FINA4380
Algo Trading Strategies, Arbitrage and HFT
A Cointegration based Statistical Arbitrage Strategy on Cryptos Futures
Final Report

Ka Hang Chan * Wei Zhia Kua † Jialu Li ‡ Pui Shan Or §

August 13, 2022

Abstract

This project is focused on using statistical arbitrage strategy to employ price inefficiency between mean-reverting portfolios with 2 or more components using crypto futures. We are interested to test if using more components in a stationary portfolio will yield better trading profits. We use Engle-Granger test to find cointegrated pairs and Johansen test to find cointegrated portfolios with three or more tickers. We then select the top n pairs/portfolios based on p-value and half-life criteria. Backtesting results and analysis are included at the end.

*1155126925@link.cuhk.edu.hk
†1155098495@link.cuhk.edu.hk
‡1155107895@link.cuhk.edu.hk
§1155108221@link.cuhk.edu.hk

1 Introduction

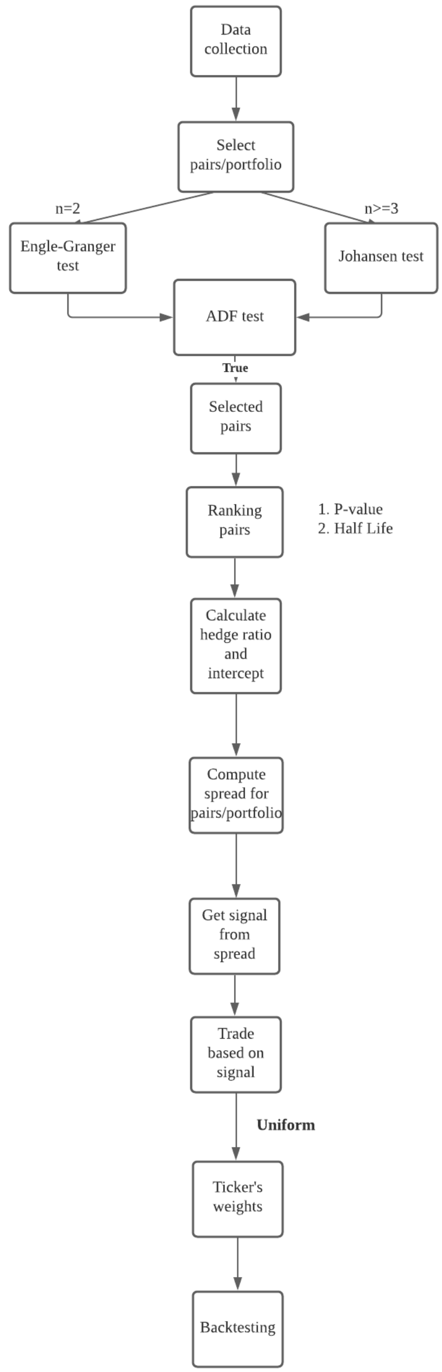
In time series analysis, cointegration refers to two or more time series that are unit-root stationary when linearly combined, where each of them are unit-root non-stationary. We hypothesize that the relationship between cointegrated cryptos is more stable (less volatile) than the crypto price processes, such that using more components to form stationary portfolios can yield a better profits. We use crypto perpetual futures to implement the strategy, which has the advantage to trade on margin and to trade both directions.

Our trading strategy uses both Engle Granger test and Johansen test to determine the cointegration pairs and portfolios. Then, we select the top 10 stationary portfolios using two different methods: T-statistics and half-life. Lastly, we use fixed beta approach to find the hedge ratios and allocate the portfolios accordingly using a uniform weighting scheme.

2 Trading Algorithm

2.1 Overview

An automated system is implemented to download the data from Binance. We then find stationary portfolios with 2, 3, 4 and 5 components using Engle-Granger test (for pairs) or Johansen test. Stationary pairs/portfolios are then selected based on the p-value from ADF test and ranked in ascending order based on two different approaches: p-value and half-life. Based on the top 10 selected portfolios, we compute the hedge ratios using a fixed beta approach, and get the spread of the pairs/portfolios to generate trading signals. A more detailed illustration on the trading strategy is shown in figure 2. Below flowchart gives a brief overview of our strategy.


Data collection
Select pairs/portfolio
n=2
n>=3
Engle-Granger test
Johansen test
ADF test
True
Selected pairs
Ranking pairs
1. P-value
2. Half Life
Calculate hedge ratio and intercept
Compute spread for pairs/portfolio
Get signal from spread
Trade based on signal
Uniform
Ticker's weights
Backtesting

2.2 Stationary pairs or stationary portfolios

2.2.1 Engle-Granger Test

Engle-Granger Test is used to determine if the price series of two assets are cointegrated. We first fit a linear regression line using Ordinary Least Squre (OLS) approach. Next, an Augmented Dickey-Fuller unit root test (ADF Test) will be conducted on the residuals of the regression. If the p-value of ADF Test is less than some specified critical size, α, those assets will be selected to form a portfolio.

2.2.2 Johansen Test

Vector error correlation form:

$$\Delta y_t = \Pi y_{t-1} + \epsilon_t$$

Johansen Test is used to determine if two or more asset price series are cointegrated. A vector error correction form is constructed (Π matrix) using log-cumulative returns. We compute eigenvalues of the Π matrix, and arrange it in a descending order. The test can be conducted based on two test statistics, namely Maximum Eigenvalue statistic and Trace Statistics, which can be computed using the eigenvalues, $\hat{\lambda}_i$.
Maximum eigenvalue statistics is denoted as:

$$\lambda_{max} = -T \ln(1 - \hat{\lambda}_{r+1}), r = 0, 1, ..., n - 1$$

where T is a factor which ensures the asymptotic distribution of the test statistics for $T \to \infty$.

Trace statistic is denoted as:

$$\lambda_{trace} = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i), r = 0, 1, ..., n - 1$$

The test is conducted sequentially with different null and alternative hypothesis as stated below. If all null hypothesis except for the last one are rejected, those assets will be selected to form a portfolio.

	λ_{max}test		Trace test	
	H_0	H_1	H_0	H_1
(i)	r = 0	r = 1	r = 0	r $\geq$ 1
(ii)	r $\leq$ 1	r = 2	r $\leq$ 1	r $\geq$ 2
.	.	.	.	.
.	r $\leq n-1$	r = n	r $\leq n-1$	r = n

2.3 Selection of top n pairs/portfolio

For diversification and risk management purposes, we select the top 10 most stationary portfolios based on two different approaches for comparison, namely p-value and half-life criteria.

2.3.1 p-value

We get the p-value from Engle-Granger test and Johansen test and arrange in ascending order to select the portfolios with the lowest n value. This intuition is that this will form the top 10 most stationary portfolios.

Algorithm 1: p-value from Johansen test

```
Input:  log cumulative returns
begin
    (1) Compute eigenvectors
    (2) Get portfolio spread by multiplying log cum returns with eigenvectors
    (3) Fit spread using ADF test
return: p-value from ADF
```

2.3.2 Half-life

For comparison, we select the portfolios that has the lowest n half-life. We model the spread using the Ornstein Uhlenbeck (OU) Process, denoted by

$$dx_t = \theta(\mu - x_t)dt + \sigma dW_t$$

where θ is the mean reversion speed. Half-life is then defined as $t_{1/2} = -log(2)/\theta$. The intuition is that portfolios that take shorter time to revert back to the mean allows us to have a higher trading frequency. However, the drawback is that a low half-life may be contributed by low volatility, which means the profitability will be restricted.

2.4 Compute hedge ratios

We have two different approaches for computing the hedge ratio, β.

For the portfolio that has only 2 components, a regression line is fitted using the log-cumulative returns of two assets. We obtain regression coefficient, β and then form a ratio vector $[1 \quad \beta]$.

For the portfolio that has 3 or more components, we directly use the eigenvectors produced by `coint johansen` and then form a ratio vector by dividing the maximum values.The ratio vector will be in a form of $[\gamma_1, 1, \gamma_2]$, where $\gamma = \frac{\beta_i}{\beta_{\max}}$

As a final step, a ratio vector will be normalised as follows.

1. Taking absolute values of all elements and sum all of them

2. All elements of the ratio vector will be divided by the sum calculated in Step 1

3 Assumptions

3.1 Strategy Assumption

- Portfolios are stationary during prediction windows with $\alpha = 0.05$,
- Maximum portfolio traded: 10,
- Number of maximum components inside a portfolio: 2, 3, 4 or 5.

3.2 Trading Assumption

Commission fee	0.04%
Leverage	10x
Auto_margin	0.1
Trading period	July - Oct 2021
Backtrader assumption	Buy on next opening price

Table 1: Trading assumption

3.3 Dataset

An automated system is implemented to download the data from Binance into csv files.

Number of tickers	24
Time Frame	30 minutes
Observation period	Apr - July 2021
Trading period	July - Oct 2021
Lookback period	4380 bars (90 days)
Prediction period	96 bars (2 days)

Table 2: Dataset

3.4 Signal Generation

Algorithm 2: Signal Generation

Input: Spread

begin

  (1) Compute signal line and stop loss line

  (2) Generate signals (short signal, long signal, short stop loss signal, long stop loss signal)

  (3) Repeat step(2) for top n portfolios to get signal matrix

  (4) Compute weighted portfolio (dim: time index * number of tickers)

return: Weighted portfolio

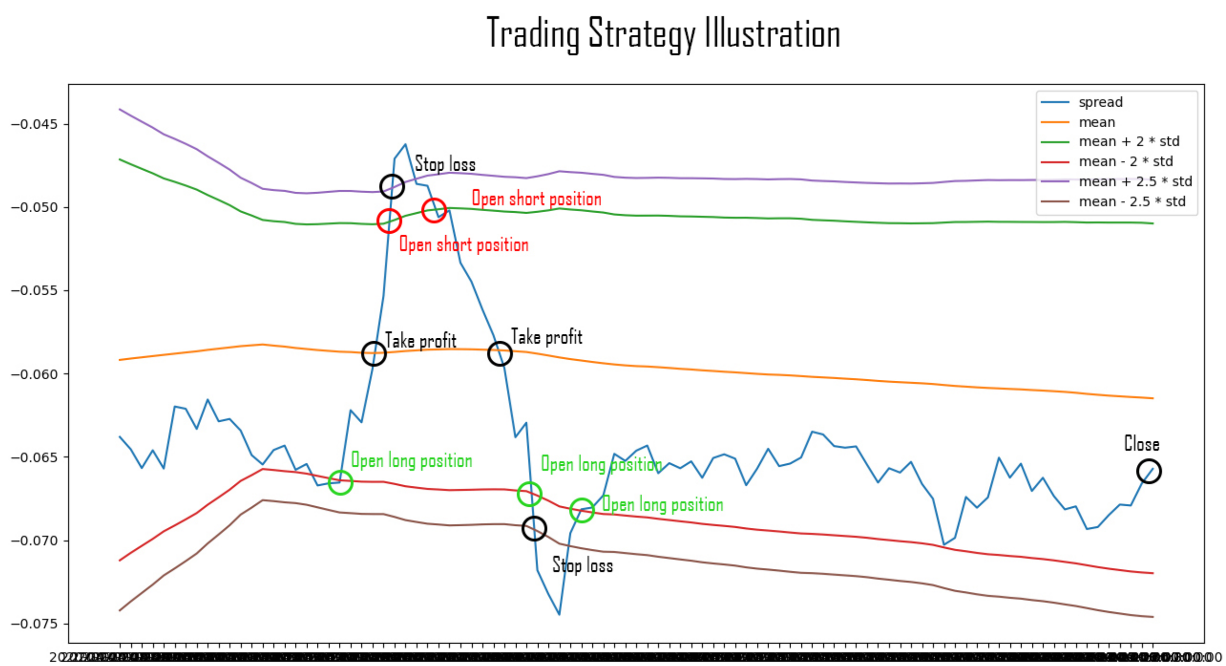


Figure 2: Trading strategy illustration

4 Backtesting Performance

Below are the returns by using different selecting criteria of portfolios based on two approaches. It's expected that a smaller t-stat and lower half-life will yield a more profitable strategy since the portfolios are more stationary. With lower half-life, we can expect to trade in higher frequency too.

	T-stat			
	2	3	4	5
Cumulative returns	12.1%	301.2%	-6.9%	-26.6%
Annual volatility	208.1%	181.9%	218.3%	157.1%
Sharpe ratio	1.103	2.957	0.981	0.229
Calmar ratio	0.631	95.663	-0.277	-1.137
Max drawdown	-57.6%	-44.1%	-61.7%	-49.9%
Daily value at risk	-25.3%	-20.8%	-26.7%	-19.7%

Table 3: T-statistics as Selecting Criteria

Above is the table of the backtesting results using t-statistics as the selecting criteria. The column number refers to the maximum component allowed per portfolio. The cumulative returns are 12.1%, 301.2%, -6.9% and -26.6% respectively. The model performs best when the maximum component equals to 3, which yields a 300% returns in 3 months.

Notice that the Calmar ratio seems to be extraordinarily high. Recall that the Calmar ratio is defined as the ratio between the annual return and the maximum drawdown. This is because Backtrader assumed that the cumulative returns for 3-month will be the same during every quarter.

Other statistics, like the annual volatility, VAR and max drawdown of our 3-component portfolio are also acceptable. We pick this portfolio as our best performance model in our project.

	Half-life			
	2	3	4	5
Cumulative returns	225.2%	-48.9%	-23.2%	-12.3%
Annual volatility	313.9%	183.9%	147.6%	155.7%
Sharpe ratio	2.472	-0.049	0.236	0.544
Calmar ratio	40.191	-1.409	-1.083	-0.606
Max drawdown	-58.3%	-59.4%	-47.2%	-49.5%
Daily value at risk	-36.5%	-23.2%	-18.5%	-19.3%

Table 4: Half-life as Selecting Criteria

Here is the table of the backtesting result of using the half life as the selecting criteria. The best performance model is using 2 as the maximum component per portfolio. Below shows the comparison of using T-stat and half-life selection criteria. Notice that the t-stat approach perform better than the half-life approach.

	T-stat, 2 Components	Half-life, 3 Components
Cumulative returns	301.2%	225.3%
Annual volatility	181.9%	313.9%
Sharpe ratio	2.957	2.472
Calmar ratio	95.663	40.191
Max drawdown	-44.1%	-58.3%
Daily value at risk	-20.8%	-36.5%

Table 5: Compare T-statistics v.s. Half-life as Selecting Criteria

As shown in the graph below, the peak of drawdown and the valley of the returns both happened in similar period. This may be due to structural changes of the price series in the portfolio, and hence it violates with the assumption of stationarity which is required for our strategy.

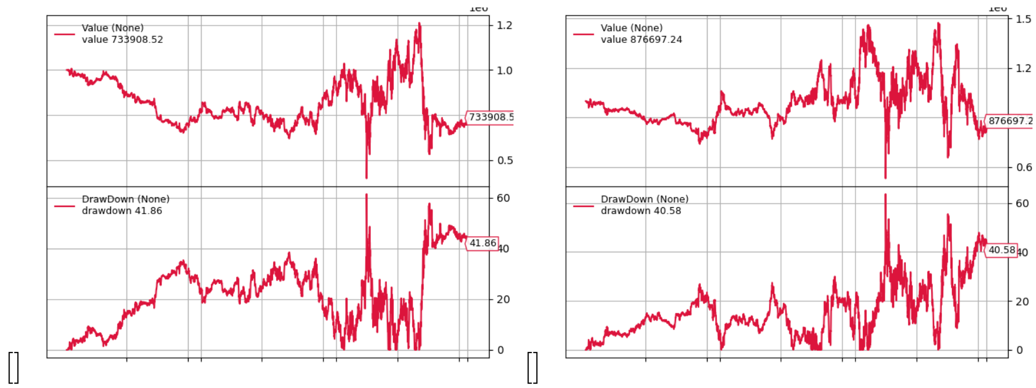


Figure 3: (a) Cumulative returns and max draw-down of portfolio with 5 components using t-test
(b) Cumulative returns and max draw-down of portfolio with 5 components using half-life

Below is the graph of the cumulative returns and maximum drawdown of portfolio when 3 components are used. The maximum drawdown is 44.1% which occured at around late September 2021.

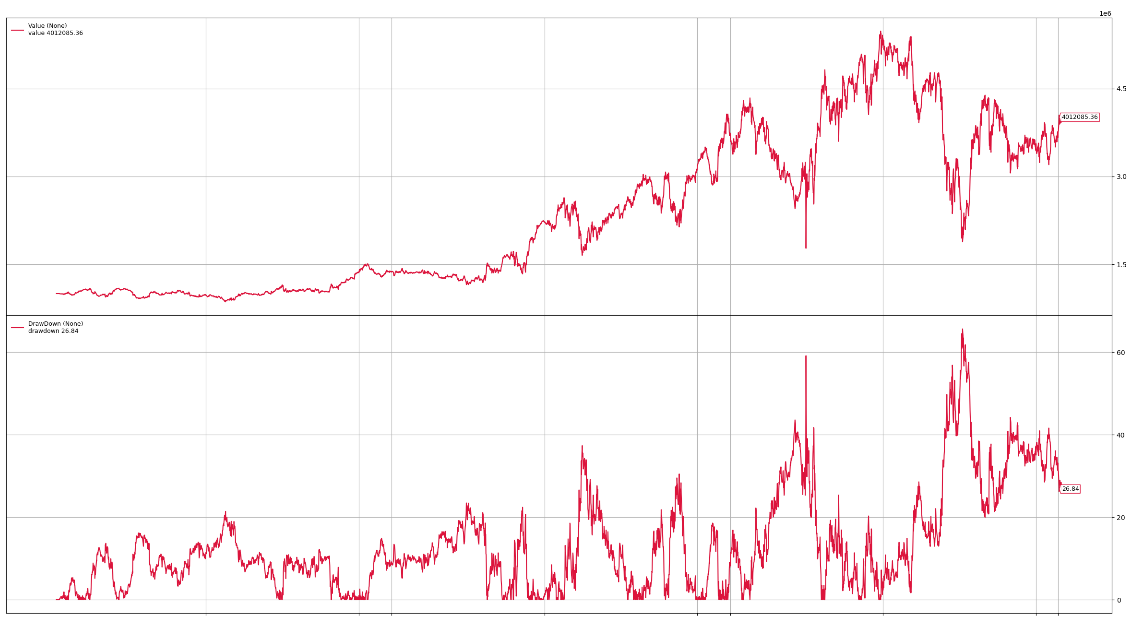
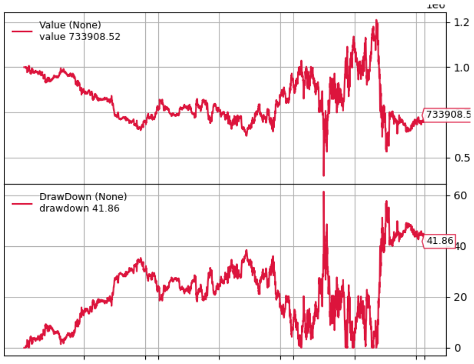


Figure 4: Cumulative returns and max draw-down of portfolio max component is 3

Lastly, we compare between the performance using two different selection criteria. We see that using t-stat yields higher returns.



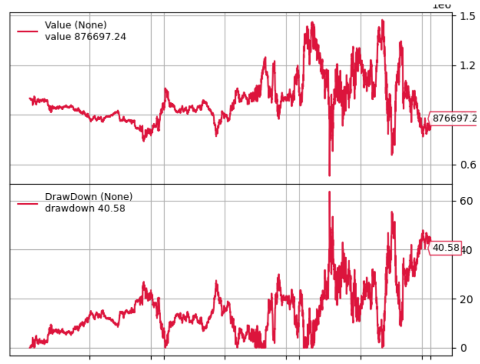


Figure 5: (a) Cumulative returns and max draw-down of portfolio with 5 components using t-test
(b) Cumulative returns and max draw-down of portfolio with 5 components using half-life

Below summarises the findings from our strategy:

- Finding(1): As max component per portfolio increases from 2 to 5, the profit shows a concave curve, with peak profit at max component = 3 when using t-stat as selecting criteria, and the the peak is reached at max component = 2 when using half-life as selecting criteria

- Finding(2): Performance of half-life selection criteria is poorer than t-stat in general

Based on the results, we conclude that using more components in a stationary portfolio will not necessarily yield a higher profit.

5 Further Improvement

- Kalman Filter

  Currently, we use a fixed beta approach. Alternatively, we can use a dynamic approach to update the hedge ratio by using Kalman filter.

- Granger Causality Test

  After selecting the portfolio, we can apply Granger Causality test to select the best regressor to fit a OLS model.

- Parameter Optimization

  Grid search and cross validation techniques could be used to find the optimal parameters.

- Test on More Tickers

  We use 24 tickers due to limited computational power, but we expect that using more tickers will potentially yield a better trading profit.

Appendix

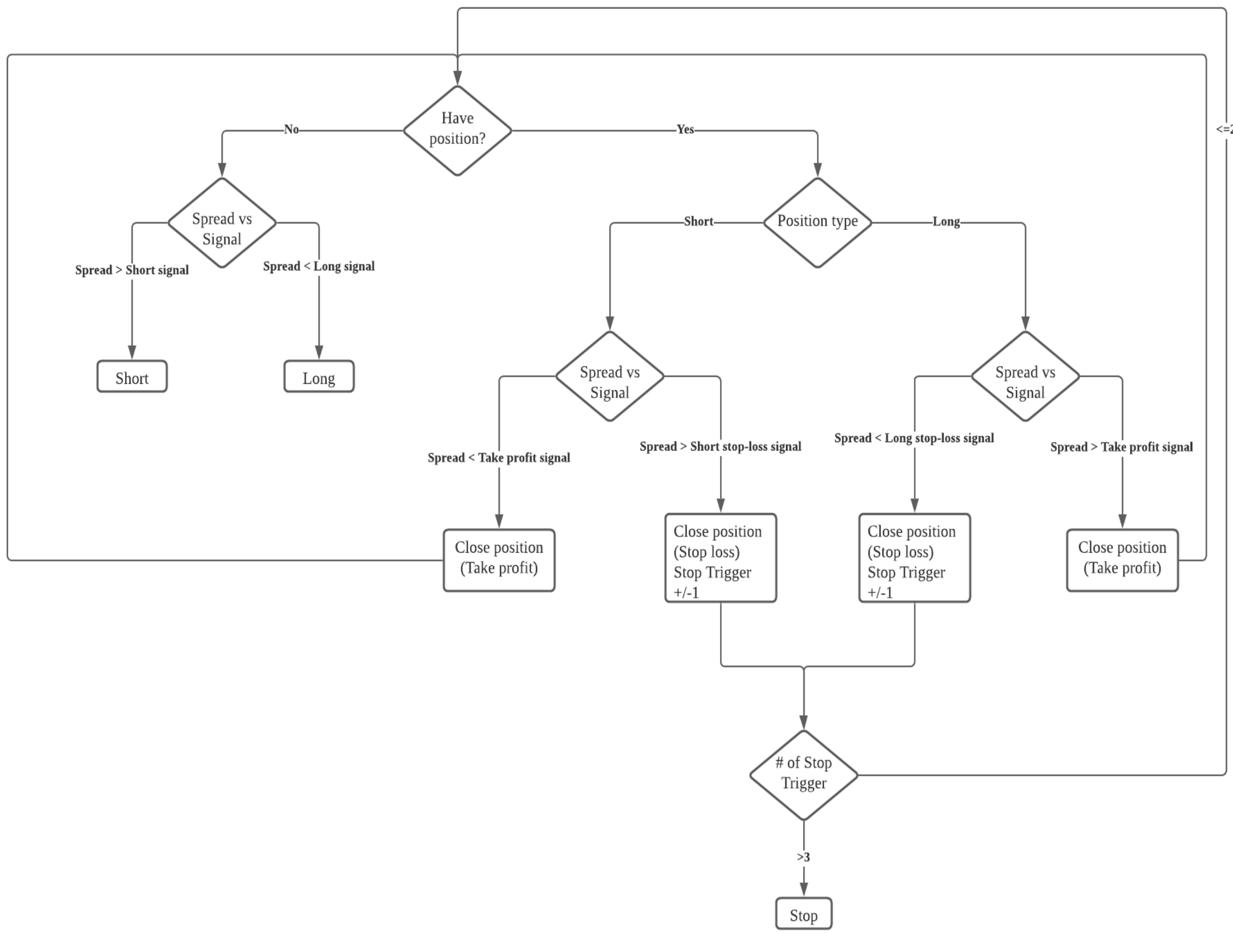


Figure 6: Signal generation flow chart

Reference

1. https://warwick.ac.uk/fac/soc/economics/staff/gboero/personal/hand3$_v$*ecm.pdf*