Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

ELA
http://math.technion.ac.il/iic/ela

OPERATOR NORMS OF WORDS FORMED FROM POSITIVE-DEFINITE MATRICES*

S.W. DRURY†

Abstract. Let $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n$ be strictly positive reals with $\alpha_1 + \alpha_2 + \cdots + \alpha_n = \beta_1 + \beta_2 + \cdots + \beta_n = s$. In this paper, the inequality

$$|||A^{\alpha_1} B^{\beta_1} A^{\alpha_2} \cdots A^{\alpha_n} B^{\beta_n}||| \le |||AB|||^s$$

when A and B are positive-definite matrices is studied. Related questions are also studied.

Key words. Positive-definite matrix, Matrix power, Operator norm, Matrix words.

AMS subject classifications. 15A45.

1. Introduction.

Let A and B be positive-definite matrices of the same shape and let W be a word formed from A and B containing p A's and q B's. Then we may ask whether necessarily $|||W||| \le |||A^p B^q|||$. Here we have denoted $|||\cdot|||$ the operator norm.

For example we have $|||(AB)^k||| \le |||A^k B^k|||$ for k a nonnegative integer. To see this we assume without loss of generality that we have the normalization $|||A^k B^k||| = 1$, which amounts to $A^k B^{2k} A^k \le_L I$ or $B^{2k} \le_L A^{-2k}$. From this it follows that $B^2 \le_L A^{-2}$ since the map $u \mapsto u^{\frac{1}{k}}$ is matrix monotone. The reader should consult [1] for facts about matrix monotone functions. We can then deduce that $|||AB||| \le 1$ and thence $|||(AB)^k||| \le 1$.

This method, which we will call the monotonicity trick, works in many situations, for example $|||A^2B^2A||| \le |||A^3B^2|||$. We normalize as before and obtain $B^4 \le_L A^{-6}$ and $A^6 \le_L B^{-4}$ whence

$$AB^2A^4B^2A \le_L AB^2B^{-\frac{8}{3}}B^2A = AB^{\frac{4}{3}}A \le_L AA^{-2}A = I$$

as required. The monotonicity trick also shows that $|||A^2B^2AB^2A^2B^2A^2||| \le |||A^7B^6|||$.

However, it does not work for all words. An example where it fails is A^2BAB^2.

More generally, and after replacing the matrices A and B by suitable positive powers we may pose the following question. Let $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n$ be strictly

* Received by the editors July 31, 2008. Accepted for publication December 29, 2008. Handling Editor: Harm Bart.

†Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, Canada H3A 2K6.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

ELA
http://math.technion.ac.il/iic/ela

positive reals with $\alpha_1+\alpha_2+\cdots+\alpha_n=\beta_1+\beta_2+\cdots+\beta_n=s$. When is it necessarily true that

$$|||A^{\alpha_1}B^{\beta_1}A^{\alpha_2}\cdots A^{\alpha_n}B^{\beta_n}||| \le |||AB|||^s \tag{1.1}$$

for all pairs of positive-definite matrices A and B of the same shape? Colloquially, we may state the main result of this article by saying that if (1.1) can be proved by the monotonicity trick, then it is true and otherwise it is false (for some choice of A and B).

$0\le\beta_n\le 1$	$\beta_n\ge 0$
$0\le\alpha_n-\beta_n\le 1$	$\alpha_n+\beta_1+\cdots+\beta_{n-1}\ge s$
$0\le\beta_{n-1}+\beta_n-\alpha_n\le 1$	$\alpha_1+\cdots+\alpha_{n-1}+\beta_{n-1}+\beta_n\ge s$
$\vdots$	$\vdots$
$0\le\beta_2+\cdots+\beta_n-\alpha_3-\cdots-\alpha_n\le 1$	$\alpha_1+\alpha_2+\beta_2+\cdots+\beta_n\ge s$
$0\le\alpha_2+\cdots+\alpha_n-\beta_2-\cdots-\beta_n\le 1$	$\beta_1+\alpha_2+\cdots+\alpha_n\ge s$
$0\le\beta_1+\cdots+\beta_n-\alpha_2-\cdots-\alpha_n\le 1$	$\alpha_1\ge 0$

Table 1

2. The Main Theorem.

Our main result is the following.

THEOREM 2.1. *The inequality* (1.1) *is true for all pairs of positive-definite matrices A and B of the same shape if and only if the inequalities in the left-hand column of Table 1 are satisfied.*

We remark that the hypothesis that $\alpha_1,\alpha_2,\ldots,\alpha_n,\beta_1,\beta_2,\ldots,\beta_n$ are strictly positive is unfortunately necessary. For example, it is easy to see that $|||AB^2A^2B|||=|||(AB)(BA)(AB)|||\le|||AB|||^3$. Now consider $AB^2A^0B^0A^2B$. Then the left inequality in $0\le\beta_2+\beta_3-\alpha_3\le 1$ fails since $\beta_2+\beta_3-\alpha_3=-1$.

Proof. Let $W=A^{\alpha_1}B^{\beta_1}A^{\alpha_2}\cdots A^{\alpha_n}B^{\beta_n}$, then it is routine to check that the inequalities necessary for the application of the monotonicity trick applied to $WW^\star$ are precisely those in the left-hand column of Table 1. If one attempts to apply the monotonicity trick applied to $W^\star W$, then one gets the same inequalities in an equivalent form and in the reverse order. For the reverse implication, we need to show that if (1.1) is true for all pairs of positive-definite matrices A and B then the inequalities in the left-hand column of Table 1 are satisfied.

To tackle the right-hand inequalities, we use infinitessimal methods. Let $B=\operatorname{diag}(b_1,\ldots,b_d)$ where $b_k=e^{p_k}$ for $k=1,\ldots,d$. Let $P=\operatorname{diag}(p_1,\ldots,p_d)$ and $A(t)=\exp(-P-tQ)$ where Q is a $d\times d$ matrix to be specified later. Then $A(0)=B^{-1}$ and

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

we have

$$A(t)^\alpha = B^{-\alpha} + t\left[\frac{e^{-\alpha p_j} - e^{-\alpha p_k}}{p_j - p_k} q_{j,k}\right]_{j,k} + O(t^2)$$

where the quotient $\dfrac{e^{-\alpha p_j} - e^{-\alpha p_k}}{p_j - p_k}$ is interpreted as a divided difference (evaluating to $-\alpha e^{-\alpha p_j}$) in case $p_j = p_k$. Now let H be an arbitrary positive-definite $d \times d$ matrix and define Q by

$$q_{j,k} = -\frac{p_j - p_k}{e^{-2p_j} - e^{-2p_k}} h_{j,k}$$

then it follows that $A(t)^2 = B^{-2} - tH + O(t^2) \leq_L B^{-2}$ for $t \geq 0$ sufficiently small. Let $W(t) = A(t)^{\alpha_1} B^{\beta_1} A(t)^{\alpha_2} \cdots A(t)^{\alpha_n} B^{\beta_n}$, then $W(0) = I$ and so that by hypothesis $|||W(t)||| \leq 1$ for $0 \leq t$ small and therefore

$$\frac{d}{dt}_{t=0} W(t)W(t)^\star \leq_L 0.$$

A calculation shows that

$$\frac{d}{dt}_{t=0} W(t)W(t)^\star = -\left[h_{j,k}\left(\sum_{\ell=1}^{n} \frac{b_j^{-\alpha_\ell} - b_k^{-\alpha_\ell}}{b_j^{-2} - b_k^{-2}}\left(b_j^{\gamma_\ell} b_k^{\alpha_\ell - \gamma_\ell} + b_k^{\gamma_\ell} b_j^{\alpha_\ell - \gamma_\ell}\right)\right)\right]_{j,k}$$

where the γ_ℓ are linear combinations of $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n$ to be specified later. It follows that the matrix M given by

$$m_{j,k} = \sum_{\ell=1}^{n} \frac{b_j^{-\alpha_\ell} - b_k^{-\alpha_\ell}}{b_j^{-2} - b_k^{-2}}\left(b_j^{\gamma_\ell} b_k^{\alpha_\ell - \gamma_\ell} + b_k^{\gamma_\ell} b_j^{\alpha_\ell - \gamma_\ell}\right)$$

is a Schur multiplier of positive-definite matrices to positive-semidefinite matrices and therefore M is itself positive-semidefinite. However, the expression $b_j^{-1} b_k^{-1} m_{j,k}$ which is also positive-semidefinite is a function of b_j/b_k for all choices of positive (b_j) and hence arises as a continuous positive-definite function on the group $(0, \infty)$ with multiplication as the group operation. Written as a positive-definite function on the line this is

$$\varphi(u) = \sum_{\ell=1}^{d} \frac{\sinh(\frac{\alpha_\ell u}{2})}{\sinh(u)} \cosh\left(\frac{\alpha_\ell - 2\gamma_\ell}{2} u\right).$$

It is well-known [3] that such a function has to be bounded, in fact we must have $|\varphi(u)| \leq \varphi(0)$. However, each of the terms in the sum is nonnegative and it follows that

$$\left|\frac{\alpha_\ell}{2}\right| + \left|\frac{\alpha_\ell - 2\gamma_\ell}{2}\right| \leq 1$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

ELA
http://math.technion.ac.il/iic/ela

or equivalently

$$\alpha_\ell - \gamma_\ell \le 1 \qquad \text{and} \qquad \gamma_\ell \le 1$$

for all $\ell = 1, 2, \ldots, n$. We have $\gamma_1 = 0$ so that $0 \le 1$ and $\alpha_1 \le 1$, $\gamma_2 = \beta_1 - \alpha_1$ so that $\beta_1 - \alpha_1 \le 1$ and $\alpha_1 + \alpha_2 - \beta_1 \le 1$ etc. obtaining the $2n-1$ right-hand inequalities from the left-hand column of Table 1 reading upwards.

Next we tackle the left-hand inequalities in the left-hand column of Table 1 and these have been rewritten in an equivalent form in the right-hand column of the table. The first and last inequalities are therefore evident. A continuity argument allows us to assume that (1.1) is true for all pairs of positive-semidefinite matrices A and B. Let us define

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & p & p \\ 0 & p & p \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} p & p & 0 \\ p & p & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

with p small and positive to be determined. Then we have

$$A^x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^{x-1}p^x & 2^{x-1}p^x \\ 0 & 2^{x-1}p^x & 2^{x-1}p^x \end{pmatrix} \qquad \text{and} \qquad B^x = \begin{pmatrix} 2^{x-1}p^x & 2^{x-1}p^x & 0 \\ 2^{x-1}p^x & 2^{x-1}p^x & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

provided that $x > 0$, but not for $x = 0$. Then

$$(AB)(AB)^\star = \begin{pmatrix} 2p^2 & 2p^3 & 2p^3 \\ 2p^3 & p^2 + 2p^4 & p^2 + 2p^4 \\ 2p^3 & p^2 + 2p^4 & p^2 + 2p^4 \end{pmatrix}.$$

It follows that

$$|||AB||| = p\sqrt{2\left(1 + p^2 + \sqrt{p^4 + 2\,p^2}\right)} = O(p).$$

Clearly W has rank two and it follows that

$$\frac{1}{2}\mathrm{tr}(W^\star W) \le |||W|||^2 \le \mathrm{tr}(W^\star W)$$

leading to $\mathrm{tr}(W^\star W) = O(p^{2s})$. Also, with p chosen small and positive, all the terms comprising $\mathrm{tr}(W^\star W)$ are nonnegative and hence each individual term in the expansion of $\mathrm{tr}(W^\star W)$ is $O(p^{2s})$. There are two types of inequalities remaining to be discussed.

Case 1. The inequality $\beta_1 + \cdots + \beta_k + \alpha_{k+1} + \cdots + \alpha_n \ge s$ for $k = 1, \ldots, n-1$. We consider the individual term $z\overline{z}$ where

$$z = A_{1,1}^{\alpha_1} B_{1,1}^{\beta_1} \cdots A_{1,1}^{\alpha_k} B_{1,2}^{\beta_k} A_{2,3}^{\alpha_{k+1}} B_{3,3}^{\beta_{k+1}} \cdots B_{3,3}^{\beta_n} = Cp^{\beta_1 + \cdots + \beta_k + \alpha_{k+1} + \cdots + \alpha_n}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

ELA
http://math.technion.ac.il/iic/ela

Since $|z|^2 = O(p^{2s})$ as $p \downarrow 0$, we get the desired inequality.

Case 2. The inequality $\alpha_1 + \cdots + \alpha_k + \beta_k + \cdots + \beta_n \geq s$ for $k = 2, \ldots, n-1$. We consider the individual term $z\overline{z}$ where

$$z = A_{3,3}^{\alpha_1} B_{3,3}^{\beta_1} \cdots B_{3,3}^{\beta_{k-1}} A_{3,2}^{\alpha_k} B_{2,1}^{\beta_k} A_{1,1}^{\alpha_{k+1}} B_{1,1}^{\beta_{k+1}} \cdots B_{1,1}^{\beta_n} = Cp^{\alpha_1+\cdots+\alpha_k+\beta_k+\cdots+\beta_n}.$$

Since $|z|^2 = O(p^{2s})$ as $p \downarrow 0$, we get the desired inequality. □

3. Remarks and Comments. In light of the strict positivity assumption, it is instructive to observe that there is a corresponding result for words of odd length.

$0 \leq \alpha_n \leq 1$	$\alpha_n \geq 0$
$0 \leq \beta_{n-1} - \alpha_n \leq 1$	$\alpha_1 + \cdots + \alpha_{n-1} + \beta_{n-1} \geq s$
$0 \leq \alpha_{n-1} + \alpha_n - \beta_{n-1} \leq 1$	$\alpha_{n-1} + \alpha_n + \beta_1 + \cdots + \beta_{n-2} \geq s$
$\vdots$	$\vdots$
$0 \leq \beta_2 + \cdots + \beta_{n-1} - \alpha_3 - \cdots - \alpha_n \leq 1$	$\alpha_1 + \alpha_2 + \beta_2 + \cdots + \beta_{n-1} \geq s$
$0 \leq \alpha_2 + \cdots + \alpha_n - \beta_2 - \cdots - \beta_{n-1} \leq 1$	$\beta_1 + \alpha_2 + \cdots + \alpha_n \geq s$
$0 \leq \beta_1 + \cdots + \beta_{n-1} - \alpha_2 - \cdots - \alpha_n \leq 1$	$\alpha_1 \geq 0$

Table 2

THEOREM 3.1. *Let $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_{n-1}$ be strictly positive reals with $\alpha_1 + \alpha_2 + \cdots + \alpha_n = \beta_1 + \beta_2 + \cdots + \beta_{n-1} = s$. Then*

$$|||A^{\alpha_1} B^{\beta_1} A^{\alpha_2} \cdots B^{\beta_{n-1}} A^{\alpha_n}||| \leq |||AB|||^s$$

is true for all pairs of positive-definite matrices A and B of the same shape if and only if the inequalities in the left-hand column of Table 2 are satisfied.

We omit the proof since it is very similar to the proof of Theorem 2.1.

A case of some interest is Theorem 2.1 for $n = 2$, $\alpha_1 = \beta_2 = x$, $\alpha_2 = \beta_1 = 2y$, i.e. when does $|||A^x B^{2y} A^{2y} B^x||| \leq |||AB|||^{x+2y}$ hold for all A and B positive-definite? An equivalent formulation is to ask when we have $A^{-y} B^x A^{-y} \leq_L B^{-y} A^x B^{-y}$ given that A and B are positive-definite and that $B \leq_L A$. Theorem 2.1 gives that this holds if and only if either of the following conditions holds

- $y = 0$ and $0 \leq x \leq 1$,
- $0 \leq x \leq 1$ and $\frac{1}{2}x \leq y \leq \frac{1}{2}(1 + x)$.

Note that the region $0 < x \leq 1$, $0 < y < \frac{1}{2}x$ is excluded. This brings into sharp focus the necessity of insisting on strict positivity in Theorems 2.1 and 3.1.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009
ELA
http://math.technion.ac.il/iic/ela

It is amusing that one can give a precise answer to the infinitessimal question in this case. Precisely, this asks whether

$$\frac{d}{dt}_{t=0}\left(B^{-y}A(t)^xB^{-y}-A(t)^{-y}B^xA(t)^{-x}\right)\geq_L 0 \tag{3.1}$$

given that $A(0)=B$ and $\dfrac{d}{dt}_{t=0}A(t)\geq_L 0$.

Following the method of Theorem 2.1, we conclude that infinitessimal question is satisfied if and only if the function φ given by

$$\varphi(u)=\frac{2\sinh(xu)-\sinh((x-2y)u)}{\sinh(u)}$$

is positive-definite on $\mathbb{R}$. According to Bochner's Theorem [3], the continuous function φ is positive-definite if and only if it is the Fourier transform of a positive measure. According to [2, page 1148] the inverse Fourier transform of $u\mapsto\dfrac{\sinh(xu)}{\sinh(u)}$ is $\xi\mapsto\dfrac{\pi\sin(\pi x)}{\cosh(\pi\xi)+\cos(\pi x)}$ in case $|x|<1$ and is a point mass at 0 in case $x=\pm1$.

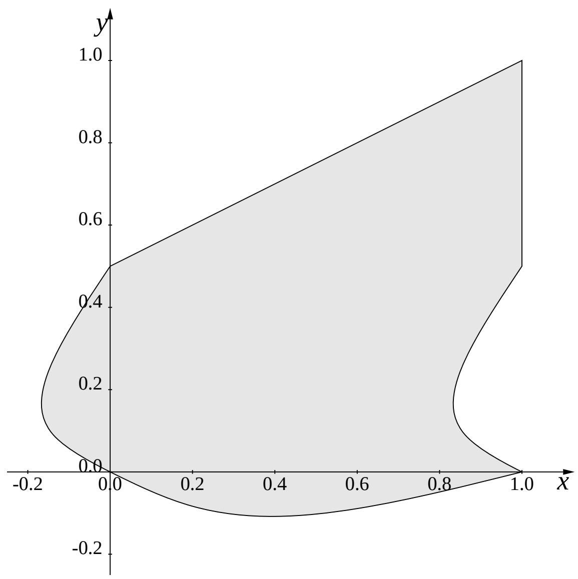

FIG. 3.1. *Region for the infinitessimal question.*

THEOREM 3.2. *Let x and y be real. Then the infinitessimal question* (3.1) *is satisfied if and only if all of the following hold.*

- $x\geq 2y-1$,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

- $x \leq 1$,
- $2\sin(\pi x) \geq \sin(\pi(x-2y))$,
- $y \geq \dfrac{x}{2} - \dfrac{1}{\pi}\arctan\left(2\tan\left(\dfrac{\pi x}{2}\right)\right)$.

The region of validity is depicted in Figure 3.1 and includes an area where $x < 0$ and an area where $y < 0$.

Proof. **Case** $|x| < 1$ **and** $|x - 2y| < 1$**.** It follows from Bochner's Theorem that we must have that for all ξ

$$2\frac{\pi\sin(\pi x)}{\cosh(\pi\xi)+\cos(\pi x)} \geq \frac{\pi\sin(\pi(x-2y))}{\cosh(\pi\xi)+\cos(\pi(x-2y))}.$$

This holds for all ξ if and only if it holds for $\xi = 0$ and also in the limit as $\xi \to \infty$. The two conditions boil down to

$$2\tan\left(\frac{\pi x}{2}\right) \geq \tan\left(\frac{\pi(x-2y)}{2}\right) \quad \text{and} \quad 2\sin(\pi x) \geq \sin(\pi(x-2y)).$$

Case $|x| = 1$ **and** $|x - 2y| = 1$**.** We see that if $|x| = 1$, then necessarily $x = 1$ or else φ cannot be nonnegative. The two points $(x, y) = (1, 1)$ and $(1, 0)$ corresponding to $|x - 2y| = 1$ are both admissible.

Case $|x| = 1$ **and** $|x - 2y| < 1$**.** Then we must have $x = 1$ and

$$\frac{\pi\sin(\pi(x-2y))}{\cosh(\pi\xi)+\cos(\pi(x-2y))} \leq 0$$

for all ξ which occurs for $x = 1$ and $\frac{1}{2} \leq y < 1$.

Case $|x| < 1$ **and** $|x - 2y| = 1$**.** Then we must have $x - 2y = -1$ for else the point mass will have a negative sign. We also require for all ξ that

$$2\frac{\pi\sin(\pi x)}{\cosh(\pi\xi)+\cos(\pi x)} \geq 0$$

which occurs for $0 \leq x < 1$.

Some additional work is required to show that these constraints reduce to the ones given in the theorem. ∎

REFERENCES

[1] William F. Donoghue, Jr. *Monotone Matrix Functions and Analytic Continuation.* Die Grundlehren der mathematischen Wissenschaften, Band 207. Springer-Verlag, New York-Heidelberg, 1974.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 18, pp. 13-20, January 2009

ELA
http://math.technion.ac.il/iic/ela

[2] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Corrected and enlarged edition edited by Alan Jeffrey. Academic Press, New York-London-Toronto, Ont., 1980.

[3] Walter Rudin. *Fourier Analysis on Groups*. Interscience, New York, 1962.