Mini-symposium Hawkes
Processus de Hawkes et leurs applications

Mini-symposium porté en partie par le projet ANR Calibration (2011 BS01 010 01)

Résumé

Les processus de Hawkes constituent une famille de processus ponctuels introduite dans les années 70 - pour modéliser à l'origine les répliques des tremblements de terre - et qui suscite un regain d'intérêt et un développement significatif dans de nombreuses applications depuis quelques années, allant des neurosciences aux réseaux sociaux en passant par la finance statistique. Le but de ce symposium est de présenter quelques unes des avancées et applications dans ce domaine.

Organisateur(s)

1. **Patricia Reynaud-Bouret**, Laboratoire JA Dieudonné, Univ. Nice Sophia Antipolis.

Liste des orateurs

1. **Julien Chevallier**, Laboratoire JA Dieudonné, Univ. Nice Sophia Antipolis
   Titre : Approche microscopique d'un système structuré en âge.
2. **Sophie Donnet**, INRA AgroParisTech
   Titre : Estimation bayésienne non paramétrique pour processus de Hawkes multidimensionnels.
3. **Thibault Jaisson**, CMAP, Ecole Polytechnique
   Titre : Estimation de noyaux de Hawkes decroissant lentement : Application à la modélisation du carnet d'ordre..
4. **Laure Sansonnet**, AgroParisTech
   Titre : Processus de Hawkes localement stationnaires.

Patricia Reynaud-Bouret, Univ. Nice Sophia Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France, `reynaudb@unice.fr`
Julien Chevallier, Univ. Nice Sophia Antipolis, CNRS, LJAD, UMR 7351, 06100 Nice, France, `Julien.CHEVALLIER@unice.fr`
Sophie Donnet, INRA - Unité MIA 518 AgroParisTech 16 rue Claude Bernard 75 231 Paris Cedex 05, `sophie.donnet@agroparistech.fr`
Thibault Jaisson, Ecole Polytechnique 91128 Palaiseau, France, `thibault.jaisson@polytechnique.edu`
Laure Sansonnet, AgroParisTech 16 rue Claude Bernard 75 231 Paris Cedex 05, `laure.sansonnet@agroparistech.fr`

Introduction

Les processus de Hawkes constituent une famille de processus ponctuels introduite dans les années 70 - pour modéliser à l'origine les répliques des tremblements de terre - et qui suscite un regain d'intérêt et un développement significatif dans de nombreuses applications depuis quelques années, allant des neurosciences aux réseaux sociaux en passant par la finance statistique. Le but de ce symposium est de présenter quelques unes des avancées et applications dans ce domaine. Voici les résumés des quatre interventions.

1 Approche microscopique d'un système structuré en âge (J.Chevallier)

On cherche ici à relier deux approches mathématiques de la modélisation de réseaux de neurones.
D'un côté, on s'intéresse au modèle introduit par K. Pakdaman, B. Perthame et D. Salort dans une série d'articles [1, 2, 3]. Ce modèle s'intéresse à la dynamique du réseau dans sa globalité, à un niveau macroscopique. Ils considèrent la densité de probabilité $n(s,t)$ de trouver, au temps t, un neurone avec un âge s, où l'âge représente le temps depuis la dernière décharge. Ils étudient le système d'équations aux dérivées partielles suivant

$$\begin{cases} \frac{\partial n(s,t)}{\partial t} + \frac{\partial n(s,t)}{\partial s} + p(s, X(t))\ n(s,t) = 0, \\ m(t) := n(s=0,t) = \int_0^{+\infty} p(s, X(t))\ n(s,t) ds, \end{cases} \tag{PPS}$$

où $X(t)$ représente l'activité globale du réseau. Pour modéliser l'intégration synaptique, on peut prendre

$$X(t) = \int_0^t d(s) m(t-s) ds,$$

où d est une fonction de délai.
De l'autre côté, on s'intéresse aux neurones de manière individuelle, à un niveau microscopique. Les temps de décharge de chaque neurone peuvent être modélisés par des processus ponctuels.
Pour être plus précis, on cherche à expliciter le lien qu'il y a entre l'intensité $\lambda(t, \mathcal{F}_{t-}^N)$ du processus ponctuel N associé à un neurone typique, et le coefficient $p(s, X(t))$ présent dans le système d'EDP.

Comme le système (PPS) code la dynamique de l'âge des neurones, la première étape de notre travail consiste à s'intéresser au processus d'âge associé à un processus ponctuel. Or, il se trouve que le résultat de "thinning" de Ogata [4] permet de représenter la dynamique du processus d'âge prévisible $(S_{t-})_{t\geq 0}$ sous la forme d'un système d'équations différentielles stochastiques. De plus, on peut montrer que la version "en espérance" de ce système est très proche du système (PPS).
Malheureusement, la correspondance n'est pas celle que l'on pouvait attendre. En particulier, le système que l'on trouve met en jeu l'espérance conditionnelle de l'intensité sachant l'âge, i.e.

$$\rho(t,s) = \mathbb{E}\left[\lambda(t, \mathcal{F}_{t-}^N) \middle| S_{t-} = s\right]$$

à la place du coefficient $p(s, X(t))$.
On peut toutefois établir un lien clair et justifié entre les deux échelles de modélisation quand l'intensité ne dépend que de l'âge.
Mais, le lien n'est pas clair quand l'intensité a une forme plus complexe (e.g. processus de Wold ou de Hawkes). La plus grande difficulté vient du calcul de l'intensité conditionnelle ρ. On ne peut pas l'expliciter pour des processus ponctuels très généraux. En revanche, on peut décrire l'intensité conditionnelle ρ via des équations implicites si le processus ponctuel sous-jacent N est un processus de Hawkes linéaire. Tout cela est possible grâce à sa décomposition sous formes de "clusters".
Pour conclure, dans le cadre d'une dynamique sous-jacente non markovienne (e.g. processus de Hawkes), notre résultat permet d'expliquer pourquoi le coefficient $p(s, X(t))$ doit prendre en compte toute la dynamique passée.

Cet exposé est basé sur un travail en collaboration avec M.J. Cacérès, M. Doumic et P. Reynaud-Bouret.

2 Estimation bayésienne non paramétrique pour processus de Hawkes multidimensionnels (S. Donnet)

Les processus de Hawkes multidimensionnels sont utilisés pour la modélisation des potentiels d'actions neuronaux. L'estimation des fonctions d'intensité permet de comprendre la structure d'interactions des neurones. L'estimation non-paramétrique de ces fonctions a été proposée par des méthodes de type LASSO dans un cadre fréquentiste. Nous nous intéressons à leur estimation non-paramétrique dans un cadre bayésien, ce qui permet entre autre d'intégrer des informations a priori. Une fois définie la loi a priori sur les fonctions d'intensités, l'échantillonnage de leur loi a posteriori ne peut être fait de façon exacte et nous devons avoir recours à des méthodes du type de Monte Carlo. Les algorithmes du type Monte Carlo Markov Chain (MCMC) sont les outils computationnels les plus largements utilisés en pratique : ils génèrent une chaîne de Markov dont la loi stationnaire est la loi d'intérêt (ici la loi a posteriori des paramètres). Cependant, ces algorithmes atteignent leurs limites d'application en pratique dans les problèmes de grande dimension (ce qui est le cas dans le cadre non-paramétrique). Nous proposons d'avoir recours à l'algorithme SMC proposé par [5] dont l'idée est d'échantillonner séquentiellement les paramètres conditionnellement à une partie des observations. Après avoir présenté ces algorithmes et leurs avantages, nous présenterons les résultats obtenus sur des données simulées de processus de Hawkes multidimensionnels.

3 Estimation de noyaux de Hawkes decroissant lentement : Application à la modélisation du carnet d'ordre. (T. Jaisson)

Nous présentons une version modifiée de l'estimation non paramétrique des noyaux de Hawkes de [6] que nous adaptons à des noyaux décroissant lentement comme observés en finance. Nous montrons que sur des simulations numériques de tailles raisonnables, cette méthode nous permet de retrouver des noyaux en loi de puissance significatifs sur au moins 6 ordres de grandeurs. Nous proposons ensuite un modèle de Hawkes en dimension 8 incorporant tout les événements associés aux premières limites du carnet d'ordre. En appliquant notre méthode d'estimation à ce modèle, nous obtenons les principaux couplages entre les ordres de marché, les ordres limites, les ordres d'annulation et les variations du prix.

4 Processus de Hawkes localement stationnaires (L. Sansonnet)

L'étude des processus de Hawkes, introduits par Hawkes en 1971 [7] (voir aussi [8], [9], [10] et [11]) fait le postulat de la stationnarité des fonctions d'interaction sous-jacentes. Par exemple, dans le cas d'un modèle de Hawkes univarié N (processus linéaire auto-excitant) d'intensité conditionnellement au passé

$$\lambda(t) = \nu + \int_{-\infty}^{t^-} h(t-s)\, dN(s),$$

on suppose que la fonction h est la même tout au long de l'intervalle d'observation, autrement dit elle ne dépend pas de la position temporelle du processus. Or, en pratique, cette hypothèse n'est pas raisonnable (ou sur un intervalle de temps très court). C'est pourquoi, de nombreux systèmes physiques ou biologiques sont modélisés par des processus localement stationnaires, où des fluctuations aléatoires sont produites par un mécanisme qui varie doucement dans le temps. De tels processus peuvent être approchés localement par des processus stationnaires. Le processus résultant est alors localement stationnaire sur des intervalles de temps de tailles variables. On pourra par exemple se référer à [12] pour l'étude théorique de ces processus. On définit alors un processus de Hawkes localement stationnaire (dans le cas univarié et spatial) en ayant recours à la définition par clusters du processus de Hawkes non-stationnaire, son intensité locale ainsi que son spectre de Bartlett local. Cette approche permet par exemple d'obtenir de bonnes approximations des moments d'ordre 1 et 2 du processus, via en particulier la convergence de la transformée de Laplace du processus de Hawkes non-stationnaire vers sa version locale. Ces résultats seront de plus illustrés par une étude par simulations.

Références

[1] Pakdaman, K., Perthame, B., and Salort, D. , *Dynamics of a structured neuron population*, Nonlinearity, vol. 23, no 1, 2010, p. 55.

[2] Pakdaman, K., Perthame, B., and Salort, D. , *Relaxation and self-sustained oscillations in the time elapsed neuron network model*, SIAM Journal on Applied Mathematics, vol. 73, no 3, 2013, p. 1260-1279.

[3] Pakdaman, K., Perthame, B., and Salort, D. , *Adaptation and fatigue model for neuron networks and large time asymptotics in a nonlinear fragmentation equation*, The Journal of Mathematical Neuroscience (JMN), vol. 4, no 1, 2014, p. 1-26.

[4] Ogata, Y., *On Lewis' simulation method for point processes*, Information Theory, IEEE Transactions on, vol. 27, no 1, 1981, p. 23-31.

[5] Del Moral, P., Doucet, A., and Jasra, A., *Sequential Monte Carlo Samplers.*, Journal of the Royal Statistical Society : Series B (Statistical Methodology), 68(3), 2006, p. 411-436.

[6] Bacry, E. and Muzy, J.-F., *Second order statistics characterization of Hawkes processes and non-parametric estimation*, Preprint 2014.

[7] A. G. Hawkes, *Spectra of some self-exciting and mutually exciting point processes*, Biometrika, 58(1), 1971, 83–90.

[8] D. Oakes, *The Markovian self-exciting process*, J. Appl. Probab., 12, 1975, 69–77.

[9] P. Brémaud and L. Massoulié, *Stability of nonlinear Hawkes processes*, Ann. Probab., 24(3), 1996, 1563–1588.

[10] D. J. Daley and D. Vere-Jones, *An Introduction to the Thoery of Point Processes - Volume I : Elementary Theory and Methods*, Springer, New York, 2003.

[11] P. Brémaud, L. Massoulié and A. Ridolfi, *Hawkes branching point processes without ancestors*, J. Appl. Probab., 37, 2005, 1116–1146.

[12] R. Dahlhaus, *A likelihood approximation for locally stationary processes*, Ann. Statist., 28(6), 2000, 1762–1794.