Feuille d'exercice N° 10

Approche bayésienne pour des mélanges gaussiens

Exercice 1. *Echantillonneur de Gibbs*
On considère le modèle bayésien suivant pour le mélange gaussien : Soient $K \geq 2$ et $n \in \mathbb{N}$ fixés. Soient $\alpha_i, \beta_i, \gamma_i$ et λ_i pour $i = 1, \ldots, K$ des hyperparamètres déterministes et fixés. On considère le vecteur aléatoire $(\Theta, \mathbf{X}, \mathbf{Z})$ avec $\mathbf{X} = (X_1, \ldots, X_n)$ et $\mathbf{Z} = (Z_1, \ldots, Z_n)$ généré de la façon suivante :

— On considère $\mathbf{P} = (P_1, \ldots, P_K)$ un vecteur aléatoire de la loi de Dirichlet $\mathcal{D}(\gamma_1, \ldots, \gamma_K)$ de densité

$$f(p_1, \ldots, p_K) = \frac{\Gamma(\sum_{k=1}^K \gamma_k)}{\prod_{k=1}^K \Gamma(\gamma_k)} \prod_{k=1}^K p_k^{\gamma_k - 1}, \qquad (p_1, \ldots, p_K)^T \in (0,1)^K : \sum_{k=1}^K p_k = 1.$$

Puis on considère des variables aléatoires $M_1, \ldots, M_K$ et $S_1^2, \ldots, S_K^2$ telles que les couples (M_k, S_k^2) pour $k = 1, \ldots, K$ sont mutuellement indépendants et telles que

$$S_k^2 \sim \mathcal{IG}\left(\frac{\lambda_k}{2}, \frac{\beta_k}{2}\right), \quad k = 1, \ldots, K,$$

où $\mathcal{IG}(\alpha, \beta)$ désigne la loi inverse-gamma de densité

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left\{-\frac{\beta}{x}\right\}, \qquad x > 0,$$

et, conditionnellement à $S_k^2 = s_k^2$, les M_k sont de loi normale

$$M_k | s_k^2 \sim \mathcal{N}\left(\alpha_k, \frac{s_k^2}{\lambda_k}\right), \quad k = 1, \ldots, K.$$

On pose $\mathbf{M} = (M_1, \ldots, M_K), \mathbf{S}^2 = (S_1^2, \ldots, S_K^2)$ et $\Theta = (\mathbf{P}, \mathbf{M}, \mathbf{S}^2)$.

— Conditionnellement à $\Theta = \theta = (\mathbf{p}, \mathbf{m}, \mathbf{s}^2)$ avec $\mathbf{p} = (p_1, \ldots, p_K)$, $\mathbf{m} = (m_1, \ldots, m_K)$ et $\mathbf{s}^2 = (s_1^2, \ldots, s_K^2)$, on considère $\mathbf{X}$ et $\mathbf{Z}$ tels que les couples (X_i, Z_i) pour $i = 1, \ldots, n$ sont indépendants, les Z_i sont des variables aléatoires à valeurs dans $\{1, \ldots, K\}$ avec des probabilités $\mathbb{P}(Z_i = k | \Theta = \theta) = p_k$ pour $k = 1, \ldots, K$, c'est-à-dire

$$Z_i | \theta \sim \sum_{k=1}^K p_k \delta_{\{k\}}, \quad i = 1, \ldots, n$$

et, conditionnellement à $Z_i = z_i$, X_i est de loi normale

$$X_i | z_i, \theta \sim \mathcal{N}(m_{z_i}, s_{z_i}^2), \quad i = 1, \ldots, n$$

Ceci définit un vecteur aléatoire $(\Theta, \mathbf{X}, \mathbf{Z})$ de densité

$$\pi(\theta, \mathbf{x}, \mathbf{z}) = \pi(\mathbf{p}) \left(\prod_{k=1}^{K} \pi(s_k^2) \pi(m_k | s_k^2) \right) \prod_{i=1}^{n} \pi(z_i | \theta) \pi(x_i | z_i, \theta),$$

où on note $\pi(w)$ la densité de la loi d'une variable aléatoire$W$ évaluée en la valeur w (densité par rapport à la mesure de Lebesgue si W est une v.a. continue ; par rapport à la mesure de comptage si W est une variable discrète) et $\pi(w|u)$ la densité conditionnelle de W sachant U.

1. Montrer que la loi a posterior de Θ sachant que $\mathbf{X} = \mathbf{x}$ avec $\mathbf{x} = (x_1, \ldots, x_n)$ est donnée par
$$\pi(\theta | \mathbf{x}) \propto \prod_{i=1}^{n} \left(\sum_{k=1}^{K} p_k f_{\mathcal{N}(m_k, s_k^2)}(x_i) \right) \times \pi(\theta).$$

2. L'échantillonneur de Gibbs exploite la structure données par les variables latentes $\mathbf{Z}$. Plus précisément, l'algorithme consiste à simuler successivement des lois conditionnelles $\pi(\mathbf{z}|\mathbf{x}, \theta)$ et $\pi(\theta|\mathbf{z}, \mathbf{x})$. Montrer que
$$\begin{aligned}
&\pi(\mathbf{z} | \mathbf{x}, \theta) = \prod_{i=1}^{n} p_{z_i} f_{\mathcal{N}(m_{z_i}, s_{z_i}^2)}(x_i) \qquad \text{et} \\
&\pi(\theta | \mathbf{z}, \mathbf{x}) \propto \pi(\mathbf{p} | \mathbf{z}) \times \pi(\mathbf{m}, \mathbf{s}^2 | \mathbf{z}, \mathbf{x}) \qquad \text{avec} \\
&\qquad \mathbf{P} | \mathbf{z} \sim \mathcal{D}(\gamma_1 + n_1, \ldots, \gamma_K + n_k) \quad \text{où} \quad n_k = \#\{i : z_i = k\} \quad \text{et} \\
&\qquad \pi(\mathbf{m}, \mathbf{s}^2 | \mathbf{z}, \mathbf{x}) \propto \prod_{k=1}^{K} \pi(m_k | s_k^2, \mathbf{z}, \mathbf{x}) \pi(s_k^2 | \mathbf{z}, \mathbf{x})
\end{aligned}$$
et plus précisément,
$$\begin{aligned}
M_k | s_k, \mathbf{z}, \mathbf{x} &\sim \mathcal{N}\left(\frac{\sum_{i:z_i=k} x_i + \alpha_k \lambda_k}{n_k + \lambda_k}, \frac{s_k^2}{n_k + \lambda_k} \right) \\
S_k^2 | \mathbf{z}, \mathbf{x} &\sim \mathcal{IG}\left(\frac{n_k + \lambda_k}{2}, \frac{1}{2} \left(\sum_{i:z_i=k} x_i^2 + \lambda_k \alpha_k^2 + \beta_k - \frac{(\sum_{i:z_i=k} x_i + \alpha_k \lambda_k)^2}{n_k + \lambda_k} \right) \right).
\end{aligned}$$