Lecture 14

Wednesday, November 05, 2014
3:29 PM

Classical-Quantum Metrics:

Shannon entropy of POVM: Let $p(x) = \mathrm{Tr}(\Lambda_x \rho)$

Entropy of POVM is $H(X)$.

Accessible Information of Ensemble:

Given $\mathcal{E} = \{(p(x), \rho_x)\}$

$$I_{acc}(\mathcal{E}) = \max_{\{\Lambda_y\}} I(X;Y)$$

Note, the specific ensemble is important. Not just a function of density.

Classical Mutual Information:

$$I_c(\rho^{AB}) = \max_{\{\Lambda_x^A\}\{\Lambda_y^B\}} I(X;Y) \qquad \text{where } p(x,y) = \mathrm{Tr}\left((\Lambda_x^A \otimes \Lambda_y^B)\rho^{AB}\right)$$

Always maximized by rank 1 POVM.

How do we define quantum conditional entropy?

One idea: Condition on measurement outcome on A.

$$H(B|A)_\pi = \sum_x p(x)\, H(\rho_x^B)$$

x indicates which measurement outcome was obtained by measuring A.

Prob. of measurement.

Which measurement should be used? The best one?

$$\underset{\pi}{\text{minimize}}\; H(B|A)_\pi$$

Instead, let $\boxed{H(B|A) = H(A,B) - H(A)}$

In the classical case, both definitions are the same.

In fact, in a classical-quantum system, all quantities are similar to classical quantities.

$$\rho^{XA} = \sum_x p(x)\, |x\rangle\langle x|^X \otimes \rho_x^A$$

Joint Entropy:

$$H(A,B) = -\mathrm{Tr}\left(\rho^{AB} \log \rho^{AB}\right)$$

Something funny happens: $H(A,B) \not\geq H(A)$

Why? Think of von Neumann entropy as the minimum Shannon entropy over measurements.

Consider an entangled pure state: $H(A,B) = 0$

$H(A) > 0 \leftarrow$ Why?

1.) In fact, $H(A) = H(B)$ for a pure state.

Proof: Schmidt Decomposition

$$|\psi\rangle^{AB} = \sum_i \sqrt{\lambda_i}\, |i\rangle^A |i\rangle^B$$

$$\Rightarrow \rho^A = \sum_i \lambda_i |i\rangle\langle i|^A \qquad \rho^B = \sum_i \lambda_i |i\rangle\langle i|^B$$

$$H(A) = H(\{\lambda_i\}) = H(B)$$

Consider ρ^A with purification $|\psi\rangle^{RA}$: $H(R) = H(A)$

In fact, they have the same spectrum.

Corollary to 1) For a pure state on (A,B,C)

$$H(A) = H(B,C)$$
$$H(B) = H(A,C)$$
$$H(C) = H(A,B)$$

Joint entropy of product state is additive:

$$H(\rho \otimes \sigma) = H(\rho) + H(\sigma)$$

Classical-Quantum State:

$$H(X,A) = \underbrace{H(X)}_{H(p(x))} + \underbrace{\sum_x p(x)\, H(\rho_x^A)}_{\text{Conditional entropy}}$$

<u>Proof:</u> Consider the spectral decomposition $\rho_x^A = \sum_a p(a|x)\, |\varphi_{x,a}\rangle\langle\varphi_{x,a}|^A$

First calculate $H(X)$ and $H(\rho_x^A)$

Notice: $\rho^X = \sum_x p(x)\, |x\rangle\langle x|^X$

$$\Rightarrow H(X) = \sum_x p(x) \log \frac{1}{p(x)}$$

Also, $H(\rho_x^A) = \sum_a p(a|x) \log \frac{1}{p(a|x)}$

$$\rho^{XA} = \sum_{x,a} p(x)\, p(a|x)\; |x\rangle\langle x|^X \otimes |\varphi_{x,a}\rangle\langle\varphi_{x,a}|^A$$

Notice that $|x\rangle^X |\varphi_{x,a}\rangle^A$ is an orthonormal basis

$$\Rightarrow H(\rho^{XA}) = H\big(p(x)\, p(a|x)\big)$$

$$= H(X) + \sum_x p(x)\, H(\rho_x^A)$$

Back to Conditional Entropy:

Define $H(A|B) \triangleq H(A,B) - H(B)$

We've already seen that this can be negative.

<u>Operational significance:</u>

Classical result: $H(X|Y)$ is minimum lossless compression rate for X if Y is known to the decoder,

Quantum result: $H(A|B)$ is minimum quantum compression rate for transferring A's state if B is available to the decoder, <u>assuming unlimited classical communication.</u>

A negative value means you end up with a net gain in quantum communication.

Specifically, you end up with entanglement. Recall $[qq] = [q \to q]$ with unlimit $[c \to c]$

Simple example: $|\Phi^+\rangle^{AB}$ is an ebit (Bell state)

$$H(A|B) = -1 \text{ bits}$$

Because $|\Phi^+\rangle^{AB}$ is a known pure state,
the decoder can produce a copy without communication.
Furthermore, $|\Phi^+\rangle^{AB}$ can be used for one $[q \to q]$ in the future (teleportation).

Coherent Information:

$$I(A \rangle B) \triangleq -H(A|B)$$

To be defined.

$$\Rightarrow I(A \rangle B) = H(B) - H(A,B) = I(A;B) - H(A)$$

Bounds on conditional entropy.

$$H(A|B) \leq H(A) \quad \leftarrow \text{ Will use relative entropy to show this.}$$

$$H(A|B) \geq -H(A) \iff I(A \rangle B) \leq H(A)$$

$$I(A \rangle B) = H(A|E) \leq H(A)$$

purification

Shown in problem set.

$$\boxed{|H(A|B)| \leq H(A)}$$

Quantum Mutual Information:

$$I(A;B)_\rho = H(A)_\rho + H(B)_\rho - H(A,B)_\rho$$

Quantum

Conditional Mutual Information:

$$I(A;B|C)_\rho = H(A|C)_\rho + H(B|C)_\rho - H(A,B|C)_\rho$$

All classical chain rules still apply (by definition)

$$\boxed{\text{Th}^{\underline{m}} \qquad I(A;B|C) \geq 0}$$

← Much of information theory rests on this.

Corollary: $I(A;B) \geq 0$ ← Already claimed above (i.e. $H(A|B) \leq H(A)$)

Quantum Relative Entropy:

$$D(\rho\|\sigma) \triangleq \mathrm{Tr}\left(\rho \log \rho - \rho \log \sigma\right)$$

Properties:

1.) $D(\rho\|\sigma) \geq 0$

2.) If $\mathrm{supp}(\rho) \cap \mathrm{supp}(\sigma)^\perp \neq \{0\}$

$D(\rho\|\sigma) = \infty$

Proof of 1:

Let $\rho = \sum_x p(x) |\varphi_x\rangle\langle\varphi_x|$

$\sigma = \sum_{x'} q(x') |\psi_{x'}\rangle\langle\psi_{x'}|$

$$D(\rho\|\sigma) = \mathrm{Tr}\left(\rho \log \rho - \rho \log \sigma\right)$$

$$= \sum_x \langle\varphi_x| \left(\rho \log \rho - \rho \log \sigma\right) |\varphi_x\rangle$$

$$= \sum_x p(x) \log p(x) - \sum_x p(x) \sum_{x'} |\langle\varphi_{x'}|\psi_x\rangle|^2 \log q(x')$$

This is because the bases are different.

Notice that $|\langle\varphi_{x'}|\psi_x\rangle|^2$ is doubly-stochastic.

That is $\sum_{x'} |\langle\varphi_{x'}|\psi_x\rangle|^2 = 1 \quad \forall x$

and $\sum_{x} |\langle\varphi_{x'}|\psi_x\rangle|^2 = 1 \quad \forall x'$

$\Rightarrow p(x'|x) = |\langle\varphi_{x'}|\psi_x\rangle|^2$ is a valid conditional distribution.

and $r(x') = \sum_{x'} q(x') |\langle\varphi_{x'}|\psi_x\rangle|^2$ is a valid pdf

Opposite interpretations

$$\Rightarrow -\sum_x p(x) \sum_{x'} |\langle\varphi_{x'}|\psi_x\rangle|^2 \log q(x') = -\sum_x p(x) \underbrace{\sum_{x'} p(x'|x)}_{\text{Expected value}} \log q(x')$$

$$\geq -\sum_x p(x) \log\left(\sum_{x'} p(x'|x) \log q(x')\right) \qquad \text{Jensen}$$

$$= -\sum_x p(x) \log\left(\sum_{x'} |\langle\varphi_{x'}|\psi_x\rangle|^2 q(x')\right)$$

$$= -\sum_x p(x) \log r(x)$$

$$\Rightarrow D(\rho\|\sigma) \geq \sum_x p(x) \log p(x) - \sum_x p(x) \log r(x)$$

$$= \sum_x p(x) \log \frac{p(x)}{r(x)} = D(p(x)\|r(x)) \geq 0 \qquad \square$$

<u>Consequences:</u> $I(A;B) = D(\rho^{AB}\|\rho^A \otimes \rho^B)$ ← Show in problem set.

$\Rightarrow I(A;B) \geq 0$

<u>Monotonicity of Relative Entropy:</u> (Data-Processing Ineq.)

$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))$ ← We encountered this for trace distance.

<u>Proof</u>: Trick is to reduce to special case: $D(\rho^{AB}\|\sigma^{AB}) \geq D(\rho^A\|\sigma^A)$ Assumption 1.

Believable but tedious to prove. (Appendix)

$D(\rho\|\sigma) \geq D(\rho \otimes |0\rangle\langle 0|^E \| \sigma \otimes |0\rangle\langle 0|^E)$ ←

$$D(\rho\|\sigma) \geq D\left(\rho \otimes |0\rangle\langle 0|^E \,\|\, \sigma \otimes |0\rangle\langle 0|^E\right)$$

$$\boxed{D(\rho_1^A \otimes \rho_2^B \,\|\, \sigma_1^A \otimes \sigma_2^B) = D(\rho_1\|\sigma_1) + D(\rho_2\|\sigma_2)}$$

$$= D\left(U(\rho \otimes |0\rangle\langle 0|^E)U^\dagger \,\|\, U(\rho \otimes |0\rangle\langle 0|^E)U^\dagger\right)$$

Properties

$$\boxed{\text{Relative Entropy is unitarily invariant}}$$

$$\geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \quad \text{under assumption 1.}$$

<u>Consequences:</u>

$$\underbrace{D(\rho^{ABC}\|\rho^A \otimes \rho^{BC})}_{= I(A;BC)} \geq \underbrace{D(\rho^{AB}\|\rho^A \otimes \rho^B)}_{= I(A;B)}$$

$$\Rightarrow I(A;C|B) \geq 0$$

Quantum Data Processing Ineq.:

Let $\rho^{AB'} = (I^A \otimes \mathcal{N}^{B\to B'})(\rho^{AB})$

$$I(A;B) \geq I(A;B') \qquad \left(\text{Equivalently: } I(A\rangle B) \geq I(A\rangle B')\right)$$