Intraday Trading Volume Modeling and Forecasting (PD3-14)

Chen, Ran

Prof. Palomar, Daniel

Project Overview

1. Objectives
The explosion of algorithmic trading has been one of the most remarkable trends in the financial industry in recent years. The key ingredient of algorithmic trading is the prediction of intraday trading volume. This Final Year Thesis explored existing literature on this area, and proposed a new methodology based on stochastic process modeling to outperform the existing models.

2. What Is Algorithmic Trading?
Algorithmic trading is the process of using electronic platforms that utilize very advanced mathematical models to make automated transaction decisions in financial markets. The underlying models attempt to determine the optimal time for placing an order so that the least amount of impact on a stock price will be caused.

Algorithmic trading provides a more systematic approach to active trading than methods based on a human trader's intuition or instinct.

3. Why is Volume So Important?
Indeed, intraday trading volume is becoming a key of the most important research areas in the financial industry, as it is essential for high frequency and algorithmic trading strategies. The major reason of volume precision is to minimize the transaction cost and reduce market impact during the execution process via optimally placing orders, thereby improving the performance of automated trading strategies and minimizing execution risk, especially in the cases of large blocks of stocks and illiquid trades.

Methodology

1. Forecasting Model: Kalman Filter
The general assumption of our proposed model is that the intraday trading volume can be decomposed into three components: daily average component, intraday periodic component and intraday non-periodic component. We considered the logarithm of volume as observation and the logarithm of these components as hidden state. Therefore, we can apply Kalman Filter to model them.

2. Inference: Expectation-Maximization Algorithm
Expectation-Maximization (EM) Algorithm is used as the parameter estimation methodology in our project. It is an iterative method for finding maximum likelihood estimates of parameters. The EM iteration alternates between performing an expectation (E) step, which constructs the lower bound of the expectation of log-likelihood using the current estimation of parameters, and a maximization (M) step, which computes parameters maximizing that lower bound found from E-step. These parameter estimates are then used to determine the distribution in the next E-step. Through this iterative way of estimation, the parameters are able to gradually converge to their true values and are ready to be used in Kalman Filter Modeling.

Results

1. Hidden Components Estimation
The results of the hidden state estimation from the synthetic experiments are shown above, from which we can see that these two hidden components can be precisely estimated. For the intraday periodic component, it is treated as parameters and can be inferred accurately by EM Algorithm.

2. Raw Trading Volume Prediction
We conducted empirical experiments using real data of Exchange-Traded Funds (ETFs) and SPY. The results below are the out-of-sample prediction results. The left figure is the trend of the predicted values and the right figure is the trend of the residuals. The prediction error with real data is in the table below.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Kalman</th>
<th>Filter</th>
<th>ARMMA</th>
<th>Delayed</th>
<th>Rolling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic</td>
<td>0.41</td>
<td>0.47</td>
<td>0.51</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>Real Result (SPY)</td>
<td>0.25</td>
<td>0.35</td>
<td>0.38</td>
<td>0.69</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion

From the synthetic experiments and empirical experiments, we reached our final conclusion that this proposed methodology using Kalman Filter and EM Algorithm is able to significantly outperform existing forecasting models and delivers more accurate predictions.