Polynomial Regression in Trading Systems

Lucas Brown

December 2025

1 Introduction

Polynomial regression offers a natural extension of linear modeling by allowing nonlinear trends to be approximated through higher-order terms while preserving mathematical interpretability. Unlike deep learning or machine learning models, polynomial regression provides explicit functional forms that enable statistical analysis and, more importantly, transparent decision-making for beginners learning algorithmic trading.

A recent paper by Gil Cohen, *Polynomial Moving Regression Band Stocks Trading System*, has demonstrated that polynomial moving regression bands, constructed by fitting polynomial curves to historical price data and defining upper and lower confidence bands using standard deviation thresholds, can be used to generate profitable momentum-based trading signals. Motivated by these findings, this project aims to design, implement, and analyze a polynomial regression–based modeling framework in Python.

2 Developing Regression Models

2.1 General Regression Models in Trading

Financial price series are inherently nonlinear, exhibiting behaviors such as curvature, acceleration, and transitions that cannot be adequately captured by linear trend models. While linear regression provides a useful baseline for modeling directional movement, it fails to represent the infinitely complex structural dynamics that characterize real-world markets. As a result, more flexible yet interpretable modeling frameworks are required for effective trend estimation in trading systems.

Polynomial regression does just this by allowing nonlinear trends to be approximated through higher-order terms while retaining a closed-form mathematical structure. Unlike machine learning approaches, polynomial models provide explicit functional representations of price behavior, enabling interpretability, statistical diagnostics, and controlled complexity. These properties make polynomial regression particularly well-suited for systematic trading research, where transparency and reproducibility are essential.

In these applications, polynomial regression is not applied globally across an entire price history but instead within a rolling-window framework. By recalibrating the model over a fixed-length window of recent observations, the estimated trend adapts continuously to local market conditions. This localized fitting allows the model to respond to evolving price dynamics while avoiding the instability associated with fitting high-order polynomials over long horizons.

Deviations between observed prices and the fitted polynomial trend provide a quantitative measure of short-term price dispersion. These deviations can be summarized using standard residual-based measures of variability and used to construct dynamic upper and lower bounds around the estimated trend. When interpreted as regression bands, these bounds serve as statistically meaningful thresholds that distinguish ordinary price fluctuations from significant momentum-driven movements.

2.2 Developing Functional Models

In developing a functional model for our purpose, the first variable to be determined is d, the degree of the polynomial. Too small of a degree will not showcase the predictive behavior on dynamics like curvature and accelerations, yet too high of a degree will lead to overfitting, where the model will not be able to reproduce the dynamics of the curves and accelerations. The results of Cohen's strategy recommend using a fourth-degree polynomial MRB (Moving Regression Band), which is what I deployed. The second factor which must be determined is the k value, the bandwidth multiplier. Assuming that the residuals are approximately symmetric, a small k (e.g. 0.5 - 1.0), leads to a very tight band behavior with frequent band crossings and high sensitivity to noise. This leads to high trading frequency and short holding periods. A large k results in rare band crossings and very few trades. In our scenario, the most beneficial bandwidth multiplier is a moderate value, 1.5 - 2.0. This leads to a balance between sensitive and robust behavior, and captures statistically meaningful deviations while filtering random noise.

2.3 Rolling Polynomial Regression for Trading Systems

The *rolling window* ensures that the model adapts to local market behavior, allowing the regression curve and its bands to evolve as new data becomes available. Let W denote the window length. Typical values range from 20 to 60 trading days, as shorter windows may produce unstable fits while much longer windows fail to reflect short-term market regimes. In this project, a 60-day window was selected to balance stability and reactivity.

For each day t, the regression is fit using the previous W observations. Let the corresponding time index be

$$x = 0, 1, 2, \ldots, W - 1,$$

and let y_i denote the closing price on day i. A polynomial regression of degree

d is expressed as

$$\hat{y}_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_d x_i^d.$$

The coefficients β_k are determined by ordinary least squares. Each coefficient governs a different component of the curve's structure: β_1 captures linear trend, β_2 introduces curvature, and higher-order terms allow inflection and more nuanced behavioral changes. The choice of d therefore controls the flexibility of the model. Guided by the findings of Cohen, a fourth-degree model was selected for its superior performance and balanced complexity.

After fitting the polynomial on the rolling window, the predicted value for the current day is given by

$$\hat{y}_t = f(x_t),$$

where $x_t = W$, representing the next point beyond the fitted window. The residuals over the window are

$$\varepsilon_i = y_i - \hat{y}_i,$$

and the residual dispersion is quantified via the root mean squared error:

$$\sigma_t = \sqrt{\frac{1}{W} \sum_{i=1}^{W} \varepsilon_i^2}.$$

This rolling standard deviation reflects short-term market volatility around the fitted trend.

Upper and lower polynomial regression bands are then constructed as

$$\text{Upper}_t = \hat{y}_t + k\sigma_t, \qquad \text{Lower}_t = \hat{y}_t - k\sigma_t,$$

where k is a positive constant controlling the width of the bands. Smaller values of k produce tight bands that trigger frequent signals but are highly sensitive to noise, while larger values produce wide bands that trigger only under exceptional price movement. Empirical and statistical considerations suggest a moderate choice of $k \in [1.5, 2.0]$; in this project, $k = 2$ was adopted consistent with conventional confidence-band theory and the methodology in Cohen's study.

These bands allow the construction of a momentum-based trading strategy. A long position is initiated when the closing price crosses above the upper band,

$$\text{price}_t > \text{Upper}_t,$$

indicating statistically significant upward momentum. The position is closed when the price falls below the lower band,

$$\text{price}_t < \text{Lower}_t,$$

signaling trend deterioration. Because both the regression curve and the volatility measure σ_t evolve dynamically, the resulting model adapts naturally to market conditions and provides a flexible framework for evaluating nonlinear trend-following behavior in financial time series.

3 Results

This section evaluates the performance of the polynomial regression band trading strategy relative to a passive buy-and-hold strategy.

3.1 Profitability and Comparison

To establish baseline effectiveness, we first examine whether the strategy generates positive returns and whether it outperforms a simple buy-and-hold investment.

A stock is considered profitable under the strategy if the final value of the portfolio value exceeds its initial value. Under this definition, 81 out of 86 stocks (94.2%) produced profitable trade sets, indicating that the strategy frequently captures positive price movements.

However, profitability alone does not imply superiority to passive investment. Compared directly with buy-and-hold outcomes, the strategy achieved a higher final portfolio value in only a small subset of stocks and exceeded the buy-and-hold Compound Annual GrowthRrate (CAGR) in 11 of 86 cases (12.8%). This disparity highlights that while the strategy is often profitable, it does not consistently outperform a long-term passive approach over the post-2010 sample period.

3.2 Compound Annual Growth Rate (CAGR)

CAGR measures the annualized rate at which capital grows over time and allows direct comparison across assets with different time horizons. It is defined as:

$$\mathrm{CAGR} = \frac{V_{\mathrm{final}}}{V_{\mathrm{initial}}}^{1/T} - 1$$

where V is the portfolio value and T is the number of years traded.

Across all stocks, the strategy achieves a mean CAGR of 12.06% and a median of 10.98%, compared to 21.58% and 19.51%, respectively, for buy-and-hold. The negative average CAGR difference of -9.53% reflects unfortunate underperformance.

While a small number of stocks, like ARM and META, exhibit high strategy CAGR values, these cases are exceptions rather than the norm, and rarely outperform the control. Overall, CAGR results indicate that the strategy sacrifices long-term growth in exchange for other desirable properties, such as risk control.

3.3 Maximum Drawdown

Maximum drawdown measures the largest peak-to-trough decline in portfolio value and captures downside risk in a way volatility metrics cannot. It reflects the worst loss an investor would have experienced before a recovery.

Formally, drawdown is computed as the percentage decline from the running maximum of the equity curve.

The strategy exhibits a mean maximum drawdown of -43.88%, compared to -53.08% for buy-and-hold. Importantly, drawdown is lower under the strategy in 73 out of 86 stocks (84.9%), with an average drawdown reduction of 9.21 percentage points.

This consistent reduction in downside risk suggests that the polynomial band exits frequently remove exposure during adverse price movements, even though doing so may limit participation in strong upward trends.

3.4 Risk-Adjusted Performance (Sharpe Ratio)

While drawdown captures extreme losses, the Sharpe ratio evaluates overall risk-adjusted performance by measuring excess return per unit of volatility:

$$\text{Sharpe} = \frac{\mu}{\sigma}$$

where μ is the mean daily return and σ is the standard deviation of daily returns.

The strategy achieves a mean Sharpe ratio of 0.58, compared to 0.71 for buy-and-hold. Although the strategy produces smoother equity curves in many cases, the reduction in average returns outweighs volatility improvements for most assets. As a result, the strategy only exceeds the buy-and-hold Sharpe ratio in 26 out of 86 stocks (30.2

These results indicate that while the strategy reduces drawdown, it does not consistently improve overall risk-adjusted efficiency as measured by Sharpe.

3.5 Trade-Level Win Rate

To better understand *how* the strategy generates returns, we examine the in-position daily win rate, defined as the percentage of days with positive strategy returns while a position is held. This metric reflects short-term directional accuracy rather than trade-level success.

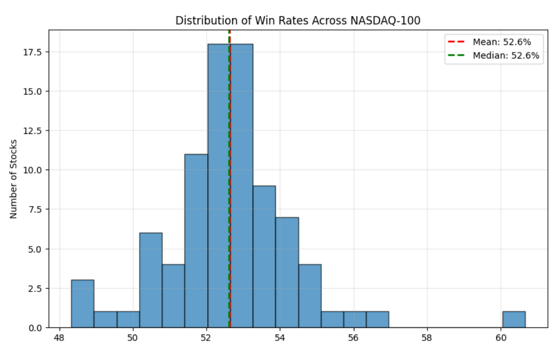

Figure 1: Distribution of Win Rates Across NASDAQ-100

Figure 1 shows the strategy exhibits a mean win rate of 52.62%, with low dispersion across assets. Over 94% of stocks maintain win rates above 50%, indicating that the strategy is directionally correct slightly more often than not.

However, win rates cluster tightly around the mean, and high win rates alone do not correspond to exceptional returns. This suggests that profitability depends more on the magnitude of gains and losses than on frequent winning.

3.6 Profit Factor and Trade Efficiency

The profit factor measures the ratio of total gains to total losses across all trades:

$$\text{Profit Factor} = \frac{\sum \text{Profits}}{\sum |\text{losses}|}$$

It is one of the most informative trade-level metrics, as it captures payoff asymmetry directly.

The strategy achieves a mean profit factor of 1.69, with 51.2% of stocks exceeding a profit factor of 1.5. Moreover, profit factor is strongly correlated with CAGR (correlation coefficient 0.818), indicating that long-term performance is driven primarily by payoff structure rather than win frequency.

Stocks with the highest profit factors also tend to exhibit superior risk-adjusted outcomes, reinforcing the importance of loss containment and selective participation.

3.7 Summary of Findings

Taken together, the results indicate that the polynomial regression band strategy functions primarily as a risk-management and mean-reversion overlay, rather than a return-maximizing strategy. While it generates broadly profitable trades and materially reduces drawdowns in most cases, it does not consistently outperform passive investment in a strongly trending market environment.

Its primary strengths lie in interpretability, controlled downside risk, and consistent trade execution, while its limitations stem from reduced exposure to prolonged price trends.

3.8 Parameter Sensitivity and Model Robustness

While the preceding analysis evaluates a fixed configuration of the polynomial regression band model, the performance of the strategy is dependent on three hyperparameters: polynomial degree d, rolling window length W, and bandwidth multiplier k. To assess the robustness and sensitivity of the band model, an exhaustive grid search (quite literally 24 hours of raw computation) was conducted across the following ranges:

$$d \in \{1, 2, \ldots, 9\}, \quad W \in \{20, 30, \ldots, 120\}, \quad k \in [0.5, 4.0] \text{ in increments of 0.1.}$$

This resulted in a total of 3,564 parameter combinations, each evaluated across the same 86-stock universe using average cumulative return as the objective metric.

The globally optimal configuration was achieved at degree $d = 8$, window length $W = 30$, and bandwidth $k = 0.9$, producing an average return of 831.82%. However, strong performance was observed across multiple polynomial degrees and window-bandwidth combinations. Table-level maxima for each degree are summarized below:

Degree	Window (days)	Bandwidth k	Avg. Return (%)
1	120	2.7	619.47
2	30	0.5	579.90
3	40	0.6	693.90
4	50	0.7	672.69
5	90	0.5	603.08
6	20	1.0	815.82
7	20	1.5	762.33
8	30	0.9	831.82
9	30	3.4	807.73

Across all parameter combinations, average returns exhibit substantial dispersion, with a mean of 382.69%, median of 366.25%, and standard deviation of 113.59%. This indicates that parameter selection meaningfully affects profitability and that naïve configurations can underperform even when the overall modeling framework is effective.

Figure 2 illustrates representative performance surfaces across bandwidth and window length for varying polynomial degrees. The surfaces reveal ridge-like regions of strong performance rather than isolated peaks, suggesting that the strategy exhibits moderate robustness within localized parameter neighborhoods but remains sensitive to large deviations from optimal regimes.

A focused heatmap for degree 4, corresponding to Cohen's recommended configuration, is shown in Figure 3. The presence of a broad plateau of elevated returns supports the use of degree-4 models as a stable compromise between flexibility and overfitting, even if higher degrees can achieve larger peak performance under carefully tuned conditions.

Overall, the parameter sweep demonstrates that while polynomial regression bands can yield strong returns under favorable configurations, performance is highly contingent on hyperparameter selection. This sensitivity highlights the importance of validation procedures and raises concerns regarding overfitting when optimizing purely on historical return metrics.

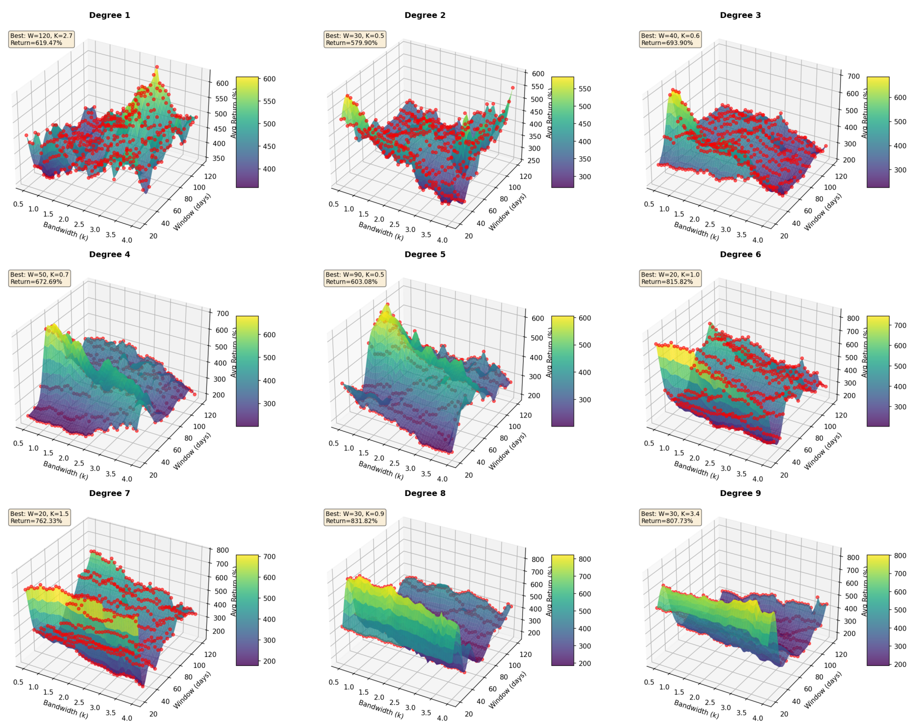

Figure 2: Average return surfaces across bandwidth and window length for polynomial degrees 1 through 9.

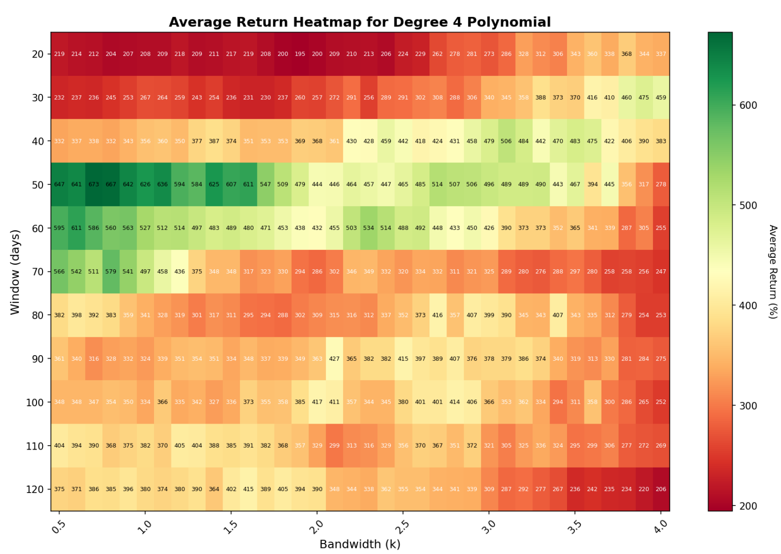

Figure 3: Average return heatmap for degree-4 polynomial regression bands across window length and bandwidth.