

Reinforcement Learning Pair Trading - A Dynamic Scaling Approach

Pratyush Kalli (12459436)

Abhay Kanwar (12457992)

Manish Patturu (12454918)

Shayan Choudhury (12455354)

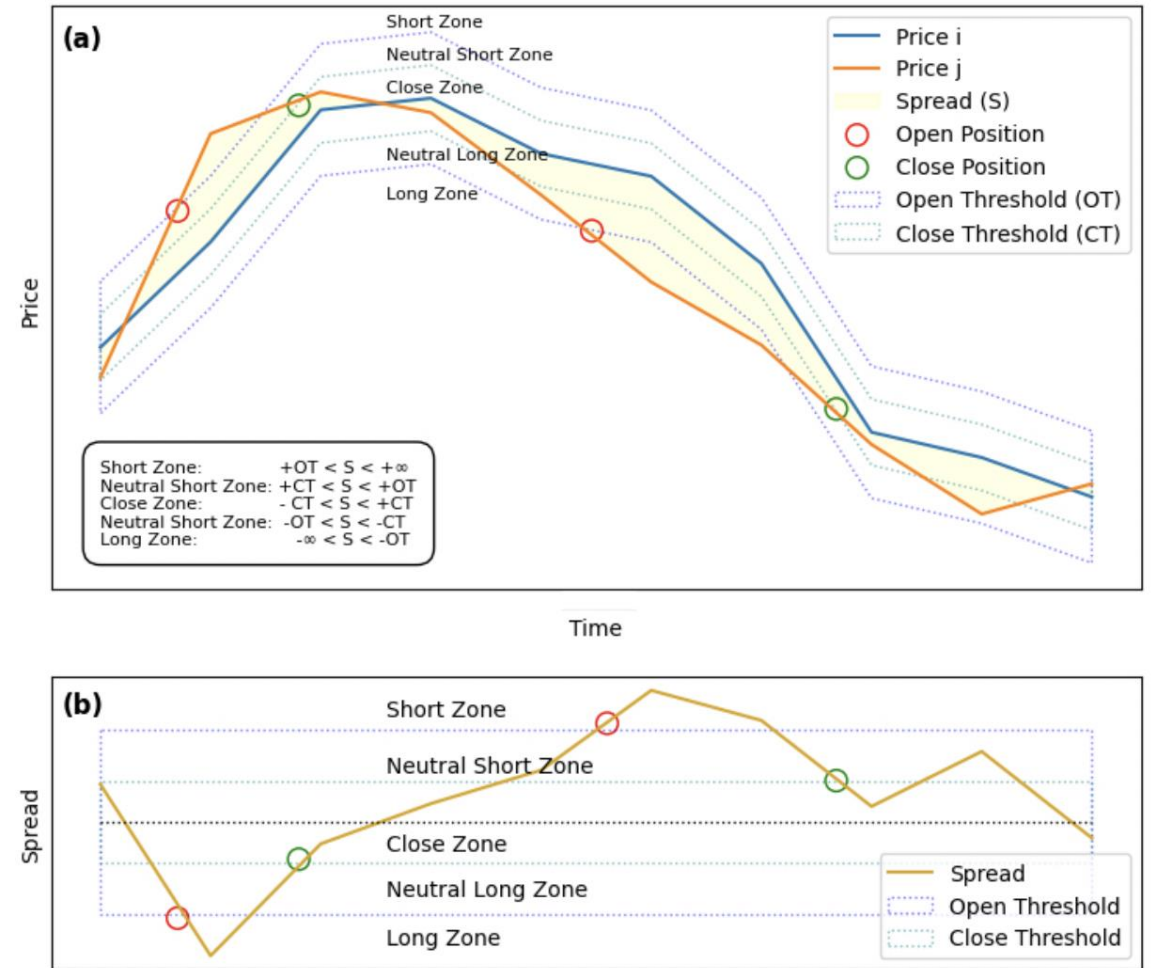
Satyam Saurabh (12452991)

Content

- ▶ Strategy Overview
 - ▶ Traditional Pair Trading
 - ▶ Reinforcement Learning Variations (Policy based, Actor-Critic)
 - ▶ Investment Securities
- ▶ Methodology
- ▶ “When” and “How much” to trade - Parameter Selection
- ▶ Evaluation Metrics

Traditional Pair Trading

- ▶ Pair trading involves two legs, where one asset is bought (long) while the other is sold (short) simultaneously.
- ▶ A long leg means buying the first asset and shorting the second, while a short leg reverses this.
- ▶ This strategy is market-neutral, as gains from one leg offset losses from the other, reducing overall market risk.



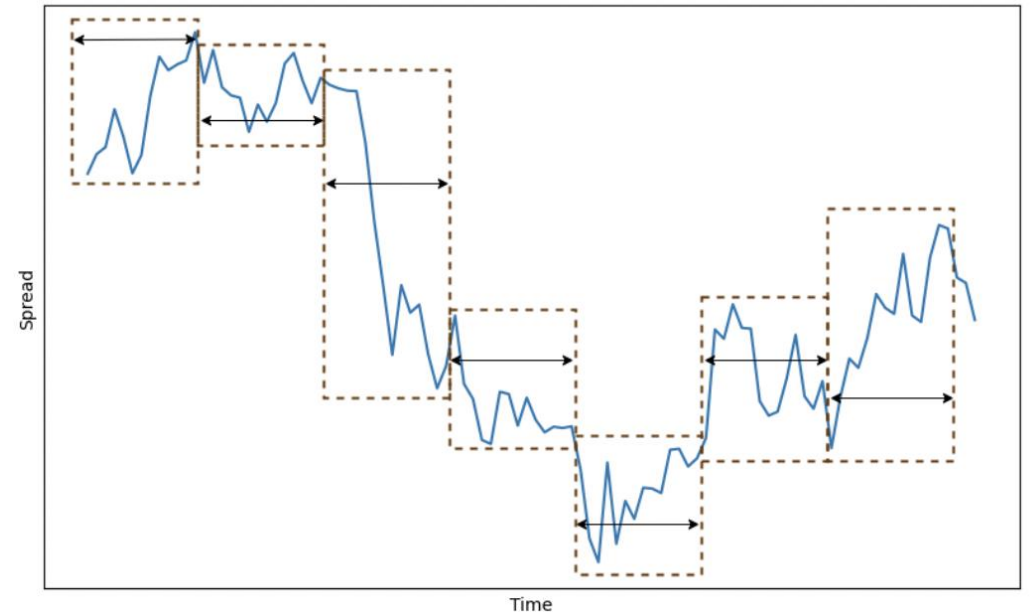
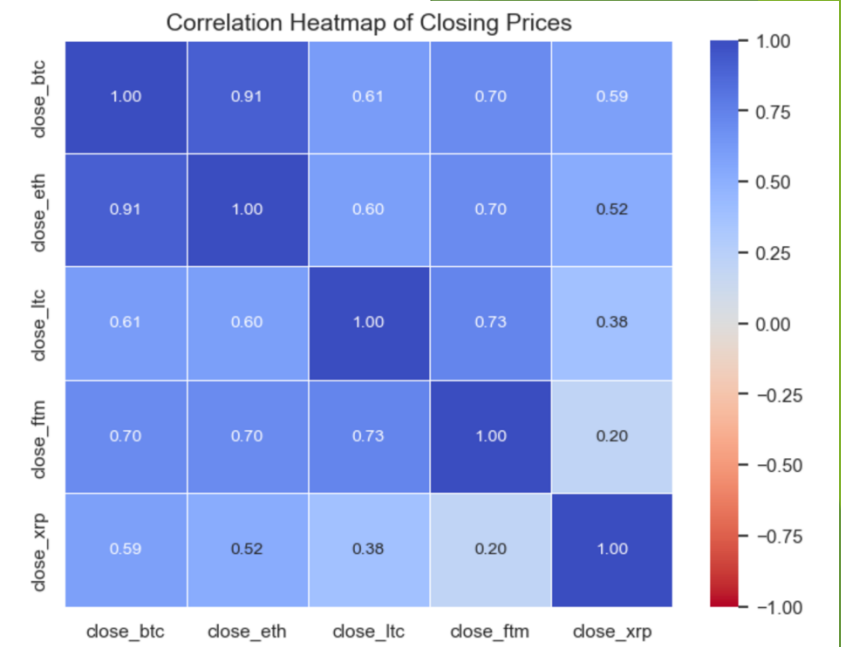
Stretched pair trading view of price distance between two assets

Strategy Overview

Features	Standard RL	RL_1 (Trade timing and direction)	RL_2 (Trade size and timing)
Description	Learns by trial and feedback, improving decisions over time	Optimizes when to trade and direction (long/short)	Optimizes trade size along with timing
Type	Value Based or Policy Based	Policy Based (Actor - Critic)	Policy Based (Actor-Critic)
Action Space	Discrete or Continuous	Discrete: Enter long-short, exit, enter short-long	Continuous: Adjust position size (-1 to 1)
Use Case	Applies in robotics, games and finance	Helps in trade entry, exit and direction decision	Dynamically adjusts position size for better risk control

Investment Securities

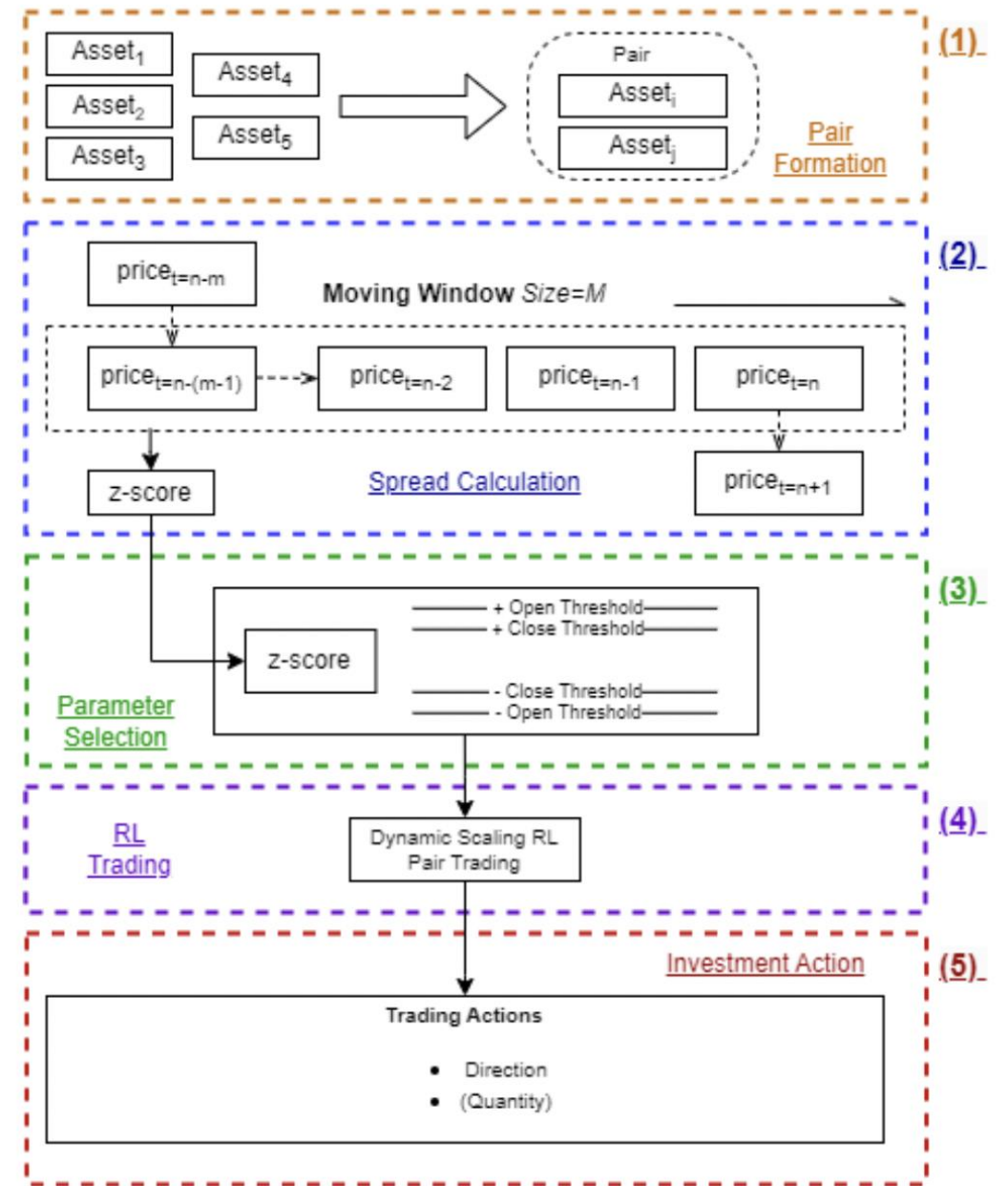
- Pairs are selected based on 2 criteria :
 - Correlation
 - Cointegration
- Moving window on historical prices ensures asset selection with strong long term statistical relationship
- High frequency and volatile assets like Cryptocurrency suit RL techniques
- BTC/USD, ETH/USD, LTC/USD, FTM/USD, XRP/USD



Window size cut for correlation and co-integration testing

Methodology

1. *Pair formation* for selecting assets to form a tradeable pair
2. *Spread calculation* utilizing the moving-window technique to extract the spread in a limited retrospective time frame
3. *Parameter selection* from an historical dataset to decide the most suitable hyperparameters
4. *RL trading* by allowing RL to decide the trading timing and quantity in pair trading
5. *Investment action* for taking the actions produced from RL trading into market execution.



“when” and “how much” to trade

Observation Space

- Information an RL agent observes
 1. Position: percentage measuring the direction of investment
 2. Spread: how much the spread has deviated from mean
 3. Zone: Comparison between z-score with threshold for signals

Action Space

- Freedom to choose an investment percentage
 1. Open Position: action of opening a position
 2. Close Position: closure of position
 3. Adjust Position: modifies an existing position by increasing the investment size

Reward Spacing

- Encourages profitable trades
 1. Portfolio Reward: profit/loss realized when closing a position
 2. Action Reward: Rewards agent for taking desired action
 3. Transaction Punishment: penalizes large position changes

Evaluation Metrics

Profitability Indicator	Description
Cumulative Return	Profit achieved during trading period
CAGR	Compound Annual Growth Rate
Sharpe Ratio	Risk-adjusted returns ratio

Activity Indicator	Description
Win/Loss Action Count	Number of winning/losing trades
Max Win/Loss Action	Maximum profit/loss per Action
Time in Market	Percentage of time invested in the market

Risk Indicator	Description
Volatility (ann.)	Annualized standard deviation of returns
Skew	Asymmetry of returns distribution
Kurtosis	“Tailedness” of returns distribution