Lecture 19

quantum entropies & information

begin w/ von Neumann entropy & related quantities

(AKA quantum entropy)

$$H(\rho_A) = H(A)_\rho = -\mathrm{Tr}[\rho_A \log_2 \rho_A]$$

for $\rho_A = \sum_{i=1}^{r} \lambda_i |\psi_i\rangle\langle\psi_i|$

$$H(\rho_A) = -\sum_{i=1}^{r} \lambda_i \log_2 \lambda_i$$

use convention that $0 \log_2 0 = 0$

b/c $\lim_{x \to 0^+} x \log x = 0$

related quantities include

conditional entropy $\quad H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$

for ρ_{AB}

coherent information:

$$I(A\rangle B) = H(B)_\rho - H(AB)_\rho = -H(A|B)_\rho$$

mutual information:

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

conditional mutual information of a tripartite state ρ_{ABC}:

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

All of these can be derived from q. relative entropy:

$$D(\rho\|\sigma) = \begin{cases} \mathrm{Tr}[\rho(\log_2\rho - \log_2\sigma)] & \text{if } \mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma) \\ +\infty & \text{else} \end{cases}$$

for ρ a density operator + σ PSD

$$D(\rho\|\sigma) = \lim_{\varepsilon \to 0^+} \mathrm{Tr}\left[\rho\left(\log_2 \rho - \log_2(\sigma + \varepsilon I)\right)\right]$$

Proof: Decompose Hilbert space as $\mathcal{H} = \mathrm{supp}(\sigma) \oplus \ker(\sigma)$

Π_σ projects onto $\mathrm{supp}(\sigma)$ & $\Pi_\sigma^\perp$ onto $\ker(\sigma)$.

blocks according to Π_σ & $\Pi_\sigma^\perp$

Then $\rho = \begin{bmatrix} \rho_{0,0} & \rho_{0,1} \\ \rho_{0,1}^\dagger & \rho_{1,1} \end{bmatrix}$, $\sigma = \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix}$

Suppose that $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$.

Then $\rho_{0,0} = \rho$, & $\rho_{0,1} = 0$, $\rho_{1,1} = 0$

$$\mathrm{Tr}[\rho \log_2(\sigma + \varepsilon I)]$$

$$= \mathrm{Tr}\left[\begin{bmatrix} \rho_{0,0} & 0 \\ 0 & 0 \end{bmatrix} \log_2 \begin{bmatrix} \sigma + \varepsilon \Pi_\sigma & 0 \\ 0 & \varepsilon \Pi_\sigma^\perp \end{bmatrix}\right]$$

$$= \mathrm{Tr}\left[\quad '' \quad \begin{bmatrix} \log(\sigma + \varepsilon \Pi_\sigma) & \\ & \log \varepsilon \Pi_\sigma^\perp \end{bmatrix}\right]$$

$$= \mathrm{Tr}\left[\rho_{0,0} \log(\sigma + \varepsilon \Pi_\sigma)\right]$$

Taking $\lim_{\varepsilon \to 0^+}$ gives $\mathrm{Tr}[\rho_{0,0} \log \sigma]$
$= \mathrm{Tr}[\rho \log \sigma]$

If $\text{supp}(\rho) \not\subseteq \text{supp}(\sigma)$

then $\rho_{1,1}$ is non-zero

($\rho_{0,1}$ could be also)

& we find that

$$\text{Tr}\{\rho \log_2(\sigma + \varepsilon I)\}$$

$$= \text{Tr}\left[\begin{bmatrix} \rho_{0,0} & \rho_{0,1} \\ \rho_{0,1}^{+} & \rho_{1,1} \end{bmatrix} \begin{bmatrix} \log(\sigma + \varepsilon \Pi_\sigma) & 0 \\ 0 & \log \varepsilon \Pi_\sigma^{\perp} \end{bmatrix}\right]$$

$$= \text{Tr}\{\rho_{0,0} \log(\sigma + \varepsilon \Pi_\sigma)\}$$

$$+ \text{Tr}\{\rho_{1,1} \log_2 \varepsilon \Pi_\sigma^{\perp}\}$$

$$\downarrow$$

$$\log_2 \varepsilon \; \underbrace{\text{Tr}\{\rho_{1,1} \Pi_\sigma^{\perp}\}}_{> 0}$$

$$\Rightarrow \lim_{\varepsilon \to 0^+} (\;''\;) = -\infty$$

& relative entropy blows up.

For $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ (spectral decomp's)

$$\sigma = \sum_k q_k |\phi_k\rangle\langle\phi_k|$$

$$D(\rho\|\sigma) = \sum_{j,k} |\langle\psi_j|\phi_k\rangle|^2 \, p_j \log_2\left(\frac{p_j}{q_k}\right)$$

can define probability distributions

$$\lambda_{j,k} = p_j |\langle\psi_j|\phi_k\rangle|^2$$

$$\mu_{j,k} = q_k |\langle\psi_j|\phi_k\rangle|^2$$

& then

$$D(\rho\|\sigma) = D(\lambda\|\mu)$$

classical relative entropy

6

Properties of Q. relative entropy

1. Isometric invariance:

$$D(\rho\|\sigma) = D(V\rho V^\dagger \| V\sigma V^\dagger)$$

for V an isometry.

2. a) If $\mathrm{Tr}[\sigma] \leq 1$, then $D(\rho\|\sigma) \geq 0$

Sps. $\mathrm{Tr}[\sigma] \leq 1$. b) then $D(\rho\|\sigma) = 0$ iff $\rho = \sigma$

c) If $\rho \leq \sigma$, then $D(\rho\|\sigma) \leq 0$.

d) If $\sigma \leq \sigma'$, then

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

3. Additivity:

$$D(\rho_1 \otimes \rho_2 \| \sigma_1 \otimes \sigma_2) = D(\rho_1\|\sigma_1) + D(\rho_2\|\sigma_2)$$

A. Direct-sum prop.:

Let $\rho_{XA} = \sum_x p(x)\, |x\rangle\langle x| \otimes \rho_A^x$ &

$$\sigma_{XA} = \sum_x q(x)\, |x\rangle\langle x| \otimes \sigma_A^x$$

then

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

Proofs: Sps. we have data-processing:

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

then $D(\rho\|\sigma) = D(V\rho V^+ \| V\sigma V^+)$

follows from

$$D(\rho\|\sigma) \geq D(V\rho V^+ \| V\sigma V^+)$$

& $D(V\rho V^+ \| V\sigma V^+) \geq D((\mathcal{R}^{V} \circ \mathcal{V})(\rho) \| (\mathcal{R}^{V} \circ \mathcal{V})(\sigma))$

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

where $\mathcal{R}^V(\cdot) = V^\dagger(\cdot)V$

$+ \mathrm{Tr}[(I - VV^\dagger)(\cdot)]\tau$

for some state τ.

2a.) $D(\rho\|\sigma) \geq D(\mathrm{Tr}[\rho] \| \mathrm{Tr}[\sigma])$

$= 1 \log \frac{1}{\mathrm{Tr}[\sigma]} \geq 0$

by assumption

2b) If $\rho = \sigma$, then $D(\rho\|\sigma) = 0$

Sps $D(\rho\|\sigma) = 0$

Later we learn that

$D(\rho\|\sigma) \geq -\log_2 F(\rho, \sigma)$

So $D(\rho\|\sigma) = 0 \Rightarrow F(\rho,\sigma) = 1$

$\Rightarrow \rho = \sigma$ (under assumption that $\mathrm{Tr}[\sigma] \leq 1$)

2c) Suppose ρ & σ positive definite

then $\rho \le \sigma \Rightarrow \log \rho \le \log \sigma$

(log op. monotone)

$\Rightarrow \mathrm{Tr}[\rho(\log\rho - \log\sigma)] \le 0$

2d) again operator monotonicity of log

others by direct calculation

Joint convexity of relative entropy

$$\sum_x p(x)\, D(\rho^x \| \sigma^x) \ge D(\bar{\rho} \| \bar{\sigma})$$

where

$$\bar{\rho} = \sum_x p(x)\, \rho^x \quad \& \quad \bar{\sigma} = \sum_x p(x)\, \sigma^x$$

<u>Proof:</u> apply data processing under Tr_X to ρ_{XA} & σ_{XA}, as well as direct-sum property

---

Other entropies from relative entropy

$$H(\rho) = -D(\rho \| I)$$

$$H(A|B)_\rho = -D(\rho_{AB} \| I_A \otimes \rho_B)$$

$$= -\inf_{\sigma_B} D(\rho_{AB} \| I_A \otimes \sigma_B)$$

$$I(A\rangle B) = D(\rho_{AB} \| I_A \otimes \rho_B)$$

$$I(A;B) = D(\rho_{AB} \| \rho_A \otimes \rho_B)$$

can also minimize over either or both

Properties of entropy:

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

$$H(\rho) = H(V\rho V^{+})$$ for V an isometry

$$H(\bar{\rho}) \geq \sum_x p(x)\, H(\rho^x)$$ (follows from joint convexity of QRE)

$$H(\rho) \geq 0$$ follows from spectral decomp., concavity, & entropy of pure state is zero

$$H(\rho) \leq \log d$$

from concavity, mixing under Heisenberg-Weyl, & evaluating for max. mixed state

$$H\left(\sum_x p(x) |x\rangle\langle x| \otimes \rho^x\right) = H(p) + \sum_x p(x) H(\rho^x)$$

Chain rules

$$H(AB|C) = H(A|C) + H(B|AC)$$

(can pick C trivial)

$$I(A;BC|D) = I(A;B|D) + I(A;C|BD)$$

(can pick D trivial)

Strong subadditivity follows from data processing for relative entropy

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

$$I(A;B|C) = H(B|C) - H(B|AC)$$

$$= -D(\rho_{BC} \| I_B \otimes \rho_C)$$

$$+ D(\rho_{ABC} \| I_B \otimes \rho_{AC})$$

then Tr_A is the channel to use in data processing to get

$$D(\rho_{ABC} \| I_B \otimes \rho_{AC}) \geq D(\rho_{BC} \| I_B \otimes \rho_C)$$