1 Orderbook Shape Process

Our orderbook shape process follows the same dynamics as the conditional intensity for the hawkes process. We have

$$\begin{cases} dk^+ = \beta_k(\theta_k - k^+)dt + n_k d\overline{M_t^+} + v_k d\overline{M_t^-} \\ dk^- = \beta_k(\theta_k - k^-)dt + v_k d\overline{M_t^+} + n_k d\overline{M_t^-} \end{cases} \tag{1}$$

Using our model for k that we will call $\hat{k}$ (refer to signal measurment, we will look at a bunch of orderbook and trade data to describe how liquidity changes as distance from the mid price changes), we seek to find the parameters in (31) to fit our emperical measurements the best. Using our work from the Hawkes Processes, I define the objective

$$\sum_{OBS}(\hat{k} - \theta_k - \sum_{t_i<t} B_i e^{-\beta(t-t_{influential})})^2 \tag{2}$$

Note that computing the objective as defined here is of quadratic complexity in the number of influential trades. In practice this objective needs to be computed recursively using the definition of dk. The intuition behind this choice of orderbook shape process is this: when a buy or sell market order arrives, other market makers are going to react to this information by place quotes closer to the mid price in anticipation of more orders that are on the way (market orders are self-exciting). Once the spike is over, market makers will withdraw their liquidity, motivating the exponential decay seen in this process.

1.1 K Process Measurement

We ultimately want to determine the impact of the arrival of orders on the probability that our quotes get filled as a function of the mid price. We need to be able to update our quotes immeadiately, so need to have some idea from history of the impact that a MO arrival has on liquidity. The first step of doing this is getting estimates for k^+ and k^- over periods that are long enough that enogh orders arrive, but short enough to allow intra-hourly changes. I did my fitting on a 5 minute interval. One way of estimating k from this data is to record hits at two deltas δ_1 and δ_2, and solve this system of equations.

$$\begin{cases} \lambda_n(\delta_1) = Ae^{-k\delta_1} \\ \lambda_n(\delta_2) = Ae^{-k\delta_2} \end{cases} \tag{3}$$

An additional way of doing this is to record fills at more deltas, and then to perform a linear regression / least squares fitting of $Ae^{-k\delta}$ to the imperical lambdas. Over each 5 minute interval, record measurements for k using this process. Over these same intervals, we also want to include several other quantities that we expect to have an impact on the value of k that we meausre. We want to record:

1. The average spread during this interval
2. Bid / ask density
3. The coefficients a, b in $v(\delta) = a\delta^{\pm} + b$, where v denotes the volume of orders present on the order book.

These are some rudimentary features that we can extract from the orderbook that describe it's shape. Intuitively, we expect $\hat{k}$ to be inversely related to spread. I also expect a roughly linear relationship between bid and ask density, and $a^{\pm}, b^{\pm}$ and $\hat{k}$. Performing a rudimentary least-squares curve fitting, I get a high R^2 of . Using these same features, a residual neural network can be fitted to predict the residuals of this model, causing a marginal but measurable improvement in R^2. Ultimately, we can perform this same procedure without the rudimentary feature extraction and feed entire orderbooks (or top n levels) of the orderbook to some neural network to predict $\hat{k}$ in a similar fashion. Once we have this model fitted, we will fit our super MLE estimator on its outputs at each instant in time.

1.2 Resources

1. Das, Sanjiv R. “The surprise element: Jumps in interest rates.” Journal of Econometrics, vol. 106, no. 1, 2002, pp. 27–65, https://doi.org/10.1016/s0304-4076(01)00085-9.

2. Rasmussen, Jakob Gulddahl. “Lecture Notes: Temporal Point Processes and the Conditional Intensity Function.” arXiv.Org, 1 June 2018, arxiv.org/abs/1806.00221.

3. Laub, Patrick J., et al. “Hawkes Processes.” arXiv.Org, 10 July 2015, arxiv.org/abs/1507.02822.

4. Aït-Sahalia, Yacine, et al. “How and When Are High-Frequency Stock Returns Predictable?” NBER, 22 Aug. 2022, www.nber.org/papers/w30366.

5. Cartea, Álvaro, et al. “Buy Low Sell High: A High Frequency Trading Perspective.” SSRN, 26 Nov. 2011, `deliverypdf.ssrn.com/delivery.php?ID=70102006811110210409411909407311510605005102600703401000 amp;EXT=pdf&INDEX=TRUE`.

6. Sophie Laruelle - Events.Chairefdd.Org, `events.chairefdd.org/wp-content/uploads/2013/06/CAHIER_MICRO_1.pdf`. Accessed 18 Oct. 2023.