EECS 769 Homework 4

Elise McEllhiney

Nov. 3, 2016

0.1 Channel with Jammer

A sender transmits a random variable X through a channel with output Y given by

$$Y = X + Z \mod 2$$

where Z is the signal inserted by a jammer. X and Z are independent. X and Z can only take binary values 0 or 1. The sender is limited in energy so that $E_X[X^2] \leq 1/2$ and the jammer is limited in energy so that $E_Z[Z^2] \leq 1/4$. The sender seeks to maximize mutual information $I(X;Y)$, while the jammer seeks to minimize it. What is the policy followed by the sender, by the jammer, and the resulting mutual information.

Since Z is independent of X, we know that:

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(X) - H(X|Y) = 1 - (2 - \frac{3}{4}\log 3) = \frac{3}{4}\log 3 - 1$$

Note: $H(X|Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \frac{3}{8}\log\frac{3}{4} + \frac{1}{8}\log\frac{1}{4} + \frac{1}{8}\log\frac{1}{4} + \frac{1}{8}\log\frac{1}{4} = 2 - \frac{3}{4}\log 3$
The sender will maximize mutual information by sending X as a uniform distribution with $p(X = 0) = 1/2$ and $p(X = 1) = 1/2$. The jammer would maximize disruption by sending Z as a uniform distribution. However, since it can send a maximum of 1 out of every 4 bits, it should do so. so it sends Z with the distribution $p(Z = 0) = 3/4$ and $p(Z = 1) = 1/4$

0.2 Statistician

One is given a communication channel with transition probabilities $p(x|y)$ and channel capacity $C = \max_{P_X} I(X;Y)$. A helpful statistician preprocesses the output by forming $\hat{Y} = g(Y)$. He claims that this will strictly improve the capacity.

(a) Show that he is wrong.

We know that he is wrong by the data processing inequality. $X \to Y \to \hat{Y}$ is a Markov chain. And as such, by the data processing inequality we know that:

$$I(X;Y) \geq I(X;\hat{Y})$$

If $\hat{p}(x)$ is the maximizing distribution, we then know:

$$C = max_{p(x)} I(X;Y) \geq I(X;Y)_{p(x)=\hat{p}(x)} \geq I(X;\hat{Y})_{p(x)=\hat{p}(x)} = max_{p(x)} I(X;\hat{Y})$$

(b) Under what conditions does he not strictly decrease the capacity?

He doesn't strictly decrease capacity if we have equality in the data processing inequality. In other words, he doesn't lose capacity if $X \to \hat{Y} \to Y$ also forms a Markov chain.

0.3 Channel with variable noise

Find the channel capacity of the following discrete memoryless channel:

$$Y = X + Z \quad X \in \{0, 1\} \quad Z \in \{0, a\}$$

where $Pr\{Z=0\} = Pr\{Z=a\} = 1/2$ and a is a non-zero integer. Assume that Z is independent of X.
Note that the channel capacity depends on the value of a.

If $a = 0:$ We know that $Y = X$. Therefore $\max I(X;Y) = \max H(X) = 1 bits/transmission$

If $a = 1:$ We know $Y = \{0,1,2\}$. This is a binary erasure with $a = 1/2$ since $Pr\{Z = a\} = 1/2$. Therefore the capacity when $a = 1$ is $1 - a = 1 - 1/2 = \frac{1}{2} bits/transmission$

If $a = -1:$ We know $Y = \{-1,0,1\}$. Again, this is a binary erasure channel with $a = 1/2$. Therefore, capacity when $a = -1$ is $1 - a = 1 - 1/2 = \frac{1}{2} bits/transmission$

If $a \neq -1,0,1:$ We know $Y = \{0,1,a,1+a\}$. Now we can see that if we know Y, we will know X and $H(X|Y) = 0$. Therefore, $\max I(X;Y) = \max H(X) = 1 bits/transmission$

0.4 Noisy typewriter channel

Consider the discrete memoryless channel

$$Y = X + Z(\mod 11)$$

where $Pr\{Z=1\} = Pr\{Z=2\} = Pr\{Z=3\} = 1/3$ and $X \in 0,1,...,10$. Assume that Z is independent of X.

(a) Find the capacity

Since Z is a uniform distribution and independent of X, we know that:

$$H(Z) = H(Z|X) = H(Y|X) = \log 3$$

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X) = \max_{p(x} H(Y) - \log 3 = \log 11 - \log 3 = \log \frac{11}{3} bits/transmission$$

(b) What is the maximizing $p^*(x)$

This is achieved when Y has a uniform distribution. Or, $p(X = i) = \frac{1}{11}$ for $i = \{0,1,...,10\}$

0.5 Erasure channel

Let $\{\mathcal{X}, p(x|y), \mathcal{Y}\}$ be a discrete memoryless channel with capacity C. Suppose this channel is immediately cascaded with an erasure channel $\{\mathcal{Y}, p(s|y), \mathcal{S}\}$ that erases α of its symbols. Specifically, $S = \{y_1, y_2, ..., y_m, e\}$, and

$$Pr\{S = y|X = x\} = \bar{\alpha}p(y|x), \quad y \in \mathcal{Y}$$

$$Pr\{S = e|X = x\} = \alpha$$

Determine the capacity of this channel

$$\begin{cases} 1 & Z = e \\ 0 & otherwise \end{cases}$$

Now we have a variable independent of X and we know $p(Z = 1) = \alpha$ and $p(Z = 0) = 1 - \alpha$

$$\begin{aligned}
I(X; S) &= H(S) - H(S|X) \\
&= H(S, Z) - H(S, Z|X) + H(Z) + H(S|Z) - H(Z|X) - H(S|X, Z) \\
&= I(X; Z) + I(S; X|Z) \\
&= 0 + \alpha I(X; S|Z = 1) + (1 - \alpha)I(X; S|Z = 0)
\end{aligned} \tag{1}$$

Since $Z = 1$ means we encountered an erasure, ie $S = e$. We know that $H(S|Z = 1) = H(S|X, Z = 1) = 0$. We also know that when $Z = 0$, $S = Y$ and therefore $I(X; S|Z = 0) = I(X; Y)$

$$I(X; S) = (1 - \alpha)I(X; Y)$$

A channel cascaded with an erasure channel will have $(1 - \alpha)$ times the capacity of the original channel (without the erasure channel).

0.6 Source and channel

We wish to encode a Bernoulli(α) process $V_1, V_2, ...$ for transmission over a binary symmetric channel with crossover probability p. Find the conditions on α and p so that the probability of error $P(\hat{V^n} \neq V^n)$ can be made to go to zero as $n \to \infty$

By the source-channel theorem we know that to have a probability of error approach 0 as $n \to \infty$ we need to have the entropy of the source be less than the channel capacity.

$$H(\alpha) + H(p) < 1 \qquad \alpha^{\alpha}(1 - \alpha)^{1-\alpha}p^{p}(1 - p)^{1-p} < \frac{1}{2}$$

0.7 Capacity

Find the capacity of

(Used in part b) Like the effective alphabet size of two combined alphabets, we'll find that $2^C = 2^{C_1} + 2^{C_2}$ when we have two channels in parallel.

$$X = \begin{cases} X_1 & \alpha \\ X_2 & 1 - \alpha \end{cases}$$

$$Z = \begin{cases} 1 & X_1 \\ 2 & X_2 \end{cases}$$

Thus, since we have disjoint channels we know that

$$X \to Y \to Z$$

$$I(X; Y, Z) = I(X|Z) + I(X; Y|Z) = I(X; Y) + I(X; Z|Y)$$

Since elements in a Markov chain are conditionally independent given earlier variables in the chain: $I(X; Z|Y) = 0$

$$I(X; Y, Z) = H(\alpha) + \alpha I(X_1; Y_1) + (1 - \alpha)I(X_2; Y_2)$$

$$C = \log(2^{C_1} + 2^{C_2})$$

(a) Two parallel BSCs

We have a symmetric channel. We know the highest capacity is achieved by a uniform distribution.

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X)$$
$$\leq \log |Y| - H(p) = 2 - H(p)$$

(b) BSC and a single symbol

I was unable to do this particular one without using the rule I found for parallel channels mentioned above

$$C_1 = 1 - H(p) \qquad C_2 = 0$$
$$2^C = 2^{C_1} + 1$$
$$C = log(2^{1-H(p)} + 1)$$

(c) BSC and a ternary channel

Since we have a weekly symmetric channel (pg 191).

$$C = \log |\mathcal{Y}| - H(\text{row of transition matrix}) = \log 5 - (\frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}) = \log 5 - \log 2 = \log\frac{5}{2}$$

We can also solve it the following way:
Ternary Channel :

$$C_1 = \max_{p(x)} I(X;Y) = \max p(x) H(Y) - H(Y|X) \leq \log 3 - H(1/2) = \log 3 - 1$$

BSC:

$$C_2 = 1 - H(p) = 0$$

In Parallel:

$$2^C = \frac{3}{2} + 1 \quad => \quad C = \log\frac{5}{2}$$

(d) Ternary channel

Since we have a weekly symmetric channel

$$C = \log |\mathcal{Y}| - H(\text{row of transition matrix}) = \log 3 - H(\frac{1}{3}, \frac{2}{3}) = \log 3 - (\log 3 - \frac{2}{3}) = \frac{2}{3}$$

The capacity is achieved by a uniform distribution of the input alphabet.