

Point Process Modeling of Limit Order Books

Oden Petersen

UNSW Mathematics and Statistics

Graph

normally modeled with time series models. eg autoregression, state models, garch

these models have had some success but don't describe the process of price formation from first principles

Limit Order Book

Markets are venues for buyers and sellers to come together to exchange financial assets

Limit order book diagram

Modeling individual order arrivals, cancellations, and trades allows us to simulate the evolution of the limit order book

Order Components

$$\left(\underbrace{-1}, \underbrace{84.1}, \underbrace{9:52}, \underbrace{2} \right)$$

+1 for a buy order The least favourable price The time at which the The maximum quantity of the product that will
 -1 for a sell order (maximum for buy, minimum for sell) order was first published be traded with this order
 at which the order can trade

Order Book Events

For order insertions, cancellations we need to model size, price, side, and time.

First three can be modeled with standard time series tools. But modeling arrival time process is nontrivial. Some parts of the day have lots of arrivals, others don't.

graph of intraday activity bucketed by minute also show mean qty and side bucketed by minute

Inhomogenous Poisson Process Model

graph of intraday activity bucketed by minute
overlay maxlikelihood spline intensity
counting process overlaid with integral $\bar{\Lambda}$ of intensity

Likelihood function for Point Process Models

Explain why the loglikelihood formula is reasonable

Why integral of intensity should be iid exponential
Show that it isn't \rightarrow motivate autoregressive intensity

Hawkes Processes

Formula for intensity

kernel is linear combination of exponential functions. can approximate analytic functions on $[0, t]$ uniformly well with enough exponential components by weierstrass approximation theorem.

Can treat events as coming from a mixture of multiple point processes. As is common with mixture models, use an EM algorithm to fit.

Explain how to get branching matrix.

State Dependence

Book imbalance vs time until next buy event / time until next sell event

Motivate state dependence

Formula for state dependence - discrete and continuous

Maybe some more details on book imbalance

Point Process Simulation

Detail Ogata algorithm

Original paper uses markov state switching. I can use time series model for orders and limit order book simulation to make this a bit more sophisticated.

Model Results

Parametric bootstrap

IC by model

crossvalidation loglikelihood

actual vs expected counts

residual distribution and autocorrelation and ks pval table

replication number and distance to average parent for various kernels

Price Impact Function

Graph of impact vs size
VWAP vs TWAP vs

EM and branching matrix details

Future directions

Options markets

Metaorder modelling via random kernels each day