S-72.340 INFORMATION THEORY

Exercise 10, November 22, 2000 Solutions

1. The transition probabilities $P(y|x)$ of the Z channel are $P(0|0) = 1 - \epsilon$, $P(1|0) = \epsilon$ and $P(1|1) = 1$. The upper bound for $\bar{P}_{e,m}$ is given by

$$\begin{aligned}
&\left\{ \sum_{j=0}^{1} \left(\sum_{k=0}^{1} Q(k) \sqrt{P(j|k)} \right)^2 \right\}^N \\
&= \left\{ \left(Q(0)\sqrt{1-\epsilon} \right)^2 + \left(Q(0)\sqrt{\epsilon} + Q(1) \right)^2 \right\}^N \\
&= \left\{ Q(0)^2 + Q(1)^2 + 2Q(0)\,Q(1)\sqrt{\epsilon} \right\}^N \\
&= \left\{ 2(1-\sqrt{\epsilon})Q(0)^2 - 2(1-\sqrt{\epsilon})Q(0) + 1 \right\}^N .
\end{aligned}$$

This expression is clearly minimized when the function of $Q(0)$ inside the curly braces is minimized. This happens when its derivative is 0. This leads to

$$4(1-\sqrt{\epsilon})Q(0) - 2(1-\sqrt{\epsilon}) = 0 \Rightarrow Q(0) = \frac{1}{2},$$

and the tightest upper bound for $\bar{P}_{e,m}$ is

$$\left\{ 2(1-\sqrt{\epsilon}) \cdot \frac{1}{4} - 2(1-\sqrt{\epsilon}) \cdot \frac{1}{2} + 1 \right\}^N = \left(\frac{1+\sqrt{\epsilon}}{2} \right)^N .$$

2. The transition probabilities $P(y|x)$ of the binary erasure channel are given by

$$P(0|0) = P(1|1) = 1 - \epsilon \quad \text{and} \quad P(E|0) = P(E|1) = \epsilon\,.$$

The upper bound for $\bar{P}_{e,m}$ is given by

$$\begin{aligned}
&\left\{\sum_j \left(\sum_{k=0}^{1} Q(k)\sqrt{P(j|k)}\right)^2\right\}^N \\
&= \left\{\left(Q(0)^2 + Q(1)^2\right)(1-\epsilon) + \epsilon\right\}^N \\
&= \left\{2(1-\epsilon)\left(Q(0)^2 - Q(0)\right) + 1\right\}^N .
\end{aligned}$$

In the same way as in Problem 1, we find that the bound is minimized when $Q(0) = Q(1) = 1/2$, and the minimal upper bound is now equal to

$$\left(\frac{1+\epsilon}{2}\right)^N .$$

3. The function $E_0(\rho, \mathbf{Q})$ is given by

$$\begin{aligned}
E_0(\rho, \mathbf{Q}) &= -\ln \sum_j \left[\sum_k Q(k) P(j|k)^{1/(1+\rho)}\right]^{1+\rho} \\
&= -\ln\left[Q(0)^{1+\rho}(1-\epsilon) + \epsilon + Q(1)^{1+\rho}(1-\epsilon)\right] \\
&= -\ln\left\{(1-\epsilon)\left[Q(0)^{1+\rho} + (1-Q(0))^{1+\rho}\right] + \epsilon\right\}.
\end{aligned}$$

Assume that ρ has some fixed value in $[0,1]$. Then, $E_0(\rho, \mathbf{Q})$ is maximized when the expression

$$Q(0)^{1+\rho} + (1-Q(0))^{1+\rho}$$

is minimized over the interval $0 \leq Q(0) \leq 1$. First, we notice that at 0 and 1 the function is equal to 1. By setting the derivative as 0, we get

$$(1+\rho)Q(0)^{\rho} - (1+\rho)(1-Q(0))^{\rho} = 0 \Rightarrow Q(0) = 1 - Q(0) \Rightarrow Q(0) = \frac{1}{2}\,.$$

The value of the function at $Q(0) = 1/2$ is $(1/2)^{\rho} \leq 1$ when $\rho \in [0,1]$. Hence, no matter what the precise value of ρ is, $\hat{\mathbf{Q}} = (1/2, 1/2)$ is always the distribution that maximizes $E_0(\rho, \mathbf{Q})$.

The random coding exponent is given by

$$\begin{aligned} E_r(R) &= \max_{0\leq\rho\leq 1} \max_{\mathbf{Q}} [E_0(\rho, \mathbf{Q}) - \rho R] \\ &= \max_{0\leq\rho\leq 1} [E_0(\rho, \hat{\mathbf{Q}}) - \rho R] \,. \end{aligned}$$

However, it seems next to impossible to solve $E_r(R)$ explicitly, and we must settle for a parametric presentation, where both R and $E_r(R)$ are given as functions of the parameter $\rho \in [0,1]$. The derivation begins from the function

$$E_0(\rho, \hat{\mathbf{Q}}) = -\ln\left[(1-\epsilon)\cdot\left(\frac{1}{2}\right)^{\rho} + \epsilon\right] .$$

Now we get

$$R = \frac{\partial E_0(\rho, \hat{\mathbf{Q}})}{\partial \rho} = \frac{\log 2 \cdot (1-\epsilon)(1/2)^{\rho}}{(1-\epsilon)(1/2)^{\rho} + \epsilon}$$

and

$$\begin{aligned} E_r(R) &= E_0(\rho, \hat{\mathbf{Q}}) - \rho \frac{\partial E_0(\rho, \hat{\mathbf{Q}})}{\partial \rho} \\ &= -\ln\left[(1-\epsilon)\cdot\left(\frac{1}{2}\right)^{\rho} + \epsilon\right] - \frac{\log 2 \cdot (1-\epsilon)(1/2)^{\rho}}{(1-\epsilon)(1/2)^{\rho} + \epsilon} \,. \end{aligned}$$

Figure 1 shows $E_r(R)$ as a function of R when $\epsilon = 0.1$.

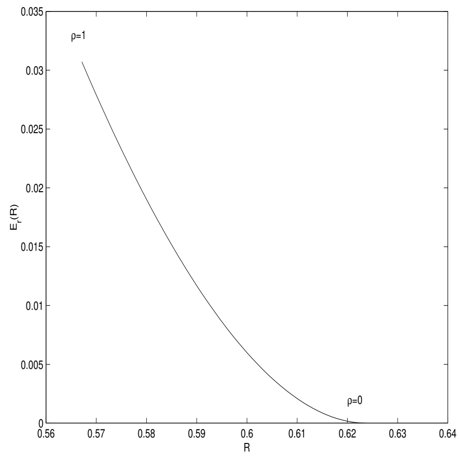

Figure 1: The random coding exponent for a binary erasure channel

4. The function $E_0(\rho, \mathbf{Q})$ is given by

$$E_0(\rho, \mathbf{Q}) = -\ln \sum_{j=0}^{4} \left[\sum_{k=0}^{4} Q(k) P(j|k)^{1/(1+\rho)} \right]^{1+\rho}$$

$$\begin{aligned} = -\ln \Bigg[& \frac{1}{2}(Q(0)+Q(4))^{1+\rho} + \frac{1}{2}(Q(0)+Q(1))^{1+\rho} \\ & + \frac{1}{2}(Q(1)+Q(2))^{1+\rho} + \frac{1}{2}(Q(2)+Q(3))^{1+\rho} \\ & + \frac{1}{2}(Q(3)+Q(4))^{1+\rho} \Bigg] . \end{aligned}$$

Let us again first maximize $E_0(\rho, \mathbf{Q})$ over $\mathbf{Q}$. This is equivalent to minimizing the function

$$\begin{aligned} g(\rho, \mathbf{Q}) = & (Q(0)+Q(4))^{1+\rho} + (Q(0)+Q(1))^{1+\rho} \\ & + (Q(1)+Q(2))^{1+\rho} + (Q(2)+Q(3))^{1+\rho} \\ & + (Q(3)+Q(4))^{1+\rho} . \end{aligned}$$

This is a convex function of $\mathbf{Q}$, since it is a sum of convex functions of linear functions of $\mathbf{Q}$. Hence, the necessary and sufficient conditions on the probability vector $\mathbf{Q}$ to minimize $g(\rho, \mathbf{Q})$ are:

$$\frac{\partial g}{\partial Q(k)} \geq \lambda$$

for all k, with equality if $Q(k) > 0$. For example,

$$\frac{\partial g}{\partial Q(0)} = (1+\rho)(Q(0)+Q(4))^{\rho} + (1+\rho)(Q(0)+Q(1))^{\rho} .$$

From this set of equations, it follows that

$$Q(0) = Q(1) = \cdots = Q(4) = \frac{1}{5}$$

is the probability vector $\hat{\mathbf{Q}}$ that maximizes $E_0(\rho, \mathbf{Q})$, and

$$\begin{aligned} E_r(R) &= \max_{0 \leq \rho \leq 1} [E_0(\rho, \hat{\mathbf{Q}}) - \rho R] = \max_{0 \leq \rho \leq 1} [-\ln \left(\frac{2}{5}\right)^{\rho} - \rho R] \\ &= \max_{0 \leq \rho \leq 1} \left[(\ln \frac{5}{2} - R)\rho \right] = \ln \frac{5}{2} - R , \end{aligned}$$

when $R < \ln(5/2)$.

5. Consider an arbitrary DMC with a binary input, where the input symbol is 0 or 1, and the output symbol varies in the range $0, 1, \ldots, J-1$. If $\rho = 1$, we get

$$\begin{aligned} E_0(1, \mathbf{Q}) &= -\ln \sum_{j=0}^{J-1} \left[\sum_{k=0}^{1} Q(k) P(j|k)^{1/2} \right]^2 \\ &= -\ln \sum_{j=0}^{J-1} \left[Q(0)\sqrt{P(j|0)} + [1 - Q(0)]\sqrt{P(j|1)} \right]^2 \\ &= -\ln f(Q(0)) \,. \end{aligned}$$

Clearly, $E_0(1, \mathbf{Q})$ is maximized when $f(Q(0))$ is minimized.

$$\begin{aligned} f(Q(0)) &= \left\{ \sum_{j=0}^{J-1} \left[P(j|0) - 2\sqrt{P(j|0)P(j|1)} + P(j|1) \right] \right\} Q(0)^2 \\ &\quad + \left\{ \sum_{j=0}^{J-1} \left[2\sqrt{P(j|0)P(j|1)} - 2P(j|1) \right] \right\} Q(0) + P(j|1) \\ &= \left[2 - 2\sum_{j=0}^{J-1} \sqrt{P(j|0)P(j|1)} \right] Q(0)^2 \\ &\quad + \left[2\sum_{j=0}^{J-1} \sqrt{P(j|0)P(j|1)} - 2 \right] Q(0) + P(j|1) \,. \end{aligned}$$

The function $f(Q(0))$ is clearly convex, since the second degree coefficient is positive. Hence, it has a (global) minimum at the point where its derivative is zero. By differentiating with respect to $Q(0)$ and setting the derivative as zero, we get

$$Q(0) = Q(1) = \frac{1}{2} \,.$$