Creating a Model to Track Recessions in the United States

Lilly Peacor

University of Nevada, Reno
Department of Mathematics and Statistics
STAT 753 Final Project

Abstract

In this study, we model the US economy using previous data of recessions [**5**] and bond rates [**3, 4**]. The chosen model is a 4 state discrete-time Markov chain. This Markov chain will give us a transition matrix which allows us to simulate future recessions and bond rate changes. The stationary distribution allows us to determine future likelihood of recession, and the rate of convergence tells us how long until we are able to reasonably use this stationary distribution to make predictions.

This model aids in showing how strong of a predictor of recession having the short bond rate exceed the long bond rate, and our analysis shows that it is a strong predictor. In historical data, when the short bond rate exceeds the long bond rate, a recession follows shortly after. The economy spends the majority of time in a state where the long rate is greater than the short rate and is not in recession, but it is almost guaranteed that shortly after a yield curve inversion, defined here as the short rate exceeding the long rate, a recession will happen.

Currently, we are in an extended period of the short bond rate staying greater than the long bond rate, and our simulation shows that in almost all cases, we will have at least one period of recession by the year 2030.

1: Introduction

Tracking and anticipating recessions can be quite difficult. In fact, there are no strictly defined rules used to govern the determination of if the United States is in a recession [**1**]. According to The White House, The National Bureau of Economic Research (NBER) Business Cycle Committee commonly uses metrics such as real personal income less government transfers, employment, consumer spending, and industrial production to name a few, but these aren't all. In many cases, the Business Cycle Committee is unable to call a recession as it begins due to a lag in available data, thus it is important that economists find other ways to predict a recession happening in the near future.

The economists at NBER, as well as other private and public economists rely heavily on publicly available data to help predict current and future economic declines. One such source of this data is the Federal Reserve Economic Data (FRED) [**2**]. FRED is a publicly available database which contains 824,000 US and international time series all related to economies, and these time series are compiled from 114 different sources. This provides a convenient place to locate necessary data for stochastic analysis, all from trusted sources.

Some important data sets from FRED used in tracking recessions include the DGS1 [**3**], DGS10 [**4**], and USREC [**5**] data sets. The DGS1 tracks at a daily rate the market yield on U.S. Treasury Securities at 1-year constant maturity, quoted on investment basis. The DGS10 is similar, but tracks the market yield on a 10-year constant maturity basis. These two data sets have data points for every day, with the exception of national holidays and weekends. The USREC data set has monthly data points which state whether NBER determines that the US economy is in a period of expansion or recession. Many economists use the DGS1 and DGS10 as ways to anticipate a recession. If the short bond rate from the DGS1 is greater than the long bond rate from the DGS10, it is predicted that a recession will occur in the near future [**6**].

1.1: Goal

In this analysis, we intend to model the US economy as a 4 state discrete time Markov chain using data from FRED in order to forecast future recessions based on the US bond rates. The states are as follows:

1. Short bond rate $>$ long bond rate and the US is in recession
2. Short bond rate $<$ long bond rate and the US is in recession
3. Short bond rate $>$ long bond rate and the US is not in recession
4. Short bond rate $<$ long bond rate and the US is not in recession

We will find the transition matrix for this Markov chain through computing, conduct theoretical calculations to find the stationary distribution and rate of convergence for this Markov chain, and, using this model, simulate the future 20 years.

2: Methods

2.1: Computing

We pulled the data from the FRED website on 4/18/2024, giving us data spanning from January 1962 through March 2024. The USREC dataset was pulled with one point per month through that timespan, and the DGS1 and DSG10 sets contained daily points excluding holidays and weekends for the entirety of the period.

The analysis was completed using the Python 3 programming language, utilizing the numpy, pandas, and matplotlib.pyplot libraries. The data was imported to a Python script, and Python was used to clean and match the time series together. The daily data from the DGS1 and DGS10 data sets had multiple missing entries due to holidays and weekends, so those entries, initially represented by a '.' were changed to the numpy 'NaN' for ease of processing later. All missing entries were then dropped, and the data type was converted to a float. The DGS1, DGS10, and USREC data sets then had the date column set to the pandas Date_Time object.

A pandas dataframe 'groupby' function was then used to change the datasets to monthly averages, moving the data column to the index. This had no effect on the USREC data aside from matching the index to the indices of the DGS1 and DGS10 data. At this point, we had 3 data frames, and began using merge functions to align them over the date multi-index. This new dataframe was now a discrete time series with an entry in each variable for every month from January 1962 through March 2024.

A new column was formed to indicate if the short bond rate (DGS1) was greater than the long bond rate (DGS10) using a simple loop, and this column was added to the dataframe. A chart was then created to visualize the data using pyplot.

Utilizing the new Boolean variable alongside the USREC variable, we were able to find the states of our Markov chain **(fig. 1)** for every time point. The states are listed below:

1. Short bond rate > long bond rate and the US is in recession
2. Short bond rate < long bond rate and the US is in recession

3. Short bond rate > long bond rate and the US is not in recession

4. Short bond rate < long bond rate and the US is not in recession

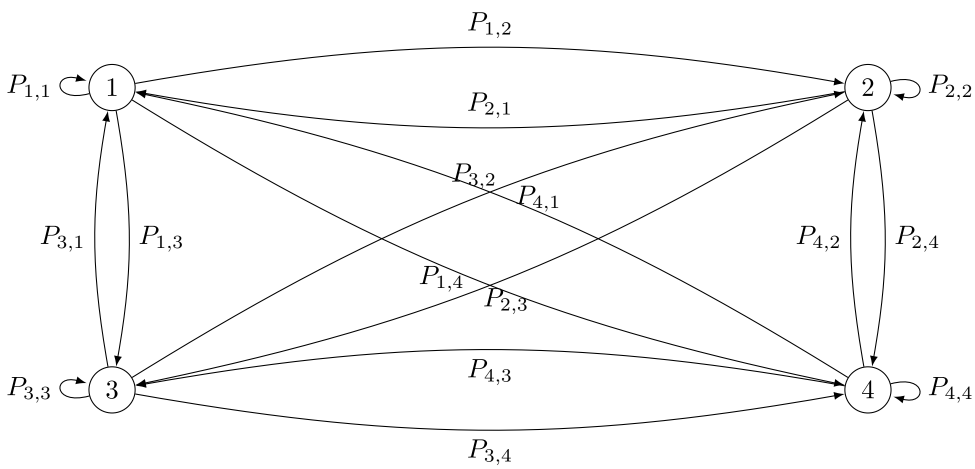


Figure 1: Markov chain showing states and transition probabilities as entries in the transition matrix P

A 4x4 matrix was constructed with counts of transitions from state to state, and the total transitions from each state were counted separately as entries in an array. These were used to construct the transition matrix (**fig. 2**) for the Markov chain.

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & P_{1,3} & P_{1,4} \\ P_{2,1} & P_{2,2} & P_{2,3} & P_{2,4} \\ P_{3,1} & P_{3,2} & P_{3,3} & P_{3,4} \\ P_{4,1} & P_{4,2} & P_{4,3} & P_{4,4} \end{bmatrix}$$

Figure 2: Transition matrix for the Markov chain represented in Fig. 1

This transition matrix was then used to generate 4 simulations of the next 20 years of the Markov chain with monthly data points. These simulations were then plotted using pyplot, each showing the entirety of data, including the time series from before, as well as simulated data, with a dashed line included showing where historical data ends and simulated data begins.

2.2: Theoretical Calcualtions

The transition matrix of the Markov chain found using computing (P) was then used to conduct theoretical analysis to find the stationary distribution (π) and rate of convergence to this stationary distribution. To find these, first the row-eigenvalues and eigenvectors of P were calculated. The eigenvalue of 1 was used to find π, which is the eigenvector corresponding to that eigenvalue after scaling to be of length 1, and the eigenvalue satisfying the inequality

$$\Lambda = max(|\lambda|) < 1$$

was designated the rate of convergence when raised to the n^{th} power [**7**].

3: Results

The time series data visualization below (**fig. 3**) charts the movement of the DGS1 and DGS10 variables over the entire time frame, as well as including lines for the USREC variable and Boolean variable. Figure 3 has shading in red wherever the short rate exceeds the long rate, and shading in blue where the USREC variable indicates a recession occurred. the white zones indicate that we are in state 4, the red zones indicate state 3, the blue zones indicate state 2, and the purple zones indicate state 1.

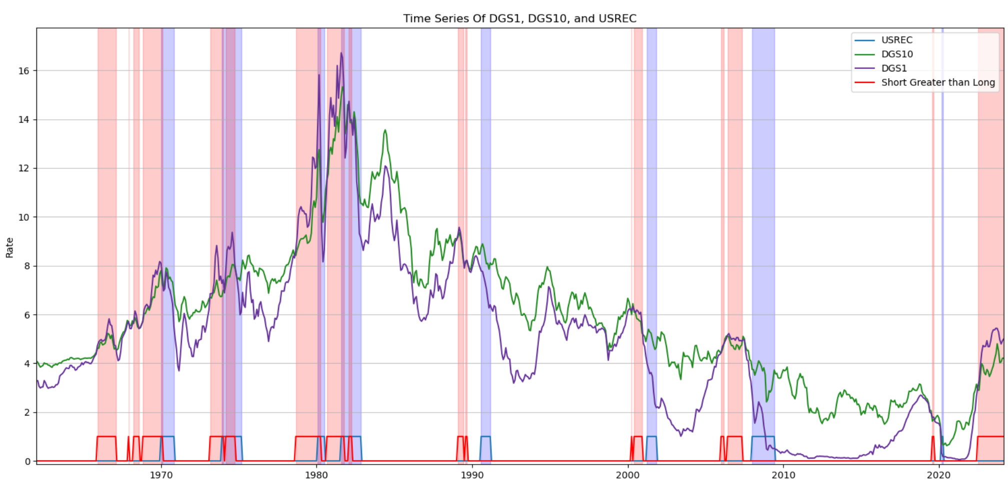


Figure 3: Time Series plot of all variables

The transition probability matrix for the Markov chain (**fig. 4**) shows the empirical probabilities of the process to move from each state to another based on historical data. Each probability was reduced to 3 significant digits. Note that a 0 in any position indicates that the matrix never moved from the state represented by the row number to the state represented by the column number.

$$P = \begin{bmatrix} 0.714 & 0.286 & 0 & 0 \\ 0.031 & 0.844 & 0 & 0.125 \\ 0.032 & 0 & 0.887 & 0.081 \\ 0 & 0.007 & 0.028 & 0.965 \end{bmatrix}$$

Figure 4: Transition matrix P

The presence of zeros in this transition matrix enables us to simplify figure 1. This new figure includes transition probabilities as well (**fig. 5**).

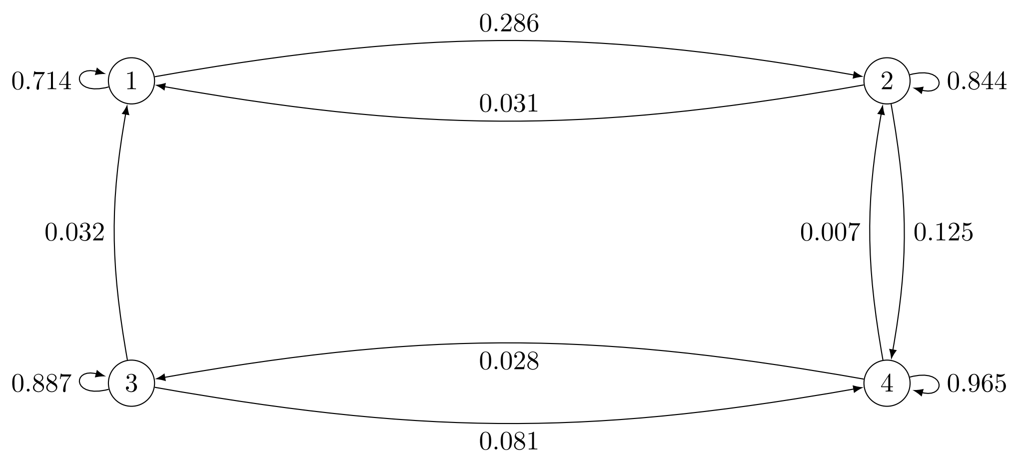


Figure 5: Markov chain showing states and transition probabilities as entries in the transition matrix P

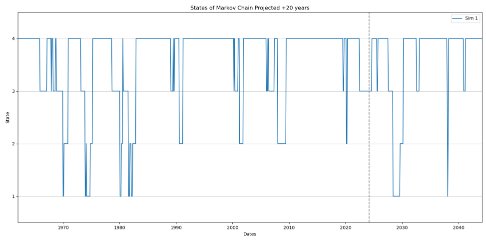
We also are able to visualize the movement of the Markov chain, as shown in **fig. 6 (a)-(d)**. These figures include the historical movement between states, as well as simulated movement, beginning April of 2024. This transition from historical to simulated data points is represented by a vertical dashed line.

We are able to use these visualizations to assist us in analyzing the data set as a whole alongside the transition probabilities.

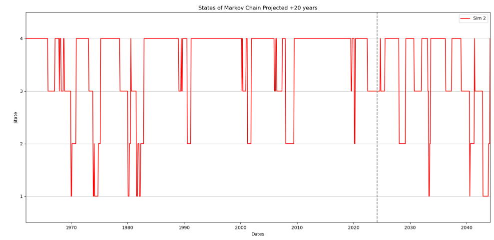
The row eigenvalues of this matrix come out to

$$\lambda_1 = 1, \lambda_2 = 0.874 + 0.031i, \lambda_3 = 0.874 - 0.031i, \lambda_4 = 0.662$$

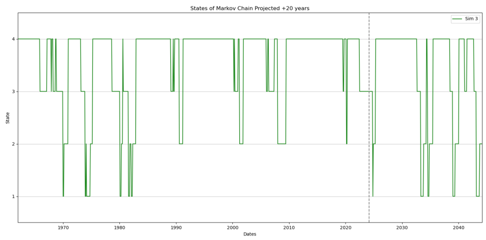
As two of these eigenvalues are imaginary, we must then find the magnitude of those eigenvalues to determine the rate of convergence:



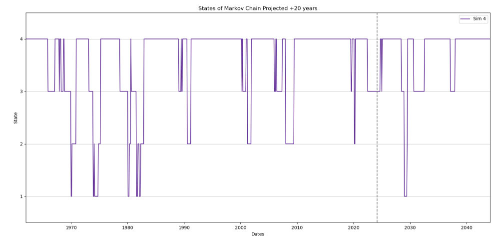
(a)



(b)



(c)



(d)

Figure 6: 4 Representations of the movement of the Markov chain

$$|\lambda_2| = \sqrt{0.874^2 + 0.031^2} = 0.875$$

We can clearly see here that we will obtain the rate of convergence, Λ^n from λ_2, thus

$$\Lambda^n = |\lambda_2|^n = 0.875^n \tag{1}$$

is the rate at which we will approach the stationary distribution.

The eigenvector associated with $\lambda_1 = 1$ is

$$e_1 = \begin{bmatrix} -0.039 & -0.115 & -0.239 & -0.963 \end{bmatrix}$$

We then normalize to length 1 by dividing each component by the sum of components:

$$-0.039 - 0.115 - 0.239 - 0.963 = -1.356$$

$$\frac{e_1}{-1.356} = \pi = \begin{bmatrix} 0.029 & 0.085 & 0.176 & 0.710 \end{bmatrix} \tag{2}$$

This vector π is the stationary distribution, representing the probability that at any time step n, we will be in that state.

4: Discussion

Prior to any analysis, the generation of Figure 3 enables us to see that, in general, it is correct to presume that a recession is coming following periods where the short bond rate exceeds the long bond rate. We note this visually by seeing that every time there is a column of red, it is shortly followed, if not overlapped by, a column of blue, indicating recession. It is interesting to note that at the end of the chart, we see that to date for some time the short bond rate has been greatly exceeding the long bond rate, which, given historical data trends, tells us that we should be expecting a recession in the coming months. This can be difficult to see using the Markov chain, as it is not uncommon for the US to have a period of no recession with the long rate exceeding the short rate prior to a recession being called according to USREC data.

Using the transition probability matrix of the Markov chain, we are able to see a clear picture of the movement between states. Looking along the diagonal of P that the vast majority of the time, this process prefers to stay in the state it is in. We can see that we will never move from not being in recession with the long rate greater than the short rate, directly to being in recession with a yield curve inversion. It is more common for the economy to go from state 4 to the yield curve inverting with no recession than to go directly to recession. We can see that when starting from state 3, we will never go directly to being in recession with the long rate greater than the short rate. From state 3 we will more commonly go to recession with an inverted yield curve than to no recession and no inverted yield curve. This helps to confirm the above statement that the short rate being above the long rate is a good indicator of near-future recession. Generally, once in recession we will see the short rate fall back below the long rate if it was there at the beginning, then move out of recession, as shown by $P_{2,4} > P_{2,1}$, though it is not unheard of for the rate to invert again, it is substantially less common. When we are in a state of having the short bond rate higher than the long bond rate, this is the least stable state, as shown by $P_{1,1}$ being substantially less than $P_{2,2}$, $P_{3,3}$, and $P_{4,4}$. This state will only move to being in recession with the long rate greater than the short rate, never leaving recession directly from that state.

The stationary distribution represented in Equation (2) tells us a lot about the probability of being in any specific state at one specific point in time. It shows us that most of the time, with a likelihood of 71%, we will be in state 4 with no recession and the long rate greater than the short rate. Otherwise, in decreasing probability, we will be in state 3 with the yield curve inverted and no recession (17.6%), then state 2: in recession with the long rate greater than the short rate (8.5%), and finally in a recession with the yield curve inverted (2.9%). We can confirm this visually by looking at figure 3. It is important to note that with the rate of convergence seen in Equation (1), Λ^n, it took some time to reach this point, but with the position of $n = 748$, we are extremely close to this stationary distribution.

Moving on to view Figure 6, we can see 4 possible trajectories for the future. With our current state of having an inverted yield curve, it is expected that we will have a recession in the coming months, if not few years, and this is confirmed by the simulations. All 4 show at least 1 recession prior to the year 2030, and most showing significant volatility thereafter.

Future analysis of this data is warranted. This should include using the transition matrix to find the distribution at step n for all n, determining mean time spent in each state when stating from another state, and potentially constructing a continuous time Markov chain model to determine exit rates from each state to aid us in predicting how long we can expect the economy to hold up in times of expansion, as well as the amount of time we can expect to be in economic downturns. It is important and necessary to continuously conduct simulations based on new data, as this transition matrix may not hold true as time moves on and more data is available to us.

It would be beneficial to look for other markers of future recession, similar to utilizing the difference between the short bond rate and long bond rate. This will enable us to have a stronger model to predict future economic disaster with higher accuracy.

References

1. The White House, How Do Economissts Determine Whether the Economy Is in a Recession?, July 21, 2022; https://www.whitehouse.gov/cea/written-materials/2022/07/21/how-do-economists-determine-whether-the-economy-is-in-a-recession/

2. Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/

3. Board of Governors of the Federal Reserve System (US), Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity, Quoted on an Investment Basis [DGS1], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DGS1, April 18, 2024.

4. Board of Governors of the Federal Reserve System (US), Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis [DGS10], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DGS10, April 18, 2024.

5. Federal Reserve Bank of St. Louis, NBER based Recession Indicators for the United States from the Period following the Peak through the Trough [USREC], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/USREC, April 18, 2024.

6. Andrey Sarantsev, Quantitative Finance Projects, April 2024, Canvas Web Page; https://webcampus.unr.edu/courses/106400/pages/quantitative-finance-projects

7. Andrey Sarantsev, Introduction to Probability Theory and Stochastic Processes for Undergraduate Students, February 28, 2023; https://asarantsev.github.io/WebArchive/notes.pdf