Bastien LE CHENADEC
bastien.le-chenadec@eleves.enpc.fr

TP 4
Computational statistics

20 décembre 2023

Exercice 1 Adaptive Metropolis-Hastings within Gibbs sampler

A - Metropolis-Hastings within Gibbs sampler

1. Sur le notebook.
2. Sur le notebook.
3. Sur le notebook.

B - Adaptive Metropolis-Hastings within Gibbs sampler

4. Sur le notebook.
5. Sur le notebook.

Exercice 2 Sampling from multimodal distributions

1. Sur le notebook.
2. Sur le notebook.
3. Sur le notebook.
4. Sur le notebook.

Exercice 3 Bayesian analysis of a one-way random effects model

1. Par la formule de Bayes :

$$P(X,\mu,\sigma,\tau|Y) = \frac{P(Y|X,\mu,\sigma,\tau)P(X,\mu,\sigma,\tau)}{P(Y)}$$

On sait que $Y_{ij} = X_i + \varepsilon_{ij} \sim \mathcal{N}(X_i, \tau^2)$. On a donc :

$$P(Y|X,\mu,\sigma,\tau) \propto \prod_{i=1}^{N}\prod_{j=1}^{k_i} \frac{1}{\tau} \exp\left(-\frac{1}{2\tau^2}(y_{ij}-x_i)^2\right)$$

Toujours par la formules de Bayes, on sait que $P(X,\mu,\sigma,\tau) \propto P(X|\mu,\sigma,\tau)P(\mu,\sigma,\tau)$, et :

$$P(X|\mu,\sigma,\tau) \propto \prod_{i=1}^{N} \frac{1}{\sigma} \exp(-\frac{1}{\sigma^2}(x_i-\mu)^2)$$

De plus $P(\mu,\sigma,\tau)=\pi_{\text{prior}}(\mu,\sigma,\tau)$. En regroupant tous les termes on a :

$$P(X,\mu,\sigma,\tau|Y)\propto \tau^{-k-2(1+\gamma)}\sigma^{-N-2(1+\alpha)}\prod_{i=1}^{N}\prod_{j=1}^{k_i}\exp\left(-\frac{1}{2\tau^2}(y_{ij}-x_i)^2\right)$$
$$\times\prod_{i=1}^{N}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\times\exp\left(-\beta\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\right)$$

où on a noté $k=\sum_{i=1}^{N}k_i$.

2. On a besoin de déterminer les lois a posteriori $\pi_{\text{pos}}(X|\mu,\sigma,\tau)$, $\pi_{\text{pos}}(\mu|X,\sigma,\tau)$, $\pi_{\text{pos}}(\sigma|X,\mu,\tau)$ et $\pi_{\text{pos}}(\tau|X,\mu,\sigma)$. Il suffit de conditionner la vraissemblance qu'on vient de calculer :

$$\pi_{\text{pos}}(X|\mu,\sigma,\tau)\propto\prod_{i=1}^{N}\prod_{j=1}^{k_i}\exp\left(-\frac{1}{2\tau^2}(y_{ij}-x_i)^2\right)\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)$$
$$\pi_{\text{pos}}(\mu|X,\sigma,\tau)\propto\prod_{i=1}^{N}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)$$
$$\pi_{\text{pos}}(\sigma|X,\mu,\tau)\propto\sigma^{-N-2(1+\alpha)}\exp\left(-\beta\frac{1}{\sigma^2}\right)\prod_{i=1}^{N}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)$$
$$\pi_{\text{pos}}(\tau|X,\mu,\sigma)\propto\tau^{-k-2(1+\gamma)}\prod_{i=1}^{N}\prod_{j=1}^{k_i}\exp\left(-\frac{1}{2\tau^2}(y_{ij}-x_i)^2\right)\exp\left(-\beta\frac{1}{\tau^2}\right)$$

On peut maintenant identifier les lois a posteriori :

$$X_i|\mu,\sigma,\tau\sim\mathcal{N}\left(\frac{\sigma^2\sum_{j=1}^{k_i}y_{ij}+\tau^2\mu}{k_i\sigma^2+\tau^2},\frac{\sigma^2\tau^2}{k_i\sigma^2+\tau^2}\right)$$
$$\mu|X,\sigma,\tau\sim\mathcal{N}\left(\frac{1}{N}\sum_{i=1}^{N}x_i,\frac{\sigma^2}{N}\right)$$
$$\sigma^2|X,\mu,\tau\sim\Gamma^{-1}\left(\frac{N}{2}+\alpha,\beta+\frac{1}{2}\sum_{i=1}^{N}(x_i-\mu)^2\right)$$
$$\tau^2|X,\mu,\sigma\sim\Gamma^{-1}\left(\frac{k}{2}+\gamma,\beta+\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{k_i}(x_i-y_{ij})^2\right)$$

Pour l'identification des lois normales, on se souviendra des propriétés de linéarité : si $X_1\sim\mathcal{N}(\mu_1,\sigma_1^2)$ et $X_2\sim\mathcal{N}(\mu_2,\sigma_2^2)$ sont indépendantes, alors $X_1+X_2\sim\mathcal{N}(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

3. On procède comme pour la question précédente en conditionnant la vraissemblance. Pour le bloc (X,μ) on a (on développe le carré en ne gardant que les termes en x_i et μ) :

$$\pi_{\text{pos}}(X,\mu|\sigma,\tau)\propto\prod_{i=1}^{N}exp\left(-\frac{\sum_{j=1}^{k_i}(x_i-y_{ij})^2}{2\tau^2}-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$
$$\propto\prod_{i=1}^{N}exp\left(-\frac{\sum_{j=1}^{k_i}x_i^2-2x_iy_{ij}+y_{ij}^2}{2\tau^2}-\frac{x_i^2-2x_i\mu+\mu^2}{2\sigma^2}\right)$$
$$\propto\prod_{i=1}^{N}exp\left(-\frac{1}{2}\left(\frac{(k_i\sigma^2+\tau^2)x_i^2}{\tau^2\sigma^2}-\frac{2\sum_{j=1}^{k_i}x_iy_{ij}}{\tau^2}-\frac{2x_i\mu}{\sigma^2}+\frac{\mu^2}{\sigma^2}\right)\right)$$

Si on pose $\tilde{x}_i = \frac{\sum_{j=1}^{k_i} y_{ij}}{\tau^2}$ et $\Sigma^{-1} = \begin{pmatrix} \frac{k_1\sigma^2+\tau^2}{\sigma^2\tau^2} & 0 & 0 & \frac{-1}{\sigma^2} \\ 0 & \ddots & \vdots & \vdots \\ 0 & ... & \frac{k_N\sigma^2+\tau^2}{\sigma^2\tau^2} & \frac{-1}{\sigma^2} \\ \frac{-1}{\sigma^2} & \cdots & \frac{-1}{\sigma^2} & \frac{N}{\sigma^2} \end{pmatrix}$ alors on obtient :

$$\pi_{\text{pos}}(X, \mu|\sigma, \tau) \propto \exp\left(-\frac{1}{2}\left[\begin{pmatrix} X \\ \mu \end{pmatrix} - \begin{pmatrix} \tilde{X} \\ 0 \end{pmatrix}\right]^T \Sigma^{-1} \left[\begin{pmatrix} X \\ \mu \end{pmatrix} - \begin{pmatrix} \tilde{X} \\ 0 \end{pmatrix}\right]\right)$$

Donc $X, \mu|\sigma, \tau \sim \mathcal{N}\left(\begin{pmatrix} \tilde{X} \\ 0 \end{pmatrix}, \Sigma\right)$.

4. Théoriquement le Gibbs par blocs devrait être plus rapide que le Gibbs classique car il fait seulement trois tirages à chaque itération. Dans notre cas en particulier le Gibbs classique doit faire $N+3$ tirages avec N qui devient grand. En termes qualitatifs, Gibbs par blocs donne de bons résultats quand les variables sont fortement corrélées entre elles. Dans notre cas, on peut s'attendre à ce que les X_i soient fortement corrélés entre eux car ils sont tirés de la même loi normale.

5. Sur le notebook.