Information Theory

Problem Sheet 3

Notation: x, $\mathbf{x}$, $\mathbf{X}$ are scalar, vector and matrix random variables respectively.

1. The Z channel. The Z-channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by the following matrix:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \quad x, y \in \{0,1\}$$

Find the capacity of the Z-channel and the maximizing input probability distribution.

2. Calculate the capacity of the following channel with probability transition matrix

$$Q = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix} \quad x, y \in \{0,1,2\}$$

3. Differential entropy. Evaluate the differential entropy $h(X) = -\int f \ln f$ for the following:
(a) The exponential density $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.
(b) The sum of X_1 and X_2; where X_1 and X_2 are independent normal random variables with means μ_i and variances σ_i^2 for $i = 1, 2$.

4. Parallel channels and waterfilling. Consider a pair of parallel Gaussian channels, i.e.,

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

where

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N\left(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right)$$

and there is a power constraint $E(X_1^2 + X_2^2) \leq 2P$. Assume that $\sigma_1^2 \geq \sigma_2^2$. At what power does the channel stop behaving like a single channel with noise variance σ_2^2 and begin behaving like a pair of channels?

5. Rate-distortion function. Let

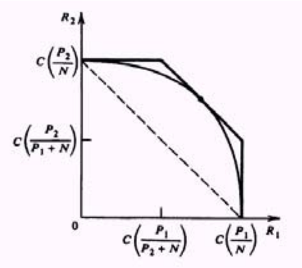
$$D(R) = \min_{p(\hat{x}|x): I(X;\hat{X}) \leq R} Ed(X, \hat{X})$$

be the rate-distortion function.
(a) Is $D(R)$ a decreasing or increasing function of R?
(b) Is $D(R)$ convex or concave in R?

6. Describe the capacity region of a two-user multiple access channel. Interpret the corner points (i.e., why can one of the users achieve the capacity as if the other user were absent?) Verify the following equality for the corner point:

$$C\left(\frac{P_1}{N}\right) + C\left(\frac{P_2}{P_1 + N}\right) = C\left(\frac{P_1 + P_2}{N}\right)$$

where $C(x) = (\log(1+x))/2$ is the capacity function.



7. Slepian-Wolf region for binary sources. Let X_i be i.i.d. Bernoulli(p). Let Z_i be i.i.d. Bernoulli(r), and let Z be independent of X. Finally, let Y = X⊕Z (mod 2 addition). Let X be described at rate R_1 and Y be described at rate R_2: What region of rates allows recovery of X, Y with probability of error tending to zero?

8. Describe the capacity region of a two-user Gaussian broadcast channel and sketch it.

Information Theory

Solution Sheet 3

1. First we express I(X;Y), the mutual information between the input and output of the Z-channel, as a function of x = Pr(X = 1):

$$H(Y\mid X) = \mathrm{P}(X = 0)\cdot 0 + \mathrm{P}(X = 1)\cdot 1 = x$$
$$H(Y) = H(P(Y=1)) = H(x/2)$$
$$I(X;Y) = H(Y) - H(Y\mid X) = H(x/2) - x$$

Taking the derivative, we have

$$\frac{dI(X;Y)}{dx} = \frac{1}{2}\log_2\frac{1-x/2}{x/2} - 1$$

which is equal to zero for x = 2/5. So the capacity of the Z-channel in bits is H(1/5) – 2/5 = 0.722 – 0.4 = 0.322.

2. Since the channel is symmetric,

$$C = \log|Y| - H(Q(1,:)) = \log_2 3 - 1 = 0.58 \text{ bits}$$

3. (a) log(e/λ) bits.

$$h(X) = -\int_0^\infty f(x)\log f(x)dx = -\log e\int_0^\infty f(x)\ln f(x)dx$$
$$= -\log e\int_0^\infty \lambda e^{-\lambda x}(\ln\lambda - \lambda x)dx = -\log e(\ln\lambda - \int_0^\infty \lambda x e^{-\lambda x}d\lambda x)$$
$$= -\log e(\ln\lambda - 1) = -\log e\cdot\ln(\lambda/e) = \log(e/\lambda) \text{ bits}$$

(b) The sum of two normal random variables is also normal, so applying the result derived the class for the normal distribution,

$$h(f) = \frac{1}{2}\log 2\pi e(\sigma_1^2 + \sigma_2^2) \text{ bits}$$

4. By the result of Lecture 14, it follows that we will put all the signal power into the channel with less noise until the total power of noise + signal in that channel equals the noise power in the other channel. After that, we will split any additional power "evenly" (in the sense that the power of noise + signal in one channel is equal to that in the other channel) between the two channels. Thus the combined channel begins to behave like a pair of parallel channels when the signal power is equal to the difference of the two noise powers, i.e., when $2P = \sigma_1^2 - \sigma_2^2$.

5. (a) Decreasing.

(b) Convex.

6. Capacity region

$$R_1 < C\left(\frac{P_1}{N}\right)$$
$$R_2 < C\left(\frac{P_2}{N}\right)$$
$$R_1 + R_1 < C\left(\frac{P_1 + P_2}{N}\right)$$

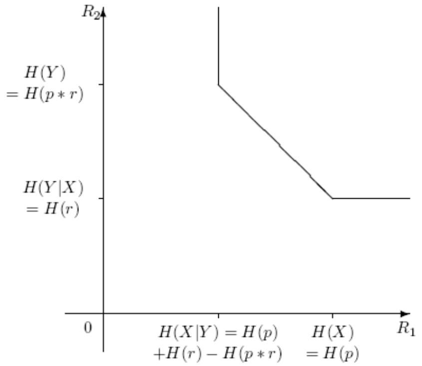
Interpretation: onion-peeling.
Verification:

$$\begin{aligned} C(\frac{P_1+P_2}{N}) &= \frac{1}{2}\log(1+\frac{P_1+P_2}{N}) \\ &= \frac{1}{2}\log(\frac{N+P_1+P_2}{N}) \\ &= \frac{1}{2}\log(\frac{N+P_1+P_2}{N+P_1}\cdot\frac{N+P_1}{N}) \\ &= \frac{1}{2}\log(\frac{N+P_1+P_2}{N+P_1}) + \frac{1}{2}\log(\frac{N+P_1}{N}) \\ &= C(\frac{P_2}{P_1+N}) + C(\frac{P}{N_1}) \end{aligned}$$

7. The Slepian-Wolf region is given by

$$R_1 \ge H(X\mid Y)$$
$$R_2 \ge H(Y\mid X)$$
$$R_1 + R_2 \ge H(X,Y)$$

X = Bernoulli(*p*). *Y* = *X*⊕*Z*, *Z* = Bernoulli(*r*). Then *Y* = Bernoulli(*p***r*), where *p***r* = *p*(1-*r*)+*r*(1-*p*). *H*(*X*) = *H*(*p*). *H*(*Y*) = *H*(*p***r*), *H*(*X*,*Y*) = *H*(*X*,*Z*) = *H*(*X*)+*H*(*Z*) = *H*(*p*)+*H*(*r*). Hence *H*(*Y*|*X*) = *H*(*r*) and *H*(*X*|*Y*) = *H*(*p*)+*H*(*r*)- *H*(*p***r*).
This region is shown in the next page.



8. The capacity region is

$$R_1 \leq C\left(\frac{\alpha P}{N_1}\right)$$

$$R_2 \leq C\left(\frac{(1-\alpha)P}{\alpha P + N_2}\right)$$

where $0 \leq \alpha \leq 1$. It looks like this: