Chaînes de Markov, Génération Aléatoire, et Couplage Depuis le Passé (8)

Petite École de Combinatoire

P. Duchon

LaBRI

The story so far…

On a :

- montré qu'une chaîne de Markov converge vers la distribution voulue;

The story so far…

On a :

- montré qu'une chaîne de Markov converge vers la distribution voulue;
- vu des méthodes pour évaluer la vitesse à laquelle a lieu cette convergence (estimation du nombre de pas de simulation nécessaires pour une simulation à ϵ près de la distribution limite).

The story so far...

On a :

- montré qu'une chaîne de Markov converge vers la distribution voulue;
- vu des méthodes pour évaluer la vitesse à laquelle a lieu cette convergence (estimation du nombre de pas de simulation nécessaires pour une simulation à ϵ près de la distribution limite).

Couplage Venu du Passé (CFTP - Coupling From the Past) : une technique permettant (parfois) de simuler exactement la distribution limite (Propp, Wilson '96)

Couplage par fonctions de mise à jour

Une façon bien pratique de visualiser le couplage : on simule la chaîne de base par une procédure de mise à jour :

- Tirer une fonction $F : \Omega \to \Omega$, aléatoire, suivant une certaine distribution;

Couplage par fonctions de mise à jour

Une façon bien pratique de visualiser le couplage : on simule la chaîne de base par une procédure de mise à jour :

- Tirer une fonction $F : \Omega \to \Omega$, aléatoire, suivant une certaine distribution;
- Si l'état courant est x, le nouvel état est $x' = F(x)$

Couplage par fonctions de mise à jour

Une façon bien pratique de visualiser le couplage : on simule la chaîne de base par une procédure de mise à jour :

- Tirer une fonction $F : \Omega \to \Omega$, aléatoire, suivant une certaine distribution;
- Si l'état courant est x, le nouvel état est $x' = F(x)$

La loi sur les fonctions doit satisfaire :

$$\forall (x, y) \in \Omega^2, \Pr(F(x) = y) = p_{xy}$$

Couplage par fonctions de mise à jour

Une façon bien pratique de visualiser le couplage : on simule la chaîne de base par une procédure de mise à jour :

- Tirer une fonction $F : \Omega \to \Omega$, aléatoire, suivant une certaine distribution;
- Si l'état courant est x, le nouvel état est $x' = F(x)$

Pour le même prix, on a un couplage généralisé : une copie $(X_n^{(u)})_n$ de la chaîne partant de chaque état u, avec automatiquement la condition de convergence

$$X_n^{(u)} = X_n^{(v)} \Rightarrow X_{n+1}^{(u)} = X_{n+1}^{(v)}$$

Remarque : fonctions de mise à jour

On peut toujours simuler la chaîne par une loi sur les fonctions de Ω vers Ω, par exemple en définissant la loi par :

- les images $(F(u))_{u\in\Omega}$ sont indépendantes;

Remarque : fonctions de mise à jour

On peut toujours simuler la chaîne par une loi sur les fonctions de Ω vers Ω, par exemple en définissant la loi par :

- les images $(F(u))_{u\in\Omega}$ sont indépendantes;
- chaque $F(u)$ est distribuée selon $(p_{u,.})$

Remarque : fonctions de mise à jour

On peut toujours simuler la chaîne par une loi sur les fonctions de Ω vers Ω, par exemple en définissant la loi par :

- les images $(F(u))_{u\in\Omega}$ sont indépendantes;
- chaque $F(u)$ est distribuée selon $(p_{u,.})$

C'est juste, en général, une mauvaise idée : tirer une fonction aléatoire ne sera pas réalisable en pratique; on a besoin d'une loi (sur les fonctions) qu'on sache simuler facilement (au moins, le nombre de fonctions possibles doit être raisonnable).

Exemple : marche sur un cycle

$\Omega = \{1, 2, \ldots, N\}$; $p_{i,i} = p_{i,i+1} = 1/2$

Exemple : marche sur un cycle

$\Omega = \{1, 2, \ldots, N\}$; $p_{i,i} = p_{i,i+1} = 1/2$
On peut réaliser un couplage généralisé de différentes manières

Exemple : marche sur un cycle

$\Omega = \{1, 2, \dots, N\}$; $p_{i,i} = p_{i,i+1} = 1/2$
On peut réaliser un couplage généralisé de différentes manières

- Chaque image $F(i)$ choisie independamment parmi $\{i, i+1\}$: 2^N fonctions possibles;

Exemple : marche sur un cycle

$\Omega = \{1, 2, \ldots, N\}$; $p_{i,i} = p_{i,i+1} = 1/2$
On peut réaliser un couplage généralisé de différentes manières

- Chaque image $F(i)$ choisie independamment parmi $\{i, i+1\}$: 2^N fonctions possibles;
- Tirage à pile ou face : pile, $F(i) = i$ pour tout i; face, $F(i) = i + 1$ pour tout i (2 fonctions possibles);

Exemple : marche sur un cycle

$\Omega = \{1, 2, \ldots, N\}$; $p_{i,i} = p_{i,i+1} = 1/2$
On peut réaliser un couplage généralisé de différentes manières

- Chaque image $F(i)$ choisie independamment parmi $\{i, i+1\}$: 2^N fonctions possibles;
- Tirage à pile ou face : pile, $F(i) = i$ pour tout i; face, $F(i) = i + 1$ pour tout i (2 fonctions possibles);
- Tirage à pile ou face : pile, $F(i) = i$ pour i pair, $F(i) = i + 1$ pour i impair; face, le contraire (2 fonctions possibles).

Exemple : marche sur un cycle

$\Omega = \{1, 2, \ldots, N\}$; $p_{i,i} = p_{i,i+1} = 1/2$
On peut réaliser un couplage généralisé de différentes manières

- Chaque image $F(i)$ choisie independamment parmi $\{i, i+1\}$: 2^N fonctions possibles;
- Tirage à pile ou face : pile, $F(i) = i$ pour tout i; face, $F(i) = i + 1$ pour tout i (2 fonctions possibles);
- Tirage à pile ou face : pile, $F(i) = i$ pour i pair, $F(i) = i + 1$ pour i impair; face, le contraire (2 fonctions possibles).

Seule la troisième est une “bonne” proposition; en règle générale, il faut éviter que les fonctions de mise à jour soient bijectives.

Simulation exacte, mais idiote

Pour tout entier n, on obtient une simulation exacte de la distribution limite π, par:

- tirer n fonctions aléatoires indépendantes $F_1, \dots, F_n$;
- calculer $G = F_n \circ F_{n-1} \circ \dots \circ F_1$;
- choisir u selon la loi π;
- retourner $G(u)$.

Simulation exacte, mais idiote

Pour tout entier n, on obtient une simulation exacte de la distribution limite π, par:

- tirer n fonctions aléatoires indépendantes $F_1, \ldots, F_n$;
- calculer $G = F_n \circ F_{n-1} \circ \ldots \circ F_1$;
- choisir u selon la loi π;
- retourner $G(u)$.

(C'est idiot : si on sait tirer u selon la loi π, on n'a pas besoin d'un algorithme plus compliqué pour simuler la loi)

Oui, mais. . .

Si la loi de F est bien choisie, pour n assez grand, la fonction G sera très probablement une fonction constante; pas besoin de choisir astucieusement u, il suffit de retourner $G(u_0)$ pour un u_0 quelconque (puisque le résultat ne dépend pas de u).

Oui, mais...

Si la loi de F est bien choisie, pour n assez grand, la fonction G sera très probablement une fonction constante; pas besoin de choisir astucieusement u, il suffit de retourner $G(u_0)$ pour un u_0 quelconque (puisque le résultat ne dépend pas de u).

(Il y a un piège...)

Remarque sur la coalescence

(Coalescence : la composée finit par être constante)
De deux choses l'une :

- soit, pour tout N, la composée de N fonctions aléatoires a probabilité 0 d'être constante;

Remarque sur la coalescence

(Coalescence : la composée finit par être constante)
De deux choses l'une :

- soit, pour tout N, la composée de N fonctions aléatoires a probabilité 0 d'être constante;
- soit cette coalescence se produit avec probabilité 1, et l'espérance du nombre de fonctions à composer est finie.

Remarque sur la coalescence

(Coalescence : la composée finit par être constante)
De deux choses l'une :

- soit, pour tout N, la composée de N fonctions aléatoires a probabilité 0 d'être constante;
- soit cette coalescence se produit avec probabilité 1, et l'espérance du nombre de fonctions à composer est finie.

Preuve : si la probabilité n'est pas nulle, alors la composée de N_0 fonctions a probabilité $\epsilon > 0$ d'être constante, et, chaque fois qu'on compose N_0 nouvelles fonctions, on a cette même probabilité de coalescence; le temps avant coalescence a donc une espérance d'au plus $N_0/\epsilon < +\infty$.

Couplage vers le futur

(**Idée :** Si on attend assez longtemps, le couplage va bien finir par être coalescent)

Couplage vers le futur

(**Idée :** Si on attend assez longtemps, le couplage va bien finir par être coalescent)

- $G \leftarrow I$
- Tant que G n'est pas constante, $G \leftarrow \text{RandomF}() \circ G$
- Retourner $G(u_0)$

Couplage vers le futur

(**Idée :** Si on attend assez longtemps, le couplage va bien finir par être coalescent)

- $G \leftarrow I$, $u \leftarrow u_0$
- Tant que G n'est pas constante, $F \leftarrow$ RandomF$()$; $G \leftarrow F \circ G$; $u \leftarrow F(u)$
- Retourner u

Couplage vers le futur

(**Idée :** Si on attend assez longtemps, le couplage va bien finir par être coalescent)

- $G \leftarrow I$, $u \leftarrow u_0$
- Tant que G n'est pas constante, $F \leftarrow$ RandomF$()$; $G \leftarrow F \circ G$; $u \leftarrow F(u)$
- Retourner u

C'est un couplage vers le futur : après n étapes, $G = G_n = F_n \circ \ldots \circ F_1$; $G_n(u) =$ RandomF$()(G_{n-1}(u))$.

Couplage venu du passé

- $G \leftarrow I$
- Tant que G n'est pas constante, $G \leftarrow G \circ \text{RandomF}()$
- Retourner $G(u_0)$

Couplage venu du passé

- $G \leftarrow I$
- Tant que G n'est pas constante, $G \leftarrow G \circ \text{RandomF}()$
- Retourner $G(u_0)$

C'est un couplage venu du passé : $G_n = F_1 \circ \ldots \circ F_n$; en renumérotant les F_i en F_{-i}, $G_n = F_{-1} \circ F_{-2} \circ \ldots \circ F_{-n}$.

Couplage venu du passé

- $G \leftarrow I$
- Tant que G n'est pas constante, $G \leftarrow G \circ \text{RandomF}()$
- Retourner $G(u_0)$

C'est un couplage venu du passé : $G_n = F_1 \circ \ldots \circ F_n$; en renumérotant les F_i en F_{-i}, $G_n = F_{-1} \circ F_{-2} \circ \ldots \circ F_{-n}$. $G_n(u) = G_{n-1}(\text{RandomF}(u))$: pour calculer l'image d'un point, la composition se fait dans le mauvais sens. . .

Couplage venu du passé

- $G \leftarrow I$
- Tant que G n'est pas constante, $G \leftarrow G \circ \text{RandomF}()$
- Retourner $G(u_0)$

C'est un couplage venu du passé : $G_n = F_1 \circ \ldots \circ F_n$; en renumérotant les F_i en F_{-i}, $G_n = F_{-1} \circ F_{-2} \circ \ldots \circ F_{-n}$. $G_n(u) = G_{n-1}(\text{RandomF}(u))$: pour calculer l'image d'un point, la composition se fait dans le mauvais sens...
Idée : Si on prend un couplage qui tourne depuis infiniment longtemps, il est forcément coalescent au temps 0; on va simplement s'arranger pour calculer où il est au temps 0.

Le piège. . .

- Le couplage vers le futur ne simule pas (en général) la loi stationnaire;

Le piège…

- Le couplage vers le futur ne simule pas (en général) la loi stationnaire;
- le couplage venu du passé la simule (pour peu que la coalescence soit possible, donc quasi-sûre).

Un exemple

$\Omega = \{1,2,3,4,5\}$, $p_{i,i+1} = p, p_{i,i-1} = 1-p$ (sauf $p_{5,5} = p$, $p_{1,1} = 1-p$)

Un exemple

$\Omega = \{1, 2, 3, 4, 5\}$, $p_{i,i+1} = p, p_{i,i-1} = 1 - p$ (sauf $p_{5,5} = p$, $p_{1,1} = 1 - p$)
Couplage réalisé par 2 fonctions :

$$\begin{aligned} F_-(i) &= \max(1, i-1) \;\; (1-p) \\ F_+(i) &= \min(5, i+1) \;\; (p) \end{aligned}$$

Un exemple

$\Omega = \{1, 2, 3, 4, 5\}$, $p_{i,i+1} = p, p_{i,i-1} = 1 - p$ (sauf $p_{5,5} = p$, $p_{1,1} = 1 - p$)
Couplage réalisé par 2 fonctions :

$$\begin{array}{rcl} F_-(i) & = & \max(1, i-1) \ \ (1-p) \\ F_+(i) & = & \min(5, i+1) \ \ (p) \end{array}$$

Ce couplage finira par être coalescent : chaque fois qu'on compose 4 fonctions, on a probabilité $p^4 + (1-p)^4$ d'avoir 4 fois F_- ou 4 fois F_+

Un exemple

$\Omega = \{1,2,3,4,5\}$, $p_{i,i+1} = p, p_{i,i-1} = 1-p$ (sauf $p_{5,5} = p$, $p_{1,1} = 1-p$)
Couplage réalisé par 2 fonctions :

$$\begin{array}{rcl} F_-(i) & = & \max(1, i-1) \;\; (1-p) \\ F_+(i) & = & \min(5, i+1) \;\; (p) \end{array}$$

Ce couplage finira par être coalescent : chaque fois qu'on compose 4 fonctions, on a probabilité $p^4 + (1-p)^4$ d'avoir 4 fois F_- ou 4 fois F_+
(Pour $p = 1/2$, le nombre moyen de compositions avant coalescence est d'au plus 32)

Pourquoi le CFTP est correct

On considère une suite doublement infinie de fonctions aléatoires indépendantes $(F_n)_{n\in\mathbb{Z}}$, et on pose ($n < m$)

$$G_{n,m} = F_{m-1} \circ F_{m-2} \circ \ldots \circ F_n$$

Pourquoi le CFTP est correct

On considère une suite doublement infinie de fonctions aléatoires indépendantes $(F_n)_{n\in\mathbb{Z}}$, et on pose ($n < m$)

$$G_{n,m} = F_{m-1} \circ F_{m-2} \circ \ldots \circ F_n$$

- En tant que fonction aléatoire, $G_{n,m}$ laisse la loi stationnaire inchangée;

Pourquoi le CFTP est correct

On considère une suite doublement infinie de fonctions aléatoires indépendantes $(F_n)_{n \in \mathbb{Z}}$, et on pose ($n < m$)

$$G_{n,m} = F_{m-1} \circ F_{m-2} \circ \ldots \circ F_n$$

- En tant que fonction aléatoire, $G_{n,m}$ laisse la loi stationnaire inchangée;
- avec probabilité 1, il existe $n < 0$ tel que $G_{n,0}$ est une fonction constante;

Pourquoi le CFTP est correct

On considère une suite doublement infinie de fonctions aléatoires indépendantes $(F_n)_{n\in\mathbb{Z}}$, et on pose ($n < m$)

$$G_{n,m} = F_{m-1} \circ F_{m-2} \circ \ldots \circ F_n$$

- En tant que fonction aléatoire, $G_{n,m}$ laisse la loi stationnaire inchangée;
- avec probabilité 1, il existe $n < 0$ tel que $G_{n,0}$ est une fonction constante;
- si $G_{n,0}$ est constante, $G_{n',0} = G_{n,0}$ pour tout $n' < n$;

Pourquoi le CFTP est correct

On considère une suite doublement infinie de fonctions aléatoires indépendantes $(F_n)_{n\in\mathbb{Z}}$, et on pose ($n < m$)

$$G_{n,m} = F_{m-1} \circ F_{m-2} \circ \ldots \circ F_n$$

- En tant que fonction aléatoire, $G_{n,m}$ laisse la loi stationnaire inchangée;
- avec probabilité 1, il existe $n < 0$ tel que $G_{n,0}$ est une fonction constante;
- si $G_{n,0}$ est constante, $G_{n',0} = G_{n,0}$ pour tout $n' < n$;
- donc, pour tout u,

$$\lim_{n\to+\infty} \mathbb{P}(G_{-n,0} = u) = \pi_u$$

Versions pratiques du CFTP

- On n'a pas besoin de calculer complètement $G_{n,0}$, seulement de détecter (éventuellement avec retard) que $G_{n,0}$ est constante;

(les cas faciles sont de ce type)

Versions pratiques du CFTP

- On n'a pas besoin de calculer complètement $G_{n,0}$, seulement de détecter (éventuellement avec retard) que $G_{n,0}$ est constante;
- **CFTP Monotone :** dans tous les cas où Ω est un ensemble ordonné avec des éléments minimaux et des éléments maximaux, et que les fonctions F sont croissantes, la coalescence est équivalente à $G(u) = G(v)$ pour tous les extremaux;

(les cas faciles sont de ce type)

Versions pratiques du CFTP

- On n'a pas besoin de calculer complètement $G_{n,0}$, seulement de détecter (éventuellement avec retard) que $G_{n,0}$ est constante;
- **CFTP Monotone :** dans tous les cas où Ω est un ensemble ordonné avec des éléments minimaux et des éléments maximaux, et que les fonctions F sont croissantes, la coalescence est équivalente à $G(u) = G(v)$ pour tous les extremaux;
- en particulier, si Ω est un treillis distributif fini (maximum et minimum uniques), il suffit de calculer $G_{n,0}(\max)$ et $G_{n,0}(\min)$; (les cas faciles sont de ce type)

Versions pratiques du CFTP

- On n'a pas besoin de calculer complètement $G_{n,0}$, seulement de détecter (éventuellement avec retard) que $G_{n,0}$ est constante;
- **CFTP Monotone :** dans tous les cas où Ω est un ensemble ordonné avec des éléments minimaux et des éléments maximaux, et que les fonctions F sont croissantes, la coalescence est équivalente à $G(u) = G(v)$ pour tous les extremaux;
- en particulier, si Ω est un treillis distributif fini (maximum et minimum uniques), il suffit de calculer $G_{n,0}(\max)$ et $G_{n,0}(\min)$; (les cas faciles sont de ce type)
- **binary-backoff CFTP :** on calcule $G_{-2^k,0}$ pour $k = 1, 2, \ldots$, en stockant les fonctions F_n de manière à les réutiliser; ainsi, la composition se fait “dans le bon sens”