An extension of Kantorovich inequality to $n$-operators via the geometric mean by Ando–Li–Mathias

Takeaki Yamazaki

Department of Mathematics, Kanagawa University, Yokohama 221-8686, Japan

Received 14 September 2005; accepted 17 December 2005
Available online 3 February 2006
Submitted by C.-K. Li

Abstract

In this paper, we shall extend Kantorovich inequality. This is an estimate by using the geometric mean of $n$-operators which have been defined by Ando–Li–Mathias in [T. Ando, C. K. Li, R. Mathias, Geometric means, Linear Algebra Appl. 385 (2004) 305–334]. As a related result, we obtain a converse of arithmetic–geometric means inequality of $n$-operators via Kantorovich constant.

© 2006 Elsevier Inc. All rights reserved.

AMS classification: 47A63; 47A64; 47A30

Keywords: Kantorovich inequality; Geometric mean of $n$-operators; Arithmetic–geometric means inequality; Specht’s ratio

1. Introduction

In what follows a capital letter means a bounded linear operator on a complex Hilbert space $\mathcal{H}$. An operator $T$ is said to be positive if $\langle Tx, x \rangle \geq 0$ holds for all $x \in \mathcal{H}$. For an operator $T$ such that $0 < m I \leq T \leq M I$, the following inequality is called “Kantorovich inequality” [6,7]:

$$\langle Tx, x \rangle \langle T^{-1}x, x \rangle \leq \frac{(m + M)^2}{4mM} \quad \text{for } \|x\| = 1. \tag{1.1}$$

We call the constant $\frac{(m + M)^2}{4mM}$ Kantorovich constant. (1.1) is closely related to properties of convex functions, and many authors have given many results and comments [3,5,9,10,12]. It is well known that (1.1) is equivalent to the following form by replacing $x$ with $\frac{T^\frac{1}{2} x}{\|T^\frac{1}{2} x\|}$ in (1.1):

E-mail address: yamazt26@kanagawa-u.ac.jp

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
\[ \langle T^2 x, x \rangle \leq \frac{(m + M)^2}{4mM} \langle Tx, x \rangle^2 \quad \text{for } \|x\| = 1. \]  
(1.1')

For positive invertible operators \( A \) and \( B \), the geometric mean \( A_B^\# B \) of \( A \) and \( B \) is defined as follows [8]:

\[ A_B^\# B = A^{1/2} (A^{-1/2} B A^{-1/2})^{1/2} A^{1/2}. \]

\( A_B^\# B \) is an extension of the geometric mean \( \sqrt{ab} \) of positive numbers \( a \) and \( b \). It is well known that Kantorovich inequality is equivalent to the following inequality [2]: Let \( A \) and \( B \) be positive invertible operators whose spectrums are contained in \([m, M]\) with \( 0 < m < M \). Then

\[ \langle Ax, x \rangle \langle Bx, x \rangle \leq \frac{(m + M)^2}{4mM} \langle A_B^\# Bx, x \rangle^2 \quad \text{for } x \in H. \]  
(1.2)

In this paper, we call it “Kantorovich inequality of 2-operators”.

Very recently, as an extension of \( A_B^\# B \), the geometric mean \( G(A_1, A_2, \ldots, A_n) \) of \( n \)-tuples of positive invertible operators \( A_i \) has been defined by Ando et al. [1] as follows.

**Definition 1** (Geometric mean of \( n \)-operators [1]). Let \( A_i \) be positive invertible operators for \( i = 1, 2, \ldots, n \). Then the geometric mean \( G(A_1, A_2, \ldots, A_n) \) is defined by induction as follows:

(i) \( G(A_1, A_2) = A_1^\# A_2 \).

(ii) Assume that the geometric mean of any \( n - 1 \)-tuple of operators is defined. Let

\[ G((A_j)_{j \neq i}) = G(A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n), \]

and let sequences \( \{A_i^{(r)}\}_{r=0}^{\infty} \) be \( A_i^{(0)} = A_i \) and \( A_i^{(r)} = G(A_{j}^{(r-1)})_{j \neq i} \). If there exists \( \lim_{r \to \infty} A_i^{(r)} \), and it does not depend on \( i \), then we define the geometric mean of \( n \)-operators as

\[ \lim_{r \to \infty} A_i^{(r)} = G(A_1, A_2, \ldots, A_n). \]

In [1], it has been shown that for any positive invertible operators \( A_i \) for \( i = 1, 2, \ldots, n \), there exists \( \lim_{r \to \infty} A_i^{(r)} \) and

\[ \lim_{r \to \infty} A_i^{(r)} = G(A_1, A_2, \ldots, A_n), \]

uniformly. In fact, they have shown it for \( n \)-matrices in [1]. But by their proof, we can understand that the result can be extended to Hilbert space operators.

The geometric mean defined above has the following properties in [1]:

(P1) Consistency with scalars. If \( A_i \) commute with each other, then

\[ G(A_1, A_2, \ldots, A_n) = (A_1 A_2 \cdots A_n)^{1/2}. \]

(P2) Joint homogeneity. For positive numbers \( s_i \),

\[ G(s_1 A_1, s_2 A_2, \ldots, s_n A_n) = (s_1 s_2 \cdots s_n)^{1/2} G(A_1, A_2, \ldots, A_n). \]

(P3) Permutation invariance. For any permutation \( \pi(A_1, A_2, \ldots, A_n) \) of \( (A_1, A_2, \ldots, A_n) \),

\[ G(\pi(A_1, A_2, \ldots, A_n)) = G(A_1, A_2, \ldots, A_n). \]
(P4) Monotonicity. If $A_i \geq B_i > 0$, then $G(A_1, A_2, \ldots, A_n) \geq G(B_1, B_2, \ldots, B_n)$.

(P5) Continuity above. For each $i$, if $\{A_i,k\}_{k=1}^\infty$ are monotonic decreasing sequences converging to $A_i$ as $k \to \infty$, respectively, then
\[ \lim_{k \to \infty} G(A_{1,k}, A_{2,k}, \ldots, A_{n,k}) = G(A_1, A_2, \ldots, A_n). \]

(P6) Congruence invariance. For an invertible operator $S$,
\[ G(S^* A_1 S, S^* A_2 S, \ldots, S^* A_n S) = S^* G(A_1, A_2, \ldots, A_n) S. \]

(P7) Joint concavity. The map $(A_1, A_2, \ldots, A_n) \mapsto G(A_1, A_2, \ldots, A_n)$ is jointly concave, i.e., for $0 < \lambda < 1$,
\[
G(\lambda A_1 + (1-\lambda) B_1, \lambda A_2 + (1-\lambda) B_2, \ldots, \lambda A_n + (1-\lambda) B_n) \\
\geq \lambda G(A_1, A_2, \ldots, A_n) + (1-\lambda) G(B_1, B_2, \ldots, B_n).
\]

(P8) Self-duality. $G(A_1, A_2, \ldots, A_n) = G(A_1^{-1}, A_2^{-1}, \ldots, A_n^{-1})^{-1}$.

(P9) Determinant identity. For positive invertible matrices $A_i$,
\[ \det(A_1, A_2, \ldots, A_n) = (\det A_1 \cdot \det A_2 \cdots \det A_n)^{1/n}. \]

Moreover, $G(A_1, A_2, \ldots, A_n)$ satisfies the arithmetic–geometric means inequality:
\[ G(A_1, A_2, \ldots, A_n) \leq \frac{A_1 + A_2 + \cdots + A_n}{n}. \]

For positive numbers $a_i$, as a converse of arithmetic–geometric means inequality, the following inequality [11] is known: For positive numbers $a_i$ with $0 < m \leq a_i \leq M$,
\[ \frac{a_1 + a_2 + \cdots + a_n}{n} \leq S_h \sqrt{a_1 a_2 \cdots a_n} \tag{1.3} \]
holds, where $h = \frac{M}{m} > 1$ and $S_h = \frac{(h-1)h^{1/h}}{e \log h}$. We call $S_h$ the Specht’s ratio, and there are a lot of properties of Kantorovich constant and Specht’s ratio in [3–5]. We remark that Specht’s ratio in (1.3) is the optimal constant.

In this paper, we shall give an extension of Kantorovich inequality of 2-operators to that of $n$-operators via geometric mean by Ando–Li–Mathias. As a related result of it, we shall discuss an extension of (1.3). These results are estimates via Kantorovich constant. Next, we shall show more precise estimations of them in the 3-tuples of operators case.

2. Main results

**Theorem 2.1.** Let $A_i$ be positive operators for $i = 1, 2, \ldots, n$ satisfying $0 < m I \leq A_i \leq MI$ with $m < M$. Then
\[ \frac{A_1 + A_2 + \cdots + A_n}{n} \leq \left\{ \frac{(m + M)^2}{4mM} \right\}^{\frac{n-1}{2}} G(A_1, A_2, \ldots, A_n). \]

**Theorem 2.2.** Let $A_i$ be positive operators for $i = 1, 2, \ldots, n$ satisfying $0 < m I \leq A_i \leq MI$ with $0 < m < M$. Then
\[ \langle A_1 x, x \rangle \langle A_2 x, x \rangle \cdots \langle A_n x, x \rangle \leq \left\{ \frac{(m + M)^2}{4mM} \right\}^{\frac{n(n-1)}{2}} \langle G(A_1, A_2, \ldots, A_n) x, x \rangle^n \]

holds for all \( x \in \mathcal{H} \).

**Remark.** In [1], the following inequality has been already shown: For positive invertible operators \( A_i \):
\[ \langle G(A_1, A_2, \ldots, A_n) x, x \rangle^n \leq \langle A_1 x, x \rangle \langle A_2 x, x \rangle \cdots \langle A_n x, x \rangle. \]

Hence Theorem 2.2 is a converse of the above inequality.

For positive invertible operators \( A \) and \( B \), let
\[ R(A, B) = \max \{ r(A^{-1} B), r(B^{-1} A) \}, \]
where \( r(T) \) means the spectral radius of \( T \). \( R(A, B) \) was defined in [1], and many nice properties of \( R(A, B) \) were shown as follows: For positive invertible operators \( A, B \) and \( C \),

(i) \( R(A, C) \leq R(A, B) R(B, C) \) (triangle inequality).

(ii) \( R(A, B) \geq 1 \), and \( R(A, B) = 1 \) iff \( A = B \).

(iii) \( \| A - B \| \leq (R(A, B) - 1) \| A \| \).

Moreover, the following inequality holds: For positive invertible operators \( A_i \) and \( B_i \), \( i = 1, 2, \ldots, n \),
\[ R(G(A_1, A_2, \ldots, A_n), G(B_1, B_2, \ldots, B_n)) \leq \left\{ \prod_{i=1}^{n} R(A_i, B_i) \right\}^{\frac{1}{n}}. \]

Especially,
\[ R(A_i^{(1)}, A_k^{(1)}) = R(G((A_j)_{j \neq i}), G((A_j)_{j \neq k})) \leq R(A_i, A_k)^{\frac{1}{n-1}} \quad (2.1) \]
holds.

To prove the above theorems, we shall show the following lemma.

**Lemma 2.3.** Let \( A_i \) be positive invertible operators for \( i = 1, 2, \ldots, n \), and \( h = \max_{i,j} R(A_i, A_j) \). Then
\[ \frac{A_1 + A_2 + \cdots + A_n}{n} \leq \left( \frac{1 + h}{2\sqrt{h}} \right)^{n-1} G(A_1, A_2, \ldots, A_n). \]

**Proof.** We will prove it by induction on \( n \).

In case \( n = 2 \). Let \( X = A_1^{-\frac{1}{2}} A_2 A_1^{-\frac{1}{2}} \), and
\[ X = \int \lambda \, dE_{\lambda} \]
be the spectral decomposition of \( X \). Since \( h = R(A_1, A_2) \), then we have \( \frac{1}{h} \leq X \leq h \) and
\[ \frac{1 + X}{2} = \int \frac{1 + \lambda}{2} \, dE_{\lambda} = \int \frac{1 + \lambda}{2\sqrt{\lambda}} \sqrt{\lambda} \, dE_{\lambda} \leq \int \frac{1 + h}{2\sqrt{h}} \sqrt{\lambda} \, dE_{\lambda} = \frac{1 + h}{2\sqrt{h}} X^{\frac{1}{2}}. \]
Hence we have
\[
1 + \frac{A_1^{-1/2} A_2 A_1^{-1/2}}{2} \leq \frac{1 + h}{2\sqrt{h}} \left( A_1^{-1/2} A_2 A_1^{-1/2} \right)^{1/2}.
\]

Multiplying \( A_1^{1/2} \) to both sides of this inequality we have
\[
\frac{A_1 + A_2}{2} \leq \frac{1 + h}{2\sqrt{h}} A_1 A_2 = \frac{1 + h}{2\sqrt{h}} G(A_1, A_2).
\]

Assume that Lemma 2.3 holds for \( n - 1 \). We have to prove the case \( n \). For positive integer \( r \), we define \( A_i^{(r)}, h_r \) and \( K_r \) as follows:
\[
A_i^{(0)} = A_i \quad \text{and} \quad A_i^{(r)} = G\left( \left( A_j^{(r-1)} \right)_{j \neq i} \right),
\]
\[
h_0 = h \quad \text{and} \quad h_r = \max_{i,j} R\left( A_i^{(r)}, A_j^{(r)} \right),
\]
\[
K_r = \frac{1 + h_r}{2\sqrt{h_r}}.
\]

Then by the induction hypothesis on \( n \), we have
\[
\frac{1}{n} \sum_{i=1}^{n} A_i = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{1}{n-1} \sum_{j \neq i} A_j \right)
\]
\[
\leq \frac{1}{n} \sum_{i=1}^{n} K_0^{n-2} A_i^{(1)}
\]
\[
= K_0^{n-2} \frac{1}{n} \sum_{i=1}^{n} A_i^{(1)}
\]
\[
\leq (K_0 K_1)^{n-2} \frac{1}{n} \sum_{i=1}^{n} A_i^{(2)}
\]
\[
\vdots
\]
\[
\leq (K_0 K_1 \cdots K_r)^{n-2} \frac{1}{n} \sum_{i=1}^{n} A_i^{(r+1)}.
\]

Since
\[
\lim_{r \to \infty} A_i^{(r)} = G(A_1, A_2, \ldots, A_n) \quad \text{for} \quad i = 1, 2, \ldots, n,
\]
we have
\[
\lim_{r \to \infty} \frac{1}{n} \sum_{i=1}^{n} A_i^{(r+1)} = G(A_1, A_2, \ldots, A_n).
\]

So we have only to prove the following inequality:
\[
\limsup_{r \to \infty} K_0 K_1 \cdots K_r \leq K_0^{n-2}.
\]
By (2.1), we have

\[ 1 \leq h_r \leq h_{r-1}^{\frac{1}{n-1}} \leq \cdots \leq h_0^{\left(\frac{1}{n-r}\right)^r}. \]

Since

\[ \frac{1}{2} \left( \frac{1}{x} + x \right) \leq \frac{1}{2} \left( \frac{1}{y^\alpha} + y^\alpha \right) \leq \left\{ \frac{1}{2} \left( \frac{1}{y} + y \right) \right\}^\alpha \]

holds for \( 1 \leq x \leq y^\alpha \) and \( \alpha \in (0, 1] \), we have

\[ K_r = \frac{1 + h_r}{2\sqrt{h_r}} = \frac{1}{2} \left( \frac{1}{\sqrt{h_r}} + \sqrt{h_r} \right) \leq \left\{ \frac{1}{2} \left( \frac{1}{\sqrt{h_0}} + \sqrt{h_0} \right) \right\}^{\left(\frac{1}{n-r}\right)^r} = K_0^{\left(\frac{1}{n-r}\right)^r}. \]

Therefore we obtain

\[ K_0 K_1 \cdots K_r \leq K_0^{1 + \frac{1}{n-1} + \cdots + \left(\frac{1}{n-r}\right)^r} \to K_0^{\frac{n-1}{n}} \text{ as } r \to \infty. \]

Hence we have

\[ \frac{A_1 + A_2 + \cdots + A_n}{n} \leq \left( \frac{1 + h}{2\sqrt{h}} \right)^{n-1} G(A_1, A_2, \ldots, A_n). \]

This completes the proof. \( \square \)

**Proof of Theorem 2.1.** By putting \( h = \frac{M}{m} \) in Lemma 2.3, we obtain Theorem 2.1. \( \square \)

**Proof of Theorem 2.2.** By using Theorem 2.1 and arithmetic–geometric means inequality, we have

\[ \prod_{i=1}^n \langle A_i, x, x \rangle^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^n \langle A_i, x, x \rangle = \langle \frac{1}{n} \sum_{i=1}^n A_i, x, x \rangle \leq \left\{ \frac{(m + M)^2}{4mM} \right\}^{\frac{n-1}{2}} G(A_1, A_2, \ldots, A_n) x, x) \]

This completes the proof. \( \square \)

### 3. More precise estimations

In this section, we shall give more precise estimations than the results shown in Section 2 in the 3-tuples of operators case.

**Theorem 3.1.** Let \( A, B, C \) be positive operators whose spectra are contained in \([m, M]\) with \( 0 < m < M \). Then
\[
\frac{A + B + C}{3} \leq \frac{h^2 - 1}{2h \log h} G(A, B, C),
\]
where \( h = \frac{M}{m} > 1 \).

**Proof.** As in the proof of Lemma 2.3, we have
\[
\frac{A + B + C}{3} \leq K_0 K_1 \cdots K_r \frac{A^{(r+1)} + B^{(r+1)} + C^{(r+1)}}{3},
\]
where
\[
K_r = \frac{h_r + 1}{2\sqrt{h_r}} \quad \text{and} \quad h_r = \max \{ R(A^{(r)}, B^{(r)}), R(B^{(r)}, C^{(r)}), R(C^{(r)}, A^{(r)}) \}.
\]
By (2.1), \( 1 \leq h_r \leq h_{r-1}^\frac{1}{2} \leq \cdots \leq h^\frac{1}{r} \), and we obtain
\[
K_r = \frac{1}{2} \left( \frac{1}{h_r^\frac{1}{2}} + h_r^\frac{1}{2} \right) \leq \frac{1}{2} \left( \frac{1}{h^\frac{1}{2^{r+1}}} + h^\frac{1}{2^{r+1}} \right) = \frac{h^\frac{1}{r} + 1}{2h^\frac{1}{2^{r+1}}},
\]
Hence we have
\[
K_0 K_1 \cdots K_r \leq \frac{h + 1}{2h^\frac{1}{2}} \cdot \frac{h^\frac{1}{2} + 1}{2h^\frac{1}{2}} \cdots \frac{h^\frac{1}{2^{r+1}} + 1}{2h^\frac{1}{2^{r+1}}} = \frac{h^2 - 1}{2h \log h}\]
\[
\rightarrow \frac{h^2 - 1}{2h \log h} \quad \text{as} \quad r \rightarrow \infty,
\]
where the limit is given by \( \lim_{n \rightarrow \infty} n(h^\frac{1}{n} - 1) = \log h \).
This completes the proof. \( \square \)

**Theorem 3.2.** Let \( A, B, C \) be positive invertible operators whose spectra are contained in \([m, M]\) with \( 0 < m < M \). Then
\[
\langle Ax, x \rangle \langle Bx, x \rangle \langleCx, x \rangle \leq \left( \frac{h^2 - 1}{2h \log h} \right)^3 \langle G(A, B, C)x, x \rangle^3,
\]
where \( h = \frac{M}{m} > 1 \).

Theorem 3.2 is easily obtained in the same way as the proof of Theorem 2.2.

**Remark.** In Theorem 3.1, we obtain a more precise constant \( \frac{h^2 - 1}{2h \log h} \) than Theorem 2.1. However, this is not less than the Specht’s ratio in (1.3) as follows: First of all, we shall show
\[
f(h) = (h - 1) \log(h + 1) - (h - 1) \log 2 - h \log h + (h - 1) \geq 0 \quad \text{for} \quad h \geq 1. \quad (3.1)
\]
By easy calculation, we have

\[ f'(h) = \log(h + 1) - \log h - \frac{2}{h + 1} + 1 - \log 2, \]

\[ f''(h) = \frac{h - 1}{h(h + 1)^2} \geq 0 \quad \text{for} \quad h \geq 1. \]

Since \( f'(1) = 0, f(h) \geq 0 \) holds for \( h \geq 1 \). Then by \( f(1) = 0 \), we have (3.1).

Next, (3.1) is equivalent to

\[ \frac{h}{h - 1} \log h - 1 \leq \log \left( \frac{h + 1}{2} \right), \]

i.e.,

\[ \frac{h^{\frac{1}{e}}}{e} \leq \frac{h + 1}{2h} \quad \text{for} \quad h \geq 1. \]

Hence we obtain

\[ S_h = \frac{h - 1}{\log h} \cdot \frac{h^{\frac{1}{e}}}{e} \leq \frac{h - 1}{\log h} \cdot \frac{h + 1}{2h} = \frac{h^2 - 1}{2h \log h}. \]

Acknowledgments

The author is grateful to the referee for his careful reading of the manuscript and kind many comments.

References