Pairs Trading via Cointegration & Z-Score Signals

Author: HARSH BYJASH | Date: 2025-10-01

1. Objective

Demonstrate a statistical arbitrage approach by identifying a cointegrated pair, constructing a spread, and trading mean reversion using z-score thresholds.

2. Data

Synthetic cointegrated prices for two assets (X, Y) from 2020–2024. File: data/cointegrated_pair.csv

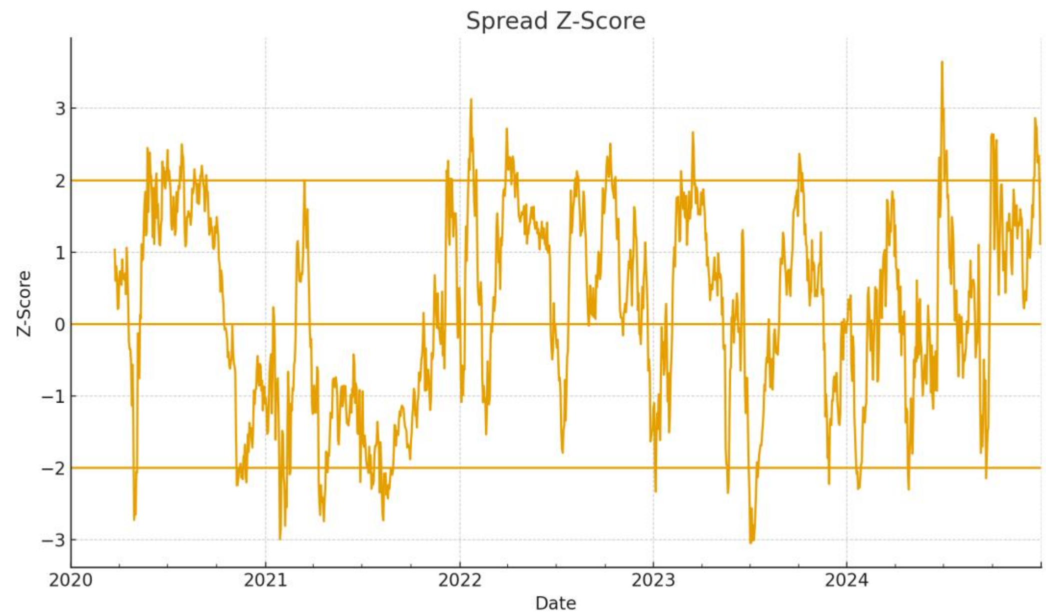
3. Method

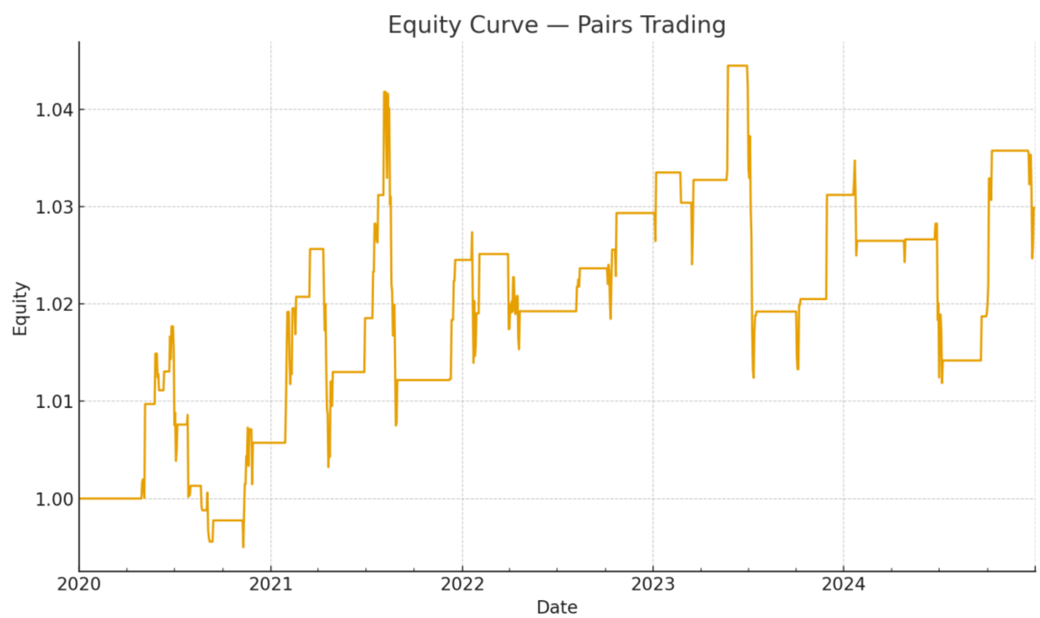
Estimated beta with OLS (Y ~ beta*X). Computed spread = Y - beta*X, standardized via a 60-day rolling z-score. Entered long spread when $z < -2$, short when $z > +2$; flattened when $|z| < 0.5$. Approximated PnL with returns of (Y - beta*X).

4. Results

CAGR	Vol	Sharpe	MaxDD
0.57%	2.51%	0.23	-3.29%

Figures:





5. Interpretation

The z-score entry/exit rules captured mean-reversion episodes in the synthetic pair. Performance depends on stability of cointegration and thresholds; turnover and execution costs will reduce results.

6. Limitations

Synthetic data and simplified execution. Real-world pairs require rolling beta, stationarity tests, and transaction cost modeling.

7. Next Steps

Test on real pairs (e.g., KO/PEP, XOM/CVX), add rolling regression, include hard stops, and run out-of-sample walk-forward validation.