Nonparametric Estimation of Hawkes Causality

Massil Achab†, Emmanuel Bacry†, Stéphane Gaïffas†, Iacopo Mastromatteo* and Jean-François Muzy†,‡

† Ecole Polytechnique, * Capital Fund Management, ‡ Université de Corse

Summary

We propose a new method to measure causal relaships between nodes (users) based on their activity timestamps. We answer to the question *How many events of type j does an event of type i cause?*.

Introduction

A **temporal point process** is a collection of points $Z = \{\tau_1, \tau_2, \dots\}$ randomly located on $\mathbb{R}^+$. Another representation is based on the counting process $N_t = \sum_{\tau \in Z} \mathbb{1}_{\tau \le t}$.

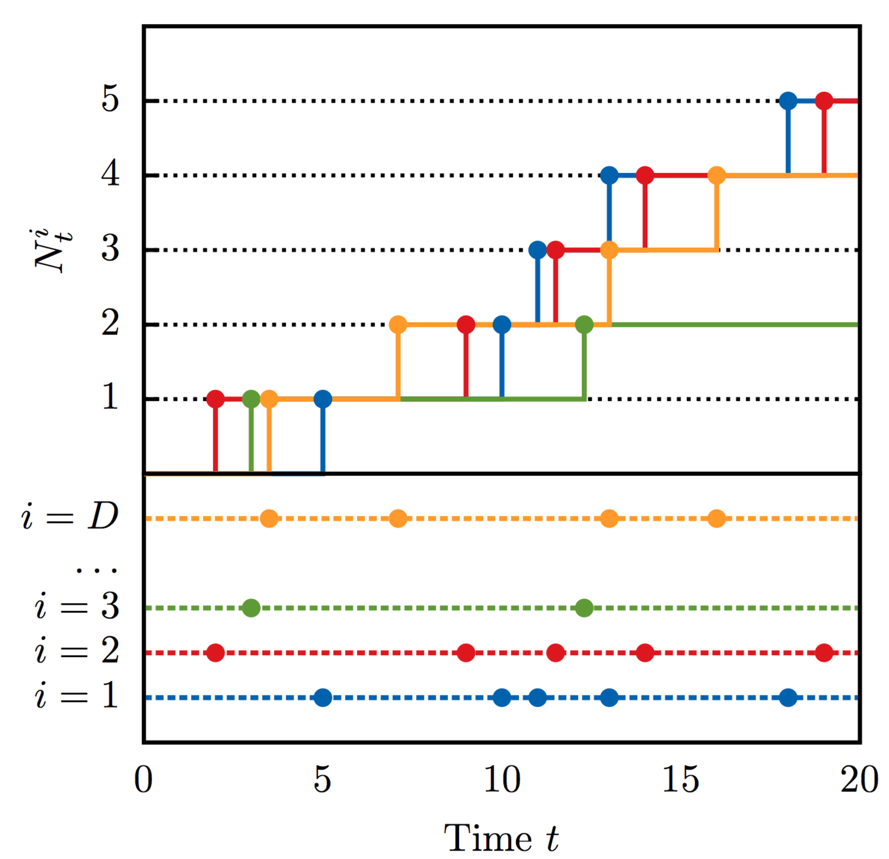

Figure 1: Example of multivariate point process

The vector of stochastic intensities $\boldsymbol{\lambda}_t = [\lambda_t^1 \cdots \lambda_t^D]^\top$ associated with the multivariate counting process $\boldsymbol{N}_t$ is defined as

$$\lambda_t^i = \lim_{dt \to 0} \frac{\mathbb{P}(N_{t+dt}^i - N_t^i = 1 | \mathcal{F}_t)}{dt}.$$

$\boldsymbol{N}_t$ is called a **Hawkes point process** if the stochastic intensities can be written as

$$\lambda_t^i = \mu^i + \sum_{j=1}^{D} \sum_{\tau_j \in Z^j} \phi^{ij}(t - \tau_j),$$

where $\mu^i \in \mathbb{R}^+$ is an exogenous intensity and ϕ^{ij} are positive, integrable and causal (with support in $\mathbb{R}^+$) functions called *kernels* encoding the impact of an action by j on the activity of i.

Hawkes Causality

Instead of trying to estimate the kernels ϕ^{ij}, we focus on the direct estimation of their *integrals*:

$$g^{ij} = \int_0^{+\infty} \phi^{ij}(u) du. \qquad (1)$$

The matrix $\boldsymbol{G}$ can be seen as a weighted adjacency matrix. Using Hawkes process representation as a Poisson cluster process [1], we introduce the counting function $N_t^{i \leftarrow j}$ that counts the number of events of i whose direct ancestor is an event of j. We know that

$$\mathbb{E}[dN_t^{i \leftarrow j}] = g^{ij} \mathbb{E}[dN_t^j] = g^{ij} \Lambda^j dt. \qquad (2)$$

We refer to *learning the kernels' integrals* as *uncovering Hawkes causality* since each integral encodes the number of events directly caused from a node to another node, as described at Eq. (2). This is also linked to Granger causality (see the paper for more details).

Integrated Cumulants

A general formula for the cumulants of Hawkes processes is provided in [2] (cumulants are centered moments, up to the third order). The integrals of these cumulants are related to $\boldsymbol{R} = (\boldsymbol{I} - \boldsymbol{G})^{-1}$. For the purpose of our method, we only need to consider cumulants up to the third order.

$$\Lambda^i = \sum_{m=1}^{D} R^{im} \mu^m \qquad (3)$$

$$C^{ij} = \sum_{m=1}^{D} \Lambda^m R^{im} R^{jm} \qquad (4)$$

$$S^{ijk} = \sum_{m=1}^{D} (R^{im} R^{jm} C^{km} + R^{im} C^{jm} R^{km} + C^{im} R^{jm} R^{km} - 2\Lambda^m R^{im} R^{jm} R^{km}). \qquad (5)$$

We designed asymptotically unbiased estimators of the integrated cumulants above, see the paper for details.

Nonparametric Estimation

To recover $\boldsymbol{G}$, we minimize the following criterion using SGD:

$$\mathcal{L}(\boldsymbol{R}) = (1-\kappa) || \boldsymbol{S_c}(\boldsymbol{R}) - \widehat{\boldsymbol{S_c}} ||_2^2 + \kappa || \boldsymbol{C}(\boldsymbol{R}) - \widehat{\boldsymbol{C}} ||_2^2, \qquad (6)$$

Our method is called **NPHC** (for Nonparametric Hawkes Cumulants).

Consistency of NPHC

Suppose that $(\boldsymbol{N_t})$ is observed on $\mathbb{R}^+$ and assume that

- $g_0(\boldsymbol{R}) = 0$ if and only if $\boldsymbol{R} = \boldsymbol{R}_0$;
- $\boldsymbol{R} \in \Theta$, which is a compact set;
- the spectral radius of the kernel norm matrix satisfies $||\boldsymbol{G}_0|| < 1$;
- $H_T \to \infty$ and $H_T^2 / T \to 0$.

Then

$$\widehat{\boldsymbol{G}}_T = \boldsymbol{I} - \left(\arg\min_{\boldsymbol{R} \in \Theta} \mathcal{L}_T(\boldsymbol{R}) \right)^{-1} \xrightarrow{\mathbb{P}} \boldsymbol{G}_0.$$

Results

We consider the following 8-dimensional point process, already introduced in [3], which models the dynamic of an order book:

$$N_t = (P_t^+, P_t^-, T_t^a, T_t^b, L_t^a, L_t^b, C_t^a, C_t^b)$$

Each dimension counts the number of events before t:

- P^+ (P^-): upwards (downward) mid-price move triggered by any order.
- T^a (T^b): market order at the ask (bid) that does not move the price.
- L^a (L^b): limit order at the ask (bid) that does not move the price.
- C^a (C^b): cancel order at the ask (bid) that does not move the price.

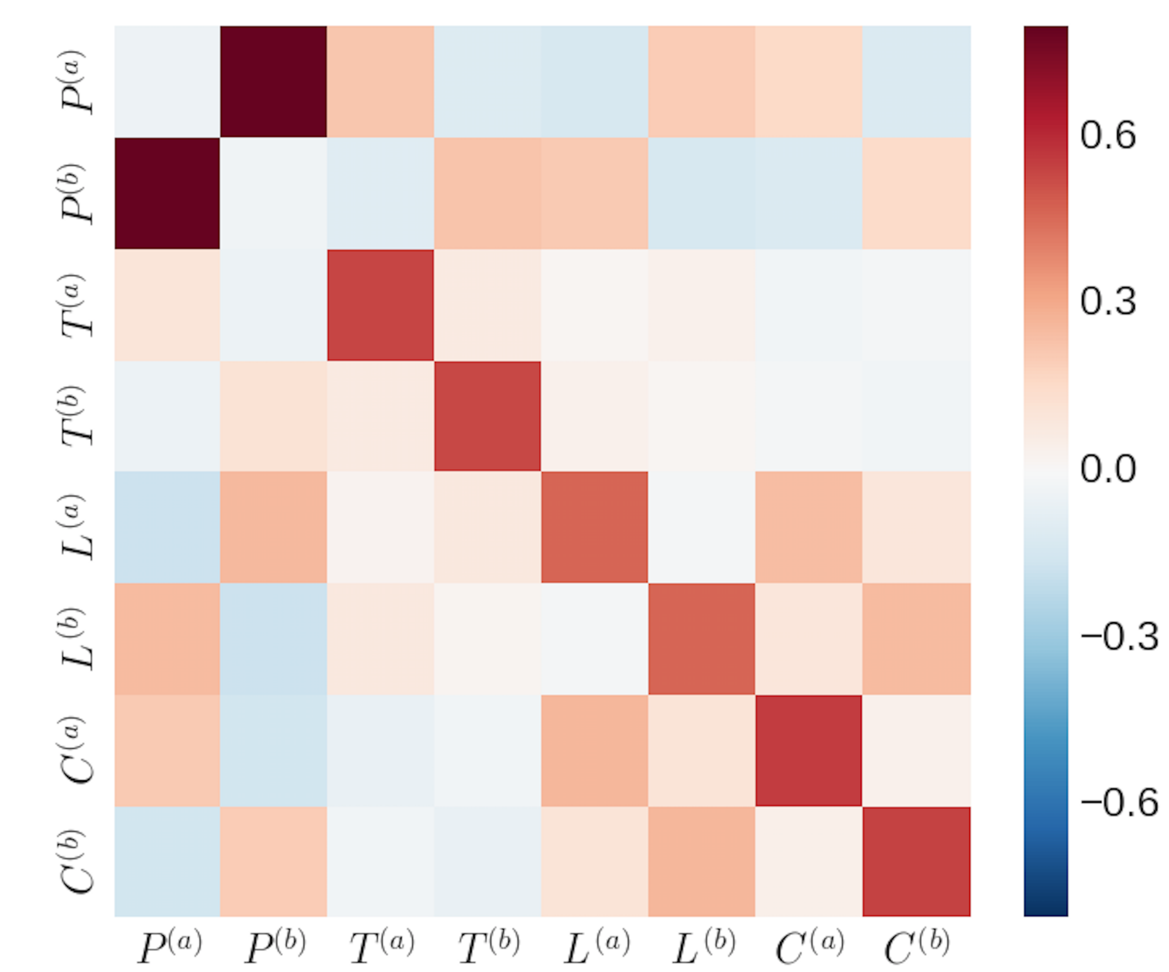

Figure 2: Estimated $\widehat{\boldsymbol{G}}$ on DAX order book data.

References

[1] A. G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *Journal of Applied Probability*, pages 493–503, 1974.

[2] S. Jovanović, J. Hertz, and S. Rotter. Cumulants of hawkes point processes. *Physical Review E*, 91(4):042802, 2015.

[3] E. Bacry, T. Jaisson, and J.-F. Muzy. Estimation of slowly decreasing hawkes kernels: application to high-frequency order book dynamics. *Quantitative Finance*, 16(8):1179–1201, 2016.

Acknowledgements

This research benefited from the support of the *École Polytechnique fund raising - Data Science Initiative*. The authors want to thank Marcello Rambaldi for fruitful discussions on order book data's experiments.

CMAP

DATA SCIENCE INITIATIVE