Cointegration Pairs Strategy Model

Objective

The objective of this assignment is to develop a pairs trading strategy for the cryptocurrency market.

Table of Contents

1. Introduction
2. Cointegration Pairs Trading Strategy
3. Data Collection
4. Testing for Cointegration
5. Strategy Development
6. Backtesting
7. Results
8. Use in Practice
9. Literature List

Introduction

This project aims to develop a pairs trading strategy by identifying cointegrated pairs of cryptocurrencies and leveraging their mean-reverting behavior. The strategy involves collecting historical data, testing for cointegration, developing trading rules, and backtesting the strategy to evaluate its performance. For this project, I had one week to work on it and had the following requirements: “Use Python and the following libraries: `ccxt`, `vectorbt`, and/or `backtrader`. The timeframe of data collection needs to be on 1 minute.” All the CSVs used for the code are also shared within the Jupyter Notebook.

Cointegration Pairs Trading Strategy

Cointegration is a term widely recognized in econometrics. It all started in 1981 when Granger (Econometrician) published a paper about cointegration in the Journal of Econometrics. In the upcoming period, Granger collaborated with Engle (Economist) to release a new paper about cointegration in 1985. The entire econometric community validated and acknowledged the concept of cointegration in 1987, and it was published by Econometrica (Journal of the Econometric Society) (Meuriot, 2016). Engle & Granger (1987) define a vector of time series as cointegrated if each element is stationary only after differencing, while linear combinations of them are themselves stationary. The concept of cointegration is used to capture the notion that nonstationary variables may possess long-run equilibrium relationships and thus have a tendency to move together (Blake & Fomby, 1997).

Cointegration is a test used in pairs trading, a popular strategy mostly used by hedge funds that can be applied in different financial markets and was introduced in the ‘80s by Gerry Bamberger. The quantitative group was led by Nunzio Tartaglia at Morgan Stanley (Bookstaber, 2007). The idea behind pairs trading is to identify a pair of stocks with a high degree of correlation in price movements that will tend to behave similarly in the future (Mohandas, 2023). A typical form of pairs trading involves selling (short) the stock at a relatively high price and buying (long) the other stock at a relatively low price (diverge), expecting that the higher-priced stock will decline and the lower-priced stock will rise in the future (converge) (Clegg & Krauß, 2017). The price gap (spread) of the two stocks acts as a signal to open and close positions on a certain threshold. When the position is closed on the desired threshold, the eventual profit is collected (Huang et al., 2015). Cointegration-based pairs trading is particularly notable due to its superior profitability results compared to other strategies (Huck & Afawubo, 2015; Rad et al., 2016; Blázquez et al., 2018).

Data Collection

For this assignment, crypto data was collected using the `ccxt` package to retrieve data from ByBit via the API. The data collected contains the top ~100 cryptocurrencies by market cap, excluding stable coins and cryptocurrencies not paired with USDT on the ByBit exchange. Larger market cap cryptocurrencies were chosen as they tend to have more stable prices, making it easier to identify and exploit cointegration relationships. The data collected for each cryptocurrency for every minute from 2023-06-12 00:00 till 2024-06-12 23:59 includes:

- Close price
- Min price
- Max price

Testing for Cointegration

The first step is to test for a cointegrated relationship between the cryptos. We used the Johansen Test, developed by Johansen in 1991, to test for cointegration between time series. We store unique pairs with a P value of <0.05, indicating they are cointegrated.

Strategy Development

After storing the unique cointegrated pairs, we calculate the normalized spread, a measure of the difference between the prices of two cointegrated assets, adjusted by a factor (typically the cointegration coefficient). We calculate four different spreads using the min and max price to ensure no trading signal is missed.

Next, we calculate the Z-score, a statistical measure describing the position of a value relative to the mean of a group of values. The Z-score standardizes the spread and identifies overbought or oversold conditions. We calculate four different Z-scores using the min and max price to ensure no trading signal is missed.

The open and close positions strategy is based on previous studies (Mohandas, 2023; Jansen, 2018):

- Long the spread (long Crypto X and short Crypto Y) whenever the Z-score is below –2.
- Short the spread (short Crypto X and long Crypto Y) whenever the Z-score is above 2.
- Exit positions when the Z-score hits zero (mean).
- Stop-loss of the positions when the spread reaches a Z-score of ± 4.

For this assignment, we use three timeframes for different strategies:

1. 1 month
2. 6 months
3. 12 months

Backtesting

Backtesting was performed using the Backtrader package. Each timeframe was backtested independently, using the same trading strategy settings.

Results

In the results section, the outcomes of the backtesting will be evaluated using consistent metrics. The outcomes for each timeframe will be compared to determine which timeframe yields the best results.

Use in Practice

The strategies can be used in practice with real-time data, but this is out of scope for this assignment. The results and P values ($P<0.01$ / $P>0.01$ ~ $P<0.05$) can be used to optimize risk management. For example, a pair with a P value of 0.008 may have a higher stake than a pair with a P value of 0.029, optimizing risk management based on statistical significance.

Literature List

- Blake, N. S., & Fomby, T. B. (1997). Threshold cointegration. International Economic Review, 38(3), 627. https://doi.org/10.2307/2527284
- Blázquez, M. C., De La Cruz, C., & Román, C. P. (2018). Pairs trading techniques: An empirical contrast. European Research On Management And Business Economics, 24(3), 160–167. https://doi.org/10.1016/j.iedeen.2018.05.002
- Bookstaber, R. (2017). The End of Theory: Financial Crises, the Failure of Economics, and the Sweep of Human Interaction. https://muse.jhu.edu/book/64609/
- Clegg, M., & Krauß, C. (2017). Pairs trading with partial cointegration. Quantitative Finance, 18(1), 121–138. https://doi.org/10.1080/14697688.2017.1370122

- Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica, 55(2), 251. https://doi.org/10.2307/1913236
- Huang, C., Hsu, C., Chen, C., Chang, B. R., & Li, C. (2015). An Intelligent Model for Pairs Trading Using Genetic Algorithms. Computational Intelligence And Neuroscience, 2015, 1–10. https://doi.org/10.1155/2015/939606
- Huang, Z., & Martin, F. F. (2017). Optimal pairs trading strategies in a cointegration framework. RePEc: Research Papers in Economics. https://EconPapers.repec.org/RePEc:hal:wpaper:halshs-01566803
- Huck, N. (2010). Pairs trading and outranking: The multi-step-ahead forecasting case. European Journal Of Operational Research, 207(3), 1702–1716. https://doi.org/10.1016/j.ejor.2010.06.043
- Huck, N., & Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior? RePEc: Research Papers in Economics. https://EconPapers.repec.org/RePEc:hal:journl:hal-01508010
- Meuriot, V. (2016). The concept of cointegration: the decisive meeting between Hendry and Granger (1975). Cahiers D'économie Politique, n° 68(1), 91–118. https://doi.org/10.3917/cep.068.0091
- Mohandas, G. (2023). Cointegration-Based Pairs Trading Strategy in Commodity Markets: A Novel Approach to Statistical Arbitrage. International Journal Of Science And Research, 12(11), 1731–1735. https://doi.org/10.21275/sr231123060916
- Rad, H., Low, R. K. Y., & Faff, R. W. (2016). The profitability of pairs trading strategies: distance, cointegration and copula methods. Quantitative Finance, 16(10), 1541–1558. https://doi.org/10.1080/14697688.2016.1164337