Universität Duisburg-Essen
Fachbereich VWL
Lehrstuhl Ökonometrie

Winter 2019/20

Dr. Yannick Hoga   Thilo Reinschlüssel

Multivariate Time Series Analysis
Exercise Sheet 6

Exercise 1: Model Selection – Review

a) Is the MSE scale-invariant? (Does it matter if you use percentage points instead of decimals?)

b) What is the fundamental trade-off which information criteria are supposed to balance?

c) Does a linear transformation affect the value of the information criteria? Does it also influence the locations of the minima (from which we pick the optimal order)?

d) Finally: Are OLS standard errors scale invariant?

Exercise 2: Simplification and Forecasting – Macroeconomic Data

Reconsider Exercise 2 from Exercise Sheet 5. Again, please import/load the dataset 'us_macrodata.Rda' into your workspace and compute the growth rates of the variables appearing non-stationary. There are still 5 variables—CPI, Real GDP, the unemployment rate, general private investment and the debt-to-GDP ratio. All series have been sampled quarterly and were seasonally adjusted before downloaded from *FRED*.

a) Fit a VAR(p) model according to the Hannan-Quinn information criterion.

b) Use the estimated coefficients and the associated standard errors to compute the t-statistics ($H_0 : \phi_{p,ij} = 0$ vs $H_1 : \phi_{p,ij} \neq 0$) for each coefficient separately. Then count how many coefficients are <u>not</u> significantly different from zero at the 5% level.

c) Estimate the refined model using the command `refVAR` by setting a threshold corresponding to the 5% level from b).

   i) How many variables have been set to 0?

   ii) Does the number coincide with your count in task b)?

   iii) What may be the reason for the two numbers differing? *(Hint: Slide 4-21)*

d) Compare the values of all information criteria offered to you both for the 'ordinary' VAR and the refined VAR model. Which model is best? Is the recommendation unanimous?

e) Proceed to compare the MSEs of the 'ordinary' VAR model and the refined model. Is the model picked by the information criteria again superior? Explain your results.

f) Calculate the numbers of coefficients estimated both for the 'ordinary' model and the refined model. Then perform a Ljung-Box test on the residuals of both models.

   i) Do the models absorb the dynamics in the data completely?

   ii) Explain the massive differences of the two tests at $m = \{3, 4\}$.
   Hint: Mind the option 'adj' using the command mq!

g) Now estimate a VAR(1) model (with intercept). How does it compare to the VAR(4) model in terms of MSE?

h) Lastly, compute the forecasts' MSEs (referred to as MSFE) for both models using the command `VARpred`. Please use a forecast horizon h of 10.

   i) Does the model superior in f) still prevail at every h?

   ii) Explain what conceptual difference between MSE and MSFE drives the resuls in i).

   iii) To which values do the forecasts converge to if $h \to \infty$?

   Hint: `VARpred` *objects offer the RMSE based on the forecasts.*

Exercise 3: Simplification and Forecasting – Exchange Rates

Reconsider Exercise 3 from Exercise Sheet 5. Please reuse the file 'quandl_fx_download.R' from last time to download and prepare the data. (Feel free register at `www.quandl.com` to get an API key for unlimited downloads.)

a) Fit a VAR(1) model to the data regardless of the information criteria.

b) Now fit the refined model based on your VAR(1) setting the threshold to 1.96. How many coefficients have been set to zero?

c) Compare the MSEs of the 'ordinary' model and the refined model.

d) Thirdly, estimate a VAR(0) with intercept by regression. Compare its MSFE with the forecast errors of the 'ordinary' VAR(1) model regarding a forecast horizon $h = 10$.
   Hint: Write down the loss function and find the FOC's trivial solution.

e) How do your results in d) align with the insights you gained in exercise 2h)?

f) What can you do to reproduce the findings from 2h) in this setting?
   Hint: Degrees of freedom.

Exercise 4: Forecast Errors

Show that Equation (5.1) in the lecture slides implies that

$$\mathsf{E}[\hat{z}^{(i)}_{T,T+h} - z_{T+h}]^2 \geq \mathsf{E}[z^{(i)}_T(h) - z_{T+h}]^2,$$

where $\hat{z}^{(i)}_{T,T+h}$ and $z^{(i)}_T(h)$ denote the i-th components of the respective forecasts ($i = 1, \ldots, K$) for the observation z_{T+h}. This means that the optimal *univariate* forecasts are simply the components of the optimal *multivariate* forecast $z_T(h)$.

This exercise sheet will be discussed in the tutorial on Wednesday, 27 November 2019