Gravier Thomas

Professeure : Stéphanie Allassonnière

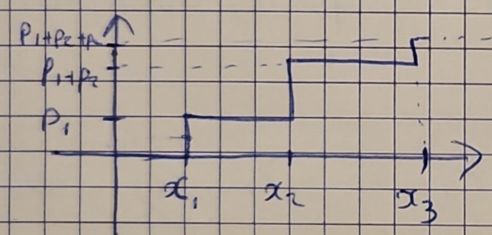
TP n°2 : Expectation - Maximisation algorithm

Exercice n°1 : Discrete distributions

1) Soit $n \in \mathbb{N}^*$ et $X = \{x_1, \dots, x_n\}$ un ensemble de n réels distinct
Soit $(p_i)_{i \in [\![1,n]\!]}$ une suite de réel positif tel que $\sum_{i=1}^{n} p_i = 1$

On pose $U \sim \mathcal{U}([0,1[)$ une variable aléatoire discrète
Soit X la variable aléatoire discrète dont la distribution est donné par les probabilités sur les singletons $P(X = x_i) = p_i$

Soit la fonction de répartition de X $\mathfrak{F}_r(x) = P(X \leqslant x)$



L'inverse de cette fonction de répartition sera $\mathfrak{F}_r^{-1}(u) := \begin{cases} x_1 & \text{si } 0 < u \leqslant p_1 \\ x_2 & \text{si } p_1 \leqslant u < p_1 + \dots \\ x_i & \text{si } u \in I_i \end{cases}$

avec $I_i = \left\{ x, \sum_{j=1}^{i} p_i < x \leqslant \sum_{j=1}^{i+1} p_i \right\}$

L'algorithme est :

1) On ordonne X
2) On tire u dans la loi uniforme
3) On calcul la somme cumulé des probabilité
$\hookrightarrow$ on s'arrête si $u \in I_i$
4) On retourne $x_i = \mathfrak{F}_r^{-1}(u)$

Exercice n°2 : Gaussian mixture model and EM algorithm

1) Tout d'abord nous clarifions les notations
~~On~~ On observe des variables aléatoire X_i qui suivent un mélange gaussien. Pour chaque X_i on suppose qu'il existe une variable $z_i \in \mathbb{R}^m$ qui est un one hot vector et qui encode la classe de X_i

On est dans un modèle de mélange gaussien donc $\forall k \in [\![1,n]\!], \forall j \in [\![1,m]\!]$

$X_i \mid \theta, \{z_{ik} = 1\} \sim \mathcal{N}(\mu_k, \Sigma_k)$ μ_k moyenne de la classe k
Σ_k matrice de covariance de la classe k

~~On note $P(z_{ik}=1) = \alpha_j$~~
On note $\mathbb{P}(z_{ih} = 1) = \alpha_h$ la probabilité que X_i appartienne à la classe h. ~~On le note donc un~~

Les paramètres $\theta = (\mu_1, \dots, \mu_m, \Sigma_1, \dots, \Sigma_m, \alpha_1, \dots, \alpha_m)$

De plus
$$\begin{aligned} \mathcal{L}((x_i)_i, \theta) &= \log p((x_i)_i \mid \theta) \\ &\overset{iid}{=} \log \prod_{i=1}^{n} p(x_i \mid \theta) \\ &= \sum_{i=1}^{n} \log p(x_i \mid \theta) \\ &= \boxed{\sum_{i=1}^{n} \log \sum_{h=1}^{m} \alpha_h \mathcal{N}(x_i; \mu_h, \Sigma_h)} \end{aligned}$$

où $\mathcal{N}(x_i; \mu_h, \Sigma_h)$ est la densité d'une loi normal multivariée appliqué en x_i.