Notes from Trading Systems

Hardik Rajpal

May 11, 2023

1 What To Expect

Trading techniques, not trading systems, are of the sort:

- System-Deduction-based trading
- Trade because your gut says so
- Trade because my friend says so
- Trade because the guy on TV says so
- Trade because my broker says so

A trading system (1) takes some inputs from the market, (2) runs specific logic, implemented by you and (3) gives you an output. The onus of deciding whether or not to make the trade is still on the human trader.
The module plans to discuss:

1. Pair trading (not related to that thing you suggested to your wife!)
2. Volatility based Delta hedging
3. Calendar spreads
4. Momentum strategy (Portfolio approach)

No trading system is complete without backtesting, which is not covered by this module.

2 Pair Trading Logic

If two entities A and B (listed on the stock market) are related (have similar business landscapes and environments), their stock prices are expected to behave similarly (barring any anomalies). If we observe that SP(A) moves up by X% on a particular day, we expected SP(B) to move up by y% too. If at the moment, this expected change has not occurred, we say that B has the cheaper stock and A has expensive stock.
Cheaper Stock: The upward change in its price is less than that of the stock prices of entities similar to it. It is undervalued.
Expensive Stock: The upward change in its price is more than that of stock prices of entities similar to it. It is overvalued.
Anomaly: An event that triggers the changes in stock prices of two closely related entities to deviate. Anomalies in stock prices give us an opportunity to trade.
The bulk of pair trading revolves around:

1. Identifying the relationship between two *pair* stocks.
2. Quantifying their relationship.
3. Tracking the behaviour of this relationship on a daily basis.

4. Looking for anomalies in the price behaviour.

Two popular techniques to define the relationship between two stocks are based on **price spreads and ratios** and **linear regression**.
Note that pair trading is not a market neutral strategy, because although you may be long and short, you are long and short on two different stocks. To be market neutral, you need to long and short on the same stock at the same time. In the "calendar spread",youare long and short on the same stock, expiring on two different dates.
Pair trading can also be called 'Relative Value Trading', as we are trying to profit from the relative value of the two securities. It's also called 'Statisical Arbitrage'. (Arbitrage refers to the simultaneous buying and selling of commodities to benefit from the difference in their prices.)
Note that undervalued and overvalued are relative terms. How this relative value is calculated is discussed next.

3 Pair Trading Method 1: Correlation

3.1 Tracking Pairs

Here comes some jargon:

- **Stock Value:** What someone is willing to pay for it; what the stock is worth.
- **Stock Price:** The stock's current value to buyers and sellers.
- **Spreads:** The word refers to the difference between the closing values of two stocks (in the pair trading world).

  $$\text{spread} = \text{CV(A) - CV(B)}$$

  Historical spread refers to spread data collected on a daily basis.
- **Differential:** It refers to the difference in closing prices of two stocks.

  $$\text{differential} = \text{CP(A) - CP(B)}$$

  Differentials are not great to track pairs on an intraday basis; they're best used as an end-of-day basis. Spreads can be used to track pairs on an intraday basis.
- **Ratio:** It's the stock price of A by the stock price of B.

  $$\text{ratio} = \text{SP(A)/SP(B)}$$

- **Divergence:** If the ratio or spread between two stocks is expected to move apart (the graph moves up), we call it divergence. Then, we try to make money by setting up a divergence trade.
- **Convergence:** If the ratio of spread is expected to decrease (the graph goes down), we call it convergence. We try setting up a convergence trade.

The question to make money then is, of course, is the pair going to diverge or converge? Even before this, how do we qualify two stocks as a pair? The second question's answer is correlation. The sign of the correlation tells us how the direction of changes are related and the magnitude tells us how strongly. Note that correlation allows for drawing probabilistic conclusions.
The correlation data only makes sense if the data is stationary around the mean; the dataset sticks close to the average values.

Not clear about the difference between spread and differential.

3.2 Pair Stats

3.2.1 Correlation and its Types

To estimate the correlation between the stock prices of two entities, say A and B, ensure that you have the same number of data points for both A and B (and at the same points in time). Also, the data should be "cleaned" for corporate actions such as bonus, splits, etc. We estimate daily returns for both stocks:

daily return at day t = (CP(t)/CP(t-1)) - 1

absolute per day change = CP(t) - CP(t-1)

We can use the `=Correl(col1, col2)` function in excel suitably. Note that the correlation function is commutative. the correlation calculated on the closing prices of stocks is called "closed correlation."

3.2.2 Statistic Variables

Mean: The arithmetic average. $\overline{x} = \sum_{i\in[n]} x_i/n$
Median: The middle number if n is odd ($x_{(n+1)/2}$), else it's $\frac{x_{n/2}+x_{(n/2)+1}}{2}$.
Mode: The most frequently occuring value: $x_M = \text{argmax}_{x_i} freq(x_i)$

3.3 Pre Trade Setup

3.3.1 Normal Distro

Note that the variables spread, differential and ratio are expected to be normally distributed. Here are some normal distribution facts (100/100):

- Within the 1st standard deviation lies 68% of the data.
- Within the 2nd standard deviation lies 95% of the data.
- Within the 3rd standard deviation lies 99.7% of the data.

3.3.2 Descriptive Stats

Excel functions `average()`, `median()`, `mode.mult()` are useful in the pursuit of mean, median and mode. Standard deviation is the most common measure variability and is used to determine the volatility of stock markets or other investments. (Excel: `Stdev.p()`)

$$\sigma = \sqrt{\sum_{i\in[n]} (x_i - \overline{x})^2} \tag{1}$$

Absolute deviation is also known as mean absolute deviation (MAD).

$$MAD = \sum_{i\in[n]} |x_i - \overline{x}| \tag{2}$$

MAD is seldom preferred because of the complications brought along with the absolute value. (Excel: `avedev()`) The mean, median, mode, standard deviation and absolute deviation are collectively known as **the basic descriptive statistics**.

3.3.3 The Standard Deviation Table

Standard Deviation			
	Spread	Differential	Ratio
3			
2			
1			
Mean	**0.06**	**228.52**	**1.87**
-1			
-2			
-3			

We use the normal distribution as a basis to deduce probabilistic conclusions like:

- If the current data point are at a far end of the bell shape, say 3σ, the next data point is likely to be closer to the mean with probability 99.7%. We say with 99.7% confidence that there is only 0.3% chance for the point to be further away from the mean.
- On second thoughts, shouldn't the probability of return be $99.7 + 0.15 = 99.85\%$ and the probability of it going higher be $0.3/2 = 0.15\%$?

3.4 The Density Curve

3.4.1 Selecting the Variable

We need to pick one of spread, differential and ratio to avoid getting confused by conflicting signals. It's up to the trader to choose the variable they are most comforable with. We go ahead with the ratio in the discussions ahead.

3.4.2 The Trade Trigger

Wherever the selected variable (ratio) is today, there is a great chance (quantifiable) that it will return to the running mean over the next few days. This phenomenon is known in the hood as **mean reversion** or **reversion to mean**. Think of the ratio as an abstract form of stock: we say that we can "short" (sell) the ratio when it's significantly above the mean (only to buy it back when it returns to the mean, hence profiting) or "long" (buy) it when it is significantly below the mean (only to sell it back when it returns, hence profiting).
Key Trigger to Trade: The ratio's current value w.r.t. its mean.
Above the mean $\Longrightarrow$ short the ratio!
Below the mean $\Longrightarrow$ long the ratio!
The only question that remains is:

$\boxed{\text{How far away from the mean do we wait for the ratio to go before acting?}}$

3.4.3 The Density Curve

We need to initiate a trade only when we are reasonably ceratin that the ratio will slide down to the mean value, **as quickly as possible.**. Here we put our faith in the good ol' Normal Distribution's properties. If the ratio is 3 SDs away from the mean, the chance of it going further away is 0.3%; the "chance of it reverting to the mean" is 99.7%. Using this idea, we filter out opportunities to initiate trades only at points where the likelihood of mean reversion is high.
As the trigger to trade depnds on the ratio and its standard deviation, we need to track the daily standard deviation of the ratio as opposed to the ratio itself. This is achieved by tracking the 'Density Curve' of the ratio.
The density curve is the continous counterpart of the relative frequency histogram; the bin sizes tend to zero. This is the cumulative distribution function, cdf(x). In Excel, the function NORM.DIST function is useful for this:

NORM.DIST(x, μ, σ, isCumulative)
if(isCumulative):
    NORM.DIST(x, μ, σ, isCumulative) $\equiv$ cdf(x;μ, σ)
else:
    NORM.DIST(x, μ, σ, isCumulative) $\equiv$ pdf(x;μ, σ)

3.5 The Pair Trade

Here are some density curve (cdf(x)) facts (100/100):

- cdf(x_{curr}) = 0.997 $\implies$ $x_{curr} = \mu + 3\sigma$
- cdf(x_{curr}) = 0.974 $\implies$ $x_{curr} = \mu + 2\sigma$
- cdf(x_{curr}) = 0.84 $\implies$ $x_{curr} = \mu + \sigma$
- cdf(x_{curr}) = 0.16 $\implies$ $x_{curr} = \mu - \sigma$
- cdf(x_{curr}) = 0.025 $\implies$ $x_{curr} = \mu - 2\sigma$
- cdf(x_{curr}) = 0.003 $\implies$ $x_{curr} = \mu - 3\sigma$

General thresholds for trading short and long are given below. The target value of the density curve is the value you want the density curve to go to after you make the trade for considerable profit. Note: the stoploss is the value of the density curve (or more generally, the parameter we are observing) at which it is considered smart to pull out of the trade.

- cdf(x_{curr}) $\in$ (0.003,0.025) $\implies$ long. (Stoploss=0.003 or higher, target = 0.25 or lower)
- cdf(x_{curr}) $\in$ (0.975,0.997) $\implies$ short. (Stoploss=0.997 or higher, target = 0.975 or lower)

In concrete terms, if the ratio is Stock A/Stock B,

- long trade $\implies$ buy Stock A and sell Stock B.
- short trade $\implies$ sell Stock A and buy Stock B.

Remember it as the numerator being the dominating term and long (buy) $\implies$ buy the numerator while short (sell) $\implies$ sell the numerator. The denominator always suffers the alternative fate.

3.5.1 Real Talk

- For a given pair, one can expect at most 2-3 signals in a year; we need to observe multiple pairs to continuously find profit opportunities.
- We have not discussed neutrality of both the positions, which is apparently a key angle.
- As pair trading is a margin money guzzler, one needs to have sufficient funds to pair trade.

4 Pair Trading Method 2: Statistical Arbitrage

This is also known as **Relative Value Trading**.

4.1 Straight Line Equation

Straigh line equation: $y = mx + \epsilon$

4.2 Linear Regression

Given $\{x_i, y_i\}_{i\in[n]}$, we estimate m and ϵ.

$$m = \frac{(n\sum_{i\in[n]} x_i y_i) - (\sum_{i\in[n]} x_i)(\sum_{i\in[n]} y_i)}{(n\sum_{i\in[n]} x_i^2) - (\sum_{i\in[n]} x_i)^2} \tag{3}$$

$$\epsilon = \frac{\sum_{i\in[n]} y_i - m\sum_{i\in[n]} x_i}{n} \tag{4}$$

$$\text{residual} = y_i - (mx_i + \epsilon)$$

4.3 The Error Ratio

When x and y are replaced by two stock prices, deciding which stock depends on the other is important. This is determined by three things:

1. Standard Error
2. Standard Error of intercept
3. The ratio of the above two variables (aka the error ratio)

The said ratio helps quantify how strongly an linear regression equation predicts the dependent variable.

4.3.1 Residuals

The aim of the regression analysis here is not to predict the prices, but to identify a pattern in the residuals of the predicted linear relation. Once said pattern is identified, we can work backwords to construct a trade; a trade which involves buying and selling two stocks, hence qualifying it as a pair trade.

4.3.2 Standard Error and that of the Intercept

$$\text{standard error} = \text{standard deviation of residuals} = \sqrt{(\sum_{i\in[n]} x_i)^2 - \sum_{i\in[n]} (x_i^2)}$$

The standard error of the intercept is a measure of how much the intercept itself can vary. It's the variance of the intercept.

4.3.3 Deciding the dependence

We pick x and y labels for two stocks so that the error ratio is minimzed.

4.4 The ADF test

To say two time series are "co-integrated" means that they move together and deviations, if any, are temporary or can be attributed to a stray event, and one can expect the two time series to converge and move together again. To check if two stocks are co-integrated, we run a linear regression analysis on the them and obtain the residuals; the two stocks are co-integrated if the residuals are stationary. Of course, two stocks that are co-integrated make for a pair ripe for tracking trading opportunities.

If the residuals of a regression analysis are not staionary, the regression relation shouldn't be used. Speculation and setup of trades on a co-integrated time series is a lot more meaningful and independent of market direction. **The ADF test** is used the check if the residuals are stationary or not.

4.4.1 Stationary series

We define three statistical conditions for a time series:

1. The mean of the time series should be within a "tight" range.
2. The standard deviation of the series should be within a range.
3. There should be no autocorrelation within the series: f(n) should not depend on f(i) for any i less than n.

Note that conditions 1 and 2 are checked by splitting the series into 3 parts, evaluating the mean and standard deviation of the parts separately and verifying that they lie within a range. Condition 3 is verified by considering two sub-series: 1....n-k and k+1...n, and finding the ("k-lag") correlation between these two; a small value of correlation implies there is no autocorrelation.
We say a pair exhibits complete stationarity (or has stationary residuals) if its residuals satisfy all three conditions. Weak stationarity is when at least one of these conditions are statisfied; otherwise, the residuals are non-stationary. In pair trading, we are only interested in completely stationary pairs.

4.4.2 The ADF test

It's the Augmented Dickey-Fuller test to determine the stationarity of a time series; we use it to determine the stationarity of the residuals series. The test outputs a probability of the series NOT being stationary; this is called 'the P value'. A time series is stationary if the P value is less than 0.05. The ADF test is complicated; so, use a plugin or something.

5 Trade Identification

We extend the linear equation to include the residual term:

$$y = mx + c + \epsilon \tag{5}$$

The slope tells us how many stocks of x would equal the price of y. Going long on y by d and going short on x by md, "hedges away the directional risk" associated with the pair. But we're still left with $c + \epsilon$. With good reason, the lower the intercept and its standard error, the better.
As the residual is a stationary time series, we can apply the properties of the normal distribution and trigger a trade only when it hits the upper of lower standard deviation.

- Go long on the pair (buy y, sell x) when $\epsilon \approxeq -2\sigma$
- Go short on the pair (sell y, buy x) when $\epsilon \approxeq 2\sigma$

6 Live Examples

The report of pair data mentions five variables:

1. Intercept (c)
2. Beta (m)
3. ADF Value (P)
4. Std_err: this is the ratio of today's residual over the standard error of the residual, which is reported in the regression output.

   $$\text{std\_err} = \frac{\text{today's residual}}{\text{std. err of residuals}}$$

   The std_err captures the number of standard deviations the residual is (the factor used for triggering trade).

5. Sigma: The standard error of the residuals from the regression output.

To remain **beta neutral**, we trade beta stocks of x for every stock of y. It's worth noting that we have no idea as to which stock will move to lower the difference that has risen, but by remaining we **beta neutral**, we ensure ourselves a profit so long as the two stocks eventually converge.

$\boxed{\text{What is a zscore?}}$

6.1 Position Sizing

Lot sizes of stocks constraint the choices of quantity that we can make in trading. TODO: understand what these are and summarize here.
The intercept represents the part of a stock price that is not "explained" or captured by the regression analysis' predicted dependence. If a large fraction of the price is unexplained, say $(y_i/c) \approxeq 80\%$, the trade is not trustable.

7 Calendar Spreads