Quantum

Lecture 6

- Shannon information
- Quantum information
- Distance measures

Shannon Entropy and Information

The Shannon entropy for a discrete variable X with alphabet $\mathcal{X}$ and pmf $p(x) = \Pr(X = x)$

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x)$$

average amount of uncertainty removed when observing the value of X = information gained when observing X

It holds that

$$0 \leq H(X) \leq \log |\mathcal{X}|$$

$= 0$ only if $p(x) = 1$ for some x

$= \log |\mathcal{X}|$ only if $p(x) = 1/|\mathcal{X}|$

Join entropy of $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, $p(x,y) = \Pr(X = x, Y = y)$

$$H(X,Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log p(x,y)$$

Conditional entropy of Y given $X = x$

$$H(Y|X = x) = -\sum_{y\in\mathcal{Y}} p(y|x)\log p(y|x)$$

Conditional entropy of Y given X

$$H(Y|X) = \sum_{x\in\mathcal{X}} p(x)H(Y|X = x)$$

Chain rule

$$H(X,Y) = H(Y|X) + H(X)$$

Relative entropy between the pmf's $p(\cdot)$ and $q(\cdot)$

$$D(p\|q) = \sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)}$$

$D(p\|q) \geq 0$ with $= 0$ only if $p(x) = q(x)$

Mutual information

$$I(X;Y) = D\big(p(x,y)\|p(x)p(y)\big) = \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$$

information about X obtained when observing Y (and vice versa)

$I(X;Y) \geq 0$ with $= 0$ only if $p(x,y) = p(x)p(y)$

Data processing inequality

$$X \to Y \to Z \implies I(X;Z) \le I(X;Y)$$

In particular,

$$I(X;f(Y)) \le I(X;Y)$$

$\Rightarrow$ no clever manipulation of the data can extract additional information that is not already present in the data itself

Quantum Entropy and Information

An ensemble $\{p_i, |\psi_i\rangle\}$, and with $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$

The quantum or Von Neumann entropy of ρ

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho) = -\sum_i \lambda_i \log \lambda_i$$

where $\{\lambda_i\}$ are the eigenvalues of ρ

$S(\rho) \ge 0$ with $= 0$ only if ρ is a pure state ($p_i = 1$ for some i)

In a d-dimensional space ($d \le \infty$)

$$S(\rho) \le d$$

with $= d$ only if $\{|\psi_i\rangle\}$ is an orthonormal set of size d and all p_i's are equal, i.e. a ρ is a completely mixed state

The (quantum) relative entropy between two states ρ and σ

$$S(\rho\|\sigma) = \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma)$$

$S(\rho\|\sigma) \geq 0$ with $= 0$ only if $\rho = \sigma$

For the composition of two systems $\mathcal{A}$ and $\mathcal{B}$ and a state $\rho^{AB} \in \mathcal{A} \otimes \mathcal{B}$, the joint entropy is $S(\rho^{AB})$

In the special case $\rho^{AB} = \rho \otimes \sigma$, we get

$$S(\rho^{AB}) = S(\rho) + S(\sigma)$$

c.f. $H(X, Y) = H(X) + H(Y)$ iff X and Y *independent*

In general, let $\rho_A = \mathrm{Tr}_{\mathcal{B}}\rho^{AB}$ and $\rho_B = \mathrm{Tr}_{\mathcal{A}}\rho^{AB}$

Conditional entropy

$$S(\rho_A|\rho_B) = S(\rho^{AB}) - S(\rho_B)$$

and mutual information

$$S(\rho_A; \rho_B) = S(\rho_A) + S(\rho_B) - S(\rho^{AB})$$

While $H(X|Y) \geq 0$, we have:
$S(\rho_B|\rho_A) < 0$ if (and only if) ρ^{AB} is entangled (has rank > 1)

It also holds that

$$S(\rho^{AB}) \leq S(\rho_A) + S(\rho_B)$$

with $=$ only if $\rho^{AB} = \rho_A \otimes \rho_B$. Furthermore

$$S(\rho^{AB}) \geq |S(\rho_A) - S(\rho_B)|$$

For three systems $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, we have

$$S(\rho_A) + S(\rho_B) \leq S(\rho^{AC}) + S(\rho^{BC})$$
$$S(\rho^{ABC}) + S(\rho_B) \leq S(\rho^{AB}) + S(\rho^{BC})$$

(where $\rho^{AB} = \mathrm{Tr}_{\mathcal{C}}\rho^{ABC}$, etc.)

Implications,

conditioning reduces entropy, $S(\rho_A|\rho^{BC}) \leq S(\rho_A|\rho_B)$
adding a system increases information, $S(\rho_A;\rho_B) \leq S(\rho_A;\rho^{BC})$

Quantum data processing inequality

For a composite system $\mathcal{A} \otimes \mathcal{B}$, if $\mathcal{E}$ is a trace-preserving quantum operation on $\mathcal{B}$, mapping ρ^{AB} to σ^{AB}, then

$$S(\rho_A;\rho_B) \geq S(\sigma_A;\sigma_B)$$

Tracing out subsystems decreases relative entropy

$$S(\rho^A\|\sigma^A) \leq S(\rho^{AB}\|\sigma^{AB})$$

Consider a discrete rv $X \in \mathcal{X}$ with pmf $p(x)$, and let $\{|e(x)\rangle\}$ be a basis for the $|\mathcal{X}|$-dimensional Hilbert space $\mathcal{H}$. Then we can "embed" the classical variable X in the quantum system $\mathcal{H}$ as

$$\sum_{x \in \mathcal{X}} p(x)|e(x)\rangle\langle e(x)|$$

Given a collection of $|\mathcal{X}|$ quantum states $\sigma(x)$, we can also define the mixed classical-quantum state

$$\sum_{x \in \mathcal{X}} p(x)|e(x)\rangle\langle e(x)| \otimes \sigma(x)$$

The joint (quantum) entropy of this classical-quantum state is

$$H(X) + \sum_{x \in \mathcal{X}} p(x) S(\sigma(x))$$

Classical Distance Measures

Two classical pmf's, $p(x)$ and $q(x)$ for a variable $x \in \mathcal{X}$

L_1 distance,

$$\|p(x) - q(x)\| = \sum_{x \in \mathcal{X}} |p(x) - q(x)|$$

For $A \subseteq \mathcal{X}$, let $p(A) = \sum_{x \in A} p(x)$ (and similarly for q), then

$$\max_{A \subseteq \mathcal{X}} (p(A) - q(A)) = \frac{1}{2}\|p(x) - q(x)\| = V(p, q)$$

the variational distance

Pinsker's inequality

$$D(p\|q) \geq \frac{1}{2\ln 2}\|p-q\|$$

For a discrete or continuous variable X, let $M(s) = E[\exp(sX)]$, then for all $s \geq 0$ we have the Chernoff bound

$$\Pr(X \geq a) \leq e^{-sa}M(s)$$

According to the Neyman–Pearson lemma, the optimal test between two (discrete) distributions p and q is of the form

$$\text{decide } p \text{ if } \ln\frac{p(x)}{q(x)} \geq \alpha$$

Thus,

$$\Pr(\text{decide } p|q \text{ is true}) = \Pr\left(\ln\frac{p(x)}{q(x)} \geq \alpha \Big| q\right) \leq e^{-s\alpha}E\left[\left(\frac{p}{q}\right)^s |q\right]$$

With $\alpha = 0$, and choosing $s = 1/2$

$$\Pr(\text{decide } p|q \text{ is true}) = \Pr(\text{decide } q|p \text{ is true}) \leq F(p,q)$$

where (assuming discrete variables)

$$F(p,q) = \sum_x \sqrt{p(x)q(x)}$$

is the fidelity of (p,q)

The entity $-\ln F(p,q)$ is called the Bhattacharyya distance

Distance Between Quantum States

The trace distance between ρ and σ

$$V(\rho, \sigma) = \frac{1}{2}\mathrm{Tr}|\rho - \sigma|$$

The fidelity of ρ and σ

$$F(\rho, \sigma) = \mathrm{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$$

If $\mathcal{E}$ is trace-preserving, then

$$V(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq V(\rho, \sigma)$$

and

$$F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma)$$

It always holds that

$$1 - \sqrt{F(\rho, \sigma)} \leq V(\rho, \sigma) \leq \sqrt{1 - (F(\rho, \sigma))^2}$$

$$\Rightarrow F(\rho, \sigma) = 1 \iff V(\rho, \sigma) = 0 \iff \rho = \sigma$$