

## TP 3 : Méthodes bayésiennes, MCMC

Ce TP est une introduction aux méthodes bayésiennes dans le contexte des modèles de mélange gaussien, et en particulier à l'échantillonneur de Gibbs et à l'algorithme Metropolis-Hastings.

Je ne crois pas avoir donné en cours la principale référence que je vous propose pour ce chapitre : il s'agit de Marin and Robert (2014).

L'objet de ce TP est d'écrire et de coder vous-mêmes chacun des algorithmes étudiés dans le cas d'un modèle de mélange très simple. Bien sûr, il existe des packages dans lesquels ces algorithmes sont déjà codés. L'avantage de le faire vous-mêmes est de bien les comprendre (pensez donc à réfléchir à ce que vous faites à chaque étape, à rester en mode « critique »). Vous pourrez utiliser les packages existants pour des situations plus compliquées, fortes et forts de ce que vous aurez compris dans ce cas simple.

Nous nous intéressons au modèle de mélange suivant :

$$\mathcal{M} = \{\pi\phi(\cdot; \mu_1, 1) + (1 - \pi)\phi(\cdot; \mu_2, 1) : \mu_1 \in \mathbb{R}, \mu_2 \in \mathbb{R}\},$$

avec  $\phi$  la densité gaussienne sur  $\mathbb{R}$  et  $\pi \neq \frac{1}{2}$  connu.

La loi a priori de  $(\mu_1, \mu_2)$  est donnée par :  $\mu_1 \sim \mathcal{N}(\delta_1, \frac{1}{\lambda})$ ,  $\mu_2 \sim \mathcal{N}(\delta_2, \frac{1}{\lambda})$ ,  $\delta_1, \delta_2 \in \mathbb{R}$  et  $\lambda > 0$ , avec  $\mu_1$  et  $\mu_2$  indépendantes.

**Échantillonneur de Gibbs** Écrire l'algorithme de l'échantillonneur de Gibbs pour la simulation de la loi a posteriori pour le modèle  $\mathcal{M}$ .

Indice : montrez au passage que les lois normales forment une famille de lois conjuguées pour chacune des espérances du modèle (en fait dans le modèle complet associé au modèle  $\mathcal{M}$ ). Profitez-en pour interpréter les paramètres de la loi a posteriori pour chaque espérance (dans le modèle complet) : que pouvez-vous en dire ?

Codez l'algorithme que vous avez écrit, pour une initialisation et une loi a priori sur les paramètres données et un nombre d'itérations fixé. Vérifiez qu'il semble bien fonctionner, quitte à bien choisir la loi a priori ( $\delta_1$ ,  $\delta_2$  et  $\lambda$ ) et l'initialisation ( $\mu_1^0$  et  $\mu_2^0$ ) ! Vous pourrez constater que cet algorithme, un peu comme l'algorithme EM, court le risque de rester coincé à proximité d'un maximum local.

**Algorithme de Metropolis-Hastings** Écrire l'algorithme de Metropolis-Hastings pour la simulation de la loi a posteriori pour le modèle  $\mathcal{M}$ .

Vous pourrez choisir comme loi de proposition une loi normale qu'il conviendra donc de centrer en la valeur courante des paramètres et pour laquelle on pourra choisir une matrice de covariance diagonale (on simule des propositions indépendantes pour  $\mu_1$  et  $\mu_2$ ) :

$$\begin{pmatrix} \zeta^2 & 0 \\ 0 & \zeta^2 \end{pmatrix} \text{ avec } \zeta^2 > 0 \text{ (qui se prononce « zeta »!).}$$

Codez l'algorithme que vous avez écrit, pour une initialisation et une loi a priori sur les paramètres données et un nombre d'itérations fixé. Vérifiez qu'il semble bien fonctionner, quitte à bien choisir les hyperparamètres ( $\zeta$ ), la loi a priori ( $\delta_1$ ,  $\delta_2$  et  $\lambda$ ) et l'initialisation ( $\mu_1^0$  et  $\mu_2^0$ ) ! Vous pourrez constater que la difficulté avec cet algorithme est de bien calibrer la proportion d'acceptation des paramètres proposés (vous pouvez suivre la valeur de cette proportion), ce qui revient ici à bien choisir  $\zeta$ .

**Comparaison des algorithmes** Simulez un jeu de données, par exemple selon un mélange à deux composantes gaussiennes.

Appliquez un nombre fixé d'itérations de chacun des algorithmes que vous avez codés (par exemple en choisissant  $\mu_1^0$  et  $\mu_2^0$  de manière aléatoire uniformément parmi les observations et en tout cas en utilisant la même initialisation pour les deux algorithmes afin de bien comprendre ce qui les différencie). Étudiez et comparez leur comportement. Pensez à évaluer l'effet d'un éventuel tour de chauffe.

### Quelques idées pour diagnostiquer le fonctionnement des algorithmes

Vous pourrez par exemple dans chaque cas :

1. Représenter, au gré des itérations de l'algorithme :
  - l'évolution de la vraisemblance ;
  - les espérances simulées ;
  - la classe affectée à chacune des observations, dans le cas de l'échantillonneur de Gibbs ;
  - l'évolution du taux d'acceptation au grès des itérations, dans le cas de l'algorithme de Metropolis-Hastings.
2. Représenter par exemple des boîtes à moustaches des vraisemblances obtenues ou des histogrammes des paramètres simulés lors des itérations de l'algorithme, après un éventuel « tour de chauffe » que vous prendrez soin d'éliminer.

Bonus : il pourra être intéressant de comparer les valeurs de la vraisemblance et les estimations obtenues par ces méthodes avec celles obtenues par maximum de vraisemblance via l'algorithme EM.

## Références

Marin, J.-M. and Robert, C. P. (2014). *Bayesian essentials with R*, volume 48. Springer.