EBC4222
Descriptive and Predictive Analytics
Tutorial 4 Exercises: Vector AutoRegression - VAR models

Roselinde Kessels
r.kessels@maastrichtuniversity.nl

2 May 2024

For the first seven questions, consider the following VAR(1) model for two data series x_t and y_t for observations $t = 1, \ldots, T$:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$
$$\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim \text{NID}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix} \right)$$

where $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ for $t = 1, \ldots, T$ are the error terms, and $\delta_1, \delta_2, \theta_{11}, \theta_{12}, \theta_{21}, \theta_{22}, \sigma_1^2, \sigma_{1,2}^2$ and σ_2^2 are the model parameters. In addition, NID(μ, Σ) denotes an independent (multivariate) normal distribution with mean μ and variance-covariance matrix Σ.

1. Show that the model leads to the below Autoregressive Distributed Lag (ADL) models for series x_t and y_t and clearly state the correlation assumptions in the error terms:

$$y_t = \delta_1 + \theta_{11} y_{t-1} + \theta_{12} x_{t-1} + \varepsilon_{1,t},$$
$$x_t = \delta_2 + \theta_{21} y_{t-1} + \theta_{22} x_{t-1} + \varepsilon_{2,t},$$

2. Provide a reason to apply the VAR model in the question instead of the two separate ADL models for series x_t and y_t.

3. Explain whether we need stationarity assumptions for the series x_t and y_t in order to apply the VAR model in the question.

4. Extend the VAR model in the question to allow for deterministic trends in the two variables.

5. Suppose that we estimate the model in the question and find that the residuals for series y_t have non-zero autocorrelation. I.e. ACF and PACF of the estimated error terms $\hat{\varepsilon}_{1,t}$ are not zero. Provide a method that can be used to solve the non-zero correlation problem in the residuals.

6. The purpose in this question is to compare a VAR(1) model and a VAR(2) model in terms of their in-sample fit for data x_t and y_t. The tools to compare two models are listed below. For each model comparison tool, state whether this approach can be used to VAR(1) and VAR(2) models. Clearly explain your reasoning.

   (a) General to specific model selection using t-tests or F-tests,

   (b) Choosing the model with highest adjusted R^2,

   (c) Choosing the model based on BIC or AIC,

   (d) Misspecification tests for the residuals.

7. Explain the 'curse of dimensionality' in unrestricted VAR models. For this purpose, calculate how many parameters need to be estimated for an unrestricted VAR(1) model with an intercept, for 10 series.

8. We will analyze the macroeconomic indicators of Canada. These data are provided in the *R* package `vars` and include four economic indicators: unemployment rate `U`, production index `prod`, real wages `rw` and civil employment rate `e`. See https://cran.r-project.org/web/packages/vars/vars.pdf

   (a) Plot the time series of the four economic indicators in Canada data and check each of them for stationarity using the ADF test. Make the time series stationary if needed by first differencing.

   (b) Estimate a VAR(1) model (without a trend or constant) for all (differenced) series. Comment on the estimation output for the unemployment `U`. What are the effects of the other variables on unemployment?

   (c) Estimate a VAR(1) model with a constant and a VAR(1) model with a constant and a trend for all (differenced) series. Compare the adjusted R^2 of the unemployment equations with that in part (b). Which model is better in explaining Canadian unemployment rates?

   (d) An important consideration is whether real wages affect the rest of the economic indicators such as unemployment and production. Perform a Granger Causality test to assess whether real wages `rw` are the causes of the changes in the remaining variables.

   (e) Estimate and plot the impulse responses of the variables `U`, `prod` and `e` to a one unit change in real wages `rw`. Comment on the impulse response plots.