FRTN70 - Project in systems, control and learning

Fall 2021, 7.5 hp

Title

Trading Swedish stocks using optimization and learning

Recommended prerequisites

- Familiarity with optimization
- An introductory course in statistics or machine learning
- No previous courses or experience in finance is required

Project description

There are a plethora of financial instrument one can trade - stocks, bonds, currencies, cryptocurrencies, commodities, options, futures, etc. In this project, we restrict ourselves to the stocks that constitute the OMX Stockholm 30 stock market index[1]. The goal is to manage a portfolio by picking among these stocks. We do this by simulating trades using historical data in a backtesting[2] environment.

Fundamentally, there are three factors that determine the success of a systematic trading strategy:

1. Information advantage - plenty of data with a high signal-to-noise ratio
2. Superior models - having models that extract the most signal from the data
3. Computational advantage - the computational infrastructure to efficiently create and execute on predictions by the models

Data collection and cleaning

Relevant data sources include price data, order book data, macroeconomic data, financial statements, text data from e.g. news articles and twitter posts, images such as satellite images, etc. In this project, we start looking at daily open, high, low and close prices, and volume data for the stocks in the OMXS30 stock market index.

Collection: The data can be found on e.g. Yahoo Finance[3] for free by searching each company name and downloading it as a `csv`-file. We recommend using `DataFrames.jl` or `pandas` in `Julia` or `Python`, respectively, to store and manipulate the data.

Cleaning: There might exist missing data, outliers or other anomalies. Those cases need to be handled appropriately.

[1] `http://www.nasdaqomxnordic.com/index/index_info?Instrument=SE0000337842`
[2] `https://www.investopedia.com/terms/b/backtesting.asp`
[3] `https://finance.yahoo.com/`

Trading environment

We will use the basic method, or trading environment, found in Boyd et al. (2017) for evaluating the performance of our trading strategies. The method allows us to trade off expected return, risk, transaction cost and holding cost by solving a convex optimization problem each time we rebalance our portfolio. In this project we consider daily rebalancing.

First, the group needs to decide which holding and trading constraints they wish to employ. This will result in a particular convex optimization problem. Groups are welcome to implement their own solver with help from the supervisor. However, groups are also allowed to use common frameworks for solving convex optimization problems, e.g. `Convex.jl` or `CVXPY` in `Julia` or `Python`, respectively. These frameworks make it very easy to specify and modify convex optimization problems in just a handful of lines of easy to understand code.

Boyd et al. (2017) presents a single-period version and a generalization in a multi-period version. The single-period version traces back to Markowitz (1952); the multi-period version trace back to model predictive control. However, the method does not address the critical issue of predictions or forecasts of future quantities and must be dealt with separately. This is were statistics and machine learning becomes useful in this task.

Predictions

The prediction of portfolio risk, transaction cost and holding cost are comparatively easy quantities to estimate. The return of stocks on the other hand are notoriously hard to predict. In fact, asset managers essentially build their business based on their highly proprietary return forecasts. The forecasts are generated using many methods, ranging from analyst predictions to sophisticated techniques from statistics and machine learning.

This part of the project is open-ended. Students are encouraged to explore new own ideas or look into the literature. However, we would like to raise a finger of warning. Many novice practitioners achieve exceptional risk-adjusted returns in backtesting, by either violating causality or grossly overfitting to the historical period, but with horrible forward performance.

Examination

Project meetings are held regularly during the course. In the project the students must search for knowledge and information independently. The projects results and experiences are reported both in written and oral form.

Project supervisor

Manu Upadhyaya, `manu.upadhyaya@control.lth.se`

References

Boyd, S., Busseti, E., Diamond, S., Kahn, R. N., Koh, K., Nystrup, P. and Speth, J. (2017), 'Multi-period trading via convex optimization.'.

Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.