Exploring Regime-adaptive Linear Models for Time Series Forecasting

Jianzhi Wang

April 2 2024

1 Introduction

Linear models are used ubiquitously in prediction. In the setting of time series, data typically comes sequentially and will be correlated with the past observations. Techniques in linear models have been developed to deal with the sequential nature of the data, such as the Gauss Updating Formula. In this project, I further explore other related linear models, such as ridge regression, indicator regime regression, and switching autoregressive model, that help to enhance one-step ahead prediction.

For the dataset, I chose to look into the U.S. unemployment rate. Unemployment rate is known to be a lagging indicator, influenced by other macroeconomic factors such as inflation and interest rate. A survey of models used to model unemployment rate includes the Markov Switching Model[1][2][3]; however, those models typically focus on a univariate analysis. There were also studies done to compare causes of unemployment between various states in the United States[4]. Hence, the research problem is to predict U.S. unemployment rate, in a one-step ahead manner, using various macroeconomic indicators as covariates. The macroeconomic indicators consists the Gross Domestic Product (GDP), Consumer Price Index (CPI) and interest rate.

In this project, I will investigate the following problems:

- Is interest rate a useful predictor of U.S. unemployment rate?
- Is it possible to predict one-month-ahead unemployment rate?
- What are the models to fit unemployment rate?

2 Data

2.1 Description

The dataset comprises multiple columns, all of which are obtained from FRED. The columns are curated to consist various macroeconomic metrics. The unemployment rate spans January 1948 to March 2024. However, only the part after July 1954 is used, due to the unavailability of other covariates earlier.

- `DATE`
- `UNRATE`[5]: the unemployment rate, in percent, seasonally adjusted
- `FEDFUNDS`[6]: Federal Funds Effective Rate, in percent, not seasonally adjusted
- `GDP`[7]: Gross Domestic Product, released quarterly.
- `CPIAUCSL`[8]: Consumer Price Index for all urban consumers

2.2 Exploratory Data Analysis

Firstly, the data is split into train ($n = 700$) and test set ($n = 134$). The train set consists of data from July 1954 to October 2012.

Figure 1 shows the graph of the unemployment data in the training set. Visually, one can see that the U.S. unemployment follows a cyclical pattern. However, the pattern does not have a fixed period. Moreover, it consists steeper rises and gentler falls. Both the rising and faller process seem linear. Thus, it will be ideal to model it by two regimes, provided that the change point can be accurately predicted. This prompts the use of the switching autoregressive model, which will be described in the later sections.

Figure 2 shows the graph of the differenced unemployment rate in the training set. Due to the precision of measurement, the unemployment data is recorded to the nearest 0.1 percent. The ADF test for stationarity gave a p-value of $0.01 < 0.05$, hence we reject the null hypothesis that the differenced time series is non-stationary.

Therefore, we proceed with the autocorrelation and partial autocorrelation analysis. Figures 3 and 4 show the autocorrelation and partial autocorrelation plot of `diff` respectively. As expected, the difference array show significant autocorrelation and partial autocorrelation. I chose to focus on the partial autocorrelation, which showed significant coefficients at lag 1, 2, 3, 4, 12, and 24. The partial autocorrelation at 12 and 24 are somewhat expected as they indicate annual trends. This motivates the use of AR(4) as a baseline model.

Lastly, the correlation matrix between various variables is computed in Table 1. As shown, `diff` is more correlated with `FEDFUNDS` as compared to other covariates. This can be explained by the Phillips curve in economic theory, which claims that unemployment and inflation has a negative relationship. Since

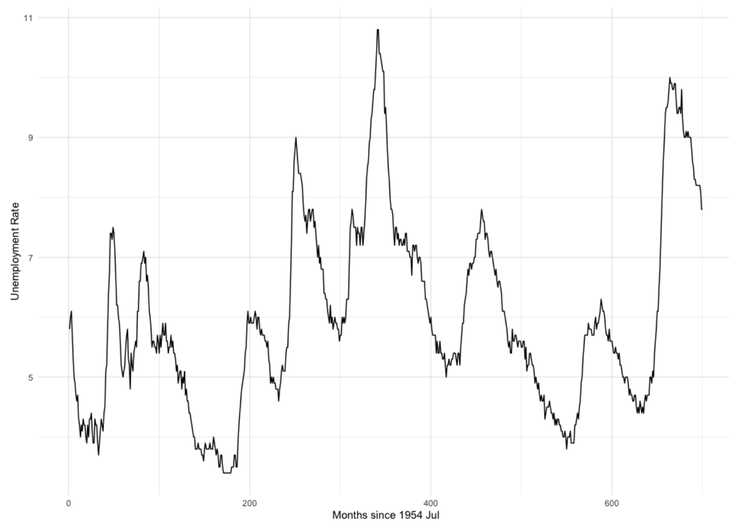

Figure 1: Graph of U.S. Unemployment from 1954 Jul to 2012 Oct

interest rate and inflation also has a negative relationship, there is a positive relationship between unemployment rate (`diff`) and interest rate (`FEDFUNDS`). Inspired by this, the significance of `FEDFUNDS` as a covariate will be investigated.

	diff	UNRATE	FEDFUNDS	GDP	CPIAUCSL
diff	1.000	-0.052	0.111	0.025	0.020
UNRATE		1.000	0.041	0.201	0.250
FEDFUNDS			1.000	-0.328	-0.220
GDP				1.000	0.981
CPIAUCSL					1.000

Table 1: Correlation Matrix

3 Effect of Interest Rate on Unemployment

The common view that an increased interest rate will lead to lower inflation due to the increased cost of borrowing. And by Philips curve, lower inflation is associated with higher unemployment rate. This implies a positive relationship between interest rate and unemployment rate.

We consider the following three models.

- $\texttt{diff}_t \sim \mathbf{1} + \texttt{FEDFUNDS}_t$
- $\texttt{diff}_t \sim \mathbf{1} + \texttt{FEDFUNDS}_t + \texttt{GDP}_t + \texttt{CPIAUCSL}_t + \texttt{UNRATE}_t$

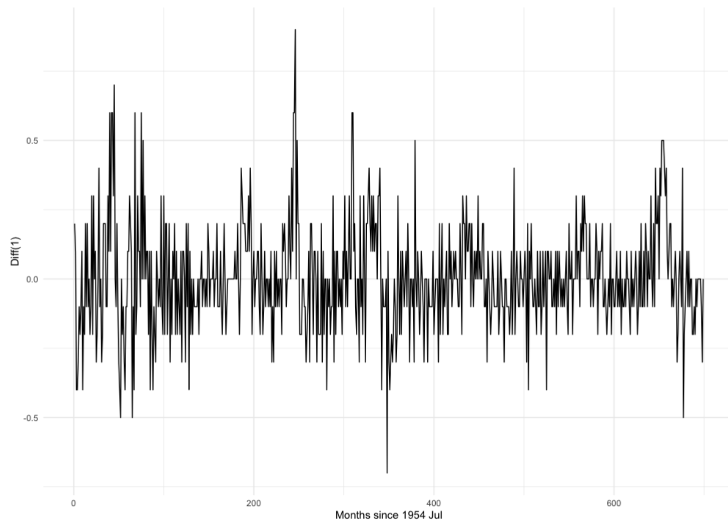

Figure 2: Plot of One-step Difference: diff(1)

- $\texttt{diff}_t \sim \mathbf{1} + \texttt{FEDFUNDS}_t + \texttt{GDP}_t + \texttt{CPIAUCSL}_t + \texttt{UNRATE}_t + \texttt{diff}_{t-1} + \texttt{diff}_{t-2} + \texttt{diff}_{t-3} + \texttt{diff}_{t-4} + \texttt{diff}_{t-12} + \texttt{diff}_{t-24}$

The purpose is to investigate the effect of `FEDFUNDS` on `diff`, while controlling for other variables. This allows us to isolate the effect of `FEDFUNDS`.

The results from all three regressions gave similar coefficients and standard errors (see Table 2). However, when the HC0 standard error is calculated, it is smaller in the model that included other control covariates. Firstly, this means that under the heteroskedastic assumption, `FEDFUNDS` has a more pronounced effect due to the smaller standard error (and hence, a relatively larger p-value). Another conclusion is that the linear model might be misspecified, resulting in the difference between the normal standard error and the HC0 standard error.

Model	Coefficient ($\times 10^3$)	SE ($\times 10^3$)	HC0 SE ($\times 10^3$)
1	7.259	4.094	4.602
2	8.198	4.920	2.745
3	7.872	5.147	2.938

Table 2: Coefficient and SE of `FEDFUNDS` Across Models

The results also show that `FEDFUNDS` is slightly less important than the lags (specifically `lag1`, `lag2`, `lag3`, `lag4`). This could possibly explain why many unemployment studies online focused on a univariate time series analysis. Perhaps the effects of other macroeconomic indictors take time to propagate to unemployment, thus the previous lags were sufficient covariates to explain most of the variance. Nonetheless, we keep the other covariates in later analysis, since

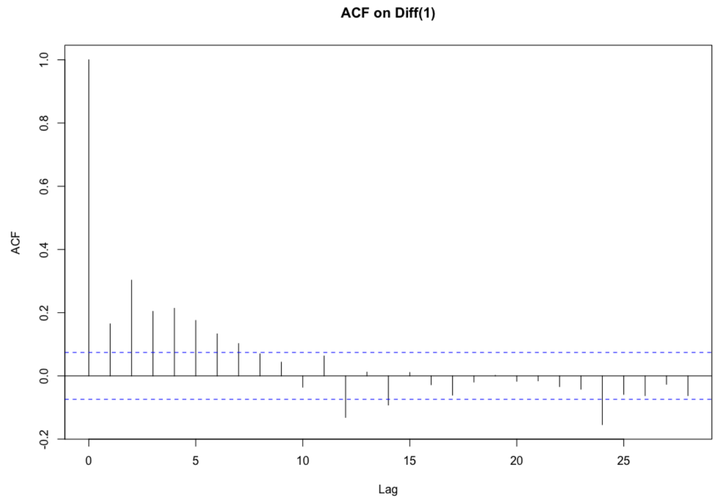

Figure 3: Autocorrelation Plot of diff(1)

we are focused on prediction.

4 Prediction

To perform prediction, I split the data into train (`n.train`=700) and test set. Unfortunately, the test set included the COVID-19 period, so I have in addition a truncated test set that cutoffs before the COVID-19 period. I still keep the full test set as a measure of robustness of the prediction method.

4.1 Baseline Models

We consider three baseline models. The first is the naive prediction of using just the previous unemployment rate to predict the next. The second is a naive linear prediction by propagating the difference. The last one is to use AR(4), motivated by the PACF plot. All of these baseline models have autoregressive flavour.

1. $\hat{\texttt{diff}}_t = 0$
2. $\hat{\texttt{diff}}_t = \texttt{UNRATE}_t - \texttt{UNRATE}_{t-1}$
3. $\hat{\texttt{diff}}_t \sim \text{AR}(4)$

Table 3 shows the baseline results. Figures 5, 6, 7 show the prediction plots. Surprising, we see that the most naive model actually achieved the lowest prediction error (0.027) in both pre-COVID-19 period and the total period. This

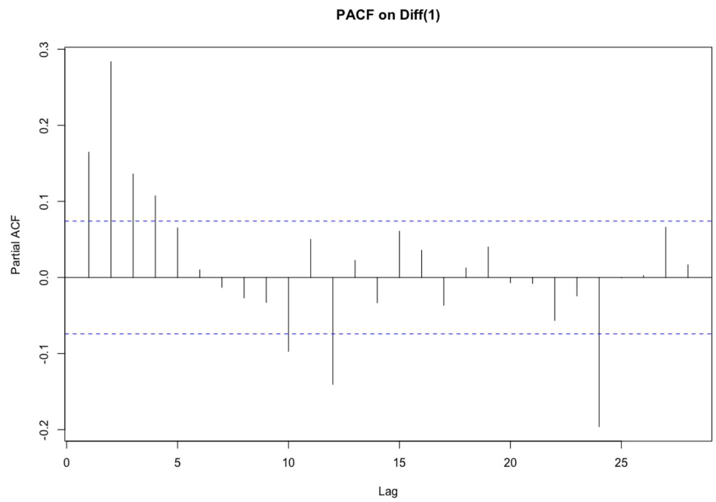

Figure 4: Partial Autocorrelation Plot of diff(1)

indicates that having only a univariate time series analysis may result in higher prediction error.

Model	MSE (with COVID-19)	MSE (w/o COVID-19)
1	0.931	0.027
2	1.843	0.057
3	1.089	0.031

Table 3: Coefficient and SE of `FEDFUNDS` Across Models

4.2 Ridge Regression

Next, I performed ridge regression using the following two models.

- $\hat{\texttt{diff}}_t \sim \mathbf{1} + \texttt{FEDFUNDS}_t + \texttt{GDP}_t + \texttt{CPIAUCSL}_t + \texttt{UNRATE}_t$
- $\hat{\texttt{diff}}_t \sim \mathbf{1} + \texttt{FEDFUNDS}_t + \texttt{GDP}_t + \texttt{CPIAUCSL}_t + \texttt{UNRATE}_t + \texttt{diff}_{t-1} + \texttt{diff}_{t-2} + \texttt{diff}_{t-3} + \texttt{diff}_{t-4} + \texttt{diff}_{t-12} + \texttt{diff}_{t-24}$

Table 4 shows the ridge coefficients in the two models. We see that after including the past lags, the coefficients of previous covariates had large changes, especially `UNRATE` and `GDP` which underwent a change in magnitude. This highlights the importance of including in the lags as covariates.

Table 5 shows the optimal ridge hyperparameters and the prediction error of both methods.

We see that with a lower prediction error of 0.02651, we have beat the baseline

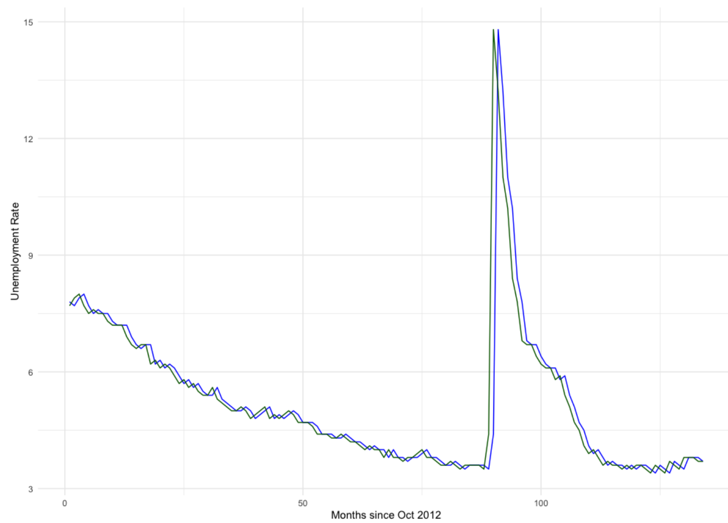

Figure 5: Baseline Model 1: $\hat{\texttt{diff}}_t = 0$

models with the second ridge regression model. Furthermore, the coefficient of FEDFUNDS is 8.26×10^{-3}, which is slightly larger than the values in Section 3. This could mean that the effect of interest rate on unemployment is higher than as indicated by the baseline models.

4.3 Indicator for Regime Switches

This idea is motivated by the two different kinds of behaviour, specifically the steep rise and gentle fall, shown by the U.S. unemployment data. The method of indicators has been used ubiquitously in linear regression, including in change-point and change-slope problems. The use of indicator for regime switches extends those ideas.

One desired property of one-step ahead prediction is how fast it can react to regime changes. Therefore, a suitable detection via indicator is of the form $\mathbb{1}\{\texttt{UNRATE}_t - \texttt{UNRATE}_{t-c} > \lambda\}$ where $c > 0$ and λ are tuning parameters. This effectively specifies two regimes, where $\texttt{diff}_t$ is fitted to different linear models depending on the regime.

The selection of λ and c were done using a grid search. For each pair of (λ, c), a ridge regression with the optimal λ_{ridge} is fitted. Then, (λ^*, c^*) is selected to be the pair that minimised the training error. It turns out that $(\lambda^*, c^*) = (0.1, 5)$. Figure 8 shows the classification of each time point into the regimes. Visually, the classification proceeded well - the upslopes and downslopes were mostly classified correctly.

Figure 9 shows the prediction on the test set as well as the regime predictions.

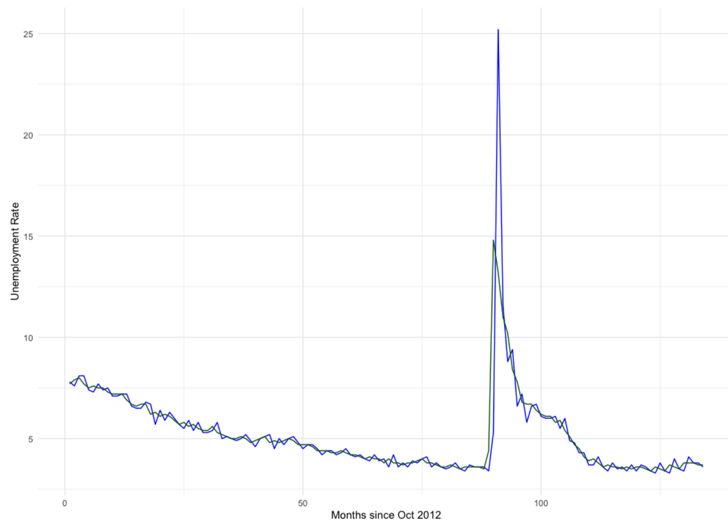

Figure 6: Baseline Model 2: $\hat{\texttt{diff}}_t = \texttt{UNRATE}_t - \texttt{UNRATE}_{t-1}$

We see that the classification ability of the indicator regime regression on the training set carried over to the test set. During downslope periods, the regime stayed mostly at 2. The regime prediction quickly changed to 1 during the COVID-19 period and swiftly went back to 2 in the following downslope.

Table 6 shows the ridge regression hyperparameters of each regime as well as the train and test errors. The train and test errors were calculated by a weighted average over both regimes. The indicator regime regression method beats all three baseline predictions as well as naive ridge regression, achieving a test error of 0.0245.

In conclusion, indicator regime regression achieved the lowest prediction error thus far. In my opinion, this model exploited the structure of the unemployment data, relying on the fact that it has two visually distinct regimes. Also, the choice of $c = 5$ helps to reduce the noise when deciding regime changes - this is because two data points which are very close to each other may be more susceptible to noise, which lowers their signal in dictating regime changes.

5 Additional Models

5.1 Switching Autoregressive Model

Lastly, we investigate the switching autoregressive model, which is a methodology proposed by Hamilton[9] and exposited in Shumway and Stoffer[10]. This model assumes a latent process $(X_t)_t$ which indicates the regime. To model the U.S. unemployment rate, a natural choice is a two-state regime i.e. take

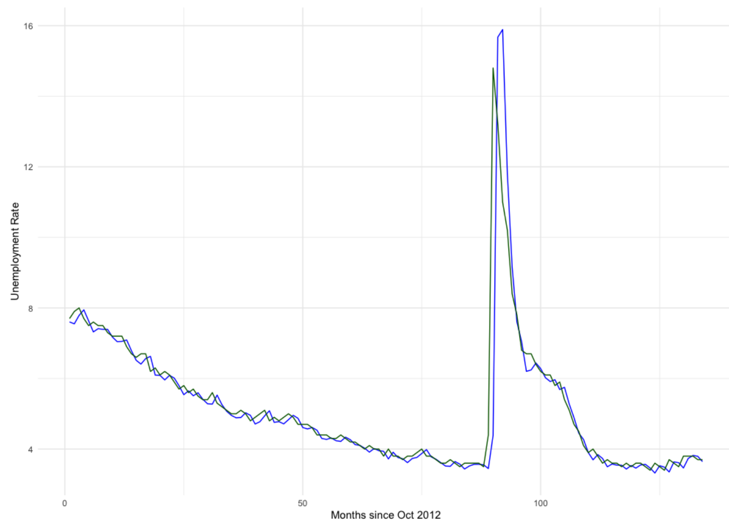

Figure 7: Baseline Model 3: $\hat{\texttt{diff}}_t \sim \text{AR}(4)$

$X_t \in \{1, 2\}\, \forall t$. The latent process $(X_t)_t$ evolves by a transition matrix. Each regime correspond to a linear autoregressive model of $\texttt{diff}_t$. The fitting of parameters and coefficients are done by the maximum likelihood principle. The R package `MSwM`[11] is used for this method.

With the switching autoregressive model, I tried the following three sets of covariates.

1. $\hat{\texttt{diff}}_t \sim \mathbf{1} + \texttt{diff}_{t-1}$
2. $\hat{\texttt{diff}}_t \sim \mathbf{1} + \texttt{diff}_{t-1} + \texttt{diff}_{t-2} + \texttt{diff}_{t-3} + \texttt{diff}_{t-4}$
3. $\hat{\texttt{diff}}_t \sim \mathbf{1} + \texttt{FEDFUNDS}_t + \texttt{GDP}_t + \texttt{CPIAUCSL}_t + \texttt{UNRATE}_t + \texttt{diff}_{t-1} + \texttt{diff}_{t-2} + \texttt{diff}_{t-3} + \texttt{diff}_{t-4} + \texttt{diff}_{t-12} + \texttt{diff}_{t-24}$

Figures 10 and 11 show the switching between regimes as the time series progresses. In the difference plots, the regimes seem to correspond to regions of lower and higher variances. When projected into the U.S. unemployment graph (13, 14, 15), we see that those regions correspond to the upslopes and downslopes. From Figure 12, we also note that this model requires a choice of threshold to determine which state the system is in. In what follows, I picked the threshold to be 0.5 arbitrarily.

Prediction is again performed one step at a time. The test set is processed sequentially. Once a prediction occurs and the true diff_t is revealed, the entire switching autoregressive model is refitted to the new data set in order to perform the next prediction.

Table 7 shows the results of the various switching autoregressive models. We

	Model 1	Model 2
(Intercept)	6.83×10^{-3}	7.11×10^{-2}
`UNRATE`	-7.23×10^{-3}	5.51×10^{-2}
`FEDFUNDS`	9.39×10^{-3}	8.26×10^{-3}
`GDP`	1.50×10^{-5}	6.93×10^{-6}
`CPIAUCSL`	-8.68×10^{-4}	-2.73×10^{-4}
`lag1`		5.54×10^{-2}
`lag2`		-1.09×10^{-2}
`lag3`		-3.17×10^{-2}
`lag4`		-6.85×10^{-2}
`lag12`		-1.42×10^{-2}
`lag24`		-5.66×10^{-3}

Table 4: Train and Test Error of Ridge Regression

Model	λ^*	Train Error	Test Error
1	0.002126	0.03547	0.03039
2	0.004144	0.03069	0.02651

Table 5: Train and Test Error of Ridge Regression

see that all three models beat the other models we have experimented so far. Even Model 1, with $\texttt{diff}_{t-1}$ as its sole covariate, bested the indicator regime regression. Notably, Model 3, which included all covariates, did not achieve the lowest prediction error. This may constitute as more favourable evidence for a univariate time series analysis of U.S. unemployment data.

Figures 13, 14, 15 show the regime changes and the predictions made by the three methods. We see that in Model 3, there is more switching effects happening in the train set, which may be interpreted as overfitting.

A possible reason why the switching autoregressive model performed better than the indicator regime model is due to the underlying Markov Chain structure which is used to model the state. Maximum likelihood penalises frequently switching between the states. In Figure 9, it is shown that the indicator regime regression still had some switching of states when applied to the test set. This is not seen in the switching autoregressive model.

6 Conclusion

In conclusion, interest rate is a relatively important covariate for predicting future unemployment rate. However, it pales in effect when compared to the lags.

This project shows that prediction error can be improved via ridge regression.

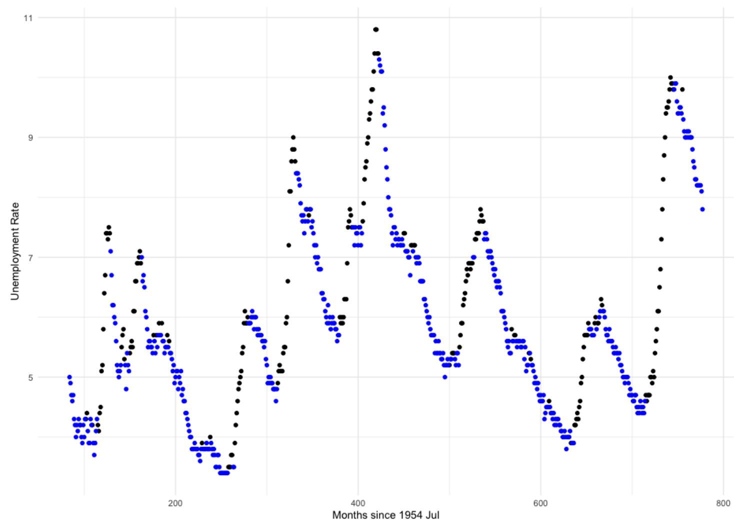

Figure 8: Plot of U.S. Unemployment of Each Regime

Model	$(\lambda_1^*, \lambda_2^*)$	Train Error	Test Error
1	(0.007516, 0.004144)	0.02921	0.02454

Table 6: Train and Test Error of Indicator Regime Regression

Model	Test Error
1	0.0245
2	0.0231
3	0.0240

Table 7: Train and Test Error of Switching Autoregressive Model

Furthermore, prediction error can be significantly improved from baseline models by exploiting the structure of the data. In the U.S. unemployment rate dataset, the visually distinct regimes motivated the modelling of the regimes, through indicator regime regression and switching autoregressive model. Under these two models, the investigation of the covariates show that a univariate time series analysis is justifiable, since the models with only the lags as their covariates performed almost equivalently, if not better, than the full model.

Although this project seemed to focus on time series data, many methodologies still fell in the realm of linear models, hence relating this project back to class.

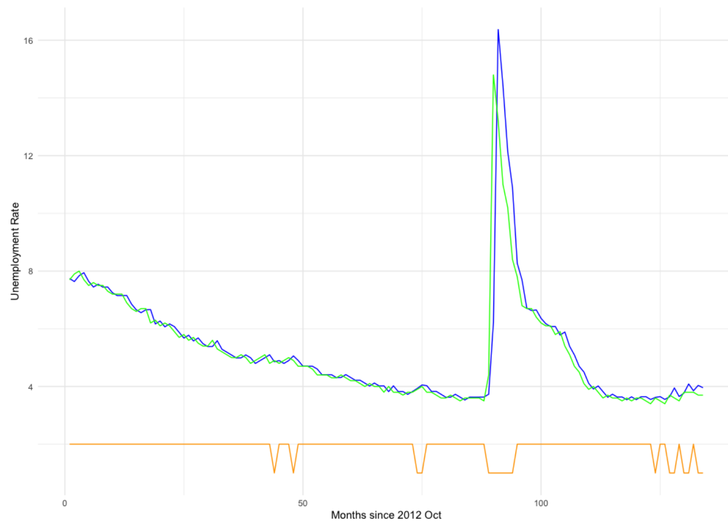

Figure 9: Plot of Indicator Regime Regression Prediction (regime in orange)

References

[1] Stephen M. Goldfeld and Richard E. Quandt. A markov model for switching regressions. *Journal of Econometrics*, 1(1):3–15, 1973.

[2] M McGrane. A markov-switching model of the unemployment rate. *Congressional Budget Office*, 2022.

[3] M Perlin. Estimation, simulation and forecasting of a markov switching regression. *Annals of Mathematics*, 54(2):286–295, 2007.

[4] W Yu. What explains the variation in unemployment rates across the united states? *UCLA Anderson Forecast*, 2022.

[5] U.S. Bureau of Labor Statistics. Unemployment rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis, April 23 2024.

[6] Board of Governors of the Federal Reserve System (US). Federal funds effective rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis, April 23 2024.

[7] U.S. Bureau of Economic Analysis. Gross domestic product [GDP], retrieved from FRED, Federal Reserve Bank of St. Louis, April 23 2024.

[8] U.S. Bureau of Labor Statistics. Consumer price index for all urban consumers: All items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis, April 23 2024.

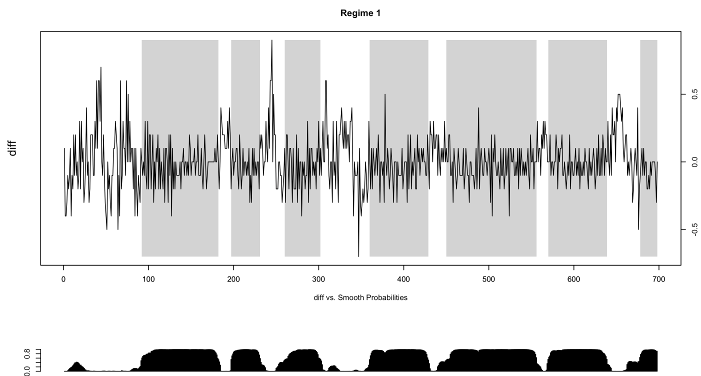

Figure 10: Plot of Regime 1 in the Switching Autoregressive Model

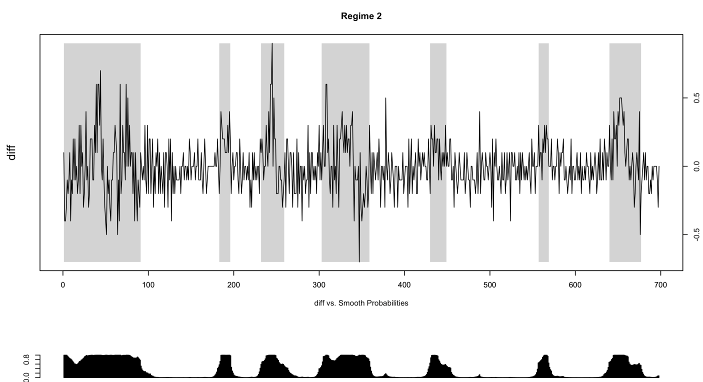

Figure 11: Plot of Regime 2 in the Switching Autoregressive Model

[9] James D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.

[10] Robert H. Shumway and David S. Stoffer. *Time Series Analysis and Its Applications*. Springer, 2000.

[11] Alberto Lopez-Moreno Josep A. Sanchez-Espigares. MSwM: Fitting markov switching models univariate autoregressive markov switching models for linear and generalized models.

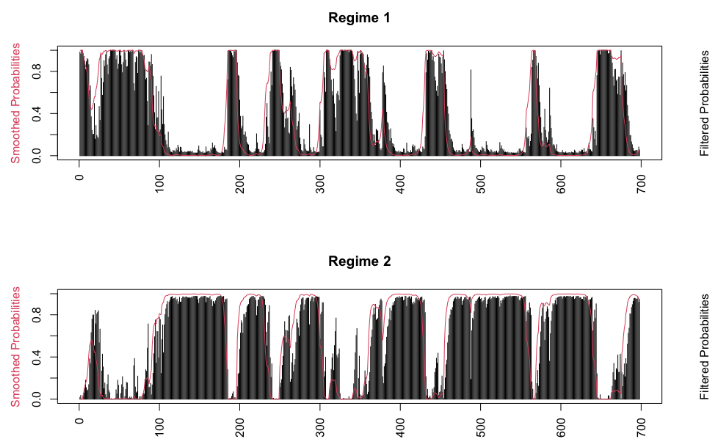

Figure 12: Plot of Regime Changes and the Threshold in the Switching Autoregressive Model

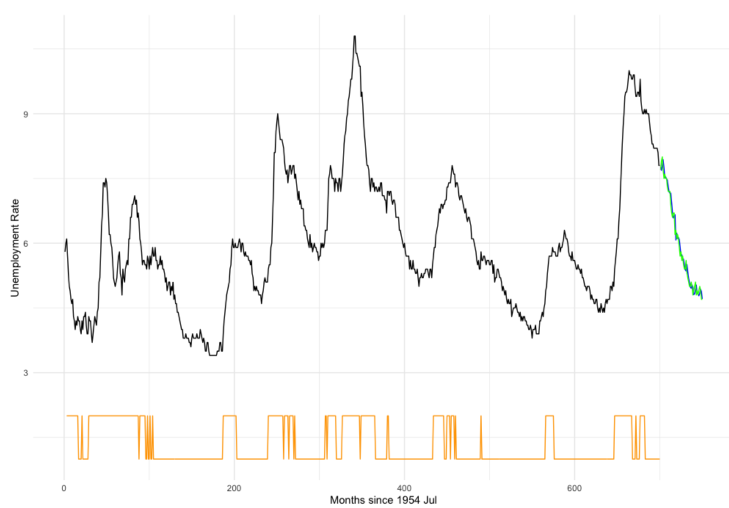

Figure 13: Plot of Regime Changes in the Switching Autoregressive Model

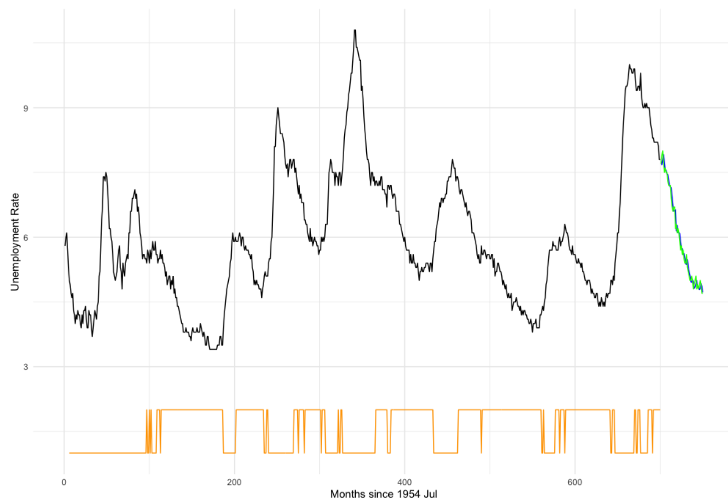

Figure 14: Plot of Regime Changes in the Switching Autoregressive Model

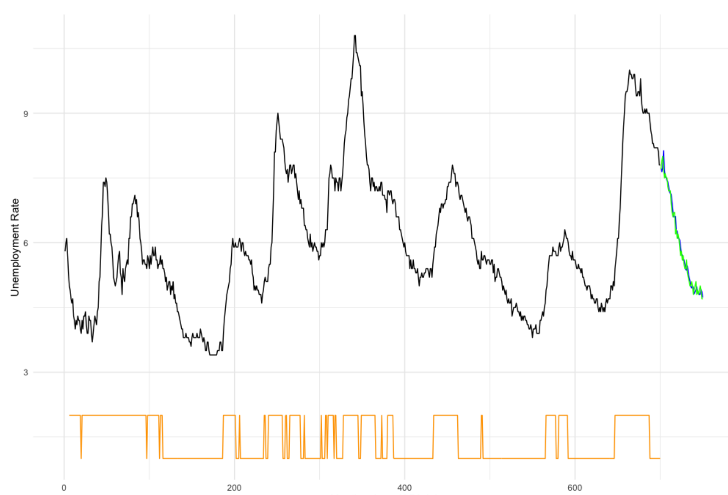

Figure 15: Plot of Regime Changes in the Switching Autoregressive Model