E2 201 (Aug–Dec 2014)

Homework Assignment 7

Discussion: Friday, Nov. 14 Quiz: Friday, Nov. 21

This assignment consists of two pages.

1. [Gallager, Problem 7.6; see also Problem 9.8 in Cover & Thomas, 2nd ed.]

   We have 10 parallel, independent, additive Gaussian noise channels

   $$Y_i = X_i + Z_i$$

   with $Z_i \sim \mathcal{N}(0, i^2)$, for $i = 1, 2, \ldots, 10$. The inputs to the 10 channels must satisfy an overall power constraint

   $$\sum_{i=1}^{10} \frac{1}{i} P_i \leq P,$$

   where P_i is the power allocated to the ith channel.

   (a) Give an expression for the capacity of this parallel channel.

   (b) Until what value of P does the overall channel act like a single channel (i.e., all power is allocated to a single channel)? For what range of P does power get allocated to all 10 channels?

2. (a) Consider a memoryless channel $\mathbf{Y} = \mathbf{X} + \mathbf{Z}$, with input and output alphabets $\mathbb{R}^3$, and additive Gaussian noise $\mathbf{Z}$ with mean $\mathbf{0}$ and covariance matrix

   $$K = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

   The noise $\mathbf{Z}$ is independent of $\mathbf{X}$. Determine the capacity of the channel under an average input power constraint $\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x}_i\|^2 \leq P$, and give an explicit coding scheme for achieving capacity.

   (b) Suppose that we have a channel as in part (a), except that the noise $\mathbf{Z}$ is **not** specified to be Gaussian. The channel description otherwise remains the same, including the mean and covariance matrix of the noise. Is this information sufficient to determine the capacity of the channel under an average input power constraint $\frac{1}{n}\sum_{i=1}^{n} \|\mathbf{x}_i\|^2 \leq P$? If YES, determine the capacity, and give an explicit coding scheme for achieving capacity. If NO, then give a non-zero lower bound on the capacity, and give an explicit coding scheme that achieves the lower bound.

3. [Based on Problem 9.14, Cover & Thomas, 2nd ed.] Consider a memoryless channel $Y = X + Z$ with input and output alphabets $\mathbb{R}$, and additive noise Z independent of X, defined as follows:

   $$Z = \begin{cases} 0 & \text{with probability } p \\ Z^* & \text{with probability } 1 - p \end{cases}$$

   where $p > 0$ and $Z^* \sim \mathcal{N}(0, \sigma^2)$. Determine the capacity of this channel under an average input power constraint P. Also, give a coding scheme (encoder + decoder) for achieving capacity.

4. Consider a k-dimensional Gaussian source $\mathbf{X} \sim \mathcal{N}_k(0, K)$, with $k \times k$ covariance matrix K. Let $d : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}_+$ be the squared-error distortion measure: for $\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}^k$, $d(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|^2$. Show that the information rate-distortion function $R^{(I)}(D)$, defined as

   $$\min_{f(\hat{\mathbf{x}}|\mathbf{x}) : \mathbb{E}d(\mathbf{X}, \hat{\mathbf{X}}) \leq D} I(\mathbf{X}\,;\,\hat{\mathbf{X}}),$$

evaluates to

$$R^{(I)}(D) = \begin{cases} \sum_{j=1}^{k} \frac{1}{2} \log \frac{\lambda_j}{D_j} & \text{if } 0 \leq D < \text{trace}(K) \\ 0 & \text{if } D \geq \text{trace}(K) \end{cases}$$

where $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of K, and $D_j = \min\{\theta, \lambda_j\}$ with θ chosen so that $\sum_{j=1}^{k} D_j = D$. [*Hint*: First prove this for $K = \text{diag}(\sigma_1^2, \ldots, \sigma_k^2)$. Then, for general K, use diagonalization $K = Q\Lambda Q^T$.]

5. Consider the rate-distortion problem with source and reconstruction alphabets $\mathcal{X} = \hat{\mathcal{X}} = \{1, 2, \ldots, m\}$ for some $m \geq 2$, and Hamming distortion measure

$$d(x, \hat{x}) = \begin{cases} 0 & \text{if } x = \hat{x} \\ 1 & \text{if } x \neq \hat{x} \end{cases}$$

   (a) For any source random variable X taking values in the alphabet $\mathcal{X}$, show that the rate-distortion function $R(D)$ can be lower bounded as follows:

$$R(D) \geq H(X) - D \log(|\mathcal{X}| - 1) - H(D)$$

   for $0 \leq D \leq D^*$, where as usual, $D^* = \min_{\hat{x} \in \hat{\mathcal{X}}} \mathbb{E}[d(X, \hat{x})]$.

   (b) Determine $R(D)$ for a source random variable X that is uniformly distributed over the alphabet $\mathcal{X}$.

6. Consider a discrete memoryless Bernoulli($\frac{1}{2}$) source X, with Hamming distortion measure d. A sequence of n bits, X^n, produced by the source is to be encoded so that its reconstruction $\hat{X}^n$ satisfies $\mathbb{E}d(X^n, \hat{X}^n) \leq D$. (Recall that $d(X^n, \hat{X}^n) = \frac{1}{n} \sum_{i=1}^{n} d(X_i, \hat{X}_i)$.) Two coding schemes are proposed.

   Scheme A: Whenever the source produces a '0', it is flipped to a '1' with probability ρ (it stays as '0' with probability $1 - \rho$), where the value of ρ is chosen in accordance with the maximum allowed expected distortion D. Whenever the source produces a '1', the bit remains unchanged. In this manner, the original symmetric source is converted into an asymmetric source. Next, this asymmetric source is *losslessly* encoded, using R_A bits/source symbol. (So as to be unambiguous, R_A is the average number of coded bits per source symbol.) The reconstruction (or decoding) is performed by reconstructing (losslessly) the asymmetric source. Let $R_A(D)$ denote the minimum rate R_A needed by this scheme to yield an expected distortion at most D.

   Scheme B: Given a source sequence of length n, only the first nR_B bits are losslessly encoded. The decoding consists of reconstructing (losslessly) these first nR_B bits and padding them with 0s to obtain a sequence of total length n. Let $R_B(D)$ denote the minimum number of bits/source symbol needed by this scheme to yield an expected distortion not exceeding D.

   (a) Determine explicitly $R_A(D)$ as a function of D, $0 \leq D \leq \frac{1}{2}$.

   (b) Determine explicitly $R_B(D)$ as a function of D, $0 \leq D \leq \frac{1}{2}$.

   (c) Compare Schemes A and B by plotting $R_A(D)$ and $R_B(D)$ in the same figure. Include also in this figure the rate-distortion function $R(D)$ of the given source.