JENSEN INEQUALITY IS A COMPLEMENT TO KANTOROVICH INEQUALITY

MASATOSHI FUJII AND MASAHIRO NAKAMURA

Received December 28, 2004; revised January 11, 2005

Dedicated to Professor Hisaharu Umegaki on his 80th Birthday

Abstract. Usually the celebrated Kantorovich inequality is regarded as a complementary inequality of the Jensen inequality. In this note, we point out that "monotonicity and linearity" of integral imply the Kantorovich inequality. As a consequence, the Jensen inequality is complementary to the Kantorovich inequality in this sense.

1. We first cite the following inequalities: Let $0 < m < M$. Then

$$\int_I t d\mu(t) \cdot \int_I \frac{1}{t} d\mu(t) \leq \frac{(M+m)^2}{4Mm} \tag{1}$$

for all probability measures μ on $I = [m, M]$, and

$$(Ax, x)(A^{-1}x, x) \leq \frac{(M+m)^2}{4Mm} \tag{2}$$

for all positive invertible operators A on a Hilbert space H with $0 < m \leq A \leq M$ and unit vectors $x \in H$. Many authors gave proofs of the inequality, among them we here cite early papers [8] and [14], for example. Note that (1) and (2) are equivalent and (2) is called the Kantorovich inequality. Usually they are regarded as complementary inequalities of

$$1 \leq \int_I t d\mu(t) \cdot \int_I \frac{1}{t} d\mu(t) \quad \text{and} \quad 1 \leq (Ax, x)(A^{-1}x, x) \tag{3}$$

respectively, where μ (resp. A and x) is as in (1) (resp. (2)).

By the way, (3) is a special case of the Jensen inequality, i.e., if f is a real-valued continuous convex function on I, then

$$f((Ax, x)) \leq (f(A)x, x) \tag{4}$$

for all selfadjoint operators A on a Hilbert space H whose spectra are included in I and unit vectors $x \in H$. From the veiwpoint of (4), Mond-Pečarić and Ando considered, as generalizations of the Kantorovich inequality, complementary inequalities of the Jensen inequality, see [11] and also [13]. We here cite Furuta's textbook [5] as a pertinent reference to Kantorovich inequalities.

Reviewing the Kantorovich inequality, we shall point out in this note that the Jensen inequality is complementary to the Kantorovich inequality. To do this, we shall set up two elementary and simple axioms "monotonicity and linearity", by which we approach to the Kantorovich inequality and its generalizations. Incidentally such an idea can be seen in the discussion of [17] by Takahasi, Tsukada, Tanahashi and Ogiwara.

2000 *Mathematics Subject Classification.* 47A63 and 47B15.
Key words and phrases. Kantorovich inequality, Jensen inequality and Mond-Pecaric method in operator inequality.

2. At the beginning of this section, we cite two simple properties on integral used below. Let $C(I)$ be the algebra of all real-valued continuous functions on $I = [m, M]$ and μ a probability measure on I. Then

(E1) *If $f \le g$ for $f, g \in C(I)$, then* $\int_I f(t)d\mu(t) \le \int_I g(t)d\mu(t)$.

(E2) *If g is linear, i.e., $g(t) = at + b$ for some $a, b \in \mathbb{R}$ and $\int_I td\mu(t) = s$, then $\int_I g(t)d\mu(t) = g(s)$.*

It is clear that (E1) and (E2) are satisfied. In addition, they are true for a unital positive linear functional on $C(I)$, $E[f] = E_{A,x}[f] = (f(A)x, x)$ for $f \in C(I)$, where A and x are as in (2). That is, $E[f]$ satisfies

(E1) *If $f \le g$ for $f, g \in C(I)$, then* $E[f] \le E[g]$, and

(E2) *If g is linear, i.e., $g(t) = at + b$ for some $a, b \in \mathbb{R}$ and $E[t] = s$, then $E[g] = g(s)$.*

They also hold for $E_{A,\phi}(f) = \phi(f(A))$ where ϕ is an arbitrary state on the C^*-algebra genetated by A and the identity. Anyway, (E1) and (E2) are quite trivial, but essential.

We here apply them to the Kantorovich inequality (1): Let $L = L_{\frac{1}{t}}$ be the function corresponding to the segment connecting the points $(m, \frac{1}{m})$ and $(M, \frac{1}{M})$, and put $s_0 = \int_I td\mu(t)$. Then we have $\frac{1}{t} \le L(t)$ for $t \in [m, M]$ and $s_0 \in I$. Hence (E1) and (E2) imply

$$\int_I td\mu(t) \cdot \int_I \frac{1}{t} d\mu(t) \le \int_I td\mu(t) \cdot \int_I L(t)d\mu(t) = L(s_0)s_0.$$

Since $L(m)m = L(M)M = 1$, the quadratic polynomial $L(s)s$ attains its maximum at the midpoint $s = \frac{M+m}{2}$ of m and M. Hence we have

$$\max_{s\in[m,M]} L(s)s = L\left(\frac{M+m}{2}\right)\frac{M+m}{2} = \frac{\frac{1}{m}+\frac{1}{M}}{2} \cdot \frac{M+m}{2} = \frac{(M+m)^2}{4Mm},$$

so that we obtain (1).

In the above proof of (1), (E1) and (E2) worked under the following facts actually: The function $f(t) = \frac{1}{t}$ satisfies

(i) *$f(t) \le L_f(t)$ for all $t \in I$*, where L_f is the function corresponding to the segment connecting the points $(m, f(m))$ and $(M, f(M))$, and

(ii) *$f(m)m = f(M)M$.*

The latter (ii) says that the function $L_f(s)s$ has the same value at $s = m, M$. As a consequence, we have the following estimation which is a generalization of (2).

Theorem 1. *If $f \in C(I)$ satisfies (i) and (ii) cited above and $f(m) > f(M)$, then*

$$(f(A)x, x)(Ax, x) \le \frac{(f(M) + f(m))(M + m)}{4}$$

for all selfadjoint operators A on a Hilbert space H with $m \le A \le M$ and unit vectors $x \in H$.

Proof. Since $s_0 = (Ax, x) \in [m, M]$ and $E[f] = (f(A)x, x) \le E[L_f] = L_f(s_0)$ by (E2) , it follows from the assumption (ii) that

$$(f(A)x, x)(Ax, x) \le L_f(s_0)s_0 \le \max_{s\in[m,M]} L_f(s)s$$

$$= L_f\left(\frac{M+m}{2}\right)\frac{M+m}{2} = \frac{(f(M) + f(m))(M + m)}{4}.$$

3. Next we apply them to generalized Kantorovich inequalities due to Furuta [4] and [7], cf. also [2] and [1]. Following him [6], we denote by $K(m, M, p)$ the generalized Kantorovich constant for $m < M$ and $p \in \mathbb{R}$:

$$K(m, M, p) = \frac{mM^p - Mm^p}{(p-1)(M-m)} \left\{ \frac{(p-1)(M^p - m^p)}{p(mM^p - Mm^p)} \right\}^p .$$

We note that $K(m, M, -1) = K(m, M, 2) = \frac{(M+m)^2}{4Mm}$ is the original Kantorovich constant.

Theorem 2. *Let A be a positive operator on a Hilbert space H with $0 < m \le A \le M$, and let x a unit vector in H. If $p \notin [0, 1)$, then*

$$(A^p x, x) \le K(m, M, p)(Ax, x)^p.$$

On the other hand, if $p \in (0, 1)$, then

$$(A^p x, x) \ge K(m, M, p)(Ax, x)^p.$$

Proof. It is similar to that of the Kantorovich inequality. We put $E[f] = (f(A)x, x)$ for $f \in C(I)$. Then it satisfies (E1) and (E2) and in particular $s_0 = E[t] = (Ax, x) \in [m, M]$. Put $L_p = L_{t^p}$, i.e., the linear function connecting the points (m, m^p) and (M, M^p), or $L_p(t) = \mu_p t + \nu_p$ where $\mu_p = \frac{M^p - m^p}{M - m}$ and $\nu_p = \frac{M^p m - m^p M}{M - m}$.

If $p \notin [0, 1)$, then $t^p \le L_p(t)$ for $t \in [m, M]$ and so

$$E[t^p]E[t]^{-p} \le E[L_p]E[t]^{-p} = L_p(s_0)s_0^{-p}$$

by (E1) and (E2). Since $\frac{p\nu_p}{(1-p)\mu_p}$ is a unique solution of $\frac{d}{ds}L_p(s)s^{-p} = 0$ and contained in $[m, M]$, we have

$$\max_{s \in [m,M]} L_p(s)s^{-p} = L_p\left(\frac{p\nu_p}{(1-p)\mu_p}\right)\left(\frac{(1-p)\mu_p}{p\nu_p}\right)^p = K(m, M, p),$$

that is, we obtain the former.

Next, if $p \in (0, 1)$, then

$$E[t^p]E[t]^{-p} \ge E[L_p]E[t]^{-p} = L_p(s_0)s_0^{-p}$$

for $s_0 = E[t] \in [m, M]$ by (E1) and (E2). Since

$$\min_{s \in [m,M]} L_p(s)s^{-p} = K(m, M, p),$$

we obtain the latter.

4. The following theorem is proposed in [10] as a generalization of the Mond-Pečarić inequality [11] which is a typical application of the Mond-Pečarić method [15].

Theorem 3. *Let A_j be positive operators on a Hilbert space H with $0 < m \le A_j \le M$ ($j = 1, 2, \cdots, k$) and $x_1, \cdots, x_k$ vectors in H with $\sum \|x_j\|^2 = 1$. Let f and g be in $C(I)$ where $I = [m, M]$, and U and V intervals such that $U \supseteq f(I)$ and $V \supseteq g(I)$. If f is convex and a real-valued function $F(u, v)$ on $U \times V$ is non-decreasing in u, then*

$$F\left(\sum (f(A_j)x_j, x_j), g\left(\sum (A_j x_j, x_j)\right)\right) \le \max_{t \in I} F(\mu_f(t - m) + f(m), g(t)),$$

where $\mu_f = \frac{f(M) - f(m)}{M - m}$.

Proof. Put $L = L_f$, that is, $L(t) = \mu_f t + \nu_f$ where $\nu_f = \frac{Mf(m)-mf(M)}{M-m}$. Since $E[f] = \sum(f(A_j)x_j, x_j)$ satisfies (E1) and (E2), we have $E[f] \le E[L] = L(s_0) = \mu_f s_0 + \nu_f$ where $s_0 = E[t] = \sum(A_j x_j, x_j) \in I$. Therefore it follows that

$$\begin{aligned} &F(\sum(f(A_j)x_j, x_j), g(\sum(A_j x_j, x_j))) = F(E[f], g(E[t]) \\ &\le F(\mu_f s_0 + \nu_f, g(s_0)) \le \max_{t\in I} F(\mu_f(t-m) + f(m), g(t)), \end{aligned}$$

because $\mu_f s_0 + \nu_f = \mu_f(s_0 - m) + f(m)$.

Remark 4. *In Theorem 2, If $F(u,v)$ is non-increasing, then we have the following inequality with reverse order to the above,* [10]*:*

$$F(\sum(f(A_j)x_j, x_j), g(\sum(A_j x_j, x_j))) \ge \min_{t\in I} F(\mu_f(t-m) + f(m), g(t)).$$

5. It is known that the Kantorovich inequality (2) is generalized as follows: If Φ is a unital positive linear map between C^*-algebras, then

$$\Phi(A^{-1}) \le \frac{(M+m)^2}{4Mm}\Phi(A)^{-1} \tag{5}$$

for all positive invertible elements A with $0 < m \le A \le M$, see [12] and also [13]. Incidentally we note that it follows from (2) and (3) for states, i.e.,

$$1 \le \phi(A)\phi(A^{-1}) \le \frac{(M+m)^2}{4Mm} \tag{6}$$

for all positive invertible operators A with $0 < m \le A \le M$ and states ϕ on the C^*-algebra generated by A, see [13]. Defining $\phi(X) = \psi(\Phi(X))$ for an arbitrary state ψ, it follows from (6) that

$$\begin{aligned} \psi(\Phi(A^{-1})) = \phi(A^{-1}) &\le \frac{(M+m)^2}{4Mm}\phi(A)^{-1} = \frac{(M+m)^2}{4Mm}\psi(\Phi(A))^{-1} \\ &\le \frac{(M+m)^2}{4Mm}\psi(\Phi(A)^{-1}) \end{aligned}$$

by the Jensen inequality $\psi(X)^{-1} \le \psi(X^{-1})$ for $X > 0$. Hence (5) is proved.

As in the proof of (5), Theorem 2 can be generalized to positive linear maps between C^*-algebras, in which we need the Hölder-McCarthy inequality: If A is a positive operator and ψ is a state on a C^*-algebra containing A, then

$$\psi(A)^p \le \psi(A^p) \ (p \notin [0,1]) \text{ and } \psi(A)^p \ge \psi(A^p) \ (p \in [0,1])$$

Corollary 5. *Let A be a positive operator in a C^*-algebra with $0 < m \le A \le M$ and Φ a unital positive linear map. If $p \notin [0,1)$, then*

$$\Phi(A^p) \le K(m,M,p)\Phi(A)^p.$$

On the other hand, if $p \in (0,1)$, then

$$\Phi(A^p) \ge K(m,M,p)\Phi(A)^p.$$

In addition, we point out that (5) implies the result [12; Theorem 1] which is a variant of (5): If Φ is a unital positive linear map betwee C^*-algebras, then

$$\Phi(A) \ \sharp \ \Phi(A^{-1}) \le \frac{M+m}{2\sqrt{Mm}}$$

for all positive invertible elements A with $0 < m \le A \le M$, where $\sharp$ is the geometric mean in the sense of Kubo-Ando [9], i.e., for positive invertible operators A and B,

$$A \sharp B = A^{\frac{1}{2}}(A^{\frac{-1}{2}}BA^{\frac{-1}{2}})^{\frac{1}{2}}A^{\frac{1}{2}}.$$

As a matter of fact, the monotonicity of the geometric mean implies that

$$\Phi(A) \sharp \Phi(A^{-1}) \le \Phi(A) \sharp \frac{(M+m)^2}{4Mm}\Phi(A)^{-1} = \frac{M+m}{2\sqrt{Mm}}$$

because $\Phi(A)$ and $\Phi(A)^{-1}$ commute.

As well-known, the geometric mean $\sharp$ is generalized to the α-geometric mean $\sharp_\alpha$; for positive invertible operators A and B

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

We now point out that our previous result [2;Theorem 4] is a corollary of Theorem 2. As a matter of fact, it can be expressed by the following simpler form, by which it could be understood as a noncommutative version of the Pólya-Szegö inequality, see Greub-Rheinboldt [8, Theorem 2] and also [3, Theorem 3]:

Theorem 6. *Let A and B be positive operators on a Hilbert space H with $0 < m_1 \le A \le M_1$ and $0 < m_2 \le B \le M_2$. Then for $\alpha \in [0,1]$ and $x \in H$*

$$((B \sharp_\alpha A)x,x) \le (Ax,x)^\alpha (Bx,x)^{1-\alpha} \le K^\alpha((B \sharp_\alpha A)x,x),$$

where $K = K((\frac{m_1}{M_2})^\alpha, (\frac{M_1}{m_2})^\alpha, \frac{1}{\alpha})$.

To prove this, we have to modify Theorem 2: If $0 < m \le X \le M$ and $p > 1$, then

$$(Xy,y) \le (X^py,y)^{\frac{1}{p}}\|y\|^{2(1-\frac{1}{p})} \le K(m,M,p)^{\frac{1}{p}}(Xy,y)$$

for all $y \in H$. We apply it by taking $X = (B^{\frac{-1}{2}}AB^{\frac{-1}{2}})^\alpha$, $y = B^{\frac{1}{2}}x$ and $\alpha = \frac{1}{p}$. Then

$$(X^py,y) = (Ax,x),\ \|y\|^2 = (Bx,x) \text{ and } (Xy,y) = ((B \sharp_\alpha A)x,x)$$

Finally the constant $K = K((\frac{m_1}{M_2})^\alpha, (\frac{M_1}{m_2})^\alpha, \frac{1}{\alpha})$ follows from

$$\frac{m_1}{M_2} \le m_1B^{-1} \le B^{\frac{-1}{2}}AB^{\frac{-1}{2}} \le M_1B^{-1} \le \frac{M_1}{m_2}.$$

6. Related to (5), we recall a conditional expectation introduced by Umegaki [18] and [19], which is an important tool to study von Neumann algebras. One of its simple examples is the diagonalization due to von Neumann. Let D be the diagonalization of the matrix algebra $M_2(\mathbb{C})$ of all 2×2 matrices onto the diagonal subalgebra. Then $D[A]D[A^{-1}]$ is a scalar multiple of the identity matrix, say a scalar simply. Thus we may consider the extremal case $D[A]D[A^{-1}] = \frac{(M+m)^2}{4Mm}$, where $\{m,\ M\}$ are propervalues of A. Namely we have the following.

Remark 7. *Let $A \in M_2(\mathbb{C})$ be a positive invertible matrix with the propervalues $\{m,\ M\}$. Then $D[A]$ is a scalar if and only if*

$$D[A]D[A^{-1}] = \frac{(M+m)^2}{4Mm}. \tag{7}$$

In fact, let $A = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$ where $a,\ c > 0$ and $|A| = ac - |b|^2 > 0$. Then

$$D[A]D[A^{-1}] = \frac{ac}{|A|} \quad \text{and} \quad \frac{(M+m)^2}{4Mm} = \frac{(a+c)^2}{4|A|}.$$

Therefore (7) holds if and only if

$$\frac{ac}{|A|} = \frac{(a+c)^2}{4|A|},$$

i.e., $a = c$ by $|A| > 0$, or equivalently, $D[A] = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$.

More precisely, such a matrix A can be expressed as

$$A = \begin{pmatrix} \frac{M+m}{2} & \frac{M-m}{2}e^{i\theta} \\ \frac{M-m}{2}e^{-i\theta} & \frac{M+m}{2} \end{pmatrix}$$

for some $\theta \in \mathbb{R}$. As a matter of fact, since $|A| = Mm$ and $(2a)^2 = (M+m)^2$, we have $a = c = \frac{M+m}{2}$. Moreover $|b| = \frac{M-m}{2}$ follows from $(\frac{M+m}{2})^2 - |b|^2 = a^2 - |b|^2 = |A| = Mm$.

7. Finally we pose a proof along with Rennie [16] to the fact: If $0 < m \le A \le M$, then

$$\phi(A^2) \le \frac{(M+m)^2}{4Mm}\phi(A)^2$$

for any states ϕ of a C^*-algebra containing A. Actually, since $0 \ge (A-m)(A-M)$, we have $(M+m)A \ge A^2 + Mm$, so that

$$\frac{M+m}{2}\phi(A) \ge \frac{\phi(A^2)+Mm}{2} \ge [\phi(A^2)Mm]^{\frac{1}{2}}.$$

Hence we have the conclusion.

References

[1] J.I.Fujii, M.Fujii, Y.Seo and M.Tominaga, *On generalized Kantorovich inequalities*, Banach and Function Spaces, Edited by M.Kato and L.Maligranda, Yokohama Publishers, Yokohama, 2004, 205-213.

[2] M.Fujii, S.Izumino, R.Nakamoto and Y.Seo, *Operator inequalities related to Cauthy-Schwarz and Hölder-McCarthy inequalities*, Nihonkai Math. J., 8(1997), 117-122.

[3] M.Fujii, E.Kamei and A.Matsumoto, *Parameterized Kantorovich inequality for pssitive operators*, Nihonkai Math. J., 6(1995), 129-134.

[4] T.Furuta, *Operator inequalities asociated with Hölder-McCarthy and Kantorovich inequalities* , J. Inequal. Appl., 2(1998), 137-148.

[5] T.Furuta, *Invitation to Linear Operators*, Taylor and Francis, London and New York, 2001.

[6] T.Furuta, *Specht ratio $S(1)$ can be expressed by Kantorovich constant $K(p)$: $S(1) = \exp[K'(1)]$ and its application*, Math. Inequal. Appl., 6(2003), 521-530.

[7] T.Furuta, *Basic properties of the generalized Kantorovich constant $K(h,p) = \frac{(h^p-h)}{(p-1)(h-1)}(\frac{(p-1)}{p}\frac{(h^p-1)}{(h^p-h)})^p$ and its application*, Acta Sci. Math. Szeged, 70(2004), 319-337.

[8] W.Greub and W.Rheinboldt, *On a generalization of an inequality of L.V.Kantorovich*, Proc. Amer. Math. Soc., 10(1959), 407-415.

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

[10] J.Micic, Y.Seo, S.-E. Takahasi and M.Tominaga, *Inequalities of Furuta and Mond-Pecaric*, Math. Inequal. Appl., 2(1999), 83-111.

[11] B.Mond and J.E.Pečarić, *Convex inequalities in Hilbert space*, Houston J. Math., 19(1993), 405-420.

[12] B.Mond and J.E.Pečarić, *Bounds for Jensen's inequality for several operators*, Houston J. Math., 20(1994), 645-651.

[13] R.Nakamoto and M.Nakamura, *Operator mean and Kantorovich inequality*, Math. Japon., 44(1996), 495-498.

[14] M.Nakamura, *A remark on a paper of Greub and Reinboldt*, Proc. Japan Acad., 36(1960), 198-199.

[15] J.E.Pečarić, T.Furuta, Y.Seo and J.Micic, *Mond-Pečarić Method in Operator Inequalities*, in preperation.

[16] B.Rennie, *An inequality which includes that of Kantorovich*, Amer. Math. Monthly, 70(1963), 982.

[17] S.-E.Takahasi, M.Tsukada, K.Tanahashi and T.Ogiwara, *An inverse type of Jensen's inequality*, Math. Japon., 50(1999), 85-91.

[18] H.Umegaki, *Conditional expectation in an operator algebra, II*, Tohoku Math. J., 8(1956), 86-100.

[19] H.Umegaki, *Conditional expectation in an operator algebra, III*, Kodai Math. Sem. Rep., 11(1959), 51-64.

Department of Mathematics, Osaka Kyoiku University, Asahigaoka, Kashiwara, Osaka 582-8582, Japan.

E-mail address: mfujii@@cc.osaka-kyoiku.ac.jp