Structural Vector Autoregression

Time Series Econometrics

Impulse Response Functions, Forecasts

- Errors across equations in the system are correlated.
  - $Cov(\epsilon_{it}\epsilon_{jt}) \neq 0$ in general.
  - Therefore a shock to ϵ_{it} contains information about all the other ϵ_{jt}'s, and so is not orthogonal to them in reality.
  - When we calculate the simple derivative, we cannot hold the other ϵ_{jt}'s constant without controlling for this.

Orthogonalized Impulse Response Functions

- If the equations are appropriately ordered, then shock of ϵ_{1t} has information about contemporaneous shocks $\epsilon_{2t}, \ldots \epsilon_{nt}$
- We might want to take account of this extra information if we learn ϵ_{1t} but have not yet learned the value of the other shocks.
- Another way to ask the question:
  - Suppose we have only period $t-1$ information $(\mathbf{y_{t-1}}, \mathbf{y_{t-2}}, \ldots)$.
  - Then suddenly we learn the value of one variable, say y_{1t}.
  - How does learning only y_{1t} change our forecast of $\mathbf{y_{t+s}}$?

Orthogonalized Impulse Response Functions

- Our VAR says:
- $\mathbf{y_{t+s}} = \mathbf{c_s} + \epsilon_{\mathbf{t+s}} + \Psi_1 \epsilon_{\mathbf{t+s-1}} + \cdots + \Psi_{s-1} \epsilon_{\mathbf{t+1}} + \Psi_s \mathbf{y_t} + \mathbf{F_{12}^{(s)}} \mathbf{y_{t-1}} + \cdots + \mathbf{F_{1p}^{(s)}} \mathbf{y_{t-p+1}}$
- For forecasting, $E(\epsilon_{\mathbf{t+k}}) = \mathbf{0}$, linear projection of $\mathbf{y_{t+s}}$ on period t information is
- $\hat{\mathbf{y}}_{\mathbf{t+s|t}} = \mathbf{c_s} + \Psi_s \mathbf{y_t} + \mathbf{F_{12}^{(s)}} \mathbf{y_{t-1}} + \cdots + \mathbf{F_{1p}^{(s)}} \mathbf{y_{t-p+1}}$
- Row i of this is our forecast of $y_{i,t+s}$ based on period t information.

Orthogonalized Impulse Response Functions

- What happens if we only knew period $t-1$ information and then learned ϵ_{1t},
  - ϵ_{1t} is correlated with $\epsilon_{2t}, \ldots, \epsilon_{nt}$ through the first column of Ω.
- Suppose our linear projection of $\mathbf{y_{t+s}}$ is based only on knowing ϵ_{1t} and all the $t-1$ information:
- $\hat{E}(\mathbf{y_{t+s}}|\epsilon_{1t}, \mathbf{y_{t-1}}, \mathbf{y_{t-2}}, \ldots) =$
  $\mathbf{c_s} + \Psi_s \hat{E}(\epsilon_{\mathbf{t}}|\epsilon_{1t}, \mathbf{y_{t-1}}, \mathbf{y_{t-2}}, \ldots) + \mathbf{F_{12}^{(s)}} \mathbf{y_{t-1}} + \cdots + \mathbf{F_{1p}^{(s)}} \mathbf{y_{t-p+1}}$
  - We use ϵ_{1t} to guess the rest of the period t information.

Orthogonalized Impulse Response Functions

- Let $\mathbf{a_1}\epsilon_{1t} = \hat{E}(\epsilon_\mathbf{t}|\epsilon_{1t})$
- $\hat{E}(\mathbf{y_{t+s}}|\epsilon_{1t}, \mathbf{y_{t-1}}, \mathbf{y_{t-2}}, \ldots) = \mathbf{c_s} + \Psi_s \mathbf{a_1}\epsilon_{1t} + \mathbf{F^{(s)}_{12}}\mathbf{y_{t-1}} + \cdots + \mathbf{F^{(s)}_{1p}}\mathbf{y_{t-p+1}}$
- So the effect of y_{1t} on our forecast of $\mathbf{y_{t+s}}$ if y_{1t} is all we know at period t is
  - $\frac{\partial \hat{E}(\mathbf{y_{t+s}}|\epsilon_{1t}, \mathbf{y_{t-1}}, \mathbf{y_{t-2}}, \ldots)}{\partial y_{1t}} = \Psi_s \mathbf{a_1}$
  - $\mathbf{a_1} = \frac{\partial \hat{E}(\epsilon_\mathbf{t}|\epsilon_{1t})}{\partial \epsilon_{1t}}$
- How do we estimate this?

Orthogonalized Impulse Response Functions

- Notice that $\mathbf{a_1}$ accounts for how ϵ_{1t} is correlated with $\epsilon_{2t}, \ldots, \epsilon_{nt}$ through the first column of Ω.
- Recall $\hat{\Omega} = \frac{1}{T} \sum \hat{\epsilon}_\mathbf{t} \hat{\epsilon}'_\mathbf{t}$
- We can factor this as $\hat{\Omega} = \hat{A}\hat{D}\hat{A}'$
- $\hat{A}$ is a lower triangular matrix with ones on the principle diagonal, and $\hat{D}$ is diagonal with all elements positive.
- $\hat{\mathbf{a}}_\mathbf{1}$ is the first column of $\hat{A}$

Orthogonalized Impulse Response Functions

- We could take the Cholesky factorization of $\hat{\Omega}$:
  - $\hat{\Omega} = \hat{P}\hat{P}' = \hat{A}\hat{D}\hat{A}' = \hat{A}\hat{D}^{\frac{1}{2}}\hat{D}^{\frac{1}{2}}\hat{A}'$
  - $\hat{P}$ is lower triangular with positive elements (standard deviations) along principle diagonal.
  - Let $\hat{D}^{\frac{1}{2}}$ be the diagonal matrix whose diagonal is the diagonal of $\hat{P}$ (and all other elements of $\hat{D}$ are zero).
  - Then $\hat{A} = \hat{P}\hat{D}^{-\frac{1}{2}}$

Orthogonalized Impulse Response Functions

- The orthogonalized impulse response function is $\hat{\Psi}_s \hat{\mathbf{a}}_1$.
  - Many statistical packages report $\hat{\Psi}_s \hat{\mathbf{p}}_1$ where $\hat{\mathbf{p}}_1$ is the first column of $\hat{P}$
  - This only differs by the scale.
  - $\hat{\Psi}_s \hat{\mathbf{a}}_1$ is the effect of a one-unit increase in y_{1t}
  - $\hat{\Psi}_s \hat{\mathbf{p}}_1$ is the effect of a one-standard deviation increase in y_{1t}.

Orthogonalized Impulse Response Functions

- Now what happens once I've learned y_{1t}, if I then learn y_{2t} but not y_{3t}, etc.?
  - $\hat{\Psi}_s \hat{\mathbf{a}}_2$ where $\hat{\mathbf{a}}_2$ is the second column of $\hat{A}$.
  - or use $\hat{\Psi}_s \hat{\mathbf{p}}_2$ in terms of standard deviations.
- More generally, these are called "recursively orthogonalized impulse response functions"
  - $\hat{\Psi}_s \hat{\mathbf{A}}$
- The value of this is different depending on the ordering of the equations.
  - How you order them depends on what question you want to answer.
  - Typically order them in the order in which you think information would be made known or in which you think effects happen.

Structural VAR

- VARs we have looked at so far have been “reduced form”: each equation in the system is a function of lags only: of its own value, and lags of other variables
- Maybe we can use/test structural information about contemporaneous relationships.
- This increases the number of parameters we need to estimate - only okay if the number of parameters is not too large to be recovered from reduced form estimates.
  - If the number of parameters is too large, we may impose some structural parameters in order to estimate others.

Structural VAR

- Consider a macroeconomic relationship
  - y_{1t} =real GDP growth
  - y_{2t} =inflation
  - y_{3t} =fed funds rate
  - y_{4t} =rate of growth of M2

Structural VAR

- Current spending (GDP growth) depends only on past shocks
  - $y_{1t} = k_1 + B_1^{(1,.)}\mathbf{y_{t-1}} + \cdots + u_{1t}$
- Inflation has a Phillips curve relation between spending and inflation:
  - $y_{2t} = k_2 + b_0^{(2,1)}y_{1t} + B_1^{(2,.)}\mathbf{y_{t-1}} + \cdots + u_{2t}$
- Federal reserve responds to current output growth and inflation:
  - $y_{3t} = k_3 + b_0^{(3,1)}y_{1t} + b_0^{(3,2)}y_{2t} + B_1^{(3,.)}\mathbf{y_{t-1}} + \cdots + u_{3t}$
- Money demand depends on current output, inflation, and interest rates:
  - $y_{3t} = k_3 + b_0^{(4,1)}y_{1t} + b_0^{(4,2)}y_{2t} + b_0^{(4,3)}y_{3t} + B_1^{(3,.)}\mathbf{y_{t-1}} + \cdots + u_{4t}$

Structural VAR

- We can stack these equations into a vector dynamic structural model:
  - $\mathbf{B_0 y_t} = \mathbf{k} + \mathbf{B_1 y_{t-1}} + \cdots + \mathbf{B_p y_{t-p}} + \mathbf{u_t}$
  - Any dynamic structural model that has a linear approximation can be written in this way.
  - $\mathbf{B_0} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ b_0^{(2,1)} & 1 & 0 & 0 \\ b_0^{(3,1)} & b_0^{(3,2)} & 1 & 0 \\ b_0^{(4,1)} & b_0^{(4,2)} & b_0^{(4,3)} & 1 \end{bmatrix}$
  - The lower triangular form is a special case for this model

Structural VAR

- If we premultiply by $\mathbf{B_0^{-1}}$, we get the reduced form:
  - $\underset{nX1}{\mathbf{y_t}} = \mathbf{c} + \Phi_1 \mathbf{y_{t-1}} + ... + \Phi_p \mathbf{y_{t-p}} + \epsilon_\mathbf{t}$
  - $\mathbf{c} = \mathbf{B_0^{-1}k}$
  - $\Phi_j = \mathbf{B_0^{-1}B_j}$
  - $\epsilon_\mathbf{t} = \mathbf{B_0^{-1}u_t}$

Structural VAR

- The reduced form VAR is just another representation of a linearized dynamic structural model.
  - If we knew the structural parameters $\mathbf{B}$ and shocks $\mathbf{u}$ we could calculate the VAR parameters Φ and ϵ and the associated impulse response functions.
  - $\frac{\partial \mathbf{y_{t+m}}}{\partial \mathbf{u_t'}} = \frac{\partial \mathbf{y_{t+m}}}{\partial \epsilon_\mathbf{t}'} \frac{\partial \epsilon_\mathbf{t}}{\partial \mathbf{u_t'}} = \Psi_m \mathbf{B_0}^{-1}$

Structural VAR

- So can we estimate the reduced form and recover the structural parameters?
  - Reduced form VAR parameters: $n + n(n+1)/2 + n^2 p$
  - Structural VAR parameters: $n + n(n+1)/2 + n^2 p + n^2$
  - Need enough restrictions on $\mathbf{B_0}$ (e.g., zeros for some elements) and $E(\mathbf{u_t u_t'})$ (e.g., diagonal matrix) in order for their to be a one-to-one mapping.
  - The earlier lower triangular form, with diagonal $E(\mathbf{u_t u_t'})$, is one such example.

Structural VAR

- If there is a one-to-one mapping, then estimate reduced form and figure out the appropriate transformation of Φ and Ω
  - For example in the case above:
  - $\epsilon_{\mathbf{t}} = \mathbf{B_0^{-1} u_t}$
  - $\Omega = \mathbf{B_0^{-1} D (B_0^{-1})'}$
  - Find triangular factorization of $\hat{\Omega}$ as before.
  - When the structural equation is lower triangular, this is the same as the recursively orthogonalized impulse response function.

Structural VAR

- For non-recursive structural models:
  - Let $E(\mathbf{u_t u_t'}) = \mathbf{D}$
  - Estimate reduced form by OLS, put enough restrictions on $\mathbf{D}, \mathbf{B_0}, \ldots, \mathbf{B_p}$ so that there is a unique mapping from $\Omega, \Phi_1, \ldots, \Phi_p$.
  - Typical approach: let $\mathbf{B_1}, \ldots, \mathbf{B_p}$ be unrestricted.
  - $E(\epsilon_\mathbf{t} \epsilon_\mathbf{t}') = \Omega = \mathbf{B_0^{-1} D (B_0^{-1})'}$
  - $\hat{\Omega}$ has $n(n+1)/2$ distinct elements. so we can have just that many unknowns in $\mathbf{B_0}$ and $\mathbf{D}$
  - Need to impose as many structural restrictions from theory, long run equilibrium, etc., to reduce parameters to a number that can be recovered from the reduced form (or estimated directly using MLE).