À propos des chaînes de Markov discrètes

X MAP 432 PC 17

Automne 2013

Soit $X = (X_n)_{n\geq 0}$ une suite de v.a. sur E au plus dénombrable. Les trois propriétés suivantes sont **équivalentes**, et on dit alors que X est une **chaîne de Markov**.

1. Il existe une **suite de noyaux ou matrices de transition** [1] $(P_n)_{n\geq 0}$ sur E telle que pour tout $n \geq 0$ et tous $x_0, \ldots, x_{n+1}$ dans E,

$$\mathbb{P}(X_0 = x_0, \ldots, X_{n+1} = x_{n+1}) = \mathbb{P}(X_0 = x_0)P_0(x_0, x_1) \cdots P_n(x_n, x_{n+1}),$$

   Note : la suite aléatoire X est une v.a. à valeurs dans $E^{\mathbb{N}}$ muni de la tribu engendrée par les cylindres, et sa loi est entièrement caractérisée par ses lois marginales de dimension finie, par exemple par la donnée pour tout $n \geq 0$ et tous $x_0, \ldots, x_{n+1} \in E$ de $\mathbb{P}(X_0 = x_0, \ldots, X_{n+1} = x_{n+1})$. Ainsi, lorsque X est une chaîne de Markov, la loi de X dans $E^{\mathbb{N}}$ a une **structure de dépendance spéciale, paramétrée par la loi initiale et les noyaux de transition** (tous égaux ssi la chaîne est homogène).

2. **Conditionnellement au présent fixé et au passé, le futur ne dépend que du présent** : pour tout $n \geq 1$ (instant présent), tout $x \in E$ (valeur du présent fixée), tout $B \in \sigma(X_0, \ldots, X_{n-1})$ (passé), tout $A \in \sigma(X_{n+1}, X_{n+2}, \ldots)$ (futur),

$$\mathbb{P}(A \,|\, X_n = x, B) = \mathbb{P}(A \,|\, X_n = x).$$

   Exemple : $B = \{X_0 = x_0, \ldots, X_{n-1} = x_{n-1}\}$ et $A = \{X_{n+1} = x_{n+1}\}$, qui donne

$$\mathbb{P}(X_{n+1} = x_{n+1} \,|\, X_n = x, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} \,|\, X_n = x).$$

   Cette **propriété d'absence de mémoire d'ordre** 1 correspond à l'ordre 1 de la récurrence aléatoire ci-dessous, ainsi qu'à l'ordre 1 de l'épinglage ci-dessous.

3. **Le passé et le futur sont indépendants conditionnellement au présent fixé** : pour tout $n \geq 1$ (instant présent), tout $x \in E$ (valeur du présent fixée), tout $B \in \sigma(X_0, \ldots, X_{n-1})$ (passé), tout $A \in \sigma(X_{n+1}, X_{n+2}, \ldots)$ (futur), on a

$$\mathbb{P}(B \cap A \,|\, X_n = x) = \mathbb{P}(B \,|\, X_n = x)\mathbb{P}(A \,|\, X_n = x);$$

   Exemple : $B = \{X_{n-1} = y\}$ et $A = \{X_{n+1} = z\}$, ce qui donne

$$\mathbb{P}(X_{n-1} = y, X_{n+1} = z \,|\, X_n = x) = \mathbb{P}(X_{n-1} = y \,|\, X_n = x)\mathbb{P}(X_{n+1} = z \,|\, X_n = x).$$

   On peut imaginer le conditionnement par X_n comme un épinglage de la trajectoire X. Les deux bouts restants (passé et futur) fluctuent de manière indépendante !

1. C'est-à-dire que pour tout $n \geq 0$, $P_n : E \times E \to [0, 1]$ et $\sum_{y\in E} P_n(x, y) = 1$ pour tout $x \in E$. Il faut concevoir P_n comme un tableau où chaque ligne est une loi sur E. Il y a une ligne par état, et cette ligne est la loi sur E qui décrit comment fabriquer aléatoirement l'état suivant de la chaîne.

$$\mathrm{Loi}(X_{n+1} \,|\, X_n = x) = P_n(x, \cdot) = \sum_{y\in E} P_n(x, y)\delta_y.$$

Suites récurrentes aléatoires

Soit X_0 une v.a. à valeur dans E, $(U_n)_{n\geq 1}$ des v.a. i.i.d. à valeurs dans un espace mesuré F, indépendantes de X_0, et $(f_n)_{n\geq 1}$ des applications mesurables de $E \times F$ dans E. Alors la **suite récurrente aléatoire du premier ordre** $X = (X_n)_{n\geq 0}$ définie par

$$X_{n+1} = f_n(X_n, U_{n+1})$$

est une chaîne de Markov, dont les noyaux de transition sont donnés par

$$P_n(x,y) = \mathbb{P}(X_{n+1} = y \,|\, X_n = x) = \mathbb{P}(f_n(x, U_1) = y).$$

La v.a. X_{n+1} est $\sigma(X_0, \ldots, X_n, U_{n+1}) = \sigma(X_0, U_1, \ldots, U_{n+1})$ mesurable.

Réciproquement, si $X = (X_n)_{n\geq 0}$ est une chaîne de Markov sur E de noyaux de transition $(P_n)_{n\geq 0}$, alors X a la même loi que la suite récurrente aléatoire du premier ordre $Y = (Y_n)_{n\geq 0}$ définie par $Y_{n+1} = f_n(Y_n, U_{n+1})$ pour tout $n \geq 0$, où

- Y_0 a la même loi que X_0 ;
- $(U_n)_{n\geq 1}$ i.i.d. de loi uniforme sur $F = [0,1]$, indépendantes de Y_0 ;
- pour tout $x \in E$ et $n \geq 0$, on a, pour tout $z \in F = [0,1]$, $f_n(x,z) = y$ si $z \in I_{x,y}$, où $[0,1] = \cup_{y\in E} I_{x,y}$ est une partition de l'intervalle $[0,1]$ vérifiant $|I_{x,y}| = P_n(x,y)$.

Cette formulation récursive permet souvent de reconnaître les chaînes de Markov dans la nature, riche en systèmes dynamiques. Cela fournit également un algorithme pour simuler les chaînes de Markov à partir de leur loi initiale et de leurs noyaux de transition. Cela permet enfin de mettre au point des couplages, qui utilisent la même suite $(U_n)_{n\geq 1}$.

Équations d'évolution linéaires

Les noyaux de transition, vus comme des matrices, agissent par multiplication à gauche sur les mesures sur E, vues comme des vecteurs lignes, et agissent par multiplication à droite sur fonctions de E dans $\mathbb{R}$, vues comme des vecteurs colonne. En d'autres termes :

$$\mu P f = \sum_{x,y\in E} \mu(x) P(x,y) f(y).$$

Soit $X = (X_n)_{n\geq 0}$ une chaîne de Markov sur E de noyaux de transition $(P_n)_{n\geq 0}$. Notons μ la loi de X_0. Alors pour tout $n \geq 0$ et toute fonction positive ou bornée $f : E \to \mathbb{R}$,

$$\mathbb{E}(f(X_n)) = \sum_{x_0,\ldots,x_n\in E} \mathbb{P}(X_0 = x_0) P_0(x_0,x_1) \cdots P_n(x_{n-1},x_n) f(x_n) = \mu P_0 \cdots P_{n-1} f,$$

où le produit est celui des matrices. Supposons que X est homogène de noyau P. Alors X_n a pour loi $\mu P_0 \cdots P_{n-1} = \mu P^n$ où la puissance est celle des matrices. La loi de X est décrite par la loi initiale μ et le semi-groupe[2] $(P^n)_{n\geq 0}$, avec la convention $P^0 = I$. Ceci montre que la formulation des chaînes de Markov homogènes comme suites récurrentes aléatoires du premier ordre cache en fait une **équation d'évolution linéaire de mesures de probabilité** : $\mu_{n+1} = \mu_n P$, avec $\mu_0 = \mu$, ou encore $\partial_n \mu_n := \mu_{n+1} - \mu_n = (P - I)\mu_n =: L\mu_n$. Lorsque la chaîne est une marche aléatoire, alors P est tridiagonale et $L = P - I$ est un opérateur laplacien discret. On retrouve un analogue discret de l'équation de la chaleur $\partial_t u(t,x) = \Delta_x u(t,x)$, qui décrit l'évolution de la loi du mouvement Brownien, processus de Markov à temps et espace continus, limite d'échelle de la marche aléatoire simple.

Espace d'état

L'espace d'état E n'a pas de structure particulière ici. Bien qu'il soit identifiable à une partie de $\mathbb{N}$ car dénombrable, il n'est pas muni d'une distance, d'un ordre, ou d'une structure additive ou multiplicative. Exemples : individus dans une population, configurations d'un système complexe, etc. Ainsi, en général $\mathbb{E}(X_n)$ n'a pas de sens, mais $\mathbb{E}(f(X_n))$ a un sens pour toute fonction positive ou bornée $f : E \to \mathbb{R}$ («fonction test» ou «observable»). Exemple : $f = \alpha \mathbf{1}_A$ avec $A \subset E$, qui donne $\mathbb{E}(f(X_n)) = \alpha \mathbb{P}(X_n \in A)$.

2. On dit parfois que la matrice $L = P - I$ est le générateur du semi-groupe $(P^n)_{n\geq 0}$.