Locally stationary Hawkes processes

François Roueff
LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay

Talk based on Roueff, von Sachs, and Sansonnet [2016]

Outline

Hawkes processes and applications

A Hawkes process is a self-exciting point process with linear conditional intensity (see Hawkes [1971]).

Hawkes processes and applications

A Hawkes process is a self-exciting point process with linear conditional intensity (see Hawkes [1971]).

Applications of this model include:

- seismology Ogata [1988],
- genomics Reynaud-Bouret and Schbath [2010],
- neuroscience Reynaud-Bouret et al. [2013],
- finance: microstructure dynamics for high-frequency data Bowsher [2007], Bacry et al. [2015, 2013]. . .

Preliminary remarks and conventions

- A point process N is identified with a random measure with discrete support: $N = \sum_k \delta_{T_k}$, where δ_t is the Dirac measure at point t.

Preliminary remarks and conventions

- A point process N is identified with a random measure with discrete support: $N = \sum_k \delta_{T_k}$, where δ_t is the Dirac measure at point t.
- We will consider point processes on $\mathbb{R}^\ell$ (often with $\ell = 1$).

Preliminary remarks and conventions

- A point process N is identified with a random measure with discrete support: $N = \sum_k \delta_{T_k}$, where δ_t is the Dirac measure at point t.
- We will consider point processes on $\mathbb{R}^\ell$ (often with $\ell = 1$).
- We do not consider marked or multivariate point processes.

Preliminary remarks and conventions

- A point process N is identified with a random measure with discrete support: $N = \sum_k \delta_{T_k}$, where δ_t is the Dirac measure at point t.
- We will consider point processes on $\mathbb{R}^\ell$ (often with $\ell = 1$).
- We do not consider marked or multivariate point processes.
- For a test function g, we denote
$$N(g) = \int g \,\mathrm{d}N = \sum_k g(T_k).$$

Preliminary remarks and conventions

- A point process N is identified with a random measure with discrete support: $N = \sum_k \delta_{T_k}$, where δ_t is the Dirac measure at point t.
- We will consider point processes on $\mathbb{R}^\ell$ (often with $\ell = 1$).
- We do not consider marked or multivariate point processes.
- For a test function g, we denote
  $$N(g) = \int g \,\mathrm{d}N = \sum_k g(T_k).$$
- (Some) notation on functional norms:
  - The L^q-norm of g is denoted by $|g|_q$ for $q \in [1;\infty]$, and

Preliminary remarks and conventions

- A point process N is identified with a random measure with discrete support: $N = \sum_k \delta_{T_k}$, where δ_t is the Dirac measure at point t.
- We will consider point processes on $\mathbb{R}^\ell$ (often with $\ell = 1$).
- We do not consider marked or multivariate point processes.
- For a test function g, we denote
  $$N(g) = \int g \, \mathrm{d}N = \sum_k g(T_k).$$
- (Some) notation on functional norms:
  - The L^q-norm of g is denoted by $|g|_q$ for $q \in [1;\infty]$, and
  - We will use a polynomial weighted L^1 norm depending on some $\beta > 0$, $|h|_{(\beta)} := \left|h \times |\cdot|^\beta\right|_1 = \int |h(s)|\,|s|^\beta \,\mathrm{d}s$.

Stationary Hawkes processes

- A stationary Hawkes process on $\mathbb{R}$ is often defined as a point process $N = \sum_k \delta_{T_k}$ on the real line with conditional intensity

$$\lambda(t) = \lambda_c + \int_{-\infty}^{t^-} p(t-s)\, N(\mathrm{d}s) = \lambda_c + \sum_{T_i < t} p(t - T_i)\ , \quad (1)$$

where $\lambda_c > 0$ and $p : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\int p < 1$.

Stationary Hawkes processes

- A stationary Hawkes process on $\mathbb{R}$ is often defined as a point process $N = \sum_k \delta_{T_k}$ on the real line with conditional intensity

$$\lambda(t) = \lambda_c + \int_{-\infty}^{t^-} p(t-s)\, N(\mathrm{d}s) = \lambda_c + \sum_{T_i < t} p(t - T_i)\ , \quad (1)$$

  where $\lambda_c > 0$ and $p : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\int p < 1$.
- A description as a Cluster process (Hawkes and Oakes [1974]) can be used to define a more general class of spatial stationary Hawkes processes in $\mathbb{R}^\ell$:

Stationary Hawkes processes

- A stationary Hawkes process on $\mathbb{R}$ is often defined as a point process $N = \sum_k \delta_{T_k}$ on the real line with conditional intensity

$$\lambda(t) = \lambda_c + \int_{-\infty}^{t^-} p(t-s)\, N(\mathrm{d}s) = \lambda_c + \sum_{T_i < t} p(t - T_i)\ , \quad (1)$$

  where $\lambda_c > 0$ and $p : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\int p < 1$.
- A description as a Cluster process (Hawkes and Oakes [1974]) can be used to define a more general class of spatial stationary Hawkes processes in $\mathbb{R}^\ell$:
  - Start with a center process $N_c = \sum_k \delta_{t_k^c}$, taken as a Poisson point process (PPP) with intensity λ_c.

Stationary Hawkes processes

- A stationary Hawkes process on $\mathbb{R}$ is often defined as a point process $N = \sum_k \delta_{T_k}$ on the real line with conditional intensity

$$\lambda(t) = \lambda_c + \int_{-\infty}^{t^-} p(t-s)\, N(\mathrm{d}s) = \lambda_c + \sum_{T_i<t} p(t-T_i)\ , \quad (1)$$

  where $\lambda_c > 0$ and $p : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\int p < 1$.
- A description as a Cluster process (Hawkes and Oakes [1974]) can be used to define a more general class of spatial stationary Hawkes processes in $\mathbb{R}^\ell$:
  - Start with a center process $N_c = \sum_k \delta_{t_k^c}$, taken as a Poisson point process (PPP) with intensity λ_c.
  - Given N_c, generate independent component processes or clusters $N(\cdot|t_k^c)$ with descendants propagating through a fertility measure μ,

Stationary Hawkes processes

- A stationary Hawkes process on $\mathbb{R}$ is often defined as a point process $N = \sum_k \delta_{T_k}$ on the real line with conditional intensity

$$\lambda(t) = \lambda_c + \int_{-\infty}^{t^-} p(t-s)\, N(\mathrm{d}s) = \lambda_c + \sum_{T_i < t} p(t - T_i)\ , \quad (1)$$

where $\lambda_c > 0$ and $p : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\int p < 1$.
- A description as a Cluster process (Hawkes and Oakes [1974]) can be used to define a more general class of spatial stationary Hawkes processes in $\mathbb{R}^\ell$:
  - Start with a center process $N_c = \sum_k \delta_{t_k^c}$, taken as a Poisson point process (PPP) with intensity λ_c.
  - Given N_c, generate independent component processes or clusters $N(\cdot|t_k^c)$ with descendants propagating through a fertility measure μ,
  - The superposition $N = \int N(\cdot|t^c)\, N_c(\mathrm{d}t^c) = \sum_k N(\cdot|t_k^c)$ defines a spatial Hawkes process .

Stationary Hawkes processes : generation of clusters

- ▷ Offspring process: given a location $t \in \mathbb{R}^\ell$, we let $M(\cdot|t)$ be a PPP with control measure $\mu(\cdot - t)$.

Stationary Hawkes processes : generation of clusters

- Offspring process: given a location $t \in \mathbb{R}^\ell$, we let $M(\cdot|t)$ be a PPP with control measure $\mu(\cdot - t)$.
- Define successive generations of descendants as follows :

Stationary Hawkes processes : generation of clusters

- Offspring process: given a location $t \in \mathbb{R}^\ell$, we let $M(\cdot|t)$ be a PPP with control measure $\mu(\cdot - t)$.
- Define successive generations of descendants as follows :
  Generation 0 : $N^{(0)}(\cdot|t^c) = \delta_{t^c}$.

Stationary Hawkes processes : generation of clusters

- Offspring process: given a location $t \in \mathbb{R}^\ell$, we let $M(\cdot|t)$ be a PPP with control measure $\mu(\cdot - t)$.
- Define successive generations of descendants as follows :

  Generation 0 : $N^{(0)}(\cdot|t^c) = \delta_{t^c}$.
  Generation 1 : $N^{(1)}(\cdot|t^c) = M(\cdot|t^c)$.

Stationary Hawkes processes : generation of clusters

- Offspring process: given a location $t \in \mathbb{R}^\ell$, we let $M(\cdot|t)$ be a PPP with control measure $\mu(\cdot - t)$.
- Define successive generations of descendants as follows :

  Generation 0 : $N^{(0)}(\cdot|t^c) = \delta_{t^c}$.
  Generation 1 : $N^{(1)}(\cdot|t^c) = M(\cdot|t^c)$.

$$\vdots$$

  Generation n+1 : Denoting $N^{(n)}(\cdot|t^c) = \sum_i \delta_{t_i^{(n)}}$,

$$N^{(n+1)}(\cdot|t^c) = \int M(\cdot|s)\, N^{(n)}(\mathrm{d}s|t^c) = \sum_i M(\cdot|t_i^{(n)})$$

$$\vdots$$

Stationary Hawkes processes : generation of clusters

- Offspring process: given a location $t \in \mathbb{R}^\ell$, we let $M(\cdot|t)$ be a PPP with control measure $\mu(\cdot - t)$.
- Define successive generations of descendants as follows :

  Generation 0 : $N^{(0)}(\cdot|t^c) = \delta_{t^c}$.
  Generation 1 : $N^{(1)}(\cdot|t^c) = M(\cdot|t^c)$.

  $$\vdots$$

  Generation n+1 : Denoting $N^{(n)}(\cdot|t^c) = \sum_i \delta_{t_i^{(n)}}$,

  $$N^{(n+1)}(\cdot|t^c) = \int M(\cdot|s)\, N^{(n)}(\mathrm{d}s|t^c) = \sum_i M(\cdot|t_i^{(n)})$$

  $$\vdots$$

- and finally set $N(\cdot|t^c) = \sum_{k \geqslant 0} N^{(k)}(\cdot|t^c)$.

Stationary Hawkes processes

- Each cluster $N(\cdot|t^c)$ has a finite mean number of points if and only if $\int \mathrm{d}\mu < 1$.
- The distribution of $N(\cdot - t^c|t^c)$ does not depend on t^c.
- Since N_c is a homogeneous PPP, the resulting process N is a stationary finite intensity point process.
- The conditional intensity $\lambda(t)$ in (1) is obtained by setting $\mu(\mathrm{d}t) = p(t)\mathrm{d}t$.

First and second order moments of stationary Hawkes processes

- The mean intensity m_1 of N is given by $m_1 = \dfrac{\lambda_c}{1 - \int \mathrm{d}\mu}$.
- The Bartlett spectrum of N defined by

$$\mathrm{Cov}\left(N(g_1), N(g_2)\right) = \int \hat{g}_1(\omega)\overline{\hat{g}_2(\omega)}\ \Gamma(\mathrm{d}\omega)$$

with

$$\hat{g}_j(\omega) = \int g_j(t)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}t\ , \quad j = 1, 2\ ,$$

is given by

$$\Gamma(\mathrm{d}\omega) = \frac{m_1}{2\pi}\left|1 - \int \mathrm{e}^{-\mathrm{i}t\omega}\ \mu(\mathrm{d}t)\right|^{-2}\ \mathrm{d}\omega$$

A parametric example

Consider a specific class of Gamma-shaped fertility functions p, depending on parameters $\zeta, \delta, \eta, \theta > 0$

$$p(t) = \zeta \, \frac{\theta^\eta}{\Gamma(\eta)} \, (t - \delta)_+^{\eta - 1} \, \mathrm{e}^{-\theta(t-\delta)} \; .$$

Then the stationarity assumption simply reads

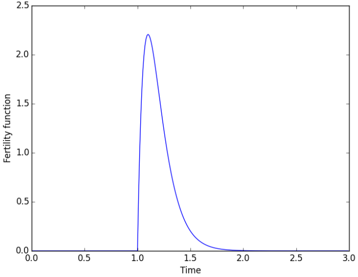
$$\int p = \zeta < 1 \; .$$

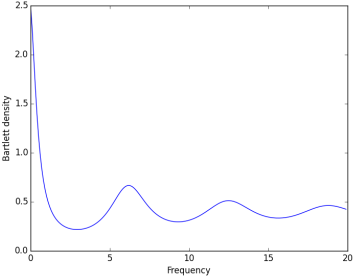
Moreover,

$$\hat{p}(\omega) = \int p(t) \, \mathrm{e}^{-\mathrm{i} t \omega} \, \mathrm{d}t = \zeta \mathrm{e}^{-\mathrm{i}\delta\omega} \, (1 + \mathrm{i}\omega/\theta)^{-\eta} \; .$$

Bartlett spectrum of a Hawkes process with Gamma shape fertility

A positive δ induces a periodic phenomenon in the self-excitation.
Here we take $\zeta = 0.6$, $\delta = 1$, $\eta = 2$, $\theta = 10$.





Non-stationary modeling

- In many applications, it is interesting to investigate whether the statistical behavior evolves in time or space.

Non-stationary modeling

- In many applications, it is interesting to investigate whether the statistical behavior evolves in time or space.
- For instance, one can think of a Gamma shape fertility with time-varying parameters.

Non-stationary modeling

- In many applications, it is interesting to investigate whether the statistical behavior evolves in time or space.
- For instance, one can think of a Gamma shape fertility with time-varying parameters.
- In the latter case, the Bartlett spectrum should also evolves along the time, resulting in a time-frequency analysis.

Non-stationary modeling

- In many applications, it is interesting to investigate whether the statistical behavior evolves in time or space.
- For instance, one can think of a Gamma shape fertility with time-varying parameters.
- In the latter case, the Bartlett spectrum should also evolves along the time, resulting in a time-frequency analysis.
- Our goal is to provide a reasonable modeling approach for a comprehensive statistical analysis in this non-stationary context.

Non-stationary Hawkes processes, using conditional intensity

- Non-stationary Hawkes processes often refer to a process initiated in the empty state at the origin,

$$\lambda(t) = \lambda_c + \int_{\underline{0}}^{t^-} p(t-s)\, N(\mathrm{d}s)\,, \quad t \geqslant 0\,.$$

- In Chen and Hall [2013], additional non-stationarity is introduced by letting λ_c depend on time,

$$\lambda(t) = \lambda_c\underline{(t)} + \int_0^{t^-} p(t-s)\, N(\mathrm{d}s)\,, \quad t \geqslant 0\,.$$

- We propose the more general form

$$\lambda(t) = \lambda_c\underline{(t)} + \int_{-\infty}^{t^-} p(t-s\underline{;\,t})\, N(\mathrm{d}s)\,.$$

Non-stationary Hawkes processes, using clusters

Similarly, in the cluster construction, one can

- Replace the constant baseline intensity λ_c of N_c by a baseline intensity function $t \mapsto \lambda_c(t)$.
- Replace the control measure $\mu(\cdot - s)$ of the offspring process $M(\cdot|s)$ by "any" measure $\mu(\cdot|s)$.

Non-stationary Hawkes processes, using clusters

Similarly, in the cluster construction, one can

- Replace the constant baseline intensity λ_c of N_c by a baseline intensity function $t \mapsto \lambda_c(t)$.
- Replace the control measure $\mu(\cdot - s)$ of the offspring process $M(\cdot|s)$ by "any" measure $\mu(\cdot|s)$.

Question

What should replace the condition

$$\int \mu < 1$$

used to guaranty the existence of a finite mean intensity stationary Hawkes process ?

Non-stationary Hawkes processes, stability

Consider the two following conditions :

(C-1) $\zeta_1 = \sup_s \int \mu(\mathrm{d}t|s) < 1.$

(C-2) $\int \mu(\cdot|s)\mathrm{d}s$ admits a density bounded by $\zeta_1 < 1$.

Then if $\mu(\cdot|s) = \mu(\cdot - s)$, they both are equivalent to $\int \mathrm{d}\mu < 1$.

Non-stationary Hawkes processes, stability

Consider the two following conditions :

(C-1) $\zeta_1 = \sup_s \int \mu(\mathrm{d}t|s) < 1.$

(C-2) $\int \mu(\cdot|s)\mathrm{d}s$ admits a density bounded by $\zeta_1 < 1$.

Then if $\mu(\cdot|s) = \mu(\cdot - s)$, they both are equivalent to $\int \mathrm{d}\mu < 1$.

It turns out that

- (C-1) implies that $\mathbb{E} \int N(\mathrm{d}s|t^c) \leqslant (1 - \zeta_1)^{-1}$ (each cluster has a uniformly bounded finite mean number of points).

Non-stationary Hawkes processes, stability

Consider the two following conditions :

(C-1) $\zeta_1 = \sup_s \int \mu(\mathrm{d}t|s) < 1.$

(C-2) $\int \mu(\cdot|s)\mathrm{d}s$ admits a density bounded by $\zeta_1 < 1$.

Then if $\mu(\cdot|s) = \mu(\cdot - s)$, they both are equivalent to $\int \mathrm{d}\mu < 1$.

It turns out that

- (C-1) implies that $\mathbb{E} \int N(\mathrm{d}s|t^c) \leqslant (1 - \zeta_1)^{-1}$ (each cluster has a uniformly bounded finite mean number of points).
- (C-2) and $|\lambda_c|_\infty < \infty$ imply that N admits a finite mean intensity bounded by $|\lambda_c|_\infty / (1 - \zeta_1)$.

Non-stationary Hawkes processes, density assumption

Suppose that the measure $\mu(\cdot|s)$ admits a density, written as

$$t \mapsto d(t-s;s) \; .$$

Then (C-1) reads $\zeta_1 := \sup_s \int d(r;s)\mathrm{d}r < 1$.

Non-stationary Hawkes processes, density assumption

Suppose that the measure $\mu(\cdot|s)$ admits a density, written as

$$t \mapsto d(t-s;s) =: p(t-s;t) \; .$$

Then (C-1) reads $\zeta_1 := \sup_s \int d(r;s)\mathrm{d}r < 1$.

While (C-2) reads $\zeta_1 := \sup_t \int p(r;t)\mathrm{d}r < 1$.

Non-stationary Hawkes processes, density assumption

Suppose that the measure $\mu(\cdot|s)$ admits a density, written as

$$t \mapsto d(t-s;s) =: p(t-s;t) \ .$$

Then ▸ (C-1) reads $\zeta_1 := \sup_s \int d(r;s)\mathrm{d}r < 1$.

While ▸ (C-2) reads $\zeta_1 := \sup_t \int p(r;t)\mathrm{d}r < 1$.

Definition (Non-stationary Hawkes processes)

We call a non-stationary Hawkes process a process defined as N above under the condition

$$\text{(NS-1)} \quad \zeta_1 := \sup_t \int p(r;t)\mathrm{d}r < 1 \quad \text{and} \quad |\lambda_c|_\infty < \infty \ ,$$

which implies that N has finite mean intensity bounded by $|\lambda_c|_\infty / (1-\zeta_1)$.

Introduction

Non-stationary Hawkes processes

Locally stationary Hawkes processes

Numerical experiments

References

Parallel with time series

A Hawkes process with fertility function p is often paralleled with an autoregressive (AR) time series with parameter θ, since in both cases, conditional intensity/expectation are linear functions of the past, resulting in similar Bartlett/spectral density formula

$$\frac{m_1}{2\pi}\left|1-\int p(t)\,\mathrm{e}^{-\mathrm{i}t\omega}\,\mathrm{d}t\right|^{-2} \quad / \quad \frac{\sigma^2}{2\pi}\left|1-\sum_k \theta_k \mathrm{e}^{-\mathrm{i}\lambda k}\right|^{-2}$$

(here expressed as functions of ω and λ, resp.)

Parallel with time series

A Hawkes process with fertility function p is often paralleled with an autoregressive (AR) time series with parameter θ, since in both cases, conditional intensity/expectation are linear functions of the past, resulting in similar Bartlett/spectral density formula

$$\frac{m_1}{2\pi}\left|1-\int p(t)\,\mathrm{e}^{-\mathrm{i}t\omega}\,\mathrm{d}t\right|^{-2} \quad / \quad \frac{\sigma^2}{2\pi}\left|1-\sum_k \theta_k \mathrm{e}^{-\mathrm{i}\lambda k}\right|^{-2}$$

(here expressed as functions of ω and λ, resp.)

We can draw the same parallel between the previously introduced non-stationary Hawkes process and time-varying AR (TVAR) processes defined by the equation

$$X_t = \mu(t) + \sum_k \theta_k(t) X_{t-k} + \sigma(t)\epsilon_t \,.$$

Asymptotic setting: locally stationary TVAR processes

We propose to follow the approach used in Dahlhaus [1996] for time series, using rescaled time-varying parameters σ, μ and θ,

$$X_t = \mu(t) + \sum_k \theta_k(t) X_{t-k} + \sigma(t) \epsilon_t .$$

Asymptotic setting: locally stationary TVAR processes

We propose to follow the approach used in Dahlhaus [1996] for time series, using rescaled time-varying parameters σ, μ and θ,

$$X_{t,T} = \mu(t/T) + \sum_k \theta_k(t/T) X_{t-k,T} + \sigma(t/T)\epsilon_t \ .$$

We thus get a collection of processes $(X_{t,T})_{t \in \mathbb{Z}}$, $T > 0$.

Asymptotic setting: locally stationary TVAR processes

We propose to follow the approach used in Dahlhaus [1996] for time series, using rescaled time-varying parameters σ, μ and θ,

$$X_{t,T} = \mu(t/T) + \sum_k \theta_k(t/T) X_{t-k,T} + \sigma(t/T)\epsilon_t \ .$$

We thus get a collection of processes $(X_{t,T})_{t\in\mathbb{Z}}$, $T > 0$.

Two important consequences as $T \to \infty$

- A T-sample $X_{1,T}, \ldots, X_{T,T}$ basically involves the parameter function $u \mapsto (\sigma(u), \mu(u), \theta(u))$ on a fixed interval $u \in [0, 1]$, allowing us for a consistent estimation of these parameters.
- If $u \mapsto (\sigma(u), \mu(u), \theta(u))$ is smooth, a subsample $X_{t,T}$ with indices t such that $t/T \simeq u \in (0, 1)$, can be approximated by a stationary AR process with parameter $(\sigma(u), \mu(u), \theta(u))$.

Asymptotic setting: locally stationary Hawkes processes

Definition (Locally stationary Hawkes processes)

A locally stationary Hawkes process with

- local baseline intensity $\lambda_c^{<\mathsf{LS}>} : \mathbb{R}^\ell \to \mathbb{R}_+$ and
- local fertility function $p^{<\mathsf{LS}>} : \mathbb{R}^\ell \times \mathbb{R}^\ell \to \mathbb{R}_+$

is a collection $(N_T)_{T>0}$ such that, for all $T > 0$, N_T is a non-stationary Hawkes processes with baseline intensity $t \mapsto \lambda_c^{<\mathsf{LS}>}(t/T)$ and fertility function $(r,t) \mapsto p^{<\mathsf{LS}>}(r;t/T)$.

Asymptotic setting: locally stationary Hawkes processes

Definition (Locally stationary Hawkes processes)

A locally stationary Hawkes process with

- local baseline intensity $\lambda_c^{<\text{LS}>} : \mathbb{R}^\ell \to \mathbb{R}_+$ and
- local fertility function $p^{<\text{LS}>} : \mathbb{R}^\ell \times \mathbb{R}^\ell \to \mathbb{R}_+$

is a collection $(N_T)_{T>0}$ such that, for all $T > 0$, N_T is a non-stationary Hawkes processes with baseline intensity $t \mapsto \lambda_c^{<\text{LS}>}(\underline{t/T})$ and fertility function $(r, t) \mapsto p^{<\text{LS}>}(r; \underline{t/T})$.

- For a given real location t, the scaled location $u = t/T$ is typically called an absolute location.
- At each u, we denote by $N(\cdot; u)$ a stationary Hawakes process with baseline intensity $\lambda_c^{<\text{LS}>}(u)$ and fertility function $p^{<\text{LS}>}(\cdot; u)$.

Assumptions on local baseline intensity and local fertility function

Condition ▸ (NS-1) becomes

$$\text{(LS-1)} \quad \zeta_1^{<\text{LS}>} := \sup_u \int p^{<\text{LS}>}(r; u)\, \mathrm{d}r < 1 \quad \text{and} \quad \left|\lambda_c^{<\text{LS}>}\right|_\infty < \infty\,.$$

Note that it does not involve T.

Assumptions on local baseline intensity and local fertility function

Condition ▸ (NS-1) becomes

$$\text{(LS-1)} \quad \zeta_1^{<\mathsf{LS}>} := \sup_u \int p^{<\mathsf{LS}>}(r;u)\,\mathrm{d}r < 1 \quad \text{and} \quad \left|\lambda_c^{<\mathsf{LS}>}\right|_\infty < \infty\,.$$

Note that it does not involve T.

More regularity assumptions:

- (LS-2): β-Hölder type smoothness condition on $\lambda_c^{<\mathsf{LS}>}(u)$
- (LS-3): β-Hölder type smoothness conditions on $p^{<\mathsf{LS}>}(t,u)$ w.r.t. its second argument u.
- (LS-4): some β-power decay condition on $p^{<\mathsf{LS}>}(\cdot;u)$ uniformly bounded w.r.t. u.

Laplace functional

We use Laplace functionals to show that, for a given absolute location u, as $T \to \infty$, in the neighborhood of Tu,

$$N_T \text{ can be approximated by } N(\cdot; u).$$

For all $T > 0$ and $u \in \mathbb{R}^\ell$, letting g denote some test function,

- the Laplace functional of N_T is denoted by

$$\mathcal{L}_T(g) = \mathbb{E}\left[\exp N_T(g)\right] .$$

- the Laplace functional of $N(\cdot; u)$ is denoted by

$$\mathcal{L}(g; u) = \mathbb{E}\left[\exp N(g; u)\right] .$$

Hence we compare $\mathcal{L}_T \circ S^{-Tu}$ with $\mathcal{L}(\cdot; u)$, where

$$S^{-Tu} g : t \mapsto g(t - Tu) .$$

Approximation results

- A typical result is that if $|g|_1, |g|_\infty, |g|_{(\beta)} \leqslant 1$, then

$$\left|\mathcal{L}_T \circ S^{-Tu}(g) - \mathcal{L}(g;u)\right| = O\left(T^{-\beta}\right) ,$$

where constants in O-terms only depend on $\lambda_c^{<\mathsf{LS}>}$ and $p^{<\mathsf{LS}>}$.

Approximation results

- A typical result is that if $|g|_1, |g|_\infty, |g|_{(\beta)} \leqslant 1$, then

$$\left|\mathcal{L}_T \circ S^{-Tu}(g) - \mathcal{L}(g;u)\right| = O\left(T^{-\beta}\right) ,$$

  where constants in O-terms only depend on $\lambda_c^{<\mathrm{LS}>}$ and $p^{<\mathrm{LS}>}$.
- In fact, under suitable conditions, the result holds uniformly in z for functions $g = g(\cdot; z)$ which are holomorphic w.r.t. z, see [Roueff et al., 2016, Theorem 2].

Approximation results

- A typical result is that if $|g|_1, |g|_\infty, |g|_{(\beta)} \leqslant 1$, then

$$\left|\mathcal{L}_T \circ S^{-Tu}(g) - \mathcal{L}(g; u)\right| = O\left(T^{-\beta}\right) ,$$

where constants in O-terms only depend on $\lambda_c^{<\text{LS}>}$ and $p^{<\text{LS}>}$.

- In fact, under suitable conditions, the result holds uniformly in z for functions $g = g(\cdot; z)$ which are holomorphic w.r.t. z, see [Roueff et al., 2016, Theorem 2].
- Consequently, we obtain, if $|g_j|_1, |g_j|_\infty, |g_j|_{(\beta)} \leqslant 1$ for $j = 1, \ldots, m$, then

$$\left|\mathrm{Cum}\left(\{N_T \circ S^{-Tu}(g_j)\}_{1\leqslant j\leqslant m}\right) - \mathrm{Cum}\left(\{N(g_j; u)\}_{1\leqslant j\leqslant m}\right)\right| = O\left(T^{-\beta}\right).$$

Example of application: approximation of local mean density

- For all $T > 0$ and $t \in \mathbb{R}^\ell$, let $m_{1\,T}(t)$ denote the mean density function of N_T.
- For all $u \in \mathbb{R}^\ell$ Let $m_1^{<\mathsf{LS}>}(u)$ denote the local mean intensity, that is, the mean intensity of $N(\cdot; u)$, which is given by

$$m_1^{<\mathsf{LS}>}(u) = \frac{\lambda_c^{<\mathsf{LS}>}(u)}{1 - \int p^{<\mathsf{LS}>}(\cdot; u)} .$$

- The previous result with $m = 1$ implies:

$$\sup_{t\,:|t - Tu| \leqslant b} \left| m_{1\,T}(t) - m_1^{<\mathsf{LS}>}(u) \right| = O\left((1 + b^\beta) T^{-\beta} \right)$$

Another application : time-frequency analysis

Suppose that $\ell = 1$. We can define and approximate the local Bartlett spectrum as follows :

- For all $u \in \mathbb{R}^\ell$, let $\gamma^{<\mathsf{LS}>}(\omega; u)$ denote the local Bartlett density, that is, the Bartlett density of $N(\cdot; u)$ which is

$$\gamma^{<\mathsf{LS}>}(\omega; u) = \frac{m_1^{<\mathsf{LS}>}(u)}{2\pi} \left|1 - \widehat{p^{<\mathsf{LS}>}}(\omega; u)\right|^{-2}$$

  where

$$\widehat{p^{<\mathsf{LS}>}}(\omega; u) = \int p^{<\mathsf{LS}>}(t; u)\, \mathrm{e}^{-\mathrm{i}\omega t}\, \mathrm{d}t$$

- The cumulent approximation result with $m = 2$ implies:

$$\mathrm{Var}\left(N_T(S^{-Tu} g)\right) = \int |\hat{g}(\omega)|^2 \gamma^{<\mathsf{LS}>}(\omega; u)\, \mathrm{d}\omega + O(T^{-\beta})$$

- Kernel estimation of $\gamma^{<\mathsf{LS}>}(\omega; u)$ can thus be achieved by an empirical estimate of $\mathrm{Var}\left(N_T(S^{-Tu} g)\right)$ for a well chosen g.

Simulation of a locally stationary Hawkes process

If $\ell = 1$ and $p^{<\mathsf{LS}>}(\cdot; u)$ is supported on $\mathbb{R}_+$ for all u, we can use that N_T has conditional intensity given by

$$\lambda_T(t) = \lambda_c^{<\mathsf{LS}>}(t/T) + \int p^{<\mathsf{LS}>}(t-s; t/T)\, N_T(\mathrm{d}s)\ .$$

Use Ogata's modified thinning algorithm Ogata [1981] to simulate the non-stationary Hawkes process N_T on the interval $[0, T]$.

Simulated examples

- ▷ We take a time-constant baseline intensity $\lambda_c^{<\mathrm{LS}>}(u) = 1/2$.
- ▷ The local fertility function $p^{<\mathrm{LS}>}(\cdot; u)$ is set as a Gamma-shaped function with time varying parameters.

Simulated examples

- We take a time-constant baseline intensity $\lambda_c^{<\mathsf{LS}>}(u) = 1/2$.
- The local fertility function $p^{<\mathsf{LS}>}(\cdot; u)$ is set as a Gamma-shaped function with time varying parameters.

- Example 1 [Time varying scale $\theta(u)$ and overall fertility $\zeta(u)$]:

$$p^{<\mathsf{LS}>}(s; u) = \zeta(u)\ \theta(u)\mathrm{e}^{-\theta(u)s}\mathbb{1}_{s>0}\ ,$$

with $\zeta(u)$, $\theta(u)$ of cosine form.

Simulated examples

- We take a time-constant baseline intensity $\lambda_c^{<\mathsf{LS}>}(u) = 1/2$.
- The local fertility function $p^{<\mathsf{LS}>}(\cdot; u)$ is set as a Gamma-shaped function with time varying parameters.

- Example 1 [Time varying scale $\theta(u)$ and overall fertility $\zeta(u)$]:

$$p^{<\mathsf{LS}>}(s; u) = \zeta(u)\ \theta(u)\mathrm{e}^{-\theta(u)s}\mathbb{1}_{s>0}\ ,$$

  with $\zeta(u)$, $\theta(u)$ of cosine form.
- Example 2 [Time varying delay $\delta(u)$]:

$$p^{<\mathsf{LS}>}(s; u) = \frac{1}{2}(s - \delta(u))_+\mathrm{e}^{-(s-\delta(u))}$$

  with $\delta(u) = (6 - 10u) \times \mathbb{1}_{[0;1/2]}(u) + (10u - 4) \times \mathbb{1}_{(1/2;1]}(u)$ inducing a periodic phenomenon in the self-excitation.

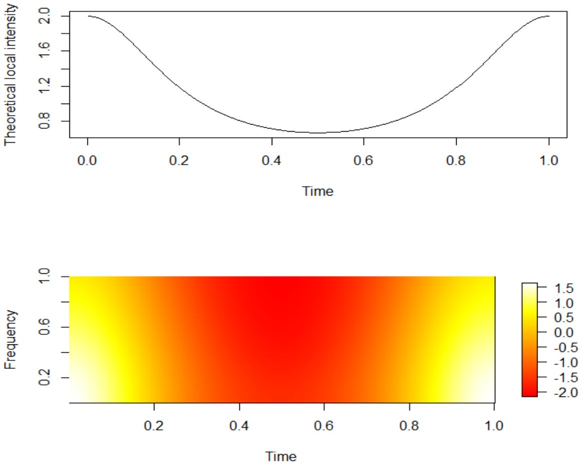


Figure: *Theoretical local mean density (top) and Bartlett spectrum (bottom) for Example 1.*

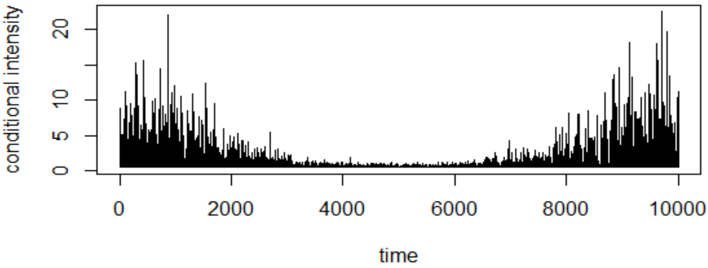


Figure: *Conditional intensity function of a simulated Hawkes process following Example 1, with* $T = 10000$.

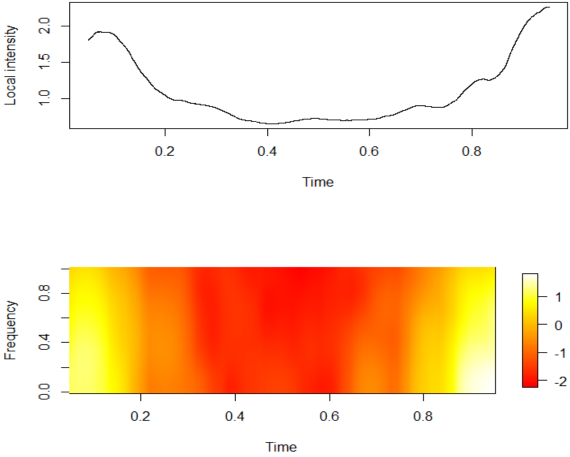


Figure: *Estimation of the local mean density (top) and of the local Bartlett spectrum (bottom) for Example 1.*

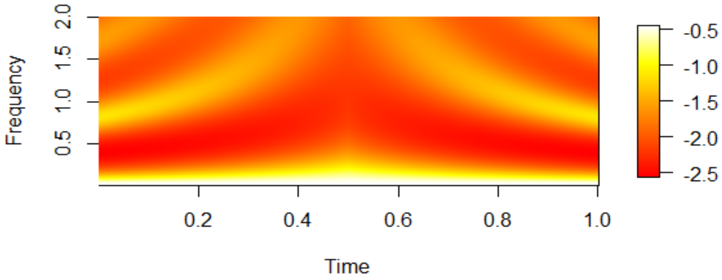


Figure: *Theoretical local Bartlett spectrum for Example 2 (local mean density being constant over time).*

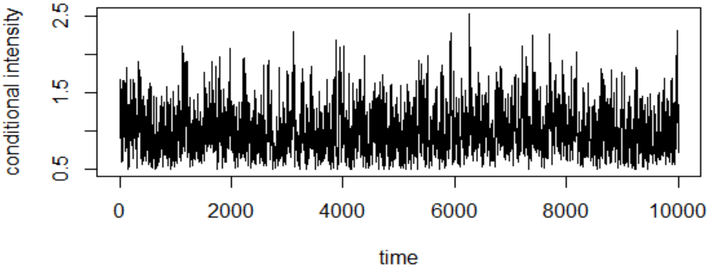


Figure: *Conditional intensity function of a simulated Hawkes process following Example 2, with* $T = 10000$.

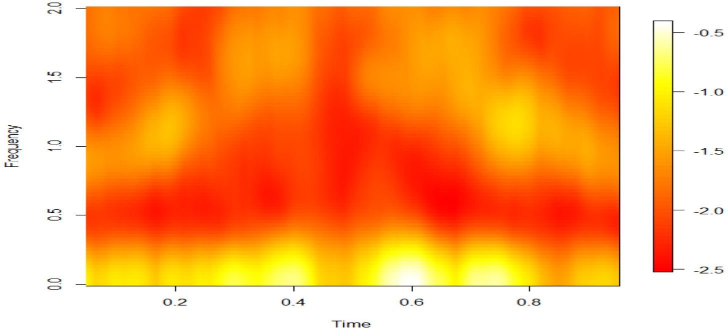


Figure: *Estimation of the local Bartlett spectrum for Example 2.*

Conclusion

- Self-exciting point processes ("Hawkes" processes) are somewhat similar to autoregressive processes in time series.

Conclusion

- Self-exciting point processes ("Hawkes" processes) are somewhat similar to autoregressive processes in time series.
- Unfortunately, unlike linear locally stationary time series, we cannot rely on a simple representation such as the TVMA(∞) (or spectral representation).

Conclusion

- Self-exciting point processes ("Hawkes" processes) are somewhat similar to autoregressive processes in time series.
- Unfortunately, unlike linear locally stationary time series, we cannot rely on a simple representation such as the TVMA(∞) (or spectral representation).
- Local stationary approximations can still be obtained using local Laplace transforms and its derivatives (control of cumulants).

Conclusion

- Self-exciting point processes ("Hawkes" processes) are somewhat similar to autoregressive processes in time series.
- Unfortunately, unlike linear locally stationary time series, we cannot rely on a simple representation such as the TVMA(∞) (or spectral representation).
- Local stationary approximations can still be obtained using local Laplace transforms and its derivatives (control of cumulants).
- Definitions of (local) mean intensity and a local Bartlett spectrum follow.

Conclusion

- Self-exciting point processes ("Hawkes" processes) are somewhat similar to autoregressive processes in time series.
- Unfortunately, unlike linear locally stationary time series, we cannot rely on a simple representation such as the TVMA(∞) (or spectral representation).
- Local stationary approximations can still be obtained using local Laplace transforms and its derivatives (control of cumulants).
- Definitions of (local) mean intensity and a local Bartlett spectrum follow.
- Work in progress: Asymptotic estimation theory.

References I

E. Bacry, S. Delattre, M. Hoffmann, and J.F. Muzy. Some limits for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.*, 123(7):2475–2499, 2013.

E. Bacry, I. Mastromatteo, and J-F. Muzy. Hawkes processes in finance. *ArXiv e-prints*, (1502.04592v2), 2015.

C.G. Bowsher. Modelling security market events in continuous time: Intensity based, multivariate point process models. *J. Econometrics*, 141(2):876–912, Dec 2007.

F. Chen and P. Hall. Inference for a non-stationary self-exciting point process with an application in ulta-high frequency financial data modeling. *J. Appl. Probab.*, 50:1006–1024, 2013.

R. Dahlhaus. On the Kullback-Leibler information divergence of locally stationary processes. *Stochastic Process. Appl.*, 62: 139–168, 1996.

References II

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

A.G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3): 493–503, 1974.

Y. Ogata. On Lewis' simulation method for point processes. *Information Theory, IEEE Transactions on*, 27(1):23–31, Jan 1981.

Y. Ogata. Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *J. Amer. Statist. Assoc.*, 83 (401):9–27, Mar 1988.

P. Reynaud-Bouret and S. Schbath. Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Statist.*, 38(5):2781–2822, 2010.

References III

P. Reynaud-Bouret, V. Rivoirard, and C. Tuleau-Malot. Inference of functional connectivity in neurosciences via Hawkes processes. In *1st IEEE Global Conference on Signal and Information Processing, Austin, Texas, USA*, 2013.

François Roueff, Rainer von Sachs, and Laure Sansonnet. Locally stationary hawkes processes. *Stochastic Processes and their Applications*, 126(6):1710 – 1743, 2016. ISSN 0304-4149. doi: http://dx.doi.org/10.1016/j.spa.2015.12.003. URL http://www.sciencedirect.com/science/article/pii/S0304414915003075.