Chaines de Markov

Clément Pellegrini

clement.pellegrini@math.univ-toulouse.fr

Institut de Mathématiques de Toulouse,
Equipe de Statistique et Probabilité,
Bureau 220 Bâtiment 1R1

Définitions

Dans tout ce qui suit on travaille sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Dans tout ce qui suit E désignera un ensemble fini ou dénombrable.

Definition

Soit (X_n) une suite de v.a à valeurs dans un espace E. On dit que (X_n) est une chaine de Markov CM si pour tout $n \in \mathbb{N}$ pour tout $(i_0, i_1, \ldots, i_{n-1}, i, j) \in E^{n+1}$ on a

$$\mathbb{P}[X_{n+1} = j | X_n = i, \ldots, X_0 = i_0] = \mathbb{P}[X_{n+1} = j | X_n = i].$$

Si de plus $\mathbb{P}[X_{n+1} = j | X_n = i] = \mathbb{P}[X_1 = j | X_0 = i]$ pour tout n alors on dit que la chaine de Markov est homogène CMH.

On peut voir qu'une CMH (X_n) peut toujours s'écrire sous la forme

$$X_{n+1} = f(X_n, \epsilon_n), n \in \mathbb{N}$$

avec (ϵ_n) une suite de v.a telle que ϵ_n et X_n sont indépendantes. C'est un moyen pratique de montrer qu'une suite (X_n) est une CM.

Definition

Soit (X_n) une CMH, la matrice P définie par

$$P(i,j) = \mathbb{P}[X_1 = j | X_0 = i], i, j \in E$$

s'appelle la matrice de transition de (X_n)

Définitions

Proposition

La loi d'une CMH (X_n) *est entièrement déterminée par sa matrice de transition et sa loi initiale* $\mu_0(i) = \mathbb{P}[X_0 = i], i \in E$ *on a*

$$\mathbb{P}[X_n = i_n, \ldots, X_0 = i_0] = P(i_{n-1}, i_n) \ldots P(i_0, i_1)\mu_0(i_0)$$

pour tout $i_n, \ldots, i_0$ *élément de E.*

Proposition

Chapman Kolmogorov: pour tout $i, j \in E$ *et pour tout* n, m

$$\mathbb{P}[X_{n+m} = j | X_0 = i_0] = \sum_{k \in E} \mathbb{P}[X_m = j | X_0 = k]\mathbb{P}[X_n = k | X_0 = i_0]$$

En particulier $\mathbb{P}[X_n = j | X_0 = i] = P^n(i, j)$ *et*

$$X_n \sim \mu P^n, n \in \mathbb{N}$$

Définitions

Definition

Soit (X_n) une CMH, on pose

$$\mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n).$$

Un temps d'arrêt τ est une v.a $\Omega \to \mathbb{N} \cup \{+\infty\}$ tel que pour tout $n \in \mathbb{N}$

$$\{\tau = n\} \in \mathcal{F}_n$$

ou de manière équivalente

$$\{\tau \leqslant n\} \in \mathcal{F}_n$$

Définitions

Théorème

MARKOV FAIBLE. Soit (X_n) une CMH et soit $m \in \mathbb{N}$. Conditionnellement à l'évenement $\{X_m = i\}$ la suite de v.a $(X_n + m)$ est une CMH de matrice de transition P et de loi initiale δ_i et indépendante de $(X_0, \ldots, X_m)$

$$\mathbb{P}[X_{n+m} = i_n, \ldots, X_{m+1} = i_1, X_m = i_0 \cap A | X_m = i] = P(i_{n-1}, i_n) \ldots P(i_0, i_1)\delta_{i,i_0}\,\mathbb{P}[A|X_m = i]$$

pour tout $A \in \mathcal{F}_n$.

Définitions

Théorème

MARKOV FORT. Soit (X_n) une CMH et soit τ un temps d'arrêt. Conditionnellement à l'évenement $\{X_\tau = i\}$ et $\{\tau < \infty\}$, la suite de v.a $(X_n + \tau)$ est une CMH de matrice de transition P et de loi initiale δ_i et indépendante de $(X_0, \ldots, X_\tau)$

Ici $\mathcal{F}_\tau$ va désigner la tribu définie par

$$A \in \mathcal{F}_\tau <=> \forall n \in \mathbb{N}, \quad A \cap \{\tau = n\} \in \mathcal{F}_n$$

$$\mathbb{P}[X_{n+\tau} = i_n, \ldots, X_{m+\tau} = i_1, X_\tau = i_0 \cap A | X_\tau = i, \tau < +\infty] = \\ P(i_{n-1}, i_n) \ldots P(i_0, i_1) \delta_{i,i_0} \mathbb{P}[A | X_m = i, \tau < +\infty]$$

pour tout $A \in \mathcal{F}_\tau$.

Définitions

Definition

Soient i et j deux éléments de E on dit que j est accessible à partir de i si il existe $n \in \mathbb{N}$ tel que

$$P^n(i,j) > 0$$

on écrit $i \to j$. On dit que i et j communiquent et on note $i \leftrightarrow$ si $i \to j$ et $j \to i$.

En particulier $i \to j$ ssi il existe $i_0, \ldots, i_n$ avec $i_0 = i$ et $i_n = j$ tel que

$$\mathbb{P}[X_n = j, \ldots, X_1 = i_1 | X_0 = j] = P(i_{n-1}, j) \ldots P(j, i_1) > 0$$

En d'autres termes il existe un chemin de probabilité strictement positive qui mêne de i à j.

Définitions

La relation $\leftrightarrow$ est une relation d'équivalence et donc partitionne E en classe

Definition

On dit qu'une CMH est iréductible si elle n'a qu'une classe d'équivalence

- Pour déterminer les classes d'équivalence on travaille sur le graphe de la CM.
- Marche aléatoire sur $\mathbb{Z}$

Définitions

Definition

Un état i est dit récurrent si partant de i la CM revient $\mathbb{P}_i$ p.s en i. Un état non récurrent est dit transient. On pose

$$\tau_i = \inf\{n \geqslant 1 | X_n = i\}$$

alors i est récurrent si $\mathbb{P}_i[\tau_i < \infty] = \mathbb{P}[\tau_i < \infty | X_0 = i] = 1$ et i est transient si $\mathbb{P}_i[\tau_i < \infty] = \mathbb{P}[\tau_i < \infty | X_0 = i] < 1$

Définitions

Definition

On définit le nombre de visite du site i par

$$N_i = \sum_{n \in \mathbb{N}} \mathbf{1}_{X_n = i}$$

On a par exemple

$$\mathbb{E}_j[N_i] = \sum P^n(j, i)$$

On a le résultat suivant

Proposition

$$\mathbb{P}_j(N_i \geqslant n + 1) = \mathbb{P}_j(\tau_i < +\infty)\mathbb{P}_i(N_i \geqslant n)$$

Définitions

Proposition

Les conditions suivantes sont équivalentes

1. *i est récurrent*
2. $\mathbb{P}_i(N_i = +\infty) = 1$
3. $\sum_n P^n(i,i) = \infty$

Les conditions suivantes sont équivalentes

1. *i est transient*
2. $\mathbb{P}_i(N_i = +\infty) = 0$ *et la v.a* N_i *suit une loi géométrique de paramètres* $\mathbb{P}_i(\tau_i < \infty)$
3. $\sum_n P^n(i,i) < \infty$

$$\mathbb{E}_i(N_i) = \sum_n P^n(i,i)$$

Définitions

Proposition

- *La récurrence et la transience sont des propriétés de classe*
- *Tous les états d'une même classe récurrente sont visités $\mathbb{P}_j$ presque sûrement pour tout état j de la classe*

$$\mathbb{P}_j(\tau_i < \infty) = \mathbb{P}_i(N_i = \infty) = 1$$

pour tout i et j dans la classe de récurrence.

On ne sort pas d'une classe de récurrence

Définitions

Proposition

Si E est fini alors la CM admet au moins une classe de récurrence. En particulier toute CM irréductible sur un espace d'états finis est récurrente

$$\begin{aligned}\mathbb{E}_j[N_i] &= \sum_n \mathbb{P}_j(N_i \geqslant n+1) = \sum_n \mathbb{P}_j(\tau_i < \infty)\mathbb{P}_i(N_i \geqslant n) \\ &= P_j(\tau_i < \infty)(1 + \mathbb{E}_i[N_i])\end{aligned}$$

- Si i est transient $\mathbb{E}_j[N_i] < \infty$
- $\infty = \mathbb{E}_j[\sum_i N_i] = \sum_i \mathbb{E}_j[N_i]$ si E est de cardinal fini et tous les états j transient la dernière quantité devrait être finie.

Définitions

Concernant le comportement en temps long quelles sont les questions que l'on peut se pose

- Convergence en loi

$$\mathbb{P}_i[X_n = j] = P^n(i,j) \underset{n}{\rightarrow} \mu(i,j)$$

- Moyenne empirique

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k=j} \underset{n}{\rightarrow} \lambda(i,j), \quad \mathbb{P}_i \ p.s$$

- On peut voir que si ces deux limites existent alors elles sont égales

Définitions

- Si j est transient alors pour tout i

$$\mathbb{E}_i[N_j] < \infty$$

- Or on a

$$\mathbb{E}_i[N_j] = \sum_n P^n(i,j)$$

- Donc

$$P^n(i,j) \to 0$$

- De plus $N_j < \infty$ est fini $\mathbb{P}_i$ p.s d'où

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k=j} \underset{n}{\to} 0 \quad \mathbb{P}_i \ p.s$$

- CCL: on peut se restreindre au classes récurrentes et donc supposer la CM irréductible (car on ne sort pas d'une classe de récurrence)

Définitions

Definition

On dit qu'une mesure π est invariante pour la matrice stochastique P si

$$\pi P = \pi$$

Proposition

Si π est est un mesure de probabilité invariante pour la matrice stochastique P et si $X_0 \sim \pi$ alors

$$X_n \sim \pi, \forall n \in \mathbb{N}$$

En effet rappelons que $X_n \sim \pi P^n$ et si $\pi P = \pi$ alors $X_n \sim \pi$.
Rq: si $P^n(i,j) \to L_i(j)$ alors la mesure de probabilité $L_i(.)$ est invariante.

Définitions

Theorem

Toute CMH irréductible admet une MESURE invariante STRICTEMENT POSITIVE sur E et toutes les MESURES invariantes sont proportionnelles entre elles

- Cette mesure peut être de masse infinie si E est infini dénombrable
- Si E est fini, une CMH irréductible admet une unique mesure de probabilité invariante

Définitions

Definition

Soit i un état récurrent et τ_i le temps de premier retour (donc fini p.s)

- i est dit récurrent positif si $\mathbb{E}_i[\tau_i] < \infty$
- i est dit récurrent nul si $\mathbb{E}_i[\tau_i] = \infty$

Théorème

Soit X une CMH irréd alors les propositions suivantes sont équivalentes

- *il existe un état i récurrent positif*
- *tous les états sont récurrents positifs*
- *la CM admet une unique probabilité invariante donnée par*

$$\pi(i) = \frac{1}{\mathbb{E}_i[\tau_i]}$$

Définitions

Théorème

Une CMH irréductible à espace d'état fini admet une unique probabilité invariante et donc tous ses états sont récurrents positifs

- La connaissance de la mesure invariante permet de calculer $\mathbb{E}_i[\tau_i]$
- Si E est fini on résoud $\pi P = \pi$ et on prend un vecteur à coordonnées positives et on le normalise. Il y en existe forcément un. Quelle est la dimension de l'espace propre associé à 1 pour P^t
- Une mesure λ est dite P réversible si

$$\lambda(i)P(i,j) = \lambda(j)P(j,i)$$

  une telle mesure est nécessairement invariante

Définitions

Definition

On appelle période d'un état

$$d(i) = pgcd\{n \geqslant 1 | P^n(i,i) > 0\}$$

Proposition

Tous les états d'une même classe ont même période

Definition

On dit qu'une classe est apériodique si l'un des états de cette classe à une période 1

Définitions

Proposition

Soit i un état de période 1 alors

- *il existe $N(i)$ telle que pour tout $n \geqslant N(i)$: $P^n(i,i) > 0$*
- *pour tout $j \leftrightarrow i$ il existe $N(i,j)$ telle que pour tout $n \geqslant N(i,j)$: $P^n(i,j) > 0$*

- Pour E fini une CMH irréductible est apériodique si et seulement si il existe n tel que P^n n'a que des coefficients strictement positifs.

Définitions

Théorème

Soit X une CMH irréductible apériodique pour laquelle il existe une probabilité invariante π alors

$$P^n(i,j) \to \pi(j)$$

pour tout état i. Ainsi pour toute mesure initiale λ

$$\mathbb{P}_\lambda(X_n = j) \to \pi(j)$$

- Si X est une CMH irréductible apériodique et récurrente positif alors on a CONVERGENCE EN LOI
- Si E est fini et si X est une CMH irréductible et apériodique alors il ya forcément convergence en loi

Définitions

Théorème

Soit X une CMH irréductible alors pour toute mesure initiale λ et pour tout état j

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k=j} \to \frac{1}{\mathbb{E}_j[\tau_j]} \quad \mathbb{P}_\lambda \ p.s$$

- Si E est fini et si X est une CMH irréductible alors

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{X_k=j} \to \pi(j) \quad \mathbb{P}_\lambda \ p.s$$

  où π est l'unique probabilité invariante
- C'est ce résultat que l'on utilise dans les simulations pour approcher la mesure invariante.

Définitions

Théorème

Soit X une CMH irréductible et récurrente positive alors pour toute mesure initiale λ et pour toute fonction mesurable bornée on a

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \to \int f d\pi \quad \mathbb{P}_\lambda \ p.s$$

- Méthode de Monte Carlo

On considère E fini

Proposition

Soit P une matrice stochastique alors P admet 1 comme valeur propre et toutes ses valeurs propres sont de module plsu petit ou égale à 1

- Si P est diagonalisable et que 1 est sa valeur propre de module 1 et qu'elle est simple alors il est facile de voir que

$$P^n(i,j) \to \pi_j$$

Définitions

Proposition

Toute CMH à espace d'états finis admet une mesure de probabilité invariante

- Soit μ_0 une mesure initiale, on considère

$$f_n(j) = \frac{1}{n}\sum_{k=1}^{n} \mu_0 P^k$$

C'est une suite à valeurs dans le complexe des vecteurs de probabilité qui est compact donc admet uen sous-suite convergente

Définitions

Proposition

Soit P une matrice stochastique et soit λ une valeur propre de module 1 *de P^t et soit v un vecteur propre associé. Alors $|v|^t = (|v_1|, \ldots, |v_d|)$ est un vecteur propre associés à la valeur propre* 1 *pour P^t.*

Proposition

Soit P une matrice stochastique associé à une CMH irréductible et apériodique alors P n'a que 1 *comme v.p de module* 1 *et l'espace propre associé est de dimension 1*

Définitions

- Pour générer une CM à espace d'états finis on peut utiliser la fonction

$$grand(N,' markov', P, x0)$$

  Cela génère une trajectoire de taille N et de point de départ $x0$
- Sinon on peut générer une v.a discrète à chaque étape. Si à l'étape n on est dans létat i on calcule $sum(cumsum(P(i,:)) < rand()) + 1$ permet de détrminer l'état suivant.
- Quand E est infini on adapte souvent la méthode précédente