Norm Inequalities for Positive Operators

RAJENDRA BHATIA[1] and FUAD KITTANEH[2]
[1]*Indian Statistical Institute, New Delhi, 110016, India*
[2]*Department of Mathematics, University of Jordan, Amman, Jordan*

(Received: 2 June 1997)

Abstract. Let A, B be positive operators on a Hilbert space, z any complex number, m any positive integer, and $|||\cdot|||$ any unitarily invariant norm. We show that $|||A+zB||| \leqslant |||A+|z|B|||$ and $|||A^m - B^m||| \leqslant |||(A+B)^m|||$. Some related inequalities are also obtained.

Mathematics Subject Classifications (1991): 47A30, 47B10, 47B15, 47B20.

Key words: positive operator, unitarily invariant norm, majorisation.

1. Introduction

Inequalities for sums and products of positive operators and for their norms, traces, determinants and eigenvalues are of wide interest in analysis and physics. Several such inequalities may be found in [1, 3, 4, 6] and in the many papers cited therein.

The aim of this Letter is to prove some basic inequalities that augment these results.

Let $\mathcal{B}(\mathcal{H})$ be the space of all bounded linear operators on a complex separable Hilbert space $\mathcal{H}$. Apart from the usual operator norm $\|T\|$ on this space, we consider the *unitarily invariant* or *symmetric* norms $|||T|||$. Each of these is defined on an ideal in $\mathcal{B}(\mathcal{H})$, and it will be implicitly understood that when we talk of $|||T|||$, then the operator T is in this ideal.

Our first theorem is:

THEOREM 1. *Let A, B be positive operators and let z be any complex number. Then*

$$|||A-|z|B||| \leqslant |||A-zB||| \leqslant |||A+|z|B||| \tag{1}$$

for every unitarily invariant norm.

We will use this theorem to prove the following:

THEOREM 2. *Let A, B be positive operators and let m be any positive integer. Then*

$$|||A^m+B^m||| \leqslant |||(A+B)^m|||, \tag{2}$$

for every unitarily invariant norm.

Theorem 2 suggests an obvious question: what happens to inequality (2) when m is replaced by a positive real number? We will answer this for some special cases.

2. Proofs of the Main Results

Let us recall some basic facts that we will use. These may be found in the references cited above.

If T is a compact operator, we denote by $s_1(T) \geqslant s_2(T) \geqslant \cdots$ the eigenvalues of $(T^*T)^{1/2}$. These are called the *singular values* of T. A maximum principle due to Ky Fan says that, for $k = 1, 2, \ldots$, we have

$$\sum_{j=1}^{k} s_j(T) = \max \sum_{j=1}^{k} |\langle x_j, Ty_j \rangle|, \tag{3}$$

where the maximum is taken over all choices of orthonormal vectors $\{x_1, \ldots, x_k\}$ and $\{y_1, \ldots, y_k\}$. If T is Hermitian, we can choose $x_j = y_j$. This is a generalisation of the statement that for any operator T, we have $||T|| = \sup_{||x||=||y||=1} |\langle x, Ty \rangle|$, and if T is Hermitian, we can choose $x = y$ in this supremum.

Another theorem, also due to Ky Fan, and called the *Dominance Principle*, says that if S and T are any two operators, then the inequality $|||S||| \leqslant |||T|||$ is valid for all unitarily invariant norms if and only if

$$\sum_{j=1}^{k} s_j(S) \leqslant \sum_{j=1}^{k} s_j(T), \tag{4}$$

for all $k = 1, 2, \ldots$. In the standard notation for majorisation [1, 4], the family of inequalities (4) is written as

$$\{s_j(S)\} \prec_w \{s_j(T)\}. \tag{5}$$

The symbol $\prec_w$ stands for *weak majorisation.*

Proof of Theorem 1. Let A, B be compact positive operators. By Fan's Maximum Principle, there exist orthonormal vectors $e_1, e_2, \ldots$, and $f_1, f_2, \ldots$, such that for $k = 1, 2, \ldots$, we have

$$\begin{aligned}\sum_{j=1}^{k} s_j(A - zB) &= \sum_{j=1}^{k} |\langle e_j, (A + zB) f_j \rangle| \\ &\leqslant \sum_{j=1}^{k} \{ |\langle e_j, Af_j \rangle| - |z|\, |\langle e_j, Bf_j \rangle| \}.\end{aligned}$$

Using the Cauchy–Schwarz inequality, we get

$$\sum_{j=1}^{k} s_j(A - zB)$$

$$\leqslant \sum_{j=1}^{k} \{[\langle e_j, Ae_j\rangle\langle f_j, Af_j\rangle]^{1/2} + |z|[\langle e_j, Be_j\rangle\langle f_j, Bf_j\rangle]^{1/2}\}.$$

Now using the arithmetic-geometric mean inequality, we get

$$\sum_{j=1}^{k} s_j(A - zB)$$

$$\leqslant \frac{1}{2}\sum_{j=1}^{k}\{\langle e_j, Ae_j\rangle + \langle f_j, Af_j\rangle + |z|\langle e_j, Be_j\rangle + |z|\langle f_j, Bf_j\rangle\}$$

$$= \frac{1}{2}\sum_{j=1}^{k}\{\langle e_j, (A + |z|B)e_j\rangle + \langle f_j, (A + |z|B)f_j\rangle\}.$$

Another application of Fan's Maximum Principle leads to the inequality

$$\sum_{j=1}^{k} s_j(A - zB) \leqslant \sum_{j=1}^{k} s_j(A + |z|B).$$

This is true for all $k = 1, 2, \ldots$. Hence, by the Fan Dominance Principle, $|||A + zB||| \leqslant |||A + |z|B|||$ for all unitarily invariant norms. This proves the second inequality in (1).

The proof of the first inequality is simpler. Since $A - |z|B$ is Hermitian, we can find orthonormal vectors $e_1, e_2, \ldots$ such that for $k = 1, 2, \ldots$, we have

$$\begin{aligned}\sum_{j=1}^{k} s_j(A - |z|B) &= \sum_{j=1}^{k} |\langle e_j, (A - |z|B)e_j\rangle| \\ &= \sum_{j=1}^{k} |\langle e_j, Ae_j\rangle - |z|\langle e_j, Be_j\rangle| \\ &\leqslant \sum_{j=1}^{k} |\langle e_j, (A + zB)e_j\rangle|.\end{aligned}$$

The last inequality above is a consequence of the statement $|x - |z|y| \leqslant |x + zy|$ for all positive numbers x, y and complex numbers z. Now using the two Fan principles again, we obtain the first inequality in (1) from this.

For noncompact operators, only the operator norm needs to be considered. The same proof works, except that instead of a maximum we have a supremum. □

In the proof of Theorem 2 we will need the following result.

PROPOSITION 1. *Let X be any operator and let Y be a positive operator. If $|||X||| \leqslant |||Y|||$ for all unitarily invariant norms, then for all positive integers m, we have $|||X^m||| \leqslant |||Y^m|||$.*

Proof. The inequality $|||X||| \leqslant |||Y|||$ is equivalent to the weak majorisation $\{s_j(X)\} \prec_w \{s_j(Y)\}$. The function $f(t) = t^m$, $m \geqslant 1$, is convex and monotonically increasing on $[0, \infty)$. Hence, it preserves weak majorisation [1, p. 42], [4, p. 116]. In other words, $\{s_j^m(X)\} \prec_w \{s_j^m(Y)\}$. Since Y is positive, $s_j^m(Y) = s_j(Y^m)$. So, to prove the proposition we need to prove that

$$\{s_j(X^m)\} \prec_w \{s_j^m(X)\}. \tag{6}$$

This can be proved by a well-known argument that we indicate briefly for the reader's convenience.

Write the inequality $\|X^m\| \leqslant \|X\|^m$ as $s_1(X^m) \leqslant s_1^m(X)$. Replace X by its kth antisymmetric tensor product $\wedge^k X$. This gives

$$\prod_{j=1}^{k} s_j(X^m) \leqslant \prod_{j=1}^{k} s_j^m(X), \tag{7}$$

for all $k = 1, 2, \ldots$. Taking logarithms converts the products in (7) to sums and then, using the fact that $f(t) = e^t$ is convex and monotonically increasing, one gets (6). See [1, pp. 42, 72] for details. □

Proof of Theorem 2. Let ω be a primitive mth root of unity. Then we have the identity

$$A^m + B^m = \frac{1}{m}\{(A+B)^m + (A+\omega B)^m + \cdots + (A+\omega^{m-1}B)^m\}, \tag{8}$$

for any two operators A, B. Hence,

$$|||A^m + B^m||| \leqslant \frac{1}{m}\{|||(A+B)^m||| + |||(A+\omega B)^m||| + \cdots + |||(A+\omega^{m-1}B)^m|||\}.$$

Now, suppose A and B are positive. Then for each $j = 0, 1, \ldots, m-1$, we have, by Theorem 1,

$$|||A + \omega^j B||| \leqslant |||A - B|||,$$

and, hence, by Proposition 1

$$|||(A + \omega^j B)^m||| \leqslant |||(A + B)^m|||.$$

Combining the last three inequalities, we obtain the inequality (2). □

3. Further Results and Remarks

In this section, we will prove analogues of the inequality (2) when m is replaced by a positive real number r, but the norm is the operator norm or the trace norm.

THEOREM 3. *Let A, B be positive operators, then*

$$||A^r + B^r|| \leqslant ||(A + B)^r|| \quad \textit{for } 1 \leqslant r < \infty, \tag{9}$$

$$||A^r + B^r|| \geqslant ||(A + B)^r|| \quad \textit{for } 0 \leqslant r \leqslant 1. \tag{10}$$

Proof. Note that $||(A + B)^r|| = ||A + B||^r$. Let m be any positive integer and let Ω_m be the set of all real numbers r, $1 \leqslant r \leqslant m$, for which the inequality (9) is true. We will show that Ω_m is a convex set. Since Ω_m contains the points 1 and m, this will prove the inequality (9). In our proof, we will use the fact that $||T|| = ||T^*T||^{1/2} = ||TT^*||^{1/2}$ for all operators T.

Suppose $r, s \in \Omega_m$. Let $t = (t + s)/2$. Note that

$$\begin{bmatrix} A^{r/2} & B^{r/2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A^{s/2} & 0 \\ B^{s/2} & 0 \end{bmatrix} = \begin{bmatrix} A^t + B^t & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence

$$\begin{aligned} ||A^t + B^t|| &\leqslant \left\| \begin{pmatrix} A^{r/2} & B^{r/2} \\ 0 & 0 \end{pmatrix} \right\| \left\| \begin{pmatrix} A^{s/2} & 0 \\ B^{s/2} & 0 \end{pmatrix} \right\| \\ &= ||A^r + B^r||^{1/2} ||A^s + B^s||^{1/2} \\ &\leqslant ||A + B||^{r/2} ||A + B||^{s/2} \\ &= ||A + B||^t \\ &= ||(A + B)^t||. \end{aligned}$$

This shows that $t \in \Omega_m$. Hence, the inequality (9) is established.

Let $0 < r \leqslant 1$. Then from (9) we have $||A^{1/r} + B^{1/r}|| \leqslant ||A + B||^{1/r}$. Replacing A and B by A^r and B^r, we obtain the inequality (10). □

For the trace norm, the analogue of Theorem 3 is known to be true. Lemma 2.6 of McCarthy [5], for example, can be translated to say that for positive operators A, B

$$||A^r + B^r||_1 \leqslant ||(A + B)^r||_1, \quad \text{for } 1 \leqslant r < \infty, \tag{11}$$

$$||A^r + B^r||_1 \geqslant ||(A + B)^r||_1, \quad \text{for } 0 \leqslant r \leqslant 1. \tag{12}$$

A simple proof of these inequalities was given in [2]. From this, we can conclude a little more. For simplicity write $A \oplus B$ for the 2×2 operator matrix $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$. If A and B are positive, then Theorem 1 in [2] says that

$$\{s_j(A \oplus B)\} \prec \{s_j((A + B) \oplus 0)\},$$

where the symbol $\prec$ stands for *majorisation* [1, 4]. So, if f is any convex function on $[0, \infty)$, using Corollary II.3.4 in [1], we get

$$\{f(s_j(A \oplus B))\} \prec_w \{f(s_j((A + B) \oplus 0))\}.$$

If, further, $f(0) = 0$, this gives

$$\{s_j(f(A) \oplus f(B))\} \prec_w \{s_j(f(A + B) \oplus 0)\}.$$

In other words, we have

$$\left|\left|\left| \begin{bmatrix} f(A) & 0 \\ 0 & f(B) \end{bmatrix} \right|\right|\right| \leqslant \left|\left|\left| \begin{bmatrix} f(A + B) & 0 \\ 0 & 0 \end{bmatrix} \right|\right|\right|, \tag{13}$$

for any convex function on $[0, \infty)$ with $f(0) = 0$. If f is a concave function on $[0, \infty)$ with $f(0) = 0$, the inequality (13) is reversed. The inequalities (11) and (12) are obtained from this by choosing $f(t) = t^r$.

It is natural to conjecture that the result of Theorem 3 is valid for all unitarily invariant norms. More generally, we conjecture that if f is any *operator monotone function* on $[0, \infty)$ with $f(0) = 0$, then for all positive operators A, B, and for all unitarily invariant norms, we have

$$|||f(A + B)||| \leqslant |||f(A) + f(B)|||.$$

There are several inequalities in the literature that are akin to this. Some of them can be found in [1] where the reader can also find references to the original papers.

Acknowledgement

The second author thanks the Third World Academy of Sciences and the Indian Statistical Institute for sponsoring his visit to New Delhi, during which a part of this work was done. The authors are grateful to M. D. Choi for illuminating conversations.

References

1. Bhatia, R.: *Matrix Analysis*, Springer, New York, 1997.
2. Bhatia, R. and Holbrook, J. A. R.: On the Clarkson–McCarthy inequalities, *Math. Ann*. **281** (1988), 7–12.
3. Gohberg, I. C. and Krein, M. G.: *Introduction to the Theory of Linear Nonselfadjoint Operators*, Amer. Math. Soc. Transl. Ser. 2, Vol. 18; Amer. Math. Soc., Providence, 1969.
4. Marshall, A. W. and Olkin, I.: *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.
5. McCarthy, C. A.: c_p, *Israel J. Math.* **5** (1967), 249–271.
6. Simon, B.: *Trace Ideals and Their Applications*, Cambridge University Press, 1979.