

École Nationale Supérieure de Techniques Avancées  
module : Commande des Systèmes

examen du cours B7–1

“Filtrage bayésien et approximation particulière”

vendredi 23 octobre 2015, 15:45 à 17:15

PROBLÈME

L’objectif de ce problème est d’étudier deux approximations de type particulière pour le calcul de la probabilité de dépassement

$$P_n = \mathbb{P}[V(X_n) \geq a] ,$$

portant sur l’état final  $X_n$  d’une chaîne de Markov  $(X_0, X_1, \dots, X_n)$  à valeurs dans  $E$ , caractérisée par

- la distribution initiale  $\eta(dx)$ ,
- le noyau de transition  $Q_k(x, dx')$  pour tout  $k = 1, \dots, n$ .

Cette formulation apparemment restrictive ne permet à première vue de considérer que des dépassements de niveau à l’instant final.

- (i) **Quitte à considérer une chaîne de Markov augmentée, incluant aussi le maximum courant de la suite  $(V(X_0), V(X_1), \dots, V(X_n))$ , montrer que cette formulation permet en fait de traiter le calcul de la probabilité de dépassement plus générale**

$$P_{0:n} = \mathbb{P}[V(X_k) \geq a, \text{ pour un certain } k = 0, 1, \dots, n] .$$

---

SOLUTION

---

On pose  $X_k^{\max} = (X_k, M_k)$  avec  $M_k = \max(V(X_0), V(X_1), \dots, V(X_k))$  pour tout  $k = 0, 1, \dots, n$ .

On vérifie que toute fonction des variables  $(X_0^{\max}, X_1^{\max}, \dots, X_{k-1}^{\max})$  peut s’exprimer comme une fonction des variables  $(X_0, X_1, \dots, X_{k-1})$  et la réciproque est évidente, de

sorte que la tribu engendrée par  $X_{0:k-1} = (X_0, X_1, \dots, X_{k-1})$  coïncide avec la tribu engendrée par  $X_{0:k-1}^{\max} = (X_0^{\max}, X_1^{\max}, \dots, X_{k-1}^{\max})$ . Compte tenu que  $M_k = \max(M_{k-1}, V(X_k))$  et que  $M_{k-1}$  est une fonction des variables  $(X_0, X_1, \dots, X_{k-1})$  par définition, on obtient

$$\begin{aligned} \mathbb{E}[\phi(X_k, M_k) \mid X_{0:k-1}^{\max}] &= \mathbb{E}[\phi(X_k, M_k) \mid X_{0:k-1}] \\ &= \mathbb{E}[\phi(X_k, \max(M_{k-1}, V(X_k))) \mid X_{0:k-1}] \\ &= \int_E \phi(x', \max(M_{k-1}, V(x'))) \mathbb{P}[X_k \in dx' \mid X_{0:k-1}] \\ &= \int_E \phi(x', \max(M_{k-1}, V(x'))) \mathbb{P}[X_k \in dx' \mid X_{k-1}] , \end{aligned}$$

pour toute fonction mesurable bornée  $\phi$  définie sur l'ensemble produit  $E \times \mathbb{R}$ , et le résultat ne dépend que de  $X_{k-1}^{\max} = (X_{k-1}, M_{k-1})$ , d'où le caractère markovien.

On remarque que  $V(X_k) \geq a$  pour un certain  $k = 0, 1, \dots, n$  si et seulement si  $\max(V(X_0), V(X_1), \dots, V(X_n)) \geq a$ , c'est-à-dire si et seulement si  $M_n \geq a$ , de sorte que la probabilité de dépassement plus générale

$$P_{0:n} = \mathbb{P}[V(X_k) \geq a, \text{ pour un certain } k = 0, 1, \dots, n] = \mathbb{P}[M_n \geq a] ,$$

peut aussi s'interpréter comme une probabilité de dépassement portant sur l'état final d'une chaîne de Markov augmentée.

□

Les deux méthodes proposées dans ce problème reposent sur la remarque suivante, justifiée par la théorie des grandes déviations : pour évaluer la probabilité de dépassement  $P_n$ , une stratégie efficace consiste à utiliser la mesure de probabilité  $\mathbb{P}^\lambda$ , absolument continue par rapport à la mesure de probabilité  $\mathbb{P}$ , et définie par sa dérivée de Radon–Nikodym

$$\frac{d\mathbb{P}^\lambda}{d\mathbb{P}} = \exp\{\lambda (V(X_n) - \Lambda_n(\lambda))\} \quad \text{avec} \quad \Lambda_n(\lambda) = \log \mathbb{E}[\exp\{\lambda V(X_n)\}] .$$

Intuitivement, cette mesure de probabilité  $\mathbb{P}^\lambda$  favorise (pondère davantage) les trajectoires telles que  $V(X_n)$  est grand, qui sont justement les trajectoires susceptibles d'entrer dans l'ensemble critique  $\{x \in E, V(x) \geq a\}$ . Le choix du paramètre  $\lambda$  est une question importante, mais ne sera pas abordé ici : on considérera donc que  $\lambda$  est donné.

## PREMIÈRE MÉTHODE

On introduit les fonctions positives

$$g_k^+(x') = \exp\{\lambda V(x')\} \quad \text{et} \quad g_k^-(x') = \exp\{-\lambda V(x')\} ,$$

définies sur  $E$  pour tout  $k = 0, 1, \dots, n$ . On pose

$$T_k(x_0, x_1, \dots, x_k) = \prod_{p=0}^k g_p^-(x_p),$$

pour tout  $x_0, x_1, \dots, x_k \in E$  et pour tout  $k = 0, 1, \dots, n$ .

(ii) **Montrer que la probabilité de dépassement peut s'exprimer comme**

$$P_n = \mathbb{E}[\mathbf{1}_{(V(X_n) \geq a)} T_n(X_0, X_1, \dots, X_n) \prod_{k=0}^n g_k^+(X_k)],$$

---

SOLUTION

---

Clairement

$$g_k^-(x') g_k^+(x') = \exp\{-\lambda V(x')\} \exp\{\lambda V(x')\} = 1,$$

pour tout  $x' \in E$  et pour tout  $k = 0, 1, \dots, n$ , de sorte que

$$T_n(x_0, x_1, \dots, x_n) \prod_{k=0}^n g_k^+(x_k) = \prod_{k=0}^n g_k^-(x_k) \prod_{k=0}^n g_k^+(x_k) = 1,$$

pour tout  $x_0, x_1, \dots, x_n \in E$ .

□

On pose  $X_k^{\text{mf}} = (X_k, W_k)$  avec  $W_k = T_k(X_0, X_1, \dots, X_k)$  pour tout  $k = 0, 1, \dots, n$ .

(iii) **Montrer que la suite  $(X_0^{\text{mf}}, X_1^{\text{mf}}, \dots, X_n^{\text{mf}})$  forme une chaîne de Markov à valeurs dans l'ensemble produit  $E \times [0, \infty)$ , caractérisée par**

- la distribution initiale définie par

$$\eta_0^{\text{mf}}(dx, dv) = \mathbb{P}[X_0 \in dx, W_0 \in dv] = \eta(dx) \delta_{g_0^-(x)}(dv),$$

- et le noyau de transition défini par

$$\begin{aligned} Q_k^{\text{mf}}(x, v, dx', dv') &= \mathbb{P}[X_k \in dx', W_k \in dv' \mid X_{k-1} = x, W_{k-1} = v] \\ &= Q_k(x, dx') \delta_{v g_k^-(x')} (dv'), \end{aligned}$$

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

On vérifie que toute fonction des variables  $(X_0^{\text{mf}}, X_1^{\text{mf}}, \dots, X_{k-1}^{\text{mf}})$  peut s'exprimer comme une fonction des variables  $(X_0, X_1, \dots, X_{k-1})$  et la réciproque est évidente, de sorte que la tribu engendrée par  $X_{0:k-1} = (X_0, X_1, \dots, X_{k-1})$  coïncide avec la tribu engendrée par  $X_{0:k-1}^{\text{mf}} = (X_0^{\text{mf}}, X_1^{\text{mf}}, \dots, X_{k-1}^{\text{mf}})$ . Compte tenu que  $W_k = W_{k-1} g_k^-(X_k)$  et que  $W_{k-1}$  est une fonction des variables  $(X_0, X_1, \dots, X_{k-1})$  par définition, on obtient

$$\begin{aligned} \mathbb{E}[\phi(X_k, W_k) \mid X_{0:k-1}^{\text{mf}}] &= \mathbb{E}[\phi(X_k, W_k) \mid X_{0:k-1}] \\ &= \mathbb{E}[\phi(X_k, W_{k-1} g_k^-(X_k)) \mid X_{0:k-1}] \\ &= \int_E \phi(x', W_{k-1} g_k^-(x')) \mathbb{P}[X_k \in dx' \mid X_{0:k-1}] \\ &= \int_E \phi(x', W_{k-1} g_k^-(x')) \mathbb{P}[X_k \in dx' \mid X_{k-1}] , \end{aligned}$$

pour toute fonction mesurable bornée  $\phi$  définie sur l'ensemble produit  $E \times [0, \infty)$ , et le résultat ne dépend que de  $X_{k-1}^{\text{mf}} = (X_{k-1}, W_{k-1})$ , d'où le caractère markovien. Il suffit de vérifier ensuite que

$$\begin{aligned} \mathbb{E}[\phi(X_0, W_0)] &= \mathbb{E}[\phi(X_0, g_0^-(X_0))] \\ &= \int_E \phi(x, g_0^-(x)) \eta(dx) \\ &= \int_E \int_0^\infty \phi(x, v) \eta(dx) \delta_{g_0^-(x)}(dv) , \end{aligned}$$

et

$$\begin{aligned} \mathbb{E}[\phi(X_k, W_k) \mid X_{k-1} = x, W_{k-1} = v] &= \mathbb{E}[\phi(X_k, W_{k-1} g_k^-(X_k)) \mid X_{k-1} = x, W_{k-1} = v] \\ &= \int_E \phi(x', v g_k^-(x')) \mathbb{P}[X_k \in dx' \mid X_{k-1} = x] \\ &= \int_E \phi(x', v g_k^-(x')) Q_k(x, dx') \\ &= \int_E \int_0^\infty \phi(x', v') Q_k(x, dx') \delta_{v g_k^-(x')} (dv') , \end{aligned}$$

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

---

□

Clairement la probabilité de dépassement peut s'exprimer comme

$$P_n = \mathbb{E}[\mathbf{1}_{(V(X_n) \geq a)} W_n \prod_{k=0}^n g_k^+(X_k)] .$$

On introduit les distributions de Feynman–Kac non-normalisées, et les distributions de Feynman–Kac normalisées associées, définies par

$$\langle \gamma_k^{\text{mf}}, F \rangle = \mathbb{E}[F(X_k, W_k) \prod_{p=0}^k g_p^+(X_p)] \quad \text{et} \quad \langle \mu_k^{\text{mf}}, F \rangle = \frac{\langle \gamma_k^{\text{mf}}, F \rangle}{\langle \gamma_k^{\text{mf}}, 1 \rangle} , \quad (\star)$$

pour toute fonction  $F$  définie sur l'ensemble produit  $E \times [0, \infty)$ , pour tout  $k = 0, 1, \dots, n$ .

- (iv) **Montrer que la probabilité de dépassement peut s'exprimer comme  $P_n = \langle \gamma_n^{\text{mf}}, F_a \rangle$  pour une fonction  $F_a$  particulière dont on donnera l'expression, définie sur l'ensemble produit  $E \times [0, \infty)$ .**

---

SOLUTION

---

Il suffit de considérer la fonction définie par

$$F_a(x, v) = \mathbf{1}_{(V(x) \geq a)} v ,$$

pour tout  $x \in E$  et pour tout  $v \geq 0$ .

---

□

- (v) **Mettre en œuvre l'algorithme SIR et expliciter l'approximation particulière obtenue pour les distributions de Feynman–Kac définies en  $(\star)$ . En déduire une première approximation particulière de la probabilité de dépassement  $P_n$ . Décrire de manière intuitive le comportement qualitatif de l'algorithme ainsi obtenu.**

---

SOLUTION

---

On part de la représentation

$$\begin{aligned} P_n &= \langle \gamma_n^{\text{mf}}, F_a \rangle \\ &= \langle \mu_n^{\text{mf}}, F_a \rangle \langle \gamma_n^{\text{mf}}, 1 \rangle = \frac{\langle \eta_n^{\text{mf}}, g_n^+ F_a \rangle}{\langle \eta_n^{\text{mf}}, g_n^+ \rangle} \prod_{k=0}^n \langle \eta_k^{\text{mf}}, g_k^+ \rangle = \langle \eta_n^{\text{mf}}, g_n^+ F_a \rangle \prod_{k=0}^{n-1} \langle \eta_k^{\text{mf}}, g_k^+ \rangle , \end{aligned}$$

et on recherche une approximation des distributions normalisées  $\eta_k^{\text{mf}}$  et  $\mu_k^{\text{mf}}$  par des distributions empiriques (éventuellement pondérées) de la forme

$$\eta_k^{\text{mf},N} = \frac{1}{N} \sum_{i=1}^N \delta_{(\xi_k^i, v_k^i)} \quad \text{et} \quad \mu_k^{\text{mf},N} = \sum_{i=1}^N w_k^i \delta_{(\xi_k^i, v_k^i)},$$

respectivement, pour tout  $k = 0, 1, \dots, n$ . Les étapes successives de l'algorithme SIR sont décrites de la manière suivante.

Pour  $k = 0$ , indépendamment pour  $i = 1, \dots, N$

- on simule  $z_0^i = (\xi_0^i, v_0^i)$  selon  $\eta_0^{\text{mf}}(dx, dv)$ , c'est-à-dire que on simule  $\xi_0^i$  selon  $\eta(dx)$  et on pose  $v_0^i = g_0^-(\xi_0^i)$ ,
- on évalue le poids  $w_0^i \propto g_0^+(\xi_0^i)$ ,
- on approxime la constante de normalisation  $\alpha_0 = \langle \eta_0^{\text{mf}}, g_0^+ \rangle$  par la moyenne empirique

$$\alpha_0^N = \frac{1}{N} \sum_{i=1}^N g_0^+(\xi_0^i).$$

Pour  $k = 1, \dots, n$ , indépendamment pour  $i = 1, \dots, N$

- on choisit un individu  $\widehat{z}_{k-1}^i = (\widehat{\xi}_{k-1}^i, \widehat{v}_{k-1}^i)$  au sein de la population  $(z_{k-1}^1, \dots, z_{k-1}^N)$  et en fonction des poids respectifs  $(w_{k-1}^1, \dots, w_{k-1}^N)$ ,
- on simule  $z_k^i = (\xi_k^i, v_k^i)$  selon  $Q_k^{\text{mf}}(\widehat{\xi}_{k-1}^i, \widehat{v}_{k-1}^i, dx', dv')$ , c'est-à-dire que on simule  $\xi_k^i$  selon  $Q_k(\widehat{\xi}_{k-1}^i, dx')$  et on pose  $v_k^i = \widehat{v}_{k-1}^i g_k^-(\xi_k^i)$ ,
- on évalue le poids  $w_k^i \propto g_k^+(\xi_k^i)$ ,
- on approxime la constante de normalisation  $\alpha_k = \langle \eta_k^{\text{mf}}, g_k^+ \rangle$  par la moyenne empirique

$$\alpha_k^N = \frac{1}{N} \sum_{i=1}^N g_k^+(\xi_k^i).$$

Finalement, pour  $k = n$ , on approxime la moyenne  $\beta_n = \langle \eta_n^{\text{mf}}, g_n^+ F_a \rangle$  par la moyenne empirique

$$\beta_n^N = \frac{1}{N} \sum_{i=1}^N g_n^+(\xi_n^i) F_a(\xi_n^i, v_n^i).$$

En utilisant l'expression explicite des différentes fonctions, les étapes successives de l'algorithme sont reformulées de la manière suivante.

Pour  $k = 0$ , indépendamment pour  $i = 1, \dots, N$

- on simule  $\xi_0^i$  selon  $\eta(dx)$  et on pose  $v_0^i = \exp\{-\lambda V(\xi_0^i)\}$ ,
- on évalue le poids  $w_0^i \propto \exp\{\lambda V(\xi_0^i)\}$ ,
- on approxime la constante de normalisation  $\alpha_0$  par la moyenne empirique

$$\alpha_0^N = \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_0^i)\} .$$

Pour  $k = 1, \dots, n$ , indépendamment pour  $i = 1, \dots, N$

- on choisit un individu  $\widehat{z}_{k-1}^i = (\widehat{\xi}_{k-1}^i, \widehat{v}_{k-1}^i)$  au sein de la population  $(z_{k-1}^1, \dots, z_{k-1}^N)$  et en fonction des poids respectifs  $(w_{k-1}^1, \dots, w_{k-1}^N)$ ,
- on simule  $\xi_k^i$  selon  $Q_k(\widehat{\xi}_{k-1}^i, dx')$  et on pose  $v_k^i = \widehat{v}_{k-1}^i \exp\{-\lambda V(\xi_k^i)\}$ ,
- on évalue le poids  $w_k^i \propto \exp\{\lambda V(\xi_k^i)\}$ ,
- on approxime la constante de normalisation  $\alpha_k$  par la moyenne empirique

$$\alpha_k^N = \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_k^i)\} .$$

Finalement, pour  $k = n$ , on approxime la moyenne  $\beta_n$  par la moyenne empirique

$$\beta_n^N = \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_n^i)\} \mathbf{1}_{(V(\xi_n^i) \geq a)} v_n^i ,$$

et la probabilité de dépassement  $P_n = \alpha_0 \alpha_1 \dots \alpha_{n-1} \beta_n$  par  $P_n^N = \alpha_0^N \alpha_1^N \dots \alpha_{n-1}^N \beta_n^N$ , c'est-à-dire par

$$P_n^N = \prod_{k=0}^{n-1} \left[ \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_k^i)\} \right] \\ \times \left[ \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_n^i)\} \mathbf{1}_{(V(\xi_n^i) \geq a)} v_n^i \right] .$$

Intuitivement, cet algorithme favorise à chaque instant (sélectionne davantage) les particules qui donnent une grande valeur à la fonction  $V$ , qui sont justement les particules susceptibles de dépasser le niveau  $a$ . En contre-partie, et pour compenser le biais apporté par cette stratégie de sélection, chaque particule est affectée à l'instant final d'un poids qui est évalué de manière multiplicative le long de la trajectoire dont elle est issue : l'effet de ce poids est limité à la pondération et ne contribue pas à la sélection.

□

DEUXIÈME MÉTHODE

On introduit les fonctions positives

$$g_0(x, x') = g_0(x') = \exp\{\lambda V(x')\} \quad \text{et} \quad g_k(x, x') = \exp\{\lambda (V(x') - V(x))\},$$

définies sur  $E \times E$  pour tout  $k = 0, 1, \dots, n$ .

(vi) **Montrer que la probabilité de dépassement peut s'exprimer comme**

$$P_n = \mathbb{E}[\theta_a(X_n) \prod_{k=0}^n g_k(X_{k-1}, X_k)],$$

avec

$$\theta_a(x) = \mathbf{1}_{(V(x) \geq a)} \exp\{-\lambda V(x)\},$$

pour tout  $x \in E$ .

---

SOLUTION

---

Clairement

$$\prod_{k=0}^n g_k(x_{k-1}, x_k) = \exp\{\lambda V(x_n)\},$$

comme produit télescopique, de sorte que

$$\begin{aligned} \theta_a(x_n) \prod_{k=0}^n g_k(x_{k-1}, x_k) &= \mathbf{1}_{(V(x_n) \geq a)} \exp\{-\lambda V(x_n)\} \exp\{\lambda V(x_n)\} \\ &= \mathbf{1}_{(V(x_n) \geq a)}, \end{aligned}$$

pour tout  $x_0, x_1, \dots, x_n \in E$ .

□

On introduit les distributions de Feynman–Kac non-normalisées, et les distributions de Feynman–Kac normalisées associées, définies par

$$\langle \gamma_k, \phi \rangle = \mathbb{E}[\phi(X_k) \prod_{p=0}^k g_p(X_{p-1}, X_p)] \quad \text{et} \quad \langle \mu_k, \phi \rangle = \frac{\langle \gamma_k, \phi \rangle}{\langle \gamma_k, 1 \rangle}, \quad (**)$$

pour toute fonction  $\phi$  définie sur  $E$ , pour tout  $k = 0, 1, \dots, n$ .

Clairement, la probabilité de dépassement peut s'exprimer comme  $P_n = \langle \gamma_n, \theta_a \rangle$ .



- (vii) Mettre en œuvre l’algorithme SIR et expliciter l’approximation particulière obtenue pour les distributions de Feynman–Kac définies en (\*\*). En déduire une deuxième approximation particulière de la probabilité de dépassement  $P_n$ . Décrire de manière intuitive le comportement qualitatif de l’algorithme ainsi obtenu.

---

SOLUTION

---

On remarque que les fonctions de pondération / sélection dépendent ici de l’état de la chaîne de Markov à deux instants successifs, et on reformule le problème en terme d’une chaîne de Markov à valeur transitions.

On pose  $X_0^{\text{tr}} = X_0$  et  $X_k^{\text{tr}} = (X_{k-1}, X_k)$  pour tout  $k = 1, \dots, n$ , et on rappelle que la suite  $(X_0^{\text{tr}}, X_1^{\text{tr}}, \dots, X_n^{\text{tr}})$  forme une chaîne de Markov à valeurs dans  $E \times E$ , caractérisée par

- la distribution initiale définie par

$$\eta_0^{\text{tr}}(dx) = \eta(dx) ,$$

- et le noyau de transition défini par

$$Q_k^{\text{tr}}(x_1, x_2, dx'_1, dx'_2) = \delta_{x_2}(dx'_1) Q_k(x'_1, dx'_2) ,$$

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

On introduit les distributions de Feynman–Kac non-normalisées, et les distributions de Feynman–Kac normalisées associées, définies par

$$\langle \gamma_k^{\text{tr}}, f \rangle = \mathbb{E}[f(X_k^{\text{tr}}) \prod_{p=0}^k g_p(X_p^{\text{tr}})] \quad \text{et} \quad \langle \mu_k^{\text{tr}}, f \rangle = \frac{\langle \gamma_k^{\text{tr}}, f \rangle}{\langle \gamma_k^{\text{tr}}, 1 \rangle} ,$$

pour toute fonction  $f$  définie sur  $E \times E$ , pour tout  $k = 0, 1, \dots, n$ .

On part de la représentation

$$P_n = \langle \gamma_n, \theta_a \rangle = \langle \gamma_n^{\text{tr}}, \theta_a \circ \pi \rangle$$

$$= \langle \mu_n^{\text{tr}}, \theta_a \circ \pi \rangle \langle \gamma_n^{\text{tr}}, 1 \rangle = \frac{\langle \eta_n^{\text{tr}}, g_n \theta_a \circ \pi \rangle}{\langle \eta_n^{\text{tr}}, g_n \rangle} \prod_{k=0}^n \langle \eta_k^{\text{tr}}, g_k \rangle = \langle \eta_n^{\text{tr}}, g_n \theta_a \circ \pi \rangle \prod_{k=0}^{n-1} \langle \eta_k^{\text{tr}}, g_k \rangle ,$$

où l’application  $\pi$  définie sur  $E \times E$  désigne la projection sur la deuxième coordonnée (correspondant à l’état d’arrivée de la transition), et on recherche une approximation

des distributions normalisées  $\eta_k^{\text{tr}}$  et  $\mu_k^{\text{tr}}$  par des distributions empiriques (éventuellement pondérées) de la forme

$$\eta_k^{\text{tr},N} = \frac{1}{N} \sum_{i=1}^N \delta(\xi_k^{1,i}, \xi_k^{2,i}) \quad \text{et} \quad \mu_k^{\text{tr},N} = \sum_{i=1}^N w_k^i \delta(\xi_k^{1,i}, \xi_k^{2,i}),$$

respectivement, pour tout  $k = 0, 1, \dots, n$ . Les étapes successives de l'algorithme SIR sont décrites de la manière suivante.

Pour  $k = 0$ , indépendamment pour  $i = 1, \dots, N$

- on simule  $\xi_0^{\text{tr},i} = \xi_0^i$  selon  $\eta_0^{\text{tr}}(dx)$  c'est-à-dire que on simule  $\xi_0^i$  selon  $\eta(dx)$ ,
- on évalue le poids  $w_0^i \propto g_0(\xi_0^{\text{tr},i})$  c'est-à-dire que on pose  $w_0^i \propto g_0(\xi_0^i)$ ,
- on approxime la constante de normalisation  $\alpha_0 = \langle \eta_0^{\text{tr}}, g_0 \rangle$  par la moyenne empirique

$$\alpha_0^N = \frac{1}{N} \sum_{i=1}^N g_0(\xi_0^i).$$

Pour  $k = 1, \dots, n$ , indépendamment pour  $i = 1, \dots, N$

- on choisit un individu  $\widehat{\xi}_{k-1}^{\text{tr},i} = (\widehat{\xi}_{k-1}^{1,i}, \widehat{\xi}_{k-1}^{2,i})$  au sein de la population  $(\xi_{k-1}^{\text{tr},1}, \dots, \xi_{k-1}^{\text{tr},N})$  et en fonction des poids respectifs  $(w_{k-1}^1, \dots, w_{k-1}^N)$
- on simule  $\xi_k^{\text{tr},i} = (\xi_k^{1,i}, \xi_k^{2,i})$  selon  $Q_k^{\text{tr}}(\widehat{\xi}_{k-1}^{1,i}, \widehat{\xi}_{k-1}^{2,i}, dx'_1, dx'_2)$ , c'est-à-dire que on pose  $\xi_k^{1,i} = \widehat{\xi}_{k-1}^{2,i}$  et on simule  $\xi_k^{2,i}$  selon  $Q_k(\widehat{\xi}_{k-1}^{2,i}, dx')$ ,
- on évalue le poids  $w_k^i \propto g_k(\xi_k^{\text{tr},i})$ , c'est-à-dire que on pose  $w_k^i \propto g_k(\widehat{\xi}_{k-1}^{2,i}, \xi_k^{2,i})$ ,
- on approxime la constante de normalisation  $\alpha_k = \langle \eta_k^{\text{tr}}, g_k \rangle$  par la moyenne empirique

$$\alpha_k^N = \frac{1}{N} \sum_{i=1}^N g_k(\widehat{\xi}_{k-1}^{2,i}, \xi_k^{2,i}).$$

Finalement, pour  $k = n$ , on approxime la moyenne  $\beta_n = \langle \eta_n^{\text{tr}}, g_n \theta_a \circ \pi \rangle$  par la moyenne empirique

$$\beta_n^N = \frac{1}{N} \sum_{i=1}^N g_n(\widehat{\xi}_{n-1}^{2,i}, \xi_n^{2,i}) \theta_a(\xi_n^{2,i}).$$

En remarquant que toutes les approximations obtenues s'expriment en terme de l'état d'arrivée seulement des transitions, les étapes successives de l'algorithme peuvent être reformulées de la manière suivante.

Pour  $k = 0$ , indépendamment pour  $i = 1, \dots, N$

- on simule  $\xi_0^i$  selon  $\eta(dx)$ ,
- on évalue le poids  $w_0^i \propto g_0(\xi_0^i)$ ,
- on approxime la constante de normalisation  $\alpha_0$  par la moyenne empirique

$$\alpha_0^N = \frac{1}{N} \sum_{i=1}^N g_0(\xi_0^i) .$$

Pour  $k = 1, \dots, n$ , indépendamment pour  $i = 1, \dots, N$

- on choisit un individu  $\widehat{\xi}_{k-1}^i$  au sein de la population  $(\xi_{k-1}^1, \dots, \xi_{k-1}^N)$  et en fonction des poids respectifs  $(w_{k-1}^1, \dots, w_{k-1}^N)$
- on simule  $\xi_k^i$  selon  $Q_k(\widehat{\xi}_k^i, dx')$ ,
- on évalue le poids  $w_k^i \propto g_k(\widehat{\xi}_k^i, \xi_k^i)$ ,
- on approxime la constante de normalisation  $\alpha_k$  par la moyenne empirique

$$\alpha_k^N = \frac{1}{N} \sum_{i=1}^N g_k(\widehat{\xi}_k^i, \xi_k^i) .$$

Finalement, pour  $k = n$ , on approxime la moyenne  $\beta_n$  par la moyenne empirique

$$\beta_n^N = \frac{1}{N} \sum_{i=1}^N g_n(\widehat{\xi}_n^i, \xi_n^i) \theta_a(\xi_n^i) .$$

En utilisant l'expression explicite des différentes fonctions, les étapes successives de l'algorithme sont finalement reformulées de la manière suivante.

Pour  $k = 0$ , indépendamment pour  $i = 1, \dots, N$

- on simule  $\xi_0^i$  selon  $\eta(dx)$ ,
- on évalue le poids  $w_0^i \propto \exp\{\lambda V(\xi_0^i)\}$ ,
- on approxime la constante de normalisation  $\alpha_0$  par la moyenne empirique

$$\alpha_0^N = \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_0^i)\} .$$

Pour  $k = 1, \dots, n$ , indépendamment pour  $i = 1, \dots, N$

- on choisit un individu  $\widehat{\xi}_{k-1}^i$  au sein de la population  $(\xi_{k-1}^1, \dots, \xi_{k-1}^N)$  et en fonction des poids respectifs  $(w_{k-1}^1, \dots, w_{k-1}^N)$ ,
- on simule  $\xi_k^i$  selon  $Q_k(\widehat{\xi}_{k-1}^i, dx')$ ,
- on évalue le poids  $w_k^i \propto \exp\{\lambda(V(\xi_k^i) - V(\widehat{\xi}_{k-1}^i))\}$ ,
- on approxime la constante de normalisation  $\alpha_k$  par la moyenne empirique

$$\alpha_k^N = \frac{1}{N} \sum_{i=1}^N \exp\{\lambda(V(\xi_k^i) - V(\widehat{\xi}_{k-1}^i))\},$$

Finalement, pour  $k = n$ , on approxime la moyenne  $\beta_n$  par la moyenne empirique

$$\begin{aligned} \beta_n^N &= \frac{1}{N} \sum_{i=1}^N \exp\{\lambda(V(\xi_n^i) - V(\widehat{\xi}_{n-1}^i))\} \mathbf{1}_{(V(\xi_n^i) \geq a)} \exp\{-\lambda V(\xi_n^i)\} \\ &= \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(V(\xi_n^i) \geq a)} \exp\{-\lambda V(\widehat{\xi}_{n-1}^i)\}, \end{aligned}$$

et la probabilité de dépassement  $P_n = \alpha_0 \alpha_1 \dots \alpha_{n-1} \beta_n$  par  $P_n^N = \alpha_0^N \alpha_1^N \dots \alpha_{n-1}^N \beta_n^N$ , c'est-à-dire par

$$\begin{aligned} P_n^N &= \left[ \frac{1}{N} \sum_{i=1}^N \exp\{\lambda V(\xi_0^i)\} \right] \times \prod_{k=1}^{n-1} \left[ \frac{1}{N} \sum_{i=1}^N \exp\{\lambda(V(\xi_k^i) - V(\widehat{\xi}_{k-1}^i))\} \right] \\ &\quad \times \left[ \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(V(\xi_n^i) \geq a)} \exp\{-\lambda V(\widehat{\xi}_{n-1}^i)\} \right]. \end{aligned}$$

Intuitivement, cet algorithme favorise à chaque instant (sélectionne davantage) les particules qui permettent d'augmenter la valeur de la fonction  $V$  par rapport à l'instant précédent, En contre-partie, et pour compenser le biais apporté par cette stratégie de sélection, chaque particule est affectée à l'instant final d'un poids qui est évalué pour l'état de départ seulement de la dernière transition : l'effet de ce poids est limité à la pondération et ne contribue pas à la sélection. □

Ces deux méthodes ont été proposées, et leurs performances respectives ont été étudiées et comparées par Del Moral et Garnier [2]. Leur stratégie commune est d'introduire artificiellement (là où il n'y avait initialement aucun poids, ou dit autrement seulement un poids égal à 1) des poids

- dont le produit est égal à 1, c'est-à-dire qui réalisent une *partition de l'unité*,
- et dont une fraction est utilisée opportunément pour sélectionner des échantillons, tandis que la fraction restante est utilisée pour pondérer les échantillons ainsi sélectionnés.

Les deux méthodes diffèrent dans leur manière de regrouper les poids dédiés à la sélection. La deuxième méthode a été utilisée et mise-en-œuvre par Carmona, Fouque et Vestal [1] pour évaluer numériquement un risque de crédit.

## Références bibliographiques

- [1] René A. Carmona, Jean-Pierre Fouque, and Douglas Vestal. Interacting particle systems for the computation of rare credit portfolio losses. *Finance and Stochastics*, 13(4 (Special issue on Computational Methods in Finance (Part II))):613–633, September 2009.
- [2] Pierre Del Moral and Josselin Garnier. Genealogical particle analysis of rare events. *The Annals of Applied Probability*, 15(4):2496–2534, November 2005.