Lecture 16

Thursday, November 13, 2014
8:30 AM

Ch. 12

Information quantities for channel $\mathcal{N}^{A\to B}$

Classical Mutual Information: $I(\mathcal{N}) = \max_{p(a)} I(A;B)$ ← Channel gives conditional distribution.

- Concave in $p(a)$: $I(A;B) = H(B) - H(B|A)$

  $H(B)$: concave composed with linear; $H(B|A)$: linear

$\Rightarrow$ Optimization is easy ("convex optimization problem")

Holevo Information: $\chi(\mathcal{N}) = \max_{\rho^{XA} \text{ seperable}} I(X;B)$

- Okay to require $\rho^{XA} = \sum_x p(x)\, |x\rangle\langle x|^X \otimes |\varphi_x\rangle\langle\varphi_x|^A$

  $|x\rangle$: orthogonal; $|\varphi_x\rangle$: pure state.

- Concave in $p(x)$. ← Consider mixtures as conditioning input on classical information Y

  Markov chain $Y-X-B \Rightarrow I(X;B) \geq I(X;B|Y)$.

- Difficult to optimize because choice of pure states $|\varphi_x\rangle$

- Proof that pure states suffice:

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

  Spectral decomposition of ρ_x: $\rho_x = \sum_y p(y|x)\, |\varphi_{xy}\rangle\langle\varphi_{xy}|$

  Define $\rho^{XYA} = \sum_{x,y} p(x)p(y|x)\, |x\rangle\langle x|^X \otimes |y\rangle\langle y|^Y \otimes |\varphi_{xy}\rangle\langle\varphi_{xy}|^A$ ← pure states

  Notice that $\mathrm{Tr}_Y(\rho^{XYA}) = \rho^{XA}$

  $\Rightarrow I(XY;A) \geq I(X;A)$

- Proof that $|x\rangle$ should be orthogonal pure state.

  Can always define channel $|x\rangle \to \sigma_x$ if $\{|x\rangle\}$ orthogonal.

  Data Processing ineq.

Mutual Information: $I(\mathcal{N}) = \max_{\rho^{XA}} I(X;B)$

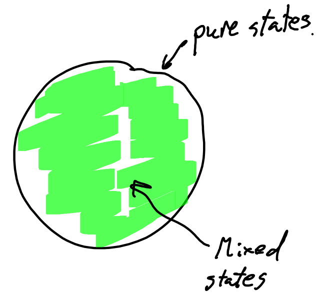
- Okay to require ρ^{XA} pure.

- Concave in input density (in an unusual way) pure states.

- Okay to require $\rho^{\cdot\cdot}$ pure.
- Concave in input density (in an unusual way)

$$\sum_y p(y)\, I(X;B)_{\rho_y} \le I(X;B)_\sigma$$

where σ^{XA} is purification of $\sum_y p(y)\rho_y^A$



Proof that pure state suffice:

Let $|\psi\rangle^{X'YA}$ be purification of ρ^{XA}

$$I(X'Y;B) \ge I(X;B) \qquad \text{Let } X = (X';Y')$$

Coherent Information: $\quad Q(N) = \max\limits_{\rho^{XA}_{pure}} I(X>B)$

$Q(N) \ge 0$ even though $I(X>B)$ can be < 0.

Proof: $\rho^{XA} = \rho^X \otimes \rho^A \leftarrow$ both pure

$$\Rightarrow \rho^{XB} = \rho^X \otimes N(\rho^A)$$

$$\Rightarrow I(X>B) = H(B) - H(X,B)$$

$$= H(B) - \cancel{H(X)}^{0} - H(B) = 0.$$

Not a convex opt. problem in general.

If channel is degradable, concave in input state.

($\Rightarrow$ not optimized by pure state)

Additivity:

Super-additivity:

$$I(N_1 \otimes N_2) \ge I(N_1) + I(N_2)$$

$$\chi(N_1 \otimes N_2) \ge \chi(N_1) + \chi(N_2)$$

$$Q(N_1 \otimes N_2) \ge Q(N_1) + Q(N_2)$$

$$P(N_1 \otimes N_2) \ge P(N_1) + P(N_1)$$

Proof: Classical

$$I(N^{X_1 \to Y_1} \otimes N^{X_2 \to Y_2}) = \max$$

$$I\left(\mathcal{N}_1^{X_1\to Y_1}\otimes \mathcal{N}_2^{X_2\to Y_2}\right) = \max_{p(x_1,x_2)} I(X_1X_2\,;Y_1Y_2)$$

$$\geq \max_{p(x_1)p(x_2)} I(X_1X_2\,;Y_1Y_2)$$

$$= \max_{p(x_1)p(x_2)} H(X_1X_2) + H(Y_1Y_2) - H(X_1X_2Y_1Y_2)$$

$$= \max_{p(x_1)p(x_2)} H(X_1) + H(X_2) + H(Y_1) + H(Y_2) - H(X_1Y_1) - H(X_2Y_2)$$

$$= \max_{p(x_1)p(x_2)} I(X_1;Y_1) + I(X_2;Y_2)$$

$$= I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

Quantum:

$$I\left(\mathcal{N}_1^{A_1\to B_1}\otimes \mathcal{N}_2^{A_2\to B_2}\right) = \max_{\rho^{XA_1A_2}} I(X\,;B_1B_2)$$

$$\geq \max_{\rho^{X_1A_1}\otimes\rho^{X_2A_2}} I(X_1X_2\,;B_1B_2)$$

$$= \max_{\rho^{X_1A_1}\otimes\rho^{X_2A_2}} I(X_1;B_1) + I(X_2;B_2)$$

$$= I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

Same proof for other quantities.

Sub-additivity:

$$I(\mathcal{N}_1\otimes\mathcal{N}_2) = I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

$$\chi(\mathcal{N}_1\otimes\mathcal{N}_2) = \chi(\mathcal{N}_1) + \chi(\mathcal{N}_2) \quad \text{if entanglement-breaking}$$

$$Q(\mathcal{N}_1\otimes\mathcal{N}_2) = Q(\mathcal{N}_1) + Q(\mathcal{N}_2) \quad \text{if degradable}$$

$$P(\mathcal{N}_1\otimes\mathcal{N}_2) = P(\mathcal{N}_1) + P(\mathcal{N}_2) \quad \text{if degradable}$$

(In fact, $P=Q$ if degradable, In general $P\geq Q$)

Proofs: Classical:

$$I\left(\mathcal{N}_1^{X_1\to Y_1}\otimes \mathcal{N}_2^{X_2\to Y_2}\right) = \max_{p(x_1,x_2)} I(X_1X_2\,;Y_1Y_2)$$

$$= \max_{p(x_1,x_2)} I(X_1;Y_1) + \cancelto{0}{I(X_2;Y_1|X_1)} + I(X_1X_2;Y_2|Y_1)$$

Notice: $Y_1 - X_1 - X_2 - Y_2$

Notice: $Y_1' - X_1 - X_2 - Y_2$

$$H(Y_2|Y_1) - H(Y_2|X_1 X_2 Y_1)$$
$$= H(Y_2|Y_1) - H(Y_2|X_2)$$
$$\leq H(Y_2) - H(Y_2|X_2)$$

$$I(\mathcal{N}_1 \otimes \mathcal{N}_2) \leq \max_{p(x_1 x_2)} I(X_1;Y_1) + I(X_2;Y_2)$$
$$\leq I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

Quantum M.I.

Consider isometric extensions $\mathcal{U}_{\mathcal{N}_1}^{A_1 \to B_1 E_1}$ and $\mathcal{U}_{\mathcal{N}_1}^{A_2 \to B_2 E_2}$

For any pure state $\rho^{XA_1A_2}$

$$I(X;B_1B_2)_\psi = H(X)_\psi + H(B_1B_2)_\psi - H(XB_1B_2)_\psi$$
$$= H(B_1B_2E_1E_2)_\psi + H(B_1B_2)_\psi - H(E_1E_2)_\psi$$
$$= H(B_1B_2|E_1E_2)_\psi + H(B_1B_2)_\psi$$
$$\leq H(B_1|E_1E_2)_\psi + H(B_2|E_1E_2) + H(B_1) + H(B_2)$$
$$\leq H(B_1|E_1)_\psi + H(B_2|E_2) + H(B_1) + H(B_2)$$
$$= H(B_1E_1)_\sigma + H(B_1) - H(E_1)$$
$$+ H(B_2E_2)_\theta + H(B_2) - H(E_2)$$
$$= H(XA_2)_\sigma + H(B_1)_\sigma - H(XA_2B_1)_\sigma$$
$$+ H(XA_1)_\theta + H(B_2)_\theta - H(XA_1B_2)_\theta$$
$$= I(XA_2;B_1) + I(XA_1;B_2)$$
$$\leq I(\mathcal{N}_1) + I(\mathcal{N}_2)$$

Density operators:

$$\sigma^{XB_1E_1A_2} = \mathcal{U}_{\mathcal{N}_1}^{A_1\to B_1E_1}(\rho^{XA_1A_2})$$
$$\theta^{XA_1B_2E_2} = \mathcal{U}_{\mathcal{N}_2}(\rho^{XA_1A_2})$$
$$\psi^{XB_1E_1B_2E_2} = \mathcal{U}_{\mathcal{N}_1} \otimes \mathcal{U}_{\mathcal{N}_1}(\rho)$$

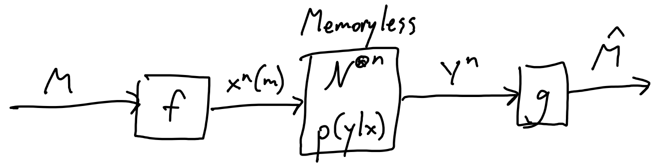
All pure and consistant.

Why do we care about additivity?

Converse proofs:

Memoryless communication setting $\mathcal{N}^{\otimes n}$

Classical Capacity:



Message : $M \sim \mathrm{Unif}[2^{nR}]$

Encoder : $f : [2^{nR}] \to \mathcal{X}^n$

Decoder : $g : \mathcal{Y}^n \to [2^{nR}]$

Def. Rate R achievable if $\forall \varepsilon > 0 \ \exists n, f, g : \ P(M \neq \hat{M}) < \varepsilon$

Def. Capacity : $C = \sup\{R \text{ achievable}\}$

$$\boxed{\text{Capacity Theorem (Shannon)}: \quad C = I(\mathcal{N})}$$

Notice : If $I(\mathcal{N})$ were not additive, this could not be true.

Conside super-symbols

Converse Proof:

$$\begin{aligned}
nR &= H(M) \\
&= I(M;\hat{M}) + H(M|\hat{M}) && \text{Small Fano's inequality} \\
&\leq I(X^n;Y^n) && \text{by Data Processing inequality} \\
& && M - X^n \to Y^n - \hat{M} \\
&\leq I(\mathcal{N}^{\otimes n}) \\
&= n\, I(\mathcal{N}) && \text{additivity}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow R &\leq I(\mathcal{N}) + \frac{1}{n}\left(nR\varepsilon + h(\varepsilon)\right) \\
&\leq I(\mathcal{N}) + \varepsilon R + h(\varepsilon)
\end{aligned}$$

True for all $\varepsilon > 0$. $\quad \Rightarrow R \leq \inf_{\varepsilon > 0}\left(I(\mathcal{N}) + \varepsilon R + h(\varepsilon)\right) = I(\mathcal{N}) \quad \square$