Recap

- trace distance between quantum states ρ, σ : $\frac{1}{2}\|\rho - \sigma\|_1$
- Variational expression: $\frac{1}{2}\|\rho - \sigma\|_1 = \max_{0 \leq \Lambda \leq \mathbb{1}} \operatorname{tr} \Lambda(\rho - \sigma)$
- Optimal measurement: $\Pi = \{\rho - \sigma \geq 0\}$ $\left(X = \sum_i \lambda_i |i\rangle\langle i| \Rightarrow \{X \geq 0\} = \sum_{i : \lambda_i \geq 0} |i\rangle\langle i| \right)$
- Data-processing: $\frac{1}{2}\|\rho - \sigma\|_1 \geq \frac{1}{2}\|\mathcal{N}(\rho) - \mathcal{N}(\sigma)\|_1$ for quantum channel $\mathcal{N}$ (noise makes states <u>less</u> distinguishable)
- Binary state discrimination: two fundamental errors

  type-I error: $\alpha = \Pr(\sigma | \rho) = \operatorname{tr}(\mathbb{1} - \Lambda)\rho$

  type-II error: $\beta = \Pr(\rho | \sigma) = \operatorname{tr} \Lambda \sigma$

  measurement:

  $\Lambda \longleftrightarrow \rho$

  $\mathbb{1} - \Lambda \longleftrightarrow \sigma$

- Symmetric hypothesis testing: $\min \alpha + \beta$ $\longrightarrow$ trace distance
- Asymmetric hypothesis testing: keep α bounded, minimize β:

  fix $\varepsilon > 0$: $\beta^*(\varepsilon) = \min\{\operatorname{tr} \Lambda\sigma : 0 \leq \Lambda \leq \mathbb{1}, \operatorname{tr}(\mathbb{1} - \Lambda)\rho \leq \varepsilon\}$

- How does $\beta^*(\varepsilon)$ behave in the asymptotic setting $\rho \to \rho^{\otimes n}, \sigma \to \sigma^{\otimes n}$?
- Relative entropy: $D(\rho\|\sigma) = \begin{cases} \operatorname{tr} \rho(\log \rho - \log \sigma) & \operatorname{supp} \rho \subseteq \operatorname{supp} \sigma \\ \infty & \text{else} \end{cases}$
- Quantum Stein's lemma: $\lim_{n \to \infty} \frac{1}{n} \log \beta_n^*(\varepsilon) = -D(\rho\|\sigma)$

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

§1.3 Properties of relative entropy

Recall: measures of distinguishability (between quantum states) should be monotonic under quantum channels

("noise cannot make two states more dist.)

→ data processing inequality

Thm 5 Let ρ be a quantum state, $\sigma \geq 0$, and $\mathcal{N}$ a quantum channel:

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))$$

Data-processing inequality for relative entropy.

We defer the proof to later, and first discuss some properties of $D(\cdot\|\cdot)$

Prop 6 i) Let ρ be a classical state, and $\sigma \geq 0$ be classical.

$\rho = \sum_x p_x |x\rangle\langle x|$, $\sigma = \sum_x q_x |x\rangle\langle x|$:

$$D(\rho\|\sigma) = \sum_x p_x \log \frac{p_x}{q_x} = D(p\|q)$$

Kullback-Leibler divergence

ii) Let ρ, σ be quantum states, then $D(\rho\|\sigma) \geq 0$,

and $D(\rho\|\sigma) = 0$ iff $\rho = \sigma$.

iii) $D(\rho\|\sigma) = D(V\rho V^\dagger \| V\sigma V^\dagger)$ for V an isometry $(V^\dagger V = \mathbb{1})$

iv) classical-quantum states $\rho_{XA} = \sum_x p_x |x\rangle\langle x|_X \otimes \rho_A^x$

$\sigma_{XA} = \sum_x p_x |x\rangle\langle x|_X \otimes \sigma_A^x$ $\quad$ (ρ_A^x states, $\sigma_A^x \geq 0$):

$$D(\rho_{XA} \| \sigma_{XA}) = \sum_x p_x \, D(\rho_A^x \| \sigma_A^x)$$

v) Joint convexity: Let $\{\rho_x\}$ be states, $\{\sigma_x\}$ with $\sigma_x \geq 0$, and $\{\lambda_x\}$ a prob. dist.

Then, $\quad D\left(\sum_x \lambda_x \rho_x \| \sum_x \lambda_x \sigma_x\right) \leq \sum_x \lambda_x \, D(\rho_x \| \sigma_x)$

vi) Let ρ be a state, $\sigma, \sigma' \geq 0$ with $\sigma \leq \sigma'$, then

$$D(\rho\|\sigma) \geq D(\rho\|\sigma').$$

<u>Proof</u>: i) by definition of $D(\cdot\|\cdot)$

ii) ρ, σ states: use DPI w.r.t. $\mathcal{N} = \mathrm{tr}$

$$D(\rho\|\sigma) \geq D(\mathrm{tr}(\rho) \| \mathrm{tr}(\sigma))$$

$$= D(1\|1) = 0 \quad \rightarrow \checkmark$$

$D(\rho\|\sigma) = 0$ iff $\rho = \sigma$ will be proved later.

iii) to prove: $D(\rho\|\sigma) = D(V\rho V^\dagger \| V\sigma V^\dagger)$ for V with $V^\dagger V = \mathbb{1}$.

there is an easy proof for $D(\cdot\|\cdot)$, but the following generalizes to other quantities:

1) $V\cdot V^\dagger$: $D(\rho\|\sigma) \geq D(V\rho V^\dagger \| V\sigma V^\dagger)$ by DPI

$V^\dagger \cdot V$ is CP but not trace-preserving, since $\operatorname{Im} V \subsetneq K$

Solution: $\Pi = VV^\dagger$ projection onto $\operatorname{Im} V$

define $W: \mathcal{B}(K) \to \mathcal{B}(\mathcal{H})$ $\qquad |0\rangle \in \mathcal{H}$

$$W(X) = V^\dagger X V + \operatorname{tr}\left((\mathbb{1}-\Pi)X\right)|0\rangle\langle 0|$$

$$\operatorname{tr} W(X) = \underbrace{\operatorname{tr} V^\dagger X V}_{\operatorname{tr} VV^\dagger X = \operatorname{tr} \Pi X} + \operatorname{tr}((\mathbb{1}-\Pi)X) = \operatorname{tr} X.$$

$$W(VYV^\dagger) = Y$$

2) $D(\rho\|\sigma) \geq D(V\rho V^\dagger \| V\sigma V^\dagger) \geq D(W(V\rho V^\dagger) \| W(V\sigma V^\dagger))$

$= D(\rho\|\sigma) \quad \to$ iii)

iv) cq-states: $\rho_{XA} = \sum_x p_x |x\rangle\langle x|_X \otimes \rho_A^x$

$$= \bigoplus_x p_x \rho_A^x = \begin{pmatrix} \boxed{p_1 \rho_A^1} & 0 & \\ 0 & \boxed{p_2 \rho_A^2} & 0 \\ & 0 & \ddots \end{pmatrix}$$

$D(\rho\|\sigma) = \operatorname{tr} \rho (\log \rho - \log \sigma)$

$$\log \rho_{XA} = \begin{pmatrix} \log(p_1 \rho_A^1) & & \\ & \log(p_2 \rho_A^2) & \\ & & \ddots \end{pmatrix}$$

$$\rho_{XA} = \begin{pmatrix} p_1\rho_A^1 & & \\ & p_2\rho_A^2 & \\ & & \ddots \end{pmatrix} \Rightarrow \log\rho_{XA} = \begin{pmatrix} \log(p_1\rho_A^1) & & \\ & \log(p_2\rho_A^2) & \\ & & \ddots \end{pmatrix}$$

$$\log\rho_{XA} - \log\sigma_{XA} = \sum_x |x\rangle\langle x|_X \otimes \left(\log\rho_A^x - \log\sigma_A^x\right)$$

$$\sigma_{XA} = \sum_x p_x |x\rangle\langle x| \otimes \sigma_A^x$$

$$D(\rho_{XA}\|\sigma_{XA}) = \mathrm{tr}\,\rho_{XA}\left(\log\rho_{XA} - \log\sigma_{XA}\right)$$

$$= \mathrm{tr}\left(\left(\sum_x p_x \underline{|x\rangle\langle x|} \otimes \rho_A^x\right)\left(\sum_y \underline{|y\rangle\langle y|} \otimes \left(\log\rho_A^y - \log\sigma_A^y\right)\right)\right)$$

$$= \sum_x p_x \,\mathrm{tr}\left(|x\rangle\langle x| \otimes \rho_A^x\left(\log\rho_A^x - \log\sigma_A^x\right)\right)$$

$$= \sum_x p_x D(\rho_A^x\|\sigma_A^x). \quad \to \checkmark$$

v) to show: $D\left(\sum_x \lambda_x \rho_x \,\middle\|\, \sum_x \lambda_x \sigma_x\right) \le \sum_x p_x D(\rho_x\|\sigma_x)$

use iv) with $\rho_{XA} = \sum_x \lambda_x |x\rangle\langle x| \otimes \rho_A^x$, $\sigma_{XA} = \sum_x \lambda_x |x\rangle\langle x| \otimes \sigma_A^x$

$$\sum_x \lambda_x D(\rho_A^x\|\sigma_A^x) \overset{(iv)}{=} D(\rho_{XA}\|\sigma_{XA})$$

$$\underset{\text{wrt. } \mathrm{tr}_X}{\overset{DPI}{\geq}} D(\rho_A\|\sigma_A) = D\left(\sum_x \lambda_x \rho_A^x \,\middle\|\, \sum_x \lambda_x \sigma_A^x\right).$$

vi) If $\sigma \le \sigma'$, then $D(\rho\|\sigma) \ge D(\rho\|\sigma')$.

using (iv): $D(\rho\|\sigma) = D\left(\rho\otimes|0\rangle\langle 0| \,\middle\|\, \sigma\otimes|0\rangle\langle 0| + \overbrace{(\sigma'-\sigma)}^{\ge 0}\otimes|1\rangle\langle 1|\right)$

$$D(\rho\|\sigma) = D(\rho \otimes |0\rangle\langle 0| \,\|\, \sigma \otimes |0\rangle\langle 0| + \underbrace{(\sigma' - \sigma)}_{\geq 0} \otimes |1\rangle\langle 1|)$$

$$D(\rho\|\sigma) \overset{DPI}{\geq} D(\rho \| \sigma + \sigma' - \sigma) = D(\rho\|\sigma').$$ $\square$

Next goal: prove Thm 5 (data-processing)

Detour 1: Functions on operators and operator convexity

·) Let $A \in B(\mathcal{H})$ be Hermitian with spectral decomposition

$$A = \sum_i \lambda_i |i\rangle\langle i| \quad (\lambda_i \in \mathbb{R}, \langle i|j\rangle = \delta_{ij})$$

·) Let $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$, be such that $\underline{\operatorname{spec} A \subseteq I}$, then we set

$$f(A) = \sum_i f(\lambda_i) |i\rangle\langle i|$$

In words, $f(A)$ is a Hermitian op. with the same eigenbasis as A and spectrum $\{f(\lambda_i)\}_i$. In particular, $[A, f(A)] = 0$

·) Example: matrix logarithm $\leftarrow \log : \mathbb{R}^+ \to \mathbb{R}$

$$\rho = \sum_i \lambda_i |i\rangle\langle i| \Rightarrow \log\rho = \sum_i \log\lambda_i |i\rangle\langle i|$$

·) Examples: $\eta(t) = t\log t$: $\lim_{t\to 0} \eta(t) = 0 \Rightarrow 0\log 0 \equiv 0$

$\rho \geq 0 \Rightarrow \rho\log\rho$ (remember: $S(\rho) = -\operatorname{tr}\rho\log\rho$)

·) Let $V : \mathcal{H} \to \mathcal{K}$ be an isometry, $V^\dagger V = \mathbb{1}$, then

$$\underline{\underline{f(VAV^\dagger) = V f(A) V^\dagger}}$$