TS226 - Théorie de l'information
année 2019/2020

Romain Tajan

1 Canal à effacements binaires

On s'intéresse, dans cet exercice au canal à effacement binaire donné en Figure 1.

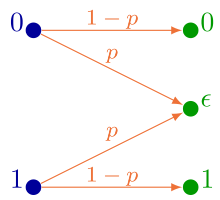

FIGURE 1 – Canal à effacements binaires

Question 1. Montrer que la capacité de ce canal est $C(p) = 1 - p$. Dans cette question, nous noterons $\mathbb{P}(X = 1) = q$ et donc $\mathbb{P}(X = 0) = 1 - q$. Aussi, nous noterons $h_2(p)$ la fonction $h_2(p) = -p\log_2(p) - (1-p)\log_2(1-p)$.

Par définition (voir cours de théorie de l'information) nous avons

$$\mathbb{I}(X;Y) = \mathbb{H}(X) - \mathbb{H}(X|Y)$$

avec $\mathbb{H}(X|Y) = \sum_{y\in\{0,\epsilon,1\}} p(y)\mathbb{H}(X|Y=y)$ et $\mathbb{H}(X|Y=y) = -\sum_{x\in\{0,1\}} p(x|y)\log_2 p(x|y)$.

Les termes $p_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y)$ peuvent être calculés en utilisant la règle de Bayes :

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{p_{Y|X}(y|x)p_X(x)}{p_Y(y)}$$

- **Calcul de $p_{X|Y}(0|0)$ utilisant la règle précédente** : nous avons $p_{Y|X}(0|0) = 1-p$, $p_X(0) = 1-q$, de plus $p_Y(0) = p_{Y|X}(0|0)p_X(0) + p_{Y|X}(0|1)p_X(1) = (1-p)(1-q)$ d'où finalement
$$p_{X|Y}(0|0) = 1$$
- **Calcul de $p_{X|Y}(0|\epsilon)$ utilisant la règle précédente** : nous avons $p_{Y|X}(\epsilon|0) = p$, $p_X(0) = 1-q$ de plus $p_Y(\epsilon) = p_{Y|X}(\epsilon|0)p_X(0) + p_{Y|X}(\epsilon|1)p_X(1) = p(1-q) + pq = p$ d'où finalement
$$p_{X|Y}(0|\epsilon) = 1 - q$$
.

- **Calcul de $p_{X|Y}(0|1)$ utilisant la règle précédente** : nous avons $p_{Y|X}(1|0) = 0$, d'où finalement
$$p_{X|Y}(0|1) = 0$$

Nous obtenons donc les entropies suivantes :
- $H(X|Y = 0) = h_2(p_{X|Y}(0|0)) = h_2(1) = 0$
- $H(X|Y = \epsilon) = h_2(p_{X|Y}(0|\epsilon)) = h_2(1 - q) = h_2(q)$
- $H(X|Y = 1) = h_2(p_{X|Y}(0|1)) = h_2(0) = 0$

Finalement nous obtenons l'information mutuelle comme

$$\mathbb{I}(X;Y) = \mathbb{H}(X) - \mathbb{H}(X|Y) = h_2(q) - p_Y(\epsilon)H(X|Y = \epsilon) = (1 - p)h_2(q)$$

En étudiant la fonction $q \mapsto h_2(q)$ nous avons que cette fonction est maximale pour $q = 1/2$ et que $h_2(0.5) = 1$. La capacité est donc $C(p) = 1 - p$.

Question 2. Pour chaque valeur de p, donner la distribution d'entrée permettant d'atteindre $C(p)$. Cette valeur de capacité est atteinte pour $q = 1/2$ quelque soit p.

Question 3. Pour quelle(s) valeurs de p a-t'on $C(p) = 0$? Nous avons $C(p) = 0$ si et seulement si $p = 1$.

2 Canal en Z

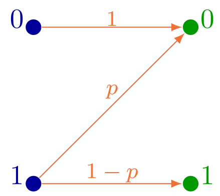

Figure 2 – Canal en Z

Considérons le canal à entrées/sorties binaires, dont les probabilités de transitions sont données dans la figure 2. Ce modèle de canal, dit canal en Z, est très utilisé dans le contexte des communications optiques.

Question 1. Donner l'expression de la capacité du canal en Z en fonction de p. On utilise la même démarche que pour le canal précédent à la différence près que maintenant nous utiliserons la formule
$$\mathbb{I}(X;Y) = \mathbb{H}(Y) - \mathbb{H}(Y|X)$$

Y étant une variable aléatoire binaire, $\mathbb{H}(Y) = h_2(p_Y(0))$ où $p_Y(0)$ provient de la loi des probabilités totales i.e. $p_Y(0) = p_{Y|X}(0|0)p_X(0) + p_{Y|X}(0|1)p_X(1) = (1 - q) + pq = 1 - (1 - p)q$, d'où
$$\mathbb{H}(Y) = h_2((1 - p)q)$$

De plus nous avons $\mathbb{H}(Y|X) = p_X(0)\mathbb{H}(Y|X = 0) + p_X(1)\mathbb{H}(Y|X = 1)$ d'où $\mathbb{H}(Y|X) = (1 - q)h_2(p_{Y|X}(0|0)) + qh_2(p_{Y|X}(0|1))$ et finalement $\mathbb{H}(Y|X) = qh_2(p)$. Ce qui donne

$$\mathbb{I}(X;Y) = h_2((1 - p)q) - qh_2(p)$$

La fonction h_2 étant une fonction continue sur $[0,1]$ le théorème de Weierstrass suffit à démontrer existence d'une valeur $q^\star$ maximisant $\mathbb{I}(X;Y)$, de plus

$$C(p) = h_2((1-p)q^\star) - q^\star h_2(p)$$

Question 2. Donner la distribution d'entrée atteignant la capacité pour $p = 1/2$. Ce résultat est-il intuitif, au vu de la structure du canal ?

Il nous reste à déterminer la distribution d'entrée $p_X(x)$ qui maximise $\mathbb{I}(X;Y)$. Pour cela, nous allons dériver $\mathbb{I}(X;Y)$ par rapport à q. On remarque que pour cela, $\frac{dh_2}{dx} = \log_2 \frac{1-x}{x}$. Nous obtenons donc

$$\frac{\partial \mathbb{I}(X;Y)}{\partial q} = (1-p)\log_2 \frac{1-(1-p)q}{(1-p)q} - h_2(p)$$

On en déduit de l'équation $\frac{\partial \mathbb{I}(X;Y)}{\partial q} = 0$ que

$$q^\star = \frac{1}{(1-p)(1+2^{h_2(p)/(1-p)})}$$

soit pour $p = 1/2$, $q^\star = \frac{2}{5}$.

Ce résultat est assez logique, en effet le symbole 0 étant envoyé sans erreur il est assez logique que la stratégie de codage optimale envoie plus de 0 que de 1.

3 Canal symétrique

Soit un canal $(\mathcal{X}, \mathcal{Y}, p_{Y|X}(y|x))$ tel que $\mathcal{X} = \{x_0, \dots x_{N_x-1}\}$ et $\mathcal{Y} = \{y_0, \dots y_{N_y-1}\}$. La matrice de transition T de ce canal est la matrice de taille $N_x \times N_y$ ayant pour terme général $T_{i,j} = p_{Y|X}(y_j|x_i)$. Un canal est dit symétrique si ça matrice de transition vérifie les propriétés suivantes :

1. Chaque ligne est obtenue comme une permutation des autres lignes.
2. Si toutes les sommes des colonnes sont égales.

Question 1. Montrer que pour un canal symétrique, la capacité majorée par

$$C \leq \log_2(N_y) - h(\mathbf{r})$$

où $\mathbf{r} = [r_0, r_1, \dots]$ est une ligne de la matrice de transition T et $h(\mathbf{r}) = -\sum_j r_j \log_2 r_j$.

Pour démontrer cette question on part de l'équation

$$\mathbb{I}(X;Y) = \mathbb{H}(Y) - \mathbb{H}(Y|X) = \mathbb{H}(Y) - \sum_{i=0}^{N_x-1} p(x_i)\mathbb{H}(Y|X=x_i)$$

où $\mathbb{H}(Y|X=x_i) = -\sum_{j=0}^{N_y-1} p_{Y|X}(y_j|x_i)\log_2(p_{Y|X}(y_j|x_i)) = h_2(\mathbf{r}_i)$ où $\mathbf{r}_i$ représente la $i^{\text{ème}}$ ligne de la matrice T. Pour un canal symétrique les lignes sont toutes des permutations de la première ligne r_0. Il suffit de remarquer que la fonction $h_2(\mathbf{r})$ est invariante par permutation des éléments de $\mathbf{r}$ pour conclure que

$$\mathbb{I}(X;Y) = \mathbb{H}(Y) - \sum_{i=0}^{N_x-1} p(x_i)h_2(\mathbf{r}_i) = \mathbb{H}(Y) - h_2(\mathbf{r}_0)\sum_{i=0}^{N_x-1} p(x_i) = \mathbb{H}(Y) - h_2(\mathbf{r}_0)$$

où la première égalité provient du fait que la ligne i est une permutation de la ligne 0.

Il convient de remarquer que l'identité précédemment démontrée est indépendante de la loi d'entrée $p_X(x)$. De plus $\mathcal{Y}$ étant un alphabet fini de taille N_y, $\mathbb{H} \leq \log_2 N_y$. Pour toute distribution $p_X(x)$ nous obtenons l'inégalité

$$\mathbb{I}(X;Y) \leq \log_2 N_y - h_2(\mathbf{r}_0)$$

et finalement

$$C = \sup_{p_X(x)} \mathbb{I}(X;Y) \leq \log_2 N_y - h_2(\mathbf{r}_0)$$

Question 2. Montrer que pour un canal symétrique, la capacité est

$$C = \log_2(N_y) - h(\mathbf{r})$$

pour une entrée uniforme.

Le cas d'égalité de l'inégalité précédente est obtenu si Y suit une loi uniforme. Or, pour une distribution d'entrée uniforme nous avons $\forall i \in \{0,\ldots,N_x-1\}, p_X(x_i) = N_x^{-1}$ donc

$$p_Y(y_j) = \sum_{i=0}^{N_x-1} p(x_i)p_{Y|X}(y_j|x_i) = \frac{1}{N_x}\sum_{i=0}^{N_x-1} p_{Y|X}(y_j|x_i) = \frac{A}{N_x}$$

où la dernière égalité est obtenue en utilisant la seconde hypothèse (les colonnes de T possèdent toutes la même somme) et où $A = \sum_{i=0}^{N_x-1} p_{Y|X}(y_0|x_i)$ est la somme de la première ligne. Étant donné que $\sum_j p_Y(y_j) = 1$, on en déduit que $A = \frac{N_x}{N_y}$ et que Y suit une loi uniforme.

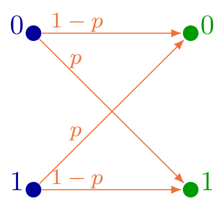

FIGURE 3 – Canal à binaire symétrique

Question 3. Le canal binaire symétrique donné en figure 3 est-il symétrique ?

Question 4. Montrer que le canal binaire symétrique donné en figure 3 est symétrique. La matrice T pour ce canal est la suivante

$$T = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

La somme des colonnes fait 1 et la seconde ligne est bien une permutation de la première ligne. Ce canal est donc symétrique.

Question 5. En déduire la capacité du canal binaire symétrique. Nous avons donc que la capacité pour ce canal est obtenue pour une entrée uniforme et de plus

$$C(p) = \log_2(2) - h([p, 1-p]) = 1 - h_2(p)$$

Question 6. Le canal en Z donné en figure 2 est-il symétrique ? La matrice T pour le canal en Z est la suivante

$$T = \begin{bmatrix} 1 & 0 \\ p & 1-p \end{bmatrix}$$

La seconde ligne n'est pas une permutation de la première suffit à démontrer que ce canal n'est pas symétrique.

Question 7. Le canal BEC donné en figure 1 est-il symétrique ? La matrice T pour le canal en BEC est la suivante

$$T = \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$

Le fait que la somme des colonnes n'est pas constante suffit à démontrer que ce canal n'est pas symétrique.

4 Canaux AWGN parallèles

On considère un cas général de canaux parallèle composé de K canaux gaussiens indépendants donnés dans la Figure 4 :

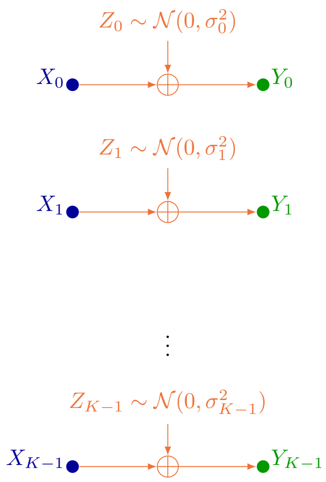

FIGURE 4 – Canaux gaussiens parallèles

où $Z_0, \ldots, Z_{K-1}$ sont indépendants et $Z_k \sim \mathcal{N}_{\mathbb{R}}(0, \sigma_k^2)$. En considérant l'entrée comme le vecteur $\mathbf{X} = (X_0, \ldots, X_{K-1})$ à composantes décorrélées, et la sortie comme $\mathbf{Y} = (Y_0, \ldots, Y_{K-1})$, la capacité d'un tel canal, pour des puissances d'entrées fixées $P_0, \ldots, P_{K-1}$, est donnée par

$$C = \sup_{p(\mathbf{x}) : \mathbb{V}(X_k) \le P_k} \mathbb{I}(\mathbf{X}, \mathbf{Y})$$

où le supremum est pris sur toutes les densités de probabilité sur $\mathbb{R}^K$ telles que $\mathbb{V}(X_k) \le P_k$ pour $k = 0, \ldots K-1$. Calculer explicitement cette capacité, en précisant une distribution d'entrée qui permet de l'atteindre.

Pour ce canal d'entrée et de sortie dans $\mathbb{R}^K$ nous avons

$$\mathbb{I}(\mathbf{Y}; \mathbf{X}) = \mathbb{H}(\mathbf{Y}) - \mathbb{H}(\mathbf{Y}|\mathbf{X})$$

où $\mathbb{H}(\mathbf{Y}|\mathbf{X}) = \int_{\mathbb{R}^K} \mathbb{H}(\mathbf{Y}|\mathbf{x}) p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$ avec $\mathbb{H}(\mathbf{Y}|\mathbf{x}) = -\int_{\mathbb{R}^K} p_{Y|X}(\mathbf{y}|\mathbf{x}) \log p_{Y|X}(\mathbf{y}|\mathbf{x}) d\mathbf{y}$. Les variables aléatoire Z_k étant indépendantes

$$p_{Y|X}(\mathbf{y}|\mathbf{x}) = p_Z(\mathbf{y} - \mathbf{x}) = \prod_{k=0}^{K-1} p_{Z_k}(y_k - x_k) = \prod_{k=0}^{K-1} \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{1}{2\sigma_k^2}(y_k - x_k)^2}$$

En généralisant la propriété de l'entropie vue en cours pour les couples de variables aléatoires indépendantes aux vecteurs de variables aléatoires indépendantes, on a $H(\mathbf{X}) = \sum_{k=0}^{K-1} H(X_k)$

$$\mathbb{H}(\mathbf{Y}|\mathbf{x}) = \sum_{k=0}^{K-1} \frac{1}{2} \log 2\pi e \sigma_k^2 = \mathbb{H}(\mathbf{Y}|\mathbf{X})$$

la dernière égalité étant obtenue car le terme central ne dépend pas de $\mathbf{x}$.

En généralisant la majoration de l'entropie jointe par la somme des entropies marginales, nous obtenons

$$\mathbb{H}(\mathbf{Y}) \leq \sum_{k=0}^{K-1} \mathbb{H}(Y_k)$$

ou le cas d'égalité est obtenu si et seulement si les variables aléatoires Y_k sont indépendantes. L'entropie de chaque $H(Y_k)$ peut maintenant être majorée par l'entropie d'une variable aléatoire Gaussienne de même variance que Y_k (notée $\sigma_{y_k}^2$)

$$\mathbb{H}(\mathbf{Y}) \leq \sum_{k=0}^{K-1} \frac{1}{2} \log 2\pi e \sigma_{y_k}^2$$

De plus $Y_k = X_k + Z_k$ avec Z_k indépendante de X_k d'où $\sigma_{y_k}^2 = \sigma_{x_k}^2 + \sigma_k^2 \leq P_k + \sigma_k^2$. On en déduit la majoration suivante

$$\mathbb{I}(\mathbf{X}, \mathbf{Y}) \leq \frac{1}{2} \sum_{k=0}^{K-1} \log \left(2\pi e(P_k + \sigma_k^2)\right) - \log \left(2\pi e \sigma_k^2\right) = \frac{1}{2} \sum_{k=0}^{K-1} \log \left(1 + \frac{P_k}{\sigma_k^2}\right)$$

De cette inégalité, on obtient

$$C \leq \frac{1}{2} \sum_{k=0}^{K-1} \log \left(1 + \frac{P_k}{\sigma_k^2}\right)$$

avec égalité si et seulement si les variables aléatoires Y_k sont indépendantes et gaussiennes de variance $\sigma_{y_k}^2 = P_k + \sigma_k^2$.

Supposons maintenant $X_k \sim \mathcal{N}_{\mathbb{R}}(0, P_k)$, X_k indépendantes entre elles, alors $Y_k = X_k + Z_k$ seront indépendantes entre elles car les X_k et les Z_k sont indépendantes entre elles et les Z_k sont indépendantes des X_k. De plus, comme $X_k \sim \mathcal{N}_{\mathbb{R}}(0, P_k)$ et $Z_k \sim \mathcal{N}_{\mathbb{R}}(0, \sigma_k^2)$, on a $Z_k \sim \mathcal{N}_{\mathbb{R}}(0, P_k + \sigma_k^2)$.

Ce qui permet de conclure que la capacité de canaux AWGN parallèles est la somme des capacités des canaux AWGN individuels

$$C = \frac{1}{2} \sum_{k=0}^{K-1} \log \left(1 + \frac{P_k}{\sigma_k^2}\right)$$

cette capacité est atteinte pour des entrées gaussiennes indépendantes, de puissance P_k.