Chaînes de Markov, Génération Aléatoire, et Couplage Depuis le Passé

Petite École de Combinatoire

P. Duchon

LaBRI

Chaînes de Markov

- Qu'est-ce qu'une chaîne de Markov ?
- Comment ça se définit ?
- À quoi ça peut servir ?
- Et le CFTP, ça marche comment ?

Chaîne de Markov

V ensemble fini ou dénombrable
Définition : Une *chaîne de Markov à valeurs dans V* est une suite $(X_n)_{n\geq 0}$ de variables aléatoires à valeurs dans V, et telle que, pour tous $n, x_0, \ldots, x_n, y$,

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = y | X_0 = x_0, \ldots, X_n = x_n) \\
&= \mathbb{P}(X_{n+1} = y | X_n = x_n) \\
&= f(n, x_n, y)
\end{aligned}$$

La chaîne est *homogène* si f ne dépend pas de n.

Idée: la connaissance de la trajectoire $(X_0, \ldots, X_n)$ n'est pas utile pour prédire X_{n+1} : connaître la trajectoire n'apporte pas plus d'information que de connaître la seule valeur de X_n (et de n si la chaîne n'est pas homogène).

Un exemple non homogène

N un entier fixé, w un chemin de Dyck aléatoire (uniforme) de longueur $2N$; on note $h_n(w)$ la hauteur finale du n-ème pas de w, et on pose

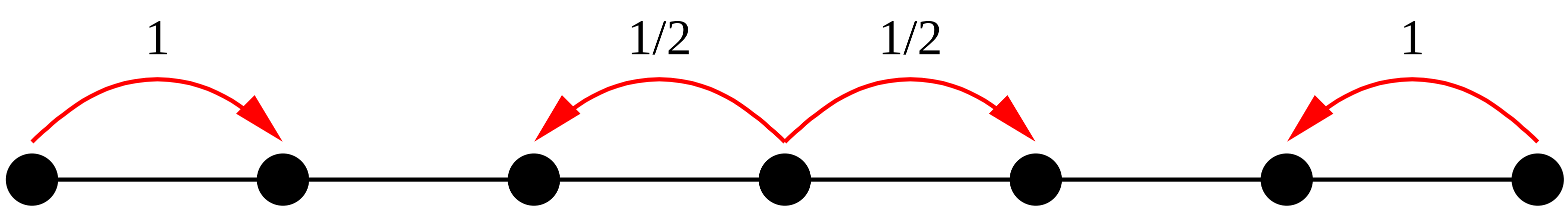
$$X_n = \begin{cases} 0 & \text{si } n \geq 2N \\ h_n(w) & \text{si } n < 2N \end{cases}$$

(X_n) est une chaîne de Markov (non homogène) : la probabilité, dans un chemin de Dyck aléatoire de longueur $2N$, que le pas qui suit un préfixe donné soit un pas montant ($h_n(w) = h_{n-1}(w) + 1$) est la même pour tous les préfixes de Dyck de même longueur finissant à même hauteur.

Un exemple homogène

Marche aléatoire non biaisée sur un segment :
$V = \{1, \ldots, N\}$, f définie par

$$
\begin{array}{rcll}
f(k, k+1) & = & 1/2 & \text{pour } 1 < k < N \\
f(k, k-1) & = & 1/2 & \text{pour } 1 < k < N \\
f(1, 2) & = & f(N, N-1) = 1 & \\
f(k, \ell) & = & 0 & \text{sinon}
\end{array}
$$



Monsieur Jourdain…

En fait, une chaîne de Markov homogène n'est rien d'autre qu'une marche aléatoire (biaisée, en général) sur un graphe orienté (pondéré, à poids positifs) simple (avec boucles), où les poids des arcs sont les probabilités de transition (d'où la condition : la somme des poids des arcs sortant d'un sommet doit être 1).

Matrice de transition

- $\mathbf{X}$ une chaîne homogène : sa matrice de transition est M :
$$M_{xy} = f(x,y) = \mathbb{P}(X_{t+1} = y | X_t = x)$$
C'est la matrice d'adjacence (pondérée) du graphe!
- Itération/Puissances : M^n a pour coefficents
$$p_{xy}^{(n)} = \mathbb{P}(X_{t+n} = y | X_t = x) \tag{1}$$
(somme des poids des chemins de x à y, de longueur n)
- Si la loi de probabilités de X_0 est $\pi = (\pi_x)_{x \in V}$, la loi de X_n n'est autre que
$$\pi^{(n)} = \pi . M^n$$
(combinaison linéaire de (1))

Définition d'une chaîne de Markov

Pour définir une chaîne de Markov, on doit normalement définir

- la distribution au temps 0 (loi de X_0)
- la matrice (loi) de transition.

Dans la pratique, beaucoup de propriétés ne dépendent pas (ou peu) de la loi initiale, mais essentiellement des lois de transition; on parle de la chaîne partant de u ou la chaîne partant de v, si elles ont la même matrice de transition.

Irréductibilité

Accessibilité : *Un état v est accessible à partir d'un état u si la chaîne, partant de u, a probabilité strictement positive de passer par v.*

Irréductibilité

Accessibilité : *Un état v est accessible à partir d'un état u si la chaîne, partant de u, a probabilité strictement positive de passer par v.*
C'est équivalent à : il existe un n tel que $\mathbb{P}(X_n = v | X_0 = u) > 0$, c'est-à-dire qu'il existe un chemin de u à v dans le graphe

Irréductibilité

Accessibilité : *Un état* v *est accessible à partir d'un état* u *si la chaîne, partant de* u*, a probabilité strictement positive de passer par* v*.*
(accessibilité dans le graphe)

Irréductibilité

Accessibilité : *Un état v est accessible à partir d'un état u si la chaîne, partant de u, a probabilité strictement positive de passer par v.*
(accessibilité dans le graphe)

Irréductibilité : *Les classes d'équivalence pour la relation de double accessibilité, sont appelées ensembles irréductibles d'états (de sommets).*

Irréductibilité

Accessibilité : *Un état v est accessible à partir d'un état u si la chaîne, partant de u, a probabilité strictement positive de passer par v.*
(accessibilité dans le graphe)

Irréductibilité : *Les classes d'équivalence pour la relation de double accessibilité, sont appelées ensembles irréductibles d'états (de sommets).*
(composantes fortement connexes du graphe)

Irréductibilité

Accessibilité : *Un état v est accessible à partir d'un état u si la chaîne, partant de u, a probabilité strictement positive de passer par v.*
(accessibilité dans le graphe)

Irréductibilité : *Les classes d'équivalence pour la relation de double accessibilité, sont appelées ensembles irréductibles d'états (de sommets).*
(composantes fortement connexes du graphe)

Chaîne irréductible : *La chaîne est dite irréductible si l'ensemble de tous les états est irréductible.*

Irréductibilité

Accessibilité : *Un état v est accessible à partir d'un état u si la chaîne, partant de u, a probabilité strictement positive de passer par v.*
(accessibilité dans le graphe)

Irréductibilité : *Les classes d'équivalence pour la relation de double accessibilité, sont appelées ensembles irréductibles d'états (de sommets).*
(composantes fortement connexes du graphe)

Chaîne irréductible : *La chaîne est dite irréductible si l'ensemble de tous les états est irréductible.*
(graphe fortement connexe)

Récurrence

Définition (état et chaîne récurrents) : Un état u est récurrent, si la probabilité que la chaîne, partant de u, repasse par u, est 1. Dans le cas contraire, il est transient. Une chaîne est récurrente si tous ses états sont récurrents.

Définition (temps de premier passage) : Le temps de premier passage en v est

$$T_v = \min\{n > 0 : X_n = v\} \cup \{+\infty\}$$

Un état est récurrent, si et seulement si, pour la chaîne partant de cet état, le temps de premier retour en cet état est fini avec probabilité 1.

Récurrence (suite)

On note
$$q_{uv}^{(n)} = \mathbb{P}(T_v = n | X_0 = u),$$
la probabilité que le premier passage en v (à un temps > 0) de la chaîne partie de u, ait lieu au temps n, et $P_{uv}(x), Q_{uv}(x)$ les séries génératrices :
$$P_{uv}(x) = \sum_{n \geq 0} p_{uv}^{(n)} x^n, \quad Q_{uv}(x) = \sum_{n > 0} q_{uv}^{(n)} x^n.$$

Remarque : $Q_{uv}(1)$ est la probabilité de passer par v en partant de u; $P_{uv}(1)$ est l'espérance du nombre de passages par v, en partant de u.

Récurrence (suite)

Par concaténation des chemins :

$$P_{uv}(x) = \delta_{uv} + Q_{uv}(x)P_{vv}(x)$$

$u = v$: la série $P_{uu}(x)$ converge en $x = 1$, si et seulement si $Q_{uu}(x) < 1$, c'est-à-dire, si u est transient.

Conséquence :

- si v est récurrent, alors si v est accessible à partir de u (ie $Q_{uv}(1) > 0$), on a $P_{uv}(1) = +\infty$;
- si v est transient, alors $P_{uv}(1) < +\infty$ pour tout u.

Récurrence (suite)

Proposition : Soit u un état récurrent, et N_u le nombre total de passages en u. Alors $\mathbb{P}(N_u = +\infty | X_0 = u) = 1$.

Preuve : Soit $g(k, n)$ la probabilité que la chaîne (partant de u) ne passe pas par u entre les temps k (exclu) et $k + n$ (inclus). À k fixé, g est une fonction décroissante de n, et sa limite $\overline{g}(k)$ est la probabilité de ne pas passer par u au-delà du temps k. La propriété de Markov donne :

$$g(N, n) = \sum_{k=0}^{N} \mathbb{P}(X_k = u) g(0, N - k + n) \leq (N + 1) g(0, n).$$

Or $\lim_n g(0, n) = 0$ (u est récurrent), donc $\forall N, \overline{g}(N) = 0$. On a donc

$$\mathbb{P}(N_u < +\infty) = \lim_N \mathbb{P}(N_u \leq N) = 0$$

Récurrence (encore)

Proposition : Tout état accessible à partir d'un état récurrent est également récurrent, et se trouve dans la même composante fortement connexe.

Preuve : Soit n_0 tel que $p_{uv}^{(n)} = a > 0$. Si u n'était pas accessible à partir de v, N_u ne pourrait pas être supérieur à n_0 avec probabilité supérieure à $1 - a$.

Pour la récurrence de v : prenons n_1 tel que $p_{vu}^{(n_1)} > 0$; on a forcément

$$p_{vv}^{(n_0+n+n_1)} \geq p_{vu}^{(n_0)} p_{uu}^{(n)} p_{uv}^{(n_1)}$$

En sommant sur n, il vient

$$P_{vv}(1) \geq \sum_n p_{vv}^{(n_0+n+n_1)} \geq +\infty$$

Récurrence (fin!)

Théorème : Les classes récurrentes sont forcément des puits du graphe des composantes fortement connexes.

Récurrence (fin!)

Théorème : Les classes récurrentes sont forcément des puits du graphe des composantes fortement connexes.

Proposition : Toute chaîne homogène d'espace d'états fini, a au moins un état récurrent.

Récurrence (fin!)

Théorème : Les classes récurrentes sont forcément des puits du graphe des composantes fortement connexes.

Proposition : Toute chaîne homogène d'espace d'états fini, a au moins un état récurrent.

Preuve : $\sum_v p_{uv}^{(n)} = 1$, donc $\sum_{n,v} p_{uv}^{(n)} = +\infty$, donc l'un au moins des $\sum_n p_{uv}^{(n)}$ est infini: v est donc récurrent.

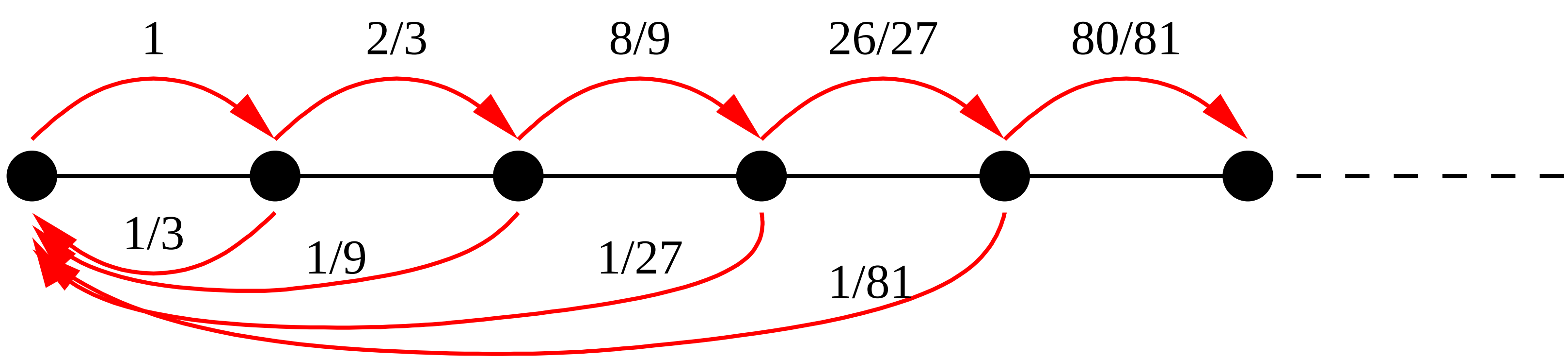
Récurrence (fin!)

Théorème : Les classes récurrentes sont forcément des puits du graphe des composantes fortement connexes.

Proposition : Toute chaîne homogène d'espace d'états fini, a au moins un état récurrent.

Preuve : $\sum_v p_{uv}^{(n)} = 1$, donc $\sum_{n,v} p_{uv}^{(n)} = +\infty$, donc l'un au moins des $\sum_n p_{uv}^{(n)}$ est infini: v est donc récurrent.

Note : ce serait faux pour un espace d'états dénombrable!



Récurrence (fin!)

Théorème : Les classes récurrentes sont forcément des puits du graphe des composantes fortement connexes.

Proposition : Toute chaîne homogène d'espace d'états fini, a au moins un état récurrent.

Preuve : $\sum_v p_{uv}^{(n)} = 1$, donc $\sum_{n,v} p_{uv}^{(n)} = +\infty$, donc l'un au moins des $\sum_n p_{uv}^{(n)}$ est infini: v est donc récurrent.

Corollaire : pour un graphe fini, connexe, tous les états sont récurrents si et seulement si le graphe est fortement connexe

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

- Existe-t-il une loi stationnaire ?

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

- Existe-t-il une loi stationnaire ?
- Si oui, est-elle unique ?

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

- Existe-t-il une loi stationnaire ?
- Si oui, est-elle unique ?
- A-t-on alors automatiquement $\lim_n \mathbb{P}(X_n = u) = \pi_u$?

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

- Existe-t-il une loi stationnaire ? Oui si V fini
- Si oui, est-elle unique ?
- A-t-on alors automatiquement $\lim_n \mathbb{P}(X_n = u) = \pi_u$?

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

- Existe-t-il une loi stationnaire ? Oui si V fini
- Si oui, est-elle unique ? Oui si irréductible
- A-t-on alors automatiquement $\lim_n \mathbb{P}(X_n = u) = \pi_u$?

Lois stationnaires

Une loi de probabilités π sur V est stationnaire pour une chaîne de Markov (pour sa matrice de transition) si, lorsque l'état au temps 0 est distribué selon π, l'état au temps 1 est également distribué selon π.
En d'autres termes, $\pi.M = \pi$, et on a donc $\pi.M^n = \pi$ (la chaîne partant de la distribution π est stationnaire).

- Existe-t-il une loi stationnaire ? Oui si V fini
- Si oui, est-elle unique ? Oui si irréductible
- A-t-on alors automatiquement $\lim_n \mathbb{P}(X_n = u) = \pi_u$? Oui si apériodique