ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 19 Information Theory and Coding

Solutions to Homework 9 November 27, 2011

PROBLEM 1. From symmetry, the capacity achieving distribution has to be $p(1) = p(-1) = \alpha$, and $p(0) = 1 - 2\alpha$ for some α. The cost constraint translates to

$$p(1) + p(-1) = 2\alpha \leq \beta.$$

Computing $I(X;Y)$ we get that

$$I(X;Y) = 2\alpha.$$

So, we want that α is the largest possible:

$$\alpha = \begin{cases} \beta/2, & \text{if } \beta \leq 1 \\ 1/2, & \text{if } \beta > 1 \end{cases}.$$

Hence, the capacity is

$$C = \begin{cases} \beta, & \text{if } \beta \leq 1 \\ 1, & \text{if } \beta > 1 \end{cases}.$$

PROBLEM 2.

(a) We know that log is concave, the sum of concave functions is concave again.

(b) The Kuhn-Tucker conditions are

$$\frac{\frac{1}{\sigma_i^2}}{1 + \frac{p_i}{\sigma_i^2}} = \mu, \text{ for } p_i > 0,$$
$$\frac{\frac{1}{\sigma_i^2}}{1 + \frac{p_i}{\sigma_i^2}} \leq \mu, \text{ for } p_i = 0,$$

With rearranging the terms we can get the desired form with $\lambda = \frac{1}{\mu}$.

(c) It directly follows from part (b).

(d) Water-filling: start with $\lambda = 0$ and start increasing λ until $\sum_i p_i = 1$.

PROBLEM 3. Since X and Z are both in the interval $[-1, 1]$, their sum $X + Z$ lies in the interval $[-2, +2]$. If we could *choose* the distribution of $X + Z$ as we wished (without the constraint that it has to be the sum of two independent random variables, one of which is uniform) we would have chosen it to be uniform on the interval $[-2, +2]$ to have the largest entropy. Observe now that if we choose X as the random variable that equals $+1$ with probability $1/2$ and -1 with probability $1/2$, then $X + Z$ is uniform in $[-2, +2]$ and thus this distribution maximizes the entropy. An alternate derivation is as follows: note that since X and Z are independent, the moment generating functions of the random variables involved satisfy $E[e^{s(X+Z)}] = E[e^{sX}]E[e^{sX}]$. Now, we know that $E[e^{sZ}] = \int e^{sz} f_Z(z)\, dz = \int_{-1}^{+1} \frac{1}{2} e^{sz}\, dz = [e^s - e^{-s}]/(2s)$. Similarly, if we want $X + Z$ to be uniform on $[-2, 2]$, we can compute $E[e^{s(X+Z)}] = [e^{2s} - e^{-2s}]/(4s)$. This then requires $E[e^{sX}] = \frac{1}{2}[e^{2s} - e^{-2s}]/[e^s - e^{-s}] =$

$\frac{1}{2}[e^s + e^{-s}]$ which is the moment generating function of a random variable which takes on the values $+1$ and -1, each with probability $1/2$.

Similarly, under the constraint XZ lies in the interval $[-1, +1]$, and the best we could hope is that XZ is uniform on this interval. But this can be achieved by making sure that X only takes on the values $+1$ or -1.

PROBLEM 4.

$$\begin{aligned}
h(X) &= \frac{1}{2}\log(2\pi e\sigma_x^2) \\
h(Y) &= \frac{1}{2}\log(2\pi e\sigma_y^2) \\
h(X,Y) &= \frac{1}{2}\log((2\pi e)^2\det(K)) = \frac{1}{2}\log((2\pi e)^2(\sigma_x^2\sigma_y^2 - \rho^2\sigma_x^2\sigma_y^2) \\
I(X,Y) &= h(X) + h(Y) - h(X,Y) = \frac{1}{2}\log\frac{1}{1-\rho^2}
\end{aligned}$$

Note that the result does not depend on σ_x, σ_y, which says that normalization does not change the mutual information.

PROBLEM 5. *First Method:*

(a) It suffices to note that $H(X|Y) = H(X + f(Y)|Y)$ for any function f.

(b) Since among all random variables with a given variance the gaussian maximizes the entropy, we have

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

(c) From (a) and (b) we have

$$\begin{aligned}
I(X;Y) &= H(X) - H(X - \alpha Y|Y) \\
&\geq H(X) - H(X - \alpha Y) \\
&\geq H(X) - \frac{1}{2}\log 2\pi e E((X-\alpha Y)^2) \ .
\end{aligned}$$

(d) We have that $\frac{dE((X-\alpha Y)^2)}{d\alpha} = 0$ is equivalent to $E(Y(X - \alpha Y)) = 0$. Hence $\frac{dE((X-\alpha Y)^2)}{d\alpha}$ is equal to zero for $\alpha = \alpha^* = \frac{E(XY)}{E(Y^2)}$. Now on the one hand $E(XY) = E(X(X+Z)) = E(X^2)+E(XZ)$ and because of the independence between X and Z and the fact that Z has zero mean we have that $E(XZ) = 0$, and hence $E(XY) = P$. On the other hand $E(Y^2) = E((X+Z)^2) = E(X^2) + 2E(XZ) + E(Z^2) = P + 0 + \sigma^2$. Therefore $\alpha^* = P/(P + \sigma^2)$.

Then observing that $E((X-\alpha Y)^2)$ is a convex function of α we deduce that $E((X - \alpha Y)^2)$ is minimized for $\alpha = \alpha^*$. Finally an easy computation yields to $E((X - \alpha^* Y)^2) = \frac{\sigma^2 P}{\sigma^2 + P}$.

(e) Since X is gaussian from (c) and (d) we deduce that

$$\begin{aligned}
I(X;Y) &\geq \frac{1}{2}\log 2\pi e P - \frac{1}{2}\log 2\pi e\frac{\sigma^2 P}{\sigma^2 + P} \\
&= \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right) \ . \qquad (1)
\end{aligned}$$

with equality if and only if Z is gaussian with covariance σ^2.

Second Method:

(a) This is by the definition of mutual information once we note that $p_{Y|X}(y|x) = p_Z(y-x)$.

(b) Note that $p_X(x)p_Z(y-x)$ is simply the joint distribution of (x, y), and thus the integral

$$\iint p_X(x)p_Z(y-x)\ln\frac{\mathcal{N}_{\sigma^2}(y-x)}{\mathcal{N}_{\sigma^2+P}(y)}\,dxdy.$$

is an expectation, namely

$$E\ln\frac{\mathcal{N}_{\sigma^2}(Y-X)}{\mathcal{N}_{\sigma^2+P}(Y)}.$$

Substituting the formula for $\mathcal{N}$, this in turn, is

$$\begin{aligned}
&E\ln\frac{\mathcal{N}_{\sigma^2}(Y-X)}{\mathcal{N}_{\sigma^2+P}(Y)} \\
&\quad= \frac{1}{2}\ln\left(1+P/\sigma^2\right) + \frac{1}{2(\sigma^2+P)}E[Y^2] - \frac{1}{2\sigma^2}E[(Y-X)^2] \\
&\quad= \frac{1}{2}\ln\left(1+P/\sigma^2\right) + \frac{1}{2(\sigma^2+P)}E[(X+Z)^2] - \frac{1}{2\sigma^2}E[Z^2] \\
&\quad= \frac{1}{2}\ln\left(1+P/\sigma^2\right) + \frac{1}{2(\sigma^2+P)}E[X^2+Z^2+2XZ] - \frac{1}{2} \\
&\quad= \frac{1}{2}\ln\left(1+P/\sigma^2\right) + \frac{1}{2(\sigma^2+P)}(P+\sigma^2+0) - \frac{1}{2} \\
&\quad= \frac{1}{2}\ln\left(1+P/\sigma^2\right)
\end{aligned}$$

(c) The steps we need to justify read

$$\begin{aligned}
\ln(1+P/\sigma^2) - I(X;Y) &= \iint p_X(x)p_Z(y-x)\ln\frac{\mathcal{N}_{\sigma^2}(y-x)p_Y(y)}{\mathcal{N}_{\sigma^2+P}(y)p_Z(y-x)}\,dxdy \\
&\le \iint \frac{p_X(x)\mathcal{N}_{\sigma^2}(y-x)p_Y(y)}{\mathcal{N}_{\sigma^2+P}(y)}\,dxdy - 1 \\
&= \int p_Y(y)\,dy - 1 \\
&= 0.
\end{aligned}$$

The first equality is by substitution of parts (a) and (b). The inequality is by $\ln(x) \le x-1$. The next equality is by noting that

$$\int p_X(x)\mathcal{N}_{\sigma^2}(y-x)dx = (p_X * \mathcal{N}_{\sigma^2})(y) = (\mathcal{N}_P * \mathcal{N}_{\sigma^2})(y) = \mathcal{N}_{P+\sigma^2}(y).$$

The last equality is because any density function integrates to 1.

(d) The conclusion is made by noting that the right hand side of the first equality in (c) is equal to zero if $p_Z = \mathcal{N}_{\sigma^2}$.