Homework 5 Solutions

Chapter 7

20) **A channel with two independent looks at Y.** Let Y_1 and Y_2 be conditionally independent and conditionally identically distributed given X.

a) Show $I(X;Y_1,Y_2) = 2I(X;Y_1) - I(Y_1,Y_2)$.

b) Conclude that the capacity of the channel

$$X \longrightarrow \boxed{} \longrightarrow (Y_1, Y_2)$$

is less than twice the capacity of the channel

$$X \longrightarrow \boxed{} \longrightarrow Y_1$$

20) *A channel with two independent looks at Y*

a)

$$\begin{aligned}
I(X;Y_1,Y_2) &= H(Y_1,Y_2) - H(Y_1,Y_2|X) && (670)\\
&= H(Y_1) + H(Y_2) - I(Y_1;Y_2) - H(Y_1|X) - H(Y_2|X) && (671)\\
&\quad \text{(since } Y_1 \text{ and } Y_2 \text{ are conditionally independent given } X\text{)} && (672)\\
&= I(X;Y_1) + I(X;Y_2) - I(Y_1;Y_2) && (673)\\
&= 2I(X;Y_1) - I(Y_1;Y_2) \quad \text{(since } Y_1 \text{ and } Y_2 \text{ are conditionally identi-. cally distributed)} && (674)
\end{aligned}$$

b) The capacity of the single look channel $X \rightarrow Y_1$ is

$$C_1 = \max_{p(x)} I(X;Y_1). \tag{675}$$

The capacity of the channel $X \rightarrow (Y_1, Y_2)$ is

$$\begin{aligned}
C_2 &= \max_{p(x)} I(X;Y_1,Y_2) && (676)\\
&= \max_{p(x)} 2I(X;Y_1) - I(Y_1;Y_2) && (677)\\
&\leq \max_{p(x)} 2I(X;Y_1) && (678)\\
&= 2C_1. && (679)
\end{aligned}$$

Hence, two independent looks cannot be more than twice as good as one look.

26) **Noisy typewriter.** Consider the channel with $x, y \in \{0,1,2,3\}$ and transition probabilities $p(y|x)$ given by the following matrix:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

a) Find the capacity of this channel.

b) Define the random variable $z = g(y)$ where

$$g(y) = \begin{cases} A & \text{if} \quad y \in \{0,1\} \\ B & \text{if} \quad y \in \{2,3\} \end{cases}.$$

For the following two PMFs for x, compute $I(X;Z)$

i)

$$p(x) = \begin{cases} \frac{1}{2} & \text{if} \quad x \in \{1,3\} \\ 0 & \text{if} \quad x \in \{0,2\} \end{cases}$$

ii)

$$p(x) = \begin{cases} 0 & \text{if} \quad x \in \{1,3\} \\ \frac{1}{2} & \text{if} \quad x \in \{0,2\} \end{cases}$$

c) Find the capacity of the channel between x and z, specifically where $x \in \{0,1,2,3\}$, $z \in \{A,B\}$, and the transition probabilities $P(z|x)$ are given by

$$p(Z = z|X = x) = \sum_{g(y_0)=z} P(Y = y_0|X = x)$$

d) For the X distribution of part i. of b, does $X \to Z \to Y$ form a Markov chain?

26) *Noisy typewriter*

a) This is a noisy typewriter channel with 4 inputs, and is also a symmetric channel. Capacity of the channel by Theorem 7.2.1 is $\log 4 - 1 = 1$ bit per transmission.

b) i) The resulting conditional distribution of Z given X is

$$\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

If

$$p(x) = \begin{cases} \frac{1}{2} & \text{if} \quad x \in \{1,3\} \\ 0 & \text{if} \quad x \in \{0,2\} \end{cases}$$

then it is easy to calculate $H(Z|X) = 0$, and $I(X;Z) = 1$. If

$$p(x) = \begin{cases} 0 & \text{if} \quad x \in \{1,3\} \\ \frac{1}{2} & \text{if} \quad x \in \{0,2\} \end{cases}$$

then $H(Z|X) = 1$ and $I(X;Y) = 0$.

ii) Since $I(X;Z) \leq H(Z) \leq 1$, the capacity of the channel is 1, achieved by the input distribution

$$p(x) = \begin{cases} \frac{1}{2} & \text{if} \quad x \in \{1,3\} \\ 0 & \text{if} \quad x \in \{0,2\} \end{cases}$$

c) For the input distribution that achieves capacity, $X \leftrightarrow Z$ is a one-to-one function, and hence $p(x,z) = 1$ or 0. We can therefore see the that $p(x,y,z) = p(z,x)p(y|x,z) = p(z,x)p(y|z)$, and hence $X \to Z \to Y$ forms a Markov chain.

29) **Binary multiplier channel.**

a) Consider the discrete memoryless channel $Y = XZ$ where X and Z are independent binary random variables that take on values 0 and 1. Let $P(Z = 1) = \alpha$. Find the capacity of this channel and the maximizing distribution on X.

b) Now suppose the receiver can observe Z as well as Y. What is the capacity?

29) *Binary Multiplier Channel* (Repeat of problem 7.23)

a) Let $P(X = 1) = p$. Then $P(Y = 1) = P(X = 1)P(Z = 1) = \alpha p$.

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) - P(X = 1)H(Z) \\ &= H(\alpha p) - pH(\alpha) \end{aligned}$$

We find that $p^* = \frac{1}{\alpha(2^{\frac{H(\alpha)}{\alpha}}+1)}$ maximizes $I(X;Y)$. The capacity is calculated to be $\log(2^{\frac{H(\alpha)}{\alpha}}+1) - \frac{H(\alpha)}{\alpha}$.

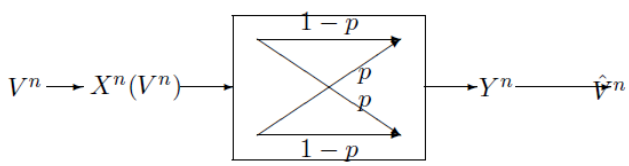
b) Let $P(X = 1) = p$. Then

$$\begin{aligned} I(X;Y,Z) &= I(X;Z) + I(X;Y|Z) \\ &= H(Y|Z) - H(Y|X,Z) \\ &= H(Y|Z) \\ &= \alpha H(p) \end{aligned}$$

The expression is maximized for $p = 1/2$, resulting in $C = \alpha$. Intuitively, we can only get X through when Z is 1, which happens α of the time.

31) **Source and channel.**

We wish to encode a Bernoulli(α) process $V_1, V_2, \ldots$ for transmission over a binary symmetric channel with crossover probability p.



Find 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 $n \longrightarrow \infty$.

31) *Source And Channel*

Suppose we want to send a binary i.i.d. Bernoulli(α) source over a binary symmetric channel with error probability p.

By the source-channel separation theorem, in order to achieve an error rate that vanishes asymptotically, $P(\hat{V}^n \neq V^n) \to 0$, we need the entropy of the source to be smaller than the capacity of the channel. In this case this translates to

$$H(\alpha) + H(p) < 1,$$

or, equivalently,

$$\alpha^\alpha(1-\alpha)^{1-\alpha}p^p(1-p)^{1-p} < \frac{1}{2}.$$

Chapter 8

2) **Concavity of determinants.** Let K_1 and K_2 be two symmetric nonnegative definite $n \times n$ matrices. Prove the result of Ky Fan [?]:

$$| \lambda K_1 + \overline{\lambda} K_2 | \geq | K_1 |^{\lambda} | K_2 |^{\overline{\lambda}}, \quad \text{for } 0 \leq \lambda \leq 1, \ \overline{\lambda} = 1 - \lambda,$$

where $| K |$ denotes the determinant of K.
Hint: Let $\mathbf{Z} = \mathbf{X}_\theta$, where $\mathbf{X}_1 \sim N(0, K_1)$, $\mathbf{X}_2 \sim N(0, K_2)$ and $\theta = \text{Bernoulli}(\lambda)$. Then use $h(\mathbf{Z} \mid \theta) \leq h(\mathbf{Z})$.

2) *Concavity of Determinants.* Let X_1 and X_2 be normally distributed n-vectors, $\mathbf{X}_i \sim \phi_{K_i}(\mathbf{x})$, $i = 1, 2$. Let the random variable θ have distribution $\Pr\{\theta = 1\} = \lambda$, $\Pr\{\theta = 2\} = 1 - \lambda$, $0 \leq \lambda \leq 1$. Let θ, $\mathbf{X}_1$, and $\mathbf{X}_2$ be independent and let $\mathbb{Z} = \mathbf{X}_\theta$. Then $\mathbb{Z}$ has covariance $K_Z = \lambda K_1 + (1 - \lambda) K_2$. However, $\mathbb{Z}$ will not be multivariate normal. However, since a normal distribution maximizes the entropy for a given variance, we have

$$\frac{1}{2} \ln(2\pi e)^n |\lambda K_1 + (1 - \lambda) K_2| \geq h(\mathbb{Z}) \geq h(\mathbb{Z}|\theta) = \lambda \frac{1}{2} \ln(2\pi e)^n |K_1| + (1 - \lambda) \frac{1}{2} \ln(2\pi e)^n |K_2| \ .$$

Thus

$$|\lambda K_1 + (1 - \lambda) K_2| \geq |K_1|^{\lambda} |K_2|^{1-\lambda} \ , \tag{698}$$

as desired.

6) **Variational inequality:** Verify, for positive random variables X, that

$$\log E_P(X) = \sup_Q \left[E_Q(\log X) - D(Q||P) \right] \tag{687}$$

where $E_P(X) = \sum_x x P(x)$ and $D(Q||P) = \sum_x Q(x) \log \frac{Q(x)}{P(x)}$, and the supremum is over all $Q(x) \geq 0$, $\sum Q(x) = 1$. It is enough to extremize $J(Q) = E_Q \ln X - D(Q||P) + \lambda(\sum Q(x) - 1)$.

6) *Variational inequality*
Using the calculus of variations to extremize

$$J(Q) = \sum_x q(x) \ln x - \sum_x q(x) \ln \frac{q(x)}{p(x)} + \lambda(\sum_x q(x) - 1) \tag{706}$$

we differentiate with respect to $q(x)$ to obtain

$$\frac{\partial J}{\partial q(x)} = \ln x - \ln \frac{q(x)}{p(x)} - 1 + \lambda = 0 \tag{707}$$

or

$$q(x) = c' x p(x) \tag{708}$$

where c' has to be chosen to satisfy the constraint, $\sum_x q(x) = 1$. Thus

$$c' = \frac{1}{\sum_x x p(x)} \tag{709}$$

Substituting this in the expression for J, we obtain

$$\begin{aligned} J^* &= \sum_x c' x p(x) \ln x - \sum_x c' x p(x) \ln \frac{c' x p(x)}{p(x)} && (710) \\ &= -\ln c' + \sum_x c' x p(x) \ln x - \sum_x c' x p(x) \ln x && (711) \\ &= \ln \sum_x x p(x) && (712) \end{aligned}$$

To verify this is indeed a maximum value, we use the standard technique of writing it as a relative entropy. Thus

$$\begin{aligned} \ln \sum_x x p(x) - \sum_x q(x) \ln x + \sum_x q(x) \ln \frac{q(x)}{p(x)} &= \sum_x q(x) \ln \frac{q(x)}{\frac{x p(x)}{\sum_y y p(y)}} && (713) \\ &= D(q \| p') && (714) \\ &\geq 0 && (715) \end{aligned}$$

Thus

$$\ln \sum_x x p(x) = \sup_Q \left(E_Q \ln(X) - D(Q \| P) \right) \tag{716}$$

This is a special case of a general relationship that is a key in the theory of large deviations.