Indian J. Pure Appl. Math., **41**(1): 99-111, February 2010

MODULUS OF CONTINUITY OF THE MATRIX ABSOLUTE VALUE

Rajendra Bhatia

Indian Statistical Institute, New Delhi 110 016, India
e-mail: rbh@isid.ac.in

Abstract Lipschitz continuity of the matrix absolute value $|A| = (A^*A)^{1/2}$ is studied. Let A and B be invertible, and let $M_1 = \max(\|A\|, \|B\|)$, $M_2 = \max(\|A^{-1}\|, \|B^{-1}\|)$. Then it is shown that

$$\| \, |A| - |B| \, \| \leq (1 + \log \ M_1 M_2) \ \|A - B\|.$$

A proof is given for the well-known theorem that there is a constant $c(n)$ such that for any two $n \times n$ matrices A and B $\| \, |A| - |B| \, \| \leq c(n) \ \|A - B\|$ and the best constant in this inequality is $O(\log \ n)$.

Key words Matrix absolute value, perturbation bound, commutator, triangular truncation, Schur product.

1. Introduction

Let $\mathbb{M}(n)$ be the space of $n \times n$ complex matrices. Every element A of $\mathbb{M}(n)$ has a polar decomposition $A = UP$ in which U is unitary and P is positive (semidefinite). The matrix $P = (A^*A)^{1/2}$ is called the *absolute value* of A and is denoted as $|A|$. For computational and theoretical reasons it is of interest to know how well behaved the map $f(A) = |A|$ is.

The two most used norms on $\mathbb{M}(n)$ are the *operator norm (spectral norm)* and the *Hilbert-Schmidt norm (Frobenius norm)* defined, respectively, as

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = s_1(A), \tag{1}$$

$$\|A\|_2 = (\text{tr } A^*A)^{1/2} = \left(\sum_{i,j} \|a_{ij}\|^2 \right)^{1/2} = \left(\sum_j s_j^2(A) \right)^{1/2}, \tag{2}$$

where $s_1(A) \geq s_2(A) \geq \cdots \geq s_n(A)$ are the singular values of A. Clearly

$$\|A\| \leq \|A\|_2 \leq \sqrt{n} \, \|A\|. \tag{3}$$

For several reasons it is much easier to deal with the Frobenius norm, and so it is for this problem.

The question discussed in this mainly expository note is whether we have an inequality of the type

$$\| \, |A| \, - |B| \, \| \leq C \, \|A - B\|, \tag{4}$$

and, if so, what is the "best" factor C that occurs here. If we allow C to depend on A and B we have the following inequalities, the proofs of all of which are given in the book [2]:

$$\| \, |A| \, - |B| \, \| \leq \left(1 + \frac{2 \, \min(\|A\|, \|B\|)}{\|A^{-1}\|^{-1} + \|B^{-1}\|^{-1}}\right) \|A - B\|, \tag{5}$$

$$\| \, |A| \, - |B| \, \| \leq \frac{2}{\pi}\left(2 + \log \frac{\|A\| + \|B\|}{\|A - B\|}\right) \|A - B\|, \tag{6}$$

$$\| \, |A| \, - |B| \, \| \leq \mathrm{cond}(A)\|A - B\| + O\left(\|A - B\|^2\right). \tag{7}$$

See Exercise VII.5.3, Theorem X.2.5 and Theorem X.3.6 in [2]. The inequality (5) can be deduced from the main theorem in [16], and here it is assumed that A and B are invertible; (6) is due to Kato [13]. The inequality (7) is slightly different from the kind in (4); here it is assumed that A is invertible, B is close to A, and $\mathrm{cond}(A)$ stands for the condition number $\|A\| \, \|A^{-1}\|$. This inequality and its refinements were proved in [5], [6].

In Section 5 we prove another inequality of this sort:

Theorem 1.1. *Let A and B be nonsingular $n \times n$ matrices. Then*

$$\| \, |A| \, - |B| \, \| \leq (1 + \log \, \mathrm{cond}(A \oplus B)) \, \|A - B\|. \tag{8}$$

Here $A \oplus B$ stands for the block matrix $\begin{bmatrix} A & O \\ O & B \end{bmatrix}$. So

$$\mathrm{cond}(A \oplus B) = \max(\|A\|, \|B\|) \max(\|A^{-1}\|, \|B^{-1}\|). \tag{9}$$

These inequalities are "local" in that they involve not just $\|A - B\|$ on the right hand side but factors that depend on $\|A\|$ and $\|B\|$, and worse perhaps $\|A^{-1}\|$ and $\|B^{-1}\|$. For the Frobenius norm there is a "global" inequality.

$$\| \, |A| \, - |B| \, \|_2 \leq \sqrt{2} \, \|A - B\|_2. \tag{10}$$

The factor $\sqrt{2}$ here is best possible. This inequality (for compact Hilbert-Schmidt operators) was first proved by Araki and Yamagami [1]. A much simpler proof (of a stronger inequality) was given by Kittaneh [14]. This proof is reproduced in [2, pp. 214-215]. From (10) it follows that

$$\| \, |A| \, - |B| \, \| \leq \sqrt{2n} \, \|A - B\|. \tag{11}$$

This does give an inequality of the form (4) but with a factor growing with the dimension. It turns out that this dependence on n cannot entirely be dispensed with. However, a much better inequality can be proved. We have

$$\| \, |A| \, - |B| \, \| \leq C(n) \|A - B\|, \tag{12}$$

where $C(n) = O(\log\, n)$, and this is best possible. This result, first obtained in connection with problems coming from physics, is of interest to theoretical computer scientists and numerical analysts as well. In Section 4 we give a complete, self-contained and simple proof of this fact.

This problem has close connections with a seemingly unrelated matter. Let $\triangle_n$ be the *triangluar truncation* operator on $\mathbb{M}(n)$; i.e. $\triangle_n(A)$ is the upper triangular matrix with entries a_{ij} for all $i \leq j$. Clearly $\| \triangle_n \, (A)\|_2 \leq \|A\|_2$. However $\| \triangle_n \, (A)\|$ can be larger than $\|A\|$, and it is known that

$$\| \triangle_n \, (A)\| \leq C'(n) \|A\|, \tag{13}$$

where $C'(n) = O(\log\, n)$.

When the norm $\| \cdot \|$ is replaced by any of the Schatten p-norms, $1 < p < \infty$, the factor $C'(n)$ in (13) can be replaced by a factor γ_p that depends on p but is independent of n. Using a beautiful and ingenious argument, Davies [11] has used this fact to obtain an estimate like (12) for these norms with a factor independent of n in place of $C(n)$. The proof given here is essentially taken from [11]. The only novelty is that for the inequality (13) we give a rather simple proof using elementary Fourier analysis. That idea is taken from my earlier paper [7]. The bridge between the estimates (13) and (12) is a theorem on Schur multiplier norms that is of independent interest. This can be stated in another form as follows. Let A be a positive definite matrix and let X be any matrix. Then

$$\|AX - XA\| \leq C''(n) \, \|AX + XA\|, \tag{14}$$

where $C''(n) = O(\log\, n)$. We prove the following theorem, on which is based our proof of Theorem 1.1.

Theorem 1.2. *Let A be a positive definite matrix and let X be any matrix. Then*

$$\|AX - XA\| \leq \frac{1}{2} \, \log\, \left(\|A\| \, \|A^{-1}\|\right) \|AX + XA\|. \tag{15}$$

Though our problem is of obvious intrinsic importance, it might be interesting to mention its early history. The analyst A. P. Calderon made many contributions to the subject of singular integral operators. To prove that such operators are bounded is often a difficult task. One of Calderon's theorems is that the operator T defined on $L_2(\mathbb{R})$ as

$$Tf(x) = \frac{1}{\pi} \text{ p.v. } \int \frac{b(x) - b(y)}{(x-y)^2} \, f(y) \, dy$$

is bounded if the function b' is (essentially) bounded. This led T. Kato to the following consideration. Let $Af = if'$ and $Bf = bf$. The operator A, called the differentiation operator, is unbounded on $L_2(\mathbb{R})$; the operator $B,$ called a multiplication operator, is bounded if the function b is bounded. Evidently

$$(AB - BA)f = i(bf)' - ibf' = ib'f.$$

So if b' is a bounded function, then $AB - BA$ is a bounded operator. It turns out that the singular integral operator T can be expressed as

$$T = |A|B - B|A|.$$

So if one could show, in general for two operators A and B on a Hilbert space, that $|A|B - B|A|$ is bounded whenever $AB - BA$ is bounded, then Calderon's theorem would follow as a very special case. A counterexample to this question of Kato was given by McIntosh [18].

Kato then proved the inequality (6) and interpreted it to mean that the map $A \longmapsto |A|$ is almost Lipschitz continuous. The general question of finding conditions for functions f on $\mathbb{R}$ to be Lipschitz continuous on Hermitian operators was studied by M. Sh. Birman and his collaborators. Particularly relevant to the problem we are discussing is the work of Farforovskaya [12].

2. Trimming and Truncating of Matrices

In this section we study how replacing some diagonals of a matrix A by zeros affects $\|A\|$. The problem was studied in [7] from where we borrow the results.

For $-n < k < n$ let $\mathcal{D}_k(A)$ be the matrix obtained from A as follows: if $s = r + k$, then the (r, s) entry of $\mathcal{D}_k(A)$ is a_{rs}, and all other entries of $\mathcal{D}_k(A)$ are zero. In other words $\mathcal{D}_k(A)$ is the matrix obtained from A by replacing with zeros all entries of A except those on one of the diagonals parallel to the main diagonal. Let U_θ be the diagonal matrix $U_\theta = \operatorname{diag}\left(e^{i\theta}, e^{i2\theta}, \ldots, e^{in\theta}\right)$. Then

$$\mathcal{D}_k(A) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta} U_\theta A U_\theta^* d\theta. \tag{16}$$

Let

$$\mathcal{T}_{2k+1}(A) = \sum_{j=-k}^{k} \mathcal{D}_j(A). \tag{17}$$

This is a *band matrix* obtained from A by retaining all entries on the main diagonal and on k sub and super diagonals, and replacing the entries of A outside this band by zero. We call this a *trimming* of A. The sum

$$D_k(\theta) = \sum_{j=-k}^{k} e^{ij\theta}$$

is called the *Dirichlet kernel* and the numbers

$$L_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} |D_k(\theta)|\, d\theta$$

are called the *Lebesgue constants.* Since $\|U_\theta A U_\theta^*\| = \|A\|$ for all θ, it follows from (16) and (17) that

$$\|\mathcal{T}_{2k+1}(A)\| \leq L_k \|A\|. \tag{18}$$

It is a well-known fact of Fourier analysis [3, p.35] that

$$L_k = \frac{4}{\pi^2}\ \log\ k + O(1).$$

From the estimate (18) one can obtain an estimate for $\| \triangle_n (A)\|$ by going to 2×2 block matrices. If A is an $n \times n$ matrix, then we let

$$\widetilde{A} = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}.$$

We have $\|\widetilde{A}\| = \|A\|$. Evidently

$$\mathcal{T}_{2n+1}(\widetilde{A}) = \begin{bmatrix} 0 & \triangle_n(A) \\ \triangle_n(A)^* & 0 \end{bmatrix}.$$

Hence

$$\| \triangle_n (A)\| \leq C'(n)\|A\|,$$

where $C'(n) = O(\log\ n)$. This is the statement (13). This inequality is of the best possible kind. The famous Hilbert matrix X with entries $x_{ii} = 0$ and $x_{ij} = 1/(i-j)$ when $i \neq j$, has the properties $\|X\| \leq \pi$ and $\triangle_n(X) = O(\log\ n)$. (From what we know about the Lebesgue constants, it is possible to get more precise information about $C'(n)$ and other such expressions that follow. We leave this to the interested reader.)

3. Norms of Schur Products

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $n \times n$ matrices. Their *Schur-Hadamard product* is the matrix $A \circ B = [a_{ij}b_{ij}]$. Each A induces a linear operator on $\mathbb{M}(n)$ defined as $X \longmapsto A \circ X$. The norm of this linear operator is

$$\|A\|_{SH} = \sup_{\|X\|=1}\ \|A \circ X\|. \tag{19}$$

Generally it is difficult to evaluate this norm.

When A is a positive semidefinite matrix, a simple expression

$$\|A\|_{SH} = \max_i\ \ a_{ii}$$

was obtained by Schur. See e.g., [4, p. 16].

Let E_n be the $n \times n$ matrix all of whose entries are equal to 1, T_n the matrix obtained from E_n by replacing all entries below the main diagonal by 0, and S_n the matrix obtained by replacing all entries below the main diagonal by -1. Then

$$T_n = \frac{1}{2}\left(S_n + E_n\right). \tag{20}$$

Clearly $\|E_n\|_{SH} = 1$; and since $\triangle_n(X) = T_n \circ X$, the result of Section 2 can be restated as

$$\|T_n\|_{SH} = O(\log\, n). \tag{21}$$

So from (20) we have

$$\|S_n\|_{SH} = O(\log\, n). \tag{22}$$

Proposition 3.1. *Let $\lambda_1, \ldots, \lambda_n$ be positive real numbers and let W be the matrix with entries*

$$w_{ij} = \frac{|\lambda_i - \lambda_j|}{\lambda_i + \lambda_j}. \tag{23}$$

Then

$$\|W\|_{SH} \leq 2.$$

Proof : Using the equation

$$\begin{aligned} w_{ij} &= \frac{\lambda_i + \lambda_j - 2\, \min(\lambda_i, \lambda_j)}{\lambda_i + \lambda_j} \\ &= 1 - 2\, \frac{\min(\lambda_i, \lambda_j)}{\lambda_i + \lambda_j}, \end{aligned}$$

we can express W as the difference $W = E - Q$, where E is the matrix with all entries equal to 1, and

$$q_{ij} = \frac{2\, \min(\lambda_i, \lambda_j)}{\lambda_i + \lambda_j}.$$

The matrices $[\min(\lambda_i, \lambda_j)]$ and $\left[\frac{1}{\lambda_i + \lambda_j}\right]$ both are positive semidefinite; see [4, p.152] and [4, p.3]. Hence the matrix Q is a positive semidefinite matrix with $q_{ii} = 1$. So by Schur's theorem $\|Q\|_{SH} = 1$, and therefore $\|W\|_{SH} \leq \|E\|_{SH} + \|Q\|_{SH} \leq 2$. $\square$

Corollary 3.2. *Let $\lambda_1, \ldots, \lambda_n$ be positive real numbers and let Y be the matrix with entries*

$$y_{ij} = \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j}.$$

Then

$$\|Y\|_{SH} \leq C\, \log\, n, \tag{24}$$

where C is a constant independent of $\lambda_1, \ldots, \lambda_n$.

Proof : Assume, without loss of generality, that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then

$$y_{ij} = \frac{|\lambda_i - \lambda_j|}{\lambda_i + \lambda_j}\, \operatorname{sgn}(i - j).$$

So Y can be expressed as the product $Y = W \circ S_n$, and we have

$$\|Y\|_{SH} \leq \|W\|_{SH} \|S_n\|_{SH}.$$

The inequality (24) now follows from (22) and (23). □

The inequality (24) is of the best possible kind. Let $\alpha > 0$ and let $\lambda_i = \alpha^i$. Then

$$y_{ij} = \frac{\alpha^i - \alpha^j}{\alpha^i + \alpha^j} = \frac{\alpha^i/\alpha^j - 1}{\alpha^i/\alpha^j + 1}.$$

As $\alpha \longrightarrow \infty$, y_{ij} goes to 1 for $i > j$ and to -1 for $i < j$. The entries y_{ii} are zero. So $Y \approx S_n - I$ and $\|Y\|_{SH} \approx \|S_n\|_{SH} = O(\log\, n)$.

Corollary 3.3. *Let $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$ be positive real numbers and let X be the $n \times n$ matrix with entries*

$$x_{ij} = \frac{\lambda_i - \mu_j}{\lambda_i + \mu_j}.$$

Then

$$\|X\|_{SH} \leq C\ \log\, n, \tag{25}$$

where C is a constant independent of $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n$.

Proof : List the given $2n$ numbers as $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n$ and form the $2n \times 2n$ matrix Y as in Corollary 3.2. This matrix has the form

$$Y = \begin{bmatrix} Y_{11} & X \\ -X' & Y_{22} \end{bmatrix}.$$

From this it follows that $\|X\|_{SH} \leq \|Y\|_{SH} \leq C\ \log\, 2n$, where C is the constant in (24). So the inequality (25) is valid, now with a different C. □

Remark 3.4 : Some of these considerations can be extended to rectangular matrices. Identify an $m \times n$ matrix X with a linear operator from $\mathbb{C}^n$ into $\mathbb{C}^m$ and let $\|X\|$ be the norm of this linear operator. Let $Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix}$ in which Y_{11} and Y_{22} are, respectively $m \times m$ and $n \times n$ matrices. Then $Y_{12} = PAQ$, where P and Q are orthogonal projections, and hence $\|Y_{12}\| \leq \|Y\|$.

The Schur norm $\|X\|_{SH}$ can be defined as before. If $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_n$ are positive numbers, and X is the $m \times n$ matrix with entries.

$$x_{ij} = \frac{\lambda_i - \mu_j}{\lambda_i + \mu_j},$$

then we have $\|X\|_{SH} \leq C\ \log(m + n)$.

4. Commutators and the Absolute Value

Let $T = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix}$ be a partitioned matrix. Then $\|T_j\| \leq \|T\|$ for $1 \leq j \leq 4$, and

$$\begin{aligned} \|T\| &\leq \left\| \begin{bmatrix} T_1 & 0 \\ 0 & T_4 \end{bmatrix} \right\| + \left\| \begin{bmatrix} 0 & T_2 \\ T_3 & 0 \end{bmatrix} \right\| \\ &= \max\left(\|T_1\|, \|T_4\|\right) + \max\left(\|T_2\|, \|T_3\|\right) \\ &\leq \|T_1\| + \|T_2\| + \|T_3\| + \|T_4\|. \end{aligned}$$

Theorem 4.1. *There exists a positive number C with the following property: if A and B are operators on an n-dimensional Hilbert space $\mathcal{H}$ and A is Hermitian, then*

$$\| \, |A|B - B|A| \, \| \leq C \ \log n \ \|AB - BA\|. \tag{26}$$

Proof : Choose an orthonormal basis for $\mathcal{H}$ in which A is diagonal and has the form

$$A = \begin{bmatrix} P & O \\ O & -Q \end{bmatrix}, \tag{27}$$

where $P = \operatorname{diag}(\lambda_1, \ldots, \lambda_k)$ and $Q = \operatorname{diag}(\mu_1, \ldots, \mu_{n-k})$ both are nonnegative. In this basis let B have the partitioned form

$$B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}.$$

By adding a few zeros if necessary, we may assume $k = n-k$. This does not affect the argument that follows. It is obvious that

$$AB - BA = \begin{bmatrix} PB_1 - B_1P & PB_2 + B_2Q \\ -QB_3 - B_3P & -QB_4 + B_4Q \end{bmatrix}. \tag{28}$$

From (27) we see that

$$|A| = \begin{bmatrix} P & O \\ O & Q \end{bmatrix},$$

and hence,

$$|A|B - B|A| = \begin{bmatrix} PB_1 - B_1P & PB_2 - B_2Q \\ QB_3 - B_3P & QB_4 - B_4Q \end{bmatrix}. \tag{29}$$

We will compare the norms of the four blocks of (29) with the corresponding blocks of (28). The $(1,1)$ blocks of the two matrices are the same; the $(2,2)$ blocks differ only by a sign and, therefore, their norms are equal. The blocks in the $(1,2)$ position are related as follows:

$$PB_2 - B_2Q = \left[\frac{\lambda_i - \mu_j}{\lambda_i + \mu_j}\right] \circ (PB_2 + B_2Q).$$

So from (25) we have

$$\|PB_2 - B_2Q\| \leq C \ \log \ k \ \|PB_2 + B_2Q\|.$$

The same argument shows that

$$\|QB_3 - B_3P\| \leq C \ \log \ k \ \|QB_3 + B_3P\|.$$

Now using the relation between the norms of a block matrix and the blocks, given at the beginning of this section, we can obtain the inequality (26) (with another constant C). □

Using this we prove our main theorem.

Theorem 4.2. *There exists a positive number C such that for any two $n \times n$ matrices A and B we have*

$$\| \, |A| - |B| \, \| \leq C \;\log\, n \, \|A - B\|. \tag{30}$$

Proof : Let H and K be $n \times n$ Hermitian matrices and let

$$R = \begin{bmatrix} H & O \\ O & K \end{bmatrix}, \qquad S = \begin{bmatrix} O & I \\ O & O \end{bmatrix}.$$

From Theorem 4.1 we have

$$\| \, |R|S - S|R| \, \| \leq C \;\log\; n \, \|RS - SR\| \tag{31}$$

for some constant C. Now note that

$$|R|S - S|R| = \begin{bmatrix} O & |H| - |K| \\ O & O \end{bmatrix}, \qquad RS - SR = \begin{bmatrix} O & H - K \\ O & O \end{bmatrix}.$$

Hence from (31) we have

$$\| \, |H| - |K| \, \| \leq \; C \;\log\; n \, \|H - K\|. \tag{32}$$

For any two $n \times n$ matrices A and B let

$$H = \begin{bmatrix} O & A \\ A^* & O \end{bmatrix}, \qquad K = \begin{bmatrix} O & B \\ B^* & O \end{bmatrix}.$$

Then H and K are Hermitian, and

$$|H| = \begin{bmatrix} |A^*| & O \\ O & |A| \end{bmatrix}, \qquad |K| = \begin{bmatrix} |B^*| & O \\ O & |B| \end{bmatrix}.$$

From the inequality (32) we obtain (30) (with a different constant C). □

Remark 4.3. The key estimate for this argument is (20). This also shows why the Frobenius norm is different. Since $|y_{ij}| \leq 1$ for the matrix Y of Corollary 3.2 we have $\|Y \circ A\|_2 \leq \|A\|_2$ for all A. If we imitate the proofs of Theorem 4.1 and 4.2 we get instead of (26) the inequality

$$\| \, |A|B - B|A| \, \|_2 \leq \|AB - BA\|_2 \tag{33}$$

valid for Hermitian A and all B, and instead of (30)

$$\| \, |A^*| - |B^*| \, \|_2 + \| \, |A| - |B| \, \|_2 \leq \sqrt{2} \, \|A - B\|_2 \tag{34}$$

for all A and B. This is a stronger statement than (10) and was first proved by Kittaneh. It is known [8] that among all unitarily invariant norms the Frobenius norm is the only one that is a monotonically increasing function of the absolute values of the entries of a matrix. So this argument is very special to the Frobenius norm. The main point of Davies [11] is that for the Schatten p-norm, $1 < p < \infty$, there exists a constant γ_p, independent of n, such that $\|Y \circ A\|_p \leq \gamma_p \|A\|_p$ for all A.

Remark 4.4. Kosaki [15] has shown that for a unitarily invariant norm $|||\cdot|||$ the following conditions are equivalent:

(i) There exists a constant C_1 independent of the dimension n such that

$$|||\,|A| - |B|\,||| \leq C_1\, |||A - B|||$$

for all A and B.

(ii) There exists a constant C_2 independent of n such that

$$|||\Delta_n(A)||| \leq C_2\, |||A|||$$

for all A.

(iii) There exists a constant C_3 independent of n such that for the matrix Y of Corollary 3.2 we have

$$|||Y \circ A||| \leq C_3\, |||A|||$$

for all A.

Remark 4.5. In [17] Mathias has followed an alternative approach to obtain the estimate (22). The special structure of S_n (it is a skew circulant) can be used to obtain explicit expressions for all its eigenvalues. This information can be used to find the value of $\|S_n\|_{SH}$.

5. Some New Inequalities

Let $|||\cdot|||$ be any unitarily invariant norm on $\mathbb{M}(n)$. Let

$$|||A|||_{SH} := \sup_{|||X|||=1} |||A \circ X|||. \tag{35}$$

It is a well known fact that if A is positive semidefinite, then

$$|||A|||_{SH} = \max_i \; a_{ii}.$$

(See Exercise 2.7.12 in [4].) The following proposition should be compared with Corollary 3.2.

Proposition 5.1. *Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be positive real numbers and let Y be the matrix with entries*

$$y_{ij} = \frac{\lambda_i - \lambda_j}{\lambda_i + \lambda_j}.$$

Then

$$|||Y|||_{SH} \leq \frac{1}{2} \left(\log\, \lambda_1 - \log\, \lambda_n\right). \tag{36}$$

Proof : We write Y as the product $Y = W \circ L$, where

$$w_{ij} = \frac{2(\lambda_i - \lambda_j)}{(\log\, \lambda_i - \log\, \lambda_j)\,(\lambda_i + \lambda_j)},$$

and

$$\ell_{ij} = \frac{1}{2} \left(\log\, \lambda_i - \log\, \lambda_j\right).$$

It is known [10], [4, p. 164], that the matrix W is positive semidefinite. Since $w_{ii} = 1$ for all i, we have $|||W|||_{SH} = 1$. For each X we have $L \circ X = \frac{1}{2}\,(DX - XD)$ where D is the diagonal matrix with entries $\log\, \lambda_j$ down its diagonal. By Theorem 2 in [9] we have, therefore,

$$|||L|||_{SH} \leq \frac{1}{2} \left(\log\, \lambda_1 - \log\, \lambda_n\right).$$

Combined with the information about W, this gives the inequality (36). □

A corollary of this is the following statement for all unitarily invariant norms that includes Theorem 1.2 as a special case.

Corollary 5.2. *Let A be a positive definite matrix and let X be any matrix. Then*

$$|||AX - XA||| \leq \frac{1}{2}\, \log\, \left(\|A\|\, \|A^{-1}\|\right) |||AX + XA|||. \tag{37}$$

Proof : If $\lambda_1 \geq \cdots \geq \lambda_n$ are the eigenvalues of A, then $\|A\| = \lambda_1$ and $\|A^{-1}\| = 1/\lambda_n$. So the inequality (37) follows from (36). □

More generally, using familiar arguments we can prove

Corollary 5.3. *Let A be an $n \times n$ and B an $m \times m$ positive definite matrix. Then for every $n \times m$ matrix X we have*

$$|||AX - XB||| \leq \frac{1}{2}\, \log\, \left(\|A \oplus B\|\, \|A^{-1} \oplus B^{-1}\|\right) |||AX + XB|||. \tag{38}$$

This, in turn, leads to the following.

Theorem 5.4. *Let A be a nonsingular Hermitian $n \times n$ matrix and let B be any $n \times n$ matrix. Then*

$$|||\, |A|B - B|A|\, ||| \leq \left(1 + \log\, \|A\|\, \|A^{-1}\|\right) |||AB - BA|||. \tag{39}$$

Proof : The proof is similar to that of Theorem 4.1. We indicate how the factor on the right hand side of (39) arises. if $T = \begin{bmatrix} T_1 & T_2 \\ T_3 & T_4 \end{bmatrix}$ is a partitioned matrix, then we have

$$|||T||| \leq |||T_1 \oplus T_4||| + |||T_2||| + |||T_3|||.$$

Further $|||T_1 \oplus T_4||| \le |||T|||$, and $|||T_j||| \le |||T|||$ for $1 \le j \le 4$. Using the same notations as in the proof of Theorem 4.1 we see that the diagonal of $|A|B - B|A|$ is $\begin{bmatrix} I & O \\ O & -I \end{bmatrix}$ times the diagonal of $AB - BA$. So these two diagonal parts have the same norm. To compare the off-diagonal blocks of (28) and (29) use Corollary 5.3. We have for the (1,2) blocks

$$\begin{aligned} |||PB_2 - B_2Q||| &\le \frac{1}{2} \log \left(\|P \oplus Q\| \, \|P^{-1} \oplus Q^{-1}\|\right) |||PB_2 + B_2Q||| \\ &= \frac{1}{2} \log \left(\|A\| \, \|A^{-1}\|\right) |||PB_2 + B_2Q|||. \end{aligned}$$

The same consideration can be applied to the (2,1) blocks. Combining all this information we obtain (39). □

For the sake of brevity we introduce the quantity

$$\begin{aligned} c(A,B) &= 1 + \log \operatorname{cond}(A \oplus B) \\ &= 1 + \log \left(\|A \oplus B\| \, \|A^{-1} \oplus B^{-1}\|\right). \end{aligned} \tag{40}$$

Theorem 5.5. *Let A and B be nonsingular $n \times n$ matrices. If A and B are normal, we have*

$$|||\,|A| - |B|\,||| \le c(A,B)\ |||A - B|||. \tag{41}$$

If A and B are arbitrary, we have

$$|||\ (|A^*| - |B^*|) \oplus (|A| - |B|)\ ||| \le c(A,B)\ |||\ (A^* - B^*) \oplus (A - B)\ |||. \tag{42}$$

Proof : Follow the arguments in the proof of Theorem 4.2 and obtain from Theorem 5.4 the inequality (41) for Hermitian A and B, and (42) for all A and B. If A and B are normal, then $|A^*| = |A|$ and $|B^*| = |B|$ and (41) follows from (42). □

For the operator norm we have $\|X \oplus Y\| = \max(\|X\|, \|Y\|)$ and $\|X^* \oplus X\| = \|X\|$. So Theorem 1.1 follows from (42). For the Schatten p-norms, $1 \le p < \infty$, the inequality (42) gives

$$\|\,|A^*| - |B^*|\,\|_p^p + \|\,|A| - |B|\,\|_p^p \le 2\,c(A,B)^p\,\|A - B\|_p^p. \tag{43}$$

Finally, we remark that the results of this section remain true for operators in infinite dimensional Hilbert spaces.

Acknowledgement

This is the text of a talk given at Microsoft Research Labs, Bangalore, in July 2008.

References

1. H. Araki and S. Yamagami, An inequality for the Hilbert-Schmidt norm, *Commun. Math. Phys.*, **81** (1981), 89-98.

2. R. Bhatia, *Matrix Analysis*, Springer, 1997.

3. R. Bhatia, *Fourier Series*, 2nd Edition, Hindustan Book Agency, 2003; Mathematical Association of America, 2005.

4. R. Bhatia, *Positive Definite Matrices*, Hindustan Book Agency; Princeton University Press, 2007.

5. R. Bhatia, First and second order perturbation bounds for the operator absolute value, *Linear Algebra Appl.*, **208/209** (1994), 367-376.

6. R. Bhatia, Perturbation bounds for the operator absolute value, *Linear Algebra Appl.*, **226** (1995), 639-645.

7. R. Bhatia, Pinching, trimming, truncating and averaging of matrices, *Amer. Math. Monthly*, **107** (2000), 602-608.

8. R. Bhatia, M. -D. Choi and C. Davis, Comparing a matrix to its off-diagonal part, *Operator Theory: Advances and Applications*, **40** (1989), 151-164.

9. R. Bhatia and F. Kittaneh, Commutators, pinchings, and spectral variation, *Operators and Matrices*, **2** (2008), 143-151.

10. R. Bhatia and K. R. Parthasarathy, Positive definite functions and operator inequalities, *Bull. London Math. Soc.*, **32** (2000), 214-228.

11. E. B. Davies, Lipschitz continuity of functions of operators in the Schatten classes, *J. London Math. Soc.*, **37** (1988), 148-157.

12. Yu. B. Farforovskaya, An estimate of the norm $\|f(B) - f(A)\|$ for self-adjoint operators A and B, *J. Soviet Math.*, **14** (1980), translated from the Russian original published in 1976.

13. T. Kato, Continuity of the map $S \to |S|$ for linear operators, *Proc. Japan Acad.*, **49** (1973), 157-160.

14. F. Kittaneh, On Lipschitz functions of normal operators, *Proc. Amer. Math. Soc.*, **94** (1985), 416-418.

15. H. Kosaki, Unitarily invariant norms under which the map $A \to |A|$ is Lipschitz continuous, *Publ. Res. Inst. Math. Sci.*, **28** (1992), 299-313.

16. R. -C. Li, New perturbation bounds for the unitary polar factors, *SIAM J. Matrix Anal. Appl.*, **16** (1995), 327-332.

17. R. Mathias, The Hadamard operator norm of a circulant and applications, *SIAM J. Matrix Anal. Appl.*, **14** (1993), 1152-1167.

18. A. McIntosh, Counterexample to a question on commutators, *Proc. Amer. Math. Soc.*, **29** (1971), 337-340.