Chapitre 4 : Chaînes de Markov

Introduction

Un modèle dynamique pour lequel le futur dépend de l'état actuel et du hasard est appelé chaîne de Markov : modèle simple pour représenter quelque chose qui évolue au cours du temps.

Situations où un tel modèle est utilisé :

- finance (cours de la bourse)
- théorie du signal (filtrage)
- informatique (files d'attente dans les réseaux)

etc

Notre cadre : chaîne de Markov en temps discret et espace d'états fini

I Définition

Soit E un espace fini, typiquement $E = \{1, \dots, N\}$

E : espace d'état

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires à valeurs dans E.

définition

$(X_n)_{n \geq 0}$ est une chaîne de Markov si :

$\forall n \geq 1$, $\forall (i_0, i_1, \dots, i_{n-1}, i, j)$ de E, on a

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

Rq : cela suppose que $P(X_0 = i_0, \dots, X_n = i) > 0$

Interprétation : sachant le présent, le futur ne dépend pas du passé.

définition :

Une chaîne de Markov est dite homogène dans le temps si

$$P(X_{n+1}=j \mid X_n=i) = P(X_n=j \mid X_{n-1}=i) = \cdots = P(X_1=j \mid X_0=i)$$

$$= p_{ij}$$

probabilité de transition de i vers j

Matrice de transition $P = (p_{ij})_{1 \leq i,j \leq M}$

Rq : La connaissance de la loi initiale, à savoir $(P(X_0=i))_{i\in[1,M]}$

et des probabilités de transition permet d'écrire la loi jointe de $(X_0, \dots, X_n)$ si on a une chaîne homogène

$$P(X_0=i_0, \dots, X_n=i_n) = P(X_n=i_n \mid X_0=i_0, \dots, X_{n-1}=i_{n-1})$$

$$\times P(X_0=i_0, \dots, X_{n-1}=i_{n-1})$$

$$= P(X_n=i_n \mid X_{n-1}=i_{n-1}) \times P(X_0=i_0, \dots, X_{n-1}=i_{n-1})$$

$$= P(X_n=i_n \mid X_{n-1}=i_{n-1}) \times P(X_{n-1}=i_{n-1} \mid X_{n-2}=i_{n-2}) \cdots$$

$$\times P(X_1=i_1 \mid X_0=i_0)\, P(X_0=i_0)$$

$$= P(X_0=i_0)\, p_{i_0 i_1} \times p_{i_1 i_2} \times \cdots \times p_{i_{n-1} i_n}$$

Propriété

Toute matrice de transition $P = (p_{ij})$ vérifie :

- $\forall (i,j) \quad p_{ij} \in [0,1]$

- $\sum_{j=1}^{M} p_{ij} = 1$

. P admet la valeur propre 1 et un vecteur propre associé est $\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$

preuve :

. $$\sum_{j=1}^{N} p_{ij} = \sum_{j=1}^{N} \frac{P(X_{n+1}=j, X_n=i)}{P(X_n=i)} = \frac{1}{P(X_n=i)} \sum_{j} P(X_{n+1}=j, X_n=i)$$

$$= \frac{P(X_n=i)}{P(X_n=i)}$$

. $$P \times \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix} = \begin{pmatrix}\sum_j p_{1j} \\ \sum_j p_{2j} \\ \vdots \\ \sum_j p_{Nj}\end{pmatrix} = \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$$

Rq : Une matrice qui vérifie les deux premières propriétés est appelée matrice stochastique. Une telle matrice vérifie nécessairement la 3^e propriété.

Rq : La dimension du sous-espace propre associé à la valeur propre 1 n'est pas forcément 1

exemple. $P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ matrice stochastique.

donc 1 valeur propre

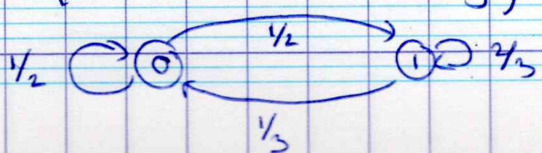
$\begin{pmatrix}1\\1\\1\\1\end{pmatrix}$ vecteur propre associé

mais $\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$ aussi !

Sur cet exemple, si on connait l'état initial, on connait l'état de tous les instants d'après !

A toute chaîne de Markov homogène peut être associé un graphe de transition :

- les sommets du graphe sont les états $1, \dots, N$
- il existe un arc entre deux états, étiqueté p_{ij}, si $p_{ij} > 0$ (arc allant de i vers j)



1) Equations de Chapman - Kolmogorov

Notation :

- $P_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$ probabilité d'aller de i à j en n coups

- $P^{(n)} = \left(P_{ij}^{(n)} \right)$: matrice de transition en n coups

- $P^{(0)} = I_N$

proposition :

$\forall n \geq 0, \quad P^{(n)} = P^n$

preuve :

Raisonnement par récurrence sur n.

- $n = 0$ ok
- supposons qu'il existe un $n \geq 0$ tq $P^{(n)} = P^n$

montrons que $P^{(n+1)} = P^{n+1}$

soit $1 \leq i,j \leq N$: $P_{ij}^{(n+1)} = P(X_{n+1} = j \mid X_0 = i)$

$$= \sum_{k=1}^{N} P(X_{n+1} = j, X_n = k \mid X_0 = i)$$

$$= \sum_{k=1}^{N} P(X_{n+1} = j \mid X_n = k, X_0 = i) P(X_n = k \mid X_0 = i)$$

$$= \sum_{k=1}^{N} P(X_{n+1} = j \mid X_n = k) P(X_n = k \mid X_0 = i)$$

$$= \sum_{k=1}^{N} P_{kj} \, P_{ik}^{(n)}$$

$$\Rightarrow P^{(n+1)} = P^{(n)} . P$$

$$= P^n . P$$

$$= P^{n+1}$$

Rq : on a aussi $P^{(m+n)} = P^{(m)} . P^{(n)}$ équation de Chapman

Rq : la position initiale peut être aléatoire.
on représente la loi initiale de X_0 par un vecteur ligne μ
$$\mu = (\mu_1, \dots, \mu_N) \quad \text{avec} \quad \mu_i = P(X_0 = i).$$

On note $P(X_n) = \left(P(X_n = 1), \dots, P(X_n = N) \right)$ la loi de X_n.

Corollaire :
Soit $(X_n)_{n \geq 0}$ une chaîne de Markov de loi initiale μ et de matrice de transition P.
$\forall n \in \mathbb{N}, \quad P(X_n) = \mu P^n$

preuve : Soit $j \in \{1, \dots, N\}$

$$P(X_n = j) = \sum_i P(X_n = j, X_0 = i) = \sum_i P(X_n = j \mid X_0 = i) P(X_0 = i)$$
$$= \sum_i \mu_i P_{ij}^{(n)}$$

Rq : La convergence en loi de X_n est équivalente à la convergence de chacune des composantes de $P(X_n)$.
Puisque $P(X_n) = \mu P^n$ il suffit de la convergence de $(P^n)_{n \geq 0}$.

Ex : graphe précédent

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}$$

Pour calculer P^n, une idée est de diagonaliser P.
On connait une valeur propre 1 et la somme des valeurs propres au n^b de 2 maximum vaut $\frac{1}{2} + \frac{2}{3} = \frac{7}{6}$.

Donc une deuxième valeur propre est $1/6$

Un vecteur propre associé à $1/6$ est $\begin{pmatrix} -1/2 \\ 1/3 \end{pmatrix}$

$\Rightarrow \quad P = Q D Q^{-1}$ avec $Q = \begin{pmatrix} 1 & -1/2 \\ 1 & 1/3 \end{pmatrix}$ $D = \begin{pmatrix} 1 & 0 \\ 0 & 1/6 \end{pmatrix}$

$$Q^{-1} = \frac{6}{5} \begin{pmatrix} 1/3 & 1/2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2/5 & 3/5 \\ -6/5 & 6/5 \end{pmatrix}$$

$$D^n \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

donc $P^n \longrightarrow \begin{pmatrix} 2/5 & 3/5 \\ 2/5 & 3/5 \end{pmatrix}$

d'où $P(X_n) = \mu P^n \longrightarrow \left(2/5 \, , \, 3/5 \right)$

Rq: si état d'occupation d'une ligne téléphonique.
2 états libre (état 0) ou occupé (état 1)
Entre 2 états successifs, une probabilité $1/2$ pour qu'un appel arrive.
Si un appel arrive quand ligne occupée, il est perdu.
La probabilité qu'une ligne se libère entre 2 instants est $1/3$

Au bout d'un certain temps, 2 chances sur 5 que la ligne soit libre.

/!\ Même si P^n converge, la loi limite de X_n n'est pas nécessairement indépendante de la loi initiale.

définition :
On appelle état absorbant un état i tel que ~~si $X_n = i$ alors $\forall m \geq n$,~~
~~$X_m = i$~~ $p_{ii} = 1$.

III Classification des états

définition :

On dit que i communique avec j ($i \to j$) s'il existe une suite d'indices $i_0 = i, i_1, \dots, i_n = j$ telle que

$$\forall k \in \{0, \dots, n-1\} \quad P_{i_k i_{k+1}} > 0$$

On peut donc aller de i à j en un certain nombre de coups.

Rq : $i \to j$ si et seulement si $\exists n > 0$ tq $P_{ij}^{(n)} > 0$

$i \to j$ et $j \to k$ alors $i \to k$ (transitivité)

Supposons que la chaîne soit initialement en l'état i ($X_0 = i$).

On pose : $T_i = \inf \{ n \geq 1 \text{ tq } X_n = i \}$

instant de premier retour en i

T_i est une variable aléatoire à valeur dans $\mathbb{N}^* \cup \{+\infty\}$

On note $r_i = P(T_i < \infty \mid X_0 = i)$

définition :

- si $r_i < 1$, alors l'état i est transitoire (état visité un n^b fini de fois)
- si $r_i = 1$, alors l'état i est récurrent (état visité un nombre infini de fois)

p^b : Bien souvent, on ne sait pas calculer r_i

proposition :

Soit $i \in \{1, \dots, N\}$, s'il existe j tel que $i \to j$ et $j \not\to i$ alors i est transitoire.

preuve :

Posons $r_{ij} = P(T_j < \infty \mid X_0 = i)$ probabilité de venir en j étant parti de i

$r_{ij} > 0$ et $r_i \leq 1 - r_{ij} < 1$

définition

- Un sous-ensemble A de E est clos ou fermé s'il est impossible d'en sortir.

$$\forall j \in A, \ \forall k \in \bar{A} \quad p_{jk} = 0$$

- Un sous-ensemble B de E est irréductible si tous ses états communiquent

$$\forall i,j \in B \quad i \rightarrow j$$

proposition :

Si un ensemble est fermé et irréductible, alors tous ses états sont récurrents.

Théorème :

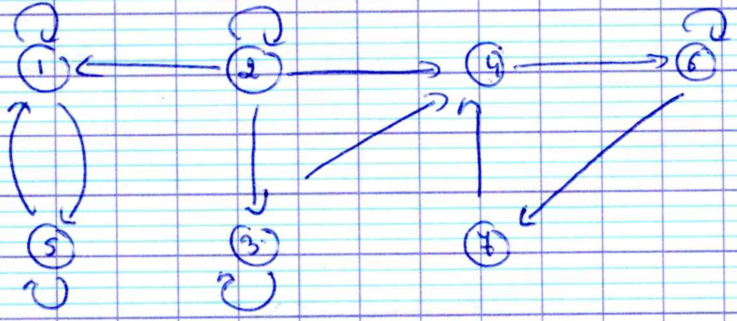
Soit E l'espace d'état, alors :

$E = T \cup R_1 \cup \dots \cup R_k$ avec T : l'ensemble des états transients

R_i : des classes de récurrence, à savoir des ensembles fermés irréductibles.

~~définition~~ :

Exemple :



$2 \rightarrow 1$ mais $1 \not\rightarrow 2$ donc 2 est transitoire

$3 \rightarrow 4$ mais $4 \not\rightarrow 3$ donc 3 est transitoire

$\{1, 5\}$: fermé et irréductible

$\{4, 6, 7\}$: fermé et irréductible

$\{1, \dots, 7\} = \{2; 3\} \cup \{1, 5\} \cup \{4, 6; 7\}$

définition :

Si E n'est formé que d'une classe de récurrence, alors la chaîne est dite irréductible.

définition :

Si E ne se compose que d'états transitoires et d'une classe de récurrence (et une seule), alors la chaîne est dite indécomposable.

IV Comportement asymptotique

objectif : distinguer les différentes situations relatives à la convergence en loi d'une chaîne de Markov

définition

La période d_i de l'état i est le plus grand entier naturel divisant tout entier n tel que $P^n(i,i) > 0$.

Autrement dit, si l'on note $Z_i = \{n \geq 1 \text{ tq } P^n(i,i) > 0\}$

$$d_i = \text{pgcd}(Z_i)$$

Si $d_i = 1$, on dit que l'état i est apériodique.

Si tous les états sont apériodiques, on dit que la chaîne est apériodique.

propriétés :

- si $p_{ii} > 0$, alors i est apériodique
- ~~si i est apériodique~~,
- si i et j communiquent, alors i et j ont la même période.
- si la chaîne est irréductible, tous les états ont même période.

définition :

Un vecteur de probabilité ligne π est une loi stationnaire (ou invariante ou d'équilibre) d'une chaîne de Markov de matrice de transition P si

$$\pi P = \pi$$

Lemme : Soit π une loi invariante.

Si $X_0 \sim \pi$, alors $\forall n \geq 0$, $X_n \sim \pi$

propriété :

Pour toute matrice de transition P, il existe au moins une loi ~~inv~~ invariante π.

proposition :

Si π est une loi invariant d'une chaîne de Markov, pour tout état transitoire i, on a $\pi_i = 0$.

proposition :

Soit une chaîne de Markov irréductible. Alors la loi stationnaire est unique et pour tout i, $\pi_i > 0$.

Théorème :

Si la chaîne de Markov est irréductible et apériodique, de loi invariante π, alors :

$$\forall i,j \in \{1,..,N\} \quad p_{ij}^{(n)} \longrightarrow \pi_j$$

En particulier, pour toute loi initiale μ, la loi de X_n converge vers π

$$\mu P^n \longrightarrow \pi$$

Théorème :

Soit (X_n) une chaîne de Markov irréductible, de loi stationnaire π.

Soit f $E \to \mathbb{R}$ une f^{ct}, alors :

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow[p.s]{} \sum_{i=1}^{N} f_i \pi_i$$

Corollaire :

Soit (X_n) une chaîne de Markov irréductible de loi stationnaire π.

Le temps relatif passé par une trajectoire de la chaîne dans ~~l~~ l'état i converge presque sûrement vers π

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{X_k = i} \xrightarrow[p.s]{} \pi_i$$