

Simulation parfaite des chaînes de Markov

Ana Bušić

INRIA - ENS

`http://www.di.ens.fr/~busic/`

`ana.busic@inria.fr`

Master AMIS - UVSQ

Versailles, octobre 2016

Système à événements discrets

SED : $(\mathcal{X}, \pi^0, A, p, \cdot)$

- ▶ \mathcal{X} est un **espace d'états** fini.
- ▶ π^0 est une **distribution initiale** :
 $\pi_x^0 \geq 0$ est la probabilité que le système soit dans l'état $x \in \mathcal{X}$ à l'instant 0.
- ▶ A est un **ensemble d'événements** fini.
- ▶ p est une mesure de probabilité sur A :
 $p_a > 0$ est la **probabilité de l'événement** $a \in A$.
- ▶ \cdot est un opérateur qui décrit l'action des événements dans A sur les états dans \mathcal{X} - **fonction de transition** :

$$\cdot : \mathcal{X} \times A \rightarrow \mathcal{X}, \quad (x, a) \mapsto x \cdot a.$$

Fonction de transition

Extension aux séquences (mots) d'événements fini :

$$\cdot : \mathcal{X} \times A^n \rightarrow \mathcal{X}, n \in \mathbb{N}.$$

Soit $a_{1 \rightarrow n} \stackrel{\text{def}}{=} a_1 \dots a_n$ dans A^n :

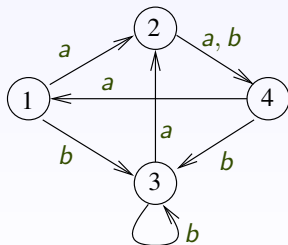
$$(x, a_{1 \rightarrow n}) \mapsto x \cdot a_{1 \rightarrow n} \stackrel{\text{def}}{=} x \cdot a_1 \cdot a_2 \cdot \dots \cdot a_n.$$

Évolution du système

Sur n étapes :

1. On choisit l'état initial X_0 selon la distribution π^0 .
2. Pour $i = 1$ à n faire :
 - ▶ Choisir un événement $a_i \in A$ selon p
 - ▶ $X_i := X_{i-1} \cdot a_i = X_0 \cdot a_{1 \rightarrow i}$

Exemple :



Soit $p_a = 1/3$, $p_b = 2/3$, et
 $\pi^0 = (1/4, 1/4, 1/4, 1/4)$.

Une **trajectoire** possible :

1-3-3-2-4-1-3-3-...

en partant de l'état 1 et pour la
séquence d'événements

bbababb...

SED vs. chaîne de Markov

Un SED $\mathcal{A} = (\mathcal{X}, \pi^0, A, p, \cdot)$ induit une chaîne de Markov
(en temps discret homogène) :

▶ chaîne de Markov

- ▶ Soient $\{a_n\}_{n \in \mathbb{N}}$ une suite *i.i.d.* d'événements de A distribués selon p et X_0 distribué selon π^0 .
- ▶ Alors $\{X_n \stackrel{\text{def}}{=} X_0 \cdot a_{1 \rightarrow n}\}_{n \in \mathbb{N}}$ est une chaîne de Markov avec la matrice de transition P :

$$\text{pour tout } x, y \text{ dans } \mathcal{X}, \quad P_{x,y} = \sum_{a \in A : x \cdot a = y} p_a. \quad (1)$$

SED vs. chaîne de Markov

Un SED $\mathcal{A} = (\mathcal{X}, \pi^0, A, p, \cdot)$ induit une chaîne de Markov
(en temps discret homogène) :

▶ chaîne de Markov

- ▶ Soient $\{a_n\}_{n \in \mathbb{N}}$ une suite *i.i.d.* d'événements de A distribués selon p et X_0 distribué selon π^0 .
- ▶ Alors $\{X_n \stackrel{\text{def}}{=} X_0 \cdot a_{1 \rightarrow n}\}_{n \in \mathbb{N}}$ est une chaîne de Markov avec la matrice de transition P :

$$\text{pour tout } x, y \text{ dans } \mathcal{X}, \quad P_{x,y} = \sum_{a \in A : x \cdot a = y} p_a. \quad (1)$$

De plus, pour toute chaîne de Markov finie avec la distribution initiale π^0 et matrice de transition P il existe un SED $\mathcal{A} = (\mathcal{X}, \pi^0, A, p, \cdot)$ vérifiant (1). Pas unique.

SED vs. chaîne de Markov

Un SED $\mathcal{A} = (\mathcal{X}, \pi^0, A, p, \cdot)$ induit une chaîne de Markov
(en temps discret homogène) :

▶ chaîne de Markov

- ▶ Soient $\{a_n\}_{n \in \mathbb{N}}$ une suite *i.i.d.* d'événements de A distribués selon p et X_0 distribué selon π^0 .
- ▶ Alors $\{X_n \stackrel{\text{def}}{=} X_0 \cdot a_{1 \rightarrow n}\}_{n \in \mathbb{N}}$ est une chaîne de Markov avec la matrice de transition P :

$$\text{pour tout } x, y \text{ dans } \mathcal{X}, \quad P_{x,y} = \sum_{a \in A : x \cdot a = y} p_a. \quad (1)$$

De plus, pour toute chaîne de Markov finie avec la distribution initiale π^0 et matrice de transition P il existe un SED $\mathcal{A} = (\mathcal{X}, \pi^0, A, p, \cdot)$ vérifiant (1). **Pas unique.**

Si $|\mathcal{X}| = N$, il existe un SED avec au plus N^2 événements.

Echantionnage de la distribution stationnaire

$\{X_n\}_{n \in \mathbb{N}}$ une chaîne irréductible et apériodique.

Question

Comment échantillonner sa distribution stationnaire π ?

Echantionnage de la distribution stationnaire

$\{X_n\}_{n \in \mathbb{N}}$ une chaîne irréductible et apériodique.

Question

Comment échantillonner sa distribution stationnaire π ?

Solution : résoudre le système linéaire $\pi = \pi P$ pour calculer π , puis utiliser l'échantillonnage d'une distribution discrète..

Complexité : $O(N^3)$ (avec $N = |\mathcal{X}|$).

Echantionnage de la distribution stationnaire

$\{X_n\}_{n \in \mathbb{N}}$ une chaîne irréductible et apériodique.

Question

Comment échantillonner sa distribution stationnaire π ?

Solution : résoudre le système linéaire $\pi = \pi P$ pour calculer π , puis utiliser l'échantillonnage d'une distribution discrète..

Complexité : $O(N^3)$ (avec $N = |\mathcal{X}|$).

Question

Comment éviter le calcul de π (et même P) ?

Simulation Monte-Carlo (MCMC)

Algorithme :

- ▶ On choisit l'état initial X_0 selon π^0 .
- ▶ Pour $i = 1$ jusqu'à n :
 - ▶ On choisit a_i selon p .
 - ▶ $X_i = X_{i-1} \cdot a_i$.

Sortie : un échantillon de $\pi^n = \pi^0 P^n$.

Complexité : $O(\mathcal{C}n)$, où \mathcal{C} est la complexité du calcul de la fonction de transition \cdot .

Remarque : échantillonnage d'une distribution discrète en $\Theta(1)$ en utilisant la méthode de alias [Walker, 77]. ▶ alias

Inconvénient : approximation.

Estimation d'erreur est difficile : dépend de la deuxième valeur propre de P - difficile à calculer [Brémaud, Asmussen and Glynn].

Simulation parfaite (exacte)

Objectif :

- ▶ des échantillons distribués selon la distribution π exactement sans jamais la calculer (ni P).
- ▶ temps d'arrêt fini.

Propp et Wilson (1996) ont proposé d'utiliser le **couplage depuis le passé** comme une technique effective de simulation parfaite.

Couplage depuis le passé [Borovkov 75, Glynn 96] - travaux d'un intérêt essentiellement théorique et existentiel.

Couplage vers l'avant

On considère une suite a_1, \dots, a_n, \dots d'événements i.i.d. selon la distribution p et l'ensemble des trajectoires qu'elle induit en partant de tous les états du SED.

Notation : $E \cdot a = \{x \cdot a : x \in E\}$.

On pose $E_0 = \mathcal{X}$, $E_1 = \mathcal{X} \cdot a_1$, \dots , $E_n = E_{n-1} \cdot a_n = \mathcal{X} \cdot a_{1 \rightarrow n}$.

Couplage vers l'avant

On considère une suite a_1, \dots, a_n, \dots d'événements i.i.d. selon la distribution p et l'ensemble des trajectoires qu'elle induit en partant de tous les états du SED.

Notation : $E \cdot a = \{x \cdot a : x \in E\}$.

On pose $E_0 = \mathcal{X}$, $E_1 = \mathcal{X} \cdot a_1$, \dots , $E_n = E_{n-1} \cdot a_n = \mathcal{X} \cdot a_{1 \rightarrow n}$.

Par construction les tailles des ensembles E_n sont décroissantes, donc il existe $\ell = \lim_{n \rightarrow \infty} |E_n|$.

Définition

Le SED a la *propriété de couplage (couple)* si $\ell = 1$ pour presque toutes les suites d'événements a_1, \dots, a_n, \dots

Le *temps de couplage* est alors $\tau^f \stackrel{\text{def}}{=} \min\{n \in \mathbb{N} \text{ t.q. } |E_n| = 1\}$.

Couplage vers l'avant(II)

Théorème

Un SED couple si et seulement si il admet un mot couplant (ou synchronisant).

La variable aléatoire τ^f a une queue de distribution géométrique et son espérance est finie :

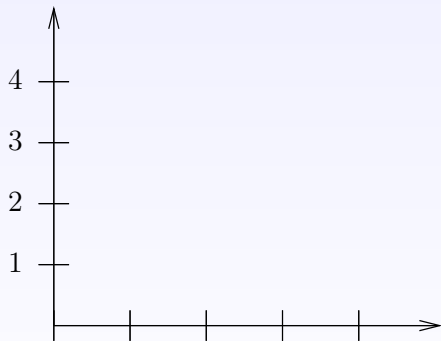
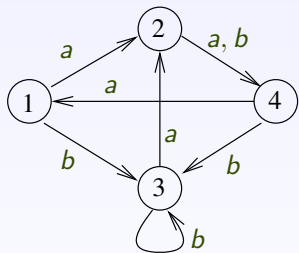
soit $a_1 \cdots a_k$ un mot couplant, alors

$$P(\tau^f > n.k) \leq (1 - p_{a_1} \cdots p_{a_k})^n$$

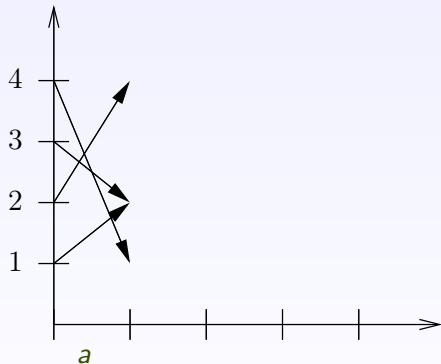
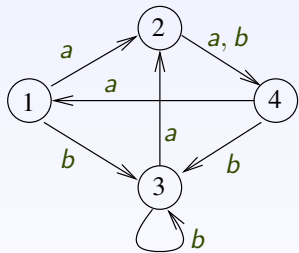
et

$$\mathbb{E}\tau^f \leq k/(p_{a_1} \cdots p_{a_k}).$$

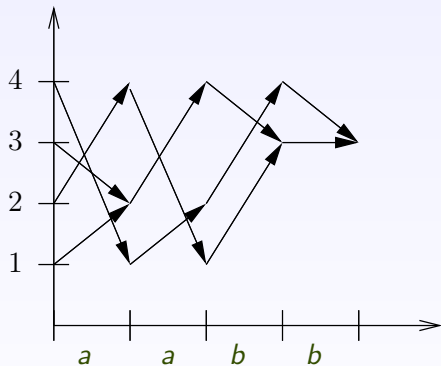
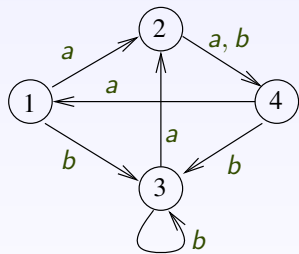
Couplage vers l'avant(III)



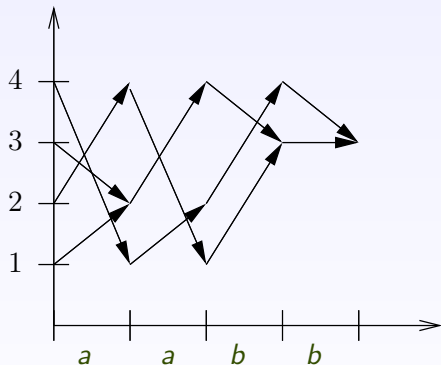
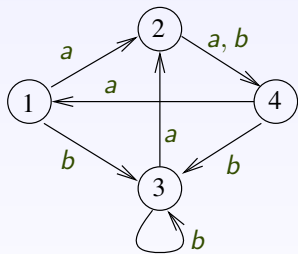
Couplage vers l'avant(III)



Couplage vers l'avant(III)



Couplage vers l'avant(III)

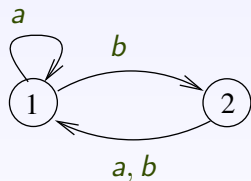


Couplage en état 3.

Questions

Question

Couplage vers l'avant donne-t-il un échantillon stationnaire ?

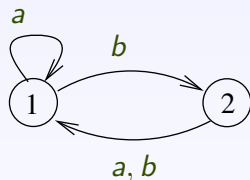


$$p_a = p_b = 0.5$$

Questions

Question

Couplage vers l'avant donne-t-il un échantillon stationnaire ?

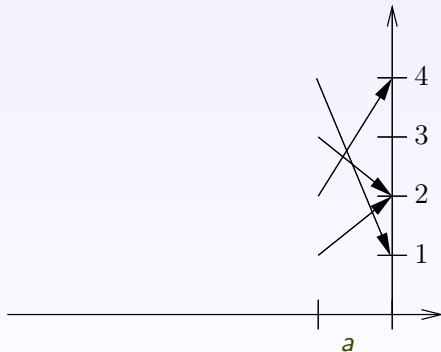
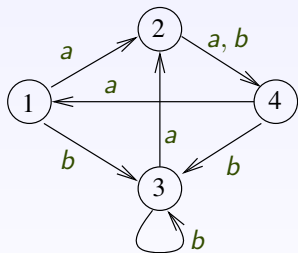


$$p_a = p_b = 0.5$$

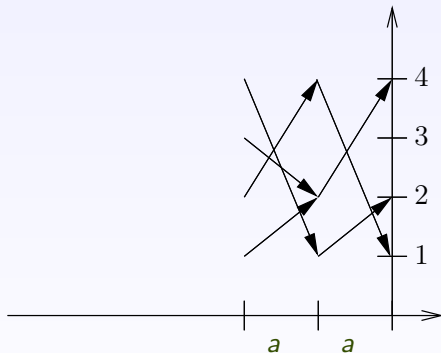
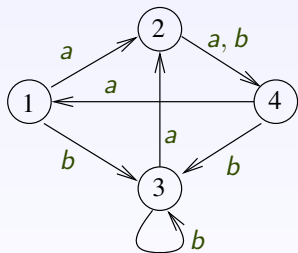
Question

Un SED dont la chaîne de Markov est irréductible et apériodique couple-t-il toujours ?

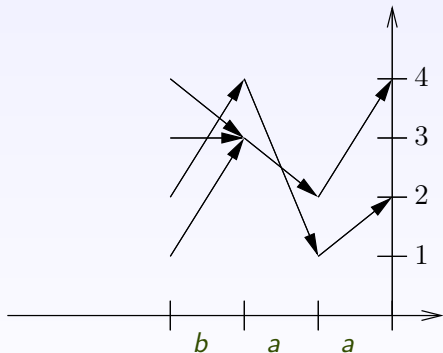
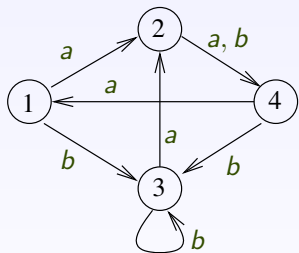
Couplage depuis le passé : l'idée



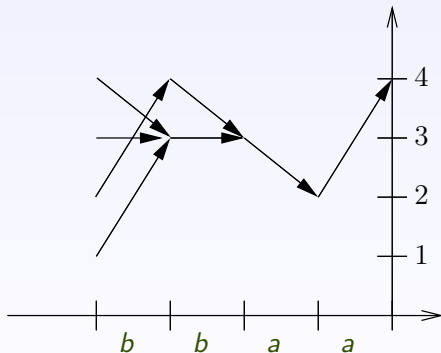
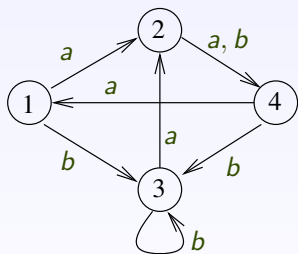
Couplage depuis le passé : l'idée



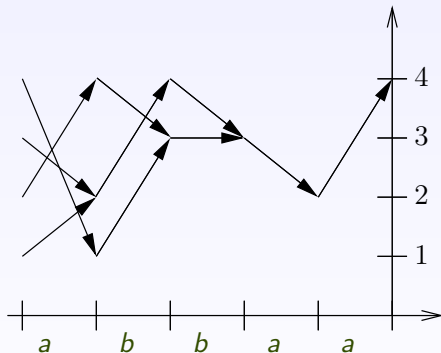
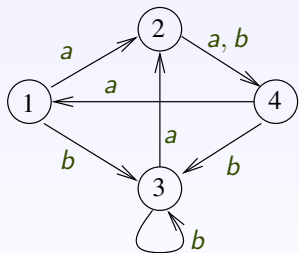
Couplage depuis le passé : l'idée



Couplage depuis le passé : l'idée



Couplage depuis le passé : l'idée



Couplage depuis le passé

$a_1, a_2, \dots, a_n, \dots$ une suite *i.i.d.* d'événements distribués selon p .

On regarde l'ensemble des trajectoires depuis le passé :

$$S_0 = \mathcal{X}, S_1 = \mathcal{X} \cdot a_1, S_2 = \mathcal{X} \cdot a_2 a_1,$$

$$S_n \stackrel{\text{def}}{=} \mathcal{X} \cdot a_{n \rightarrow 1} \stackrel{\text{def}}{=} \mathcal{X} \cdot a_n \cdot a_{n-1} \dots a_1.$$

Couplage depuis le passé

$a_1, a_2, \dots, a_n, \dots$ une suite *i.i.d.* d'événements distribués selon p .

On regarde l'ensemble des trajectoires depuis le passé :

$$S_0 = \mathcal{X}, S_1 = \mathcal{X} \cdot a_1, S_2 = \mathcal{X} \cdot a_2 a_1, \\ S_n \stackrel{\text{def}}{=} \mathcal{X} \cdot a_{n \rightarrow 1} \stackrel{\text{def}}{=} \mathcal{X} \cdot a_n \cdot a_{n-1} \dots a_1.$$

Par construction, on a $S_n \subseteq S_{n-1} \subseteq \dots \subseteq S_1 \subseteq \mathcal{X}$.

Il existe donc un ensemble limite S .

Si le SED est couplant, alors S est réduit à un seul élément p.s.. On définit alors le temps de couplage depuis le passé :

$$\tau^b \stackrel{\text{def}}{=} \min\{n \in \mathbb{N} \text{ t.q. } |S_n| = 1\}.$$

Couplage depuis le passé

$a_1, a_2, \dots, a_n, \dots$ une suite *i.i.d.* d'événements distribués selon p .

On regarde l'ensemble des trajectoires depuis le passé :

$$S_0 = \mathcal{X}, S_1 = \mathcal{X} \cdot a_1, S_2 = \mathcal{X} \cdot a_2 a_1, \\ S_n \stackrel{\text{def}}{=} \mathcal{X} \cdot a_{n \rightarrow 1} \stackrel{\text{def}}{=} \mathcal{X} \cdot a_n \cdot a_{n-1} \dots a_1.$$

Par construction, on a $S_n \subseteq S_{n-1} \subseteq \dots \subseteq S_1 \subseteq \mathcal{X}$.

Il existe donc un ensemble limite S .

Si le SED est couplant, alors S est réduit à un seul élément p.s.. On définit alors le temps de couplage depuis le passé :

$$\tau^b \stackrel{\text{def}}{=} \min\{n \in \mathbb{N} \text{ t.q. } |S_n| = 1\}.$$

Comme la probabilité d'occurrence de $a_n \dots a_1$ et celle de $a_1 \dots a_n$ sont les mêmes : $\mathbb{P}(a_n \dots a_1) = p_{a_n} \dots p_{a_1} = \mathbb{P}(a_1 \dots a_n)$, alors

$$\tau^f \stackrel{D}{=} \tau^b.$$

Couplage depuis le passé (II)

Théorème

- ▶ Si SED est couplant, $\mathbb{E}(\tau^b) < \infty$.
- ▶ Soit Y (l'unique) état dans S_{τ^b} . Alors Y est distribué selon π .

Couplage depuis le passé (II)

Théorème

- ▶ Si SED est couplant, $\mathbb{E}(\tau^b) < \infty$.
- ▶ Soit Y (l'unique) état dans S_{τ^b} . Alors Y est distribué selon π .

Démonstration.

- ▶ Puisque $\tau^f \stackrel{D}{=} \tau^b$, $\mathbb{E}(\tau^b) = \mathbb{E}(\tau^f) < \infty$.

Couplage depuis le passé (II)

Théorème

- ▶ Si SED est couplant, $\mathbb{E}(\tau^b) < \infty$.
- ▶ Soit Y (l'unique) état dans S_{τ^b} . Alors Y est distribué selon π .

Démonstration.

- ▶ Puisque $\tau^f \stackrel{D}{=} \tau^b$, $\mathbb{E}(\tau^b) = \mathbb{E}(\tau^f) < \infty$.
- ▶ On doit montrer que : $\mathbb{P}(Y = x) = \pi_x, \forall x \in \mathcal{X}$. Soit $x \in \mathcal{X}$ un état fixé. Il suffit de montrer que pour tout $\epsilon > 0$,

$$|\mathbb{P}(Y = x) - \pi_x| \leq \epsilon.$$

Démonstration (suite)

Soit $\epsilon > 0$ fixé. Puisque $\tau^b < \infty$ p.s., il existe un m suffisamment grand t.q.

$$\mathbb{P}(|S_m| = 1) \geq 1 - \epsilon.$$

Démonstration (suite)

Soit $\epsilon > 0$ fixé. Puisque $\tau^b < \infty$ p.s., il existe un m suffisamment grand t.q.

$$\mathbb{P}(|S_m| = 1) \geq 1 - \epsilon.$$

Fixons m . Soit l'état X choisi selon π et observons la trajectoire en partant de X à l'instant $-m$, en utilisant la même suite d'événements a_m, \dots, a_1 . Notons par $\tilde{Y} = X \cdot a_{m \rightarrow 1}$.

Démonstration (suite)

Soit $\epsilon > 0$ fixé. Puisque $\tau^b < \infty$ p.s., il existe un m suffisamment grand t.q.

$$\mathbb{P}(|S_m| = 1) \geq 1 - \epsilon.$$

Fixons m . Soit l'état X choisi selon π et observons la trajectoire en partant de X à l'instant $-m$, en utilisant la même suite d'événements a_m, \dots, a_1 . Notons par $\tilde{Y} = X \cdot a_{m \rightarrow 1}$.

Puisque X est stationnaire, \tilde{Y} l'est aussi. De plus,

$$\mathbb{P}(Y \neq \tilde{Y}) \leq \epsilon.$$

Démonstration (suite)

Nous avons :

$$\begin{aligned}\mathbb{P}(Y = x) - \pi_x &= \mathbb{P}(Y = x) - \mathbb{P}(\tilde{Y} = x) \\ &\leq \mathbb{P}(Y = x, \tilde{Y} \neq x) \\ &\leq \mathbb{P}(Y \neq \tilde{Y}) \leq \epsilon.\end{aligned}$$

Démonstration (suite)

Nous avons :

$$\begin{aligned}\mathbb{P}(Y = x) - \pi_x &= \mathbb{P}(Y = x) - \mathbb{P}(\tilde{Y} = x) \\ &\leq \mathbb{P}(Y = x, \tilde{Y} \neq x) \\ &\leq \mathbb{P}(Y \neq \tilde{Y}) \leq \epsilon.\end{aligned}$$

Aussi :

$$\begin{aligned}\pi_x - \mathbb{P}(Y = x) &= \mathbb{P}(\tilde{Y} = x) - \mathbb{P}(Y = x) \\ &\leq \mathbb{P}(\tilde{Y} = x, Y \neq x) \\ &\leq \mathbb{P}(Y \neq \tilde{Y}) \leq \epsilon.\end{aligned}$$

Démonstration (suite)

Nous avons :

$$\begin{aligned}\mathbb{P}(Y = x) - \pi_x &= \mathbb{P}(Y = x) - \mathbb{P}(\tilde{Y} = x) \\ &\leq \mathbb{P}(Y = x, \tilde{Y} \neq x) \\ &\leq \mathbb{P}(Y \neq \tilde{Y}) \leq \epsilon.\end{aligned}$$

Aussi :

$$\begin{aligned}\pi_x - \mathbb{P}(Y = x) &= \mathbb{P}(\tilde{Y} = x) - \mathbb{P}(Y = x) \\ &\leq \mathbb{P}(\tilde{Y} = x, Y \neq x) \\ &\leq \mathbb{P}(Y \neq \tilde{Y}) \leq \epsilon.\end{aligned}$$

Donc,

$$|\mathbb{P}(Y = x) - \pi_x| \leq \epsilon.$$

Algorithme de simulation parfaite (PSA)

Data: $(\mathcal{X}, A, \cdot, p)$.

Result: Un état de \mathcal{X} distribué selon π .

begin

$n := 1$;

foreach état $x \in \mathcal{X}$ **do**

$T(x) := x$;

repeat

 Tirer a selon p ;

foreach état $x \in \mathcal{X}$ **do**

$S(x) := T(x \cdot a)$;

$T := S$;

until $|\{T(x), x \in \mathcal{X}\}| = 1$;

return $T(x_1)$ (x_1 un élément dans \mathcal{X})

Algorithme de simulation parfaite (PSA)

Data: $(\mathcal{X}, A, \cdot, p)$.

Result: Un état de \mathcal{X} distribué selon π .

begin

$n := 1$;

foreach état $x \in \mathcal{X}$ **do**

$T(x) := x$;

repeat

 Tirer a selon p ;

foreach état $x \in \mathcal{X}$ **do**

$S(x) := T(x \cdot a)$;

$T := S$;

until $|\{T(x), x \in \mathcal{X}\}| = 1$;

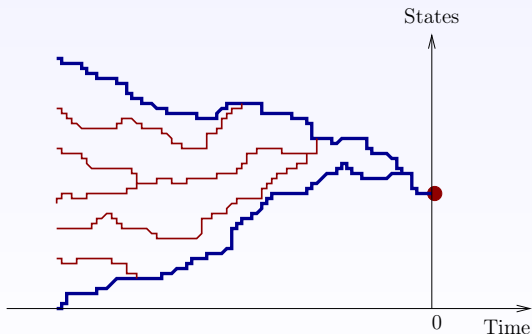
return $T(x_1)$ (x_1 un élément dans \mathcal{X})

Inconvenient : La complexité $O(\mathcal{C} \times |\mathcal{X}| \times \tau^b)$, avec \mathcal{C} la complexité du calcul de la fonction de transition (\cdot) .

SED monotone

On suppose que l'espace d'état est muni d'une relation d'ordre (\preceq) et que la fonction de transition de l'automate est monotone :

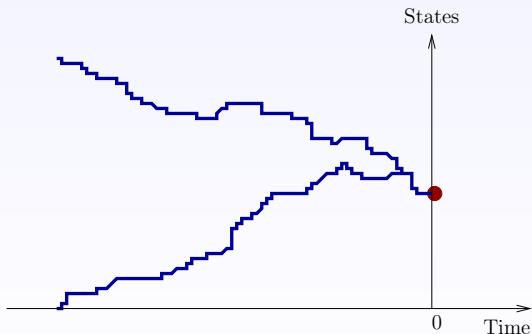
$$x \preceq y \Rightarrow \forall a \in A, x \cdot a \preceq y \cdot a.$$



SED monotone

On suppose que l'espace d'état est muni d'une relation d'ordre (\preceq) et que la fonction de transition de l'automate est monotone :

$$x \preceq y \Rightarrow \forall a \in A, x \cdot a \preceq y \cdot a.$$



SED monotone (II)

Soit E l'ensemble des éléments extrémaux de \mathcal{X} pour \preceq .

On note $\bar{U} \stackrel{\text{def}}{=} \{x \in \mathcal{X} \text{ t.q. } \exists u, v \in U \mid u \preceq x \preceq v\}$.

Par définition, $\mathcal{X} = \bar{E}$.

SED monotone (II)

Soit E l'ensemble des éléments extrémaux de \mathcal{X} pour \preceq .

On note $\overline{U} \stackrel{\text{def}}{=} \{x \in \mathcal{X} \text{ t.q. } \exists u, v \in U \mid u \preceq x \preceq v\}$.

Par définition, $\mathcal{X} = \overline{E}$.

Par monotonie, $\overline{U} \cdot a \subseteq \overline{U \cdot a}$,

et donc, pour tout mot a_n, \dots, a_1 ,

$$E \cdot a_{n \rightarrow 1} \subseteq \mathcal{X} \cdot a_{n \rightarrow 1} \subseteq \overline{E \cdot a_{n \rightarrow 1}}.$$

Toutes les trajectoires sont “capturées” entre les trajectoires extrémales.

SED monotone (II)

Soit E l'ensemble des éléments extrémaux de \mathcal{X} pour \preceq .

On note $\overline{U} \stackrel{\text{def}}{=} \{x \in \mathcal{X} \text{ t.q. } \exists u, v \in U \mid u \preceq x \preceq v\}$.

Par définition, $\mathcal{X} = \overline{E}$.

Par monotonie, $\overline{U} \cdot a \subseteq \overline{U \cdot a}$,

et donc, pour tout mot a_n, \dots, a_1 ,

$$E \cdot a_{n \rightarrow 1} \subseteq \mathcal{X} \cdot a_{n \rightarrow 1} \subseteq \overline{E \cdot a_{n \rightarrow 1}}.$$

Toutes les trajectoires sont “capturées” entre les trajectoires extrémales.

Si a_n, \dots, a_1 est un mot couplant pour E ,

$$E \cdot a_{n \rightarrow 1} = \mathcal{X} \cdot a_{n \rightarrow 1} = \overline{E \cdot a_{n \rightarrow 1}}.$$

Exemple

file M/M/1/3

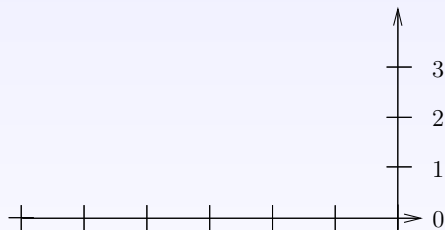
Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée

avec $\Lambda = \lambda + \mu$.



Exemple

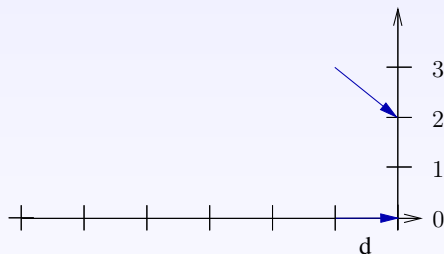
file M/M/1/3

Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée
avec $\Lambda = \lambda + \mu$.



Exemple

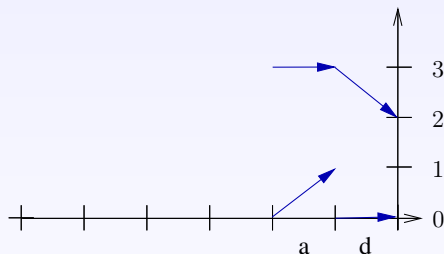
file M/M/1/3

Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée
avec $\Lambda = \lambda + \mu$.



Exemple

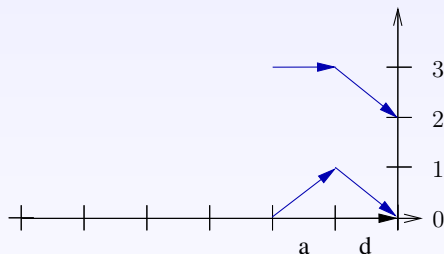
file M/M/1/3

Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée
avec $\Lambda = \lambda + \mu$.



Exemple

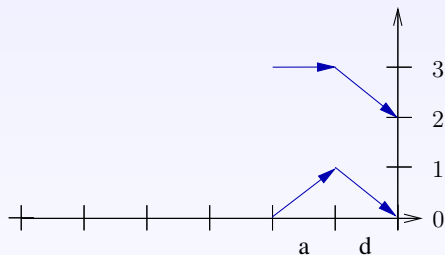
file M/M/1/3

Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée
avec $\Lambda = \lambda + \mu$.



Exemple

file M/M/1/3

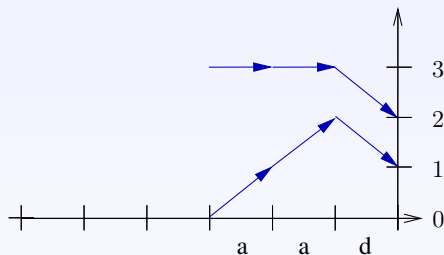
Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée

avec $\Lambda = \lambda + \mu$.



Exemple

file M/M/1/3

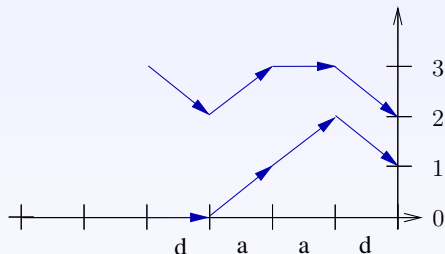
Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée

avec $\Lambda = \lambda + \mu$.



Exemple

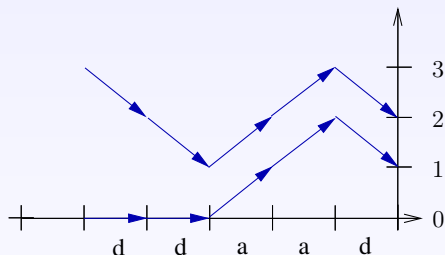
file M/M/1/3

Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée
avec $\Lambda = \lambda + \mu$.



Exemple

file M/M/1/3

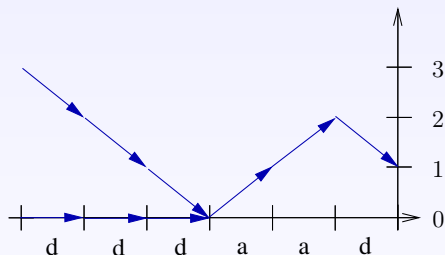
Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée

avec $\Lambda = \lambda + \mu$.



Exemple

file M/M/1/3

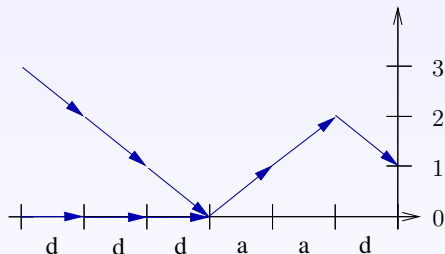
Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée

avec $\Lambda = \lambda + \mu$.



Inconvénient : Complexité $O((\tau^b)^2 \times \mathcal{C} \times |E|)$
(comparé à $O(\tau^b \times \mathcal{C} \times |\mathcal{X}|)$).

Exemple

file M/M/1/3

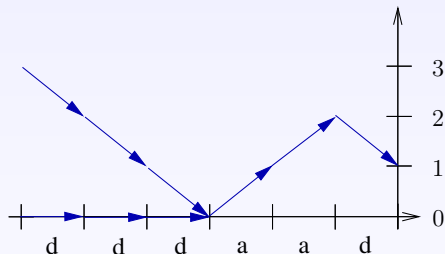
Événements :

a - arrivée (λ)

d - départ (μ)

Chaîne uniformisée

avec $\Lambda = \lambda + \mu$.



Inconvénient : Complexité $O((\tau^b)^2 \times \mathcal{C} \times |E|)$

(comparé à $O(\tau^b \times \mathcal{C} \times |\mathcal{X}|)$).

Solution : **doublement de l'intervalle.**

Algorithme de simulation parfaite monotone

Data: $(\mathcal{X}, A, \cdot, p)$, $E = \{ \text{états extrémaux} \}$.

Result: $x \in \mathcal{X}$ généré selon π

begin

$n = 1$;

repeat

$S := E$;

for $i = n$ **downto** $\lfloor n/2 \rfloor + 1$ **do**

 └ Tirer a_i selon p

for $i = n$ **downto** 1 **do**

 └ $S := S \cdot a_i$;

$n := 2n$;

until $|S| = 1$;

return x *unique état dans* S ;

Questions

Question

Quelle est la complexité de l'algorithme PSA-monotone ?

Questions

Question

Quelle est la complexité de l'algorithme PSA-monotone ?

Question

L'algorithme PSA-monotone nécessite de garder en mémoire la suite a_1, a_2, \dots et de les réutiliser à chaque doublement de l'intervalle. Que se passe-t-il si on utilise des nouveaux tirages (indépendants) à chaque fois ?

Questions

Question

Quelle est la complexité de l'algorithme PSA-monotone ?

Question

L'algorithme PSA-monotone nécessite de garder en mémoire la suite a_1, a_2, \dots et de les réutiliser à chaque doublement de l'intervalle. Que se passe-t-il si on utilise des nouveaux tirages (indépendants) à chaque fois ?

Question

On suppose que la fonction de transition de l'automate est vérifiée pour chaque $a \in A$:

- ▶ $x \preceq y \Rightarrow x \cdot a \preceq y \cdot a$ (événement monotone) ou
- ▶ $x \preceq y \Rightarrow x \cdot a \succeq y \cdot a$ (événement anti-monotone).

Montrer qu'on peut utiliser l'algorithme PSA-monotone pour échantillonner π .

Chaînes de Markov

- ▶ Une **chaîne de Markov** (en temps discret) sur un espace fini \mathcal{X} est une suite de variables aléatoires X_0, X_1, \dots sur \mathcal{X} telles que pour tout $n \in \mathbb{N}$, et tout $i_n, i_{n-1}, \dots, i_0 \in \mathcal{X}$:

$$\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}).$$

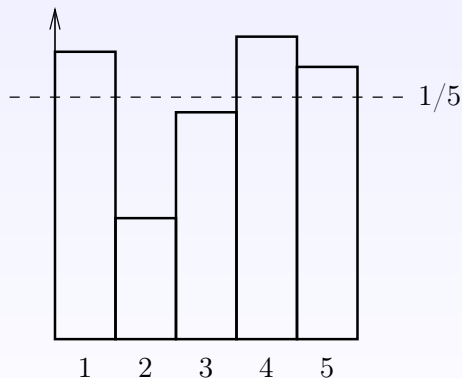
- ▶ Si $\mathbb{P}(X_n = j | X_{n-1} = i) = P_{i,j}$, $\forall n$, alors on dit que la chaîne est homogène et on appelle P la **matrice de transition**.
- ▶ Notation : π^n ($\pi_i^n \stackrel{\text{def}}{=} \mathbb{P}(X_n = i)$) la distribution de X_n . Alors,

$$\pi^n = \pi^{n-1} P.$$

- ▶ Si la chaîne est irréductible et apériodique, alors il existe la distribution limite $\pi = \lim_{n \rightarrow \infty} \pi^n$. Elle vérifie $\pi P = \pi$ et on l'appelle la **distribution stationnaire**.

Échantillonnage par table d'alias (Walker 1977)

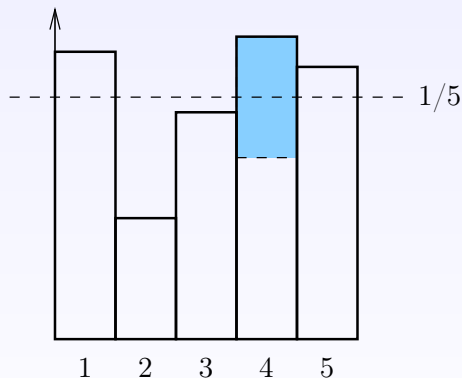
← retour



On commence par construire un histogramme de μ .

Échantillonnage par table d'alias (Walker 1977)

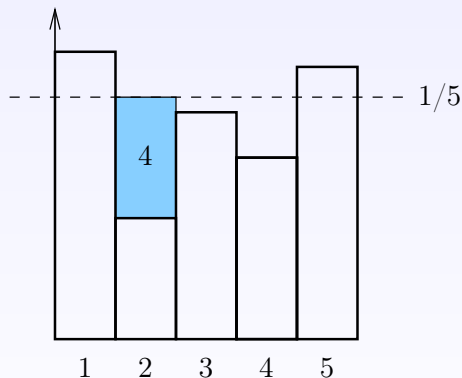
← retour



On remplit la colonne la plus basse jusqu'au niveau $1/N$ en puisant dans la colonne la plus haute.

Échantillonnage par table d'alias (Walker 1977)

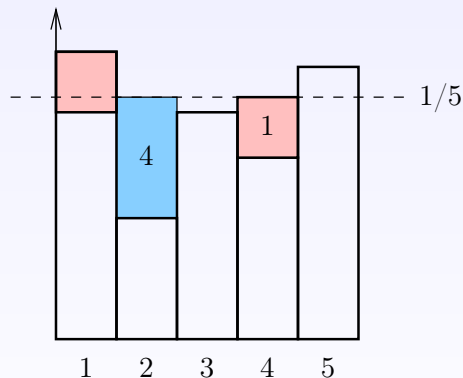
← retour



On remplit la colonne la plus basse jusqu'au niveau $1/N$ en puisant dans la colonne la plus haute.

Échantillonnage par table d'alias (Walker 1977)

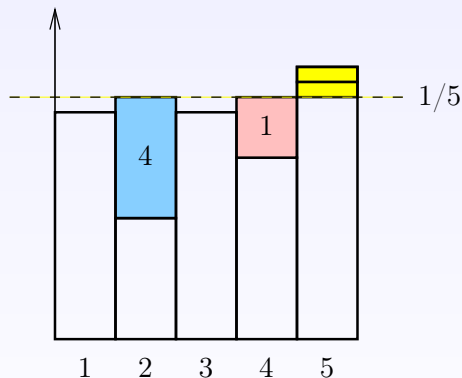
← retour



On recommence tant que toutes les colonnes ne sont pas exactement au niveau $1/N$. Au-delà du seuil $S[i]$, la colonne i correspond à la valeur $V[i]$.

Échantillonnage par table d'alias (Walker 1977)

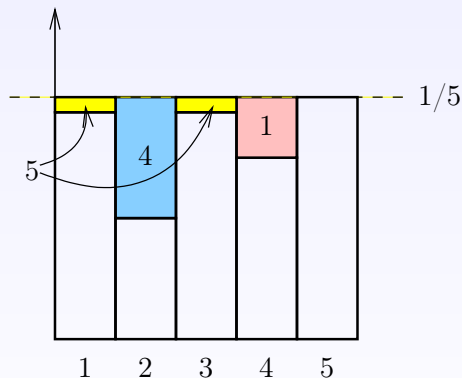
← retour



On recommence tant que toutes les colonnes ne sont pas exactement au niveau $1/N$. Au-delà du seuil $S[i]$, la colonne i correspond à la valeur $V[i]$.

Échantillonnage par table d'alias (Walker 1977)

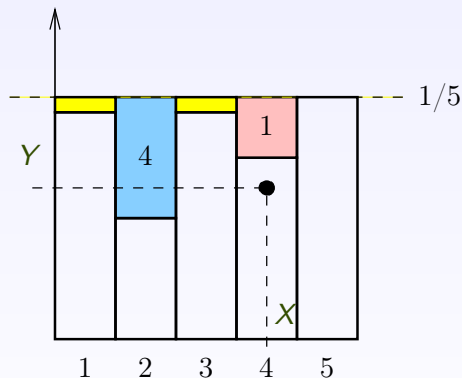
← retour



On recommence tant que toutes les colonnes ne sont pas exactement au niveau $1/N$. Au-delà du seuil $S[i]$, la colonne i correspond à la valeur $V[i]$.

Échantillonnage par table d'alias (Walker 1977)

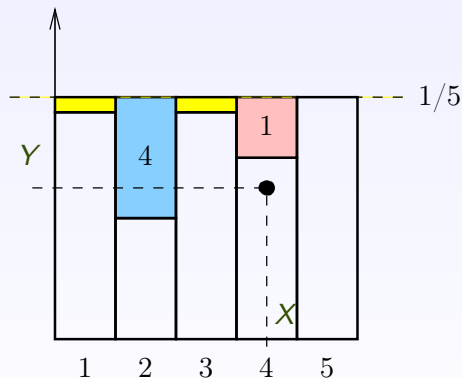
← retour



On tire X et Y deux variables uniformes, indépendantes : X dans $\{1, \dots, N\}$, Y dans $[0, 1/N]$. Par construction des tables S et V , la quantité $X\mathbf{1}_{Y < S[X]} + V[X]\mathbf{1}_{Y \geq S[X]}$ est distribuée selon μ .

Échantillonnage par table d'alias (Walker 1977)

← retour



Une fois les tableaux S et V construits, le tirage se fait en temps constant.