Mean-Reversion – A Pairs Trading Model with FOX and FOXA

Peter Kokalov

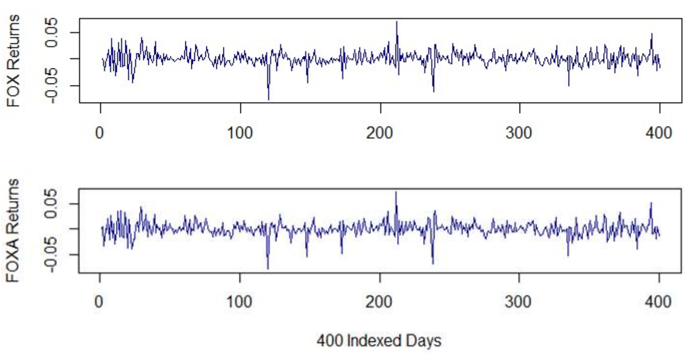
1. Pairs Trading – An Introduction:

Pairs trading is a market neutral, stat-arb trading strategy that tracks the spread of two supposedly related securities. If the relationship between these securities weakens, a trader can make a bet that the spread will eventually converge - the basic assumption being that prices tend to move towards some historical average. In this report, I will discuss some methods for modelling share prices and how the pairs trade can be a profitable trading algorithm.

Take a generalized case of two securities X_t and Y_t and suppose that X_t is showing above-average returns for some defined period of k previous days. Let R_t^x and R_t^y denote the return for our given securities. The spread is calculated as $R_t^x - \beta R_t^y$ where β is a carefully chosen constant and may change over time. If we have some fundamental reason to believe the spread possesses mean-reverting behavior, it can be expected that $R_t^x - \beta R_t^y$ will return to some historical equilibrium. This is sometimes referred to as "exhibiting memory". Here we will assume that the returns of two stocks satisfy the SDE:

$$\frac{dR_t^1}{R_t^1} = \alpha dt + \beta \frac{dR_t^2}{R_t^2} + dXt$$

where X_t is the residual at time t and the intercept term α is negligible compared to the fluctuations of either process and can typically be ignored.



2. Geometric Brownian Motion and OU-Process:

A geometric Brownian Motion (GBM) is a time-continuous stochastic process that is most often used to model share price. A stochastic process Y_t satisfies the SDE

$$dY_t = \mu Y_t dt + \sigma Y_t dW_t,$$

which can then be written as

$$\frac{dY_t}{Y_t} = \mu dt + \sigma dW_t$$

where μ is the percentage drift, σ is the volatility, and W_t is a Weiner Process given by mean = 0, and variance = t. Using Itō's Lemma, and defining $X_t = e^{Y_t}$, we write use an expanded Taylor Series and drop high orders,

$$d(\ln X_t) = \frac{dY_t}{Yt} - \frac{1}{2}\frac{1}{Y_t^2}(dY_t)^2.$$

Substituting for $(dY_t)^2$ and $\frac{dY_t}{Yt}$, we get

$$d(\ln X_t) = \mu dt - \frac{\sigma^2}{2}dt + \sigma dW_t$$

and integrating directly with respect to time,

$$\ln \frac{X_t}{X_0} = \int_0^t (\mu - \frac{\sigma^2}{2})dt + \int_0^t \sigma dW_t$$

$$= (\mu - \frac{\sigma^2}{2}t) + \sigma W_t$$

$$\frac{X_t}{X_0} = e^{(\mu - \frac{\sigma^2}{2}) + \sigma W_t}.$$

Therefore, the price of some share X_t at a given time t is can be written as

$$X_t = X_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$

Geometric Brownian Motion is of particular interest when modelling mean-reverting processes because of the relationship between the drift term and the equation X_t itself. In this paper, when trying to model the spread, which we assume to be mean-reverting (testing discussed in Data Preprocessing), we will use the Ornstein Uhlenbeck or OU-process. Such a process describes a large Brownian particle moving under the influence of friction. This behavior can be defined as exhibiting stationarity, which is the underlying trait of any mean-reverting process. The OU-process is defined by the SDE

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t$$

where X_t is the residual term at time t, W_t is a Weiner process, and σ is the standard deviation. Applying Itō's Lemma again, and defining the function $X_t = X_t e^{\theta t}$, we get

$$d(X_t e^{\theta t}) = (\theta e^{\theta t} + e^{\theta t}\theta(\mu - X_t))dt + e^{\theta t}\sigma dW_t)$$

and directly integrating with respect to time, we have

$$X_t e^{\theta t} - X_0 = \theta\mu \int_0^t e^{\theta s}\, ds + \sigma \int_0^t e^{\theta s} dW_s,$$

$$X_t = X_0 e^{-\theta t} + \theta(1 - e^{-\theta t}) + \sigma \int_0^t e^{2\theta(s-t)} dW_s.$$

3. Linear Regression for Parameter Estimation:

Now, note that the discretized OU-process can be understood as an autoregressive, AR(1) process:

$$X_t = c + \varphi X_{t-1} + \varepsilon_t,$$

where ε_t is assumed to be $iid \sim N(0, \sigma^2)$. To put the whole mess of equations together, we'd have the OU-process simplify to the regression model

$$X_{t+1} = a + \beta X_t + \varepsilon_{t+1},$$

where we can obtain our parameters a, β, and $\text{var}(\varepsilon_t)$ from the parameters given by our solution to the stochastic differential equation for the OU-process!

$$a = \theta(1 - e^{-\theta \Delta t})$$

$$\beta = e^{-\theta \Delta t}$$

$$\text{var}(\varepsilon_t) = \frac{\sigma^2}{2\theta}(1 - e^{-2\Delta\theta})$$

Our residuals X_t are obtained from running regression of the returns from one stock on the returns of another followed by another regression to estimate the parameters.

$$\theta = -log(\beta) \cdot 252$$

$$\mu = \frac{\text{a}}{1 - \text{b}}$$

$$\sigma = \sqrt{\frac{\text{var}(\varepsilon) \cdot 2\theta}{1 - \beta^2}}$$

And the equilibrium value of the standard deviation given as

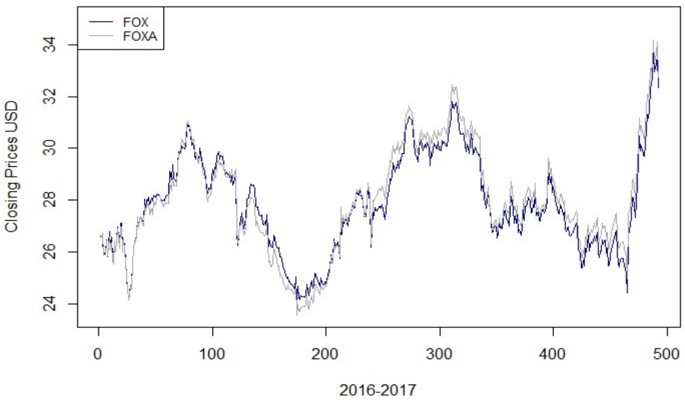
Mean-Reversion – A Pairs Trading Model with FOX and FOXA

Peter Kokalov

$$\sigma_{\text{eq}} = \sqrt{\frac{var(\varepsilon)}{1-\beta^2}}$$

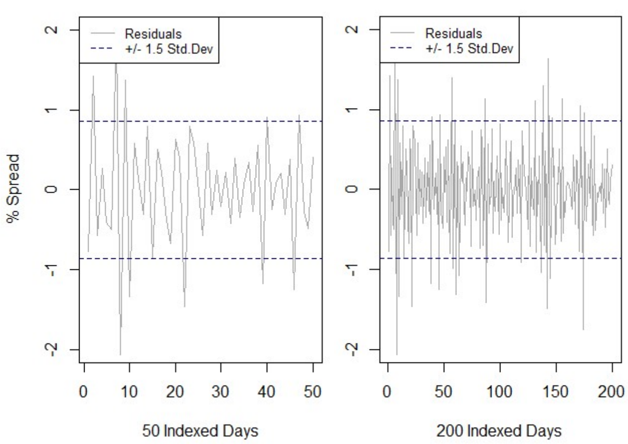
4. Data Preprocessing:

All data was obtained using AlphaVantage's API (alphavantage.co). There were 2016 data periods where the time-frame was intraday. Our training data consisted of trading days from Dec 10, 2009 to Dec 31, 2015, and our backtest period was from Jan 4, 2016 to Dec 13, 2017 - a little under 2 years. Daily closing prices were used as our price input.



5. Stationarity Testing:

When selecting the pairs, I looked for two stocks that had similar fundamental properties and exhibited strong stationarity. Engel and Granger first stated the concept of cointegration. A pair of assets X_t and Y_t are said to be cointegrated *I(1)* if there exists α and b such that the linear combination $Z_t = Y_t - \alpha X_t - b$ is *I(0)*. In other words, the spread is stationary, or mean-reverting. In this project, the Augmented-Dickey Fuller (ADF) Test was used to test for stationarity in the spread. The ADF test has a null-hypothesis that the spread Z_t has a unit-root and the alternative that the process is stationary. An example of order-1 integration $Z_t = Z_{t-1} + \varepsilon_t$ which is the case of a random walk, would be integrated *I(1)* as $\Delta Z = Z_t - Z_{t-1} = \varepsilon_t$. Through my analysis, I decided NASDAQ:FOX and NASDAQ:FOXA would be good candidates.



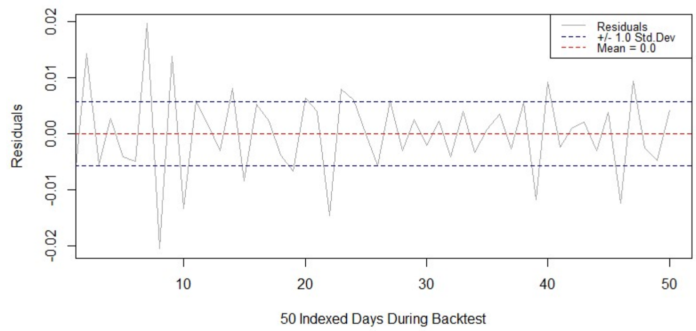
Working with the spread we obtained from our regression analysis, we can formulate a Z-Score with our σ_{eq}, X_t, and θ. We denote our Z-score as

$$z = \frac{X_t - \theta}{\sigma_{eq}}$$

which will measure the number of standard deviations our spread has deviated from it's "mean". Our trades will correspond to the buying and selling (or vise-versa) of \$1 of FOX and \$$\beta$ of FOXA. The thresholds I used for the decision-making was

open short position = open long position for $z = 1.0$

close short position = close long position for $z = 0.0$.



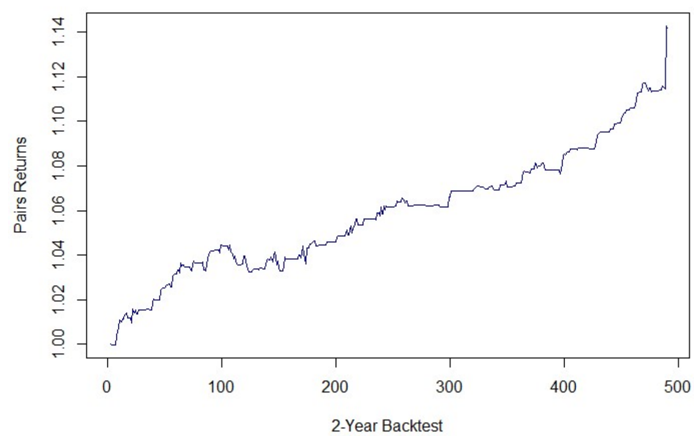
Throughout the backtest period, our β, or "hedge ratio" was updated on a rolling 25-day basis. From the analysis, it seems that the strength of the reversion grew almost monotonically over time!

6. Results:

The goal of the algorithm was profitability. Although I did not account for the transaction cost of each trade in this project, the FOX – FOXA pair was profitable!

Portfolio Return	Sharpe Ratio	Max Drawdown
14.28%	2.55%	1.20%

In the table shown above, all three metrics span from Jan 4, 2016 to Dec 13, 2017. The cumulative returns of the portfolio are shown below.



7. References:

[1] En.wikipedia.org. (2017). *Pairs trade*. [online] Available at: https://en.wikipedia.org/wiki/Pairs_trade [Accessed 20 Dec. 2017].

[2] www1.maths.ox.ac.uk. (2017). http://www1.maths.ox.ac.uk/system/files/attachments/Thierfelder_v1.08.pdf [Accessed 20 Dec. 2017].

[3] Huang, C., Hsu, C., Chen, C., Chang, B. and Li, C. (2017). *An Intelligent Model for Pairs Trading Using Genetic Algorithms*.

Mean-Reversion – A Pairs Trading Model with FOX and FOXA

Peter Kokalov

[4] Engle, Robert and Granger, C.W.J. *Co-Integration and Error Correction: Representation, Estimation, and Testing*. Econometrica, Vol. 55. 1987.