02417: Time Series Analysis

Rune Grønborg Junker
Speing 2023

DTU Compute
Technical University of Denmark

Assignment 3: The Danish housing market

In this assignment we will consider house prices throughout Denmark. To improve predictions the inflation rate (change in consumer price index) and the interest rate from the Danish national treasury are included. The data contains the average sales price per square meter in quarterly resolution. The average prices are given for all of Denmark and averaged within the capital region, Sealand, Middle Jutland and Rural areas (defined as southern Denmark and north Jutland). Beware that the data for the house prices ends in 2022 Q3, while the inflation and interest rate continues until 2023 Q1. The data file "A4Data.csv" contains all the data, only NA's after 2023 Q2 until 2024 Q1. All data is obtained from Danmarks Statistik (`https://www.dst.dk/da/Statistik/emner/oekonomi`).
For part one of the assignment you will find a suitable ARIMA(X) model for the average price in Denmark. Then, in part two you will instead find a MARIMAX model the models the four regions of Denmark simultanously. I expect some differences in the obtained models, depending on what strategy you use for model selection.
It is the intention that you use the R-package "marima" for the second part of the assignment - other software packages may have similar functionality that you can use, but it might be significantly more difficult. marima can be installed in R by: `install.packages("marima")`.

Part 1 - Univariate models

Question 4.1: Presenting the data Plot the quarterly average sales prices in Denmark and the additional variables, interest rate and inflation rate. Analyze the behaviour of the data, such as whether or not it is stationary, and comment on how this will affect the model building. Consider transformations of the individual series already at this point.

Question 4.2: ACF and PACF Estimate and plot the autocorrelation function and the partial autocorrelation function of the house prices and if relevant also for transformations of the data. Estimate and plot the cross-correlation function between the four time series of the house prices. You do not have to make tables of the estimates.

Argue for an initial guess for a model structure based on the ACF and PACF.

Question 4.3: Univariate model selection Find suitable univariate ARIMA model for house prices. It is up to you how you do so, but you have to explain your method. You do not have to show all intermediary plots that you have used.

Question 4.4: Residual diagnostics For you final model, determine whether the residuals can be considered white noise and normal distributed.

Question 4.5: Forecasting the future house prices - I Use your ARIMA to forecast the house prices for each quarter until 2024 Q1, i.e. 6 time steps beyond your trainning data. Make a plot of the predictions, including training data if relevant, and a 95% prediction interval.

Question 4.6: External inputs Extend your ARIMA to an ARIMAX model where interest rate and/or inflation rate is used as an external regressor. Be careful when using arima from R. It will difference your external regressors (xreg) the same way as it is differences your observations. See exercise from week 7. Comment on whether the external regressor(s) are worthwhile to keep in the model.

Question 4.7: Forecasting the future house prices - II Use your ARIMAX to forecast the house prices for each quarter until 2024 Q1, i.e. 6 time steps beyond your training data, assuming that the inflation and interest rates will stay at the same level as 2023 Q1. Make a plot of the predictions, including training data if relevant, and a 95% prediction interval.

Question 4.8: Conclusions - I Argue why you trust or do not trust your estimated model for forecasting the house prices in Denmark. Explain whether it is economically wise to buy a house in Denmark right now according to your models.

Part 2 - Multivariate models

Question 4.9: Re-presenting the data Plot the quarterly average sales prices for each of the four regions, and comment on how their behaviour will influence the model building.

Question 4.10: ACF and PACF Plot the multivariate ACF and PACF for the prices in the four regions, and use these to argue for a MARIMAX model structure.

Question 4.11: Multivariate model selection Find suitable univariate MARIMAX model for house prices. It is up to you how you do so, but you have to explain your method. You do not have to show all intermediary plots that you have used. Determine whether inflation and interest rates should be used as simple regression variables (reg.var in marima package) or as indepdent variables (indep in marima package).

Question 4.12: Multivariate explanation Present your final model along with the estimated parameters in a way that a reader who has never heard of MARIMA models, but who is familiar with ARIMA models can understand it. Interpret the parameter estimates, and explain how your model believes the different regions to influence each other.

Question 4.13: Forecasting the future house prices - III Use your MARIMAX model to make 1-step predictions for the whole data set and forecast the future house prices until 2024 Q1 with 95% predictions intervals. Make a plot of the 1-step predictions and forecasts of the future together with the training data. Zoom in if this makes the behaviour of your model more clear.

Question 4.8: Conclusions - II Argue why you trust or do not trust your estimated MARIMAX model for forecasting the house prices in Denmark. Explain where, if any place, that it is the best economic decision to buy a house in Denmark according to your model.