A note on unitarily invariant matrix norms

Wenxuan Ding, Chi-Kwong Li*, and Yuqiao Li
Department of Mathematics, College of William and Mary, Williamsburg, VA 23187.
Email: wding@email.wm.edu, ckli@math.wm.edu, yli53@email.wm.edu

Abstract

Given a unitarily invariant matrix norm $\|\cdot\|$ on $n \times n$ (real or complex) matrices, characterization is given for matrices A and B such that $\|AB\| = \|A\|\|B\|$.

AMS Classification. 15A60.

Key words. Unitarily invariant matrix norm.

1 Introduction

Let M_n be the set of $n \times n$ matrices over $\mathbb{F}$, where $\mathbb{F}$ is the real or complex field. Denote by A^* the conjugate transpose of $A \in M_n$, where A^* reduces to the transpose if $\mathbb{F}$ is real. The Frobenius norm on M_n is defined by $\|A\|_F = (\mathrm{tr}\, A^*A)^{1/2}$. It is known and not hard to prove that $\|\cdot\|_F$ is a matrix norm, i.e., $\|AB\|_F \le \|A\|_F\|B\|_F$ for any $A, B \in M_n$. In [5, Problem I.11 (9)], the author raised the question of studying $A, B \in M_n$ such that $\|AB\|_F = \|A\|_F\|B\|_F$. In this case, the equality holds if and only if $A = xy^*$ and $B = yz^*$ for some vectors $x, y, z \in \mathbb{F}^n$. In fact, the same conclusion holds for the Schatten p-norm, $p \ge 1$, defined by

$$S_p(A) = (\mathrm{tr}\, |A|^p)^{1/p} = (\sum_{j=1}^n s_j(A)^p)^{1/p},$$

where $|A|$ is the positive semidefinite matrix satisfying $|A|^2 = A^*A$ and $s_1(A) \ge \cdots \ge s_n(A)$ are the eigenvalues of $|A|$, known as the singular values of A; see [4]. In other words, for $A, B \in M_n$, $S_p(AB) = S_p(A)S_p(B)$ if and only if $A = xy^*$ and $B = yz^*$ for some vectors $x, y, z \in \mathbb{F}^n$. These results will follow from our general theorem (Theorem 2.4) in the next section.

Note that $S_2(A)$ reduces to $\|A\|_F$. One may consider $p \to \infty$, and define $S_\infty(A) = s_1(A)$, the largest singular value of A. In such a case, if $A = xy^*$ and $B = yz^*$ for some vectors $x, y, z \in \mathbb{F}^n$, then $S_\infty(AB) = S_\infty(A)S_\infty(B)$. But the converse is not true. For example, $1 = S_\infty(AB) = S_\infty(A)S_\infty(B)$ for any $A, B \in U_n$, where U_n is the set of unitary matrices in M_n.

A norm $\|\cdot\|$ on M_n is unitarily invariant if $\|UAV\| = \|A\|$ for any $U, V \in U_n$, and $A \in M_n$; see [2, 3] for some general background. In this note, we characterize matrices $A, B \in M_n$ such that $\|AB\| = \|A\|\|B\|$ for a unitarily invariant matrix norm.

*Research supported by the Simons Foundation Grant 351047.

2 Results and Proofs

By [2, Theorem 2.1], we have the following.

Lemma 2.1 *A unitarily invariant norm on M_n is a matrix norm if and only if any one of the following conditions holds.*

(a) $\|uv^*\| \ge 1$ *for some (or for all) unit vectors* $u, v \in \mathbb{F}^n$. (b) $\|A\| \ge s_1(A)$ *for all* $A \in M_n$.

We also need the following known facts. We give short proofs of them for easy reference.

Lemma 2.2 *Let $\|\cdot\|$ be a unitarily invariant matrix norm on M_n.*

(a) *If $C \in M_n$ satisfies $s_1(C) \le 1$, then C is a convex combination of matrices in U_n.*

(b) *If $A, B \in M_n$, then $\|AB\| \le s_1(A)\|B\|$ and $\|AB\| \le s_1(B)\|A\|$.*

Proof. (a) By the singular value decomposition theorem, there are $U, V \in U_n$ such that $UCV = \mathrm{diag}\,(c_1, \dots, c_n)$ with $1 \ge c_1 \ge \cdots \ge c_n \ge 0$. We will show by induction on n that UCV is a convex combination of $C_1, \dots, C_k \in U_n$. Then C will be a convex combination of $U^*C_1V^*, \dots, U^*C_kV^* \in U_n$.

Let $t = c_1 \in [0,1]$ so that $(1-t)/2, (1+t)/2 \in [0,1]$. If $n = 1$, then $C = \frac{1+t}{2}C_1 + \frac{1-t}{2}C_2$ with $C_1 = [1], C_2 = [-1] \in U_1$. The result holds.

If $n > 1$, $UCV = \frac{1+t}{2}([1] \oplus R) + \frac{1-t}{2}([-1] \oplus R)$ with $R = \mathrm{diag}\,(c_2, \dots, c_n)$. By induction assumption, R is a convex combination of $T_1, \dots, T_k \in U_{n-1}$, say, $R = q_1T_1 + \cdots + q_kT_k$. Then

$$UCV = \sum_{j=1}^{k} \frac{1+t}{2} q_j([1] \oplus T_j) + \sum_{j=1}^{k} \frac{1-t}{2} q_j([-1] \oplus T_j)$$

is a convex combination of matrices in U_n. The result follows.

(b) Suppose $A, B \in M_n$. Assume A and B are nonzero to avoid trivial considerations. Then by (a), there are $A_1, \dots, A_k \in U_n$ and $p_1, \dots, p_k > 0$ with $\sum_{j=1}^k p_j = 1$ such that $A/s_1(A) = \sum_{j=1}^k p_jA_j$. Thus,

$$\|AB/s_1(A)\| = \|(\sum_{j=1}^{k} p_jA_j)B\| \le \sum_{j=1}^{k} p_j\|A_jB\| = \sum_{j=1}^{k} \|B\| = \|B\|.$$

Hence, $\|AB\| \le s_1(A)\|B\|$. Similarly, we can prove that $\|AB\| \le s_1(B)\|A\|$. □

By Lemmas 2.1 and 2.2, we can deduce the following.

Theorem 2.3 *Suppose $\|\cdot\|$ is a unitarily invariant matrix norm on M_n. There exist nonzero $A, B \in M_n$ such that $\|AB\| = \|A\|\|B\|$ if and only if $\|uv^*\| = 1$ for some (or for all) unit vectors $u, v \in \mathbb{F}^n$.*

Proof. By condition (a) of Lemma 2.1, we have $\|uv^*\| \ge 1$ for some (or for all) unit vectors $u, v \in \mathbb{F}^n$. Suppose there are $A, B \in M_n$ such that $\|AB\| = \|A\|\|B\|$. Then by Lemma 2.1 (b) and Lemma 2.2 (b),

$$\|A\|\|B\| = \|AB\| \le s_1(A)\|B\| \le \|A\|\|B\|.$$

Thus,

$$\|A\| = s_1(A) = \|s_1(A)u_1v_1^*\| \le \|A\|$$

if A has singular value decomposition $A = \sum_{j=1}^n s_j(A)u_jv_j^*$, where $\{u_1, \dots, u_n\}, \{v_1, \dots, v_n\} \subseteq \mathbb{F}^n$ are orthonormal sets. Thus, $1 = \|u_1v_1^*\| = \|uv^*\|$ for any unit vectors u, v, as $u_1v_1^*$ and uv^* have the same singular values $1, 0, \dots, 0$.

Conversely, if $\|uv^*\| = 1$ for unit vectors $u, v \in \mathbb{F}^n$, then $\|xy^*\| = 1$ for any unit vectors x, y as uv^* and xy^* have the same singular values. Moreover, for any unit vector $u \in \mathbb{F}^n$ and $A = B = uu^*$ we have $\|AB\| = \|A\|\|B\|$. □

By Theorem 2.3, to study $\|AB\| = \|A\|\|B\|$ for unitarily invariant matrix norms $\|\cdot\|$, we may focus on those $\|\cdot\|$ such that $\|uv^*\| = 1$ for unit vectors u, v. From the discussion in Section 1, we need to understand the difference between the S_∞ norm and S_p norms for other $p \in [1, \infty)$. It turns out that the major difference is the following. For $p \in [1, \infty)$, a matrix $A \in M_n$ satisfies $S_p(A) = s_1(A)$ if and only if A has rank one; for $p = \infty$, there are many matrices with rank larger than 1 satisfying $S_p(A) = s_1(A)$. Based on this insight and the previous lemmas, we have the following.

Theorem 2.4 *Suppose $\|\cdot\|$ is a unitarily invariant norm on M_n such that $\mathcal{R} \subseteq \mathcal{S}$, where*

$$\mathcal{R} = \{uv^* : u, v \in \mathbb{F}^n \text{ are unit vectors}\} \quad \text{and} \quad \mathcal{S} = \{X \in M_n : \|X\| = s_1(X) = 1\}.$$

Let $A, B \in M_n$ be nonzero matrices. Then $\|AB\| = \|A\|\|B\|$ if and only if $s_1(AB) = \|A\|\|B\|$, equivalently,

$$A/s_1(A), B/s_1(B) \in \mathcal{S} \quad \text{and} \quad s_1(AB) = s_1(A)s_1(B).$$

In case $\mathcal{R} = \mathcal{S}$, $\|AB\| = \|A\|\|B\|$ if and only if $A = xy^*$ and $B = yz^*$ for some vectors $x, y, z \in \mathbb{F}^n$.

Several remarks are in order before we present the proof of Theorem 2.4. By Theorem 2.3, the assumption $\mathcal{R} \subseteq \mathcal{S}$ ensures the existence of $X, Y \in M_n$ satisfying $\|XY\| = \|X\|\|Y\|$. In general, for any $X, Y \in M_n$ we have

$$s_1(XY) \le \|XY\| \le \|X\|\|Y\|.$$

The theorem asserts that the second inequality becomes an equality if and only if both inequalities become equalities. Moreover, all matrices $X, Y \in M_n$ satisfying $\|XY\| = \|X\|\|Y\|$ are nonnegative multiples of matrices in $\mathcal{S}$ such that $s_1(XY) = s_1(X)s_1(Y)$. Evidently, this equality holds if and only if there is a unit vector $x \in \mathbb{F}^n$ such that $\|Yx\| = s_1(Y)$ such that $\|X(Yx)\| = s_1(X)s_1(Y)$. Let $N(C)$ be the null space of $C \in M_n$, i.e., $N(C) = \{x \in \mathbb{F}^n : Cx = 0\}$. Then $s_1(XY) = s_1(X)s_1(Y)$ if and only if $N(X^*X - s_1(X)^2I) \cap \{Yz : z \in N(Y^*Y - s_1(Y)^2I)\} \ne \emptyset$.

Proof of Theorem 2.4. Suppose A, B satisfy $\|AB\| = \|A\|\|B\|$. By Lemma 2.2 (b),

$$\|A\|\|B\| = \|AB\| \le s_1(A)\|B\| \le \|A\|\|B\| \quad \text{and} \quad \|A\|\|B\| = \|AB\| \le \|A\|s_1(B) \le \|A\|\|B\|.$$

Thus, $\|A\| = s_1(A)$ and $\|B\| = s_1(B)$. Consequently,

$$1 \le s_1(AB)/\|AB\| = s_1(AB)/(\|A\|\|B\|) \le s_1(A/\|A\|)s_1(B/\|B\|) = 1.$$

It follows that $s_1(AB) = \|AB\| = s_1(A)s_1(B) = \|A\|\|B\|$.

If $s_1(AB) = \|A\|\|B\|$, then $s_1(AB) \le \|AB\| \le \|A\|\|B\| = s_1(AB)$, and hence $\|AB\| = \|A\|\|B\|$. Furthermore, $s_1(AB) = \|A\|\|B\|$ if and only if

$$s_1(AB) \le s_1(A)s_1(B) \le \|A\|\|B\| = s_1(AB).$$

so that $s_1(AB) = s_1(A)s_1(B)$, $s_1(A) = \|A\|$ and $s_1(B) = \|B\|$.

For the last assertion, if $A = xy^*$ and $B = yz^*$ for some nonzero vectors $x, y, z \in \mathbb{F}^n$, then $A/\|A\|, B/\|B\| \in \mathcal{R} = \mathcal{S}$. Since $\|xy^*\| = s_1(xy^*) = \|x\|\|y\|$ and $\|yz^*\| = s_1(yz^*) = \|y\|\|z\|$,

$$s_1(AB) = s_1(xy^*yz^*) = \|x\|\|y\|^2\|z\| = s_1(A)s_1(B).$$

Conversely, if $A, B \in M_n$ satisfy $\|AB\| = \|A\|\|B\|$. Then $A/\|A\|, B/\|B\| \in \mathcal{S}$ so that $A = xy^*$ and $B = uv^*$ for some nonzero vectors $x, y, u, v \in \mathbb{F}^n$. Furthermore,

$$|y^*u|\|x\|\|v\| = s_1(AB) = s_1(A)s_1(B) = \|x\|\|y\|\|u\|\|v\|.$$

So, $|y^*u| = \|y\|\|u\|$. Hence, $u = \xi y$ for some $\xi \in \mathbb{F}$, and $B = yz^*$ with $z = \bar{\xi}v$. □

We give some examples to illustrate Theorem 2.4.

Example 2.5 An important class of unitarily invariant norm on M_n is the Ky-Fan k-norm, $1 \le k \le n$, defined by $F_k(A) = \sum_{j=1}^k s_j(A)$; see [1] and also [2] for background. When $k = 1, n$, we get $S_\infty(A)$ and $S_1(A)$. One can check that for $1 < k \le n$, $\mathcal{R} = \mathcal{S}$. Hence, for $1 < k \le n$, two matrices $A, B \in M_n$ satisfy $F_k(AB) = F_k(A)F_k(B)$ if and only if $A = xy^*$ and $B = yz^*$ for some vectors $x, y, z \in \mathbb{F}^n$.

One can consider the dual Ky-Fan k-norm defined by $F_k^D(A) = \max\{s_1(A), \sum_{j=1}^n s_j(A)/k\}$ on M_n with $n > 1$. Define $\mathcal{R}$ and $\mathcal{S}$ as in Theorem 2.4. If $k > 1$, then $\mathcal{R} \subsetneq \mathcal{S}$. In such a case, if $A = \text{diag}\,(1, 1, 0, \ldots, 0) \in \mathcal{S} \setminus \mathcal{R}$, then $F_k^D(A^2) = F_k^D(A)F_k^D(A) = 1$.

Finally, we remark that our results and proofs can be extended to norm ideals of completely continuous operators equipped with unitarily invariant norms; see [4].

References

[1] K. Fan, On a theorem of Weyl concerning eigenvalues of linear transformations. Proceedings of the National Academy of Sciences of U.S.A. 35 (1949), 652-655.

[2] C.K. Li and N.K. Tsing, On unitarily invariant norms and related results, Linear and Multilinear Algebra 20 (1987), 107-119.

[3] L. Mirsky, Symmetric gauge functions and unitarily invariant norms, Quart. J. Math. Oxford (Ser. 2) 11 (1960) 50-59.

[4] R. Schatten, Norm Ideals of Completely Continuous Operators, Springer, Berlin, 1960.

[5] G. Strang, Linear Algebra and Learning from Data, Wellesley-Cambridge, 2019.