Q-Q plot from car library

02417: Time Series Analysis

Sebastian Wolf and Lasse Engbo Christiansen
Spring 2019

DTU Compute
Technical University of Denmark

Assignment 3: ARIMAX model for building data

In this assignment, we consider the same simplified energy balance for a building as we did in Assignment 1.

$$\Phi_h = H_{tot}(T_i - T_e) + gA_{sol}I_{sol} + \varepsilon, \tag{1}$$

where the quantities are

- Φ_h heating power inside the building [W]
- T_i indoor temperature (controlled at a constant setpoint of 25 $^\circ C$)
- T_e outdoor temperature [$\deg C$]
- I_{sol} solar irradiation [Wm^{-2}]
- H_{tot} sum of total transmission and ventilation loss coefficient [W/K]
- gA_{sol} parameter which is the product of: g solar transmittance of the transparent facade elements and their total area A_{sol} [m^2]
- ε noise

However, this time we take temporal dynamics of the inputs and the output into account.

The data is provided in `house_data_30min.csv` and is sampled every 30 minutes, i.e. a smaller time scale than in Assignment 1. The data set includes the following columns:

t: Date and time of the observation

Ph: The heating [W]

Te: The outdoor temperature [$degC$]

Isol: The solar radiation [W/m^2]

You should not use the observations for the last 2 hours for estimations - only for comparisons. In other words: The last four observations are to be left out of the training data and only used for testing model predictions.

Question 3.1: Plotting Read the data and plot the given quantities as a function of time. Indicate the test data in the plots. Comment on the evolution of the values over time.

Question 3.2: Correlation structure

1. Plot the autocorrelation function of *Ph* and the cross-correlation function of *Ph* and the model inputs. Comment on the plots.

2. What can you say about a model for *Ph* at this stage?

Question 3.3: ARIMA model Finding a suitable ARIMA model for the heating level.

1. Try to find a suitable ARIMA model for the heating level *Ph*. Start off with a simple model and successviely add more terms. Develop the model using graphical and numerical indicators (Including tests). Do not rely on one single indicator.

   In case the residuals of your final are not satisfactory, describe what you have tried and argue why it is not possible to find a better model.

   You don't have to present a full residual analysis for every single model that you have looked at. However, you should present some intermittent models in order to argue for the choices that you have made.
2. Predict the four observations that were left out. Present the predictions both in a plot and in a table.

Question 3.4: ARIMAX model Next, we try to improve the model by adding the external covariates *Te* and Isol. Do not only use the current values of these variables, but do also include lagged versions.

1. Make a model selection by adding/removing the external variables in a reasonable way.

   Again, if the residuals of the final model are not satisfactory, i.e. that they are far from white noise, argue why this is the case and why the model cannot be improved by using the given data.
   Hint: Add extra columns for the lagged versions of *Te* and Isol to the data frame.
2. Predict the four observations that were left out. Present the predictions both in a plot and in a table.

Question 3.5: Comparison Compare the models from the two previous questions.