Note on a Concavity Theorem of Lieb

Yao-Liang Yu
University of Alberta

January 9, 2011

Notation: All matrices are assumed to be Hermitian, if not otherwise stated. Define the partial ordering $A \succcurlyeq B$ if $A - B$ is positive semidefinite (PSD). $\mathcal{S}_+ = \{A : A \succcurlyeq 0\}$ is the PSD cone, and $\mathrm{Tr}(A)$ denotes the trace of A. $\mathbb{I}$ is the identity matrix with appropriate sizes.

The following theorem is due to Lieb [2]:

Theorem 1 *Fix a Hermitian matrix H, then $f(A) = \mathrm{Tr}\exp(H + \log A)$ is concave on $\mathcal{S}_+$.*

This theorem has been proven very useful in obtaining sharp non-commutative concentration inequalities by Tropp [4]. Tropp [5] recently also gave an elementary proof of Theorem 1 by the variational representation:

$$\mathrm{Tr}\exp(H + \log A) \;=\; \max_{X \succ 0} \; \mathrm{Tr}(XH) - S(X; A) + \mathrm{Tr}X, \tag{1}$$

where $S(X; A) := \mathrm{Tr}(X \log X - X \log A)$ is the quantum relative information. It is well-known that $S(X; A)$ is jointly convex in X and A; for one elegant proof, see [1]. The concavity claim of Theorem 1 then follows immediately from (1). We noted in passing that there is no necessity in assuming the attainability of the maximum in (1).

[3] contained a related theorem:

Theorem 2

$$\log \mathrm{Tr}\exp(H + \log A) \;=\; \max_{X \succ 0, \mathrm{Tr}X = 1} \; \mathrm{Tr}(XH) - S(X; A). \tag{2}$$

On the other hand, if $X \succ 0$ and $\mathrm{Tr}X = 1$, then

$$S(X; A) \;=\; \max_{H} \; \mathrm{Tr}(XH) - \log \mathrm{Tr}\exp(H + \log A)). \tag{3}$$

From (2), it is again immediate that $\log \mathrm{Tr}\exp(H + \log A)$ is concave on $\mathcal{S}_+$ (as a function of A). This also follows from Theorem 1 by the composition rule. On the other hand, being the pointwise maximum of affine functions, $\log \mathrm{Tr}\exp(H + \log A)$ is convex as a function of H. Similar conclusion can be drawn for $\mathrm{Tr}\exp(H + \log A)$, however, this time, the convexity of it follows from that of $\log \mathrm{Tr}\exp(H + \log A)$ by the composition rule. Somewhat surprisingly, there is a mutual implication between (1) and (2), which we prove now:

Proof: (1) $\Rightarrow$ (2):

$$\begin{aligned}
\max_{X \succ 0, \mathrm{Tr}X = 1} \mathrm{Tr}(XH) - S(X; A) &= \max_{X \succ 0, \mathrm{Tr}X = 1} \mathrm{Tr}(XH) - S(X; A) + \mathrm{Tr}X - 1 \\
\text{(introduce Lagrangian multiplier } \lambda) &= \max_{X \succ 0} \min_{\lambda} \; \mathrm{Tr}(XH) - S(X; A) + \mathrm{Tr}X - 1 + \lambda(\mathrm{Tr}X - 1) \\
\text{(strong duality)} &= \min_{\lambda} \max_{X \succ 0} \; \mathrm{Tr}[X(H + \lambda\mathbb{I})] - S(X; A) + \mathrm{Tr}X - 1 - \lambda \\
(1) &= \min_{\lambda} \; \mathrm{Tr}\exp(H + \lambda\mathbb{I} + \log A) - 1 - \lambda \\
&= \min_{\lambda} \; e^{\lambda} \cdot \mathrm{Tr}\exp(H + \log A) - 1 - \lambda \\
&= \log \mathrm{Tr}\exp(H + \log A),
\end{aligned}$$

where the last equality is due to the Legendre-Fenchel duality: $\log \sigma = \max_{\lambda > 0} \lambda\sigma - 1 - \log \lambda$.

(2) $\Rightarrow$ (1):

$$\begin{aligned}
\max_{X\succ 0}\ \mathrm{Tr}(XH) - S(X;A) + \mathrm{Tr}X &= \max_{X=\lambda\tilde{X},\tilde{X}\succ 0,\mathrm{Tr}\tilde{X}=1,\lambda>0} \mathrm{Tr}(XH) - S(X;A) + \mathrm{Tr}X \\
&= \max_{\lambda>0}\ \max_{\tilde{X}\succ 0,\mathrm{Tr}\tilde{X}=1} \mathrm{Tr}(\lambda\tilde{X}H) - S(\lambda\tilde{X};A) + \mathrm{Tr}\lambda\tilde{X} \\
&= \max_{\lambda>0}\ \max_{\tilde{X}\succ 0,\mathrm{Tr}\tilde{X}=1} \lambda\cdot\left[\mathrm{Tr}(\tilde{X}H) - S(\tilde{X};A)\right] + \lambda - \lambda\log\lambda \\
(2) \quad &= \max_{\lambda>0} \lambda\cdot\log\mathrm{Tr}\exp(H+\log A) + \lambda - \lambda\log\lambda \\
&= \mathrm{Tr}\exp(H+\log(A)),
\end{aligned}$$

where the last equality is due to the Legendre-Fenchel duality: $\exp(\sigma) = \max_{\lambda>0}\lambda\sigma - (\lambda\log\lambda - \lambda)$. ■

The proof of (2) in [3] is also elementary, hence combining it with the arguments above should give another (elementary?) proof of (1); and conversely, (1) can be used to give yet another proof of (2).

Notice that (3) is just the Legendre-Fenchel dual of (2). Similarly, from (1) we can obtain:

$$S(X;A) = \max_{H}\ \mathrm{Tr}(HX) - \mathrm{Tr}\exp(H+\log A) + \mathrm{Tr}X. \tag{4}$$

Given our experience above, it is not surprising anymore that there is a similar mutual implication between (3) and (4). We omit the details.

References

[1] Edward G. Effros. A matrix convexity approach to some celebrated quantum inequalities. *Proceedings of National Academy of Science*, (4):1006–1008, 2009.

[2] E. H. Lieb. Convex trace functions and the wigner-yanase-dyson conjecture. *Advances in Mathematics*, pages 267–288, 1973.

[3] Dènes Petz. A survey of certain trace inequalities. In *Functional Analysis and Operator Theory, Vol. 30 of Banach Center Publications*, pages 287–298. 1994.

[4] Joel. A. Tropp. User-friendly tail bounds for matrix martingales, 2010.

[5] Joel. A. Tropp. From joint convexity of quantum relative entropy to a concavity theorem of lieb, 2011.