[Joint convexity of the quantum relative entropy]

Theorem (Lieb). For every Hermitian matrix H, the map $A \mapsto \operatorname{Tr}(e^{H + \log A})$ is concave on positive matrices.

A - random matrix $\quad \mathbb{E}\operatorname{Tr}(e^{H+\log A}) \le \operatorname{Tr}(e^{H + \log \mathbb{E}A})$

$$a \mapsto e^{h + \log a} = a \cdot e^{h}$$

$p, q \in \mathbb{R}^n_+$ (prob. distr.) on $\{1, 2, \dots, n\}$

$$D(p \| q) := \sum_{i=1}^{n} \Big(p_i \log \frac{p_i}{q_i} + \underbrace{(q_i - p_i)}_{=0 \text{ for prob.}} \Big)$$

q = prior distr.

p = actual distr.

$X_1, X_2 \cdots X_n \longrightarrow$

$X_i \sim q$ iid

Stein's Lemma: Given $X_1, X_2, \dots, X_n \sim p$ or $X_1, \dots, X_n \sim q$

null hypothesis

min prob you answer q when $X_1, \dots, X_n \sim p$

(*) [answer q w/ prob. $\to 1$ as $n \to \infty$

opt. $\mathbb{P}[\text{error}] = \underbrace{e^{-n(D(p\|q) + o(1))}}$

$$D(p \| q) := \sum_{i=1}^{n} \Big(p_i \log \frac{p_i}{q_i} + (q_i - p_i) \Big)$$

Joint convexity: $\overbrace{(p, q)}^{2n} \longmapsto D(p \| q)$ convex

For all $p_1, p_2, q_1, q_2 \in \mathbb{R}^n_+$ prob. distrs and $t \in [0,1]$

$$D(\underbrace{tp_1+(1-t)p_2}_{P} \,\|\, \underbrace{tq_1+(1-t)q_2}_{Q}) \le t\,D(p_1\|q_1) + (1-t)D(p_2\|q_2)$$

$X_1,\dots,X_n \sim tp_1+(1-t)p_2$ $\quad$ $X_1,\dots,X_n \sim tq_1+(1-t)q_2$

w/ prob t $\quad$ w/ prob $1-t$

$$\begin{cases} X_1,\dots,X_n \sim p_1 \text{ vs. } X_1,\dots,X_n \sim q_1 \\ X_1,\dots,X_n \sim p_2 \text{ vs. } X_1,\dots,X_n \sim q_2 \end{cases}$$

$$\boxed{D(p\|q) := \sum_{i=1}^n \left(p_i \log\frac{p_i}{q_i} + (q_i - p_i)\right)}$$

$\lim_{\varepsilon\to 0} \varepsilon \cdot \log \varepsilon = 0$

$0 \cdot \log 0 = 0$

$$\phi(p) = \sum_i (p_i \log p_i - p_i) \quad \text{(negative entropy)}$$

$\frac{1}{p_i} \ge 0$ for $p_i > 0$

$$[\nabla\phi(p)] = \log(p_i)$$

convexity of ϕ

$$D(p\|q) = \phi(p) - \phi(q) - \langle \nabla\phi(q), p-q\rangle \ge 0 \quad \forall p,q \in \mathbb{R}^n_+$$

Suffices: $\boxed{g : (a,b) \mapsto a \log\frac{a}{b} + (b-a)}$ is convex

$\mathbb{R}^2_+$

$$\nabla^2 g(a,b) \succeq 0 \qquad \nabla^2 g(a,b) = \begin{pmatrix} 1/a & -1/b \\ -1/b & a/b^2 \end{pmatrix}$$

$$\left.\begin{array}{l} \mathrm{tr} \ge 0 \Rightarrow \lambda_1+\lambda_2 \ge 0 \\ \det = 0 \Rightarrow \lambda_1\lambda_2 = 0 \end{array}\right\} \Rightarrow \lambda_1, \lambda_2 \ge 0$$

$A \in M_{n^2}$

$\boxed{A \mapsto A \log A}$

$$\boxed{\Phi(A) = \mathrm{Tr}(A\log A - A)}$$

$A = \mathrm{diag}(p_1,\dots,p_n)$

$\sum_i p_i \log p_i - p_i$

$$\boxed{\nabla \Phi(A) = \log A}$$

$A \succ 0$

Lemma: $f : (0,\infty) \to \mathbb{R}_+$ convex $\Rightarrow A \mapsto \mathrm{Tr}(f(A))$ convex on $A \succ 0$

$$\begin{bmatrix} \boxed{a_{11}\log a_{11} - a_{11}} & & 0 \\ & \ddots & \\ 0 & & a_{nn}\log a_{nn} - a_n \end{bmatrix}$$

$\log a_{11}$

$$S(A\|B) := \Phi(A) - \Phi(B) - \langle \nabla\Phi(B), A-B\rangle$$
$$= \mathrm{Tr}(A(\log A - \log B) - (A-B)) \geq 0$$

$$\underbrace{p_i \log \frac{p_i}{q_i}}_{A\log(B^{-1/2}AB^{-1/2})} = p_i(\log p_i - \log q_i)$$

$$S(A\|B) = \mathrm{Tr}(A(\log A - \log B) - (A-B)) \qquad A, B \succeq 0$$

Goal: Jointly convex

$$H[A,B] = (A^{-1} + B^{-1})^{-1} \quad \text{jointly concave}$$

f op. convex:

$f(tA + (1-t)B) \preceq t f(A) + (1-t) f(B)$

Lemma: $(X,B) \mapsto XB^{-1}X^*$ is jointly convex on $M_n(\mathbb{C}) \times H^n_{++}$

$(x,y) \mapsto x^2/y$ is convex for $y > 0$.

Claim: $A, B \succ 0$, $X \in M_n(\mathbb{C})$

$$\begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \succeq 0 \iff A \succeq XB^{-1}X^*$$

$$\begin{bmatrix} A - XB^{-1}X^* & 0 \\ 0 & B \end{bmatrix} = \underbrace{\begin{bmatrix} I & -XB^{-1} \\ 0 & I \end{bmatrix}}_{U} \begin{bmatrix} A & X \\ X^* & B \end{bmatrix} \underbrace{\begin{bmatrix} I & 0 \\ -B^{-1}X^* & I \end{bmatrix}}_{U^*}$$

PSD $\iff A \succeq X B^{-1} X^*$

Lemma: $(X, B) \mapsto X B^{-1} X^*$ is jointly convex on $M_n(\mathbb{C}) \times H^n_{++}$

midpoint op. convex

$$\left(\frac{X_1 + X_2}{2}\right)\left(\frac{B_1 + B_2}{2}\right)^{-1}\left(\frac{X_1 + X_2}{2}\right)^* \preceq \underbrace{\frac{X_1 B_1^{-1} X_1^* + X_2 B_2^{-1} X_2^*}{2}}_{A}$$

$$\underbrace{\begin{bmatrix} X_1 B_1^{-1} X_1^* & X_1 \\ X_1^* & B_1 \end{bmatrix}}_{M_1}, \underbrace{\begin{bmatrix} X_2 B_2^{-1} X_2^* & X_2 \\ X_2^* & B_2 \end{bmatrix}}_{M_2} \succeq 0$$

$$\begin{bmatrix} \overbrace{\frac{X_1 B_1^{-1} X_1^* + X_2 B_2^{-1} X_2^*}{2}}^{A} & \overbrace{\frac{X_1 + X_2}{2}}^{X} \\ \underbrace{\left(\frac{X_1 + X_2}{2}\right)^*}_{X^*} & \underbrace{\frac{B_1 + B_2}{2}}_{B} \end{bmatrix} \succeq 0$$

Lemma: $(X, B) \mapsto X B^{-1} X^*$ is jointly convex on $M_n(\mathbb{C}) \times H^n_{++}$

$\Psi(X, B) = X B^{-1} X^*$

$H[A, B]$ jointly concave

$\downarrow$ (?)

$S(A \| B)$ jointly convex

$\downarrow$ (?)

Lieb's concavity thm

$H[A,B] = (A^{-1}+B^{-1})^{-1}$ jointly concave

$$\boxed{H[A,B] \stackrel{?}{=} A - \boxed{A(A+B)^{-1}A}}$$

$$= A - \boxed{\Psi(A, A+B)}$$

$(A,B) \mapsto -\Psi(A, A+B)$ joint concave

$$\boxed{(A^{-1}+B^{-1})^{-1} = A - A(A+B)^{-1}A}$$

$\underline{A,B > 0}$

$\checkmark \quad I = (A+B)^{-1}(A+B)$

Lemma: $(X,B) \mapsto XB^{-1}X^*$ is jointly convex on $M_n(\mathbb{C}) \times H^n_{++}$

$\checkmark \quad I - (A+B)^{-1}A = (A+B)^{-1}B$

$\Rightarrow \quad A - A(A+B)^{-1}A = \underline{A(A+B)^{-1}B}$

$$\boxed{A(A+B)^{-1}B = (A^{-1}+B^{-1})^{-1}}$$

$\Leftrightarrow \quad (A+B)^{-1}B = \underline{A^{-1}(A^{-1}+B^{-1})^{-1}}$

$$(A+B)^{-1}\underline{B}(A^{-1}+\underline{B^{-1}})A$$

$$= (A+B)^{-1}(B+A) = I$$

$$\underline{A^p} = c_p \int_0^\infty t^p \underline{(tI+A)^{-1}} dt$$

$$\underbrace{A^{-1/2}(A^{-1}+B^{-1})^{-1}A^{-1/2}} = A^{-1/2}(A - A(A+B)^{-1}A)$$

$$(I + A^{1/2}B^{-1}A^{1/2})^{-1} = I - A^{1/2}(A+B)^{-1}A^{1/2}$$

$$= I - (I + A^{-1/2}BA^{-1/2})^{-1}$$

$$(I + X)^{-1} = I - (I + X^{-1})^{-1}$$

$$\frac{1}{1+x} = 1 - \frac{1}{1+\frac{1}{x}} = 1 - \frac{x}{x+1}$$

$$= \frac{1}{1+x} \quad \checkmark$$