BEEM012 – Homework Exercises: Week 1

Julian Dyer

Note - This assignment is not going to be graded

1 Manipulating Time Series Data

The first thing we will do is learn some basic tools to manipulate time series data. We will learn how to load time series data and put it into a format where we can easily take lags and calculate growth rates, as well as plotting time series.

1.1 Install Libraries

The first step is to make sure we have all the libraries needed to do time series analysis. Run the code under Section 1 in the R script to install the following models:

- `AER` – This contains a number of useful sample datasets and utilities from Hanck et al. (*note:* If you're having trouble installing this package, check your version of R is at least R 3.5)
- `readxl` – this will be helpful for handling data in Excel format
- `dynlm` – This is a useful package for running time series regressions
- `xts` – This package has useful tools for manipulating timereis data and creating lagged variables.
- `zoo` – This package has nice tools for plotting Ordered Observations, and especially time series.

Once you have run this set of lines once, you can then comment them out so it doesn't re-install these modules every time you run your code.

1.2 Load UK Macroeconomic Data

The next step is to load our macroeconomic time series data from the UK. For now, we will start with just the GDP time series. The first thing we need to do is tell R where our files are located, then load the xls file of data. To use our time series tools we need to make sure that R can interpret the data as a time series, and this means reformatting our data so that the "Date" column is interpreted so that the Years, Months, Quarters, etc. are properly understood. This requires looking at the data you are using and adjusting the format. Here, our dates are stored in the format `1960 Q2`. Can you use this to understand why we use `format = ''%Y Q%q''`?

Now we will plot our time series, to get some idea of the patterns in our GDP data. Here we are looking at quarterly GDP per capita in the UK, taking into account a measure of inflation. First, let's just look at the plots of quarterly GDP per capita over time, as well as the log of this value over time.

1.3 Generating Lags and Differences

As you know, lags and changes over time are going to be crucial for time series econometrics. We will therefore use the `lag()` and `diff()` operators to generate lags and differences, and then plot them to check our understanding.

2 Computing the Autocorrelation and Autocovariance

Now that we know how to generate lags, let's now compute a sample estimate of autocorrelation for quarterly GDP growth from 1990 to 2005. First we're going to compute these by hand, then we will use a function that calculates them in one step.

We will use the following sample approximation:

$$\hat{\rho}_j = \frac{\widehat{\text{cov}(Y_t, Y_{t-j})}}{\widehat{\text{Var}(Y_t)}} \tag{1}$$

Now, recall that if we take the 1990 to 2005 subsample of GDP, and then take lags, we will have to drop the first observation of GDP growth, because we cannot observe the lag for the first period where we compute GDP growth. Next, we check that our manually computed estimate is consistent with that computed in one convenient step by the `acf()` tool, which can compute autocorrelations for

multiple lags. What can you say from looking at these autocorrelations? Do they suggest that past GDP growth is useful for predicting future growth?

Having computed the autocorrelation of GDP growth from 1990-2005, compute this again for periods 1965-1990 and from 2005-2020. What can you say about the strength of the autocorrelations during these periods? During what time period did past GDP growth most strongly predict future growth?

3 Autoregressions

So now that we've looked at serial correlation, we will now explore these patterns in a regression framework. To do this, we first need to set up our data.

3.1 First-order Autoregression: Manual Lags

As you recall, we can't use the first observation in the time series in an autoregression, so we drop this item from the observations for the growth rate. Similarly, we drop the final observation fromis data, we estimate a linear model using the `lm()` model and test coefficients. We first do this with the first-order autoregression. Given the magnitude and sign of the first-order autocorrelation, what are you expecting for this regression coefficient?

3.2 Autoregressions using the `lag()` operator

We can also use the `lag()` operator to do this easily in one line. Try using this to estimate the first order autoregression and check that it gives the same coefficient as above. Once you have done this, modify the code to run the AR(2) and AR(3) models. For how many quarters do the lags have (individually) statistically significant effects on GDP growth?

We can also use the `ar.ols()` command to compute our autoregressions automatically. For now, we will avoid using this module as it is more of a 'black-box' that automatically selects the best options for you, and now we want to ensure you understand all the moving parts when estimating autoregressive models.

4 Forecasting and Forecast Error

Now that we know how to estimate linear autoregressive models, we can try forecasting to future data and see how they perform. Recall that to predict the outcome in period Y_{T+1} we use the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$:

$$Y_{T+1|T} = \hat{\beta}_0 \cdot 1 + \hat{\beta}_1 \cdot Y_T$$

Modify the code I provided for forecasting with the AR(1) model, and now use the AR(2) and AR(3) models to forecast the GDP growth in Q1 of 2006. Which model is the most accurate?

5 Autoregressive Distributed Lag Model

We can also introduce other potentially lagged variables to predict our variable of interest. Now, we are going to add information on yield spreads (the difference between long-term securities and short-term securities) to our regressions, using data on yields for UK 10-year and 90-day rates from the Bank of England. From the ADL(2,1) model, what is the relationship between term spreads and quarterly growth rates for the period 1990-2005?

Now, take the code provided and estimate the ADL(2,2) model and the ADL(3,3) model.

Finally, put it all together by loading in the UK quarterly unemployment rate from the `UK_UnemploymentRate.xls` spreadsheet, set up the time series and estimate the ADL21 model including term spreads and the unemployment rate. Is the relationship between the quarterly unemployment rate and GDP growth positive or negative? Is it statistically significant?

Another package that may come in handy is the `dynlm()` package, which allows you to setup the start and end dates in one line with the full time series as input, without having to manually select the time period. This module, however, works with `ts` or `zoo` data types, so we use the `as.zoo()` tag to put this in the correct format. Some features of `dynlm()` will be useful depending on the task, so it is good to be familiar with how it works.