FINAL PROJECT

1. Introduction

The aim of the final project is for students to put together the weekly homework into a single framework. This note outlines the various steps that make up the framework, the baseline expected from the students and suggested directions to add to the basic approach. Note that students should not attempt all bells and whistles simultaneously. Instead, emphasis should be placed on implementing two or three properly, and communicating clearly the effect of the added complexity.

Dates and deadlines:

(a) Presentation on the afternoons of 25 Apr or 26 Apr
(b) Deadline to submit the report by 23 Apr 23:59pm

2. Project Steps

2.1. **Data scope.** First, students should ready their data in a way to quickly traverse their data for various fitting and simulation tasks. This involves splitting data into in and out of sample sets, and establishing which files can be merged or should be kept separate to allow for faster computation.

The baseline involves fitting model parameters using 20 stocks over one month of in-sample data. The model is the next month's data over the same 20 stocks.

To expand, do the study on a rolling basis over the full universe of stocks. This bell and whistle is strongly recommended to achieve more robust statistical metrics for the final simulation.

2.2. **Impact model fitting.** Second, students should fit multiple price impact models on their in-sample data using the public trading tape.

The baseline involves fitting the OW model and one more advanced model, e.g., the AFS or the reduced form.

To expand, fit a non-parametric model and add an appropriate regularization. For instance, one can fit a different model parameter λ_i per stock i and add a regularization term $||\lambda_i - \bar{\lambda}||^2$ where $\bar{\lambda}$ is the parameter for the universal model. Note that such a regularization requires a further sample, leading to three samples: the training sample (in-sample), the testing sample (regularization sample), and the validation sample (out of sample). Care should be taken in explaining how the regularization meta-parameter was tuned.

2.3. **Backtest engine.** Third, students should implement a backtest engine that accepts generic trades rather than just optimal ones.

The baseline implements a version of Waelbroeck's simulator based on the fitted price impact model that resets positions and impact states at the start of each day.

To expand, implement a backtest engine that tracks impact states and positions over multiple days. Please take into account the possibility of stock splits. You can either hard-code a list of affected stocks and dates, or detect them when prices and volumes jump in opposite directions. For instance, prices may double, daily volumes may half, leaving the overall daily trade notional unchanged. Exclude stocks with splits.

2.4. **Synthetic alphas.** Fourth, students should implement synthetic look-ahead alphas for the alpha level and decay.

The baseline implements the same synthetic alpha as in the homework.

To expand, implement an additional synthetic alpha for the overnight return.

2.5. **Optimal trading strategy.** Fifth, students should implement trading strategies based on the alpha signal.

The baseline implements the optimal strategy for the OW model and the fitted model.

To expand, implement the optimal trading strategy with a synthetic overnight alpha. Simulate trading strategies with and without the intra-day alpha.

2.6. **Performance metrics.** Sixth, students should build a performance report for each strategy.

The baseline outlines the alpha's correlation with returns, as well as key portfolio metrics of the strategy: expected daily P&L, Sharpe ratio, transaction costs, max daily drawdown, max impact dislocation.

To expand, make the performance report visually as appealing and informative as possible.

2.7. **Sensitivity analysis and stress testing.** Seventh, students should leverage their framework to answer questions about the strategies.

The baseline outlines the various degrees of freedom, e.g., choice of impact model, strength of alpha signal and highlights how these affect the performance of the strategy.

To expand, simulate the answer to questions such as: "What if I use the wrong impact model?" "What if my signal has a latency of one minute?" "What if I am forced to liquidate my position mid-way through the day in a single block trade?"

3. Format of Assessment

The report is due on the weekend before the scheduled presentation (Sunday 23:59PM). It is expected to polished and presentable, with all the sections mentioned, but does not need to be very technical. For each section, if the procedure is standard, only a simple example is needed; if the procedure is beyond the basics, more outputs (images / sample path / summary statistics / any notion of size for the object) are needed. No code needs to be attached to the report, but the use of signature (what is the input/output) is encouraged to explain the important functions. At the start of each section, it is expected that the students indicate the ones who contributed to this section, and also use a few sentences to explain what is done (basic or beyond the basics). The students are not expected to do a lot of extensions: Two extensions (beyond the basics) done carefully is better than five extensions done not very carefully.

The oral presentation is expected to be more technical. The main content expected is what is done beyond the basics. During the presentation, the students will be asked questions (individually) about the part(s) of coursework they have contributed to. The ones contributed to the code should be ready to talk about or describe the code during the presentation.