cours SOD333

Filtrage Bayésien et Approximation Particulaire

Simulation d'Évènements Rares

François Le Gland
INRIA Rennes et IRMAR

people.rennes.inria.fr/Francois.Le_Gland/ensta/

15 octobre 2021

Simulation d'évènements rares

- motivation
- splitting (principe)
- splitting (modélisation)
- distributions de Feynman–Kac multi–niveaux

Approximation particulaire

Motivation : ATM

air–traffic management (ATM) étude dans le cadre des projets européens HYBRIDGE (programme IST) et iFLY (programme Aerospace), partenaires : National Aerospace Laboratory (NLR), Twente University, DSNA/DTI, INRIA, etc.

- les plans de vols nominaux respectent la distance de séparation minimale autorisée (5 miles nautiques, environ 9260 mètres)
- la présence de perturbations aléatoires, essentiellement dues aux conditions atmosphériques, mais dues aussi aux incidents mécaniques, aux erreurs humaines, etc., fait que la distance de séparation réelle peut devenir plus petite que la distance de séparation minimale autorisée
- risque de conflit / collision non nul, quoique très petit
- objectif : évaluer si la conception des plans de vol nominaux peut être assouplie, si davantage d'autonomie peut être accordée aux pilotes, de façon à accroître le trafic, mais sans compromettre la sécurité

Splitting : principe

processus de Markov fort à temps continu $\{X(t)\}$ à valeurs dans E
temps d'atteinte d'une région critique $B \subset E$

$$T_B = \inf\{t \geq 0 \, : \, X(t) \in B\}$$

par exemple

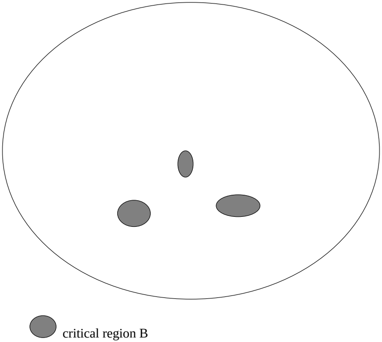
$$B = \{x \in E \, : \, u(x) \geq c\} \quad \text{auquel cas} \quad \mathbb{P}[T_B \leq T] = \mathbb{P}[\sup_{0 \leq t \leq T} u(X(t)) \geq c]$$

objectif : calculer numériquement

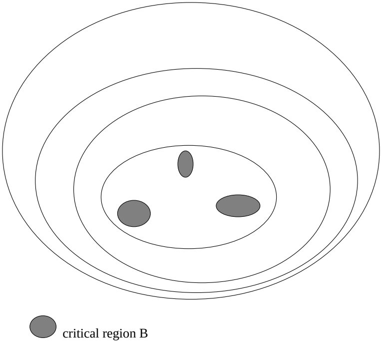
$$P_B = \mathbb{P}[T_B \leq T] \ll 1 \quad \text{et} \quad \mathbb{E}[f(X(t)\,,\, 0 \leq t \leq T_B) \mid T_B \leq T]$$

avec T un instant déterministe ou bien un temps d'arrêt p.s.–fini

- probabilité de l'évènement rare mais critique $(T_B \leq T)$
- et distribution de probabilité des trajectoires critiques $(X(t)\,,\, 0 \leq t \leq T_B)$



en pratique, aucune des trajectoires simulées n'atteindra jamais la région critique, et la méthode de Monte Carlo naïve échouera



splitting (ou branchement multi–niveaux) : sélectionner les trajectoires qui réussissent à atteindre les niveaux intermédiaires (moins critiques), et les autoriser à poursuivre

- branchement donné
- effort donné
- succès donné

quelques problèmes

- à niveaux intermédiaires donnés, variance asymptotique des estimateurs
- forme des régions intermédiaires (fonction d'importance)
- seuils (niveaux) pour les régions intermédiaires
- nombre de régions intermédiaires

Splitting : modélisation

suite emboîtée de régions intermédiaires de plus en plus critiques

$$E \equiv A_0(t) \supset A_1(t) \supset \cdots \supset A_n(t) \equiv B \qquad \text{pour tout } 0 \leq t \leq T$$

par exemple, en termes d'une fonction d'importance

$$A_k(t) = \{x \in E \,:\, u(t,x) \geq c_k\}$$

temps d'atteinte associés (avec la convention $\inf \emptyset = +\infty$)

$$\begin{aligned} T_k &= \inf\{t \geq 0 \,:\, X(t) \in A_k(t)\} \\ &= \inf\{t \geq 0 \,:\, u(t, X(t)) \geq c_k\} \qquad \text{[par exemple]} \end{aligned}$$

par construction

$$0 = T_0 \leq T_1 \leq \cdots T_k \leq \cdots \leq T_n = T_B$$

en introduisant

$$A_k = \bigcup_{t \in [0,T]} \{t\} \times A_k(t)$$

de sorte que

$$(t,x) \in A_k \qquad \text{ssi} \qquad x \in A_k(t)$$

on a la formulation équivalente

$$\begin{aligned} T_k &= \inf\{t \geq 0 \,:\, X(t) \in A_k(t)\} \\ &= \inf\{t \geq 0 \,:\, (t, X(t)) \in A_k\} \end{aligned}$$

et nécessairement

$$A_0 \supset A_1 \supset \cdots \supset A_n \equiv [0,T] \times B$$

factorisation triviale

$$
\begin{aligned}
P_B = \mathbb{P}[T_B \leq T] &= \mathbb{P}[T_n \leq T] \\
&= \mathbb{P}[T_n \leq T, \cdots, T_1 \leq T] \\
&= \prod_{k=1}^{n} \underbrace{\mathbb{P}[T_k \leq T \mid T_{k-1} \leq T]}_{p_k}
\end{aligned}
$$

en pratique, les probabilités de transition p_k sont inconnues et doivent être apprises modélisation : on définit la v.a. $\Xi_k = (T_k \wedge T, X(T_k \wedge T))$ à valeurs dans l'ensemble produit $E' = [0, \infty) \times E$

la suite $\{\Xi_k\}$ est une chaîne de Markov, indexée par les niveaux $k = 0, 1 \cdots n$, caractérisée par

- la distribution de probabilité initiale $\delta_0 \times \eta_0$
- et le noyau de transition

$$\begin{aligned} Q_k(t, x, dt', dx') \; &= \; \mathbb{P}[T_k \wedge T \in dt', X(T_k \wedge T) \in dx' \mid \\ &\qquad T_{k-1} \wedge T = t, X(T_{k-1} \wedge T) = x] \end{aligned}$$

fonction de sélection

$$g_k(t, x) = 1_{(x \in A_k(t))} \quad \text{de sorte que} \quad g_k(\Xi_k) = 1_{(T_k \leq T)}$$

en effet

$$T_k \leq T \quad \text{ssi} \quad X(T_k \wedge T) \in A_k(T_k \wedge T) \quad \text{ssi} \quad g_k(\Xi_k) = 1$$

modélisation : on définit la v.a. $\Xi_k = (T_k \wedge T, X(T_k \wedge T))$ à valeurs dans l'ensemble produit $E' = [0, \infty) \times E$
la suite $\{\Xi_k\}$ est une chaîne de Markov, indexée par les niveaux $k = 0, 1 \cdots n$, caractérisée par

- la distribution de probabilité initiale $\delta_0 \times \eta_0$
- et le noyau de transition

$$Q_k(t, x, dt', dx') \;=\; \mathbb{P}[T_k \wedge T \in dt', X(T_k \wedge T) \in dx' \mid T_{k-1} \wedge T = t, X(T_{k-1} \wedge T) = x]$$

fonction de sélection

$$g_k(t, x) = 1_{(u(t,x) \geq c_k)} \quad \text{de sorte que} \quad g_k(\Xi_k) = 1_{(T_k \leq T)}$$

en effet

$$T_k \leq T \quad \text{ssi} \quad u(T_k \wedge T, X(T_k \wedge T)) \geq c_k \quad \text{ssi} \quad g_k(\Xi_k) = 1$$

comment simuler une v.a. Ξ distribuée selon $Q_k(t,x,dt',dx')$?
on remarque que

$$\begin{aligned} Q_k(t,x,dt',dx') &= \mathbb{P}[T_k \wedge T \in dt', X(T_k \wedge T) \in dx' \mid \\ &\qquad T_{k-1} \wedge T = t, X(T_{k-1} \wedge T) = x] \\ &= \mathbb{P}[T_k \wedge T \in dt', X(T_k \wedge T) \in dx' \mid X(t) = x] \end{aligned}$$

il suffit donc (i) de simuler une trajectoire $\{X(s)\}$ partant de x à l'instant t

- jusqu'à ce qu'elle atteigne la région intermédiaire suivante A_k, i.e. jusqu'à l'instant

$$T_k = \inf\{s \geq t \,:\, X(s) \in A_k(s)\} = \inf\{s \geq t \,:\, u(s, X(s)) \geq c_k\}$$

- ou bien jusqu'à l'instant final T

selon que l'un ou l'autre instant arrive en premier
et (ii) de poser $\Xi = (T_k \wedge T, X(T_k \wedge T))$

Distributions de Feynman–Kac multi–niveaux

interprétation des probabilités d'évènements rares en termes de distributions de Feynman–Kac non–normalisées

$$\langle \gamma_k, \phi \rangle = \mathbb{E}[\phi(T_k, X(T_k))\, 1_{(T_k \leq T)}] = \mathbb{E}[\phi(\Xi_k)\, g_k(\Xi_k)]$$

en particulier pour $\phi \equiv 1$ (constante de normalisation)

$$\langle \gamma_k, 1 \rangle = \mathbb{P}[T_k \leq T]$$

de sorte que (distribution normalisée)

$$\langle \mu_k, \phi \rangle = \frac{\langle \gamma_k, \phi \rangle}{\langle \gamma_k, 1 \rangle} = \mathbb{E}[\phi(T_k, X(T_k)) \mid T_k \leq T] = \mathbb{E}[\phi(\Xi_k) \mid T_k \leq T]$$

interprétation comme distribution d'entrée dans le niveau A_k

$$\mu_k(dt, dx) = \mathbb{P}[T_k \in dt, X(T_k) \in dx \mid T_k \leq T]$$

de même, en introduisant

$$\langle \eta_k, \phi \rangle = \mathbb{E}[\phi(T_k \wedge T, X(T_k \wedge T)) \mid T_{k-1} \leq T\,] = \mathbb{E}[\phi(\Xi_k) \mid T_{k-1} \leq T\,]$$

avec l'interprétation

$$\eta_k(dt, dx) = \mathbb{P}[T_k \wedge T \in dt, X(T_k \wedge T) \in dx \mid T_{k-1} \leq T\,]$$

on remarque que

$$\begin{aligned}
\langle \gamma_k, \phi \rangle &= \mathbb{E}[\phi(\Xi_k)\, g_k(\Xi_k)] \\
&= \mathbb{E}[\phi(\Xi_k)\, 1_{(\,T_k \,\leq\, T)}\,] \\
&= \mathbb{E}[\phi(\Xi_k)\, 1_{(\,T_{k-1} \,\leq\, T\,,\, T_k \,\leq\, T)}\,] \\
&= \mathbb{E}[\phi(\Xi_k)\, g_k(\Xi_k)\, 1_{(\,T_{k-1} \,\leq\, T)}\,] \\
&= \underbrace{\mathbb{E}[\phi(\Xi_k)\, g_k(\Xi_k) \mid T_{k-1} \leq T]}_{\langle \eta_k,\, g_k\, \phi \rangle}\, \underbrace{\mathbb{P}[T_{k-1} \leq T]}_{\langle \gamma_{k-1},\, 1 \rangle}
\end{aligned}$$

en particulier pour $\phi \equiv 1$, on obtient la probabilité de transition

$$\langle \eta_k, g_k \rangle = \mathbb{E}[g_k(\Xi_k) \mid T_{k-1} \leq T] = \mathbb{P}[T_k \leq T \mid T_{k-1} \leq T] = p_k$$

et la relation de récurrence

$$\langle \gamma_k, 1 \rangle = \langle \eta_k, g_k \rangle \, \langle \gamma_{k-1}, 1 \rangle$$

et en itérant cette relation

$$P_B = \mathbb{P}[T_B \leq T] = \mathbb{P}[T_n \leq T] = \langle \gamma_n, 1 \rangle = \prod_{k=1}^{n} \langle \eta_k, g_k \rangle$$

en divisant membre à membre, on obtient également

$$\langle \mu_k, \phi \rangle = \frac{\langle \gamma_k, \phi \rangle}{\langle \gamma_k, 1 \rangle} = \frac{\langle \eta_k, g_k \, \phi \rangle}{\langle \eta_k, g_k \rangle}$$

i.e.

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

on remarque aussi que

$$\langle \eta_k , \phi \rangle = \frac{\mathbb{E}[\phi(\Xi_k)\, 1_{(T_{k-1} \leq T)}]}{\mathbb{P}[T_{k-1} \leq T]} = \frac{\mathbb{E}[\phi(\Xi_k)\, g_{k-1}(\Xi_{k-1})]}{\mathbb{P}[T_{k-1} \leq T]}$$

et d'après la propriété de Markov

$$\begin{aligned} \mathbb{E}[\phi(\Xi_k)\, g_{k-1}(\Xi_{k-1})] &= \mathbb{E}[\mathbb{E}[\phi(\Xi_k) \mid \Xi_{k-1}]\, g_{k-1}(\Xi_{k-1})] \\ &= \mathbb{E}[Q_k\, \phi(\Xi_{k-1})\, g_{k-1}(\Xi_{k-1})] \\ &= \mathbb{E}[Q_k\, \phi(\Xi_{k-1})\, 1_{(T_{k-1} \leq T)}] \end{aligned}$$

de sorte que

$$\begin{aligned} \langle \eta_k , \phi \rangle &= \frac{\mathbb{E}[Q_k\, \phi(\Xi_{k-1})\, 1_{(T_{k-1} \leq T)}]}{\mathbb{P}[T_{k-1} \leq T]} \\ &= \mathbb{E}[Q_k\, \phi(\Xi_{k-1}) \mid T_{k-1} \leq T] = \langle \mu_{k-1} , Q_k\, \phi \rangle = \langle \mu_{k-1}\, Q_k , \phi \rangle \end{aligned}$$

i.e.

$$\eta_k = \mu_{k-1}\, Q_k$$

en résumé, la distribution d'entrée

$$\mu_k(dt, dx) = \mathbb{P}[T_k \in dt, X(T_k) \in dx \mid T_k \leq T]$$

dans le niveau A_k vérifie une relation de récurrence à la Feynman–Kac

$$\mu_{k-1} \xrightarrow{\text{exploration}} \eta_k = \mu_{k-1}\, Q_k \xrightarrow{\text{vérification}} \mu_k = g_k \cdot \eta_k$$

les méthodes particulaires fournissent des approximations numériques pour

- la probabilité de l'évènement rare

$$P_B = \mathbb{P}[T_B \leq T] = \langle \gamma_n, 1 \rangle$$

- la probabilité de transition d'un niveau vers le niveau suivant

$$p_k = \mathbb{P}[T_k \leq T \mid T_{k-1} \leq T] = \langle \eta_k, g_k \rangle$$

- la distribution d'entrée dans un niveau

$$\mathbb{P}[T_k \in dt, X(T_k) \in dx \mid T_k \leq T] = \mu_k(dt, dx)$$

nombreuses estimations et résultats asymptotiques quand le nombre de particules tend vers l'infini
attention : les fonctions de sélection sont seulement positives ou nulles

Proposition soit τ un temps d'arrêt

$$\mathbb{E}[F \mid \tau \leq T] = \int_0^T \int_E \mathbb{E}_{t,x}[F] \, \mathbb{P}[\tau \in dt, X(\tau) \in dx \mid \tau \leq T]$$

pour toute variable aléatoire F mesurable par rapport à la tribu $\sigma(X(t), t \geq \tau)$

Preuve

$$\mathbb{E}[F \mid \tau \leq T] = \frac{\mathbb{E}[F \, 1_{(\tau \leq T)}]}{\mathbb{P}[\tau \leq T]} = \frac{\mathbb{E}[\, \mathbb{E}[F \mid \mathcal{F}_\tau] \, 1_{(\tau \leq T)}]}{\mathbb{P}[\tau \leq T]}$$

et d'après la propriété de Markov forte

$$\mathbb{E}[F \mid \mathcal{F}_\tau] = \mathbb{E}_{\tau, X(\tau)}[F]$$

de sorte que

$$\mathbb{E}[F \mid \tau \leq T] = \frac{\mathbb{E}[\, \mathbb{E}_{\tau, X(\tau)}[F] \, 1_{(\tau \leq T)}]}{\mathbb{P}[\tau \leq T]} = \mathbb{E}[\, \mathbb{E}_{\tau, X(\tau)}[F] \mid \tau \leq T]$$

$$= \int_0^T \int_E \mathbb{E}_{t,x}[F] \, \mathbb{P}[\tau \in dt, X(\tau) \in dx \mid \tau \leq T] \qquad \square$$

Exemple pour $F = \phi(T_k \wedge T, X(T_k \wedge T))$ et pour $\tau = T_{k-1}$, on obtient la preuve alternative

$$\begin{aligned}
\langle \eta_k, \phi \rangle &= \mathbb{E}[\phi(T_k \wedge T, X(T_k \wedge T)) \mid T_{k-1} \leq T] \\
&= \int_0^T \int_E \mathbb{E}_{t,x}[\phi(T_k \wedge T, X(T_k \wedge T))] \\
&\qquad\qquad \mathbb{P}[T_{k-1} \in dt, X(T_{k-1}) \in dx \mid T_{k-1} \leq T] \\
&= \int_0^T \int_E Q_k \, \phi(t,x) \, \mu_{k-1}(dt, dx) \\
&= \langle \mu_{k-1}, Q_k \, \phi \rangle = \langle \mu_{k-1} \, Q_k, \phi \rangle
\end{aligned}$$

en effet, on rappelle que

$$\mathbb{P}[T_{k-1} \in dt, X(T_{k-1}) \in dx \mid T_{k-1} \leq T] = \mu_{k-1}(dt, dx)$$

et

$$\begin{aligned} Q_k(t, x, dt', dx') &= \mathbb{P}[T_k \wedge T \in dt', X(T_k \wedge T) \in dx' \mid X(t) = x] \\ &= \mathbb{P}_{t,x}[T_k \wedge T \in dt', X(T_k \wedge T) \in dx'] \end{aligned}$$

de sorte que

$$\mathbb{E}_{t,x}[\phi(T_k \wedge T, X(T_k \wedge T))] = Q_k\, \phi(t, x)$$

Exemple pour $F = 1_{\left(\sup_{T_{k-1} \leq t \leq T} u(t, X(t)) \geq c\right)}$ et pour $\tau = T_{k-1}$, on obtient

$$
\begin{aligned}
F_k(c) &= \mathbb{P}[\sup_{T_{k-1} \leq t \leq T} u(t, X(t)) \geq c \mid T_{k-1} \leq T] \\
&= \int_0^T \int_E \mathbb{P}_{s,x}[\sup_{s \leq t \leq T} u(t, X(t)) \geq c] \, \mu_{k-1}(ds, dx)
\end{aligned}
$$

Simulation d'évènements rares

Approximation particulaire

- niveaux fixés
- illustration: mouvement brownien avec dérive
- théorèmes limites
- choix adaptatif des niveaux

Niveaux fixés

relation de récurrence à la Feynman–Kac

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

approximation particulaire pondérée de la forme

$$\eta_k \approx \eta_k^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\Xi_k^i} \quad \text{et} \quad \mu_k \approx \mu_k^N = \sum_{i=1}^{N} w_k^i\, \delta_{\Xi_k^i} \quad \text{avec} \quad \sum_{i=1}^{N} w_k^i = 1$$

où chaque particule–fin d'excursion est de la forme
$\Xi_k^i = (T_k^i \wedge T, X^i(T_k^i \wedge T))$

en partant de

$$\eta_k \approx \eta_k^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\Xi_k^i}$$

on remarque que

$$g_k \cdot \eta_k^N = \frac{\displaystyle\sum_{i=1}^{N} g_k(\Xi_k^i)\, \delta_{\Xi_k^i}}{\displaystyle\sum_{j=1}^{N} g_k(\Xi_k^j)} = \frac{1}{|I_k^N|} \sum_{i \in I_k^N} \delta_{\Xi_k^i}$$

et

$$\langle \eta_k^N, g_k \rangle = \frac{1}{N} \sum_{i=1}^{N} g_k(\Xi_k^i) = \frac{|I_k^N|}{N}$$

où

$$I_k^N = \{ i = 1 \cdots N \,:\, g_k(\Xi_k^i) = 1 \} = \{ i = 1 \cdots N \,:\, T_k^i \leq T \}$$

correspond aux particules qui ont atteint A_k avant l'instant final T

approximation particulaire des distributions d'entrée

$$\mu_k = g_k \cdot \eta_k \approx g_k \cdot \eta_k^N = \frac{1}{|I_k^N|} \sum_{i \in I_k^N} \delta_{\Xi_k^i}$$

interprétation comme

distribution de probabilité empirique des particules qui ont atteint A_k avant l'instant final T

approximation particulaire des probabilités de transition

$$p_k = \langle \eta_k, g_k \rangle \approx \langle \eta_k^N, g_k \rangle = \frac{1}{N} \sum_{i=1}^{N} g_k(\Xi_k^i) = \frac{|I_k^N|}{N}$$

interprétation comme

proportion des particules qui ont atteint A_k avant l'instant final T

approximation particulaire de la probabilité de l'évènement rare

$$P_B = \mathbb{P}[T_B \leq T] = \prod_{k=1}^{n} \langle \eta_k, g_k \rangle \approx \prod_{k=1}^{n} \langle \eta_k^N, g_k \rangle = \prod_{k=1}^{n} \frac{|I_k^N|}{N}$$

algorithme SIR (*sampling with importance resampling*)

- pour $k = 0$, indépendamment pour tout $i = 1 \cdots N$
  on simule une v.a. Ξ_0^i distribuée selon $\delta_0 \times \eta_0$, et on définit

$$w_0^i = \frac{1}{N}$$

- 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 $Q_k(\widehat{\Xi}_{k-1}^i, dt', dx')$

  et on définit

$$w_k^i = \frac{g_k(\Xi_k^i)}{\displaystyle\sum_{j=1}^N g_k(\Xi_k^j)}$$

- pour $k = 0$, indépendamment pour tout $i = 1 \cdots N$
on simule une v.a. $\Xi_0^i = (0, X_0^i)$ distribuée selon $\delta_0 \times \eta_0$, et on définit

$$I_0^N = \{1 \cdots N\}$$

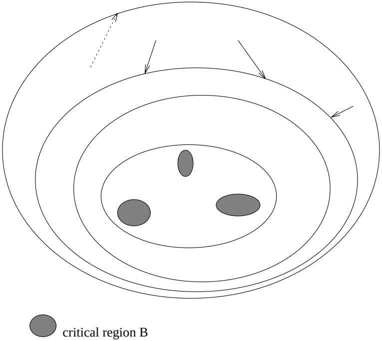
- pour tout $k = 1 \cdots n$, indépendamment pour tout $i = 1 \cdots N$
  - on sélectionne un individu $\widehat{\Xi}_{k-1}^i = (\widehat{T}_{k-1}^i, \widehat{X}_{k-1}^i)$ au sein de la population $(\Xi_{k-1}^i \,,\, i \in I_{k-1}^N)$
  - on simule une trajectoire $\{X^i(t)\}$ partant de $\widehat{X}_{k-1}^i$ à l'instant $\widehat{T}_{k-1}^i$, jusqu'à ce qu'elle atteigne A_k, i.e. jusqu'à l'instant

    $$T_k^i = \inf\{t \geq T_{k-1}^i \,:\, X^i(t) \in A_k(t)\} = \inf\{t \geq T_{k-1}^i \,:\, u(t, X^i(t)) \geq c_k\}$$

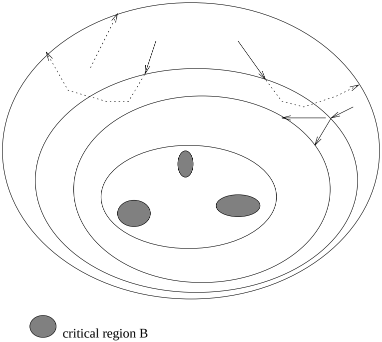
    ou bien jusqu'à l'instant final T, selon que l'un ou l'autre instant arrive en premier
  - on pose $\Xi_k^i = (T_k^i \wedge T, X^i(T_k^i \wedge T))$

  et on définit

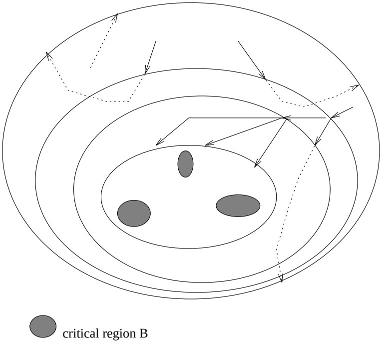
$$I_k^N = \{i = 1 \cdots N \,:\, g_k(\Xi_k^i) = 1\} = \{i = 1 \cdots N \,:\, T_k^i \leq T\}$$



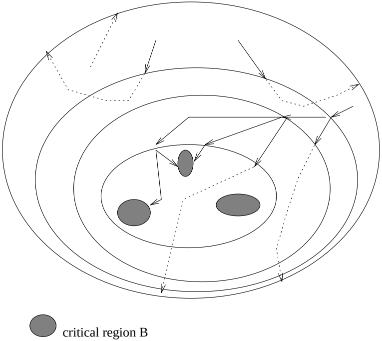
entre deux niveaux, les particules explorent l'espace d'état
en imitant le comportement de la chaîne de Markov de départ $\{X(t)\}$



entre deux niveaux, les particules explorent l'espace d'état
en imitant le comportement de la chaîne de Markov de départ $\{X(t)\}$
les trajectoires qui réussissent à atteindre le niveau suivant avant
l'instant T sont sélectionnés, les autres trajectoires sont éliminées



entre deux niveaux, les particules explorent l'espace d'état
en imitant le comportement de la chaîne de Markov de départ $\{X(t)\}$
les trajectoires qui réussissent à atteindre le niveau suivant avant
l'instant T sont sélectionnés, les autres trajectoires sont éliminées



entre deux niveaux, les particules explorent l'espace d'état

en imitant le comportement de la chaîne de Markov de départ $\{X(t)\}$

les trajectoires qui réussissent à atteindre le niveau suivant avant l'instant T sont sélectionnés, les autres trajectoires sont éliminées

il peut arriver qu'aucune trajectoire ne réussisse à atteindre le niveau suivant avant l'instant T !

Théorèmes limites

durée de vie du système de particules

$$\tau^N = \inf\{k \geq 0 \,:\, |I_k^N| = 0\} = \inf\{k \geq 0 \,:\, I_k^N = \emptyset\}$$

probabilité d'extinction exponentiellement petite

$$\mathbb{P}[\tau^N \leq n] \leq a'_n \exp\{-a_n N\}$$

pour des constantes positives $a'_n > 0$ et $a_n > 0$
approximation particulaire de la probabilité de l'évènement rare, définie sur l'évènement de non extinction

$$P_B = \mathbb{P}[T_B \leq T] = \langle \gamma_n, 1 \rangle \approx \langle \gamma_n^N, 1 \rangle = \prod_{k=1}^{n} \langle \eta_k^N, g_k \rangle = \prod_{k=1}^{n} \frac{|I_k^N|}{N}$$

théorème central limite : quand $N \uparrow \infty$

$$\sqrt{N} \, [\, \frac{\langle \gamma_n^N, 1 \rangle}{P_B} - 1 \,] \, 1_{(\tau^N > n)} \Longrightarrow \mathcal{N}(0, \sigma_n^2)$$

variance asymptotique (cas particulier d'une fonction de sélection binaire)

$$\sigma_n^2 = \sum_{k=0}^{n} \frac{\mathrm{var}(g_k\, R_{k+1:n}\, 1, \eta_k)}{\langle \eta_k, g_k\, R_{k+1:n}\, 1\rangle^2}$$

$$= \sum_{k=0}^{n} (\frac{1}{p_k} - 1) + \sum_{k=0}^{n} \frac{1}{p_k}\; \frac{\mathrm{var}(R_{k+1:n}\, 1, \mu_k)}{\langle \mu_k, R_{k+1:n}\, 1\rangle^2}$$

où $p_k = \langle \eta_k, g_k \rangle$ et où

$$R_{k+1:n}\, 1(t,x) = \mathbb{E}[\prod_{p=k+1}^{n} g_p(\Xi_p) \mid \Xi_k = (t,x)]$$

$$= \mathbb{P}[T_n \leq T \mid T_k = t, X(T_k) = x]$$

$$= \mathbb{P}[T_B \leq T \mid X(t) = x]$$

ne dépend finalement pas du niveau k

variance asymptotique (suite)

$$\sigma_n^2 = \sum_{k=1}^{n} (\frac{1}{p_k} - 1) + \sum_{k=1}^{n} \frac{1}{p_k} \, \frac{\mathrm{var}(u_B, \mu_k)}{\langle \mu_k, u_B \rangle^2}$$

où

$$u_B(t,x) = \mathbb{P}[T_B \leq T \mid X(t) = x]$$

par définition, et où la distribution d'entrée

$$\mu_k(dt, dx) = \mathbb{P}[T_k \in dt, X(T_k) \in dx \mid T_k \leq T]$$

est portée par l'ensemble $A_k \subset [0, \infty) \times E$, par construction, de sorte que

$$\begin{aligned} \mathrm{var}(u_B, \mu_k) &= \int_0^\infty \int_E u_B^2(t,x) \, \mu_k(dt, dx) - | \int_0^\infty \int_E u_B(t,x) \, \mu_k(dt, dx)|^2 \\ &= \int_{A_k} u_B^2(t,x) \, \mu_k(dt, dx) - | \int_{A_k} u_B(t,x) \, \mu_k(dt, dx)|^2 \end{aligned}$$

pour le choix particulier des régions intermédiaires

$$A_k = \{(t,x) \in [0,\infty) \times E \,:\, u_B(t,x) \geq c_k\}$$

et si la distribution d'entrée

$$\mu_k(dt,dx) = \mathbb{P}[T_k \in dt, X(T_k) \in dx \mid T_k \leq T]$$

est portée par l'ensemble de niveau (avec égalité)

$$\{(t,x) \in [0,\infty) \times E \,:\, u_B(t,x) = c_k\} \subset A_k$$

par exemple si les trajectoires du processus $\{X(t)\,,\, t \geq 0\}$ sont continues — alors la fonction $u_B(t,x)$ est constante sur le support de la distribution d'entrée, et

$$\mathrm{var}(u_B,\mu_k) = \int_{A_k} u_B^2(t,x)\,\mu_k(dt,dx) - |\int_{A_k} u_B(t,x)\,\mu_k(dt,dx)|^2 = 0$$

pour ce choix particulier, la variance asymptotique se réduit à l'expression

$$\sigma_n^2 = \sum_{k=1}^{n}(\frac{1}{p_k} - 1) \ll \frac{1}{P_B} - 1$$

minimiser le premier terme de la variance asymptotique

$$\sum_{k=1}^{n} (\frac{1}{p_k} - 1) \qquad \text{sous la contrainte} \qquad p_1 \cdots p_n = P_B$$

conduit à choisir les probabilités de transition toutes égales entre elles

$$p_1 = \cdots = p_n = P_B^{1/n}$$

auquel cas

$$\frac{1}{P_B} - 1 \geq \sigma_n^2 = n\,(\frac{1}{P_B^{1/n}} - 1) \geq -\log P_B$$

avec décroissance monotone vers la limite $-\log P_B$ quand le nombre n de niveaux augmente (on rappelle que $P_B \ll 1$)

Choix adaptatif des niveaux

idéalement, les niveaux doivent être choisis de telle sorte que les probabilités de transition

$$p_k = \mathbb{P}[T_k \leq T \mid T_{k-1} \leq T]$$

soient toutes égales entre elles (c'est–à–dire ne dépendent pas de l'indice), mais cette prescription définit les niveaux de manière implicite seulement, en termes de quantiles : en effet

$$\mathbb{P}[T_k \leq T \mid T_{k-1} \leq T] = \mathbb{P}[\sup_{0 \leq t \leq T} u(t, X(t)) \geq c_k \mid T_{k-1} \leq T]$$

$$= \mathbb{P}[\sup_{T_{k-1} \leq t \leq T} u(t, X(t)) \geq c_k \mid T_{k-1} \leq T] = F_k(c_k)$$

avec la fonction de répartition définie par

$$F_k(c) = \mathbb{P}[\sup_{T_{k-1} \leq t \leq T} u(t, X(t)) \geq c \mid T_{k-1} \leq T]$$

$$= \int_0^T \int_E \mathbb{P}_{s,x}[\sup_{s \leq t \leq T} u(t, X(t)) \geq c]\, \mu_{k-1}(ds, dx)$$

pour garantir que
$$\mathbb{P}[T_k \leq T \mid T_{k-1} \leq T] = q$$
où la valeur q est fixée, on doit donc choisir le seuil comme le quantile $c_k = F_k^{-1}(q)$

en pratique, les fonctions de répartition F_k et leurs quantiles $F_k^{-1}(q)$ sont inconnus et doivent être appris

idée : puisqu'au final la probabilité de transition p_k sera estimée comme la proportion des particules qui ont atteint A_k avant l'instant final T, on se donne une probabilité de succès q et on choisit le seuil de telle sorte que $\lfloor q\,N \rfloor$ trajectoires parmi N aient dépassé ce seuil algorithme idéalisé

- pour $k = 0$, indépendamment pour tout $i = 1 \cdots N$ on simule une v.a. $\Xi_0^i = (0, X_0^i)$ distribuée selon $\delta_0 \times \eta_0$, et on définit

$$I_0^N = \{1 \cdots N\}$$

- pour tout $k = 1 \cdots n$, indépendamment pour tout $i = 1 \cdots N$
  - on sélectionne un individu $\widehat{\Xi}_{k-1}^i = (\widehat{T}_{k-1}^i, \widehat{X}_{k-1}^i)$ au sein de la population $(\Xi_{k-1}^i \,,\, i \in I_{k-1}^N)$
  - on simule une trajectoire $\{X^i(t)\}$ partant de $\widehat{X}_{k-1}^i$ à l'instant $\widehat{T}_{k-1}^i$, et jusqu'à l'instant final T, et on définit

$$M_k^i = \sup_{\widehat{T}_{k-1}^i \le t \le T} u(t, X^i(t))$$

  on ordonne selon le rang décroissant

$$M_k^{(1)} \ge M_k^{(2)} \ge \cdots \ge M_k^{(N)}$$

  et on définit le seuil comme le quantile empirique $c_k^N = M_k^{(\lfloor q\, N \rfloor)}$

pour tout $i = 1 \cdots N$, on pose

$$T_k^i = \inf\{t \geq T_{k-1}^i \,:\, u(t, X^i(t)) \geq c_k^N\} \quad \text{et} \quad \Xi_k^i = (T_k^i \wedge T, X^i(T_k^i \wedge T))$$

et on définit

$$I_k^N = \{i = 1 \cdots N \,:\, M_k^i \geq c_k^N\} = \{i = 1 \cdots N \,:\, T_k^i \leq T\}$$

par construction

$$|I_k^N| = \lfloor q\, N \rfloor$$

de sorte que

$$p_k = \mathbb{P}[T_k \leq T \mid T_{k-1} \leq T] \approx \frac{\lfloor q\, N \rfloor}{N} \approx q$$

interprétation comme

proportion des particules qui ont atteint A_k^N avant l'instant final T

où la région

$$A_k^N = \{(t, x) \in [0, \infty) \times E \,:\, u(t, x) \geq c_k^N\}$$

correspond au seuil empirique c_k^N

Remarque évaluer la valeur maximum M_k^i et conserver l'instant et la position où ce maximum a été atteint est facile, et peut être effectué en–ligne

en revanche, si une valeur (comme le seuil empirique c_k^N par exemple) est seulement spécifiée après coup, c'est–à–dire après que toute la trajectoire ait été simulée, alors

- décider si cette valeur particulière a été atteinte (ou pas) est facile : il suffit de comparer cette valeur avec le maximum M_k^i
- mais retrouver quand (à quel instant) et où (dans quelle position) cette valeur particulière a été atteinte le cas échéant, est seulement possible si toute la trajectoire à été conservée, ce qui peut réclamer un espace–mémoire considérable

en d'autres termes, prétendre évaluer

$$T_k^i = \inf\{t \geq T_{k-1}^i \,:\, u(t, X^i(t)) \geq c_k^N\} \quad \text{et} \quad X_k^i = X^i(T_k^i)$$

pour tout $i \in I_k^N$, n'est pas réaliste algorithme pratique en deux étapes

- pour $k = 0$, indépendamment pour tout $i = 1 \cdots N$
  on simule une v.a. $\Xi_0^i = (0, X_0^i)$ distribuée selon $\delta_0 \times \eta_0$, et on définit

$$I_0^N = \{1 \cdots N\}$$

- pour tout $k = 1 \cdots n$, (1ère étape) indépendamment pour tout $i = 1 \cdots N$
  - on sélectionne un individu $\widehat{\Xi}_{k-1}^i = (\widehat{T}_{k-1}^i, \widehat{X}_{k-1}^i)$ au sein de la population $(\Xi_{k-1}^i \,,\, i \in I_{k-1}^N)$
  - on simule une trajectoire $\{X^i(t)\}$ partant de $\widehat{X}_{k-1}^i$ à l'instant $\widehat{T}_{k-1}^i$, et jusqu'à l'instant final T, et on définit

$$M_k^i = \sup_{\widehat{T}_{k-1}^i \leq t \leq T} u(t, X^i(t))$$

  on ordonne selon le rang décroissant

$$M_k^{(1)} \geq M_k^{(2)} \geq \cdots \geq M_k^{(N)}$$

  et on définit le seuil comme le quantile empirique $c_k^N = M_k^{(\lfloor q\, N \rfloor)}$

on oublie les particules pilotes et on ne conserve que le seuil c_k^N (2ème étape) indépendamment pour tout $i = 1 \cdots N$

- on sélectionne un individu $\widehat{\Xi}_{k-1}^i = (\widehat{T}_{k-1}^i, \widehat{X}_{k-1}^i)$ au sein de la population $(\Xi_{k-1}^i \, , \, i \in I_{k-1}^N)$
- on simule une trajectoire $\{X^i(t)\}$ partant de $\widehat{X}_{k-1}^i$ à l'instant $\widehat{T}_{k-1}^i$, jusqu'à ce qu'elle atteigne A_k^N, i.e. jusqu'à l'instant

  $$T_k^i = \inf\{t \geq T_{k-1}^i \, : \, (t, X^i(t)) \in A_k^N\} = \inf\{t \geq T_{k-1}^i \, : \, u(t, X^i(t)) \geq c_k^N\}$$

  ou bien jusqu'à l'instant final T, selon que l'un ou l'autre instant arrive en premier
- on pose $\Xi_k^i = (T_k^i \wedge T, X^i(T_k^i \wedge T))$

et on définit

$$I_k^N = \{i = 1 \cdots N \, : \, g_k(\Xi_k^i) = 1\} = \{i = 1 \cdots N \, : \, T_k^i \leq T\}$$

approximation particulaire des probabilités de transition

$$p_k = \mathbb{P}[T_k \leq T \mid T_{k-1} \leq T] \approx \frac{|I_k^N|}{N} \neq \frac{\lfloor q\, N \rfloor}{N}$$

interprétation comme

proportion des particules qui ont atteint A_k^N avant l'instant final T

où la région

$$A_k^N = \{(t,x) \in [0,\infty) \times E \,:\, u(t,x) \geq c_k^N\}$$

correspond au seuil empirique c_k^N