	ANNÉE UNIVERSITAIRE 2017/2018	
université de BORDEAUX	**UE 4TMS802EX** **Simulation stochastique et méthodes bayésiennes pour le traitement du signal** **Devoir surveillé** **Date : 02/05/2017 Heure : 14h30 Durée : 3h00** Documents non autorisés.	Collège Sciences & Technologies

Important : les pages 2 à 3 sont à compléter et à rendre avec votre copie !

1 Approximation de loi a posteriori par MCMC

Dans cet exercice on considère un modèle bayésien standard : on note $X = (X_1, ..., X_n)$ un échantillon iid (indépendance conditionnelle) de n variables aléatoires réelles dont la loi $f(x|\theta)$ dépend d'un paramètre $\theta \in \mathbb{R}$. On note également $\pi(\theta)$ la loi a priori sur le paramètre θ.

On considère dans cette partie le modèle bayésien ci-dessous :

$$f(x_i|\theta) \sim \mathcal{N}(\theta, \sigma^2) \quad \text{et} \quad \pi(\theta) \sim \mathcal{U}([a, b]),$$

avec σ un paramètre connu et a, b deux hyperparamètres.

On rappelle que l'algorithme de Metropolis-Hasting vise à simuler une chaîne de Markov $(Y_j)_{j\geqslant 0}$ dont la loi stationnaire est $\pi(\theta|X)$ pour le modèle ci-dessus. Etant donné que la chaîne se trouve dans l'état Y_j, une itération de l'algorithme de Metropolis-Hasting consiste en les étapes suivantes (avec une loi de proposition symétrique) :

1. $Y^{new} = Y_j + Z$, où Z est une variable aléatoire centrée, indépendante de Y_j, de loi symétrique autour de zéro (c'est-à-dire que Z et $-Z$ ont même loi),
2. soit $U \sim \mathcal{U}([0, 1])$: si $U \leqslant \min\left(\frac{\pi(Y^{new}|X)}{\pi(Y_j|X)}, 1\right)$ alors $Y_{j+1} = Y^{new}$, et sinon $Y_{j+1} = Y_j$.

1. Que se passe-t-il si la loi de Z n'est pas centrée ? Il est juste demandé d'expliquer ce qui va changer dans l'algorithme, sans donner les formules.
2. Pour ce modèle bayésien, vous devez compléter le code R ci-dessous (en remplaçant les ......) pour implémenter une méthode de simulation de la loi a posteriori par l'algorithme de Metropolis-Hasting en utilisant uniquement la loi du modèle $f(x|\theta)$ et la loi a priori $\pi(\theta)$. Peut-on prendre n'importe qu'elle initialisation pour la chaîne de Markov ? Expliquez dans votre copie votre choix d'initialisation.

   Indications - On pourra utiliser les fonctions `dnorm` et `dunif` pour le calcul des densités a priori et du modèle. On rappelle l'utilisation de ces fonctions pour le calcul d'une densité ou d'une probabilité :

   - si $Z \sim \mathcal{N}(0, 1)$ alors la valeur de la densité f de Z au point $x \in \mathbb{R}$ est donnée par
     `dnorm(x, mean = 0, sd = 1)`,
   - si $Z \sim \mathcal{U}(a, b)$ avec $a = 1$ et $b = 2$ alors la valeur de la densité f de Z au point $x \in \mathbb{R}$ est donnée par
     `dunif(x,min = 1, max = 2)`.

   Dans les codes ci-dessus, si $x = (x_1, \ldots, x_n)$ est un vecteur de dimension n, alors ces fonctions renvoient le vecteur des valeurs de la densité (ou probabilité) de Z aux points $(x_1, \ldots, x_n)$. Enfin on rappelle que la somme (respectivement le produit) des éléments d'un vecteur x se calcule avec la fonction `sum(x)` (respectivement `prod(x)`).

Nom :
Prénom :

```
#################################
# Début du code à compléter
#################################

# Paramètres
n = 20

# Donnees
theta1 = 3 # Vrai valeur de theta inconnue
sigma = 1
X1 = rnorm(n,theta1,sigma)

# A priori
a = 0 # min loi a priori
b = 5 # max loi a priori


Nb = 50000 # Nbre d'iteration de l'algo de MH

# MCMC
theta_post = rep(0,Nb) # Stockage des valeurs de la Chaine de Markov

y = ................. # Initialisation
for (j in ................. ){
  # Loi de proposition
  ynew = .................

  # Acceptation rejet

  r =  ......................................................
       ......................................................
       ......................................................
       ......................................................

  if (................. < r) {

    y = .................

  }

   theta_post[j] = .................
}
```

3. Vous devez compléter le code R ci-dessous (en remplaçant les ......) pour déterminer une valeur approchée de l'estimateur bayésien $\hat{\theta}(X) = \int_{\mathbb{R}} \theta\pi(\theta|X)d\theta$ à partir des simulations de la loi a posteriori par MCMC. Expliquez sur votre copie à quoi sert la variable `Nb0`.

```
# Burn-in period
Nb0 = 1001

# Valeur aprochée de l'estimateur bayésien
theta_post_final = theta_post1[Nb0:..................]

theta_approx1 = .....................................................

################################
# Fin du code à compléter
################################
```

2 Modèle hiérarchique

On s'intéresse à deux échantillons Y_1 et Y_2 distincts de taille n_1 et n_2. Pour ces deux échantillons on considère le modèle hiérarchique suivant :

— $Y_{i,j} = \theta_i + \varepsilon_{i,j}$ pour $i \in \{1, 2\}$ et $1 \leqslant j \leqslant n_i$, avec $\varepsilon_{i,j}$ des variables aléatoires iid de loi $\mathcal{N}(0, 1)$.

— On suppose que les θ_i suivent une loi $\mathcal{B}(p)$ et sont indépendantes (conditionnellement à p).

— On suppose que $p \sim \mathcal{U}([0, 1])$.

1. Montrer que $\theta_1 \sim \mathcal{B}(1/2)$.
2. Calculer la loi de θ_1 sachant $(Y_{1,j})_{1 \leqslant j \leqslant n_1}$. On pourra remarquer que l'on a une loi de Bernoulli.
3. Montrer que l'estimateur du maximum a posteriori de θ_1 vaut
$$\mathbb{1}_{\frac{\sum_{j=1}^{n_1} Y_{1,j}}{n_1} > 1/2}.$$
4. Montrer que $\mathbb{P}((\theta_1, \theta_2) = (0, 0)) = \mathbb{P}((\theta_1, \theta_2) = (1, 1)) = 1/3$. Donner la loi de (θ_1, θ_2). Est-ce que θ_1 et θ_2 sont indépendantes ?
5. Calculer la loi de (θ_1, θ_2) sachant $(Y_{1,j})_{1 \leqslant j \leqslant n_1}$ et $(Y_{2,j})_{1 \leqslant j \leqslant n_2}$.

3 Estimation et traitement des images

On considère le modèle d'observation d'image suivant :

$$v = u_0 + e \tag{1}$$

où v, u_0, e sont à valeur dans $\mathbb{R}^{n \times n}$. L'image u_0 est à estimer. Le bruit e est une suite de v.a. iid de loi $\mathcal{N}(0, \sigma)$. On considère les a priori $\pi_1 \sim \mathcal{N}(0, \mu_1)$ sur u_0, et $\pi_2 \sim \mathcal{N}(0, \mu_2)$ sur ∇u_0, où ∇ est le gradient discret d'une image.

1. Montrer comment interpréter π_2 comme un a priori sur u_0. Exprimer le **modèle bayésien** sur v pour ces deux a priori.
2. Dans les deux cas, donner la forme variationnelle du problème d'estimation du maximum a posteriori (MAP) et proposer un paramètre λ déterminant la force de l'a priori.

3. En R, on dispose de l'observation `v` et de la vérité terrain `u0` . On dispose de deux fonctions `MAP1(lambda,u)` et `MAP2(lambda,u)` estimant le MAP de paramètre `lambda` pour une observation `u` de `u0`. Proposer un code R calculant et affichant l'erreur quadratique des deux estimateurs de u_0 à partir des obsrvations v en fonction de `lambda`.
4. Dessiner la forme que ce dernier graphique doit avoir pour un estimateur qui permet d'effectuer du débruitage. Expliquer.

4 Simulation d'une loi vectorielle

On considère la densité sur $\mathbb{R}^2$ donnée par

$$f(x,y) = C \exp\left(-\frac{(1+y)(1+x^2)}{2}\right) \mathbb{1}_{y>0}.$$

1. Montrer que f est bien une densité de probabilité sur $\mathbb{R}^2$. Il n'est pas nécessaire de calculer explicitement la constante C.
2. Soit (X,Y) un vecteur aléatoire de densité f. Donner la loi de X sachant Y et la loi de Y sachant X. X et Y sont-elles indépendantes ?
3. En déduire une façon de simuler des tirages de vecteurs aléatoires de densité f. Il est demandé de décrire en pseudo-code l'algorithme utilisé. On suppose que vous avez accès aux générateurs pseudo-aléatoires des lois de probabilité usuelles.