Lec 22: The Guassian Channel

8th Nov 2017

Recap

- Relative entropy $D(p\|q)$

$I(X;Y)$

$I(X^{\Delta};Y^{\Delta}) \longrightarrow I(X;Y)$

- $I(X;Y)$ $\quad \begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\underline{0}, \begin{bmatrix} \sigma^2, & \rho\sigma^2 \\ \rho\sigma^2, & \sigma^2 \end{bmatrix}\right)$

- $D(f\|g) \geq 0$
- $I(X;Y) \geq 0$
- $h(X|Y) \leq h(X)$

$h(X+a) = h(X)$

$h(aX) = h(X) + \log|a|$

$h(A\underline{X}) = h(\underline{X}) + \log|A|$

$\underline{X}$ with $E[\underline{X}] = 0$, $E[\underline{X}\,\underline{X}^T] = \underline{K}$

$\Rightarrow h(\underline{X}) \leq \frac{1}{2}\log(2\pi e)^n |K|$

Equality iff $X \sim N(0, K)$

Theorem: let X be a r.v and $\hat{X}$ be any estimate of X.

Then $E[(X-\hat{X})^2] \geq \frac{1}{2\pi e} e^{2h(X)}$ $\qquad X \to \boxed{\text{Estimate}} \to \hat{X}$

Pf: $\quad E\left[(X-\hat{X})^2\right] \geq \min_{\hat{X}} E\left[(X-\hat{X})^2\right] \overset{?}{=} E\left[(X-E[X])^2\right]$

$= \sigma_X^2$

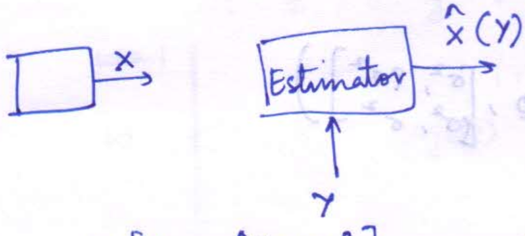
But if $Y \sim \eta(0, \sigma_X^2)$, we know

$h(Y) = \frac{1}{2}\log 2\pi e\sigma_X^2 \quad \therefore \sigma_X^2 = \frac{1}{2\pi e} e^{2h(Y)}$

$$\Rightarrow E[(X-\hat{X})^2] \geq \frac{1}{2\pi e} e^{2h(Y)}$$

$$\geq \frac{1}{2\pi e} e^{2h(X)}$$

Next consider



Corollary: $E[(X-\hat{X}(Y))^2] \geq \frac{1}{2\pi e} e^{2h(X|Y)}$

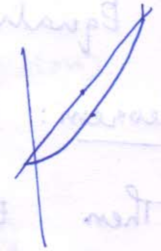
Pf: $$E[(X-\hat{X}(Y))^2] = \iint f(x,y)(x-\hat{x}(y))^2 \, dx \, dy$$

$$= \int_y f(y) \left[\int_x f(x|y) (x-\hat{x}(y))^2 dx\right] dy$$

$$\geq \int_y f(y) \left[\int_x (x - E(X|Y=y))^2 f(x|y) \, dx\right] dy$$

$$= \int_y f(y) \, \sigma^2_{X|Y=y} \, dy$$

$$\left(\sigma^2_{X|Y=y} \geq \frac{1}{2\pi e} e^{2h(X|Y=y)}\right)$$



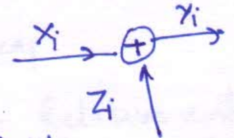
$$\geq \int_y \frac{f(y)}{2\pi e} e^{2h(X|Y=y)} dy$$

Using Jensens inequality

$$\geq \cancel{\int_y \frac{f(y)}{2\pi e}} \frac{1}{2\pi e} e^{\int_y f(y) \, 2h(X|Y=y) \, dy}$$

$$\geq \frac{1}{2\pi e} e^{2h(X|Y)}.$$

The Guassian channel



$\{Z_i\}$ iid $\sim \mathcal{N}(0, N)$

Assume $E[X_i^2] \le P$ (Power constraint)

Defn: The information capacity of the guassian channel is given by

$$C = \max_{\substack{f(x) \text{ st} \\ E[X^2] \le P}} I(X;Y)$$

We can compute C as follows

$$I(X;Y) = h(Y) - h(Y|X)$$

$$E[Y^2] = E[X^2] + E[Z^2] \le P + N$$

$$\therefore h(Y) \le \frac{1}{2} \log 2\pi e \,(P+N)$$

$$\Rightarrow I(X;Y) \le \frac{1}{2} \log 2\pi e\,(P+N) - h(X+Z|X)$$

$$= \frac{1}{2} \log 2\pi e\,(P+N) - h(Z|X)$$ [translation doesn't change diff entropy]

$$= \frac{1}{2} \log 2\pi e\,(P+N) - h(Z)$$

$$= \frac{1}{2} \log 2\pi e\,(P+N) - \frac{1}{2}\log(2\pi e) N$$

$$= \frac{1}{2} \log \frac{P+N}{N} = \frac{1}{2} \log\left(1 + \frac{P}{N}\right) = \underline{\frac{1}{2}\log(1+SNR)}$$

But by choosing $X \sim \mathcal{N}(0, P)$, we achieve this value of $I(X;Y)$

$$\therefore C = \frac{1}{2} \log\left(1 + \frac{P}{N}\right)$$

Plan to show this has operational meaning.

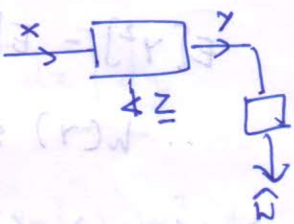
Defn : An (M,n) code for the Guassian channel under power constraint P consists of a) an index set $\{1, 2, .. M\}$

(b) an encoding function

$$x : \{1, 2, \cdots M\} \longrightarrow \mathcal{X}^n$$

$$\begin{array}{c c} & \begin{array}{ccccccc} 1 & 2 & \cdots & j & \cdots & \cdots & n \end{array} \\ \begin{array}{c} w=1 \\ \\ w \\ \\ w=M \end{array} & \begin{bmatrix} & & & & \\ & & & & \\ & & x_j(w) & & \\ & & & & \\ & & & & \end{bmatrix} \end{array}$$

where $\frac{1}{n} \sum_{i=1}^{n} x_i^2 \leq P \xrightarrow{\text{channel use}}$ power constraint.



(c) a decoding function

$$g : \mathcal{Y}^n \longrightarrow \hat{W} \in \{1, 2, \cdots M\}.$$

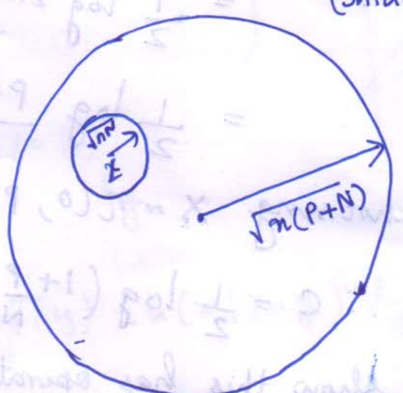
$$\lambda_w = Pr(E \mid W = w)$$

$$\lambda^{(n)} = \max_w \lambda_w \quad , \quad P_e^{(n)} = \frac{1}{M} \sum_{w=1}^{M} \lambda_w$$

Geometric view

(Intuition).

Vol of a sphere in n dimension $= c_n r^n$



Recieved vector lies in this sphere with high probability

$$\therefore \frac{c_n \left(\sqrt{n(P+N)}\right)^n}{c_n \left(\sqrt{nN}\right)^n}$$

$$= \left(\sqrt{1+\frac{P}{N}}\right)^n$$

$$= 2^{\frac{n}{2} \log\left(1+\frac{P}{N}\right)}$$

$$\therefore R \simeq \frac{1}{2} \log\left(1+\frac{P}{N}\right)$$

We will next show that using random coding arguments, the existence of a code that achieves this capacity.

$$
\begin{array}{c|c}
 & 1 \quad 2 \; \cdots \; j \; \cdots \; n \\
\begin{array}{c} 1 \\ 2 \\ \vdots \\ w \\ \vdots \\ M \end{array} & \left[\begin{array}{c} \\ \\ X_j(w) \\ \\ \\ \end{array} \right]
\end{array}
\quad \text{at random}
$$

Populate the codebook by sampling from $N(0, P-\epsilon)$ to generate each one of the Mn entries in the codebook.

<u>Decoding</u>:

$$
\hat{W} = \hat{w} \quad \text{if} \quad \begin{cases} (\underline{x}(\hat{w}), \underline{y}) \in A_\epsilon^{(n)} \\ \text{and} \\ (\underline{x}(w), \underline{y}) \notin A_\epsilon^{(n)} \\ \quad \text{for } w \neq \hat{w} \end{cases}
$$

Declare error if $\hat{W} \neq W$, Also declare error if $\frac{1}{n}\sum_{j=1}^{n} x_j^2(W) > P$.

~~Also declare e~~ let E_0 be the event that $\frac{1}{n}\sum_{j=1}^{n} X_j^2(w) > P$, let E_w be the event that $(\underline{X}(w), Y) \in A_\epsilon^{(n)}$.

$$Pr(\mathcal{E}) = \int f(\underline{c}) \, Pr(\mathcal{E}|\underline{c}) \, d\underline{c} \qquad \underline{c} = \text{codebook (Mn entries)}$$

$$= \int f(\underline{c}) \sum_{w=1}^{M} \frac{1}{M} Pr(\mathcal{E}|\underline{c}, W=w) \, d\underline{c}$$

$$= \frac{1}{M}\sum_{w=1}^{M} \underbrace{\int f(\underline{c}) \, Pr(\mathcal{E}|\underline{c}, W=w) \, d\underline{c}}_{\text{independant of } w} \qquad w=1$$

$\Rightarrow \Pr(\mathcal{E}) = \Pr(\mathcal{E} \mid W=1)$

Assume $W=1$ from here on.

$$\Pr(\mathcal{E} \mid W=1) = \Pr(E_0 \cup E_1^c \cup E_2 \cup \cdots E_M)$$
$$\le \Pr(E_0) + \Pr(E_1^c) + \sum_{j=2}^{M} \Pr(E_j)$$

Now

$\Pr(E_0) \le \epsilon$ since $\frac{1}{n}\sum_{j=1}^{n} x_j^2(\omega) \to P-\epsilon$

and $\Pr\left(\left| \frac{1}{n}\sum_{j=1}^{n} x_j^2(\omega) - (P-\epsilon) \right| > \epsilon \right) \le \epsilon$

$\therefore \Pr\left(\left| \frac{1}{n}\sum_{j=1}^{n} x_j^2(\omega) \right| > P \right) \le \epsilon$.

$\Rightarrow \Pr(E_0) \le \epsilon$

By typicality theory $\Pr(E_1^c) \le \epsilon$ for n-large enough

$\Pr((\underline{X}(1), \underline{Y}) \in A_\epsilon^{(n)}) \ge (1-\epsilon)$ for n-large enough.

Can show as in the discrete case that:

for $\omega \ne 1$ $\Pr((\underline{X}(\omega), Y) \in A_\epsilon^{(n)}) \le 2^{-n(I(X;Y)-3\epsilon)}$

$$\Rightarrow \Pr(\mathcal{E}) \le \epsilon + \epsilon + (M-1)\, 2^{-n(I(X;Y)-3\epsilon)}$$

Now substitute $M = 2^{nR}$

$$\le 2\epsilon + 2^{-n(I(X;Y)-R-3\epsilon)}$$

$$\le 3\epsilon$$

for n sufficiently large as long as

$$R < I(X;Y)$$

Note $P_e^{(n)} \le \epsilon$.

$\therefore$ for half of the code words, $\Pr(\mathcal{E} \mid W=\omega) \le 2\epsilon$ and every codeword satisfies power constraint. $\therefore$ By picking those code words alone $\frac{\log \lceil 2^{nR}/2 \rceil}{n} \to R$.

Hence $I(X;Y) = \frac{1}{2}\log\left(1+\frac{P-\epsilon}{N}\right)$

$$\longrightarrow \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

$\therefore$ we can get arbitrarily close to ~~I(X;Y~~

$$C = \frac{1}{2}\log\left(1+\frac{P}{N}\right)$$ using such a random

codebook. ~~and every~~