ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 18 Information Theory and Coding

Homework 9 November 20, 2012

PROBLEM 1. A discrete memoryless channel has three input symbols: $\{-1; 0; 1\}$, and two output symbols: $\{1; -1\}$. The transition probabilities are

$$p(-1|-1) = p(1|1) = 1, \quad p(1|0) = p(-1|0) = 0.5.$$

Find the capacity of this channel with cost constraint β, if the cost function is $b(x) = x^2$.

PROBLEM 2. For given positive numbers $\sigma_1^2, \ldots, \sigma_K^2$ define the function

$$f(p_1, \ldots, p_K) = \sum_{i=1}^{K} \log(1 + p_i/\sigma_i^2)$$

on the simplex $\{(p_1, \ldots, p_K) : p_i \geq 0, \sum_i p_i = 1\}$.

(a) Show that f is concave.

(b) Write the Kuhn-Tucker conditions for the p that maximizes $f(p)$; show that they can be equivalently written as "there exists λ, such that

$$\begin{aligned} p_i &= \lambda - \sigma_i^2, \quad \text{for } i \text{ for which } p_i > 0 \\ 0 &\geq \lambda - \sigma_i^2, \quad \text{for } i \text{ for which } p_i = 0\text{''}. \end{aligned}$$

(c) Show that the maximizing p can be written in the form

$$p_i = (\lambda - \sigma_i^2)^+$$

for some λ where $a^+ = \max\{0, a\}$.

(d) Given what we have shown so far, the maximization of f is reduced to finding the right λ. Describe a procedure to find this λ.

PROBLEM 3. Suppose Z is uniformly distributed on $[-1, 1]$, and X is a random variable, independent of Z, constrainted to take values in $[-1, 1]$. What distribution for X maximizes the entropy of $X + Z$? What distribution of X maximizes the entropy of XZ?

PROBLEM 4. Random variables X and Y are correlated Gaussian variables:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; K = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \right).$$

Find $I(X; Y)$.

PROBLEM 5. Consider an additive noise channel with input $x \in \mathbb{R}$, and output

$$Y = x + Z$$

where Z is any real random variable independent of the input x, has zero mean and variance equal to σ^2.

In this problem we prove in two different ways that the Gaussian channel has the smallest capacity among all additive noise channels of a given noise variance. Let $\mathcal{N}_{\sigma^2}$ denote the Gaussian density with zero mean and variance σ^2.

First Method: Let X be a random variable with density $\mathcal{N}_P$.

(a) Show that

$$I(X;Y) = H(X) - H(X - \alpha Y|Y)$$

for any $\alpha \in \mathbb{R}$.

(b) Observe that

$$H(X - \alpha Y) \leq \frac{1}{2} \log 2\pi e E((X - \alpha Y)^2)$$

for any $\alpha \in \mathbb{R}$.

(c) Deduce from (*a*) and (*b*) that

$$I(X;Y) \geq H(X) - \frac{1}{2} \log 2\pi e E((X - \alpha Y)^2)$$

for any $\alpha \in \mathbb{R}$.

(d) Show that

$$E((X - \alpha Y)^2) \geq \frac{\sigma^2 P}{\sigma^2 + P}$$

with equality if and only if $\alpha = \frac{P}{P+\sigma^2}$.

(e) Deduce from (*c*) and (*d*) that

$$I(X;Y) \geq \frac{1}{2} \log \left(1 + \frac{P}{\sigma^2}\right)$$

and conclude that the Gaussian channel has the smallest capacity among all additive noise channels of a given noise variance.

Second Method:

(a) Denote the input probability density by p_X. Verify that

$$I(X;Y) = \iint p_X(x) p_Z(y - x) \ln \frac{p_Z(y - x)}{p_Y(y)}\, dx dy \quad \text{nats.}$$

where p_Y is the probability density of the output when the input has density p_X.

(b) Now set $p_X = \mathcal{N}_P$. Verify that

$$\tfrac{1}{2} \ln\bigl(1 + P/\sigma^2\bigr) = \iint p_X(x) p_Z(y - x) \ln \frac{\mathcal{N}_{\sigma^2}(y - x)}{\mathcal{N}_{P+\sigma^2}(y)}\, dx dy.$$

(c) Still with $p_X = \mathcal{N}_P$, show that

$$\tfrac{1}{2} \ln(1 + P/\sigma^2) - I(X;Y) \leq 0.$$

[Hint: use (a) and (b) and $\ln t \leq t - 1$.]

(d) Show that an additive noise channel with noise variance σ^2 and input power P has capacity at least $\frac{1}{2}\log_2(1 + P/\sigma^2)$ bits per channel use. Conclude that the Gaussian channel has the smallest capacity among all additive noise channels of a given noise variance.