Valuable Algorithms:
A Comparison of Financial and Computational Performance of Stock Trading Algorithms

Hautala, Jesse

Howard, Gregory

MacAvoy, Maxwell

Okara, Chinemeremma

November 23, 2022

Context

In recent years there has been an explosion of algorithmic trading activity of equities in public markets. This is clearly visible through books, like Scott Patterson's The Quants, automated trading platforms, like Composer, and so-called "Robo-Advisors", like Wealthfront. One of the first moments where it became clear to the public that algorithmic trading had created a new paradigm for the public markets was when stock indices collapsed and rebounded at a striking pace on May 6, 2010, in what came to be known as the "Flash Crash". In the aftermath of this event, regulators actually banned the use of three types of trading algorithms (i.e. spoofing, layering, and front-running algorithms).[1]

It has been shown that the majority of hedge funds, who are reliant on these quantitative trading techniques, fail to outperform the broader S&P 500 index.[2] However, there is no doubt that, if one were to simply buy that same index at its local minima and sell at its local maxima, one can outperform the index itself.

---

[1] Bates, John. *Traders Magazine Online News.* "Post Flash, Regulators Still Use Bicycles to Catch Ferraris." Source.
[2] Perry, Mark. *AEI.* "The SP 500 Index Out-performed Hedge Funds over the Last 10 Years. And It Wasn't Even Close." Source.

Question

So our question is simple. With our newfound knowledge of algorithms, can we do better? Or more explicitly:

> What is the best algorithm for trading the S&P 500 index and how does the time complexity of this algorithm compare to alternatives?

Motivation

Collins

The New York Stock Market is globally the biggest, with a market capitalization of $26.2 trillion. Interestingly, about 145 million American adults own stock. Over the years, it has been pleasant to hear how algorithms are changing the complexity of trading to just zeroes and ones (BUY or SELL). It is interestingly good to work on a project not just for the justification of this course but for the adventures of knowing how algorithms are used to determine or predict price movement. I am geared toward this project because it has been long overdue in learning the nitty gritty of algorithm-based trading.

Greg

I've been trading stocks for several years and am lucky enough to be paying for this graduate education largely with gains from one highly successful and outsized stock trade between November 2019 and January 2020. To date, I have largely considered trades based on fundamental research of the financial and operational situation of a given company. However, I'm interested in learning more about automated trading based solely on price movement. The word on the street is that, when done right, such algorithms can produce consistent and compounding returns. So I want to learn more about what would make a trading algorithm the right algorithm. This project excites me, because it provides a strong basis to explore that curiosity and apply the methods of this course directly to a real-world application.

Jesse

Having experience in sound production and recording (mostly music), where audio channels consist of one-dimensional time series data, I would like to increase my familiarity with tools and techniques for analyzing and processing one-dimensional time series data. I am especially interested to observe any intersection of relevant techniques between the two superficially disparate domains. For example, a simple low-pass filter could be implemented as a moving average and the ARIMA model uses moving averages (a form of convolution) to forecast. As a student, I look forward to learning more about the stock trading domain as an example of a one-dimensional, non-stationary time series dataset.

Max

Our group decided to focus our final project and general research question around stock trading algorithms. We set out with one major goal in mind; work on something that may be actually useful in daily life and that allows us to all participate and show our skills learned in this class. The stock market is present in our lives as well as the lives of nearly 60% of Americans who are directly invested in it and any sort of algorithm we create could in theory better direct our future investing.

The stock market is a relatively easily understood space and therefore we are able to spend more time working on the technical aspect of the project instead of working to learn an entirely new area. We decided that a great way to keep all group members involved and build a little healthy competition was to have all four of us develop our own algorithm and pit them against each other.

I have been involved in the stock market since I was 18 and it has been a focus of interest for me for many years. At the beginning of college I worked with an Alumni to backtest and develop an Index fund based off of the number of patents a company had filed for. We did not use any types of real algorithms in our work but I hope to be able to transfer some of the ideas from that project into this final project.

Scope

We intend to analyze performance differences among a set of trading algorithms. We will assess the time/memory performance of applying these algorithms and also the real world outcomes, or financial performance. In the interest of healthy competition and distribution of work, we've decided to each develop multiple algorithms, at least one per contributor, and evaluate relative performance financially (in USD) and computationally (time and memory requirements).

To limit complexity, we will work with daily index level pricing based on the prices recorded at the open and close of the market. We will assume our algorithm is capable of trading exactly at these recorded numbers, even though we know trading instantaneously is not quite possible and that fluctuation at the open and close is significant. We will analyze USD performance of each trading algorithm over the past 36 months of data, and assume this historic analogue is the best available test of future performance.

Note: A critical limitation is that our models will only consume *historical* data on which to make buy/sell decisions. This is fundamental, but is worth stating nonetheless.

Description

We have not yet selected a data source, but we have briefly discussed some alternative models (e.g. ARIMA) and the rules by which these algorithms will be applied (see "Scope" section above).