Limit Order Book Dynamics

Our goal is to use the dynamics of the Limit Order Book (LOB) as an indicator for high-frequency stock price movement, thus enabling statistical arbitrage. Formally, we will the study limit order book imbalance process, $I(t)$, and the stock price process, $S(t)$, and attempt to establish a stochastic relationship $\dot{S} = f(S, I, t)$. We will then attempt to derive an optimal trading strategy based on the observed relationship.

Recap Next Steps

1. Complete in-sample backtesting of the 'naive' trading strategies.
2. Formulate stochastic control problem
3. Extra Reading: Bellman Equations, MDP, Partially Observable MDP

In-Sample Backtesting of Naive Trading Strategies

As a refresher:

We are a considering a CTMC for the joint distribution $(I(t), \Delta S(t))$ where $I(t) \in \{1, 2, \ldots, \#_{bins}\}$ is the bin corresponding to imbalance averaged over the interval $[t - \Delta t_I, t]$, and $\Delta S(t) = \text{sign}(S(t+\Delta t_S) - S(t)) \in \{-1, 0, 1\}$, considered individually for the best bid and best ask prices. The pair $(I(t), \Delta S(t))$ was then reduced into one dimension with a simple encoding.

From the resulting timeseries we estimated a generator matrix $\boldsymbol{G}$ and used it to obtain a one-step transition probability matrix $\boldsymbol{P} = e^{\boldsymbol{G}\Delta t_I}$. The entries of $\boldsymbol{P}$ contain the conditional probabilities $\mathbb{P}\left[\rho_{curr}, \Delta S_{curr} \mid \rho_{prev}, \Delta S_{prev}\right]$, from which we can solve for the probability of now seeing a given price change (ΔS_{curr}) conditional on the current imbalance, the previous imbalance, and the previous price change.

For example, one such conditional probability matrix $\boldsymbol{P_C}$ (using 3 imbalance bins) was:

$$\begin{array}{rc}
 & \overbrace{\hspace{9em}}^{\rho_n = 1} \quad \overbrace{\hspace{9em}}^{\rho_n = 2} \quad \overbrace{\hspace{9em}}^{\rho_n = 3} \\
\begin{array}{r} \Delta S_n{<}0\rightarrow \\ \Delta S_n{=}0\rightarrow \\ \Delta S_n{>}0\rightarrow \end{array} &
\begin{bmatrix}
.67 & .05 & .04 & .01 & .03 & .04 & .00 & .05 & .05 & .02 & .50 & .12 & .01 & .00 & .02 & .05 & .01 & .02 & .00 & .00 & .52 & .00 & .01 & .00 & .00 & .00 & .00 \\
.33 & .95 & .96 & .99 & .97 & .96 & .41 & .93 & .95 & .96 & .49 & .87 & .98 & .99 & .97 & .91 & .48 & .96 & .98 & .95 & .47 & .95 & .96 & .93 & .98 & .88 & .34 \\
.00 & .00 & .00 & .00 & .00 & .00 & .58 & .02 & .00 & .02 & .01 & .00 & .01 & .01 & .01 & .05 & .51 & .01 & .02 & .04 & .01 & .05 & .03 & .02 & .02 & .12 & .66
\end{bmatrix} \\
 & \underbrace{\hspace{3em}}_{\Delta S_{n-1} < 0} \underbrace{\hspace{3em}}_{} \underbrace{\hspace{3em}}_{\Delta S_{n-1} > 0} \underbrace{\hspace{3em}}_{} \underbrace{\hspace{3em}}_{\Delta S_{n-1} = 0} \underbrace{\hspace{3em}}_{} \underbrace{\hspace{3em}}_{} \underbrace{\hspace{3em}}_{} \underbrace{\hspace{3em}}_{}
\end{array}$$

Immediately evident from $\boldsymbol{P_C}$ is that in most cases we are expecting no price change. In fact, the only cases in which the probability of a price change is > 0.5 show evidence of *momentum*; for example, the way to interpret the value in row 1, column 1 is: if $\rho_{prev} = \rho_{curr} = 1$ and previously we saw a downward price change, then we expect to again see a downward price change. In fact, the best way to summarize the matrix is:

$$\mathbb{P}\left[\Delta S_{curr} = \Delta S_{prev} \mid \rho_{prev} = \rho_{curr}\right] > 0.5$$

We backtested a number of naive trading strategies, outlined below, based on this observation.

Algorithm 1 Naive Trading Strategy

$cash = 0$
$asset = 0$
for $t = 2$: `length`$(timeseries)$ **do**
  if $\mathbb{P}\left[\Delta S_{curr} < 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{prev}\right] > 0.5$ **then**
    $cash \mathrel{+}= data.BuyPrice(t)$
    $asset \mathrel{-}= 1$
  else if $\mathbb{P}\left[\Delta S_{curr} > 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{prev}\right] > 0.5$ **then**
    $cash \mathrel{-}= data.SellPrice(t)$
    $asset \mathrel{+}= 1$
  end if
end for
if $asset > 0$ **then**
  $cash \mathrel{+}= asset \times data.BuyPrice(t)$
else if $asset < 0$ **then**
  $cash \mathrel{+}= asset \times data.SellPrice(t)$
end if

Algorithm 2 Naive+ Trading Strategy

$cash = 0$
$asset = 0$
for $t = 2$: `length`$(timeseries)$ **do**
  if $\mathbb{P}\left[\Delta S_{curr} < 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{prev}\right] > 0.5$ **then**
    $cash \mathrel{+}= data.BuyPrice(t)$
    $asset \mathrel{-}= 1$
  else if $\mathbb{P}\left[\Delta S_{curr} > 0 \mid \rho_{curr}, \rho_{prev}, \Delta S_{prev}\right] > 0.5$ **then**
    $cash \mathrel{-}= data.SellPrice(t)$
    $asset \mathrel{+}= 1$
  end if
end for
if $asset > 0$ **then**
  $cash \mathrel{+}= asset \times data.BuyPrice(t)$
else if $asset < 0$ **then**
  $cash \mathrel{+}= asset \times data.SellPrice(t)$
end if