POINT PROCESS MODELLING OF A LIMIT ORDER BOOK

Oden Petersen

Supervisor: Dr. Tom Stindl

School of Mathematics and Statistics
UNSW Sydney

September 2024

Submitted in partial fulfillment of the requirements of the degree of
Bachelor of Advanced Mathematics with Honours

Plagiarism statement

I declare that this thesis is my own work, except where acknowledged, and has not been submitted for academic credit elsewhere.

I acknowledge that the assessor of this thesis may, for the purpose of assessing it:

- Reproduce it and provide a copy to another member of the University; and/or,
- Communicate a copy of it to a plagiarism checking service (which may then retain a copy of it on its database for the purpose of future plagiarism checking).

I certify that I have read and understood the University Rules in respect of Student Academic Misconduct, and am aware of any potential plagiarism penalties which may apply.

By signing this declaration I am agreeing to the statements and conditions above.

Signed: ______________________________ Date: ____________________

Acknowledgements

Abstract

This is the abstract

Contents

Chapter 1

Introduction

On a single day in 2023, the US stock market saw an average of around $500 billion dollars worth of shares traded on various exchanges and other venues, exceeding the annual GDP of a typical European country [7]. Modern securities markets facilitate the exchange of shares and other financial assets at extremely high frequency, as a result of aggressive investment in specialised networking hardware, custom-made computer architectures, and high-throughput machine learning systems. In facilitating capital flows for the global economy, financial markets have at the same time managed to claim an increasing fraction of resources and attention, with economic consequences that are not yet fully understood [1].

Despite many mysteries and open questions about the origins and dynamics of market phenomena, the financial sector itself has readily adapted to the increasing scale and complexity of the markets in which it operates. The practical design of exchange rules and trading systems has in large part been an empirical endeavour on the part of market participants, operators, and regulators. Many phenomena have been observed to emerge in an apparently decentralised fashion from the application of exploitative heuristics and predictive algorithms that interact with exchanges and aggregate to produce desirable outcomes. While users of trading strategies aim to maintain acceptable risk levels while generating profits over the long term, market operators and regulators are tasked with the design of incentive mechanisms that exploit this self-interested behaviour to improve market outcomes, including reduced transaction costs, fast and accurate incorporation of external information (such as economic news or earnings reports), and adherence to various concepts of fairness, propriety, and legality.

In this thesis, I describe and extend prior work from the empirical market microstructure literature, making extensive use of the state-dependent Hawkes process model for event arrivals. I begin with a conceptual overview of the trading mechanism, and formalise the mathematical tools that will be used to construct and describe variations on the basic Hawkes process model. I continue with a summary of existing literature on point processes as applied to market data, including both mathematical foundations and empirical findings. Next, I explore techniques to reduce the computational burden of parametric inference for point processes on large datasets. Finally, empirical applications and findings are discussed, including applications of generative modeling to a variety of open problems in market microstructure.

1.1 Limit Order Books

Intraday trading allows participants to respond to exogenous news and endogenous market events in a manner that maintains acceptable levels of risk and generates profits over the long term. This activity has a significant influence on the formation of market prices and plays a key role in reducing transaction costs while increasing the speed at which large institutions can control their exposure to various financial risks and opportunities.

Securities exchanges facilitate automated matching of buyers and sellers at prices favourable to both. Understanding the dynamics of this exchange process at a high degree of resolution can provide insights into the design of automated *matching engines* that produce desirable market behaviour, as well as insights into the design of *trading strategies* that exploit the dynamics of the exchange process to generate profits.

A matching engine is tasked with receiving and acting on various messages from market participants indicating their intent to buy or sell a particular security with particular conditions. As a result of this process, trades may be formed that match buyers and sellers at a mutually agreeable price and quantity. Trade reports are broadcast to relevant participants and may be used for the purposes of risk management and forecasting, as well as for the ultimate transfer of the assets that have been traded (which often occurs after trading hours).

A typical matching engine permits two kinds of incoming messages, known as *order insertion* and *order cancellation*, and maintains an internal state consisting of a single data structure, known as a *limit order book*.

An order insertion message indicates a participant's willingness to buy (or sell) some quantity of a particular asset at or below (respectively, above) a particular price, and results in the addition of an *order* to the limit order book $\mathcal{L}$.

Conversely, an order cancellation message results in the removal of a particular order from the limit order book, either in part (by reducing the remaining volume associated with the order) or in full (by removing the order entirely from the book). It might be good to include a statistic about roughly how large cancellation rates are, to highlight the importance of this message type.

Formally, a limit order book can be defined as a set of tuples ("orders") of the form

$$(\text{side}, \text{price}, \text{time}, \text{size}) = (s, p, t, q) \in \{-1, 1\} \times P \times T \times Q.$$

Each published order represents an intention to buy (or sell) some quantity of an asset at a maximum (respectively, minimum) price, as illustrated in figure **??**.

$$\left(\underbrace{-1}_{\substack{\text{+1 for a buy order}\\ \text{-1 for a sell order}}}, \underbrace{84.1}_{\substack{\text{The least favourable price}\\ \text{(maximum for buy,}\\ \text{minimum for sell)}\\ \text{at which the order can trade}}}, \underbrace{9:52}_{\substack{\text{The time at which the}\\ \text{order was first published}}}, \underbrace{2}_{\substack{\text{The maximum quantity}\\ \text{of the product that will}\\ \text{be traded with this order}}}\right)$$

Figure 1.1: Components of an example limit order

1.2 The Matching Algorithm

1.2.1 The Bid and Ask

When an order insertion message is received, the exchange will attempt to form trades with the existing orders in the book such that as much volume as possible is matched. Any volume that cannot be matched will be added to the orders already in the book.

As a result of this matching, the total volume posted to the book may be depleted over time, even in the absence of cancellations. Any time a buy order has a price equal to or greater than that of a sell order, volume will be matched and a trade will occur. Consequently, the most competitive buy price (known as the *bid*) will always be less than the most competitive sell price (known as the *ask*), i.e.

$$\mathrm{bid}_{\mathcal{L}} = \max_{\{p:(1,p,t,q)\in\mathcal{L}\}} p \leq \min_{\{p:(-1,p,t,q)\in\mathcal{L}\}} p = \mathrm{ask}_{\mathcal{L}}.$$

An excessively high bid price will be depleted by traders seeking to sell the product at a premium to its true value. Conversely, an excessively low ask price will be pushed up by participants hoping to buy at a discount. It is therefore common to regard the bid and ask as lower and upper bounds respectively on the consensus fair price of the product.

Motivated by this, the *midprice* is a naive point estimate for the consensus fair price, defined by

$$\mathrm{mid}_{\mathcal{L}} = \frac{1}{2}\left(\mathrm{bid}_{\mathcal{L}} + \mathrm{ask}_{\mathcal{L}}\right).$$

Many other proxies for the consensus fair price exist, and it is common to use these as prediction targets in the construction of trading signals. I might discuss them later

1.2.2 Liquidity

At any point in time, the contents of the limit order book represent trading opportunities presented to all market participants. The abundance of these opportunities, also known as *liquidity*, represents a positive externality insofar as it allows impatient traders (*liquidity takers*) to buy or sell products precisely under those

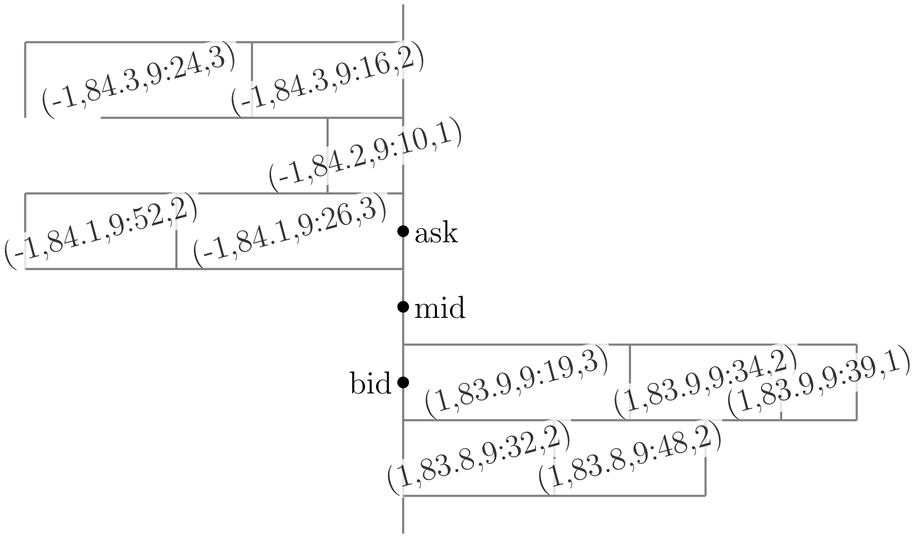


Figure 1.2: An example order book, arranged by order price and time

circumstances where it is favourable to them. Conversely, order publishers (*liquidity providers*) must adhere to the terms of any trade resulting from a posted order, regardless of whether it is in their interests.

One common measure for market liquidity is the *bid-ask spread*, defined as

$$\mathrm{spread}_{\mathcal{L}} = \mathrm{ask}_{\mathcal{L}} - \mathrm{bid}_{\mathcal{L}}.$$

Notably, if $\mathrm{mid}_{\mathcal{L}}$ is taken to represent the fair value of the product, the *half-spread*, i.e. $\frac{1}{2}\mathrm{spread}_{\mathcal{L}}$, represents the premium paid to liquidity providers by participants trading against the posted bid or ask.

1.2.3 Price-Time Priority

1.3 Price Evolution

Queue depletions are related to price changes [3]

Evolution of the limit order book over time → motivate the emphasis on modeling arrival times

1.4 Point Process Models of Book Updates

With the goal of modelling the arrival time process, we now turn to an introductory overview of point processes, and highlight some key variants described in existing literature that are relevant to this application.

For each extension: - show some example realisations for various parameter choices: barcode and intensity plot

1.4.1 Point Process Fundamentals

Need to define somewhere what non-explosiveness corresponds to formally

Given a sample space Ω with probability measure $\mathbb{P}$, a point process is any increasing random sequence of times

$$\mathcal{T} : \Omega \to {\mathbb{R}_{\geq 0}}^{\mathbb{N}}.$$

These times are often referred to as *event times*, with the implication that the sequence represents the times at which some event of interest occurs (for instance, the arrival of messages sent to a matching engine).
For any point process, there exists a corresponding nondecreasing random function $N_{\mathcal{T}} : \mathbb{R}_{\geq 0} \times \Omega \to \mathbb{N}$ known as the *counting process*, that gives the number of times having occurred before or at a given time. This is defined as

$$N_{\mathcal{T}}(t, \omega) = |\{i \in \mathbb{N} | \mathcal{T}(\omega)_i \leq t\}|.$$

All (or perhaps many) of the point processes considered in this thesis will additionally be *adapted point processes*, in the sense that $N_{\mathcal{T}}$ is a stochastic process adapted with respect to some filtration $\mathcal{F}$.

Perhaps I should define here what a filtration is?

A point process $\mathcal{T}$ may also be *marked*, in which case it is associated with one or more random sequences of marks $\mathcal{E} : \Omega \to E^{\mathbb{N}}$ drawn from a set E, typically representing additional information about each event. A common special case of this is when E is finite, in which case we may refer to the elements of E as *event types*, and refer to $\mathcal{T}$ as a *multivariate point process* with a *multivariate counting process*

$$\tilde{N}_{\mathcal{T},\mathcal{E}} : \mathbb{R}_{\geq 0} \times \omega \to \mathbb{N}^{|E|}$$

$$N_{\mathcal{T},\mathcal{E}}(t, \omega)_e = |\{i \in \mathbb{N} | \mathcal{T}(\omega)_i \leq t, \mathcal{E}(\omega)_i = e\}|.$$

1.4.2 The Expectation Measure

Let $B(\mathbb{R})$ be the Borel σ-algebra on $\mathbb{R}$. Does the σ-algebra need to be this restricted, or can it just be $\mathcal{P}(\mathbb{R})$?
For any point process $\mathcal{T}$ and for any $\omega \in \Omega$, we can define the *counting measure* Λ_ω as

$$\Lambda_\omega : B(\mathbb{R}) \to \mathbb{N}$$

$$\Lambda_\omega(S) = |S \cap \mathcal{T}(\omega)|.$$

Taking the expectation of the counting measure with respect to $\mathbb{P}$ (optionally conditional on some sub-σ-algebra of Ω) gives the *expectation measure* $\Lambda(S)$.

`https://en.wikipedia.org/wiki/Campbell%27s_theorem_(probability)`

For the sake of intuition, it should probably be shown here that $d\Lambda = d(\mathbb{E}N)$

It is also important to define the expectation measure with respect to a sigma-algebra, and in particular one from the filtration.

1.4.3 Intensity Functions

These can be formulated using the Radon-Nikodym derivative of Λ with respect to the lebesgue measure on time, i.e. $d\Lambda = \lambda dt$. For intuition, this is the arrival rate.

Existence and uniqueness shown in https://projecteuclid.org/journals/annals-of-probability/volume-24/issue-3/Stability-of-nonlinear-Hawkes-processes/10.1214/aop/106572 under restrictions on the kernel

1.5 Types of Point Processes

1.5.1 *Poisson Processes*

Probably the simplest modeling assumption we can make is to assume that the arrival rate of the process is a constant conditional on any sigma-algebra in the filtration to which the process is adapted.

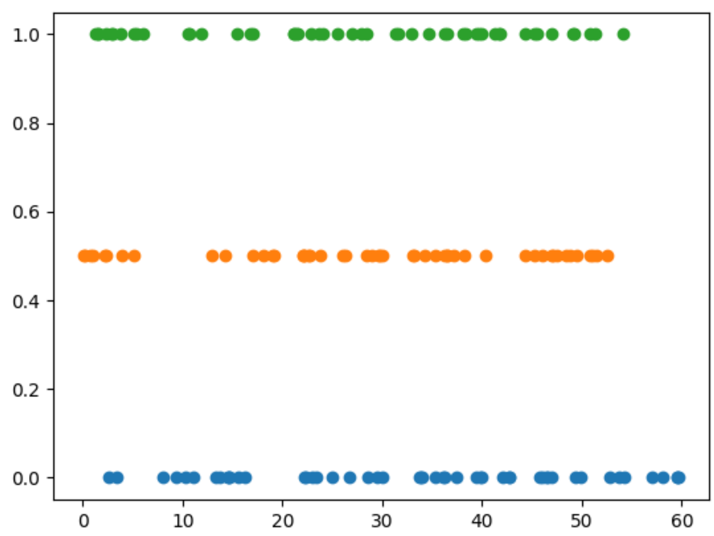
A Poisson process is a point process for which the following two properties hold.

1. For any collection of disjoint Borel sets $B \subset B(\mathbb{R})$, their couting measures $\{\Lambda_w(b) : b \in B\}$ are independent random variables.
2. For any Borel set b with finite Lebesgue measure $\mu(b)$, we have $\Lambda(b) = \lambda\mu(b) \propto \mu(b)$ for some fixed $\lambda \in \mathbb{R}$. Put differently, we have $d\Lambda = \lambda dt$.

This definition might require a reference.

Multivariate case

There is a link here to the sante fe order book model, described in the Bouchaud source, which is an example of the application of poisson arrivals.



I think I should change this to be separate plots. Can show the intensity function for each; e.g. for inhomogenous it's the same for every realisation, but for hawkes process the intensity function is itself random.

1.5.2 *Inhomogeneous Poisson Process*

An inhomogeneous Poisson process is a generalisation of the Poisson process for which the second condition is replaced by the weaker requirement that every Borel set b with finite Lebesgue measure $\mu(b)$ has a finite expectation measure. Is this weaker version even needed, or is it sufficient to just drop the condition entirely?

1.5.3 Residuals

1.5.4 Autoregressive Intensity & Hawkes Processes

Intuitively, self-excitation or self-inhibition; mutually exciting/inhibiting

Original hawkes process paper https://academic.oup.com/biomet/article-abstract/58/1/83/

immigration/birth interpretation branching matrix & complete data likelihood

1.5.5 Hawkes Processes

Long-run mean

$$\bar{\lambda} = \lim_{t\to\infty} \mathbb{E}\left[\lambda(t)\right]$$

$$\bar{\lambda} = \nu + \bar{\lambda}\int_0^\infty k(t)dt$$

Stability requires $\int_0^\infty k(t)dt < 1$. Not sure if sufficient condition.

Multidimensional:

$$\bar{\lambda} = \nu + K\bar{\lambda}$$

$$\bar{\lambda} = (I - K)^{-1}\nu$$

Requires $|\det K| < 1$ (why? and what about if there are directions orthogonal to ν?)

According to [3] discussion near their equation 9.16, hawkes processes are fully determined by the first two moments of the intensity function

1.5.6 State Dependence

Moriaru-Patrichi and Pakkanen find that conditioning on the state of the order book improves the fit of the model. (How?)

They consider conditioning on only the state of the triggering event (their 2.3), as well as conditioning on both the state of the triggering event and the current state (their 4.2).

The first of these:

$$\tilde{\lambda}_{ex}(t) = \phi_e(X(t), x)\left(\nu_e + \sum_{e'\in\mathcal{E}, x'\in\mathcal{X}} \int_{[0,t)} k_{e',e}(t-s, x')d\tilde{N}_{e'x'}(s)\right), t \geq 0, e \in \mathcal{E}, x \in \mathcal{X}.$$

The second:

$$\tilde{\lambda}_{ex}(t) = \nu_{ex} + \sum_{e'\in\mathcal{E}, x'\in\mathcal{X}} \int_{[0,t)} k_{e'x'ex}(t-s)d\tilde{N}_{e'x'}(s), t \geq 0, e \in \mathcal{E}, x \in \mathcal{X}.$$

1. markov switching 2. linear state space model / kalman filter (we'll see if i get around to figuring out how to work this)

1.5.7 Regression on marks

Multiply the kernel by $\exp(X\cdot\beta)$ where X is a vector of mark variables and β are coefficients. transfer functions should be separable, see [6.25 Definition] in https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/151886/eth-1112-02.pdf

1.5.8 Adding structure to the event space

For a finite number of event types (e.g. 10 book levels) this just corresponds to constraining the coefficient matrices.

For a continuous event space it is a bit more complicated

Another way to look at this is just having marks but the distribution of the marks is not constant - this works too. This approach is nice because it admits Kalman filtering. (Is it possible to mix Kalman filtering in event time, volume time, and wall time? Maybe better to just use wall time though.)

1.5.9 Hidden Events

Seems relevant https://www.cambridge.org/core/journals/advances-in-applied-probability/arti elementary-derivation-of-moments-of-hawkes-processes/79A7355542F08087C8AE828C664A304

Hawkes Process Inference with Missing Data https://aaai.org/papers/12116-hawkes-process-inference-with-missing-data/ makes use of monte carlo EM algorithm https://www.jstor.org/stable/2290005?seq=1

Hawkes processes with hidden marks https://www.tandfonline.com/doi/full/10.1080/135184

1.6 the most general model

State-dependent multi-kernel multivariate hawkes process with increasing kernel components and possible inhibition.

Nonparametric estimate of nu, sim study for u-shaped true nu. Probably splines, maybe kde if its fast enough. Compare to quadratic regression

The marks evolve according to state-space model involving various model variables, but have no causal impact on the point process (though obviously there is state-dependence, and the state may include the marks).

The state might be generalised to a kalman-filter-ish model [15]

$$
\begin{aligned}
IRF_{e,e',k}(s,t) &= \alpha_{e,e',k} e^{X(s)\cdot c_{e,e',k}} \exp\left(-\beta_{e,e',k}(t-s)\right) \\
\lambda_{e,e',k}(t,\omega) &= \int_0^t IRF_{e,e',k}(s,t)\, d\Lambda_{e',\omega}(s) \\
\lambda_e(t,\omega) &= \sum_{e'} \sum_k \lambda_{e,e',k}(t,\omega) + \text{spline intensity}(t)
\end{aligned}
$$

Chapter 2
Estimation

2.1 MLE

- What is a likelihood function
  - Derive likelihood function for the model

$$\ell(\alpha, \beta, c) = \sum_e \left(\int_0^T \log(\lambda_e(t)) d\Lambda_{e,\omega}(t) - \int_0^T \lambda_e(t) dt \right)$$

Clearly λ_e must be positive at the event times.
At stationary points, loglikelihood has zero derivatives.

$$\begin{aligned}
\frac{\partial \ell(\alpha, \beta, c)}{\partial \alpha_{e,e',k}} &= \int_0^T \frac{\partial \lambda_e(t)}{\partial \alpha_{e,e',k}} \frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - \int_0^T \frac{\partial \lambda_e(t)}{\partial \alpha_{e,e',k}} dt \\
&= \int_0^T \frac{\partial \lambda_e(t)}{\partial \alpha_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right) \\
&= \int_0^T \frac{\partial \lambda_{e,e',k}(t)}{\partial \alpha_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right) \\
&= \int_0^T \frac{\lambda_{e,e',k}(t)}{\alpha_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell(\alpha, \beta, c)}{\partial \beta_{e,e',k}} &= \int_0^T \frac{\partial \lambda_e(t)}{\partial \beta_{e,e',k}} \frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - \int_0^T \frac{\partial \lambda_e(t)}{\partial \beta_{e,e',k}} dt \\
&= \int_0^T \frac{\partial \lambda_e(t)}{\partial \beta_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right) \\
&= \int_0^T \frac{\partial \lambda_{e,e',k}(t)}{\partial \beta_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right) \\
&= \int_0^T \left(\int_0^t (s-t) IRF_{e,e',k}(s,t) d\Lambda_{e',\omega}(s) \right) \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right)
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial \ell(\alpha,\beta,c)}{\partial c_{e,e',k}} &= \int_0^T \frac{\partial \lambda_e(t)}{\partial c_{e,e',k}} \frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - \int_0^T \frac{\partial \lambda_e(t)}{\partial c_{e,e',k}} dt \\
&= \int_0^T \frac{\partial \lambda_e(t)}{\partial c_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right) \\
&= \int_0^T \frac{\partial \lambda_{e,e',k}(t)}{\partial c_{e,e',k}} \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right) \\
&= \int_0^T \left(\int_0^t X(s) IRF_{e,e',k}(s,t) d\Lambda_{e',\omega}(s) \right) \left(\frac{1}{\lambda_e(t)} d\Lambda_{e,\omega}(t) - dt \right)
\end{aligned}
$$

Explain why no explicit solution

Iterative methods: find an iteration whose fixed points are stationary points. Is there a way to show that this converges?

2.2 EM

2.2.1 Information Theory and EM

Can cite original dempster paper? As well as the one that corrects the errors in that paper

2.2.2 EM for Hawkes Processes

Branching matrix vs exo/endo classification. Usefulness for multikernel and other limitations of gradient based MLE.

$$
B_{e,e',k}(s,t) = \frac{IRF_{e,e',k}(s,t)}{\lambda_e(t)}
$$

$$
\sum_{e'} \sum_k \int_0^t B_{e,e',k}(s,t) d\Lambda_{e',\omega}(s) = 1
$$

Branching matrix:

$$
B_{i,j,k} = \frac{IRF_{e_j,e_i,k}(t_i,t_j)}{\lambda_{e_j}(t_j)}
$$

- Visualisation of branching matrix. How does it evolve throughout the fitting procedure? - Quasi EM approximation - Closed form constant time M step - Negative probabilities. Conditions for kernel nonnegativity? (Probably not tractable)

2.2.3 Generalised Fixed Point Methods

Is it a contraction mapping?

2.2.4 *Method of Scoring (Newton's Method)*

2.3 Uncertainty quantification for MLE

Asymptotic normality

Parametric Bootstrap

2.4 Inference for the mark process

2.5 Computational Concerns

Sensor fusion for parallelisation

Momentum (analyse autocorrelation of parameter changes throughout the learning process)

Exploiting sparsity [14]

2.6 Bayesian Approach

- Uncertainty quantification - Less likely to get stuck in local optima

2.7 Model Selection

Performance of information criteria for selection of Hawkes process models of financial data https://www.tandfonline.com/doi/full/10.1080/14697688.2017.1403140

Chapter 3

Generative Sampling

3.1 Simulation Methods

Immigration-birth interpretation [13]

Ogata thinning [13]

Do these have different time complexity? Memory complexity?

3.2 Simulation Study of Estimation Methods

Convergence analysis for simple models (eg univariate)

3.3 Impulse Response Function

Causal analysis, price impact of orders.

Chapter 4
Application to KOSPI/SPY Data

Hawkes processes and their applications to finance: a review https://www.tandfonline.com/doi/
https://www.tandfonline.com/journals/rejf20/collections/Hawkes-Processes-in-Finance

- Basic poisson model
- Univariate hawkes
- Multivariate hawkes - what are the event types?
- State-dependent hawkes - which state variables are relevant?
- Full model
- Replicate findings from [13]
- Hidden events (& events on different exchanges) - either poisson distributed or more complex
- Modeling changes in the entire order book
- Market impact (are there any datasets on this? square-root law, other common findings. power law impact for hawkes processes is explicitly studied here `https://arxiv.org/pdf/1805.07134`)
- Midprice change prediction/explanation - explicit formula or simulation?
- Realised volatility prediction
- Correlated products (with low beta, preferably - or see what is done in literature studies of correlated products)
- Options (if I can get data) - would give lots of (nonlinearly) correlated products. Can estimate the correlation between products at any point in time using factor loadings & historical factor correlations. Here is one source: `https://www.nber.org/system/files/working`$_{p}apers/w29369/w29369.pdf.Optionspri$ `https://arxiv.org/abs/1602.03043`$Closedformforoptimalexecutionofsignals(incluc$ *impact*)
- This paper arxiv.org/pdf/2401.11495 shows a functional limit theorem for hawkes processes behaving as integrated CIR processes. These are a popular volatility model so this makes sense!

`https://quant.stackexchange.com/questions/59593/what-are-some-currently-oper`
`https://papers.ssrn.com/sol3/papers.cfm?abstract`$_{i}d = 4844711 OptionPricingUsingH$

`A slightly depressing jump model: intraday volatility pattern simulation https://www.tandfonline.com/doi/full/10.1080/14697688.2017.1403139 Implementat and evaluation of the Heston-Queue-Hawkes option pricing model https://uu.diva-`

4.0.1 Clustering Ratio

`Variance of interevent time divided by expectation. Bouchaud p 165. Should this section be elsewhere? Does this mean something for residuals too?`

Databento info: CME: ES MES OPRA: SPX - Index options SPY - Index ETF options https://databento.medium.com/getting-futures-tick-sizes-and-notiona

Chapter 5

Conclusion

This is the conclusion

References

[1] Palley, T. I., *Financialization: What It Is and Why It Matters*, Working Paper No. 525, The Levy Economics Institute, December 2007. Paper presented at the conference on \Finance-led Capitalism? Macroeconomic Effects of Changes in the Financial Sector," sponsored by the Hans Boeckler Foundation, Berlin, Germany, October 26{27, 2007. Available at: https://www.levyinstitute.org/pubs/wp_525.pdf

[2] Balakrishnan, S., Wainwright, M. J., and Yu, B., *Statistical guarantees for the EM algorithm: From population to sample-based analysis*, *The Annals of Statistics*, 45(1) (2017), 77--120. doi = 10.1214/16-AOS1435 various facts about EM algorithm in general, useful in particular because it shows the general principle that likelihood gradient = EM gradient. balakrishnan, wainwright, yu page 82 as cited at https://stats.stackexchange.com/questions/45652/what-is-the-difference-be

[3] Bouchaud, J.-P., Bonart, J., Donier, J., and Gould, M., *Trades, Quotes and Prices: Financial Markets Under the Microscope*, Cambridge University Press, 2018. ISBN = 9781107156050

[4] Chen, F. and Stindl, T., *Direct Likelihood Evaluation for the Renewal Hawkes Process*, *Journal of Computational and Graphical Statistics*, 27(1) (2018), 119--131. doi = 10.1080/10618600.2017.1341324

[5] Daley, D. J. and Vere-Jones, D., *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*, 2nd edition, Springer, 2003. doi = 10.1007/b97277

[6] Embrechts, P., Liniger, T., and Lin, L., *Multivariate Hawkes processes: An application to financial data*, Journal of Applied Probability, 48(A) (2011), 367--378. doi = 10.1239/jap/1318940477

[7] Financial Industry Regulatory Authority, *2024 Industry Snapshot: Market Data*, FINRA, 2024. Available at: https://www.finra.org/media-center/reports-studies/2024-industry-snapsho Accessed: 2 September 2024.

[8] Jamshidian, M. and Jennrich, R. I., *Acceleration of the EM Algorithm by Using Quasi-Newton Methods*, *Journal of the Royal Statistical Society. Series B (Methodological)*, 59(3) (1997), 569--587. Available at: http://www.jstor.org/stable/2346010

[9] Jamshidian, M. and Jennrich, R. I., *Standard errors for EM estimation*, *Biometrika*, 89(1) (2002), 63--75. doi = 10.1111/1467-9868.00230

[10] Jiang, A. Z. and Rodriguez, A., *Improvements on Scalable Stochastic Bayesian Inference Methods for Multivariate Hawkes Processes*, *Statistics and Computing*, 34, article 85 (2024). doi = 10.1007/s11222-024-10392-x

[11] Laub, P. J., Lee, Y., and Taimre, T., *The Elements of Hawkes Processes*, Springer, 2021. doi = 10.1007/978-3-030-84639-8

[12] Lewis, E. and Mohler, G., *A nonparametric EM algorithm for multiscale Hawkes processes*, *Journal of Nonparametric Statistics*, 1(1) (2011), 1--20. one example of EM algorithm applied to point processes. may or may not deserve a direct citation, but can see if anything cited here is useful

[13] Morariu-Patrichi, M. and Pakkanen, M. S., *State-Dependent Hawkes Processes and Their Application to Limit Order Book Modelling*, Quantitative Finance, 22(3) (2022), 563--583 doi = 10.1080/14697688.2021.1983199

[14] Nickel, M. and Le, M., *Learning Multivariate Hawkes Processes at Scale*, arXiv preprint arXiv:2002.12501 [cs.LG], 2020. doi = 10.48550/arXiv.2002.12501

[15] Smith, A. C. and Brown, E. N., *Estimating a state-space model from point process observations*, *Neural Computation*, 15(5) (2003), 965--991. doi = 10.1162/089976603765202622

[16] Veen, A. and Schoenberg, F. P., *Estimation of Space-Time Branching Process Models in Seismology Using an EM-Type Algorithm*, *Journal of the American Statistical Association*, 103(June) (2008), 614--624. doi = 10.1198/016214508000000148

[17] Zhou, F., Li, Z., Fan, X., Wang, Y., Sowmya, A., and Chen, F., *Efficient inference for nonparametric Hawkes processes using auxiliary latent variables*, *The Journal of Machine Learning Research*, 21(1) (2020), 9745--9775.
https://www.idescat.cat/sort/sort461/46.1.1.Worrall-etal.pdf
https://www.tandfonline.com/doi/full/10.1080/10618600.2022.2050247
https://www.santafe.edu/research/results/working-papers/studies-of-the-li
https://arxiv.org/pdf/1903.03223
https://hal.science/hal-01121711v5/document
https://link.springer.com/chapter/10.1007/978-88-470-1766-5$_4$

https://www.jstor.org/stable/2334319
https://www.tandfonline.com/doi/full/10.1080/14697688.2021.1983199
https://www.tandfonline.com/doi/full/10.1080/10618600.2017.1341324
https://www.researchgate.net/publication/4742983$_{E}stimation_{o}f_{S}pace{T}ime_{B}ranching_{P}$
https://pubmed.ncbi.nlm.nih.gov/12803953/
http://paleo.sscnet.ucla.edu/Lewis-Molher-EM$_{P}reprint.pdf$
https://opus.lib.uts.edu.au/bitstream/10453/145676/2/19-930.pdf
https://projecteuclid.org/journals/annals-of-statistics/volume-45/issue-1/Stat
https://www.tandfonline.com/doi/full/10.1080/1351847X.2021.1917441

https://projecteuclid.org/journals/annals-of-probability/volume-24/issue-3/Stal
Data Sketching for Large-Scale Kalman Filtering
https://arxiv.org/pdf/1606.08136