An Inequality for the trace of matrix products, using absolute values

Bernhard Baumgartner[1]

Fakultät für Physik, Universität Wien
Boltzmanngasse 5, A-1090 Vienna, Austria

July 1, 2011

Abstract

The absolute value of matrices is used to give inequalities for the trace of products. An application gives a very short proof of the tracial matrix Hölder inequality.

Keywords: matrix, absolute value, trace inequality, Hölder inequality

MSC: 39B42, 15A45 , 47A50, PACS: 02.10.Yn

[1] Bernhard.Baumgartner@univie.ac.at

1 The baby inequality and its application

1 THEOREM. ***Using absolute values:*** *Consider two complex $m \times n$ matrices A, B and their absolute values, $|A| = (A^*A)^{1/2}$, $\quad |A^*| = (AA^*)^{1/2}$. Then*

$$|\mathrm{Tr}\, A^*B| \le (\mathrm{Tr}\, |A|\cdot|B|)^{1/2} \cdot (\mathrm{Tr}\, |A^*|\cdot|B^*|)^{1/2} \tag{1}$$

Proof. Let e_i be a basis made of normalized eigenvectors of A^*A, with eigenvalues $a_i^2 \ge 0$. For $a_i \ne 0$, $f_i = a_i^{-1} A e_i$ are normalized eigenvectors of AA^*, obeying $A^* f_i = a_i e_i$. Eventually, to make a full basis, this set has to be completed by introducing eigenvectors of AA^* with eigenvalue 0. Analogously, there are basis-sets of normalized vectors g_j and h_j, such that $Bg_j = b_j h_j$, $B^* h_j = b_j g_j$, with $b_j \ge 0$. With these vectors we get

$$|\mathrm{Tr}\, A^*B| \;=\; \left| \sum_{i,j} a_i b_j \langle g_j, e_i\rangle \langle f_i, h_j\rangle \right|. \tag{2}$$

Applying the Cauchy-Schwarz inequality gives

$$|\mathrm{Tr}\, A^*B| \;\le\; \left(\sum_{i,j} a_i b_j \langle e_i, g_j\rangle \langle g_j, e_i\rangle \right)^{1/2} \left(\sum_{i,j} a_i b_j \langle f_i, h_j\rangle \langle h_j, f_i\rangle \right)^{1/2} \tag{3}$$

$$= \; (\mathrm{Tr}\, |A|\cdot|B|)^{1/2} \, (\mathrm{Tr}\, |A^*|\cdot|B^*|)^{1/2} \tag{4}$$

In the last step the identities $|A| e_i = a_i e_i$, $|A^*| f_i = a_i f_i$, $|B| g_j = b_j g_j$, and $|B^*| h_j = b_j h_j$ have been used. ☐

We remark that the inequality is sharp. In case both matrices A and B have rank one, it becomes an equality. An example is given with 2×2 matrices

$$A \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \tag{5}$$

$$\text{with} \quad |B| \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad |B^*| \;=\; \sqrt{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix};$$

so $\quad \mathrm{Tr}\, A^*B = 1, \quad \mathrm{Tr}\, |A|\cdot|B| = 1/\sqrt{2}, \quad \mathrm{Tr}\, |A^*|\cdot|B^*| = \sqrt{2}.$

This example shows also, that the distinction between the absolute values $|B|$ and $|B^*|$ is necessary. Only if both A and B are normal matrices, equation (1) becomes $|\mathrm{Tr}\, A^*B| \le (\mathrm{Tr}\, |A|\cdot|B|)$.

The following application was at the origin of my investigations, searching for a simple proof of the Hölder inequality for matrices and operators. (For other proofs, see f.e. [MBR72, RS75, RB97, EC09])

2 THEOREM. ***Matrix Hölder Inequality:*** *Consider two $m \times m$ matrices A, B and their absolute values, then*

$$|\mathrm{Tr}\, A^*B| \le (\mathrm{Tr}\, |A|^p)^{1/p} \cdot (\mathrm{Tr}\, |B|^q)^{1/q}, \qquad 1 \le p, q \le \infty, \quad p^{-1} + q^{-1} = 1 \tag{6}$$

Proof. Using the same notation as in the proof of Theorem 1, we note that $\operatorname{Tr}|A|^p = \operatorname{Tr}|A^*|^p = \sum_i a_i^p$. So, the Hölder Inequality for the left hand side of (1) is proven, if it holds for each factor on the right hand side. There we have normal operators. For these we can use the classical Hölder Inequality for weighted sums, followed by using completeness of the basis sets $\{e_i\}$ and $\{g_j\}$:

$$\operatorname{Tr}|A|\cdot|B| = \sum_{i,j} a_i b_j \, |\langle e_i, g_j\rangle|^2 \;\leq\; \left(\sum_{i,j} a_i^p \, |\langle e_i, g_j\rangle|^2\right)^{1/p} \cdot \left(\sum_{i,j} b_j^q \, |\langle e_i, g_j\rangle|^2\right)^{1/q}$$

$$= \left(\sum_i a_i^p\right)^{1/p} \cdot \left(\sum_j b_j^q\right)^{1/q} \;=\; (\operatorname{Tr}|A|^p)^{1/p} \cdot (\operatorname{Tr}|B|^q)^{1/q}.$$

Analogously for $\operatorname{Tr}|A^*|\cdot|B^*|$. □

2 Generalizations

There are possibilities to generalize the baby inequality. It can grow by: Inserting extra matrices, one $m \times m$ another one $n \times n$; using different exponents for the different absolute values; going into vector spaces with infinite dimension.

One can insert extra matrices M and N:

3 THEOREM. ***Inequality with intermediate matrices:***

$$|\operatorname{Tr} MA^*NB| \leq (\operatorname{Tr} M\,|A|\,M^*\,|B|\,)^{1/2} \cdot (\operatorname{Tr} N^*\,|A^*|\,N\,|B^*|\,)^{1/2} \tag{7}$$

Proof. Just insert the extra matrices in the right places in the inner products appearing in (2) and (3): Replace $\langle g_j, e_i\rangle$ by $\langle g_j, Me_i\rangle$ and $\langle f_i, h_j\rangle$ by $\langle f_i, Nh_j\rangle$. □

There is the possibility to consider different exponents:

4 THEOREM. ***Inequality with exponents:***

$$|\operatorname{Tr} A^*B| \leq \left(\operatorname{Tr}|A|^\alpha\,|B|^\beta\,\right)^{1/2} \cdot \left(\operatorname{Tr}|A^*|^{2-\alpha}\,|B^*|^{2-\beta}\right)^{1/2}, \qquad 0 \leq \alpha, \beta \leq 2, \tag{8}$$

with $0^0 = 0$*, so that* $|A|^0 =$ *projector onto range*$(|A|)$.

Proof. Modify (2) and (3) as

$$|\operatorname{Tr} A^*B| = \left|\sum_{i,j} \varphi_{i,j} \cdot \psi_{i,j}\right| \leq \left(\sum_{i,j} |\varphi_{i,j}|^2 \cdot |\psi_{i,j}|^2\right)^{1/2}, \tag{9}$$

with $\varphi_{i,j} = a_i^{\alpha/2} b_j^{\beta/2} \langle e_i, g_j\rangle$, $\psi_{i,j} = a_i^{1-\alpha/2} b_j^{1-\beta/2} \langle h_j, f_i\rangle$. Observe, that for $a_i = 0$ the matrix elements involving e_i or f_i are just absent. The same holds for b_j, g_j, h_j. □

Extensions into infinite dimensions can be done in different ways. I present the following result:

5 THEOREM. ***Inequality for two Hilbert Schmidt class operators:*** *Let A and B be operators from Hilbert space $\mathcal{H}$ to the Hilbert space $\mathcal{K}$, with the properties* $\mathrm{Tr}_{\mathcal{H}}\, A^*A = \mathrm{Tr}_{\mathcal{K}}\, A{\cdot}A^* < \infty$ *and* $\mathrm{Tr}_{\mathcal{H}}\, B^*B = \mathrm{Tr}_{\mathcal{K}}\, B{\cdot}B^* < \infty$*. Then*

$$|\mathrm{Tr}_{\mathcal{H}}\, A^*B| \leq (\mathrm{Tr}_{\mathcal{H}}\, |A|\,|B|\,)^{1/2} \cdot (\mathrm{Tr}_{\mathcal{K}}\, |A^*|\,|B^*|)^{1/2} \tag{10}$$

Proof. As in the proof of Theorem 1 we use the singular values a_i and basis sets e_i, f_i, here $e_i \in \mathcal{H}$, $f_i \in \mathcal{K}$, such that $Ae_i = a_i f_i$, and analogously $Bg_i = b_i h_i$. In Dirac's notation $A = \sum_i |f_i\rangle a_i \langle e_i|$ and $B = \sum_i |h_i\rangle b_i \langle g_i|$. A being in the Hilbert-Schmidt class means

$$\mathrm{Tr}_{\mathcal{H}}\, A^*A = \mathrm{Tr}_{\mathcal{K}}\, A{\cdot}A^* = \mathrm{Tr}_{\mathcal{H}}\, |A|^2 = \mathrm{Tr}_{\mathcal{K}}\, |A^*|^2 = \left(\sum_i a_i^2\right)^{1/2} < \infty,$$

and the analogue for B. Introducing the operators with finite rank

$$A_N = \sum_i^N |f_i\rangle|a_i\langle e_i| \qquad \text{and} \qquad B_N = \sum_i^N |h_i\rangle|b_i\langle g_i|, \tag{11}$$

one can apply Theorem 1 to the matrices which represent these operators, giving the inequality (10) for A_N and B_N. One observes the convergences in norm:

$$\|A{-}A_N\| = \|A^*{-}A_N^*\| = \|\,|A|{-}|A_N|\,\| = \|\,|A^*|{-}|A_N^*|\,\| = \left(\sum_{N+1}^{\infty} a_i^2\right)^{1/2} \to_{N\to\infty} 0,$$

and the same for B. The Hilbert-Schmidt inner products are jointly norm-continuous in both factors, so each side of the inequality (10) for A_N and B_N converges as $N \to \infty$, giving the same inequality without the N as an index. □

Applications are new proofs for Hölder type inequalities used in mathematical physics. They will be discussed in a following article.

3 Comparison with another use of absolute values

The product $A^*{\cdot}B$ can be represented as

$$A^*{\cdot}B = U{\cdot}|A^*|{\cdot}|B^*|{\cdot}V^*, \tag{12}$$

by extending the $m \times n$ matrices to $N \times N$ matrices, where $N = \max\{m, n\}$, and using the polar decompositions $A^* = U \cdot |A^*|$, $B^* = V \cdot |B^*|$. (Equivalently, one can stay with the $m \times n$ matrices and use isometries U and V instead of unitary operators.) This equality implies that the set of singular values of $A^* \cdot B$ is identical to that of $|A^*| \cdot |B^*|$. (Eventually, when staying with $m \neq n$, the numbers of zeroes are different.) So, all the invariant norms, see [RB97], are identical. One identity is for the operator norm

$$\|A^* \cdot B\| = \| \, |A^*| \cdot |B^*| \, \|, \tag{13}$$

another one gives

$$\mathrm{Tr}\, |A^* \cdot B| = \mathrm{Tr}\, | \, (|A^*| \cdot |B^*|) \, |. \tag{14}$$

Together with $|\mathrm{Tr} M| \leq \mathrm{Tr}\, |M|$, which holds for each matrix, we get

$$|\mathrm{Tr}\, A^* \cdot B| \leq \mathrm{Tr}\, | \, (|A^*| \cdot |B^*|) \, |. \tag{15}$$

This inequality is not as sharp as the baby inequality given in Theorem 1.

References

[MBR72] Mary Beth Ruskai: *Inequalities for Traces on von Neumann Algebras* Commun. Math. Phys. **26**, 280 – 289, (1972)

[RS75] Michael Reed and Barry Simon: *Methods of Modern Mathematical Physics II* (Academic Press, New York) 1975

[RB97] Rajendra Bhatia: *Matrix Analysis*. Graduate Texts in Mathematics **169**, (Springer Verlag, New York) 1997

[EC09] Eric A. Carlen: *Trace Inequalities and Quantum Entropy: An introductory course.* Lecture Course given at "Entropy and the Quantum" - Tucson, Arizona, March 16-20, 2009 http://www.mathphys.org/AZschool/material/AZ09-carlen.pdf