Théorie de l'information et du codage

TD n^{o}5 – Canaux et capacités + entropie différentielle

Entropie différentielle

On termine ici le td sur l'entropie différentielle. On définit l'entropie différentielle de la variable aléatoire X par :

$$H(X) = \int_{x\in\mathbb{R}} f(x)\log\frac{1}{f(x)}dx,$$

quand cette quantité existe, où f est la densité de X sur $\mathbb{R}$.

Pour rappel, les propriétés suivantes, démontrées dans le cas discret, sont encore valables dans le cas continu :

1. $D(f||g) \geq 0$, avec égalité si et seulement si $f = g$ presque partout ;
2. $I(X;Y) \geq 0$, avec égalité si et seulement si X et Y sont indépendants ;
3. $H(X|Y) \leq H(X)$, avec égalité si et seulement si X et Y sont indépendants ;
4. $H(X_1,\ldots,X_p) = \sum_{i=1}^{p} H(X_i|X_1,\ldots,X_{i-1})$.

Soit $\mu \in \mathbb{R}^d$ et K une matrice $d\times d$ symétrique réelle définie positive. On dit que le vecteur aléatoire $d-$dimensionnel $X = (X_1,\ldots,X_d)$ suit une loi gaussienne de moyenne μ et de matrice de covariance K (on note $X \sim \mathcal{N}(\mu,K)$) s'il admet la densité suivante sur $\mathbb{R}^d$:

$$x \mapsto \frac{1}{\sqrt{(2\pi)^d|K|}}e^{-\frac{1}{2}\,{}^{\mathsf{T}}(x-\mu)K^{-1}(x-\mu)},$$

où $|K|$ désigne le déterminant de la matrice K.

Question 1. *Calculer l'entropie différentielle d'un tel vecteur et montrer qu'à covariance K fixée, aucun autre vecteur aléatoire $d-$dimensionnel n'a une entropie aussi élevée.*

Un peu d'algèbre linéaire

Pour terminer, retrouvons deux résultats classiques d'algèbre linéaire à l'aide de l'entropie.

Question 2 (Inégalité d'Hadammard). *Montrer que le déterminant d'une matrice symétrique réelle positive est toujours inférieur au produit des éléments diagonaux, et donner une condition nécessaire et suffisante pour qu'il y ait égalité.*

Question 3 (log-concavité du déterminant). *Soit K_1 et K_2 de matrices $d\times d$ réelles symétriques définies positives, et soit $0\leq\theta\leq 1$. Montrer que :*

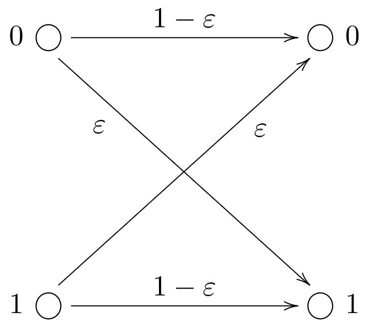
$$|\theta K_1 + (1-\theta)K_2| \geq |K_1|^{\theta}|K_2|^{1-\theta}.$$

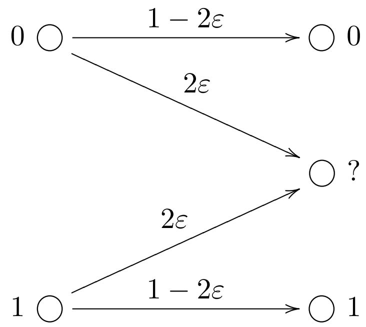
Indication : on pourra introduire trois v.a. indépendantes $X_1 \sim \mathcal{N}(0, K_1)$, $X_2 \sim \mathcal{N}(0, K_2)$ et $U \sim$ Bernoulli(θ), et majorer l'entropie différentielle de la variable aléatoire

$$Z = \begin{cases} X_1 & \textit{si } U = 0, \\ X_2 & \textit{si } U = 1. \end{cases}$$

Quelques calculs de capacité de canaux

Question 4 (Canal binaire symétrique et canal à effacement). *Calculer puis comparer les capacités des deux canaux suivants :*





Quelle est la capacité du canal obtenu en les mettant en série ?

Question 5 (Canal $n-$aire symétrique avec coût). *On considère le canal n-aire dont la matrice de transition est :*

$$\begin{pmatrix} q & p & \dots & p \\ p & q & \dots & p \\ \vdots & \vdots & \ddots & \vdots \\ p & p & \dots & q \end{pmatrix}.$$

Étant donnée une fonction de coût b (positive), déterminer la fonction capacité-coût du canal :

$$C(\beta) = \max \left\{ I(X,Y); \mathbb{E}[b(X)] \leq \beta \right\}.$$

Indication : Montrer dans un premier temps que le problème se ramène à maximiser l'entropie de la v.a. Y sous une certaine contrainte de coût que l'on explicitera.

Question 6 (Canal de Zorro). *Le canal de Zorro est le canal binaire de matrice de transition* $\begin{pmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{pmatrix}$. *Calculer sa capacité.*

Question 7 (Canal gaussien). *Le canal gaussien a pour alphabet d'entrée $\mathcal{A}_X = \mathbb{R}$ et pour alphabet de sortie $\mathcal{A}_Y = \mathbb{R}$. Sa réponse $Y_1, Y_2, \cdots$ à une séquence d'entrée $X_1, X_2, \cdots$ est simplement donnée par $Y_i = X_i + Z_i$, où $Z_1, Z_2, \dots$ sont des variables aléatoires i.i.d. de loi $\mathcal{N}(0, \sigma^2)$. Le paramètre $\sigma > 0$ représente l'intensité moyenne du bruit dans le canal. Par ailleurs, le coût de transmission $b(x)$ d'un symbole $x \in \mathbb{R}$ à travers ce canal est simplement son énergie x^2.*
De la même manière que dans le cas discret, on peut définir la fonction capacité-coût :

$$C(\beta) = \max \left\{ I(X,Y); X \textit{ v.a. réelle à densité telle que } \mathbb{E}[b(X)] \leq \beta \right\},$$

$$\begin{tikzcd}
& Z \sim \mathcal{N}(0,\sigma^2) \arrow[d] & \\
X \arrow[r] & \bigoplus \arrow[r] & Y
\end{tikzcd}$$

et l'on peut montrer que $C(\beta)$ représente le nombre maximum de bits d'information non-erronée qu'il est possible de transmettre en moyenne à chaque utilisation du canal avec une puissance moyenne de transmission limitée à β. Montrer que :

$$C(\beta) = \frac{1}{2} \log \left(1 + \frac{\beta}{\sigma^2} \right).$$

Question 8 (Source gaussienne)**.** *La source gaussienne a pour alphabet-source $\mathcal{A}_U = \mathbb{R}$: les symboles $U_1, U_2, \cdots$ produits sont des variables aléatoires i.i.d. de loi $\mathcal{N}(0, \sigma^2)$. Le paramètre $\sigma > 0$ représente la puissance moyenne d'émission de la source. La mesure de distorsion choisie ici est $d \colon (x, y) \mapsto (x - y)^2$. On peut alors définir la fonction taux-distorsion de cette source :*

$$R(\delta) = \min \left\{ I(U, V); V \textit{ v.a. réelle à densité telle que } \mathbb{E}[d(U, V)] \leq \delta \right\}.$$

Comme dans le cas discret, on peut montrer que $R(\delta)$ représente le nombre minimal de bits nécessaires en moyenne pour décrire chaque symbole émis par la source avec une erreur moyenne d'au plus δ. Montrer que :

$$R(\delta) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{\delta} & \textit{si } \delta \leq \sigma^2, \\ 0 & \textit{si } \delta \geq \sigma^2. \end{cases}$$