CS229r - IT ∈ CS

2/14/19

LECTURE 06

TODAY

"CHANNEL CODING"
- DEFINITIONS
- BINARY SYMMETRIC CHANNEL
- GENERAL CHANNELS

—— x ——

Next few lectures: Error-Correction (with "random errors")

General Channel of Communication (memoryless)

$$X \longrightarrow \boxed{} \longrightarrow Y$$

- given by $P_{Y|X}$ given by $\Omega_X \times \Omega_Y$ matrix

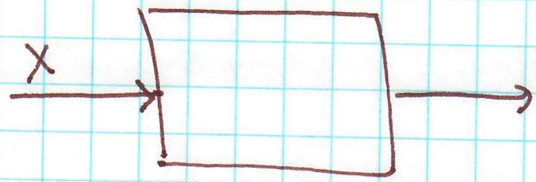
- $P_{Y|X}(\alpha, \beta) = \Pr[Y = \beta \mid X = \alpha]$.

n i.i.d. uses of channel.

How much "information" per use?

Simple Example

BSC(p) [Binary Symmetric Channel]



$$Y = \begin{cases} X & \text{w.p. } 1-p \\ 1-X & \text{w.p. } p \end{cases}$$

"Capacity": rate at which information can be pushed through

Formally: encoding + decoding functions E_n, D_n achieve rate R ~~if~~ with error ϵ

1. $E_n : \{0,1\}^{Rn} \to \Omega_X^n$
2. $D_n : \Omega_Y^n \to \{0,1\}^{Rn}$
3. $\Pr[\text{Decoding failure}] = \Pr_{\substack{m \leftarrow \text{Unif}(\{0,1\}^{Rn}) \\ Y \sim P_{Y|X=E(m)}}} [D(Y) \neq m]^{\epsilon}_{=o(1)}$

Capacity of channel $P_{Y|X}$

$$= \underset{R}{\sup}\ \left\{ \exists\ \cancel{E_n}\ \lim_{\epsilon \to 0} \lim_{n \to \infty} \left\{ \exists\, E_n, D_n \text{ of rate } R \text{ \& error } \epsilon \right\} \right\}$$

Connections to Information Theory

1. $\text{Capacity}(P_{Y|X}) = \max_{P_X} \{ I(X;Y) \}$!

[operational view of Information].
[theorem ... to be proved later].

2. Information Theory gives "best" algorithms + codes !!

Today 1

Special Case : $BSC(p)$

$$\text{Capacity} = 1 - h(p) \qquad [P_X = \text{Bern}(\tfrac{1}{2})].$$

$$= H(\text{Bern}(\tfrac{1}{2})) - H(Y|X)$$

$$= 1 - H(\text{Bern}(p))$$

$$= 1 - h(p).$$

Proof of 1 for $BSC(p)$.

1. $\forall \epsilon$, for suff large n, $R \geq 1 - h(p) - \epsilon$. ; $k = (R \cdot n)$

Pick $E_n : \{0,1\}^k \to \{0,1\}^n$ at random

D_n = max. likelihood decoding.

Lemma:

$$\Pr_{E_n, m, Y | E_n(m)} \left[D_n(Y) \neq m \right] \leq \epsilon$$

$$\left[\Rightarrow \exists E_n \; \Pr_{m, Y | E_n(m)} \left[\quad \cdots \quad \right] \leq \epsilon \right]$$

Proof: Error events

① $\Pr_{Y|X} [\underbrace{\Delta(Y, E_n(m)) \geq (p+\epsilon)}_{E1}] \leq \exp(-\epsilon^2 n)$

$\Uparrow$

Chernoff Bounds

$\forall E_n(m)$, $Y | E_n(m)$

② $\Pr_{E_n'} \left[\exists m' \neq m ; \; \Delta(E_n'(m'), Y) \leq (p+\epsilon) n \right]$

$$\leq 2^k \cdot \binom{n}{(p+\epsilon)n} \cdot \frac{1}{2^n}$$

$$\approx 2^k \cdot 2^{H(p) \cdot n} \cdot 2^{-n}$$

$$= \exp(-\epsilon n) . \boxtimes$$

if E1 or E2 don't happen then decoding right

$\boxtimes$

General Channel

Fix P_X

- Pick $E_n : \{0,1\}^{\cancel{R(\epsilon)n}\ Rn} \to \Omega_X^n$

  by picking $E_n(m)_i \sim P_X$ ind. for all (m,i).

- Decoding (Y)

  if $\exists$ unique ~~$X \in \Omega_X^n$~~ $m \in \{0,1\}^{Rn}$ s.t. for $X = E(m)$

  s.t. (1) X is P_X^n typical

  $\left[\Pr[X] \approx \cancel{\tfrac{1}{}}\ 2^{-H(P_X) n} \right]$

  & (2) (X,Y) is P_{XY}^n typical

  $$\Pr[XY] \approx \frac{1}{2^{H(P_{XY})n}}.$$

  Output m.

  else error.

Analysis

I Two types of errors

E1 X not typical, Y not typical, (X, Y) not jointly typical

E2 $(E(m'), Y)$ jointly typical, for some $m' \neq m$.

—— x ——

1. $\Pr[E1] \to 0$ by AEP

2. E2? Key + useful lemma.

—— x ——

Lemma: Let P, Q be distributions over Ω.

$$\Pr_{Z \sim P^n}\left[Z \text{ typical for } Q^n \right] \leq 2^{-D(Q \| P) \cdot n}$$

—— x ——

In our case

$(E(m'), Y)$ drawn from $P_X^n \times P_Y^n$

$$\Pr\left[(E(m'), Y) \text{ typical for } P_{XY}^n \right]$$

$$\leq 2^{-D(P_{XY} \| P_X \times P_Y) n}$$

$$= 2^{-I(Y; X) n}$$

$\Rightarrow$ Can take union bound over $2^{I(x;y)n}$ many m's.

$\Rightarrow$ Rate $\geq I(x;y)$!

Can optimize over P_x to get

$$\text{Capacity} \geq \sup_{P_x} \{ I(x;y) \}$$

<u>Converse Coding Theorem</u>:

<u>V1</u>: if Pr[decoding failure] $\to 0$ then ~~R ≤ Capacity~~

~~Capacity~~ Rate $\leq \sup_{P_x} \{ I(x;y) \}$

<u>V2</u>: for BSC(p) : if Rate $= \sup \{ I(x;y) \} + \epsilon$ then

Pr[decoding failure] $\geq 1 - \exp(-n)$.

[V2 much stronger quantitatively; but ~~holds~~ being shown only for BSC(p).]

Proof of V1 : (uses Fano's Inequality)

have $m \longrightarrow X^n \longrightarrow Y^n \longrightarrow \hat{m}$ – a Markov Chain.

① $I(X^n; Y^n) \leq$ ~~nee~~. $n \cdot \sup_{P_X} \{ I(X;Y) \}$

② $H(m) = nR = H(m|\hat{m}) + I(m; \hat{m})$

$\leq H(m|\hat{m}) + I(X^n; Y^n)$ [DPI]

$\leq H(m|\hat{m}) + n \cdot C$

need to bound $H(m|\hat{m})$

Fano: $H(m|\hat{m}) \leq h(\Pr[m \neq \hat{m}]) + \Pr[m \neq \hat{m}] \cdot nR$

$\leq 1 + o(nR)$

$\Rightarrow nR(1 - o(1)) \leq nC$

$R(1 - o(1)) \leq C$

$\Rightarrow R \leq C$ in the limit

$\boxtimes$

V2: Exercise / Pset.