GR5010: Intro to the Mathematics of Finance

**Mean Reversion Pairs Trading**

Anand Kannan, Michael Copeland, Michael Demian, Yiran Xia

horizontal line

# Abstract

Pairs trading is a market neutral strategy based on trading cointegrated stocks. Given that this co-movement in price is well established, it is possible to create a market or even sector neutral trading strategy offering diversification from traditional equity investments. By determining the cointegration levels of these pairs of equities and analyzing to see if these pairs will stay in sync, we can trade off temporary aberrations in the difference of their price ratios as they revert to their long term “true” levels. We investigated first how to find which stocks are most cointegrated and then examined relative price movements to a simple moving average of their price differences. In this implementation the Hurst exponent was used to measure the long-term memory and autocorrelation of the spread series. Using the method and stocks chosen, it yielded a return of approximately 10% with a Sharpe Ratio of 2.38. Although a successful model, it can be improved by testing in various environments and with different parameters.

## Introduction

As widely recognized to be the “ancestor” of statistical arbitrage, a strategy that aims to profit from pricing inefficiencies in a group of securities, pairs trading is one of the fundamental trading strategies that can work as a stepping stone for people to better understand the world of systematic trading. Pairs trading allows investors to speculate on the relative relations between stocks, which is much more discernible and empirically driven. This type of strategy is more advantageous compared to traditional investment decision process that involves analyzing the fundamentals of the companies, only attempting to assess the ever-mysterious fair value of these stocks. Originated in the mid 80s, pairs trading has evolved constantly, in terms of generating trading signals, with technology advancement and introduction of more statistic models. However, the keystone for the strategy should remain the same, which is using cointegration tests to single out historically cointegrated stock pairs with similar characteristics and take opposite positions when there are statistically significant discrepancies between the price spread of the two stocks. This can yield equity gains that are uncorrelated to the larger equity market, allowing for diversification with a strong alpha. This diversification has been shown countless times to improve gains, while minimizing risk to portfolio.

**Methodology/Model**

The pairs trading strategy is a well explored model. Given two stocks and , we expect them to adhere to the following theoretical mean reverting process:

The traditional pairs trading model takes the following form: given two stocks, and , which are both I(1) series, meaning that both series are non-stationary, then if the two stocks are cointegrated, there exists some relationship such that

Where is the relationship between the two stocks and is a stationary series.

However, rather than following this model explicitly by running regressions to determine the linear relationship between two stocks, the implementation presented here follows more simply just the intuition and meaning of cointegration. What this means is that, the spread between cointegrated stock pairs should have a long-term mean. If the spread of the two stocks deviate from this long-term mean, it is speculated that they will revert back at some point in the near future. Given an appropriate look back period, the simple moving average of the spread of the two stocks should suffice to serve as a representation of such long-term mean.

The primary assumption driving this strategy is that the two stocks are driven by the same fundamental risk factors. This in turn implies returns for each stock have the same common driving factors such that prices are cointegrated. As a result, we expect the stocks to remain cointegrated by following a mean reverting process as previously stated. However, this assumption does not always hold and is the primary flaw with pairs trading. Instead, cointegrated pairs may break and as a result major losses can be incurred if the long/short position is held for too long.

To amend our model to help overcome the issue of pairs breaking, we implement a Hurst exponent. The Hurst exponent is a test that quantifies whether a time series regresses to a mean or develops a trend away from the mean. The Hurst exponent *H* can be expressed as:

Where is the range of the first n cumulative deviations from the mean, is the standard deviation, is a constant, is the number of observations and is the Hurst exponent. The Hurst exponent is constructed such that, if , there is a tendency for the time series to regress towards the mean. If , the time series is uncorrelated. Finally, if , the time series will break from the mean. The construction of the Hurst exponent provides a simple numeric test to determine whether a given pair is breaking from cointegration and therefore can help filter out trades that may prove harmful to a pair trading strategy.

## Data

The trading strategy is built on Quantopian and utilizes Quantopian data sets. Considering the vast universe of investable public companies, there might be a large number of stock pairs that pass cointegration tests without fundamental reasons behind to support such cointegrations. As a result, instead of directly performing cointegration tests on the S&P 500, we have subdivided our universe into sectors based on Morningstar’s Sensitive Equities Category using the integrated APIs available in Quantopian. These stocks represent companies that are sensitive to the economy to pick up on changes, but not overly so causing too great of swings. This universe was chosen to ensure that the selected stocks had some alpha guaranteeing that deviations from the mean occurred. These deviations create opportunities for arbitrage, while at the same time not with too much alpha such that the cointegrations would break. Sectors fulfilling these requirements included the following: communication services, energy, industrials and technology. By going after stocks pairs in different sectors, the portfolio should be even further diversified. In addition, in order to ensure liquidity in the portfolio, a threshold of $10 billion market cap is applied to exclude possibly illiquid stocks, and to avoid possible interference in pairing a very small market cap with a large one.

A sample period of January 1, 2017 to January 1, 2018 was chosen for the formation of pairs. The pairs are as follows:

* Communications
  + AT&T Inc. (T) and Albermale Corporation (ALB)
  + Liberty Global plc (LBTY\_A) and SBA Communications Corporation (SBAC)
* Energy
  + Exxon Mobil Corporation (XOM) and Cheniere Energy (LNG)
  + Hess Corporation (HES) and Anadarko Petroleum Corporation (APC)
* Industrials
  + Raytheon Company (RTN) and Lockheed Martin Corporation (LMT)
  + Northrop Grumman Corporation (NOC) and Rockwell Automation, Inc. (ROK)
  + Caterpillar Inc. (CAT) and Dover Corporation (DOV)
* Technology
  + Autodesk, Inc. (ADSK) and Workday, Inc. (WDAY)
  + Adobe Inc. (ADBE) and Salesforce.com, Inc. (CRM)
  + Corning Incorporated (GLW) and Alphabet Inc. (GOOGL)

## Implementation

With the given pairs, the portfolio was constructed using equal weightings for every pair (i.e. 10% capital allocation per pair) from initial capital of $100,000. Two calculations were done using rolling windows (i.e. the look back period) to generate signals to enter certain positions. First, the spread between the stock prices of a given pair were calculated over a rolling window of variable lengths (see Figure 1 for chosen window lengths). The Z-score is calculated daily based on the rolling window, and if the outputted value exceeded |2| (i.e. ), the deviation is significant enough for an arbitrage opportunity to be present and a long-short position is entered in line with the logic of pairs trading. Secondly, the hurst exponent is calculated over a variable rolling window (see Figure 2 for chosen lengths). If the calculated exponent exceeds or equals 0.5 (i.e. ), the trade is not entered.

Rebalancing and trading for the strategy occurs daily, meaning the previously mentioned calculations are performed daily. To determine when to exit positions, we track the Z-score of the pair. If the value of the Z-score drops below |0.5| (i.e. ), then the pair is determined to have returned to its running mean and the position is exited. It is important to note that the Hurst exponent is only used to filter potential trades but does not serve as a trigger to exit positions already taken. This is because doing so could trigger exiting profitable positions too early.

The test period was chosen as January 1, 2018 to December 7, 2018. Rebalancing and trades are executed daily during this time period as previously mentioned. Additional trading considerations were also implemented, including: volume limit of 2.5% and price impact constant of 0.1 to avoid possible slippage and nonideal trading price, as well as transactions costs of 75 bps to mimic real-life trading.

## Empirical Results

We tested multiple sets of parameters to see how certain adjustments affected the overall performance of the strategy.

Figure 1: Metrics Sans Hurst Exponent

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Lookback Days** | **15** | **20** | **25** | **30** | **40** | **50** |
| **Cumulative Returns** | 7.6% | 11.8% | 9.3% | 8.1% | 10.4% | 9.8% |
| **Sharpe Ratio** | 1.46 | 2.25 | 1.72 | 1.46 | 1.76 | 1.77 |
| **Max Drawdown** | -3.8% | -2.8% | -3.8% | -4.4% | -5.0% | -4.2% |
| **Daily VaR (95%)** | -0.7% | -0.6% | -0.7% | -0.7% | -0.7% | -0.7% |
| **Beta** | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.06 |

Figure 1 reflects the results of the pairs trading strategy without the use of the Hurst exponent. This was done to see how introducing the Hurst exponent would change the strategy’s performance, with said results presented later. In Figure 1, the base pairs trading strategy yields strong results, with the best performing parameterization with a lookback window (used to calculate the moving average) of 20 days yielding 11.8% cumulative returns. However, across all parameterizations it is interesting to note the relatively high Sharpe ratio. The low beta is also important to note as it signals that the strategy is performing as expected as pairs trading strategies are intended to be market neutral.

Figure 2: Metrics with Hurst Exponent

|  |  |  |  |
| --- | --- | --- | --- |
| **Lookback Days** | **20** | **25** | **30** |
| **Hurst Exponent** | 0.5 | 0.5 | 0.5 |
| **Hurst Lookback** | 15 | 20 | 25 |
| **Cumulative Return** | 10% | 8.8% | 2.1% |
| **Sharpe Ratio** | 2.38 | 1.66 | 0.48 |
| **Max Drawdown** | -1.3% | -4.2% | -3.0% |
| **Daily VaR (95%)** | -0.5% | -0.7% | -0.6% |
| **Beta** | 0.05 | 0.02 | 0.01 |

Figure 2 shows the results of a subset of parameterizations with the Hurst exponent implemented. The net effect of the Hurst exponent appears to lower cumulative returns but also improved risk metrics, such as increasing Sharpe ratio and lowering maximum drawdown. This is indicating that the Hurst exponent is filtering out certain trades as per the implementation. However, the lower cumulative returns may indicate that said filtered trades were not necessarily harmful to the portfolio. This may be due to a multitude of factors such as the test sample or the specific implementation of the Hurst exponent.

The following figures present different aspects of the best performing strategy (i.e. without Hurst exponent and lookback of 20 days):

Figure 3: Best Performing Strategy Cumulative Returns (relative to SPY Benchmark)



Figure 4: Best Performing Strategy Daily Returns

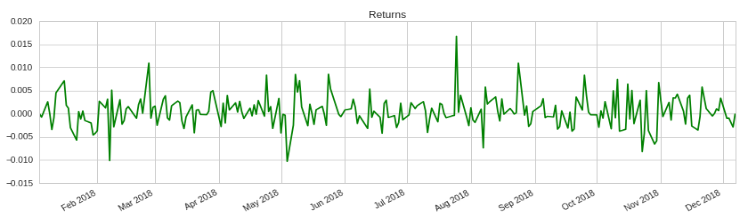


Figure 5: Maximum Drawdown of Best Performing Strategy

## 

## Extensions

Despite yielding interesting results, there remains many ways to optimize the proposed strategy. Dynamic security selection and portfolio weighting is one such area where improvements can be made. The fixed and equal weighting as well as constant portfolio composition may lead this strategy to fail over longer periods of time. Altering how the pairs are found using principal component analysis (PCA) may also improve the strategy by better identifying pairs that will remain cointegrated over longer periods of time. Finally, performing a more sophisticated and robust calculation of the Hurst exponent may provide a better signal to prevent trading of certain stock pairs that may lose cointegration. One such improvement is using detrended fluctuation analysis as it provides a similar signal as the Hurst exponent, but can be applied to non-stationary time series.

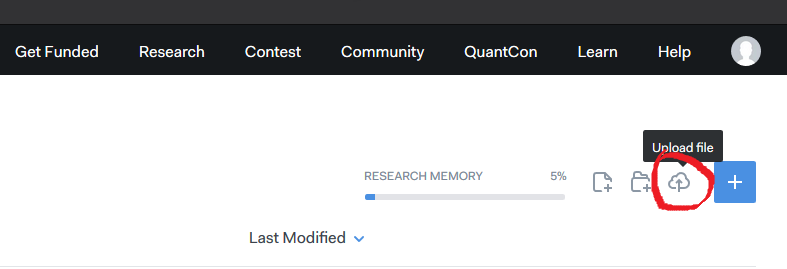
## Conclusion

The implementation presented showcases the basic intuition of statistical arbitrage, to profit from short-term deviations from long-term trends. As is shown in the implementation, although it does not fully represent real-life trading, a fairly simple strategy with a simple trading signal has surpassed typical market benchmarks. Yet this strategy reflects only one permutation of many possible strategies and parameterizations, showing the depth and scope of the field of quantitative finance.

## References

* Auquan, T. (2017, December 19). Pairs Trading using Data-Driven Techniques: Simple Trading Strategies Part 3. Retrieved December 21, 2018, from <https://medium.com/auquan/pairs-trading-data-science-7dbedafcfe5a>
* Avellaneda, M., & Lee, J. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance,10*(7), 761-782. doi:10.1080/14697680903124632
* Chan, E. (2016, April 07). Quantitative Trading. Retrieved December 21, 2018, from http://epchan.blogspot.com/2016/04/mean-reversion-momentum-and-volatility.html?m=1
* Polak, P. (n.d.). *Statistical Inference and Time-Series Modelling*. Lecture presented in Columbia University, New York.

## Directions For Accessing and Running Code

1. Create an account on Quantopian: <https://www.quantopian.com>
2. For cointegration notebook:
   1. Click on Research -> Notebooks tab in website navigation bar
   2. Take attached .ipynb file ‘Intro To MAFN Pairs Mean Reversion.ipynb’ file and upload it via the Upload file button as pictured
   3. Run notebook as normal Jupyter notebook
   4. \*Note that the industry code and market cap was varied to pick the top choices above a particular market cap from each industry
3. For code and backtest:
   1. See the following link and click on clone algorithm to get to the IDE to run. <https://www.quantopian.com/posts/intro-to-mafn-best-hurst-ak-mc-md-yx>