ORIE 5741 Project: Regime Detection

Yifu Wang, Zhe Zhou, Chengxi Shan

May 9, 2024

Abstract

This project focus on employ both unsupervised learning and supervised learning for detecting and predicting the market regime in the market. We first use K-Means clustering algorithm to detect the market regime hidden in the market where we use elbow method to determine the optimal number of cluster number. Next we construct macroeconomic data for predicting the next market regime. By integrating the two step model into portfolio optimizing for better parameter estimation, we demonstrate the power of incorporating predicting market regime in investment framework.

1 Introduction

It is well acknowledged that there has been regime shifting in the market, as people would often categorized as "bull market" and "bear market". In different market condition, the dynamic of the market asset can varies significantly. For instance, it is well establish that in different market condition the estimation of return and correlation of asset can varies. In a study by Cao, Sun, Ma, and Liu, they examine the concept of asymmetric correlations in financial markets. They explore how asset returns exhibit different behaviors under varying market conditions, emphasizing the impact on hedging and portfolio diversification [1].
However, there is no direct indicator out there to show what regime the market is current in, thus multiple method has been proposed for detecting the market cluster hidden in the plain sight. Academic has attempts to employs principal component analysis to extract features that capture significant variance in economic data, leading to a low-dimensional representation used for regime detection [2]. There is also study applies Hidden Markov Models (HMM) to identify market regimes in the US stock market and develops an investment strategy that adapts to these regimes [3].
In the project, we propose a two step approach of market regime detection, where we first determine the market cluster and then use predictive model to forecast the future market regime. Such approach is considered a common approach as been adopted by various study. For instance, Werge explores the use of Hidden Markov Models (HMMs) to predict market regimes across different asset classes. The study emphasizes the importance of selecting an optimal number of hidden states in the model to balance complexity and performance[4]. In a similar manner, we use K-Means clustering algorithm to find the market regime and use elbow method to find the optimal cluster number.
In order to demonstrate the significance of market regime detection and prediction, we incorporate the two step framework into the classic Markowitz portfolio optimization. As the asset dynamic differs drastically in different market structure, estimating the mean and variance from different market regime data can lead to misleading signals. We show that our two step market regime appraoch both provide better alpha generation and robust risk management.

2 Data

In this section we discuss the data we use in this project, including two main part, the asset data we use in our portfolio and the feature data we use to train the preidctive model. The data source of our data majorly come from Yahoo Finance Federal Reserve Economic Data (FRED).

2.1 Asset Data

The dataset we use in this project is the SPDR sector ETFs. The SPDR Sector ETFs are a series of exchange-traded funds (ETFs) that allow investors to target specific sectors of the economy. Each ETF focuses on a particular sector, enabling investors to gain exposure to industries such as technology, healthcare, financials, consumer discretionary, energy, and others without having to buy shares in every individual company within that sector. There are total of 11 sector ETFs and we use 9 of them for the rest has limited data points.

Table 1 is the full 11 sector ETFs and its inception date, we can see the Real Estate and Communication Services sector contain limited data for the whole dataset to functional, and therefore we only consider the rest of the sector in our project. The data range in this project is from 1999-01-01 to 2023-12-31.

Sector	**Symbol**	**Inception**
Communication Services	XLC	2018-06-19
Consumer Discretionary	XLY	1998-12-22
Consumer Staples	XLP	1998-12-22
Energy	XLE	1998-12-22
Financials	XLF	1998-12-22
Health Care	XLV	1998-12-22
Industrials	XLI	1998-12-22
Information Technology	XLK	1998-12-22
Materials	XLB	1998-12-22
Real Estate	XLRE	2015-10-08
Utilities	XLU	1998-12-22

Table 1: SPDR sector ETFs

2.2 Feature Data

Our models are trained based on both macroeconomic variables and financial turbulence indicators.

2.2.1 Feature Construction - Macroeconomic Factors

We choose 14 macroeconomic variables and categorize them into 5 factors. Here are the details about the 5 factors:

1. Inflation. Inflation influences expectations and can significantly affect factor performance. It affects interest rates, bond yields, and the appeal of value stocks compared to other investments. The models consider the annualized changes in core CPI over the past one and ten years as inflation indicators.

2. GDP Growth / Business Cycle. GDP growth is vital as it indicates business profit cycles, with businesses often seeing faster profit growth during economic expansions. This benefits value stocks, which perform well due to operational leverage. Models use US real GDP growth to

assess economic expansions and include the Composite Leading Indicators, Coincident Economic Activity Index, and metrics like Average Weeks Unemployed, Unit Labor Costs, and the consumer credit to personal income ratio for economic indexing.

3. Financial Conditions. In financial markets, factor performance is significantly affected by financial conditions such as changes in interest rates, loan availability, and market volatility. Models use the Chicago Fed National Financial Conditions Risk Subindex, Credit Subindex, and Leverage Subindex as key indicators to assess these conditions.

4. Monetary Policy Expectations. Policy expectations significantly impact factor performance in financial markets, affecting investor behavior and asset prices. Shifts in expectations about fiscal, monetary, or regulatory policies can influence factors like growth, value, and momentum. To measure these expectations, models use indicators such as the Federal Funds Effective Rate and the Volatility Index (VIX).

5. Equity Earning Yield. The equity earning yield has a big impact on valuations that are based on projected future earnings for companies. The Cyclically Adjusted PE Ratio (CAPE Ratio) represents this factor.

Combining these five factors into one comprehensive data frame without any missing values, it includes 320 observations starting in March 1997 and ending in October 2023.

2.2.2 Feature Construction - Financial Turbulence

Kritzman and Li (2010) [5] introduced a novel method to identify systemic risk in financial markets by defining the concept of financial turbulence. This concept describes situations where asset prices significantly deviate from their usual patterns, featuring intense price fluctuations, a breakdown in correlations between assets that typically move together, and unexpected alignment among previously uncorrelated assets. Recognizing financial turbulence enhances our understanding of market behavior and the risks of widespread market disruptions.
Financial turbulence is measured as:

$$d_t = (y_t - \mu)\Sigma^{-1}(y_t - \mu)'$$

where

- d_t = turbulence for a particular time period t
- y_t = vector of asset returns for period t
- μ = sample average vector of historical returns
- Σ = sample covariance matrix of historical returns

Following the methodology of Kritzman and Li, we utilized monthly returns from ten SP 500 sectors (including information technology, energy, financials, healthcare, consumer staples, consumer discretionary, utilities, industrials, telecommunications, and materials) along with returns from Treasury securities of various maturities (1 year, 2 years, and 10 years) to develop our series on financial turbulence, as demonstrated in Figure 1. The financial turbulence series has 302 observations starting from February 1999 and ending in March 2024. Our model acknowledges that accurately identifying market risk regimes requires ongoing analysis of market conditions over time, rather than solely focusing on current situations.

3 Methods

The market drawdown of SP500, shown in Figure 2, starts on February 1987 and ends in January 2024, which yields a total of 443 observations.

3.1 K-Means Cluster

We employed the K-means clustering algorithm, a popular unsupervised machine learning method, to investigate equity market drawdown data. K-means clustering segregates data points into K clusters (Hartigan and Wong, 1979) [6], assigning each data point to the cluster whose mean is closest to it. From a mathematical perspective, given N data points $\{x_1, x_2, ..., x_n\}$, the goal of K-means clustering is to identify K cluster centroids $\{\mu_1, \mu_2, ..., \mu_k\}$ that minimize the sum of squared distances within each cluster:

$$minimize = \Sigma_i \Sigma_j \|x_i - \mu_j\|^2$$

where

$$\|x_i - \mu_j\|$$

represents the Euclidean distance between the data point x_i and the cluster centroid μ_j.
As stated earlier, we started by using K-Means clustering to divide the SP 500 drawdown data into different clusters. We used the Elbow approach to find the ideal number of clusters, as shown in Figure 3, and found that three clusters was the best option for clustering the data.
This three-cluster technique (cluster 0 represents the "normal market risk regime," cluster 1 represents the "market correction regime," and cluster 2 represents the "bear market regime") effectively distinguishes between the different market risk regimes based on the market drawdown data, as demonstrated in Figure 4.
As shown in Table 2, 63.57% of drawdowns are labeled as Cluster 0, representing the normal market risk regime; 26.47% of drawdowns are labeled as Cluster 1, representing the market correction regime; and 9.95% of drawdowns are labeled as Cluster 2, representing the bear market regime.

Clusters	**Percentage**
Cluster 0 (normal market risk regime)	63.57%
Cluster 1 (market correction regime)	26.47%
Cluster 2 (bear market regime)	9.95%

Table 2: Percentage of Each Cluster

3.2 Synthetic Minority Over-Sampling Technique

Imbalanced datasets present significant challenges in machine learning classification tasks. Most classification algorithms tend to favor majority classes, as these models are predominantly trained on more frequently labeled data. This bias often results in suboptimal performance when classifying minority labeled data. One popular techniques to address such problem is Synthetic Minority Oversampling Technique (SMOTE), which is developed by Chawla, Bowyer, Hall, and Kegelmeyer (2011) [7]. In our project, we specifically apply the package `imblearn` to use the SMOTE method on our dataset.
SMOTE is a popular approach that synthesizes new minority class instances by interpolating existing samples. The technique is based on the concept of generating synthetic data points along the line segments connecting nearest neighbors in the feature space. Mathematically, given a minority class instance x_i and its K nearest neighbors, SMOTE creates new synthetic instances $x_$new by randomly selecting a neighbor x_j and computing:

$$x\_\text{new} = x_i + \alpha * (x_j - x_i)$$

where α is a random value between 0 and 1 . SMOTE effectively augments the minority class, providing the model with more diverse examples and mitigating the class imbalance problem.
Considering that the majority of our data (63.57%) is classed as normal market risk, 26.47% as market correction, and 9.95% as bear market regime, we decided to resolve the class imbalance by using the synthetic minority oversampling technique (SMOTE). We tried to balance the dataset by creating artificial examples for the minority classes.

3.3 Time Series Cross Validation

Time series cross-validation is a critical technique used in predictive modeling for time-dependent data. It differs fundamentally from traditional cross-validation methods due to the sequential nature of time series data.
Time series cross-validation offers several theoretical advantages that are pivotal for effective model evaluation:

- Prevents Look-Ahead Bias: By ensuring that the model only uses past data to predict future outcomes, it eliminates the risk of look-ahead bias, where future information is inadvertently used in the model training process.
- Accurately Assesses Model Stability: It evaluates a model's stability and performance over time, providing insights into how the model reacts to new data as it becomes available.
- Reflects Real-World Performance: By mimicking the actual application scenario of time series forecasting, this method provides a more accurate assessment of how the model will perform in practice.

3.4 XGBoost Classifier

XGBoost is a sophisticated and scalable implementation of the gradient boosting framework, designed to enhance speed and performance. It operates by building an ensemble of decision trees sequentially, with each tree aiming to correct the errors of the previous ones, thereby improving accuracy progressively. The process involves using gradient descent to minimize a regularized loss function, which includes L1 and L2 regularization to prevent overfitting.
In XGBoost, each new tree addresses the residuals or errors from prior trees, continuously refining the model's predictions towards the true values. The loss function integral to XGBoost includes a training loss, which evaluates the model's predictive accuracy on training data, and a regularization term that controls complexity, ensuring the model remains applicable to new, unseen data.
To predict next market regime, we apply XGBoost to macroeconomic feature. Our goal was to determine which model performs better in terms of various scoring metrics, including F1 Score, Accuracy, Precision, and Recall. We employed a TimeSeriesSplit with 5 splits for cross-validation to ensure the temporal integrity of our financial data.
Pipelines were constructed using XGBoost classifiers. For the XGBoost classifier, we also considered the learning rate as a parameter. The parameter grids for these models were as follows:

- XGBoost Parameters:
  - Number of Estimators: 100, 200
  - Max Depth: 6, 10
  - Learning Rate: 0.01, 0.1

The scoring metrics included F1 Score, accuracy, precision, and recall. Both models were refitted on the best parameters based on the F1 Score.

	XGBoost
F1 Score	0.898
Accuracy	0.910
Precision	0.900
Recall	0.897

Table 3: Comparison Between Random Forest and XGBoost

The detailed scores are shown on table 3:
After retraining the best XGBoost model on the resampled training data and evaluating it on the test set, the model achieve a F1 Score of 0.776 and accuracy of 0.818

4 Portfolio Optimization

To show the efficacy of the regime detection and prediction we now integrate the model into portfolio optimization. We mainly have 3 steps in this project: (1) Application of clustering algorithms, such as K-Means, to identify prevailing market regimes. (2) Utilization of macroeconomic factors to develop features for a classification algorithm, which forecasts the subsequent month's market regime. (3) Employment of the predictive model to ascertain the forthcoming period's regime, followed by the application of the most recent analogous regime data to estimate parameters in the optimization process.
We formulate the optimization problem as follow:

$$\begin{aligned} \text{maximize} \quad & \mu^T(w_0 + w_1) - cw_1 \\ & - \gamma(w_0 + w_1)^T \Sigma (w_0 + w_1) \\ \text{subject to} \quad & \mathbf{1}^T(w_0 + w_1) = 1 \end{aligned}$$

Where w_0 is the initial weight vector at the beginning of optimization period, which is a zero vector at the start of the optimization. w_1 is the next period weight vector which we get from optimization. c is the transaction cost parameter, here we set to $c = 0.5\%$ as a convention to calculate transaction cost. γ is the risk aversion coefficient which depends on investor's risk profile.

5 Backtesting Framework

Following is the whole backtesting framework for our project:

1. Split the data into training and testing data to prevent data leaking.
2. Using the K-Means clustering algorithm to determine the market regime.
3. Training a Classification model (Random Forest/ XGBoost) with macroeconomic feature to predict the next period market regime
4. At each period for optimization, use the model to determine the next period market regime.
5. With the forecasted market regime, say i, look back the same amount of data points in the past and use the data points that has the same market regime as forecasted and estimate the return (μ_i) and covariance matrix (Σ_i).
6. If the look back period has no same regime data points as the predicted future regime, we thinks the market would have a extreme shifting and we would sell all the position and buy the risk free asset (Assuming 5% per year)

7. Using the calculated μ_i and Σ_i, perform a Markowitz portfolio optimization and obtain the optimal portfolio weights w_1 and update $w_0 = w_0 + w1$ for next period.

8. Repeat the rebalancing procedure until the end.

6 Results & Discussion

6.1 Portfolio Performance

The comparative analysis between the classical Markowitz portfolio optimization and our proposed regime-detection-enhanced portfolio is demonstrated through their respective performance plots. It is evident from these plots that the incorporation of regime detection markedly enhances portfolio performance. Particularly, the drawdown analysis during the 2020 COVID-19 market crash reveals a significantly minimized drawdown in the regime-detection portfolio compared to its classical counterpart. This reduction in drawdown underscores the regime detection model's efficacy in anticipating market fluctuations and adapting the portfolio configuration accordingly, thereby showcasing a promising feature of future market regime awareness.
Further, the performance metrics table substantiates the outperformance of the regime detection model over the classical optimization strategy. This outperformance is mainly attributed to the returns aspect, where the mean return of the regime detection portfolio is approximately twice that of the classical model. Meanwhile, other metrics such as standard deviation and maximum drawdown remain comparable between the two models. The analysis of both the net value and the performance metrics indicates that the regime detection strategy not only offers enhanced alpha generation but also improves risk management. Consequently, these findings suggest that the regime detection model merits further investigation for its potential applications in portfolio management.

Metric	**Regime**	**Classic**
Mean (%)	0.080	0.035
Std (%)	0.237	0.183
MaxDD (%)	-0.930	-1.052
Calmar	1.030	0.395
Sharpe	1.169	0.653

Table 4: Performance Metrics

6.2 Future Improvement

We introduce a novel framework aimed at enhancing the robustness of parameter estimation across various market regimes. Despite initial progress, further development is necessary to ensure the framework's reliability for practical trading applications.

6.3 Regime Detection and Clustering Data

In this study, we applied the K-Means clustering algorithm to drawdown data from the S&P 500 index. Although the results were promising, relying solely on drawdowns for market indices analysis presents limitations due to the exclusion of other pertinent market data. An innovative approach has been demonstrated by Two Sigma, which employs a Gaussian Mixture Model to analyze multiple factor data. This method allows for the clustering of a more dynamic and diverse array of information, potentially yielding more significant and robust market regime classifications.

6.4 Regime Prediction and Classification Models

Our comparative analysis of the Random Forest and XGBoost models led to the selection of XGBoost due to its superior performance. Nevertheless, for enhanced robustness in prediction, advanced models such as neural networks and deep learning may be more appropriate. Furthermore, meticulous hyperparameter tuning is essential prior to deploying these strategies in live settings. In this project, we segmented the data into training and testing sets, using the former solely for training purposes. As an improvement, we propose employing a rolling methodology, continuously integrating new data into the model throughout the optimization process to refine its accuracy.

6.5 Signal Construction and Risk Estimation

This project initially employed a simple mean-variance optimization using past means and covariances to estimate future return signals and risk metrics. However, a more sophisticated approach could involve using factors rather than returns as trading signals. Numerous studies have indicated that factors, which exhibit cyclical performance, can generate more substantial alpha and be effectively integrated into a market regime detection framework. Additionally, rather than utilizing covariance as a risk indicator, alternative methodologies such as applying a Barra factor model or leveraging machine learning techniques to predict future covariances could provide a more accurate risk assessment, particularly when combined with regime detection models.

7 Bibliography

References

[1] Cao, L., Sun, R., Ma, T., & Liu, C. (2023). On asymmetric correlations and their applications in financial markets. Journal of Risk and Financial Management, 16(3), 187.

[2] Ammann, M., et al. (2020). A hybrid learning approach to detecting regime switches in financial markets. Journal of Financial Data Science. doi:10.3390/jfds2020045.

[3] Wang, M., Lin, Y.-H., & Mikhelson, I. (2020). Regime-switching factor investing with hidden Markov models. Journal of Risk Financial Management, 13(12), 311. doi:10.3390/jrfm13120311.

[4] Werge, N. (2021). Predicting risk-adjusted returns using an asset independent regime-switching model. Expert Systems with Applications, 184, 115576.

[5] Kritzman, M., & Li, Y. (2010). Skulls, financial turbulence, and risk management. Financial Analysts Journal, 66(5), 30-41.

[6] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1), 100-108.

[7] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2011). SMOTE: Synthetic minority over-sampling technique. The Journal of Artificial Intelligence Research, 16, 321-357.

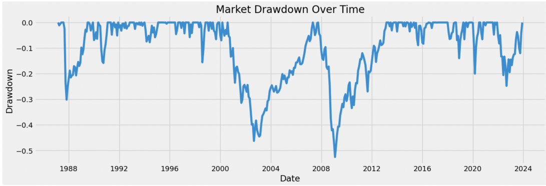

Figure 1: Market Drawdown Over Time

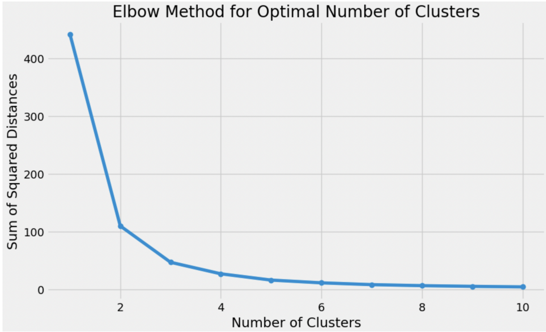

Figure 2: Elbow Method for Optimal Number of Clusters

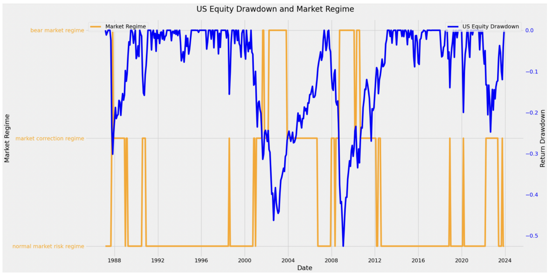

Figure 3: US Equity Drawdown and Clustering

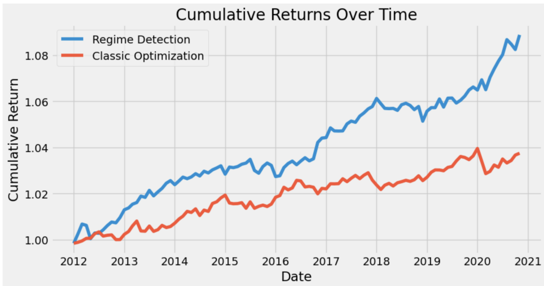

Figure 4: Portfolio Performance

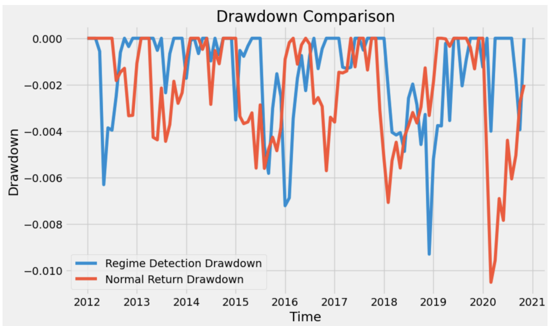

Figure 5: Drawdown of Portfolio