MOORE-PENROSE INVERSE AND OPERATOR MEAN

Jun Ichi Fujii

Dedicated to the memory of the late professor Takayuki Furuta

Received June 27, 2017 ; revised July 16, 2017

Abstract. Recently the geometric operator mean is extended to the multi-variable one; the Karcher mean. Including these multivarible means, we discuss a construction method by the Moore-Penrose inverse. The key concept is the orthogonality of operator means.

1 Introduction. Let m be *an operator mean* in the sense of Kubo-Ando [12] which is defined by a positive operator monotone function f_{m} on the half interval $(0,\infty)$ with $f_{\mathrm{m}}(1)=1$;

$$A \,\mathrm{m}\, B = A^{\frac{1}{2}} f_{\mathrm{m}}\left(A^{-\frac{1}{2}} B A^{-\frac{1}{2}}\right) A^{\frac{1}{2}}$$

for positive invertible operators A and B on a Hilbert space. Thus every operator mean can be constructed by a numerical function $f_{\mathrm{m}}(x) = 1 \,\mathrm{m}\, x$ which is called the representing function of m. Among common properties for operator means, we pay attention to the *orthogonality*:

$$(A_1 \oplus A_2) \,\mathrm{m}\, (B_1 \oplus B_2) = (A_1 \,\mathrm{m}\, B_1) \oplus (A_2 \,\mathrm{m}\, B_2)$$

and the *transformer inequality*:

$$T^*(A \,\mathrm{m}\, B)T \le (T^*AT) \,\mathrm{m}\, (T^*BT).$$

Recall that the Karcher mean $X = \mathsf{G}(\omega_j; A_j)$ for invertible $A_j \ge 0$ with a weight $\{\omega_j\}$ is defined as a unique solution of the *Karcher equation* [11, 13, 14]:

$$\sum_j \omega_j S(X|A_j) = \sum_j \omega_j X^{\frac{1}{2}} \log\left(X^{-\frac{1}{2}} A_j X^{-\frac{1}{2}}\right) X^{\frac{1}{2}} = 0.$$

We extend it to non-invertible case in [11], which is an extension of the weighted geometric mean. Moreover in [11], we extended such multi-variable *operator mean* $\mathrm{M}(A_j) = \mathrm{M}(\omega_j; A_1, ..., A_n)$ including the Karcher mean: Define an (n-variable) *general operator mean* $\mathrm{M}(\omega_j; A_j)$ as an n-ary operation on positive invertible operators on $\mathcal{H}$ satisfying the following properties where each weight ω_j is assumed to be positive here:

(M1) **transformer equality**: $T^*\,\mathrm{M}(\omega_j; A_j)T = \mathrm{M}(\omega_j; T^*A_jT)$ for all invertible T.

(M1') **homogeneity**: $\mathrm{M}(\omega_j; tA_j) = t\,\mathrm{M}(\omega_j; A_j)$ for $t > 0$.

(M2) **normalization**: $\mathrm{M}(\omega_j; A) = A$.

(M3) **monotonicity**: $A_j \le B_j$ implies $\mathrm{M}(\omega_j; A_j) \le \mathrm{M}(\omega_j; B_j)$.

2000 *Mathematics Subject Classification.* 47A64, 47A63 ,15A09.
Key words and phrases. operator mean, kernel inclusion, transformer inequality, generalized inverse.

(M4) **sub-additivity**: $\mathrm{M}(\omega_j; A_j + B_j) \geq \mathrm{M}(\omega_j; A_j) + \mathrm{M}(\omega_j; B_j)$.

(M5) **adjoint sub-additivity**: $\mathrm{M}(\omega_j; A_j : B_j) \leq \mathrm{M}(\omega_j; A_j) : \mathrm{M}(\omega_j; B_j)$.

(M6) **orthogonality**: $\mathrm{M}(\omega_j; \bigoplus_m A_j^{(m)}) = \bigoplus_m \mathrm{M}(\omega_j; A_j^{(m)})$.

Here : stands for the parallel sum defined by

$$A : B = \left(A^{-1} + B^{-1}\right)^{-1}.$$

In addition, we can define

$$\mathrm{M}(\omega_j; A_j) = \operatorname*{s-lim}_{\varepsilon \to 0} \mathrm{M}(\omega_j; (A_j + \varepsilon))$$

for (non-invertible) positive operators A_j where the above properties preserve, which includes our extended Karcher mean. Also for $t \in [0,1]$, note that

(M7) **joint concavity:** $\mathrm{M}(\omega_j; (1-t)A_j + tB_j) \leq (1-t)\,\mathrm{M}(\omega_j; A_j) + t\,\mathrm{M}(\omega_j; B_j)$

follows from the sub-additivity and homogeneity. Here we pay attention to the orthogonality for operator means as in the below.

On the other hand, for the parallel sum (the half of the harmonic mean), rephrasing them into the harmonic mean, we have

$$A \,\mathrm{h}\, B = A \left(\frac{A+B}{2}\right)^{\dagger} B$$

if $A+B$ has the generalized inverse [1]. Incidentally the Moore-Penrose generalized inverse † for operators was discussed in [9, 15]: It is known that if $\operatorname{ran} X$ is closed, then $\operatorname{ran} X^*$, $\operatorname{ran} XX^*$ and $\operatorname{ran} X^*X$ are also closed, and $(X^*X)^{\dagger} = \left(X^*X\big|_{\operatorname{ran} X^*}\right)^{-1} \oplus 0_{(\operatorname{ran} X^*)^{\perp}}$ and $X^{\dagger} = (X^*X)^{\dagger}X^* = X^*(XX^*)^{\dagger}$.

In this note, we observe operator means from the viewpoint of the generalized inverse, which includes our extended version of the Karcher mean. We discuss the constructing formulae for operator means using the Moore-Penrose inverses if they exist:

$$A^{\frac{1}{2}}(I \,\mathrm{m}\, A^{\dagger\frac{1}{2}} B A^{\dagger\frac{1}{2}}) A^{\frac{1}{2}} \qquad \text{or} \qquad B^{\frac{1}{2}}(B^{\dagger\frac{1}{2}} A B^{\dagger\frac{1}{2}} \,\mathrm{m}\, I) B^{\frac{1}{2}}.$$

Our equality condition [6] for the transformer iequality shows that it represents the operator mean $A \,\mathrm{m}\, B$ if $\ker A \subset \ker B$ or $\ker A \supset \ker B$ respectively. We also show that they are not less than the original one if the kernel of the mean $A \,\mathrm{m}\, B$ includes those for A and B.

2 Transformer equality.

In [6], we gave an equality condition for transformer inequality for certain means:

Theorem F . *If* $\ker T^* \subset \ker A \cap \ker B$, *then* $T^*(A \,\mathrm{m}\, B)T = (T^*AT) \,\mathrm{m}\, (T^*BT)$ *for an operator mean* m.

This assures the Izumino construction of operator means: Let $R = (A+B)^{\frac{1}{2}}$, then, there exist the derivatives D and E with $A^{\frac{1}{2}} = RD$ and $B^{\frac{1}{2}} = RE$ by the range inclusion theorem in [3, 4]. So we have $D^*D + E^*E = I_{\overline{\operatorname{ran} R}}$ and an operator mean is reduced into the commutative case [6]:

$$A \,\mathrm{m}\, B = R(D^*D \,\mathrm{m}\, E^*E)R,$$

which is a space-free version of the Pusz-Woronowics means [16, 17].

But the original proof of the above was based on the integral representation of operator means, so that we cannot extend the equality in Theorem F to multi-variable means. Under the closedness of the ranges for operators, we show the equality for our extended (multi-variable) operator means including the Karcher operator mean:

Theorem 1. *Let* $\mathrm{M}(A_j) = \mathrm{M}(\omega_j; A_1, .., A_n)$ *be an operator mean* (*satisfying the orthogonality*). *If an operator* T *on* H *satisfies* $\ker T^* \subset \bigcap_j \ker A_j$ *and* $\operatorname{ran} T$ *is closed, then the transformer equality holds:*

$$T^* \mathrm{M}(A_j)T = \mathrm{M}(T^* A_j T).$$

Proof. Note that $\operatorname{ran} T^*$ is also closed. Recall that $P = TT^\dagger$ and $Q = T^\dagger T$ are projections onto $\operatorname{ran} T$ and $\operatorname{ran} T^*$ respectively, see e.g. [9, 15]. By the assumption $\operatorname{ran} T^\perp = \ker T^* \subset \ker A_j$, we have $PA_jP = A_j$ for all j. Also $QT^*A_jTQ = T^*A_jT$ implies $QM(T^*A_jT)Q = M(T^*A_jT)$ for all j by the orthogonality. Then we have

$$\begin{aligned} T^* \mathrm{M}(A_j)T \le \mathrm{M}(T^* A_j T) &= Q\,\mathrm{M}(TA_jT)Q = T^*T^{\dagger *}\,\mathrm{M}(T^*A_jT)T^\dagger T \\ &\le T^*\,\mathrm{M}\big(T^{\dagger *}T^*A_jTT^\dagger\big)T = T^*\,\mathrm{M}(PA_jP)T = T^*\,\mathrm{M}(A_j)T, \end{aligned}$$

which shows the required equality. □

Remark. The assumption $\ker T^* \subset \bigcap_j \ker A_j$ in the above is equivalent to $\operatorname{ran} T \supset \bigvee_j \operatorname{ran} A_j$ under the closedness of operators.

Corollary 2. *Let* m *be an* (2*-variable*) *operator mean. If* $\ker A \subset \ker B$ *and* $\operatorname{ran} A$ *is closed, then*

$$A\,\mathrm{m}\,B = A^{\frac12}(I\,\mathrm{m}\,A^{\dagger\frac12}BA^{\dagger\frac12})A^{\frac12} = A^{\frac12}f_{\mathrm{m}}(A^{\dagger\frac12}BA^{\dagger\frac12})A^{\frac12}.$$

Remark. Contrastively we have

$$A\,\mathrm{m}\,B = B^{\frac12}(B^{\dagger\frac12}AB^{\dagger\frac12}\,\mathrm{m}\,I)B^{\frac12}$$

if $\ker B \subset \ker A$ and $\operatorname{ran} B$ is closed.

3 Means satisfying the kernel condition. Initiated by [5], we observe the kernel conditions for operator means, see also [7, 8]:

$$\ker A\,\mathrm{m}\,B \supset \ker A \vee \ker B \tag{1}$$

if and only if $1\,\mathrm{m}\,0 = 0\,\mathrm{m}\,1 = 0$. The geometric or harmonic mean satisfies this, while the arithmetic mean does not. In [11], we showed $\ker A\#B = \ker A \vee \ker B$. Moreover, based on this property, we introduced the Karcher mean $X = \mathsf{G}(\omega_j; A_j)$ for non-invertible positive operators A_j under this kernel condition: $\ker X = \vee_j \ker A_j$.

For invertible operators, we have two expressions:

$$A\,\mathrm{m}\,B = A^{\frac12}(I\,\mathrm{m}\,A^{-\frac12}BA^{-\frac12})A^{\frac12} = B^{\frac12}(B^{-\frac12}AB^{-\frac12}\,\mathrm{m}\,I)B^{\frac12}. \tag{2}$$

Then we discuss the means where the inverses in (2) are exchanged into the Moore-Penrose inverse:

Theorem 3. *Let* m *be an operator mean satisfying the above kernel condition* (1). *If* $\operatorname{ran} A$ (*resp.* $\operatorname{ran} B$) *is closed, then*

$$A\,\mathrm{m}\,B \le A^{\frac12}(I\,\mathrm{m}\,A^{\dagger\frac12}BA^{\dagger\frac12})A^{\frac12} \qquad \Big(\textit{resp.} \;\; \le B^{\frac12}(B^{\dagger\frac12}AB^{\dagger\frac12}\,\mathrm{m}\,I)B^{\frac12}\Big).$$

Proof. Let P be the projections onto $(\ker A)^\perp$, that is, $P = A^\dagger A = A^{\dagger\frac{1}{2}}A^{\frac{1}{2}}$. The kernel condition shows $\operatorname{ran} A \,\mathrm{m}\, B \subset \operatorname{ran} P$ and hence Theorem 1 implies

$$\begin{aligned} A \,\mathrm{m}\, B &= P(A \,\mathrm{m}\, B)P \\ &\le (PAP)\,\mathrm{m}\,(PBP) = A\,\mathrm{m}(A^{\frac{1}{2}}A^{\dagger\frac{1}{2}}BA^{\dagger\frac{1}{2}}A^{\frac{1}{2}}) \\ &= A^{\frac{1}{2}}\left(P\,\mathrm{m}(A^{\dagger\frac{1}{2}}AA^{\dagger\frac{1}{2}})\right)A^{\frac{1}{2}} \le A^{\frac{1}{2}}\left(I\,\mathrm{m}(A^{\dagger\frac{1}{2}}BA^{\dagger\frac{1}{2}})\right)A^{\frac{1}{2}}. \end{aligned}$$

Similarly we have the other case. □

Remark. The kernel condition (1) is necessary in the above theorem. In fact, the arithmetic mean $A\nabla B = (A+B)/2$ does not satisfy (1). Let $P(= B^{\frac{1}{2}})$ be a projection that does not commute with A. Then $PAP \not\geq A$, so that

$$P(P^\dagger AP^\dagger \nabla I)P = PAP\nabla P = \frac{PAP+P}{2} \not\geq \frac{A+P}{2} = A\nabla B.$$

The difference in the inequality in the above theorem is somewhat larger than we expected as in the following examples:

Example. For $0 < a < 1$, we define a positive-definite matrix A and a projection P : Put

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}^2 = \begin{pmatrix} 1+a^2 & 2a \\ 2a & 1+a^2 \end{pmatrix}.$$

Then we have $A^{-\frac{1}{2}} = A^{\dagger\frac{1}{2}} = \dfrac{1}{1-a^2}\begin{pmatrix} 1 & -a \\ -a & 1 \end{pmatrix}$ and

$$P^{\frac{1}{2}}\sqrt{P^{\dagger\frac{1}{2}}AP^{\dagger\frac{1}{2}}}P^{\frac{1}{2}} = P\sqrt{PAP}P = \sqrt{1+a^2}P\,(\ge P).$$

On the other hand,

$$A^{\dagger\frac{1}{2}}PA^{\dagger\frac{1}{2}} = \frac{1}{(1-a^2)^2}\begin{pmatrix} 1 & -a \\ -a & a^2 \end{pmatrix} = \frac{1+a^2}{(1-a^2)^2}Q,$$

where $Q = \dfrac{1}{1+a^2}\begin{pmatrix} 1 & -a \\ -a & a^2 \end{pmatrix}$ is a rank 1 projection. Hence we have

$$A\#P = P\#A = A^{\frac{1}{2}}\sqrt{A^{\dagger\frac{1}{2}}PA^{\dagger\frac{1}{2}}}A^{\frac{1}{2}} = \frac{\sqrt{1+a^2}}{1-a^2}A^{\frac{1}{2}}QA^{\frac{1}{2}} = \frac{1-a^2}{\sqrt{1+a^2}}P\ (\le P).$$

These differences are under the kernel inclusion as in Corollary 2.

To see a general case, we put $B = \begin{pmatrix} 1 & b \\ b & 1 \end{pmatrix}^2$ for $0 < b < 1$. For $X = P \oplus B$, $Y = A \oplus P$, the orthogonality shows

$$X\#Y = (P\#A) \oplus (B\#P) = \frac{1-a^2}{\sqrt{1+a^2}}P \oplus \frac{1-b^2}{\sqrt{1+b^2}}P.$$

Thus we have

$$X\#Y \le P \oplus \frac{1-b^2}{\sqrt{1+b^2}}P \equiv M_1 \quad \text{and} \quad X\#Y \le \frac{1-a^2}{\sqrt{1+a^2}}P \oplus P \equiv M_2,$$

while

$$X^{\frac{1}{2}}\sqrt{X^{\dagger\frac{1}{2}}YX^{\dagger\frac{1}{2}}}X^{\frac{1}{2}} = \sqrt{1+a^2}P \oplus \frac{1-b^2}{\sqrt{1+b^2}}P \geq M_1 \quad \text{and}$$
$$Y^{\frac{1}{2}}\sqrt{Y^{\dagger\frac{1}{2}}XY^{\dagger\frac{1}{2}}}Y^{\frac{1}{2}} = \frac{1-a^2}{\sqrt{1+a^2}}P \oplus \sqrt{1+b^2}P \geq M_2.$$

Acknowledgement. This study is partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (C), JSPS KAKENHI Grant Number JP 16K05253.

References

[1] W.N.Anderson and R.J.Duffin, Series of parallel addition of matrices, J. Math. Anal. Appl., **26**(1969), 576–594.

[2] T.Ando, "Topics on operator inequalities", Hokkaido Univ. Lecture Note, 1978.

[3] R.G.Douglas, On majorization, factorization and range inclusion of operators in Hilbert space, Proc. Amer. Math. Soc., **17**(1966), 413-416.

[4] P.A.Fillmore and J.P.Williams, On operator ranges, Adv. in Math., **7**(1971), 254–281.

[5] J.I.Fujii, Initial conditions on operator monotone functions, Math. Japon., **23**(1979), 667–669.

[6] J.I.Fujii, Izumino's view of operator means, Math. Japon., **33**(1988), 671–675.

[7] J.I.Fujii, Operator means and the relative operator entropy. Operator theory and complex analysis (Sapporo, 1991), 161–172, Oper. Theory Adv. Appl., **59**, Birkhäuser, Basel, 1992.

[8] J.I.Fujii, Operator means and range inclusion. Linear Algebra Appl., **170**(1992), 137–146.

[9] C.W.Groetsch, "Generalized Inverses of Linear Operators: Representation and Approximation", Marcel Dekker, Inc., 1977.

[10] J.I.Fujii, M.Fujii and Y.Seo, An extension of the Kubo -Ando theory: Solidarities, Math. Japon., **35**(1990), 509–512.

[11] J.I.Fujii and Y.Seo, The relative operator entropy and the Karcher mean, to appear in Linear Algebra Appl..

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

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

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

[15] M.Ould-Ali and B.Messirdi, On closed range operators in Hilbert space, Int. J. Alg., **4**(2010), 953–958.

[16] W.Pusz and S.L.Woronowicz, Functional calculus for sesquilinear forms and the purification map, Rep. on Math. Phys., **8**(1975), 159–170.

[17] W.Pusz and S.L.Woronowicz, Form convex functions and the WYDL and other inequalities, Let. in Math. Phys., **2**(1978), 505–512.

Communicated by *Masatoshi Fujii*

Department of Educational Collaboration (Science, Mathematics, Information), Osaka Kyoiku University, Asahigaoka, Kashiwara, Osaka 582-8582, Japan.

E-mail address, fujii@cc.osaka-kyoiku.ac.jp