Recap

- Binary state discrimination:

  given two quantum states ρ_0, ρ_1 with equal prob. $\frac{1}{2}$ each

- Goal: Decide which one you received

- Strategy: measure state $\to$ 2-element POVM $\Lambda = \{\Lambda_0, \Lambda_1\}$

  where $\Lambda_0, \Lambda_1 \geq 0$, $\Lambda_0 + \Lambda_1 = \mathbb{1}$ (i.e. $\Lambda_1 = \mathbb{1} - \Lambda_0$)

- Unknown state $\sigma \in \{\rho_0, \rho_1\}$:

  outcome "0" $\leftrightarrow \rho_0$ with probability $\mathrm{tr}(\Lambda_0 \sigma)$

  outcome "1" $\leftrightarrow \rho_1$ with probability $\mathrm{tr}\big((\mathbb{1} - \Lambda_0)\sigma\big)$

- Success probability: $$p_{succ} = \max_{\Lambda \text{ POVM}} \frac{1}{2} \Pr(\rho_0|\rho_0) + \frac{1}{2} \Pr(\rho_1|\rho_1)$$

  $$= \frac{1}{2}\Big(1 + \max_{0 \leq \Lambda \leq \mathbb{1}} \mathrm{tr}\, \Lambda (\rho_0 - \rho_1)\Big)$$

- trace distance: $$\max_{0 \leq \Lambda \leq \mathbb{1}} \mathrm{tr}\, \Lambda(\rho_0 - \rho_1) = \frac{1}{2} \|\rho_0 - \rho_1\|_1 \qquad (*)$$

  with trace norm: $\|X\|_1 = \mathrm{tr}\sqrt{X^\dagger X} = \sum_{i=1}^{d} s_i(X)$

- $s_i(X)$ are the singular values of $X \leftrightarrow$ eigenvalues of $|X| = \sqrt{X^\dagger X}$

Corollary to Lemma 1: Optimal measurement Λ in

$$\frac{1}{2}\|\rho_0-\rho_1\|_1 = \max_{0\leq\Lambda\leq\mathbb{1}} \operatorname{tr}\Lambda(\rho_0-\rho_1)$$

is given by $\Pi = \{\rho_0-\rho_1 \geq 0\}$ (projective measurement)

For Hermitian X with spectral decomposition $X=\sum_i \lambda_i |i\rangle\langle i|$,

$$\{X\geq 0\} = \sum_{i:\lambda_i\geq 0} |i\rangle\langle i| \quad (\text{similarly for } >, <, \leq)$$

Recall: $p_{succ} = \frac{1}{2}\left(1+\frac{1}{2}\|\rho_0-\rho_1\|_1\right)$

·) If $\frac{1}{2}\|\rho_0-\rho_1\|_1 = 1$, then $p_{succ}=1$

·) If $\frac{1}{2}\|\rho_0-\rho_1\|_1 = 0$, then $p_{succ}=\frac{1}{2}$

=> trace distance between two states is a measure of distinguishability

$\boxed{\text{Prop 2}}$ Let $T(\rho_0,\rho_1) = \frac{1}{2}\|\rho_0-\rho_1\|_1$.

For any two states ρ_0,ρ_1 and a quantum channel $\mathcal{N}$,

$$T(\rho_0,\rho_1) \geq T(\mathcal{N}(\rho_0),\mathcal{N}(\rho_1)).$$

"Data-processing inequality"

Proof: Lemma 1: $\frac{1}{2}\|\rho_0-\rho_1\|_1 = \max_{0\le\Lambda\le\mathbb{1}} \operatorname{tr}\Lambda(\rho_0-\rho_1)$

Let $\Lambda\ge 0$ with $\Lambda\le\mathbb{1}$ be optimal for $\frac{1}{2}\|\mathcal{N}(\rho_0)-\mathcal{N}(\rho_1)\|_1$:

$$\begin{aligned}\frac{1}{2}\|\mathcal{N}(\rho_0)-\mathcal{N}(\rho_1)\|_1 &= \operatorname{tr}\Lambda\left(\mathcal{N}(\rho_0)-\mathcal{N}(\rho_1)\right)\\ &= \operatorname{tr}\Lambda\,\mathcal{N}(\rho_0-\rho_1)\\ &= \operatorname{tr}\mathcal{N}^\dagger(\Lambda)(\rho_0-\rho_1) \quad \text{by definition of the adjoint } \mathcal{N}^\dagger\end{aligned}$$

want to show: i) $\mathcal{N}^\dagger(\Lambda)\ge 0$

ii) $\mathcal{N}^\dagger(\Lambda)\le\mathbb{1}$

i) $\mathcal{N}^\dagger$ is CP (because $\mathcal{N}$ is CP), in particular it's positive:

$$\Lambda\ge 0 \Rightarrow \mathcal{N}^\dagger(\Lambda)\ge 0 \quad ✓$$

ii) $\mathcal{N}^\dagger$ is unital: $\mathcal{N}^\dagger(\mathbb{1})=\mathbb{1}$: $\mathcal{N}^\dagger(\mathbb{1}-\Lambda)\ge 0$

$\Lambda\le\mathbb{1} \Leftrightarrow \mathbb{1}-\Lambda\ge 0$

$$\left.\begin{aligned}&\mathcal{N}^\dagger(\mathbb{1}-\Lambda)\ge 0\\ &= \mathcal{N}^\dagger(\mathbb{1})-\mathcal{N}^\dagger(\Lambda)\\ &= \mathbb{1}-\mathcal{N}^\dagger(\Lambda)\end{aligned}\right\}\ \mathcal{N}^\dagger(\Lambda)\le\mathbb{1}.$$

i)+ii): $\mathcal{N}^\dagger(\Lambda)$ is feasible in $\max_{0\le K\le\mathbb{1}} \operatorname{tr}K(\rho_0-\rho_1)$

$\square$

$$\Rightarrow \frac{1}{2}\|\rho_0-\rho_1\|_1 = \max_{0\le K\le\mathbb{1}} \operatorname{tr}K(\rho_0-\rho_1) \ge \operatorname{tr}\mathcal{N}^\dagger(\Lambda)(\rho_0-\rho_1) = \frac{1}{2}\|\mathcal{N}(\rho_0)-\mathcal{N}(\rho_1)\|_1$$

$$\frac{1}{2}\|\rho_0-\rho_1\|_1 \geq \frac{1}{2}\|\mathcal{N}(\rho_0)-\mathcal{N}(\rho_1)\|_1$$

§ 1.2 Error analysis and hypothesis testing

$$P_{succ} = \frac{1}{2} \Pr(\rho_0|\rho_0) + \frac{1}{2} \Pr(\rho_1|\rho_1)$$

error: $P_{error} = 1 - P_{succ} = \frac{1}{2}\Big(\Pr(\rho_1|\rho_0) + \Pr(\rho_0|\rho_1)\Big)$

"different errors"

Hypothesis testing: null hypothesis H_0 (ρ_0)

alternative hypothesis H_1 (ρ_1)

α = type-I error: "infer ρ_1 when you have ρ_0" false reject

β = type-II error: "infer ρ_0 when you have ρ_1" false accept

trade-off between these two errors

·) symmetric hypothesis testing : $\min(\alpha+\beta)$ $\longrightarrow$ discussion above,

("Bayesian") leads to trace dist.

·) asymmetric hypothesis testing

assume that type-I error is constant (and small),

$$\alpha = \varepsilon \quad \text{for some chosen } \varepsilon > 0.$$

Q: how small can I make β under this constraint?

Realistic settings: avoid false negatives at all costs!

$\boxed{\text{Def 3}}$ Let ρ, σ be two quantum states ($\rho \leftrightarrow$ null hypothesis, $\sigma \leftrightarrow$ alt. hypothesis), and let Λ with $0 \leq \Lambda \leq \mathbb{1}$ be a test operator defining a 2-element POVM $\{\underset{\substack{\uparrow\\ \rho}}{\Lambda}, \underset{\substack{\uparrow\\ \sigma}}{\mathbb{1} - \Lambda}\}$

$$\alpha(\Lambda) = \operatorname{tr} \rho (\mathbb{1} - \Lambda) \qquad \text{type-I error}$$

$$\beta(\Lambda) = \operatorname{tr} \sigma \Lambda \qquad \text{type-II error}$$

Information-theoretic setting: $\rho^{\otimes n}$ vs. $\sigma^{\otimes n}$, $n \to \infty$

$$\alpha_n(\Lambda_n) = \operatorname{tr} \rho^{\otimes n} (\mathbb{1} - \Lambda_n)$$

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

$$\Lambda_n \in \mathcal{B}(\mathcal{H}^{\otimes n}), \ \Lambda_n \geq 0, \quad \Lambda_n \leq \mathbb{1}_{\mathcal{H}}^{\otimes n}$$

$$\beta_n^*(\varepsilon) = \min \{ \beta_n(\Lambda_n) : 0 \leq \Lambda_n \leq \mathbb{1}, \alpha_n(\Lambda_n) \leq \varepsilon \} \quad \text{for } \varepsilon > 0.$$

Q: how does $\beta_n^*(\varepsilon)$ behave as $n \to \infty$? i.e., $\beta_n^*(\varepsilon) = f(n) \to 0$, what is f?

$\boxed{\text{Def 4}}$ a) For a linear operator $X \in \mathcal{B}(\mathcal{H})$, the support of X is defined as $\operatorname{supp} X = (\ker X)^\perp$.

If X is Hermitian with spectral decomposition $X = \sum_i \lambda_i |i\rangle\langle i|$ ($\lambda_i \in \mathbb{R}$, $\langle i|j\rangle = \delta_{ij}$), then $\operatorname{supp} X = \operatorname{span}\{|i\rangle : \lambda_i \neq 0\}$.

The projection onto $\operatorname{supp} X$ is given by

$$\sum_{i:\lambda_i \neq 0} |i\rangle\langle i| = \lim_{\alpha \to 0} X^\alpha = X^0$$

b) Let $\rho \geq 0$, $\operatorname{tr}\rho = 1$, $\sigma \geq 0$. Then the relative entropy $D(\rho\|\sigma)$ is defined as

$$D(\rho\|\sigma) = \begin{cases} \operatorname{tr}(\rho\log\rho - \rho\log\sigma) & \text{if } \operatorname{supp}\rho \subseteq \operatorname{supp}\sigma \\ \infty & \text{else.} \end{cases}$$

$\boxed{\text{Quantum Stein's lemma (Hiai/Petz, Ogawa/Nagaoka)}}$ $\forall \varepsilon > 0$,

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

Intuitively: $\beta_n^*(\varepsilon) \approx \exp(-n D(\rho\|\sigma))$

$\to$ measure of distinguishability in the asymmetric setting.

The larger $D(\rho\|\sigma)$, the better I can distinguish ρ and σ

(decay of β_n^*, optimal type-II error,

if type-I error is bounded/small)

Asymmetric hypothesis testing setting:

$D(\rho\|\sigma) \neq D(\sigma\|\rho)$ $\Rightarrow$ not a metric in the math. sense

But: If ρ, σ are states, then $D(\rho\|\sigma) \geq 0$ and $= 0$ iff $\rho = \sigma$.