

Chaîne de Markov - 2

Pierre-Henri WUILLEMIN

Licence d'Informatique – Université Paris 6

Classements des états

▶ Définition (Périodicité)

Soit $\mathcal{K}_i = \left\{ n \geq 0 \text{ tel que } P_{ii}^{(n)} > 0 \right\}$,

Un état i est périodique si et seulement si $\mathcal{K}_i \neq \emptyset$ et $\text{pgcd}(\mathcal{K}_i) \neq 1$.
La période de i est alors $k_i = \text{pgcd}(\mathcal{K}_i)$.

▶ 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 transitoire)

Pour tout état i , avec $X_0 = i$,

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

NB : transitoire=transitoire



Classement des états, suite

▶ Définition

On note et nomme :

- la probabilité que le premier retour en i soit en n étapes : $f_{ii}^{(n)} = P(\tau_i = n)$
- la probabilité de revenir en i : $f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)} = P(\tau_i < \infty)$
- le temps moyen⁽¹⁾ de retour en i : $M_i = \sum_{n=1}^{\infty} (n \cdot f_{ii}^{(n)})$

(1) quand ce calcul a un sens, $= \infty$ sinon.

▶ Définition (État récurrent, état transitoire)

Pour tout état i , avec $X_0 = i$,

- i transitoire $\iff f_{ii} < 1$
- i récurrent $\iff f_{ii} = 1$
- i récurrent nul si $M_i = \infty$
- i récurrent positif (ou non nul) si $M_i < \infty$

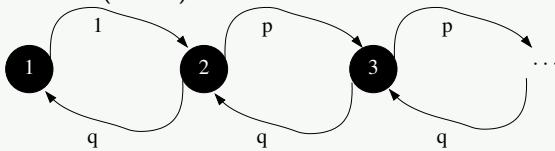
Note : Périodicité, récurrence et transience sont des propriétés de classe :
si $i \leftrightarrow j$ alors i et j sont forcément du même type.



Transience, récurrence, récurrence positive

Exemple

Soit la chaîne de Markov (infinie) suivante :



Chaîne de Markov irréductible donc tous les états de même type.

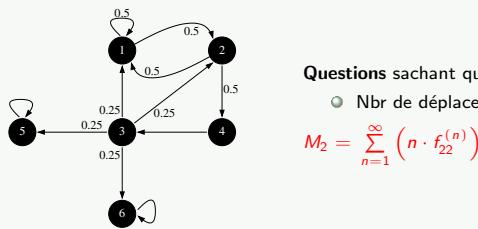
- Si $p > q$ alors $f_{ii} < 1$: état transients
- Si $p = q$ alors $f_{ii} = 1$ mais $M_i = \infty$: état récurrent nul
- Si $p < q$ alors $f_{ii} = 1$ mais $M_i < \infty$: état récurrent positif

Propriétés

Toute chaîne de Markov, à états finis, homogène a au moins un état récurrent.
En particulier, toute chaîne, homogène, à états finis, irréductible sur un espace d'états finis est récurrente positive.

Suite de la souris

souris et labyrinthe



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

- Nbr de déplacements moyens pour revenir en 2 ?

$$M_2 = \sum_{n=1}^{\infty} (n \cdot f_{22}^{(n)})$$

- $f_{22}^{(1)} = P_{22} = 0$

- $f_{22}^{(n)} = P_{21} \cdot P_{12}^{(n-1)} + P_{24} \cdot P_{42}^{(n-1)}$ (cf. Chapman-Kolmogorov)



Étude en régime permanent

Ce qui nous intéresse ici est le comportement de la chaîne de Markov si on laisse se dérouler le processus durant un temps très important.

Que peut-on dire de la position du système ? Suit-il une loi de probabilité particulière ?

En notant $\pi^{(n)}$ le vecteur de probabilité du système à l'instant n , on se rappelle que :

$$\pi^{(n+1)} = \pi^{(n)} \cdot P = \pi^{(0)} \cdot P^n$$

► Définition (distribution de probabilité invariante)

Une distribution de probabilité est **invariante** pour la chaîne de Markov si et seulement si elle s'écrit comme le vecteur π et :

$$\pi = \pi \cdot P$$

i.e. : π est un vecteur propre de P^T pour la valeur propre 1

En supposant que $(\pi^{(n)})_{n \in \mathbb{N}}$ converge vers π^* alors :

$$\pi^* = \lim_{n \rightarrow \infty} \pi^{(n)} = \pi^{(0)} \cdot \lim_{n \rightarrow \infty} P^n = \pi^{(0)} \cdot P^*$$

Propriété

$(\pi^{(n)})_{n \in \mathbb{N}}$ converge vers π^* indépendamment de $\pi^{(0)}$ si et seulement si $\lim_{n \rightarrow \infty} P^{(n)} = P^*$, matrice dont toutes les lignes sont égales entre elles (et égales à π^*).



Ergodicité

▶ Définition (Chaîne de Markov ergodique)

Une chaîne de Markov est ergodique si et seulement si elle est irréductible, apériodique et récurrente positive.

Théorème (théorème ergodique)

Une chaîne de Markov ergodique est telle que $(\pi^{(n)})_{n \in \mathbb{N}}$ converge, quelque soit $\pi^{(0)}$, vers π^* vérifiant :

$$\begin{cases} \pi^* \cdot P = \pi^* \\ \pi^* \cdot 1 = 1 \end{cases}$$

De plus,

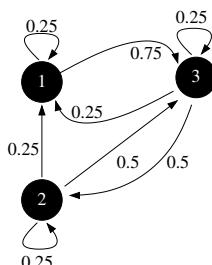
$$\pi_j^* = \frac{1}{M_j}$$

Autrement dit, la proportion des instants où la chaîne se trouve dans l'état j tend vers π_j^* avec probabilité 1.
Pour presque toutes les trajectoires, la moyenne temporelle est identique à la moyenne spatiale.



Exemple 1

$$P = \begin{pmatrix} 0.25 & 0 & 0.75 \\ 0.25 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{pmatrix}$$



	π_1	π_2	π_3
$n = 0$	1	0	0
$n = 1$	0.25	0	0.75
$n = 2$	0.25	0.375	0.375
$n = 3$	0.25	0.28125	0.46875
$n = 4$	0.25	0.30469	0.44531
$n = 5$	0.25	0.29883	0.45117
$n = 6$	0.25	0.30029	0.44971
$n = 7$	0.25	0.29993	0.45007
$n = 8$	0.25	0.30002	0.44998
$n = 9$	0.25	0.30000	0.45000
...	0.25	0.30000	0.45000

- irréductible, apériodique, à état fini \Rightarrow récurrente positive.

$$\bullet P^* = \lim_{n \rightarrow \infty} P^n = \begin{pmatrix} 0.25 & 0.3 & 0.45 \\ 0.25 & 0.3 & 0.45 \\ 0.25 & 0.3 & 0.45 \end{pmatrix}$$

$$\bullet \pi^* = [0.25, 0.3, 0.45]$$



Exemple 2

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$



	π_1	π_2
$n = 0$	1	0
$n = 1$	0	1
$n = 2$	1	0
$n = 3$	0	1
$n = 4$	1	0
...	0	1

- irréductible, périodique,

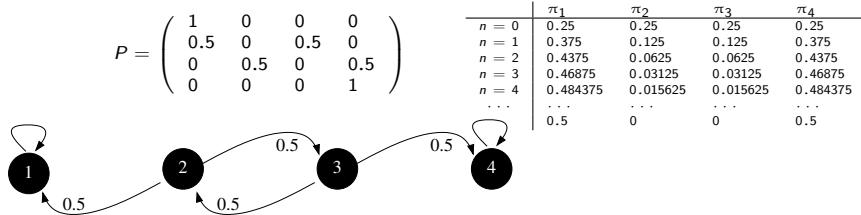
$$\bullet P^{2k} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ et } P^{2k+1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} : \text{ pas de } P^*.$$

- $\pi = [0.5, 0.5]$ est bien une distribution invariante.

- Aucune convergence vers π , sauf si $\pi^{(0)} = \pi$ (processus stationnaire)



Exemple 3



- **réductible** composantes irréductibles : {1},{23},{4}. absorbants : {1} et {4}

- $P^* = \lim_{n \rightarrow \infty} P^n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.67 & 0 & 0 & 0.33 \\ 0.33 & 0 & 0 & 0.67 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

- $\pi^* = \pi^{(0)} \cdot P^* = _1 + \frac{2}{3}\pi_2 + \frac{1}{3}\pi_3, 0, 0, \pi_4 + \frac{2}{3}\pi_3 + \frac{1}{3}\pi_2$

⚠ Dépend de $\pi^{(0)}$!



Utilisation des chaînes de Markov : introduction à MCMC

Problème : Estimer $E_P(f)$

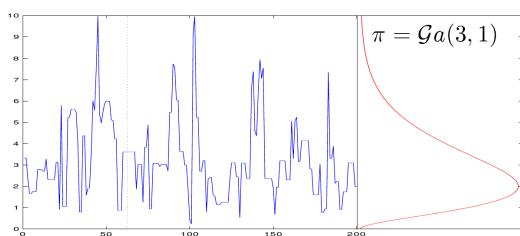
La solution théorique : $E_P(f) = \sum_x f(x) \cdot P(x)$.

Comment faire sur un espace de grande taille, difficile à énumérer ?

Les méthodes MCMC créent une longue chaîne de Markov ergodique $(X_n)_{n \in \mathbb{N}}$, dont la loi π^* est la loi P requise. On peut alors utiliser les différents X_n comme des v.a. distribués suivant P :

$$\text{Si } X_n \sim P, \hat{E}_P(f) = \frac{1}{N} \sum_n f(x_n).$$

Convergence d'après la loi des grands nombres : il faut que N soit suffisamment grand.



Historique

- Les méthodes MCMC sont apparues dans les années 50-60 pour la physique statistique. [Metropolis et al. 1953]

- 1970 : article précurseur de Hastings.

- 1984 : Échantilleur de Gibbs [Geman and Geman, 1984]

- 1990 : apparition des méthodes MCMC dans la littérature statistique et d'analyse du signal [Gelfand and Smith, 1990]



Extension des chaînes de Markov : Chaîne de Markov cachée

Les modèles de Markov cachés (HMM) sont une autre évolution possible des chaînes de Markov. Ces nouveaux modèles se basent cette fois sur deux processus stochastiques dépendants l'un de l'autre.
L'état du système n'est plus directement observable ; il est caché par un processus d'observation.

Un ou deux dés ?

Un joueur peut lancer un dé (avec un résultat de 1 à 6) ou deux dés (avec un résultat de 2 à 12). Sachant le lancer choisi, il est donc facile de prédire le résultat attendu. Si, maintenant, le lanceur se trouve derrière un rideau, et annonce simplement le résultat, comment peut-on déterminer s'il a lancé 1 ou 2 dés ? À partir d'une séquence d'observations, quelle est la séquence d'états correspondante ?

modélisation

Une chaîne de Markov cachée est définie par :

- S : ensemble d'états
- $P(X_t | X_{t-1})$: la matrice de transition pour l'état
- $P(O | X)$: la distribution de probabilité de l'observation, sachant l'état.

