Quantitative Pairs Trading Development Report (Day 36 – Day 40)

Executive Summary

This report documents **Days 36 through 40** of the *Quantitative Pairs Trading Development Project*, focusing on the statistical foundation, model enhancement, and practical implementation of a cointegration-based trading strategy.

The objective of this development phase was to progress from theoretical understanding to the construction of a **fully adaptive and risk-managed trading framework**. Over five consecutive days, the work evolved from foundational econometric testing (Engle–Granger and Johansen methodologies) to sophisticated model refinements involving **Vector Error Correction Models (VECM)**, **dynamic hedge ratio estimation**, and **realistic trade simulations** that account for market frictions.

Each stage systematically addressed a critical layer of the trading architecture:

- **Day 36–37:** Establishing robust statistical relationships and validating cointegration.
- **Day 38–39:** Implementing, testing, and refining the trading logic through rolling backtests and performance tracking.
- **Day 40:** Integrating advanced multivariate modeling and dynamic hedging to align with real-world trading dynamics.

Collectively, these developments represent a significant step toward a **scalable, data-driven pairs trading system** capable of adapting to changing market conditions while maintaining disciplined risk management and performance transparency.

Day 36: Cointegration Foundations via Engle–Granger

Day 36 established the foundational statistical requirement for pairs trading: identifying a long-term equilibrium relationship between two non-stationary price series.

Methodology: Engle–Granger (EG) Two-Step Cointegration Test

1. **Step 1: OLS Regression**

The regression

$$y_t = \alpha + \beta x_t + \epsilon_t$$

was performed to determine the linear relationship.

- The coefficient **β** serves as the hedge ratio.
- The residual **ϵ**□ defines the “spread.”

2. **Step 2: Stationarity Test**
   The residuals (**ϵ**□) were tested for stationarity using the **Augmented Dickey–Fuller (ADF)** test.
   - If residuals are stationary (I(0)), the assets are considered **cointegrated**, meaning the spread reverts to a stable mean.

Key Findings and Trading Relevance

- Successful identification of cointegrated pairs:
  - **ICICIBANK.NS – HDFCBANK.NS**
  - **HDFCBANK.NS – ^NSEI**
  - **HDFCBANK.NS – GLD**
- Indicates suitability for **mean-reversion strategies**.
- Highlighted the importance of calculating the **half-life of mean reversion** — a critical metric for position sizing and optimal holding periods.

Day 37: Advanced Cointegration Testing via Johansen Multivariate Framework

The limitations of the bivariate EG approach were addressed by implementing the **Johansen cointegration framework**.

Methodology and Advantages

1. **Lag Selection**
    - Automated optimal lag selection using **AIC/BIC** criteria on differenced price data.
    - Ensures appropriate model complexity.
2. **Multivariate Testing**
    - Uses **trace** and **maximum eigenvalue (max-eig)** statistics to determine the **cointegration rank (r)** — i.e., the number of cointegrating relationships.
3. **Output**
    - Extracts the **normalized cointegrating vector (β)**, defining the stationary spread's linear combination.

Actionable Insights

- Strong consensus pairs: **GLD–HDFCBANK.NS**, **GLD–RELIANCE.NS**.
- "Mixed evidence" pairs (where tests conflicted) were flagged for **visual inspection** and **robustness checks**.

Day 38: Classic Pairs Trading Framework Implementation

Building upon confirmed cointegration relationships, Day 38 implemented a **basic systematic trading logic**.

Implementation Steps

1. **Spread Construction**

$$S_t = \mathrm{StockA} - \beta \cdot \mathrm{StockB}$$

2. **Normalization**
    - The spread was standardized into a **Z-score**, enabling consistent entry/exit thresholds.
3. **Trading Rules**

- **Entry:**
  - Go **Short** the spread when $z > 2$
  - Go **Long** the spread when $z < -2$
- **Exit:**
  - Close the position when $z \to 0$ (mean reversion).

Framework Limitations

- Static hedge ratio (**β**) estimated over entire history.
- Arbitrary **±2σ** thresholds.
- No inclusion of **transaction costs** or **slippage**.

Day 39: Rolling Backtest and Trade Tracking Enhancement

To improve robustness, Day 39 introduced **adaptive methodologies** and **realistic trade management**.

1. Rolling Backtest

- The hedge ratio (**β**), spread mean, and standard deviation were recalculated using a **rolling window** (e.g., 252 days).
- Ensures adaptability to time-varying cointegration.

2. Enhanced Trade Management

- **Risk Limits:**
  - *max_holding_days*, *target_pnl*, and *stop_loss* to manage open positions.
- **Exit Reason Logging:**
  - Trades categorized by exit type:

        - **Z-Exit**, **Hit Target**, **Stop Loss**, **Time Limit**.

3. Performance Metrics

- Outputs included:
    - **Cumulative PnL**
    - **Drawdown analysis**
    - **Hit ratio**

Day 40: VAR/VECM Modeling and Dynamic Hedging

Day 40 integrated advanced **time-series modeling** and **realistic simulation features** to deepen analytical precision.

1. VECM/VAR Modeling

- The analytical function (**cointF()**) employed:
    - **VECM** if cointegration detected ($r \geq 1$).
    - **VAR** (on differenced data) otherwise.
- Captures long-run equilibrium corrections and enables **joint asset forecasting**.

2. Dynamic Beta Implementation

- **Rolling OLS:** Moving-window β estimation.
- **Kalman Filter:**
    - Recursive, time-varying β updates.
    - Ideal for intraday or regime-shifting markets.

3. Real-World Frictions

- **Transaction Costs:** Fixed *cost_per_trade*.
- **Slippage:** Percentage-based *slippage_pct*, proportional to spread magnitude.
- **Net PnL:** Computed after all frictions.

4. Drawdown Enforcement

- A **Drawdown Enforcement Threshold** (percentage of historical max drawdown) was introduced.
- Forces all positions flat and halts trading if breached, ensuring **capital preservation**.

Conclusion

The five-day development cycle from **Day 36 to Day 40** marked a decisive evolution in the quantitative pairs trading framework — transforming it from a theoretical construct into a robust, data-driven trading system.

The journey began with the establishment of **cointegration foundations** through the Engle–Granger methodology, ensuring that selected asset pairs shared stable long-term relationships suitable for mean reversion strategies. Building upon this, the **Johansen framework** expanded analytical depth by incorporating multivariate testing, enabling the identification of multiple cointegrating vectors and improving statistical reliability.

The project then transitioned from theory to practice with the **implementation of a systematic trading framework**, featuring Z-score normalization and rule-based entries and exits. Through the introduction of **rolling backtests** and **enhanced trade management**, the system evolved into a dynamic model capable of adapting to time-varying market structures while maintaining clear risk and performance tracking.

Finally, the integration of **VECM/VAR modeling** and **dynamic hedge ratio estimation** (via rolling OLS and Kalman filtering) positioned the model closer to real-world trading conditions. Incorporation of transaction costs, slippage, and drawdown enforcement further possibly strengthen its realism and capital protection mechanisms.

Overall, this development phase successfully established a **scalable, statistically grounded, and risk-aware framework** for pairs trading. The next phase will focus on **multi-pair portfolio optimization, live market calibration, and execution efficiency**, with the goal of transitioning from research to fully automated deployment.