

ESTIMATION BAYÉSIENNE NON-PARAMÉTRIQUE POUR LES PROCESSUS DE HAWKES

Sophie Donnet ¹, Vincent Rivoirard ² & Judith Rousseau ³

¹ *UMR MIA 518, Agroparistech, 16 rue Claude Bernard, 75005 Paris, France.
sophie.donnet@agroparistech.fr*

² *Cérémade, Université Paris-Dauphine, Place du Maréchal Delattre de Tassigny, 75016
Paris, France. rivoirard@ceremade.dauphine.fr*

³ *Cérémade, Université Paris-Dauphine, Place du Maréchal Delattre de Tassigny, 75016
Paris, France. rousseau@ceremade.dauphine.fr*

Résumé. Les processus de Hawkes multidimensionnels sont utilisés pour la modélisation des potentiels d'actions neuronales. 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. Pour cela, nous mettons en place des algorithmes du type Sequential Monte Carlo Sampler, particulièrement adaptés à ces processus ponctuels.

Mots-clés. Processus de Hawkes, estimation bayésienne, Sequential Monte Carlo, Reversible jump,

Abstract. Multidimensional Hawkes processes are used to modelise multivariate neuron spike data. The estimation of intensity functions allows to understand the neuronal interaction structure. In a non-parametric frequentist framework, LASSO estimators have been proposed in the literature. In this work, we propose a Bayesian non-parametric estimation. We sample the posterior distribution through a Sequential Monte Carlo algorithm, well adapted to point processes.

Keywords. Hawkes processes, Bayesian estimation, Sequential Monte Carlo, Reversible jump

1 Introduction

L'activité neuronale se manifeste par des impulsions électriques très brèves appelées *potentiels d'actions* dont l'information essentielle réside dans les instants d'apparition. Afin de modéliser ces instants de pic d'activité neuronale, on utilise les processus ponctuels sur la demi-droite réelle. Parmi les processus ponctuels, le processus de Poisson génère des temps inter-événements indépendants. Afin de modéliser un phénomène d'auto-excitation, les processus de Hawkes permettent de générer des processus avec un effet d'agrégation des événements dans le temps. Dans leur version multi-dimensionnelle, les processus

de Hawkes permettent la modélisation des phénomènes d'inter-excitation ou d'inhibition entre neurones. La loi des processus de Hawkes est décrite par leur fonction d'intensité conditionnelle. Dans ce travail, nous cherchons à estimer de façon non-paramétrique ces quantités. Hansen et al. (2014) ont proposé un estimateur du type LASSO. Dans ce travail, nous nous plaçons dans un cadre bayésien. Après avoir spécifié une loi a priori sur une certaine classe de fonctions, nous proposons d'échantillonner la loi a posteriori par des méthodes du type Sequential Monte Carlo. Le papier est organisé comme suit. Dans la section 2, nous décrivons le processus de Hawkes multidimensionnel. Nous proposons une loi a priori sur les paramètres dans la partie 3. Dans la section 4, nous décrivons l'algorithme utilisé pour échantillonner la loi a posteriori. Les méthodes proposées sont appliquées sur données simulées dans la section 5.

2 Processus de Hawkes mutlidimensionnel

2.1 Processus de Hawkes unidimensionnel

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Soit \mathbf{N} un processus ponctuel. La loi du processus ponctuel peut-être caractérisée par son intensité. Soit (\mathcal{F}_t) une filtration adaptée au processus, l'intensité conditionnelle est définie par

$$\lambda^*(t) = \lambda(t|\mathcal{F}_t) = \lim_{h \rightarrow 0} \mathbb{P}(\text{Un évènement ait lieu dans } [t, t+h] | \mathcal{F}_t).$$

Dans le cas du processus de Poisson l'intensité conditionnelle ne dépend pas du passé du processus : $\lambda^*(t) = \lambda(t)$. Le processus de Hawkes prend en compte le passé du processus de la façon suivante :

$$\lambda^*(t) = \left\{ \nu + \sum_{t_i < t} h(t - t_i) \right\}_+$$

où ν est le taux de saut spontané et h est appelée *fonction d'interaction*. $\{\cdot\}_+$ garantie la positivité de l'intensité si h prend des valeurs négatives (inhibition). Ce processus avec auto-excitation génère (dans le cas où h est positive) un phénomène d'aggrégation des évènements (dans le temps).

2.2 Processus de Hawkes multidimensionnel

Pour M neurones en interaction, notons $\mathbf{N}^{(m)}$ le processus de sauts observé sur le neurone m ($1 \leq m \leq M$) et $t_1^{(m)}, \dots, t_{n_m}^{(m)}$ les instants de saut sur le neurone m . L'intensité conditionnelle s'écrit :

$$\lambda^{*(m)}(t) = \left\{ \nu^{(m)} + \sum_{\ell=1}^M \sum_{t_i^{(\ell)} < t} h_\ell^{(m)}(t - t_i^{(\ell)}) \right\}_+$$

où $h_\ell^{(m)}$ est la fonction d'excitation de m par ℓ . Afin de modéliser les inhibitions entre neurones, $h_\ell^{(m)}$ peut être amenée à prendre des valeurs négatives.

L'existence d'un régime stationnaire est garantie si le rayon spectral de la matrice $\mathbf{I} = \left(\int_0^\infty \|h_\ell^{(m)}(u)\| du \right)_{1 \leq \ell, m \leq M}$ est strictement inférieur à 1. Dans la suite, nous supposons que les fonctions d'interactions ont des supports bornés. Par soucis de simplification, nous notons $[0, s_{max}]$ le support commun aux fonctions d'interaction.

3 Paramétrisation et loi a priori

Nous cherchons à estimer $\theta = \left\{ \{\nu^{(m)}\}_{m=1\dots M}, \{h_\ell^{(m)}\}_{(\ell, m) \in \{1, \dots, M\}^2} \right\}$. Dans la suite, nous supposons que les $h_\ell^{(m)}$ sont des fonctions constantes par morceaux :

$$h_\ell^{(m)}(u) = \sum_{k=1}^{K^{\ell m}} \alpha_k^{\ell m} \mathbb{I}_{[s_k^{\ell m}, s_{k+1}^{\ell m}]}(u)$$

avec $(\alpha_k^{\ell m})_{k=1\dots K^{\ell m}} \in \mathbb{R}$ et $s_{K^{\ell m}+1} = s_{max}$. Nous considérons la reparamétrisation suivante :

$$\alpha_k^{\ell m} = I^{\ell m} \frac{\delta_k^{\ell m} A_k^{\ell m}}{\sum_{r=1}^{K^{\ell m}} \delta_r^{\ell m} A_r^{\ell m}} \frac{1}{s_{k+1}^{\ell m} - s_k^{\ell m}}, \quad \text{pour } k = 1 \dots K^{\ell m}$$

de sorte que $\int_0^\infty |h_\ell^{(m)}(u)| du = I^{\ell m}$. Nous proposons les lois a priori suivantes.

- Les amplitudes $(A_k^{\ell m})_{k=1\dots K^{\ell m}}$ sont distribuées selon une loi $\Gamma(a_A, b_A)$
- Les $(\delta_k^{\ell m})_{k=1\dots K^{\ell m}}$ sont à valeurs dans $\{-1, 0, 1\}$, de sorte que les plateaux de $h_\ell^{(m)}$ peuvent valoir exactement 0. En particulier, en imposant le dernier palier $\alpha_{K^{\ell m}}^{\ell m} = 0$, nous pouvons estimer le support de $h_\ell^{(m)}$ (s_{max} servant alors de borne supérieure pour le support).

$$\mathbb{P} [\delta_k^{\ell m} = i] = \pi_i \quad \text{avec } \pi_0, \pi_1, \pi_{-1} \in]0, 1[.$$

- $\left(\frac{s_1^{\ell m} - 0}{s_{max}}, \frac{s_2^{\ell m} - s_1^{\ell m}}{s_{max}}, \dots, \frac{s_{max} - s_{K^{\ell m}}^{\ell m}}{s_{max}} \right) \sim \text{Dirichlet}(a_s)$, pour tout $(\ell, m) \in \{1, \dots, M\}$
- $\nu^{(m)} \sim \Gamma(a_\nu, b_\nu), \forall m = 1 \dots M$
- Les nombres de paliers $(K_{lm})_{1 \leq \ell, m \leq M}$ sont distribués selon une loi de Poisson de paramètre μ_K (potentiellement tronquée).
- Les intégrales des fonctions d'interaction $\mathbf{I} = (I^{\ell m})_{1 \leq \ell, m \leq M}$ sont distribuées selon M^2 $\Gamma(a_I, b_I)$ indépendantes, restreintes à l'espace des matrices \mathbf{I} dont le rayon spectral est strictement plus petit que 1.

4 Échantillonnage de la loi a posteriori

Nous nous plaçons dans un cadre bayésien et cherchons à échantillonner la loi a posteriori. Supposons qu'on observe le processus \mathbf{N} sur M neurones dans l'intervalle $[\tau_0, \tau_1]$. Notons n_m^* le nombre de sauts sur le neurone m , et $t_1^{(m)}, \dots, t_{n_m^*}^{(m)}$ les instants de saut sur le neurone m , pour tout $m = 1 \dots M$. La loi a posteriori est $\pi(\theta|\mathbf{N}) \propto \pi(\theta)\ell(\mathbf{N};\theta)$ où $\ell(\mathbf{N};\theta)$ est la vraisemblance :

$$\ell(\mathbf{N};\theta) = \exp \left[\sum_{m=1}^M \sum_{j=1}^{n_m^*} \log \lambda^{(m)*}(t_j^{(m)}) - \int_{\tau_0}^{\tau_1} \lambda^{(m)*}(u) du \right] \quad (1)$$

Remarque 1 Notons que les intensités conditionnelles $\lambda^{(m)*}$ se calculent en prenant en compte tous les événements passés, sur tous les neurones.

Nous cherchons à fournir un échantillon $\{\theta^{(i)}\}_{i=1 \dots N}$ simulé sous la loi a posteriori. La simulation directe (impossible) sous la loi a posteriori est remplacée par la simulation sous une loi d'importance $\theta^{(i)} \sim_{i.i.d.} \eta$, à chaque particule $\theta^{(i)}$ est alors associé un poids $W^{(i)}$ tel que

$$W^{(i)} = \frac{w^{(i)}}{\sum_{i=1}^N w^{(i)}} \quad \text{avec} \quad w^{(i)} = \frac{\pi(\theta^{(i)})\ell(\mathbf{N};\theta^{(i)})}{\eta(\theta^{(i)})}$$

En pratique le choix de η est difficile puisque cette loi doit être une "bonne" approximation de la loi a posteriori. En effet, pour un mauvais choix de η , de nombreuses particules auront un poids nul et les estimateurs obtenus à partir de l'échantillon seront de mauvaise qualité. Le Sequential Monte Carlo (SMC) permet de construire séquentiellement cette loi d'importance en utilisant à chaque itération d'une partie des observations (Del Moral et al., 2006). Dans le cas particulier des processus de Hawkes multidimensionnels, nous réécrivons la vraisemblance de la façon suivante :

$$\ell(\mathbf{N};\theta) = \exp \left[\left(\sum_{l=1}^{n^*} \lambda^{(m_l)*}(t_l) - \int_{t_l^-}^{t_l} \lambda^{(m_l)*}(u) du \right) - \sum_{m=1}^M \int_{t_{n_m^*}^{(m)}}^{\tau_1} \lambda^{(m)*}(u) du \right]$$

où $n^* = \sum_{m=1}^M n_m^*$ est le nombre total de sauts et $(t_l)_{l=1 \dots n^*}$ est la séquence réordonnée des instants de saut $(t_j^{(m)})_{j=1 \dots n_m^*, m=1 \dots M}$. Pour tout l , m_l est l'indice du neurone où a eu lieu le saut de l'instant t_l et t_l^- est le saut précédent t_l sur ce même neurone. Définissons

$$\ell_n(\mathbf{N};\theta) = \exp \left[\sum_{l=1}^n \lambda^{(m_l)*}(t_l) - \int_{t_l^-}^{t_l} \lambda^{(m_l)*}(u) du \right], \quad \forall n = 1 \dots n^*$$

et $\ell_{n^*+1}(\mathbf{N};\theta) = \ell(\mathbf{N};\theta)$. Nous proposons de séquentiellement échantillonner

$$\pi_n(\theta) = \frac{\ell_n(\mathbf{N};\theta)\pi(\theta)}{p_n(\mathbf{N})} = \frac{\gamma_n(\theta)}{Z_n}.$$

L'algorithme SMC proposé par [Del Moral et al. \(2006\)](#) s'écrit comme suit :

Algorithme SMC

1. *Initialisation* : $n = 0$. $\forall i = 1 \dots N$, $\theta_0^{(i)} \sim_{i.i.d} \pi(\cdot)$
 2. *A l'itération* $n = 1 \dots n^* + 1$:
 - Calculer : $\tilde{w}_{n-1}^{(i)} = \frac{\gamma_n(\theta_{n-1}^{(i)})}{\gamma_{n-1}(\theta_{n-1}^{(i)})} = \lambda^{(m_n)*}(t_n, \theta_{n-1}^{(i)}) e^{-\int_{t_n}^{t_n} \lambda^{(m_n)*}(u, \theta_{n-1}^{(i)}) du}$ et $W_n^{(i)} = \frac{W_{n-1}^{(i)} \tilde{w}_{n-1}^{(i)}}{\sum_{j=1}^N W_{n-1}^{(j)} \tilde{w}_{n-1}^{(j)}}$.
 - *Re-échantillonnage* : si $ESS = \frac{1}{\sum_{i=1}^N (W_n^{(i)})^2} < T$, régénérer les particules selon $\sum_{i=1}^N W_n^{(i)} \delta_{\theta_n^{(i)}}$ et poser $W_n^{(i)} = 1/N$
 - *Perturber les particules* : $\forall i = 1 \dots N$, $\theta_n^{(i)} \sim_{i.i.d} K_n(\theta_{n-1}^{(i)}, \theta_n^{(i)})$ où K_n est un noyau de loi stationnaire π_n .
-

En d'autres termes, un premier jeu de N paramètres est simulé selon la loi a priori. Ces particules sont pondérées selon la vraisemblance de la première observation (t_1). A l'itération n , si la taille effective de l'échantillon (ESS) est sous un seuil limite, les particules sont échantillonnées avec remise dans ce même échantillon, de façon à privilégier les particules de fort poids. Les particules sont ensuite perturbées selon un noyau de loi stationnaire π_n , l'idée étant que π_n est relativement proche de π_{n-1} (puisque de l'un à l'autre nous n'avons ajouté qu'une observation) et si la chaîne générée par K_n se déplace vite, alors on peut penser que nous aurons simulé nos particules sous une loi proche de la loi "idéale" π_n . Si ce n'est pas le cas, la correction apportée lors du calcul des poids permet de toute façon de garantir l'exactitude de l'algorithme.

A l'issue de l'algorithme, les particules pondérées fournissent un échantillon généré exactement sous la loi a posteriori. Afin de perturber les particules selon un noyau K_n de loi stationnaire π_n , nous utilisons des techniques de Monte Carlo Markov Chain. Dans ce cas particulier, nous utilisons un algorithme du type Reversible Jumps de façon à estimer le nombre de palliers de chaque fonction d'interaction.

5 Résultats numériques

Les méthodes d'estimation sont testées sur données simulées. Nous considérons les paramètres suivants : $M = 2$ neurones, $\nu^{(1)} = \nu^{(2)} = 20$. Les fonctions d'interaction ont pour support maximal $[0, 0.04]$. Le processus est observé sur $[1, 3]$. Pour ce jeu de données nous avons $n_1^* = 215$ et $n_2^* = 117$. L'algorithme SMC est appliqué avec 1000 particules. Les estimateurs des fonctions h tracés sur la Figure 1 donnent un premier aperçu de l'efficacité de l'algorithme.

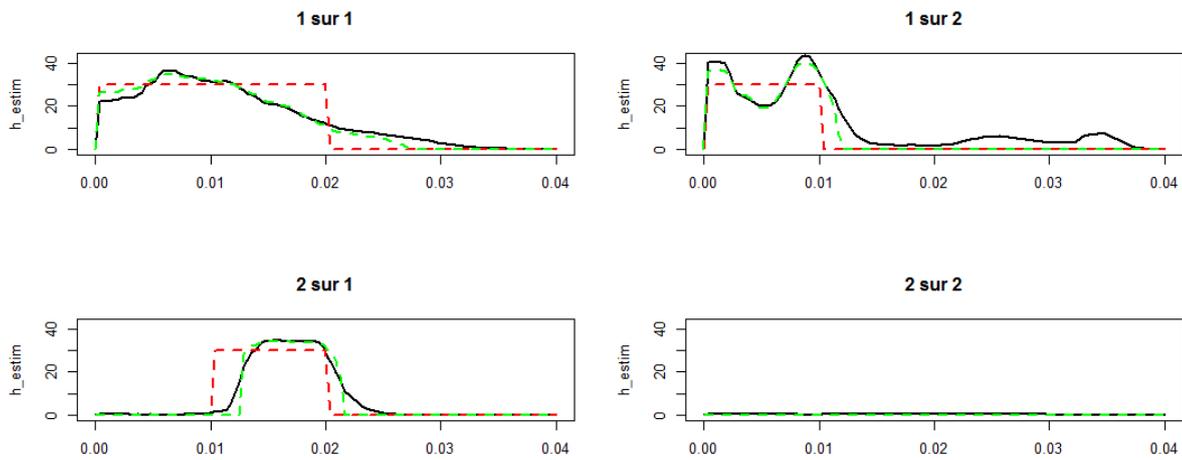


FIGURE 1 – Estimation des fonctions d’interactions. En rouge : vraie $h_{\ell}^{(m)}$. En vert : médiane a posteriori. En noir moyenne a posteriori (point par point)

Références

- Del Moral, P., Doucet, A., and Jasra, A. (2006). Sequential monte carlo samplers. Journal of the Royal Statistical Society : Series B (Statistical Methodology), 68(3) :411–436.
- Hansen, N. R., Reynaud-Bouret, P., and Rivoirard, V. (2014). Lasso and probabilistic inequalities for multivariate point processes. to appear in, Bernoulli, to appear.