Quantum Spread: A Cointegration-Based Pairs Trading Framework

Dhruvkumar R Pansuriya
Roll Number: 230101071
Department of Computer Science and Engineering

Pairs Trading and Cointegration Strategy

A Statistical Arbitrage Approach to Market-Neutral Trading

1 Project Overview and Methodology

1.1 Problem Statement and Objective

Pairs trading represents a sophisticated market-neutral trading strategy that enables traders to profit from various market conditions, including uptrends, downtrends, and sideways movements. Classified under statistical arbitrage, this strategy involves taking opposing long and short positions in two highly cointegrated assets. The primary objective of this project is to design and implement a robust low-frequency pairs trading strategy that leverages cointegration theory to exploit relative price movements between asset pairs.

The strategy's market-neutral nature provides significant advantages by reducing exposure to systematic market risk while capitalizing on the mean-reverting properties of cointegrated asset pairs. This approach allows for consistent profit generation regardless of overall market direction, making it particularly valuable in volatile market environments.

1.2 Pair Identification and Selection Framework

The foundation of successful pairs trading lies in the rigorous identification of cointegrated asset pairs. For this project, I employed the Engle-Granger two-step cointegration methodology, which is considered the gold standard for detecting long-term equilibrium relationships between financial time series.

Data Collection and Preprocessing: The analysis focused on 25 carefully selected IT sector stocks, chosen for their similar business models and market exposures. Historical adjusted closing price data was obtained using the Yahoo Finance API for the selection period spanning April 2019 to March 2022. This three-year window provided sufficient data points for robust statistical analysis while capturing various market cycles.

Cointegration Testing Process:

- Comprehensive pairwise cointegration tests were performed on all possible combinations of the 25 stocks, resulting in 300 unique pairs
- The Engle-Granger two-step method was applied to each pair, first testing for unit roots in individual series, then testing the residuals from the cointegrating regression
- Only pairs with cointegration test p-values less than 0.05 were retained, ensuring statistical significance at the 95% confidence level
- Final pair selection involved ranking all statistically significant pairs by ascending p-values and selecting the top 4-5 pairs with the strongest cointegration relationships

1.3 Trading Strategy Architecture

Spread Construction and Signal Generation: The core of the trading strategy revolves around monitoring the spread between cointegrated pairs. The spread is calculated as:

$$\text{spread}_t = P_t^A - \beta P_t^B \tag{1}$$

where β represents the hedge ratio estimated through ordinary least squares regression, ensuring the spread exhibits stationary properties.

To standardize signals across different pairs, the spread is converted to a Z-score:

$$Z_t = \frac{\text{spread}_t - \mu_t}{\sigma_t} \tag{2}$$

where μ_t and σ_t represent the rolling mean and standard deviation of the spread, computed over a 30-day lookback window to capture recent market dynamics.

Entry and Exit Logic: The trading signals are generated based on Z-score thresholds that have been optimized through historical analysis:

- **Long A, Short B Entry:** Triggered when $Z_t < -1$, indicating Stock A is undervalued relative to Stock B
- **Short A, Long B Entry:** Triggered when $Z_t > 1$, indicating Stock A is overvalued relative to Stock B
- **Exit Condition:** Positions are closed when $|Z_t| < 0.25$, capturing the mean reversion while avoiding whipsaw trades

2 Implementation and Risk Management

2.1 Portfolio Construction and Position Sizing

The strategy implementation began with an initial capital allocation of INR 1,00,000, strategically divided equally among the selected cointegrated pairs to ensure diversification. Position sizes are calculated dynamically based on available capital and current asset prices, with the number of shares determined by:

$$\text{Shares}_A = \frac{\text{Allocated Capital}}{P_t^A \times (1 + \beta)} \tag{3}$$

$$\text{Shares}_B = \beta \times \text{Shares}_A \tag{4}$$

This approach maintains the hedge ratio while ensuring market neutrality. The portfolio undergoes rebalancing on every signal event to preserve neutral exposure to market direction and optimize capital utilization.

2.2 Risk Management Framework

Comprehensive risk management measures were implemented to protect capital and ensure strategy robustness:

- **Stop-Loss Implementation:** A 5% stop-loss threshold is applied to each position to limit maximum loss per trade
- **Capital Allocation:** Equal capital distribution across pairs prevents over-concentration in any single relationship
- **Dynamic Hedging:** Position sizes are continuously adjusted to maintain market neutrality as prices evolve
- **Cointegration Monitoring:** Regular reassessment of pair relationships ensures continued statistical validity

2.3 Backtesting Framework and Performance Analysis

Extensive backtesting was conducted over the period from April 2022 to March 2025, providing a comprehensive three-year out-of-sample evaluation. The backtesting framework meticulously logged every trade with detailed records including entry dates, exit dates, profit and loss calculations, and cumulative portfolio impact.

Key Performance Metrics:

- **Cumulative Return:** Measures overall portfolio performance across the entire backtesting period
- **Sharpe Ratio:** Evaluates risk-adjusted returns using annualized return and volatility metrics
- **Maximum Drawdown:** Quantifies the largest peak-to-trough loss experienced during the strategy execution

- **Trade Frequency:** Monitors the total number of executed trades to assess strategy activity levels
- **Win Rate:** Calculates the percentage of profitable trades to evaluate strategy consistency

2.4 Visualization and Analysis Tools

The project incorporates comprehensive visualization components to facilitate strategy understanding and performance evaluation:

- **Equity Curve Analysis:** Detailed plotting of cumulative returns over time to visualize strategy performance evolution
- **Spread Evolution Tracking:** Time series plots showing spread behavior and mean reversion patterns
- **Z-Score Band Visualization:** Charts displaying Z-score movements with entry/exit signal annotations
- **Trade Analysis:** Detailed entry and exit point markers overlaid on price charts for trade validation

3 Conclusion and Future Enhancements

The developed pairs trading strategy demonstrates significant promise in generating consistent returns through systematic exploitation of cointegrated relationships. The rigorous pair selection process, combined with disciplined risk management and comprehensive backtesting, validates the effectiveness of this statistical arbitrage approach.

The strategy's market-neutral nature provides valuable diversification benefits while the mean-reverting properties of selected pairs offer reliable profit opportunities. Future enhancements may include machine learning-based pair selection, dynamic threshold optimization, and expansion to additional sectors for improved diversification.

Code Implementation: The complete implementation is available in the accompanying Jupyter notebook, featuring modular design for easy customization and extension. All visualizations and detailed trade logs are included for comprehensive strategy analysis.

References:

- QuantInsti: `https://blog.quantinsti.com/pairs-trading-basics/`
- Hudson Thames: `https://hudsonthames.org/an-introduction-to-cointegration/`
- Educational Video Lecture - `https://youtu.be/g-qvFjvyqcs?t=266`