Lab 5: ARIMA models and decomposition methods

Introduction to Time Series Analysis

Name:
This lab is to be done in class (completed outside of class if need be). You can collaborate with your classmates, but you must identify their names above, and you must submit **your own** lab as an knitted pdf file. To answer the questions, display the results and write your answers if asked.

1. For the United States GDP data (from `global_economy`):

- If necessary, find a suitable Box-Cox transformation for the data;
- Fit a suitable ARIMA model to the transformed data;
- Try some other plausible models by experimenting with the orders chosen;
- Choose what you think is the best model and check the residual diagnostics;
- Produce forecasts of your fitted model. Do the forecasts look reasonable?
- Compare the results with what you would obtain using ETS() (with no transformation).

2. For the `fma::condmilk` series:

- Do the data need transforming? If so, find a suitable transformation.
- Are the data stationary? If not, find an appropriate differencing which yields stationary data.
- Identify a couple of ARIMA models that might be useful in describing the time series.
- Which of your models is the best according to their AIC values?
- Estimate the parameters of your best model and do diagnostic testing on the residuals. Do the residuals resemble white noise? If not, try to find another ARIMA model which fits better.
- Forecast the next 24 months of data using your preferred model.
- Compare the forecasts obtained using `ets()`.

3. Consider `fpp2::austourists`, the quarterly number of international tourists to Australia for the period 1999–2010.

- Describe the time plot.
- What can you learn from the ACF graph?
- What can you learn from the PACF graph?
- Produce plots of the seasonally differenced data $(1 - B^4)y_t$. What model do these graphs suggest?
- Does `ARIMA()` give the same model that you chose? If not, which model do you think is better?
- Write the model you think is better in terms of the backshift operator.

4. Consider `fpp2::usmelec`, the total net generation of electricity (in billion kilowatt hours) by the U.S. electric industry (monthly for the period January 1973 – June 2013). In general there are two peaks per year: in mid-summer and mid-winter.

- Examine the 12-month moving average of this series to see what kind of trend is involved.
- Do the data need transforming? If so, find a suitable transformation.
- Are the data stationary? If not, find an appropriate differencing which yields stationary data.
- Identify a couple of ARIMA models that might be useful in describing the time series. Which of your models is the best according to their AIC values?
- Estimate the parameters of your best model and do diagnostic testing on the residuals. Do the residuals resemble white noise? If not, try to find another ARIMA model which fits better.
- Forecast the next 15 years of electricity generation by the U.S. electric industry. Get the latest figures from the https://www.eia.gov/totalenergy/data/monthly/#electricity to check the accuracy of your

forecasts.

- Eventually, the prediction intervals are so wide that the forecasts are not particularly useful. How many years of forecasts do you think are sufficiently accurate to be usable?

5. Undergraduate students only: show that ETS (A, N, N) is equivalent to ARIMA$(0,1,1)$ with $\theta_1 = \alpha - 1$.
6. Graduate students only: show that ETS (A, A_d, N) is equivalent to ARIMA$(1,1,2)$ with $\phi_1 = \phi$, $\theta_1 = \alpha + \phi\beta - 1 - \phi$, and $\theta_2 = (1-\alpha)\phi$.
7. Undergraduate students only: show that ETS (A, N, A) is equivalent to ARIMA$(0,0,m)(0,1,0)_m$.
8. Graduate students only: show that ETS (A, A, A) is equivalent to ARIMA$(0,1,m+1)(0,1,0)_m$.