
ÉCHANTILLONNEUR DE GIBBS

Warm-up

Soit (X, Y) un couple de variables de loi gaussienne centrée de matrice de covariance

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

où $\rho \in (0, 1)$. Écrire un échantillonneur de Gibbs permettant de simuler approximativement la loi de (X, Y) .

Introduisons le vecteur $Z = X - \rho Y$. Alors,

$$\text{Cov}(Z, Y) = \text{Cov}(X, Y) - \rho \text{Cov}(Y, Y) = 0.$$

Ainsi, puisque le vecteur (Y, Z) est gaussien, Y et Z sont indépendantes. On écrit ensuite

$$\mathbb{E}[X|Y] = \mathbb{E}[Z + \rho Y|Y] = \mathbb{E}[Z] + \rho Y = \rho Y.$$

et

$$\mathbb{V}[X|Y] = \mathbb{E}[(X - \mathbb{E}[X|Y])^2|Y] = \mathbb{E}[(X - \rho Y)^2|Y] = \mathbb{E}[Z^2|Y] = \mathbb{V}[Z] = 1 - \rho^2.$$

La loi conditionnelle de X sachant Y , notée $\pi_{X|Y}(\cdot|Y)$, est donc une gaussienne de moyenne ρY et de variance $1 - \rho^2$. On obtient un résultat similaire pour la loi de Y sachant X , notée $\pi_{Y|X}(\cdot|X)$, par symétrie. Ainsi une itération d'un échantillonneur de Gibbs lorsque l'état courant est (X_k, Y_k) serait :

- Simuler $X_{k+1} \sim \pi_{X|Y}(\cdot|Y_k)$.
- Simuler $Y_{k+1} \sim \pi_{Y|X}(\cdot|X_{k+1})$.

Échantillonneur pour un mélange gaussien

Soit $K \geq 2$ et $n \geq 1$. On considère le vecteur aléatoire (θ, X, Z) où $X = (X_1, \dots, X_n)$ et $Z = (Z_1, \dots, Z_n)$ ayant la loi suivante.

- On simule $p = (p_1, \dots, p_K)$ un vecteur ayant la loi de densité proportionnelle à (loi de Dirichlet) $p \mapsto \prod_{k=1}^K p_k^{\gamma_k - 1}$.
- On simule $s_{1:K}^2$, mutuellement indépendantes, et telles que pour tout $1 \leq k \leq K$, s_k^2 a une loi inverse-gamma de paramètres, $\lambda_k/2$ et $\beta_k/2$, i.e. de densité proportionnelle à $u \mapsto u^{-\lambda_k/2-1} \exp(-\beta_k/(2u))$ sur \mathbb{R}_+^* .
- Pour tout $1 \leq k \leq K$, la loi conditionnelle de m_k sachant s_k^2 est gaussienne de moyenne α_k et de variance s_k^2/λ_k .
- Conditionnellement à $\theta = (p_1, \dots, p_K, m_1, \dots, m_K, s_1^2, \dots, s_K^2)$, les $(Z_i, X_i)_{1 \leq i \leq n}$ sont indépendantes et telles que :
 - pour tout $1 \leq k \leq K$, $\mathbb{P}(Z_i = k|\theta) = p_k$;

- conditionnellement à θ et Z_i , X_i suit une loi gaussienne de moyenne m_{Z_i} et de variance $s_{Z_i}^2$.

La densité jointe peut alors s'écrire :

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

où $\pi(w_1 | w_2)$ est une notation générique pour la densité de la loi conditionnelle de la variable W_1 sachant W_2 .

1. Montrer que la loi a posteriori de θ s'écrit :

$$\pi(\theta | x) \propto \pi(\theta) \prod_{i=1}^n \left(\sum_{k=1}^K p_k \varphi_{m_k, s_k^2}(x_i) \right),$$

où φ_{m_k, s_k^2} est la densité gaussienne de moyenne m_k et de variance s_k^2 .

Pour écrire cette loi conditionnelle, il suffit d'écrire que la densité conditionnelle est proportionnelle 'a la densité jointe et ensuite de ne conserver que les quantités qui dépendent de θ . On obtient alors,

$$\begin{aligned} \pi(\theta | x) &\propto \pi(\theta, x) \propto \pi(\theta) \pi(x | \theta), \\ &\propto \pi(\theta) \prod_{i=1}^n \pi(x_i | \theta), \\ &\propto \pi(\theta) \prod_{i=1}^n \left(\sum_{k=1}^K p_k \varphi_{m_k, s_k^2}(x_i) \right). \end{aligned}$$

2. Écrire la densité de la loi conditionnelle de Z sachant (X, θ) .

En procédant comme à la question précédente, on obtient

$$\begin{aligned} \pi(z | \theta, x) &\propto \pi(\theta, z, x) \propto \pi(z | \theta) \pi(x | z, \theta), \\ &\propto \prod_{i=1}^n \pi(z_i | \theta) \pi(x_i | z_i, \theta), \\ &\propto \prod_{i=1}^n \left(\sum_{k=1}^K \mathbb{1}_{z_i=k} p_k \varphi_{m_k, s_k^2} \right). \end{aligned}$$

3. Écrire la densité de la loi conditionnelle de θ sachant (Z, X) .

La densité de la loi conditionnelle de θ sachant (Z, X) s'écrit

$$\pi(\theta | z, x) \propto \pi(\theta, z, x) \propto \pi(\theta) \pi(z | \theta) \pi(x | z, \theta) \propto \pi(p | z) \pi(m, s^2 | x, z).$$

On peut ensuite calculer chacune de ces deux densités. Tout d'abord,

$$\begin{aligned} \pi(p | z) &\propto \pi(\theta, z, x) \propto \pi(p) \pi(z | p), \\ &\propto \prod_{k=1}^K p_k^{\gamma_k - 1} \prod_{i=1}^n p_{z_i} \propto \prod_{k=1}^K p_k^{\gamma_k + n_k - 1}, \end{aligned}$$

où n_k est le nombre de z_i égaux à k . Ainsi, $\pi(p|z)$ est la densité de la loi de Dirichlet de paramètres $(\gamma_1 + n_1, \dots, \gamma_K + n_K)$. D'autre part,

$$\begin{aligned} \pi(m, s^2|x, z) &\propto \pi(s^2)\pi(m|s^2)\pi(z|m, s^2)\pi(x|z, m, s^2), \\ &\propto \left\{ \prod_{i=1}^n s_{z_i}^{-1} \exp\left(-\frac{(x_i - m_{z_i})^2}{2s_{z_i}^2}\right) \right\} \left\{ \prod_{k=1}^K (s_k^2)^{-\lambda_k/2-1} \exp(-\beta_k/(2s_k^2)) \right\} \\ &\quad \times \prod_{k=1}^K s_k^{-1} \exp\left(-\frac{\lambda_k(m_k - \alpha_k)^2}{2s_k^2}\right), \\ &\propto \prod_{k=1}^K (s_k^2)^{-n_k/2-\lambda_k/2-3/2} \exp(-\beta_k/(2s_k^2)) \\ &\quad \times \prod_{k=1}^K \exp\left(\frac{-1}{2s_k^2} \left(\lambda_k(m_k - \alpha_k)^2 + \sum_{i=1}^n \mathbb{1}_{z_i=k} (x_i - m_k)^2 \right)\right). \end{aligned}$$

On remarque alors que

$$\begin{aligned} &\lambda_k(m_k - \alpha_k)^2 + \sum_{i=1}^n \mathbb{1}_{z_i=k} (x_i - m_k)^2 \\ &= (n_k + \lambda_k)m_k^2 - 2m_k \left(\sum_{i=1}^n \mathbb{1}_{z_i=k} x_i + \lambda_k \alpha_k \right) + \lambda_k \alpha_k^2 + \sum_{i=1}^n \mathbb{1}_{z_i=k} x_i^2, \\ &= (n_k + \lambda_k) \left(m_k - \frac{\sum_{i=1}^n \mathbb{1}_{z_i=k} x_i + \lambda_k \alpha_k}{n_k + \lambda_k} \right)^2 - \frac{(\sum_{i=1}^n \mathbb{1}_{z_i=k} x_i + \lambda_k \alpha_k)^2}{n_k + \lambda_k} + \lambda_k \alpha_k^2 + \sum_{i=1}^n \mathbb{1}_{z_i=k} x_i^2, \\ &= (n_k + \lambda_k) (m_k - \tau_k)^2 - \frac{(\sum_{i=1}^n \mathbb{1}_{z_i=k} x_i + \lambda_k \alpha_k)^2}{n_k + \lambda_k} + \lambda_k \alpha_k^2 + \sum_{i=1}^n \mathbb{1}_{z_i=k} x_i^2, \end{aligned}$$

où $\tau_k = (\sum_{i=1}^n \mathbb{1}_{z_i=k} x_i + \lambda_k \alpha_k) / (n_k + \lambda_k)$. Ainsi,

$$\begin{aligned} \pi(m, s^2|x, z) &\propto \prod_{k=1}^K (s_k^2)^{-n_k/2-\lambda_k/2-3/2} \exp\left(-\frac{\rho_k}{2s_k^2}\right) \\ &\quad \times \prod_{k=1}^K \exp\left(\frac{-(n_k + \lambda_k)}{2s_k^2} (m_k - \tau_k)^2\right), \end{aligned}$$

où

$$\rho_k = \beta_k + \lambda_k \alpha_k^2 + \sum_{i=1}^n \mathbb{1}_{z_i=k} x_i^2 - \frac{(\sum_{i=1}^n \mathbb{1}_{z_i=k} x_i + \lambda_k \alpha_k)^2}{n_k + \lambda_k}.$$

Finalement, sous $\pi(m, s^2|x, z)$, les s_k^2 sont indépendants de loi $\mathcal{IG}((n_k + \lambda_k)/2; \rho_k/2)$ et les m_k sont indépendants conditionnellement aux s_k^2 et de loi $\mathcal{N}(\tau_k; s_k^2/(n_k + \lambda_k))$, $1 \leq k \leq K$.

4. Écrire le pseudo-code de l'échantillonneur de Gibbs.

Il suffit, à chaque itération, de simuler la variable sélectionnée conditionnellement aux autres en utilisant les questions précédentes.