Eloka Obi
Applied Time Series Analysis
5/10/2020
Barut

Final Project for Stats 4181

Intro:

This final report is centered on what might affect GDP as a percentage change for a particular region or place. The data that was derived was from the federal reserve bank of St.Louis. We are trying to find how variables like unemployment and CPI(Consumer Price Index) affect GDP percentage change growth. CPI is virtually the same type of measurement as GDP. CPI is a measure that examines the weighted average of prices of a basket of consumer goods and services, such as transportation, food, and medical care. These two-time series were integral and significant in our model determining how strong and significant our model was with our response variable. Working in time series analysis, our model must be manipulated in order to accurately assess how our model is autocorrelative and partially autocorrelative. Our dataset was reduced down to four variables that we assessed would be compatible with a Generalized Least Squares model worked with throughout this entire analysis. This dataset required minor cleaning as the dataset was already very manageable to work with at first. I went through multiple steps in order to find the best model for our analysis in accordance with the project. A strenuous effort was put forth and a great deal of consulting in trying to create a model that would contain many values for the coefficients and truly valid p-values for our model. I hypothesized that the GDP would increase with a decrease in unemployment and an increase in CPI. Economic theory states that with economic growth we see higher growth rates in our GDP with lower rates of unemployment and growth in CPI. Because the model is differenced, we expect there to be results for our model that explain the differenced variables for our model.

Output/Results:

A lag was attached in the creation of our variables found that with our variables that there was a significant correlation with our ACF points falling within our 95% significance level. Some of our points with our ACF for unemployment did lie outside of the 95% confidence interval. The difference variables gave us different outputs for our autocorrelation This was the same with our PACF as well.

An Arima function was used for our fixed coefficients for our residuals in order to find the right model for the residuals. The dynlm is used for fitting linear regressions, like lm, except L is used to control the lag. An ARIMA function was used in order to find significant coefficients in the residuals of our linear model. Using an ARIMA with these parameters can make our coefficients all negative or related to the moving average. From our AIC (Akaike information criterion) we obtain an understanding of how significant our ARIMA model is. Our AIC for our first one is 795.83. The next ARIMA model has an AIC of 795.25. The next two ARIMA models are the same as the previous two models but with higher AICs of 861 plus compared to our previous models that have AICs of less than 800. Our residuals for this model are much higher than the past two ARIMA models because the last two are different variables and not the original variables.

This was the best GLM model that was found between our different models and our original model with original variables. The Generalized Least Square model is best used for estimating the unknown parameters in a model when they are highly correlated with the residuals. We see that all of our values for our coefficient are negative in our difference model except GDP, which is our response variable. Our GDP as a percent change is measured differently from our regular GDP. This is why we have negative observations in our model. It seems from our model that both variables of unemployment and CPI are negative and are significant in their values. When unemployment decreases, we expect our Gross Domestic Product to increase. This is not necessarily the same as the consumer price index. These p-values that are shown are significant and are not too high for us to cancel out. Both of our p-values are higher than the 0.05 threshold which means both of the variables are significant to our model. It was found that our response variable did not pass the 0.05 threshold. So it seems that

For our original model, we see the same effect of the GDPs coefficient being positive and the unemployment and CPI being negative. It doesn't seem like our model changes that much. Our p-values have virtually the same ballpark with our difference model meaning that our whole model is significant.

Since we find a high correlation between our differenced variables, according to AIC theory, we found that this model would be the best for giving us the best explanation of how our variables are correlated, so we can understand its impacts on GDP. The coefficients of our model for our variables tell the truth on the impact on the response variable for our Generalized Least Squares model.

Conclusion:

It was found that our response variable did not pass the 0.05 threshold. So it seems that there is not a relationship between our response variable and the rest of the model. So it seems that the percent change of GDP truly does not have virtually any type of effect with unemployment and the CPI. We found that the Generalized Least Square of our different model was the best model to use according to our AIC and the significance of our coefficients

```
unemployment <- as.vector(datagdp12$UNRATE)
unemployment1 <- diff(unemployment)
gdp <- as.vector(datagdp12$GDPC1_PCH)
gdp1 <- diff(gdp)
cpi <- as.vector(datagdp12$CPIAUCSL_PCH)
cpi1 <- diff(cpi)
diff_data <- data.frame(unemployment1,gdp1,cpi1)
datagdp$id <- seq(1,292,1)
diff_data$id<- seq(2,292,1)
final <- merge(datagdp12,diff_data)
plot(acf(unemployment1))
plot(acf(unemployment))
plot(acf(gdp1))
plot(acf(gdp))
plot(acf(cpi))
plot(acf(cpi1))
plot(acf(diff_data))
plot(pacf(unemployment1))
plot(pacf(unemployment))
plot(pacf(gdp))
plot(pacf(gdp1))
plot(pacf(cpi))
plot(pacf(cpi1))
plot(pacf(diff_data))
ccf(unemployment,gdp, lag=1)
ccf(cpi,gdp)
ccf(unemployment1,gdp1)
ccf(cpi1, gdp1)
```

```
ccf(gdp,gdp1)
ccf(unemployment,unemployment1)
ccf(cpi,cpi1)
ccf(gdp1,unemployment)
ccf(gdp1,cpi)
C.u ~ L(C.e,5) + L(C.e,0)
install.packages("dynlm")
dfm <- (datagap12 ~ L(datagap12, 5) + L(datagap12, 0))
(acf(dfm))
```