ITESO, Universidad
Jesuita de Guadalajara

PROJECT 4

PAIRS TRADING WITH COINTEGRATION & KALMAN FILTERS

MICROSTRUCTURE AND TRADING SYSTEMS

LUIS FELIPE GÓMEZ ESTRADA

ARANTZA GOMEZ HARO GAMBOA AND JAVIER ALEJANDRO FAJARDO LÓPEZ

NOVEMBER 15, 2025

ÍNDICE

Introduction

This project develops a financial strategy based on pairs trading, which seeks to identify pairs of stocks that maintain a historical equilibrium relationship through cointegration testing. When the relative price between the two assets deviates from this equilibrium, a long position is taken in the undervalued asset and a short position in the overvalued asset, keeping the trading strategy market neutral.

To dynamically adjust the hedge ratio and generate more precise trading signals, Kalman filters are used to update the model parameters in real time as new market information becomes available. This approach allows the hedge ratio to evolve continuously rather than remain fixed, improving the strategy's ability to adapt to changing market conditions.

Additionally, the cointegration relationship between the two selected assets is confirmed using both the Engle–Granger method and Johansen's Cointegration Test, ensuring statistical robustness when identifying valid trading pairs. The spread between the assets is modeled and monitored so that deviations from its mean can be translated into entry and exit decisions.

Overall, this project combines statistical analysis with Kalman filtering to construct a market-neutral trading framework capable of responding effectively to continuous market changes and evolving price dynamics. The integration of dynamic hedge estimation and rigorous cointegration validation aims to improve the consistency and reliability of trading signals, providing a structured approach to identifying and exploiting temporary pricing inefficiencies within historically stable asset relationships.

Strategy Description

Pairs Trading Approach

Pairs trading is a market neutral strategy that focuses on the relationship between two different assets rather than on predicting the overall direction of the market, since this can be highly uncertain and difficult to forecast. Instead of relying on price trends or movements, the strategy identifies pairs of assets that historically move together in a stable long-term relationship. When the price relationship between the assets temporarily diverges, it is expected that the spread between them will revert to its historical mean. This concept, known as mean reversion, is the core mechanism that enables potential profit in pairs trading.

Pairs trading involves taking two positions simultaneously:

- A long position in the asset that appears undervalued relative to its historical relationship.
- A short position in the asset that appears overvalued.

As the spread converges back toward its equilibrium level, both positions are closed, capturing the profit generated by the price correction. Because the long and short positions hedge each other, the strategy remains market neutral, meaning that performance is less dependent on the overall direction of the market.

A key challenge in pairs trading is ensuring that the relationship between the assets is stable and not simply the result of temporary correlation. For this reason, cointegration testing is used to identify pairs with a statistically validated long-term equilibrium, which strengthens the reliability of the mean reversion signal and improves the robustness of the strategy.

Pairs trading offers several benefits, one of its main ones being market neutrality. This means that returns depend on the relative price movement between two assets instead of on broad market trends. This can provide certain protection and security during periods of high volatility. Additionally, because the strategy focuses on spread behavior rather than directional price movement, it can serve as a diversifying component within a broader portfolio. When implemented correctly with a stable long-term relationship between assets, pairs trading can generate consistent profit opportunities with comparatively lower risk than common directional strategies.

With this said, the effectiveness of this approach depends heavily on the validity of the statistical relationship being modeled. There is model risk, as the cointegration observed historically may weaken during the trading period. On the other hand, transaction costs and borrowing fees can also significantly reduce the profitability of the trading strategy. Lastly, there is also liquidity risk if one of the assets becomes difficult to trade at the expected price level, making it challenging to enter or exit positions correctly.

Cointegration

In pairs trading, it is not enough for two assets to simply have a high correlation, because this only reflects short-term strength and direction of a linear relationship and does not ensure that their prices will maintain a stable long-term relationship. Correlation can fluctuate during periods of market stress, leading to divergences that may not revert, resulting in potential losses.

On the other hand, cointegration indicates a long-term relationship between the prices. This demonstrates that, although the prices may drift apart in the short run, they are bound by a long-term equilibrium relationship. When the spread between their prices deviates from this equilibrium, it is expected to eventually revert to its historical average. This cointegration and mean reversion property is what enables the pairs trading strategy to function as an opportunity of statistical arbitrage, where profits are obtained from temporary price deviations between the two assets rather than by predicting the direction of the market.

For cointegration to be held, three conditions must be satisfied:

- Each asset price series must be individually non – stationary, meaning they may follow certain trends over time rather than fluctuating around a constant mean.
- The linear combination of the two price series, in this case the spread, must be stable over time. This means that, although the prices may drift apart in the short – term, they will eventually return to their historical average.
- Cointegration usually arises when the assets share some type of economic relationship, like belonging to the same sector. This ensures that the equilibrium relationship is economically meaningful.

With this said, to validate the existence of cointegration between the assets selected, two statistical tests were used:

- Engle – Granger Test: This test is used to estimate the relationship between the two price series and verifies whether the resulting spread is stationary.
- Johansen Test: This test allows the detection of one or more cointegration vectors between multiple time series.

By confirming cointegration between the assets, the strategy increases the likelihood that deviations from the historical mean will revert, strengthening the reliability of entry and exit trading signals.

Dynamic Hedging with Kalman Filters

In traditional pairs trading, the hedge ratio, which determines how much of one asset is bought relative to how much of the other is shorted, is estimated using a linear regression. However, this assumes that the relationship between the two assets remains constant over time. Market conditions change and evolve, which would mean that a constant hedge ratio is not effective and may lead to incorrect trade signals.

To prevent the use of a constant hedge ratio, this pairs trading strategy makes use of two Kalman Filters, which are the following:

- Kalman Filter 1: This filter dynamically estimates the hedge ratio between the two assets. By updating the ratio in real time as new market data arrives, it allows the strategy to adjust the size of the long and short positions continuously. This dynamic adjustment improves the accuracy of the hedging and ensures that the positions remain market neutral even when the relationship between the asset's shifts.
- Kalman Filter 2: The second filter is applied to estimate the parameters of the Vector Error Correction Model, also known as VECM. This captures long-term equilibrium relationship between the assets and generates the normalized spread used to produce trading signals. By updating these parameters in real time, the filter allows this trading strategy to react to evolving market conditions while maintaining the validity of mean reversion signals. In other words, this Kalman Filter ensures that the strategy correctly identifies when the spread between the assets has deviated from its equilibrium and when it is likely to revert, providing reliable entry and exit points for trades.

By using both Kalman Filters, the strategy achieves dynamic hedging. This dual approach allows the pairs trading strategy to remain responsive to changing market conditions, improving the accuracy and profitability of the overall model.

Expected Market Conditions for Strategy Success

For a pairs trading strategy to be successful, the market environment must support the stability and mean-reverting behavior of the spread between the two selected assets. While the strategy is market neutral and does not rely on predicting the overall market direction, its performance does improve when the following conditions are true:

- There must be a certain level of volatility, otherwise temporary divergences in prices will not exist or will be too small to generate profitable trading opportunities. However, volatility must not be excessively high, as extreme price movements can prevent mean reversion.

- The assets chosen must remain liquid, allowing long and short positions to be entered and exited efficiently. If this is not followed, transaction costs can reduce returns significantly, as trades would need to be executed at unfavorable prices due to the market conditions.
- The relationship between the assets must remain economically meaningful. For example, the assets chosen should be influenced by similar market fundamentals, whether they belong to the same sector or being substitutable goods. If these conditions change, such as business model shifts, the spread may fail to be mean reverting.
- Lastly, the general market conditions should not be dominated by prolonged trends affecting both assets in the same direction. During such periods, the spread may not revert as expected, as the market may push both prices in the same direction for extended periods. This increases the risk of persistent spread divergence and weakens the effectiveness of statistical arbitrage signals.

With all of this said, this pairs trading strategy succeeds in environments where spreads are mean reverting, liquidity is sufficient, and volatility provides trading opportunities without destabilizing the price relationship.

Pairs Selection Methodology

Correlation

As mentioned previously, correlation measures the short – term linear relationship between the return of two assets. While correlation is useful for identifying assets that move togther in short – term, it does not guarantee that their prices will maintain a stable long – term relationship. For this reason, correlation is used as an initial screening criteria rather than the sole basis for pair selection. With this said, it is important to note the following:

- A correlation close to 1 indicates that the two assets tend to move in the same direction.
- A correlation close to 0 suggests little to no linear relationship.
- A correlation close to -1 indicates that the assets move in opposite directions.

For this project, a correlation threshold of 0.6 was used to select pairs for further analysis, this means that the correlation must be greater than this value to pass the test. This means that only asset pairs with a historically strong positive relationship were considered. By applying this filter, we ensure that the candidate pairs show meaningful correlation in the short – term, which increases the likelihood that any deviations in their spread are profitable through mean reversion strategies.

Engle – Granger Method

Once the correlated pairs are identified, they are then subjected to more rigorous testing using cointegration methods to confirm the stability of their long-term relationship. With this said, once correlated pairs are identified, the Engle – Granger Method is applied to analyze the spread of the two assets. The spread, which is the linear combination of the two asset prices, is tested for stationarity by using the Augmented Dickey Fuller test, which is also known as the ADF test. The following steps were taken:

- A linear regression is run between the two asset prices, where one asset is treated as the dependent variable and the other as the independent variable.
- The residuals from this regression represent the spread between the two assets.
- The ADF test is applied to the residuals to check for stationarity. A low p-value, in this case below 0.05, indicates that the spread is stationary, confirming that deviations from the long-term relationship are temporary and will likely revert to the mean.

By doing this, it ensures that the selected pairs exhibit mean reverting behavior, which is fundamental for executing a pairs trading strategy effectively. The Engle – Granger Method provides a way to test if there is cointegration between two assets, meaning that their prices share a stable long – term relationship despite short – term fluctuations. However, this approach is limited to analyzing only two assets at a time and can detect only one cointegration relationship.

Johansen's Cointegration Method

As mentioned previously, while the Engle – Granger Method is effective for testing cointegration between two assets, it is limited when analyzing multiple time series or when there may be more than one cointegration relationship. The Johansen Method addresses these limitations by allowing the detection of multiple cointegration vectors among several assets simultaneously.

This method uses a Vector Autoregression framework (VAR), which models the joint dynamics of the asset prices while accounting for their past values. By analyzing this, the Johansen Method is able to identify linear combinations of asset prices that are stationary over time. These stationary combinations, known as cointegration vectors, indicate long – term equilibrium relationships between the assets.

This approach is particularly useful in pairs trading for constructing spreads that capture mean reverting behavior across more complex sets of assets and for generating robust trading signals based on long – term equilibrium deviations.

To implement the Johansen Method, the analysis involves estimating the number of cointegration relationships present among the selected assets. This is done by using two test statistics, which are trace statistic and the maximum eigenvalue statistic. The hypothesis used in these tests are the following:

- Trace Statistics:
  - H_0: $K = K_0$ (there are K_0 cointegrating vectors)
  - H_a: $K > K_0$ (there are more than K_0 cointegrating vectors)

- Maximum Eigenvalue Statistic:
  - H_0: $K = K_0$ (there are K_0 cointegrating vectors)
  - H_a: $K = K_0 + 1$ (there are at least $K_0 + 1$ cointegrating vectors)

These statistics help determine the number of stationary linear combinations of asset prices, which are the cointegration vectors, that indicate stable long – term relationships. Identifying these vectors is crucial for constructing spreads that show mean reverting behavior and for geanerating reliable trading signals in a pairs trading strategy. To determine whether the pair passes the Johansen test, the trace statistic is compared with a critical value at a given confidence level. If the trace statistic is higher than the critical value, the asset pair passes the test and is considered cointegrated, so it is labeled with a 1. On the other hand, if the trace statistic is lower than the critical value, the pair is not cointegrated and is labeled with a 0.

Statistical Evidence: Correlation & Cointegration

Sector	Tickers	Correlation	PValue_ADF	Johansen Pass
ENERGY	XOM - EOG	0.62	0.0008	1
ENERGY	XOM - CVX	0.72	0.045	0
CONSUMER DISCRETIONARY	NKE - LOW	0.60	0.0009	1
CONSUMER DISCRETIONARY	NKE - HD	0.61	0.07	0
FINANCIALS	SCHW - BK	0.87	0.001	1
FINANCIALS	BLK - MS	0.78	0.01	0
FOOD AND BEVERAGE	HSY - PEP	0.54	0.06	0
FOOD ANS BEVERAGE	MNST- PEP	0.44	0.14	0
FINANCIALS	MS - SCHW	0.84	0.01	1
REAL STATE	AMT - PLD	0.67	0.002	1

With the previous table, we can evaluate the statistical evidence regarding the presence of long – term equilibrium relationships and correlation among the various asset pairs selected. With this information, we are able to distinguish between asset pairs that simply move together in the short – term and those that demonstrate a stable, long – term relationship that is suitable for pairs trading.

The table summarizes, for different sectors, the ticker pairs analyzed along with their corresponding correlation coefficients and the results of the Engle–Granger and Johansen cointegration tests. For a pair to be suitable for this project, the correlation between them must be greater than 0.60, their p-value obtained from the Engle – Granger Method must be less than 0.05, and the result from the Johansen Method must be 1.

It can be observed that not all asset pairs exhibit a high correlation. For example, from the Consumer Staples sector, Monster Beverage Corp (MNST) and PepsiCo Inc (PEP) resulted in a correlation of 0.44, which does not surpass the established threshold. Therefore, this pair does not qualify for further analysis, regardless of their Engle – Granger and Johansen test results.

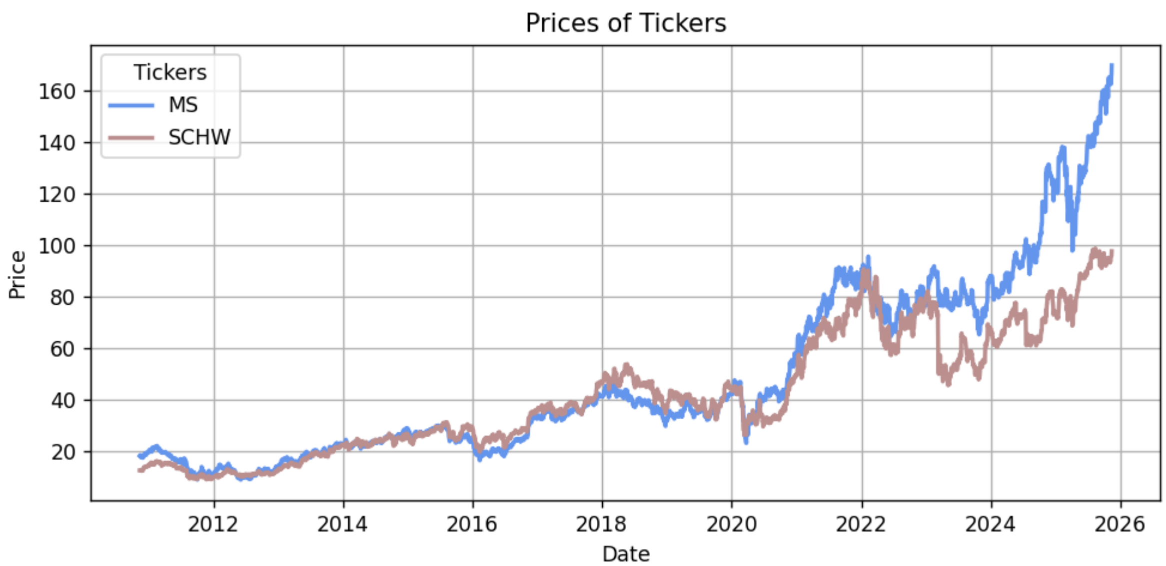
On the other hand, there are pairs that do display a strong correlation but do not show evidence of cointegration. This is the case of Nike Inc (NKE) and Home Depot Inc (HD), which have a correlation of 0.61. Although this suggests that they move together in the short – term, both the Engle – Granger and Johansen tests indicate the absence of a stable long – term equilibrium relationship. This implies that the spread between their prices does not revert to a mean over time, and therefore, this pair is not suitable for this type of trading strategy.

On the contrary, there are asset pairs that do satisfy all three criteria. This is the case of Morgan Stanley (MS) and Charles Schwab Corporation (SCHW). These assets have a correlation of 0.84, a P – Value of 0.01, and the result from the Johansen Method is 1. These results confirm both strong short – term correlation and the presence of a stable long – term equilibrium relationship. This makes the pair a strong candidate for the implementation of a pairs trading strategy.

Identifying such pairs is crucial, as the profitability of the strategy depends on price deviations reverting to their long–term equilibrium. Therefore, the pairs that meet all the selection criteria will be carried forward to the next stage of the analysis, where their spreads and trading performance will be evaluated.

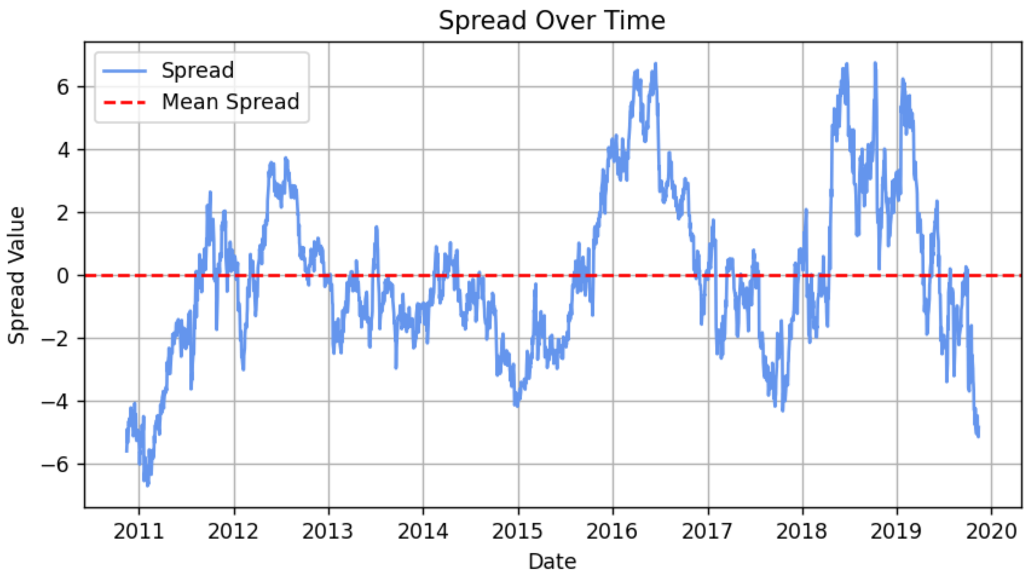
Price Relationships and Spread Evolution

Price Relationships



In the previous chart, the relationship between the prices of both assets, MS and SCHW, can be observed. From 2012 to 2023, their prices followed a very similar pattern, at times one was slightly higher than the other, but both moved in the same direction and on a comparable scale. After 2023, MS's price began to increase more rapidly than SCHW's, creating a gap of approximately $60 between them. Despite this divergence, which can be attributed to differences in company management and broader macroeconomic factors, both assets have continued to exhibit similar movements and overall trends, confirming the strong correlation that exists between the two.

Spread Evolution



This graph shows the spread between both assets in training data. The spread exhibits oscillatory movements around the red line, which represents its mean. These fluctuations, and the way the spread consistently reverts to the mean despite the assets' high volatility, indicate a long-term equilibrium relationship, which is a key condition for pairs trading. The positive and negative peaks represent moments when trading opportunities arise, since regardless of whether the spread is above or below its mean, it tends to revert to it. This mean-reverting behavior can be used to open and close positions profitably.

Selected Pair

After thoroughly analyzing multiple stock pairs, the combination of Morgan Stanley (MS) and Charles Schwab Corporation (SCHW) was selected as the optimal pair for this trading strategy. The selection process was based on three key points:

- Correlation between the assets
- P – Value obtained from the Engle – Granger test
- Results from the Johansen method

The results of all three criteria confirmed a strong long – term equilibrium relationship between both assets. The correlation coefficient between MS and SCHW indicated a high level of co – movement, suggesting that both assets respond similarly to market movements.

In terms of economic relation between these two companies, both Morgan Stanley and Charles Schwab Corporation operate within the financial services sectors, sharing the exposure to similar macroeconomic factors such as interest rate fluctuations, market volatility, and investment activity. Both companies generate income through wealth management and advisory services, meaning that their financial performance tends to move in relation to the overall health of the capital markets. On the other hand, both companies cater to very similar, if not the same client segments, such as individual and institutional investors. These great similarities in business models and customer bases contribute to these assets having similar reactions to changes in market conditions.

With this said, the Morgan Stanley and Charles Schwab Corporation pair was chosen not only due to its strong statistical cointegration and correlation, but also because of its clear economic relation, making it a robust candidate for this mean – reversion pairs strategy.

Sequential Decision Analysis Framework

Kalman Filter 1

1. State

$$S_t = (R_t, I_t, B_t)$$

where:

- $R_t: Physical\ Variables$
- $I_t: Exogenous\ Information$
- $B_t: Beliefs$

$$S_t = (w_t, P_t^x, P_t^Y)$$

where:

- $P_t^x: Current\ market\ price\ asset\ x$
- $P_t^Y: Current\ Price\ asset\ Y$
- $w_t: [w_0, w_1]\ Kalman\ filter\ state\ vector$

2. *Decision Variable*

$$X^\pi(S_t) = x_t$$

$$x_t = K_t(y_t - \hat{y}_t)$$

where:

- $K_t: Kalman\ gain\ matrix$
- $(y_t - \hat{y}_t): Prediction\ error$

3. *Exogenous Information*

$$W_{t+1} = \left(\hat{P}_{t+1}^x, \hat{P}_{t+1}^Y \right)$$

where:

- $W_{t+1}: New\ stock\ prices\ arriving\ between\ times\ t\ and\ t+1$

4. *Transition Function*

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

where:

- $S_{t+1}: State\ of\ the\ system\ at\ the\ time\ t+1$
- $S^M: Transition\ Function$

5. *Objective Function*

$$\min_{\pi} E\left\{\sum_{t=0}^{\pi} C_t\big(S_t, X^{\pi}(S_t)\big) \mid S_0\right\}$$

$$\min_{\pi} E[(y_t - \hat{y}_t)^2]$$

Six Step Modeling Process

1. *Narrative:*

To perform a pairs trading strategy, it is essential to determine the optimal hedge ratio between the assets each day. As new prices arrive, the Kalman Filter is used to update the relationship between both series in real time. This update incorporates incoming information while maintaining a measure of uncertainty through the covariance matrix.

At each step, the model adjusts the parameter that defines the relationship between the assets based on both the prediction and the new observed data. As new information becomes available, the Kalman Filter updates the state through a transition function. The objective is to minimize the prediction error and maintain a reliable and profitable estimate of the hedge ratio.

2. *Core Elements:*

- Metrics: Minimize the prediction error variance
- Decision: Update the parameter vector w_t
- Source of uncertainty: Random price fluctuation
- Exogenous Information: Arrival of new market prices $\left(\hat{P}_t^x, \hat{P}_t^Y\right)$

3. *Mathematical Model*

1) State

$$S_t = (w_t, P_t^x, P_t^Y)$$

2) Decision Variable

$$x_t = K_t(y_t - \hat{y}_t)$$

3) Exogenous Information

$$W_{t+1} = \left(\hat{P}_{t+1}^x , \hat{P}_{t+1}^Y \right)$$

4) Transition Function

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

5) Objective Function

$$\min_{\pi} E\left\{\sum_{t=0}^{\pi} C_t\big(S_t, X^\pi(S_t)\big) \mid S_0\right\}$$

$$\min_{\pi} E[(y_t - \hat{y}_t)^2]$$

4. Uncertainty Model

Since this process does not involve a random variable or any source of uncertainty, an uncertainty model does not apply.

5. Designing Policies

$$w_t = w_{t-1} + K_t(y_t - \hat{y}_t)$$

6. Evaluating Policies

Since there is no valuation component involved in this Kalman filter, an evaluation policy does not apply.

Kalman Filter 2

1. State

$$S_t = (R_t, I_t, B_t)$$

- $R_t: Physical\ Variables$
- $I_t: Exogenous\ Information$
- $B_t: Beliefs$

$$S_t = \left(w_t, (P_t^x, P_t^Y,)\right)$$

- $P_t^x: Current\ market\ price\ asset\ x$
- $P_t^Y: Current\ Price\ asset\ Y$
- $w_t: [eig_1\ \ eig_2]\ Kalman\ filter\ state\ vector$

2. Decision Variable

$$X^\pi(S_t) = x_t$$

$$x_t = K_t(y_t - \hat{y}_t)$$

- $K_t: Kalman\ gain\ matrix$
- $(y_t - \hat{y}_t): Prediction\ error$

3. *Exogenous Information*

$$W_{t+1} = (\hat{P}^{Y}_{t+1}, \quad \hat{P}^{x}_{t+1}, \hat{w}_t)$$

where:

- $W_{t+1}: New\ stock\ prices\ and\ eigenvector\ values\ between\ t\ and\ t+1$

4. *Transition Function*

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

where:

- $S_{t+1}: State\ of\ the\ system\ at\ the\ time\ t+1$
- $S^M: Transition\ Function$

5. *Objective Function*

$$\min_{\pi} E\left\{\sum_{t=0}^{\pi} C_t\big(S_t, X^{\pi}(S_t)\big) \mid S_0\right\}$$

$$\min_{\pi} E[(y_t - \hat{y}_t)^2]$$

Six Step Modeling Process

1. *Narrative:*

The model must estimate the relationship of long-term equilibrium between assets on a daily basis. Although this relationship can be obtained using a historical window through the Johansen test, the equilibrium changes with the evolving market conditions, making continuous updating necessary.

To achieve this, the model uses a second Kalman Filter that adjusts the eigenvector coefficients each time new prices and new eigenvector values become available. In this way, the model adapts day by day to changing market conditions and the daily noise that may arise.

2. *Core Elements:*

- Metrics: Minimize mean squared error between observed cointegration signal and its predicted value
- Decision: Update the parameter vector w_t
- Uncertainties: Random movements in prices and instability of the cointegration relationship
- Exogenous Information: New observed market prices $\left(\hat{P}_t^x\ \hat{P}_t^Y\right)$ and the eigenvector values $\widehat{w}_t$

3. *Mathematical Model*

1) State

$$S_t = (w_t, (P_t^x, P_t^Y,))$$

2) Decision Variable

$$x_t = K_t(y_t - \hat{y}_t)$$

3) Exogenous Information

$$W_{t+1} = (\hat{P}_{t+1}^x, \quad \hat{P}_{t+1}^Y, \widehat{w}_t)$$

4) Transition Function

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

5) Objective Function

$$\min_{\pi} E\left\{\sum_{t=0}^{\pi} C_t\big(S_t, X^{\pi}(S_t)\big) \mid S_0\right\}$$

$$\min_{\pi} E[(y_t - \hat{y}_t)^2]$$

4. *Uncertainty Model*

Since this process does not involve a random variable or any source of uncertainty, an uncertainty model does not apply

5. *Designing Policies*

$$w_t = w_{t-1} + K_t(y_t - \hat{y}_t)$$

6. *Evaluating Policies*

Since there is no valuation component involved in this Kalman Filter, an evaluation policy does not apply.

Kalman Gain

The Kalman Gain is a component through which the Kalman Filter determines how much the model should adjust based on the newly arriving information. At each moment, the filter combines the previous estimate from the model with the data coming from the market. If the incoming data is consistent, the filter assigns greater weight to it, but if the data contains too much noise, the filter relies more on the previous estimate. In other words, the Kalman Gain controls the model´s learning speed, adjusting according whether the data is stable or noisy. The formula for the Kalman Gain is the following:

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R)^{-1}$$

where:

- $K_t: Kalman\ gain\ at\ t$
- $P_{t|t-1}: The\ predicted\ covariance$
- $H_t^T: The\ observation\ Matrix$
- $R: The\ observation\ covanriance\ matrix$

In both filters, this dynamic allows the model to adjust its estimates to real market changes without overreacting to fluctuations caused solely by noise.

Q and R Matrix

In both filters, the same Q matrix was used, which is the following:

$$Q = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}$$

This matrix indicates that, for both Kalman Filters, the parameters are allowed to change over time, but not abruptly. It permits only moderate variation in the estimated parameters. Higher values in the Q matrix would generate excessive oscillations and lead to overfitting to daily noise. On the other hand, smaller values would make the model react too slowly to market changes. The chosen value of 0.01 represents a balance, as it allows the parameters to adapt continuously to new market conditions without compromising stability.

As with Q, the same R value was used for both filters, and it is the following:

$$R = [0.0001]$$

This assumes that the prices and the VECM are relatively reliable and do not contain excessive noise. This small value ensures that the Kalman Gain assigns more weight to the new incoming market data, allowing the estimates to react more sensitively to new observations. However, it is not so small that it would cause instability or abrupt changes. In other words, the chosen value of 0.0001 indicates that the model trusts the observed data while maintaining a level of smoothing to avoid overreacting and producing unnecessary variations.

Entire Model

1. State

$$S_t = (R_t, I_t, B_t)$$

- $R_t: Physical\ Variables$

- $I_t: Exogenous\ Information$
- $B_t: Beliefs$

$$S_t = \left(w_{1_t}, P_t^x, P_t^Y, v_t, \theta\ \right)$$

- $P_t^x: Current\ market\ price\ asset\ x$
- $P_t^Y: Current\ Price\ asset\ Y$
- $w_{1_t} = Hedge\ ratio\ obtained\ from\ Kalman\ 1\ at\ time\ t$
- $v_t = Vector\ Error\ Correction\ Model\ (VECM)\ at\ time\ t$

  $v_t = eig_{1_t} * P_t^Y + eig_{2_t} * P_t^X$
- $\theta = Theta\ threshold\ used\ to\ trade\ signals$

2. Decision Variable

$$X^\pi(S_t) = x_t$$

$$X^\pi(S_t) = \begin{cases} \varphi(VECM) > \ \theta, Open\ position \\ \varphi(VECM) < -\theta, Open\ position \\ |\varphi(VECM)| < \ 0.05, Close\ position \\ Otherwise \end{cases}$$

$$F^\pi(S_t) = C\left(S_{t,} X^\pi(S_t), \theta \mid S_0\right)$$

$$\theta = [0.3, 1.0]$$

3. Exogenous Information

$$W_{t+1} = \left(\hat{w}_{1_t}, \hat{P}_t^x, \hat{P}_t^Y, \hat{v}_t\right)$$

where:

- $W_{t+1}: New\ stock\ prices, hedge\ ratio\ and\ VECM\ between\ t\ and\ t+1$

4. Transition Function

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

where:

- $S_{t+1}: State\ of\ the\ system\ at\ the\ time\ t+1$
- $S^M: Transition\ Function$

5. *Objective Function*

$$\max_{\pi} E\left\{\sum_{t=0}^{\pi} C_t\big(S_t, X^{\pi}(S_t)\big) \mid S_0\right\}$$

$$\max_{\pi}\ (Portfolio\ Value\ (\pi))$$

Six Step Modeling Process

1. *Narrative:*

The complete pairs trading model must determine each day whether it is convenient to open, close, or hold a position between two cointegrated assets. To make this decision, the model dynamically updates both the hedge ratio and the cointegration vector using two different Kalman Filters, and it calculates the deviation from equilibrium using the normalized VECM.

If this deviation exceeds the threshold, the model opens a position in expectation of mean reversion, closing it later to take profit. Since prices and the relationship between assets change continuously under uncertainty, the model operates as a sequential decision process in which estimations and action evolve as new data arrives.

2. *Core Elements:*

- Metrics: Maxime portfolio value at the end of testing data
- Decision: Generation of trading signals based on standardization deviations of equilibrium spread
- Uncertainties: Estimation noise in Kalman filters weights and estimation noise in cointegration parameters

- Exogenous Information: New observed market prices $\left(\hat{P}_t^x\ \hat{P}_t^Y\right)$, hedge ratio $({w_1}_t)$ and VECM (v_t).

3. *Mathematical Model*

1) State

$$S_t = \left({w_1}_t, P_t^x, P_t^Y, v_t, \theta\right)$$

2) Decision Variable

$$X^\pi(S_t) = \begin{cases} \varphi(VECM) > \theta, Open\ position \\ \varphi(VECM) < -\theta, Open\ position \\ |\varphi(VECM)| < 0.05, Close\ position \\ Otherwise \end{cases}$$

3) Exogenous Information

$$W_{t+1} = \left(\hat{w}_{1_t}, \hat{P}_t^x, \hat{P}_t^Y, \hat{v}_t\right)$$

4) Transition Function

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

5) Objective Function

$$\max_\pi E\left\{\sum_{t=0}^{\pi} C_t\left(S_t, X^\pi(S_t)\right) \mid S_0\right\}$$

$$\max_\pi\ (Portfolio\ Value\ (\pi))$$

4. Uncertainty Model

Since this process does not involve a random variable or any source of uncertainty, an uncertainty model does not apply

5. Designing Policies

$$X^{\pi}(S_t) = \begin{cases} \varphi(VECM) > \theta, Open\ position \\ \varphi(VECM) < -\theta, Open\ position \\ |\varphi(VECM)| < 0.05, Close\ position \\ Otherwise \end{cases}$$

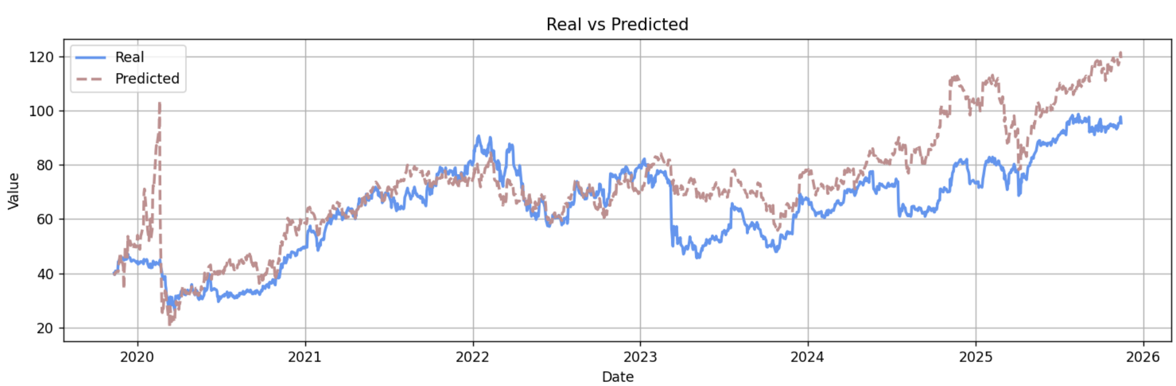
6. Evaluating Policies

$$Portfolio\ Value\ profit = V_{t-1} + \sum \left(\Delta P_t^{(i)} * Position_t^{(i)} \right) - Costs_t$$

Worked Example Showing State Evolution Over Several Periods

To show the state evolution over the test period, three charts were created which compare the price of the asset SCHW, the VECM, and the hedge ratio, each obtained from both Kalman Filters.

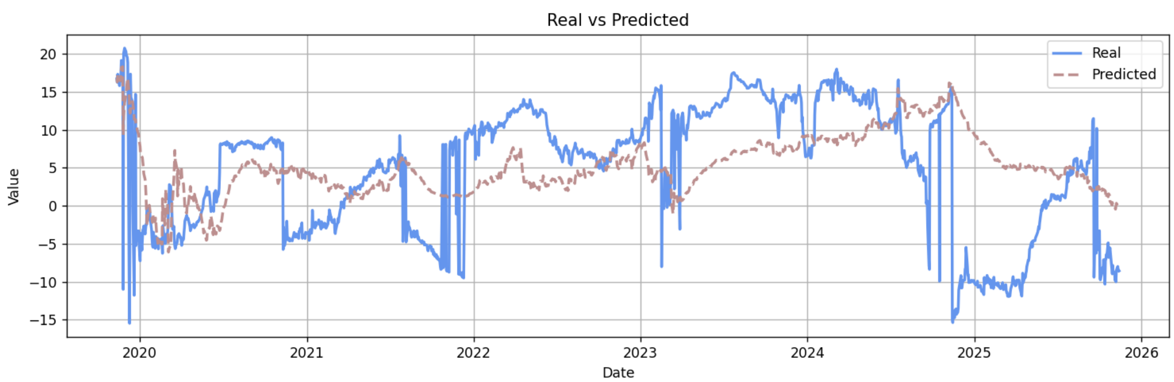
Second Price Real vs Estimated



The previous graph shows the real price of the asset SCHW compared with its own estimated price. To confirm that the trading model is working correctly, the estimated price does not need to be identical to the real price, but it should follow the same pattern. In this case, except for what happened in 2020 due to the market conditions caused by the COVID – 19 pandemic, both lines are quite similar, moving in the same direction and maintaining the same general trend.

The fact that both lines remain close to each other indicates that the intercept and hedge ratio are appropriately capturing the relationship between the assets. Additionally, because the lines do not diverge significantly, this suggests that the parameters have been correctly estimated and that there is limited noise affecting the model. Overall, this chart is a positive indication that the trading model can operate efficiently.

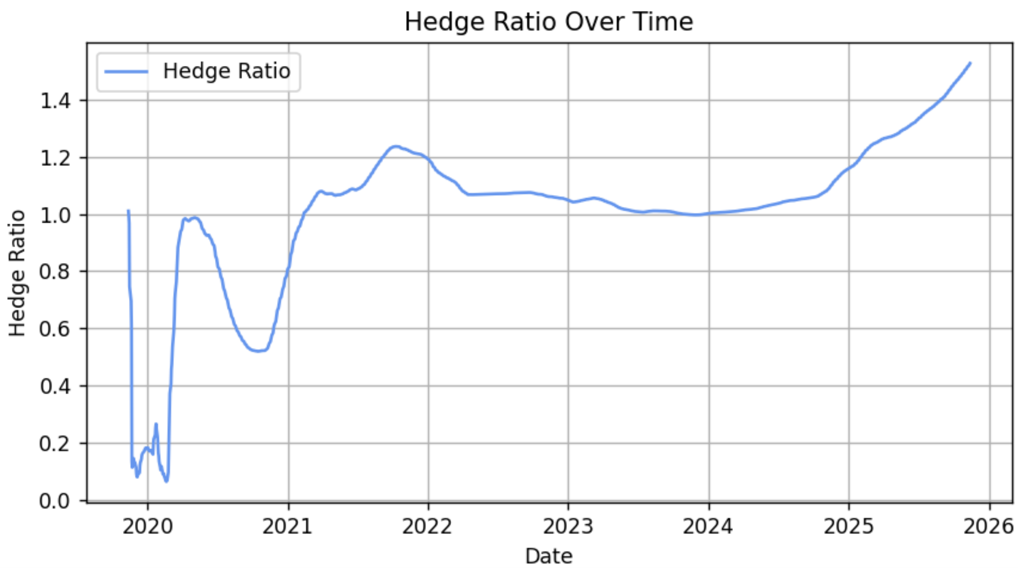
VECM real vs estimated



This chart shows the evolution over time of the real VECM, calculated from the Johansen test, compared with the estimated VECM obtained through the Kalman Filter. Like the chart of the real price versus the estimated price, when both lines remain close to each other, it indicates that the filter is correctly capturing the cointegration relationship. If the lines were to move in opposite directions, it would suggest a high level of noise, making it difficult for the model to achieve its objective.

In this case, both lines display a similar trend, and the estimated VECM closely follows the movements of the real VECM. This demonstrates that the filter is effectively reducing noise and generating a stable VECM, which strengthens the reliability of the trading signals.

Hedge Ratio



In the previous chart, it is possible to observe the evolution of the hedge ratio between the assets over time. In this case, the independent asset is SCHW and the dependent asset is MS. At the beginning of the test period, the hedge ratio was low, meaning that for each trade involving SCHW, the required number of MS shares was almost the same. As time progressed, the hedge ratio began to increase, and since 2023 it has remained stable within a range of approximately 1 to 1.4.

Analyzing this continuous updating of the hedge ratio is important because it helps identify the periods in which the hedge ratio increases and decreases and how this relates to the market conditions. This makes it easier to understand the trading signals and the reasons why they occur at specific points in time.

Kalman Filter Implementation

Initialization Procedures

For both Kalman filters, an initialization scheme based on standard practices for financial time series was used. First, an initial state vector was defined using values previously calculated from the training data. This allows the filter to adjust more quickly to the structure of the data. In addition, both the observation noise matrix and the process noise matrix were set based on an analysis of volatility and cointegration. This approach enables a progressive improvement in model fitting and, consequently, better signal prediction. As a result, it contributes to the development of a profitable portfolio while minimizing the risks associated with investing in financial assets.

Parameter Estimation Methodology

Both Kalman filters, the one used to update the hedge ratio and the one applied to update the VECM, which assists in generating trading signals, operate following the same standard Kalman filter procedure. This procedure consists of two sequential stages: prediction and updating, which can be mathematically expressed as follows:

First stage: Prediction

$$\widehat{w}_{t|t-1} = F\widehat{w}_{t|t-1}$$

$$P_{t|t-1} = FP_{t|t-1}F^T + Q$$

Second stage: Updating

$$K_t = P_{t|t-1}H^T(H_tP_{t|t-1}H^T + R)^{-1}$$

$$\widehat{w}_{t|t} = \widehat{w}_{t|t-1} + K_t(y_t - H_t\widehat{w}_{t|t-1})$$

$$P_{t|t} = (I - K_tH_t)\,P_{t|t-1}$$

Although the general structure is the same for both Kalman filters, the interpretation of the state variables, the incorporation of exogenous information, and the specification of process and observation noises vary depending on the specific objective of each model. This approach allows for a unified methodological framework that remains adaptable and effective across multiple scenarios.

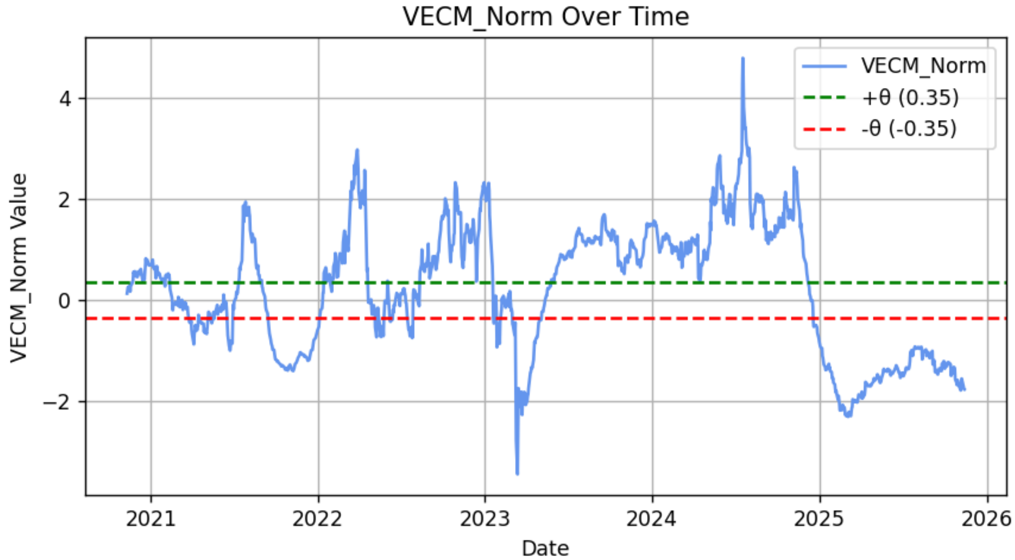
Reestimation Schedule and Validation Approach

Both filters are re-estimated continuously as new market observations arrive. This sequential updating allows the filter to keep their parameters aligned to the current market context and have a good adaptation to market movements. Additionally, the parameters are periodically recalibrated to prevent them from drifting over time and ensure better adjustment in the long-term equilibrium relationship.

For the validation process, two general perspectives were considered. First, the coherence of the signals was evaluated, in this case, whether the signals exhibited consistent behavior with the equilibrium assumptions and the objectives of the model. Second, the performance of the filters was assessed using several trade statistics, such as the number of trades, win rate, average win/loss ratio, and profit factor. These metrics provide evidence that the filters contribute to generating reliable and profitable trading signals.

Convergence Analysis and Filter Stability

The stability of both Kalman filters is evaluated by analyzing the behavior of their estimates and volatility over time. In both cases, the parameters converge progressively until they reach consistent statistical values, without exhibiting abrupt jumps or drastic changes. This behavior indicates that the filters operate within a stable range.



The evolution of the normalized VECM is shown in the previous graph. This series exhibits a clearly mean-reverting dynamic around zero, with oscillations that activate the entry and exit zones defined by the theta threshold. This behavior confirms that the filters do not diverge, instead, they incorporate new information appropriately and maintain a stable interaction between both Kalman filters. Such stability ensures that the model can generate economically viable and profitable trading signals

Trading Strategy Logic

Z – Score Definition using VECM

Once the cointegrated pair was selected, the Vector Error Correction Model (VECM) was estimated to obtain the spread's deviation from its long – term equilibrium relationship. The VECM residual captures the short – term imbalance between the two assets.

To be able to compare the deviations from the equilibrium consistently over time, the residual was standardized. This was done by calculating its mean and standard deviation over a rolling window of 252 days. The current residual was then transformed into a Z - Score, which expresses how many standard deviations it lies away from its recent average.

This normalization allows the spread to be evaluated on a common scale. A Z – Score close to zero indicates that the spread is aligned with its long – term equilibrium, while large values, whether they be positive or negative, suggest temporary mispricing. Positive Z – Scores indicate that the spread is wider than expected, while negative Z – Scores indicate that the spread is narrower. In other words, when the Z – Score reaches very high or low levels, it signals potential overbought or oversold conditions in the spread, which are used as entry signals for the trading strategy.

Optimal Entry and Exit Z – Score Policy

The policy for opening and closing positions is the following:

$$X^{\pi}(S_t) = \begin{cases} \varphi(VECM) > \theta, Open\ position \\ \varphi(VECM) < -\theta, Open\ position \\ |\varphi(VECM)| < 0.05, Close\ position \\ Otherwise \end{cases}$$

It was necessary to determine a value for theta that maximizes the portfolio's performance. To identify this value, several theta values within a defined range were tested to observe which one produced the highest profit. The range used in this test was $\theta = [0.3\ , 1.0]$. Each theta was evaluated using test data to assess its trading performance under different market conditions. For every tested theta, the resulting portfolio value was recorded and compared, allowing the selection of the optimal threshold.

THETA VALUES	PORTFOLIO VALUES
1	$1,181,895.31
0.9	$1,103,060.05
0.8	$1,079,905.70
0.7	$1,098,169.33
0.6	$1,187,562.42
0.5	$1,172,356.78
0.4	$1,200,515.22
0.3	$1,225,403.62

THETA VALUES	PORTFOLIO VALUES
0.39	$1,222,339.22
0.38	$1,212,593.73
0.37	$1,226,039.05
0.36	$1,227,969.13
0.35	$1,227,969.25
0.34	$1,233,218.61
0.33	$1,244,204.16
0.32	$1,244,203.81
0.31	$1,234,436.24

After testing each value from 0.3 to 1, it was observed that the best portfolio performance occurred at a value 0.3 and 0.4. To obtain a more precise estimate of the theta value that maximizes the portfolio value, additional tests were conducted within a narrower range of $\theta = [0.31, 0.39]$. After this second test, the results showed that the theta threshold that yields the highest portfolio value was 0.33. With this said, a threshold of 0.33 was selected for calculating the final portfolio value and all metrics.

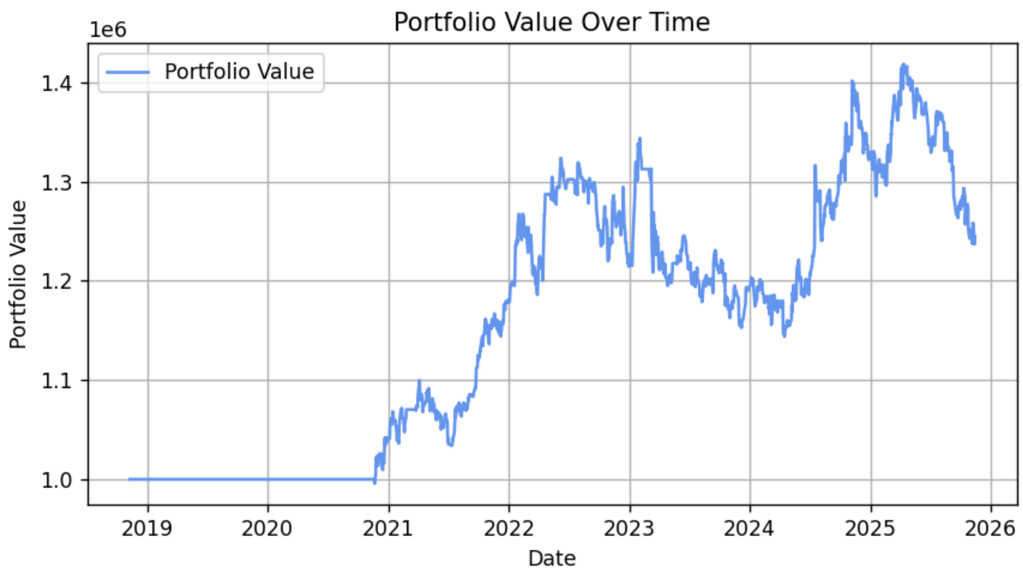
Commissions and Borrow Rates

Transaction costs were incorporated directly into the backtesting process to ensure realistic performance evaluation. Each time a position was opened or closed, a commission cost of 0.125% per trade was deducted from the current cash balance. On the other hand, short positions considered a daily borrow cost. Th borrow cost considered in this trading strategy was an annualized rate of 0.25%.

Including these costs is essential because pairs trading can consist of frequent trading activity, and even small fees may accumulate over time. By accounting for commissions and borrow rates directly in the model, this strategy avoids overestimating returns and ensures that the performance results are realistic and more comparable to real market outcomes.

Results and Performance Metrics

Equity Curve Plot



The previous graph shows the evolution of the portfolio value over time, from 2019 to 2026. The chart exhibits a clear upward trend from 2019 until 2023, indicating a period of strong and consistent growth. This positive trajectory was briefly interrupted by a temporary decline, which was soon followed by the portfolio's recovery. After 2024, the portfolio value resumed its upward trend. This pattern suggests that the trading strategy was able to generate consistent profits over time while effectively adapting to changing market conditions, demonstrating robustness through the implementation of the pairs trading strategy. When finalizing this period, the portfolio value was $1,244,203.58, with a return of 24.42%.

Performance Metrics

Sharpe Ratio

The Sharpe Ratio is a performance metric that compares the investment's return and profitability with its risk. In this case, a high Sharpe Ratio indicates that the returns are consistent when comparing it to the volatility or risk that is assumed. On the other hand, if the Sharpe Ratio is a low value, it indicates that the returns do not compensate for the risk that is taken when executing the trading strategy. For this reason, the objective is to obtain a positive Sharpe Ratio, as this would demonstrate that the strategy is balanced with returns that outweigh the risk level taken. With this said, the formula for this performance metric is the following:

$$Sharpe\ Ratio = \frac{R_p}{\sigma_p}$$

R_p: Average portfolio return

σ_p: Standard deviation of the portfolio returns (level of risk)

Sortino Ratio

Sortino Ratio is a performance metric that is very similar to Sharpe Ratio, but in this case, this metric only considers the downside volatility that is associated. This metric evaluated how well the strategy compensates for the negative returns instead of the general fluctuations of the asset. In this case, a higher value of this metric is better, and it indicates that the investment generates a good profit considering the downside risk. A negative value would mean that the strategy is not efficient since it does not effectively compensate for the risk taken. The formula for this metric is the following:

$$Sortino\ Ratio = \frac{R_p}{\sigma_d}$$

R_p: Average portfolio return

σ_d: Standard deviation of the negative portfolio returns (downside deviation)

Maximum Drawdown

The Maximum Drawdown represents the largest observed drop from a portfolio's peak value to its lowest point before a new high point is achieved. In other words, this metric measures the biggest possible loss that can occur during certain period. In this case, the objective is to obtain a lower value, since this would indicate that the strategy is more stable and exposed to less severe losses. A larger value would mean a higher potential risk. The ideal range for the result of this metric would be lower than 10%, but a value between 10% - 25% is still acceptable. The formula for this performance metric is the following:

$$Maximum\ Drawdown = max\left(\frac{Peak\ Value - \ Through\ Value}{Peak\ Value}\right) * 100$$

$Peak\ Value$: Highest portfolio value before the decline

$Through\ Value$: Lowest portfolio value during the decline

This performance metric is multiplied by 100 because this value should be a percentage.

Calmar Ratio

Calmar Ratio is a performance metric that compares the average annual return of the portfolio to its maximum drawdown, which was explained previously. In other words, this metric shows how effectively the trading strategy obtains consistent profits without having large drawdowns, reflecting better risk management. Unlike the Maximum Drawdown, in this case, a larger value above 1 is expected, since this would indicate a good balance between return and risk of drawdowns. The formula for this metric is the following:

$$Calmar\ Ratio = \frac{R_p}{MDD}$$

R_p: Average annual portfolio return

MDD: Maximum Drawdown

Metric Results

After calculating the performance metrics for the implemented pairs trading strategy, the following results were obtained:

METRICS	SHARPE RATIO	SORTINO RATIO	MAX DRAWADOWN	CALMAR RATIO
VALUES	0.40	0.68	15%	0.23

The Sharpe Ratio of 0.40 indicates that the strategy generates positive returns relative to the total risk taken, although at a moderate level. This means that while the strategy is profitable, its performance could be further optimized to improve the balance between risk and return. Nevertheless, maintaining a positive Sharpe value confirms that the trading strategy effectively produces returns.

The Sortino Ratio, with a value of 0.68, evaluates returns considering only downside volatility. The fact that the Sortino Ratio is higher than the Sharpe Ratio suggests that most of the observed fluctuations stemmed from positive returns rather than losses. This indicates efficient risk management, where downside movements are limited or infrequent.

The Maximum Drawdown of 15% shows that the portfolio experienced a moderate decline from its peak value. This percentage falls within an acceptable range, suggesting that the strategy is relatively stable and capable of recovering from temporary losses.

Lastly, the Calmar Ratio of 0.23 reflects the relationship between annualized returns and the maximum drawdown. This value shows that, although the strategy generates positive returns, these gains are relatively small compared to its largest loss. In other words, the recovery after major declines tends to be gradual.

Overall, these metrics indicate that the implemented strategy is stable and achieves controlled profitability, prioritizing risk management and long-term consistency. While there is room for improvement, the results confirm that the strategy is capable of maintaining profitability under changing market conditions.

Trade Statistics

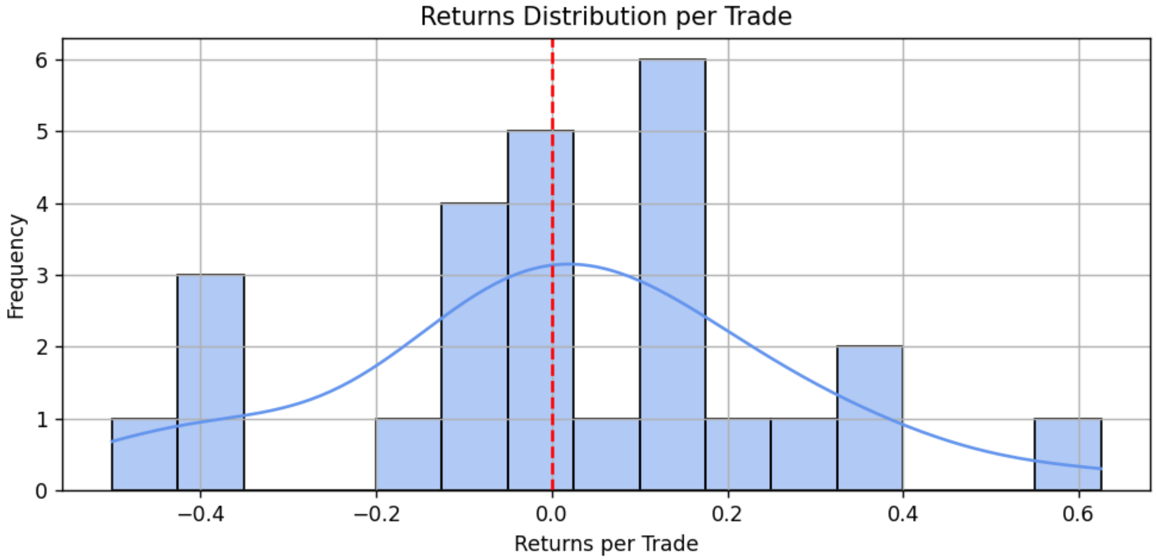
STATISTICS	VALUES
TOTAL TRADES	26
WIN	13
LOSSES	13
WIN RATE	50.00%
WIN/LOSS RATIO	1.2761
PROFIT FACTOR	1.2761

A total of 26 trades were carried out during the strategy's testing period. Out of these, 13 were profitable and the remaining 13 resulted in losses, leading to a win rate of 50%. This balanced proportion between winning and losing trades indicates that the strategy maintains consistent performance without any strong bias toward either side.

On the other hand, the Win / Loss Ratio of approximately 1.27 indicates that, on average, the gains from the winning trades are about 27% higher than the losses from the losing trades. This means that even with a moderate win rate, the profitability per winning positions compensates for the losses effectively.

In terms of the profit factor, this was approximately 1.27. This further supports the previous conclusion. In other words, for every dollar that the strategy loses, the strategy earns $1.27 in profits. By obtaining a profit factor greater than 1, it is demonstrated that this trading strategy is profitable overall, although there is room to better these values.

Out of Sample Performance on Validation Period



The previous graph demonstrates the distribution of returns per trade during the out of sample validation period, or in other words, the frequency with which different levels of returns occurred across all executed trades. Most trades cluster around small positive returns that are close to zero, indicating that the strategy consistently captures modest profits from short – term price deviations. With this said, the distribution is skewed to the right, meaning that, while most trades obtain limited gains, a smaller number of trades achieve higher returns. The left tail is relatively short, suggesting that large losses are infrequent and effectively controlled during this phase. This contributes significantly to the overall profitability of this pairs trading strategy. On the other hand, the relatively short left tail indicates that large losses are not common. Overall, the out of sample represents how the strategy generates stable profits and capitalizes on occasional high return opportunities.

Cost Analysis

COSTS	VALUES
BORROW COST	$55.02
COMMISSION	$16,019.04

As mentioned previously, two main types of costs were consider, which were the borrow costs and commissions. The borrow cost, which represents the expense of short selling positions, resulted in $55.02. On the other hand, the commission costs represented a much larger portion of the total expenses, ammounting to $16,019.04. These costs were obtained from the excution of multiple trades throughout the duration of the strategy, reflecting the cumulative effect of trading activity. While the borrow cost was relatively insignificant, the commision expenses had a more noticeable impact on the overall profitability.

This demonstrates the importance of managing transaction frecuency and optimizing trade execution to reduce unecessary costs. Even though the strategy proved profitable, minimizing commissions through efficent trade management could further enhance returns.

Conclusions

Key Findings and Strategy Viability

The results of the implemented pairs trading strategy demonstrate that the approach is viable and capable of generating stable and consistent profits under changing market conditions. The strong cointegration between Morgan Stanley and Charles Schwab Corporation, combined with a good correlation and a mean reverting behavior, allowed the model to produce reliable trading signals that effectively profit from temporary price divergences.

The equity curve confirms a persistent upward trend throughout most of the period, with the portfolio closing at a value of $1,244,203.58, representing a return of 24.42%. On the other hand, the performance metrics, like Sharpe and Sortino Ratio, indicate that the strategy maintains a good balance between return and risk. Also, the Maximum Drawdowns remains within an acceptable range. Additionally, the trade statistics reinforce these findings, showing a solid win rate, a positive Win / Loss Ratio, and a profit factor greater than one.

The out – of – sample validation further supports these conclusions. The distribution of returns per trade reveals right skewed profits, limited downside risk, and consistent capture of small gains. This demonstrates a strong capability of generalization.

Profitable After Costs

In terms of the trading costs, this pairs trading strategy remains profitable even after accounting for all expenses associated with the transactions. Both the borrow cost and the commissions were incorporated into the final portfolio valuation, ensuring that the performance reflects more realistic market conditions.

Despite these costs, the strategy maintained a strong upward trend, and the profit remained significantly positive. This indicates that the trading signals generated are efficient and overcome operational costs. The fact that the profitability consists after costs reinforces the robustness of the strategy.

Potential Improvements

Although the strategy delivered solid results, there are areas where improvements could further enhance performance and robustness. For example, implementing a frequent recalibration of the model’s theta parameter to help maintain certain accuracy as market conditions evolve. With this, the strategy would be able to adapt more effectively to shifts in volatility or changes in the spread, ensuring that the entry and exit signals remain aligned with the current market dynamics.

On the other hand, optimizing transaction costs can improve the performance of the strategy. Since commissions represent a significant part of the total expenses, reducing unnecessary trade frequency could increase net profitability.

Lastly, expanding the strategy to include multiple cointegrated pairs to have diversification across sectors could reduce dependency on a single spread. This would lead to smoother, better returns and improve the overall robustness of the investment approach.

Bibliography

- Chen, J. (n.d.). *Pairs trading strategy: Definition, benefits, and examples*. Investopedia. https://www.investopedia.com/terms/p/pairstrade.asp
- Chen, J. (n.d.-b). *What is mean reversion, and how do investors use it?*. Investopedia. https://www.investopedia.com/terms/m/meanreversion.asp
- *¿Qué es backtesting en trading? cómo hacerle backtesting a una estrategia de trading*. IG. (n.d.-b). https://www.ig.com/es/estrategias-de-trading/backtesting-como-testear-una-estrategia-de-trading-220720
- Engle Granger test - statistics how to. (n.d.-a). https://www.statisticshowto.com/engle-granger-test/
- *Cointegration*. Corporate Finance Institute. (2023, November 21). https://corporatefinanceinstitute.com/resources/data-science/cointegration/
- *What we do: Solutions and services*. Morgan Stanley. (n.d.). https://www.morganstanley.com/what-we-do
- Chapter T Utorial: The Kalman filter t on y lacey . . in tro duction the. (n.d.-a). https://web.mit.edu/kirtley/kirtley/binlustuff/literature/control/Kalman%20filter.pdf
- Schwab.com. (n.d.). *Charles Schwab: A modern approach to investing & retirement*. Schwab Brokerage. https://www.schwab.com/
- Sequential decision analytics - castle - princeton university. (n.d.-e). https://castle.princeton.edu/sda/
- Rankia. (s.f.). *Ratio Calmar* https://www.rankia.com/diccionario/fondos-inversion/ratio-calmar
- Finect. (s.f.). *¿Qué es el ratio de Sharpe?* https://www.finect.com/usuario/sandraalquezar/articulos/que-es-el-ratio-de-sharpe
- Brixx. (s.f.). *Definición del cálculo del índice de Sortino*. https://brixx.com/es/Definici%25C3%25B3n-del-c%25C3%25A1lculo-del-%25C3%25ADndice-de-Sortino/

- Wall Street Prep. (s.f.). *Maximum drawdown (MDD)*. https://www.wallstreetprep.com/knowledge/maximum-drawdown-mdd/