Sujet de Thèse

— **Titre :** Statistique des processus de Markov déterministes par morceaux pour la mobilité.
— **Unité de recherche :** IRMAR, UMR-6625
— **Thème :** Statistique, Probabilités
— **Mots clefs :** Processus de Markov déterministes par morceaux, estimation non-paramétrique, théorèmes limites, environnement aléatoire, application à des questions environnementales
— **Équipe d'encadrement :**
1. Krell Nathalie, nathalie.krell@univ-rennes2.fr, Université de Rennes 2, Maîtresse de conférence (HDR)
2. Ronan Le Guével, ronan.leguevel@univ-rennes2.fr, Université de Rennes 2, Maître de conférence (HDR)

Processus de Markov déterministes par morceaux

Un processus de Markov déterministe par morceaux (PDMP pour l'anglais piecewise-deterministic Markov process) [3, 4] est un processus à temps continu $(X_t)_{t\in\mathbb{R}_+}$ évoluant de manière déterministe pour presque tout t mais subissant ponctuellement des sauts aléatoires à des instants aléatoires. Un tel processus défini sur un ouvert E de $\mathbb{R}^d$ est décrit par un triplet (λ, Q, Φ) où $\lambda : E \to \mathbb{R}_+$ est le taux de saut, $Q : \overline{E} \times \mathcal{B}(E) \to [0,1]$ est le noyau de transition et $\Phi : \mathbb{R} \times E \to E$ est le flot auquel obéit la dynamique déterministe. Si on note T_n les instants de saut du processus (avec $T_0 = 0$), la dynamique du processus $(X_t)_{t\in\mathbb{R}_+}$ entre les instants T_n et T_{n+1}, conditionnellement à X_{T_n}, s'écrit ainsi :

— Pour tout $T_n \leq t < T_{n+1}$, $X_t = \Phi(t - T_n|X_{T_n})$;
— La position $X_{T_{n+1}}$ en T_{n+1} est sélectionnée selon le noyau Q, i.e.

$$\mathbb{E}\left[\varphi(X_{T_{n+1}}) \,|\, X_{T_n}, T_{n+1} - T_n\right] = \int_E Q(\Phi(T_{n+1} - T_n|X_{T_n}), dy)\varphi(y).$$

Il reste alors à décrire la loi des instants de saut, $S_{n+1} = T_{n+1} - T_n$,

$$\mathbb{P}(S_{n+1} > t \,|\, X_{T_n} = x) = \exp\left(-\int_0^t \lambda(\Phi(s|x))ds\right) \mathbb{1}_{\{0\leq t<t^\star(x)\}},$$

où $t^\star : E \to \mathbb{R}_+ \cup \{+\infty\}$ désigne le temps déterministe d'atteinte de la frontière de E suivant le flot.

Objectif de la thèse

La **statistique des PDMPs** a déjà été étudiée par exemple dans le contexte de la **modélisation pour des questions environnementales**, notamment en écologie [1, 7]. L'objectif de cette thèse sera de l'appliquer à une problématique de modélisation de la mobilité des personnes en milieu urbain. Dans ce cadre, on étudiera un PDMP spatial à valeur dans $\mathbb{R}^2$ décrivant le déplacement d'une personne au cours du temps. On introduira également un processus stochastique $(E_t)_{t\in\mathbb{R}_+}$ d'environnement, affectant la dynamique du processus à travers une ou plusieurs de ses caractéristiques. Ce processus environnemental peut à la fois provenir de facteurs exogènes (e.g. le prix de l'essence, la météo) ou endogènes (e.g. une maladie, un congé).

Nous nous poserons alors les questions suivantes :

— Peut-on décrire la loi du processus pour des variables environnementales observées ou non ?

— Quel est l'effet de la variable environnementale sur chacune des caractéristiques locales du processus ?
— Pour quelles valeurs de la variable environnementale le processus obéit-il à une loi cible correspondant à une dynamique souhaitable pour le système écologique modélisé ?

Démarche statistique

Pour répondre à ces questions, on pourra considérer deux schémas d'observations : une trajectoire en temps long d'une part, et plusieurs individus évoluant sur une durée déterminée d'autre part. À partir de ces observations, on cherchera dans un premier temps à estimer de manière précise les paramètres du processus en environnement déterministe, notamment en affinant des résultats récents de la littérature sur l'estimation du taux de saut [2, 5].

Dans un deuxième temps, on se demandera comment calculer la loi du processus couplé avec l'environnement aléatoire, mais également conditionnellement à ce dernier. On regardera en particulier comment se propagent les erreurs d'estimation dans ces calculs. Le problème inverse consistant à déterminer la variable environnementale sous laquelle le processus obéit à une loi donnée sera abordé par des techniques numériques (MCMC par exemple). Un modèle plus complexe intégrant des interactions entre individus pourra être abordé dans un dernier temps, ce qui nécessitera d'adapter les techniques d'estimation comme dans [6].

Les résultats théoriques obtenus seront validés par des simulations numériques, et confrontés à des données réelles de la littérature, comme les données GPS de circulation automobile dans le bassin rennais (données NeoVya).

Références

[1] Candy Abboud, Rachid Senoussi, and Samuel Soubeyrand. *Piecewise-deterministic Markov Processes for Spatio-temporal Population Dynamics*, chapter 7, pages 209–255. John Wiley & Sons, Ltd, 2018.

[2] Romain Azaïs and Aurélie Muller-Gueudin. Optimal choice among a class of nonparametric estimators of the jump rate for piecewise-deterministic Markov processes. *Electronic Journal of Statistics*, 10(2) :3648 – 3692, 2016.

[3] Azaïs, Romain, Bardet, Jean-Baptiste, Génadot, Alexandre, Krell, Nathalie, and Zitt, Pierre-André. Piecewise deterministic markov process — recent results. *ESAIM : Proc.*, 44 :276–290, 2014.

[4] M. H. A. Davis. Piecewise-deterministic markov processes : A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society : Series B (Methodological)*, 46(3) :353–376, 1984.

[5] Nathalie Krell and Émeline Schmisser. Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes. *Bernoulli*, 27(4) :2362 – 2388, 2021.

[6] Frédéric Lavancier and Ronan Le Guével. Spatial birth-death-move processes : basic properties and estimation of their intensity functions. *Journal of the Royal Statistical Society (B)*, 83(4) :798–825, 2021.

[7] Arnaud Personne. *Dynamique du modèle de Moran en environnement aléatoire.* Theses, Université Clermont Auvergne [2017-2020], December 2019.