Stein’s Lemma in the Quantum Hypothesis Testing

1.T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," in IEEE Transactions on Information Theory, vol. 46, no. 7, pp. 2428-2433, Nov. 2000, doi: 10.1109/18.887855.
2.Hiai, F., Petz, D. The proper formula for relative entropy and its asymptotics in quantum probability. *Commun.Math. Phys.* 143, 99–114 (1991). https://doi.org/10.1007/BF02100287

Haodong Yang

Hypothesis Testing
(Recap)

- Let $X_1, X_2, ..., X_n$ be $i.i.d$ ~ $Q(X)$.
- $H_1 : Q = P_1$
- $H_2 : Q = P_2$
- Given $X^n = (X_1, X_2, ..., X_n)$, We need to decide which hypothesis is true, Or equivalently to say
- $\hat{H} = H_1$ if $g(X^n) = 1$. $\hat{H} = H_2$ if $g(X^n) = 2$.

Error Type

- Type 1 error : $H = H_1, \hat{H} = H_2$ (False alarm)
- Type 2 error: $H = H_2, \hat{H} = H_1$ (Miss detection)
- $\alpha = P(\hat{H} = H_2 \,|\, H = H_1)$
- $\beta = P(\hat{H} = H_1 \,|\, H = H_2)$

Question: What is Quantum version of Hypothesis Testing.

Quantum Hypothesis testing

- The hypothesis testing problem of two quantum states (Density Operator).

Density Operator

- density matrices, also called density operators, which conceptually take the role of the state vectors, as they encode all the (accessible) information about a quantum mechanical system.

Density Matrix

Properties

- Density operator : $\rho = \sum_i p_i |\rho_i\rangle\langle\rho_i|$
- The expectation value of an observable A in a state, represented by a density matrix ρ , is given by $< A >_\rho = tr(\rho A)$

Quantum Hypothesis Testing

Problem Description

- Let $B(H)$ be the set of linear operators on H.
  $S(H) = \{\rho \in B(H) \,|\, \rho = \rho^* \geq 0, tr\{\rho\} = 1\}$.

- Null hypothesis $\rho \in S(H)$ versus alternative hypothesis $\sigma \in S(H)$.

- Decide which hypothesis is true, $\rho^{\otimes n}$ or $\sigma^{\otimes n}$,and the decision is given by a two-valued quantum measurement $\{A_n, I - A_n\}$ $(A_n \in B(H^{\otimes n}), 0 \leq A_n \leq I)$

- A_n corresponds acceptance of $\rho^{\otimes n}$ and $I - A_n$ corresponds acceptance of $\sigma^{\otimes n}$.

Types of Errors

$$\alpha_n(A_n) = tr(\rho^{\otimes n}(I - A_n))$$

$$\beta_n(A_n) = tr(\sigma^{\otimes n} A_n)$$

$\alpha_n(A_n)$ is the error probability of the acceptance of $\sigma^{\otimes n}$ when $\rho^{\otimes n}$ is true. (Type 1)

$\beta_n(A_n)$ is the error probability of the converse situation. (Type 2)

Asymmetric v.s. Symmetric case

- Asymmetric case :
  $\beta_n^*(\epsilon) = min\{\beta_n(A_n)\,|\,A_n \in B(H^{\otimes n}), 0 \leq A_n \leq I, \alpha_n(A_n) \leq \epsilon\}$

- Symmetric case:
  $p_{err}^*(p, q) = min\{p \cdot \alpha_n(A_n) + q \cdot \beta_n(A_n)\,|\,A_n \in B(H^{\otimes n}), 0 \leq A_n \leq I, p + q = 1\}$

Weak Converse property

versus

Strong Converse property

Asymmetric Quantum Hypothesis Testing

Weak converse property

$$\lim_{n\to\infty} \sup \frac{1}{n} \log \beta_n^*(\epsilon) \leq -D(\rho\,||\,\sigma) \qquad (1)$$

$$-\frac{1}{1-\epsilon} D(\rho\,||\,\sigma) \leq \lim_{n\to\infty} \inf \frac{1}{n} \log \beta_n^*(\epsilon) \qquad (2)$$

$$D(\rho\,||\,\sigma) = tr(\rho(\log \rho - \log \sigma))$$

Quantum Data Processing Inequality(QDPI)

For any quantum channel $\mathscr{E} : \mathscr{H}_A \to \mathscr{H}_B$, and density operator $\rho, \sigma \in \mathscr{D}(\mathscr{H}_A)$:

$$D(\mathscr{E}(\rho) \,||\, \mathscr{E}(\sigma)) \leq D(\rho \,||\, \sigma)$$

Proof

$$D(\rho^{\otimes n} || \sigma^{\otimes n})$$

$$\geq \alpha_n(A_n)\log\frac{\alpha_n(A_n)}{1-\beta_n(A_n)} + (1-\alpha_n(A_n))\log\frac{1-\alpha_n(A_n)}{\beta_n(A_n)} \cdots \text{(QDPI)}$$

$$\geq \alpha_n(A_n)(\log\alpha_n(A_n) - \log(1-\beta_n(A_n)) + (1-\alpha_n(A_n))(\log(1-\alpha_n(A_n)) - \log\beta_n(A_n))$$

$$\geq \alpha_n(A_n)(\log(\alpha_n(A_n))) + (1-\alpha_n(A_n))(\log(1-\alpha_n(A_n))) - \alpha_n(A_n)\log(1-\beta_n(A_n)) - (1-\alpha_n(A_n))\log\beta_n(A_n)$$

$$\geq -H(\alpha_n(A_n)) - \alpha_n(A_n)\log(1-\beta_n(A_n)) - (1-\alpha_n(A_n))\log\beta_n(A_n)$$

$$\geq -log2 - (1-\alpha_n(A_n))\log\beta_n(A_n) \ \cdots \text{(1)}$$

Continue

From equation (1):

We have :

$$(1-\alpha_n(A_n))\frac{1}{n}\log\beta_n(A_n) \geq -\frac{\log 2}{n} - D(\rho\,||\,\sigma) \quad \cdots \text{(2)}$$

From equation (2) and $(1-\alpha_n(A_n) \leq 1-\epsilon)$ we have:

$$(1-\epsilon)\frac{1}{n}\log\beta_n(A_n) \geq -\frac{\log 2}{n} - D(\rho\,||\,\sigma) \quad \cdots \text{(3)}$$

From equation (3) when $n \to \infty$:

$$-\frac{1}{1-\epsilon}D(\rho\,||\,\sigma) \leq \liminf_{n\to\infty}\frac{1}{n}\log\beta_n^*(\epsilon)$$

Asymmetric Quantum Hypothesis Testing

Weak converse property

$$(1-\alpha_n(A_n))\frac{1}{n}\log\beta_n(A_n) \geq -\frac{\log 2}{n} - D(\rho||\sigma) \quad \cdots \text{(2)}$$

From equation (2), setting $\beta_n(A_n) \leq e^{-nr}$ for $r > D(\rho||\sigma)$. Then we will get:

$$1-\alpha_n(A_n) \leq \frac{-log2 - nD(\rho||\sigma)}{\log\beta_n}$$

$$\Rightarrow \alpha_n(A_n) - 1 \geq \frac{log2 + nD(\rho||\sigma)}{\log\beta_n} \geq \frac{\log 2 + nD(\rho||\sigma)}{-nr}$$

$$\Rightarrow \alpha_n(A_n) \geq \frac{nD(\rho||\sigma)}{-nr} + 1 = \frac{D(\rho||\sigma)}{-r} + 1 > 0 \quad \text{(Since } r > D(\rho||\sigma)\text{)}$$

Asymmetric Quantum Hypothesis Testing

Weak converse property

Theorem: if $\beta_n(A_n) \leq e^{-nr}(r > D(\rho || \sigma))$,then $\alpha_n(A_n)$ does not go to zero as $n \to \infty$.

Asymmetric Quantum Hypothesis Testing

Strong Converse property

What we want to show:

if $\beta_n(A_n) \leq e^{-nr}(r > D(\rho \,||\, \sigma))$,then $\alpha_n(A_n)$ goes to one as $n \to \infty$.

Asymmetric Quantum Hypothesis Testing

Strong Converse property

Lemma 1: For any test A_n, we have :

$$tr(\rho^{\otimes n} - e^{n\lambda}\sigma^{\otimes n}X_{n,\lambda}) \geq tr(\rho^{\otimes n} - e^{n\lambda}\sigma^{\otimes n}A_n)$$

Theorem 1: For any test A_n and any $\lambda \in R$, we have:

$$1 - \alpha_n(A_n) \leq e^{-n\varphi(\lambda)} + e^{n\lambda}\beta_n(A_n)$$

Where

$$\varphi(\lambda) = \max_{0 \leq s \leq 1} \{\lambda s - \psi(s)\}$$

$$\psi(s) = \log tr(\rho^{1+s}\sigma^{-s})$$

Eigen(spectral)-decomposition:

$$\rho^{\otimes n} - e^{n\lambda}\sigma^{\otimes n} = \sum_j \mu_{n,j}E_{n,j}$$

$$X_{n,\lambda} = \sum_{j \in D_n} E_{n,j}$$ where $D_n = \{j \,|\, \mu_{n,j} > 0\}$

Lemma 1 proof

$$
\begin{aligned}
tr((\rho^{\otimes n} - e^{n\lambda}\sigma^{\otimes n})A_n) &= \sum_j \mu_{n,j} tr(E_{n,j}A_n) \\
&\leq \sum_{j \in D_n} \mu_{n,j} tr(E_{n,j}A_n) \\
&\leq \sum_{j \in D_n} \mu_{n,j} tr(E_{n,j}) \\
& tr((\rho^{\otimes n} - e^{n\lambda}\sigma^{\otimes n})X_{n,j})
\end{aligned}
$$

Convexity of $\psi(s)$

$$\varphi(\lambda) = \max_{0 \le s \le 1} \{\lambda s - \psi(s)\}$$

$$\psi(s) = \log tr(\rho^{1+s}\sigma^{-s})$$

Let $u = \rho^{1+s}\sigma^{-s}$, Then $\psi(s) = \log tr(u)$

$$\psi'(s) = \frac{1}{tr(u)}\frac{d}{ds}(tr(u))$$

$$\frac{d}{ds}tr(u) = \frac{d}{ds}(tr(\rho^{1+s}\sigma^{-s}))$$

$$= tr(\frac{d}{ds}(\rho^{1+s}\sigma^{-s}))$$

$$= tr(\rho^{1+s}\sigma^{-s}(\log\rho - \log\sigma))$$

Continue the proof

Combine all of them we got:

$$\psi'(s) = e^{-\psi(s)} tr(\rho^{1+s}\sigma^{-s}(\log \rho - \log \sigma))$$

Let $A = \log \rho - \log \sigma - \psi'(s)$

$$\psi''(s) = e^{-\psi(s)} tr(\rho^{1+s} A \sigma^{-s} A)$$

$$= e^{-\psi(s)} tr((\rho^{\frac{1+s}{2}} A \sigma^{-\frac{s}{2}})(\rho^{\frac{1+s}{2}} A \sigma^{-\frac{s}{2}})^*)$$

$$> 0 \quad \cdots\cdots \text{(a)}$$

Observation

More observation:

(1) $\psi(0) = 0$

(2) $\psi'(0) = D(\rho || \sigma)$

(3) $\varphi(\lambda) > 0$ if $\lambda > D(\rho || \sigma)$

(4) $s^* = \arg \max_{0 \leq s \leq 1} \{\lambda s - \psi(s)\} \iff \psi'(s^*) = \lambda$ if $D(\rho || \sigma) \leq \lambda \leq \psi'(1)$

Proof of Theorem 1

We define two probability distribution: $p_n = \{p_{n,j}\}, q = \{q_{n,j}\}$

$p_{n,j} = tr(\rho^{\otimes n} E_{n,j}), q_{n,j} = tr(\sigma^{\otimes n} E_{n,j})$ (Since $E_{n,j}$ are eigen-decomposition they are orthogonal to each other (i.e. $E_{n,j}E_{n,k} = 0$ if $j \neq k$))

$$\mu_{n,j} tr(E_{n,j}) = p_{n,j} - e^{n\lambda} q_{n,j}$$

$$\mu_{n,j} \geq 0 \iff p_{n,j} - e^{n\lambda} q_{n,j} \geq 0$$

$$D_n = \{j \,|\, 0 \leq \forall s \leq 1, e^{-n\lambda s} p_{n,j}^{s} q_{n,j}^{-s} \geq 1\}$$

Continue

A function f is said to be matrix convex of order n if for all n x n Hermitian matrices A and B and for all real numbers $0 \leq \lambda \leq 1$:

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

$$tr(\rho^{\otimes n} X_{n,j}) = \sum_{j \in D_n} tr(\rho^{\otimes n} E_{n,j})$$

$$= \sum_{j \in D_n} p_{n,j}$$

$$\leq \sum_{j \in D_n} p_{n,j} \cdot e^{-n\lambda s} p_{n,j}^{s} q_{n,j}^{-s}$$

$$\leq e^{-n\lambda s} \sum_{j} p_{n,j}^{1+s} q_{n,j}^{-s} \quad \text{(convex of q)}$$

$$\leq e^{-n\lambda s} tr((\rho^{\otimes n})^{1+s} (\sigma^{\otimes n})^{-s}) \cdots \text{(4)}$$

Note:

$f(u) = u^{-s} (0 \leq s \leq 1)$ is a convex function

Detail

$$tr((\rho^{\otimes n})^{1+s}(\sigma^{\otimes n})^{-s}) = tr(\sum_j E_{n,j}(\rho^{\otimes n})^{1+s}E_{n,j}(\sigma^{\otimes n})^{-s})$$

$$\geq \sum_j tr(E_{n,j}(\rho^{\otimes n})^{1+s})tr(E_{n,j}(\sigma^{\otimes n})^{-s})$$

Continue

We have :

$$tr(\rho^{\otimes n} X_{n,j}) \leq e^{-n(\lambda s - \psi(s))}$$

Hence we know that :

$tr(\rho^{\otimes n} X_{n,j}) \leq e^{-n\varphi(\lambda)}$ by taking the maximum

Finally we have:

$$\begin{aligned} 1 - \alpha_n(A_n) &= tr(\rho^{\otimes n} A_n) \\ &\leq tr(\rho^{\otimes n} - e^{n\lambda}\sigma^{\otimes n})X_{n,j} + e^{n\lambda} tr(\sigma^{\otimes n} A_n) \\ &\leq tr(\rho^{\otimes n} X_{n,j}) + e^{n\lambda} tr(\sigma^{\otimes n} A_n) \\ &\leq e^{-n\varphi(\lambda)} + e^{n\lambda}\beta_n(A_n) \quad \text{Proved!} \end{aligned}$$

Quantum Stein's lemma

Theorem 2: For any $0 \leq \epsilon < 1$ it holds that :

$$\lim_{n \to \infty} \frac{1}{n} \log \beta_n^*(\epsilon) = -D(\rho || \sigma)$$

Proof of theorem 2

From theorem 1 we know that :

$$1 - \epsilon \leq 1 - \alpha_n(A_n) \leq e^{-n\varphi(\lambda)} + e^{n\lambda}\beta_n(A_n)$$

$$\implies \beta_n(A_n) \geq e^{-n\lambda}(1 - \epsilon - e^{-n\varphi(\lambda)})$$

Let $\lambda = D(\rho\,||\,\sigma + \delta)(\delta > 0)$, From property of $\varphi(\lambda)$ we know that $\varphi(\lambda) > 0$ in this case

Hence $1 - \epsilon - e^{n\varphi(\lambda)} > 0$ for n sufficiently large.

This implies : $\dfrac{1}{n}\log \beta_n^*(\epsilon) \geq -\lambda + \dfrac{1}{n}\log(1 - \epsilon - e^{-n\varphi(\lambda)})$

$\implies \lim_{n\to\infty} \inf \geq -D(\rho\,||\,\sigma) - \delta$ for $\forall \delta > 0$

Strong Converse

Theorem 3: For any test A_n, if

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n(A_n) \leq -r, \quad \cdots\cdots(13)$$

Then

$$\limsup_{n\to\infty} \frac{1}{n} \log(1 - \alpha_n(A_n)) \leq -\varphi(\lambda^*) \quad \cdots\cdots (14)$$

Where λ^* is a real number which satisfies $\varphi(\lambda^*) = r - \lambda^*$. Moreover, $\varphi(\lambda^*)$ is a represented as:

$$\varphi(\lambda^*) = \max_{0\leq s\leq 1} \{\frac{s}{1+s}r - \frac{1}{1+s}\psi(s)\}. \quad \cdots\cdots (15)$$

Proof of (14)

For all $\delta > 0$, there exists n_0 such that:

$$\beta_n(A_n) \leq e^{-n(r-\delta)}, \forall n \geq n_0 \text{ from (13).}$$

Put $\lambda = \lambda*$ in theorem 1, we have:

$$1 - \alpha_n(A_n) \leq e^{-n\varphi(\lambda*)} + e^{-n(r-\lambda*-\delta)}, \forall n \geq n_0$$

$$\iff 1 - \alpha_n(A_n) \leq 2e^{-n(\varphi(\lambda*)-\delta)}$$

$$\iff \limsup_{n\to\infty} \frac{1}{n} \log(1 - \alpha_n(A_n)) \leq -\varphi(\lambda*) + \delta$$

Since δ is arbitrary, (14) has been proved!

Proof of (15)

Suppose $\psi'(0) \leq r \leq 2\psi'(1) - \psi(1)$,

Define: $u(r) = \varphi(\lambda^*) = \max_{0 \leq s \leq 1} \{s\lambda^* - \psi(s)\} = r - \lambda^*$

Using observation (4) $\lambda^* = \psi'(s^*)$ we mentioned before:

$u(r) = s^*\psi'(s^*) - \psi(s^*)$ where $r = (s^* + 1)\psi'(s^*) - \psi(s^*)$

Combine both we got $u(r) = \dfrac{s^*}{s^* + 1} r - \dfrac{1}{s^* + 1}\psi(s^*)$.

Continue

Let's see the derivative of function:

$$g(s) = \frac{s}{s+1}r - \frac{1}{s+1}\psi(s)$$

$$g'(s) = \frac{1}{(s+1)^2}(r + \psi(s) - (1+s)\psi'(s))$$

We just need to see how function $h(s) = r + \psi(s) - (1+s)\psi'(s)$ works.

$h'(s) = -(1+s)\psi''(s) \leq 0$ by equation (a), which says that the sign of

$g'(s)$ changes at most once.

Therefore $g(s)$ get maximum value at $h(s) = 0 \iff r = (s+1)\psi'(s) - \psi(s)$

So we got $u(r) = \max_{0 \leq s \leq 1} g(s)$

Clearly if $r \geq 2\psi'(1) - \psi(1)$:

$$\varphi(\lambda^*) = \frac{1}{2}r - \frac{1}{2}\psi(1) = g(1) = \max_{0 \leq s \leq 1} g(s)$$

If $r \leq \psi'(0)$:

$$\varphi(\lambda^*) = 0 = g(0) = \max_{0 \leq s \leq 1} g(s)$$

Strong Converse

Corollary 1: For any test A_n, if

$$\limsup_{n\to\infty} \frac{1}{n} \log \beta_n(A_n) < -D(\rho\,||\,\sigma)$$

Then $\alpha_n(A_n)$ goes to one exponentially .

Proof: set $r = -D(\rho\,||\,\sigma) - \delta, \forall \delta > 0$.

We will get $1 - \alpha_n(A_n) \leq 2^{-n(D(\rho||\sigma)-\delta-\lambda^*)}$ by theorem 3.