ECON 30401: *Lecture 5*

Nicky Grant

Autumn semester, 2017

Outline of today's lecture

System Modeling (Topic 4)

- Motivation
- VAR models
    - vector white noise
    - basic framework of VAR models

So far in course have considered *univariate* time series models, *e.g.* AR(1)

$$y_t = \phi y_{t-1} + \varepsilon_t$$

Now consider time series models for vector processes, such as bivariate Vector Autoregressive (VAR) model of order 1,

$$\mathbf{y}_t = \mathbf{\Phi} \mathbf{y}_{t-1} + \varepsilon_t$$

where now

$$\mathbf{y}_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix}, \ \mathbf{\Phi} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{bmatrix}, \ \varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

Note:

- in both models y_t depends on its lagged value and error
- in VAR, we have

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

$\Rightarrow$ $y_{1,t}$, say, depends on $y_{1,t-1}$ and $y_{2,t-1}$.

VAR models introduced by Christopher Sims in 1980 paper entitled "Macroeconomics and Reality", and have become very popular in macroeconometrics.

VARs are an example of systems modeling.

Systems modeling is both important in econometrics and also crucial to development of the subject.

So we start by describing how VAR's fit into evolution of ideas about systems modeling in econometrics.

Background: the Great Depression in the 1930's and economic theories of the business cycle.

Jan Tinbergen:

- 1937: model of the Dutch economy, system of 16 equations + some accounting identities (31 variables).
- 1939: model for the US macroeconomy, system of 48 equations (71 variables)
- awarded the inaugural Nobel prize in economics in 1969, (with Ragnar Frisch)

Alfred Cowles was investment counsellor who realized he did not understand US macroeconomy following 1929 Wall Street Crash.

Cowles Commission for Research in Economics:

- started in Colorado Springs (1932), moved to University of Chicago (1939), moved to Yale University (1950)
- made fundamental contributions to the statistical theory for systems of linear equations
- 1950's-1960's: Lawrence Klein developed ever more sophisticated econometric models of the US macroeconomy.

Klein's Model 1:

$$\begin{aligned}
C_t &= \alpha_0 + \alpha_1 P_t + \alpha_2 P_{t-1} + \alpha_3 (W_t^p + W_t^g) + u_{1,t} \\
I_t &= \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \beta_3 K_{t-1} + u_{2,t} \\
W_t^p &= \gamma_0 + \gamma_1 X_t + \gamma_2 X_{t-1} + \gamma_3 A_t + u_{3,t} \\
X_t &= C_t + I_t + G_t \\
P_t &= X_t - T_t - W_t^p \\
K_t &= K_{t-1} + I_t
\end{aligned}$$

where: C_t = consumption, I_t = investment, W_t^p = private wages, X_t = total output, P_t = private profits, K_t =capital stock, G_t = government non-wage expenditure, T_t = indirect business taxes plus net exports, W_t^g = government wage expenditure, A_t =a time trend

To understand the interpretation of these models, note variables divided into three categories:

- *endogenous variables*: C_t, I_t, W_t^p, P_t, X_t and K_t.
- *predetermined variables*: K_{t-1}, P_{t-1} and X_{t-1}.
- *exogenous variables*: G_t, T_t, W_t^g and A_t.

Klein's model I is a linear simultaneous equations model (LSEM).

Estimation of LSEM depends crucially on the division of variables into these categories. → Sims motivation for VAR's.

Linear Simultaneous Equation Models

$$\begin{aligned} y_{1,t} + \beta_1 y_{2,t} &= \gamma_1 x_{1,t} + \gamma_2 x_{2,t} + u_{1,t} \\ \beta_2 y_{1,t} + y_{2,t} &= \gamma_3 x_{1,t} + u_{2,t} \end{aligned}$$

- $\mathbf{y}_t = (y_{1,t}, y_{2,t})' \sim$ endogenous variables
- $\mathbf{x}_t = (x_{1,t}, x_{2,t})' \sim$ exogenous variables
- $\mathbf{u}_t = (u_{1,t}, u_{2,t})' \sim$ error terms
- parameters are $(\beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3)$

This is known as the *structural form* of system.

Linear Simultaneous Equation Models

Another important representation of system is known as the reduced form.

To obtain the reduced form, rewrite the structural form in matrix notation as:

$$B\mathbf{y}_t = \Gamma\mathbf{x}_t + \mathbf{u}_t$$

$$B = \begin{bmatrix} 1 & \beta_1 \\ \beta_2 & 1 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_3 & 0 \end{bmatrix}.$$

Assume B is nonsingular $\Rightarrow$

$$\begin{aligned} \mathbf{y}_t &= B^{-1}\Gamma\mathbf{x}_t + B^{-1}\mathbf{u}_t, \\ &= \Pi\mathbf{x}_t + \mathbf{v}_t, \qquad \text{say.} \end{aligned}$$

which is the reduced form.

Linear Simultaneous Equation Models

Fundamental question: can we estimate B, Γ from a sample $\{\mathbf{y}_t', \mathbf{x}_t'\}$?

Answer: depends on the mapping from reduced form parameters to the structural form parameters $\Pi = B^{-1}\Gamma$.

Cowles Commission called this issue of identification and devised rules for checking if estimation is possible. These depend on:

- number of exogenous variables in system
- exclusion restrictions

Critiques of LSEM of macroeconomy

Klein's models worked well in 1950's - 1960's but started to break down in 1970's.

Economists and econometricians became critical of this approach:

- Robert Lucas (1976) "Policy critique": parameters of LSEM are functions of economic environment and so cannot be assumed constant when model used for economic policy analysis.
- Sims (1980): identification never holds, instead use models for large blocks of equations (for sectors of economy) and in blocks all variables endogenous and each depends on past of all others ⇒ VAR(p) model for $k \times 1$ vector y_t:

$$\mathbf{y}_t = \Theta_1 \mathbf{y}_{t-1} + \Theta_2 \mathbf{y}_{t-2} + \ldots + \Theta_p \mathbf{y}_{t-p} + \varepsilon_t$$

Structural VARs

Popularity of Sim's methods → Structural VAR (SVAR) model:

$$\Psi_0 \mathbf{y}_t \;=\; \Psi_1 \mathbf{y}_{t-1} + \Psi_2 \mathbf{y}_{t-2} + \ldots + \Psi_P \mathbf{y}_{t-P} + \mathbf{u}_t$$

Example: simple New Keynesian macroeconomic model,

$$\begin{aligned}
\tilde{y}_t &= \beta_{1,1}\tilde{y}_{t-1} + \beta_{1,2}\pi_t + \beta_{1,3}r_{t-1} + u_{1,t} \\
\pi_t &= \beta_{2,1}\pi_{t-1} + \beta_{2,2}\tilde{y}_t + \beta_{2,3}r_t + \beta_{2,4}r_{t-1} + u_{2,t} \\
r_t &= \beta_{3,1}r_{t-1} + \beta_{3,2}\tilde{y}_{t-1} + \beta_{3,3}\pi_t + u_{3,t}
\end{aligned}$$

where $\tilde{y}_t$ is output gap, π_t is the inflation rate, r_t is the short-term interest rate (intercepts omitted)

Structural VARs

This is SVAR with $P = 1$ and

$$\mathbf{y}_t = \begin{bmatrix} \tilde{y}_t \\ \pi_t \\ r_t \end{bmatrix}, \qquad \mathbf{u}_t = \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix},$$

$$\Psi_0 = \begin{bmatrix} 1 & -\beta_{1,2} & 0 \\ -\beta_{2,2} & 1 & -\beta_{2,3} \\ 0 & -\beta_{3,3} & 1 \end{bmatrix}, \qquad \Psi_1 = \begin{bmatrix} \beta_{1,1} & 0 & \beta_{1,3} \\ 0 & \beta_{2,1} & \beta_{2,4} \\ \beta_{3,2} & 0 & \beta_{3,1} \end{bmatrix}$$

But same identification issues occur in SVAR as in LSEM.

But can gain important insights about dynamic relationship of variables from (reduced form) VAR models so these will be our primary focus in this course.

VAR Analysis

- Many univariate concepts extend to *VAR*
- Stationarity is crucial, as for *AR*
- Need to extend concept of *white noise* to define *VAR*
- Representations of *VAR* important
  - Have vector *MA* (*VMA*) representation
  - But also other forms

First two moments of random vectors

Let $\mathbf{y}_t$ be $k \times 1$ random vector with i^{th} element $y_{i,t}$.

Mean: $E[\mathbf{y}_t]$

- is $k \times 1$ vector with i^{th} element $E[y_{i,t}]$;

Variance-covariance matrix: $Var[\mathbf{y}_t]$

- is a $k \times k$ matrix with $(i,j)^{th}$ element $Cov[y_{i,t}, y_{j,t}]$;
- diagonal elements are variances of $y_{i,t}$ for $i = 1, 2, \ldots k$;

Autocovariance matrices: $Cov[\mathbf{y}_t, \mathbf{y}_s]$

- $(i,j)^{th}$ element $Cov[y_{i,t}, y_{j,s}]$

Variance-Covariance Matrices

Using $E[\cdot]$, we can define: $Var[\mathbf{y}_t] = E[\mathbf{y}_t - E(\mathbf{y}_t)][\mathbf{y}_t - E(\mathbf{y}_t)]'$

For bivariate case

$$\begin{aligned} Var[\mathbf{y}_t] &= E\begin{bmatrix} y_{1,t} - E(y_{1,t}) \\ y_{2,t} - E(y_{2,t}) \end{bmatrix} \begin{bmatrix} y_{1,t} - E(y_{1,t}) & y_{2,t} - E(y_{2,t}) \end{bmatrix} \\ &= \begin{bmatrix} var(y_{1,t}) & cov(y_{1,t}, y_{2,t}) \\ cov(y_{1,t}, y_{2,t}) & var(y_{2,t}) \end{bmatrix} \end{aligned}$$

Variance-Covariance matrix is symmetric & positive definite

- Positive definiteness ensures variance of any linear combination of elements is positive

Vector White Noise

- Disturbance ε_t in $AR(p)$ is white noise
- $VAR(P)$ has **vector white noise** disturbance ε_t
- Vector white noise process ε_t $(k \times 1)$ satisfies:

$$\begin{aligned} E[\varepsilon_t] &= \mathbf{0} \quad \text{all } t \\ Var[\varepsilon_t] &= \Sigma \quad \text{all } t \\ Cov[\varepsilon_t, \varepsilon_s'] &= \mathbf{0} \quad \text{all } t \,\&\, s,\ \ s \neq t \end{aligned}$$

  - Zero auto & cross-covariances: $Cov[\varepsilon_{i,t}, \varepsilon_{j,s}] = 0$, $t \neq s$, all i, j
  - But $Cov[\varepsilon_{i,t}, \varepsilon_{j,t}] = \sigma_{i,j} \neq 0$, in general
  - Σ is a positive definite covariance matrix
    - Disturbances in different equations typically correlated at t
    - But not across different time periods

VAR(2) Example: Lag Operator Notation

Bivariate $VAR(2)$

$$\begin{aligned}\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} \\ + \begin{bmatrix} \phi_{11}^{(2)} & \phi_{12}^{(2)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}\end{aligned}$$

Define following matrices (& corresponding vectors)

$$\Phi_1 = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{bmatrix}, \quad \Phi_2 = \begin{bmatrix} \phi_{11}^{(2)} & \phi_{12}^{(2)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} \end{bmatrix}$$

In matrices

$$\begin{aligned} \mathbf{y}_t &= \alpha + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \varepsilon_t \\ (1 - \Phi_1 L - \Phi_2 L^2)\mathbf{y}_t &= \alpha + \varepsilon_t \end{aligned}$$

Lag Operator Notation: VAR(P)

- $VAR(P)$ for $k \times 1$ $\mathbf{y}_t$:

$$\mathbf{y}_t = \alpha + \Phi_1 \mathbf{y}_{t-1} + \ldots + \Phi_P \mathbf{y}_{t-P} + \varepsilon_t$$

  with ε_t vector white noise

- Write using lag operator as

$$\begin{aligned} \Phi(L)\mathbf{y}_t &= \alpha + \varepsilon_t \\ \Phi(L) &= \mathbf{I}_k - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_P L^P \end{aligned}$$

- Polynomial $\Phi(L)$ captures all dynamic interactions
  - Both within and across variables
  - No dynamics left in disturbances
  - But contemporaneous covariances permitted: $E(\varepsilon_t \varepsilon_t') = \Sigma$

VAR Stationarity Definition

- *VAR* process is stationary if:

$$
\begin{aligned}
E(\mathbf{y}_t) &= \mu \quad \text{all } t \\
E(\mathbf{y}_t - \mu)(\mathbf{y}_t - \mu)' &= \Gamma_0 \quad \text{all } t \\
E(\mathbf{y}_t - \mu)(\mathbf{y}_{t-\ell} - \mu)' &= \Gamma_\ell \quad \text{all } t \text{ \& any } \ell = 1, 2, \ldots
\end{aligned}
$$

- Here:
  - Γ_0 is the covariance matrix of $\mathbf{y}_t$
  - Γ_ℓ is the **autocovariance matrix** at lag ℓ
- Technically definition is second-order stationarity but we simply refer to process as stationary

Autocovariance Matrix: Example

- With $k = 2$

$$\Gamma_\ell = E\begin{bmatrix} y_{1t}-\mu_1 \\ y_{2t}-\mu_2 \end{bmatrix}\begin{bmatrix} y_{1,t-\ell}-\mu_1 & y_{2,t-\ell}-\mu_2 \end{bmatrix}$$

$$= \begin{bmatrix} E(y_{1t}-\mu_1)(y_{1,t-\ell}-\mu_1) & E(y_{1t}-\mu_1)(y_{2,t-\ell}-\mu_2) \\ E(y_{2t}-\mu_2)(y_{1,t-\ell}-\mu_1) & E(y_{2t}-\mu_2)(y_{2,t-\ell}-\mu_2) \end{bmatrix}$$

Diagonal terms are autocovariances at lag ℓ

Off-diagonal terms are cross-covariances at lag ℓ

- Note: Γ_ℓ is NOT (generally) symmetric for $\ell \neq 0$

VMA Representation for VAR(1)

- Stationary $VAR(1)$:

  $$\mathbf{y}_t = \alpha + \Phi_1 \mathbf{y}_{t-1} + \varepsilon_t$$

  successively substituting on right-hand side

  $$\begin{aligned} \mathbf{y}_t &= \alpha + \Phi_1\{\alpha + \Phi_1 \mathbf{y}_{t-2} + \varepsilon_{t-1}\} + \varepsilon_t \\ &= \alpha(\mathbf{I}_k + \Phi_1 + \Phi_1^2 + \ldots) + \varepsilon_t + \Phi_1 \varepsilon_{t-1} + \Phi_1^2 \varepsilon_{t-2} + \ldots \end{aligned}$$

  using $\Phi_1^j \to \mathbf{0}$ as $j \to \infty$ (a consequence of stationarity)

  - Gives vector moving average, $VMA(\infty)$ representation of $VAR(1)$

- Equivalently, using lag operator:

  $$\begin{aligned} (\mathbf{I}_k - \Phi_1 L)\mathbf{y}_t &= \alpha + \varepsilon_t \\ \mathbf{y}_t &= (\mathbf{I}_k - \Phi_1 L)^{-1}\{\alpha + \varepsilon_t\} \\ &= (\mathbf{I}_k - \Phi_1)^{-1}\alpha + (\mathbf{I}_k + \Phi_1 L + \Phi_1^2 L^2 + \ldots)\varepsilon_t \\ &= \mu + \Theta(L)\varepsilon_t \end{aligned}$$

VMA Representation for VAR(P)

- *VAR*(1) generalises to *VAR*(*P*)

$$\begin{aligned}\mathbf{y}_t &= \alpha + \Phi_1 \mathbf{y}_{t-1} + \ldots + \Phi_P \mathbf{y}_{t-P} + \varepsilon_t \\ \Phi(L)\mathbf{y}_t &= \alpha + \varepsilon_t, \quad \Phi(L) = \mathbf{I}_k - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_P L^P\end{aligned}$$

  With stationarity, can successively substitute for lagged $\mathbf{y}_t$ on right-hand side

- Equivalent to inverting *VAR* coefficient polynomial in L

$$\begin{aligned}\mathbf{y}_t &= [\Phi(L)]^{-1}\{\alpha + \varepsilon_t\} \\ &= [\Phi(1)]^{-1}\alpha + [\Phi(L)]^{-1}\varepsilon_t \\ &= \mu + \Theta(L)\varepsilon_t\end{aligned}$$

- Coefficients Θ_i ($i = 1, 2, \ldots$) of $VMA(\infty)$ are functions of *VAR* coefficients $\Phi_1, \ldots, \Phi_P$

Next Week

- Continue with VAR
- Generalise other aspects of *AR* processes
  - Require *VARMA* representation
  - Then examine stationarity condition for *VAR*
  - Prediction, estimation, test for residual autocorrelation, etc!