Linear Algebra and its Applications 431 (2009) 2154–2159

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

A matrix reverse Hölder inequality

Jean-Christophe Bourin [a], Eun-Young Lee [b,*], Masatoshi Fujii [c], Yuki Seo [d]

[a] Laboratoire de mathématiques, Université de Franche-Comté, 16, route de Gray, 25 000 Besançon, France
[b] Department of Mathematics, Kyungpook National University, Taegu 702-701, Republic of Korea
[c] Department of Mathematics, Osaka Kyoiku University, Kashiwara, Osaka 582, Japan
[d] Faculty of Engineering, Shibaura Institute of Technology, 307 Fukakusa, Minuma-ku, Saitama-City, Saitama 337-8570, Japan

ARTICLE INFO

Article history:
Received 23 December 2008
Accepted 5 July 2009
Available online 12 August 2009

Submitted by C.K. Li

AMS classification:
47A30

Keywords:
Positive linear maps
Matrix geometric mean
Hölder inequality
Reverse inequalities

ABSTRACT

A matrix reverse Hölder inequality is given. This result is a counterpart to the concavity property of matrix weighted geometric means. It extends a scalar inequality due to Gheorghiu and contains several Kantorovich type inequalities.

1. Introduction

Let $\alpha \in [0, 1]$. The weighted geometric mean $a \sharp_\alpha b = a^{1-\alpha} b^\alpha$ of two positive numbers a, b is a concave operation. Letting $\alpha = 1/q$ and $1 - \alpha = 1/p$, this statement is equivalent to the numerical Hölder inequality

$$\sum a_i^{1/p} b_i^{1/q} \leqslant \left(\sum a_i\right)^{1/p} \left(\sum b_i\right)^{1/q} \tag{1}$$

for positive numbers a_i, b_i, $(i = 1, \ldots, m)$. The weighted geometric mean of positive definite matrices shares similar properties. Let us first recall some basic facts about the non-weighted case. Let $A, B, \ldots, Z$ be $n \times n$ matrices, or operators on an n-dimensional space $\mathcal{H}$. For $A, B > 0$ (positive definite), their geometric mean $A \sharp B$ is defined by two quite natural requirements:

* Corresponding author.
E-mail addresses: jcbourin@univ-fcomte.fr (J.-C. Bourin), eylee89@knu.ac.kr (E.-Y. Lee), mfujii@cc.osaka-kyoiku.ac.jp (M. Fujii), yukis@sic.shibaura-it.ac.jp (Y. Seo).

0024-3795/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2009.07.010

1. $AB = BA$ implies $A \sharp B = \sqrt{AB}$,
2. $(X^*AX) \sharp (X^*BX) = X^*(A \sharp B)X$ for any invertible X.

Then, we must have

$$A \sharp B = A^{1/2}(I \sharp A^{-1/2}BA^{-1/2})A^{1/2} = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \tag{2}$$

so that $A \sharp B$ should be solution of the Ricatti equation $ZA^{-1}Z = B, Z > 0$, or equivalently to $ZB^{-1}Z = A$. Hence, $A \sharp B$ can be defined by (2) and $A \sharp B = B \sharp A$. Since $f(t) = t^{1/2}$ is operator monotone, $A \sharp B$ is operator increasing. Remarkable properties of the geometric mean are a maximal characterization by Pusz-Woronovicz [8] and its immediate concavity corollary:

Theorem 1. *Let* $A, B > 0$. *Then* $A \sharp B = \max \left\{ X > 0 \mid \begin{pmatrix} A & X \\ X & B \end{pmatrix} \geqslant 0 \right\}$.

Corollary 1. *The geometric mean* $A \sharp B$ *is concave on pairs of positive definite matrices. Equivalently, for positive definite matrices* $\{A_i\}_{i=1}^m$ *and* $\{B_i\}_{i=1}^m$,

$$\sum A_i \sharp B_i \leqslant \left(\sum A_i\right) \sharp \left(\sum B_i\right).$$

Note that Corollary 1 is a matrix version of the Cauchy–Schwarz inequality. Bhatia's book [3] is a good reference on the matrix geometric mean.

For $A, B > 0$, their weighted geometric mean $A \sharp_\alpha B$ may also be defined by two quite natural requirements:

1. $AB = BA$ implies $A \sharp_\alpha B = A^{1-\alpha}B^\alpha$,
2. $(X^*AX) \sharp_\alpha (X^*BX) = X^*(A \sharp_\alpha B)X$ for any invertible X.

Then, we must have

$$A \sharp_\alpha B = A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}$$

so that $A \sharp_{1/2} B = A \sharp B$. The above definition is coherent with (2) in the sense that expected compatibility conditions like

$$A \sharp_{1/4} B = A \sharp (A \sharp B)$$

hold. Since the geometric mean is operator monotone and concave, it then follows that the weighted geometric means are also monotone and concave. The concavity property yields the matrix version of the Hölder inequality (1):

Corollary 2. *Let* $q > 1$. *For positive definite matrices* $\{A_i\}_{i=1}^m$ *and* $\{B_i\}_{i=1}^m$,

$$\sum A_i \sharp_{1/q} B_i \leqslant \left(\sum A_i\right) \sharp_{1/q} \left(\sum B_i\right).$$

We will show in Section 3 a matrix reverse Hölder inequality companion to Corollary 2. This shall gives us the opportunity to review, in the next section, some elegant reverse results related to the most classical inequalities.

2. Reverse Cauchy–Schwarz inequality

The next results for sums, more generally for positive linear maps $\Phi : M_n(\mathbb{C}) \longrightarrow M_k(\mathbb{C})$ hold [6]:

Proposition 2. *Let* $A_i, B_i > 0, i = 1, \ldots, m$, *with* $cA_i \geqslant B_i \geqslant dA_i$ *and* $c, d > 0$, *and let* $w = c/d$. *Then*

$$\left(\sum A_i\right) \sharp \left(\sum B_i\right) \leqslant \frac{w^{1/4} + w^{-1/4}}{2} \sum A_i \sharp B_i.$$

Theorem 3. *Let $A, B > 0$ with $cA \geqslant B \geqslant dA$ and $c,\ d > 0$, let $w = c/d$ and let Φ be a positive linear map. Then*

$$\Phi(A) \sharp \Phi(B) \leqslant \frac{w^{1/4} + w^{-1/4}}{2} \Phi(A \sharp B).$$

To see that Theorem 3 implies Proposition 2, take $\Phi : M_{nk}(\mathbb{C}) \longrightarrow M_k(\mathbb{C})$ defined by

$$\Phi\left(\begin{pmatrix} X_{1,1} & & \\ & \ddots & \\ & & X_{n,n} \end{pmatrix}\right) = X_{1,1} + \cdots + X_{n,n}$$

and apply Φ to $A = \mathrm{diag}(A_1, \ldots, A_n)$ and $B = \mathrm{diag}(B_1, \ldots, B_n)$. In the same way, Corollary 1 is a special case of Ando's inequality, see [1],

$$\Phi(A \sharp B) \leqslant \Phi(A) \sharp \Phi(B).$$

Since positive linear maps are regarded as a matrix version of integrals, this is a genuine matrix Cauchy–Schwarz inequality. Cassel inequality, a reverse to the Cauchy–Schwarz one, can be stated as follows.

Let Ω be a probability space. Let $f(\omega)$ and $g(\omega)$ be measurable functions on Ω such that $c \geqslant f(\omega)/g(\omega) \geqslant d$ for some $c,\ d > 0$. Then

$$\sqrt{\mathbb{E}(f)\mathbb{E}(g)} \leqslant \frac{(c/d)^{1/4} + (d/c)^{1/4}}{2} \mathbb{E}\left(\sqrt{fg}\right).$$

Here, $\mathbb{E}$ stands for the expectation. Thus Theorem 3 is a matrix Cassel inequality. An interesting remark follows from Cassel inequality. Given real numbers $a_i, i = 1, \ldots, n$, we denote by $a_i^{\downarrow}, i = 1, \ldots, n$, their non-increasing rearrangement. As a reverse result to a basic rearrangement inequality, we have:

Let $a_i, b_i, i = 1, \ldots, m$ be positive numbers such that $r \geqslant a_i/b_i \geqslant s$ for some $r \geqslant s > 0$. Then

$$\sum a_i^{\downarrow} b_i^{\downarrow} \leqslant \frac{r+s}{2\sqrt{rs}} \sum a_i b_i. \tag{3}$$

This statement follows from the Cauchy–Schwarz inequality

$$\sum a_i^{\downarrow} b_i^{\downarrow} \leqslant \left(\sum a_i^2 \sum b_i^2\right)^{1/2}$$

and application of Cassel inequality with $c = r^2, d = s^2$. Cassel inequality also contains a classical inequality: Squaring both of its sides and using $\mathbb{E}^2\left(\sqrt{fg}\right) \leqslant \mathbb{E}(fg)$ we get

$$\mathbb{E}(f)\mathbb{E}(g) \leqslant \frac{(\sqrt{c} + \sqrt{d})^2}{4\sqrt{cd}} \mathbb{E}(fg).$$

The special case, by letting $g(\omega) = 1/f(\omega)$, is the Kantorovich inequality:

Let $f(\omega)$ be a measurable function such that $a \geqslant f(w) \geqslant b > 0$. Then,

$$\mathbb{E}(1/f) \leqslant \frac{(a+b)^2}{4ab} 1/\mathbb{E}(f) \tag{4}$$

and, equivalently,

$$\mathbb{E}(f^2) \leqslant \frac{(a+b)^2}{4ab} \mathbb{E}^2(f). \tag{5}$$

The constant in (4) and its square root occur in several natural matrix inequalities. For instance, given $I \geqslant A \geqslant 0$ and $Z > 0$, we have

$$AZA \leqslant \frac{(a+b)^2}{4ab} Z$$

where a, b are the extremal eigenvalues of Z. Another striking and very recent example is a matrix version of (3):

Proposition 4 [4]. *Let $A,\ B \geqslant 0$ such that $rA \geqslant B \geqslant sA$ for some $r, s > 0$. Then,*

$$\sum \lambda_i(A)\lambda_i(B) \leqslant \frac{r+s}{2\sqrt{rs}}\,\mathrm{Tr}\,AB.$$

Here $\lambda_i(\cdots), i = 1, 2, \ldots$ stand for the eigenvalues arranged in non-increasing order. Thus Proposition 4 is a reverse statement to the famous von-Neumann Trace Inequality. It also improves an earlier result from [6].

3. Reverse Hölder inequality

Let h be a unit vector and let $Z > 0$ with the largest eigenvalue a and the smallest one b. Jensen's inequality for convex functions $t \longrightarrow t^p$ when $p > 1$ or $p < 0$ admits the following reverse statement:

$$(Z^p h, h) \leqslant K(a, b, p)(Zh, h)^p \tag{6}$$

where

$$K(a, b, p) = \frac{a^p b - ab^p}{(p-1)(a-b)} \left(\frac{p-1}{p}\frac{a^p - b^p}{a^p b - ab^p}\right)^p$$

is the (generalized) Kantorovich or Ky-Fan [5] constant. Thus $K(a, b, 2) = K(a, b, -1)$ is the constant occurring in (4) and (5). The constant $K(a, b, p)$ only depends on the condition number $w = a/b$ and can be written as

$$K(w, p) = \frac{w^p - w}{(p-1)(w-1)} \left(\frac{p-1}{p}\frac{w^p - 1}{w^p - w}\right)^p.$$

Since $K(w, p)$ can be defined for all p and satisfies $K(w, p) = K(w, 1-p)$ it may be convenient to use the following symmetric form of the Kantorovich constant:

Let α be any real number and let $\beta = 1 - \alpha$. Then,

$$K(w, \alpha) = \left(\frac{w^\alpha - 1}{\alpha}\right)^\alpha \left(\frac{w^\beta - 1}{\beta}\right)^\beta \frac{w^{\alpha\beta}}{w-1}.$$

The reverse inequality (6) for convex power functions holds for all unital positive linear maps Φ, as noted by Li and Mathias in [7]:

If $Z > 0$ has condition number w, Φ is a unital positive linear map and $p > 1$ or $p < 0$, then

$$\Phi(Z^p) \leqslant K(w, p)\Phi(Z)^p.$$

For concave power functions $t \to t^\alpha$, $1 > \alpha > 0$, the constant $K(w, \alpha)$ plays a similar role:

Let $Z > 0$ with a condition number w and let $1 > \alpha > 0$. Then, for all unital positive linear maps Φ,

$$\Phi(Z)^\alpha \leqslant \frac{1}{K(w, \alpha)}\Phi(Z^\alpha), \tag{7}$$

also see [7].

Now we can state a reverse Hölder inequality involving the Kantorovich constant. It is a reverse statement to Corollary 2.

Proposition 5. *Let $A_i, B_i > 0, i = 1, \ldots, m$ such that $cA_i \geqslant B_i \geqslant dA_i$ for some $0 < d \leqslant c$, and let $w = c/d$. Then, if $0 < \alpha < 1$,*

$$\left(\sum A_i\right) \sharp_\alpha \left(\sum B_i\right) \leqslant \frac{1}{K(w, \alpha)} \sum A_i \sharp_\alpha B_i.$$

When $\alpha = \frac{1}{2}$ in Proposition 5, the constant is (of course)

$$\frac{w^{1/4} + w^{-1/4}}{2}$$

and we get Proposition 2. Exactly as Proposition 2 follows from Theorem 3, we can derive Proposition 5 from:

Theorem 6. *Let A, $B > 0$ with $cA \geqslant B \geqslant dA$ for some c, $d > 0$ and $w = c/d$, and let Φ be a positive linear map. Then*

$$\Phi(A) \sharp \Phi(B) \leqslant \frac{1}{K(w, \alpha)} \Phi(A \sharp B).$$

This result is a reverse of Ando inequality and it contains (7) by letting $B = I$. However, to prove Theorem 6, we need as a lemma the most elementary case of (7):

Lemma 7. *Let $Z > 0$ with extremal eigenvalues a, b and condition number $w = a/b$. Then, for all $0 < \alpha < 1$ and all unit vectors h,*

$$(Zh, h)^{\alpha} \leqslant K(w, \alpha)^{-1} (Z^{\alpha} h, h).$$

(Futhermore this inequality is sharp.)

Proof. Put $\mu = \frac{a^{\alpha} - b^{\alpha}}{a - b}$ and $\nu = \frac{ab^{\alpha} - a^{\alpha} b}{a - b}$. Since $y = t^{\alpha}$ is concave for $0 < \alpha < 1$, then for the line $\mu t + \nu$ crossing t^{α} at $t = b$ and $t = a$, we have

$$K(a, b, \alpha) t^{\alpha} \leqslant \mu t + \nu \leqslant t^{\alpha}$$

on $[b, a]$. In fact, $F(t) = \mu t + \nu - K(a, b, \alpha) t^{\alpha}$ is a convex function with minimum at $t_0 = \frac{\alpha}{1-\alpha} \frac{a^{\alpha} b - ab^{\alpha}}{a^{\alpha} - b^{\alpha}}$ since $F'(t_0) = 0$ (and $t_0 \in [b, a]$). Thus we have

$$K(a, b, \alpha)(Zh, h)^{\alpha} \leqslant \mu(Zh, h) + \nu = ((\mu Z + \nu)h, h) \leqslant (Z^{\alpha} h, h)$$

for all unit vector h, so that we have the required inequality.

Since $\mu t + \nu < t^{\alpha}$ for $t \in (b, a)$, the equality $(Z^{\alpha} h, h) = ((\mu Z + \nu)h, h)$ holds if and only if is h is a linear combination of eigenvalues corresponding to a and b. Moreover, the only zero of F is t_0. Hence the equality $\mu(Zh, h) + \nu = K(a, b, \alpha)(Zh, h)^{\alpha}$ holds if and only if $(Zh, h) = t_0$. $\square$

We will actually use the following variation of Lemma 6:

Lemma 8. *Let $A, B > 0$ with $cA \geqslant B \geqslant dA$ for some scalars $0 < d \leqslant c$ and $w = c/d$. Then, for all vectors h and all $0 < \alpha < 1$,*

$$(A \sharp_{\alpha} Bh, h) \leqslant (Ah, h)^{1-\alpha} (Bh, h)^{\alpha} \leqslant K(w, \alpha)^{-1} (A \sharp_{\alpha} Bh, h)$$

Proof. Let $Z > 0$ with $aI \geqslant Z \geqslant bI$ and $w = a/b$. By concavity of $t \longrightarrow t^{\alpha}$ and Lemma 7, we have

$$K(w, \alpha)(Zx, x)^{\alpha} (x, x)^{1-\alpha} \leqslant (Z^{\alpha} x, x) \leqslant (Zx, x)^{\alpha} (x, x)^{1-\alpha}$$

for every x. Replacing Z and x by $A^{-1/2} B A^{-1/2}$ and $A^{1/2} h$ respectively gives the lemma. $\square$

We turn to the proof of Theorem 6.

Proof. Suppose that Φ is a vector state: $\Phi(A) = (Ah, h)$ for a vector h. By Lemma 8, it follows that

$$\begin{aligned}
\Phi(A) \sharp_{\alpha} \Phi(B) &= (Ah, h)^{1-\alpha} (Bh, h)^{\alpha} \\
&\leqslant K(w, \alpha)^{-1} (A \sharp_{\alpha} Bh, h) \\
&= K(w, \alpha)^{-1} \Phi(A \sharp_{\alpha} B).
\end{aligned}$$

Now consider the case of a general positive linear map Φ. Let h be any vector. Then, by the first inequality of Lemma 8, it follows that

$$(\Phi(A) \sharp_\alpha \Phi(B)h, h) \leqslant (\Phi(A)h, h)^{1-\alpha}(\Phi(B)h, h)^\alpha = \Psi(A)^{1-\alpha}\Psi(B)^\alpha,$$

where Ψ is defined by $\Psi(A) = (\Phi(A)h, h)$.

Since Ψ is a positive linear functional on $M_n(\mathbb{C})$, there exists $X \geqslant 0$ such that $\Phi(A) = \operatorname{Tr} AX$. Hence, if $\pi(A) : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is the left multiplication by A, we can write

$$\Psi(A) = \langle h, \pi(A)h \rangle,$$

where the inner product $\langle \cdot, \cdot \rangle$ is the canonical inner product on $M_n(\mathbb{C})$ and $h = X^{1/2}$. Since $cA \geqslant B \geqslant dA$ implies $c\pi(A) \geqslant \pi(B) \geqslant d\pi(A)$, the second inequality of Lemma 8 yields

$$\Psi(A)^{1-\alpha}\Psi(B)^\alpha \leqslant K(w, \alpha)^{-1}\Psi(A \sharp_\alpha B).$$

Combining with the previous inequality we have

$$(\Phi(A) \sharp_\alpha \Phi(B)h, h) \leqslant K(w, \alpha)^{-1}(\Phi(A \sharp_\alpha B)h, h)$$

for every h. $\square$

Proposition 5 for scalars can be stated as a reverse numerical Hölder inequality as follows:

Corollary 3. *Let $\{a_k\}_{k=1}^n$ and $\{b_k\}_{k=1}^n$ be n-tuples of positive numbers and let $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. If M, m are two positive numbers such that $0 < m \leqslant a_k/b_k^{q/p} \leqslant M (k = 1, \ldots, n)$, then*

$$\left(\sum a_k^p\right)^{1/p}\left(\sum b_k^q\right)^{1/q} \leqslant \lambda \sum_{k=1}^n a_k b_k$$

where

$$\lambda = \frac{M^p - m^p}{p^{1/p}q^{1/q}(M-m)^{1/p}(M^p m - m^p M)^{1/q}}.$$

It turns out that this inequality goes back to a paper written in French in 1933 by Gheorgiu. This result, without proof, and its reference are in the survey book [9, p. 124]. Hence, Cassel's inequality (1951), Ky Fan constant and reverse inequalities (1966) were already known or implicit in Gheorgiu's paper.

In this paper, the "swandwich condition" for positive operators,

$$cA \geqslant B \geqslant dA$$

is the key for all statements. This condition is also the natural one for several forthcoming, related results. For instance, interesting rearrangement inequalities for unitarily invariant norms are considered in [2,4].

References

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard products, Linear Algebra Appl. 26 (1979) 203–241.
[2] T. Ando, Lee's approach revisited, unpublished manuscript.
[3] R. Bhatia, Positive Definite Matrices, Princeton University Press, Princeton, 2007.
[4] J.-C. Bourin, Eun-Young Lee, Some reverse matrix inequalities, unpublished manuscript.
[5] Ky Fan, Some Matrix Inequalities, vol. 29, Abh. Math. Sem. Univ., Hamburg, 1966, pp. 185–196.
[6] Eun-Young Lee, A matrix reverse Cauchy–Schwarz inequality, Linear Algebra Appl. 430 (2009) 805–810.
[7] C.-K. Li, R. Mathias, Matrix inequalities involving a positive linear map, Linear and Multilinear Algebra 41 (1996) 221–231.
[8] W. Pusz, S.L. Woronowicz, Functional calculus for sesquilinear forms and the purification maps, Rep. Math. Phys. 8 (1975) 159–170.
[9] D.S. Mitrinović, J.E. Pečarić, A.M. Fink, Classical and New Inequalities in Analysis, Kluwer Academic Publishers.