Pramāṇa, Vol. 4, No. 5, 1975, pp 242–245. © Printed in India.

Hölder's inequality for matrices

C L MEHTA
Department of Physics, Indian Institute of Technology, New Delhi 110029

MS received 2 April 1975

Abstract. We prove that for arbitrary $n \times n$ matrices $A_1, A_2, \cdots, A_m$ and for positive real numbers $p_1, p_2, \cdots, p_m$ with $p_1^{-1} + p_2^{-1} + \cdots + p_m^{-1} = 1$, the inequality

$$|\operatorname{Tr}(A_1 A_2 \cdots A_m)^2| \leqslant \prod_{k=1}^{m} [\operatorname{Tr}(A_k^\dagger A_k)^{p_k}] p_k^{-1}$$

holds.

Keywords. Matrix sums; Hölder's inequality.

The inequality

$$|\sum_{i=1}^{n} u_i^{(1)} u_i^{(2)} \ldots u_i^{(m)}| \leqslant \prod_{k=1}^{m} [\sum_{i=1}^{n} |u_i^{(k)}|^{p_k}]^{p_k^{-1}}, \tag{1}$$

where $u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(m)}$, $(i = 1, 2, \ldots, n)$ are arbitrary complex numbers and $p_1, p_2, \ldots, p_m$ are positive real numbers with

$$p_1^{-1} + p_2^{-1} + \ldots + p_m^{-1} = 1$$

is well known in literature and is called Hölder's inequality (Beckenbach and Bellman 1961).

In this paper we prove an inequality analogous to (1) involving arbitrary matrices. Our result is expressed by the following:

THEOREM: Let $A_1, A_2, \ldots, A_m$ be a set of arbitrary $(n \times n)$ matrices and let $p_1, p_2, \ldots, p_m$ be positive real numbers with $\Sigma_{k=1}^{m} p_k^{-1} = 1$, m being any positive integer. If dagger denotes the Hermitian adjoint, then

$$|\operatorname{Tr}(A_1 A_2, \ldots A_m)^2| \leqslant \prod_{k=1}^{m} [\operatorname{Tr}(A_k^\dagger A_k)^{p_k}]^{p_k^{-1}}. \tag{2}$$

Some other related results are considered by Mehta (1968).

Proof of the theorem is based on two lemmas:

LEMMA 1. For an arbitrary $(n \times n)$ matrix X,

$$|\operatorname{Tr} X^2| \leqslant \operatorname{Tr}(XX^\dagger); \tag{3}$$

LEMMA 2. Let $A_1, A_2, \ldots, A_m$ be a set of arbitrary $(n \times n)$ matrices and for each k, $(k = 1, 2, \ldots, m)$, let the eigenvalues $a_1^{(k)}, a_2^{(k)}, \ldots, a_n^{(k)}$ of the matrix $A_k A_k^\dagger$ (or that of $A_k^\dagger A_k$) be so arranged that

$$a_1^{(k)} \geqslant a_2^{(k)} \geqslant \ldots \geqslant a_n^{(k)}, \tag{4}$$

Then the inequality

$$\sum_{j=1}^{l} \langle x_j \mid A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger \mid x_j \rangle \leqslant \sum_{j=1}^{l} a_j^{(1)} a_j^{(2)} \ldots a_j^{(m)} \tag{5}$$

holds. Here $\mid x_j \rangle$, $j = 1, 2, \ldots, l$, $(l \leqslant n)$ are arbitrary orthonormal vectors in the n-dimensional space.

Lemma 1 is proved by Schur (1909) and is a special case of a more general result due to Weyl (1949).

We now proceed to prove lemma 2. Let $\mid \phi_i \rangle$, $(i = 1, 2, \ldots, n)$, be an orthonormal set of eigenvectors of the matrix $A_m A_m^\dagger$

$$A_m A_m^\dagger \mid \phi_i \rangle = a_i^{(m)} \mid \phi_i \rangle. \tag{6}$$

For each j, we write

$$\begin{aligned}
&\langle x_j \mid A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger \mid x_j \rangle \\
&= \sum_{i=1}^{n} \langle x_j \mid A_1 \ldots A_{m-1} a_i^{(m)} \mid \phi_i \rangle \langle \phi_i \mid A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle \\
&= a_l^{(m)} \sum_{i=1}^{n} \mid \langle x_j \mid A_1 \ldots A_{m-1} \mid \phi_i \rangle \mid^2 \\
&\quad + \sum_{i=1}^{l} (a_i^{(m)} - a_l^{(m)}) \mid \langle x_j \mid A_1 \ldots A_{m-1} \mid \phi_i \rangle \mid^2 \\
&\quad + \sum_{i=l+1}^{n} (a_i^{(m)} - a_l^{(m)}) \mid \langle x_j \mid A_1 \ldots A_{m-1} \mid \phi_i \rangle \mid^2.
\end{aligned} \tag{7}$$

Since for $i \geqslant l$, $a_i^{(m)} - a_l^{(m)} \leqslant 0$, it follows that

$$\begin{aligned}
\langle x_j \mid A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger \mid x_j \rangle &\leqslant a_l^{(m)} \langle x_j \mid A_1 \ldots A_{m-1} A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle \\
&\quad + \sum_{i=1}^{l} (a_i^{(m)} - a_l^{(m)}) \mid \langle x_j \mid A_1 \ldots A_{m-1} \mid \phi_i \rangle \mid^2,
\end{aligned} \tag{8}$$

from which on summing over j, we obtain

$$\begin{aligned}
\sum_{j=1}^{l} \langle x_j \mid A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger \mid x_j \rangle &\leqslant a_l^{(m)} \sum_{j=1}^{l} \langle x_j \mid A_1 \ldots A_{m-1} A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle \\
&\quad + \sum_{i=1}^{l} (a_i^{(m)} - a_l^{(m)}) \sum_{j=1}^{l} \langle \phi_i \mid A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle \langle x_j \mid A_1 \ldots A_{m-1} \mid \phi_i \rangle.
\end{aligned} \tag{9}$$

Further, since $\mid x_j \rangle$ are orthonormal

$$\begin{aligned}
&\sum_{j=1}^{l} \langle \phi_i \mid A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle \langle x_j \mid A_1 \ldots A_{m-1} \mid \phi_i \rangle \\
&\quad \leqslant \langle \phi_i \mid A_{m-1}^\dagger \ldots A_1^\dagger A_1 \ldots A_{m-1} \mid \phi_i \rangle.
\end{aligned} \tag{10}$$

From (9) and (10) it follows that

$$\sum_{j=1}^{l} \langle x_j \mid A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger \mid x_j \rangle \leqslant a_l^{(m)} \sum_{j=1}^{l} \langle x_j \mid A_1 \ldots A_{m-1} A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle$$

$$+ \sum_{j=1}^{l} (a_j^{(m)} - a_l^{(m)}) \langle \phi_j \mid A_{m-1}^\dagger \ldots A_1^\dagger A_1 \ldots A_{m-1} \mid \phi_j \rangle. \tag{11}$$

We now apply the method of induction. Setting

$m = 1$ in (11), we obtain

$$\sum_{j=1}^{l} \langle x_j \mid A_1 A_1^\dagger \mid x_j \rangle \leqslant \sum_{j=1}^{l} a_j^{(1)}, \tag{12}$$

so that the inequality (5) is obviously true for $m = 1$. Inequality (12) is a special case of a result due to Fan (1949). Our proof is similar to that given by him.

Next assume that (5) is true for $m - 1$, *i.e.*, that

$$\sum_{j=1}^{l} \langle x_j \mid A_1 \ldots A_{m-1} A_{m-1}^\dagger \ldots A_1^\dagger \mid x_j \rangle \leqslant \sum_{j=1}^{l} a_j^{(1)} a_j^{(2)} \ldots a_j^{(m-1)}, \quad (l \leqslant n). \tag{13}$$

Substituting $\mid \phi_j \rangle$ for $\mid x_j \rangle$ and using the fact that the eigenvalue spectrum of $A^\dagger A$ is identical to that of $AA^\dagger$ we can rewrite (13) as

$$\sum_{j=1}^{k} \langle \phi_j \mid A_{m-1}^\dagger \ldots A_1^\dagger A_1 \ldots A_{m-1} \mid \phi_j \rangle \leqslant \sum_{j=1}^{k} a_j^{(1)} a_j^{(2)} \ldots a_j^{(m-1)}. \tag{14}$$

Let us multiply (14) by $(a_k^{(m)} - a_{k+1}^{(m)})$ and sum over k from $k = 1$ to $k = l - 1$, $(l \leqslant n)$. We then obtain

$$\sum_{j=1}^{l} (a_j^{m} - a_l^{(m)}) \langle \phi_j \mid A_{m-1}^\dagger \ldots A_1^\dagger A_1 \ldots A_{m-1} \mid \phi_j \rangle$$

$$\leqslant \sum_{j=1}^{l} (a_j^{(m)} - a_l^{(m)}) a_j^{(1)} a_j^{(2)} \ldots a_j^{(m-1)}. \tag{15}$$

From (11), (13) and (15) it follows that the inequality (5) is true for m. This completes the proof of lemma 2.

If we set $l = n$ in (5) and also use Hölder's inequality (1), we obtain

$$\begin{aligned} \operatorname{Tr}(A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger) &\leqslant \sum_{j=1}^{m} a_j^{(1)} a_j^{(2)} \ldots a_j^{(m)} \\ &\leqslant \prod_{k=1}^{m} \Big[\sum_{j=1}^{n} (a_j^{(k)})^{p_k} \Big]^{p_k^{-1}} \\ &= \prod_{k=1}^{m} [\operatorname{Tr}(A_k^\dagger A_k)^{p_k}]^{p_k^{-1}}, \end{aligned} \tag{16}$$

where $p_1, p_2, \ldots, p_m$ are positive numbers with $\sum_{k=1}^{m} p_k^{-1} = 1$.

Also if we set $X = A_1 A_2 \ldots A_m$ in (3) and use (16), we obtain the required result expressed by the theorem, *viz.*, the inequality

$$
\begin{aligned}
|\operatorname{Tr}(A_1 A_2 \ldots A_m)^2| &\leqslant \operatorname{Tr}(A_1 \ldots A_m A_m^\dagger \ldots A_1^\dagger) \\
&\leqslant \prod_{k=1}^{m} [\operatorname{Tr}(A_k^\dagger A_k)^{p_k}]^{p_k^{-1}} .
\end{aligned} \tag{17}
$$

In this paper we have considered the case when $A_1, A_2, \ldots, A_m$ are arbitrary $(n \times n)$ matrices. However, the same results and proofs hold also in the cases when $A_1, A_2, \ldots, A_m$ are arbitrary completely continuous linear operators in a Hilbert space or when they are continuous kernels of linear integral equations.

References

Beckenbach E F and Bellman R 1961 *Inequalities* (Springer-Verlag, Berlin) p. 20

Fan K 1949 *Proc. Natl. Acad. Sci., U.S.A.* **35** 652

Mehta C L 1968 *J. Math. Phys.* **9** 693

Schur I 1909 *Math. Ann.* **66** 488

Weyl H 1949 *Proc. Natl. Acad. Sci., U.S.A.* **35** 408