Dynamic Hedging with the Kalman Filter for Statistical Arbitrage

Mohammed El Azhar

July 15, 2025

Abstract

This project demonstrates the significant performance improvement achieved by using a dynamic hedging strategy over a conventional static approach in a statistical arbitrage framework. The core challenge was to hedge the market risk (beta) of a volatile underlying portfolio. We compared two methods: a static hedge ratio calculated with Ordinary Least Squares (OLS) and a dynamic hedge ratio continuously updated using a Kalman filter. The results are unequivocal: the static hedge failed to generate consistent returns and suffered significant losses. In stark contrast, the dynamic Kalman filter strategy produced consistent, positive monthly returns, resulting in a smooth equity curve and a total cumulative profit over the same period. This demonstrates that adapting to changing market relationships in real-time is crucial for the viability of this arbitrage strategy.

Contents

Chapter 1

Theoretical Framework: Finding Tradable Pairs

While momentum based strategies capitalize on the prices trends, the basis of this strategy is exploiting mean reversion, a phenomenon where a time series return to its historical average. While most individual financial assets are not mean-reverting and typically follow random walks, it is possible to construct a portfolio of assets whose combined value *is* mean-reverting [1]. This chapter outlines the methodology used to identify such portfolios.

1.1 Manufacturing Stationarity through Cointegration

The technique for creating a tradable, mean-reverting series is **cointegration**. Two or more non-stationary price series (e.g., $P_t^{(A)}$ and $P_t^{(B)}$) are cointegrated if a linear combination of them is stationary. This synthetic, stationary series is known as the **spread**, S_t, and is formed as follows:

$$S_t = P_t^{(A)} - \beta \cdot P_t^{(B)} \tag{1.1}$$

Here, β is the hedge ratio that makes the spread stationary. This spread becomes the primary instrument for our trading strategy.

1.2 Statistical Validation: The Augmented Dickey-Fuller (ADF) Test

We must statistically verify that a spread is stationary. The **Augmented Dickey-Fuller (ADF) test** is employed for this purpose. The test evaluates if the change in the spread's value, ΔS_t, depends on its prior value, S_{t-1}, by fitting the regression model:

$$\Delta S_t = \lambda S_{t-1} + \sum_{k=1}^{p} \phi_k \, \Delta S_{t-k} + \varepsilon_t \tag{1.2}$$

The null hypothesis is that $\boldsymbol{\lambda = 0}$, which signifies a non-stationary random walk. Our goal is to find pairs where we can **reject the null hypothesis** at a high confidence level (e.g., p-value $<$ 0.05). A statistically significant λ confirms mean reversion, making the pair a candidate for trading.

1.3 The Pair Selection Pipeline

Our practical procedure to identify tradable pairs involved the following steps:

1. **Correlation Screening:** First, we filtered a universe of assets to identify pairs with a high historical price correlation. This serves as an initial screen for assets that move together. While correlation and cointegration are ways to see similarities between random variables, they aren't necessarily related, we can have cointegrated pairs that aren't correlated, here we just add the correlation filter because we want to look at correlated pairs, it isn't mandatory.

2. **Cointegration Validation:** Next, for each correlated pair, we constructed the optimal spread and performed the ADF test.

3. **Final Selection:** Only pairs that passed the ADF test at a significant p-value were selected as the basis for the market-neutral strategies detailed in the subsequent chapters.

Chapter 2

Data Universe & Back-test Parameters

2.1 Asset Universe

We consider a broad, highly liquid U.S. equity universe of $N = 73$ tickers spanning large-cap technology, financials, health-care, industrials, and consumer staples (full list in Table 2.1). Symbols are the Yahoo/NASDAQ convention (e.g. `BRK-B` for Berkshire Hathaway class B).

AAPL, MSFT, AMZN, NVDA, GOOGL, GOOG, META, BRK-B, LLY, V, JPM, XOM, WMT, UNH, MA, PG, JNJ, HD, COST, MRK, ABBV, CVX, CRM, PEP, KO, ADBE, BAC, MCD, CSCO, TMO, ACN, AVGO, LIN, NEE, WFC, DIS, TXN, ORCL, CMCSA, PM, RTX, HON, AMGN, UNP, PFE, INTC, CAT, LOW, IBM, GE, GS, BA, DE, NOW, ELV, SPGI, AXP, ISRG, AMD, BKNG, BLK, SBUX, UPS, LMT, PLD, C, AMT, CI, GILD, TJX, PYPL, MO, T

Table 2.1: Universe of equities used for pair generation.

2.2 Training and Testing Windows

$$\text{Train: 1 Jan 2017} \le t \le \text{31 Dec 2021}, \qquad \text{Test: 1 Jan 2022} \le t \le \text{“today”}.$$

All pair selection (correlation screen and ADF cointegration test) is performed strictly on the *training* slice to avoid look-ahead bias. The hedge ratios are then applied out-of-sample to the test period.

2.3 Signal Parameters

- Entry Z-score threshold: $z_{\text{entry}} = \mathbf{1.5}$
- Exit Z-score threshold: $z_{\text{exit}} = \mathbf{0.5}$ (symmetric; absolute value used)
- Rolling window for mean and standard deviation: $w = 20$ trading days

These hyper-parameters were chosen heuristically; a grid-search over the training set could yield more reliable values. We explored with other values and we can easily double the performance by simply choosing a better pair of values.

Chapter 3

Hedging Methodologies

To neutralize market risk, we construct a hedge by taking an opposing position in a correlated asset. The effectiveness of the hedge depends entirely on the accuracy of the **hedge ratio (β)**. We tested two distinct approaches for calculating this ratio.

3.1 Approach 1: The Static Hedge (OLS)

A common baseline approach is to assume the relationship between the portfolio and the hedging instrument is stable over time. We calculate a single, fixed hedge ratio using an Ordinary Least Squares (OLS) regression on a historical training period. The linear model is:

$$P_t^{(\text{portfolio})} = \alpha + \beta \cdot P_t^{(\text{hedge})} + \epsilon_t \tag{3.1}$$

This fixed $\hat{\beta}$ is then used for the entire out-of-sample trading period. While simple to implement, this method is brittle and cannot adapt if the underlying market relationship changes.

3.2 Approach 2: The Dynamic Hedge (Kalman Filter)

Financial markets are non-stationary, and relationships between assets evolve. The Kalman filter is a powerful recursive algorithm ideal for this environment. It treats the hedge ratio (β_t) and intercept (α_t) as hidden states that evolve over time in a state-space model.

State Equation The hedge parameters follow a random walk, allowing them to adapt each day. Let $\theta_t = [\beta_t, \alpha_t]^\top$.

$$\theta_t = \theta_{t-1} + \omega_t, \quad \omega_t \sim \mathcal{N}(0, Q) \tag{3.2}$$

Measurement Equation The observed portfolio price is a function of the current state and the price of the hedging instrument, x_t. Let $y_t = P_t^{(\text{portfolio})}$.

$$y_t = [x_t \quad 1] \cdot \theta_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, R) \tag{3.3}$$

At each time step, the filter predicts the new hedge ratio and then updates its belief based on the actual observed portfolio price. This allows the hedge to adapt in real-time to changing market dynamics.

3.3 Results

The following figure displays the hedge ratio estimated from the linear regression model, who had only seen the training set samples, as well as by the Kalman filtering which has updated knowledge every day.

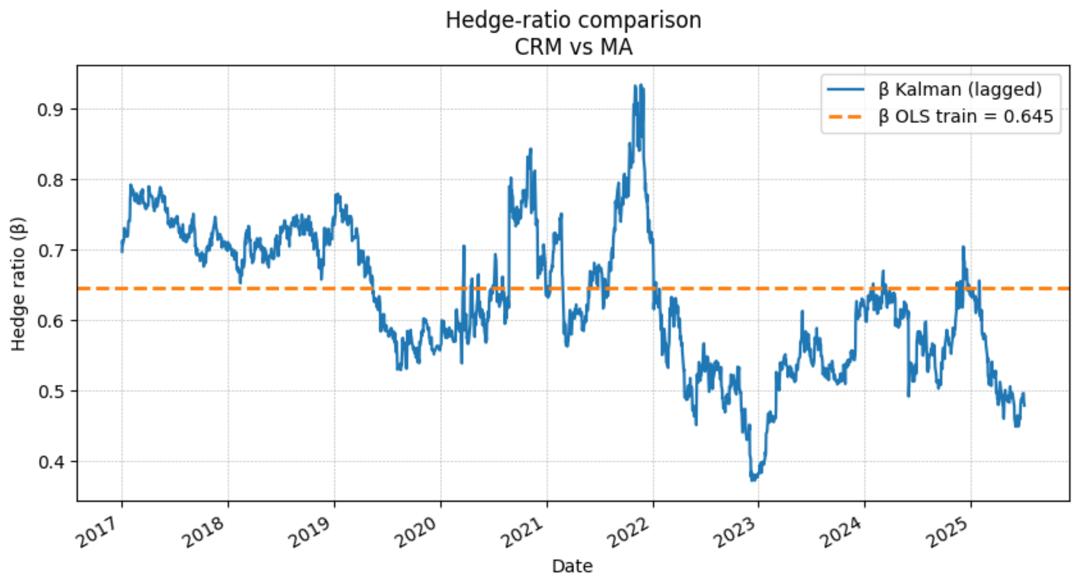

Figure 3.1: Hedge ratio estimated by both static and dynamic strategies. The dynamic strategy(blue) shows flexibility and richness that stem from the sequential knowledge update and the Kalman filtering model's complexity, while the static strategy (orange) shows naive prediction of a mean that is no longer relevant.

Chapter 4

Trading Rules & Signal Logic

At each close t we compute the *normalised spread*

$$z_t \;=\; \frac{y_t^{(1)} - \beta_{t|t-1}\, y_t^{(2)} - \alpha_{t|t-1}}{(1+\beta_{t|t-1})}$$

where $\beta_{t|t-1}, \alpha_{t|t-1}$ are the *predicted* Kalman-filter parameters [2]. A more sophisticated way is to use the Bollinger bands [3] but in our case we will remain extremely simple. The following table shows the trading signals we infer from the standardized spread. It goes without saying that when the standardized spread is between z_{entry} and z_{exit}, we hold our previous position until a new signal flag is updated.

State	**Condition at t**	**Action for bar $t \to t+1$**
Flat	$z_t < -\, z_{\text{entry}}$	**Enter long spread**: buy \$1 of leg 1, short $\beta_{t\|t-1}$ \$ of leg 2.
Flat	$z_t > +\, z_{\text{entry}}$	**Enter short spread**: short \$1 of leg 1, buy $\beta_{t\|t-1}$ \$ of leg 2.
Long	$z_t \geq -\, z_{\text{exit}}$	**Exit to flat**.
Short	$z_t \leq +\, z_{\text{exit}}$	**Exit to flat**.
Any	Stop-loss (optional)	If $\|z_t\|$ exceeds a pre-set maximum (e.g. 5σ), force an immediate exit next bar.

Table 4.1: Signal logic used in both static and dynamic hedge setups. All trades are submitted after the close and executed at the next session's open, ensuring no same-bar look-ahead.

The next figure illustrates a pairs trading strategy between the co-moving assets MA and V, where the **dynamic hedge ratio** is calculated via a Kalman filter. The core of the strategy is detailed in the bottom plot, which shows the mean-reverting **Z-score** of the asset spread. Trading signals are generated based on this score: a **short entry** is triggered when the Z-score is unusually high , while a **long entry** is triggered when it is unusually low . Positions are subsequently closed not on a full return to the mean, but on a partial reversion, a technique designed to efficiently capture profits. More profit could be extracted from the pairs trading strategy if we were to apply a dynamic threshold for the signals, but we won't delve in that. However that subject has been treated in the litterature [4].

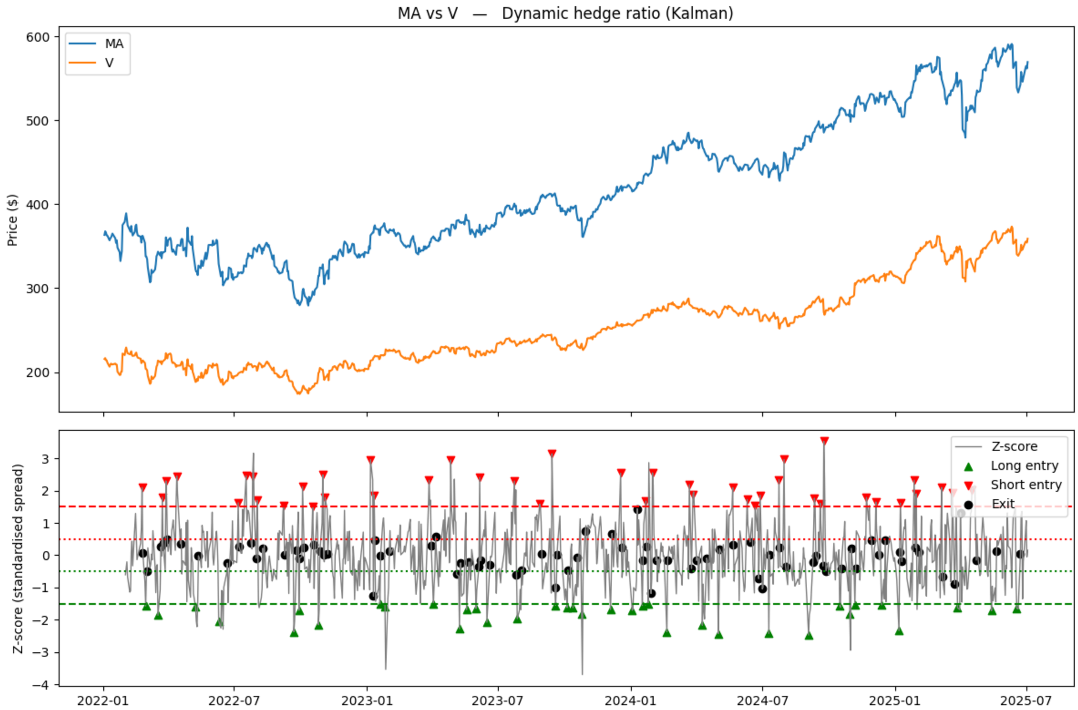

Figure 4.1: Standardized spread for the pair MA(Mastercard) and V (visa)

Daily P&L is booked as $\mathrm{P\&L}_t = p_{t-1}\big(s_t - s_{t-1}\big)$, where $p_{t-1} \in \{-1, 0, +1\}$ is the position held during the bar. We will see in the next chapter how position sizing can act as an efficient risk manager with pairs trading, as maybe it's not so great an idea to buy one full share of the first leg and β of the second, .

Chapter 5

Performance Analysis & Results

The out-of-sample performance reveals a dramatic difference between the two hedging strategies.

5.1 Equity Growth: A Tale of Two Hedges

The most telling result is the comparison of the cumulative P&L from each strategy. As seen in Figure 5.1, the static hedge's equity curve is volatile, trends sideways, and ultimately results in a net loss. It failed to effectively neutralize risk or generate profit.

Conversely, the dynamic hedge powered by the Kalman filter produced a remarkably smooth and consistent upward-trending equity curve. This visual evidence strongly supports the hypothesis that adapting the hedge ratio is critical for success.

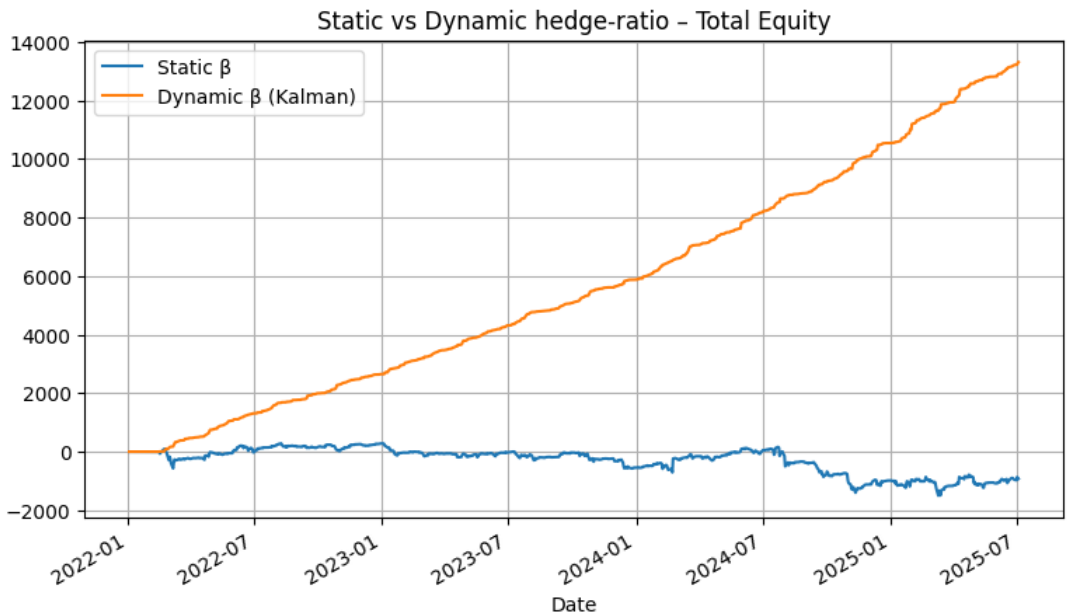

Figure 5.1: Cumulative P&L of Static vs. Dynamic Hedging. The dynamic strategy (orange) shows clear, consistent profitability, while the static strategy (blue) fails.

5.2 P&L Consistency: Monthly Performance Breakdown

A granular look at the monthly P&L heatmaps further explains the divergence in performance. The static hedge (Figure 5.2) was highly inconsistent, suffering from several periods of severe losses which erased any gains. In contrast, the dynamic hedge (Figure 5.3) was profitable in every single month of the backtest, demonstrating its robustness across different market conditions.

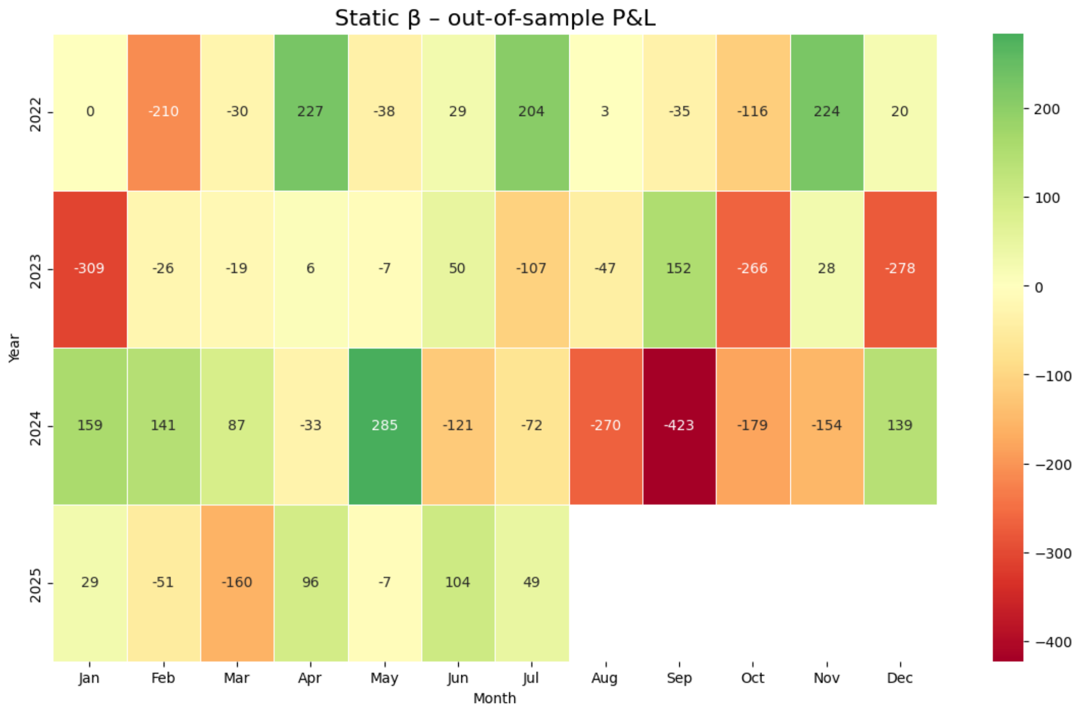

Figure 5.2: Monthly P&L from the Static Hedge. The strategy is characterized by inconsistency and large monthly drawdowns.

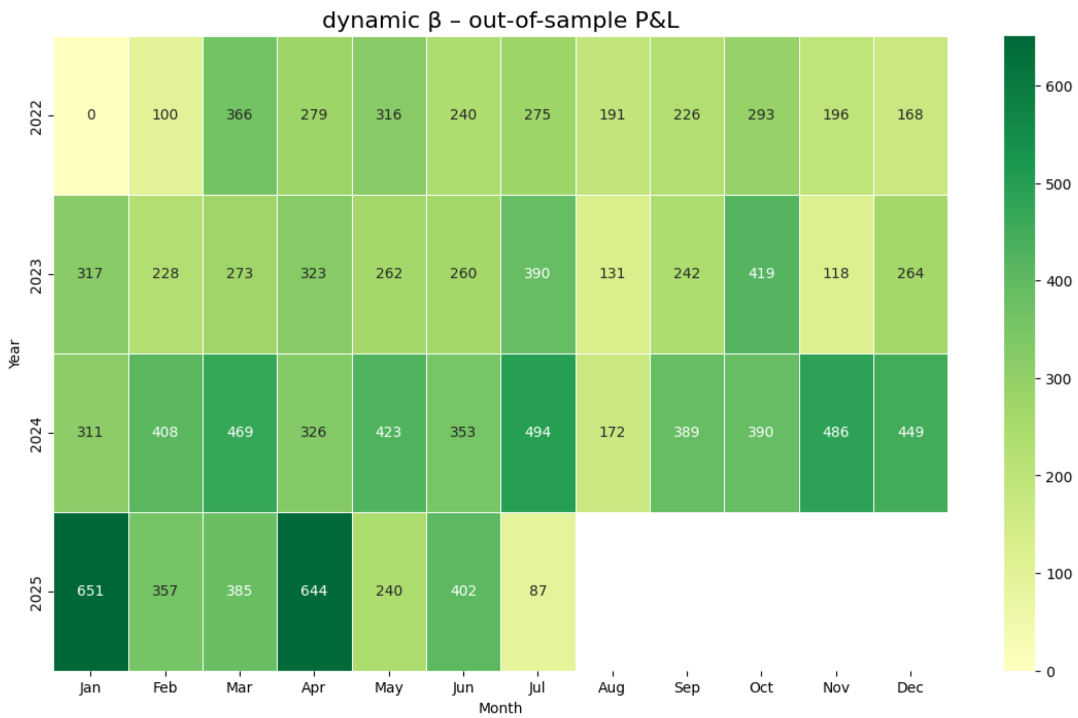

Figure 5.3: Monthly P&L from the Dynamic Hedge. The strategy delivers consistent and significant profits each month.

Chapter 6

Running a Budget-Constrained Portfolio

So far we reported raw dollar P&L generated by trading *one* share-pair per signal (Chapter 5). That view is useful for gauging *signal quality* but ignores the question every practitioner ultimately faces:

> **How much capital must I commit, and what return on that capital do I earn?**

6.1 Capital–based rescaling

To manage risk and prevent over-concentration of capital in any single trade, a dynamic position sizing mechanism is implemented. Instead of trading a fixed unit of one share of asset y against β_t shares of asset x, this mechanism calculates a scaling factor, S_t, to ensure the gross notional value of the position does not exceed a predefined capital limit. The final position taken is a fraction (or the full size) of this standard unit, determined by the available risk budget. The process is as follows:

1. Define the Per-Pair Capital Cap ($C_{\mathbf{cap}}$) This is a constant determined before the backtest begins, representing the maximum capital exposure allowed for any single pair.

$$C_{\text{cap}} = \frac{\text{Total Portfolio Budget}}{\text{Maximum Number of Pairs}} \tag{6.1}$$

2. Calculate the Gross Notional Value (N_t) For a standard 1-unit position initiated at time t, the gross notional value is the sum of the absolute market values of both legs of the pair. This is calculated using the prices at $t+1$ and the hedge ratio from time t (β_t) to estimate the next day's capital exposure.

$$N_t = |y_t^{(1)}| + |\beta_t \cdot y_t^{(2)}| \tag{6.2}$$

3. Calculate the Scaling Factor (S_t) The scaling factor S_t is calculated to reduce the position size only if the required notional value N_t exceeds the allowed cap C_{cap}. It is defined as the minimum of 1 and the ratio of the cap to the required capital.

$$S_t = \min\left(1, \frac{C_{\text{cap}}}{N_t}\right) \tag{6.3}$$

If the required capital is less than or equal to the cap ($N_t \leq C_{\text{cap}}$), then $S_t = 1$, and the position is not scaled down. Conversely, if the required capital exceeds the cap ($N_t > C_{\text{cap}}$), then $S_t < 1$, and the position is scaled down proportionally.

4. Determine the Final Position Size (Pos_t) The final, budget-constrained position size is the raw trading signal (Sig_t, which is either +1 for long or -1 for short) multiplied by the scaling factor S_t.

$$Pos_t = Sig_t \cdot S_t \tag{6.4}$$

This means that the actual number of shares traded is no longer a fixed unit, but a scaled amount. If the trading signal is to long the spread ($Sig_t = +1$), the strategy will buy $|Pos_t|$ shares of asset y and simultaneously sell $|Pos_t| \cdot \beta_t$ shares of asset x. This ensures the gross notional value of the position adheres to the predefined risk limit, C_{cap}.

6.2 Portfolio Profit and Loss Calculation

Once the scaled position size for each pair is determined, the next step is to calculate the daily profit and loss (P&L) and aggregate these results to construct the portfolio's overall equity curve.

Daily P&L for a Single Pair The daily P&L for a single pair is determined by the change in the spread's value from one day to the next, multiplied by the position taken on the previous day. A one-day lag is applied to the position to reflect the practical constraint of entering a trade based on the previous day's signal. The spread itself is defined as the deviation from the dynamically hedged relationship, as introduced in chapter 4.

$$\text{Spread}_t = \frac{y_t - (\beta_{t-1} x_t + \alpha_{t-1})}{1 + \beta_{t-1}} \tag{6.5}$$

The daily P&L for a single pair, i, is then calculated as the product of the lagged, scaled position and the daily change (difference) in its spread.

$$\text{P\&L}_{i,t} = Pos_{i,t-1} \cdot (\text{Spread}_{i,t} - \text{Spread}_{i,t-1}) \tag{6.6}$$

Total Portfolio P&L The total P&L for the entire portfolio at time t is the sum of the daily P&L values from all M pairs being traded on that day. This aggregation provides a single daily return figure for the overall strategy.

$$\text{P\&L}_{\text{Total},t} = \sum_{i=1}^{M} \text{P\&L}_{i,t} \tag{6.7}$$

Equity Curve Construction Finally, the portfolio's equity curve, which represents the cumulative growth of capital over time, is constructed. Starting with an initial capital base, C_0, the equity at time t is the equity from the previous day multiplied by one plus the daily portfolio return (the total P&L divided by the capital base).

$$\text{Equity}_t = \text{Equity}_{t-1} \cdot \left(1 + \frac{\text{P\&L}_{\text{Total},t}}{C_0}\right) \tag{6.8}$$

This can also be expressed in a cumulative product forms.

$$\text{Equity}_T = C_0 \cdot \prod_{t=1}^{T} \left(1 + \frac{\text{P\&L}_{\text{Total},t}}{C_0}\right) \tag{6.9}$$

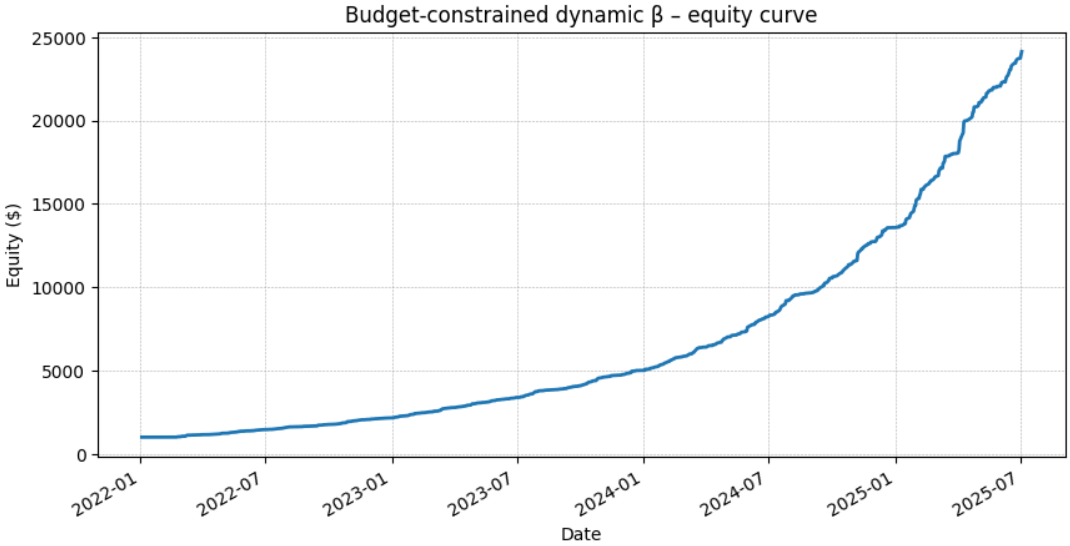

Figure 6.1: Equity curve with a capital constraint. The budgeted version starts at $1 000 and compounds to roughly $25 000

6.3 Portfolio Profit and Loss Calculation

Chapter 7

Conclusion

This project successfully demonstrates that for statistical arbitrage strategies dependent on hedging, **the method of calculating the hedge ratio is paramount**. A simple, static hedge based on a historical regression proved inadequate and unprofitable in a live trading environment. By employing a Kalman filter to dynamically update the hedge ratio on a daily basis, we were able to:

1. Effectively neutralize market risk.
2. Transform a volatile portfolio into a consistent source of alpha.
3. Generate substantial profits with a remarkably smooth equity curve.

The clear superiority of the dynamic approach validates the thesis that successfully navigating modern, non-stationary markets requires adaptive models that can respond to evolving relationships between assets.

Bibliography

[1] Chan, E. P. (2013). *Algorithmic Trading: Winning Strategies and Their Rationale*. John Wiley & Sons.

[2] Palomar, D. P. (2025). *Portfolio Optimization: Theory and Application, Chapter : Pairs Trading Portfolios*. Cambridge University Press.

[3] Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. John Wiley & Sons.

[4] Franck Martin, Zhe Huang. (2017). *Optimal pairs trading strategies in a cointegration framework.*