Class: Time series and Anomaly Detection , Master Quantitative Economics

Professor: MADALINA OLTEANU

Unveiling Market Dynamics: A Hierarchical Hidden Markov Model Analysis of the DAX and S&P 500 Indices

Duy-Tung TRAN

December 2024

Table of Contents

1. Introduction

Background

Financial markets are among the most dynamic and complex systems, which can exhibit various mechanisms of stability, growth, rising, and falling yields. Classification of such states, including moderate growth and falling or rising trends known as bullish or bearish markets, is vital for investment purposes and risk management. Despite the broad utilization of traditional statistical models such as autoregressive models and simple Hidden Markov Models for the temporal analysis of financial markets, they often fail to consider the multiple scales at which markets operate. For this reason, advanced models capable of identifying market dynamics compounding different temporal regimes need to be applied.

Research Objective

The article by Oelschläger, L., & Adam, T. (2023) entitled ***Detecting bearish and bullish markets using Hierarchical Hidden Markov Models*** introduces the Hierarchical Hidden Markov Model that could be used as a novel tool for financial markets temporal analysis. The innovation is achieved through extending the standard Hidden Markov Models to include a coarse scale that represents long-term trends and a fine scale that considers short-term fluctuations. The article aims to show how the systems could be modeled using HHMMs and how this model is advantageous compared to using the standard HMMs and apply the methodology to the real financial data, including DAX and S&P 500 indices.

Relevance and Importance

Hierarchical analysis is beneficial in financial markets, as fluctuations occur on different time frames. Introducing a multi-scale perspective allows for recognizing the markets' leaders and laggards in terms of risk and profit, protecting the investor's portfolio from adverse effects. The applicability of such a tool in the financial markets becomes evident during market crises, enabling the investor to predict the shifts in the lurched markets and avoid the potential losses.

Objective of This Report

The objective of this report is to summarize the paper and its key findings and methods, reproduce the results presented by the author using his datasets, and evaluate the method's potential performance in determining the financial market regimes. The report concludes with the analysis of the HHMMs' strengths and limitations and the ways to improve them.

2. Methodology

2.1 Introduction of Hierarchical Hidden Markov Model (HHMM)

Hierarchical Hidden Markov Models (HHMM) take the standard HMM and simply add a hierarchy of states to it, with 2 levels. To exploit both coarse-level and fine-level states, we use the hierarchical hidden Markov model (HHMM), where a coarse-state HMM controls the evolution of the cluster of fine-state HMMs to represent the different underlying processes at the fine level of detail. This enables modeling long-term trends (i.e., coarse-level states) and short-term fluctuations (i.e., fine-level states), making HHMMs ideal to be employed in modeling financial markets, as macroeconomic factors may impact micro-level behaviors.

At the coarse level, those HHMMs characterize predominant market regimes i.e. bullish (growing), bearish (dying), and neutral (steady state). States are governed by a coarse-level transition matrix, which describes the probabilities of transitioning between regimes. At the coarse level, HHMMs model regimes of different trading behaviors (e.g., liquidity provision vs. taking liquidity), while at the fine level, each course-level state is modeled by its own fine level HMM which captures the high and low volatility periods within each of these regimes.

The main ingredients of HHMMs are:

- Transition Matrices: The transition matrix at coarse-level determines transitions between regimes (in term of probability), whereas fine-level transition matrix describes transitions among regimes.
- Emission Probabilities: The emission probabilities model how likely it is to observe particular data points given the current hidden state (often using Gaussian distributions).
- Initial State Probabilities: These set the probabilities of being in any coarse and fine state at the initial moment.

2.2 Implementation Steps

Data Preprocessing:

- The dataset includes the daily closing price, Log-returns are computed daily. This conversion makes variance stable and stationary.
- We then normalize the log-returns to get rid of any differences in scale and increase the comparability across regimes.

Model Initialization:

- HHMM is initiated with a fixed number of states: Three coarse-level states indicative of bullish, bearish, and neutral market regimes.
- A high level, low level for each coarse state.
- Transition matrices, emission (means and variances), and initial state probabilities are randomly initialized

Parameter Estimation (**Baum-Welch Algorithm**): Model parameters are estimated using Expectation-Maximization (EM) framework as:

- E step : Forward backward probabilities at coarse and fine level are calculated as to estimate the probability of each state at each time step.
- In the M-step, the transition matrices and emission parameters are updated with the transition and emission probabilities, which maximizes the probability of the observable data.

The **Viterbi Algorithm** is used to recover the most probable sequence of states:

- On the coarse level, states are decoded to find the market regimes (e.g., bullish, bearish).
- On this fine level, states are decoded into visible short-term volatility clusters or micro-clusters for each regime.

2.3 Visualization and Evaluation

We perform the following steps to interpret and validate the results:

The log-returns are shown, with the underlying regime transitions and short-term fluctuations revealed by plotting the decoded coarse and fine-level states across time. These visualizations are also increasingly correlated with major financial events, like the 2008 financial crisis, as further validation of the model's accuracy.

We evaluate the fit of the HHMM to the data based on a log-likelihood estimate of the model, and verify that no significant patterns remain underexplained by the model, with pseudo-residuals.

2.4 Tools and Libraries

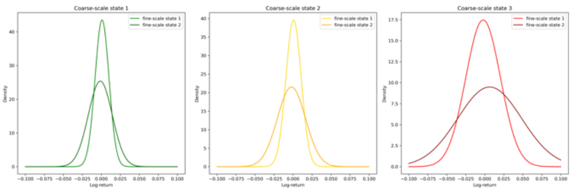
HHMM is implemented in Python, and makes use of the following libraries:

- hmmlearn: To implement and train HMMs, extended into Hierarchical structures
- numpy and pandas: For numerical computations and data preprocessing.
- matplotlib: For plot time series and visualize state transitions.

3. Results

In this section, we discuss the results of the Hierarchical Hidden Markov Model applied to the daily log-returns of the DAX index and S&P 500 index from the years 2000 to 2020. Through the HHMM framework, coarse-scale states (reflecting long-term trends) and fine-scale states (representing short-term fluctuations) could be identified, allowing the identification of significant market regimes, as well as transitions between them. The study gives an overview of key market behaviors around pivotal times like the financial crisis of 2008 and the periods of recovery since then.

3.1 Results for the DAX Index



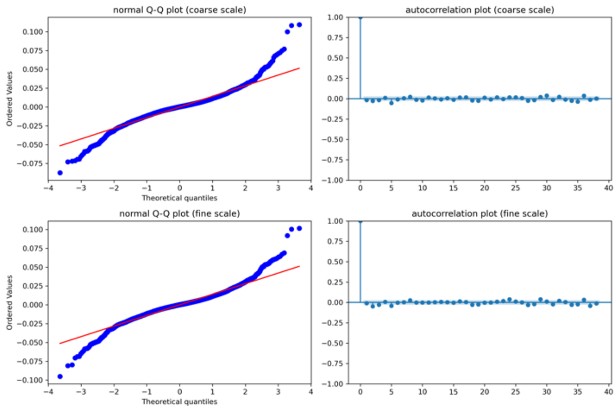


Figure 1: *State-dependent log-return distributions highlight the distinct volatility and return characteristics across states.*

Figure 2: *Pseudo-residuals exhibit low autocorrelation and follow a near-normal distribution, validating the model fit.*

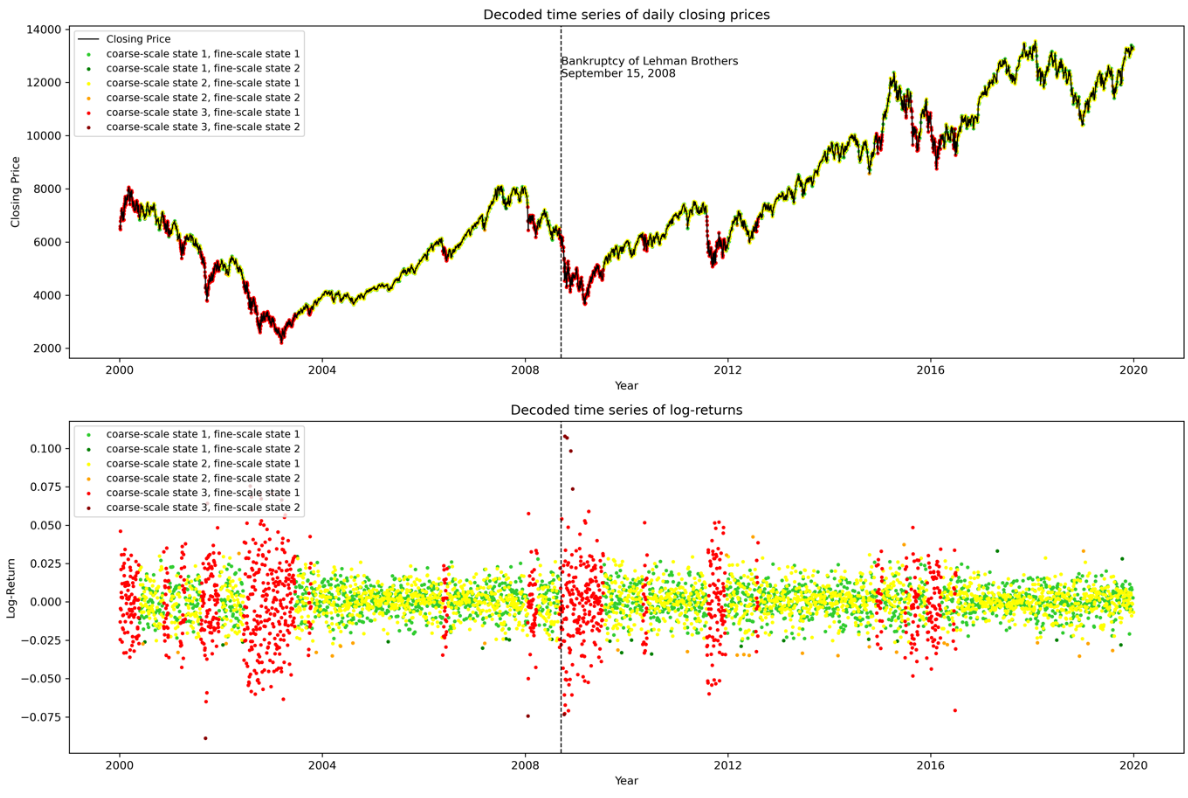


Figure 3: *Decoded time series clearly distinguish between the bullish, neutral, and bearish market phases for the DAX.*

3.1.1 Coarse-Scale Analysis for the DAX Index

The HHMM identified three distinct market regimes in the DAX index, which appear to be in line with known market behaviors and historical events. During the **bullish market (Coarse State 1**), persistent upward trends, the low volatility, and positive returns prevailed. This state prevailed during extended periods of growth, like 2004–2007 and 2013–2020. Transitioning into this state was especially common immediately following post-crisis recoveries (Figure 3 Decoded Time Series).

Coarse State 2: Neutral Market — Stable & Moderately Volatile Periods With Mixed Returns. Such phases often took place in transition, as in 2010–2012, when the DAX moved from bearish to bullish trends. This is a state of intermediate market sentiment that also serves as a harbinger: either recovery or decline.

The bearish market (Coarse_State 3) was characterized by sharp declines, high volatility, and negative returns. This regime dominated the 2008 financial crisis, which saw the collapse of Lehman Brothers. The red markers bunching together in Figure 3 reflect this dramatic downturn in the market, underscoring the model's success in detecting notable shocks in the financial market.

3.1.2 DAX Index Fine-Scale Analysis

In a bullish market regime, fine-scale states represented small consolidations or accelerations within the larger bull trend. For example, during 2013–2020, temporary downturns at the market level were well modeled while maintaining the broader bullish trend.

Conversely, the bearish market regime showed fine-scale states that delineated between extreme volatility and subsequent stabilization phases. This is especially the case in 2008 during the financial crisis where the most chaotic of times were followed by such a period of calm at first, then transitioned into a more neutral or bullish state. These dynamics are shown in Figure 1 (State-Dependent Log-Return Distributions), with variation in intensity in bearish regimes.

3.2 Results for the S&P 500 Index

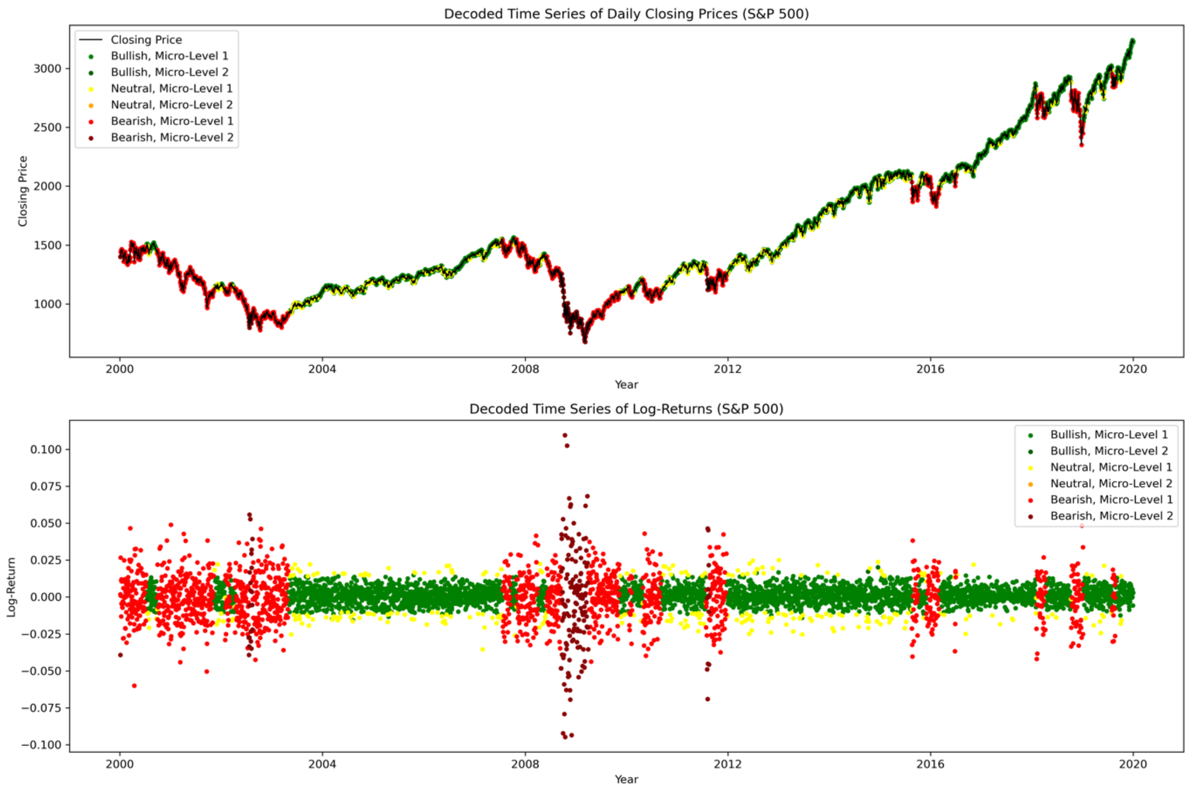


Figure 4: Decoded time series of daily closing prices and log-returns for the S&P 500 illustrates clear market segmentation into bullish, neutral, and bearish states.

3.2.1 Narrow Focus for the S&P 500 Index

In order to assess the HHMM generalizability, the model was applied on the S&P 500 index, showing similar and generalizable consistent market regimes in line with the dynamics established on the DAX index. The bullish market was marked with gains and low returns, indicating long-term positive momentum. This regime was especially dominant during the post-2008 financial crisis recovery, most notably from 2013–2020.

Neutral market covered transition periods like stable prices and moderate volatility. Such periods were often observed during recoveries or stagnations after major market movements, reflecting intermediate stages of the market cycle.

The bearish market is characterized by a significant drop in the price and the volatility of securities, most remarkably during the financial crisis of 2008. As visualized in Figure 4 (Decoded Time Series for the S&P 500), this shows how effectively the HHMM was able to segregate these periods, suggesting that it captures serious market disruptions.

3.2.2 The S&P500 Index: Fine-Scale Analyses

Fine-scale states offered greater resolution, capturing short-term dynamics within coarse-scale regimes. In the context of a bullish market, fine-scale states effectively collected temporary fluctuations without disturbing the broader pattern of persistent growth. We saw reflection of this in the long positive trend for the S&P 500, covering the years between 2013 and 2020.

In the bear market regime, the fine scale states served to distinguish phases of extreme volatility from subsequent stabilisation phases. For instance, in the context of the 2008 financial crisis, the model was able to separate the most severe market drops from the preceding waves of recoveries from which it had transitioned into neutral or bullish regimes. Such granularity shows the ability of the HHMM to better identify market behaviors and adds to its practical value.

4. Discussion

The use of Hierarchical Hidden Markov Models (HHMMs) in financial markets demonstrates their capability to capture intricate temporal relationships. HHMMs provide the clearest insight into regimes by demarcating long-term trends from short-term fluctuations. The subsequent findings of this study, applied to the DAX and S&P 500 indices, highlight the capabilities, as well as the limitations of the HHMM framework when applied to historical market events and dynamics.

4.1 Strengths

HHMMs have several advantages that make them a very appropriate model for financial time series analysis:

Comprehensive regimes detection: HHMMs detected three macro-level states of the market—bullish, bearish and neutral—and captured short-term fluctuations within fine-scale states. Such two-level study details market trends and volatility. For example, HHMMs successfully captured the 2008 financial crisis, including the collapse and recovery of the market.

Relevance to Historical Events: Decoded states closely map onto key historical market events, like the Lehman Brothers collapse, or extended growth periods like 2013–2020. The alignment strengthens the interpretability of the model, and also boosts its applicability to the real world.

Adaptability and Flexibility: This research puts forward the HHMM framework, which is resilient to different types of datasets and market conditions, and thus is able to operate for both the DAX and S&P 500 indices. Moreover, its hierarchical receptive field allows the model to catch temporal correlation in a hierarchical manner, which is largely ignored by traditional methods.

Enhanced Visualization: The system of states explored at coarse and fine levels (Figures 1–4) offer a clear and intuitive interpretation of state transitions with observable market phenomena. The model should be equipped with adequate visual tools for decision-making.

4.2 Limitations

While HHMMs have several advantages, there are also challenges that may restrict their wider application:

Dependence on Data Quality: The performance of the model is highly sensitive to the quality of input features. State detection and parameter estimation are disturbed due to missing values or noise in financial time series.

One of the stationarity and Gaussian emissions assumptions: HHMMs assume stationarity for each regime and normally-distributed emissions. But financial data usually involve heavy tails and nonstationary dynamics that these assumptions do not fully reflect (Stylized Facts).

Computational Complexity: Training HHMMs is CPU-heavy, especially for large datasets. The EM process being iterative and the hierarchical structure causing a substantial increase in resource requirements significantly limits scalability.

Fixed State Numbers: It introduces rigidity because you have to predefine how many coarse and fine states will exist. For example, if we select three coarse states and two fine states, it may not reflect the actual market insights, resulting in not necessarily the optimal solution.

4.3 Comparison with Other Models

Versus Traditional HMMs: Conventional HMMs are simple, but they cannot describe hierarchical relations and multi-scale dynamics. HHMM significantly outperforms themselves across long-term trends and short-term fluctuations.

Comparison with ARIMA and GARCH Models: Although ARIMA and GARCH models are well-suited to forecasting and volatility analysis, they offer few insights into discrete market regimes. HHMMs serve to fill the gap by locating state transitions and hierarchical dependencies.

4.4 Future Directions

In order to make HHMMs more useful and applicable, I propose some avenues of improvement:

Alternative to Gaussian Emissions: It also may benefit from adopting heavy-tailed distributions in the likelihood, such as t-distributions, which increased the model's robustness to extreme market events and outliers.

Dynamic State Selection: This could be improved by automating the determination of state numbers using techniques such as the Bayesian Information Criterion (BIC) or cross-validation, potentially increasing adaptability and accuracy of the model.

Inclusion of External Elements: As another note, integrating some macroeconomics indicators (e.g. GDP growth, interest rates) or sentiment data that could potentially give greater scope of the movements that the market may take.

5. Conclusion

With the hierarchical structure in HHMM model stability can be stablished for long-term trends and short-term fluctuations hence it is emerging as a powerful model for financial time series analysis. In this study, the use of HHMMs on the DAX and S&P 500 indices yielded significant information about market dynamics and are consistent with notable historical events, including the 2008 financial crisis.

More importantly, the results highlight the efficacy of HHMMs in identifying bullish, neutral, and bearish market phases, as well as capturing the nuances of short-term volatility in each regime. HHMMs, which handle hierarchical relationships well, outperform classic models, like standard HMMs and ARIMA/GARCH, which typically fail when faced with multi-scale dynamics. And indeed, the decoded states closely corresponded with significant financial events, demonstrating how interpretable and relevant the model is for financial analysis and decision making.

These strengths notwithstanding, HHMMs are not without limitations. High-quality data, computational complexity, and the need to specify a priori the number of states in the model limit their applicability. The stationarity and Gaussianity assumptions, although helping with the modelling, do not capture well the heavy tails and time-varying behavior prevalent in such data. These limitations indicate that there is room for improvement both in model flexibility and efficiency.

There are many avenues for improving the HHMM framework moving forward. However, allowing non-Gaussian-distributed emissions would improve the model's ability for capturing extreme events (e.g. financial crises), and dynamic selection of the number of states would increase flexibility and adaptability. In addition, incorporating macroeconomic indicators or sentiment data could add depth to the analysis and offer a more holistic perspective on market dynamics. However, there are some solutions for scalability that involve improvements in computational methods, like variational inference or parallel computing, which may help HHMMs overcome the limitations imposed by the computational requirements of parameters given large datasets or high-dimensional time series.

The HHMM is a valuable tool for seeking to remain at the cutting edge of financial analysis and would be especially useful for ones with passion in data science.