Lectures 4-6

- Joint Typicality
- Channel Capacity Theorem
- Joint Source-Channel Coding
- Channels with Feedback
- Channels with Memory
- Entropy Rate of Stochastic Processes

Joint Typicality

$(\mathrm{X}, \mathrm{Y}) \sim \mathrm{p}(\mathrm{x}, \mathrm{y})$

AEP: $-\frac{1}{\mathrm{n}} \log p((X_1, Y_1), \ldots, (X_n, Y_n)) \rightarrow H(X, Y)$

Typical set $A_\varepsilon^{(n)}$: sequences $((x_1, y_1), \ldots, (x_n, y_n))$:

$2^{-\mathrm{n}(\mathrm{H}(\mathrm{X})+\varepsilon)} \leq p(x_1, \ldots, x_n) \leq 2^{-\mathrm{n}(\mathrm{H}(\mathrm{X})-\varepsilon)}$

$2^{-\mathrm{n}(\mathrm{H}(\mathrm{Y})+\varepsilon)} \leq p(y_1, \ldots, y_n) \leq 2^{-\mathrm{n}(\mathrm{H}(\mathrm{Y})-\varepsilon)}$

$2^{-\mathrm{n}(\mathrm{H}(\mathrm{X},\mathrm{Y})+\varepsilon)} \leq p((x_1, y_1), \ldots, (x_n, y_n)) \leq 2^{-\mathrm{n}(\mathrm{H}(\mathrm{X},\mathrm{Y})-\varepsilon)}$

Properties

1. $\mathbf{P((X^n, Y^n) \in A_\varepsilon^{(n)}) \to 1 \ \text{as} \ n \to \infty}$

2. $\mathbf{| A_\varepsilon^{(n)} | \le 2^{n(H(X,Y)+\varepsilon)}}$

3. **If** $(\tilde{\mathbf{X}}^n, \tilde{\mathbf{Y}}^n) \sim p(x^n)p(y^n)$, **then**

$$\mathbf{Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}) \le 2^{-n(I(X;Y)-3\varepsilon)}}$$

$$\mathbf{Pr((\tilde{X}^n, \tilde{Y}^n) \in A_\varepsilon^{(n)}) \ge (1-\varepsilon)2^{-n(I(X;Y)+3\varepsilon)}}$$

4. $x^n \in \mathbf{A}_\varepsilon^{(n)}, \quad \mathbf{A}_\varepsilon^{(n)}(Y^n, x^n) = \{y^n : (x^n, y^n) \in \mathbf{A}_\varepsilon^{(n)}\}$

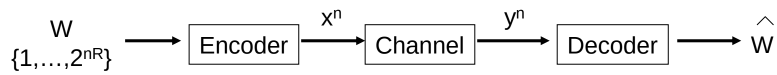
$$| \mathbf{A}_\varepsilon^{(n)}(Y^n, x^n) | \le 2^{n(H(Y|X)+2\varepsilon)}$$

Discrete Memoryless Channel

$$p(y_i \mid x^i, y^{i-1}) = p(y_i \mid x_i) \quad i = 1, \ldots, n$$

$$\Rightarrow p(y^n \mid x^n) = \prod_{i=1}^{n} p(y_i \mid x_i)$$

System Model



- Encoder assigns codeword $x^n(w)$ to each message
- Decoder assigns an index $\hat{w}$ to each received vector y^n

$$\text{Avg block error rate} = P_e^n = 2^{-nR} \sum_{w=1}^{2^{nR}} P(\hat{W} \neq W \mid W = w)$$

- Rate R achievable if there exist a sequence of $(n, 2^{nR})$ codes with P_e -> 0 as n -> infinity
- Capacity = Supremum of all achievable rates

Channel Capacity

- $C = \max_{p(x)} I(X;Y)$
- Achievability: Random coding with typical sequence decoding
- Converse: Data processing inequality and Fano's inequality

Joint Source-Channel Coding

- When can source be transmitted over a channel?
- Thm: $H(X) < C$
- Achievability: Typical sequences
- Converse: Similar to channel coding converse
- Implication: Joint source/channel coding does not increase capacity

Channels with Feedback

- Conditional distribution $p(y^n|x^n)$ no longer a product distribution
- Original converse altered slightly to show feedback does not increase capacity
  - Bound $I(W;Y^n)$ instead of $I(X^n;Y^n)$
- Feedback can make encoding/decoding easier, improve convergence of prob. error
  - Example: Erasure channel

Channels with Memory

- General capacity formula:

$$C_k = \max_{p(x^k)} \frac{1}{k} I(X^k; Y^k)$$

$$C = \lim_{k\to\infty} C_k$$

- Achievability: Random Coding
- Converse: Fano’s Inequality & Data-Processing
- Infinite letter characterization
  - Generally can not perform maximization

Entropy of Random Processes

- For stationary processes:

$$H(X) = \lim_{n\to\infty} \frac{1}{n} H(X_1,\ldots,X_n)$$

$$= \lim_{n\to\infty} H(X_n \mid X_1,\ldots,X_{n-1})$$

- For stationary & ergodic processes, AEP holds:

$$-\frac{1}{n}\log p(x_1,\ldots x_n) \to H(X) \quad \textbf{almost everywhere}$$