Capstone Project Proposal II

Yield Curve Regime Clustering & Interest Rate Forecasting

Team: Sajan Arora & Karan Badlani
Date: November 7, 2025

1. Introduction

The U.S. Treasury yield curve represents interest rates across multiple government bond maturities (e.g., 3-month, 2-year, 10-year, 30-year). Its shape serves as a key indicator of macroeconomic conditions. For instance, a normal upward-sloping yield curve suggests stable economic outlook, while an inverted curve—where short-term yields exceed long-term yields—has historically preceded recessions.

However, yield curve shapes vary over time and do not always fit simple categories. This motivates a data-driven approach to identify distinct yield curve regimes, understand their characteristics, and analyze how these regimes relate to future interest rate movements.

Goal. This project aims to (1) identify yield curve regimes using clustering, and (2) develop forecasting models to predict either future yield values or regime transitions.

Research Questions:

- What distinct yield curve regimes emerge through unsupervised clustering of historical yield curve snapshots?
- Does regime information improve forecasting of interest rates (e.g., the 10-year Treasury yield)?
- Can we predict regime transitions (e.g., entering or leaving inverted yield curve states)?

2. Data Sources

We will use weekly U.S. Treasury constant maturity yields available through the Federal Reserve Bank of St. Louis (FRED) API. Expected maturities include:

Maturity	FRED Series Code
3-Month	TB3MS
2-Year	DGS2
10-Year	DGS10
30-Year	DGS30

We will collect at least 25 years of data to capture multiple macroeconomic cycles.

3. Research Methodology

3.1 Data Preparation

- Download weekly yield data from FRED API.
- Clean missing values using forward-fill or interpolation.
- Construct weekly yield curve "snapshots" as vectors, e.g., $[y_{3m}, y_{2y}, y_{10y}, y_{30y}]$.
- Optionally compute yield spreads (e.g., $y_{10y} - y_{2y}$).

3.2 Clustering Yield Curve Regimes

- Apply clustering (e.g., K-Means or Gaussian Mixture Models) on the yield snapshots.
- Use PCA or t-SNE for visualization and interpretability.
- Interpret resulting clusters as economic regimes (e.g., normal, flat, inverted).

3.3 Forecasting Models

We will explore two predictive tasks:

A. Yield Forecasting (Regression)

- Forecast the 10-year yield using ARIMA, VAR, or Random Forest Regression.
- Evaluate whether regime labels improve accuracy.

B. Regime Transition Prediction (Classification)

- Predict whether the yield curve will shift into a new regime (e.g., inverted) within 4–12 weeks.
- Models: Logistic Regression, Gradient Boosted Trees.
- Evaluate using precision, recall, and confusion matrices.

4. Feasibility, Timeline, and Risks

Since this is a 1-month project, we adopt the following schedule:

Week	Milestone
1	Data collection, cleaning, and yield snapshot construction
2	Clustering experiments and regime interpretation
3	Forecast model implementation and evaluation
4	Final analysis, visualization, and report preparation

Risks and Mitigation:

- Time-series nonstationarity: apply differencing or yield spread analysis.
- Overfitting forecasting models: rely on walk-forward validation.
- If deep learning models are too complex, remain with ARIMA/VAR baselines.

Success Criteria:

- Identification of interpretable yield curve regimes.
- Forecast performance exceeding naive baselines.
- Clear visualization and explanation of regime dynamics.

5. Expected Contributions

This project will provide:

- A data-driven classification of historical yield curve regimes.
- Empirical insight into how regime structure affects interest rate dynamics.
- A framework for forecasting regime shifts, including inversion events linked to recession risk.

References

[1] Federal Reserve Bank of St. Louis (FRED). Treasury Constant Maturity Rates. https://fred.stlouisfed.org/

[2] Estrella, A., & Mishkin, F. S. (1996). *The Yield Curve as a Predictor of U.S. Recessions*. Federal Reserve Bank of New York. https://web.archive.org/web/20150908020815/https://www.newyorkfed.org/research/current_issues/ci12-2/ci12-2.html

[3] Diebold, F. X., & Rudebusch, G. D. (2013). *Yield Curve Modeling and Forecasting*. NBER Working Paper No. 16714. https://www.nber.org/papers/w16714