Pairs Trading Model in R

A Statistical Arbitrage Strategy

Group members:
FOUKI Houssam, AMRAOUI Hassan, CONSOLI Lorenzo,
BADARU Oluwafisayomi, MAISURADZE Alexander

Professor Pierre JACOB
Forecasting & Predictive Analytics
ESSEC Business School - CentraleSupélec

Outline

1. **Basic Idea of Pairs Trading**
2. **Testing for Cointegration**
    - **The idea of Pairs Trading based on Cointegration**
    - **Pairs Selection**
    - **Least Squares (LS) regression**
    - **Augmented Dickey-Fuller (ADF) test**
    - **Phillips-Perron (PP) test**
3. **Trading Strategy Design**
    - **Trading the Spread**
    - **Pairs Trading Performance (Simulation)**
4. **Pros and Cons of Pairs Trading**
5. **Other Pairs Trading Implementations**

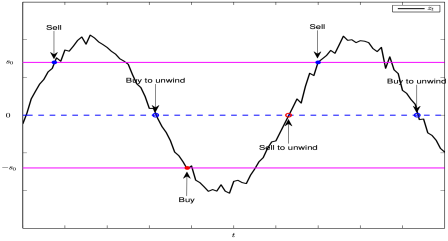
Basic Idea of Pairs Trading

Trading the spread

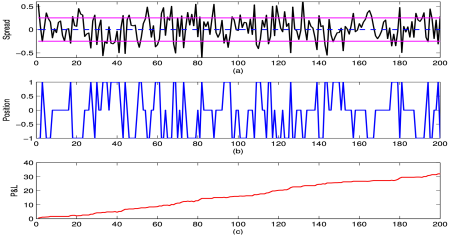
The idea behind pairs trading is to **short-sell the relatively overvalued stocks** and **buy the relatively undervalued stocks**, and also unwind the position when they are relatively fairly valued.

Illustration on how to trade the spread

$$z_t = x_t - \gamma y_t$$



Statistical arbitrage can be used in practice with profits



Outline

1. Basic Idea of Pairs Trading
2. **Testing for Cointegration**
   - **The idea of Pairs Trading based on Cointegration**
   - Pairs Selection
   - Least Squares (LS) regression
   - Augmented Dickey-Fuller (ADF) test
   - Phillips-Perron (PP) test
3. Trading Strategy Design
   - Trading the Spread
   - Pairs Trading Performance (Simulation)
4. Pros and Cons of Pairs Trading
5. Other Pairs Trading Implementations

The Idea of Pairs Trading based on Cointegration

Cointegration

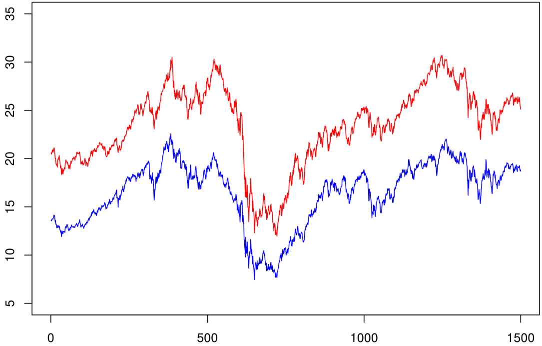
Consider a pair of time series, both of which are **non-stationary**. If we take a particular linear combination of these series, it can sometimes lead to a stationary series. Such a pair of series would then be termed **Cointegrated**.

Idea

While it may be difficult to predict individual stocks, it may be easier to predict **relative behavior** of stocks

Let x_t and y_t be two non-stationary $I(1)$ time series, with $\beta_0, \beta_1 \in \mathbb{R}$ constants. If the combined $y_t = \beta_0 + \beta_1 x_t + \epsilon_t$ series is stationary then we say that x_t and y_t are cointegrated.

Note: x and y here are clearly not stationary, but they seem to move together. In fact, they are cointegrated: $y_t - \beta_1 x_t - \beta_0$ should be stationary.



Illustrative Example

A drunk man is wandering the streets (random walk) with a dog. Both paths of man and dog are nonstationary and difficult to predict, but **the distance between them is mean-reverting and stationary.**

Outline

1. Basic Idea of Pairs Trading
2. **Testing for Cointegration**
   - The idea of Pairs Trading based on Cointegration
   - **Pairs Selection**
   - Least Squares (LS) regression
   - Augmented Dickey-Fuller (ADF) test
   - Phillips-Perron (PP) test
3. Trading Strategy Design
   - Trading the Spread
   - Pairs Trading Performance (Simulation)
4. Pros and Cons of Pairs Trading
5. Other Pairs Trading Implementations

Pairs selection: Normalized Price Distance

Design of Pairs Trading Strategy

In practice, pairs trading contains three main steps:

- **Pairs selection**: identify stock pairs that could potentially be cointegrated.
- **Cointegration test**: test whether the identified stock pairs are indeed cointegrated or not.
- **Trading strategy design**: study the spread dynamics and design proper trading rules.

- Normalized price distance (as a rough proxy to measure cointegration):

$$NPD = \sum_{t=1}^{T} (\tilde{p}_{1t} - \tilde{p}_{2t})^2$$

  where the normalized price $\tilde{p}_{1t}$ of stock 1 is given by $\tilde{p}_{1t} = p_{1t}/p_{10}$. The normalized prices of stock 2 defined similarly: $\tilde{p}_{2t} = p_{2t}/p_{20}$
- Now we can basically compute the NPD for all the possible combination of pairs and select some pairs with smallest NPD as the potentially cointegrated pairs.
- One can use a more refined measure of cointegration (more computationally demanding).

Pairs selection: Normalized Price Distance

Results were generated and we picked the pair with the lowest NPD: **('JPM', 'GS')**. This comes as no surprise since both stocks are finance related and these companies can be said to be competitors. $(Log-)$prices of both stocks are shown below:

Outline

1. Basic Idea of Pairs Trading
2. **Testing for Cointegration**
   - The idea of Pairs Trading based on Cointegration
   - Pairs Selection
   - **Least Squares (LS) regression**
   - Augmented Dickey-Fuller (ADF) test
   - Phillips-Perron (PP) test
3. Trading Strategy Design
   - Trading the Spread
   - Pairs Trading Performance (Simulation)
4. Pros and Cons of Pairs Trading
5. Other Pairs Trading Implementations

Least Squares (LS) regression

- If the spread z_t is stationary, it can be written as:

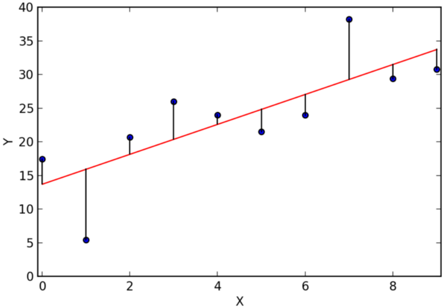
$$z_t = x_t - \gamma y_t = \mu + \epsilon_t$$

  where μ represents the equilibrium value and ϵ_t is a zero-mean residual.

- Equivalently, it can be written as:

$$x_t = \mu + \gamma y_t + \epsilon_t$$

  which now has the typical form of linear regression.



- Least Squares (LS) regression over T observations:

$$\min_{\mu,\gamma} \sum_{t=1}^{T} (x_t - (\mu + \gamma y_t))^2$$

- LS regression is used to estimate the parameters μ and γ, obtaining the estimates $\hat{\mu}$ and $\hat{\gamma}$.

Least Squares (LS) regression

LS Regression Results					
Coefficients:					
	Estimate	Std. Err	t-value	Pr. (>| t |)	
(Intercept)	2.4507	0.015	160.779	0.0000	***
JPM	0.6565	0.004	162.954	0.0000	***

Significant Codes: 0 '***'

From this we can establish the spread as:

$$z_t = log(GS)_t - (0.6565)log(JPM)_t$$

Now that we have the spread equation ready, we can run the Augmented Dickey-Fuller test and other unit roots statistical tests to test for Cointegration.

Outline

1. Basic Idea of Pairs Trading
2. **Testing for Cointegration**
   - The idea of Pairs Trading based on Cointegration
   - Pairs Selection
   - Least Squares (LS) regression
   - **Augmented Dickey-Fuller (ADF) test**
   - Phillips-Perron (PP) test
3. Trading Strategy Design
   - Trading the Spread
   - Pairs Trading Performance (Simulation)
4. Pros and Cons of Pairs Trading
5. Other Pairs Trading Implementations

Augmented Dickey-Fuller (ADF) test

Mathematically, the ADF is based on the idea of testing for the presence of a unit root in an autoregressive time series sample. It makes use of the fact that if a price series possesses mean reversion, then the next price level will be proportional to the current price level. A linear lag model of order is used for the time series:

$$\Delta y_t = \alpha + \beta_t + \gamma y_{t-1} + \delta_1 \Delta_{t-1} + ... + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

Where α is a constant, β represents the coefficient of a temporal trend and $\Delta y_t = y(t) - y(t-1)$. The role of the ADF hypothesis test is to consider the null hypothesis that $\gamma = 0$, which would indicate with $\alpha = \beta = 0$ that **the process is a random walk and thus non mean reverting.** The first task is to calculate the test statistic (ADF_τ), which is given by the sample proportionality constant $\hat{\gamma}$ divided by the standard error of the sample proportionality constant:

$$ADF_\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

The test statistic is a negative number and thus in order to be significant beyond the critical values, the number must be **less than the critical values.**

Augmented Dickey-Fuller (ADF) test

ADF Test Result (Residuals)		
p-value	0.01494	
Critical Value	-3.2985	At 1%: -3.4316
		At 5%: -2.8621
		At 10%: -2.5671

The p-value is small and hence we have evidence to **reject the null hypothesis that the series possesses a unit root**. This means that the spread is stationary and hence **the J.P Morgan and Goldman Sachs stocks are indeed cointegrated**. Now we try the Phillips-Perron Unit root test.

Outline

1. Basic Idea of Pairs Training

2. **Testing for Cointegration**
   - The idea of Pairs Trading based on Cointegration
   - Pairs Selection
   - Least Squares (LS) regression
   - Augmented Dickey-Fuller (ADF) test
   - **Phillips-Perron (PP) test**

3. Trading Strategy Design
   - Trading the Spread
   - Pairs Trading Performance (Simulation)

4. Pros and Cons of Pairs Trading

5. Other Pairs Trading Implementations

Phillips-Perron (PP) test

The ADF test is based on the assumption that the error terms have constant variance and are statistically independent. The PP test, however, which was developed as a generalization of the ADF test, has a milder assumption regarding the error terms. The regression test takes the following $AR(1)$ form:

$$y_t = a_0 + a_1 y_{t1} + \epsilon_t$$

where the null is that $a_1 = 1$ and the alternative that $a_1 < 1$. Rejecting the null will indicate that y_t does not have a unit root and is therefore stationary.

Whereas the ADF test adds lagged differentiated terms to handle higher-order correlations, the PP test modifies the coefficient a_1 from the $AR(1)$ regression for the serial correlation in ϵ_t. The derivation of the PP-test is beyond the scope of this thesis.

Phillips-Perron (PP) test

Phillips-Perron (PP) Test Result		
p-value	0.040	
Critical Value	-2.944	At 1%: -3.43
		At 5%: -2.8621
		At 10%: -2.5671

At a 95% confidence level, Phillips-Perron test is below the critical value, therefore it is safe to **reject the null hypothesis that the spread contains a unit root** and hence assume that the spread was generated by a **stationary process**.

Outline

1. Basic Idea of Pairs Trading
2. Testing for Cointegration
   - The idea of Pairs Trading based on Cointegration
   - Pairs Selection
   - Least Squares (LS) regression
   - Augmented Dickey-Fuller (ADF) test
   - Phillips-Perron (PP) test
3. **Trading Strategy Design**
   - **Trading the Spread**
   - Pairs Trading Performance (Simulation)
4. Pros and Cons of Pairs Trading
5. Other Pairs Trading Implementations

Trading the spread

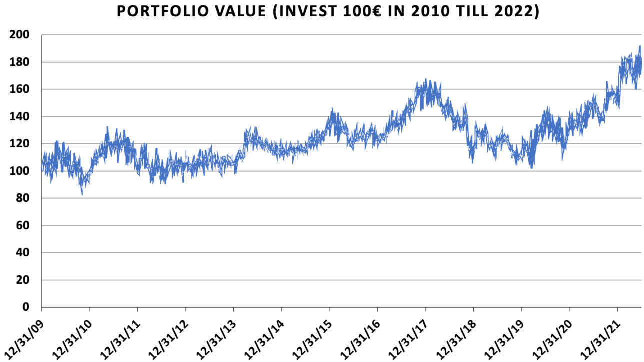
From our spread results, we standardize the spread series and set upper and lower boundaries to ± 1. Thus, our trading signals occur at $+1$ and -1. The strategy is the following:

- When the spread is above $+1$, then we sell the spread which means that we sell GS and buy JPM with units corresponding to cointegrating coefficient as explained above.
- When the spread is below -1, then we buy the spread which means that we buy GS and sell JPM with units corresponding to cointegrating coefficient as explained above.
- Whenever the spread converges again to 0, then we close the position.

Outline

1. Basic Idea of Pairs Trading
2. Testing for Cointegration
   - The idea of Pairs Trading based on Cointegration
   - Pairs Selection
   - Least Squares (LS) regression
   - Augmented Dickey-Fuller (ADF) test
   - Phillips-Perron (PP) test
3. **Trading Strategy Design**
   - Trading the Spread
   - **Pairs Trading Performance (Simulation)**
4. Pros and Cons of Pairs Trading
5. Other Pairs Trading Implementations

Pairs Trading Performance (Simulation)



Strategy return	**5%**
Beta	**-0.011875**

There are still fluctuations which shows that although the strategy is market-neutral, an investor is still exposed to **idiosyncratic risk** and the strategy is only market neutral as it provides a **return of 5% at an almost zero risk exposure** which is comparable to the market returns.

Pros and Cons of Pairs Trading

Pros

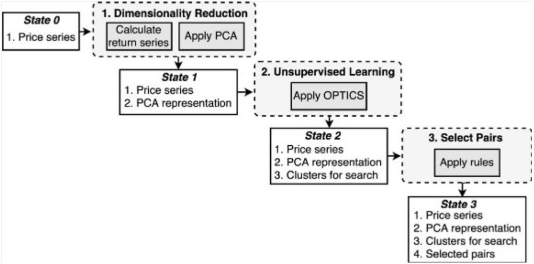
- Able To Mitigate Potential loses and Risks (one underperforms other possibly taking loses
- Profit Earning : regardless of the market
- Trader is completely hedged

Cons

- High commission charges (double of the fee of standard trade)
- Price Filling : Low margins-¿ large quantities . Risk of filling the stock orders at the desired price
- Hard to find rewarding pairs, because of the growing data and that it relies on high statistical correlation

Another example of Pairs Trading Implementation ML

- Limit space search PCA (High Dim –irrelevant features and curse
- Density based clustering algorithm
- Cointegration : Engle and Granger
- Hurst exponent 0-0.5 TS - mean-reverting
- Coherence Between mean-reversion time and trading period
- Threshold by Quantiles deciles or profit mix optimization
- Artificial Neural Networks: LSTM (Sequential dependency, History)



$$\{\alpha_S, \alpha_L\} = \underset{q}{\operatorname{argmax}} R^{\text{val}}\left(q\right),$$

$$q \in \left[\left\{Q_{f^-(x)}\left(0.20\right), Q_{f^+(x)}\left(0.80\right)\right\}\left\{Q_{f^-(x)}\left(0.10\right), Q_{f^+(x)}\left(0.90\right)\right\}\right]$$

where R^{val} is the return obtained in the validation set.

Another example of Pairs Trading Deep Reinforcement Learning

- DQN
- Produces Q function which maximizes profit
- Pairs-dickey fuller is in [0,0.05],std/mean¿0.5
- Returns next state and profit
- Updates input weights (ADAM optimizer) and repeats
- OPENAI gym environment

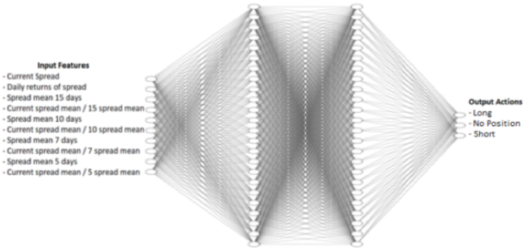


Figure 10: DQN NN structure

$$training\ rewards = action \times spread\ returns \times negative\ returns\ multiplier$$

$$testing\ rewards = action \times spread\ returns$$