Chapter 1

Fundamental inequalities and Mond-Pečarić method

In this chapter, we have given a very brief and rapid review of some basic topics in Jensen's inequality for positive linear maps and the Kantorovich inequality for several types. We present some basic ideas and the viewpoints of the Mond-Pečarić method for convex functions.

1.1 Classical Jensen's inequality

In this section, we introduce a classical Jensen's inequality associated with a convex function, and naturally extend it to an operator version. First we introduce some notations.

If a complex vector space H having the inner product is complete with respect to the distance $d(x, y) = \|x - y\|$ defined by the norm $\|x\| := (x, x)^{1/2}$, then H is called a Hilbert space. A linear operator A on a Hilbert space H is said to be bounded if

$$\|A\| := \sup\{\|Ax\| : \|x\| \leq 1, x \in H\} < \infty.$$

Then $\|A\|$ is said to be the operator norm of A. The adjoint operator A^* of A is defined by $(Ax, y) = (x, A^*y)$ for $x, y \in H$. Then it follows that $\|A\| = \|A^*\| = \|A^*A\|^{1/2}$. In an algebra of all linear operators ($H \to H$) on a Hilbert space H with the operator norm, we denote by

$\mathcal{B}(H)$ a semi-algebra of all bounded (i.e., continuous) linear operators on H. The spectrum of an operator A is the set

$$\mathsf{Sp}(A) = \{\lambda \in \mathbb{C} : A - \lambda 1_H \quad \text{is not invertible in } \mathcal{B}(H)\}.$$

The spectrum $\mathsf{Sp}(A)$ is nonempty and compact. A bounded linear operator A on a Hilbert space H is said to be selfadjoint if $A = A^*$. An operator $A \in \mathcal{B}(H)$ is selfadjoint if and only if $(Ax, x) \in \mathbb{R}$ for every vector $x \in H$. We denote by $\mathcal{B}_h(H)$ a semi-space of all selfadjoint operators in $\mathcal{B}(H)$.

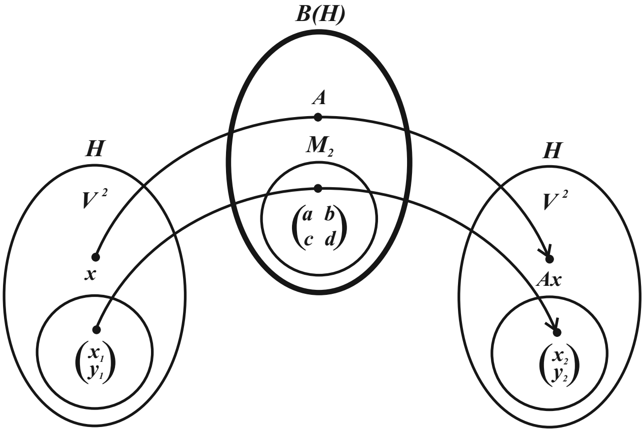

Figure 1.1: Graphic chart of space $\mathcal{B}(H)$

We introduce a partial order in $\mathcal{B}_h(H)$ as follows:

Definition 1.1 *An operator $A \in \mathcal{B}_h(H)$ is* **positive semi-definite** , (*simply,* **positive**) *and we write $A \geq 0$, if $(Ax, x) \geq 0$ for every vector $x \in H$. An operator $A \in \mathcal{B}(H)$ is positive if and only if $A = B^*B$ for some operator $B \in \mathcal{B}(H)$.*

For operators $A, B \in \mathcal{B}_h(H)$ we write $A \leq B$ (or $B \geq A$) if $B - A \geq 0$, i.e., $(Bx, x) \geq (Ax, x)$ for every vector $x \in H$. We call it the **operator order**. *In particular, for some scalars m and M, we write $m1_H \leq A \leq M1_H$ if $m \leq (Ax, x) \leq M$ for every unit vector $x \in H$. Notice that for a selfadjoint operator A, $\mathsf{Sp}(A) \subset [m, M]$ implies $m1_H \leq A \leq M1_H$.*

A positive semi-definite operator $A \in \mathcal{B}_h(H)$ is **positive definite** (**strictly positive**) *and we write $A > 0$, if there is a real number $m > 0$ such that $A \geq m1_H$.*

We denote by $\mathcal{B}^+(H)$ the set of all positive operators and $\mathcal{B}^{++}(H)$ the set of all strictly positive operators (or positive invertible operators) in $\mathcal{B}_h(H)$. The set $\mathcal{B}^+(H)$ is the convex cone contained in $\mathcal{B}_h(H)$.

Now, we review the continuous functional calculus. A rudimentary functional calculus for an operator A can be defined as follows: For a polynomial $p(t) = \sum_{j=0}^{k} \alpha_j t^j$, define

$$p(A) = \alpha_0 1_H + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_k A^k.$$

The mapping $p \to p(A)$ is a homomorphism from the algebra of polynomials to the algebra of operators. The extension of this map to larger algebras of functions is really significant in operator theory.

Let A be a selfadjoint operator on a Hilbert space H. Then the Gelfand map establishes a $*$-isometrically isomorphism Φ between the set $C(\mathrm{Sp}(A))$ of all continuous functions on $\mathrm{Sp}(A)$ and C*-algebra $C^*(A)$ generated by A and the identity operator 1_H on H as follows: For $f, g \in C(\mathrm{Sp}(A))$ and $\alpha, \beta \in \mathbb{C}$

(i) $\Phi(\alpha f + \beta g) = \alpha\Phi(f) + \beta\Phi(g)$.

(ii) $\Phi(fg) = \Phi(f)\Phi(g)$ and $\Phi(\overline{f}) = \Phi(f)^*$.

(iii) $\|\Phi(f)\| = \|f\|(:= \sup_{t\in \mathrm{Sp}(A)} |f(t)|)$.

(iv) $\Phi(f_0) = 1_H$ and $\Phi(f_1) = A$, where $f_0(t) = 1$ and $f_1(t) = t$.

With this notation, we define

$$f(A) = \Phi(f)$$

for all $f \in C(\mathrm{Sp}(A))$ and we call it the continuous functional calculus for a selfadjoint operator A. This map is an extension of $p(A)$ for a polynomial p. The continuous functional calculus is applicable. For example, if A is a positive operator and $f_{1/2}(t) = \sqrt{t}$, then $A^{1/2} = f_{1/2}(A)$. If A is a selfadjoint operator and $f(t)$ is a real valued continuous function on $\mathrm{Sp}(A)$ such that $f(t) \geq 0$ on $\mathrm{Sp}(A)$, then $f(A) \geq 0$, i.e., $f(A)$ is a positive operator. Moreover, if $g(t)$ is a real valued continuous function on $\mathrm{Sp}(A)$ such that $f(t) \geq g(t)$ on $\mathrm{Sp}(A)$, then $f(A) \geq g(A)$.

Next, we shall introduce a spectral decomposition theorem for selfadjoint, bounded linear operators on a Hilbert space H. For the sake of convenience, we recall the following well known diagonalization of Hermitian matrices in matrix theory.

If A is a Hermitian $k \times k$ matrix, then there exists a unitary matrix U (i.e., $U^*U = UU^* = 1_k$) such that

$$A = U^*\Lambda U, \tag{1.1}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_k)$ and the $\lambda_i (\in \mathbb{R})$ are the eigenvalues of A. If we put

$$\begin{aligned} E_1 &= U^*\mathrm{diag}(1,0,\cdots,0)U, \quad E_2 = U^*\mathrm{diag}(1,1,0,\cdots,0)U \\ &\cdots \\ E_k &= U^*\mathrm{diag}(1,1,\cdots,1)U, \end{aligned}$$

then (1.1) can be rewritten as follows:

$$A = \lambda_1 E_1 + \lambda_2(E_2 - E_1) + \cdots + \lambda_k(E_k - E_{k-1}) = \sum_{j=1}^{k} \lambda_j \Delta E_j, \tag{1.2}$$

where $\Delta E_j = E_j - E_{j-1}$ and $E_0 = 0$. If $f(t)$ is a real valued continuous function on the spectrum $\mathsf{Sp}(A)$, then $f(A)$ may be defined by

$$f(A) = \sum_{j=1}^{k} f(\lambda_j)\Delta E_j. \tag{1.3}$$

This result can be generalized to selfadjoint operators on a Hilbert space H.

Let A be a selfadjoint operator on a Hilbert space H and $f(t)$ a real valued continuous function defined on an interval $[m, M]$, where $m = \inf_{\|x\|=1}(Ax, x)$ and $M = \max_{\|x\|=1}(Ax, x)$. Then A can be expressed as follows:

$$A = \int_{m-0}^{M} \lambda \, dE_\lambda \tag{1.4}$$

where $\{E_\lambda : \lambda \in R\}$ is a family of projections such that $E_\lambda \le E_\mu$ if $\lambda \le \mu$, $E_{\lambda+0} = E_\lambda$, $E_{-\infty} = 0$ and $E_\infty = 1_H$. Since a selfadjoint operator A on a Hilbert space H is an extension of a selfadjoint matrix, (1.4) can be naturally considered as an extension of (1.2). Therefore, we have an extension of (1.3) under the above situation as follows:

$$f(A) = \int_{m-0}^{M} f(\lambda) dE_\lambda. \tag{1.5}$$

Next, we shall introduce a classical Jensen's inequality as an inequality associated with a convex function:

Theorem 1.1 (CLASSICAL JENSEN'S INEQUALITY) *If $f(t)$ is a convex function on an interval $[m, M]$ for some scalars $m < M$, then for every $x_1, x_2, \cdots, x_k \in [m, M]$ and every positive real numbers $t_1, t_2, \cdots, t_k$ with $\sum_{j=1}^{k} t_j = 1$,*

$$f\left(\sum_{j=1}^{k} t_j x_j\right) \le \sum_{j=1}^{k} t_j f(x_j). \tag{1.6}$$

Proof. Since $f(t)$ is convex, then for each point $(s, f(s))$ there exists a real number l such that

$$l(x-s) + f(s) \le f(x) \quad \text{for all} \quad x \in [m, M]. \tag{1.7}$$

Put $s_0 = \sum_{j=1}^{k} t_j x_j \in [m, M]$, then it follows from (1.7) that

$$l(x_j - s_0) + f(s_0) \le f(x_j) \quad \text{for} \quad j = 1, 2, \cdots, k.$$

Multiplying this inequality with $t_j \in \mathbb{R}_+$ and summing of j we have

$$\sum_{j=1}^{k} t_j(l(x_j - s_0) + f(s_0)) \le \sum_{j=1}^{k} t_j f(x_j).$$

Since

$$\sum_{j=1}^{k} t_j(l(x_j - s_0) + f(s_0)) = l\left(\sum_{j=1}^{k} t_j x_j - s_0 \sum_{j=1}^{k} t_j\right) + f(s_0) = f(s_0),$$

we have a desired inequality. □

We rephrase it under matrix situation. If we put

$$A = \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} \sqrt{t_1} \\ \vdots \\ \sqrt{t_n} \end{pmatrix},$$

then a classical Jensen's inequality (1.6) in Theorem 1.1 is expressed as

$$f((Ax, x)) \leq (f(A)x, x) \qquad \text{for every unit vector } x.$$

The following theorem is an operator version of Theorem 1.1 (classical Jensen's inequality).

Theorem 1.2 *Let $A \in \mathcal{B}_h(H)$ be a selfadjoint operator with $\mathrm{Sp}(A) \subset [m, M]$ for some scalars $m < M$. If $f(t)$ is a convex function on $[m, M]$, then*

$$f((Ax, x)) \leq (f(A)x, x) \tag{1.8}$$

for every unit vector $x \in H$.

Proof. If we put $s = (Ax, x)$, then $m \leq s \leq M$. For a given $\epsilon > 0$, there exist a straight line $l(t)$ such that (i) $l(t) \leq f(t)$ for all $t \in [m, M]$ and (ii) $l(s) \geq f(s) - \varepsilon$. Then (i) implies $l(A) \leq f(A)$. Hence we have

$$(f(A)x, x) \geq (l(A)x, x) = l(s) \geq f(s) - \varepsilon$$

for every unit vector $x \in H$. Since ε is an arbitrary, we have $(f(A)x, x) \geq f((Ax, x))$. □

The following theorem is a multiple operator version of Theorem 1.2:

Theorem 1.3 *Let $A_j \in \mathcal{B}_h(H)$ be selfadjoint operators with $\mathrm{Sp}(A_j) \subset [m, M]$ $(j = 1, 2, \cdots, k)$ for some scalars $m < M$. Let $x_1, x_2, \cdots, x_k \in H$ be any finite number of vectors such that $\sum_{j=1}^{k} \|x_j\|^2 = 1$. If $f(t)$ is a convex function on $[m, M]$, then*

$$f\left(\sum_{j=1}^{k} (A_j x_j, x_j)\right) \leq \sum_{j=1}^{k} (f(A_j)x_j, x_j). \tag{1.9}$$

Proof. If we put

$$\tilde{A} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix} \quad \text{and} \quad \tilde{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix},$$

then we have $\mathrm{Sp}(\tilde{A}) \subset [m,M]$, $\|\tilde{x}\| = 1$ and $\sum_{j=1}^{k}(A_j x_j, x_j) = (\tilde{A}\tilde{x}, \tilde{x})$. It follows from Theorem 1.2 that $f((\tilde{A}\tilde{x},\tilde{x})) \le (f(\tilde{A})\tilde{x},\tilde{x})$ and hence we have (1.9). □

As a special case of Theorem 1.2, we have the following Hölder-McCarthy inequality.

Theorem 1.4 (HÖLDER-MCCARTHY INEQUALITY) *Let $A \in \mathcal{B}_h(H)$ be a positive operator on a Hilbert space H. Then*

(i) $(A^r x, x) \ge (Ax, x)^r$ for all $r > 1$ and every unit vector $x \in H$.

(ii) $(A^r x, x) \le (Ax, x)^r$ for all $0 < r < 1$ and every unit vector $x \in H$.

(iii) If A is invertible, then $(A^r x, x) \ge (Ax, x)^r$ for all $r < 0$ and every unit vector $x \in H$.

Proof. Since the power function $f(t) = t^r$ is convex for $r > 1$ or $r < 0$, and concave for $0 < r < 1$, this theorem follows from Theorem 1.2. □

1.2 Operator convexity

In this section, we consider another operator version of a classical Jensen's inequality (1.6) in Theorem 1.1. We rephrase it under another matrix situation. If we put

$$A = \begin{pmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \sqrt{t_1} & 0 & \cdots & 0 \\ \vdots & & & \\ \sqrt{t_n} & 0 & \cdots & 0 \end{pmatrix},$$

then a classic Jensen's inequality is expressed as

$$f(V^*AV) \le V^* f(A)V.$$

The formulation offers a fresh insight into the noncommutative case. Its noncommutative version is considered in various way. We shall start with the following definition.

Definition 1.2 *A real valued continuous function $f(t)$ on an interval I is said to be* **operator convex** *(resp.* **operator concave***) if*

$$f((1-\lambda)A + \lambda B) \le (1-\lambda)f(A) + \lambda f(B) \tag{1.10}$$

(resp.

$$f((1-\lambda)A+\lambda B) \geq (1-\lambda)f(A)+\lambda f(B)) \tag{1.11}$$

for all $\lambda \in [0,1]$ *and for every selfadjoint operator A and B on a Hilbert space H whose spectra are contained in I. Also, the condition (1.10) can be replaced by the more special condition*

$$f\left(\frac{A+B}{2}\right) \leq \frac{f(A)+f(B)}{2}. \tag{1.12}$$

Notice that a function f is operator concave if $-f$ *is operator convex.*

A real valued continuous function $f(t)$ *on an interval I is said to be* **operator monotone** *if it is monotone with respect to the operator order, i.e.,*

$$A \leq B \quad \text{with } \mathsf{Sp}(A), \mathsf{Sp}(B) \subset I \quad \text{implies} \quad f(A) \leq f(B).$$

Before we present basic examples of such functions, we prove some lemmas needed later.

Lemma 1.5 *If* $A \in \mathcal{B}^+(H)$ *is positive, then* $X^*AX \geq 0$ *for every* $X \in \mathcal{B}(H)$.

Proof. For every vector $x \in H$, $(X^*AXx, x) = (AXx, Xx) \geq 0$. □

Lemma 1.6 *If* $A \in \mathcal{B}_h(H)$ *is selfadjoint and U is unitary, i.e.* $U^*U = UU^* = 1_H$, *then* $f(U^*AU) = U^*f(A)U$ *for every* $f \in C(\mathsf{Sp}(A))$.

Proof. Put $B = U^*AU$, then B is selfadjoint and $\mathsf{Sp}(B) = \mathsf{Sp}(A)$. Since $B^m = U^*A^mU$ for every integer $m \geq 0$, we have $p(B) = U^*p(A)U$ for every polynomial $p(t)$. Since there exist polynomials $\{p_j\}$ such that $\|f - p_j\| \mapsto 0$ as $j \to \infty$ for a given $f \in C(\mathsf{Sp}(A))$, we have

$$\begin{aligned}\|f(U^*AU) \;-\; U^*f(A)U\| &\leq \|f(U^*AU) - p_j(U^*AU)\| \\ &+ \|p_j(U^*AU) - U^*p_j(A)U\| + \|U^*p_j(A)U - U^*f(A)U\| \mapsto 0\end{aligned}$$

as $j \to \infty$ and so $f(U^*AU) = U^*f(A)U$. □

Lemma 1.7 *If* $A \in \mathcal{B}(H)$ *and* $f \in C([0, \|A\|^2])$, *then* $Af(A^*A) = f(AA^*)A$.

Proof. Since $A(A^*A)^n = (AA^*)^nA$ for every integer $n \geq 0$, we have $Ap(A^*A) = p(AA^*)A$ for every polynomial $p(t)$. Since there exist polynomials $\{p_j\}$ such that $\|f - p_j\| \mapsto 0$ as $j \to \infty$ for a given $f \in C([0, \|A\|^2])$, we obtain $Af(A^*A) = f(AA^*)A$. □

Now, we study basic examples of such functions.

Example 1.1 *The function* $f(t) = \alpha + \beta t$ *is operator monotone on every interval for all* $\alpha \in \mathbb{R}$ *and* $\beta \geq 0$. *It is operator convex for all* $\alpha, \beta \in R$.

Example 1.2 *If f, g are operator monotone, and if α, β are positive real numbers, then $\alpha f + \beta g$ is also operator monotone. If f_n are operator monotone and $f_n(t) \to f(t)$ as $n \to \infty$, then f is also operator monotone.*

Example 1.3 *The function $f(t) = t^2$ on $[0, \infty)$ is not operator monotone though it is monotone increasing. As a matter of fact, if we put*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

then $A \geq B$ and $A^2 \not\geq B^2$ since

$$A^2 - B^2 = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix} \not\geq 0$$

Example 1.4 *The function $f(t) = t^2$ is operator convex on every interval. To see it, for any selfadjoint operators A and B,*

$$\frac{A^2 + B^2}{2} - \left(\frac{A+B}{2}\right)^2 = \frac{1}{4}(A^2 + B^2 - AB - BA) = \frac{1}{4}(A - B)^2 \geq 0.$$

This shows that the function $f(t) = \alpha t^2 + \beta t + \gamma$ is operator convex for all $\beta, \gamma \in R$, $\alpha \geq 0$.

Example 1.5 *The function $f(t) = t^3$ on $[0, \infty)$ is not operator convex though it is convex on $[0, \infty)$. In fact, if we put*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

then we have

$$\frac{A^3 + B^3}{2} - \left(\frac{A+B}{2}\right)^3 = \frac{1}{4}\begin{pmatrix} 11 & 9 \\ 9 & 7 \end{pmatrix} \not\geq 0.$$

Example 1.6 *The function $f(t) = \frac{1}{t}$ is operator convex on $(0, \infty)$ and $g(t) = -\frac{1}{t}$ is operator monotone on $(0, \infty)$. In fact, for any positive invertible operators A and B*

$$\begin{aligned}
&\frac{A^{-1} + B^{-1}}{2} - \left(\frac{A+B}{2}\right)^{-1} \\
&= \frac{A^{-1} + B^{-1} - 4(A(A^{-1} + B^{-1})B)^{-1}}{2} \\
&= \frac{A^{-1} + B^{-1} - 4B^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}}{2} \\
&= \frac{(A^{-1} + B^{-1} - 2B^{-1})(A^{-1} + B^{-1})^{-1}(2A^{-1} - (A^{-1} + B^{-1}))}{2} \\
&= \frac{(A^{-1} - B^{-1})(A^{-1} + B^{-1})^{-1}(A^{-1} - B^{-1})}{2} \geq 0.
\end{aligned}$$

The last inequality holds by Lemma 1.5.

This fact shows that $f(t) = \frac{1}{t}$ is operator convex.

Next, let $A \geq B \geq 0$. Then $1_H \geq A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. Taking inverse both sides, we have $I_H \leq A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}}$ and hence $A^{-1} \leq B^{-1}$. Therefore it follows that $-A^{-1} \geq -B^{-1}$ and hence $g(t) = -\frac{1}{t}$ is operator monotone on $(0, \infty)$.

To relate this, we introduce the following famous Löwner-Heinz inequality established in 1934.

Theorem 1.8 (LÖWNER-HEINZ INEQUALITY) *Let A and B be positive operators on a Hilbert space H. If $A \geq B \geq 0$, then $A^r \geq B^r$ for all $r \in [0, 1]$.*

We need some elementary results in operator theory in order to prove it. The spectral radius of an operator A is defined as

$$r(A) = \max\{|\lambda| : \lambda \in \mathrm{Sp}(A)\}.$$

Notice that $r(A) \leq \|A\|$ and $r(A) = \|A\|$ if A is a selfadjoint operator. Moreover, it follows that $r(AB) = r(BA)$ for all $A, B \in B(H)$, since $\mathrm{Sp}(AB)\backslash\{0\} = \mathrm{Sp}(BA)\backslash\{0\}$. Also, if A is positive, then $A \leq 1_H$ if and only if $r(A) \leq 1$. An operator A is a contraction ($\|A\| \leq 1$) if and only if $A^*A \leq 1_H$.

Proof of Theorem 1.8. Let $A \geq B \geq 0$. Suppose that A is invertible. Put

$$\Delta = \{r \in R : A^r \geq B^r\}.$$

Then the set Δ is closed since $r \to A^r, B^r$ are norm continuous and $0 \in \Delta$ obviously. The hypothesis $A \geq B \geq 0$ ensures $1 \in \Delta$. Therefore, to prove $[0, 1] \subset \Delta$ is sufficient to show that $r, s \in \Delta$ implies $\frac{r+s}{2} \in \Delta$.

If $r \in \Delta$, then $1_H \geq A^{-\frac{r}{2}}B^rA^{-\frac{r}{2}} = \left(B^{\frac{r}{2}}A^{-\frac{r}{2}}\right)^* \left(B^{\frac{r}{2}}A^{-\frac{r}{2}}\right)$ and hence

$$\left\|B^{\frac{r}{2}}A^{-\frac{r}{2}}\right\| \leq 1.$$

By the same argument, if $s \in \Delta$, then $\left\|B^{\frac{s}{2}}A^{-\frac{s}{2}}\right\| \leq 1$.

So, we have

$$\begin{aligned}
&\left\|A^{\frac{-(r+s)}{4}}B^{\frac{r+s}{2}}A^{\frac{-(r+s)}{4}}\right\| \\
&= r\left(A^{\frac{-(r+s)}{4}}B^{\frac{r+s}{2}}A^{\frac{-(r+s)}{4}}\right) \qquad \text{by} \quad A^{\frac{-(r+s)}{4}}B^{\frac{r+s}{2}}A^{\frac{-(r+s)}{4}} \quad \text{is positive} \\
&= r\left(A^{\frac{r-s}{4}}A^{\frac{-(r+s)}{4}}B^{\frac{r+s}{2}}A^{\frac{-(r+s)}{4}}A^{\frac{s-r}{4}}\right) \qquad \text{by} \quad r(ST) = r(TS) \\
&= r\left(A^{\frac{-s}{2}}B^{\frac{r+s}{2}}A^{\frac{-r}{2}}\right) \\
&\leq \left\|A^{\frac{-s}{2}}B^{\frac{r+s}{2}}A^{\frac{-r}{2}}\right\| \qquad \text{by} \quad r(X) \leq \|X\| \\
&\leq \left\|B^{\frac{r}{2}}A^{-\frac{r}{2}}\right\| \left\|B^{\frac{s}{2}}A^{-\frac{s}{2}}\right\| \leq 1.
\end{aligned}$$

Therefore we have

$$A^{\frac{-(r+s)}{4}} B^{\frac{r+s}{2}} A^{\frac{-(r+s)}{4}} \leq 1_H$$

and hence

$$A^{\frac{r+s}{2}} \geq B^{\frac{r+s}{2}}, \quad \text{i.e.,} \quad \frac{r+s}{2} \in \Delta.$$

This fact shows the theorem under the assumption that A is invertible.

Suppose that A is not invertible. For each $\varepsilon > 0$, $A+\varepsilon 1_H$ is invertible and $A+\varepsilon 1_H \geq B$. Therefore it follows from above argument that

$$(A+\varepsilon 1_H)^r \geq B^r \qquad \text{for all} \quad 0 \leq r \leq 1.$$

By letting $\varepsilon \to 0$, we have the desired inequality $A^r \geq B^r$. □

Now, we go back to Jensen's inequality. We show some characterizations of operator convexity and operator monotonicity based on the ideas due to Hansen-Pedersen. This leads to some conditions equivalent to Jensen's inequality.

Theorem 1.9 (JENSEN'S OPERATOR INEQUALITY) *Let H and K be Hilbert space. Let f be a real valued continuous function on an interval I. Let A and A_j be selfadjoint operators on H with spectra contained in I ($j = 1, 2, \cdots, k$). Then the following conditions are mutually equivalent:*

(i) *f is operator convex on I.*

(ii) *$f(C^*AC) \leq C^*f(A)C$ for every $A \in \mathcal{B}_h(H)$ and isometry $C \in \mathcal{B}(K,H)$, i.e., $C^*C = 1_K$.*

(iii) *$f(C^*AC) \leq C^*f(A)C$ for every $A \in \mathcal{B}_h(H)$ and isometry $C \in \mathcal{B}(H)$.*

(iv) *$f\left(\sum_{j=1}^k C_j^*A_jC_j\right) \leq \sum_{j=1}^k C_j^*f(A_j)C_j$ for every $A_j \in \mathcal{B}_h(H)$ and $C_j \in \mathcal{B}(K,H)$ with $\sum_{j=1}^k C_j^*C_j = 1_K$ ($j = 1, \cdots, k$).*

(v) *$f\left(\sum_{j=1}^k C_j^*A_jC_j\right) \leq \sum_{j=1}^k C_j^*f(A_j)C_j$ for every $A_j \in \mathcal{B}_h(H)$ and $C_j \in \mathcal{B}(H)$ with $\sum_{j=1}^k C_j^*C_j = 1_H$ ($j = 1, \cdots, k$).*

(vi) *$f\left(\sum_{j=1}^k P_jA_jP_j\right) \leq \sum_{j=1}^k P_jf(A_j)P_j$ for every $A_j \in \mathcal{B}_h(H)$ and projection $P_j \in \mathcal{B}_h(H)$ with $\sum_{j=1}^k P_j = 1_H$ ($j = 1, \cdots, k$).*

Proof. (i) $\Rightarrow$ (ii): Let $X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \in \mathcal{B}_h(H \oplus K)$ for some selfadjoint operator $B \in \mathcal{B}_h(K)$ with $\sigma(B) \subset I$ and

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

where $D = \sqrt{1_H - CC^*}$. Since $C^*D = \sqrt{1_K - C^*C}C^* = 0 \in \mathcal{B}_h(H,K)$ and $DC = C\sqrt{1_K - C^*C} = 0 \in \mathcal{B}_h(K,H)$, it follows that both U and V are unitary operators of $K \oplus H$ onto $H \oplus K$. Then

$$U^*XU = \begin{pmatrix} C^*AC & C^*AD \\ DAC & DAD + CBC^* \end{pmatrix}$$

and

$$V^*XV = \begin{pmatrix} C^*AC & -C^*AD \\ -DAC & DAD + CBC^* \end{pmatrix}.$$

So, we have

$$\begin{pmatrix} C^*AC & 0 \\ 0 & D^*AD + CBC^* \end{pmatrix} = \frac{U^*XU + V^*XV}{2}.$$

Hence, it follows from the operator convexity of f and Lemma 1.6 that

$$\begin{aligned} \begin{pmatrix} f(C^*AC) & 0 \\ 0 & f(D^*AD + CBC^*) \end{pmatrix} &= f\begin{pmatrix} C^*AC & 0 \\ 0 & D^*AD + CBC^* \end{pmatrix} \\ &= f\left(\frac{U^*XU + V^*XV}{2}\right) \\ &\le \frac{f(U^*XU) + f(V^*XV)}{2} = \frac{U^*f(X)U + V^*f(X)V}{2} \\ &= \begin{pmatrix} C^*f(A)C & 0 \\ 0 & D^*f(A)D + Cf(B)C^* \end{pmatrix}. \end{aligned}$$

Thus we have $f(C^*AC) \le C^*f(A)C$ by seeing the (1,1)-components.

$(ii) \Rightarrow (iv)$: Let

$$X = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix} \in \mathcal{B}_h(H \oplus \cdots \oplus H),\ \tilde{C} = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_k \end{pmatrix} \in \mathcal{B}(K, H \oplus \cdots \oplus H).$$

Then $\tilde{C}^*\tilde{C} = 1_K$ and hence it follow from (ii) that

$$f(\sum_{j=1}^k C_j^*A_jC_j) = f(\tilde{C}^*X\tilde{C}) \le \tilde{C}^*f(X)\tilde{C} = \sum_{j=1}^k C_j^*f(A_j)C_j.$$

$(iv) \Rightarrow (vi)$: Obviously.

$(vi) \Rightarrow (i)$: Let A and B be selfadjoint operators with spectrum in I and let $0 \le t \le 1$. Let $X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, $P = \begin{pmatrix} 1_H & 0 \\ 0 & 0 \end{pmatrix}$ and $U = \begin{pmatrix} \sqrt{1-t} & -\sqrt{t} \\ \sqrt{t} & \sqrt{1-t} \end{pmatrix}$. Then U is a unitary operator on $H \oplus H$. Thus we have

$$\begin{aligned} \begin{pmatrix} f((1-t)A + tB) & 0 \\ 0 & f(tA + (1-t)B) \end{pmatrix} &= f(PU^*XUP + (1_{H\otimes H} - P)U^*XU(1_{H\otimes H} - P)) \\ &\le Pf(U^*XU)P + (1_{H\otimes H} - P)f(U^*XU)(1_{H\otimes H} - P) \\ &= PU^*f(X)UP + (1_{H\otimes H} - P)U^*f(X)U(1_{H\otimes H} - P) \\ &= \begin{pmatrix} (1-t)f(A) + tf(B) & 0 \\ 0 & tf(A) + (1-t)f(B) \end{pmatrix}. \end{aligned}$$

Hence f is operator convex on I by seeing the (1,1)-components.

Therefore, we proved the implications $(i) \Rightarrow (ii) \Rightarrow (iv) \Rightarrow (vi) \Rightarrow (i)$.

To complete the proof, we need the implication $(iii) \Rightarrow (v)$ because it is non-trivial in $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (i)$.

$(iii) \Rightarrow (v)$: We only show the case of $k = 2$, which is essential. Let

$$X = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & A_2 & \\ 0 & & & \ddots \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ C_2 & 0 & \cdots & \\ 0 & 1_H & 0 & \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Then C is isometry in $\mathcal{B}(H \oplus H \oplus \cdots)$, i.e., $C^*C = 1_{H \oplus H \oplus \cdots}$. Hence it follows from (iii) that

$$\begin{aligned}
&\begin{pmatrix} f(C_1^*A_1C_1 + C_2^*A_2C_2) & & \\ & f(A_2) & \\ & & \ddots \end{pmatrix} \\
&= f(C^*XC) \le C^*f(X)C \\
&= \begin{pmatrix} C_1^*f(A_1)C_1 + C_2^*f(A_2)C_2 & & \\ & f(A_2) & \\ & & \ddots \end{pmatrix}.
\end{aligned}$$

Thus we have $f(C_1^*A_1C_1 + C_2^*A_2C_2) \le C_1^*f(A_1)C_1 + C_2^*f(A_2)C_2$ by seeing the (1,1)-components. □

By Theorem 1.9, we show the following Hansen-Pedersen type Jensen's inequality.

Theorem 1.10 (HANSEN-PEDERSEN-JENSEN'S INEQUALITY) *Let I be an interval containing 0 and let f be a real valued continuous function defined on I. Let A and A_j be selfadjoint operators on H with spectra contained in I ($j = 1, 2, \cdots, k$). Then the following conditions are mutually equivalent:*

(i) f is operator convex on I and $f(0) \le 0$.

(ii) $f(C^*AC) \le C^*f(A)C$ for every $A \in \mathcal{B}_h(H)$ and contraction $C \in \mathcal{B}(H)$, i.e., $C^*C \le 1_H$.

(iii) $f(\sum_{j=1}^k C_j^*A_jC_j) \le \sum_{j=1}^k C_j^*f(A_j)C_j$ for every $A \in \mathcal{B}_h(H)$ and $C_j \in \mathcal{B}(H)$ with $\sum_{j=1}^k C_j^*C_j \le 1_H$

(iv) $f(PAP) \le Pf(A)P$ for every $A \in \mathcal{B}_h(H)$ and projection P.

Proof. $(i) \Rightarrow (ii)$: Suppose that f is operator convex and $f(0) \le 0$. For every contraction C, put $D = \sqrt{1_H - C^*C}$. Since $C^*C + D^*D = 1_H$, it follows from (v) of Theorem 1.9 that

$$\begin{aligned}
f(C^*AC) &= f(C^*AC + D^*0D) \\
&\le C^*f(A)C + D^*f(0)D = C^*f(A)C \quad \text{by } f(0) \le 0
\end{aligned}$$

and hence we have (ii).

(ii) ⇒ *(iii)*: Put X and $\tilde{C}$ as in the proof *(ii)* ⇒ *(iv)* of Theorem 1.9, then $\tilde{C}^*\tilde{C} \le 1_H$ and hence we have (iii).

(iii) ⇒ *(iv)*: obviously.

(iv) ⇒ *(i)*: Under the same situation in the proof *(iv)* ⇒ *(i)* of Theorem 1.9, we have

$$\begin{aligned} \begin{pmatrix} f((1-t)A+tB) & 0 \\ 0 & f(0) \end{pmatrix} &= f(PU^*XUP) \\ &\le PU^*f(X)UP = \begin{pmatrix} (1-t)f(A)+tf(B) & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Hence f is operator convex and $f(0) \le 0$. □

Theorem 1.11 *Let $f \in \mathcal{C}([0,\infty)$. If $f(t) \le 0$ for all $t \in [0,\infty)$, then conditions (i)–(vi) in Theorems 1.9 are again equivalent to the following condition*

(vii) $-f$ is an operator monotone function.

Proof. Suppose that f is operator convex. Let $A, B \in \mathcal{B}(H)$, $0 \le A \le B$. Then for any $0 < \lambda < 1$ we can write

$$\lambda B = \lambda A + (1-\lambda)\frac{\lambda}{1-\lambda}(B-A).$$

Since f is operator convex, we have

$$f(\lambda B) \le \lambda f(A) + (1-\lambda)f\left(\frac{\lambda}{1-\lambda}(B-A)\right).$$

Since $-f(X)$ is positive for every positive operator X, it follows that $f(\lambda B) \le \lambda f(A)$. Letting λ tend to 1, we have $f(B) \le f(A)$. Hence $-f$ is operator monotone.

Conversely, suppose that $-f$ is operator monotone. Let $C \in \mathcal{B}(H)$ be an isometry. Consider the unitary operator U on $H \oplus H$ given by

$$U = \begin{pmatrix} C & -D \\ 0 & C^* \end{pmatrix}$$

where $D = \sqrt{1_H - CC^*}$. We put

$$X = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{B}_h(H \oplus H)$$

and note that

$$U^*XU = \begin{pmatrix} C^*AC & -C^*AD \\ -DAC & DAD \end{pmatrix}.$$

Choose now a constant $\varepsilon > 0$ and set

$$Y = \begin{pmatrix} C^*AC + \varepsilon 1_H & 0 \\ 0 & 2\lambda 1_H \end{pmatrix},$$

where λ is a positive constant to be fixed later. We observe that

$$\begin{aligned} Y - U^*XU &= \begin{pmatrix} \varepsilon 1_H & C^*AD \\ DAC & 2\lambda 1_H - DAD \end{pmatrix} \\ &\geq \begin{pmatrix} \varepsilon 1_H & F \\ F^* & \lambda 1_H \end{pmatrix} \qquad \text{for} \quad \lambda 1_H \geq DAD, \end{aligned}$$

where $F = C^*AD$. Furthermore let $\xi, \eta \in H$, then

$$\begin{aligned} &\left(\begin{pmatrix} \varepsilon 1_H & F \\ F^* & \lambda 1_H \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right) \\ &= \varepsilon\|\xi\|^2 + (F\xi, \eta) + (F^*\xi, \eta) + \lambda\|\eta\|^2 \\ &\geq \varepsilon\|\xi\|^2 - 2\|F\|\|\xi\|\|\eta\| + \lambda\|\eta\|^2 \\ &\geq 0 \qquad \text{for} \quad \lambda \geq \frac{\|F\|^2}{\varepsilon}. \end{aligned}$$

For a sufficiently large λ we thus obtain

$$U^*XU \leq Y$$

and consequently the operator monotonicity of $-f$ implies

$$U^*f(X)U = f(U^*XU) \geq f(Y)$$

or written as matrices

$$\begin{pmatrix} C^*f(A)C & -C^*f(A)D \\ -Df(A)C^* & Df(A)D \end{pmatrix} \geq \begin{pmatrix} f(C^*AC + \varepsilon 1_H) & 0 \\ 0 & f(2\lambda 1_H) \end{pmatrix}.$$

In particular we have $C^*f(A)C \geq f(C^*AC + \varepsilon 1_H)$. Letting ε tend to 0, we get the conclusion of the theorem. □

Corollary 1.12 *Let f be a real valued continuous function mapping the positive half line $[0, \infty)$ into itself. Then f is operator monotone if and only if f is operator concave.*

Theorem 1.13 *Let $f \in \mathcal{C}([0, r))$ and $r \leq \infty$. Then the following conditions are mutually equivalent.*

(i) f is operator convex and $f(0) \leq 0$.

(ii) The function $t \mapsto \frac{f(t)}{t}$ is operator monotone on $(0, r)$.

Proof. Suppose that f is operator convex. Let $A, B \in \mathcal{B}_h(H)$ be selfadjoint operators with $\mathsf{Sp}(A), \mathsf{Sp}(B) \subset (0, r)$ and $A \leq B$. Then A and B are invertible. If we put $C = B^{-1/2}A^{1/2}$, then $CC^* = B^{-1/2}AB^{-1/2} \leq 1_H$ and $\|C\| \leq 1$. Since $A = C^*BC$, it follows from (ii) in Theorem 1.10 that

$$f(A) = f(C^*BC) \leq C^*f(B)C = A^{1/2}B^{-1/2}f(B)B^{-1/2}A^{1/2}$$

and hence $A^{-1/2}f(A)A^{-1/2} \le B^{-1/2}f(B)B^{-1/2}$. Therefore we have $A^{-1}f(A) \le B^{-1}f(B)$ and $f(t)/t$ is operator monotone.

Conversely, suppose that $f(t)/t$ is operator monotone on $(0, r)$. Since $f(t)/t \le f(\beta)/\beta$ for $0 < t < \beta \le r$, we have $f(t) \le (f(\beta)/\beta)t$. Letting $t \mapsto 0$, we have $f(0) \le 0$. We will show that f satisfies the condition (iv) of Theorem 1.10. Let P be any projection and let A be any positive operator with spectrum in $(0, r)$ and $\mathrm{Sp}((1+\varepsilon)A) \subset (0, r)$ for a sufficiently small $\varepsilon > 0$. Put $P_\varepsilon = P + \varepsilon 1_H$ and $X_\varepsilon = P_\varepsilon^{\frac{1}{2}} A^{\frac{1}{2}}$. Since $P_\varepsilon \le (1+\varepsilon)1_H$, we have $A^{\frac{1}{2}} P_\varepsilon A^{\frac{1}{2}} \le (1+\varepsilon)A$. Since

$$\begin{aligned} \left(\frac{f}{t}\right)\left(A^{\frac{1}{2}}P_\varepsilon A^{\frac{1}{2}}\right) &= f\left(A^{\frac{1}{2}}P_\varepsilon A^{\frac{1}{2}}\right)\left(A^{\frac{1}{2}}P_\varepsilon A^{\frac{1}{2}}\right)^{-1} \\ &= f(X_\varepsilon^* X_\varepsilon)(X_\varepsilon^* X_\varepsilon)^{-1} \\ &= X_\varepsilon^{-1} f(X_\varepsilon X_\varepsilon^*) X_\varepsilon X_\varepsilon^{-1} X_\varepsilon^{*-1} \\ &= X_\varepsilon^{-1} f(X_\varepsilon X_\varepsilon^*) X_\varepsilon^{*-1}, \end{aligned}$$

it follows from the operator monotonicity of $f(t)/t$ that

$$\begin{aligned} X_\varepsilon^{-1} f(X_\varepsilon X_\varepsilon^*) X_\varepsilon^{*-1} &= \left(\frac{f}{t}\right)\left(A^{\frac{1}{2}}P_\varepsilon A^{\frac{1}{2}}\right) \\ &\le \left(\frac{f}{t}\right)((1+\varepsilon)A) \\ &= (1+\varepsilon)^{-1} A^{\frac{1}{2}} f((1+\varepsilon)A) A^{\frac{1}{2}}. \end{aligned}$$

Therefore we obtain

$$\begin{aligned} f(X_\varepsilon X_\varepsilon^*) &\le (1+\varepsilon)^{-1} X_\varepsilon A^{\frac{1}{2}} f((1+\varepsilon)A) A^{\frac{1}{2}} X_\varepsilon^* \\ &= (1+\varepsilon)^{-1} P_\varepsilon f((1+\varepsilon)A) P_\varepsilon. \end{aligned}$$

Let $\varepsilon \to 0$. This gives $X_\varepsilon X_\varepsilon^* \to A^{\frac{1}{2}} P A^{\frac{1}{2}}$ and hence we have $f(APA) \le Pf(A)P$ as desired. □

From the previous theorem we obtain the following corollary.

Corollary 1.14 *Let $f \in C([0, \infty))$ and $f > 0$. The function f is operator monotone if and only if the function $t/f(t)$ is operator monotone.*

Proof. Suppose that f is operator monotone. Since $-f$ is operator convex, it follows from Theorem 1.13 that $-f(t)/t$ is operator monotone on $(0, \infty)$. Hence $t/f(t) = -(-f(t)/t)^{-1}$ is operator monotone on $(0, \infty)$. By the continuity of f, we have the desired result.

Conversely, suppose that $t/f(t)$ is operator monotone. If we put $g(t) = -t/f(t)$, then $g(t) \ge 0$ and by Theorem 1.11 the operator monotonicity of $-g(t)$ implies the operator convexity of $g(t)$ and $g(0) \le 0$. It follows from Theorem 1.13 that $g(t)/t = -1/f(t)$ is operator monotone on $(0, \infty)$ and this fact is equivalent to the operator monotonicity of $f(t)$. □