LECTURE 13

Office Hours Change: Wed. 4-5. (only for this week)

Last time:

- Midterm.
- Channel capacity as the limit of reliable communications.
- Blahut-Arimoto Algorithm.

Lecture outline

- More on source coding
- Joint Source-Channel coding.

Midterm Problem

$\underline{V}^n$ is a sequence of i.i.d. random variables. The entropy of each symbol is $H(V)$. Encode into binary sequence $\underline{U}$ with average length nR bits. A decoder decodes as $\hat{\underline{V}}^n$.

a): error $\triangleq \{\underline{V}^n \neq \hat{\underline{V}}^n\}$.

$$
\begin{aligned}
nR &\geq H(\hat{\underline{V}}^n) \\
&\geq H(\hat{\underline{V}}^n) - H(\hat{\underline{V}}^n|\underline{V}^n) \\
&= H(\underline{V}^n) - H(\underline{V}^n|\hat{\underline{V}}^n) \\
&= H(\underline{V}^n) - H(E, \underline{V}^n|\hat{\underline{V}}^n) \\
&= H(\underline{V}^n) - H(E|\hat{\underline{V}}^n) - H(\underline{V}^n|\hat{\underline{V}}^n, E) \\
&\geq nH(V) - 1 - H(\underline{V}^n|\hat{\underline{V}}^n, E) \\
&= nH(V) - 1 - P_e H(\underline{V}^n|\hat{\underline{V}}^n, E = 1) \\
&\qquad -(1 - P_e) H(\underline{V}^n|\hat{\underline{V}}^n, E = 0) \\
&\geq nH(V) - 1 - P_e H(\underline{V}^n)
\end{aligned}
$$

b): error $\triangleq \{D(\underline{V}^n, \underline{\widehat{V}}^n) > r\}$

The difference:

- Given $E = 1$, $H(\underline{V}^n | \underline{\widehat{V}}^n, E = 1)$ is slightly smaller.
- Given $E = 0$, $H(\underline{V}^n | \underline{\widehat{V}}^n, E = 0) \neq 0$.

$$
\begin{aligned}
nR &\geq nH(V) - 1 - H(\underline{V}^n | \underline{\widehat{V}}^n, E) \\
&\geq nH(V) - 1 - P_e H(\underline{V}^n) - (1 - P_e) \log |S_r^{(n)}|
\end{aligned}
$$

$$
P_e \geq 1 - \frac{nR + 1}{nH(V) - \log |S_r^{(n)}|}
$$

To drive $P_e \to 0$,

$$
R \geq H(V) - \frac{1}{n} \log |S_r^{(n)}|
$$

Source Coding as Sphere Covering

- Lossless coding: want each sequence $\underline{v}^n$ to be mapped into a distinct binary sequence. Need $2^{n \log |\mathcal{V}|}$ binary sequences.
- AEP: there are $2^{nH(V)} << 2^{n \log |\mathcal{V}|}$ typical sequences.
  - we can encode the atypical sequences with long codewords, which does not affect the average codeword length.
  - or we can ignore the atypical sequences in $\mathcal{V}^n$, and use only $2^{nH(V)}$ binary sequences, while the error probability approaches 0 as $n \to \infty$.

- Allowing some distortion: want each typical sequence $\underline{v}^n$ to be within r distance from some coded sequence.

- Use spheres of radius r to cover the set of typical sequences $A_\epsilon^{(n)}$. This requires $\frac{|A_\epsilon^{(n)}|}{|S_r^{(n)}|}$ binary sequences.

- Required rate

$$R = \frac{1}{n} \log \frac{|A_\epsilon^{(n)}|}{|S_r^{(n)}|} = H(V) - \frac{1}{n} \log |S_r^{(n)}|$$

- Covering $A_\epsilon^{(n)}$ with spheres of radius r.

Compare Source and Channel Coding

Notations

- An i.i.d. source sequence $\underline{V}^n$.
- Source coder $\mathcal{V}^n \to \{0,1\}^*$, map source sequences into binary sequences $\underline{U}^*$.
- Channel coder $\{0,1\}^* \to \mathcal{X}^m$, map binary sequence $\underline{U}^*$ into transmitted sequence $\underline{X}^m$.
- Receiver receives $\underline{Y}^m$.
- Channel decoder recovers $\widehat{\underline{U}}^*$.
- Source decoder guess $\widehat{\underline{V}}^n$.

- Source coding: use less bits to represent every possible source sequence– sphere covering in $\mathcal{V}^n$.
- Channel coding: use m transmitted symbols to represent as many as possible bits, while having small probability of error – sphere packing in $\mathcal{X}^m$.

Compare Source and Channel Coding

- Source coding reduces redundancy, use the most compact form to represent information.

$$|\mathcal{V}|^n \to 2^{nH(V)}$$

Example: Bern(p) sequence: n bits $\to nH(p)$ bits.

Example: Random process with correlation.

- Channel coding adds redundancy, to ensure reliable transmission.

$$2^{mI(X;Y)} \to 2^{mH(X)}$$

Example: BSC: $m(1 - H(\epsilon))$ bits from the input binary sequence $\to m$ bits as m transmitted symbols.

Joint Source Channel Coding

- Overall goal: transmit i.i.d. source $\underline{V}^n$ over the a DMC $P_{y|x}$ reliably:

$$P_e^{(n)} = P(\underline{V}^n \neq \widehat{\underline{V}}^n) \to 0$$

- This can be done whenever $H(V) < C$

$$\begin{aligned} P_e^{(n)} \;\leq\; & P(\underline{V}^n \text{ is not correctly encoded}) \\ & + P(\underline{U}^{nR} \text{ is not correctly recovered}) \end{aligned}$$

both the error probabilities $\to 0$ if $n \to \infty$, and $H(V) < R < C$.

- **Important** source and channel coding can be done separately.

Converse of Source-Channel Coding Theorem

For i.i.d. source $\underline{V}^n$ and DMC $P_{Y|X}$,

$$\underline{V}^n \to \underline{X}^n \to \underline{Y}^n \to \underline{\hat{V}}^n$$

, take $m = n$.

Joint source-channel encoder and decoder:

$$\begin{aligned} \underline{X}^n(\underline{V}^n) &: \mathcal{V}^n \to \mathcal{X}^n \\ \underline{\hat{V}}^n(\underline{Y}^n) &: \mathcal{Y}^n \to \mathcal{V}^n \end{aligned}$$

Probability of error

$$P_e^{(n)} = P(\underline{V}^n \neq \underline{\hat{V}}^n)$$

Question Is it possible to reliably transmit a source with $H(V) > C$?

Fano's inequality:

$$H(\underline{V}^n|\underline{\widehat{V}}^n) \leq 1 + P_e^{(n)} n \log|\mathcal{V}|$$

Now

$$\begin{aligned} H(V) &= \frac{1}{n}H(\underline{V}^n) \\ &= \frac{1}{n}H(\underline{V}^n|\underline{\widehat{V}}^n) + \frac{1}{n}I(\underline{V}^n;\underline{\widehat{V}}^n) \\ &\leq \frac{1}{n}(1 + P_e^{(n)} n \log|\mathcal{V}|) + \frac{1}{n}I(\underline{X}^n;\underline{Y}^n) \\ &\leq \frac{1}{n} + P_e^{(n)} \log|\mathcal{V}| + C \end{aligned}$$

$$P_e^{(n)} \to 0 \Rightarrow H(V) \leq C$$

Applies for very general cases:

- Source with memory
- Channel with memory

When Separation Theorem does NOT hold

- Channel is correlated with the source
- Multiple access with correlated sources
- Broadcasting channels.

Example: