<table>
<thead>
<tr>
<th>Title</th>
<th>Jensen's operator inequality and its application (Advanced Topics of Information Science and Functional Analysis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Fujii, Jun-Ichi; Fujii, Masatoshi; Nakamoto, Ritsuo</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 1396: 85-93</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2004-10</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/25970">http://hdl.handle.net/2433/25970</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>
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.,
$$f(tA + (1-t)B) \geq tf(A) + (1-t)f(B) \quad \text{for} \quad t \in [0,1] \quad \text{and} \quad A, B \in S(I),$$
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

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

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 $f_m(t) = 1 + mt$ for $t > 0$. Thus $f_m^*(t) = t m1$ 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 $A, B \geq 0$,

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

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

$$(A \log A + B \log B)(A+B)^{-1} (A \log A + B \log B) \leq A(\log A)^2 + B(\log B)^2.$$

**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

$$(C^*XC + D^*YD)^2 \leq C^*X^2C + D^*Y^2D,$$

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

\[ \text{Tr} \left( I(A|B)I(B|A) \right) \leq 0 \]

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

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

Now a direct calculation shows that

\[
\text{Tr} \left( ((A+B)(A(\log A)^2 + B(\log B)^2) - (A \log A + B \log B)^2 \right) = \text{Tr} \left( AB(\log B)^2 + BA(\log A)^2 - 2(A \log A)(B \log B) \right).
\]

On the other hand, we have

\[
\text{Tr} \left( I(A|B)I(B|A) \right) = \text{Tr} \left( A(\log A)B \log B - A(\log A)B \log A - A(\log B)B \log B + A(\log B)B \log A \right) = \text{Tr} \left( 2A(\log A)B \log B - BA(\log A)^2 - AB(\log B)^2 \right).
\]

Noting by (2), we have

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

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
\[
\log A = \frac{\log(3 + \sqrt{8})}{2\sqrt{8}} \left( \frac{\sqrt{8} + 2}{2} \frac{2}{\sqrt{8} - 2} \right) + \frac{\log(3 - \sqrt{8})}{2\sqrt{8}} \left( -2 \frac{\sqrt{8} - 2}{\sqrt{8} + 2} \right),
\]
\[
\log B = \frac{\log(3 + \sqrt{3})}{2\sqrt{3}} \left( \frac{\sqrt{3} + 2}{1} \frac{1}{\sqrt{3} - 2} \right) + \frac{\log(3 - \sqrt{3})}{2\sqrt{3}} \left( -1 \frac{\sqrt{3} - 2}{\sqrt{3} + 2} \right)
\]
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
\[
A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}.
\]

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 $I$:

(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 $\sigma(A_{k}) \subset I$ 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 $\sigma(A) \subset I$ and isometries $C$.

(iii) $P f(PAP + s(1 - P)) \leq P f(A) P$ for all selfadjoint operators $A$ with $\sigma(A) \subset I$, scalars $s \in I$ 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 $I$, $A$ or $A_{k}$ a selfadjoint operator with $\sigma(A), \sigma(A_{k}) \subset I$, and $H$ or $K$ a Hilbert space. Then the following conditions are mutually equivalent:

(i) (1) $f$ is operator convex on $I$.

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

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

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

**Proof.** (i)$\Rightarrow$(ii): Take $B = B^{*} \in B(K)$ with $\sigma(B) \subset I$. For $P = \sqrt{1_{H} - CC^{*}}$, putting $X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in B(H \oplus K)$, $U = \begin{pmatrix} C & P \\ 0 & -C^{*} \end{pmatrix}$, $V = \begin{pmatrix} C & -P \\ 0 & C^{*} \end{pmatrix} \in B(K \oplus H, H \oplus K)$,
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
\[
\begin{pmatrix} f(C^*AC) & 0 \\ 0 & f(PAP + CBC^*) \end{pmatrix} = \begin{pmatrix} C^*AC & 0 \\ 0 & PAP + CBC^* \end{pmatrix}
\leq \begin{pmatrix} f(U^*XU) + f(V^*XV) \\ \frac{U^*F(X)U + V^*f(X)V}{2} \end{pmatrix} = \begin{pmatrix} C^*f(A)C & 0 \\ 0 & Pf(A)P + Cf(B)C \end{pmatrix}.
\]
Thus we have (ii) by seeing the \((1, 1)\)-components.

(ii)\(\Rightarrow\) (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
\[
f\left( \sum_{k=1}^n C_k^* A_k C_k \right) = f\left( \tilde{C}^*\tilde{A}\tilde{C} \right) \leq \tilde{C}^* f(\tilde{A})\tilde{C} = \sum_{k=1}^n C_k^* f(A_k) C_k.
\]

(iii)\(\Rightarrow\) (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 \( A' = \sum_k t_k E_k \) where \( \{E_k\} \) is a decomposition of the unit \( 1_\mathcal{A} \). Since \( \sum_k \Phi(E_k) = 1_\mathcal{B} \) 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

$$\begin{pmatrix} f((1-t)A + tB) \\ f((1-t)B + tA) \end{pmatrix} = f(PU^*XUP + (1-P)U^*XU(1-P)) \leq PU^*f(X)UP + (1-P)U^*f(X)U(1-P)) = \begin{pmatrix} (1-t)f(A) + tf(B) \\ (1-t)f(B) + tf(A) \end{pmatrix},$$

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') $\Rightarrow$ (iii') $\Rightarrow$ (iv) $\Rightarrow$ (i).

Modifying the proof in [7], we can show (ii') $\Rightarrow$ (iii'). We may assume $n = 2$. Putting

$$\tilde{X} = \begin{pmatrix} A_1 \\ A_2 \\ A_2 \\ \vdots \end{pmatrix}, \quad \tilde{V} = \begin{pmatrix} C_1 & 0 & \cdots \\ C_2 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \in B(H \oplus H \oplus \cdots),$$

we have $\tilde{V}^*\tilde{V} = 1$ and

$$\begin{pmatrix} f(C_1^*A_1C_1 + C_2^*A_2C_2) \\ f(A_2) \\ \vdots \end{pmatrix} = f(\tilde{V}^*\tilde{X}\tilde{V}) \leq \tilde{V}^*f(\tilde{X})\tilde{V} = \begin{pmatrix} C_1^*f(A_1)C_1 + C_2^*f(A_2)C_2 \\ f(A_2) \end{pmatrix}.$$

**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