Time Series Analysis
Statistical Arbitrage

April 15, 2020

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Administrative

1. Presentation Logistics
   - Schedule has been published!
   - 10-15 minutes with 5 minutes between
   - I will be strict on time
   - Points for class participation and asking questions

2. HWs
   - HW2 will be graded this weekend
   - HW3 (not graded) will be published this weekend

Lecture 10 Agenda

1. Guest Lecture on Time Series: Q McCallum
2. Statistical Arbitrage

Q McCallum
Time Series Analysis

Statistical Arbitrage

Statistical Arbitrage

In this lecture we will present an overview of statistical arbitrage and pairs trading strategies. We will review the concept of cointegration and examine its application to construction and implementation of basic pairs trading strategies.

History of Statistical Arbitrage

- The first practice of statistical pairs trading is attributed to Morgan Stanley in the mid 1980s (quant group led by Nunzio Tartaglia), where *quantitative arbitrage strategies* using state-of-the-art statistical techniques were developed and executed seamlessly through *automated* trading systems.
- At that time, trading systems of this kind were considered the cutting edge of technology.
- One of the techniques they used for trading involved trading securities in **pairs**. The process involved identifying pairs of securities whose prices tended to move together.
- Whenever an anomaly in the relationship was noticed, the pair would be traded with the idea that the anomaly would correct itself.
- This came to be known on the street as “pairs trading”. Since that time the pairs trading strategy has since increased in popularity and has become a common trading strategy used by hedge funds and institutional investors.

Motivation

- The general theme for investing in the marketplace from a *valuation* point of view is to sell overvalued securities and buy the undervalued ones.
- However, it is possible to determine that a security is overvalued or undervalued only if we also know the true value of the security in absolute terms. But this is generally very hard to do.
- Pairs trading attempts to resolve this using the idea of *relative pricing*; that is, if two securities have similar characteristics, then the prices of both securities must be more or less the same.
- The specific prices of the securities are not of importance. What is important is that prices of the two securities be the same.
- If the prices happen to be different, it could be that one of the securities is overpriced, the other security is underpriced, or the mispricing is a combination of both.

Pairs Trading

- Pairs trading strategy involves *selling* the higher-priced security and *buying* the lower-priced security with the idea that the mispricing will *correct* itself in the future.
- The mispricing between the two securities is captured by the notion of *spread*. The greater the spread, the higher the magnitude of mispricing and greater the profit potential of the pairs trading strategy.
- Often a *long–short* position in the two securities is constructed such that it has a negligible *beta* and therefore minimal exposure to the market.
- As a result, the returns from the trade are *uncorrelated* to market returns, a feature typical of *market neutral* strategies.

Pairs Trading

- The key to success in pairs trading lies in the *identification* of security pairs.
- In a study by Gatev et al., a purely *empirical* approach based entirely on the historical price movement of securities was compared to those created by *randomly* pairing the securities.
- The difference in the returns between the two groups was found to be positive and statistically significant.
- However, unlike the purely empirical approach of Gatev et al., the methodology most traders employ to identify profitable pair trades involves *fundamental* valuation methodologies based on *fundamentals* of the firm, which are further validated by *empirical* models and use of data analysis.
- A common theoretical explanation for the co-movement of security prices stems from arbitrage pricing theory (APT).

Arbitrage Pricing Theory

According to APT, if two securities have the same risk factor exposures, then the expected return of the two securities for a given time frame is the same. Let the the price of securities A and B at time t be $p_t^A$and $p_t^B$and at time t+i be $p_{t+1}^A$and p_{t+1}^B, respectively.

The i-period return for the two securities s given as $\log\left(\frac{p_{t+1}^A}{p_t^A}\right)$ and $\log\left(\frac{p_{t+1}^B}{p_t^B}\right)$.

- Assuming we have the prices of both securities at the current time t, the return on both securities is expected to be the same in all future times.
- Since the increment to the logarithm of the prices at the current time must be about the same for both the securities at all time instances in the future, the time series of the logarithm of the two prices must move together.
- We can formalize the idea of co-movement further as the formal concept has been well developed in the field of econometrics. We discuss it in the following section on cointegration.

Cointegration

- A common econometric approach to handling *nonstationary* time series is to *transform* it into a stationary time series by *differencing*.
- By extension, when analyzing multivariate time series where each of the component series is nonstationary, it would then make sense to difference each component and then subject them to the same examination.
- However, this turns out *not* always to be the case. For multivariate series even though component series are nonstationary, it may be that specific *linear combination* of them is *stationary*; that is, the two series move together in lockstep.
- This phenomenon is referred to as *cointegration* (Engle and Granger).

Definition: Cointegration

Let y_t and x_t be two *nonstationary* time series. If there exists a real number γ such that the series $y_t - \gamma x_t$ is stationary, then the two series are said to be *cointegrated*.

- Real life examples of cointegration abound in economics.
- In fact, the first demonstration and tests of cointegration involved economic variable pairs like consumption and in come, short-term and long-term rates, the money supply and GDP, etc.

Granger Representation Theorem

- The explanation for cointegration dynamics is generally captured by the notion of *error correction*. The idea behind error correction is that cointegrated systems have a *long-run equilibrium*; that is, the long-run mean of the linear combination of the two time series.
- If there is a deviation from the long-run mean, then one or both time series adjust themselves to restore the long-run equilibrium.
- The formal theorem stating that error correction and cointegration are essentially equivalent representations is called the *Granger representation theorem*.
- We shall not attempt to discuss the proof of the theorem, but simply present here the error correction representation.

Error-Correcting Representation

Error-Correcting Representation - Let ε_{x_t} be the white noise process corresponding to time series $\{x_t\}$ and ε_{y_t} be the white noise process corresponding to the time series $\{y_t\}$. The error correction representation is then given by:

$$y_t - y_{t-1} = \alpha_y (y_{t-1} - \gamma x_{t-1}) + \varepsilon_{y_t}$$

(EQ 1)

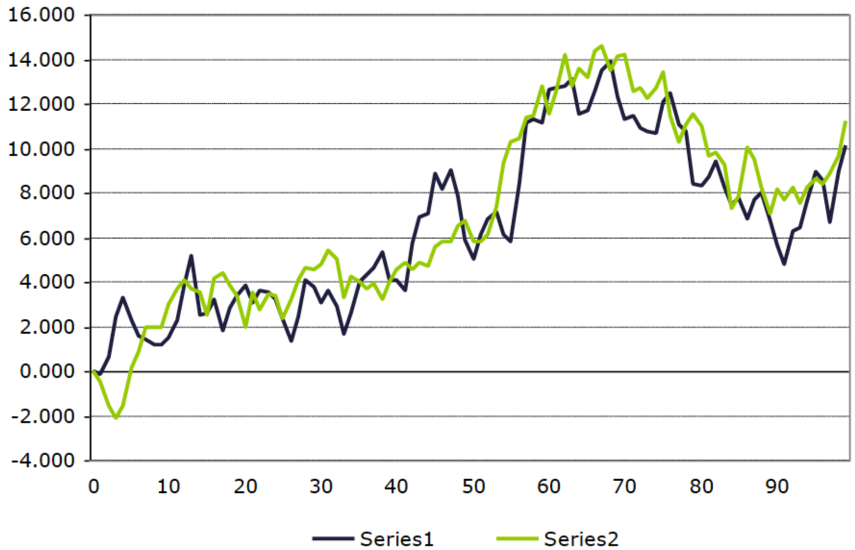
$$x_t - x_{t-1} = \alpha_x (y_{t-1} - \gamma x_{t-1}) + \varepsilon_{x_t}$$

The left-hand side is the usual increment to the time series at each time step. The right-hand side is, however, the sum of two expressions, the *error correction* part and the *white noise* part. The term $y_{t-1} - \gamma x_{t-1}$ in the error correction part is a representative of the deviation from the long-run equilibrium (equilibrium value is zero in this case), and γ is the coefficient of cointegration. α_y is the *error correction rate*, indicative of the speed with which the time series y_t corrects itself to maintain equilibrium (and the same hold for x). Thus, as the two series evolve with time, deviations from the long-run equilibrium are caused by white noise, and these deviations are subsequently corrected in future time steps.

Example

A Simple Example: Let us illustrate that the idea of error correction does indeed lead to a stationary time series for the spread. Two independent white noise series with zero mean and unit standard deviation were generated to represent ε_{x_t} and ε_{y_t}, respectively. The other values were set as $\alpha_y = -0.2$, $\alpha_x = 0.2$, and $\gamma = 1.0$.

Note that it is important to have the two coefficients α_y and α_x set to *opposite signs* for error-correcting behavior. The values for the two time series and were then generated using the simulated data and the equations from the error correction representation. A plot of the two series is shown in Figure 1 below.



'igure 1: Example of Cointegrated Time Series

Example

Subsequently, the spread at each time instance is calculated using the known value for γ.
A plot of the spread series and its autocorrelation structure (for lags 1 through 20) is shown in Figure 2 and 3 below.

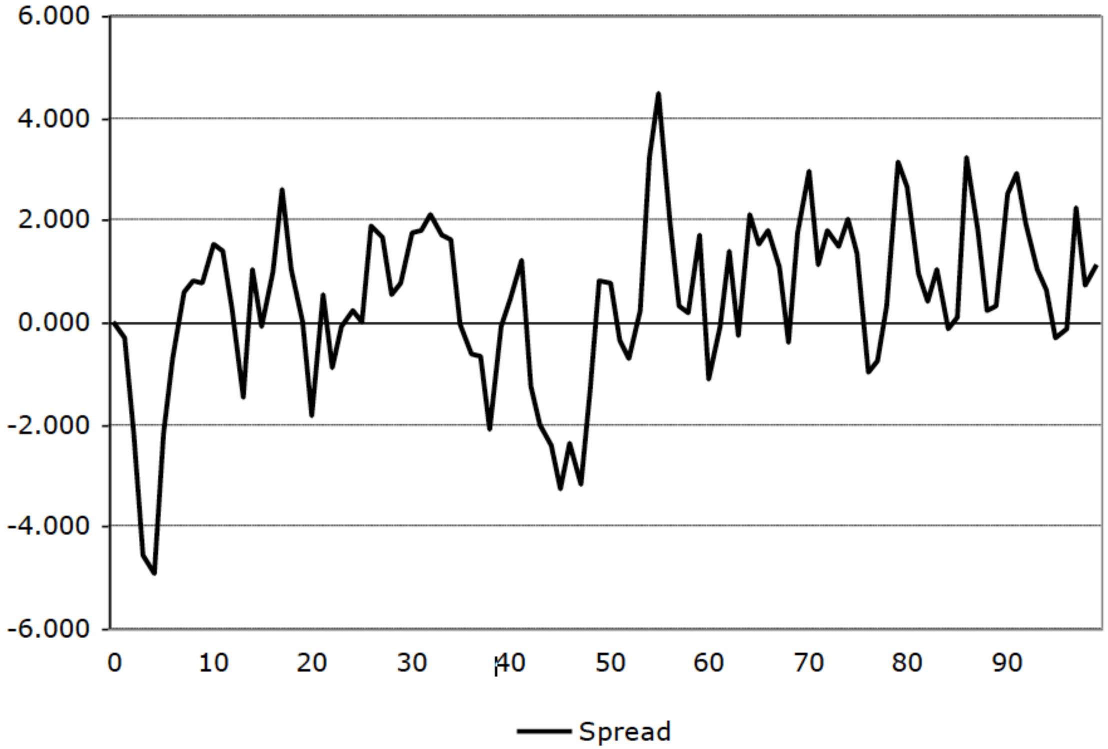


Figure 2: Spread Between Cointegrated Time Series

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Example

It is easy to appreciate from the autocorrelation function that the spread series is indeed stationary.

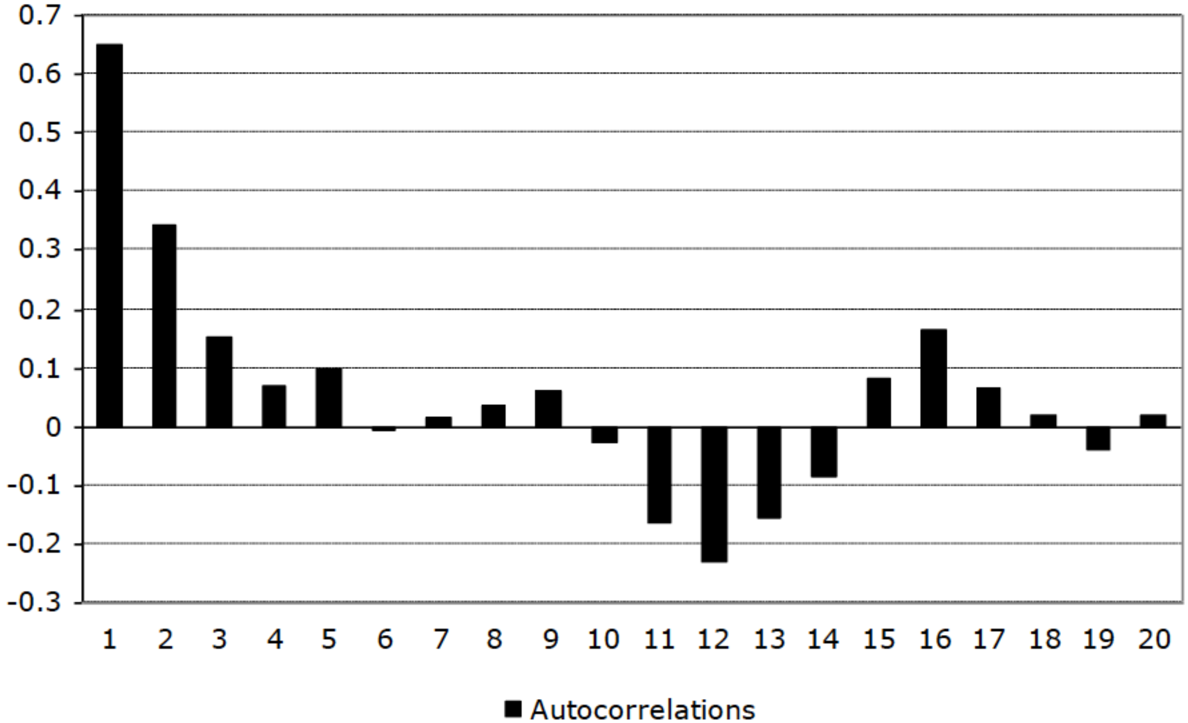


Figure 3: Autocorrelations of the Spread with Lag 1 through 20

Cointegration continued

A more direct approach to model cointegration is attributed to Stock and Watson, called the common trends model. The primary idea of the common trends model is that of a time series being expressed as a simple sum of two component time series: a stationary component and a nonstationary component. If two series are cointegrated, then the cointegrating linear composition acts to nullify the nonstationary components, leaving only the stationary components. To see what we mean, consider two time series

$$y_t = n_{y_t} + \varepsilon_{y_t}$$

(EQ 2)

$$z_t = n_{z_t} + \varepsilon_{z_t}$$

where n_{y_t} and n_{x_t} are the random walk (nonstationary) components of the two time series, and ε_{y_t} and ε_{x_t} are the stationary components of the time series. Also, let the linear combination $y_t - \gamma z_t$ be the cointegrating combination that results in a stationary time series. Expanding the linear combination and rearranging some terms, we have

$$y_t - \gamma z_t = \left(n_{y_t} - \gamma n_{z_t}\right) + \left(\varepsilon_{y_t} - \gamma \varepsilon_{z_t}\right) \qquad \text{(EQ 3)}$$

Cointegration continued

- If the combination in Equation 3 must be stationary, the nonstationary component must be zero, implying that $n_{y_t} = \gamma n_{z_t}$, or the trend component of one series must be a scalar *multiple* of the trend component in the other series.
- Therefore, for two series to be cointegrated, the trends must be identical up to a scalar.
- The transformation to stationarity is typically achieved using cointegration ideas and strict parity relationships.
- Needless to say, this approach appears as a recurrent theme in the design of trading strategies across all asset classes.

Cointegration continued

In this next section, we fit the cointegration model to the logarithm of stock prices.

- For the cointegration model to apply, we would require the logarithm of stock prices to be a nonstationary series.
- The assumption that the logarithm of stock prices is a random walk (read as nonstationary) is a rather standard one.
- It has been used fairly extensively in option-pricing models with satisfactory results.
- We are therefore good on that assumption and ready to proceed further.

Cointegration continued

Let us say that two stocks A and B are cointegrated with the nonstationary time series corresponding to them being $\{\log(p_t^A)\}$ and $\{\log(p_t^B)\}$, respectively. Applying the error correction representation described here, we have

$$\log(p_t^A) - \log(p_{t-1}^A) = \alpha_A \log(p_{t-1}^A) - \gamma \log(p_{t-1}^B) + \varepsilon_A$$

$$\log(p_t^B) - \log(p_{t-1}^B) = \alpha_B \log(p_{t-1}^A) - \gamma \log(p_{t-1}^B) + \varepsilon_B \qquad \text{(EQ 4)}$$

The parameters that uniquely determine the model are the cointegration coefficient g and the two error correction constants α_A and α_B. Therefore, estimating the model involves determining the appropriate values for α_A and α_B and γ. The left-hand side of the Equations 4 is the return of the stocks in the current time period. On the right-hand side, note the expression for the long-run equilibrium, $\log(p_{t-1}^A) - \gamma \log(p_{t-1}^B)$, in both the equations. In words, it is the scaled difference of the logarithm of price. Incidentally, this coincides with what we termed the *spread* in our earlier discussion. Also notice that the subscripts for stock prices in the expression for the long-run equilibrium is $t-1$.

Cointegration continued

- The past deviation from equilibrium plays a role in deciding the next point in the time series. Therefore, knowledge of the past realizations may be used to give us an edge in predicting the increments to the logarithm of prices; that is, returns. This is important and exciting(!)
- Even though both stocks follow a log-normal process, one can eke out some predictability in their returns based on past realizations.
- Thus, one can attempt to trade either of the stocks in the pair based on predictions using the estimated values from the error correction representation.
- Let's focus on the cointegration part of the representation theorem. This is the assertion that the time series of the long-run equilibrium (also termed *spread* in our case) is stationary and mean reverting.
- If we can associate the time series of the spread to a portfolio, we can consider trading the portfolio based on our prediction of the time series values.
- Consider a portfolio with long one share of A and short γ shares of B. The return of the portfolio for a given time period is given as:

$$\log\left(p_{t+i}^{A} / p_{t}^{A}\right) - \gamma \log\left(p_{t+i}^{B} / p_{t}^{B}\right) \qquad \text{(EQ 5)}$$

Shortcomings

- The shortcomings of statistical arbitrage strategies are easy to see; often enough, detected statistical relationships are random or “spurious” and have little predictive staying power.
- Yet other statistical relationships, those validated by academic research in economics and finance, have consistently produced positive results for many traders.
- Thorough understanding of economic theory helps quantitative analysts distinguish between solid and arbitrary relationships and, in turn, improves the profitability of trading operations that use stat-arb methodology.

Shortcomings

In addition to the issues embedded in the estimation of statistical relationships, statistical arbitrage strategies are influenced by numerous adverse market conditions.

- The strategies face a **positive probability of bankruptcy** of the parties issuing one or both of the selected financial instruments. Tough market conditions, an unexpected change in regulation, or terrorist events can destroy credible public companies overnight.
- **Transaction costs** may wipe out all the profitability of stat-arb trading, particularly for investors deploying high leverage or limited capital.
- The **bid-ask spread** may be wide enough to cancel any gains obtained from the strategy.
- Finally, the pair's performance may be determined by the **sizes of the chosen stocks along with other market frictions**—for example, price jumps in response to earnings announcements.

Careful measurement and management of risks, however, can deliver high stat-arb profitability. Gatev, Goetzmann, and Rouwenhorst (2006) document that the out-of-sample back tests conducted on the daily equity data from 1967 to 1997 using their stat-arb strategy delivered Sharpe ratios well in excess of 4.

Dynamic Adjustment

- Stat-arb strategies can be trained to dynamically adjust to changing market conditions.
- The mean of the variable under consideration, to which the identified statistical relationships are assumed to tend, can be computed as a moving weighted average with the latest observations being given more weight than the earliest observations in the computation window.
- Similarly, the standard deviation used in computations can be computed using a limited number of the most recent observations, reflecting the latest economic environment.

Statistical Arbitrage Characteristics

For our purposes:

- Short term
- Mean-reverting
- Risk Neutrality
- Large (sometimes) portfolios of diversified securities
  - Neutralize risk across some factors
  - But not all!!
  - (otherwise no excess returns)
- High co-integration (high correlation not enough – more on this later)
- Non-stationary time series (Kalman filter as test, etc.)

Intraday Stat Arb Strategies

- Pairs Strategy
- Generalized two-asset model
- Stocks vs. correlated baskets
- Index vs. cash (premium)

Pairs – Relative Value

- 2 instruments, <u>risk neutral</u> (e.g., one long, one short)
- Determine statistical relationship
- Establish a position when the relationship diverges sufficiently
- Reduce / eliminate the position when the the relationship converges
- Not just stocks: FX, bonds, commodities, cross, product. pretty much anything you can find a correlation between so long as you can maintain risk neutrality.
- Can be generalized to 2 *asset baskets:*
  - *Stock vs. stock*
  - *Stock vs. an index (future, ETF, etc.)*
  - *Stock vs. some other correlated basket*

Pairs Construction

1. Identify co-integrated pairs and establish the relationship
2. Establish the spread between the two securities as a stationary / stable process
3. Strategy rule
    a. When spread widens, buy the cheap asset, sell the expensive asset
        1. Sell the expensive asset
        2. Trade amounts should maintain risk neutrality (usually equal $ amounts)

    b. When the spread reverts, close the position

Index Arbitrage

1. Trade the index components vs. the index future
2. $FV_{future} = Fv_{cash\ index}$ + (cost of carry – dividends)
3. Future may trade at a premium (usually) or a discount
4. Arbitrage opportunity: calculate the “true” premium and then trade when there is sufficient deviation from the expected premium
5. May also be useful as a momentum signal (future as a leading predictor of index changes)
6. Premium assumptions assume constant cost of carry and dividends. Changes to either of these must be considered in strategy (e.g. major index component dividend announcements, FOMC rate changes or other economic news affecting interest rates, etc.)

Volatility Arbitrage

Attempts to profit from the difference between the forecasted future price-volatility of an asset and the implied volatility of options on the asset.

1. Forecast the volatility, for one or more options
2. Measure implied volatility from options pricing
3. If volatility is cheap (implied is lower than forecast):
    - Buy a long call
    - Short the underlying to remain delta-neutral
4. If volatility is expensive
    - Short a put
    - Buy the underlying
5. Complications
    - Forecast construction (obviously)
    - Maintaining delta hedge can reduce / eliminate arb profits
    - Timing of the reversion (though tends to happen quickly in liquid options)