QUELQUES ÉLÉMENTS SUR LES PROCESSUS DE MARKOV

CLÉMENT FOUCART

L'objectif principal de cette note est de faire un résumé "express" des résultats de la théorie des processus de Markov que nous utiliserons pour l'étude des processus de branchement. La plupart des résultats sont classiques et enseignés lors d'un cours de processus de Markov en temps et espace continus (typiquement celui du premier semestre du master de proba). Les résultats sont énoncés sans démonstration. La notion de problème martingale est abordée et un résultat général sur la forme des générateurs est énoncé. Pour un traitement rigoureux nous renvoyons aux références suivantes
1) Cours de T. Duquesne (Markov I).
2) Rogers, Williams : Diffusions, Markov Processes and Martingales Volume 1. Foundations.
3) Revuz Yor : Continuous Martingales and Brownian motion : Chapitre III
4) Ethier Kurtz : Markov Processes : Characterization and convergences.

1. Processus, semi-groupe et lois fini-dimensionnelles

Soit E un espace polonais (métrique, séparable, complet) et X un processus à valeurs dans E, adapté, c'est-à-dire une famille de variables aléatoires $(X_t, t \in T)$ définie sur un même espace de probabilité filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ telle que X_t est mesurable par rapport à $\mathcal{F}_t$. Dans le cas des CSBPs, on prendra $T = [0, \infty)$ et $E = [0, \infty]$.

Etant donné un processus $(X_t, t \geq 0)$ (adapté, on le supposera toujours). La propriété de Markov simple affirme que le comportement futur du processus après tout temps fixé s ne dépend du passé qu'à travers la valeur du processus à l'instant s : pour tout A borélien de E, tous $t > s \geq 0$:

$$\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t \in A|X_s) =: \phi(X_s) \tag{1}$$

où $\phi(x) =: P_{s,t}(x, A)$ est la probabilité partant de x à l'instant s d'atteindre A à l'instant t. Le processus est homogène si cette probabilité ne dépend que de $t-s$, i.e. $P_{s,t}(x, \cdot) = P_{0,t-s}(x, \cdot) =: P_{t-s}(x, \cdot)$. On appelle noyau de transition la famille $(P_t(x, \cdot), x \in E, t \in T)$ des lois des X_t sous $\mathbb{P}_x$.

A tout processus homogène, on associe la famille d'opérateurs agissant sur les fonctions boréliennes (positives ou bornées) $(P_t)_{t \geq 0}$, définie par

$$\mathbb{E}\left(f(X_t)|X_s\right) = P_{t-s}f(X_s) \quad \mathbb{P} - \text{p.s.}$$

On note ν la loi de X_0, c'est la loi initiale de X. Lorsque $\nu = \delta_x$ avec $x \in E$, on note $\mathbb{P}_x$ la mesure de probabilité sur Ω telle que $X_0 = x$ p.s. et on a pour tout $x \in E$,

$$\mathbb{E}\left(f(X_t)|X_0 = x\right) = P_t f(x).$$

La propriété de Markov (1), avec l'homogénéité implique la propriété de semi-groupe : $\forall f$ borélienne bornée ou positive

$$P_{t+s}f = P_t(P_s f).$$

En effet pour tout $x \in E$:

$$P_{t+s}f(x) = \mathbb{E}\big(\mathbb{E}(f(X_{t+s})|X_t)|X_0 = x\big) = \mathbb{E}\big(P_s f(X_t)|X_0 = x\big) = P_t(P_s f)(x).$$

On appelle (P_t) le semi-groupe de X. On note $\mathbb{E}_x$ l'espérance sous $\mathbb{P}_x$ et l'on a

$$P_t f(x) = \mathbb{E}_x[f(X_t)] = \int f(y) P_t(x, \mathrm{d}y)$$

Le noyau de transition et la loi initiale caractérisent les lois fini-dimensionnelles du processus X, à savoir les lois des vecteurs aléatoires $(X_0, X_{t_1}, \cdots, X_{t_n})$ avec $t_1 \leq t_2 \leq \cdots \leq t_n$, de la façon suivante : pour tous $A_0, A_1, \cdots, A_n \in \mathcal{B}_E$:

$$(2) \quad \mathbb{P}(X_0 \in A_0, X_{t_1} \in A_1, \cdots, X_{t_n} \in A_n)$$
$$= \int_{A_0} \underbrace{\mathbb{P}(X_0 \in \mathrm{d}x_0)}_{\nu(\mathrm{d}x_0)} \int_{A_1} P_{t_1}(x_0, \mathrm{d}x_1) \int_{A_2} P_{t_2-t-1}(x_1, \mathrm{d}x_2) \cdots \int_{A_n} P_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n).$$

On en déduit que le semigroupe et la loi initiale caractérisent la loi du processus X, vu comme variable aléatoire à valeurs dans E^T.

Théorème 1.1. *Réciproquement étant donné un semi-groupe $(P_t, t \geq 0)$ défini sur les fonctions boréliennes de E, si celui-ci est positif, i.e. $f \geq 0 \Longrightarrow P_t f \geq 0$ et $P_t 1 = 1$ pour tout $t \in T$, alors on peut construire un processus de Markov X (i.e. un espace de probabilité filtré sur lequel X est défini) à valeurs dans E dont les lois de fini-dimensionnelles satisfont* (2).

Preuve : Appliquer le théorème d'extension de Daniell-Kolmogorov.

Remarque 1.1. *Il suffit en fait que le semigroupe $(P_t, t \geq 0)$ soit défini sur $\mathcal{C}_b$ (espace des fonctions continues bornées) ou $\mathcal{C}_0$ (espace des fonctions continues telles que $f(\infty) = 0$) pour avoir un noyau de transition $P_t(x, \cdot)$. (l'existence du noyau est donnée par un théorème de représentation de Riesz).*

2. Processus de Feller, régularité et propriété de Markov forte

Définition 2.1 (Semi-groupe de Feller)**.** *Le semigroupe (P_t) est Feller si*

$$P_t \mathcal{C}_0 \subset \mathcal{C}_0 \text{ et } P_t f \underset{t \to 0+}{\longrightarrow} f \text{ uniformément}$$

Remarque 2.1. *Un théorème assure en fait que la convergence uniforme ci dessus est équivalente à la convergence simple lorsque $P_t \mathcal{C}_0 \subset \mathcal{C}_0$.*

On dit d'un processus qu'il est Feller si son semigroupe est Feller.

Théorème 2.1. *Soit X un processus de Feller, X admet une modification (ou version) continue à droite avec des limites à gauches (càdlàg), c'est-à-dire que sur le même espace de probabilité, il existe $(\tilde{X}_t, t \geq 0)$ de même loi que X telle que $\forall t \in \mathbb{R}_+$, $\mathbb{P}(X_t = \tilde{X}_t) = 1$ et $\mathbb{P}$-ps :*

$$\forall t \in \mathbb{R}_+, \ \tilde{X}_{t+} = \tilde{X}_t \text{ et } \forall t \in \mathbb{R}_+^\star, \ \tilde{X}_{t-} := \lim_{\epsilon \to 0^+} \tilde{X}_{t-\epsilon} \text{ existe.}$$

En particulier, la mesure de probabilité $\mathbb{P}_x$, loi du processus X issu de x, est portée par $\mathcal{D}(\mathbb{R}_+, E)$ l'espace de Skorokhod.

Lorsqu'un processus est Feller, on prendra toujours la version càdlàg. On a également le théorème suivant. En complétant la filtration naturelle de X (voir Chapitre III, pages 92-93, Revuz Yor), le processus de Feller satisfait la propriété de Markov forte. On notera la filtration naturelle continue à droite et complétée de X encore $(\mathcal{F}_t)$.

Théorème 2.2. *Pour tout temps d'arrêt T fini presque-sûrement, X_T est $\mathcal{F}_T$-mesurable et*

$$\mathbb{E}_x\big(F(X_{t+T}, t \geq 0)|\mathcal{F}_T\big) = \mathbb{E}\big(F(X_{T+t}, t \geq 0)|X_T\big) \quad \mathbb{P}_x - \textit{presque-sûrement.}$$

pour toute fonction F mesurable bornée ou positive et tout $x \in E$.

De façon plus concise, et en laissant T possiblement prendre la valeur ∞, on a que sur l'événement $\{T < \infty\}$,

$$\mathbb{E}_x(f(X_{t+T})|\mathcal{F}_T) = \mathbb{E}_{X_T}(f(X_t))$$

et sous la probabilité $\mathbb{P}(\cdot|X_T = x)$, le processus translaté $(X_{t+T}, t \geq 0)$ a même loi que $\mathbb{P}_x$ et est indépendant de $\mathcal{F}_T$.

3. Générateurs et problèmes martingales

Etant donné un processus de Feller X, on appelle générateur de X, l'opérateur $\mathcal{A}$ défini par

$$\mathcal{A}f := \lim_{t \downarrow 0} \frac{1}{t}\big(P_t f - f\big) \text{ au sens de la convergence uniforme i.e. dans } \big(\mathcal{C}_0, \|\cdot\|_\infty\big)$$

sur le domaine $\mathcal{D}_X$ défini intrinsèquement comme l'ensemble des fonctions appartenant à $\mathcal{C}_0$ pour lesquelles la limite existe.

Remarque 3.1. *Un théorème général affirme à nouveau que la convergence uniforme est équivalente à celle ponctuelle pour les processus de Feller. Voir Theorem 1.33 dans Primer on Feller semigroups, Böttcher, Schilling, Wang (2013).*

Martingale de Dynkin

$$\forall f \in \mathcal{D}_X, \Big(f(X_t) - \int_0^t \mathcal{A}f(X_s)\mathrm{d}s\Big)_{t\geq 0} \text{ est une martingale (continue à droite).}$$

Réciproquement, si $f, g \in \mathcal{C}_0$ sont telles que

$$\Big(f(X_t) - \textstyle\int_0^t g(X_s)\mathrm{d}s\Big)_{t\geq 0} \text{ est une martingale}$$

alors

$$f \in \mathcal{D}_X \text{ et } \mathcal{A}f = g.$$

Il est très difficile en général de connaître $\mathcal{D}_X$. Nous utiliserons dans la suite une notion plus souple, appelée problème-martingale (approche de Stroock-Varadhan). Soit $\mathcal{A}$ un opérateur, $\mathcal{D}$ une classe de fonctions régulières sur laquelle $\mathcal{A}$ fait sens.

Définition 3.1. *Un processus de Markov X est solution du problème martingale (PM) : $(\mathcal{A}, \mathcal{D})$ si*

$$\forall f \in \mathcal{D}, \Big(f(X_t) - \int_0^t \mathcal{A}f(X_s)\mathrm{d}s\Big)_{t\geq 0} \textit{ est une martingale.}$$

*On dit que le problème-martingale est **bien posé** s'il y a unicité en loi du processus de Markov qui le satisfait.*

Exemple 3.1. *Le PM $(\mathcal{A}, \mathcal{D}_X)$ est bien posé. Mais parfois pour caractériser la loi du processus il suffit d'avoir le PM $(\mathcal{A}, \mathcal{D})$ pour une classe simple de fonctions $\mathcal{D} \subset \mathcal{D}_X$.*

Remarque 3.2. *Les éléments présentés ici sur les problèmes-martingales ne sont que les prémisses d'une théorie générale, où par exemple le processus solution du PM n'est pas supposé Markovien. En particulier, un résultat important affirme que si le PM est bien posé alors le processus solution est automatiquement Markovien. Pour simplifier nous ne rentrons pas dans cette théorie, qui nécessiterait un développement plus long, et renvoyons au livre d'Ethier et Kurtz (Chapitre 4).*

A quoi ressemble l'opérateur $\mathcal{A}$ en général ?

Le théorème ci-dessous donne la forme générale du générateur d'un processus de Feller.

Théorème 3.1 (Courrège 1966)**.** *Soit X un processus de Feller à valeurs réelles (pour simplifier). Si $\mathcal{C}_c^2$ (les fonctions C^2 à support compact), est inclus dans $\mathcal{D}_X$ alors le générateur $\mathcal{A}$ du processus X est de la forme :*

$$\begin{aligned}\mathcal{A}g(x) &= \frac{1}{2}a(x)g''(x) + b(x)g'(x) - c(x)g(x) \quad (\mathcal{A}^{\mathrm{loc}}) \\ &\quad + \int_{\mathbb{R}} \big(g(x+y) - g(x) - yg'(x)\mathbb{1}_{\{|y|\leq 1\}}\big)\,\nu(x,\mathrm{d}y) \quad (\mathcal{A}^{\mathrm{nonloc}}),\end{aligned}$$

où $a \geq 0, c \geq 0$, et b sont des fonctions et pour tout $x \in \mathbb{R}$, $\nu(x, \mathrm{d}y)$ est une mesure borélienne sur $\mathbb{R} \setminus \{0\}$ telle que $\int y^2 \mathbb{1}_{\{|y|\leq 1\}}\nu(x, \mathrm{d}y) < \infty$.

— $\mathcal{A}^{\text{loc}}$ est le terme diffusif (partie locale) avec un terme de meurtre c, on reconnait le générateur d'une diffusion (i.e. d'un processus de Markov **continu**). Dans le cas de la diffusion de Feller vue dans le cours, $c, b \equiv 0$ et $a(x) = \sigma^2 x$.
— $\mathcal{A}^{\text{nonloc}}$ est un terme non local (même si $f = g$ sur un voisinage de x, alors en général $\mathcal{A}^{\text{nonloc}} f(x) \neq \mathcal{A}^{\text{nonloc}} g(x)$ contrairement à $\mathcal{A}^{\text{loc}}$). **Le deuxième terme se traduit trajectoriellement par des sauts dans le processus, mais comprendre cela précisément nécessite un travail conséquent.**

Remarque 3.3. *Il est naturel de s'interroger sur le terme,* ***dit de compensation*** $-yg'(x)\mathbb{1}_{\{|y|\leq 1\}}\cdots$

Le théorème 3.1 dépasse largement le cadre du cours [1] et ne sera pas utilisé en tant que tel dans l'étude des processus de branchement. Nous étudierons et interpréterons le générateur des CSBPs directement et retrouverons un opérateur de cette forme.

1. voir N. Jacob, Pseudo-differential operators and Markov Processes