02417: Time Series Analysis

Lasse Engbo Christiansen
DTU Compute
Spring 2018
Technical University of Denmark

Assignment 1: Forecasting diesel consumption

In Denmark diesel engines have dominated the market for heavy vehicles. It has also been cost effective to have a diesel car if your annual milage is sufficiently high. This assignment focuses on predicting the annual consumption of diesel in Denmark.
The data comes from `www.statistikbanken.dk` and is part of the Environmental-Economic Accounts. The file `A1_diesel.txt` contains 2 columns:

1: The year that the observation is from.

2: The consumption of diesel in tonnes.

You should not use the observations from 2014 through 2016 (Last three observations) for estimations - only for comparisons.

Question 1.1: Plot the diesel consumption as a function of time.

Question 1.2: Does the global mean value and standard deviation of the diesel consumtion give a reasonable representation of the data? (The answer should be elaborated.)

Question 1.3: Let us consider a model for all the data (except the last three years).

Formulate a GLM model in form of a simple linear regression model for all the data. Estimate the model parameters and plot your fit. Would this model be useful for making predictions of the future diesel consumption?

Question 1.4: Now we will consider methods which considers data more locally. Use simple exponential smoothing with $\lambda = 0.8$ to predict the diesel consumption. Plot the data and the corresponding one step predictions for all N. Make a table with the predictions for the three years that were left out.

Comment on the results.

Question 1.5: Use a local linear trend model to predict the diesel consumption. Plot the data and the corresponding one step predictions for all N. Make a table with the predictions for the three years that were left out. Again use $\lambda = 0.8$.

Comment on the results.

(Hint: Consult the example in Sec. 3.6 of the lecture notes).

Question 1.6: Find an optimal value of the forgetting factor for use in the local trend model suggested in the previous question. (Optimize 1-step predictions. And disregard the first 5 1-step predictions as burn in period.)

Question 1.7: Comment on the results. Which model do you prefer? Do you trust the forecasts? Do you have ideas for improving the forecast method?

Note: Simple exponential smoothing does not include a measure of uncertainty. However it can be established through a local constant trend model.
Note: If you choose to provide a prediction interval (PI) as your measure of uncertainty then make it a 95% PI.