

TD 4 : ECHANTILLONNEUR DE GIBBS

EXERCICE 1 (Cas d'école pour Gibbs)

On considère la densité

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

où C est la constante de normalisation.

1. Soient $(x, y) \in \mathbb{R}^2$. Quelle est la loi de X sachant $Y = y$? Et celle de Y sachant $X = x$?
2. En déduire un échantillonneur de Gibbs séquentiel de loi cible f produisant une chaîne de Markov $((X_n, Y_n))_{n \geq 0}$.
3. La chaîne $((X_n, Y_n))_{n \geq 0}$ ainsi construite vérifie-t-elle la convergence vers f en variation totale et le Théorème ergodique?

EXERCICE 2 (Somme d'exponentielles contrainte)

Soient $X_1, \dots, X_d$ des variables exponentielles indépendantes de paramètres $\lambda_1, \dots, \lambda_d > 0$ donnés. Soit $S = X_1 + \dots + X_d$ leur somme et $s > 0$ une constante fixée.

1. Donner la densité π du vecteur aléatoire $(X_1, \dots, X_d)$ sachant $S \geq s$.
2. Soit $x_{-1} := (x_2, \dots, x_d) \in \mathbb{R}_+^{d-1}$ fixé. Déterminer la densité conditionnelle $\pi(\cdot | x_{-1})$ de $\text{Loi}(X_1 | X_{-1} = x_{-1}, S \geq s)$. Quelle loi reconnaissez-vous?
3. En déduire un échantillonneur de Gibbs séquentiel de loi cible π .
4. Le support $\mathcal{S}(\pi)$ de la loi jointe π correspond-il au produit cartésien des supports des marginales?
5. On veut montrer que la chaîne (X_n) ainsi construite, de noyau noté P , est irréductible. Pour simplifier les notations, on suppose $d = 2$ et on considère un borélien B de $\mathbb{R}_+^2$ tel que $\pi(B) > 0$.
 - (a) Soit $x \in \mathcal{S}(\pi)$. Rappeler la densité du noyau P par rapport à la mesure de Lebesgue, c'est-à-dire la fonction $p(x, x') = p((x_1, x_2), (x'_1, x'_2))$ telle que $P(x, dx') = p(x, x')dx'$.
 - (b) Soient $x = (x_1, x_2) \in \mathcal{S}(\pi) \cap (\mathbb{R}_+ \times [s, \infty])$ et B un borélien de $\mathbb{R}^2$ tel que $\pi(B) > 0$. Montrer que $P(x, B) > 0$.
 - (c) En déduire que, pour tout $x = (x_1, x_2) \in \mathcal{S}(\pi)$ et tout borélien B de $\mathbb{R}_+^2$ tel que $\pi(B) > 0$, on a $P^2(x, B) > 0$.
6. La chaîne (X_n) ainsi construite vérifie-t-elle la convergence vers π en variation totale et le Théorème ergodique?

EXERCICE 3 (Modèle de mélange et variables latentes)

Soient $(a, b) \in (\mathbb{R}_+^*)^2$, $(m_0, m_1) \in \mathbb{R}^2$, $(\sigma_0, \sigma_1) \in (\mathbb{R}_+^*)^2$ connus et fixés. On considère le schéma bayésien suivant : la loi a priori sur θ est $\Pi = \text{Beta}(a, b)$. Sachant $\theta = \theta$, les observations $\mathbf{X} = (X_1, \dots, X_N)$ sont i.i.d. selon le mélange de gaussiennes de densité

$$p_\theta(x) = (1 - \theta)\phi_0(x) + \theta\phi_1(x) = (1 - \theta) \times \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{(x-m_0)^2}{2\sigma_0^2}} + \theta \times \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-m_1)^2}{2\sigma_1^2}}.$$

1. Déterminer la loi a posteriori, c'est-à-dire $\text{Loi}(\boldsymbol{\theta}|\mathbf{X})$. Est-ce une loi classique ?
2. On considère le schéma bayésien hiérarchique suivant : la loi a priori sur $\boldsymbol{\theta}$ est encore $\Pi = \text{Beta}(a, b)$. Puis, sachant $\boldsymbol{\theta} = \theta$, les variables $\mathbf{Z} = (Z_1, \dots, Z_N) \in \{0, 1\}^N$ sont i.i.d. selon la loi de Bernoulli $\text{Ber}(\theta)$. Enfin, sachant $\mathbf{Z} = \mathbf{z}$ et $\boldsymbol{\theta} = \theta$, les observations $\mathbf{X} = (X_1, \dots, X_N)$ sont indépendantes avec

$$\text{Loi}(X_j|\mathbf{Z} = \mathbf{z}, \boldsymbol{\theta} = \theta) = \text{Loi}(X_j|Z_j = z_j, \boldsymbol{\theta} = \theta) = \text{Loi}(X_j|Z_j = z_j) = \mathcal{N}(m_{z_j}, \sigma_{z_j}^2).$$

Les variables aléatoires Z_j sont appelées variables latentes (car non observées).

- (a) Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}_+$ une fonction test, montrer que

$$\mathbb{E}[\varphi(X_1)|\boldsymbol{\theta}] = \int_{\mathbb{R}} \varphi(x) p_{\boldsymbol{\theta}}(x) dx.$$

Qu'en déduisez-vous pour la loi a posteriori $\text{Loi}(\boldsymbol{\theta}|\mathbf{X})$ de ce schéma bayésien hiérarchique ?

- (b) Quelle est la loi de $\boldsymbol{\theta}$ sachant $\mathbf{Z}$? Quelle est la loi de $\boldsymbol{\theta}$ sachant $(\mathbf{Z}, \mathbf{X})$?
- (c) Quelle est la loi de Z_1 sachant $(\boldsymbol{\theta}, \mathbf{X})$?
- (d) Quelle est la loi de Z_1 sachant $(\boldsymbol{\theta}, \mathbf{X}, Z_2, \dots, Z_N)$?
- (e) En déduire un échantillonneur de Gibbs permettant de construire, sachant $\mathbf{X}$, une chaîne de Markov $((\boldsymbol{\theta}_n, \mathbf{Z}_n))_{n \geq 0}$ admettant pour loi stationnaire la loi a posteriori du couple $(\boldsymbol{\theta}, \mathbf{Z})$, c'est-à-dire $\text{Loi}((\boldsymbol{\theta}, \mathbf{Z})|\mathbf{X})$.
- (f) La chaîne $((\boldsymbol{\theta}_n, \mathbf{Z}_n))_{n \geq 0}$ ainsi construite est-elle irréductible ? apériodique ?
- (g) La convergence en variation totale de $\text{Loi}(\boldsymbol{\theta}_n|\mathbf{X})$ vers la loi a posteriori $\text{Loi}(\boldsymbol{\theta}|\mathbf{X})$ et le Théorème ergodique sont-ils vérifiés ?
- (h) La suite de variables aléatoires $(\boldsymbol{\theta}_n)$ est-elle une chaîne de Markov ?

EXERCICE 4 (Gibbs simultané = Gibbs erroné)

On considère sur $\mathbb{R}^2$ le couple de variables aléatoires (X, Y) où X est gaussienne standard et $Y = X + W$ avec W gaussienne standard indépendante de X .

1. Quelle est la densité de la loi jointe π du couple (X, Y) ?
2. Quelles sont les lois conditionnelles $\text{Loi}(Y|X)$ et $\text{Loi}(X|Y)$?
3. Décrire une étape de l'échantillonneur de Gibbs séquentiel pour passer du couple (X_n, Y_n) au couple (X_{n+1}, Y_{n+1}) grâce à des simulations de variables gaussiennes. Vérifier que la loi π est bien stationnaire.
4. On considère maintenant une chaîne (X_n, Y_n) dont la dynamique est la suivante : sachant (X_n, Y_n) , on simule X_{n+1} selon la loi conditionnelle $\text{Loi}(X|Y_n) = \pi(\cdot|Y_n)$ et Y_{n+1} selon la loi conditionnelle $\text{Loi}(Y|X_n) = \pi(\cdot|X_n)$. Cette façon de se déplacer s'apparente donc à un échantillonneur de Gibbs où les deux coordonnées seraient remises à jour simultanément. Montrer que dans ce cas la loi π n'est pas stationnaire.

EXERCICE 5 (Slice Sampling)

On considère la fonction

$$\pi(x, y) = \mathbf{1}_{x>0} \mathbf{1}_{0<y<\frac{1}{2}} \exp(-\sqrt{x}).$$

1. Montrer que la fonction $\pi(x) = \frac{1}{2} \exp(-\sqrt{x}) \mathbf{1}_{x>0}$ est une densité sur $\mathbb{R}$ et en déduire que $\pi(x, y)$ est une densité sur $\mathbb{R}^2$. Quelle loi reconnaissez-vous pour le couple (X, Y) de densité $\pi(x, y)$?

2. Soient $x > 0$ et $y > 0$. Quelle est la loi de Y sachant $X = x$? Quelle est la loi de X sachant $Y = y$?
3. En déduire un échantillonneur de Gibbs séquentiel pour simuler une chaîne de Markov $((X_n, Y_n))_{n \geq 0}$ admettant $\pi(x, y)$ pour loi stationnaire. On pourra également faire un dessin pour comprendre graphiquement comment fonctionne cet algorithme.
4. Montrer que $(X_n)_{n \geq 0}$ est une chaîne de Markov admettant $\pi(x)$ comme loi stationnaire.
5. (a) Partant de $X_0 = x \in \mathbb{R}_+^*$, préciser $\text{Loi}(X_1 | X_0 = x)$. En déduire que le noyau de transition de la chaîne $(X_n)_{n \geq 0}$ est $P(x, dx') = (\int \pi(y|x) \pi(x'|y) dy) dx'$.
(b) Soit un borélien B de $\mathbb{R}_+$ de mesure de Lebesgue non-nulle. Montrer que pour tout $x \in \mathbb{R}_+^*$, $P(x, B) > 0$. En déduire que la chaîne $(X_n)_{n \geq 0}$ est Lebesgue-irréductible et apériodique.
6. Montrer que si X a pour densité $\pi(x)$, alors la variable $T := \sqrt{X}$ suit une loi classique et en déduire une nouvelle méthode pour simuler X . Entre cette dernière et l'échantillonneur de Gibbs, laquelle choisissez-vous?
7. Généralisation. La méthode par échantillonnage de Gibbs introduite ci-dessus est connue sous le nom de *Slice Sampling*. Soit $\pi(x)$ une densité sur $\mathbb{R}$, expliquer à quelles conditions et de quelle façon cette méthode peut s'appliquer pour construire une chaîne de Markov de loi stationnaire π ?

EXERCICE 6 ((†) Modèle d'Ising, échantillonneur de Gibbs et dynamique de Glauber)

Soit N un entier naturel fixé. On considère le carré $\mathcal{C} = \{(i, j) \in \{1, \dots, N\}^2\}$, chaque site (i, j) ne pouvant prendre que deux valeurs $S_{ij} = \pm 1$ (son spin). À une configuration de spins $S = (S_{ij})_{1 \leq i, j \leq N}$ donnée est alors associée l'énergie $V(S)$ définie par

$$V(S) = - \sum_{(i,j) \sim (k,\ell)} S_{ij} S_{k\ell} \quad \text{avec} \quad (i, j) \sim (k, \ell) \Leftrightarrow |i - k| + |j - \ell| = 1.$$

Dans la somme définissant V interviennent donc toutes les paires de sites voisins dans la grille $\mathcal{C}$. À une température $T > 0$, on associe à V la mesure de Gibbs

$$\mu_T(S) = \frac{1}{Z_T} \exp\left(-\frac{1}{T} V(S)\right),$$

où Z_T est la constante de normalisation.

1. Soit (i_0, j_0) un site fixé dans $\mathcal{C}$. Préciser la loi conditionnelle $\mu_T(S_{i_0 j_0} | S_{-(i_0 j_0)})$.
2. En déduire un échantillonneur de Gibbs aléatoire (donc non séquentiel) pour construire une chaîne de Markov $(S_n)_{n \geq 0}$ de loi stationnaire μ_T .
3. On se place dans le cadre d'un espace d'états quelconque. Dans l'algorithme de Metropolis, au lieu de $\alpha(x, y) = \min(1, \rho(x, y))$, on considère plus généralement une fonction $h : \mathbb{R}_+ \rightarrow [0, 1]$ vérifiant $h(u) = u h(1/u)$ et la fonction d'acceptation $\alpha(x, y) = h(\rho(x, y))$. Par exemple $h(u) = u/(1+u)$ vérifie cette équation fonctionnelle.
(a) Montrer que la réversibilité de la chaîne vis-à-vis de la loi cible est encore satisfaite.
(b) On revient au modèle d'Ising. Montrer que l'échantillonneur de Gibbs ci-dessus correspond à un algorithme de Metropolis où, partant d'une configuration $S = (S_{ij})_{1 \leq i, j \leq N}$, on commence par tirer un site (i_0, j_0) au hasard (uniforme) et on propose de changer $S_{i_0 j_0}$ en $-S_{i_0 j_0}$. On note $S^{(i_0, j_0)}$ cette nouvelle configuration où un seul un spin a changé de signe. On calcule le rapport d'acceptation $\rho(S, S^{(i_0, j_0)})$, puis on décide en fonction de

$$\alpha(S, S^{(i_0, j_0)}) = h(\rho(S, S^{(i_0, j_0)})),$$

avec $h(u) = u/(1+u)$.