|.$$

The inequality

$$\|AP_2\| \leq \|A\|$$

can be proved by a very similar argument. Combining these two results gives the desired conclusion.

Let $P_1 = U\Sigma V$ be a singular value decomposition of P_1, i.e., U, V are unitary and $\Sigma = \text{diag}(\sigma(P_1))$. Define $s \in C^m$ by

$$s_j = \sigma_j(P_1) + i\sqrt{1 - \sigma_j^2(P_1)} \qquad j = 1, \ldots, m$$

and define $S = \text{diag}(s)$. Then S is unitary, since $0 \leq \sigma_j(P_1) \leq 1$ for all $j = 1, \ldots, m$, and $\Sigma = \frac{1}{2}(S + S^*)$. Using the unitary invariance of $\|\cdot\|$, the triangle inequality, and the fact that S and S^* are unitary we compute:

$$\begin{aligned} \|P_1A\| &= \|U\Sigma VA\| = \|\Sigma VA\| = \|\tfrac{1}{2}(S + S^*)VA\| \\ &\leq \tfrac{1}{2}\|SVA\| + \tfrac{1}{2}\|S^*VA\| \\ &= \tfrac{1}{2}\|A\| + \tfrac{1}{2}\|A\| = \|A\|. \end{aligned}$$

■

We are now ready to identify the cases of equality in (3.1), with a result that is an analog of the last part of Theorem 2.3.

Theorem 3.9 *Let $\|\cdot\|$ be a given unitarily invariant norm on M_n, and let $A, B \in M_{m,n}$ be given non-zero matrices. Then*

$$\|A^*B\|^2 = \|A^*A\|\|B^*B\|$$

if and only if there is a positive constant c and there are partial isometries P_1 and P_2 such that

$$AP_1 = cBP_2, \;\; \|P_1^*A^*AP_1\| = \|A^*A\|, \; \textit{and} \; \|P_2^*B^*BP_2\| = \|B^*B\|. \tag{3.11}$$

Proof: Let $\|\cdot\|$ be a unitarily invariant norm on M_n, let $A, B \in M_{m,n}$. The polar decomposition [7, Theorem 7.3.2] guarantees that there is a unitary $U \in M_n$ such that A^*BU is positive semi-definite. Use Lemma 3.7 and the Cauchy-Schwarz inequality for the Frobenius inner product to compute

$$\begin{aligned}
\|A^*B\|^2 &= \|A^*BU\|^2 \\
&= \max\{[\text{tr } CA^*BU]^2 : C \in H_n,\ C \succeq 0,\ \text{and } \|C\|^D \leq 1\} \\
&= \max\{[\text{tr } C^{1/2}A^*BUC^{1/2}]^2 : C \in H_n,\ C \succeq 0,\ \|C\|^D \leq 1\} \\
&\leq \max\{(\text{tr } C^{1/2}A^*AC^{1/2})(\text{tr } C^{1/2}U^*B^*BUC^{1/2}) : \\
&\qquad C \in H_n,\ C \succeq 0,\ \|C\|^D \leq 1\} \qquad (3.12) \\
&\leq \max\{\text{tr } C^{1/2}A^*AC^{1/2} : C \in H_n,\ C \succeq 0,\ \|C\|^D \leq 1\} \\
&\quad \cdot \max\{\text{tr } C^{1/2}U^*B^*BUC^{1/2} : C \in H_n, C \succeq 0, \|C\|^D \leq 1\} \qquad (3.13) \\
&= \max\{\text{tr } CA^*A : C \in H_n,\ C \succeq 0,\ \|C\|^D \leq 1\} \\
&\quad \cdot \max\{\text{tr } CU^*B^*BU : C \in H_n,\ C \succeq 0,\ \|C\|^D \leq 1\} \\
&= \|A^*A\|\ \|U^*B^*BU\| \\
&= \|A^*A\|\ \|B^*B\|.
\end{aligned}$$

Now suppose $\|A^*B\|^2 = \|A^*A\|\|B^*B\|$, so that the preceding inequalities must all be equalities. If inequality (3.12) is an equality, then there must be a positive semidefinite $\tilde{C} \in H_n$ such that $\|\tilde{C}\|^D \leq 1$ and

$$[\text{tr } \tilde{C}^{1/2}A^*BU\tilde{C}^{1/2}]^2 = \text{tr } (\tilde{C}^{1/2}A^*A\tilde{C}^{1/2}) \text{ tr } (\tilde{C}^{1/2}U^*B^*BU\tilde{C}^{1/2}).$$

This last condition states that equality holds in the Cauchy-Schwarz inequality, which can occur only if $A\tilde{C}^{1/2}$ and $BU\tilde{C}^{1/2}$ are dependent, i.e., there is a non-zero scalar α such that

$$A\tilde{C}^{1/2} = \alpha BU\tilde{C}^{1/2}. \qquad (3.14)$$

If inequality (3.13) is also an equality it is necessary that

$$\|A^*A\| = \text{tr } \tilde{C}^{1/2}A^*A\tilde{C}^{1/2} \qquad (3.15)$$

and

$$\|B^*B\| = \text{tr } \tilde{C}^{1/2}U^*B^*BU\tilde{C}^{1/2}. \qquad (3.16)$$

Let $E \in M_n$ be the Hermitian projection onto the range of $\tilde{C}^{1/2}$, so E is a partial isometry, $E = E^*$, and $E\tilde{C}^{1/2} = \tilde{C}^{1/2}E = \tilde{C}^{1/2}$. We now show that $c \equiv |\alpha|$ is the required positive constant and $P_1 \equiv E$, $P_2 \equiv (\alpha/|\alpha|)UE$ are the required partial isometries. By the definition of E and (3.14) we have

$$AE\tilde{C}^{1/2} = A\tilde{C}^{1/2} = \alpha BU\tilde{C}^{1/2} = \alpha BUE\tilde{C}^{1/2}$$

and hence $AE = \alpha BUE$, which is the same as $AP_1 = cBP_2$. Now use Lemma 3.8, Lemma 3.7, and (3.15) to compute

$$\begin{aligned}
\|A^*A\| &\geq \|P_1^*A^*AP_1\| \\
&= \|EA^*AE\|
\end{aligned}$$

$$
\begin{aligned}
&= \max\{\mathrm{tr}\ \ C(EA^*AE) : C \in H_n, C \succeq 0, \text{ and } \|C\|^D \leq 1\} \\
&= \max\{\mathrm{tr}\ \ C^{1/2}EA^*AEC^{1/2} : C \in H_n, C \succeq 0, \text{ and } \|C\|^D \leq 1\} \\
&\geq \mathrm{tr}\ \tilde{C}^{1/2}EA^*AE\tilde{C}^{1/2} \\
&= \mathrm{tr}\ \tilde{C}^{1/2}A^*A\tilde{C}^{1/2} \\
&= \|A^*A\|.
\end{aligned}
$$

Thus, $\|P_1^*A^*AP_1\| = \|A^*A\|$, as asserted. The same argument shows that $\|P_2^*B^*BP_2\| = \|B^*B\|$.

Conversely, if there is a positive constant c and there are partial isometries P_1 and P_2 such that

$$
AP_1 = cBP_2, \ \ \|P_1^*A^*AP_1\| = \|A^*A\|, \text{ and } \|P_2^*B^*BP_2\| = \|B^*B\|
$$

then by Lemma 3.8 we have

$$
\begin{aligned}
\|A^*B\|^2 &\leq \|A^*A\|\|B^*B\| \\
&= \|P_1^*A^*AP_1\|\|P_2^*B^*BP_2\| \\
&= \|P_1^*A^*cBP_2\|\|(1/c)P_1^*A^*BP_2\| \\
&= \|P_1^*A^*BP_2\|^2 \\
&\leq \|A^*B\|^2.
\end{aligned}
$$

Thus, both inequalities must be equalities. ■

4 Examples, Counterexamples and Corollaries

Although Wimmer's conjecture (1.1) is now settled, there are several interesting points to be noted. The first is that *not every* norm on M_n satisfies (1.1) and there are norms satisfying (1.1) that are *not* unitarily invariant. Moreover, the set of norms satisfying (1.1) is a convex set.

Example 4.1 Consider the l_p norms defined on $M_{m,n}$ by

$$
\begin{aligned}
\|A\|_p &\equiv (\sum_{i,j} |a_{ij}|^p)^{1/p} \qquad 1 \leq p < \infty \\
\|A\|_\infty &\equiv \max\{|a_{ij}| : 1 \leq i \leq m, 1 \leq j \leq n\}.
\end{aligned}
$$

Let $m = n = 2$ and

$$
A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \ \ B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
$$

Then (1.1) does not hold for the l_p norms when $1 \leq p < 2$ since

$$
\|A^*B\|_p^2 = 2^{4/p} \not\leq (2 \cdot 2^{1/p})(2^{1/p}) = \|A^*A\|_p\|B^*B\|_p.
$$

Example 4.2 Let $A = [a_{ij}], B = [b_{ij}] \in M_n$. The l_∞ norm $\|\cdot\|_\infty$ on $M_{m,n}$ is not unitarily invariant, but does satisfy (1.1). Let $A = [a_1 \dots a_n]$ and $B = [b_1 \dots b_n]$ be partitioned according to their columns. Then

$$
\|A^*B\|_\infty^2 = \max |a_i^*b_j|^2 \leq (\max a_i^*a_i)(\max b_j^*b_j) \leq \|A^*A\|_\infty\|B^*B\|_\infty.
$$

Notice that $\|\cdot\|_\infty$ does not satisfy condition (3.7); for example, consider

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad |A| = \sqrt{1/2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

For this choice of A we have $\|A\|_\infty = 1 \not\leq 1/\sqrt{2} = \|\,|A|\,\|_\infty$. However this does not contradict Theorem 3.4 because $\|\cdot\|_\infty$ is not unitary similarity invariant.

Example 4.3 Let $C \in M_n$ be given. The *C-numerical radius* is defined on M_n by

$$r_C(A) \equiv \max\{|\text{tr } \; CU^*AU| : U \in M_n \text{ is unitary}\}.$$

If C is not a scalar matrix and tr $C \neq 0$ then it is known [12] that $r_C(\cdot)$ is a norm on M_n. The function $r_C(\cdot)$ is unitary similarity invariant but is never unitarily invariant when $n > 1$ and $C \neq 0$ because, under these conditions, one can always construct $A \in M_n$ such that $r_C(A) \neq r_C(|A|)$. The classical numerical radius,

$$r(A) \equiv \max\{|x^*Ax| : x \in C^n \text{ and } \|x\|_2 = 1\},$$

is an example of a norm of the form $r_C(\cdot)$; it corresponds to the positive semidefinite matrix $C = [1] \oplus 0_{n-1}$. Note that

$$\begin{aligned} r_C(A) &= \max\{|\text{tr } \; CU^*AU)| : U \in M_n \text{ is unitary}\} \\ &\leq \max\{\sum_{i=1}^{n} \sigma_i(CU^*AU) : U \in M_n \text{ is unitary}\} \\ &\leq \max\{\sum_{i=1}^{n} \sigma_i(C)\sigma_i(U^*AU) : U \in M_n \text{ is unitary}\} \\ &= \sum_{i=1}^{n} \sigma_i(C)\sigma_i(A). \end{aligned} \tag{4.1}$$

The first inequality is an application of the inequality [13, Theorem 20.b.1]

$$|\text{tr } \; X| \leq \sigma_1(X) + \cdots + \sigma_n(X) \quad \text{for any } X \in M_n,$$

while the second is a special case of (3.3).

Suppose A and C are both positive semidefinite. Then there are unitary matrices $U_1, U_2 \in M_n$ such that

$$A = U_1 \,\text{diag}(\sigma(A))\, U_1^* \;\text{ and }\; C = U_2 \,\text{diag}(\sigma(C))\, U_2^*.$$

The choice $U \equiv U_1U_2^*$ in (4.1) then shows that the preceding inequalities are equalities in this case. Hence for positive semidefinite $C \in M_n$ and any $A \in M_n$ we have

$$r_C(|A|) = \sum_{i=1}^{n} \sigma_i(C)\sigma_i(A) \geq r_C(A).$$

Thus, Theorem 3.4 guarantees that whenever $C \in M_n$ is a non-scalar positive semidefinite matrix the unitary similarity invariant (but not unitarily invariant when $n > 1$) norm $r_C(\cdot)$ on M_n satisfies (1.1). The numerical radius is an example of such a norm.

Example 4.4 The *Hadamard operator norm* $\|\cdot\|_H$ on $M_{m,n}$ is

$$\|A\|_H \equiv \max\{\sigma_1(A \circ B) : B \in M_{m,n} \text{ and } \sigma_1(B) = 1\}.$$

Although $\|\cdot\|_H$ is not unitarily invariant, it satisfies not only (1.1) [1, Section 5] but also a rectangular version of (3.5) [15, Section 7.7] : for any $A \in M_{m,n}$

$$\|A\|_H = \min\{(\|B^*B\|_H \|C^*C\|_H)^{1/2} : B, C \in M_{m,n},\ B^*C = A\}. \tag{4.2}$$

If $A \succeq 0$ then it is known that $\|A\|_H = \max\{a_{ii} : i = 1, \ldots, n\}$. For general $A \in M_{m,n}$, however there is no known explicit formula to calculate $\|A\|_H$, so (4.2) may provide a useful bound, or may provide the basis for a practical algorithm.

Theorem 4.5 *Let $N_1(\cdot)$ and $N_2(\cdot)$ be given norms on M_n that satisfy the inequality* (1.1) *and let $\alpha \in [0,1]$ be given. Then $N(\cdot) \equiv \alpha N_1(\cdot) + (1-\alpha) N_2(\cdot)$ also satisfies* (1.1), *so the set of norms satisfying* (1.1) *is a convex set that does not include all norms and is strictly larger than the set of unitarily invariant norms.*

Proof: Since any convex combination of norms is a norm, we need only show that $N(\cdot)$ satisfies (1.1). Compute

$$\begin{aligned}
N(A^*B)^2 &= [\alpha N_1(A^*B) + (1-\alpha) N_2(A^*B)]^2 \\
&= \alpha^2 N_1^2(A^*B) + 2\alpha(1-\alpha) N_1(A^*B) N_2(A^*B) \\
&\quad + (1-\alpha)^2 N_2^2(A^*B) \\
&\le \alpha^2 N_1(A^*A) N_1(B^*B) \\
&\quad + 2\alpha(1-\alpha)([N_1(A^*A)N_1(B^*B)][N_2(A^*A)N_2(B^*B)])^{1/2} \\
&\quad + (1-\alpha)^2 N_2(A^*A) N_2(B^*B) \\
&= [\alpha N_1(A^*A) + (1-\alpha) N_2(A^*A)]\,[\alpha N_1(B^*B) + (1-\alpha) N_2(B^*B)] \\
&= N(A^*A) N(B^*B).
\end{aligned}$$

■

These ideas suggest a way to generate new norms on M_n. A *pre-norm* is a continuous, homogeneous, positive function on a real or complex vector space; it does not necessarily satisfy the triangle inequality.

Theorem 4.6 *Let $\|\cdot\|$ be a given norm on M_n, and define $N(A) \equiv \|A^*A\|^{1/2}$. Then $N(\cdot)$ is always a pre-norm on M_n. If the norm $\|\cdot\|$ also satisfies the inequality*

$$\|A^*B\|^2 \le \|A^*A\|\,\|B^*B\| \quad \textit{for all } A, B \in M_{m,n} \tag{4.3}$$

then $N(\cdot)$ is a norm on M_n. In particular, $N(A) \equiv \|A^*A\|^{1/2}$ is a unitarily invariant norm on M_n if $\|\cdot\|$ is a unitarily invariant norm, or if $\|\cdot\|$ is a unitary similarity invariant norm such that $\|A\| \le \|\,|A|\,\|$ for all $A \in M_n$, where $|A| \equiv (A^*A)^{1/2}$.

Proof: The function $N(A) \equiv \|A^*A\|^{1/2}$ is clearly positive, homogeneous, and continuous for any norm $\|\cdot\|$, so it is always a pre-norm on M_n. It is a straightforward computation to show that $N(\cdot)$ satisfies the triangle inequality if it satisfies the inequality (4.3). ∎

If one chooses for $\|\cdot\|$ the spectral norm, the trace norm ($\|A\| \equiv \text{tr } |A|$), the numerical radius, the l_∞ norm, or the Hadamard operator norm, then the respective norms $N(A) \equiv (\|A^*A\|)^{1/2}$ are the spectral norm, the Frobenius norm ($N(A) \equiv [\text{tr } A^*A]^{1/2}$), the spectral norm, and $N(A) \equiv$ the largest Euclidean column length in the last two cases.

See Example 4.13 for a unitary similarity invariant norm that does not satisfy the monotonicity condition at the end of Theorem 4.6.

We have the following analog of Theorem 4.6 for vector norms on C^n. Its proof is very similar to that of Theorem 4.6.

Theorem 4.7 *Let $\|\cdot\|$ be a norm on C^n such that $\|z\| \leq \|\,|z|\,\|$ for all $z = [z_i] \in C^n$, where $|z| \equiv [|z_i|]$. Then $\nu(x) \equiv (\|x \circ \bar{x}\|)^{1/2}$ is an absolute norm on C^n.*

We will now consider the converses of some of the results proved so far. First we characterize the norms $\|\cdot\|$ on C^n such that $\|\,|\cdot|\,\|$ is also a norm, and the unitary similarity invariant norms $\|\cdot\|$ on M_n such that $\|\,|\cdot|\,\|$ is a norm on M_n.

Theorem 4.8 *Let $\|\cdot\|$ be a given norm on C^n, and let $|x| \equiv [|x_i|]$ for all $x = [x_i] \in C^n$. Then $\nu(\cdot) \equiv \|\,|\cdot|\,\|$ is a norm if and only if*

$$\|u\| \leq \|v\| \quad \textit{whenever} \quad u, v \in R^n_+ \quad \textit{and} \quad u \leq v. \tag{4.4}$$

This result is Theorem 5 in [2] (see also [7, Theorem 5.5.10]) , where norms that satisfy the condition (4.4) are referred to as *monotone on the positive orthant.*

Notice that (2.8) provides a *sufficient* condition for $\|\,|\cdot|\,\|$ to be a norm on C^n (Lemma 2.4), while the condition (4.4) is both necessary and sufficient. In Example 4.10, we show that the condition (4.4) is strictly weaker than (2.8).

Theorem 4.9 *Let $\|\cdot\|$ be a given unitary similarity invariant norm on M_n, and let $|A| \equiv (A^*A)^{1/2}$ for $A \in M_n$. Then $N(A) \equiv \|\,|A|\,\|$ is always a unitarily invariant function on M_n, and it is a norm on M_n if and only if*

$$\|X\| \leq \|Y\| \quad \textit{whenever} \quad X, Y \in H_n \quad \textit{and} \quad 0 \preceq X \preceq Y. \tag{4.5}$$

Proof: The unitary invariance of $N(\cdot)$ is clear. If $N(\cdot)$ is a norm then it is unitarily invariant and agrees with $\|\cdot\|$ on the positive semidefinite matrices. The inequality (4.5) is true for any unitarily invariant norm because if $0 \preceq X \preceq Y$, then $\sigma_i(X) \leq \sigma_i(Y)$ for $i = 1, \ldots, n$; in particular, the singular values of X are weakly majorized by those of Y. Conversely, let $A, B \in M_n$ be given and assume (4.5). By the matrix-valued triangle inequality [18, Theorem 2], there are unitary $U, V \in M_n$ such that

$$|A + B| \preceq U|A|U^* + V|B|V^*$$

and hence (4.5), the ordinary triangle inequality, and the unitary similarity invariance of $\|\cdot\|$ give

$$\begin{aligned} N(A+B) &= \| \, |A+B| \, \| \\ &\leq \|U|A|U^* + V|B|V^*\| \\ &\leq \|U|A|U^*\| + \|V|B|V^*\| \\ &= \| \, |A| \, \| + \| \, |B| \, \| \\ &= N(A) + N(B). \end{aligned}$$

Since positivity and homogeneity are clear, it follows that $N(\cdot)$ is a norm. ∎

In Example 4.13 we show that there are unitary similarity invariant norms that do not satisfy the monotonicity condition (4.5).

As one might suspect from Theorem 4.8, the converses of Lemma 2.4 and Theorem 4.7 are not true. There are norms on C^n that satisfy the condition (4.4) but not (2.8).

Example 4.10 Consider the function $\|\cdot\|$ defined on C^2 by

$$\|x\| \equiv \max\{|x_1|, |x_2|, |x_1 - x_2|\}.$$

Then $\|\cdot\|$ is easily shown to be a norm, but it does not satisfy the monotonicity condition $\|x\| \leq \| \, |x| \, \|$. Consider, for example, $x = [1, -1]^T$. However, $\nu_1(x) \equiv \|x \circ \bar{x}\|^{1/2}$, and $\nu_2(x) \equiv \| \, |x| \, \|$ are both norms since

$$\nu_1(x) = \nu_2(x) = \|x\|_\infty = \max\{|x_1|, |x_2|\}.$$

Thus, $\|\cdot\|$ is a norm on C^2 that satisfies the condition (4.4) but not (2.8).

Similarly, one might suspect from Theorem 4.9 that the converses of Lemma 3.3 and Theorem 4.6 are also false. There are unitary similarity invariant norms on M_n that satisfy the condition (4.5), but not (3.7). To construct one we first prove

Lemma 4.11 *Let $\|\cdot\|$ be a given norm on the real vector space H_n. Then the function $\|\cdot\|' : M_n \to R_+$ defined by*

$$\|A\|' \equiv \max\{\|\tfrac{1}{2}(\alpha A + \bar{\alpha} A^*)\| : \alpha \in C, \text{ and } |\alpha| = 1\} \tag{4.6}$$

is a self-adjoint norm on M_n that agrees with $\|\cdot\|$ on H_n, i.e., $\|A\|' = \|A^*\|'$ for all $A \in M_n$ and $\|A\|' = \|A\|$ for all $A \in H_n$. If the given norm $\|\cdot\|$ is unitary similarity invariant on H_n, then the norm $\|\cdot\|'$ is also unitary similarity invariant on M_n.

Furthermore, the norm $\|\cdot\|'$ is minimal in the following sense: if $N(\cdot)$ is any self-adjoint norm on M_n that agrees with $\|\cdot\|$ on H_n, then $N(A) \geq \|A\|'$ for all $A \in M_n$.

Proof: The positivity and homogeneity of $\|\cdot\|'$ follow from the positivity and homogeneity of $\|\cdot\|$. For the triangle inequality, take $A, B \in M_n$ and compute

$$\|A+B\|' = \max\{\|\tfrac{1}{2}[\alpha(A+B) + \bar{\alpha}(A+B)^*]\| : \alpha \in C \;\; |\alpha| = 1\}$$

$$
\begin{aligned}
&\leq \max\{\|\tfrac{1}{2}(\alpha A + \bar{\alpha} A^*)\| + \|\tfrac{1}{2}(\alpha B + \bar{\alpha} B^*)\| : \\
&\qquad\qquad \alpha \in C, \text{ and } |\alpha| = 1\} \\
&\leq \max\{\|\tfrac{1}{2}(\alpha A + \bar{\alpha} A^*)\| : \alpha \in C, \ |\alpha| = 1\} \\
&\qquad + \max\{\|\tfrac{1}{2}(\alpha B + \bar{\alpha} B^*)\| : \alpha \in C, \ |\alpha| = 1\} \\
&= \|A\|' + \|B\|'.
\end{aligned}
$$

We now know that $\|\cdot\|'$ is a norm on M_n, and the fact that $\|A\|' = \|A^*\|'$ is immediate from the definition, as is the assertion about unitary similarity invariance. Suppose that $A \in H_n$ and $|\alpha| = 1$. Then

$$
\|\tfrac{1}{2}(\alpha A + \bar{\alpha} A^*)\| = \|\tfrac{1}{2}(\alpha A + \bar{\alpha} A)\| = \|(\text{Re } \alpha) A\| = |\text{Re } \alpha| \, \|A\| \leq \|A\|
$$

with equality for $\alpha = \pm 1$. This shows that $\|A\|' = \|A\|$ whenever $A \in H_n$.

Finally, consider the assertion about the minimality of $\|\cdot\|'$. Let $N(\cdot)$ be a given norm on M_n such that $N(A) = N(A^*)$ for all $A \in M_n$ and $N(A) = \|A\|$ for all $A \in H_n$. Then for any $\alpha \in C$ with $|\alpha| = 1$ we have

$$
N(A) = \tfrac{1}{2}[N(\alpha A) + N((\alpha A)^*) \geq N(\tfrac{1}{2}[\alpha A + \bar{\alpha} A^*]) = \|\tfrac{1}{2}[\alpha A + \bar{\alpha} A^*]\|
$$

and hence

$$
N(A) \geq \max\{\|\tfrac{1}{2}[\alpha A + \bar{\alpha} A^*]\| : \alpha \in C, \text{ and } |\alpha| = 1\} = \|A\|'. \qquad \blacksquare
$$

Example 4.12 We shall exhibit a norm that shows that the implications in Theorem 4.6 and Lemma 3.3 cannot be reversed. Let $\lambda_1(X) \geq \lambda_2(X)$ denote the algebraically ordered eigenvalues of $X \in H_2$. Define the function $\|\cdot\| : H_2 \to R_+$ by

$$
\|X\| = \max\{|\lambda_1(X)|, |\lambda_2(X)|, |\lambda_1(X) - \lambda_2(X)|\}.
$$

Notice the similarity between this function and the norm on C^2 defined in Example 4.10. One can easily verify that the function $\|\cdot\|$ is a norm on the real vector space H_n: Either use the Weyl inequalities [7, Theorem 4.3.1]

$$
\begin{aligned}
\lambda_1(A+B) &\leq \lambda_1(A) + \lambda_1(B) \\
\lambda_2(A+B) &\geq \lambda_2(A) + \lambda_2(B)
\end{aligned}
$$

for the eigenvalues of any $A, B \in H_2$, or note that the function $\|X\|$ is a Schur-convex function of the eigenvalues of X and apply Theorem 3 in [11]. Notice that $\|A\| = \lambda_1(A) = |||A|||_2$ if $A \in H_2$ is positive semidefinite.

Let the norm $\|\cdot\|'$ be derived from $\|\cdot\|$ as defined in (4.6), and use Lemma 4.11 to observe that $\|\cdot\|'$ is a unitary similarity invariant norm on M_2. For $X \in M_2$ set

$$
N(X) \equiv (\|X^*X\|')^{1/2} \ \text{ and } \ \nu(X) \equiv \| \, |X| \, \|', \ \text{ where } \ |X| \equiv (X^*X)^{1/2}.
$$

Then

$$
N(X) = (\|X^*X\|')^{1/2} = (\|X^*X\|)^{1/2} = (|||X^*X|||_2)^{1/2} = |||X|||_2
$$

and

$$\nu(X) = \| \, |X| \, \|' = \| \, |X| \, \| = \| \, |X| \, \|_2 = \|X\|_2.$$

Thus, both $N(X)$ and $\nu(X)$ are unitarily invariant norms on M_2. However, the choice

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

shows that the norm $\|\cdot\|'$ satisfies neither (4.5) nor (3.7) since $\|A^*A\|' = \|B^*B\|' = 1$, $\|A^*B\|' = 2$, $\|B\|' = 2$ and $\| \, |B| \, \|' = 1$.

Example 4.13 There is a unitary similarity invariant norm on M_2 that satisfies neither (3.7) nor (4.5). Let

$$C = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

and consider the unitary similarity invariant norm $r_C(\cdot)$ on M_2. There is no unitary U for which $U|C|U^*$ is a scalar multiple of C. By the Cauchy-Schwarz inequality for the Frobenius inner product and the definition of $r_C(\cdot)$ we have

$$|\text{tr } CU|C|U^* \,|^2 < (\text{tr } C^2)(\text{tr } U|C|^2U^*) = (\text{tr } C^2)^2 \le r_C^2(C)$$

for any unitary U, and hence $r_C(|C|) < r_C(C)$. Thus, the unitary similarity invariant norm $r_C(\cdot)$ does not satisfy (3.7). To see that $r_C(\cdot)$ does not satisfy (4.5) either, set

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and note that $0 \preceq X \preceq Y$, but $r_C(X) = 2 > 1 = r_C(Y)$.

5 Open Questions

In Theorem 3.4 we characterized the unitary similarity invariant norms on M_n for which (1.1) holds, and in Example 4.2 we showed that the norm $\|\cdot\|_\infty$, which is not unitary similarity invariant, satisfies (1.1). Is there a useful characterization of the norms on $M_{m,n}$ that satisfy (1.1) ?

In Corollary 3.2 we have shown that for any unitarily invariant norm $\|\cdot\|$ on M_n and all $A \in M_n$,

$$\|A\| = \min\{\|B^*B\|^{1/2}\|C^*C\|^{1/2} : B, C \in M_n \text{ and } B^*C = A\}. \tag{5.1}$$

If $\|\cdot\|$ is a unitary similarity invariant norm, then the right-hand side of (5.1) is a unitarily invariant function of $A \in M_n$. Thus, a unitary similarity invariant norm satisfies (5.1) if and only if it is unitarily invariant. We showed in Example 4.4 that the non-unitarily invariant norm $\|\cdot\|_H$ also obeys (5.1). How can one characterize the norms that satisfy (5.1)?

Consider the l_p norms $\|\cdot\|_p$ on $M_{m,n}$, defined in Example 4.1. We have shown that the inequality

$$\|A^*B\|_p^2 \le \|A^*A\|_p \, \|B^*B\|_p \quad \text{for all } A, B \in M_{m,n} \tag{5.2}$$

is false for $p \in [1,2)$ (Example 4.1), and true for $p = 2$ (this is the Cauchy-Schwarz inequality for the Frobenius inner product). Does (5.2) hold for $p \in (2,\infty)$? This question is partially answered in [10, Example 4.4].

For $p \geq 1$, the Schatten-p norm on M_n is defined by

$$\|A\|_{Sp} \equiv (\sum_{i=1}^{n} \sigma_i^p(A))^{1/p}.$$

If $p = 2k$ for some integer k, then the Schatten-p norm can also be defined by

$$\|A\|_{Sp} \equiv (\text{tr } (A^*A)^k)^{1/2k}.$$

From this representation it is clear that

$$\|[a_{ij}]\|_{Sp} \leq \|[|a_{ij}|]\|_{Sp} \tag{5.3}$$

whenever p is an even integer. Thus Theorem 2.5 ensures that

$$\|A \circ \bar{B}\|_{Sp}^2 \leq \|A \circ \bar{A}\|_{Sp}\|B \circ \bar{B}\|_{Sp} \text{ for all } A, B \in M_n \tag{5.4}$$

whenever p is an even integer. If $n = 2$, then (5.3) holds for all $p \geq 2$. To see that (5.3) is not true for $1 \leq p < 2$, consider

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

What are the values of p for which (5.3), and hence (5.4) holds? One can show that the answer depends on n. More generally, what are the unitarily invariant norms $\|\cdot\|$ for which

$$\|[a_{ij}]\| \quad \leq \quad \|[|a_{ij}|]\| \text{ for all } [a_{ij}] \in M_{m,n}?$$

Notice that the spectral norm and the Frobenius norm satisfy this inequality.

References

[1] T. Ando, R. A. Horn, and C. R. Johnson. The singular values of a Hadamard product: A basic inequality. *Lin. Multilin. Alg.*, 21:345–65, 1987.

[2] F. L. Bauer, J. Stoer, and C. Witzgall. Absolute and monotonic norms. *Numerische Mathematik*, 3:257–64, 1961.

[3] R. Bhatia. *Perturbation Bounds for Matrix Eigenvalues.* Pitman Research Notes in Mathematics 162. Longman Scientific and Technical, New York, 1987.

[4] R. Bhatia. Perturbation inequalities for the absolute value map in norm ideals of operators. *J. Operator Theory*, 19:129–36, 1988.

[5] K. Fan. Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Czech. J. Math.*, 12:382–400, 1962.

[6] A. Horn. On the singular values of a product of completely continuous operators. *Proc. Nat. Acad. Sci.*, 36:374–5, 1950.

[7] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, New York, 1985.

[8] R. A. Horn and C. R. Johnson. Hadamard and conventional submultiplicativity for unitarily invariant norms on matrices. *Lin. Multilin. Alg.*, 20:91–106, 1987.

[9] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, New York, 1989.

[10] R. A. Horn and R. Mathias. Cauchy-Schwarz inequalities associated with positive semidefinite matrices. Technical Report 514, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, 1989.

[11] C.-K. Li and N.-K. Tsing. Norms that are invariant under unitary similarities and the C-numerical radii. *Lin. Multilin. Alg.*, 24:209–22, 1989.

[12] M. Marcus and M. Sandy. Three elementary proofs of the Goldberg-Strauss theorem on numerical radii. *Lin. Multilin. Alg.*, 11:243–52, 1982.

[13] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications.* Academic Press, London, 1979.

[14] K. Okubo. Hölder-type inequalities for Schur products of matrices. *Lin. Alg. Appl.*, 91:13–28, 1987.

[15] V. I. Paulsen. *Completely Bounded Maps and Dilations.* Pitman Research Notes In Mathematics 146. Longman Scientific and Technical, Harlow, 1986.

[16] R. Schatten. *Norm Ideals of Completely Continuous Operators.* Springer-Verlag, berlin, 1970.

[17] J. Schur. Bermerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. *J. für Reine und Angewandte Mathematik*, 140:1–28, 1911.

[18] R. C. Thompson. Convex and concave functions of singular values of matrix sums. *Pacific J. of Math.*, 66:285–90, 1976.

[19] J. von Neumann. Some matrix inequalities and metrization of matric space. *Tomsk. Univ. Rev.*, 1:286–300, 1937. Also in *John von Neumann Collected Works* (A. H. Taub, ed.) Vol. IV, pp. 205-18, Pergamon Press, Oxford, 1962.

[20] H. Wimmer. Extremal problems for Hölder norms of matrices and realizations of linear systems. *SIAM J. Matrix Anal. Appl.*, 9:314–322, 1988.