ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 24 — Information Theory and Coding

Solutions to homework 10 — December 10, 2013

PROBLEM 1.

(a)

$$\begin{aligned} I(X;Y) &= I(X_k,K;Y_k,K) = I(K;Y_k,K) + I(X_k;Y_K,K|K) = H(K) + I(X_k;Y_k|K) \\ &= h_2(\alpha) + \mathbb{P}_K[1].I(X_k;Y_k|K=1) + \mathbb{P}_K[2]I(X_k;Y_k|K=2) \\ &= h_2(\alpha) + \alpha.I(X_1;Y_1) + (1-\alpha)I(X_2;Y_2) \end{aligned}$$

(b) The distribution of X is determined by α and by the distributions of X_1 and X_2. It is clear from the expression in (a) that for any given α, $I(X;Y)$ is maximized when $I(X_1;Y_1)$ and $I(X_2;Y_2)$ are maximized, i.e., when the distribution of X_1 (resp. X_2) achieves the capacity of P_1 (resp. P_2). We conclude that the value of α in the capacity achieving distribution is the one that maximizes the function $f(\alpha) = h_2(\alpha) + \alpha C_1 + (1-\alpha)C_2$. The derivative of f is:

$$f'(\alpha) = -\log_2(\alpha) - \frac{1}{\ln 2} + \log_2(1-\alpha) + \frac{1}{\ln 2} + C_1 - C_2 = C_1 - C_2 + \log_2\frac{1-\alpha}{\alpha}.$$

We have $f'(\alpha) = 0$ (resp. $f'(\alpha) > 0$, $f'(\alpha) < 0$) if $\alpha = \alpha^*$ (resp. $\alpha < \alpha^*$, $\alpha > \alpha^*$), where $\alpha^* = \dfrac{2^{C_1}}{2^{C_1}+2^{C_2}}$. This means that $f(\alpha)$ is maximized at $\alpha = \alpha^*$. Therefore, the capacity achieving distribution is such that $\alpha = \dfrac{2^{C_1}}{2^{C_1}+2^{C_2}}$ and X_1 (resp. X_2) achieves the capacity of the channel P_1 (resp. P_2).

(c) From (b), we have:

$$\begin{aligned} C &= -\frac{2^{C_1}}{2^{C_1}+2^{C_2}}\log_2\frac{2^{C_1}}{2^{C_1}+2^{C_2}} - \frac{2^{C_2}}{2^{C_1}+2^{C_2}}\log_2\frac{2^{C_2}}{2^{C_1}+2^{C_2}} + \frac{2^{C_1}C_1}{2^{C_1}+2^{C_2}} + \frac{2^{C_2}C_2}{2^{C_1}+2^{C_2}} \\ &= -\frac{2^{C_1}}{2^{C_1}+2^{C_2}}C_1 + \frac{2^{C_1}}{2^{C_1}+2^{C_2}}\log_2(2^{C_1}+2^{C_2}) - \frac{2^{C_2}}{2^{C_1}+2^{C_2}}C_2 \\ &\quad + \frac{2^{C_2}}{2^{C_1}+2^{C_2}}\log_2(2^{C_1}+2^{C_2}) + \frac{2^{C_1}C_1}{2^{C_1}+2^{C_2}} + \frac{2^{C_2}C_2}{2^{C_1}+2^{C_2}} \\ &= \log_2(2^{C_1}+2^{C_2}). \end{aligned}$$

Therefore, $2^C = 2^{C_1} + 2^{C_2}$.

PROBLEM 2.

(a)

$$\begin{aligned} F(p,r_p) - F(p,r) &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x)P(y|x)\log_2\frac{r_p(x|y)}{r(x|y)} \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x)P(y|x)\log_2\frac{p(x)P(x|y)}{r(x|y)\sum_{x'\in\mathcal{X}}p(x')P(y|x')} \\ &= D(P_1||P_2) \geq 0, \end{aligned}$$

where $P_1(x,y) := p(x)P(y|x)$ and $P_2(x,y) := r(x|y)\sum_{x'\in\mathcal{X}}p(x')P(y|x')$.

(b) We can rewrite $F(p, r)$ as follows:

$$F(p,r) = \Big(\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x)P(y|x)\log_2 r(x|y)\Big) + \Big(\sum_{x\in\mathcal{X}} p(x)\log_2\frac{1}{p(x)}\Big). \quad (1)$$

The first term in 1 is linear in p while the second term is strictly concave in p (since the function $t \longrightarrow t\log_2\frac{1}{t}$ is strictly concave). Therefore, $F(p, r)$ is strictly concave in p.

The first term in 1 is concave in r (since the function $\log_2$ is concave) and the second term is constant with respect to r. Therefore, $F(p, r)$ is concave in r.

(c) For every $x \in \mathcal{X}$, we have:

$$\frac{\partial F(p,r_k)}{\partial p(x)} = \sum_{y\in\mathcal{Y}} P(y|x)\log_2 r_k(x|y) + \log_2\frac{1}{p(x)} - \frac{1}{\ln 2}.$$

A probability distribution p satisfies the Kuhn-Tucker conditions if and only if there exists a real number λ such that for all $x \in \mathcal{X}$, we have $\frac{\partial F(p,r_k)}{\partial p(x)} \leq \lambda$ with equality if $p(x) > 0$. Therefore, for all $x \in \mathcal{X}$ we have:

$$\sum_{y\in\mathcal{Y}} P(y|x)\log_2 r_k(x|y) - \log_2(p(x)) \leq \lambda',$$

where $\lambda' = \lambda + \frac{1}{\ln 2}$. This shows that

$$p(x) \geq 2^{-\lambda'}\alpha_k(x).$$

If $p(x) > 0$, we have $p(x) = 2^{-\lambda'}\alpha_k(x)$, and if $p(x) = 0$ we must also have $p(x) = 2^{-\lambda'}\alpha_k(x) = 0$ since $2^{-\lambda'}2^{\sum_{y\in\mathcal{Y}} P(y|x)\log_2 r_k(x|y)} \geq 0$. We conclude that $p(x) = 2^{-\lambda'}\alpha_k(x)$ in all cases. Therefore, $1 = 2^{-\lambda'}\sum_{x\in\mathcal{X}}\alpha_k(x)$, and $\lambda' = \log_2\sum_{x\in\mathcal{X}}\alpha_k(x)$. We conclude that the only distribution that satisfies the Kuhn-Tucker conditions is the one given by $p(x) = \frac{\alpha_k(x)}{\sum_{x'\in\mathcal{X}}\alpha_k(x')}$. On the other hand, the fact that $F(p, r_k)$ is concave in p shows that it admits a maximum p_{k+1}, which has to satisfy the Kuhn-Tucker conditions. Therefore, $p_{k+1}(x) = \frac{\alpha_k(x)}{\sum_{x'\in\mathcal{X}}\alpha_k(x')}$.

(d) $C \geq F(p_{k+1}, r_{k+1})$ since $F(p_{k+1}, r_{k+1}) = I(X;Y)|_{p_X=p_{k+1}}$. This implies that $C \geq F(p_{k+1}, r_k)$ since $F(p_{k+1}, r_{k+1}) \geq F(p_{k+1}, r_k)$. On the other hand, we have

$$\begin{aligned}
&F(p_{k+1}, r_k)\\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\frac{\alpha_k(x)}{\sum_{x'\in\mathcal{X}}\alpha_k(x')}P(y|x)\log_2\frac{r_k(x|y)\sum_{x'\in\mathcal{X}}\alpha_k(x')}{\alpha_k(x)}\\
&= \sum_{x\in\mathcal{X}}\frac{\alpha_k(x)}{\sum_{x'\in\mathcal{X}}\alpha_k(x')}\Big[\sum_{y\in\mathcal{Y}}P(y|x)\log_2 r_k(x|y) - \log_2(\alpha_k(x)) + \log_2\sum_{x'\in\mathcal{X}}\alpha_k(x')\Big]\\
&= \log_2\sum_{x'\in\mathcal{X}}\alpha_k(x') + \sum_{x\in\mathcal{X}}\frac{\alpha_k(x)}{\sum_{x'\in\mathcal{X}}\alpha_k(x')}\Big[\log_2(\alpha_k(x)) - \log_2(\alpha_k(x))\Big]\\
&= \log_2\sum_{x\in\mathcal{X}}\alpha_k(x).
\end{aligned}$$

(e)

$$\log_2 \frac{\alpha_k(x)}{p_k(x)} = \log_2 \alpha_k(x) - \log_2 p_k(x) = \sum_{y\in\mathcal{Y}} P(y|x) \log_2 r_k(x|y) - \log_2 p_k(x)$$
$$= \sum_{y\in\mathcal{Y}} P(y|x) \log_2 \frac{r_k(x|y)}{p_k(x)} = \sum_{y\in\mathcal{Y}} P(y|x) \log_2 \frac{P(y|x)}{\sum_{x'\in\mathcal{X}} p_k(x')P(y|x')}.$$

(f) Given that $\log_2 \frac{\alpha_k(x)}{p_k(x)} = \sum_{y\in\mathcal{Y}} P(y|x) \log_2 \frac{P(y|x)}{\sum_{x'\in\mathcal{X}} p_k(x')P(y|x')}$, the inequality $C \leq \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{\alpha_k(x)}{p_k(x)}$ is a direct application of homework 8 problem 4.

(g) From (d) and (f), we have:

$$C - F(p_{k+1}, r_k)$$
$$\leq \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{\alpha_k(x)}{p_k(x)} - \log_2 \sum_{x\in\mathcal{X}} \alpha_k(x) = \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{\alpha_k(x)}{p_k(x)} - \log_2 \sum_{x'\in\mathcal{X}} \alpha_k(x')$$
$$= \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{\alpha_k(x)}{p_k(x) \sum_{x'\in\mathcal{X}} \alpha_k(x')} = \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{p_{k+1}(x)}{p_k(x)} \leq \max_{x\in\mathcal{X}} \log_2 \frac{p_{k+1}(x)}{p_k(x)}.$$

(h) We prove it by induction on n. The result is trivial for $n = 0$. Now assume that it is true for n, and let us prove it for $n + 1$:

$$\sum_{k=0}^{n+1} (C - F(p_{k+1}, r_k)) = C - F(p_{n+2}, r_{n+1}) + \sum_{k=0}^{n} (C - F(p_{k+1}, r_k))$$
$$\leq \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{p_{n+2}(x)}{p_{n+1}(x)} + \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{p_{n+1}(x)}{p_0(x)}$$
$$= \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{p_{n+2}(x)}{p_0(x)}.$$

On the other hand, since $p_{n+1}(x) \leq 1$ for all $x \in \mathcal{X}$, we have:

$$\sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{p_{n+1}(x)}{p_0(x)} \leq \sum_{x\in\mathcal{X}} p^*(x) \log_2 \frac{1}{1/|\mathcal{X}|} = \log_2 |\mathcal{X}|.$$

(i) The sequence $s_n = \sum_{k=0}^{n} C - F(p_{k+1}, r_k)$ is increasing and upper-bounded, thus convergent, which implies that the sequence $C - F(p_{k+1}, r_k) = s_k - s_{k-1}$ converges to zero. Therefore, $F(p_{k+1}, r_k)$ converges to C.