数理解析研究所講究録
第2033巻 2017年 135-141

SOME NORM INEQUALITIES FOR MATRIX MEANS

TAKEAKI YAMAZAKI

ABSTRACT. This report is based on [3]. Inequalities for unitarily invariant norms of power means of positive definite matrices are presented. Also Heron and Heinz means are treated.

1. INTRODUCTION

Let P_n be a set of all positive definite n–by–n matrices.

Definition 1 (Matrix mean, [6]). For $A, B \in P_n$, $\mathfrak{M}(A,B)$ is called matrix mean if it satisfies the following conditions:

(i) $A \leq C$ and $B \leq D$ imply

$$\mathfrak{M}(A,B) \leq \mathfrak{M}(C,D),$$

(ii) for $C = C^*$,

$$C\mathfrak{M}(A,B)C \leq \mathfrak{M}(CAC, CBC),$$

(iii) if $A_n \downarrow A$ and $B_n \downarrow B$, then

$$\mathfrak{M}(A_n, B_n) \downarrow \mathfrak{M}(A,B),$$

(iv) $\mathfrak{M}(I,I) = I$.

Matrix means can be characterized by matrix monotone functions as follows:

Theorem A ([6]). *For each matrix mean $\mathfrak{M}$, there exists a unique matrix monotone function $f : \mathbb{R}^+ \to \mathbb{R}^+$ such that*

$$f(x)I = \mathfrak{M}(I, xI) \quad (x \in \mathbb{R}^+)$$

and for $A, B \in P_n$, the formula

$$\mathfrak{M}(A,B) = A^{\frac{1}{2}} f(A^{\frac{-1}{2}} B A^{\frac{-1}{2}}) A^{\frac{1}{2}}$$

holds. A function f is called the representing function of a matrix mean $\mathfrak{M}$.

The weighted geometric mean of $A, B \in P_n$ is a typical example of matrix means which is defined by

$$A\sharp_\lambda B = A^{\frac{1}{2}}(A^{\frac{-1}{2}} B A^{\frac{-1}{2}})^\lambda A^{\frac{1}{2}}. \tag{1.1}$$

Especially, if $\lambda = \frac{1}{2}$, then $A\sharp B$ denotes $A\sharp_{1/2}B$. If A and B commute with each other, then

$$A\sharp_\lambda B = A^{1-\lambda}B^\lambda = \exp[(1-\lambda)\log A + \lambda \log B].$$

2010 *Mathematics Subject Classification.* Primary 47A30. Secondary 47A64.
Key words and phrases. Positive definite matrix; matrix mean; matrix monotone function; power mean; Heron mean; Heinz mean.

For the geometric mean, the following norm inequality is very famous.

Theorem B ([1]). *For $A, B \in P_n$,*

$$|||A\sharp B||| \leq |||\exp\left(\frac{\log A + \log B}{2}\right)|||$$

holds for any unitarily invariant norm $|||\cdot|||$.

In the recent years, the weighted geometric mean has been extended to the means of n-matrices. There are some definition of geometric means of n-matrices. But the following Karcher mean is known as the best one of geometric means.

Definition 2 ([7]). For $\mathbb{A} = (A_1, ..., A_m) \in P_n^m$ and $\omega = (w_1, ..., w_m) \in (0,1)^m$, s.t., $\sum_{i=1}^m w_i = 1$, the Karcher mean $\Lambda(\omega; \mathbb{A})$ is defined by unique solution $X \in P_n$ of the following matrix equation;

$$\sum_{i=1}^{m} w_i \log X^{\frac{-1}{2}} A_i X^{\frac{-1}{2}} = 0.$$

The representing function of the Karcher mean of two matrices is given by the matrix equation:

$$(1-\lambda)\log X^{-1} + \lambda t X^{-1} = 0$$

since $f(x)I = X = \Lambda(1-\lambda, \lambda; I, xI)$. It is equivalent to $f(x)I = X = x^\lambda$. Hence the Karcher mean of $A, B \in P_n$ is $A\sharp_\lambda B$. Moreover, If $\{A_1, ..., A_m\}$ is commutative, then

$$\Lambda(\omega; \mathbb{A}) = A_1^{w_1} \cdots A_m^{w_m} = \exp\left[\sum_{i=1}^{m} w_i \log A_i\right].$$

For the Karcher mean, we have an extension of Theorem B as follows.

Theorem C ([5]). *For $\mathbb{A} = (A_1, ..., A_m) \in P_n^m$ and $\omega = (w_1, ..., w_m) \in (0,1)^m$, s.t., $\sum_{i=1}^m w_i = 1$,*

$$|||\Lambda(\omega; \mathbb{A})||| \leq |||\exp\left[\sum_{i=1}^{m} w_i \log A_i\right]|||$$

holds for any unitarily invariant norm $|||\cdot|||$.

Moreover the Karcher mean is extended to the following power mean.

Definition 3 ([8]). For $\mathbb{A} = (A_1, ..., A_m) \in P_n^m$ and $\omega = (w_1, ..., w_m) \in (0,1)^m$, s.t., $\sum_{i=1}^m w_i = 1$, and $t \in [-1, 1] \setminus \{0\}$, the power mean $P_t(\omega; \mathbb{A})$ is defined by the unique solution $X \in P_n$ of the following matrix equation;

$$\sum_{i=1}^{m} w_i X \sharp_t A_i = X.$$

Power mean interpolates the arithmetic-Karcher-harmonic means, in fact, we have the arithmetic, Karcher and harmonic means by letting $t = 1$, $t \to 0$ and $t = -1$, respectively. If $\omega = (\frac{1}{m}, ..., \frac{1}{m})$, then $P_t(\mathbb{A})$ denotes $P_t(\omega; \mathbb{A})$, simply. For the 2-matrices case, the representing function of the power mean is a unique solution of the following equation.

$$(1-\lambda)X\sharp_t I + \lambda X\sharp_t(xI) = X,$$

since $f(x)I = X = P_t(1-\lambda, \lambda; I, xI)$. It is equivalent to

$$X^{1-t}\left[(1-\lambda)I + \lambda x^t I\right] = X.$$

Therefore

$$f(x)I = X = \left[(1-\lambda)I + \lambda x^t I\right]^{\frac{1}{t}}.$$

Hence for $A, B \in P_n$, $\lambda \in [0,1]$ and $t \in [-1,1] \setminus \{0\}$,

$$P_t(1-\lambda, \lambda; A, B) = A^{\frac{1}{2}}\left[(1-\lambda) + \lambda(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^t\right]^{\frac{1}{t}} A^{\frac{1}{2}}.$$

If $\{A_1, ..., A_m\}$ is commutative, then

$$P_t(\omega; \mathbb{A}) = \left(\sum_{i=1}^{m} w_i A_i^t\right)^{\frac{1}{t}}.$$

One might expect that the power mean also satisfies the similar norm inequality to Theorem C. However, we have shown an inequality for the spectral norm case only.

Theorem D ([9]). *For $\mathbb{A} = (A_1, ..., A_m) \in P_n^m$ and $\omega = (w_1, ..., w_m) \in (0,1)^m$, s.t., $\sum_{i=1}^m w_i = 1$, and $t \in [0,1]$,*

$$\|P_t(\omega; \mathbb{A})\| \leq \left\| \left(\sum_{i=1}^{m} w_i A_i^t\right)^{\frac{1}{t}} \right\|$$

holds for the spectral norm $\|\cdot\|$.

Hence, our problem is as follows:

Problem. *For $\mathbb{A} = (A_1, ..., A_m) \in P_n^m$ and $\omega = (w_1, ..., w_m) \in (0,1)^m$, s.t., $\sum_{i=1}^m w_i = 1$, and $t \in [0,1]$, does*

$$|||P_t(\omega; \mathbb{A})||| \leq \left|\left|\left| \left(\sum_{i=1}^{m} w_i A_i^t\right)^{\frac{1}{t}} \right|\right|\right|$$

hold for any unitarily invariant norm $|||\cdot|||$?

In this report, we shall treat only Schatten p-norms for discussing the above problem. Let $A \in M_n$ and $s_1(A), ..., s_n(A)$ be the singular values of A, i.e., the eigenvalues of $|A|$ such that

$$s_1(A) \geq \cdots \geq s_n(A).$$

For $1 < p$, Shatten p-norm of A is defined by

$$\|A\|_p := \left(\sum_{i=1}^{n} s_i(A)^p\right)^{\frac{1}{p}}.$$

Every Shatten p-norm is a unitarily invariant norm for $1 \leq p \leq \infty$. Especially, if $A \in P_n$, then $\|A\|_1 = \mathrm{tr}(A)$, $\|A\|_2 = [\mathrm{tr}(A^2)]^{\frac{1}{2}}$ and $\|A\|_\infty = \|A\|$ (spectral norm).

2. The power mean for 2-matrices

In this section, we shall discuss the problem in the case of 2-matrices case.

Theorem 1. *For $A, B \in P_n$,*

$$\|P_{1/2}(A,B)\|_p \leq \| \left(\frac{A^{\frac{1}{2}} + B^{\frac{1}{2}}}{2} \right)^2 \|_p$$

holds for $p = 1, 2, \infty$.

To prove Theorem 1, we will use the Furuta inequality.

Theorem E (Furuta inequality, [4]). *Let $A, B \in P_n$. If $A \geq B \geq 0$, then for each $r \geq 0$,*

$$(A^{\frac{r}{2}} A^p A^{\frac{r}{2}})^{\frac{1}{q}} \geq (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \quad \text{and} \quad (B^{\frac{r}{2}} A^p B^{\frac{r}{2}})^{\frac{1}{q}} \geq (B^{\frac{r}{2}} B^p B^{\frac{r}{2}})^{\frac{1}{q}}$$

hold for $p \geq 0$, $q \geq 1$ with $(1+r)q \geq p + r$.

Poof of Theorem 1. The case $p = \infty$ has been already shown in [9].

The case $p = 1$. Since

$$P_{1/2}(A,B) = A^{\frac{1}{2}} \left[\frac{I + (A^{\frac{-1}{2}} B A^{\frac{-1}{2}})^{\frac{1}{2}}}{2} \right]^2 A^{\frac{1}{2}} = \frac{1}{4}(A + B + 2A\sharp B)$$

and

$$\left(\frac{A^{\frac{1}{2}} + B^{\frac{1}{2}}}{2} \right)^2 = \frac{1}{4}(A + B + A^{\frac{1}{2}} B^{\frac{1}{2}} + B^{\frac{1}{2}} A^{\frac{1}{2}})$$

hold, it is enough to show

$$\operatorname{tr}(A + B + 2A\sharp B) \leq \operatorname{tr}(A + B + A^{\frac{1}{2}} B^{\frac{1}{2}} + B^{\frac{1}{2}} A^{\frac{1}{2}}).$$

It is equivalent to

$$\operatorname{tr}(A\sharp B) \leq \frac{1}{2} \operatorname{tr}(A^{\frac{1}{2}} B^{\frac{1}{2}} + B^{\frac{1}{2}} A^{\frac{1}{2}}).$$

It has been already shown in [2]. Therefore, the case $p = 1$ is proven.

The case $p = 2$. By the similar argument to the case $p = 1$, it is enough to show

$$\operatorname{tr}\left((A + B + 2A\sharp B)^2\right) \leq \operatorname{tr}\left(\left(A + B + A^{\frac{1}{2}} B^{\frac{1}{2}} + B^{\frac{1}{2}} A^{\frac{1}{2}}\right)^2\right). \tag{2.1}$$

We can calculate that

$$\operatorname{tr}\left((A + B + 2A\sharp B)^2\right) = \operatorname{tr}\left(A^2 + 2AB + B^2 + 4A(A\sharp B) + 4B(A\sharp B) + 4(A\sharp B)^2\right)$$

and

$$\operatorname{tr}\left(\left(A + B + A^{\frac{1}{2}} B^{\frac{1}{2}} + B^{\frac{1}{2}} A^{\frac{1}{2}}\right)^2\right) = \operatorname{tr}\left(A^2 + 4AB + B^2 + 4A^{\frac{3}{2}} B^{\frac{1}{2}} + 4A^{\frac{1}{2}} B^{\frac{3}{2}} + 2(A^{\frac{1}{2}} B^{\frac{1}{2}})^2\right).$$

Then (2.1) is equivalent to the following trace inequality.

$$\mathrm{tr}\left(2A(A\sharp B)+2B(A\sharp B)+2(A\sharp B)^2\right) \le \mathrm{tr}\left(AB+2A^{\frac{3}{2}}B^{\frac{1}{2}}+2A^{\frac{1}{2}}B^{\frac{3}{2}}+(A^{\frac{1}{2}}B^{\frac{1}{2}})^2\right).$$

Firstly, we shall show

$$\mathrm{tr}\left((A\sharp B)^2\right) \le \mathrm{tr}\left((A^{\frac{1}{2}}B^{\frac{1}{2}})^2\right) \le \mathrm{tr}\,(AB).$$

The first inequality follows from $|||A\sharp B||| \le |||A^{\frac{1}{4}}B^{\frac{1}{2}}A^{\frac{1}{4}}|||$ for any unitarily invariant norm in [2]. In fact,

$$\|A^{\frac{1}{4}}B^{\frac{1}{2}}A^{\frac{1}{4}}\|_2^2 = \mathrm{tr}\left((A^{\frac{1}{4}}B^{\frac{1}{2}}A^{\frac{1}{4}})^2\right) = \mathrm{tr}\left((A^{\frac{1}{2}}B^{\frac{1}{2}})^2\right)$$

holds. The second inequality follows from the Lieb-Thirring inequality, i.e.,

$$\mathrm{tr}((AB)^m) \le \mathrm{tr}(A^m B^m).$$

Next, we shall show $\mathrm{tr}(A(A\sharp B)) \le \mathrm{tr}(A^{\frac{3}{2}}B^{\frac{1}{2}})$. To prove this, we shall show

$$|||A^{\frac{1}{2}}(A\sharp B)A^{\frac{1}{2}}||| \le |||A^{\frac{3}{4}}B^{\frac{1}{2}}A^{\frac{3}{4}}|||$$

for any unitarily invariant norm. By considering the untisymmetric tensor technique, it is enough to show

$$A^{\frac{3}{4}}B^{\frac{1}{2}}A^{\frac{3}{4}} \le I \quad \Longrightarrow \quad A^{\frac{1}{2}}(A\sharp B)A^{\frac{1}{2}} \le I.$$

It is equivalent to

$$B^{\frac{1}{2}} \le A^{\frac{-3}{2}} \quad \Longrightarrow \quad (A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^{\frac{1}{2}} \le A^{-2}.$$

It follows from Theorem E. $\mathrm{tr}(B(A\sharp B)) \le \mathrm{tr}(A^{\frac{1}{2}}B^{\frac{3}{2}})$ can be shown by the same way since $A\sharp B = B\sharp A$ holds. Therefor the proof is completed. □

3. The Heron and Heinz means

In this section, we shall discuss similar norm inequalities to Theorem 1 for the Heron and Heinz means. Because these means have similar forms to the power mean $4P_{1/2}(A,B) = A+B+2A\sharp B$.

Definition 4 (Heron and Heinz means). Let A, $B \in P_n$ and $t \in [0,1]$. Then the Heron and Heinz means of A and B are defined as follows:

(i) Heron mean: $(1-t)\dfrac{A+B}{2} + tA\sharp B$,

(ii) Heinz mean: $\dfrac{A\sharp_t B + B\sharp_t A}{2}$.

If A and B commute with each other, we have

$$(1-t)\frac{A+B}{2} + tA\sharp B = (1-t)\frac{A+B}{2} + t\sqrt{AB} = (1-t)\frac{A+B}{2} + t\frac{A^{\frac{1}{2}}B^{\frac{1}{2}}+B^{\frac{1}{2}}A^{\frac{1}{2}}}{2}$$

and

$$\frac{A\sharp_t B + B\sharp_t A}{2} = \frac{A^{\frac{1}{2}}B^{\frac{1}{2}}+B^{\frac{1}{2}}A^{\frac{1}{2}}}{2}.$$

By the similar way to the proof of Theorem 1, we have the following result.

Theorem 2. *For $A, B \in P_n$ and $t \in [0,1]$,*

(i) $\left\|(1-t)\dfrac{A+B}{2} + tA\sharp B\right\|_p \leq \left\|(1-t)\dfrac{A+B}{2} + t\dfrac{A^{\frac{1}{2}}B^{\frac{1}{2}} + B^{\frac{1}{2}}A^{\frac{1}{2}}}{2}\right\|_p$,

(ii) $\|A\sharp_t B + B\sharp_t A\|_p \leq \|A^{1-t}B^t + A^tB^{1-t}\|_p$

hold for $p = 1, 2$.

4. Trace inequality for the power mean of several variables

In this section, we shall give a solution of the problem for the trace norm.

Theorem 3. *For $\mathbb{A} = (A_1, ..., A_m) \in P_n^m$ and $t \in (0,1]$,*

$$\|P_t(\mathbb{A})\|_p \leq \| \left(\frac{1}{m}\sum_{i=1}^m A_i^t\right)^{\frac{1}{t}} \|_p$$

hold for $p = 1, \infty$.

Proof. The case $p = \infty$ has been already shown in [9].

The case $p = 1$. Let $X = P_t(\mathbb{A})$. Then X satisfies

$$X = \frac{1}{m}\sum_{i=1}^m X\sharp_t A_i.$$

We have

$$\begin{aligned}
\mathrm{tr}(X) &= \mathrm{tr}\left(\frac{1}{m}\sum_{i=1}^m X\sharp_t A_i\right) \\
&= \frac{1}{m}\sum_{i=1}^m \mathrm{tr}(X\sharp_t A_i) \\
&\leq \frac{1}{m}\sum_{i=1}^m \mathrm{tr}(X^{1-t}A_i^t) \\
&= \mathrm{tr}\left(X^{1-t}\left[\frac{1}{m}\sum_{i=1}^m A_i^t\right]^{\frac{t}{t}}\right) \\
&\leq \mathrm{tr}\left((1-t)X + t\left[\frac{1}{m}\sum_{i=1}^m A_i^t\right]^{\frac{1}{t}}\right),
\end{aligned}$$

where the inequalities are obtained by

$$\mathrm{tr}(A\sharp_t B) \leq \mathrm{tr}(A^{1-t}B^t) \leq \mathrm{tr}(1-t)A + tB)$$

in [2]. Hence we have

$$\mathrm{tr}(P_t(\mathbb{A})) \leq \mathrm{tr}\left(\left[\frac{1}{m}\sum_{i=1}^m A_i^t\right]^{\frac{1}{t}}\right).$$

□

References

[1] T. Ando, F. Hiai, Log majorization and complementary Golden-Thompson type inequalities, Linear Algebra Appl. 197/198 (1994) 113–131.

[2] R. Bhatia, P. Grover, Norm inequalities related to the matrix geometric mean, Linear Algebra Appl. 473 (2012) 726–733.

[3] . R. Bhatia, Y. Lim, T. Yamazaki, Some norm inequalities for matrix means, Linear Algebra Appl. 501 (2016), 112–122.

[4] T. Furuta, $A \geq B \geq O$ assures $(B^r A^p B^r)^{1/q} \geq B^{(p+2r)/q}$ for $r \geq 0$, $p \geq 0$, $q \geq 1$ with $(1+2r)q \geq p+2r$, Proc. Amer. Math. Soc. 101 (1987) 85–88.

[5] F. Hiai, D. Petz, Riemannian metrics on positive definite matrices related to means II, Linear Algebra Appl. 436 (2012) 2117–2136.

[6] F. Kubo, T. Ando, Means of positive linear operators, Math. Ann. 246 (1980) 205–224.

[7] J. Lawson and Y. Lim, *Karcher means and Karcher equations of positive definite operators*, Trans. Amer. Math. Soc. Ser. B 1, (2014), 1–22.

[8] Y. Lim, M. Pálfia, The matrix power means and the Karcher mean, J. Funct. Anal. 262 (2012) 1498–1514.

[9] Y. Lim, T. Yamazaki, On some inequalities for the matrix power and Karcher means, Linear Algebra Appl. 438 (2013) 325–346.

Department of Electrical, Electronic and Computer Engineering, Toyo University, Kawagoe-Shi, Saitama, 350-8585, Japan.
E-mail address: t-yamazaki@toyo.jp