Chaîne de Markov

Pierre-Henri WUILLEMIN

Licence d'Informatique – Université Paris 6

Introduction

Soit un système $\mathcal{S}$ possédant un **nombre d'états finis**.
On nomme $S = \{x_1, x_2, \cdots x_s\}$ les états dans lesquels peut se retrouver $\mathcal{S}$.
À chaque instant n ($n \in \mathbb{N}$), on note X_n la variable aléatoire décrivant l'état du système à cet instant n.

Définition (Processus stochastique)

La famille des $(X_n)_{n\in\mathbb{N}}$ forment un processus stochastique.

$\{X_0, X_1, X_2, \cdots\}$ représente la *trajectoire* probabiliste de $\mathcal{S}$ dans le temps.
Naturellement, le comportement d'un processus stochastique dépend fortement des relations entre les (X_n).

Par exemple, s'il existe $f : S \longrightarrow S$ bijective telle que $X_{n+1} = f(X_n)$ alors X_n ne sera que l'image de la distribution de X_O par f^n :

$$P(X_n = x) = P(f^n(X_0) = x) = P(X_0 = f^{-n}(x))$$

Chaîne de Markov

Plus fréquemment qu'une relation fonctionnelle entre les (X_n), on peut étudier des relations d'indépendances conditionnelles entre ces différentes variables aléatoires.

Propriété de Markov

Un processus stochastique vérifie la propriété de Markov (d'ordre 1) si et seulement si :

$$P(X_n \mid X_0, \cdots, X_{n-1}) = P(X_n \mid X_{n-1})$$

(lorsque cette probabilité a un sens : i.e. $P(X_0, \cdots, X_{n-1}) > 0$)

Définition (Chaîne de Markov)

Une chaîne de Markov est un processus stochastique vérifiant la propriété de Markov (d'ordre 1).

Propriété de Markov d'ordre k

Un processus stochastique vérifie la propriété de Markov d'ordre k si et seulement si :

$$\mathcal{M}(k) : P(X_n \mid X_0, \cdots, X_{n-1}) = P(X_n \mid X_{n-k}, \cdots, X_{n-1})$$

Il est évident que $\forall k' > k, \mathcal{M}(k) \Rightarrow \mathcal{M}(k')$.
En particulier, $\forall k, \mathcal{M}(1) \Rightarrow \mathcal{M}(k)$.
Réciproquement ?

Si $(X_n)_{n\in\mathbb{N}}$, processus stochastique, vérifie $\mathcal{M}(k > 1)$ alors il existe une chaîne de Markov (multidimensionnelle) équivalent à (X_n) (qui vérifie donc $\mathcal{M}(1)$).

Démonstration : Soit $(X_n)_{n\in\mathbb{N}}$ vérifie $\mathcal{M}(k > 1)$
alors on construit $\mathfrak{X}_n = [X_{n-k+1}, \cdots, X_n]$ v.a. de dimension k.
alors

$$\begin{aligned} P(\mathfrak{X}_n \mid \mathfrak{X}_0, \cdots, \mathfrak{X}_{n-1}) &= P(X_{n-k+1}, \cdots, X_n \mid X_0, \cdots, X_{n-1}) \\ &= P(X_{n-k+1}, \cdots, X_n \mid X_{n-k}, \cdots, X_{n-1}) \\ &= P(\mathfrak{X}_n \mid \mathfrak{X}_{n-1}) \end{aligned}$$

■

Dorénavant, on ne précisera plus que la chaîne de Markov est *d'ordre 1*.

Chaîne de Markov homogène

Propriété (Homogénéité)

Une chaîne de Markov est dite homogène si

$$\forall n > 0, P(X_n \mid X_{n-1}) = P(X_1 \mid X_0)$$

Définition (Probabilité et matrice de transition)

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène, alors

- *la probabilité de transition de i à j est $p_{ij} = P(X_n = j \mid X_{n-1} = i)$*
- *la matrice de transition P est la matrice des $(p_{ij})_{i,j\in S}$ (si S est fini).*

Définition (Graphe de transition)

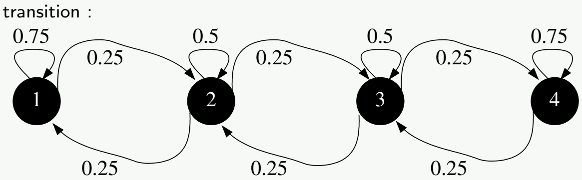
Si $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène (S fini), alors
le graphe de transition est un graphe $G = (S, E)$ orienté qui vérifie :
$(i \to j) \in E \iff p_{ij} \neq 0$

Chaîne de Markov homogène : promenade aléatoire

Mono-bestiole

une mono-bestiole se déplace dans $\{0, \cdots, s\}$ de la façon suivante : à chaque itération, elle a la probabilité 0.25 d'avancer ou de reculer et la probabilité 0.5 de rester sur place. Aux frontières, elle se cogne contre les murs et reste donc sur place. Soit X_n la v.a. de la position de la mono-bestiole à l'instant n.

- Processus stochastique ? oui
- Chaîne de Markov ? oui
- $p_{ij} = \begin{cases} 0.75 & \text{si } i=j=0 \text{ ou } i=j=s \\ 0.5 & \text{si } i=j \\ 0.25 & \text{si } |i-j|=1 \\ 0 & \text{sinon} \end{cases}$
- $s = 4$, matrice de transition $P = \begin{pmatrix} 0.75 & 0.25 & 0 & 0 \\ 0.25 & 0.5 & 0.25 & 0 \\ 0 & 0.25 & 0.5 & 0.25 \\ 0 & 0 & 0.25 & 0.75 \end{pmatrix}$
- $s = 4$, graphe de transition :

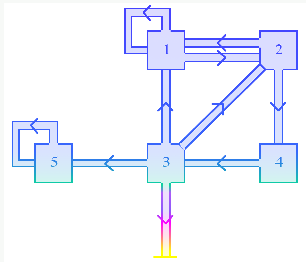


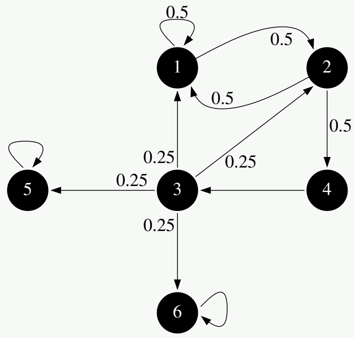
Labyrinthe (source : Bruno Baynat)

souris et labyrinthe

Une souris dans un labyrinthe :

- 5 pièces, pièce 5 sans issue ! !
- des couloirs en sens unique, pas de demi-tour
- souris sans mémoire
- loi uniforme pour le choix à chaque intersection.





Questions sachant qu'à $n = 0$, la souris est en 2 :

- $P(X_4 = 2)$?
- Combien de fois (en moyenne) en 3 avant 6 ou 5 ?
- Probabilité d'arrivée un jour en 6 ?
- Nbr de déplacement moyen pour revenir en 2 ?

Calcul de $P(X_n)$?

$$\begin{aligned}
P(X_0, \cdots X_n) &= P(X_0) \cdot P(X_1 \cdots X_n \mid X_0) \\
&= P(X_0) \cdot P(X_1 \mid X_0) \cdot P(X_2 \cdots X_n \mid X_0, X_1) \\
&= P(X_0) \cdot P(X_1 \mid X_0) \cdot P(X_2 \cdots X_n \mid X_1) \\
&= P(X_0) \cdot P(X_1 \mid X_0) \cdot P(X_2 \mid X_1) \cdot P(X_3 \cdots X_n \mid X_0, X_1, X_2)
\end{aligned}$$

Décomposition de la loi jointe dans une chaîne de Markov

Si (X_n) est une chaîne de Markov, alors

$$P(X_0, \cdots, X_n) = P(X_0) \cdot \prod_{i=1}^{n} P(X_i \mid X_{i-1})$$

$\forall j \in S, P(X_n = j) = \sum_{i \in S} P(X_n = j \mid X_{n-1} = i) \cdot P(X_{n-1} = i)$
En notant $\pi^{(n)}$ le vecteur d'état au temps n : $\pi_j^{(n)} = P(X_n = j)$.
$\forall j \in S, \pi_j^{(n)} = \sum_{i \in S} p_{ij} \cdot \pi_i^{(n-1)}$

Équation matricielle de transition

$$\pi^{(n)} = \pi^{(n-1)} \cdot P \qquad \text{et donc} \qquad \pi^{(n)} = \pi^{(0)} \cdot P^n$$

La loi d'une chaîne de Markov homogène est entièrement déterminée par $\pi^{(0)}$ et P.

Transition en m étapes

On s'intéresse ici à $P_{ij}^{(m)} = P(X_m = j \mid X_0 = i)$

Chapman-Kolmogorov

Si (X_n) est une chaîne de Markov homogène,
$P(X_{n+m} = j \mid X_0 = i) = \sum_{k \in S} P(X_n = k \mid X_0 = i) \cdot P(X_m = j \mid X_0 = k)$

Démonstration :
$P(X_{n+m} = j \mid X_0 = i) = \sum_{k \in S} P(X_n = k \mid X_0 = i) \cdot P(X_{n+m} = j \mid X_n = k, X_0 = i)$
Mais (X_n) est une chaîne de Markov, donc $X_{n+m} \perp\!\!\!\perp X_0 \,|\, X_n$
$P(X_{n+m} = j \mid X_0 = i) = \sum_{k \in S} P(X_n = k \mid X_0 = i) \cdot P(X_{n+m} = j \mid X_n = k)$
Mais (X_n) est homogène et donc invariant par translation dans le temps.
Ainsi $P(X_{n+m} = j \mid X_n = k) = P(X_m = j \mid X_0 = k)$
D'où $P(X_{n+m} = j \mid X_0 = i) = \sum_{k \in S} P(X_n = k \mid X_0 = i) \cdot P(X_m = j \mid X_0 = k)$ ∎

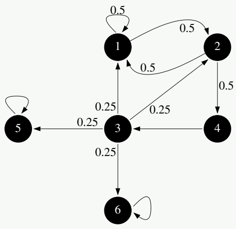
En particulier (par récurrence) :

Puissance et transition en m étapes

$$P^{(n)} = P(X_n = j \mid X_0 = i) = P^n$$

Exemples de calculs

souris et labyrinthe



Questions sachant qu'à $n = 0$, la souris est en 2 :

- $P(X_4 = 2)$?

$P(X_4 = 2) = \pi_2^{(4)}$

- $\pi^{(4)} = \pi^{(0)} \cdot P^4$ avec $\pi^{(0)} = [0, 1, 0, 0, 0, 0]$.
- $P = \begin{pmatrix} 0.5 & 0.5 & . & . & . & . \\ 0.5 & . & . & 0.5 & . & . \\ 0.25 & 0.25 & . & . & 0.25 & 0.25 \\ . & . & 1 & . & . & . \\ . & . & . & . & 1 & . \\ . & . & . & . & . & 1 \end{pmatrix}$
- $P^4 = \begin{pmatrix} 0.375 & 0.25 & 0.125 & 0.125 & 0.0625 & 0.0625 \\ 0.3125 & 0.1875 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.1875 & 0.125 & 0.0625 & 0.0625 & 0.28125 & 0.28125 \\ 0.1875 & 0.125 & 0.125 & 0.0625 & 0.25 & 0.25 \\ . & . & . & . & 1 & . \\ . & . & . & . & . & 1 \end{pmatrix}$

enfin $\pi^{(0)} \cdot P^4 = [0.3125, 0.1875, 0.125, 0.125, 0.125, 0.125]$ et $P(X_4 = 2) = 0.1875$

Classification des états

Définition (États accessibles, états communicants)

$\forall i, j \in S$, j est accessible à partir de i si et seulement si $\exists m \geq 0, P_{ij}^{(m)} > 0$
Les états i et j sont communicants si ils sont respectivement accessibles l'un par rapport à l'autre.

On note $i \leftrightarrow j$ la relation "i et j communiquent".

Classes d'états communicants

$\leftrightarrow$ est une relation d'équivalence sur S.
Les classes d'équivalence de $\leftrightarrow$ sont appelées classes d'états communicants.

Définition (Chaîne irréductible)

Une chaîne de Markov est irréductible si et seulement si elle ne possède qu'une classe d'équivalence pour $\leftrightarrow$.

Irréductibilité

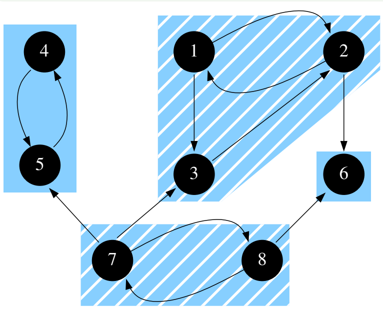
Une chaîne de Markov est irréductible $\iff \forall i \neq j \in S, \exists m \in \mathbb{N}, p_{ij}^{(m)} > 0$

Si une chaîne de Markov n'est pas irréductible, c'est que la relation $\leftrightarrow$ possède plusieurs classe d'équivalence $C_1, C_2, \ldots$.
Certaines C_i (au moins une) forme une sous-chaine de Markov, irréductible.

Définition (sous-ensemble absorbant)

Un sous-ensemble d'états $C \subset S$ est absorbant si et seulement :

- *C est une sous-chaîne de Markov.* *(: n'a pas de transition non nulle vers un état hors de C)*
- *C forme une chaîne de Markov irréductible.* *(: est une classe d'équivalence pour $\leftrightarrow$ dans S)*

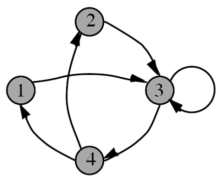


- $\{6\}$ et $\{4, 5\}$ absorbants
- $\{7, 8\}$ et $\{1, 2, 3\}$ classes d'équivalence pour $\leftrightarrow$

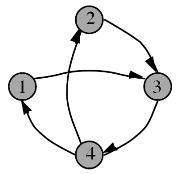
périodicité

Définition (Périodicité)

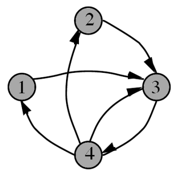
Soit $\mathcal{K} = \left\{ k \text{ tel que } \forall n \text{ non multiple de } k, P_{ii}^{(n)} = 0 \right\}$,
Un état i *est périodique si et seulement si* $1 \notin \mathcal{K}$. *La période de* i *est alors* $k_i = \max \mathcal{K}$
Si tous les états d'une chaîne de Markov sont périodiques, la chaîne de Markov est périodique, de période le PGCD des périodes de ses états (s'il est différent de 1*).*



chaîne non périodique



Chaine périodique
de période 3



Chaîne non périodique.

Récurrence et transience

Pour étudier les temps de retour à un état i, à partir des (X_n), on crée de nouvelles variables aléatoires : τ_i :

Définition (Instant de premier retour)

Pour tout état i, *avec* $X_0 = i$, $\tau_i = \begin{cases} min\{n \geq 1, X_n = i \mid X_0 = i\} \\ +\infty \text{ sinon} \end{cases}$

Définition (État récurrent, état transient)

Pour tout état i, *avec* $X_0 = i$,

- i *récurrent* $\iff P(\tau_i < \infty) = 1$
- i *transient* $\iff P(\tau_i < \infty) < 1$