The Multivariate Hawkes Process in High Dimensions:Beyond Mutual Excitation

Shizhe Chen, Ali Shojaie, Eric Shea-Brown, and Daniela Witten

March 11, 2021

- Introduction
- A Brief Overview of the Hawkes Process
- A New Approach for Analyzing the Hawkes Process
- Application: Cross-Covariance Analysis of the Hawkes Process

Challenge

Previous theoretical work on the Hawkes process is limited to a special case in which a past event can only increase the occurrence of future events, and the link function is linear. However, in neuronal networks and other real-world applications, inhibitory relationships may be present, and the link function may be non-linear.

Contributions

- Employ a thinning process representation and a coupling construction to bound the dependence coefficient of the Hawkes process.
- Establish a concentration inequality for second-order statistics of the Hawkes process.
- Apply this concentration inequality to cross-covariance analysis in the high-dimensional regime and verify the theoretical claims with simulation studies

- Introduction
- A Brief Overview of the Hawkes Process
- A New Approach for Analyzing the Hawkes Process
- Application: Cross-Covariance Analysis of the Hawkes Process

A Brief Overview of the Hawkes Process

$$\lambda_j(t) = \phi_j \left\{ \mu_j + \sum_{k=1}^{p} (\omega_{k,j} * \mathrm{d}N_k)(t) \right\}, \ j = 1, \ldots, p, \tag{1}$$

where

$$(\omega_{k,j} * \mathrm{d}N_k)(t) = \int_0^\infty \omega_{k,j}(\Delta)\mathrm{d}N_k(t-\Delta) = \sum_{i:t_{k,i} \le t} \omega_{k,j}(t - t_{k,i})$$

Assumption 1 We assume that $\phi_j(\cdot)$ is $\alpha_j - \mathit{Lipschitz}$ for $j = 1, \ldots, p$. Let Ω be a $p \times p$ matrix whose entries are $\Omega_{j,k} = \alpha_j \int_0^\infty |\omega_{k,j}(\Delta)|\mathrm{d}\Delta$ for $1 \le j, k \le p$. We assume that there exists a generic constant γ_Ω such that $\Gamma_{\max}(\Omega) \le \gamma_\Omega < 1$.

A Brief Overview of the Hawkes Process

Define the mean intensity as:

$$\Lambda_j = \mathbb{E}[\mathrm{d}N_j(t)]/\mathrm{d}t, j = 1, \ldots, p. \tag{2}$$

Define the (infinitesimal) cross-covariance $\boldsymbol{V}(\cdot) = (V_{k,j}(\cdot))_{p\times p} : \mathbb{R} \mapsto \mathbb{R}^{p\times p}$ as:

$$V_{k,j}(\Delta) \equiv \begin{cases} \mathbb{E}[\mathrm{d}N_j(t)\mathrm{d}N_k(t-\Delta)]/\mathrm{d}t\mathrm{d}(t-\Delta) - \Lambda_j\Lambda_k, & j \neq k, \\ \mathbb{E}[\mathrm{d}N_k(t)\mathrm{d}N_k(t-\Delta)]/\mathrm{d}t\mathrm{d}(t-\Delta) - \Lambda_k^2 - \Lambda_k\delta(\Delta), & j = k \end{cases} \tag{3}$$

where $\delta(\cdot)$ is the Dirac delta function.

- Introduction
- A Brief Overview of the Hawkes Process
- A New Approach for Analyzing the Hawkes Process
- Application: Cross-Covariance Analysis of the Hawkes Process

Temporal Dependence

The key idea is to represent the generalized Hawkes process using the thinning process representation to simulate data from the Hawkes process.

$$\bar{y}_{k,j} \equiv \frac{1}{T}\int_0^T \int_0^T f(t-t')\mathrm{d}N_k(t)\mathrm{d}N_j(t'), 1 \le j,k \le p \tag{4}$$

$$y_{k,j,i} \equiv \frac{1}{2\epsilon}\int_{2\epsilon(i-1)}^{2\epsilon i} \int_0^T f(t-t')\mathrm{d}N_k(t)\mathrm{d}N_j(t'), \tag{5}$$

where $f(\cdot)$ is a known function with properties to be specified later, ϵ is some small constant. $\bar{y}_{k,j}$ can be seen as the average of the sequence $\{y_{k,j,i}\}_{i=1}^{T/(2\epsilon)}$.

Temporal Dependence

Use the τ-dependence coefficient as a measure of dependence for $\{y_{k,j,i}\}_{i=1}^{T/(2\epsilon)}$.

$$\tau(\mathcal{M}, X) \equiv \mathbb{E}\left[\sup_h \left\{\left|\int h(x)\mathbb{P}_{X|\mathcal{M}}(\mathrm{d}x) - \int h(x)\mathbb{P}_X(\mathrm{d}x)\right|\right\}\right] \tag{6}$$

where $h(\cdot)$ is 1-Lipschitz function, X is a random variable, and $\mathcal{M}$ is σ-field. For a sequence $\{y_{k,j,i}\}_{i=1}^{T/(2\epsilon)}$, the temporal dependence is defined as

$$\tau_y(l) \equiv \sup_{u\in\{1,2,\ldots\}} \tau(\mathcal{H}_u^y, y_{k,j,u+l}), \tag{7}$$

where l is any positive integer, $\mathcal{H}_u^y$ is the σ-field determined by $\{y_{k,j,i}\}_{i=1}^{T/(2\epsilon)}$.

Lemma 1 Let X be an integrable random variable and $\mathcal{M}$ a σ-field defined on the same probability space. If the random variable Y has the same distribution as X, and is independent of $\mathcal{M}$, then

$$\tau(\mathcal{M}, X) \leq \mathbb{E}|X - Y|. \tag{8}$$

Then, we can get

$$\tau_y(l) \leq \sup_u \mathbb{E}\left|\tilde{y}^u_{k,j,u+l} - y_{k,j,u+l}\right| \tag{9}$$

where $\{\tilde{y}_{k,j,i}\}_{i=1}^{T/(2\epsilon)}$ is a sequence satisfying the conditions of Lemma 1.

Coupling Process Construction using Thinning

Let $N_j^{(0)}$, for $j = 1, \ldots, p$, be a homogeneous Poisson process on $\mathbb{R}^2$ with intensity 1. For $n = 1$, we construct a p-variate process $\boldsymbol{N}^{(1)}$ as

$$\mathrm{d}N_j^{(1)}(t) = N_j^{(0)}([0, \mu_j] \times \mathrm{d}t), j = 1, \ldots, p \tag{10}$$

where $N_j^{(0)}([0, \mu_j] \times \mathrm{d}t)$ is the number of points for $N_j^{(0)}$ in the area $[0, \mu_j] \times [t, t + \mathrm{d}t)$. For $n \geq 2$, we construct $\boldsymbol{N}^{(n)}$ as

$$\lambda_j^{(n)}(t) = \phi_j \left\{ \mu_j + (\boldsymbol{\omega}_{\cdot,j} * \mathrm{d}\boldsymbol{N}^{(n-1)})(t) \right\} \tag{11}$$

$$\mathrm{d}N_j^{(n)}(t) = N_j^{(0)}([0, \lambda_j^{(n)}(t)] \times \mathrm{d}t), j = 1, \ldots, p. \tag{12}$$

Coupling Process Construction using Thinning

Let $\tilde{N}_j^{(0)}$, for $j = 1, \ldots, p$, also be a homogeneous Poisson process on $\mathbb{R}^2$ with intensity 1, but independent of $N_j^{(0)}$. For $n = 1$, we construct a p-variate process $\tilde{\boldsymbol{N}}^{(1)}$ as

$$\mathrm{d}\tilde{N}_j^{(1)}(t) = \begin{cases} \tilde{N}_j^{(0)}([0, \mu_j] \times \mathrm{d}t), & t \leq z, \\ N_j^{(0)}([0, \mu_j] \times \mathrm{d}t), & t > z \end{cases} \tag{13}$$

where $z \equiv 2\epsilon u$. For $n \geq 2$, we construct $\tilde{\boldsymbol{N}}^{(n)}$ as

$$\tilde{\lambda}_j^{(n)}(t) = \phi_j \left\{ \mu_j + (\boldsymbol{\omega}_{\cdot,j} * \mathrm{d}\tilde{\boldsymbol{N}}^{(n-1)})(t) \right\} \tag{14}$$

$$\mathrm{d}\tilde{N}_j^{(n)}(t) = \begin{cases} \mathrm{d}\tilde{N}_j^{(0)}([0, \tilde{\lambda}_j^{(n)}(t)] \times \mathrm{d}t), & t \leq z, \\ \mathrm{d}N_j^{(0)}([0, \tilde{\lambda}_j^{(n)}(t)] \times \mathrm{d}t), & t > z \end{cases} \tag{15}$$

Deviation

Theorem 1 Let N be a Hawkes process with intensity of the form (1) satisfying Assumption 1. For any given $z > 0$, there exists a point process $\tilde{\boldsymbol{N}}$ satisfying the Lemma 1, called the coupling process of $\boldsymbol{N}$. Let b be a constant and define a matrix $\boldsymbol{\eta}(b) = (\eta_{j,k}(b))_{p\times p}$ with $\eta_{j,k}(b) = [\alpha_j \int_b^\infty |\omega_{k,j}(\Delta)|\mathrm{d}\Delta]$. Then, for any $u \geq 0$,

$$\mathbb{E}\left|\mathrm{d}\tilde{\boldsymbol{N}}(z+u) - \mathrm{d}N(z+u)\right|/\mathrm{d}u \preceq 2v_1\left\{\boldsymbol{\omega}^{\lfloor u/b+1\rfloor}\boldsymbol{J} + \sum_{i=1}^{\lfloor u/b+1\rfloor}\boldsymbol{\Omega}^{i-1}\boldsymbol{\eta}(b)\boldsymbol{J}\right\} \tag{16}$$

Deviation

$$
\begin{aligned}
&\mathbb{E}\left|\mathrm{d}\tilde{\boldsymbol{N}}(t')\mathrm{d}\tilde{\boldsymbol{N}}(z+u)-\mathrm{d}N(t')\mathrm{d}N(z+u)\right|/(\mathrm{d}u\mathrm{d}t') \\
&\preceq 2v_2\left\{\boldsymbol{\omega}^{\lfloor u/b+2\rfloor}\boldsymbol{J}+\sum_{i=1}^{\lfloor u/b+1\rfloor}\boldsymbol{\Omega}^i\boldsymbol{\eta}(b)\right\}\boldsymbol{J}\boldsymbol{J}^T \\
&+2v_1^2\left\{\boldsymbol{J}\boldsymbol{J}^T(\boldsymbol{\Omega}^{\lfloor \boldsymbol{u/b+1}\rfloor})^T+\sum_{i=1}^{\lfloor u/b+1\rfloor}\boldsymbol{\eta}(b)\boldsymbol{J}\boldsymbol{J}^T(\boldsymbol{\Omega}^{\boldsymbol{i}})^T\right\}
\end{aligned}
\tag{17}
$$

where $\preceq$ denotes element-wise inequality, and v_1 and v_2 are parameters that depend on $\boldsymbol{\Lambda}, \boldsymbol{V}$.

Main Result

Assumption 2 There exists a constant b_0 such that, for $b \geq b_0$ and some $r > 0$,

$$\max_j \sum_{k=1}^{p} \int_b^\infty |\omega_{k,j}(\Delta)| \mathrm{d}\Delta \leq c_1 \exp(-c_2 b^r).$$

Assumption 3 For all $j = 1, \ldots, p$, there exists a positive constant $\rho_\Omega < 1$ such that $\mathbf{\Omega}$ satisfies $\sum_{k=1}^{p} \Omega_{j,k} \leq \rho_\Omega$

Main Result

Theorem 2 Suppose that $\boldsymbol{N}$ is a Hawkes process with intensity (1), which satisfies Assumptions 1–4. Suppose also that the function $f(\cdot)$ has bounded support and $\max_x |f(x)| \le C_f$. For any positive integer l, the τ-dependence coefficient of $\{y_{j,k,i}\}_i$ introduced in (7) satisfies

$$\tau_y(l) \le a_5 \exp(-a_6 l^{r/(r+l)}) \tag{18}$$

where r is introduced in Assumption 2, and a_5 and a_6 are parameters that do not depend on p and T.

Main Result

Assumption 4 Assume that one of the following two conditions holds.

- For all $j = 1, \ldots, p$, the link functions $\phi_j(\cdot)$ in (1) are upper-bounded by a positive constant $\phi_{\max}$.
- In Assumption 2, the constant $c_1 = 0$.

Lemma 2 Suppose $\boldsymbol{N}$ is a Hawkes process with intensity (1), which satisfies Assumptions 1 and 4. For any bounded interval A, it holds that, for $j = 1, \ldots, p$, $P(N_j(A) > n) \leq \exp(1 - n/K)$ for some constant K.

Main Result

Theorem 3 Suppose that $\boldsymbol{N}$ is a Hawkes process with intensity (1), which satisfies Assumptions 1–4. Suppose also that the function $f(\cdot)$ has bounded support and $\max_x |f(x)| \leq C_f$. Then, for $1 \leq j \leq k \leq p$,

$$\mathbb{P}\left(\bigcap_{1\leq j\leq k\leq p}\left[|\bar{y}_{k,j} - \mathbb{E}\bar{y}_{k,j}| \geq c_3 T^{-(2r+1)/(5r+2)}\right]\right) \leq c_4 p^2 T \exp\left(-c_5 T^{r/(5r+2)}\right), \tag{19}$$

where r is introduced in Assumption 2, and c_3, c_4 and c_5 are parameters that do not depend on p and T.

- Introduction
- A Brief Overview of the Hawkes Process
- A New Approach for Analyzing the Hawkes Process
- Application: Cross-Covariance Analysis of the Hawkes Process

Theoretical Guarantees

As a concrete example, we consider the following smoothing estimator

$$\hat{V}_{k,j}(\Delta) = \begin{cases} (Th)^{-1} \int\int_{[0,T]^2} K\left(\frac{(t'-t)+\Delta}{h}\right) \mathrm{d}N_j(t')\mathrm{d}N_k(t) \\ -T^{-1} N_j([0,T]) \frac{1}{T} N_k([0,T]),\ j \neq k \\ \\ (Th)^{-1} \int\int_{[0,T]^2 \setminus \{t=t'\}} K\left(\frac{(t'-t)+\Delta}{h}\right) \mathrm{d}N_k(t')\mathrm{d}N_k(t) \\ -T^{-2} N_k^2([0,T]),\ j = k \end{cases} \tag{20}$$

where $K(\cdot)$ is a kernel function with bandwidth h and bounded support.

Corollary 1 Suppose that $\boldsymbol{N}$ is a Hawkes process with intensity (1), which satisfies Assumptions 1–4. Further assume that the cross-covariances $\{V_{k,j}, 1 \leq j, k \leq p\}$ are θ_0-Lipschitz functions. Let $h = c_6 T^{-(r+0.5)/(5r+2)}$ in (20) for some constant c_6. Then,

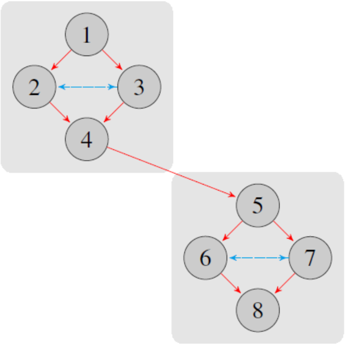
$$\mathbb{P}\left(\bigcap_{1\leq j\leq k\leq p}\left[\left\|\hat{V}_{k,j} - V_{k,j}\right\|_{2,[-B,B]} \leq c_6 T^{-\frac{r+0.5}{5r+2}}\right]\right) \geq 1 - 2c_4 p^2 T^{\frac{6r+0.5}{5r+2}} \exp(-c_5 T^{\frac{r}{5r+2}}) \tag{21}$$

where $\left\|\hat{V}_{k,j} - V_{k,j}\right\|_{2,[-B,B]}^2 \equiv \int_{-B}^{B}\left[\hat{V}_{k,j}(\Delta) - V_{k,j}(\Delta)\right]^2 \mathrm{d}\Delta$ with B a user-defined constant, c_4 and c_5 are constants introduced in Theorem 3, and c_6 depends on θ_0, B, and c_3 in Theorem 3.

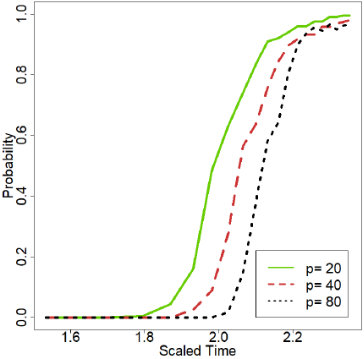
Let the intensity of the Hawkes process takes the form (1) with $\phi_j(x) = \exp(x)/[1 + \exp(x)], \mu_j = 1$, and $\omega_{k,j}(t) = a_{k,j}\gamma^2 t \exp(-\gamma t)$ for all $1 \le j, k \le p$.
Define the event A as

$$A \equiv \left\{ \bigcap_{1\le j,k\le p} \left[\left\| \hat{V}_{k,j} - \tilde{V}_{k,j} \right\|_{2,[-B,B]} \le c_6 T^{-3/14} \right] \right\} \tag{22}$$

Simulation Studies



(a) Graph



(b) Simulation results