Chap.2 : Processus stochastiques - Généralités - Chaînes de Markov

2.1 Processus stochastiques, premières définitions et exemples

Un espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$ est donné.

Définition (2.1.1)

Un processus $X = (X_t)_{t\in\Theta}$ défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $(E, \mathcal{E})$, c'est une famille de v.a. définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $(E, \mathcal{E})$, indexée par $t \in \Theta$ où Θ est un ensemble (i.e. $\forall t \in \Theta$, $X_t : \Omega \to \mathbb{E}$ est une v.a.).

Par la suite on aura le plus souvent $\Theta = \mathbb{N}$ (on parle de "processus à temps discret") ou $\Theta = \mathbb{R}_+$ (on parle de "processus à temps continu").

Pour $\omega \in \Omega$ l'application

$$X_{\cdot}(\omega) : \begin{array}{ccc} \Theta & \to & E \\ t & \mapsto & X_t(\omega) \end{array}$$

est appelée "trajectoire" du processus X (associée à l'aléa ω).

Exemple 2.1.1 : La marche aléatoire (M.A.) sur $\mathbb{Z}$, définie comme suit.
On a $(Y_i)_{i\geq 1}$ suite i.i.d. de loi Ber(p). On pose $S_0 = x \in \mathbb{Z}$. Puis
$S_n = x + \sum_{i=1}^n (2Y_i - 1)$, $n \geq 1$.
Le processus $S = (S_n)_{n\geq 0}$ ainsi défini est appelé M.A. Notons qu'on a

$$\mathbb{P}(S_{n+1} = x+1 | S_n = x) = p = 1 - \mathbb{P}(S_{n+1} = x-1 | S_n = x).$$

Si $p = 1/2$ on parle de "MA symétrique".

[Dessin et exemple de trajectoire, au tableau]

Exemple 2.1.2 : Processus de renouvellement.

Soit $(X_i)_{i\geq 1}$ suite de v.a. i.i.d t.q. $X_1 > 0$ p.s. On pose $T_0 = 0$, puis
$T_n = \sum_{i=1}^n X_i = T_{n-1} + X_n$, $\forall n \geq 1$.
Puis on définit $N_t = \max\{n \geq 0 : \ T_n \leq t\}$

[Dessin et exemple de trajectoire, au tableau]

Notons que $\forall t \geq 0$, $\forall n \geq 0$ on a $\{T_n \leq t\} = \{N_t \geq n\}$.

Les T_k's sont appelés "temps d'occurence". Les X_k's sont appelés "temps d'inter-occurence".

Le processus $N = (N_t)_{t\geq 0}$ est un "processus de renouvellement". Ses trajectoires sont entièrement déterminées par les X_k's.

Exercice 2.1.1 (Processus de Poisson) : On considère le processus de renouvellement $(N_t)_{t\geq 0}$ défini par : les X_i's sont i.i.d. de loi $\mathcal{E}(\lambda)$.
Montrer que $\forall t > 0$, $\forall k \geq 1$, $\mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$,
i.e. $N_t \sim \mathcal{P}(\lambda t)$ (loi de Poisson de paramètre λt), pour tout $t > 0$, d'où on appelle N un "processus de Poisson".

☛ Plus sur le processus de Poisson dans le chapitre 3 !

Définition (2.1.2)

Soit $X = (X_t)_{t\in\Theta}$ *un processus à valeurs dans* E *dénombrable. On dit que* X *est de Markov homogène si pour tous* $n \in \mathbb{N}^*,\ 0 \le t_1 < \ldots < t_n < t_{n+1},\ x_1, \ldots, x_{n+1} \in E$ *on a*

$$\begin{aligned}\mathbb{P}(X_{t_{n+1}} = x_{n+1}\,|\,X_{t_n} = x_n, \ldots, X_{t_1} = x_1) &= P(X_{t_{n+1}} = x_{n+1}\,|\,X_{t_n} = x_n)\\ &= p(t_{n+1} - t_n, x_n, x_{n+1})\\ &= \mathbb{P}(X_{t_{n+1}-t_n} = x_{n+1}\,|\,X_0 = x_n).\end{aligned}$$

Commentaire : Pour un processus de Markov la loi de la position à l'instant t_{n+1} dépend de la position à l'instant t_n mais pas de celle aux instants précédents.

Proposition (2.1.1)

Pour un processus de Markov homogène $X = (X_n)_{n\in\mathbb{N}}$*, indexé par* $n \in \mathbb{N}$*, la propriété de Markov homogène s'écrit de façon équivalente : pour tous* $n \in \mathbb{N}^*,\ ; x_1, \ldots, x_{n+1} \in E$

$$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1}|X_n = x_n, \ldots, X_1 = x_1) &= \mathbb{P}(X_{n+1} = x_{n+1}|x_n = x_n)\\ &= \mathbb{P}(X_1 = x_{n+1}|X_0 = x_n).\end{aligned}$$

2.2 Chaînes de Markov

Une "chaîne de Markov" (CM) c'est un processus $X = (X_n)_{n\in\mathbb{N}}$, indexé par $n \in \mathbb{N}$, à valeurs dans E dénombrable et qui est Markov homogène.

On voit par la proposition 2.1.1 que son comportement est entièrement déterminé par la donnée :
1) De la loi de sa position initiale X_0.
2) De sa "matrice de transition" (ou "matrice stochastique" ou "matrice de passage") $Q = \{Q(x,y)\}_{(x,y)\in E^2}$ qui contient les quantités

$$Q(x,y) = \mathbb{P}(X_1 = y | X_0 = x) = \mathbb{P}(X_{n+1} = y | X_n = x), \;\; \forall n \geq 0.$$

Propriété importante : les lignes d'une matrice de transition somment à 1,
i.e. $\forall x \in E$ on a $\sum_{y\in E} Q(x,y) = 1$.

☛ Exemple dans l'exercice 2 de la Fiche 3.

Proposition (2.2.1)

Soit $X = (X_n)_{n\geq 0}$ une chaîne de Markov de matrice de transition Q. Alors pour tous $n \geq 1$ et $x, y \in E$ on a

$$\mathbb{P}(X_n = y \,|\, X_0 = x) = Q^n(x, y)$$

(ici Q^n désigne simplement le produit matriciel constitué par Q multipliée $n-1$ fois par elle-même).
En particulier on a les équations de Chapman-Kolmogorov :

$$\forall x, y \in E, \;\; \forall m, n \in \mathbb{N},$$

$$\mathbb{P}(X_{m+n} = y \,|\, X_0 = x) = \textstyle\sum_{z\in E} \mathbb{P}(X_m = z | X_0 = x)\mathbb{P}(X_n = y | X_0 = z)$$

ou, écrit matriciellement,

$$\forall x, y \in E, \;\; \forall m, n \in \mathbb{N}, \quad Q^{m+n}(x, y) = \sum_{z\in E} Q^m(x, z) Q^n(z, y).$$

Preuve : [au tableau]

On va maintenant mettre en lumière un phénomène d'équilibre des chaînes de Markov.

Exemple 2.2.1 : A Griville s'il fait beau un jour, il pleut le lendemain avec probabilité $3/4$. S'il pleut un jour, il fait beau le lendemain avec probabilité $1/8$. Il en est ainsi depuis la fondation de la ville, en l'an 1200. On se propose de calculer la probabilité qu'il fasse beau demain (sans information sur le temps d'aujourd'hui).

[calculs au tableau]

Ainsi on voit que

$$Q^n \xrightarrow{n\to\infty} \begin{pmatrix} 1/7 & 6/7 \\ 1/7 & 6/7 \end{pmatrix}.$$

En d'autres termes, en temps long (i.e. pour n grand) la chaîne est en B avec proba $1/7$ et en P avec proba $6/7$, et ce *d'où qu'elle soit partie* ! En effet

$$\forall x \in \{\mathrm{B},\mathrm{P}\}, \quad Q^n(x,B) = \frac{1}{7} \quad \text{pour } n \text{ grand.}$$

La mesure $\mu = (\frac{1}{7}, \frac{6}{7})$ décrit l' "équilibre" de la chaîne.

Par ailleurs on peut faire l'observation suivante :

Si X_0 est distribuée selon μ alors

$$\forall n \geq 1, \quad \mathbb{P}(X_n = B) = \mu(B) \quad \text{et} \quad \mathbb{P}(X_n = P) = \mu(P).$$

[calcul au tableau]

Donc si $X_0 \sim \mu$ on a $X_n \sim \mu$ pour tout instant $n \geq 1$ ultérieur !

☛ la mesure μ apparaît comme la mesure "invariante" de la chaîne X.

Définition (2.2.1)

Une mesure μ sur E est appelée mesure invariante pour une CM de matrice de transition Q si

$$\forall x \in E, \quad \mu(x) = \sum_{y \in E} \mu(y) Q(y, x).$$

On note matriciellement $\mu = \mu Q$ (i.e. μ est le vecteur propre à gauche associé à la valeur propre 1).

Si $\sum_{x \in E} \mu(\{x\}) = 1$ on parle de "mesure de probabilités" invariante.

Théorème (2.2.1,Perron-Frobenius)

Etant donné une matrice de transition Q sur E espace fini, il existe au moins une mesure non triviale (i.e. non nulle), invariante pour Q.

Preuve : Admis ... ou à voir dans les bons bouquins de probas...

Définition (2.2.2)

Soit Q une matrice de transition sur E.
i) On dit que $j \in E$ est accessible depuis $i \in E$, et on note $i \to j$ s'il existe $n(i,j)$ tel que $Q^{n(i,j)}(i,j) > 0$.
ii) Si i est également accessible depuis j on dit que i et j communiquent, et on note $i \leftrightarrow j$.
iii) Une matrice de transition Q est dite irréductible (on dit de même de la CM associée) si $\forall (x,y) \in E \times E$ on a $x \leftrightarrow y$.

Théorème (2.2.2, Lemme de Kac)

Soit μ mesure de probabilités invariante pour Q irréductible sur E. Soit X CM de matrice de transition Q. Soit $x \in E$ tel que $\mu(x) \neq 0$. Alors

$$\mathbb{E}(T^{x \to x}) = \frac{1}{\mu(x)}$$

où on a noté $T^{x \to x} = \inf\{n \geq 1 : X_n^x = x\}$, le "temps de retour" en x (X^x désigne la CM X avec $X_0 = x$).

Conséquence des théorèmes 2.2.1 et 2.2.2 : Pour X CM irréductible sur E espace d'état fini il existe une unique mesure de probabilités invariante.

En effet il existe une mesure invariante μ par Perron Frobenius.

...

...

Puis on peut fabriquer $\bar{\mu}$ mesure de probabilités invariante en normalisant μ :

$$\forall x \in E, \quad \bar{\mu}(x) = \frac{\mu(x)}{\sum_{y \in E} \mu(y)}$$

(on a bien $\sum_{x \in E} \bar{\mu}(x) = 1$). Posons $c = \sum_{y \in E} \mu(y)$, on a

$$\bar{\mu} = \frac{1}{c}\mu = \frac{1}{c}\mu Q = \bar{\mu} Q,$$

donc $\bar{\mu}$ est bien invariante.

Pour finir $\bar{\mu}$ est unique grâce au lemme de Kac. En effet si il existe ν deuxième mesure de probas invariante on a

$$\bar{\mu}(x) = \frac{1}{\mathbb{E}(T^{x \to x})} = \nu(x), \quad \forall x \in E.$$

Définition (2.2.2)

Soit X une CM de matrice de transition Q sur E. On dit que X (ou Q) est apériodique si pour tout $x \in E$ on a $\mathrm{pgcd}\{n \geq 1 : \ Q^n(x,x) > 0\} = 1$.

On a les résultats de convergence généraux suivants (admis).

Théorème (2.2.3, théorème ergodique)

Soit X une CM irréductible sur E, récurrente positive (i.e. $\mathbb{E}(T^{x \to x}) < \infty$, $\forall x \in E$). On suppose que X admet une mesure de probabilités invariante μ (elle est alors unique). Alors pour toute fonction $f : E \to \mathbb{R}$ bornée on a

$$\frac{1}{n+1} \sum_{k=0}^{n} f(X_k) \xrightarrow[n \to \infty]{} \sum_{x \in E} f(x)\mu(x) = \int_E f \, d\mu.$$

Théorème (2.2.4)

Soit X une CM irréductible, apériodique, de matrice de transition Q sur E. Alors

$$\forall x, y \in E, \quad \lim_{n \to \infty} Q^n(x,y) = \frac{1}{\mathbb{E}(T^{y \to y})}.$$

Retour sur l'exemple 2.2.1 : ce qu'on a illustré avec les calculs de la dernière fois sur l'exemple de Griville c'est le théorème 2.2.4.
En effet...
☛ [Calculs au tableau]