Algorithmic Trading and Quantitative Strategies – Fourth Homework

Due date: May 8, 2018

General guidelines in regards to this homework

- Please post any questions about homework on the NYU Classes course discussion board so that all students can benefit.
- You may collaborate in groups of up to 3 people when solving the homework.
- *Important: Each student is expected to write up their own solutions in their own words. If you collaborate with others, you must list the names of the other participants on your homework submission to get full credit for your homework.*
- Submitting your homework: Submit all files to NYU Classes. No submission via email, please! Homework needs to be submitted on time for full credit. Codes need to be delivered with fully functional unit tests.

1. Cointegration

In the file cointData.csv.gz, you are provided with a matrix of stock returns where each column is a series of returns, one for each time period. Each row represents a point in time. How many cointegrated pairs can you find?

(i) Find the cointegrated pairs by performing the Granger-Engle cointegration test as discussed in class.

(ii) Efficient computation and updating for pairs trading: In Python, perform an efficient regression for the γ parameter by identifying which terms of the regression equation can be queued and used to obviate the need to re-compute the regression from scratch. The end product of your calculations should be one matrix for each cointegrated pair of returns. This matrix should include the two original price columns, as well as one column for each of the queues that allow you to perform your regression efficiently at each point in time. The final column of your matrix should be the value of γ – one value for each time period represented by a row. Please submit your final matrices as well as a separate file containing your code.

Two “no's” for 1.ii:

- No loops are allowed! If you perform this regression efficiently, you will not need to loop over your data.
- You cannot use Pythong's built in linear regression and cointegration functions. You have to build the functions you need using simple running sums. Python's built-in regression and cointegration functions will not perform the calculations needed for 1.ii efficiently.

2. Kernel Impact Model

As discussed in class, there is a mismatch between our market impact model and the type of trading we’re trying to simulate. Our trading is continuous, with our own past executions impacting our future executions. On the other hand, our rate of trading impact model is discrete. It assumes that we will wait until market impact is entirely gone before we trade again.

In fact, we can use our discrete rate of trading model to run a simulation, but only if we (i) trade at a more or less constant rate of trading throughout the day, (ii) don’t trade during the last half hour of every day, and (iii) assume that the impact of our own past executions will not significantly change our mark-to-market during the half hour non-trading period. In other words, we can use our existing model if we fully comply with its assumptions.

Note that the random walk of prices still gives us a mark-to-market during the non-trading period, but that mark-to-market overstates the impact of our own past executions until the actual close. At the close, we can assume that the closing price is the mark-to-market price of our positions and contains no residual temporary impact.

Switching to a kernel impact model

Instead of using our existing model, we will modify our rate of trading model to be a kernel impact model. The kernel impact model keeps track of the impact of our own past executions and adds the un-decayed temporary impact to the prices that our trading strategy sees in our simulation. The impact of our past executions decays over time.

To make our impact model into a kernel impact model, we will regress for the parameters of decay of temporary market impact at several points after the end of the period in which we observe an imbalance in trading.

Background reading

This is Jim Gatheral's paper on market impact:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1292353

In the paper, he considers two types of temporary impact decay, exponential and power law. Jim shows that these simple models of decay admit an arbitrage and are, therefore, inaccurate descriptions of market impact. Nevertheless, the arbitrage he describes is small and does not significantly alter the results of trading simulations that are not actively trying to perform the arbitrage. In other words, they're good enough for our purpose.

Here are two other documents on the same subject that are worth perusing.

Presentation: Three Models of Market Impact, Jim Gatheral
https://mfe.baruch.cuny.edu/wp-content/uploads/2017/05/Chicago2016OptimalExecution.pdf

Slow Decay of Temporary Market Impact
https://www.worldscientific.com/doi/pdfplus/10.1142/S2382626615500070

Procedure

We will assume that all of temporary impact is gone by the close of the following day of trading. This is a little different than the assumption we made in our rate of trading model, the assumption that all temporary impact is gone a half hour after we stop measuring imbalance. However, this assumption is safer and deprives us of only one data point per stock, not a big loss.

Suppose we are modeling the exponential decay of temporary impact. At 3:30PM, on day 1, we have temporary plus permanent impact, and at 4:00PM on day 2, we have only permanent impact. As in our current model, we can back out the temporary impact, h and the permanent impact, g. After that, we want to regress for the parameter of decay, γ.

Initially, we want to fit the following model

$$p_{(330+t)} = p_{330} + \gamma^{t/\tau} h + \varepsilon$$

This gives us a decay parameter, γ per unit time, τ for the decay of temporary impact, h.

What if this parameter is not constant over time? We can check that by simultaneously fitting the parameter at several points in time.

Assignment:

(i) Fit two models of temporary impact decay (the exponential and power law) using the same dataset as you used in Homework 2.

(ii) Propose some tests to be used to test for significance and robustness of your models. Perform these tests and describe your results.

(iii) Compare the two models in (i) with oneanother. What similarities/differences do you see?

3. LQ Regulator.

In class we showed that for the deterministic LQ regulator the value function and the optimal control are given by

$$\begin{aligned} F(x,\tau) &= \inf_u \{c(x,u) + F(Ax+Bu,\tau+1)\} \\ &= \inf_u \left\{ \begin{pmatrix} x \\ u \end{pmatrix}' \begin{pmatrix} R + A'K_{\tau+1}A & S' + A'K_{\tau+1}B \\ S + B'K_{\tau+1}A & Q + B'K_{\tau+1}B \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \right\} \end{aligned}$$

and

$$u_t = L_t x$$

where

$$L_\tau = -(Q + B'K_{\tau+1}B)^{-1}(S + B'K_{\tau+1}A)$$

Carefully derive the Riccati equation for K_t and verify the optimal control.

4. Extra Credit: Stochastic LQ Regulator

(i) Formulate the discrete time version of the Almgren and Chriss (2000) optimal execution model as a stochastic LQ regulator problem. You can assume alpha is zero.

(ii) Determine the Riccati equation and the optimal control for the problem. How does this solution compare to the solution in their paper?

(iii) Write a Python program that solves the optimal execution problem using the discrete stochastic LQ formulation from above and determine the optimal execution schedule for the following trade:

- Initial shares price: $100
- Initial holdings: 1,000,000 shares
- Liquidation time: 1 day
- Number of time periods for the liquidation: 7

Assume that the stock has an annual volatility of 22% and a drift of 7%. The daily volume in the stock is 25 million shares. You can assume that both temporary and permanent impacts are linear. Use the market impact parameters you estimated from homework 2 (or Almgren's paper) and assume the execution risk aversion parameter is 10^{-6}.

(iv) In terms of the temporary impact costs, does it matter how many time periods you use to liquidate the trade? Compare different number of time periods for the liquidation to answer this question.