MW82: Time Series Analysis, Tutorial VI

Dr. Giulia Sabbadini

DICE

WS 2023/24

Recap: Tutorial V

We did:

- Non-stationarity
- Unit root tests
- Seasonal ARIMA

Today we will go through:

- Multiple Time Series
- Stationary VAR(p)
- Granger Causality

Multiple Time Series

- So far, we had only models of single time series
- Now, we want to model interactions between multiple economic variables
- Examples:
  - Economics: GDP, consumption, investment
  - Finance: different stock markets
  - Marketing: sales across multiple products (substitution)
- Main tool:
  - Vector Autoregression (VAR) models, which are an extension of AR models to multiple variables

VAR(1) model

Assume that you are interested in two time series, y_1 and y_2.

- VAR models are an extension of AR models for univariate time series
- With AR models you had only auto-correlation:

$$
\begin{aligned}
y_{1,t} &= \phi_1 y_{1,t-1} + \varepsilon_{1,t} \\
y_{2,t} &= \phi_2 y_{2,t-1} + \varepsilon_{2,t}
\end{aligned}
$$

VAR(1) model II

Add cross-correlation terms (ϕ is α from the lecture slides).

$$y_{1,t} = \phi_{11}\, y_{1,t-1} + \phi_{12}\, y_{2,t-1} + \varepsilon_{1,t}$$
$$y_{2,t} = \phi_{21}\, y_{1,t-1} + \phi_{22}\, y_{2,t-1} + \varepsilon_{2,t}$$

In matrix notation: $y_t = \Phi_1 y_{t-1} + \varepsilon_t$, where

$$\Phi_1 = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$$

VAR(1) model II

- The vector of error terms ε_t are interpreted as *shocks* or *innovations*
- This interpretation requires $E[\varepsilon_t|\Omega_{t-1}]{=}0$
- The errors have to be serially uncorrelated $\rightarrow$ lag length p sufficiently large to ensure this
- The errors have to be independent of each other
- Gaussian white noise: $\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$
- The model dynamics are initiated by these shocks but the dynamic responses are determined by the polynomial Φ_1

VAR(p) models

- VAR models can be extended in two dimensions:
  1. More lags (p)
  2. More variables (m)
- Φ_1 has m^2 coefficients, 'curse of dimensionality'
- VAR(p) models are:

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \varepsilon_t$$

  where y_t is an $m \times 1$-vector.
- This is called the reduced form (reduced form VAR).
- It is trivial to add deterministic terms (intercepts, trends, seasonal dummies, explanatory variables); this is called (sometimes) VARX.

Stationarity and Estimation

- Similar to AR(p) models, stationarity depends on the coefficient matrix Φ.
- AR(1): $|\phi_1| < 1$ and VAR(1): All eigenvalues of matrix Φ_1 are smaller than one in absolute value.
- If we have a stationary system of variables, we can estimate the VAR model by OLS.

Choosing Variables and Lag Order

Very similar to the univariate approach:

1. Select endogenous and exogenous variables (seasonal dummies, trends, ...)
2. Estimate the model for p=$1, 2, 3, \ldots$ lags
3. Choose the model with the lowest information criteria (multivariate versions of AIC, BIC/SC)
4. Can be useful: ACF and CCF plots of time series
5. AIC tends to over-fit (too many lags), some prefer BIC/SC.

$$AIC(p) = \log|\hat{\Sigma}| + \frac{2pm^2}{T}$$

6. Check residual plots and test independence of residuals with a Portmanteau Test (multivariate version of Box tests).

Granger causality

- Within our (reduced form) VAR framework, we cannot attribute causality to our estimates
- Granger or "predictive" causality
- Time series y_1 is said to Granger-cause y_2 if lags of y_1 help predict y_2

Granger causality

- This is equivalent to checking if the VAR(p) coefficients of (lags of) y_1 in the y_2-equation are jointly zero.

$$
\begin{aligned}
y_{2,t} = &\phi_{22,1}\, y_{2,t-1} + \phi_{22,2}\, y_{2,t-2} + \cdots + \\
&\phi_{21,1}\, y_{1,t-1} + \phi_{21,2}\, y_{1,t-2} + \dots \phi_{21,p}\, y_{1,t-p}
\end{aligned}
$$

- i.e. $\phi_{21,1} = \phi_{21,2} = \cdots = \phi_{21,p} = 0$
- This can be accomplished by an F-test (H_0: not Granger causal)
- This is not real causality, just predictive ability

Impulse Response Function

- When a VAR model is used for policy analysis, the interest is not so much in the estimated parameters, but in the so-called **impulse responses**
- The idea is to track the transmission of a (structural) shock through the system over time
- An impulse response function (IRF) gives the dynamic effect of a shock $\varepsilon_{j,t}$
- That is, the IRF tells us how a shock in a variable in a certain period affects it and all the other variables in the following periods
- To measure the impact of a shock we exploit the MA-representation of the VAR(p)-system: a variable is expressed as the sum of the infinite history of past innovations

Impulse Response Function

- Problem: the error terms can be contemporaneously correlated (covariance matrix need not be diagonal)
- If ε_{1t} and ε_{2t} are correlated, we cannot consider the m=2 IRFs separately
- If they are correlated, it is not possible to trace the effect of a shock separately from the other

Impulse Response Function

- We overcome this problem by imposing a causal ordering on the variables. This is a recursive structure such that shocks to higher ordered variables affect shocks to variables ordered lower but not *vice versa* (Choloski identification, recursive identification)
- Theory enters at the interpretation stage through the imposition of the causal ordering
- Unfortunately depends on the ordering of the variables (ordering can change results)

Impulse Response Function

Consider the matrix form for VAR(1):

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ v_t \end{bmatrix}$$

or

$$z_t = \phi z_{t-1} + w_t \quad \text{(Reduced Form)}$$

where $z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}$, $\phi = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$, and $w_t = \begin{bmatrix} u_t \\ v_t \end{bmatrix}$.

with variance-covariance matrix:

$$\Omega = E(w_t w_t') = \begin{bmatrix} \sigma_u^2 & \sigma_{u,v} \\ \sigma_{u,v} & \sigma_v^2 \end{bmatrix}$$

Impulse Response Function

- Using the lag operator we can show the MA(∞) representation for the VAR(1) is:

$$z_t = w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \ldots + \phi^j w_{t-j} + \ldots$$

- The coefficient in the MA representation measures the impulse response:

$$\phi^j = \frac{dz_t}{dw_{t-j}}$$

- Note ϕ^j is a 2×2 matrix for a bivariate system.

Impulse Response Function

- More generally, denote the (m, n)-th component of ϕ^j by $\phi^j(m, n)$.
- Then $\phi^j(m, n)$ measures the response of m-th variable to the n-th error after j periods.
- For example, $\phi^3(1, 2)$ measures the response of the first variable to the error of the second variable after 3 periods.
- For a bivariate system, there are four impulse response plots.

Impulse Response Function

- In general, u_t and v_t are contemporaneously correlated (not-orthogonal), i.e., $\sigma_{u,v} \neq 0$.
- However, we can always find a lower triangular matrix A so that
$$\Omega = AA' \quad \text{(Cholesky Decomposition)}$$
- Then define a new error vector $\tilde{w}_t$ as a linear transformation of the old error vector w_t:
$$\tilde{w}_t = A^{-1} w_t$$
- By construction, the new error is orthogonal because its variance-covariance matrix is diagonal:
$$\text{var}(\tilde{w}_t) = A^{-1}\text{var}(w_t)A^{-1'} = A^{-1}\Omega A^{-1'} = A^{-1}AA'A^{-1'} = I$$

Cholesky Decomposition

Let $A = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix}$. The Cholesky Decomposition tries to solve

$$\begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} = \begin{bmatrix} \sigma_u^2 & \sigma_{u,v} \\ \sigma_{u,v} & \sigma_v^2 \end{bmatrix}$$

The solutions for a, b, c always exist and they are

$$\begin{aligned} a &= \sqrt{\sigma_u^2} \\ b &= \frac{\sigma_{u,v}}{\sqrt{\sigma_u^2}} \\ c &= \sqrt{\sigma_v^2 - \frac{\sigma_{u,v}^2}{\sigma_u^2}} \end{aligned}$$

Orthogonal IRFs

Rewrite the MA(∞) representation as

$$z_t = w_t + \phi w_{t-1} + \ldots + \phi^j w_{t-j} + \ldots$$

$$= AA^{-1} w_t + \phi AA^{-1} w_{t-1} + \ldots + \phi^j AA^{-1} w_{t-j} + \ldots$$

$$= A\tilde{w}_t + \phi A\tilde{w}_{t-1} + \ldots + \phi^j A\tilde{w}_{t-j} + \ldots$$

This implies that the impulse response to the orthogonal error $\tilde{w}_t$ after j periods is

$$j\text{-th orthogonal impulse response} = \phi^j A$$

Orthogonal IRFs

The ordering of the following model:

$$y_t = \begin{pmatrix} \text{inflation} \\ \text{growth} \\ \text{interest rate} \end{pmatrix}$$

implies that shocks to

- current inflation $\rightarrow$ current inflation, current growth & current interest rate
- current growth $\rightarrow$ current growth & current interest rate
- current interest rate $\rightarrow$ current interest rate

VAR estimation in R

R library `vars`

- estimate VAR(p): `var1 <-VAR(dataset, p = p)`
- select VAR order p by IC: `VARselect(dataset)`
- Granger-causality: `causality(var1, cause = "y1")[1]`
- Portmanteau test:
  `serial.test(var1, type = "PT.asymptotic")`
- IRF: `irf(var1,ortho = T,impulse = varname)`
- Forecasting *h* leads: `predict(var1, n.ahead=h)`

You can use `summary` for printing the results, the `forecast::Acf()` and `ccf()` functions for residuals.

Exercise I

- Load the *prod_ind* dataset. It contains the quarterly growth rate of the total industry production index for Germany and Austria from 1970Q1 (seasonally adjusted).
- Discuss the stationarity of the individual time series.
- Estimate a VAR(p) model, where p has been selected based on an information criterion
- Test the independence of the residuals and check your residual plots.
- Is *deu* Granger-causing *aut* and vice-versa? Interpret.
- Plot the orthogonal IRFs and interpret.