Operator Monotonicity and Convexity

Vjosa Blakaj
Topic 5, Matrix Analysis in Quantum Theory
May 20, 2018

**Definition 1.** A function $f : I \rightarrow \mathbb{R}$, $I \subseteq \mathbb{R}$, is said to be matrix monotone of order $n$ if it is monotone with respect to this order on $n \times n$ Hermitian matrices, i.e. if $A \leq B$ implies $f(A) \leq f(B)$ for all $A$ and $B$ whose eigenvalues are in $I$.

If $f$ is matrix monotone of order $n$ for all $n$, we say $f$ is matrix monotone or operator monotone.

**Definition 2.** A function $f : I \rightarrow \mathbb{R}$, $I \subseteq \mathbb{R}$, is said to be matrix convex of order $n$ if for all $n \times n$ Hermitian matrices $A$ and $B$ with eigenvalues in $I$ and all real numbers $0 \leq \lambda \leq 1$:

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B)$$

(1)

If $f$ is matrix convex of all orders, we say that $f$ is matrix convex or operator convex.

When $-f$ is matrix convex, then $f$ is called matrix concave.

For continuous functions we can replace (1) by the special condition:

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

(2)

Functions that satisfy condition (2) are called **mid-point operator convex**, and being continuous they are convex.

Considering the case of $1 \times 1$ matrices, we see that if $f$ is monotone or convex in the operator sense, then it must be monotone or convex in the usual sense as a function from $I$ to $\mathbb{R}$. However, the opposite is not true.

**Example 0.1.** One of the simplest examples is that of the function $f(t) = \alpha + \beta t$ which is operator monotone (on every interval) for every $\alpha \in \mathbb{R}$ and $\beta \geq 0$.

**Example 0.2.** Let $t \geq 0$ be a parameter, the function $f(x) = -(t + x)^{-1}$ is matrix monotone on $[0, \infty)$.

Indeed, let $A$ and $B$ positive matrices of the same order. Then $A_t = tI + A$ and $B_t = tI + B$ are invertible, and

$$A_t \leq B_t \iff B_t^{-\frac{1}{2}}A_tB_t^{-\frac{1}{2}} \leq I$$

(3)

Now we use a lemma, which says that whenever $0 \leq C \leq D$, $C$ and $D$ commuting matrices $\implies C^{-1} \geq D^{-1}$.

Hence, using the mentioned lemma we have: $B_t^{-\frac{1}{2}}A_t^{-1}B_t^{-\frac{1}{2}} \geq I \implies A_t^{-1} \geq B_t^{-1}$. 

1
Example 0.3. The function $f(x) = \log(x)$ is matrix monotone on $(0, \infty)$.

Using $\log(x) = \int_{0}^{\infty} \left( \frac{1}{1+t} - \frac{1}{x+t} \right) dt$ and example 0.2 we get our result. Indeed, from the previous example we have that the integrand $f_t(x) := \frac{1}{1+t} - \frac{1}{x+t}$ is matrix monotone. Now, since the set of operator monotone functions is closed under positive linear combinations, we have that:

$$\sum_{i=1}^{n} c_i f_t(i)(x)$$

is matrix monotone for any $t_i$ and positive $c_i \in \mathbb{R}$. By the fact that the integral is the limit of such functions we have that our function is matrix monotone function as being the limit of such functions.

Example 0.4. The function $f(t) = t^2$ on $[0, \infty)$ is not operator monotone, but it is operator convex on every interval.

There exist positive matrices $A$ and $B$ such that $B - A$ is positive, but $B^2 - A^2$ is not. To see this, take the matrices

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

(5)

Let us show now that it is operator convex on every interval. Indeed, for any Hermitian matrices $A$ and $B$ we get:

$$\frac{f(A) + f(B)}{2} - f\left( \frac{A + B}{2} \right) = \frac{A^2 + B^2}{2} - \left( \frac{A + B}{2} \right)^2 =$$

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

(6)

(7)

So, $f(t) = t^2$ is mid-point operator convex and by continuity we get that it is operator convex.

Example 0.5. The function $f(t) = t^3$ on $[0, \infty)$ is not operator convex.

To see this, take

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

(8)

Example 0.6. The square root function is matrix monotone.

In order to show this we consider the function

$$F(t) := \sqrt{A + tX}$$

defined for $t \in [0,1]$ and for fixed positive matrices $A$ and $X$.

If $F$ increasing, $0 \leq 1 \implies F(0) = \sqrt{A} \leq \sqrt{A + X} = F(1)$. In order to show that $F$ is increasing, it is enough to show that the eigenvalues of $F'(t)$ are positive. Consider the equality $F(t)F(t) = A + tX$, and taking the derivative on both sides with respect to $t$ we get $F'(t)F(t) + F(t)F'(t) = X$. Being a limit of hermitian matrices $F'(t)$ is hermitian. Hence,
from the spectral theorem $F'(t) = \sum \lambda_i E_i$, where $\lambda_i$ the eigenvalues and $E_i$ the projections, both depending on the value of $t$. Hence, we have:

$$\sum \lambda_i E_i F(t) + F(t) \sum \lambda_i E_i = X \quad (10)$$

Multiplying on the right and on the left by $E_j$, we have for the trace:

$$2\lambda_j \text{Tr} E_j F(t) E_j = \text{Tr} E_j X E_j \quad (11)$$

and since both traces are positive, $\lambda_j$ must be positive as well.

From the previous example we see that in order to determine whether a function $f$ is matrix monotone or not, one has to investigate the derivative of $f(A + tX)$.

**Theorem 1.** A smooth function $f:(a, b) \to \mathbb{R}$ is matrix monotone for $n \times n$ matrices if and only if the divided difference matrix $D \in M_n$ defined as:

$$D_{ij} = \begin{cases} f(t_i) - f(t_j) & \text{if } t_i - t_j \neq 0 \\ f'(t_i) & \text{if } t_i - t_j = 0 \end{cases} \quad (12)$$

is positive semidefinite for all $t_1, t_2, ..., t_n \in (a, b)$.

-[Smoothness assumption is not essential.]-

From example 0.6, we saw that the function $f(x) = x^t$ for $t = \frac{1}{2}$ is matrix monotone. In general for every $0 \leq t \leq 1$, matrix monotonicity holds:

$$0 \leq A \leq B \implies A^t \leq B^t \quad (13)$$

What happens for $t > 1$? (Homework)

The above inequality is known as the **Loewner-Heinz-Inequality**.

The theory of operator monotone functions was initiated by Karl Loewner which was followed by Fritz Kraus’ theory of operator convex functions. The following result named by Loewner gives several examples of operator monotone and operator convex functions.

**Theorem 2. Loewner-Heinz-Theorem**

- For $-1 \leq p \leq 0$, the function $f(t) = -t^p$ is operator monotone and operator concave
- For $0 \leq p \leq 1$, the function $f(t) = t^p$ is operator monotone and operator concave
- For $1 \leq p \leq 2$, the function $f(t) = t^p$ is operator convex

Furthermore the function $f(t) = \log(t)$ is operator concave and operator monotone, while the function $f(t) = t \log(t)$ is operator convex.
Actually Loewner proved more; he gave a necessary and sufficient condition for
\( f : (0, \infty) \to \mathbb{R} \) to be operator monotone.

**Remark:**
\( f : (0, \infty) \to \mathbb{R} \) is operator monotone if and only if \( f \) admits an integral representation
\[
    f(a) = \alpha + \beta a - \int_0^\infty \frac{1 - at}{t + a} \, d\mu(t)
\]
for some \( \alpha, \beta \in \mathbb{R}, \beta > 0, \) and some finite positive measure \( \mu \).

Next we define the geometric mean, and state some of its properties that will be needed later to prove Lieb’s concavity theorem.

**Definition 3.** For \( A \) and \( B \in H^+_n \), the geometric mean of \( A \) and \( B \), \( M_0(A,B) \) is defined by
\[
    M_0(A,B) = A^{\frac{1}{2}} (A^{-\frac{1}{2}} B A^{-\frac{1}{2}})^{\frac{1}{2}} A^{\frac{1}{2}} \tag{15}
\]

The map \( (A, B) \mapsto M_0(A,B) \) on \( H^+_m \times H^+_n \) is jointly concave, i.e., for \( (A_0, B_0), (A_1, B_1) \in H^+_m \times H^+_n \) and \( \lambda \in [0,1] \) one has:
\[
    M_0(\lambda A_0 + (1 - \lambda) A_1, \lambda B_0 + (1 - \lambda) B_1) \geq \lambda M_0(A_0, B_0) + (1 - \lambda) M_0(A_1, B_1) \tag{16}
\]
Such map is also monotone in each variable.

Now we prove the fundamental theorem of Lieb.

**Theorem 3. Lieb’s Concavity Theorem** For all \( m \times n \) matrices \( K \), and all \( 0 \leq q, r \leq 1 \) with \( q + r \leq 1 \) the real valued map on \( H^+_m \times H^+_n \) given by
\[
    (A, B) \mapsto \text{Tr}(K^* A^q K B^r) \tag{17}
\]
is concave.

**Proof.** Since the map \( B \mapsto \overline{B} \), using Ando’s Identity: Let \( A \in H^+_m, B \in H^+_n \) and let \( K \) be any \( m \times n \) matrix considered as a vector in \( \mathbb{C}^m \otimes \mathbb{C}^n \). Then
\[
    \langle K, (A \otimes B)K \rangle = \text{Tr}(K^* A K \overline{B}) \tag{18}
\]
shows that an equivalent formulation of Lieb’s concavity theorem is that for \( 0 \leq q, r \leq 1 \)
\[
    (A, B) \mapsto A^q \otimes B^r \tag{19}
\]
is concave from \( H^+_m \times H^+_n \) to \( H^+_m \otimes H^+_n \). Let \( \Omega \) be the subset of \( (0, \infty) \times (0, \infty) \) consisting of points \( (q, r) \) such that \( (A, B) \mapsto A^q \otimes B^r \) is concave. By Loewner’s Theorem this is true for \( (0,0), (0,1), (1,0) \), hence it suffices to show that \( \Omega \) is convex. By continuity it suffices to show that if \( (q_1, r_1), (q_2, r_2) \in \Omega \), then so is \( (q, r) := \left( \frac{q_1 + q_2}{2}, \frac{r_1 + r_2}{2} \right) \). The main point is to use the joint concavity properties of the geometric mean \( M_0 \), and tensor product properties. Notice that for such \( (p, q), (p_1, q_1), (p_2, q_2) \), we can write
\[
    A^q \otimes B^r = M_0(A^{q_1} \otimes B^{r_1}, A^{q_2} \otimes B^{r_2}) \tag{20}
\]
Indeed,

\[ M_0(A^{q_1} \otimes B^{r_1}, A^{q_2} \otimes B^{r_2}) = \]
\[ (A^{q_1} \otimes B^{r_1})^{\frac{1}{2}}[(A^{q_1} \otimes B^{r_1})^{-\frac{1}{2}} A^{q_2} \otimes B^{r_2} (A^{q_1} \otimes B^{r_1})^{\frac{1}{2}}](A^{q_1} \otimes B^{r_1})^{\frac{1}{2}} = \]
\[ (A^{q_1} \otimes B^{r_1})^{\frac{1}{2}}[(A^{q_2} \otimes B^{r_2})A^{q_1} \otimes B^{r_2}]^{\frac{1}{2}}(A^{q_1} \otimes B^{r_1})^{\frac{1}{2}} = \]
\[ (A^{q_2} \otimes B^{r_2})(A^{q_1} \otimes B^{r_2}) = A^{q_1+q_2} \otimes B^{r_1+r_2} = \]
\[ A^q \otimes B^r \]  

Now, since \((q_1, r_1), (q_2, r_2) \in \Omega\) and using the fact that the geometric mean is jointly concave and monotone in each variable we have:

\[(A + C)^{q_j} \otimes (B + D)^{r_j} \geq \frac{A^{q_j} \otimes B^{r_j} + C^{q_j} \otimes D^{r_j}}{2} \]  

for \(j = 1, 2\).

Then by monotonicity and convexity of the operator geometric mean we have:

\[ (A^q \otimes B^r)^{q_j} \otimes (B^q \otimes D^r)^{r_j} = M_0((A^q \otimes B^r)^{q_j} \otimes (B^q \otimes D^r)^{r_j}, (A^q \otimes B^r)^{q_j} \otimes (B^q \otimes D^r)^{r_j}) \]
\[ \geq M_0(A^{q_1+q_2} \otimes B^{r_1+r_2}, A^{q_1+q_2} \otimes B^{r_1+r_2}) \]
\[ \geq \frac{1}{2} M_0(A^{q_1} \otimes B^{r_1}, A^{q_2} \otimes B^{r_2}) + \frac{1}{2} M_0(C^{q_1} \otimes D^{r_1}, C^{q_2} \otimes D^{r_2}) \]
\[ = \frac{1}{2} A^q \otimes B^r + \frac{1}{2} C^q \otimes D^r \]  

This proves the midpoint concavity of \((A, B) \mapsto A^q \otimes B^r\), and now full concavity follows by continuity. Hence, \((q, r) \in \Omega\). \(\square\)

We finally show that relative quantum entropy, \(S(A||B) := Tr[A \log B] - Tr[A \log B]\), is jointly convex.

**Theorem 4.** The map \((A, B) \mapsto Tr[A \log A] - Tr[A \log B]\) from \(H^n_+ \times H^n_+\) to \(\mathbb{R}\) is jointly convex.

**Proof.** For all \(0 \leq p \leq 1\), \((A, B) \mapsto Tr[B^{1-p}A^p]\) is jointly concave (by Lieb’s concavity theorem). Now, we define the function

\[(A, B) \mapsto \frac{1}{p-1}(Tr[B^{1-p}A^p] - Tr[A]) \]

which is convex (being the negative of a concave function). Knowing that the limit of convex functions is again convex, and the fact that

\[ \lim_{p \to 1} \frac{1}{p-1}[Tr(B^{1-p}A^p) - Tr(A)] = \]

\[ \frac{1}{2} A^q \otimes B^r + \frac{1}{2} C^q \otimes D^r \]
is the derivative of the function $Tr(B^{1-p}A^p)$ at $p = 1$, we have:

$$(A, B) \mapsto Tr[-B^{1-p} \log B A^p + B^{1-p} A^p \log A]$$  \hspace{1cm} (34)

is convex, and setting $p = 1$ we get

$$(A, B) \mapsto Tr[- \log B A + A \log A]$$  \hspace{1cm} (35)

Changing the order of $\log B$ and $A$ under trace and using the linearity of trace, we have that

$$(A, B) \mapsto Tr(-A \log B + A \log A) = Tr[A \log A] - Tr[A \log B]$$  \hspace{1cm} (36)

is jointly convex.

References