SQUARE INEQUALITY AND STRONG ORDER RELATION

TSUYOSHI ANDO

This paper is dedicated to the memory of late Professor Takayuki Furuta

Communicated by M. S. Moslehian

ABSTRACT. It is well-known that for Hilbert space linear operators $0 \leq A$ and $0 \leq C$, inequality $C \leq A$ does not imply $C^2 \leq A^2$. We introduce a strong order relation $0 \leq B \lll A$, which guarantees that $C^2 \leq B^{1/2}AB^{1/2}$ for all $0 \leq C \leq B$, and that $C^2 \leq A^2$ when B commutes with A. Connections of this approach with the arithmetic-geometric mean inequality of Bhatia–Kittaneh as well as the Kantorovich constant of A are mentioned.

1. INTRODUCTION AND THEOREM

Let $B(\mathcal{H})$ denote the space of bounded linear operators on a Hilbert space $\mathcal{H}$. Throughout the paper, a capital letter means an operator in $B(\mathcal{H})$. The order relation $A \geq B$ or equivalently $B \leq A$ for $A, B \in B(\mathcal{H})$ means that both A and B are selfadjoint and $A - B$ is positive (positive semi-definite for matrices). Therefore $A \geq 0$ means that A is positive. Further, $A > 0$ means that $A \geq 0$ and A is invertible, or equivalently $A \geq \mu I$ for some $\mu > 0$, where I is the identity operator in $B(\mathcal{H})$.

It is well-known that $0 \leq C \leq A$ does not imply $C^2 \leq A^2$ in general. We look for a condition on A and B, which guarantees that

$$0 \leq C \leq B \implies C^2 \leq A^2.$$

Let us introduce a *strong order relation* $B \lll A$ for $0 \leq A, B$ as

$$B \lll A \iff PBP \leq A \text{ for all projection } P. \qquad (1.1)$$

Theorem 1.1. *If* $0 \leq C \leq B \lll A$, *then* $C^2 \leq B^{1/2}AB^{1/2}$ *and* $C^2 \leq A^2$ *whenever* $AB = BA$.

Proof. Inequality $0 \leq C \leq B$ is characterized by the relation

$$C = B^{1/2}DB^{1/2} \text{ for some } 0 \leq D \leq I. \qquad (1.2)$$

Since each $0 \leq D \leq I$ can be approximated in norm by convex combinations of projections, and since the map $D \longmapsto DBD$ is convex in the sense that

$$\{\lambda D_1 + (1-\lambda)D_2\}B\{\lambda D_1 + (1-\lambda)D_2\}$$
$$\leq \lambda D_1 B D_1 + (1-\lambda)D_2 B D_2 \text{ for all } 0 \leq \lambda \leq 1$$

we can see from (1.1) and (1.2) that

$$C^2 = B^{1/2} \cdot (DBD) \cdot B^{1/2} \leq B^{1/2}AB^{1/2}.$$

Further, $B^{1/2}AB^{1/2} \leq A^2$ when $AB = BA$. $\qquad\qquad \square$

2. STRONG ORDER RELATION

It is immediate from definition (1.1) that

$$0 \leq C \leq B \lll A \implies C \lll A, \qquad (2.1)$$

and

$$0 \leq B_j \lll A_j \ (j = 1, 2)$$
$$\implies \alpha_1 B_1 + \alpha_2 B_2 \lll \alpha_1 A_1 + \alpha_2 A_2 \text{ for all } \alpha_1, \alpha_2 \geq 0.$$

The following assertion can be verified easily

$$0 \leq A \lll A \iff A = \alpha I \text{ for some } \alpha \geq 0.$$

A little non-trivial fact is that since the square-root map $0 \leq X \longmapsto X^{1/2}$ is order-preserving (see [4, p.127])

$$0 \leq B \lll A \implies B^{1/2} \lll A^{1/2}.$$

This can be seen as follows: Since

$$PB^{1/2}P \leq (PBP)^{1/2}$$

for all $B \geq 0$ and all projections P, we can conclude that

$$0 \leq B \lll A \implies PB^{1/2}P \leq (PBP)^{1/2} \leq A^{1/2}$$
$$\implies B^{1/2} \lll A^{1/2}.$$

To see further properties of the strong order relation, given a projection P, let us consider two maps from $B(\mathcal{H})$ to $B(\mathcal{G})$ with $\mathcal{G} = \mathrm{ran}(P)$, the range space of P. First define $(\mathbf{P})X$ by

$$(\mathbf{P})X := PXP \text{ for all } X \in \mathcal{B}(\mathcal{H})$$

and second $[\mathbf{P}]X$ by

$$[\mathbf{P}]X := PXP - (PXP^\perp) \cdot (P^\perp XP^\perp)^{-1} \cdot (P^\perp XP),$$

where $P^\perp := I - P$.

The map $(\mathbf{P})$ is defined for all X while $[\mathbf{P}]$ is defined only when $P^\perp XP^\perp$ is invertible in $B(\mathcal{G}^\perp)$ where $\mathcal{G}^\perp$ is the ortho-complement of $\mathcal{G}$, or equivalently $\mathcal{G}^\perp = \operatorname{ran}(P^\perp)$.

See [1] for more details about the map $[\mathbf{P}]$. Sometimes we will abuse $(\mathbf{P})X$ and $[\mathbf{P}]X$ as if they are operators in $B(\mathcal{H})$.

It is obvious that, with $I_\mathcal{G}$ the identity operator in $B(\mathcal{G})$,

$$\mu I \le A \le \lambda I \quad \Longrightarrow \quad \mu I_\mathcal{G} \le (\mathbf{P})A \le \lambda I_\mathcal{G}. \tag{2.2}$$

A significant result is the following.

Theorem 2.1. *For all $A > 0$ and all projections P,*

$$([\mathbf{P}]A)^{-1} = (\mathbf{P})(A^{-1}) \text{ and } 0 \le [\mathbf{P}]A \le A.$$

Proof. Along the orthogonal decomposition $\mathcal{H} = \mathcal{G} \oplus \mathcal{G}^\perp$, write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{11} = PAP$, $A_{12} = PAP^\perp$, $A_{21} = P^\perp AP$ and $A_{22} = P^\perp AP^\perp$.

Everything in the assertion comes from the following decomposition:

$$
\begin{aligned}
A &= \begin{bmatrix} I_\mathcal{G} & A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix} \cdot \begin{bmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix} \cdot \begin{bmatrix} I_\mathcal{G} & 0 \\ A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix} \\
&= \begin{bmatrix} I_\mathcal{G} & A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix} \cdot \begin{bmatrix} [\mathbf{P}]A & 0 \\ 0 & (\mathbf{P}^\perp)A \end{bmatrix} \cdot \begin{bmatrix} I_\mathcal{G} & 0 \\ A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix}
\end{aligned}
$$

and the fact that both block operator matrices

$$\begin{bmatrix} I_\mathcal{G} & A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix} \text{ and } \begin{bmatrix} I_\mathcal{G} & 0 \\ A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix}$$

are invertible with respective inverses

$$\begin{bmatrix} I_\mathcal{G} & A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix}^{-1} = \begin{bmatrix} I_\mathcal{G} & -A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix}$$

and

$$\begin{bmatrix} I_\mathcal{G} & 0 \\ A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix}^{-1} = \begin{bmatrix} I_\mathcal{G} & 0 \\ -A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix}.$$

In fact

$$A^{-1} = \begin{bmatrix} I_\mathcal{G} & 0 \\ -A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix} \cdot \begin{bmatrix} ([\mathbf{P}]A)^{-1} & 0 \\ 0 & ((\mathbf{P}^\perp)A)^{-1} \end{bmatrix} \cdot \begin{bmatrix} I_\mathcal{G} & -A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix},$$

and

$$A \ge \begin{bmatrix} I_\mathcal{G} & A_{12}A_{22}^{-1} \\ 0 & I_{\mathcal{G}^\perp} \end{bmatrix} \cdot \begin{bmatrix} [\mathbf{P}]A & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} I_\mathcal{G} & 0 \\ A_{22}^{-1}A_{21} & I_{\mathcal{G}^\perp} \end{bmatrix} = [\mathbf{P}]A.$$

$\square$

Corresponding to (2.2) we have

$$\mu I \leq A \leq \lambda I \quad \implies \quad \mu I_{\mathcal{G}} \leq [\mathbf{P}]A \leq \lambda I_{\mathcal{G}}. \tag{2.3}$$

Corollary 2.2. *For $A, B > 0$,*

$$B \lll A \iff (\mathbf{P})B \leq [\mathbf{P}]A \ \text{ for all projection } P$$
$$\iff A^{-1} \lll B^{-1}.$$

3. EXAMPLES

Given A with $\mu I \leq A \leq \lambda I$ for some $0 < \mu < \lambda$, we try to find reasonable $0 \leq B$ of the form $B = \alpha I - \beta A^{-1}$ with $\alpha, \beta \geq 0$ or $= aA + b$ with $a \geq 0$ and real b for which $B \lll A$.

Theorem 3.1. *Let $\mu I \leq A \leq \lambda I$ with $0 < \mu < \lambda$ and $\alpha, \beta \geq 0$. Then validity of $0 \leq \alpha - \frac{\beta}{t} \leq t$ for all $t \in [\mu, \lambda]$ implies that $\alpha I - \beta A^{-1} \lll A$.*

Proof. Given a projection P, let $X := [\mathbf{P}]A$. Since by (2.3) $\mu I_{\mathcal{G}} \leq X \leq \lambda I_{\mathcal{G}}$ with $\mathcal{G} = \operatorname{ran}(P)$, the assumption implies

$$0 \leq \alpha I_{\mathcal{G}} - \beta X^{-1} \leq X.$$

Since $X^{-1} = (\mathbf{P})(A^{-1})$ by Theorem 2.1 , considering X and X^{-1} as operators in $B(\mathcal{H})$ we have

$$P(\alpha I - \beta A^{-1})P \leq [\mathbf{P}]A \leq A,$$

which is just the assertion. $\square$

Suppose that $\mu I \leq A \leq \lambda I$ with $0 < \mu < \lambda$ and that for $\alpha, \beta \geq 0$

$$0 \leq \alpha - \frac{\beta}{t} \leq t \ \text{ for all } t \in [\mu, \lambda],$$

or equivalently

$$\alpha\mu \leq \beta \quad and \quad h(t) := t^2 - \alpha t + \beta \geq 0 \ \text{ for all } t \in [\mu, \lambda]. \tag{3.1}$$

In this case, define a function $f_{\alpha,\beta}(t)$ by

$$f_{\alpha,\beta}(t) := \alpha - \frac{\beta}{t} \quad for \ t \in [\mu, \lambda]. \tag{3.2}$$

Next determine $a \geq 0$ and real b by the relations

$$a\mu + b = \alpha - \frac{\beta}{\mu} \quad and \quad a\lambda + b = \alpha - \frac{\beta}{\lambda}, \tag{3.3}$$

and define an affine function $g_{\alpha,\beta}(t)$ by

$$g_{\alpha,\beta}(t) := at + b \quad for \ t \in [\mu, \lambda]. \tag{3.4}$$

Corollary 3.2. *Suppose that* (3.1) *is satisfied and that $f_{\alpha,\beta}(t)$ and $g_{\alpha,\beta}(t)$ are defined according to* (3.2) *and* (3.3) *respectively. Then*

$$0 \leq g_{\alpha,\beta}(A) \leq f_{\alpha,\beta}(A) \lll A, \ \text{ so that } \ g_{\alpha,\beta}(A) \lll A.$$

Proof. Since $f_{\alpha,\beta}(t)$ is concave by (3.2) and $g_{\alpha,\beta}(t)$ is affine by (3.4), and by (3.3)

$$g_{\alpha,\beta}(\mu) = f_{\alpha,\beta}(\mu) \quad \text{and} \quad g_{\alpha,\beta}(\lambda) = f_{\alpha,\beta}(\lambda)$$

we can conclude that $g_{\alpha,\beta}(t) \le f_{\alpha,\beta}(t)$ on $[\mu, \lambda]$. Then via functional calculus and by Theorem 3.1 and implication (2.1)

$$0 \le g_{\alpha,\beta}(A) \le f_{\alpha,\beta}(A) \lll A, \quad \text{so that} \quad g_{\alpha,\beta}(A) \lll A.$$

$\square$

In the remaining part of this section, under the assumption on a pair (α, β) as in Corollary 3.2, we will investigate when the extremal cases as $f_{\alpha,\beta}(\mu) = \mu$ or $f_{\alpha,\beta}(\lambda) = \lambda$ occur.

Proposition 3.3. *If $f_{\alpha,\beta}(\mu) = \mu$, then $\mu \le \alpha \le 2\mu$ and $\beta = (\alpha-\mu)\mu$. Conversely if $\mu \le \alpha \le 2\mu$, then the pair (α, β) with $\beta := (\alpha - \mu)\mu$ satisfies condition (3.1) and $f_{\alpha,\beta}(\mu) = \mu$.*

Proof. Since the assumption $\mu = f_{\alpha,\beta}(\mu) = \alpha - \frac{\beta}{\mu}$ implies $\beta = (\alpha - \mu)\mu$, so that $\alpha \ge \mu$. Since by (3.1)

$$h(t) = (t - \mu)\{t - (\alpha - \mu)\} \ge 0 \ \text{ for all } t \in [\mu, \lambda]$$

we have $\alpha - \mu \le \mu$, that is, $\alpha \le 2\mu$.

Conversely, suppose that $\mu \le \alpha \le 2\mu$. Define $\beta := (\alpha - \mu)\mu$. Clearly $\beta \ge \alpha\mu$ and $f_{\alpha,\beta}(\mu) = \mu$. Since $\alpha - \mu \le \mu$, we have $h(t) \ge 0$ on $[\mu, \lambda]$, so that (3.1) is satisfied. $\square$

We notice the following concrete examples.

(i) When $\alpha = 2\mu$ and $\beta = \mu^2$,

$$f_{\alpha,\beta}(t) = \mu(2 - \frac{\mu}{t}) \quad \text{and} \quad g_{\alpha,\beta}(t) = \frac{\mu}{\lambda}\{t + (\lambda - \mu)\}.$$

(ii) When $\alpha = \mu$ and $\beta = 0$, $f_{\alpha,\beta}(t) = g_{\alpha,\beta}(t) = \mu$.

Proposition 3.4. *The requirement $f_{\alpha,\beta}(\lambda) = \lambda$ is possible only when $\lambda \le 2\mu$ or equivalently $2\lambda \le \frac{\lambda^2}{\lambda - \mu}$ and*

$$2\lambda \le \alpha \le \frac{\lambda^2}{\lambda - \mu} \quad \text{and} \quad \beta = \lambda(\alpha - \lambda). \tag{3.5}$$

Conversely when $\lambda \le 2\mu$, any pair (α, β) with (3.5) satisfies condition (3.1) and $f_{\alpha,\beta}(\lambda) = \lambda$.

Proof. The requirement $f_{\alpha,\beta}(\lambda) = \lambda$ implies $\beta = \lambda(\alpha - \lambda)$. On the other hand, condition (3.1)

$$(t - \lambda)\{t - (\alpha - \lambda)\} \ge 0 \ \text{ for all } t \in [\mu, \lambda]$$

implies $\alpha - \lambda \ge \lambda$, whence $\alpha \ge 2\lambda$. Again, since by (3.1) $\alpha\mu \ge \beta = \lambda(\alpha - \lambda)$, we have $\alpha \le \frac{\lambda^2}{\lambda - \mu}$, so that

$$2\lambda \le \alpha \le \frac{\lambda^2}{\lambda - \mu}.$$

The proof of the converse direction is similar. $\square$

We notice the following concrete examples.

(iii) Let $0 < \lambda \leq 2\mu$. When $\alpha := \frac{\lambda^2}{\lambda - \mu}$ and $\beta := \frac{\lambda^2 \mu}{\lambda - \mu}$,

$$f_{\alpha,\beta}(t) = \frac{\lambda^2}{\lambda - \mu}\Big\{1 - \frac{\mu}{t}\Big\} \quad and \quad g_{\alpha,\beta}(t) = \frac{\lambda}{\lambda - \mu}\{t - \mu\}.$$

(iv) Let $0 < \lambda \leq 2\mu$. When $\alpha := 2\lambda$ and $\beta := \lambda^2$,

$$f_{\alpha,\beta}(t) = \lambda\Big\{2 - \frac{\lambda}{t}\Big\} \quad and \quad g_{\alpha,\beta}(t) = \frac{\lambda}{\mu}\Big\{t - (\lambda - \mu)\Big\}.$$

4. CONNECTION WITH KNOWN RESULTS

Bhatia and Kittaneh [3] established a remarkable matrix arithmetic-geometric mean inequality. It says that for any $n \times n$ matrices $A, C \geq 0$ and any unitarily invariant norm $\|\cdot\|$ (see [2, p.91] for definition)

$$\|AC\| \leq \|\{\frac{A + C}{2}\}^2\|.$$

Taking the operator norm, this inequality is extended to the case of Hilbert space operators. Taking A^{-1} in place of A, this theorem for the operator norm says

$$C + A^{-1} \leq 2 \cdot I \quad \Longrightarrow \quad A^{-1}C^2 A^{-1} \leq I \quad \Longrightarrow \quad C^2 \leq A^2,$$

or

$$0 \leq C \leq 2 \cdot I - A^{-1} \quad \Longrightarrow \quad C^2 \leq A^2.$$

Therefore this corresponds to the case that $\alpha = 2, \beta = 1$, $\mu = \frac{1}{2}$ and any number λ with $\lambda I \geq A$.

Suppose that $0 < A$ has maximum spectrum λ and minimum spectrum μ. The numbers λ and μ can be expressed in terms of norms related to A. In fact

$$\lambda = \|A\| \quad and \quad \mu = \|A^{-1}\|^{-1}. \tag{4.1}$$

The number

$$\kappa_A := \frac{(\lambda + \mu)^2}{4\lambda\mu} \tag{4.2}$$

is called the *Kantorvich constant* of A. Then it is clear from (4.1) and (4.2) that

$$\kappa_A = \frac{(\|A\| \cdot \|A^{-1}\| + 1)^2}{4\|A\| \cdot \|A^{-1}\|}.$$

The following fact has been known (see [4, Chapter III] for more detail):

Theorem 4.1. *For $A > 0$,*

$$0 \leq C \leq A \quad \Longrightarrow \quad C^2 \leq \kappa_A \cdot A^2.$$

Let us show how this can be incorporated into our theory. The following proposition can be checked immediately.

Proposition 4.2. *When* $\alpha = \frac{4\lambda\mu}{\lambda+\mu}$ *and* $\beta = \frac{4\lambda^2\mu^2}{(\lambda+\mu)^2}$ *the pair* (α, β) *satisfies condition* (3.1) *and*

$$f_{\alpha,\beta}(t) = \frac{4\lambda\mu}{\lambda + \mu}\left\{1 - \frac{\lambda\mu}{(\lambda + \mu)t}\right\} \quad and \quad g_{\alpha,\beta}(t) = \frac{4\lambda\mu}{(\lambda + \mu)^2} \cdot t \ = \ \kappa_A^{-1}t.$$

Therefore $k_A^{-1} \cdot A \lll A.$

Now Theorem 4.1 is deduced from Proposition 4.2 and Theorem 1.1 as follows:

$$\begin{aligned} 0 \leq C \leq A \ &\implies\ \kappa_A^{-1}C \leq \kappa_A^{-1}A \lll A \\ &\implies\ \kappa_A^{-2}C^2 \leq \kappa_A^{-1}A^2 \implies\ C^2 \leq \kappa_A \cdot A^2. \end{aligned}$$

Notice that the above argument shows that

$$0 \leq C \leq \kappa_A^{-1/2} \cdot A \quad \implies \quad C^2 \leq A^2.$$

References

1. T. Ando, *Schur complements and matrix inequalities: Operator-theoretic approach*, pp. 137-162, in *The Schur complement and its applications*, edited by F. Zhang, Springer, New York, 2005.
2. R. Bhatia, *Matrix Analysis*, Springer, New York-Berlin-Heidelberg, 1997.
3. R. Bhatia and F. Kittaneh, *Notes on matrix arithmetic-geometric mean inequalities*, Linear Algebra Appl. **308** (2000), 203–211.
4. T. Furuta, *Invitation to Linear Operators*, Taylor & Francis, London and New York, 2001.

HOKKAIDO UNIVERSITY (EMERITUS), SAPPORO 060, JAPAN
E-mail address: ando@es.hokudai.ac.jp