École Nationale Supérieure de Techniques Avancées
Filière : Finance quantitative
Module : Commande des systèmes

Examen du cours B7–3
“Filtrage bayésien optimal et approximation particulaire”
Lundi 22 octobre 2007, 8:30 à 11:30

Exercice :

L'objectif de cet exercice est d'étudier une classe d'algorithmes de filtrage particulaire généralisés, connue sous le terme générique ABC (pour *approximate Bayesian computation*, en anglais), qui peut s'appliquer aux cas où

- il n'existe pas d'expression suffisamment explicite pour la fonction de vraisemblance,
- la distribution de probabilité de l'observation sachant l'état caché ne possède pas de densité (par rapport à une mesure ne dépendant pas de l'état caché), par exemple parce que le bruit d'observation n'apparaît pas de manière additive, c'est–à–dire qu'il n'existe simplement pas de fonction de vraisemblance,
- certaines composantes de l'observation sont observées exactement, sans bruit d'observation, ce qui peut s'interpréter comme une contrainte a posteriori sur l'état caché, etc.

On considère seulement le cas statique, dans l'une des deux situations suivantes

Modèle A La variable cachée X à valeurs dans E est distribuée selon $\mu(dx)$, il est facile de simuler une variable aléatoire selon la distribution $\mu(dx)$, et l'observation à valeurs dans $\mathbb{R}^d$ est reliée à l'état caché par la relation $Y = h(X)$.

Modèle B La variable cachée X à valeurs dans E et le bruit d'observation V à valeurs dans F sont indépendants, distribués selon $\mu(dx)$ et $\lambda(dv)$ respectivement, il est facile de simuler une variable aléatoire aussi bien selon la distribution $\mu(dx)$ que selon la distribution $\lambda(dv)$, et l'observation à valeurs dans $\mathbb{R}^d$ est reliée à l'état caché par la relation $Y = c(X, V)$.

Clairement, le modèle A est un cas particulier du modèle B.

(i) **Réciproquement, en posant $Z = (X, V)$, montrer que le modèle B se ramène au modèle A.**

Solution

Avec les hypothèses du modèle B, la variable $Z = (X, V)$ à valeurs dans $E \times F$ est distribuée selon $\mu(dx)\ \lambda(dv)$, il est facile de simuler une variable aléatoire selon cette distribution produit, et l'observation à valeurs dans $\mathbb{R}^d$ est reliée à l'état caché par la relation $Y = c(Z)$.

$\square$

On se contente donc dans la suite d'étudier le modèle A uniquement.

(ii) **Donner l'expression de la distribution de probabilité de l'observation Y sachant $X = x$. Que remarque–t–on ? Que peut–on dire qualitativement de la distribution conditionnelle de la variable cachée X sachant $Y = y$?**

Solution

Compte tenu de la relation $Y = h(X)$, la distribution de probabilité de l'observation Y sachant $X = x$ est la masse de Dirac en $h(x)$. Clairement, cette distribution de probabilité n'admet pas de densité par rapport à une mesure positive ne dépendant pas de la variable cachée x. D'autre part si $Y = y$, alors on sait exactement, *sans aucune incertitude*, que la variable cachée X appartient au sous–ensemble $h^{-1}(y) = \{x \in E \,:\, h(x) = y\}$.

$\square$

(iii) **Montrer comment générer un échantillon $(X_i, Y_i, i = 1, \cdots, N)$ selon la distribution de probabilité jointe de la variable cachée et de l'observation.**

Solution

Avec les hypothèses du modèle A, il suffit de simuler indépendamment pour tout $i = 1, \cdots, N$ une variable X_i selon la distribution $\mu(dx)$ et de poser $Y_i = h(X_i)$.

$\square$

On définit une première approximation de la distribution de probabilité jointe de la variable cachée et de l'observation, sous la forme de la distribution empirique

$$\mathbb{P}[X \in dx, Y \in dy] \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{(X_i, Y_i)}(dx, dy) \; .$$

On introduit ensuite une famille de noyaux régularisants : à partir d'une densité de probabilité $K(y)$ définie sur $\mathbb{R}^d$, on pose $K_\varepsilon(y) = \varepsilon^{-d} \; K(\varepsilon^{-1} \, y)$ pour tout $\varepsilon > 0$. On construit ainsi une seconde approximation de la distribution de probabilité jointe de la variable cachée et de l'observation, par convolution avec le noyau régularisant, c'est–à–dire

$$\mathbb{P}[X \in dx, Y \in dy] \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}(dx) \; K_\varepsilon(y - Y_i) \; dy \; .$$

(iv) **En déduire par marginalisation une approximation de la distribution de probabilité de l'observation, puis une approximation de la distribution de probabilité conditionnelle de la variable cachée sachant l'observation.**

Solution

À partir de l'expression de l'approximation régularisée pour la distribution de probabilité jointe de la variable cachée et de l'observation, on obtient par marginalisation (c'est–à–dire en intégrant par rapport à la variable cachée) l'approximation suivante

$$\mathbb{P}[Y \in dy] \approx \frac{1}{N} \sum_{i=1}^{N} K_\varepsilon(y - Y_i) \; dy \; ,$$

pour la distribution de probabilité de l'observation, et d'après la formule de Bayes on obtient l'approximation suivante

$$\mathbb{P}[X \in dx \mid Y = y] \approx \frac{\displaystyle \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}(dx) \; K_\varepsilon(y - Y_i)}{\displaystyle \frac{1}{N} \sum_{i=1}^{N} K_\varepsilon(y - Y_i)} = \sum_{i=1}^{N} \frac{K_\varepsilon(y - Y_i)}{\displaystyle \sum_{j=1}^{N} K_\varepsilon(y - Y_j)} \; \delta_{X_i}(dx) \; ,$$

pour la distribution de probabilité conditionnelle de la variable cachée sachant $Y = y$. On en déduit l'expression suivante

$$\mathbb{P}[X \in dx \mid Y] \approx \sum_{i=1}^{N} w_i \; \delta_{X_i}(dx) \qquad \text{avec} \qquad w_i = \frac{K_\varepsilon(Y - h(X_i))}{\displaystyle \sum_{j=1}^{N} K_\varepsilon(Y - h(X_j))} \; ,$$

pour la distribution de probabilité conditionnelle de la variable cachée sachant l'observation.

$\square$

Pour interpréter le résultat obtenu, on considère le modèle perturbé suivant : la variable cachée X à valeurs dans E et la perturbation aléatoire U à valeurs dans $\mathbb{R}^d$ sont indépendantes, distribuées selon $\mu(dx)$ et selon $K(u)\, du$ respectivement, et l'observation est reliée à l'état caché par la relation $Y = h(X) + \varepsilon\, U$.

(v) **Montrer que le modèle perturbé vérifie les hypothèses faites dans le cours, et que l'approximation obtenue en réponse à la question (iv) correspond à l'approximation par échantillonnage pondéré dans le modèle perturbé.**

Solution

Avec les hypothèses du modèle perturbé, la variable cachée X et le bruit d'observation sont indépendants, il est facile de simuler une variable aléatoire selon la distribution $\mu(dx)$, il est facile d'évaluer la densité de probabilité $K_\varepsilon(u)$ du bruit d'observation $\varepsilon\ U$, et la variable cachée est observée dans un bruit additif. Dans ce modèle, l'approximation par échantillonnage pondéré de la distribution conditionnelle de la variable cachée sachant l'observation est donc donnée par

$$\mathbb{P}[X \in dx \mid Y] \approx \sum_{i=1}^N w_i\, \delta_{X_i}(dx) \qquad \text{avec} \qquad w_i = \frac{K_\varepsilon(Y - h(X_i))}{\displaystyle\sum_{j=1}^N K_\varepsilon(Y - h(X_j))}\ ,$$

ce qui correspond à l'expression obtenue en réponse à la question (iv). En particulier si $K(u)$ est la densité uniforme sur la boule unité de $\mathbb{R}^d$, l'algorithme revient à utiliser d'abord l'approximation suivante

$$\mathbb{P}[X \in dx \mid Y] \approx \mathbb{P}[X \in dx \mid |Y - h(X)| \leq \varepsilon]\ ,$$

pour $\varepsilon > 0$ suffisamment petit, puis à en proposer une approximation par échantillonnage pondéré.

$\square$

On considère le cas particulier où $K(u)$ est la densité uniforme sur la boule unité de $\mathbb{R}^d$, et où il existe un noyau markovien $M(x, dx')$ réversible pour la distribution de probabilité $\mu(dx)$.

(vi) **En introduisant une suite décroissante $\varepsilon_1 > \cdots > \varepsilon_n > 0$ et en s'inspirant des exemples vus en cours et en travaux pratiques, proposer un algorithme particulaire pour l'approximation de la distribution conditionnelle de l'état caché sachant l'observation dans le modèle A.**

On pose

$$g_k(x) = K_{\varepsilon_k}(Y - h(x)) = \mathbf{1}_{(|Y - h(x)| \leq \varepsilon_k)} = \mathbf{1}_{(x \in A_k)} \ ,$$

où $A_k = \{x \in E \ : \ |Y - h(x)| \leq \varepsilon_k\}$ pour tout $k = 0, 1, \cdots, n$, ce qui définit une suite emboitée

$$h^{-1}(Y) \subset A_n \subset \cdots \subset A_1 \subset A_0 \ .$$

On définit alors la distribution de probabilité

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

c'est–à–dire que

$$\mu_k(dx) = \mathbb{P}[X \in dx \mid X \in A_k] = \mathbb{P}[X \in dx \mid |Y - h(X)| \leq \varepsilon_k] \ ,$$

pour tout $k = 0, 1, \cdots, n$, et on rappelle que $\mu_k = g_k \cdot \mu_{k-1}$ pour tout $k = 1, \cdots, n$. On rappelle également qu'il est possible, à partir du noyau markovien $M(x, dx')$ réversible pour la distribution de probabilité $\mu(dx)$, de construire un noyau markovien

$$M_k(x, dx') = M(x, dx') \, g_k(x') + (1 - M \, g_k(x)) \, \delta_x(dx') \ ,$$

qui laisse invariante la distribution de probabilité μ_k, pour tout $k = 0, 1, \cdots, n$. La suite $\{\mu_k \ , \ k = 0, 1, \cdots, n\}$ vérifie donc l'équation récurrente suivante

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

pour tout $k = 1, \cdots, n$, avec la condition initiale $\mu_0 = g_0 \cdot \mu$. On recherche une approximation particulaire pondérée de la forme

$$\mu_k \approx \mu_k^N = \frac{1}{N} \sum_{i=1}^N w_k^i \, \delta_{\xi_k^i} \qquad \text{avec} \qquad \sum_{i=1}^N w_k^i = 1 \ ,$$

et l'algorithme est décrit par les étapes suivantes :

initialisation : indépendamment pour tout $i = 1, \cdots, N$, on simule ξ_0^i selon $\mu(dx)$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_0^i = 1$ si $\xi_0^i \in A_0$, et $w_0^i = 0$ sinon,

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

sélection : indépendamment pour tout $i = 1, \cdots, N$, on simule τ_k^i à valeurs dans l'ensemble $\{1, \cdots, N\}$ des indices, selon le vecteur de probabilité $(w_{k-1}^1, \cdots, w_{k-1}^N)$, et on pose $\widehat{\xi}_{k-1}^i = \xi_{k-1}^{\tau_k^i}$,

mutation : indépendamment pour tout $i = 1, \cdots, N$, on simule Ξ_k^i selon $M(\widehat{\xi}_{k-1}^i, dx')$, on accepte $\xi_k^i = \Xi_k^i$ si $\Xi_k^i \in A_{k-1}$, et on conserve $\xi_k^i = \widehat{\xi}_{k-1}^i$ sinon (cette étape de mutation permet à chaque particule individuellement d'explorer l'ensemble E tout en garantissant que la population reste collectivement distribuée selon μ_{k-1}, et en particulier que chaque particule proposée reste dans le sous–ensemble A_{k-1}),

pondération : pour tout $i = 1, \cdots, N$, on pose $w_k^i = 1$ si $\xi_k^i \in A_k$, et $w_k^i = 0$ sinon.

$\square$

Problème :

L'objectif de ce problème est d'étudier une classe d'algorithmes d'approximation particulaire qui combine les algorithmes SIS et SIR, avec deux sortes de poids affectés à chaque particule

- des poids utilisés pour la redistribution, comme dans l'algorithme SIR,
- et des poids utilisés simplement pour la pondération, comme dans l'algorithme SIS.

On considère la distribution non–normalisée, et la distribution normalisée correspondante, définies sur l'ensemble E par

$$\langle \gamma_n, \phi \rangle = \int_E \cdots \int_E \phi(x_n)\; \gamma_0(dx_0) \prod_{k=1}^n R_k(x_{k-1}, dx_k) \qquad \text{et} \qquad \langle \mu_n, \phi \rangle = \frac{\langle \gamma_n, \phi \rangle}{\langle \gamma_n, 1 \rangle}\ ,$$

où $\gamma_0(dx)$ est une mesure positive (non–normalisée) et où $R_k(x, dx')$ est un noyau positif (non–normalisé) pour tout $k = 1, \cdots, n$. En toute généralité, il est toujours possible de décomposer la mesure positive (non–normalisée)

$$\gamma_0(dx) = W_0(x)\, p_0(dx)\ ,$$

en termes d'une fonction positive et d'une distribution de probabilité, et de décomposer le noyau positif (non–normalisé)

$$R_k(x, dx') = W_k(x, x')\, P_k(x, dx')\ ,$$

en termes d'une fonction positive et d'un noyau markovien pour tout $k = 1, \cdots, n$.

Si on introduit une factorisation supplémentaire des fonctions positives

$$W_0(x) = W_0^{\text{imp}}(x)\, W_0^{\text{red}}(x) \qquad \text{et} \qquad W_k(x, x') = W_k^{\text{imp}}(x, x')\, W_k^{\text{red}}(x, x')\ ,$$

pour tout $k = 1, \cdots, n$, alors on obtient la décomposition suivante

$$\begin{aligned} \gamma_0(dx) &= W_0^{\text{imp}}(x)\, W_0^{\text{red}}(x)\, p_0(dx)\ , \\ R_k(x, dx') &= W_k^{\text{imp}}(x, x')\, W_k^{\text{red}}(x, x')\, P_k(x, dx')\ , \end{aligned} \tag{1}$$

pour tout $k = 1, \cdots, n$, et l'idée consiste à exploiter cette décomposition pour proposer une approximation particulaire de la forme

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

où

- les poids $(w_k^1, \cdots, w_k^N)$ sont utilisés pour la redistribution,
- et les poids d'importance $(u_k^1, \cdots, u_k^N)$ sont utilisés simplement pour la pondération.

Une motivation possible pour cette étude est l'approximation particulaire conjointe de distributions associées à un modèle de référence et à des modèles alternatifs, sous l'hypothèse suivante d'absolue continuité

$$\gamma_0(dx) = r_0(x)\, \gamma_0^0(dx) \qquad \text{et} \qquad R_k(x, dx') = r_k(x, x')\, R_k^0(x, dx') \ ,$$

pour tout $k = 1, \cdots, n$. Dans ce cas en effet, si on introduit une décomposition de la mesure positive (non–normalisée) de référence

$$\gamma_0^0(dx) = W_0^0(x)\, p_0^0(dx) \ ,$$

en termes d'une fonction positive et d'une distribution cde probabilité, et si on introduit une décomposition du noyau positif (non–normalisé) de référence

$$R_k^0(x, dx') = W_k^0(x, x')\, P_k^0(x, dx') \ ,$$

en termes d'une fonction positive et d'un noyau markovien pour tout $k = 1, \cdots, n$, alors on obtient la décomposition suivante

$$\gamma_0(dx) = r_0(x)\, W_0^0(x)\, p_0^0(dx) \ ,$$

$$R_k(x, dx') = r_k(x, x')\, W_k^0(x, x')\, P_k^0(x, dx') \ ,$$

pour tout $k = 1, \cdots, n$, qui est clairement de la forme (1), et il est possible de proposer une approximation particulaire de la forme

$$\mu_k^N = \sum_{i=1}^N u_k^i\; w_k^{i,0}\; \delta_{\xi_k^{i,0}} \qquad \text{avec} \qquad \sum_{i=1}^N w_k^{i,0} = 1 \qquad \text{et} \qquad \sum_{i=1}^N u_k^i\; w_k^{i,0} = 1$$

où

- les positions $(\xi_k^{1,0}, \cdots, \xi_k^{N,0})$ des particules et les poids $(w_k^{1,0}, \cdots, w_k^{N,0})$ utilisés pour la redistribution dépendent du modèle de référence uniquement,
- les poids d'importance $(u_k^1, \cdots, u_k^N)$ utilisés simplement pour la pondération dépendent à la fois de modèle de référence et du modèle alternatif.

Bien entendu, les deux points de vue sont mathématiquement équivalents, à une identification près des différentes mesures positives et des différentes fonctions positives impliquées, et le choix de l'un ou l'autre point de vue dépend souvent de l'application.

Notation On utilisera souvent dans la suite les abus de notation $W_0^{\text{imp}}(x,x') = W_0^{\text{imp}}(x')$ et $W_0^{\text{red}}(x,x') = W_0^{\text{red}}(x')$.

Représentation trajectorielle

(i) **En utilisant la décomposition (1), montrer que**

$$\langle \gamma_n, \phi \rangle = \mathbb{E}[\phi(X_n) \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k) \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k)] \ ,$$

où la suite $\{X_k \, , \, k = 0, 1, \cdots, n\}$ est une chaîne de Markov à valeurs dans E, caractérisée par

- **la distribution de probabilité initiale $p_0(dx)$,**
- **et les probabilités de transition $P_k(x, dx')$ pour tout $k = 0, 1, \cdots, n$.**

Solution

En utilisant la décomposition (1), on obtient la représentation probabiliste suivante

$$\begin{aligned}
\langle \gamma_n, \phi \rangle &= \int_E \cdots \int_E \phi(x_n) \; \gamma_0(dx_0) \prod_{k=1}^n R_k(x_{k-1}, dx_k) \\
&= \int_E \cdots \int_E \phi(x_n) \; W_0^{\text{imp}}(x_0) \, W_0^{\text{red}}(x_0) \, p_0(dx_0) \\
&\qquad\qquad \prod_{k=1}^n W_k^{\text{imp}}(x_{k-1}, x_k) \, W_k^{\text{red}}(x_{k-1}, x_k) \, P_k(x_{k-1}, dx_k) \\
&= \int_E \cdots \int_E \phi(x_n) \prod_{k=0}^n W_k^{\text{imp}}(x_{k-1}, x_k) \, W_k^{\text{red}}(x_{k-1}, x_k) \\
&\qquad\qquad p_0(dx_0) \prod_{k=1}^n P_k(x_{k-1}, dx_k) \\
&= \mathbb{E}[\phi(X_n) \prod_{k=0}^n W_k^{\text{imp}}(X_{k-1}, X_k) \prod_{k=0}^n W_k^{\text{red}}(X_{k-1}, X_k)] \ ,
\end{aligned}$$

où la suite $\{X_k \, , \, k = 0, 1, \cdots, n\}$ est une chaîne de Markov à valeurs dans E, caractérisée par

- la distribution de probabilité initiale $p_0(dx)$,
- et les probabilités de transition $P_k(x, dx')$ pour tout $k = 0, 1, \cdots, n$.

□

On pose $X_k^\bullet = (X_0, \cdots, X_k) = X_{0:k}$ pour tout $k = 0, 1, \cdots, n$, et on introduit la distribution non–normalisée, et la distribution normalisée correspondante, définies sur l'ensemble trajectoriel $E_{0:n} = E \times \cdots \times E$ par

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

respectivement, où $g_k^\bullet(x_{0:k}) = W_k^{\rm red}(x_{k-1}, x_k)$ pour tout $k = 0, 1, \cdots, n$.

(ii) **Montrer que la suite $\{X_k^\bullet \ , \ k = 0, 1, \cdots, n\}$ est une chaîne de Markov à valeurs dans l'ensemble des trajectoires dans E, caractérisée par**

- **la distribution de probabilité initiale $\eta_0^\bullet(dx_0) = p_0(dx_0)$,**
- **et les probabilités de transition**

$$Q_k^\bullet(x_{0:k-1}, dx'_{0:k}) = \delta_{x_{0:k-1}}(dx'_{0:k-1}) \ P_k(x_{k-1}, dx'_k) \ ,$$

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

Solution

Clairement, l'information contenue dans $(X_0^\bullet, \cdots, X_k^\bullet)$ est équivalente à l'information contenue dans $X_k^\bullet$, pour tout $k = 0, 1, \cdots, n$, de sorte que la suite $\{X_k^\bullet \ , \ k = 0, 1, \cdots, n\}$ forme une chaîne de Markov à valeurs dans l'ensemble des trajectoires dans E. En outre

$$\begin{aligned}
&\mathbb{E}[f(X_k^\bullet) \mid X_{k-1}^\bullet = x_{0:k-1}] \\
&\quad = \mathbb{E}[f(X_0, \cdots, X_{k-1}, X_k) \mid X_{0:k-1} = x_{0:k-1}] \\
&\quad = \int_E f(x_0, \cdots, x_{k-1}, x') \ \mathbb{P}[X_k \in dx' \mid X_{0:k-1} = x_{0:k-1}] \\
&\quad = \int_E f(x_0, \cdots, x_{k-1}, x') \ \mathbb{P}[X_k \in dx' \mid X_{k-1} = x_{k-1}] \\
&\quad = \int_E f(x_0, \cdots, x_{k-1}, x') \ P_k(x_{k-1}, dx') \\
&\quad = \int_E \cdots \int_E f(x'_0, \cdots, x'_{k-1}, x'_k) \ \delta_{x_{0:k-1}}(dx'_{0:k-1}) \ P_k(x_{k-1}, dx'_k) \ ,
\end{aligned}$$

c'est–à–dire que la probabilité de transition est donnée par

$$Q_k^\bullet(x_{0:k-1}, dx'_{0:k}) = \delta_{x_{0:k-1}}(dx'_{0:k-1}) \ P_k(x_{k-1}, dx'_k) \ ,$$

pour tout $k = 1, \cdots, n$. Par ailleurs $X_0^\bullet = X_0$, de sorte que la distribution de probabilité initiale est donnée par

$$\eta_0^\bullet(dx_0) = p_0(dx_0) \ .$$

□

(iii) **En déduire, en utilisant les résultats du cours, l'équation récurrente vérifiée par la suite $\{\gamma_k^\bullet \, , \, k = 0, 1, \cdots, n\}$ des distributions non–normalisées définies sur l'ensemble des trajectoires dans E.**

Solution

À partir de la représentation probabiliste, on obtient directement l'équation récurrente

$$\gamma_k^\bullet = g_k^\bullet \ (\gamma_{k-1}^\bullet \, Q_k^\bullet) = g_k^\bullet \ (\mu_{k-1}^\bullet \, Q_k^\bullet) \ \langle \gamma_{k-1}^\bullet, 1 \rangle \ ,$$

pour tout $k = 1, \cdots, n$, avec la condition initiale $\gamma_0^\bullet = g_0^\bullet \, p_0$.

□

(iv) **Montrer que**

$$\langle \gamma_n, \phi \rangle = \langle \gamma_n^\bullet, T_n \, \phi \rangle \ ,$$

où la fonction $T_n \, \phi$ est définie sur l'ensemble trajectoriel $E_{0:n} = E \times \cdots \times E$ par

$$T_n \, \phi(x_{0:n}) = \phi(x_n) \ \prod_{k=0}^n W_k^{\rm imp}(x_{k-1}, x_k) \ .$$

Solution

Compte tenu que

$$T_n \, \phi(x_{0:n}) = \phi(x_n) \ \prod_{k=0}^n W_k^{\rm imp}(x_{k-1}, x_k) \qquad \text{et} \qquad \prod_{k=0}^n g_k^\bullet(x_{0:k}) = \prod_{k=0}^n W_k^{\rm red}(x_{k-1}, x_k) \ ,$$

par définition, et en utilisant le résultat obtenu en réponse à la question (i), il vient

$$\begin{aligned} \langle \gamma_n, \phi \rangle \ &= \ \mathbb{E}[\phi(X_n) \ \prod_{k=0}^n W_k^{\rm imp}(X_{k-1}, X_k) \ \prod_{k=0}^n W_k^{\rm red}(X_{k-1}, X_k) \,] \\ &= \ \mathbb{E}[T_n \, \phi(X_n^\bullet) \ \prod_{k=0}^n g_k^\bullet(X_k^\bullet) \,] \\ &= \ \langle \gamma_n^\bullet, T_n \, \phi \rangle \ . \end{aligned}$$

□

En d'autres termes, la distribution non–normalisée $\gamma_n^\bullet$ définie sur l'ensemble trajectoriel $E_{0:n} = E \times \cdots \times E$ englobe comme cas particulier la distribution non–normalisée γ_n.

(v) **En déduire l'expression**

- **de la constante de normalisation $\langle \gamma_n, 1 \rangle$,**
- **et de la distribution normalisée μ_n,**

en fonction de la distribution non–normalisée $\gamma_n^\bullet$ définie sur l'ensemble trajectoriel $E_{0:n} = E \times \cdots \times E$.

SOLUTION

En particulier pour $\phi(x) \equiv 1$, il vient

$$\langle \gamma_n, 1 \rangle = \langle \gamma_n^\bullet, T_n\, 1 \rangle \ ,$$

pour la constante de normalisation, et

$$\langle \mu_n, \phi \rangle = \frac{\langle \gamma_n, \phi \rangle}{\langle \gamma_n, 1 \rangle} = \frac{\langle \gamma_n^\bullet, T_n\, \phi \rangle}{\langle \gamma_n^\bullet, T_n\, 1 \rangle} = \frac{\langle \mu_n^\bullet, T_n\, \phi \rangle}{\langle \mu_n^\bullet, T_n\, 1 \rangle} \ ,$$

pour la distribution normalisée.

□

On recherche une approximation particulaire pondérée de la forme

$$\mu_k^\bullet \approx \mu_k^{\bullet,N} = \sum_{i=1}^N w_k^i \, \delta_{\xi_k^{\bullet,i}} \qquad \text{avec} \qquad \sum_{i=1}^N w_k^i = 1 \ ,$$

où $\xi_k^{\bullet,i} = (\xi_{0,k}^i, \cdots, \xi_{k,k}^i)$ est une particule dans l'ensemble trajectoriel $E_{0:k} = E \times \cdots \times E$ dont la position terminale est notée $\xi_k^i = \xi_{k,k}^i$ pour tout $i = 1, \cdots, N$.

(vi) **Décrire, en utilisant les résultats du cours, l'approximation particulaire de type bootstrap pour la distribution $\mu_k^\bullet$ définie sur l'ensemble trajectoriel $E_{0:k} = E \times \cdots \times E$. En déduire, en utilisant les expressions obtenues en réponse à la question (iv), une approximation particulaire de type bootstrap pour la distribution μ_k.**

SOLUTION

En appliquant les résultats du cours, l'algorithme est décrit par les étapes suivantes :

initialisation : indépendamment pour tout $i = 1, \cdots, N$, on simule $\xi_0^{\bullet,i}$ selon $\eta_0^\bullet(dx)$, c'est–à–dire que : on simule $\xi_{0,0}^i$ selon $p_0(dx)$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_0^i \propto g_0^\bullet(\xi_0^{\bullet,i}) = W_0^{\rm red}(\xi_{0,0}^i)$, et pour tout $k = 1, \cdots, n$

sélection : indépendamment pour tout $i = 1, \cdots, N$, on simule τ_k^i à valeurs dans l'ensemble $\{1, \cdots, N\}$ des indices, selon le vecteur de probabilité $(w_{k-1}^1, \cdots, w_{k-1}^N)$,

mutation : indépendamment pour tout $i = 1, \cdots, N$,

on simule $\xi_k^{\bullet,i}$ selon $Q_k^\bullet(\xi_{k-1}^{\bullet,\tau_k^i}, dx'_{0:k})$,

c'est–à–dire que : on pose

$$(\xi_{0,k}^i, \cdots, \xi_{k-1,k}^i) = (\xi_{0,k-1}^{\tau_k^i}, \cdots, \xi_{k-1,k-1}^{\tau_k^i}) \ ,$$

et on simule $\xi_{k,k}^i$ selon $P_k(\xi_{k-1,k-1}^{\tau_k^i}, dx')$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_k^i \propto g_k^\bullet(\xi_k^{\bullet,i}) = W_k^{\text{red}}(\xi_{k-1,k}^i, \xi_{k,k}^i)$.

En particulier pour la fonction $T_k\,\phi$ introduite dans l'énoncé de la question (iv), on a

$$T_k\,\phi(\xi_k^{\bullet,i}) = \phi(\xi_{k,k}^i) \prod_{p=0}^{k} W_p^{\text{imp}}(\xi_{p-1,k}^i, \xi_{p,k}^i) \ ,$$

de sorte que

$$\langle \mu_k^{\bullet,N}, T_k\,\phi \rangle = \frac{1}{N} \sum_{i=1}^{N} w_k^i \; T_k\,\phi(\xi_k^{\bullet,i}) = \frac{1}{N} \sum_{i=1}^{N} w_k^i \; v_k^i \; \phi(\xi_k^i) \ ,$$

pour tout $k = 0, 1, \cdots, n$, en termes de la position terminale de chaque particule trajectorielle, où par définition

$$v_0^i = W_0^{\text{imp}}(\xi_0^i) \ ,$$

et

$$\begin{aligned}
v_k^i &= \prod_{p=0}^{k} W_p^{\text{imp}}(\xi_{p-1,k}^i, \xi_{p,k}^i) \\
&= W_k^{\text{imp}}(\xi_{k-1,k}^i, \xi_{k,k}^i) \prod_{p=0}^{k-1} W_p^{\text{imp}}(\xi_{p-1,k}^i, \xi_{p,k}^i) \\
&= W_k^{\text{imp}}(\xi_{k-1,k-1}^{\tau_k^i}, \xi_{k,k}^i) \prod_{p=0}^{k-1} W_p^{\text{imp}}(\xi_{p-1,k-1}^{\tau_k^i}, \xi_{p,k-1}^{\tau_k^i}) \\
&= W_k^{\text{imp}}(\xi_{k-1}^{\tau_k^i}, \xi_k^i) \; v_{k-1}^{\tau_k^i} \ ,
\end{aligned}$$

pour tout $i = 1, \cdots, N$. On a donc, au vu de la réponse à la question (v)

$$\langle \mu_k, \phi \rangle = \frac{\langle \mu_k^\bullet, T_k\, \phi \rangle}{\langle \mu_k^\bullet, T_k\, 1 \rangle} \approx \frac{\langle \mu_k^{\bullet,N}, T_k\, \phi \rangle}{\langle \mu_k^{\bullet,N}, T_k\, 1 \rangle} = \frac{\displaystyle\sum_{i=1}^N w_k^i\, v_k^i\, \phi(\xi_k^i)}{\displaystyle\sum_{j=1}^N w_k^j\, v_k^j} = \sum_{i=1}^N w_k^i\, u_k^i\, \phi(\xi_k^i)\ ,$$

et on vérifie automatiquement la contrainte

$$\sum_{i=1}^N w_k^i\, u_k^i = 1 \qquad \text{avec le choix} \qquad u_k^i = \frac{v_k^i}{\displaystyle\sum_{j=1}^N w_k^j\, v_k^j} \qquad \text{pour tout } i = 1, \cdots, N.$$

En termes de la position terminale de chaque particule trajectorielle, l'algorithme est décrit par les étapes suivantes :

initialisation : indépendamment pour tout $i = 1, \cdots, N$,

on simule ξ_0^i selon $p_0(dx)$ et on pose $v_0^i = W_0^{\rm imp}(\xi_0^i)$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_0^i \propto W_0^{\rm red}(\xi_0^i)$,

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

sélection : indépendamment pour tout $i = 1, \cdots, N$, on simule τ_k^i à valeurs dans l'ensemble $\{1, \cdots, N\}$ des indices, selon le vecteur de probabilité $(w_{k-1}^1, \cdots, w_{k-1}^N)$,

mutation : indépendamment pour tout $i = 1, \cdots, N$,

on simule ξ_k^i selon $P_k(\xi_{k-1}^{\tau_k^i}, dx')$ et on pose $v_k^i = W_k^{\rm imp}(\xi_{k-1}^{\tau_k^i}, \xi_k^i)\; v_{k-1}^{\tau_k^i}$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_k^i \propto W_k^{\rm red}(\xi_{k-1}^{\tau_k^i}, \xi_k^i)$.

□

Représentation en terme d'une fonctionnelle multiplicative

En partant plutôt de l'hypothèse d'absolue continuité

$$\gamma_0(dx) = W_0^{\rm imp}(x)\, \gamma_0^0(dx) \qquad \text{et} \qquad R_k(x, dx') = W_k^{\rm imp}(x, x')\, R_k^0(x, dx')\ , \tag{2}$$

ce qui définit implicitement, au vu de la décomposition (1)

$$\gamma_0^0(dx) = W_0^{\rm red}(x)\, p_0(dx) \qquad \text{et} \qquad R_k^0(x,dx') = W_k^{\rm red}(x,x')\, P_k(x,dx') \ , \tag{3}$$

pour tout $k = 1, \cdots, n$, on introduit la distribution non–normalisée, et la distribution normalisée correspondante, définies sur l'ensemble produit $E \times [0,\infty)$ par

$$\langle \gamma_n^{\rm e}, F\rangle = \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\rm imp}(x_{k-1},x_k))\ \gamma_0^0(dx_0)\ \prod_{k=1}^n R_k^0(x_{k-1},dx_k) \ ,$$

et

$$\langle \mu_n^{\rm e}, F\rangle = \frac{\langle \gamma_n^{\rm e}, F\rangle}{\langle \gamma_n^{\rm e}, 1\rangle} \ ,$$

respectivement.

(vii) **Montrer que**

$$\gamma_0^{\rm e}(dx,dv) = \gamma_0^0(dx)\ \delta_{W_0^{\rm imp}(x)}(dv) \ ,$$

à l'instant initial.

Solution

Par définition

$$\langle \gamma_0^{\rm e}, F\rangle \ = \int_E F(x_0, W_0^{\rm imp}(x_0))\ \gamma_0^0(dx_0) = \int_E \int_0^\infty F(x,v)\ \delta_{W_0^{\rm imp}(x)}(dv)\ \gamma_0^0(dx) \ ,$$

pour toute fonction F définie sur l'ensemble produit $E \times [0,\infty)$.

□

Pour le modèle de référence, on introduit la distribution non–normalisée, et la distribution normalisée correspondante, définies sur E par

$$\langle \gamma_n^0, \phi\rangle = \int_E \cdots \int_E \phi(x_n)\ \gamma_0^0(dx_0)\ \prod_{k=1}^n R_k^0(x_{k-1},dx_k) \qquad \text{et} \qquad \langle \mu_n^0, \phi\rangle = \frac{\langle \gamma_n^0, \phi\rangle}{\langle \gamma_n^0, 1\rangle} \ ,$$

respectivement. Pour tout $v \in [0,\infty)$, on pose $e_0(v) \equiv 1$ et $e(v) = v$.

(viii) **Montrer que**

$$\langle \gamma_n^0, \phi\rangle = \langle \gamma_n^{\rm e}, \phi \otimes e_0\rangle \ ,$$

où la fonction $F = \phi \otimes e_0$ **est définie sur l'ensemble produit** $E \times [0,\infty)$ **par** $F(x,v) = \phi(x)$**, et montrer que**

$$\langle \gamma_n, \phi\rangle = \langle \gamma_n^{\rm e}, \phi \otimes e\rangle \ ,$$

où la fonction $F = \phi \otimes e$ **est définie sur l'ensemble produit** $E \times [0,\infty)$ **par** $F(x,v) = \phi(x)\ v$**.**

Solution

En particulier pour $F = \phi \otimes e_0$, il vient

$$\langle \gamma_n^{\mathrm{e}}, \phi \otimes e_0 \rangle = \int_E \cdots \int_E \phi(x_n)\ \gamma_0^0(dx_0)\ \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) = \langle \gamma_n^0, \phi \rangle \ ,$$

et pour $F = \phi \otimes e$, en utilisant l'hypothèse d'absolue continuité (2), il vient

$$\begin{aligned} \langle \gamma_n^{\mathrm{e}}, \phi \otimes e \rangle &= \int_E \cdots \int_E \phi(x_n)\ \prod_{k=0}^n W_k^{\mathrm{imp}}(x_{k-1}, x_k)\ \gamma_0^0(dx_0)\ \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E \phi(x_n)\ W_0^{\mathrm{imp}}(x_0)\, \gamma_0^0(dx_0)\ \prod_{k=1}^n W_k^{\mathrm{imp}}(x_{k-1}, x_k)\, R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E \phi(x_n)\ \gamma_0(dx_0)\ \prod_{k=1}^n R_k(x_{k-1}, dx_k) = \langle \gamma_n, \phi \rangle \ . \end{aligned}$$

□

En d'autres termes, la distribution non–normalisée γ_n^{e} définie sur l'ensemble produit $E \times [0, \infty)$ englobe simultanément les deux distributions non–normalisées γ_n^0 et γ_n associées respectivement au modèle de référence et au modèle alternatif.

(ix) **En déduire l'expression**

- **des deux constantes de normalisation $\langle \gamma_n^0, 1 \rangle$ et $\langle \gamma_n, 1 \rangle$,**
- **et des deux distributions normalisées μ_n^0 et μ_n,**

associées au modèle de référence et au modèle alternatif, en fonction de la distribution non–normalisée γ_n^{e} définie sur l'ensemble produit $E \times [0, \infty)$.

Solution

En particulier pour $\phi(x) \equiv 1$, il vient

$$\langle \gamma_n^0, 1 \rangle = \langle \gamma_n^{\mathrm{e}}, 1 \otimes e_0 \rangle = \langle \gamma_n^{\mathrm{e}}, 1 \rangle \ ,$$

et

$$\langle \gamma_n, 1 \rangle = \langle \gamma_n^{\mathrm{e}}, 1 \otimes e \rangle \ ,$$

pour les constantes de normalisation, et

$$\langle \mu_n^0, \phi \rangle = \frac{\langle \gamma_n^0, \phi \rangle}{\langle \gamma_n^0, 1 \rangle} = \frac{\langle \gamma_n^{\mathrm{e}}, \phi \otimes e_0 \rangle}{\langle \gamma_n^{\mathrm{e}}, 1 \rangle} = \langle \mu_n^{\mathrm{e}}, \phi \otimes e_0 \rangle \ ,$$

et

$$\langle \mu_n, \phi \rangle = \frac{\langle \gamma_n, \phi \rangle}{\langle \gamma_n, 1 \rangle} = \frac{\langle \gamma_n^{\mathrm{e}}, \phi \otimes e \rangle}{\langle \gamma_n^{\mathrm{e}}, 1 \otimes e \rangle} = \frac{\langle \mu_n^{\mathrm{e}}, \phi \otimes e \rangle}{\langle \mu_n^{\mathrm{e}}, 1 \otimes e \rangle} \ ,$$

pour les distributions normalisées.

□

On introduit le noyau positif (non–normalisé)

$$R_k^{\mathrm{e}}(x, v, dx', dv') = R_k^0(x, dx') \ \delta_{v \, W_k^{\mathrm{imp}}(x, x')}(dv') \ ,$$

défini sur l'ensemble produit $E \times [0, \infty)$ pour tout $k = 1, \cdots, n$.

(x) **Montrer que**

$$\langle \gamma_n^{\mathrm{e}}, F \rangle = \int_E \int_0^\infty \cdots \int_E \int_0^\infty F(x_n, v_n) \ \gamma_0^{\mathrm{e}}(dx_0, dv_0) \ \prod_{k=1}^n R_k^{\mathrm{e}}(x_{k-1}, v_{k-1}, dx_k, dv_k) \ ,$$

ce qui justifie l'interprétation de la distribution γ_n^{e} comme une distribution de Feynman–Kac non–normalisée.

Solution

En appliquant simplement les définitions, on obtient

$$\int_E \int_0^\infty \cdots \int_E \int_0^\infty F(x_n, v_n) \ \gamma_0^{\mathrm{e}}(dx_0, dv_0) \ \prod_{k=1}^n R_k^{\mathrm{e}}(x_{k-1}, v_{k-1}, dx_k, dv_k)$$

$$= \int_E \cdots \int_E \Big\{ \int_0^\infty \cdots \int_0^\infty F(x_n, v_n) \ \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \ \prod_{k=1}^n \delta_{v_{k-1} \, W_k^{\mathrm{imp}}(x_{k-1}, x_k)}(dv_k) \Big\}$$

$$\gamma_0^0(dx_0) \ \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \ .$$

On remarque ensuite que

$$\begin{aligned} I &= \int_0^\infty \cdots \int_0^\infty F(x_n, v_n) \ \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \ \prod_{k=1}^n \delta_{v_{k-1} \, W_k^{\mathrm{imp}}(x_{k-1}, x_k)}(dv_k) \\ &= \int_0^\infty \cdots \int_0^\infty \Big\{ \int_0^\infty F(x_n, v_n) \ \delta_{v_{n-1} \, W_n^{\mathrm{imp}}(x_{n-1}, x_n)}(dv_n) \Big\} \\ &\qquad \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \ \prod_{k=1}^{n-1} \delta_{v_{k-1} \, W_k^{\mathrm{imp}}(x_{k-1}, x_k)}(dv_k) \\ &= \int_0^\infty \cdots \int_0^\infty F(x_n, v_{n-1} \, W_n^{\mathrm{imp}}(x_{n-1}, x_n)) \\ &\qquad \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \ \prod_{k=1}^{n-1} \delta_{v_{k-1} \, W_k^{\mathrm{imp}}(x_{k-1}, x_k)}(dv_k) \ , \end{aligned}$$

et on se propose de montrer par récurrence arrière que

$$I = \int_0^\infty \cdots \int_0^\infty F(x_n, v_k \prod_{p=k+1}^n W_p^{\mathrm{imp}}(x_{p-1}, x_p))$$

$$\delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \prod_{p=1}^k \delta_{v_{p-1}\, W_p^{\mathrm{imp}}(x_{p-1}, x_p)}(dv_p) \; ,$$

pour tout $k = n, \cdots, 0, 1$. Au rang n, la propriété est vraie au vu de ce qui précède. On suppose que la propriété est vraie au rang k, et en répétant le calcul déjà effectué plus haut on remarque que

$$\begin{aligned} I &= \int_0^\infty \cdots \int_0^\infty \{ \int_0^\infty F(x_n, v_k \prod_{p=k+1}^n W_p^{\mathrm{imp}}(x_{p-1}, x_p))\; \delta_{v_{k-1}\, W_k^{\mathrm{imp}}(x_{k-1}, x_k)}(dv_k) \; \} \\ &\qquad \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \prod_{p=1}^{k-1} \delta_{v_{p-1}\, W_p^{\mathrm{imp}}(x_{p-1}, x_p)}(dv_p) \\ &= \int_0^\infty \cdots \int_0^\infty F(x_n, v_{k-1} \prod_{p=k}^n W_p^{\mathrm{imp}}(x_{p-1}, x_p)) \\ &\qquad \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \prod_{p=1}^{k-1} \delta_{v_{p-1}\, W_p^{\mathrm{imp}}(x_{p-1}, x_p)}(dv_p) \; , \end{aligned}$$

c'est–à–dire que la propriété est vraie au rang $(k-1)$. Au rang 0, on a donc

$$I = \int_0^\infty F(x_n, v_0 \prod_{p=1}^n W_p^{\mathrm{imp}}(x_{p-1}, x_p))\; \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) = F(x_n, \prod_{p=0}^n W_p^{\mathrm{imp}}(x_{p-1}, x_p)) \; ,$$

et en reportant cette expression plus haut, on obtient

$$\begin{aligned} &\int_E \int_0^\infty \cdots \int_E \int_0^\infty F(x_n, v_n)\; \gamma_0^{\mathrm{e}}(dx_0, dv_0) \prod_{k=1}^n R_k^{\mathrm{e}}(x_{k-1}, v_{k-1}, dx_k, dv_k) \\ &= \int_E \cdots \int_E \{ \int_0^\infty \cdots \int_0^\infty F(x_n, v_n)\; \delta_{W_0^{\mathrm{imp}}(x_0)}(dv_0) \prod_{k=1}^n \delta_{v_{k-1}\, W_k^{\mathrm{imp}}(x_{k-1}, x_k)}(dv_k) \; \} \\ &\qquad \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) \\ &= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\mathrm{imp}}(x_{k-1}, x_k))\; \gamma_0^0(dx_0) \prod_{k=1}^n R_k^0(x_{k-1}, dx_k) = \langle \gamma_n^{\mathrm{e}}, F \rangle \; . \end{aligned}$$

□

(xi) **En déduire, en utilisant les résultats du cours, l'équation récurrente vérifiée par la suite $\{\gamma_k^{\rm e}\,,\ k=0,1,\cdots,n\}$ des distributions non–normalisées définies sur l'ensemble produit $E \times [0,\infty)$.**

Solution

À partir de la représentation probabiliste, on obtient directement l'équation récurrente

$$\gamma_k^{\rm e} = \gamma_{k-1}^{\rm e}\, R_k^{\rm e} = (\mu_{k-1}^{\rm e}\, R_k^{\rm e})\ \langle \gamma_{k-1}^{\rm e}, 1\rangle\ ,$$

pour tout $k=1,\cdots,n$, avec la condition initiale introduite à la question (vii).

□

(xii) **En revenant à la définition, et en utilisant la décomposition (3), montrer que**

$$\langle \gamma_n^{\rm e}, F\rangle = \mathbb{E}[F(X_n, \prod_{k=0}^n W_k^{\rm imp}(X_{k-1},X_k))\ \prod_{k=0}^n W_k^{\rm red}(X_{k-1},X_k)]\ ,$$

où $\{X_k\,,\ k=0,1,\cdots,n\}$ est une chaîne de Markov à valeurs dans E, caractérisée par

- **la distribution de probabilité initiale $p_0(dx)$,**
- **et les probabilités de transition $P_k(x,dx')$, pour tout $k=1,\cdots,n$.**

Solution

En utilisant la décomposition (3), il vient

$$\begin{aligned}
\langle \gamma_n^{\rm e}, F\rangle &= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\rm imp}(x_{k-1},x_k))\ \gamma_0^0(dx_0)\ \prod_{k=1}^n R_k^0(x_{k-1},dx_k) \\
&= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\rm imp}(x_{k-1},x_k))\ W_0^{\rm red}(x_0)\, p_0(dx_0) \\
&\qquad\qquad \prod_{k=1}^n W_k^{\rm red}(x_{k-1},x_k)\, P_k(x_{k-1},dx_k) \\
&= \int_E \cdots \int_E F(x_n, \prod_{k=0}^n W_k^{\rm imp}(x_{k-1},x_k))\ \prod_{k=0}^n W_k^{\rm red}(x_{k-1},x_k) \\
&\qquad\qquad p_0(dx_0)\ \prod_{k=1}^n P_k(x_{k-1},dx_k) \\
&= \mathbb{E}[F(X_n, \prod_{k=0}^n W_k^{\rm imp}(X_{k-1},X_k))\ \prod_{k=0}^n W_k^{\rm red}(X_{k-1},X_k)]\ ,
\end{aligned}$$

où $\{X_k\,,\ k=0,1,\cdots,n\}$ est une chaîne de Markov à valeurs dans E, caractérisée par

- la distribution de probabilité initiale $p_0(dx)$,
- et les probabilités de transition $P_k(x, dx')$, pour tout $k = 1, \cdots, n$.

$\square$

On pose
$$M_0 = W_0^{\rm imp}(X_0) \qquad \text{et} \qquad M_k = W_k^{\rm imp}(X_{k-1}, X_k)\, M_{k-1} \ ,$$
pour tout $k = 1, \cdots, n$, ce qui fait de la suite $\{M_k \, , \, k = 0, 1, \cdots, n\}$ une fonctionnelle multiplicative associée à la chaîne de Markov $\{X_k \, , \, k = 0, 1, \cdots, n\}$.

(xiii) **En déduire une nouvelle représentation**
$$\langle \gamma_n^{\rm e}, F \rangle = \mathbb{E}[F(X_n, M_n) \prod_{k=0}^n W_k^{\rm red}(X_{k-1}, X_k)] \ ,$$
de la distribution non–normalisée $\gamma_n^{\rm e}$, où conjointement $\{(X_k, M_k) \, , \, k = 0, 1, \cdots, n\}$ forme une nouvelle chaîne de Markov à valeurs dans l'ensemble produit $E \times [0, \infty)$, caractérisée par

- **la distribution de probabilité initiale $p_0^{\rm e}(dx, dv) = p_0(dx)\, \delta_{W_0^{\rm imp}(x)}(dv)$,**
- **et les probabilités de transition**
$$P_k^{\rm e}(x, v, dx', dv') = P_k(x, dx')\ \delta_{v\, W_k^{\rm imp}(x, x')}(dv') \ ,$$
pour tout $k = 1, \cdots, n$.

Solution

Clairement, la variable aléatoire M_k est mesurable par rapport à $(X_0, \cdots, X_k)$ pour tout $k = 0, 1, \cdots, n$, de sorte que l'information contenue dans $(X_0, \cdots, X_k, M_0, \cdots, M_k)$ est équivalente à l'information contenue dans $(X_0, \cdots, X_k)$ seulement, pour tout $k = 0, 1, \cdots, n$. On en déduit que
$$\begin{aligned}
&\mathbb{E}[F(X_k, M_k) \mid X_0, \cdots, X_{k-1}, M_0, \cdots, M_{k-1}] \\
&\quad = \mathbb{E}[F(X_k, M_k) \mid X_0, \cdots, X_{k-1}] \\
&\quad = \mathbb{E}[F(X_k, W_k^{\rm imp}(X_{k-1}, X_k)\ M_{k-1}) \mid X_0, \cdots, X_{k-1}] \\
&\quad = \int_E F(x', W_k^{\rm imp}(X_{k-1}, x')\ M_{k-1})\ \mathbb{P}[X_k \in dx' \mid X_0, \cdots, X_{k-1}] \\
&\quad = \int_E F(x', W_k^{\rm imp}(X_{k-1}, x')\ M_{k-1})\ \mathbb{P}[X_k \in dx' \mid X_{k-1}] \ ,
\end{aligned}$$

et le résultat ne dépend que de (X_{k-1}, M_{k-1}), c'est–à–dire que la suite $\{(X_k, M_k)\, , \, k = 0, 1, \cdots, n\}$ forme une chaîne de Markov à valeurs dans l'ensemble produit $E \times [0, \infty)$. En outre

$$\mathbb{E}[F(X_k, M_k) \mid X_{k-1} = x, M_{k-1} = v]$$

$$= \int_E F(x', v\, W_k^{\rm imp}(x, x'))\; \mathbb{P}[X_k \in dx \mid X_{k-1} = x]$$

$$= \int_E F(x', v\, W_k^{\rm imp}(x, x'))\; P_k(x, dx')$$

$$= \int_E \int_0^\infty F(x', v')\; \delta_{v\, W_k^{\rm imp}(x, x')}(dv')\; P_k(x, dx')\ ,$$

c'est–à–dire que la probabilité de transition est donnée par

$$P_k^{\rm e}(x, v, dx', dv') = P_k(x, dx')\; \delta_{v\, W_k^{\rm imp}(x, x')}(dv')\ ,$$

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

$$\mathbb{E}[F(X_0, M_0)] = \mathbb{E}[F(X_0, W_0^{\rm imp}(X_0))]$$

$$= \int_E F(x, W_0^{\rm imp}(x))\; p_0(dx)$$

$$= \int_E \int_0^\infty F(x, v)\; \delta_{W_0^{\rm imp}(x)}(dv)\; p_0(dx)\ ,$$

c'est–à–dire que la distribution de probabilité initiale est donnée par

$$p_0^{\rm e}(dx, dv) = p_0(dx)\; \delta_{W_0^{\rm imp}(x)}(dv)\ .$$

□

On recherche une approximation particulaire pondérée de la forme

$$\mu_k^{\rm e} \approx \mu_k^{{\rm e},N} = \sum_{i=1}^N w_k^i\; \delta_{(\xi_k^i,\, v_k^i)} \qquad \text{avec} \qquad \sum_{i=1}^N w_k^i = 1\ ,$$

où (ξ_k^i, v_k^i) est une particule dans l'ensemble produit $E \times [0, \infty)$ pour tout $i = 1, \cdots, N$.

(xiv) **Décrire, en utilisant les résultats du cours, l'approximation particulaire de type bootstrap pour la distribution $\mu_k^{\rm e}$ définie sur l'ensemble produit $E \times [0, \infty)$. En déduire, en utilisant les expressions obtenues en réponse à la question (viii), une approximation particulaire de type bootstrap pour les deux distributions μ_k^0 et μ_k.**

Solution

En appliquant les résultats du cours, l'algorithme est décrit par les étapes suivantes :

initialisation : indépendamment pour tout $i = 1, \cdots, N$,

on simule (ξ_0^i, v_0^i) selon $p_0^{\rm e}(dx, dv)$,

c'est–à–dire que :

on simule ξ_0^i selon $p_0(dx)$ et on pose $v_0^i = W_0^{\rm imp}(\xi_0^i)$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_0^i \propto W_0^{\rm red}(\xi_0^i)$,

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

sélection : indépendamment pour tout $i = 1, \cdots, N$, on simule τ_k^i à valeurs dans l'ensemble $\{1, \cdots, N\}$ des indices, selon le vecteur de probabilité $(w_{k-1}^1, \cdots, w_{k-1}^N)$,

mutation : indépendamment pour tout $i = 1, \cdots, N$,

on simule (ξ_k^i, v_k^i) selon $P_k^{\rm e}(\xi_{k-1}^{\tau_k^i}, v_{k-1}^{\tau_k^i}, dx', dv')$,

c'est–à–dire que :

on simule ξ_k^i selon $P_k(\xi_{k-1}^{\tau_k^i}, dx')$ et on pose $v_k^i = v_{k-1}^{\tau_k^i}\, W_k^{\rm imp}(\xi_{k-1}^{\tau_k^i}, \xi_k^i)$,

pondération : pour tout $i = 1, \cdots, N$, on pose $w_k^i \propto W_k^{\rm red}(\xi_{k-1}^{\tau_k^i}, \xi_{k,k}^i)$.

En particulier pour la fonction $F = \phi \otimes e$ introduite dans l'énoncé de la question (viii), on a

$$\langle \mu_k^{{\rm e},N}, \phi \otimes e \rangle = \frac{1}{N} \sum_{i=1}^N w_k^i \; v_k^i \; \phi(\xi_k^i) \ ,$$

et au vu de la réponse à la question (viii), on obtient

$$\langle \mu_k, \phi \rangle = \frac{\langle \mu_k^{\rm e}, \phi \otimes e \rangle}{\langle \mu_k^{\rm e}, 1 \otimes e \rangle} \approx \frac{\langle \mu_k^{{\rm e},N}, \phi \otimes e \rangle}{\langle \mu_k^{{\rm e},N}, 1 \otimes e \rangle} = \frac{\displaystyle\sum_{i=1}^N w_k^i \; v_k^i \; \phi(\xi_k^i)}{\displaystyle\sum_{j=1}^N w_k^j \; v_k^j} = \sum_{i=1}^N w_k^i \; u_k^i \; \phi(\xi_k^i) \ ,$$

et on vérifie automatiquement la contrainte

$$\sum_{i=1}^N w_k^i \; u_k^i = 1 \qquad \text{avec le choix} \qquad u_k^i = \frac{v_k^i}{\displaystyle\sum_{j=1}^N w_k^j \; v_k^j} \qquad \text{pour tout } i = 1, \cdots, N.$$

$\square$