Project Proposals for MS&E 448

Spring Quarter 2019

Dr. Lisa Borland

1 Build a High Frequency Strategy

Description

The goal of this project is to design and implement a trading strategy based on limit order book data, i.e. high frequency stocks data, that exploits the specific structure of such datasets.

Initial questions

- What is a good time scale to consider: tick-by-tick or a larger time-scale (for example, can an integrated order book profile predict anything over longer horizons)?
- Use machine learning techniques or impose a fundamental relationship
- Discuss and analyze execution tactics (eg if you are aggressing, can you really get that price? How much slippage do you expect? Adverse selection?)
- Given the nature of your alpha signal, and how you expect the price to move immediately after entering an order, what is the best execution strategy to optimize the probability of fill in such a way that you minimize market impact and avoid adverse selection?
- If you can make money getting the mid price, but lose if you have to pay the spread, can you get around this by executing cleverly?
- Given multiple potential counterparty venues with different liquidity profiles, response times, rejection rates, spreads, how do you optimally route your orders?

Data and coding environment

Data and coding environment: students will have access to Thesys limit order book data and high frequency simulator.

References

Michael Kearns and Yuriy Nevmyvaka. Machine learning for market microstructure and high frequency trading. *High Frequency Trading: New Realities for Traders, Markets, and Regulators, Risk Books*, 2013
Sasha Stoikov. The micro-price: A high frequency estimator of future prices. *SSRN-id2970694*, 2018

2 Build a statistical arbitrage strategy

Description

The goal of this project is to design and implement a statistical arbitrage trading strategy. This implies identifying pairs or groups of stocks that are expected to behave similarly and use this to exploit their prices differences.

Initial questions

- Do you want to consider pairs of stocks or create clusters of similar stocks?
- How can you find similar stocks?
- What technique can you use to predict residuals returns: O-U process or other statistical techniques?
- How can you use your signal to create a portfolio that optimize for risk, transaction costs, liquidity etc?

Data and coding environment

Refer to the Guide to Financial Data in the Data section of the website. Once you will have designed a strategy, you can simulate its performance in a finance simulator such as Quantopian or cvxportfolio.

References

Marco Avellaneda and Jeong-Hyun Lee. Statistical arbitrage in the us equities market. *Quantitative Finance*, 10(7):761–782, 2010

Nicolas Huck. Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, 207(3):1702–1716, 2010

George J Miao. High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance*, 6(3):96, 2014

3 Trading strategies based on explained cross-sectional returns

Description

Under the market efficiency assumption, cross-sectional returns variance (meaning that different stocks have different returns for the same time period) is explained by the different exposure to latent source of systematic risks called risk factors. A factor model is then designed to uncover these latent risk factors from historical stocks prices and firm-level characteristics and estimate the exposure of each stock to these factors. Uisng the models out-of-sample return forecasts, stocks can be sorted into deciles. Finally, a zero-net-investment portfolio that buys the highest expected return stocks (decile 10) and sells the lowest (decile 1) can be constructed. The goal of this project is to design and implement a factor model that combines both historical stock prices and firm-level characteristics and use it to construct a portfolio based on sorted out-of-sample stock return predictions.

Data and coding environment

Refer to the Guide to Financial Data in the Data section of the website. Once you will have designed a strategy, you can simulate its performance in a finance simulator such as Quantopian or cvxportfolio.

References

Bryan Kelly, Seth Pruitt, and Yinan Su. Characteristics are covariances: A unified model of risk and return. Technical report, National Bureau of Economic Research, 2018

Shihao Gu, Bryan T Kelly, and Dacheng Xiu. Autoencoder asset pricing models. *Available at SSRN*, 2019

Luyang Chen, Jason Zhu, and Markus Pelger. Deep learning in asset pricing. 2018

4 Crypto-currencies Price Prediction Using News and Social Networks Data

Description

The goal of this project is to design and implement a model that achieves good price predictions of a chosen crypto-currency. Predictions can be made using past prices information in addition to news or social networks data such as tweet sentiments and volume or google trend data.

Data and coding environment

Crypto-currencies historical data can be accessed using the python Bitfinex API client. An example on how to use this tool to gather historical bitcoin data is documented in this article. Historical tweets related to a specific topic such as a crypto-currency can be obtained using the following python package.

References

Sean McNally, Jason Roche, and Simon Caton. Predicting the price of bitcoin using machine learning. In *2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, pages 339–343. IEEE, 2018

Young Bin Kim, Jun Gi Kim, Wook Kim, Jae Ho Im, Tae Hyeong Kim, Shin Jin Kang, and Chang Hun Kim. Predicting fluctuations in cryptocurrency transactions based on user comments and replies. *PloS one*, 11(8):e0161197, 2016

Vytautas Karalevicius, Niels Degrande, and Jochen De Weerdt. Using sentiment analysis to predict interday bitcoin price movements. *The Journal of Risk Finance*, 19(1):56–75, 2018

Muhammad Amjad and Devavrat Shah. Trading bitcoin and online time series prediction. In *NIPS 2016 Time Series Workshop*, pages 1–15, 2017

5 Short term Spot FX Price Prediction using Quotes from different Liquidity Providers

Description

The FX market is highly fragmented. There is no real notion of an order book as in equities. Each participant receives quotes from the counter-parties that they have a relationship with. At any given time, quotes from these different liquidity providers with show quotes at bid and ask that are different from each other. The LPs can also skew the mid price quote (based on their own view, or for risk management purposes.) The goal of this project is to build a hypothetical "quote book" from the data, analyze the behaviors of the the different LPs, and see if there is any information that can be detected exploiting the different views of the LPs that can predict short term price movement. Modeling techniques can include any statistical or Machine Learning technique.

Data and coding environment

FX data from the integral platform are available on the class website in the Data section.

References

Martin D. Gould, Mason A. Porter, and Sam D. Howison. Quasi-centralized limit order books

6 Reinforcement Learning based FX Trading Strategy

Description

The goal of this project is to design and implement a foreign exchange market trading strategy based on a reinforcement learning algorithm.

Data and coding environment

FX data from the integral platform are available on the class website in the Data section.

References

Michael AH Dempster and Vasco Leemans. An automated fx trading system using adaptive reinforcement learning. *Expert Systems with Applications*, 30(3):543–552, 2006

Yue Deng, Feng Bao, Youyong Kong, Zhiquan Ren, and Qionghai Dai. Deep direct reinforcement learning for financial signal representation and trading. *IEEE transactions on neural networks and learning systems*, 28(3):653–664, 2017

Martin D. Gould, Mason A. Porter, and Sam D. Howison. Quasi-centralized limit order books

7 Improving on Classical Anomalies by Combining Them and perhaps adding sentiment data

Description

Well-know market anomalies (low volatility, size, quality) exist. The goal of the project is to find alpha based on these, and to optimally combine the different signals to construct a superior portfolio. You should see if applying smarter prediction techniques (eg smarter volatility prediction for the low volatility if you can use smarter statistical/machine learning or clustering techniques either to improve on the simple anomalies or to better predict when they will work. Additionally you can include Stocktwits sentiment data.

Data and coding environment

Refer to the Guide to Financial Data in the Data section of the website. Once you will have designed a strategy, you can simulate its performance in a finance simulator such as Quantopian. Also, if you want to use stocktwits you'll need to use Quantopian.

References

Jean-Philippe Bouchaud Pierre Blanc, Rmy Chicheportiche. The fine structure of volatility feedback ii: overnight and intra-day effects
A. Beveratos G. Simon L. Laloux M. Potters J.-P. Bouchaud S. Ciliberti, Y. Lemprire. Deconstructing the low-vol anomaly
Guillaume Simon Yves Lemprire Jean-Philippe Bouchaud Stefano Ciliberti, Emmanuel Sri. The size premium in equity markets: Where is the risk?
Stefano Ciliberti Augustin Landier Guillaume Simon Jean-Philippe Bouchaud and David Thesmar. The excess returns of quality stocks: A behavioral anomaly

8 Trend-Following Strategies in Futures Markets

Description

Trend following of futures is a classical strategy that has worked for decades. Rather than prediction, it involves quickly detecting when a trend has started, and managing when to exit a trade. The goal of this project is to replicate and improve on the basic ideas, using more advanced statistical / machine learning techniques in conjunction with exploiting correlation structures within and across asset classes.

Data and coding environment

Refer to the Guide to Financial Data in the Data section of the website.

References

Lasse H. Pedersen Brian Hurst, Yao Hua Ooi. A century of evidence on trend-following investing

P. Seager M. Potters J. P. Bouchaud Y. Lemprire, C. Deremble. Two centuries of trend following

9 Project X

Description

Students may propose their own project idea.