

## Assignment 4: Multivariate Processes

In this assignment we will consider apartment prices throughout Denmark. Specifically we will consider apartment prices in the four zip codes 2000, 2800, 4000 and 4200. To improve predictions the consumer price index (CPI) and two interest rates are included. The two interest rates are a 30 year real estate bond rate (`obl30yr`) and the discount rate from Nationalbanken (`nat_tres_lend_rate`, they represent long range and short range interest rates. The data concerning house prices is obtained from Finans Danmark (<https://rkr.statbank.dk/statbank5a/SelectVarVal/Define.asp?MainTable=BM010&PLanguage=0&PXId=0&wsid=cflist> while interest rates and the consumer index is downloaded from Danmarks Statistik (<https://www.dst.dk/da/Statistik/emner/oekonomi>).

The last four observations are reserved for testing the model, with the rest being used for training.

It is the intention that you use the R-package `marima` for the last two questions - 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")`

This assignment is targeting your ability to use and understand MARIMA models. I expect your estimates and predictions to be almost identical to those that I found.

**Question 4.1: Presenting the data** Plot the prices and the additional variables. Comment on the behaviour including considerations on stationarity. Consider transformations of the individual series already at this point.

**Question 4.2: ACF, PACF and CCF** Estimate and plot the autocorrelation function and the partial autocorrelation function of the data series 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.

Comment on the structures you find.

**Question 4.3: Univariate model selection** Find suitable univariate ARIMA models for each of the price series.

You do not have to present diagnostic plots for all models. But do pick some examples and comment briefly on how the others behave.

**Question 4.4: Multivariate model estimation** Normally you would find a suitable MARIMA model based on your findings from the previous questions, but in this assignment I have chosen the model structure for you:

1. Estimate a MARIMA model using the training data where both the AR and MA part are of first order (i.e. using only lag 1), and as inputs the two kinds of interest rates non-differenced and the CPI differenced with a seasonality of 4 (yearly). Remove insignificant terms recursively, where insignificant terms are those for which the f.value

is below  $2 \cdot 1.96 = 3.92$ . This can be done using e.g. the function "step.slow" uploaded on Learn.

2. Present the parameter estimates and explain how the model works, i.e. how inputs and previous house prices are used to predict future house prices.

**Question 4.5: Predictions** Use the model you have developed for predicting the prices for the remaining four observations and include prediction intervals. Show the results in a table and a plot.

Compare with the actual data – comment on the results.

Note: Assume that the consumer price index and the interest rates are known for the test data.

**Question 4.6: Re-training the model** Re-estimate the parameters of your model, this time using all of the data including the last 4 observations.

Present the parameter estimates and make a qualitative comparison to the previous model. What have significantly changed?

**Question 4.7: What does the future hold?** Predict 4 time steps into the future (to and including the second quarter of 2022) using the new estimate of the model. Show the results in a table and a plot.

Assuming that you have the spare capital and can buy real estate in each of the four zip codes for the prices of the second quarter of 2021, in what zip codes would it be profitable to buy real estate according to your model? Describe your reasoning.

Note: Since you do not have the future consumer price index or interest rates you have to assume them to be unknown for this question.