

Estimation of slowly decreasing Hawkes kernels: Application to high frequency order book modelling

Emmanuel Bacry, Thibault Jaisson and Jean-François Muzy

7th General AMaMeF and Swissquote Conference

Wednesday 9th September 2015

Table of contents

- 1 Hawkes processes
 - Definition and basic properties
 - Causality interpretation
 - The Wiener-Hopf Equation
- 2 Estimation of Hawkes kernels
 - Review of the literature
 - The modified Wiener-Hopf method
- 3 Application to high frequency financial data
 - A Hawkes order book model
 - Estimation results

Table of contents

- 1 Hawkes processes
 - Definition and basic properties
 - Causality interpretation
 - The Wiener-Hopf Equation
- 2 Estimation of Hawkes kernels
- 3 Application to high frequency financial data

1-dimensional Hawkes process

Definition (Hawkes 1971)

A Hawkes process N is a point process on \mathbb{R} of intensity:

$$\lambda_t = \mu + \int_{-\infty}^t \phi(t-s) dN_s \quad (1)$$

$$= \mu + \sum_{J_i < t} \phi(t - J_i) \quad (2)$$

where $\mu \in \mathbb{R}_+^*$ is the exogenous intensity and ϕ is a positive kernel supported in \mathbb{R}_+ which satisfies $\int \phi < 1$ and the J_i are the points of N .

That is a simple construction of “clustered” point processes (if there has been points in a short past, there will be points in a short future).

Application of Hawkes processes

They are used to reproduce the clusterization that is observed in many fields:

- Seismology (Adamopoulos 1976).
- Neurology (Reynaud-Bouret et al. 2013).

They have recently been applied a lot in finance:

- Price changes (Hardiman et al. 2013).
- Market orders (Bacry and Muzy 2013).
- Financial contagion (Aït-Sahalia et al. 2010).

What if we have different point processes which are cross-clustered?

Multi dimensional Hawkes processes

Multi dimensional Hawkes processes are defined as self and mutually exciting point processes.

Definition

A Hawkes process $(N^i)_{i=1,\dots,D}$ is a set of point processes of intensity:

$$\lambda_t^i = \mu^i + \sum_{j=1}^D \int_{-\infty}^t \phi^{ij}(t-s) dN_s^j$$

where ϕ^{ij} encodes the influence of j over i and μ^i encodes the exogenous arrivals of i .

Our Hawkes order book model

We will use an 8-dimensional Hawkes process to jointly model 8 kinds of order book events:

- Upward and downward mid-price changes.
- Market orders at the bid and ask.
- Limit orders at the bid and ask.
- Cancel orders at the bid and ask.

This model will reproduce the cross-correlations between these kinds of events.

Stability condition

Let us denote $|\phi|$ the matrix $(\int \phi^{ij})$.

Proposition (Hawkes 1971)

The process is well defined and admits a version with stationary increments under the stability condition:

$$\rho(|\phi|) < 1$$

where $\rho(|\phi|)$ is the greatest eigenvalue of $|\phi|$.

Proposition

The average intensity of a stationary Hawkes process is

$$\Lambda = E[\lambda_t] = \mu(\mathbb{I} - |\phi|)^{-1}.$$

A population Hawkes model

Let us mention that there is a “population representation” of Hawkes processes according to which they are built as the times of arrival and birth of the following population process:

- There are individuals of type 1, ..., D .
- Migrants of type i arrive at a Poisson rate of μ^i .
- The children of type i of an individual of type j who was born or migrated in t are distributed as an inhomogeneous Poisson process of intensity $\phi^{ij}(\cdot - t)$.

Then, the intensity of the arrival of individuals of type i is

$$\lambda_t^i = \mu^i + \sum_{j=1}^D \sum_{J_k^j < t} \phi^{ij}(t - J_k^j).$$

Causality in the Hawkes model

In this model, $|\phi^{jj}|$ appears as the average number of children of type i of an individual of type j .

Similarly, $\frac{N}{\lambda^i} |\phi^{jj}|$ appears as the proportion of individuals of type i whose parent is an individual of type j and $\frac{\mu^i}{\lambda^i}$ is the proportion of exogenous i .

We thus use $|\phi^{jj}|$ as a measure of “causality” between j and i .

For example, in our Hawkes order book model, we will be able to measure the average number of bid limit orders caused by an ask market order.

Conditional laws

- The conditional law g^{ij} is defined as the probability that there is a point i at time t knowing that there has been a point j in 0

$$g^{ij}(t)dt = \mathbb{E}[dN_t^i = 1 | dN_0^j = 1] - \mathbb{1}_{i=j}\delta(t) - \Lambda^i dt.$$

- It corresponds to the cross-correlation between i and j :

$$g^{ij}(t - t')dtdt' = \mathbb{E}[dN_t^i dN_{t'}^j] / \Lambda^j - \mathbb{1}_{i=j}\delta(t)dt' - \Lambda^i dtdt'.$$

- There is a difference with the cross-causality.
- It can be measured empirically!

The Wiener-Hopf Equation

The Wiener-Hopf Equation disentangles the causality from the correlation structure.

Proposition

The conditional laws and the kernels are linked by

$$\forall t > 0, \quad g(t) = \phi(t) + g * \phi(t) \quad (3)$$

that is

$$g^{ij}(t) = \phi^{ij}(t) + \sum_{k=1}^D \int_0^{+\infty} g^{ik}(t-s) \phi^{kj}(s) ds.$$

Model independent interpretation of Wiener-Hopf

Let us consider a point process of conditional law g and intensity Λ .
The kernel ϕ such that the intensity predictor

$$\lambda_t^\phi = \mu + \int_{-\infty}^t \phi(t-s) dN_s$$

(with $\mu = \Lambda(1 - |\phi|)$) minimizes the error

$$\varepsilon(\phi) = \mathbb{E}[(\lambda_t - \lambda_t^\phi)^2]$$

is the solution of the Wiener-Hopf equation.

Table of contents

- 1 Hawkes processes
- 2 Estimation of Hawkes kernels
 - Review of the literature
 - The modified Wiener-Hopf method
- 3 Application to high frequency financial data

Likelihood based methods

Many works estimate parametric Hawkes kernels by maximizing the Hawkes likelihood as a function of the parameters of the kernels.

- Newton method (Lemonnier, Vayatis 2014).
- Majoration-Minimization algorithms (Zhu et al 2013).
- Stochastic gradient decent (work in progress with Emmanuel Bacry).

The Wiener-Hopf method

The method of Bacry and Muzy 2014 consists in finding the kernel solution to the Wiener-Hopf equation (which exists and is unique).

To do this, they approximate the convolution in

$$g = \phi + g * \phi$$

using a quadrature scheme on $[0, T]$, to obtain:

$$\tilde{g}(t_n) = \tilde{\phi}(t_n) + \sum_{k=1}^K \omega_k \tilde{g}(t_n - t_k) \tilde{\phi}(t_k)$$

where $t_k \simeq kh$ and $\omega_k \simeq h$ denote the points and the weights of the scheme.

This is a linear system which can be inverted.

Issues with this method

- This method does not work for slowly decreasing kernels.
- On financial data, Hawkes kernels present a significant behaviour over 8 decades (from $10\mu s$ to 1 000 seconds).
- The “uniform” scheme presented above would thus need around 10^8 points.
- We propose a “log-uniform” piecewise linear scheme to solve this issue.

Modification of the numerical scheme

In order to solve this issue, we consider the time grid

$$\{t_k\}_{1 \leq k \leq N} = [0, \delta T_{min}, 2\delta T_{min}, \dots, T_{min}, T_{min}e^\delta, T_{min}e^{2\delta}, \dots, T_{max}].$$

If we assume that the kernel is piecewise linear on $[t_k, t_{k+1}]$,

$$\phi(t) = \phi(t_k) + \frac{t - t_k}{t_{k+1} - t_k} (\phi(t_{k+1}) - \phi(t_k)),$$

the Wiener-Hopf system becomes

$$\begin{aligned} g(t_n) &= \phi(t_n) + \sum_{k=0}^{N-1} \phi(t_k) \int_{t_k}^{t_{k+1}} g(t_n - s) ds \\ &+ \sum_{k=0}^{N-1} (\phi(t_{k+1}) - \phi(t_k)) \int_{t_k}^{t_{k+1}} \frac{s - t_k}{t_{k+1} - t_k} g(t_n - s) ds \end{aligned}$$

Simulation results (1D)

Parameters: $T_{min} = 0.001$, $T_{max} = 1000$, $\delta = 5\%$,

$$\phi(t) = \frac{0.06}{(0.005 + t)^{1.3}} \quad , \quad \mu = 0.05 \quad \text{and} \quad T = 10^7. \quad (4)$$

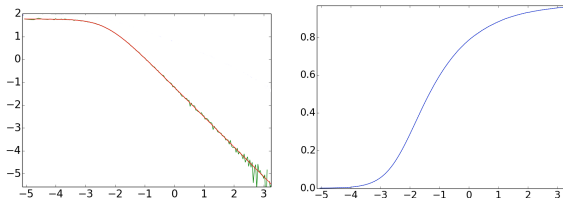


Figure : Left: Theoretical kernel (red) and estimated kernel (green) for 200 quadrature points.

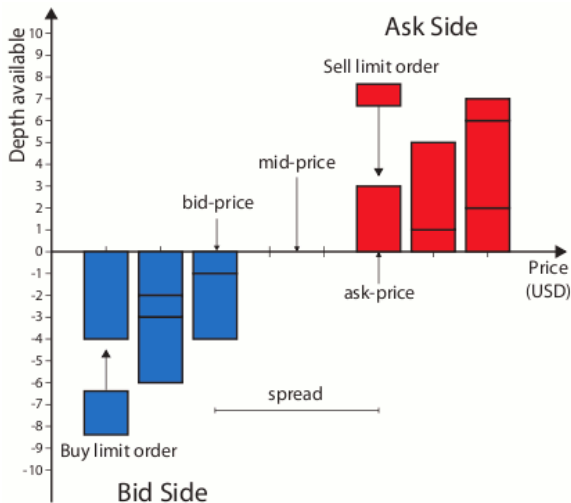
Right: Cumulated theoretical kernel $\int_0^t \phi(s) ds$ (green dots) and cumulated estimated kernel $\int_0^t \tilde{\phi}(s) ds$ (blue) as a function of $\log_{10}(t)$.

Table of contents

- 1 Hawkes processes
- 2 Estimation of Hawkes kernels
- 3 Application to high frequency financial data
 - A Hawkes order book model
 - Estimation results

The limit order book

Price changes – Market orders – Limit orders – Cancel orders



The model

We model as a Hawkes process the following 8 dimensional counting process:

$$N_t = (P_t^{(a)}, P_t^{(b)}, T_t^{(a)}, T_t^{(b)}, L_t^{(a)}, L_t^{(b)}, C_t^{(a)}, C_t^{(b)})$$

where:

- $P^{(a)}$ (resp. $P^{(b)}$) counts the number of upward (resp. downward) mid-price moves.
- $T^{(a)}$ (resp. $T^{(b)}$) counts the number of market orders at the ask (resp. bid) that do not move the price.
- $L^{(a)}$ (resp. $L^{(b)}$) counts the number of limit orders at the ask (resp. bid) that do not move the price.
- $C^{(a)}$ (resp. $C^{(b)}$) counts the number of cancel orders at the ask (resp. bid) that do not move the price.

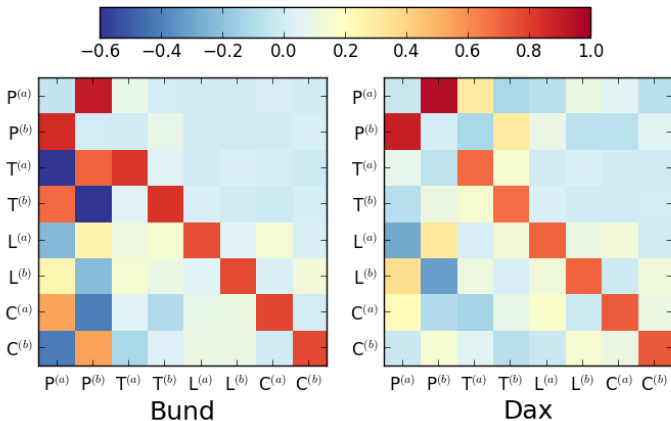
The data

The financial data used in this paper have been provided by the company QuantHouse EUROPE/ASIA.

It consists of all level-I order book data of BUND and DAX future contracts.

For every day, we only keep the most liquid maturity and we use data over one year from June 2013 to June 2014. Each file lists all the changes in the first limit (best ask or best bid) of the order book at a micro second precision.

Hawkes norms



Interpretation in terms on market dynamics

- We studied two very different assets: The Bund and the DAX.
- Orders are mostly self exciting: Trades at the buy mostly imply trades at the buy.
- Orders impact the price.
- Prices moves are mostly cross exciting: Upward price moves mostly imply downward price moves.
- This yields a certain stability on the price process.

The influence of price changes on trades

- Negative kernels correspond to an inhibitive effect.
- The tick size is the discretization step of the price.
- It characterizes the high frequency order book dynamics of an asset.
- The DAX is a small tick asset, price moves are not too significant, if the price goes upward, it means that people are buying and will keep buying.
- The BUND is a large tick asset, price moves are very significant and thus after an upward price jump, buy orders will be inefficient.

Conclusion

- Hawkes processes are a nice way to reproduce the causality between various events.
- The Wiener-Hopf systems allows to get in a model independent sense the “causality” from the “correlation”.
- We proposed a non parametric scheme to solve this system when the kernels are slowly decreasing.
- We applied this procedure to study the causality between the different market events.

Thank you for your
attention!