Jensen's operator inequality and its application

1. Introduction.

In 1980, Kubo-Ando [12] established the theory of operator means. Hansen and Hansen-Pedersen [9] considered the Jensen inequality in the frame of operator inequalities. (See also [5] and [11].) Under such situation, we discussed the invariance of the operator concavity by the transformation among functions related to operator means in [4]. As a simple application, we could prove the operator concavity of the entropy function $\eta(t) = -t \log t$ which was shown by Nakamura-Umegaki [13]. In the paper, we proposed the following characterization of the operator concavity:

**Theorem A.** Let $f$ be a continuous, real-valued function on $I = [0, r)$. Then the following conditions are mutually equivalent:

1. $f$ is operator concave on $I$, i.e.,
   $$f(tA + (1-t)B) \geq tf(A) + (1-t)f(B) \quad \text{for } t \in [0,1] \text{ and } A, B \in S(I),$$

   where $X \in S(I)$ means that $X$ is a selfadjoint operator whose spectrum is contained in $I$.

2. $f(C^*AC) \geq C^*f(A)C$ for all isometries $C$ and $A \in S(I)$.

3. $f(C^*AC + D^*BD) \geq C^*f(A)C + D^*f(B)D$ for all $C, D$ with $C^*C + D^*D = 1$ and $A, B \in S(I)$.

4. $f(PAP + P^\perp BP^\perp) \geq Pf(A)P + P^\perp f(B)P^\perp$ for all projections $P$ and $A \in S(I)$.

To show the utility of Theorem A, we review the following result in [4].

**Theorem B.** Let $f$ be a real-valued continuous function on $(0, \infty)$. Then $f$ is operator concave if and only if so is $f^*$, where $f^*(t) = tf(t^{-1})$ for $t > 0$.

In fact, suppose that $f$ is operator concave. For arbitrary positive invertible operators $A, B$ and positive numbers $s, t$ with $s^2 + t^2 = 1$, we put $E = s^2A + t^2B$ and

$$X = sA^{1/2}E^{-1/2} \quad \text{and} \quad Y = tB^{1/2}E^{-1/2}. $$
Since $X^*X + Y^*Y = 1$, it follows from Theorem A (3) that

$$f(E^{-1}) = f(X^*A^{-1}X + Y^*B^{-1}Y) \geq X^*f(A^{-1})X + Y^*f(B^{-1})Y,$$

so that

$$f^*(E) = E^{1/2}f(E^{-1})E^{1/2} \geq s^2 A^{1/2}f(A^{-1})A^{1/2} + t^2 B^{1/2}f(B^{-1})B^{1/2},$$

that is, $f^*$ is operator concave.

In addition, if we take $f(t) = \log t$, then $f^*(t) = -t \log t$. Hence, if one could the operator concavity of $\log t$, then that of the entropy function is easily obtained.

Concluding this section, we remark on the transformation $f \rightarrow f^*$. For this, we explain operator means briefly. A binary operation among positive operators on a Hilbert space $m$ is called an operator mean (connection) if it is monotone and continuous from above in each variable and satisfies the transformer inequality. The principal result is the existence of an affine-isomorphism between the classes of all operator means and all nonnegative operator monotone functions on $(0, \infty)$, which is given by $f_m(t) = 1/m t$ for $t > 0$. Thus $f_m(t) = t m 1$ is corresponding to the transpose $m^*$ of $m$, i.e., $A m^* B = B m A$.

2. Yanagi-Furuichi-Kuriyama conjecture.

In this section, we apply Theorem A to an operator inequality related to a conjecture due to Yanagi-Furuichi-Kuriyama [14]. As a matter of fact, they proposed the following trace inequality: For $A, B \geq 0$,

(1) \[ \text{Tr} \left( (A + B)^s (A(\log A)^2 + B(\log B)^2) \right) \geq \text{Tr} \left( (A + B)^{s-1} (A \log A + B \log B)^2 \right) \]

for $0 \leq s \leq 1$.

We now prove it for $s = 0$ by showing the following operator inequality:

Theorem 1. Let $A$ and $B$ be positive invertible operators on a Hilbert space. Then

$$\begin{align*}
(A \log A + B \log B)(A + B)^{-1} &\leq A(\log A)^2 + B(\log B)^2.
\end{align*}$$

Proof. It is similar to a proof of Theorem B. We put

$$C = A^{1/2}(A + B)^{-1/2} \quad \text{and} \quad D = B^{1/2}(A + B)^{-1/2}.$$

Then we have $C^*C + D^*D = 1$. We here note that the function $t^2$ is operator convex on the real line. Hence, if we put $X = \log A$ and $Y = \log B$, then it follows that

$$\begin{align*}
(C^*XC + D^*YD)^2 &\leq C^*X^2C + D^*Y^2D,
\end{align*}$$

cf. Theorem A (3). Arranging it by multiplying $(A + B)^{1/2}$ on both sides, we have the desired operator inequality.
In addition, we give a proof of (1) for $s = 1$. First of all, we note that an inequality

$$(2) \quad \text{Tr} \ (I(A|B)I(B|A)) \leq 0$$

holds for positive operators $A$ and $B$, where $I(A|B) = A \log A - A \log B$ is an operator version of Umegaki's relative entropy. Actually we have

$$\text{Tr} \ (I(A|B)I(B|A)) = \text{Tr} \ (A(\log A - \log B)B(\log B - \log A)) = -\text{Tr} \ (A^{1/2}(\log A - \log B)B(\log A - \log B)A^{1/2}) \leq 0.$$ 

Now a direct calculation shows that

$$\text{Tr} \ ((A+B)(A(\log A)^2 + B(\log B)^2)) = \text{Tr} \ (AB(\log B)^2 + BA(\log A)^2 - 2(A(\log A)(B\log B))).$$

On the other hand, we have

$$\text{Tr} \ (I(A|B)I(B|A)) = \text{Tr} \ (A(\log A)B\log B - A(\log A)B\log A - A(\log B)B\log A) = \text{Tr} \ (2A(\log A)B\log B - BA(\log A)^2 - AB(\log B)^2).$$

Noting by (2), we have

$$\text{Tr} \ ((A+B)(A(\log A)^2 + B(\log B)^2)) \geq \text{Tr} \ ((A\log A + B\log B)^2),$$

which is the inequality (1) for $s = 1$.

Next we give two examples, which show that the above problem (1) can not be solved via operator inequalities in the following sense.

**Theorem 2.** The following operator inequalities do not hold for positive invertible operators $A$ and $B$ in general:

1. $(A + B)^{1/2}(A(\log A)^2 + B(\log B)^2))(A + B)^{1/2} \geq (A\log A + B\log B)^2$.

2. $(A(\log A)^2 + B(\log B)^2))^{1/2}(A + B)(A(\log A)^2 + B(\log B)^2))^{1/2} \geq (A\log A + B\log B)^2$.

**Proof.** For the former, we take

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}.$$
Then we have
\[
\log A = \frac{\log(3 + \sqrt{8})}{2\sqrt{8}} \begin{pmatrix} \sqrt{8} + 2 & 2 \\ 2 & \sqrt{8} - 2 \end{pmatrix} + \frac{\log(3 - \sqrt{8})}{2\sqrt{8}} \begin{pmatrix} \sqrt{8} - 2 & -2 \\ -2 & \sqrt{8} + 2 \end{pmatrix},
\]
\[
\log B = \frac{\log(3 + \sqrt{3})}{2\sqrt{3}} \begin{pmatrix} \sqrt{3} + 2 & 1 \\ 1 & \sqrt{3} - 2 \end{pmatrix} + \frac{\log(3 - \sqrt{3})}{2\sqrt{3}} \begin{pmatrix} \sqrt{3} - 2 & -1 \\ -1 & \sqrt{3} + 2 \end{pmatrix}
\]
and
\[
(A + B)^{1/2} = \frac{\sqrt{11}}{10} \begin{pmatrix} 9 & 3 \\ 3 & 1 \end{pmatrix} + \frac{1}{10} \frac{\sqrt{11}}{10} \begin{pmatrix} 1 & -3 \\ -3 & 9 \end{pmatrix}.
\]
Hence
\[
X = (A + B)^{1/2}(A(\log A)^2 + B(\log B)^2)(A + B)^{1/2} - (A\log A + B\log B)^2
\]
is approximated by
\[
\begin{pmatrix} 0.2800534147 & 0.6060988713 \\ 0.6060988713 & 1.087423161 \end{pmatrix}
\]
and det $X \approx -0.06281927236 < 0$. Namely (1) does not hold for $A$ and $B$.

For the latter, we take
\[
A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}.
\]
Then we have
\[
\log A = \frac{\log 3}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}
\]
and $\log B$ is the same as the above, so that
\[
A(\log A)^2 + B(\log B)^2 = \begin{pmatrix} 15.40739329 & 5.007156201 \\ 5.007156201 & 2.62046225 \end{pmatrix}.
\]
Hence its point spectrum is \{0.8930894768, 17.13476606\} and its square root is as follows:
\[
\begin{pmatrix} 3.799679761 & 0.9847979508 \\ 0.9847979508 & 1.284770503 \end{pmatrix}.
\]
Thus the difference of the both sides
\[
\{A(\log A)^2 + B(\log B)^2\}^{1/2}(A + B)\{A(\log A)^2 + B(\log B)^2\}^{1/2} - (A\log A + B\log B)^2
\]
is approximated by
\[
\begin{pmatrix} 8.760452694 & -1.019211361 \\ -1.019211361 & -0.0425050649 \end{pmatrix}.
\]
Namely (2) does not hold for $A$ and $B$.

In a private communication with Professor Yanagi, we knew this conjecture last autumn. Very recently we were given an opportunity to read a preprint [9] by Furuta, related to Theorem 2. The authors would like to express their thanks to Professor Furuta for his kindness of sending it.


Recently, F.Hansen and G.K.Pedersen [13] reconsidered the preceding results in [12, 11] by themselves, which is along with Theorem A. (See also [10].)

Hansen-Pedersen's theorem. The following conditions are all equivalent to that $f$ is operator convex on $\mathcal{I}$:

(i) $f\left(\sum_{k=1}^{n}C_{k}^{*}A_{k}C_{k}\right) \leq \sum_{k=1}^{n}C_{k}^{*}f(A_{k})C_{k}$ for all selfadjoint $A_{k}$ with $\sigma(A_{k}) \subset \mathcal{I}$ and $C_{k}$ with $\sum_{k=1}^{n}C_{k}^{*}C_{k} = 1$.

(ii) $f(C^{*}AC) \leq C^{*}f(A)C$ for all selfadjoint $A$ with $\sigma(A) \subset \mathcal{I}$ and isometries $C$.

(iii) $Pf(PAP + s(1 - P)) \leq Pf(A)P$ for all selfadjoint operators $A$ with $\sigma(A) \subset \mathcal{I}$, scalars $s \in \mathcal{I}$ and projections $P$.

Now we synthesize Jensen's operator inequality. Among others, a theorem due to Davis [6] and Choi [5] is included as the fifth condition. (See also Ando [1].)

Theorem 3. Let $f$ be a real function on an interval $\mathcal{I}$, $A$ or $A_{k}$ a selfadjoint operator with $\sigma(A), \sigma(A_{k}) \subset \mathcal{I}$, and $H$ or $K$ a Hilbert space. Then the following conditions are mutually equivalent:

(i) (i) $f$ is operator convex on $\mathcal{I}$.

(ii) $f(C^{*}AC) \leq C^{*}f(A)C$ for all $A \in B(H)$ and isometries $C \in B(K, H)$.

(ii') $f(C^{*}AC) \leq C^{*}f(A)C$ for all $A$ and isometries $C$ in $B(H)$.

(iii) $f\left(\sum_{k=1}^{n}C_{k}^{*}A_{k}C_{k}\right) \leq \sum_{k=1}^{n}C_{k}^{*}f(A_{k})C_{k}$ for all $A_{k} \in B(H)$ and $C_{k} \in B(K, H)$ with $\sum_{k=1}^{n}C_{k}^{*}C_{k} = 1_{K}$.

(iii') $f\left(\sum_{k=1}^{n}C_{k}^{*}A_{k}C_{k}\right) \leq \sum_{k=1}^{n}C_{k}^{*}f(A_{k})C_{k}$ for all $A_{k}, C_{k} \in B(H)$ with $\sum_{k=1}^{n}C_{k}^{*}C_{k} = 1_{H}$.

(iv) $f\left(\sum_{k=1}^{n}P_{k}A_{k}P_{k}\right) \leq \sum_{k=1}^{n}P_{k}f(A_{k})P_{k}$ for all $A_{k}$ and projections $P_{k} \in B(H)$ with $\sum_{k=1}^{n}P_{k} = 1_{H}$.

(v) $f\left(\Phi(A)\right) \leq \Phi\left(f(A)\right)$ for all unital positive linear map $\Phi$ between $C^{*}$-algebras $A, B$ and all $A \in A$.

Proof. (i)⇒(ii): Take $B = B^{*} \in B(K)$ with $\sigma(B) \in \mathcal{I}$. For $P = \sqrt{1_{H} - CC^{*}}$, putting

$$X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in B(H \oplus K), \quad U = \begin{pmatrix} C & P \\ 0 & -C^{*} \end{pmatrix}, \quad V = \begin{pmatrix} C & -P \\ 0 & C^{*} \end{pmatrix} \in B(K \oplus H, H \oplus K),$$

we have

$$f\left(\sum_{k=1}^{n}C_{k}^{*}A_{k}C_{k}\right) \leq \sum_{k=1}^{n}C_{k}^{*}f(A_{k})C_{k},$$

and

$$f\left(\sum_{k=1}^{n}P_{k}A_{k}P_{k}\right) \leq \sum_{k=1}^{n}P_{k}f(A_{k})P_{k}.$$
we have

$$C^*P = \sqrt{1_K - C^*C} = 0 \in B(H, K), \quad PC = C\sqrt{1_K - C^*C} = 0 \in B(K, H),$$

so that both $U$ and $V$ are unitaries. Since

$$U^*XU = \begin{pmatrix} C^*AC & C^*AP \\ PAC & PAP + CBC^* \end{pmatrix}, \quad V^*XV = \begin{pmatrix} C^*AC & -C^*AP \\ -PAC & PAP + CBC^* \end{pmatrix},$$

then the operator convexity of $f$ implies

$$\begin{pmatrix} f(C^*AC) & 0 \\ 0 & f(PAP + CBC^*) \end{pmatrix} = f\begin{pmatrix} C^*AC & 0 \\ 0 & PAP + CBC^* \end{pmatrix} = f\left(\frac{U^*XU + V^*XV}{2}\right) \leq \frac{f(U^*XU) + f(V^*XV)}{2} = \frac{U^*F(X)U + V^*f(X)V}{2} = \begin{pmatrix} C^*f(A)C & 0 \\ 0 & Pf(A)P + Cf(B)C^* \end{pmatrix}.$$

Thus we have (ii) by seeing the $(1, 1)$-components.

(ii)⇒(iii): Putting

$$\tilde{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \in B(H \oplus \cdots \oplus H), \quad \tilde{C} = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix} \in B(K, H \oplus \cdots \oplus H),$$

we have $\tilde{C}^*\tilde{C} = 1_K$. It follows from (ii) that

$$f\left(\sum_{k=1}^{n} C_k^*A_kC_k \right) = f\left(\tilde{C}^*\tilde{A}\tilde{C}\right) \leq \tilde{C}^*f(\tilde{A})\tilde{C} = \sum_{k=1}^{n} C_k^*f(A_k)C_k.$$

(iii)⇒(v): Considering the universal enveloping von Neumann algebras and the uniquely extended linear map, we may assume that $\mathcal{A}$ is a von Neumann algebra. Thereby a selfadjoint operator $A \in \mathcal{A}$ can be approximated uniformly by a simple function $A' = \sum_k t_kE_k$ where $\{E_k\}$ is a decomposition of the unit $1_{\mathcal{A}}$. Since $\sum_k \Phi(E_k) = 1_B$ by the unitality of $\Phi$, then applying (iii) to $C_k = \sqrt{\Phi(E_k)}$, we have

$$f(\Phi(A')) = f\left(\sum_k t_k\Phi(E_k)\right) \leq \sum_k f(t_k)\Phi(E_k) = \Phi\left(\sum_k f(t_k)E_k\right) = \Phi(f(A')).$$
The continuity of $\Phi$ implies (v).
Since (v) implies (iv) obviously, we next show (iv) $\Rightarrow$ (i): Putting
\[
X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} \sqrt{1-t} & -\sqrt{t} \\ \sqrt{t} & \sqrt{1-t} \end{pmatrix},
\]
we have
\[
\left( f((1-t)A + tB) \quad f((1-t)B + tA) \right) 
\]
\[
= f(\sqrt{1-t})U^*XU + (1-P)U^*f(A)U(1-P)) 
\]
\[
\leq \left( (1-t)f(A) + tf(B) \quad (1-t)f(B) + tf(A) \right).
\]
so that $f$ is operator convex.

Consequently, we proved the equivalence of (i) - (v). To complete the proof, we need (ii)$' \Rightarrow$ (iii)$'$ because it is non-trivial in (i) $\Rightarrow$ (ii) $\Rightarrow$ (ii)$' \Rightarrow$ (ii$'$) $\Rightarrow$ (iv) $\Rightarrow$ (i).

Modifying the proof in [7], we can show (ii)$'$ $\Rightarrow$ (iii)$'$. We may assume $n = 2$. Putting
\[
\tilde{X} = \begin{pmatrix} A_1 & 0 & \ldots \\ A_2 \\ \vdots \end{pmatrix}, \quad \tilde{V} = \begin{pmatrix} C_1 & 0 & \ldots \\ C_2 & 0 & \ldots \\ 0 & 1 & 0 & \ldots \end{pmatrix} \in B(H \oplus H \oplus \cdots),
\]
we have $\tilde{V}^*\tilde{V} = 1$ and
\[
\left( f(C^*A_1C_1 + C^*_2A_2C_2) \quad f(A_2) \quad \ldots \right) = f(\tilde{V}^*\tilde{X}\tilde{V}) \leq \tilde{V}^*f(\tilde{X})\tilde{V}
\]
\[
= \begin{pmatrix} C^*_1f(A_1)C_1 + C^*_2f(A_2)C_2 \\ \ldots \end{pmatrix}.
\]

\[ \square \]

Remark 1. (1) Theorem 3 includes the above two Jensen's operator inequalities. An essential part of the proof for the Hansen-Pedersen-Jensen inequality is to show that (1) implies (2). In fact, suppose (1) and $C^*C \leq 1$. Then, putting $D = \sqrt{1 - C^*C}$, we have by (iii$'$) and $f(0) \leq 0$ that
\[
f(C^*AC + D0D) \leq C^*f(A)C + D^2f(0) \leq C^*f(A)C.
\]
(2) Note that the property either 'isometric' or 'unital' assures the spectral invariance as follows: If $m \leq A \leq M$, then $m \leq C^*AC \leq M$ and $m \leq \Phi(A) \leq M$ for any isometry $C$ and a unital positive linear map $\Phi$.

参考文献


* DEPARTMENT OF ARTS AND SCIENCES (INFORMATION SCIENCE), OSAKA KYOIKU UNIVERSITY, ASAHIGAOKA, KASHIWARA, OSAKA 582-8582, JAPAN.
E-mail address: fujii@cc.osaka-kyoiku.ac.jp

** DEPARTMENT OF MATHEMATICS, OSAKA KYOIKU UNIVERSITY, ASAHIGAOKA, KASHIWARA, OSAKA 582-8582, JAPAN.
E-mail address: mfujii@cc.osaka-kyoiku.ac.jp

*** FACULTY OF ENGINEERING, IBARAKI UNIVERSITY, HITACHI, IBARAKI 316-0033, JAPAN.
E-mail address: nakamoto@base.ibaraki.ac.jp