7: Vector Auto-Regression (VAR)

GECO 6281 Advanced Econometrics 1 (Lab)

Patrick Mokre

Fall 2019

Recapitulation

- Dealing with Endogeneity: Instrumental Variables (IV)
- Static Panel Data
- Instrumental Variables in Dynamic Panel Data: Anderson-Hsiao, Arellano-Bond, Arellano-Bover
- Seemingly Unrelated Regression (SUR)
- Auto-Regressive Distributed Lag (ARDL) and Error Correction Modeling (ECM)
- Stationarity
- Panel Stationarity

Vector Autoregression

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix} \tag{1}$$

$$\begin{aligned} y_t &= b_{11} y_{t-1} + b_{12} z_{t-1} + \epsilon_{1,t} \\ z_t &= b_{21} y_{t-1} + b_{22} z_{t-1} + \epsilon_{2,t} \end{aligned}$$

- Captures mutual dependencies between time series
- Intuitive Forecasting on all variables (ARDL/DL only allows forecasts for $y_{i,t}$)
- Takes account of the usual endogeneity between economic processes
- No simultaneity bias because only lags of the "other" variable are included

VAR: Assumptions

- y_t and z_t are dynamically related, but not contemporaneously related.
- Error terms $\epsilon_{1,t}$ and $\epsilon_{2,t}$ are contemporaneously uncorrelated
- y_t and z_t are stationary

$\Rightarrow$ Estimation by OLS is efficient, consistent, and SUR estimation does not improve efficiency. Standard Errors and Covariances can be estimated in standard form. AIC/BIC are appropriate for lag selection.

VAR with exogenous variables

Adopt **Lütkepohl** notation, because it underlies the corresponding STATA package (makes it easier to understand help files etc.)

$$y_t = AY_{t-1} + B_0 x_t + u_t \tag{2}$$

where y_t the vector of endogenous variables, A the matrix of AR coefficients, B_0 the matrix of exogenous coefficients, x_t the vector of exogenous covariates and u_t a vector of white noise disturbances ("innovations").

Intercepts are included in x_t.

VAR: Lütkepohl

Figure 1: Helmut Lütkepohl: *"I see that you have* ***not*** *adopted full matrix notation yet."*

VAR: Advanced Lütkepohl Notation

$$Y = BZ + U \tag{3}$$

$$
\begin{aligned}
Y &= (y_1, ..., y_t) & \text{Y is } K \times T \\
B &= (A, B_0) & \text{B is } K \times (Kp + M) \\
Z &= \begin{bmatrix} Y_0 & ... & Y_{T-1} \\ x_1 & ... & x_T \end{bmatrix} & \text{Z is } (Kp + M) \times T \\
U &= (u_1, ..., u_T) & \text{U is } K \times T
\end{aligned}
$$

In STATA, this model is estimated using iterative seemingly unrelated regressions (**SUR**).

Lag Selection for VAR

As for single equation auto-regressions, a variety of **goodness of fit** criteria can be calculated for VAR.

- Akaike's Final Prediction Error (FPE): the **determinant** of the average squared prediction error matrix, normalized by $((1 + m/N)/(1 - m/N))^K$ (m is the average number of coefficients between all models, K is the number of equations).
- Akaike Information Criterion (AIC): $-2\frac{LL}{T} + \frac{2t_P}{T}$ where LL is the log-likelihood of the model, and t_P is the total number of parameters in the evaluated model.
- Schwartz' Bayesian Information Criterion (BIC): $-2\frac{LL}{T} + \frac{ln(T)}{T}t_P$
- Hannan-Quin Information Criterion (HQIC): $-2\frac{LL}{T} + \frac{2ln[ln(T)]}{T}t_P$

The corresponding STATA command is `varsoc`, it can be used pre- and postestimation.

Multivariate Impulse Response Function

Impulse Response Functions (IRF) estimate how a time series reacts to a disturbance in the error terms.

Suppose that the estimated error structure ϵ is related to an underlying structural shock vector u_t.

$$\begin{aligned} \epsilon_t &= Au_t \\ E(u_t, u_t') &= I \end{aligned} \tag{4}$$

A is related to the error covariance matrix Σ:

$$\begin{aligned} \Sigma &= E[\epsilon_t \epsilon_t'] \\ &= E(Au_t u_t' A') \\ &= AE[u_t u_t']A' \\ &= AA' \end{aligned}$$

Because $\hat{\Sigma}$ can be estimated in the regression (e.g. via VAR), $\hat{A}$ can be retrieved.

In STATA, you can create post-estimation IRFs using `irf create` and `irf graph`.

Granger Causality

Granger Causality is not causality. Rather it measures which event happens first: z_t is said to "Granger cause" y_t if $(z_{t-1}, \dots, z_{t-p})$ contains information that helps predict y_t better than only $(y_{t-1}, \dots, y_{t-p})$ does.

A simple way of testing Granger Causality is to compare the tests for joint insignificance with and without $(z_{t-1}, \dots, z_{t-p})$ in predicting y_t. The corresponding STATA command is `vargranger`.

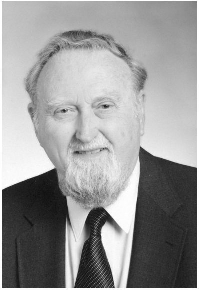

Figure 2: Clive Granger: *However, several writers stated that "of course, this is not real causality, it is only Granger causality."*

Vector Error Correction

Let x_t and y_t be two first difference stationary processes, i.e. Δy and Δx are **covariance stationary**.

According to Granger and Newbold (1974), OLS regression of y on x provides **spurious results**, i.e. t-tests suggest significance of the coefficients where there is none in the data generating process. Phillips (1986) shows this is due to the asymptotic OLS properties not holding for first difference stationary processes.

If y_t and x_t cointegrate, a regression of Δy_t on Δx_t is also misspecified.

Remember **cointegration**:

- x_t, y_t are first difference stationary
- $e_t = y_t - \alpha - \beta x_t$ is covariance stationary

Engle-Granger VECM: Intuition

Re-Define the relationship between y_t and x_t as:

$$y_t + \beta x_t = \epsilon_t \quad \epsilon_t = \epsilon_{t-1} + \xi_t \tag{5}$$

$$y_t + \alpha x_t = v_t \quad v_t = \rho v_{t-1} + \zeta_t \quad |\rho| < 1 \tag{6}$$

Here ξ_t and ζ_t are i.i.d. but mutually correlated processes responsible for the co-integration. ϵ_t is I(1), so consequently, so must be y_t and x_t.

Define $\delta = (1-\rho)/(\alpha - \beta)$ and $z_t = y_t + \alpha x_t$.

$$\Delta y_t = \beta \delta z_{t-1} + \eta_{1,t} \tag{7}$$

$$\Delta x_t = -\delta z_{t-1} + \eta_{2,t} \tag{8}$$

In $z_t = 0$, y_t and x_t are in **equilibrium**, and coefficients on z_{t-1} show how y_t and x_t react to deviations from equilibrium.

Engle-Granger: Nobel Prize Winners 2002

Figure 3: Engle and Granger: Winning the Nobel Medal in 2003 for being really careful about which relationships they call "causal" or spurious.

Engle-Granger VECM: General Case

Any VAR can be written and estimated as a VECM.

$$y_t = v_t + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \epsilon_t \tag{9}$$

$$\Delta y_t = v_t + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \epsilon_t \tag{10}$$

One last re-writing for **Johansen** maximum likelihood estimation:

$$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + v + \delta t + \epsilon_t \tag{11}$$

The important STATA commands are `varsoc` for lag selection, `vecrank` for the number of cointegrating equations and `vec` for the estimation.