

# Modeling and Forecasting Trend, Seasonality, and Cycles

Arvind Ramachandran

May 31 2019

## 1 Introduction

In this project, we will model the non-adjusted monthly unemployment rates in California and New York from January 1976 to April 2019. The time series data was downloaded from the website of the Federal Reserve Bank of St. Louis.

We will use ARIMA models on the individual time series to product a model and forecast unemployment rates into the future in both California and New York. We will then use a vector auto-regressive model to create a joint model wherein the unemployment rates for the states are dependent on past unemployment in the same state as well as in the other state.

## 2 Results

(a) The time series plots for unemployment in California and New York are shown below, along with the ACF and PACF plots for both states. The partial autocorrelation plots for both states spike at values 1 and 12, suggesting that a seasonal (one-year) and lag-1 autoregressive component. We will thus fit both time series to an seasonal ARIMA  $(1,0,0)$ ,  $(0,0,12)$  model.

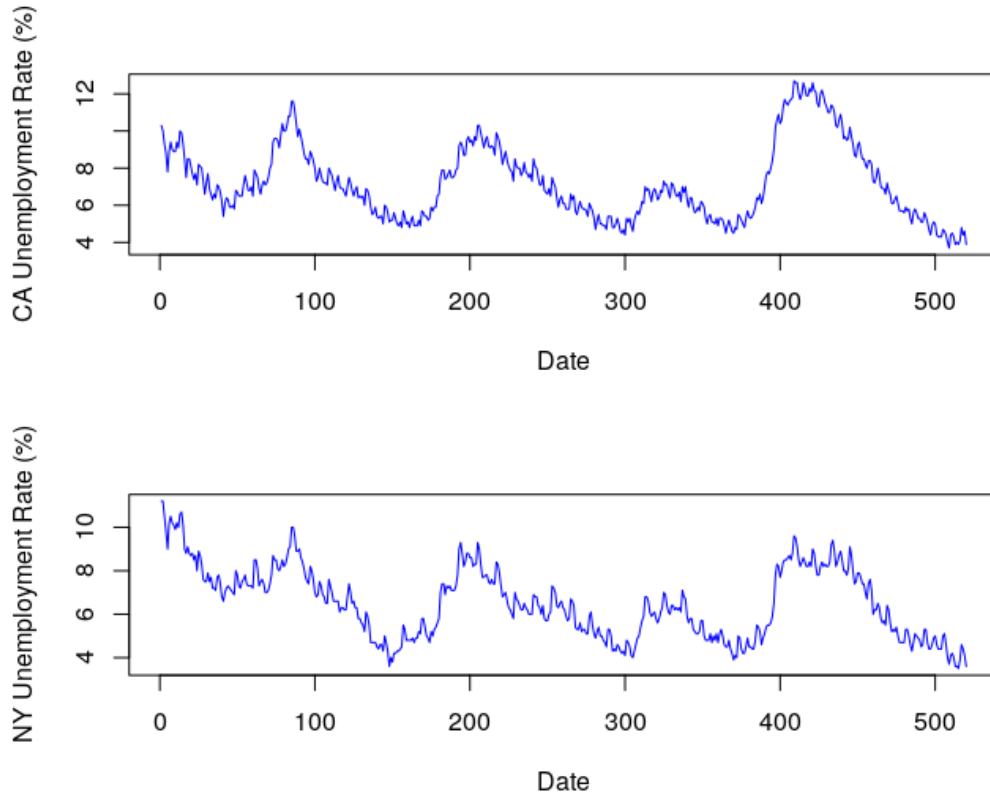


Figure 1: California and New York Unemployment Rates

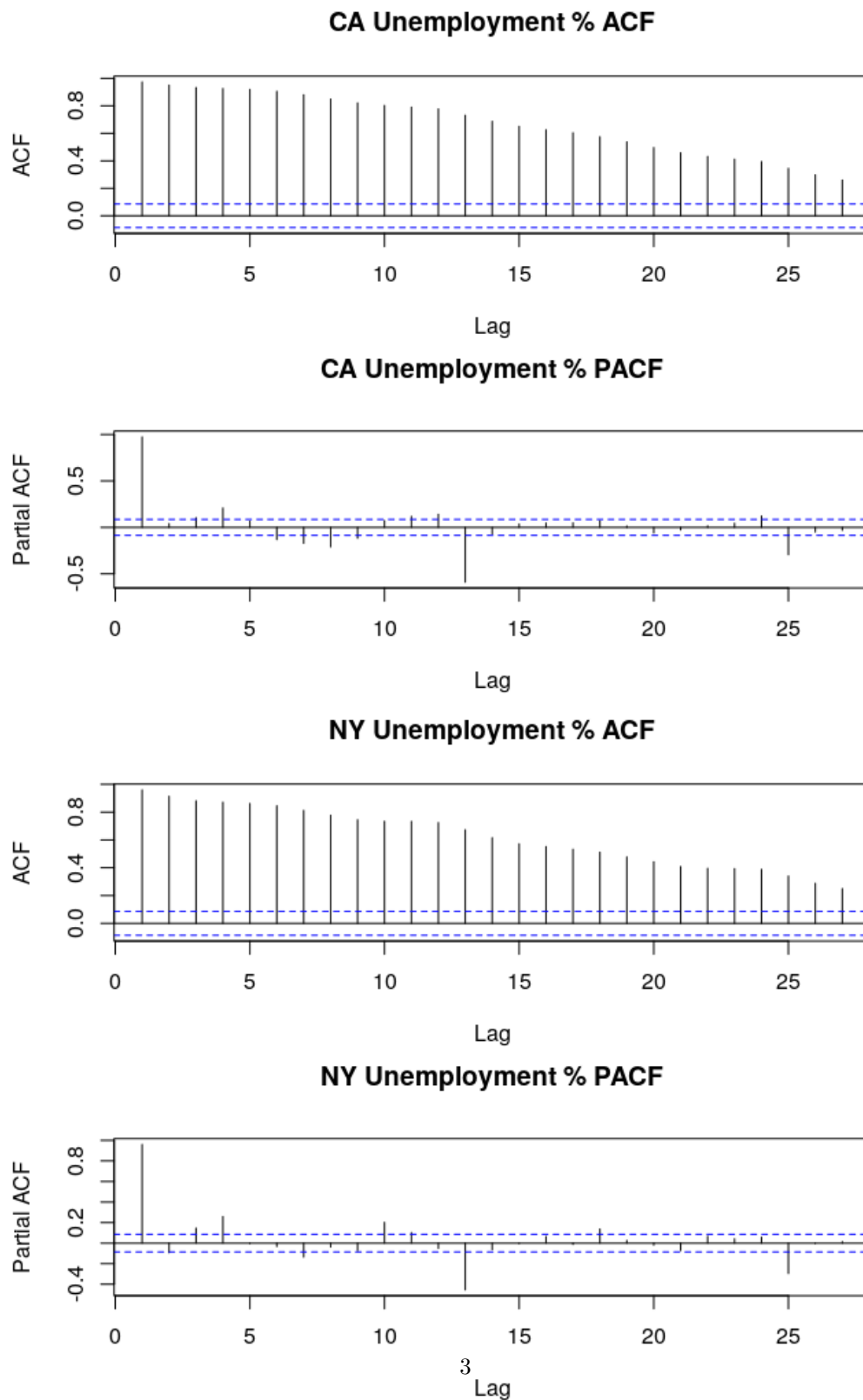


Figure 2: ACFs and PACFs

(b & c) The fit from the ARIMA (1,0,0), (0,0,12) model is shown below, along with the residuals. We notice that the model does a very good job fitting the time series, although it has not accounted for all the seasonality. I compared these with the best-fit ARIMA models, but the auto.arima models do not produce a much better fit.

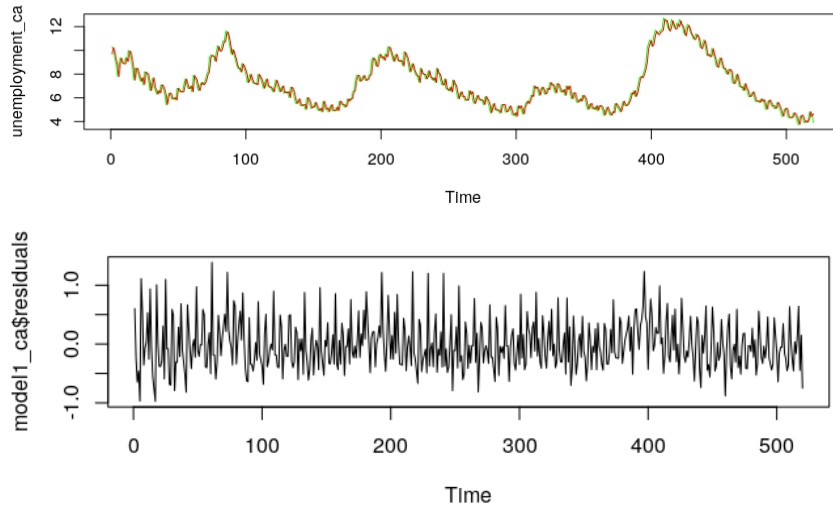


Figure 3: California ARIMA Model - In the first plot, the green line is the original time series and the red line is the model fit. The second plot shows the model residuals.

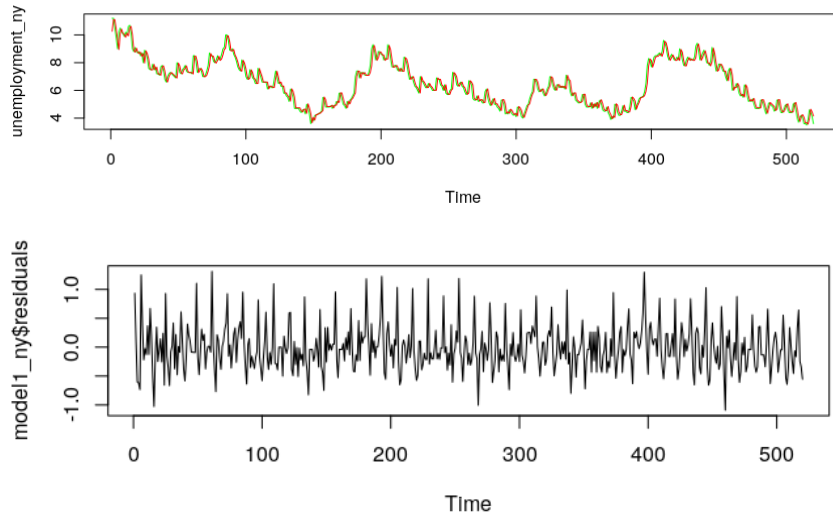


Figure 4: New York ARIMA Model - In the first plot, the green line is the original time series and the red line is the model fit. The second plot shows the model residuals.

(d) The ACF and PACF plots of the residuals show that there is some remaining autoregressive component, which indicates (as we said above) that the model has not fully fit the seasonality of the unemployment time series.

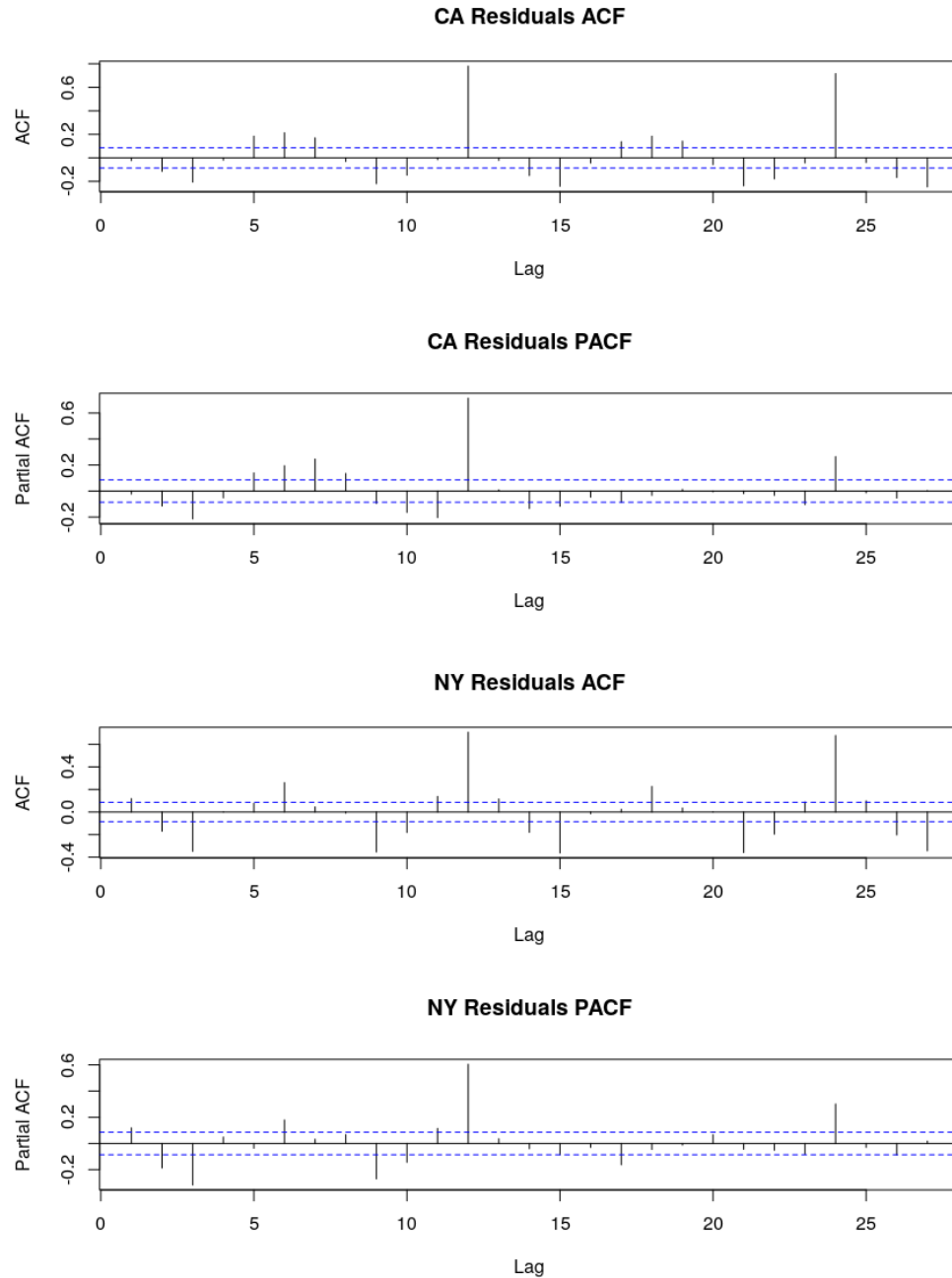


Figure 5: ARIMA Residual ACF and PACF for both states

(e) The CUSUM for both ARIMA models are shown below. The CUSUM for our model for California indicates that while the model generally does a good job, it struggles at certain points in time, around  $t = 100$ , 200 and 400-470, which correspond to the years.

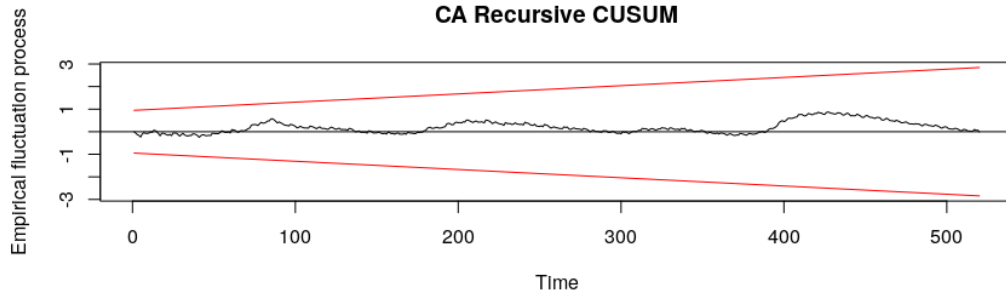


Figure 6: California ARIMA Model CUSUM

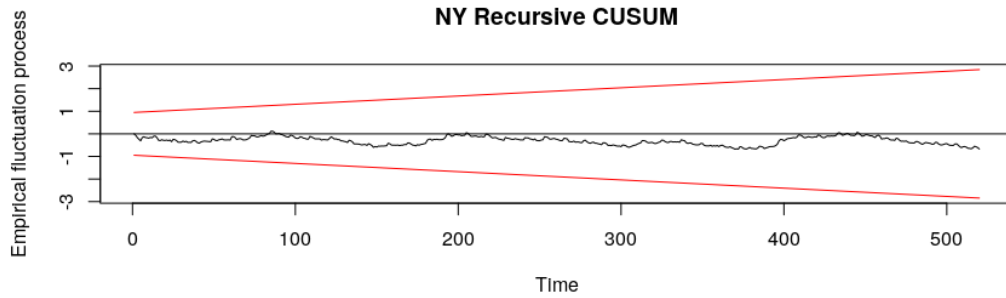


Figure 7: New York ARIMA Model CUSUM

The CUSUM for the New York model also fits well, but has trouble with the time periods  $t = 150$ , 300 and 350, which correspond to .

(f) The recursive residuals are show below:

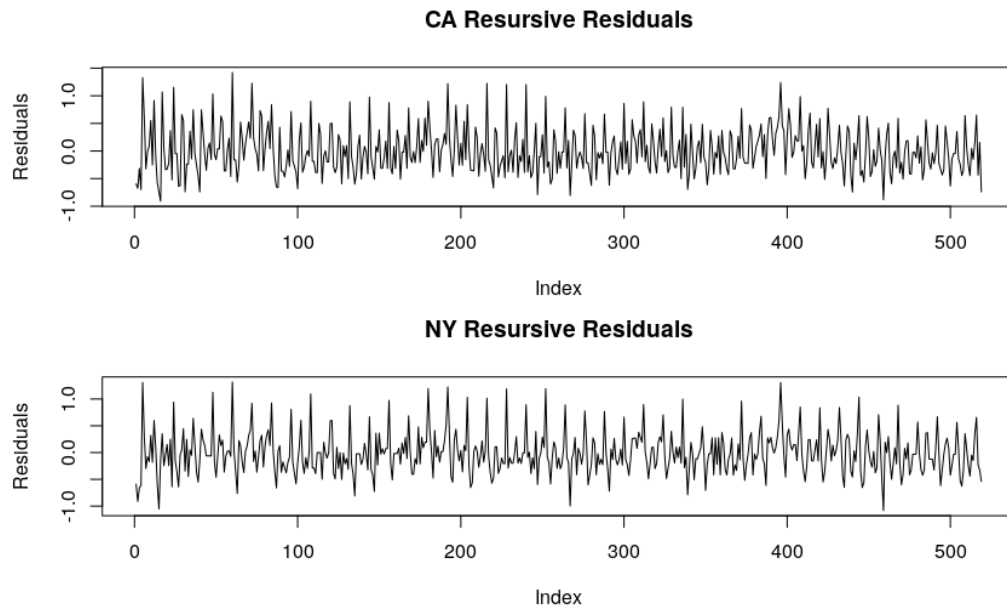


Figure 8: ARIMA Recursive Residuals

(g)

(h) The 12-month forecast and the error bands are as follows.

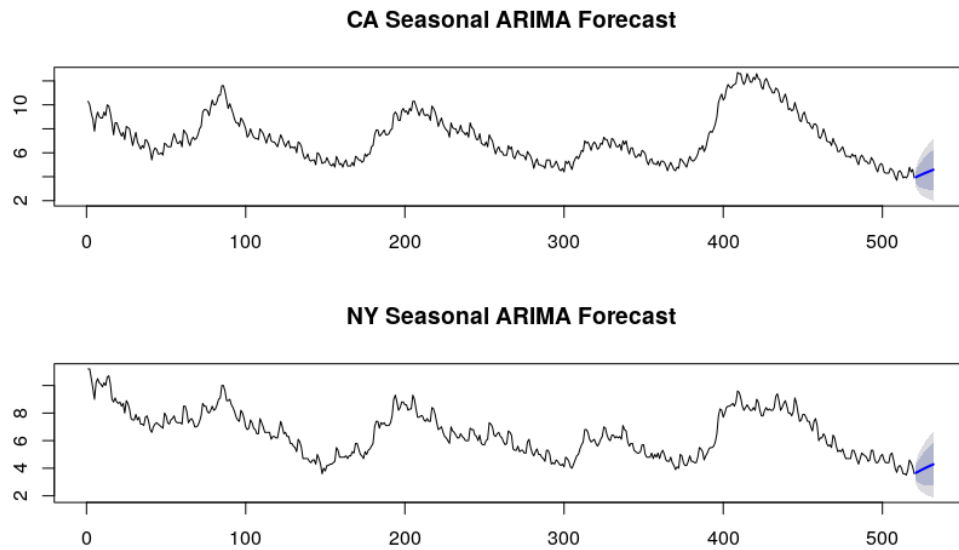


Figure 9: 12-month forecasts

(i) We observe the cross-correlation function between the unemployment time series for both states. The CCF shows that there is a auto-regressive cross-dependence between unemployment in the two states. We fit a vector auto-regressive (VAR) model with lag order 4 to the data. The model fits and residuals are shown below.



Diagram of fit and residuals for unemployment\_ca

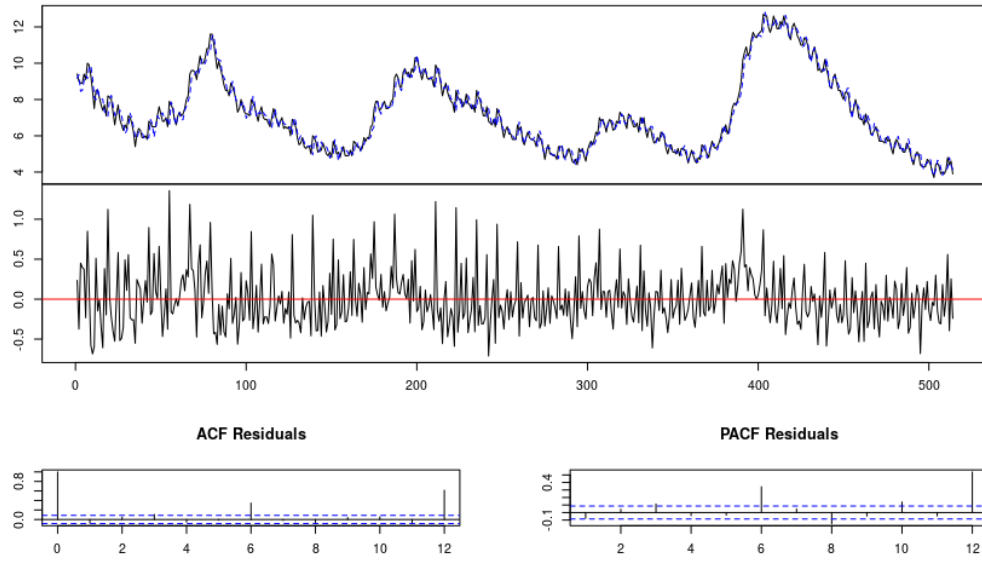


Diagram of fit and residuals for unemployment\_ny

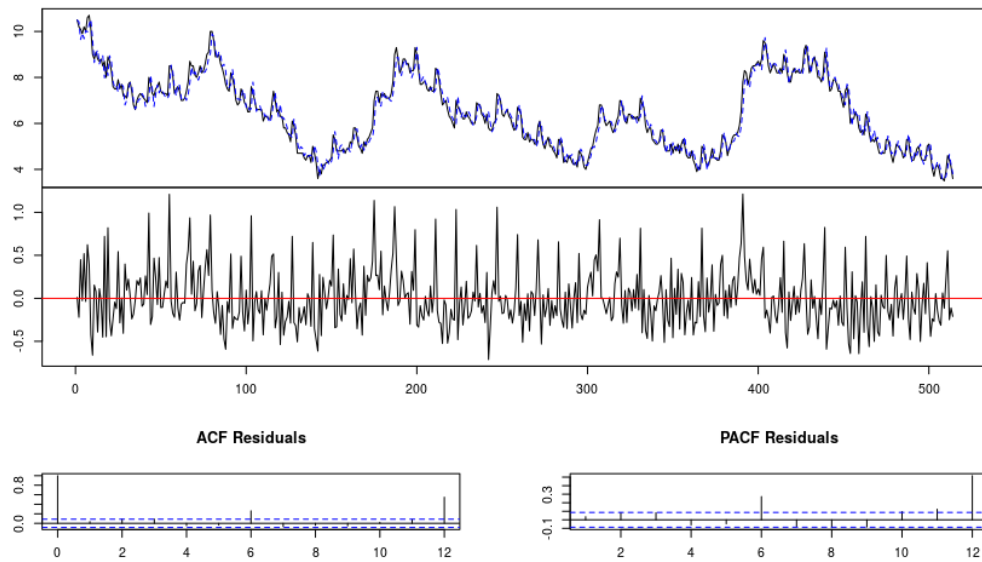


Figure 10: Vector Auto-Regression Model Fit and Residuals

(j)

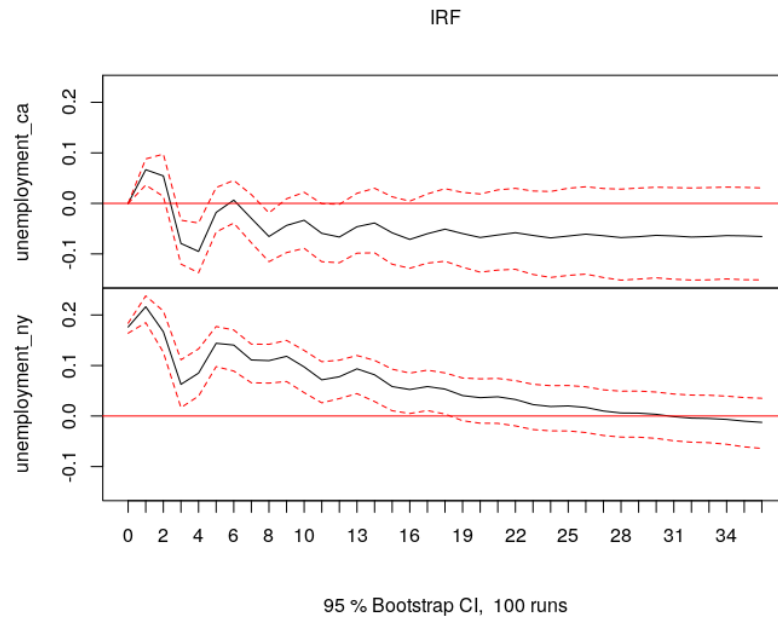


Figure 11: Vector Auto-Regression Impulse Response

(k)

```
> grangertest(unemployment_ny ~ unemployment_ca, order = 6)
Granger causality test

Model 1: unemployment_ny ~ Lags(unemployment_ny, 1:6) + Lags(unemployment_ca, 1:6)
Model 2: unemployment_ny ~ Lags(unemployment_ny, 1:6)
      Res.Df Df    F      Pr(>F)
1         501
2         507 -6 9.299 1.115e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 12: Granger-Casuality Test Results

(l)

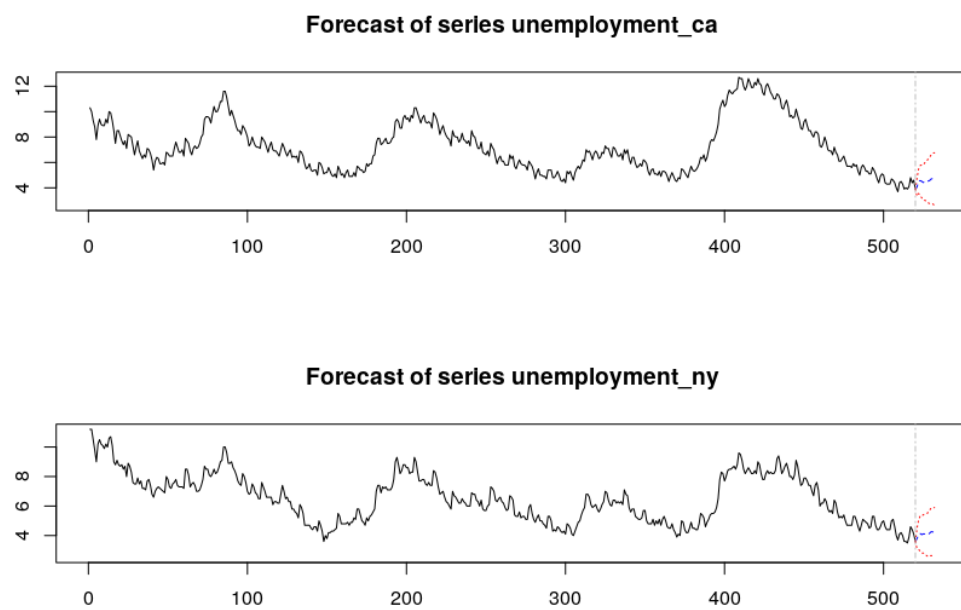


Figure 13: Vector Auto-Regression 12-month ahead forecast