

## Ch. 9

Trace Norm:

Define "trace norm" or "nuclear norm" ( $\ell_1$ -norm on singular values)

$$\|M\|_1 \equiv \text{Tr}(\sqrt{M^* M}) = \text{sum of singular values}$$

We care about Hermitian  $M$ .

$$\Rightarrow \|M\|_1 = \sum_i |\lambda_i|$$

Notice than  $\|\rho\|_1 = 1$  for density operator.

This satisfies properties of a norm:

$$1.) \|M\|_1 \geq 0 \quad \leftarrow \text{Positive Definite}$$

$$2.) \|M\|_1 = 0 \Leftrightarrow M=0$$

$$3.) \|cM\|_1 = |c| \|M\|_1 \quad \text{Homogeneous}$$

$$4.) \|M+N\|_1 \leq \|M\|_1 + \|N\|_1 \quad \text{Triangle Inequality}$$

$$\Rightarrow \text{Convex: } \|\lambda_1 M + \lambda_2 N\|_1 \leq \lambda_1 \|M\|_1 + \lambda_2 \|N\|_1 \text{ where } (\lambda_1, \lambda_2) \text{ is pmf.}$$

$$\text{Isometric Invariance: } \|U M U^\dagger\|_1 = \|M\|_1 \text{ if } U \text{ is isometry.}$$

Trace Distance:  $\|M-N\|_1$ 

$$\text{For density operators: } 0 \leq \|\rho - \sigma\|_1 \leq 2$$

↑                      ↑  
positivity            triangle inequality

Intuitively, this measures how different  $\rho$  and  $\sigma$  are.

Precisely: Largest probability difference between two measurements.

$$\underline{\text{Lemma 7.1.7: }} \|\rho - \sigma\|_1 = 2 \max_{0 \leq \lambda \leq 1} \text{Tr}(\lambda(\rho - \sigma))$$

↑                      ↑  
measurement             $\text{Tr}(\lambda\rho) - \text{Tr}(\lambda\sigma)$

Proof: Let's prove that  $\|w\|_1 = \max_{0 \leq \lambda \leq 1} \text{Tr}(\lambda w)$

Proof: Let's prove that  $\|w\|_1 = \max_{\substack{\lambda \in \Lambda \\ \text{Hermitian}}} \text{Tr}(\lambda w)$

Let  $w^+ \geq 0$  and  $w^- \geq 0$  s.t.  $w^+ - w^- = w$ .

That is,  $w^+$  takes the pos. eigenvalues  
 $w^-$  takes the neg. eigenvalues

$$\begin{aligned} \Rightarrow \text{Tr}(\lambda w) &= \text{Tr}(\lambda w^+) - \text{Tr}(\lambda w^-) \\ &\leq \text{Tr}(w^+) + \text{Tr}(w^-) = \|w\|_1 \end{aligned}$$

Choose  $\Lambda = \sum_i \text{sign}(\lambda_i) |v_i\rangle\langle v_i|$  to show equality.

For the Lemma: Notice that  $\text{Tr}(\rho - \sigma) = 0 \Rightarrow \text{Tr}[\rho - \sigma]_+ = \text{Tr}[\rho - \sigma]_-$

Factor of 2 comes from this.

Intuition: If trace distance is small, indistinguishable via measurement.

Extension: (Easy exercise)

For a POVM  $\{\Lambda_x\}$ , let  $p(x) = \text{Tr}(\Lambda_x \rho)$ ,  $q(x) = \text{Tr}(\Lambda_x \sigma)$

Then  $\|\rho - \sigma\|_1 = \max_{\{\Lambda_x\}} \|p - q\|_1$

↳ 1-1 distance of pmf.

Classical Counterpart:

Let  $P$  and  $Q$  be distributions: Total Variation Distance =  $\max_{A \in \mathcal{F}} (P(A) - Q(A))$

Property 1:  $d_{\text{TV}}(P, Q) = \frac{1}{2} \|\rho - q\|_1$  if  $p$  and  $q$  are pmf for  $P$  and  $Q$ .

Total variation is used in security (such as cryptography) to claim that two dist. are not distinguishable.

Bayesian estimation: Suppose  $X = \begin{cases} 0 & \text{w.p. } p_0 \\ 1 & \text{w.p. } p_1 \end{cases}$

$$\begin{aligned} Y &\sim P && \text{if } X=0 \\ Y &\sim Q && \text{if } X=1 \end{aligned}$$

Let  $\hat{X}$  be a function of  $Y$ .

Claim:  $P(\hat{X} \neq X) \geq \frac{1}{2} - \frac{1}{2} \|p_0 p - p_1 q\|_1$   
 Equality for best estimator.

Proof: Define  $A \subset Y$  such that  $\hat{X}(y) = 0$  iff  $y \in A$  (preimage)

$$\begin{aligned} P(\hat{X} \neq X) &= E P(\hat{X} \neq X | X) \\ &= p_0 P(\hat{X} \neq X | X=0) + p_1 P(\hat{X} \neq X | X=1) \\ &= p_0 P(A^c) + p_1 Q(A) \\ &= p_0 - \sum_{y \in A} (p_0 p(y) - p_1 q(y)) \end{aligned}$$

Notice:  $\sum_y p_0 p(y) - p_1 q(y) = p_0 - p_1$

$$\max_A \sum_{y \in A} = \sum_y [p_0 p(y) - p_1 q(y)]_+$$

$$\|p_0 p - p_1 q\|_1 = \sum_y [p_0 p(y) - p_1 q(y)]_+ + \sum_y [p_0 p(y) - p_1 q(y)]_-$$

$$2 \sum_y [\dots]_+ = p_0 - p_1 + \|p_0 p - p_1 q\|_1$$

$$\Rightarrow \max_A \sum_{y \in A} p_0 p(y) - p_1 q(y) = \frac{1}{2} (p_0 - p_1 + \|p_0 p - p_1 q\|_1)$$

$$\Rightarrow P(\hat{X} \neq X) \geq p_0 - \frac{1}{2} (p_0 - p_1 + \|p_0 p - p_1 q\|_1)$$

Special case  $p_0 = p_1 = \frac{1}{2}$ :

$$P(\hat{X} = X) \geq \frac{1}{2} - \frac{1}{4} \|p - q\|_1 = \frac{1}{2} (1 - d_{TV}(P, Q))$$

Quantum Equivalent: Let  $\{\Lambda_0, \Lambda_1\}$  be the optimal measurement to

distinguish  $\rho$  and  $\sigma$  with  $p_0$  and  $p_1$ .

$$\text{Prob. of error: } p_e = \frac{1}{2} - \frac{1}{2} \|\rho_{00} - \sigma_{00}\|_1$$

Property 2: Let  $f$  be bounded with width  $b$ .  $f: X \rightarrow [c - \frac{b}{2}, c + \frac{b}{2}]$

$$|E_p f(x) - E_\sigma f(x)| \leq b d_{\text{TV}}(\rho, \sigma) = \frac{b}{2} \|\rho - \sigma\|_1$$

This generalized Prop. 1. Let  $f$  be an indicator function.

Quantum: Restate Lemma 9.1.7 with  $\Lambda \leq bI$

Corollary 9.1.9  $\|\rho - \sigma\|_1 \geq \frac{2}{b} (\text{Tr}(\Lambda\rho) - \text{Tr}(\Lambda\sigma))$

$$\Rightarrow |\langle \Lambda \rangle_\rho - \langle \Lambda \rangle_\sigma| \leq \frac{b}{2} \|\rho - \sigma\|_1$$

Used in information theory proofs.

$\rho$  = actual distribution induced by the processing

$\sigma$  = simple distribution to analyze.

Property 3: (Triangle Inequality):  $\|\rho - \sigma\|_1 \leq \|\rho - \tau\|_1 + \|\tau - \sigma\|_1$

Lemma 9.1.8 Same for quantum.

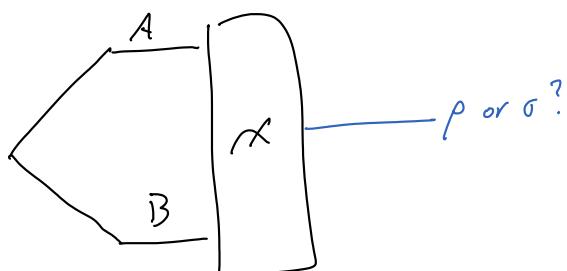
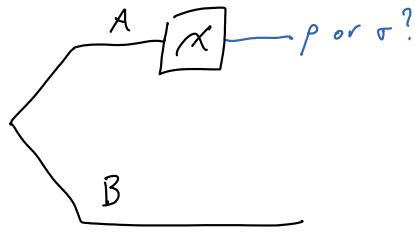
Property 4: (Monotonicity of Marginal)

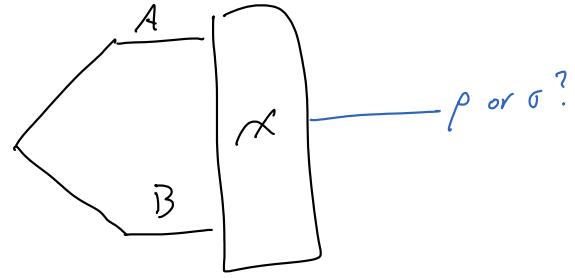
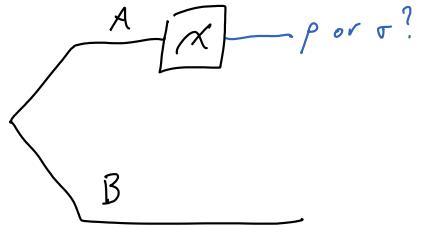
$$\|\rho(x) - \sigma(x)\|_1 \leq \|\rho(x,y) - \sigma(x,y)\|_1$$

Quantum Corollary 9.1.10:  $\|\rho^A - \sigma^A\|_1 \leq \|\rho^{AB} - \sigma^{AB}\|_1$

Proof: Use Lemma 9.1.7.

More measurements are possible on the joint system.





Property 5: (Monotonicity of Noisy Map)

$$\text{Let } p(y) = \sum_x p(x)p(y|x) \text{ and } q(y) = \sum_x q(x)p(y|x)$$

same channel

$$\text{then } \|p(y) - q(y)\|_1 \leq \|p(x) - q(x)\|_1$$

Proof:

- Show that  $\|\rho(x)p(y|x) - q(x)p(y|x)\|_1 = \|\rho(x) - q(x)\|_1$
- Use Property 4.

$$\text{Quantum: } \|N(\rho) - N(\sigma)\|_1 \leq \|\rho - \sigma\|_1$$

Proof: Consider isometric extension  $\tilde{\rho}^{AB} = U^{A \rightarrow AB} \rho$   
 $\tilde{\sigma}^{AB} = U^{A \rightarrow AB} \sigma$

$$\|\tilde{\rho}^{AB} - \tilde{\sigma}^{AB}\|_1 = \|\rho^A - \sigma^A\|_1 \text{ by isometric invariance}$$

$$\Rightarrow \|N(\rho) - N(\sigma)\|_1 = \|\tilde{\rho}^{AB} - \tilde{\sigma}^{AB}\|_1 \leq \|\rho - \sigma\|_1 \text{ by Prop. 4.}$$

Fidelity:

$$\text{Pure states: } F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2 = \langle\psi|\phi\rangle\langle\phi|\psi\rangle = \text{Tr}(|\psi\rangle\langle\phi|\phi\rangle)$$

Interpret as probability the  $|\phi\rangle$  gets measured as  $|\psi\rangle$  use  $\{|\psi\rangle\langle\psi|, I - |\psi\rangle\langle\psi|\}$

This is not a distance metric, but indicate similarity.

$$\text{If } |\psi\rangle = |\phi\rangle \text{ then } F(|\psi\rangle, |\phi\rangle) = 1$$

In general,  $F \in [0, 1]$

Let the second argument be a density:

This is tempting but will not be the general definition.

$$\begin{aligned} F(|\psi\rangle, \rho) &= E_x F(|\psi\rangle, |\varphi_x\rangle) \\ &= \sum_x p(x) |\langle \psi | \varphi_x \rangle| = \sum_x p(x) \langle \psi | \varphi_x \rangle \langle \varphi_x | \psi \rangle = \langle \psi | \rho | \psi \rangle \end{aligned}$$

Def :  $F(|\psi\rangle, \rho) = \langle \psi | \rho | \psi \rangle = \text{Tr}(|\psi\rangle \langle \psi | \rho)$

Again, this is a coincidence.