Théorie de l'information et du codage

TD n°4 – Du Discret au Continu

Nous nous sommes jusqu'ici restreints au codage de variables aléatoires discrètes. L'objet de ce TD est d'étendre la théorie au cas continu.

1 Entropie différentielle

Dans cette partie, on considère une variable aléatoire X dont la loi admet une densité f sur $\mathbb{R}$. En pratique, on ne dispose que d'un ensemble discret de symboles pour coder la source X : il est donc nécessaire de discrétiser l'espace. La manière la plus naturelle consiste à fixer un pas de quantification $\Delta > 0$ et à découper $\mathbb{R}$ en une suite d'intervalles

$$[i\Delta, (i+1)\Delta)\,, i \in \mathbb{Z}.$$

Pour chacun de ces intervalles, on choisit un représentant $x^i_\Delta \in [i\Delta, (i+1)\Delta)$, et l'on remplace la source X par son approximation discrète X_Δ définie sur l'alphabet $\{x^i_\Delta, i \in \mathbb{Z}\}$ par :

$$X_\Delta = x^i_\Delta \text{ si } X \in [i\Delta, (i+1)\Delta).$$

Puisque la variable aléatoire X_Δ est désormais discrète, on peut considérer son entropie $H(X_\Delta)$, qui correspond au nombre minimal de bits nécessaires en moyenne pour la coder.

Question 1. *On fait tendre $\Delta \to 0$: comment se comporte $H(X_\Delta)$?*

On est ainsi conduit à définir 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. Justifier précisément le nom donné à cette quantité.

Question 2. *Calculer l'entropie différentielle de X dans chacun des cas suivants :*

1. *X est uniformément distribuée sur un intervalle de longueur $0 < l < +\infty$.*
2. *X est gaussienne de moyenne $\mu \in \mathbb{R}$ et de variance $\sigma^2 \in \mathbb{R}^+$.*

2 Cas multi-dimensionnel

Dans cette partie, on considère des vecteurs aléatoires multi-dimensionnels $X = (X_1, \ldots, X_p)$ et $Y = (Y_1, \ldots, Y_q)$ et l'on suppose que la loi jointe de (X, Y) admet une densité $f_{X,Y}$ sur $\mathbb{R}^p \times \mathbb{R}^q$. En particulier, X et Y admettent respectivement sur $\mathbb{R}^p$ et $\mathbb{R}^q$ les densités

$$f_X : x \mapsto \int_{\mathbb{R}^q} f_{X,Y}(x,y) dy \quad \text{et} \quad f_Y : y \mapsto \int_{\mathbb{R}^p} f_{X,Y}(x,y) dx.$$

Question 3. *Étendre au cas continu les grandeurs suivantes, introduites dans le cas discret :*

1. $H(X) = H(X_1, \ldots, X_p)$*, l'entropie jointe de* $(X_1, \ldots, X_p)$ *;*
2. $H(X|Y)$*, l'entropie conditionnelle de* X *sachant* Y *;*
3. $I(X;Y)$*, l'information mutuelle entre* X *et* Y *;*
4. $D(f||g)$*, la divergence de Kullback-Leibler entre deux densités* f *et* g *sur* $\mathbb{R}^p$*.*

Question 4. *Étendre au cas continu les propriétés suivantes, démontrées dans le cas discret :*

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 $\mathbb{R}^d$:

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

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

Question 5. *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.*

3 Un peu d'algèbre linéaire

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

Question 6 (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 7 (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*

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