A concise survey of log-majorizations for matrices with applications to quantum information

Fumio Hiai

Tohoku University

2016, Sep. (at Ritsumeikan Univ.)

Plan

- Log-majorization basics
- Araki's log-majorization
- Ando-Hiai's log-majorization
- The Furuta inequality
- Applications to quantum divergences
- (Log-)supermajorization and anti-norms

Log-majorization basics[1 2 3]

(Weak) majorization for vectors

Let $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$, and let $(a_{[1]}, \dots, a_{[n]})$ be the decreasing rearrangement of $\mathbf{a}$.

- The weak majorization or submajorization $\mathbf{a} \prec_w \mathbf{b}$ means that

$$\sum_{i=1}^{k} a_{[i]} \le \sum_{i=1}^{k} b_{[i]}, \qquad 1 \le k \le n.$$

- The majorization $\mathbf{a} \prec \mathbf{b}$ means that $\mathbf{a} \prec_w \mathbf{b}$ and equality holds for $k = n$ in the above.

[1] A.W. Marshall, I. Olkin and B.C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer, New York, second edition, 2011.

[2] F.H., Log-majorizations and norm inequalities for exponential operators, in *Linear Operators*, J. Janas, F. H. Szafraniec and J. Zemánek (eds.), Banach Center Publications, Vol. 38, 1997, pp. 119–181.

[3] F.H., Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization, *Interdisciplinary Information Sciences* **16** (2010), 139–248.

Proposition

- $\mathbf{a} \prec \mathbf{b}$ iff $\sum_{i=1}^n f(a_i) \le \sum_{i=1}^n f(b_i)$ for any convex function f on an interval containing all a_i, b_i.
- $\mathbf{a} \prec_w \mathbf{b}$ iff $\sum_{i=1}^n f(a_i) \le \sum_{i=1}^n f(b_i)$ for any non-decreasing convex function f on an interval containing all a_i, b_i.

Proposition

- If $\mathbf{a} \prec \mathbf{b}$ and f is a convex function on an interval containing all a_i, b_i, then $f(\mathbf{a}) \prec_w f(\mathbf{b})$, where $f(\mathbf{a}) := (f(a_1), \dots, f(a_n))$.
- If $\mathbf{a} \prec_w \mathbf{b}$ and f is a non-decreasing convex function on an interval containing all a_i, b_i, then $f(\mathbf{a}) \prec_w f(\mathbf{b})$.

(Weak) log-majorization for non-negative vectors

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and $\mathbf{a}, \mathbf{b} \ge \mathbf{0}$.

- The weak log-majorization or log-submajorization $\mathbf{a} \prec_{w \log} \mathbf{b}$ means that

$$\prod_{i=1}^{k} a_{[i]} \le \prod_{i=1}^{k} b_{[i]}, \qquad 1 \le k \le n.$$

- The log-majorization $\mathbf{a} \prec_{\log} \mathbf{b}$ means that $\mathbf{a} \prec_{w \log} \mathbf{b}$ and equality holds for $k = n$ in the above.

Note

When $\mathbf{a}, \mathbf{b} > \mathbf{0}$,

$$\mathbf{a} \prec_{\log} \mathbf{b} \iff \log \mathbf{a} \prec \log \mathbf{b},$$

$$\mathbf{a} \prec_{w \log} \mathbf{b} \iff \log \mathbf{a} \prec_{w} \log \mathbf{b},$$

where $\log \mathbf{a} := (\log a_1, \ldots, \log a_n)$.

Proposition

Let $\mathbf{a}, \mathbf{b} \geq \mathbf{0}$ in $\mathbb{R}^n$, and assume that $\mathbf{a} \prec_{w \log} \mathbf{b}$. If f is a continuous non-decreasing function on $[0, \infty)$ such that $f(e^x)$ is convex, then $f(\mathbf{a}) \prec_w f(\mathbf{b})$.

Therefore,

$$\mathbf{a} \prec_{w \log} \mathbf{b} \implies \mathbf{a} \prec_w \mathbf{b} \implies \sum_{i=1}^n f(a_i) \leq \sum_{i=1}^n f(b_i)$$

for f as above.

Let $\mathbb{M}_n^{sa}$ denote the set of all Hermitian $n \times n$ matrices, and $\mathbb{M}_n^+$ the set of all positive semidefinite $n \times n$ matrices.

- For $A \in \mathbb{M}_n^{sa}$ write

$$\lambda(A) = (\lambda_1(A), \dots, \lambda_n(A))$$

  for the eigenvalues of A in decreasing order with counting multiplicities.

- For $X \in \mathbb{M}_n$ write

$$s(X) = (s_1(X), \dots, s_n(X))$$

  for the singular values of X (i.e., the eigenvalues of $|X| := (X^*X)^{1/2}$) in decreasing order with multiplicities.

(Weak) (log-)majorization for matrices

- For $A, B \in \mathbb{M}_n^{sa}$ we write $A \prec B$ (resp., $A \prec_w B$) if $\lambda(A) \prec \lambda(B)$ (resp., $\lambda(A) \prec_w \lambda(B)$).
- For $A, B \in \mathbb{M}_n^+$ we write $A \prec_{\log} B$ (resp., $A \prec_{w \log} B$) if $\lambda(A) \prec_{\log} \lambda(B)$ (resp., $\lambda(A) \prec_{w \log} \lambda(B)$).

Unitarily invariant norms

A norm $\|\cdot\|$ on $\mathbb{M}_n$ is said to be unitarily invariant (or symmetric) if

$$\|UXV\| = \|X\|$$

for all $X \in \mathbb{M}_n$ and all unitaries $U, V \in \mathbb{M}_n$. E.g.,

- for $1 \le p \le \infty$, the Schatten p-norm is
$$\|X\|_p := (\mathrm{Tr}\,|X|^p)^{1/p} = \left(\sum_{i=1}^n s_i^p(X)\right)^{1/p},$$
- for $k = 1, \dots, n$ the Ky Fan k-norm is
$$\|X\|_{(k)} := \sum_{i=1}^k s_i(X).$$

Proposition

Concerning the following conditions for $X, Y \in \mathbb{M}_n$, we have

$$\text{(i)} \iff \text{(ii)} \iff \text{(iii)} \iff \text{(iv)} \implies \text{(v)} \iff \text{(vi)} \iff \text{(vii)}.$$

(i) $|X| \prec_{w\log} |Y|$;

(ii) $\| |X|^p \|_{(k)} \le \| |Y|^p \|_{(k)}$ for every $k = 1, \dots, n$ and every $p > 0$;

(iii) $\|f(|X|)\| \le \|f(|Y|)\|$ for every symmetric norm $\|\cdot\|$ and every continuous non-decreasing function $f : [0, \infty) \to [0, \infty)$ such that $f(e^x)$ is convex on $\mathbb{R}$;

(iv) $\det f(|X|) \le \det f(|Y|)$ for every continuous non-decreasing function $f : [0, \infty) \to [0, \infty)$ such that $\log f(e^x)$ is convex on $\mathbb{R}$;

(v) $|X| \prec_w |Y|$, i.e., $\|X\|_{(k)} \le \|Y\|_{(k)}$ for every $k = 1, \dots, n$;

(vi) $\|X\| \le \|Y\|$ for every symmetric norm $\|\cdot\|$;

(vii) $\|f(|X|)\| \le \|f(|Y|)\|$ for every symmetric norm $\|\cdot\|$ and every non-decreasing convex function $f : [0, \infty) \to [0, \infty)$.

Useful variational formulas

- For $A \in \mathbb{M}_n^{sa}$ and $1 \le k \le n$,

$$\sum_{i=1}^{k} \lambda_i(A) = \max\{\mathrm{Tr}\, AP : P \text{ a projection, } \dim P = k\}.$$

- For $X \in \mathbb{M}_n$ and $1 \le k \le n$,

$$\begin{aligned} \|X\|_{(k)} &= \max\{\|XP\|_1 : P \text{ a projection, } \dim P = k\} \\ &= \min\{\|Y\|_1 + k\|Z\|_\infty : X = Y + Z\}. \end{aligned}$$

- For $A \in \mathbb{M}_n^+$ and $1 \le k \le n$,

$$\prod_{i=1}^{k} \lambda_i(A) = \max\{\det VAV^* : VV^* = I_k\}.$$

- For $X \in \mathbb{M}_n$ and $1 \le k \le n$,

$$\prod_{i=1}^{k} s_i(X) = \max\{|\det WXV^*| : VV^* = WW^* = I_k\}.$$

Anti-symmetric tensor powers

Let $\mathcal{H}$ be an n-dimensional Hilbert space (e.g., $\mathcal{H} = \mathbb{C}^n$), and $1 \le k \le n$.

- $\mathcal{H}^{\otimes k}$ is the k-fold tensor product of $\mathcal{H}$.
- For $x_1, \dots, x_k \in \mathcal{H}$ define

$$x_1 \wedge \cdots \wedge x_k := \frac{1}{\sqrt{k!}} \sum_{\pi \in S_k} (\operatorname{sgn} \pi) x_{\pi(1)} \otimes \cdots \otimes x_{\pi(k)} \quad \text{in } \mathcal{H}^{\otimes k}.$$

- The k-fold antisymmetric tensor product $\mathcal{H}^{\wedge k}$ ($\dim \mathcal{H}^{\wedge k} = \binom{n}{k}$) is defined as the subspace of $\mathcal{H}^{\otimes k}$ spanned by $\{x_1 \wedge \cdots \wedge x_k : x_i \in \mathcal{H}\}$.
- For every $X \in \mathcal{B}(\mathcal{H})$ the k-fold antisymmetric power $X^{\wedge k}$ is defined by

$$X^{\wedge k} := X^{\otimes k}|_{\mathcal{H}^{\wedge k}}.$$

Let $A \in \mathcal{B}(\mathcal{H})^+$ and $X, Y \in \mathcal{B}(\mathcal{H})$.

Lemma

- $(X^*)^{\wedge k} = (X^{\wedge k})^*$.
- $(XY)^{\wedge k} = (X^{\wedge k})(Y^{\wedge k})$.
- $|X|^{\wedge k} = |X^{\wedge k}|$.
- $A^{\wedge k} \geq 0$ and $(A^p)^{\wedge k} = (A^p)^{\wedge k}$ for all $p \geq 0$ (for all $p \in \mathbb{R}$ if A is invertible).

Lemma

$$\prod_{i=1}^{k} \lambda_i(A) = \lambda_1(A^{\wedge k}) \; (= \|A^{\wedge k}\|_\infty), \qquad (\spadesuit)$$

$$\prod_{i=1}^{k} s_i(X) = s_1(X^{\wedge k}) \; (= \|X^{\wedge k}\|_\infty).$$

Araki's log-majorization

- Golden-Thompson inequality (1965) For $H, K \in \mathbb{M}_n^{sa}$,

$$\mathrm{Tr}\, e^{H+K} \le \mathrm{Tr}\, e^H e^K.$$

- Lieb-Thirring inequality (1976) For $A, B \in \mathbb{M}_n^+$,

$$\mathrm{Tr}\,(A^{1/2}BA^{1/2})^m \le \mathrm{Tr}\, A^{m/2}B^m A^{m/2}, \qquad m = 1, 2, \dots.$$

- Araki's log-majorization (1990)[4]

$$(A^{1/2}BA^{1/2})^r \prec_{\log} A^{r/2}B^r A^{r/2}, \qquad r \ge 1,$$

$$(A^{q/2}B^q A^{q/2})^{1/q} \prec_{\log} (A^{p/2}B^p A^{p/2})^{1/p}, \qquad 0 < q < p.$$

- By the Lie-Trotter formula, for every $p > 0$,

$$e^{\log A \dotplus \log B} \prec_{\log} (A^{p/2}B^p A^{p/2})^{1/p}, \quad e^{H+K} \prec_{\log} (e^{pH/2}e^{pK}e^{pH/2})^{1/p}.$$

[4] H. Araki, On an inequality of Lieb and Thirring, *Lett. Math. Phys.* **19** (1990), 167–170.

Proof of Araki's log-majorization We may assume that A, B are invertible. First, show that

$$\|(A^{1/2}BA^{1/2})^r\|_\infty \le \|A^{r/2}B^rA^{r/2}\|_\infty, \qquad r \ge 1. \qquad (*)$$

For this, it suffices to show that

$$A^{r/2}B^rA^{r/2} \le I \implies A^{1/2}BA^{1/2} \le I,$$

equivalently, $B^r \le A^{-r} \implies B \le A^{-1}$. But this is just the Löwner-Heinz inequality. Next, apply $(*)$ to $A^{\wedge k}, B^{\wedge k}$. Since

$$((A^{1/2}BA^{1/2})^r)^{\wedge k} = ((A^{\wedge k})^{1/2}(B^{\wedge k})(A^{\wedge k})^{1/2})^r,$$

$$(A^{r/2}B^rA^{r/2})^{\wedge k} = (A^{\wedge k})^{r/2}(B^{\wedge k})^r(A^{\wedge k})^{r/2},$$

we have $\|((A^{1/2}BA^{1/2})^r)^{\wedge k}\|_\infty \le \|(A^{r/2}B^rA^{r/2})^{\wedge k}\|_\infty$, which implies by $(\spadesuit)$ that

$$\prod_{i=1}^k \lambda_i((A^{1/2}BA^{1/2})^r) \le \prod_{i=1}^k \lambda_i(A^{r/2}B^rA^{r/2}).$$

Operator means

- Associated with an operator monotone function $f \geq 0$ on $[0, \infty)$ with $f(1) = 1$, the operator mean σ_f (in the sense of Kubo-Ando, 1980) is defined by

$$A \, \sigma_f \, B := A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}$$

for $A, B \in \mathbb{M}_n^+$ with $A > 0$, and is extended to general $A, B \in \mathbb{M}_n^+$ as

$$A \, \sigma_f \, B := \lim_{\varepsilon \searrow 0} (A + \varepsilon I) \, \sigma_f \, (B + \varepsilon I).$$

- In particular, for $0 \leq \alpha \leq 1$, associated with $f(x) = x^\alpha$,

$$A \,\#_\alpha\, B := A^{1/2} (A^{-1/2} B A^{-1/2})^\alpha A^{1/2}$$

is the weighted geometric mean. The geometric mean $\# = \#_{1/2}$ was first introduced by Pusz-Woronowicz, 1975.

Ando-Hiai's log majorization

- Complementary Golden-Thompson inequality (1993)[5]
  For $A, B \in \mathbb{M}_n^+$ and $0 \le \alpha \le 1$,

$$\mathrm{Tr}\,(A^p \#_\alpha B^p)^{1/p} \le \mathrm{Tr}\ \exp\{(1-\alpha)\log A \dotplus \alpha \log B\}, \qquad p > 0.$$

- Ando-H's log-majorization (1994)[6]

$$A^r \#_\alpha B^r \prec_{\log} (A \#_\alpha B)^r, \qquad r \ge 1,$$

$$(A^p \#_\alpha B^p)^{1/p} \prec_{\log} (A^q \#_\alpha B^q)^{1/q}, \qquad 0 < q < p.$$

- By the Lie-Trotter formula, for every $p > 0$,

$$(A^p \#_\alpha B^p)^{1/p} \prec_{\log} e^{(1-\alpha)\log A \dotplus \alpha \log B}, \quad (e^{pH} \#_\alpha e^{pK})^{1/p} \prec_{\log} e^{(1-\alpha)H + \alpha K}.$$

[5] F.H. and D. Petz, The Golden-Thompson trace inequality is complemented, *Linear Algebra Appl.* **181** (1993), 153–185.

[6] T. Ando and F.H., Log majorization and complementary Golden-Thompson type inequalities, *Linear Algebra Appl.* **197** (1994), 113–131.

Proof of Ando-H's log-majorization

By continuity we may assume that A, B are invertible. Since

$$(A^r \#_\alpha B^r)^{\wedge k} = (A^{\wedge k})^r \#_\alpha (B^{\wedge k})^r, \quad ((A \#_\alpha B)^r)^{\wedge k} = ((A^{\wedge k}) \#_\alpha (B^{\wedge k}))^r,$$

it suffices to show that

$$\|A^r \#_\alpha B^r\|_\infty \le \|(A \#_\alpha B)^r\|_\infty, \qquad r \ge 1,$$

equivalently,

$$A \#_\alpha B \le I \implies A^r \#_\alpha B^r \le I.$$

When $1 \le r \le 2$, write $r = 2 - \varepsilon$ with $0 \le \varepsilon \le 1$, and let $C := A^{-1/2} B A^{-1/2}$ so that $A \#_\alpha B \le I$ implies $C^\alpha \le A^{-1}$ or $A \le C^{-\alpha}$, so $A^{1-\varepsilon} \le C^{-\alpha(1-\varepsilon)}$. We have

$$
\begin{aligned}
A^r \#_\alpha B^r &= A^{1-\frac{\varepsilon}{2}}\{A^{-1+\frac{\varepsilon}{2}}B \cdot B^{-\varepsilon} \cdot BA^{-1+\frac{\varepsilon}{2}}\}^\alpha A^{1-\frac{\varepsilon}{2}} \\
&= A^{1-\frac{\varepsilon}{2}}\{A^{-\frac{1-\varepsilon}{2}}CA^{1/2}(A^{-1/2}C^{-1}A^{-1/2})^\varepsilon A^{1/2}CA^{-\frac{1-\varepsilon}{2}}\}^\alpha A^{1-\frac{\varepsilon}{2}} \\
&= A^{1/2}\{A^{1-\varepsilon} \#_\alpha [C(A \#_\varepsilon C^{-1})C]\}A^{1/2} \\
&\le A^{1/2}\{C^{-\alpha(1-\varepsilon)} \#_\alpha [C(C^{-\alpha} \#_\varepsilon C^{-1})C]\}A^{1/2} \\
&= A^{1/2}C^\alpha A^{1/2} = A \#_\alpha B \le I.
\end{aligned}
$$

When $r > 2$, write $r = 2^m s$ with $1 \le s \le 2$. Repeating use of the above case gives

$$
\begin{aligned}
A^r \#_\alpha B^r &\prec_{w\log} (A^{2^{m-1}s} \#_\alpha B^{2^{m-1}s})^2 \prec_{w\log} \cdots \\
&\prec_{w\log} (A^s \#_\alpha B^s)^{2^m} \prec_{w\log} (A \#_\alpha B)^r.
\end{aligned}
$$

□

The Furuta inequality

- The Furuta inequality (1987) Let $A, B \in \mathbb{M}_n^+$. For $r, p \geq 0$ and $q \geq 1$ with $(1 + r)q \geq p + r$,

$$A \geq B \geq 0 \implies (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \leq A^{\frac{p+r}{q}}.$$

- The critical case is when $q = \frac{p+r}{1+r}$, i.e.,

$$A \geq B \geq 0 \implies (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1+r}{p+r}} \leq A^{1+r} \quad \text{or} \quad A^{-r} \#_{\frac{1+r}{p+r}} B^p \leq A$$

  for $p \geq 1$ and $r \geq 0$.

- Fujii-Kamei (2006)[7] showed that Ando-H's inequality implies the Furuta inequality and vice versa.

[7] M. Fujii and E. Kamei, Ando-Hiai inequality and Furuta inequality, *Linear Algebra Appl.* **416** (2006), 541–545.

Proof of Ando-H $\Longrightarrow$ Furuta

Assume that $p \geq 1$, $r \geq 0$, and $A \geq B > 0$. When $0 \leq r \leq 1$, since $A^{-r} \leq B^{-r}$,

$$B^p \#_{\frac{p}{p+r}} A^{-r} \leq B^p \#_{\frac{p}{p+r}} B^{-r} = I.$$

When $r \geq 1$,

$$B^{\frac{p}{r}} \#_{\frac{p}{p+r}} A^{-1} \leq B^{\frac{p}{r}} \#_{\frac{p}{p+r}} B^{-1} = I,$$

so Ando-H implies that $B^p \#_{\frac{p}{p+r}} A^{-r} \leq I$. We then have

$$A^{-r} \#_{\frac{1+r}{p+r}} B^p = B^p \#_{\frac{p-1}{p+r}} A^{-r} = B^p \#_{\frac{p-1}{p}} (B^p \#_{\frac{p}{p+r}} A^{-r})$$
$$\leq B^p \#_{\frac{p-1}{p}} I = B \leq A,$$

since $C \#_\alpha D = D \#_{1-\alpha} C$ and $C \#_{\alpha\beta} D = C \#_\alpha (C \#_\beta D)$. □

The Furuta inequality with negative powers: Tanahashi (1999)[8]

Let $A, B \in \mathbb{M}_n^+$ with $A > 0$. Assume that $0 < p \le 1$, $-1 \le r < 0$, and either

$$\frac{1}{2} \le q \le 1, \quad -r(1-q) \le p \le q - r(1-q),$$

or

$$0 < q < \frac{1}{2}, \quad -r(1-q) \le p \le q - r(1-q),$$
$$\frac{-r(1-q)-q}{1-2q} \le p \le \frac{-r(1-q)}{1-2q}.$$

Then

$$A \ge B \ge 0 \implies (A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q}} \le A^{\frac{p+r}{q}}, \quad \text{hence,}$$
$$(A^{\frac{r}{2}} B^p A^{\frac{r}{2}})^{\frac{1}{q'}} \le A^{\frac{p+r}{q'}} \text{ for every } q' \ge q.$$

[8] K. Tanahashi, The Furuta inequality with negative powers, *Proc. Amer. Math. Soc.* **127** (1999), 1683–1692.

Various Rényi divergences

For $A, B \in \mathbb{M}_n^+$ with $B > 0$ and for $\alpha, z > 0$, define

- $P_\alpha(A, B) := B^{1/2}(B^{-1/2}AB^{-1/2})^\alpha B^{1/2}$.
- $Q_{\alpha,z}(A, B) := (B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}})^z$.

Note P_α is the operator perspective for x^α, whose general theory has recently been developed by Effros, Hansen, and others.
$P_\alpha(A, B) = B \#_\alpha A$ when $0 \le \alpha \le 1$.

For $\alpha, z > 0$ with $\alpha \ne 1$,

- The (conventional) Rényi divergence is

$$D_\alpha(A\|B) := \frac{1}{\alpha - 1} \log \mathrm{Tr}\, A^\alpha B^{1-\alpha} = \frac{1}{\alpha - 1} \log \mathrm{Tr}\, Q_{\alpha,1}(A, B).$$

- The sandwiched Rényi divergence[9] is

$$D_\alpha^*(A\|B) := \frac{1}{\alpha-1}\log\operatorname{Tr}\left(B^{\frac{1-\alpha}{2\alpha}}AB^{\frac{1-\alpha}{2\alpha}}\right)^\alpha = \frac{1}{\alpha-1}\log\operatorname{Tr}Q_{\alpha,\alpha}(A,B).$$

- The α-z-Rényi divergence[10] [11] is

$$D_{\alpha,z}(A\|B) := \frac{1}{\alpha-1}\log\operatorname{Tr}\left(B^{\frac{1-\alpha}{2z}}A^{\frac{\alpha}{z}}B^{\frac{1-\alpha}{2z}}\right)^z = \frac{1}{\alpha-1}\log\operatorname{Tr}Q_{\alpha,z}(A,B).$$

---

[9] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr and M. Tomamichel, On quantum Rényi entropies: A new generalization and some properties, *J. Math. Phys.* **54** (2013), 122203.

[10] V. Jaksic, Y. Ogata, Y. Pautrat and C.-A. Pillet, Entropic fluctuations in quantum statistical mechanics. An Introduction, in: Quantum Theory from Small to Large Scales, August 2010, in: Lecture Notes of the Les Houches Summer School, vol. 95, Oxford University Press, 2012.

[11] K.M.R. Audenaert and N. Datta, α-z-Rényi relative entropies, *J. Math. Phys.* **56** (2015), 022202.

- The "maximal" α-Rényi divergence[12] is

$$\widehat{D}_\alpha(A\|B) := \frac{1}{\alpha-1}\log \mathrm{Tr}\, B^{1/2}(B^{-1/2}AB^{-1/2})^\alpha B^{1/2}$$
$$= \frac{1}{\alpha-1}\log \mathrm{Tr}\, P_\alpha(A,B).$$

Note

- $D_\alpha = D_{\alpha,1}$, $D_\alpha^* = D_{\alpha,\alpha}$.
- $D_\alpha = D_\alpha^* = D_{\alpha,z} = \widehat{D}_\alpha$ if $AB = BA$.
- When $\mathrm{Tr}\, A = 1$, the Umegaki relative entropy is

$$\lim_{\alpha\to1} D_\alpha(A\|B) = \lim_{\alpha\to1} D_\alpha^*(A\|B) = D(A\|B) := \mathrm{Tr}\, A(\log A - \log B),$$

and the Belavkin-Staszewski relative entropy is

$$\lim_{\alpha\to1} \widehat{D}_\alpha(A\|B) = D_{\mathrm{BS}}(A\|B) := \mathrm{Tr}\, A\log(A^{1/2}B^{-1}A^{1/2}).$$

[12] A special case of K. Matsumoto, A new quantum version of f-divergence, arXiv:1311.4722.

Applications to Rényi divergences

- When $z > z' > 0$, by Araki's log-majorization,

$$Q_{\alpha,z}(A,B) \prec_{\log} Q_{\alpha,z'}(A,B),$$

and hence

$$\begin{aligned} D_{\alpha,z}(A\|B) &\le D_{\alpha,z'}(A\|B) \quad \text{for} \;\; \alpha > 1, \\ D_{\alpha,z}(A\|B) &\ge D_{\alpha,z'}(A\|B) \quad \text{for} \;\; 0 < \alpha < 1. \end{aligned}$$

In particular, $D^*_\alpha(A\|B) \le D_\alpha(A\|B)$ for all $\alpha > 0$ with $\alpha \ne 1$.

- For $0 < \alpha \le 1$ and $z > 0$, by Araki's and Ando-H's log-majorizations together,

$$P_\alpha(A,B) = B \,\#_\alpha\, A \prec_{\log} Q_{\alpha,z}(A,B),$$

and hence $D_{\alpha,z}(A\|B) \le \widehat{D}_\alpha(A\|B)$ for $0 < \alpha < 1$ and $z > 0$.

When $\alpha \geq 1$, we have

Proposition

- If $\alpha \geq 1$ and $0 < z \leq \min\{\alpha/2, \alpha - 1\}$, then

$$P_\alpha(A, B) \prec_{\log} Q_{\alpha,z}(A, B).$$

- If $\alpha \geq 1$ and $z \geq \max\{\alpha/2, \alpha - 1\}$, then

$$Q_{\alpha,z}(A, B) \prec_{\log} P_\alpha(A, B).$$

Proof The case $\alpha = 1$ is trivial. Assume that $\alpha > 1$ and $0 < z \leq \min\{\alpha/2, \alpha - 1\}$. For the first log-majorization, it suffices to show that

$$B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}} \leq I \implies B^{1/2}(B^{-1/2}AB^{-1/2})^\alpha B^{1/2} \leq I,$$

that is,

$$A^{\frac{\alpha}{z}} \leq B^{\frac{\alpha-1}{z}} \implies (B^{-1/2}AB^{-1/2})^\alpha \leq B^{-1}.$$

Setting $\widetilde{A} := A^{\frac{\alpha}{z}}$ and $\widetilde{B} := B^{\frac{\alpha-1}{z}}$, we may prove that

$$0 \le \widetilde{A} \le \widetilde{B} \implies \left(\widetilde{B}^{\frac{z}{2(1-\alpha)}} \widetilde{A}^{\frac{z}{\alpha}} \widetilde{B}^{\frac{z}{2(1-\alpha)}}\right)^{\alpha} \le \widetilde{B}^{\frac{z}{1-\alpha}}.$$

Let

$$p := \frac{z}{\alpha}, \quad q := \frac{1}{\alpha}, \quad r := \frac{z}{1-\alpha}.$$

Then $0 < p, q \le 1$, $-1 \le r < 0$ and $\frac{p+r}{q} = \frac{z}{1-\alpha}$. Note that

$$-r(1-q) = \frac{z}{\alpha-1}\left(1 - \frac{1}{\alpha}\right) = \frac{z}{\alpha} = p \le q - r(1-q).$$

When $q < \frac{1}{2}$ and so $\alpha > 2$, we further note that

$$\frac{-r(1-q) - q}{1-2q} = \frac{z-1}{\alpha-2} \le \frac{z}{\alpha} = p \le \frac{z}{\alpha-2} = \frac{-r(1-q)}{1-2q}.$$

Hence, the first result follows from the Furuta inequality with negative powers.

Next, assume that $\alpha > 1$ and $z \geq \max\{\alpha/2, \alpha - 1\}$. For the second log-majorization, we need to show that

$$B^{1/2}(B^{-1/2}AB^{-1/2})^{\alpha}B^{1/2} \leq I \implies B^{\frac{1-\alpha}{2z}}A^{\frac{\alpha}{z}}B^{\frac{1-\alpha}{2z}} \leq I,$$

that is,

$$(B^{-1/2}AB^{-1/2})^{\alpha} \leq B^{-1} \implies A^{\frac{\alpha}{z}} \leq B^{\frac{\alpha-1}{z}}.$$

Setting $\widetilde{A} := (B^{-1/2}AB^{-1/2})^{\alpha}$ and $\widetilde{B} := B^{-1}$, we may prove that

$$0 \leq \widetilde{A} \leq \widetilde{B} \implies \left(\widetilde{B}^{-\frac{1}{2}}\widetilde{A}^{\frac{1}{\alpha}}\widetilde{B}^{-\frac{1}{2}}\right)^{\frac{\alpha}{z}} \leq \widetilde{B}^{\frac{1-\alpha}{z}}.$$

Let

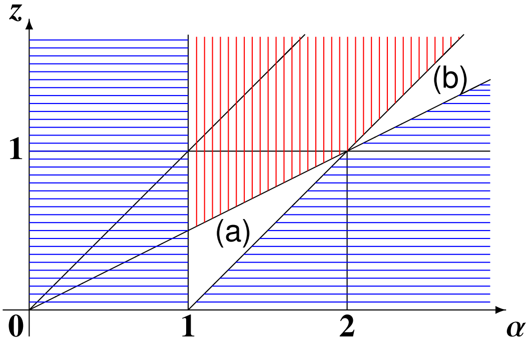
$$p := \frac{1}{\alpha}, \quad q := \frac{z}{\alpha}, \quad r := -1.$$

Since the Furuta inequality with negative powers holds for these p, q, r, the second result follows. □

In the following picture, the region of $P_\alpha \prec_{\log} Q_{\alpha,z}$ is drawn with horizontal blue lines, the region of $Q_{\alpha,z} \prec_{\log} P_\alpha$ is with vertical red lines, and the remaining regions are:

(a) $1 < \alpha < 2$ and $\alpha - 1 < z < \alpha/2$,

(b) $\alpha > 2$ and $\alpha/2 < z < \alpha - 1$.



Conjecture

For any α, z in (a) and (b), there is a pair A, B such that neither $P_\alpha(A,B) \prec_{\log} Q_{\alpha,z}(A,B)$ nor $Q_{\alpha,z}(A,B) \prec_{\log} P_\alpha(A,B)$ holds.

A partial result for the above is the following:

Proposition

Assume that $\alpha > 1$ and E is an orthogonal projection with $EB \neq BE$. Then:

- $P_\alpha(E,B) \prec_{\log} Q_{\alpha,z}(E,B)$ if and only if $z \leq \alpha - 1$.
- $Q_{\alpha,z}(E,B) \prec_{\log} P_\alpha(E,B)$ if and only if $z \geq \alpha - 1$.

Corollary

- If $0 < \alpha \le 2$ and $\alpha \ne 1$, then

$$D_\alpha^*(A\|B) \le D_\alpha(A\|B) \le \widehat{D}_\alpha(A\|B).$$

- If $\alpha \ge 2$, then

$$D_\alpha^*(A\|B) \le \widehat{D}_\alpha(A\|B) \le D_\alpha(A\|B).$$

- As $\alpha \to 1$,

$$\begin{aligned} D(A\|B) \ &\left[= D_1(A\|B) = \mathrm{Tr}\, A(\log A - \log B)\right] \\ \le D_{\mathrm{BS}}(A\|B) \ &\left[= \widehat{D}_1(A\|B) = \mathrm{Tr}\, A \log(A^{1/2}B^{-1}A^{1/2})\right]. \end{aligned}$$

$D \le D_{\mathrm{BS}}$ was first shown in[13]

[13] F.H. and D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability, *Comm. Math. Phys.* **143** (1991), 99–114.

Proposition

If A, B are not commuting and $\alpha \neq 2$, then all of the above inequalities are strict.

The proof is based on[14] [15]

[14] F.H., Equality cases in matrix norm inequalities of Golden-Thompson type, *Linear and Multilinear Algebra* **36** (1994), 239–249.

[15] F.H. and M. Mosonyi, Different quantum f-divergences and the reversibility of quantum operations, arXiv:1604.03089.

(Log-)supermajorization

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and $\mathbf{a}, \mathbf{b} \geq 0$.

- The supermajorization $\mathbf{a} \prec^w \mathbf{b}$ means that

$$\sum_{i=1}^{k} a_{[n+1-i]} \geq \sum_{i=1}^{k} b_{[n+1-i]}, \qquad 1 \leq k \leq n.$$

- The log-supermajorization $\mathbf{a} \prec^{w \log} \mathbf{b}$ means that

$$\prod_{i=1}^{k} a_{[n+1-i]} \geq \prod_{i=1}^{k} b_{[n+1-i]}, \qquad 1 \leq k \leq n.$$

Note

- When $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$,

$$\mathbf{a} \prec \mathbf{b} \iff \mathbf{a} \prec_w \mathbf{b} \iff \mathbf{a} \prec^w \mathbf{b}.$$

- When $\prod_{i=1}^n a_i = \prod_{i=1}^n b_i > 0$,

$$\mathbf{a} \prec_{\log} \mathbf{b} \iff \mathbf{a} \prec_{w \log} \mathbf{b} \iff \mathbf{a} \prec^{w \log} \mathbf{b}.$$

(Log-)supermajorization for matrices

For $A, B \in \mathbb{M}_n^+$ we write

- $A \prec^w B$ if $\lambda(A) \prec^w \lambda(B)$.
- $A \prec^{w \log} B$ if $\lambda(A) \prec^{w \log} \lambda(B)$.

Note

- $$A \prec^w B \iff -A \prec_w -B.$$
- When A, B are invertible,

$$A \prec^{w \log} B \iff \log A \prec^w \log B$$
$$\iff A^{-1} \prec_{w \log} B^{-1}.$$

Symmetric anti-norms[16] [17] [18]

Definition A symmetric anti-norm $\|\cdot\|_!$ on $\mathbb{M}_n^+$ is a non-negative continuous functional such that

1. $\|\alpha A\|_! = \alpha \|A\|_!$ for all $A \in \mathbb{M}_n^+$ and all reals $\alpha \ge 0$,
2. $\|A\|_! = \|UAU^*\|_!$ for all $A \in \mathbb{M}_n^+$ and all unitary matrices U,
3. $\|A + B\|_! \ge \|A\|_! + \|B\|_!$ for all $A, B \in \mathbb{M}_n^+$.

[16] J.-C. Bourin and F.H., Norm and anti-norm inequalities for positive semi-definite matrices, *Internat. J. Math.* **22** (2011), 1121–1138.

[17] J.-C. Bourin and F.H., Jensen and Minkowski inequalities for operator means and anti-norms, *Linear Algebra Appl.* **456** (2014), 22–53.

[18] J.-C. Bourin and F.H., Anti-norms on finite von Neumann algebras, *Publ. Res. Inst. Math. Sci.* **51** (2015), 207–235.

Definition Let $\|\cdot\|$ be a symmetric norm on $\mathbb{M}_n$ and $p > 0$. For $A \in \mathbb{M}_n^+$ define

$$\|A\|_! := \begin{cases} \|A^{-p}\|^{-1/p} & \text{if } A \text{ is invertible,} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\|\cdot\|_!$ is a symmetric anti-norm. A symmetric anti-norm $\|\cdot\|_!$ defined in this way is called a derived anti-norm.

Examples

- The Ky Fan k-anti-norm on $\mathbb{M}_n^+$ is

$$\|A\|_{\{k\}} := \sum_{j=1}^{k} \lambda_{n+1-j}(A),$$

- For $p > 0$ and $k = 1, \dots, n$,

$$\|A\|_{-p,k} := \left(\sum_{j=1}^{k} \lambda_{n+1-j}^{-p}(A) \right)^{-1/p} = \|A^{-p}\|_{(k)}^{-1/p}.$$

Examples (Cont.)

- For $k = 1, \dots, n$,

$$\Delta_k(A) := \left(\prod_{j=1}^{k} \lambda_{n+1-j}(A)\right)^{1/k}.$$

In particular, $\Delta_n = \det^{1/d}$.

Note

$$\Delta_k(A) = \lim_{p \searrow 0} \left(\frac{1}{k}\sum_{j=1}^{k} \lambda_{n+1-j}^{-p}(A)\right)^{-1/p} = \lim_{p \searrow 0} k^{1/p} \|A\|_{-p,k}.$$

Therefore, Δ_k is a limit point of the derived anti-norms.

Proposition

Concerning the following conditions for $A, B \in \mathbb{M}_n^+$, we have

$$(a) \iff (b) \iff (c) \iff (d) \implies (e) \iff (f) \iff (g) \iff (h) \iff (i).$$

(a) $A \prec^w B$, i.e., $\|A\|_{\{k\}} \ge \|B\|_{\{k\}}$ for every $k = 1, \dots, d$;

(b) $\|A\|_! \ge \|B\|_!$ for every symmetric anti-norm $\|\cdot\|_!$;

(c) $\|f(A)\|_! \ge \|f(B)\|_!$ for every symmetric anti-norm $\|\cdot\|_!$ and every continuous non-decreasing concave function $f : [0, \infty) \to [0, \infty)$;

(d) $\|f(A)\| \le \|f(B)\|$ for every symmetric norm $\|\cdot\|$ and every non-increasing convex function $f : (0, \infty) \to [0, \infty)$, where $\|f(A)\|$ for non-invertible A is defined as $\|f(A)\| := \lim_{\varepsilon \searrow 0} \|f(A + \varepsilon I)\|$;

Proposition (Cont.)

(e) $A \prec^{w \log} B$, i.e., $\Delta_k(A) \ge \Delta_k(B)$ for every $k = 1, \dots, n$;

(f) $\|A\|_{-p,k} \ge \|B\|_{-p,k}$ for every $k = 1, \dots, n$ and every $p > 0$;

(g) $\|f(A)\|_! \ge \|f(B)\|_!$ for every derived anti-norm $\|\cdot\|_!$ and every continuous non-decreasing function $f : [0, \infty) \to [0, \infty)$ such that $\log f(e^x)$ is concave on $\mathbb{R}$;

(h) $\det f(A) \ge \det f(B)$ for every continuous non-decreasing function $f : [0, \infty) \to [0, \infty)$ such that $\log f(e^x)$ is concave on $\mathbb{R}$;

(i) $\|f(A)\| \le \|f(B)\|$ for every symmetric norm $\|\cdot\|$ and every non-increasing function $f : (0, \infty) \to [0, \infty)$ such that $f(e^x)$ is convex on $\mathbb{R}$, where $\|f(A)\|$ for non-invertible A is as in (d).

Final remarks

1. The powerful method to obtain symmetric norm inequalities (in particular, trace inequalities) is weak (log-)majorization, as

$$\begin{array}{ccc} \text{weak log-majorization} & \Longleftrightarrow & \text{power symmetric norm inequality} \\ & & \|A^p\| \le \|B^p\| \text{ for all } p>0 \\ \Downarrow & & \Downarrow \\ \text{weak majorization} & \Longleftrightarrow & \text{symmetric norm inequality.} \end{array}$$

2. A counterpart of the above is the relation between (log-)supermajorization and symmetric (derived) anti-norms, as

$$\begin{array}{ccc} \text{supermajorization} & \Longleftrightarrow & \text{symmetric anti-norm inequality} \\ \Downarrow & & \Downarrow \\ \text{log-supermajorization} & \Longleftrightarrow & \text{derived anti-norm inequality} \\ & & \|A^{-p}\| \le \|B^{-p}\| \text{ for all } p>0. \end{array}$$

3. When matrix functions are made from operations of products, absolute values and powers (like $|A^p B^q \cdots|^r$), the antisymmetric power technique is quite useful to obtain log-majorizations between such matrix functions. This technique reduces log-majorizations to simple operator inequalities.

4. Important quantities in quantum information are mostly matrix trace functions. Hence, the log-majorization method is often very useful.

5. Beyond the case in 3, we have characterizations of (log-)majorizations in the following forms of logarithmic integral average of eigenvalues:

$$\lambda(A) \prec_{w\log} \exp \int_\Xi \log \lambda(B_\xi)\, d\nu(\xi) \quad \text{and} \quad \lambda(A) \prec^{w\log} \exp \int_\Xi \log \lambda(B_\xi)\, d\nu(\xi)$$

$$\Downarrow \qquad\qquad\qquad\qquad \Uparrow$$

$$\lambda(A) \prec_{w\log} \int_\Xi \lambda(B_\xi)\, d\nu(\xi) \qquad\qquad \lambda(A) \prec^{w\log} \int_\Xi \lambda(B_\xi)\, d\nu(\xi)$$

$$\Downarrow \qquad\qquad\qquad\qquad \Uparrow$$

$$\lambda(A) \prec_{w} \int_\Xi \lambda(B_\xi)\, d\nu(\xi) \qquad\qquad \lambda(A) \prec^{w} \int_\Xi \lambda(B_\xi)\, d\nu(\xi)$$

in terms of inequalities with respect to symmetric (anti-)norms [19].

[19] F.H., R. König and M. Tomamichel, Generalized log-majorization and multivariate trace inequalities, arXiv:1609.01999.

Thank you!