École Nationale Supérieure de Techniques Avancées
examen du cours SOD333
“Filtrage bayésien et approximation particulaire”

vendredi 26 octobre 2018, 13:30 à 15:30

Exercice 1

On considère les distributions normalisées μ_n et $\eta_n = \mu_n^-$ définies par

$$\langle \gamma_n, \phi \rangle = \mathbb{E}[\phi(X_n) \prod_{k=0}^{n} g_k(X_k)\,] \qquad \text{et} \qquad \langle \mu_n, \phi \rangle = \frac{\langle \gamma_n, \phi \rangle}{\langle \gamma_n, 1 \rangle} \ ,$$

et par

$$\langle \gamma_n^-, \phi \rangle = \mathbb{E}[\phi(X_n) \prod_{k=0}^{n-1} g_k(X_k)\,] \qquad \text{et} \qquad \langle \mu_n^-, \phi \rangle = \frac{\langle \gamma_n^-, \phi \rangle}{\langle \gamma_n^-, 1 \rangle} \ ,$$

respectivement, pour toute fonction mesurable bornée ϕ, où $\{X_k\}$ est une chaîne de Markov caractérisée par

- la distribution de probabilité initiale $\eta_0(dx)$,
- et les noyaux de probabilités de transition $Q_k(x, dx')$, pour tout $k = 1, \cdots, n$,

et où $g_k(x)$ sont des fonctions mesurables bornées (strictement positives) données, appelées fonctions de sélection, pour tout $k = 0, 1, \cdots, n$. On rappelle que le noyau positif $R_k(x, dx')$ est défini par

$$R_k(x, dx') = Q_k(x, dx')\, g_k(x') \ ,$$

pour tout $k = 1, \cdots, n$.

On rappelle que l'approximation particulaire *avec redistribution multinomiale* de la distribution normalisée μ_n vérifie le TCL suivant

$$\sqrt{N}\, \langle \mu_n^N - \mu_n, \phi \rangle \Longrightarrow \mathcal{N}(0, v_n(\phi)) \ ,$$

en distribution quand $N \uparrow \infty$, avec la variance asymptotique définie par

$$v_n(\phi) = \sum_{k=0}^{n} \frac{\langle \eta_k, |g_k\, R_{k+1:n}\, (\phi - \langle \mu_n, \phi \rangle)\,|^2 \,\rangle}{\langle \eta_k, g_k\, R_{k+1:n}\, 1 \rangle^2} \ ,$$

où les fonctions

$$R_{k+1:n}\,\phi(x) = R_{k+1}\cdots R_n\,\phi(x)$$

définies pour tout $k = 0, 1\cdots n$, avec la convention $R_{n+1:n}\,\phi(x) = \phi(x)$ pour $k = n$, se calculent de manière rétrograde.

L'objectif de cet exercice est d'obtenir un TCL pour l'approximation particulaire *avec redistribution multinomiale* de la distribution normalisée $\eta_n = \mu_n^-$.

(i) **Montrer que la suite γ_k^- vérifie la relation de récurrence**

$$\gamma_k^- = \gamma_{k-1}^-\,R_k^- \ ,$$

où le noyau positif $R_k^-(x,dx')$ est défini par

$$R_k^-(x,dx') = g_{k-1}(x)\,Q_k(x,dx') \ ,$$

pour tout $k = 1,\cdots,n$.

Solution

On rappelle que les distributions non–normalisées vérifient les relations linéaires

$$\gamma_k = g_k\,\gamma_k^- \qquad \text{et} \qquad \gamma_k^- = \gamma_{k-1}\,Q_k \ ,$$

soit

$$\gamma_k = g_k\,(\gamma_{k-1}\,Q_k) = \gamma_{k-1}\,R_k \ ,$$

pour tout $k = 1,\cdots,n$ avec la condition initiale $\gamma_0 = g_0\,\eta_0$. En changeant le point de vue, on a aussi

$$\gamma_k^- = \gamma_{k-1}\,Q_k \qquad \text{et} \qquad \gamma_{k-1} = g_{k-1}\,\gamma_{k-1}^- \ .$$

soit

$$\gamma_k^- = (g_{k-1}\,\gamma_{k-1}^-)\,Q_k \ ,$$

pour tout $k = 1,\cdots,n$ avec la condition initiale $\gamma_0^- = \eta_0$.

De manière équivalente $\gamma_k^- = \gamma_{k-1}^-\,R_k^-$, où le noyau positif (non–normalisé) R_k^- est défini par $R_k^-(x,dx') = g_{k-1}(x)\,Q_k(x,dx')$. On remarque en effet que

$$\begin{aligned}
\gamma_k^-(dx') &= \int_E \gamma_{k-1}(dx)\,Q_k(x,dx') \\
&= \int_E g_{k-1}(x)\,\gamma_{k-1}^-(dx)\,Q_k(x,dx') \\
&= \int_E \gamma_{k-1}^-(dx)\,g_{k-1}(x)\,Q_k(x,dx') \\
&= \int_E \gamma_{k-1}^-(dx)\,R_k^-(x,dx') \ ,
\end{aligned}$$

où les intégrales portent sur la variable $x \in E$ seulement.

□

On considère les noyaux positifs itérés, définis par leur action sur une fonction mesurable bornée ϕ arbitraire, par

$$R^-_{k+1:n}\, \phi(x) = R^-_{k+1} \cdots R^-_n\, \phi(x) \ ,$$

pour tout $k = 0, 1, \cdots, n$, avec la convention $R^-_{n+1:n}\, \phi(x) = \phi(x)$ pour $k = n$.

(ii) **Montrer (par récurrence arrière) l'identité suivante**

$$R^-_{k+1:n} = g_k\, (R_{k+1:n-1}\, Q_n) \ ,$$

pour tout $k = n-1, \cdots, 0$.

[Indication : il s'agit de montrer que $R^-_{k+1:n}(x, dx') = g_k(x)\, (R_{k+1:n-1}\, Q_n)(x, dx')$.]

Solution

Par définition, on a immédiatement

$$R^-_n(x, dx') = g_{n-1}(x)\; Q_n(x, dx') \ ,$$

c'est–à–dire que l'identité est vérifiée pour $k = (n-1)$. D'autre part, si l'identité est vérifiée à l'étape k, alors

$$\begin{aligned}
R^-_{k:n}(x, dx') &= \int_E R^-_k(x, dx'')\, R^-_{k+1:n}(x'', dx') \\
&= g_{k-1}(x) \int_E Q_k(x, dx'')\, R^-_{k+1:n}(x'', dx') \\
&= g_{k-1}(x) \int_E Q_k(x, dx'')\, g_k(x'')\; (R_{k+1:n-1}\, Q_n)(x'', dx') \\
&= g_{k-1}(x) \int_E R_k(x, dx'')\, (R_{k+1:n-1}\, Q_n)(x'', dx') \\
&= g_{k-1}(x)\; (R_k\, (R_{k+1:n-1}\, Q_n))(x, dx') \\
&= g_{k-1}(x)\; (R_{k:n-1}\, Q_n)(x, dx') \ ,
\end{aligned}$$

c'est–à–dire que l'identité est vérifiée à l'étape $(k-1)$. En particulier

$$R^-_{k+1:n}\, 1 = g_k\; R_{k+1:n-1}\, Q_n\, 1 = g_k\; R_{k+1:n-1}\, 1 \ ,$$

pour tout $k = 0, 1, \cdots, (n-1)$.

□

(iii) **Montrer que**

$$\sqrt{N}\ \langle \eta_n^N - \eta_n, \phi \rangle \Longrightarrow \mathcal{N}(0, v_n^-(\phi)) \ ,$$

en distribution quand $N \uparrow \infty$, où la variance asymptotique vérifie la relation

$$v_n^-(\phi) = v_{n-1}(Q_n\,\phi) + \mathrm{var}(\phi, \eta_n) \ .$$

Solution

On remarque que

$$\sqrt{N}\ \langle \eta_n^N - \eta_n, \phi \rangle = \sqrt{N}\ \langle \eta_n^N - \mu_{n-1}^N\,Q_n, \phi \rangle + \sqrt{N}\ \langle \mu_{n-1}^N - \mu_{n-1}, Q_n\,\phi \rangle = Z'_N + Z''_N \ ,$$

pour toute fonction ϕ mesurable bornée.

On vérifie que la v.a. Z''_N est mesurable par rapport à la tribu $\mathcal{F}_{n-1}^N$ engendrée par le système de particules jusqu'à la $(n-1)$–ème génération, et d'après le TCL rappelé dans l'énoncé, on a

$$Z''_N \Longrightarrow \mathcal{N}(0, v_{n-1}(Q_n\,\phi)) \ ,$$

en distribution quand $N \uparrow \infty$.

On remarque que

$$Z'_N = \sqrt{N}\ \langle \eta_n^N - \mu_{n-1}^N\,Q_n, \phi \rangle = \frac{1}{\sqrt{N}}\ \sum_{i=1}^N [\phi(\xi_n^i) - \langle \mu_{n-1}^N\,Q_n, \phi \rangle] \ ,$$

où conditionnellement par rapport à la tribu $\mathcal{F}_{n-1}^N$ engendrée par le système de particules jusqu'à la $(n-1)$–ème génération, les variables aléatoires $\xi_n^1, \cdots, \xi_n^N$ sont i.i.d. de distribution de probabilité commune $\mu_{n-1}^N\,Q_n$. On vérifie que conditionnellement par rapport à la tribu $\mathcal{F}_{n-1}^N$ la variable aléatoire

$$X_{i,N} = \phi(\xi_n^i) - \langle \mu_{n-1}^N\,Q_n, \phi \rangle \ ,$$

est centrée, de variance

$$s_{i,N}^2 = \mathbb{E}[\,|X_{i,N}|^2 \mid \mathcal{F}_{n-1}^N] = \mathrm{var}(\phi, \mu_{n-1}^N\,Q_n) \ ,$$

et bornée, pour tout $i = 1, \cdots, N$. Clairement

$$s_N^2 = \frac{1}{N}\ \sum_{i=1}^N s_{i,N}^2 = \mathrm{var}(\phi, \mu_{n-1}^N\,Q_n) \longrightarrow \mathrm{var}(\phi, \eta_n) \ ,$$

en probabilité quand $N \uparrow \infty$, et en suivant les lignes de la preuve du TCL vue en cours, on obtient

$$\mathbb{E}[\exp\{j\,u\ Z'_N\} \mid \mathcal{F}_{n-1}^N] \longrightarrow \exp\{-\tfrac{1}{2}\,u^2\ \mathrm{var}(\phi, \eta_n)\} \ ,$$

en probabilité quand $N \uparrow \infty$.

Il résulte du TCL *conditionnel* vu en cours que

$$\sqrt{N}\,\langle \eta_n^N - \eta_n, \phi\rangle \Longrightarrow \mathcal{N}(0, v_n^-(\phi))\ ,$$

en distribution quand $N \uparrow \infty$, avec l'expression suivante pour la variance asymptotique

$$v_n^-(\phi) = v_{n-1}(Q_n\,\phi) + \mathrm{var}(\phi, \eta_n)\ ,$$

pour toute fonction mesurable bornée ϕ.

□

(iv) **En utilisant l'identité établie à la question (ii), montrer que la variance asymptotique est aussi donnée par l'expression suivante**

$$v_n^-(\phi) = \sum_{k=0}^{n} \frac{\langle \eta_k, |R^-_{k+1:n}\,(\phi - \langle \eta_n, \phi\rangle)\,|^2\,\rangle}{\langle \eta_k, R^-_{k+1:n}\,1\rangle^2}\ .$$

Solution

Il résulte de l'expression de la variance asymptotique rappelée dans l'énoncé et de l'identité démontrée à la question (ii), que

$$\begin{aligned} v_{n-1}(Q_n\,\phi) &= \sum_{k=0}^{n-1} \frac{\langle \eta_k, |g_k\,R_{k+1:n-1}\,(Q_n\,\phi - \langle \mu_{n-1}, Q_n\,\phi\rangle)\,|^2\,\rangle}{\langle \eta_k, g_k\,R_{k+1:n-1}\,1\rangle^2} \\ &= \sum_{k=0}^{n-1} \frac{\langle \eta_k, |R^-_{k+1:n}\,(\phi - \langle \eta_n, \phi\rangle)\,|^2\,\rangle}{\langle \eta_k, R^-_{k+1:n}\,1\rangle^2}\ , \end{aligned}$$

de sorte que

$$\begin{aligned} v_n^-(\phi) &= v_{n-1}(Q_n\,\phi) + \mathrm{var}(\phi, \eta_n) \\ &= \sum_{k=0}^{n-1} \frac{\langle \eta_k, |R^-_{k+1:n}\,(\phi - \langle \eta_n, \phi\rangle)\,|^2\,\rangle}{\langle \eta_k, R^-_{k+1:n}\,1\rangle^2} + \langle \eta_n, |\phi - \langle \eta_n, \phi\rangle\,|^2\,\rangle \\ &= \sum_{k=0}^{n} \frac{\langle \eta_k, |R^-_{k+1:n}\,(\phi - \langle \eta_n, \phi\rangle)\,|^2\,\rangle}{\langle \eta_k, R^-_{k+1:n}\,1\rangle^2}\ . \end{aligned}$$

□

Exercice 2

Généralités : Soit η une distribution de probabilité sur E. Sans perte de généralité — comme le montre l'exemple des noyaux de Metropolis–Hastings introduits ci–dessous à la question (i) — on suppose l'existence d'un noyau markovien M réversible pour la distribution de probabilité η, c'est–à–dire que

$$\eta(dx)\, M(x, dx') = \eta(dx')\, M(x', dx) \ .$$

En intégrant par rapport à la variable x, on constate que $\eta\, M = \eta$, c'est–à–dire que le noyau markovien M laisse η invariante. On remarque aussi que

$$\begin{aligned}\langle \eta, u\, M\, v\rangle \ &= \int_E \eta(dx)\, u(x) \ \int_E M(x, dx')\, v(x') \\ &= \int_E \eta(dx')\, v(x') \ \int_E M(x', dx)\, u(x) = \langle \eta, v\, M\, u\rangle \ ,\end{aligned}$$

pour toute fonction u et v définies sur E.

Soit λ une mesure positive dominant η et soit Q un noyau markovien dominé par la même mesure positive λ, c'est–à–dire que

$$\eta(dx) = p(x)\, \lambda(dx) \qquad \text{et} \qquad Q(x, dx') = q(x, x')\ \lambda(dx') \ .$$

(i) **Montrer que le noyau markovien M défini par**

$$M(x, dx') = r(x, x')\ Q(x, dx') + (1 - r(x))\ \delta_x(dx') \ ,$$

avec

$$r(x, x') = \min(1, \frac{p(x')\ q(x', x)}{p(x)\ q(x, x')}) \qquad \textbf{et} \qquad r(x) = \int_E r(x, x')\ Q(x, dx') \ ,$$

pour tout $x, x' \in E$, est réversible pour la distribution de probabilité η.

Solution

Il suffit de remarquer que l'expression

$$p(x)\ r(x, x')\ q(x, x') = \min(p(x)\ q(x, x'), p(x')\ q(x', x)) \ ,$$

est inchangée quand on intervertit le rôle des variables x et x', et que

$$\begin{aligned}\int_E \int_E u(x, x')\ (1 - r(x))\ \eta(dx)\ \delta_x(dx') &= \int_E u(x, x)\ (1 - r(x))\ \eta(dx) \\ &= \int_E \int_E u(x, x')\ (1 - r(x'))\ \eta(dx')\ \delta_{x'}(dx) \ ,\end{aligned}$$

pour toute fonction u définie sur l'ensemble produit $E \times E$, de sorte que

$$(1 - r(x))\ \eta(dx)\ \delta_x(dx') = (1 - r(x'))\ \eta(dx')\ \delta_{x'}(dx)\ .$$

On en déduit que

$$\begin{aligned}
\eta(dx)\ M(x, dx') &= p(x)\ r(x, x')\ q(x, x')\, \lambda(dx)\ \lambda(dx') + p(x)\ (1 - r(x))\ \lambda(dx)\ \delta_x(dx') \\[2ex]
&= p(x')\ r(x', x)\ q(x', x)\, \lambda(dx')\ \lambda(dx) + p(x')\ (1 - r(x'))\ \lambda(dx')\ \delta_{x'}(dx) \\[2ex]
&= \eta(dx')\ M(x, dx')\ .
\end{aligned}$$

□

(ii) **Montrer que pour générer une variable aléatoire Ξ distribuée selon $M(x, dx')$, où l'état initial $x \in E$ est fixé, il suffit de générer une variable aléatoire X' selon $Q(x, dx') = q(x, x')\, dx'$, et de poser ensuite**

$$\Xi = \begin{cases} X', & \textbf{avec probabilité } r(x, X'), \\ x, & \textbf{avec probabilité } (1 - r(x, X')), \end{cases}$$

c'est–à–dire que partant de l'état $x \in E$, on accepte la proposition X' avec probabilité $r(x, X')$ et on conserve l'état initial x sinon.

Solution

On vérifie que

$$\begin{aligned}
\mathbb{E}[\phi(\Xi)] &= \mathbb{E}[\,\mathbb{E}[\phi(\Xi) \mid X']\,] \\[2ex]
&= \mathbb{E}[\,\phi(X')\ r(x, X') + \phi(x)\ (1 - r(x, X')\,] \\[2ex]
&= \int_E \phi(x')\ r(x, x')\ Q(x, dx') + \phi(x)\ \int_E (1 - r(x, x'))\ Q(x, dx') \\[2ex]
&= \int_E \phi(x')\ r(x, x')\ Q(x, dx') + \phi(x)\ (1 - r(x)) \\[2ex]
&= \int_E \phi(x')\ [\,r(x, x')\ Q(x, dx') + (1 - r(x))\ \delta_x(dx')\,] \\[2ex]
&= \int_E \phi(x')\ M(x, dx')\ ,
\end{aligned}$$

pour toute fonction mesurable bornée ϕ.

$\square$

On définit la distribution de Gibbs–Boltzmann

$$\mu^\bullet = g \cdot \eta = \frac{g\,\eta}{\langle \eta, g \rangle} \ ,$$

où g une fonction positive bornée et où l'intégrale $\langle \eta, g \rangle$ est supposée strictement positive.

(iii) **Montrer que le noyau markovien**

$$M^\bullet(x,dx') = M(x,dx')\,\frac{g(x')}{\lambda} + (1 - \frac{M\,g(x)}{\lambda})\,\delta_x(dx') \qquad \textbf{où} \qquad \lambda = \sup_{x \in E} g(x) \ ,$$

laisse invariante la distribution de Gibbs–Boltzmann $\mu^\bullet$.

Solution

Par définition

$$M^\bullet\,\phi(x) = \int_E M^\bullet(x,dx')\,\phi(x') = \frac{M(g\,\phi)(x)}{\lambda} + (1 - \frac{M\,g(x)}{\lambda})\,\phi(x) \ ,$$

d'où on déduit

$$\langle \eta, g\,(M^\bullet\,\phi) \rangle = \langle \eta, g\,\frac{M(g\,\phi)}{\lambda} \rangle + \langle \eta, g\,\phi \rangle - \langle \eta, g\,\phi\,\frac{M\,g}{\lambda} \rangle = \langle \eta, g\,\phi \rangle \ ,$$

d'après la propriété de réversibilité, et finalement

$$\langle \mu^\bullet, M^\bullet\,\phi \rangle = \frac{\langle \eta, g\,(M^\bullet\,\phi) \rangle}{\langle \eta, g \rangle} = \frac{\langle \eta, g\,\phi \rangle}{\langle \eta, g \rangle} = \langle \mu^\bullet, \phi \rangle \ ,$$

pour toute fonction mesurable bornée ϕ, c'est–à–dire que $\mu^\bullet\,M^\bullet = \mu^\bullet$.

$\square$

(iv) **Montrer que pour générer une variable aléatoire $X^\bullet$ distribuée selon $M^\bullet(x,dx')$, où $x \in E$ est fixé, il suffit de générer une variable aléatoire Ξ distribuée selon $M(x,dx')$, et de poser ensuite**

$$X^\bullet = \begin{cases} \Xi, & \textbf{avec probabilité } \dfrac{g(\Xi)}{\lambda}, \\[2ex] x, & \textbf{avec probabilité } (1 - \dfrac{g(\Xi)}{\lambda}), \end{cases}$$

c'est–à–dire que partant de l'état $x \in E$, on accepte la proposition Ξ avec probabilité $\dfrac{g(\Xi)}{\lambda}$ et on conserve l'état initial x sinon.

Solution

On vérifie que

$$\begin{aligned}
\mathbb{E}[\phi(X^\bullet)] &= \mathbb{E}[\,\mathbb{E}[\phi(X^\bullet) \mid \Xi]\,] \\
&= \mathbb{E}[\,\phi(\Xi)\,\frac{g(\Xi)}{\lambda} + \phi(x)\,(1 - \frac{g(\Xi)}{\lambda})\,] \\
&= \int_E \phi(x')\,\frac{g(x')}{\lambda}\,M(x,dx') + \phi(x)\,\int_E (1 - \frac{g(x')}{\lambda})\,M(x,dx') \\
&= \int_E \phi(x')\,\frac{g(x')}{\lambda}\,M(x,dx') + \phi(x)\,(1 - \frac{M\,g(x)}{\lambda}) \\
&= \int_E \phi(x')\,[\,\frac{g(x')}{\lambda}\,M(x,dx') + (1 - \frac{M\,g(x)}{\lambda})\,\delta_x(dx')\,] \\
&= \int_E \phi(x')\,M^\bullet(x,dx')\ ,
\end{aligned}$$

pour toute fonction mesurable bornée ϕ.

$\square$

Application à l'estimation d'une probabilité de dépassement : Soit V une fonction définie sur E et à valeurs positives réelles, et soit $c_* > 0$ un seuil strictement positif. Soit X une variable aléatoire à valeurs dans E et distribuée selon η. On souhaite

- estimer la probabilité de dépassement du seuil c_*, soit

$$p_* = \mathbb{P}[V(X) \geq c_*]\ ,$$

- générer un échantillon distribué (approximativement) selon la loi conditionnelle au dépassement du seuil c_*, soit

$$\mu_*(dx) = \mathbb{P}[X \in dx \mid V(X) \geq c_*]\ .$$

Si la probabilité $p_* \ll 1$ est très petite, il est pertinent d'introduire une suite croissante

$$0 = c_0 < c_1 < \cdots < c_k < \cdots < c_n = c_*\ ,$$

de seuils intermédiaires entre le seuil trivial $c_0 = 0$ et le seuil fixé c_*, avec l'idée de tester $(V(X) \geq c_k)$ successivement pour $k = 0, 1, \cdots, n$, au lieu de tester $(V(X) \geq c_*)$ directement.

Pour tout $k = 1, \cdots, n$ on considère la distribution de probabilité

$$\mu_k = g_k \cdot \eta = \frac{g_k\,\eta}{\langle \eta, g_k \rangle}\ ,$$

où par définition $g_k(x) = 1_{(V(x) \geq c_k)}$.

(v) **Montrer que**

$$\langle \eta, g_k \rangle = \mathbb{P}[V(X) \geq c_k] \qquad \textbf{et} \qquad \mu_k(dx) = \mathbb{P}[X \in dx \mid V(X) \geq c_k] \ ,$$

pour tout $k = 0, 1, \cdots, n$**.**

SOLUTION

Clairement

$$\langle \eta, g_k \, \phi \rangle \ = \mathbb{E}[\phi(X) \, 1_{(V(X) \, \geq \, c_k)}] = \mathbb{E}[\phi(X) \mid V(X) \geq c_k] \; \mathbb{P}[V(X) \geq c_k] \ ,$$

pour toute fonction mesurable bornée ϕ. Pour le choix particulier $\phi(x) \equiv 1$, on obtient

$$\langle \eta, g_k \rangle = \mathbb{P}[V(X) \geq c_k] \ ,$$

et en effectuant le rapport membre–à–membre, on obtient

$$\langle \mu_k, \phi \rangle = \frac{\langle \eta, g_k \, \phi \rangle}{\langle \eta, g_k \rangle} = \mathbb{E}[\phi(X) \mid V(X) \geq c_k] \ ,$$

pour toute fonction mesurable bornée ϕ.

□

(vi) **En déduire que**

$$\langle \mu_{k-1}, g_k \rangle = \mathbb{P}[V(X) \geq c_k \mid V(X) \geq c_{k-1}] \ ,$$

pour tout $k = 1, \cdots, n$**.**

SOLUTION

Clairement

$$\langle \mu_{k-1}, g_k \rangle = \mathbb{E}[g_k(X) \mid V(X) \geq c_{k-1}] = \mathbb{P}[V(X) \geq c_k \mid V(X) \geq c_{k-1}] \ .$$

□

On va donc ici

- estimer la probabilité p_* vue comme probabilité de dépassement du seuil final, soit

$$p_* = \mathbb{P}[V(X) \geq c_n] = \langle \eta, g_n \rangle \ ,$$

- générer un échantillon distribué (approximativement) selon la loi conditionnelle μ_*, vue comme loi conditionnelle au dépassement du seuil final, soit

$$\mu_*(dx) = \mathbb{P}[X \in dx \mid V(X) \geq c_n] = \mu_n(dx) \ .$$

(vii) **Montrer que**

$$g_k = g_k \, g_{k-1} \ ,$$

et en déduire que la distribution de probabilité μ_k vérifie aussi la relation

$$\mu_k = g_k \cdot \mu_{k-1} \ ,$$

en terme de la distribution de probabilité μ_{k-1}.

Solution

Si $(V(x) \geq c_k)$, alors a fortiori $(V(x) \geq c_{k-1})$ compte tenu que $c_k > c_{k-1}$, de sorte que $(V(x) \geq c_k)$ si et seulement si $(V(x) \geq c_k$ et $V(x) \geq c_{k-1})$. De manière équivalente $g_k(x) = g_k(x) \, g_{k-1}(x)$ en terme de fonctions indicatrices.

On en déduit que la distribution de Gibbs–Boltzmann $\mu_k = g_k \cdot \eta$ vérifie

$$\langle \mu_k, \phi \rangle = \frac{\langle \eta, g_k \, \phi \rangle}{\langle \eta, g_k \rangle} = \frac{\langle \eta, g_k \, g_{k-1} \, \phi \rangle}{\langle \eta, g_k \, g_{k-1} \rangle} = \frac{\langle \eta, g_{k-1} \, g_k \, \phi \rangle}{\langle \eta, g_{k-1} \rangle} \, \frac{\langle \eta, g_{k-1} \rangle}{\langle \eta, g_{k-1} \, g_k \rangle} = \frac{\langle \mu_{k-1}, g_k \, \phi \rangle}{\langle \mu_{k-1}, g_k \rangle} \ ,$$

pour toute fonction mesurable bornée ϕ, c'est–à–dire que $\mu_k = g_k \cdot \mu_{k-1}$.

□

(viii) **En utilisant les résultats préliminaires obtenus aux questions (iii) et (iv), montrer que le noyau markovien**

$$M_k(x, dx') = M(x, dx') \, g_{k-1}(x') + (\, 1 - M \, g_{k-1}(x) \,) \, \delta_x(dx') \ ,$$

laisse invariante la distribution de probabilité μ_{k-1}, et que pour simuler une variable aléatoire X_k distribuée selon $M_k(x, dx')$, où $x \in E$ est fixé, il suffit de générer une variable aléatoire Ξ distribuée selon $M(x, dx')$ et de poser ensuite : $X_k = \Xi$ si $V(\Xi) \geq c_{k-1}$, et $X_k = x$ sinon.

Solution

Compte tenu que le noyau markovien M est réversible pour la distribution de probabilité η, et compte tenu que la fonction indicatrice g_{k-1} est bornée par 1, alors il résulte de la question (iii) que le noyau markovien M_k laisse invariante la distribution de Gibbs–Boltzmann $\mu_{k-1} = g_{k-1} \cdot \eta$.

Pour simuler une variable aléatoire X_k distribuée selon $M_k(x, dx')$, où $x \in E$ est fixé, il résulte de la question (iv) qu'il suffit de générer une variable aléatoire Ξ distribuée selon $M(x, dx')$ et de poser ensuite

$$X_k = \begin{cases} \Xi, & \text{avec probabilité } g_{k-1}(\Xi), \\ x, & \text{avec probabilité } (1 - g_{k-1}(\Xi)), \end{cases}$$

c'est–à–dire, compte tenu que $g_{k-1}(x) = 1_{(V(x) \, \geq \, c_{k-1})}$, que partant de l'état $x \in E$, on accepte la proposition Ξ si $V(\Xi) \geq c_{k-1}$, et on conserve l'état initial x sinon.

□

(ix) **En déduire que la suite $\{\mu_k\}$ vérifie l'equation récurrente suivante**

$$\mu_{k-1} \xrightarrow{\text{mutation}} \eta_k = \mu_{k-1}\, M_k \xrightarrow{\text{pondération}} \mu_k = g_k \cdot \eta_k \ ,$$

et que

$$\langle \eta_k, g_k \rangle = \langle \mu_{k-1}, g_k \rangle = \mathbb{P}[V(X) \geq c_k \mid V(X) \geq c_{k-1}] \ ,$$

pour tout $k = 1, \cdots, n$.

Solution

On a montré que $\mu_k = g_k \cdot \mu_{k-1}$ et $\mu_{k-1} = \mu_{k-1}\, M_k$. De manière équivalente $\mu_k = g_k \cdot \eta_k$ en posant $\eta_k = \mu_{k-1}\, M_k$. Clairement $\langle \eta_k, g_k \rangle = \langle \mu_{k-1}, g_k \rangle$ et l'interprétation probabiliste a été obtenue à la question (vi).

□

(x) **Montrer que**

$$p_* = \langle \eta, g_n \rangle = \langle \eta, g_0 \rangle \prod_{k=1}^{n} \langle \eta_k, g_k \rangle \ .$$

Solution

En raisonnant comme à la question (vii), on a

$$\langle \eta, g_k \rangle = \langle \eta, g_k\, g_{k-1} \rangle = \frac{\langle \eta, g_{k-1}\, g_k \rangle}{\langle \eta, g_{k-1} \rangle}\ \langle \eta, g_{k-1} \rangle = \langle \mu_{k-1}, g_k \rangle\ \langle \eta, g_{k-1} \rangle = \langle \eta_k, g_k \rangle\ \langle \eta, g_{k-1} \rangle \ ,$$

compte tenu que $\eta_k = \mu_{k-1}\, M_k = \mu_{k-1}$, et en itérant on obtient

$$p_* = \langle \eta, g_n \rangle = \langle \eta, g_0 \rangle \ \prod_{k=1}^{n} \langle \eta_k, g_k \rangle \ .$$

Alternativement, en utilisant le même argument et la formule de Bayes, on obtient

$$\begin{aligned} p_* &= \mathbb{P}[V(X) \geq c_n] \\ &= \mathbb{P}[V(X) \geq c_n, \cdots, V(X) \geq c_0] \\ &= \prod_{k=1}^{n} \mathbb{P}[V(X) \geq c_k \mid V(X) \geq c_{k-1}] \times \mathbb{P}[V(X) \geq c_0] \ , \end{aligned}$$

et on reconnait les expressions

$$\mathbb{P}[V(X) \geq c_0] = \langle \eta, g_0 \rangle \qquad \text{et} \qquad \mathbb{P}[V(X) \geq c_k \mid V(X) \geq c_{k-1}] = \langle \eta_k, g_k \rangle \ ,$$

obtenues aux questions (v) et (ix) respectivement.

Avec le choix $c_0 = 0$, on a $g_0(x) \equiv 1$, auquel cas $\langle \eta, g_0 \rangle = 1$.

□

(xi) **Montrer comment mettre en œuvre l'algorithme SIR, par exemple avec *redistribution multinomiale*, pour estimer la probabilité p_* de dépassement du seuil c_* et pour générer un échantillon distribué (approximativement) selon la loi μ_* conditionnelle au dépassement du seuil c_*.**

Discuter le rôle du noyau markovien M (utilisé dans la définition du noyau markovien M_k à la question (viii)).

Solution

La mise–en–œuvre générique de l'algorithme SIR donne

- pour $k = 0$, indépendamment pour tout $i = 1 \cdots N$
  on simule une v.a. ξ_0^i distribuée selon $\eta_0(dx)$, et on définit

$$w_0^i = g_0(\xi_0^i) \; / \; [\, \sum_{j=1}^N g_0(\xi_0^j) \,] \;\; ,$$

- pour tout $k = 1 \cdots n$, indépendamment pour tout $i = 1 \cdots N$
  - on sélectionne un individu $\widehat{\xi}_{k-1}^i$ au sein de la population $(\xi_{k-1}^1, \cdots, \xi_{k-1}^N)$ et selon les poids $(w_{k-1}^1, \cdots, w_{k-1}^N)$,
  - on simule une v.a. ξ_k^i distribuée selon $M_k(\widehat{\xi}_{k-1}^i, dx')$,

  et on définit

$$w_k^i = g_k(\xi_k^i) \; / \; [\, \sum_{j=1}^N g_k(\xi_k^j) \,] \;\; ,$$

  d'où les approximations particulaires des distributions normalisées η_k et μ_k comme

$$\eta_k^N = \frac{1}{N} \; \sum_{i=1}^N \delta_{\xi_k^i} \qquad \text{et} \qquad \mu_k^N = \sum_{i=1}^N w_k^i \; \delta_{\xi_k^i} \;\; ,$$

  respectivement, et

$$\langle \eta_k^N , g_k \rangle = \frac{1}{N} \; \sum_{i=1}^N g_k(\xi_k^i) \;\; .$$

On remarque que dans la nouvelle population $(\xi_k^1, \cdots, \xi_k^N)$ tous les individus qui dépassent le nouveau seuil c_k recoivent le même poids non nul, et tous les individus qui échouent recoivent un poids nul, et on introduit l'ensemble

$$I_k^N = \{ i = 1, \cdots, N \; : \; V(\xi_k^i) \geq c_k \} = \{ i = 1, \cdots, N \; : \; w_k^i \neq 0 \} \;\; ,$$

des (indices des) individus dans la nouvelle population $(\xi_k^1, \cdots, \xi_k^N)$ qui dépassent le nouveau seuil c_k.

Par construction, seuls les individus de poids non nul dans la population $(\xi_{k-1}^1, \cdots, \xi_{k-1}^N)$ peuvent être sélectionnés, la sélection porte donc sur la sous–population $(\xi_{k-1}^N\, , \, i \in I_{k-1}^N)$, et mécaniquement tous les individus $(\widehat{\xi}_{k-1}^1, \cdots, \widehat{\xi}_{k-1}^N)$ sélectionnés dépassent le seuil c_{k-1}.

Chacun des individus sélectionnés est agité par le noyau réversible M et l'individu proposé après agitation est accepté si et seulement si lui aussi dépasse le seuil c_{k-1} (et sinon, on conserve l'individu d'origine, avant agitation, qui dépasse le seuil c_{k-1} par construction). Au final, on construit ainsi une nouvelle population $(\xi_k^1, \cdots, \xi_k^N)$ de taille N, dont tous les individus dépassent le seuil c_{k-1}, plus nombreuse et diversifiée que la population d'origine $(\xi_{k-1}^i\, , \, i \in I_{k-1}^N)$ (dont tous les individus aussi dépassent le seuil c_{k-1}).

À la lumière de cette discussion, la mise–en–œuvre de l'algorithme SIR adaptée au problème de dépassement peut se reformuler comme

- pour $k = 0$, indépendamment pour tout $i = 1 \cdots N$
  on simule une v.a. ξ_0^i distribuée selon $\eta_0(dx)$, et on définit

$$I_0^N = \{i = 1, \cdots, N \, : \, V(\xi_0^i) \geq c_0\} = \{i = 1, \cdots, N \, : \, w_0^i \neq 0\} \ ,$$

- pour tout $k = 1 \cdots n$, indépendamment pour tout $i = 1 \cdots N$
  - on sélectionne un individu $\widehat{\xi}_{k-1}^i$ au sein de la population $(\xi_{k-1}^i\, , \, i \in I_{k-1}^N)$, uniformément,
  - on propose une v.a. Ξ distribuée selon $M(\widehat{\xi}_{k-1}^i, dx')$, on évalue $V(\Xi)$ et on pose $\xi_k^i = \Xi$ si $V(\Xi) \geq c_{k-1}$ et on conserve $\xi_k^i = \widehat{\xi}_{k-1}^i$ sinon,

  et on définit

$$I_k^N = \{i = 1, \cdots, N \, : \, V(\xi_k^i) \geq c_k\} = \{i = 1, \cdots, N \, : \, w_k^i \neq 0\} \ ,$$

  d'où les approximations particulaires des distributions normalisées η_k et μ_k comme

$$\eta_k^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_k^i} \qquad \text{et} \qquad \mu_k^N = \frac{1}{|I_k^N|} \sum_{i \in I_k^N} \delta_{\xi_k^i} \ ,$$

  respectivement, et

$$\langle \eta_k^N, g_k \rangle = \frac{1}{N} \, |I_k^N| \ .$$

Clairement

$$\langle \eta_k^N, g_k \rangle = \frac{1}{N} \, |I_k^N| \qquad \text{et} \qquad \mu_k^N = \frac{1}{|I_k^N|} \sum_{i \in I_k^N} \delta_{\xi_k^i} \ ,$$

s'interprètent comme

- la fraction des individus qui dépassent le nouveau seuil c_k,
- et la distribution empirique des individus qui dépassent le nouveau seuil c_k,

pris parmi une population d'individus qui dépassent le précédent seuil c_{k-1}.

$\square$