

ESTIMATION PARAMÉTRIQUE DE SÉRIES DE HAWKES LOCALEMENT STATIONNAIRES

Felix Cheysson ^{1,2} & Gabriel Lang ² & Laurence Watier ¹

¹ *Institut Pasteur, 25-28 rue du Docteur Roux 75015 Paris, felix.cheysson@pasteur.fr*

² *AgroParisTech, 16 rue Claude Bernard 75005 Paris, felix.cheysson@agroparistech.fr*

Résumé. Les processus de Hawkes sont une famille de processus stochastiques pour lesquels l'occurrence d'un événement modifie temporairement la probabilité d'occurrence des événements futurs. Dans le cas où le comptage des événements est observé en temps discret, nous proposons une approche spectrale de l'estimation du processus de Hawkes, faisant appel au spectre de Bartlett et à l'approximation de la vraisemblance proposée par Whittle. Pour permettre l'analyse de données dont la structure de probabilité varie dans le temps, nous étendons ensuite l'approche au cadre de stationnarité locale introduit par Dahlhaus. Ces approches sont accompagnées par des jeux de simulations illustrant la qualité de l'estimation, en particulier celle de la fonction d'excitation du processus de Hawkes dans le cas où les fenêtres d'observations sont grandes.

Mots-clés. Processus de Hawkes, Spectre de Bartlett, Stationnarité locale, Séries temporelles

Abstract. Hawkes processes are a family of stochastic processes for which the occurrence of any event increases the probability of further events occurring. When count data are only observed in discrete time, we propose a spectral approach for the estimation of Hawkes processes, based on the Bartlett spectrum and the Whittle likelihood. To allow for nonstationarity in the data, we extend the approach to the local stationarity framework introduced by Dahlhaus. Simulated datasets illustrate the performances of the estimation, notably, of the Hawkes excitation function, even when the time between observations is large.

Keywords. Hawkes process, Bartlett spectrum, Local stationarity, Time series

1 Introduction

Les processus auto-excitants, éponymes de Hawkes (Hawkes, 1971a,b), forment une famille de processus stochastiques pour lesquels l'occurrence d'un événement modifie temporairement la probabilité d'occurrence d'événements futurs. D'abord d'application exclusive en sismologie (Adamopoulos, 1976; Ogata, 1988), leur utilisation s'est rapidement répandue à de nombreuses disciplines, comme la neurophysiologie (Chornoboy et al., 1988), la finance (Bacry et al., 2015), la génomique (Reynaud-Bouret and Schbath, 2010) et l'épidémiologie (Meyer et al., 2012).

Alors que l'estimation des processus de Hawkes stationnaires est aisée lorsque les temps d'occurrence des événements sont observés (Ogata, 1978, 1988; Ozaki and Ogata, 1979), elle reste problématique dans le cas où le comptage des événements est observé en temps discret. Le jeu de données se présente alors sous la forme d'une série temporelle à valeurs entières, dénombrant les occurrences du processus de Hawkes entre chaque point d'observation.

Kirchner (2016) en propose une estimation non paramétrique par approximation avec un INAR(∞) (*INteger AutoRegressive time series*; série autorégressive à valeurs entières) et montre la convergence de ce dernier vers le processus de Hawkes lorsque la taille des fenêtres d'observation tend vers zéro. Cette méthode est bien adaptée pour des événements fréquents, pour lesquels l'estimation par maximum de vraisemblance est computationnellement trop lourde (Kirchner, 2017). Dans ce cas, la taille des fenêtres est choisie selon un arbitrage entre le temps de calcul et la précision de l'estimation. Lorsque les fenêtres d'observation sont trop grandes, l'approximation par un processus INAR conduit à un biais d'estimation.

Par ailleurs, la littérature fait souvent appel aux algorithmes EM pour l'estimation de processus de Hawkes, *e.g.* lorsqu'ils sont multivariés (Olson and Carley, 2013) ou lorsque l'intensité de base est un processus de renouvellement (Wheatley et al., 2016), considérant généralement la structure de branchement comme variable latente. Dans le cadre d'un processus échantillonné à temps discret, l'approche EM consistant à considérer les temps d'arrivée comme variables latentes est malheureusement insoluble, car leur loi conditionnellement à la série de comptage n'a pas de forme explicite. Les algorithmes EM stochastiques (Celeux et al., 1995), approchant la loi conditionnelle pendant l'étape E, sont également mal adaptés, puisque les résultats usuels de convergence s'appuient sur des vraisemblances de famille exponentielle exclusivement (Delyon et al., 1999), condition non vérifiée pour les processus de Hawkes.

Nous proposons une approche spectrale de l'estimation de séries de Hawkes, en nous basant sur les travaux de Whittle (1953) et Hawkes (1971b). Nous établissons dans un premier temps la forme explicite du spectre d'une série de Hawkes, permettant d'écrire une vraisemblance de Whittle, et d'en faire une estimation paramétrique. Dans un deuxième temps, pour répondre aux besoins d'analyse de jeux de données dont la structure de probabilité varie dans le temps, nous nous intéressons à étendre cette approche aux séries non stationnaires.

2 Processus de Hawkes

L'intensité conditionnelle (parfois appelée fonction de risque) d'un processus ponctuel $N(t)$ est définie par

$$\lambda^*(t) = \frac{d \mathbb{E}[N(t) \mid \mathcal{F}_t]}{dt}$$

où $\mathcal{F}_t$ est la filtration naturelle associée à $N(t)$ et représente l'information disponible au temps t .

Le processus de Hawkes est un processus de Poisson doublement stochastique (ou *processus de Cox*) pour lequel l'intensité conditionnelle $\lambda^*(t)$ est de la forme

$$\begin{aligned}\lambda^*(t) &= \eta(t) + \int_0^t h(t-u)N(du) \\ &= \eta(t) + \sum_{t_j < t} h(t-t_j)\end{aligned}$$

où η, h sont des fonctions positives intégrables, $(t_j)_{j \in \mathbb{N}}$ sont les réalisations du processus et $\int \phi(u)N(du) = \sum \phi(t_j)$ désigne l'intégrale stochastique de Stieltjes. η est appelé l'intensité de base et h la fonction d'excitation.

Par la suite, on appellera *série de Hawkes* une série temporelle du type $\{X_t\}_{t \in \mathbb{R}} = \{N(t\Delta, (t+1)\Delta)\}_{t \in \mathbb{R}}$ ou $\{X_k\}_{k \in \mathbb{Z}} = \{N(k\Delta, (k+1)\Delta)\}_{k \in \mathbb{Z}}$, générée par le comptage du processus sur des intervalles de taille Δ (voir figure 1).

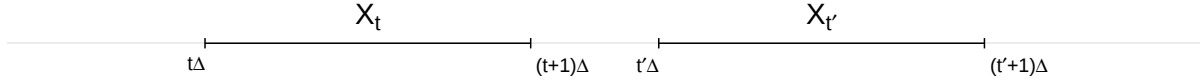


Figure 1: Une série de Hawkes $\{X_t\}_{t \in \mathbb{R}}$ avec un pas d'observation Δ .

3 Analyse spectrale

L'analyse spectrale des processus ponctuels fait appel à la notion de spectre de Bartlett (Bartlett, 1963) (i.e. la transformée de Fourier de l'autocovariance du processus), défini rigoureusement dans (Daley and Vere-Jones, 2003, Proposition 8.2.I) et rappelé ci-après dans la proposition suivante :

Proposition 1. *Soit N un processus stationnaire au second ordre sur $\mathbb{R}$, alors*

$$\text{Cov}(N(\varphi), N(\psi)) = \int_{\mathbb{R}} \tilde{\varphi}(\xi) \widetilde{\psi^*}(\xi) \Gamma(d\xi)$$

où φ et ψ sont des fonctions à décroissance rapide, $N(\phi) = \int_{\mathbb{R}} \phi(x)N(dx)$, $\psi^*(u) = \psi(-u)$, et $\widetilde{\cdot}$ désigne la transformée de Fourier : $\tilde{\varphi}(\xi) = \int_{\mathbb{R}} e^{i\xi u} \varphi(u) du$.

La mesure Γ est unique, symétrique, et est appelée le spectre de Bartlett du processus N .

Pour un processus de Hawkes stationnaire, le spectre de Bartlett admet une densité donnée par (Daley and Vere-Jones, 2003, Example 8.2(e))

$$\gamma(\xi) = \frac{m}{2\pi} |1 - H(\xi)|^{-2} \quad (1)$$

où $m = \mathbb{E}[N(0, 1]] = \frac{\eta}{1 - \int_{\mathbb{R}} h(t) dt}$ et $H(\xi) = \int_{\mathbb{R}} e^{i\xi u} h(u) du$.

Ces deux résultats permettent de donner une formulation explicite de la transformée de Fourier d'une série de Hawkes en temps continu :

Proposition 2. *Soit N un processus de Hawkes stationnaire au second ordre sur $\mathbb{R}$, et $\{X_t\}_{t \in \mathbb{R}} = \{N(t\Delta, (t+1)\Delta)\}_{t \in \mathbb{R}}$ la série temporelle associée. Alors, X_t admet pour densité spectrale la fonction*

$$\hat{\gamma}(\xi) = m \Delta \left(\frac{\sin(\xi/2)}{\xi/2} \right)^2 \left| 1 - H\left(\frac{\xi}{\Delta}\right) \right|^{-2} \quad (2)$$

Preuve. La preuve découle directement de la proposition 1 et de (1) en posant $\varphi = \mathbb{1}_{(t\Delta, (t+1)\Delta]}$ et $\psi = \mathbb{1}_{(t'\Delta, (t'+1)\Delta]}$. $\square$

Pour une série de Hawkes échantillonnée en temps discret, s'ajoute un effet de recouvrement spectral (ou *aliasing*) qui replie les composantes de haute fréquence sur un intervalle de basse fréquence :

Corollaire 1. *Soit N un processus de Hawkes stationnaire au second ordre sur $\mathbb{R}$, et $\{X_k\}_{k \in \mathbb{Z}} = \{N(k\Delta, (k+1)\Delta)\}_{k \in \mathbb{Z}}$ la série temporelle associée, échantillonnée avec un pas Δ . Alors, X_k admet pour densité spectrale la fonction*

$$\hat{\gamma}_1(\xi) = \sum_{k \in \mathbb{Z}} \hat{\gamma}(\xi + 2k\pi)$$

où $\hat{\gamma}(\cdot)$ est la fonction définie en (2).

4 Vraisemblance de Whittle

Pour une série stationnaire $\{X_k\}$ d'espérance nulle, Whittle (1953) introduit

$$\mathcal{L}_w(\theta) = \frac{1}{4\pi} \sum_{\omega \in \Omega} \left\{ \log 4\pi^2 f_\theta(\omega) + \frac{I_T(\omega)}{f_\theta(\omega)} \right\}$$

comme approximation de $-\frac{1}{T} \log$ vraisemblance Gaussienne, où $f_\theta(\cdot)$ est le spectre de la série $\{X_k\}_{1 \leq k \leq T}$, $I_T(\omega) = (2\pi T)^{-1} |\sum_{k=1}^T X_k \exp(-i\omega k)|^2$ son périodogramme et $\Omega = (2k\pi/T)_{0 \leq k < T}$. $\mathcal{L}_w(\theta)$ peut être interprétée, à une constante près, comme la distance entre la densité spectrale paramétrique $f_\theta(\omega)$ et l'estimation non paramétrique $I_T(\omega)$.

En posant $f_\theta = \hat{\gamma}_1$, l'estimation d'une série de Hawkes stationnaire découle de la minimisation de la vraisemblance de Whittle :

$$\hat{\theta} = \arg \min \mathcal{L}_w(\theta),$$

où $\hat{\theta}$ est un estimateur convergent et asymptotiquement Gaussien.

5 Processus non stationnaire

Pour un processus non stationnaire, la notion d'asymptotique change, puisque les observations futures ne donnent aucune information sur la structure de probabilité des données observées. Nous nous plaçons alors dans le cadre de stationnarité locale développé par Dahlhaus (1997), où l'asymptotique consiste à considérer que la structure locale du processus X_t est observée de plus en plus finement sur l'intervalle $[1, T]$.

Dans ce cadre, la vraisemblance de Whittle s'écrit

$$\mathcal{L}_w(\theta) = \frac{1}{4\pi} \sum_{t=1}^T \sum_{\omega \in \Omega} \left\{ \log 4\pi^2 f_\theta \left(\frac{t}{T}, \omega \right) + \frac{\tilde{I}_T(t/T, \omega)}{f_\theta(t/T, \omega)} \right\}$$

où $f_\theta(u, \cdot)$ est le spectre local au temps $t = uT$ avec $u \in [0, 1]$ et

$$\tilde{I}_T(u, \omega) = \frac{1}{2\pi} \sum_{k: 1 \leq [uT+0.5+k/2] \leq T} X_{[uT+0.5+k/2]} X_{[uT+0.5-k/2]} \exp(-i\omega k)$$

est le préperiodogramme de X_t , une version locale du periodogramme au temps $t = uT$.

Sous certaines conditions de régularité, l'estimateur résultant de la minimisation de $\mathcal{L}_w(\theta)$ est convergent et asymptotiquement Gaussien (Dahlhaus, 1997).

References

- Adamopoulos, L. (1976). Cluster models for earthquakes: Regional comparisons. *J. Int. Assoc. Math. Geol.*, 8(4):463–475.
- Bacry, E., Mastromatteo, I., and Muzy, J.-F. (2015). Hawkes Processes in Finance. *Mark. Microstruct. Liq.*, 01(01):1550005.
- Bartlett, M. S. (1963). The Spectral Analysis of Point Processes. *J. R. Stat. Soc. Ser. B*, 25(2):264–296.
- Celeux, G., Chauveau, D., and Diebolt, J. (1995). On Stochastic Versions of the EM Algorithm. Technical report, INRIA.
- Chornoboy, E. S., Schramm, L. P., and Karr, A. F. (1988). Maximum likelihood identification of neural point process systems. *Biol. Cybern.*, 59(4-5):265–275.

- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Stat.*, 25(1):1–37.
- Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods*. Probability and its Applications. Springer, New York.
- Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the EM algorithm. *Ann. Stat.*, 27(1):94–128.
- Hawkes, A. G. (1971a). Point Spectra of Some Mutually Exciting Point Processes. *J. R. Stat. Soc. Ser. B*, 33(3):438–443.
- Hawkes, A. G. (1971b). Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika*, 58(1):83–90.
- Kirchner, M. (2016). Hawkes and INAR(∞) processes. *Stoch. Process. their Appl.*, 126(8):2494–2525.
- Kirchner, M. (2017). An estimation procedure for the Hawkes process. *Quant. Financ.*, 17(4):571–595.
- Meyer, S., Elias, J., and Höhle, M. (2012). A Space-Time Conditional Intensity Model for Invasive Meningococcal Disease Occurrence. *Biometrics*, 68(2):607–616.
- Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann. Inst. Stat. Math.*, 30(1):243–261.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.*, 83(401):9–27.
- Olson, J. F. and Carley, K. M. (2013). Exact and approximate EM estimation of mutually exciting hawkes processes. *Stat. Inference Stoch. Process.*, 16(1):63–80.
- Ozaki, T. and Ogata, Y. (1979). Maximum likelihood estimation of Hawkes’ self-exciting point processes. *Ann. Inst. Stat. Math.*, 31(1):145–155.
- Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for hawkes processes; Application to genome analysis. *Ann. Stat.*, 38(5):2781–2822.
- Wheatley, S., Filimonov, V., and Sornette, D. (2016). The Hawkes process with renewal immigration & its estimation with an EM algorithm. *Comput. Stat. Data Anal.*, 94:120–135.
- Whittle, P. (1953). Estimation and information in stationary time series. *Ark. för Mat.*, 2(5):423–434.