Homework Set #10

Due TBD

Problem 1 (A MUTUAL INFORMATION GAME)

Consider the channel shown in Fig. 1

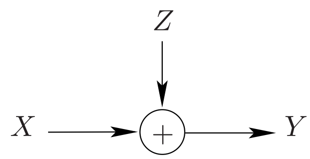

Figure 1: The channel

Throughout this problem we shall constrain the signal power

$$\mathbb{E}[X] = 0, \quad \mathbb{E}[X^2] = P, \tag{1}$$

and the noise power

$$\mathbb{E}[Z] = 0, \quad \mathbb{E}[Z^2] = N, \tag{2}$$

and assume that X and Z are independent. Let $\mathcal{F}_P$ denotes the set of distributions with constraints in (1), and $\mathcal{F}_N$ denotes the set of distributions with constraints in (2). For a given noise distribution, the channel capacity is given by $\max_{p(x)\in\mathcal{F}_P} I(X; X+Z)$.

Now, consider the following game with two players, the noise player and the signal player. The noise player chooses a distribution on Z to minimize $I(X; X+Z)$, while the signal player chooses a distribution on X to maximize $I(X; X+Z)$. In the following you are asked to show that $X^* \sim \mathcal{N}(0, P)$ and $Z^* \sim \mathcal{N}(0, N)$ satisfy the saddle point conditions

$$\begin{aligned}
\min_{p(z)\in\mathcal{F}_N} \max_{p(x)\in\mathcal{F}_P} I(X; X+Z) &= \max_{p(x)\in\mathcal{F}_P} \min_{p(z)\in\mathcal{F}_P} I(X; X+Z) \\
&= I(X^*; X^* + Z^*) \\
&= \frac{1}{2}\log\left(1 + \frac{P}{N}\right).
\end{aligned} \tag{3}$$

(a) Let $f(\cdot, \cdot)$ be an arbitrary function with two variables. Show that

$$\min_a \max_b f(a, b) \geq \max_b \min_a f(a, b).$$

(b) Show that if the noise player chooses the Gaussian noise, the best the signal player can do is choose the Gaussian distribution, *i.e.,* for all X such that $p(x) \in \mathcal{F}_P$,

$$I(X; X+Z^*) \leq I(X^*; X^* + Z^*).$$

(c) Given the fact that the signal player has chosen the Gaussian distribution, show that the worst noise is the Gaussian noise, *i.e.*, for all Z such that $p(z) \in \mathcal{F}_N$,

$$I(X^*; X^* + Z^*) \leq I(X^*; X^* + Z).$$

Hint: Verify the following chain on equalities and inequalities.

$$\begin{aligned}
I(X^*; X^* + Z^*) - I(X^*; X^* + Z) &\overset{(1)}{=} h(Y^*) - h(Z^*) - h(Y) + h(Z) \\
&\overset{(2)}{=} -\int_y f_{Y^*}(y) \log f_{Y^*}(y) dy + \int_y f_Y(y) \log f_Y(y) dy \\
&\quad + \int_z f_{Z^*}(z) \log f_{Z^*}(z) dz - \int_z f_Z(z) \log f_Z(z) dz \\
&\overset{(3)}{=} -\int_y f_Y(y) \log f_{Y^*}(y) dy + \int_y f_Y(y) \log f_Y(y) dy \\
&\quad + \int_z f_Z(z) \log f_{Z^*}(z) dz - \int_z f_Z(z) \log f_Z(z) dz \\
&\overset{(4)}{=} \int_y f_Y(y) \log \frac{f_Y(y)}{f_{Y^*}(y)} dy + \int_z f_Z(z) \log \frac{f_{Z^*}(z)}{f_Z(z)} dz \\
&\overset{(5)}{=} \int_y \int_z f_{Y,Z}(y,z) \log \frac{f_Y(y) f_{Z^*}(z)}{f_{Y^*}(y) f_Z(z)} dz dy \\
&\overset{(6)}{\leq} \log \int_y \int_z f_{Y,Z}(y,z) \frac{f_Y(y) f_{Z^*}(z)}{f_{Y^*}(y) f_Z(z)} dz dy \\
&\overset{(7)}{=} \log \int_y \int_z f_Z(z) f_{X^*}(y-z) \frac{f_Y(y) f_{Z^*}(z)}{f_{Y^*}(y) f_Z(z)} dy dz \\
&\overset{(8)}{\leq} \log \int_y \frac{f_Y(y)}{f_{Y^*}(y)} \int_z f_{X^*}(y-z) f_{Z^*}(z) dz dy \\
&\overset{(9)}{=} \log \int_y \frac{f_Y(y)}{f_{Y^*}(y)} f_{Y^*}(y) dy \\
&\overset{(10)}{=} \log 1 = 0.
\end{aligned}$$

(d) Using parts (b) and (c) conclude that

$$\min_{p(z) \in \mathcal{F}_N} \max_{p(x) \in \mathcal{F}_P} I(X; X + Z) \leq \max_{p(x) \in \mathcal{F}_P} \min_{p(z) \in \mathcal{F}_N} I(X; X + Z),$$

and combining it with the result of part (a), show (3).

Problem 2 (ERASURE DISTORTION)

Consider $X \sim$ Bernoulli $(\frac{1}{2})$, and let the distortion measure be given by the matrix

$$d(x, \hat{x}) = \begin{bmatrix} 0 & 1 & \infty \\ \infty & 1 & 0 \end{bmatrix},$$

where $\hat{x} \in \{0, E, 1\}$ and E denotes the erasure symbol. Calculate the rate distortion function for this source. Can you suggest a simple scheme to achieve any value of the rate distortion function you found?

Problem 3 (Convexity of mutual information as a function of $w(y|x)$)

Let

$$(X, Y_1) \sim p_1(x, y) = q(x)w_1(y|x)$$

and

$$(X, Y_2) \sim p_2(x, y) = q(x)w_2(y|x),$$

where X, Y_1 and Y_2 are continuous random variables. Construct

$$(X, Y_\lambda) \sim p_\lambda(x, y) = q(x)w_\lambda(y|x) = q(x)(\lambda w_1(y|x) + (1-\lambda)w_2(y|x)).$$

Prove that $I(X; Y_\lambda) \leq \lambda I(X; Y_1) + (1-\lambda)I(X; Y_2)$, *i.e.*, prove that the mutual information $I(X; Y)$ is a convex function of $w(y|x)$ for fixed $q(x)$.

Hint: Define a Bernoulli random variable Z independent of X, where $\Pr(Z = 1) = 1 - \Pr(Z = 2) = \lambda$*, and show that*

$$Y_\lambda = \begin{cases} Y_1 & \text{if } Z = 1 \\ Y_2 & \text{if } Z = 2. \end{cases}$$

Compare $I(X; Y_\lambda)$ to $I(X; Y_\lambda|Z)$.

Problem 4

Consider the case of representing m independent normal random sources $X_1 \cdots X_m$ with squared-error distortion. Assume $X_i \sim \mathcal{N}(0, \sigma_i^2)$ and assume that we are given R bits with which to represent this random vectors. Through out this problem we answer to the natural question of how to allot these R bits to minimize the total distortion. Define $d(x^m, \hat{x}^m) = \sum_{i=1}^m (x_i - \hat{x}_i)^2$ and write the natural extension of the rate distortion function for this vector case:

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

(a) Verify the following steps:

$$\begin{aligned} I(X^m; \hat{X}^m) &\geq \sum_i h(X_i) - h(X_i|\hat{X}_i) &(4)\\ &\geq \sum_i R(D_i) &(5)\\ &= \sum_i (\frac{1}{2} \log \frac{\sigma_i^2}{D_i})^+. &(6) \end{aligned}$$

What is D_i in (5)?

(b) When is the above set of inequalities tight? i.e., when do we have $I(X^m; \hat{X}^m) = \sum_{i=1}^m (\frac{1}{2} \log \frac{\sigma_i^2}{D_i})^+$?

(c) Write D in terms of D_i and reformulate $R(D)$ as a minimization problem over $\sum_i (\frac{1}{2} \log \frac{\sigma_i^2}{D_i})^+$.

(d) Use Kuhn-Tucker conditions to check that the rate distortion function is given by

$$R(D) = \sum_{i=1}^m \frac{1}{2} \log \frac{\sigma_i^2}{D_i}$$

where

$$D_i = \begin{cases} \lambda & \text{if } \lambda < \sigma_i^2 \\ \sigma_i^2 & \text{if } \lambda \geq \sigma_i^2 \end{cases}$$

where λ is such that $\sum_i D_i = D$