Today

Shannon's Coding Theorem (for symmetric channels)

Review of last lecture

Defined Capacity of discrete memoryless channel

Channel specified by \( \mathbb{X}, \mathbb{Y} \subseteq \mathbb{R} \)

\[
\begin{cases}
    P_{y|x}(y|x) & \text{for } y \in \mathbb{Y}, x \in \mathbb{X}
\end{cases}
\]

Capacity \( \text{Cap} = \max_{\mathbb{X}} \left\{ I(X;Y) \right\} \)
Example Channels

BEC
BEC C 0 0 \rightarrow 0 ? \text{ prob. } = p \rightarrow 1 ^% G_p = 1 - p

BSC
BSC 0 \rightarrow 0 \text{ Cross-over prob. } = p \rightarrow 1 ^% G_p = 1 - H(p)

Symmetric Channel (row/column permutations)
Symmetric Channel \text{ Gap. } = \frac{1}{2} \log \left( \frac{1 - H(\text{row})}{2} \right) . \text{ Gap. = uniform}

\text{ Always } p^x = \text{ uniform}

Example
Example 0 0 \rightarrow 0 \rightarrow 1 \text{ Gap ... needs to be worked out, by } p^x = \text{ uniform}
Today o realize capacity

"Design" Encoder / Decoder s.t.

\[ m \in \{0,1\}^n \quad \Rightarrow \quad x^n \xrightarrow{\text{Encoder}} \quad \overline{y}^n \xrightarrow{\text{Decoder}} \quad \hat{y} \rightarrow \quad \text{Rec.} \]

Need functions

\[ E : \{0,1\}^n \rightarrow \mathbb{R}^n_x \]

\[ D : \mathbb{R}^n_y \rightarrow \{0,1\}^n_y \]

so that for

\[ \tilde{x}^n = E(m) ; \quad \tilde{y}^n = \text{Channel} (\tilde{x}^n) ; \quad m' = \delta (\tilde{y}^n) \quad \text{we have} \quad m' = m. \]
- Can't do this even for Erasure channel?

w.p. \( p^n y^n = ?^n \to \text{no info on } \tilde{x}^n \)

- So allow prob. of error.

\[ P_e \left[ m' \neq m \right] \text{ is small} \]

- What should \( E \) be?

- What should \( D \) do?

For fixed \( E \), \( D \) may as well be:

\[ m' = \arg \max \left[ \text{Prob} \left[ m \text{ is transmitted} \right] \right] \]

\[ m = \arg \max \left[ \text{Prob} \left[ y^n \text{ is rec'd} \mid X^n = E(m) \right] \right] \text{ is trans} \]

Very hard computationally; but math. well-defined.
But how do choose $E$?

Some issues:

BEC: $m_1, \ldots, m_k$

Receive: $\{Y_j\}_{j \in S}$

$|S| \approx n R_n = (1-p)n$

Need $I(m_1, \ldots, m_k; \{Y_j\}_{j \in S}) = k$

for most $S$.

BSC: $m_1, \ldots, m_k \rightarrow x_1, \ldots, x_n$

$\text{dim.} \approx pn$

Typical $y^k$ for $X^n$
- How can we design such E's?
- Need a new one for each channel?
- BSC: can we achieve any $R > 0$?

---

Shannon's remarkable "E"

(Works only for symmetric channels)

For every $m \in \{0,1\}^k$ [let $k = R \cdot n$]

- Pick $E(m) \in \mathbb{R}^n$.

- Independent for every $m \in \{0,1\}^k$.

$\Box$ }
Main Lemma:

\[ \forall \text{ Symmetric Channel } P_{y|x} \text{ with capacity } C \]
\[ \forall R < C \]

\[ \lim_{n \to \infty} \sum_{m \in \mathbb{Z}^n} \left[ m \in D(y) \right] \to 0 \]

\[ \Downarrow \]

\[ \lim_{n \to \infty} \left\{ \min_{E \subseteq \mathbb{Z}^n} \sum_{m \in E} \left[ m \in D(y) \right] \right\} \to 0 \]

Why does this $E$ work?

in case of BEC $\Rightarrow$ spreads int. uniformly

in case of BSC $\Rightarrow$ spreads codeword evenly
How to analyse?

From receiver's perspective
see \( Y^n \rightarrow \)

tries \( X^n(1) \rightarrow Y^n \) ?
\( X^n(2) \rightarrow Y^n \)
\( X^n(m) \rightarrow Y^n \)
\vdots
\( X^n(2^k) \rightarrow Y^n \) ?

How to pick message?

What can we expect \( Y \) to look like given \( m \)?

1. \( Y^n \) is equally likely to be a
"typical" required word given X

(E.g. in BSC \[ Y_i = X_i \text{ w.p. } 1-p \]
\[ = \overline{X_i} \text{ w.p. } p \])

\[ \Rightarrow Y_i \text{ will be like } X_i \text{ in all but } p \text{ places.} \]

etc.)

2. Can it also be a typical required word for \( m' \neq m \)?

But note:

(1) \( Y^n \) is distributed uniformly over \( \Omega^n_x \)

(Using Symmetry)
2. \( E(m') \) independent of \( Y^n \)

depends on

\( E(m) \); Channel;

but not \( E(m') \)!

\[ \Pr \left[ Y^n \text{ typical for } E(m) \right] \]

\[ \leq \left| \text{Size of typical set for } E(m) \right| \]

\( \binom{2^R}{n} \)

**Defn:** for any \( \tilde{x} \in \Omega_x^n \)

*typical set* \( A_{\epsilon, \tilde{x}}^{(n)} = \left\{ \tilde{y} \mid \Pr[\tilde{y}|\tilde{x}] \approx 2^{-H(\tilde{x})} \right\} \)

***Claim***: \( |A_{\epsilon, \tilde{x}}^{(n)}| \leq 2^{(H(\tilde{x}) + \epsilon) \cdot n} \)
Capacity = \log |\Omega| - \mathbb{H}(y)

R < C \Rightarrow \quad \text{for every } \bar{y} \in A_{\epsilon}, \bar{y}(m) \quad \text{(n)}

Claim: \quad \Pr \left[ \exists m' = m \mid y \in A_{\epsilon}, \bar{y}(m') \right]

\leq 2^{K \sum_{m'} |A_{\epsilon}, \bar{y}(m')|}

\leq (2^{\Omega})^n

\rightarrow 0 \quad \text{as} \quad R < C

\epsilon \leq \frac{C - R}{3}
Ready to prove Shannon's Theorem

Events

- Pick $x^n = E(m)$ at random; receive $y^n$...

- $E_1$: $y^n \notin A^{(n)}_{\delta, E(m)}$

\[ \Pr[ E_1 ] \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty \]

- Fix $E(m)$, $y^n$

- Now pick $E(m')$ s.t. $m' \neq m$

- $E_2$: $\exists m' \neq m$ s.t. $y \in A^{(n)}_{\epsilon, E(m')}$

\[ \Pr[ E_2 ] \rightarrow 0 \]
if neither E1 nor E2 occur
then \( m = D(\mathcal{E}^n) \)

---

Proves Shannon's Coding Theorem

\[ \text{But ONLY for symmetric channel.} \]

[\text{Covers BSC}]

[\text{Noisy typewriter (Big Deal!) But not BEC}]

Why not BEC? 

E still works, but analysis needs to change
- \((S_2^n)^n\) is irrelevant...

- \(Y^n\) is not random

**Broader Cases:**

"\(E(m)\) uniform in \(\Omega_x^n\)"

is not right!

(Channel capacity may not be maximized by unit. dist on \(X^n\);
then can't be that \(E(m)\) varying from that can achieve capacity)
Next Lecture

will consider

"E" picked as follows

Let \( P \) on \( \mathbb{S}_x \) be distribution maximizing \( I(x; y) \).

Then for every \( m \in \{0, 1\}^k \), \( i \in 1 \ldots n \),

\[ E(m) \text{ are i.i.d. with dist } P. \]

To analyze this setting will consider joint dist. on \( (x^n, y^n) \)
where $X^n$ dist according to $P^n$

$Y^n$ dist according to $(P_{y|x})^n$

Deducing procedure: $ar{Y}$

if $(E(m), Y)$ typical for $\text{dist}$;

$(E(m'), Y)$ not typical for $\text{dist}$;

then deducing $\bar{Y} = m$;

else ERROR.

Will refine $E_1$ ? for this setting.

$E_2$