Structural VARs

Structural Representation

Consider the structural VAR (SVAR) model

$$\begin{aligned} y_{1t} &= \gamma_{10} - b_{12}y_{2t} + \gamma_{11}y_{1t-1} + \gamma_{12}y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= \gamma_{20} - b_{21}y_{1t} + \gamma_{21}y_{1t-1} + \gamma_{22}y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

where

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim \text{iid} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right).$$

Remarks:

- ε_{1t} and ε_{2t} are called structural errors
- In general, $\text{cov}(y_{2t}, \varepsilon_{1t}) \neq 0$ and $\text{cov}(y_{1t}, \varepsilon_{2t}) \neq 0$
- All variables are endogenous - OLS is not appropriate!

In matrix form, the model becomes

$$\begin{aligned}
&\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \\
= \ &\begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}
\end{aligned}$$

or

$$\begin{aligned}
\mathbf{B}\mathbf{y}_t &= \boldsymbol{\gamma}_0 + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\
E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] &= \mathbf{D} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}
\end{aligned}$$

In lag operator notation, the SVAR is

$$\begin{aligned}
\mathbf{B}(L)\mathbf{y}_t &= \boldsymbol{\gamma}_0 + \boldsymbol{\varepsilon}_t, \\
\mathbf{B}(L) &= \mathbf{B} - \boldsymbol{\Gamma}_1 L.
\end{aligned}$$

Reduced Form Representation

Solve for $\mathbf{y}_t$ in terms of $\mathbf{y}_{t-1}$ and $\boldsymbol{\varepsilon}_t$:

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{B}^{-1}\gamma_0 + \mathbf{B}^{-1}\mathbf{\Gamma}_1\mathbf{y}_{t-1} + \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t \\
&= \mathbf{a}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{u}_t \\
\mathbf{a}_0 &= \mathbf{B}^{-1}\gamma_0, \mathbf{A}_1 = \mathbf{B}^{-1}\mathbf{\Gamma}_1, \mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t
\end{aligned}$$

or

$$\begin{aligned}
\mathbf{A}(L)\mathbf{y}_t &= \mathbf{a}_0 + \mathbf{u}_t \\
\mathbf{A}(L) &= \mathbf{I}_2 - \mathbf{A}_1 L
\end{aligned}$$

Note that

$$\mathbf{B}^{-1} = \frac{1}{\Delta}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix},\ \Delta = \det(\mathbf{B}) = 1 - b_{12}b_{21}$$

The reduced form errors $\mathbf{u}_t$ are linear combinations of the structural errors $\boldsymbol{\varepsilon}_t$ and have covariance matrix

$$\begin{aligned}
E[\mathbf{u}_t\mathbf{u}_t'] &= \mathbf{B}^{-1}E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t']\mathbf{B}^{-1\prime} \\
&= \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1\prime} \\
&= \mathbf{\Omega}.
\end{aligned}$$

Remark: Parameters of RF may be estimated by OLS equation by equation

Identification Issues

Without some restrictions, the parameters in the SVAR are not identified. That is, given values of the reduced form parameters $\mathbf{a}_0, \mathbf{A}_1$ and $\mathbf{\Omega}$, it is not possible to uniquely solve for the structural parameters $\mathbf{B}, \boldsymbol{\gamma}_0, \mathbf{\Gamma}_1$ and $\mathbf{D}$.

- 10 structural parameters and 9 reduced form parameters

- Order condition requires at least 1 restriction on the SVAR parameters

Typical identifying restrictions include

- Zero (exclusion) restrictions on the elements of $\mathbf{B}$; e.g., $b_{12} = 0$.

- Linear restrictions on the elements of $\mathbf{B}$; e.g., $b_{12} + b_{21} = 1$.

MA Representations

Wold representation

Multiplying both sides of reduced form by $\mathbf{A}(L)^{-1} = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}$ to give

$$\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\mathbf{u}_t \\
\boldsymbol{\Psi}(L) &= (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1} \\
&= \sum_{k=0}^{\infty} \boldsymbol{\Psi}_k L^k,\ \ \boldsymbol{\Psi}_0 = \mathbf{I}_2, \boldsymbol{\Psi}_k = \mathbf{A}_1^k \\
\boldsymbol{\mu} &= \mathbf{A}(1)^{-1}\mathbf{a}_0 \\
E[\mathbf{u}_t \mathbf{u}_t'] &= \boldsymbol{\Omega}
\end{aligned}$$

Remark: Wold representation may be estimated using RF VAR estimates

Structural moving average (SMA) *representation*

SMA of $\mathbf{y}_t$ is based on an infinite moving average of the structural innovations $\boldsymbol{\varepsilon}_t$. Using $\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t$ in the Wold form gives

$$\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\mathbf{B}^{-1}\boldsymbol{\varepsilon}_t \\
&= \boldsymbol{\mu} + \boldsymbol{\Theta}(L)\boldsymbol{\varepsilon}_t \\
\boldsymbol{\Theta}(L) &= \sum_{k=0}^{\infty} \boldsymbol{\Theta}_k L^k \\
&= \boldsymbol{\Psi}(L)\mathbf{B}^{-1} \\
&= \mathbf{B}^{-1} + \boldsymbol{\Psi}_1\mathbf{B}^{-1}L + \cdots
\end{aligned}$$

That is,

$$\begin{aligned}
\boldsymbol{\Theta}_k &= \boldsymbol{\Psi}_k\mathbf{B}^{-1} = \mathbf{A}_1^k\mathbf{B}^{-1},\ k = 0, 1, \ldots. \\
\boldsymbol{\Theta}_0 &= \mathbf{B}^{-1} \neq \mathbf{I}_2
\end{aligned}$$

Example: SMA for bivariate system

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} \\ \theta_{21}^{(1)} & \theta_{22}^{(1)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t-1} \\ \varepsilon_{2t-1} \end{bmatrix} + \cdots$$

Notes

- $\mathbf{\Theta}_0 = \mathbf{B}^{-1} \neq \mathbf{I}_2$. $\mathbf{\Theta}_0$ captures initial impacts of structural shocks, and determines the contemporaneous correlation between y_{1t} and y_{2t}.

- Elements of the $\mathbf{\Theta}_k$ matrices, $\theta_{ij}^{(k)}$, give the dynamic multipliers or impulse responses of y_{1t} and y_{2t} to changes in the structural errors ε_{1t} and ε_{2t}.

Impulse Response Functions

Consider the SMA representation at time $t+s$

$$\begin{bmatrix} y_{1t+s} \\ y_{2t+s} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \cdots$$
$$+ \begin{bmatrix} \theta_{11}^{(s)} & \theta_{12}^{(s)} \\ \theta_{21}^{(s)} & \theta_{22}^{(s)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \cdots .$$

The *structural dynamic multipliers* are

$$\frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}} = \theta_{11}^{(s)}, \frac{\partial y_{1t+s}}{\partial \varepsilon_{2t}} = \theta_{12}^{(s)}$$
$$\frac{\partial y_{2t+s}}{\partial \varepsilon_{1t}} = \theta_{21}^{(s)}, \frac{\partial y_{2t+s}}{\partial \varepsilon_{2t}} = \theta_{22}^{(s)}$$

The *structural impulse response functions* (IRFs) are the plots of $\theta_{ij}^{(s)}$ vs. s for $i, j = 1, 2$. These plots summarize how unit impulses of the structural shocks at time t impact the level of y at time $t+s$ for different values of s.

Stationarity of y_t implies

$$\lim_{s \to \infty} \theta_{ij}^{(s)} = 0, \; i, j = 1, 2$$

The *long-run cumulative impact* of the structural shocks is captured by

$$\begin{aligned}
\mathbf{\Theta}(1) &= \begin{bmatrix} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{\infty} \theta_{11}^{(s)} & \sum_{s=0}^{\infty} \theta_{12}^{(s)} \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} & \sum_{s=0}^{\infty} \theta_{22}^{(s)} \end{bmatrix} \\
\mathbf{\Theta}(L) &= \begin{bmatrix} \theta_{11}(L) & \theta_{12}(L) \\ \theta_{21}(L) & \theta_{22}(L) \end{bmatrix} \\
&= \begin{bmatrix} \sum_{s=0}^{\infty} \theta_{11}^{(s)} L^s & \sum_{s=0}^{\infty} \theta_{12}^{(s)} L^s \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} L^s & \sum_{s=0}^{\infty} \theta_{22}^{(s)} L^s \end{bmatrix}
\end{aligned}$$

Digression: Dynamic Regression Models

In the SVAR every variable is engodenous. Suppose, for example, y_{2t} is strictly exogenous which implies $b_{21} = 0$ and $\gamma_{21} = 0$. Then, the first equation is an ADL(1,1)

$$\begin{aligned} y_{1t} &= \alpha + \phi y_{1t-1} + \beta_0 y_{2t} + \beta_1 y_{2t-1} + \varepsilon_{1t} \\ \text{cov}(y_{2t}, \varepsilon_{1t}) &= 0 \end{aligned}$$

In lag operator notation the equation becomes

$$\begin{aligned} \phi(L) y_{1t} &= \alpha + \beta(L) y_{2t} + \varepsilon_{1t} \\ \phi(L) &= 1 - \phi L,\ \beta(L) = \beta_0 + \beta_1 L \end{aligned}$$

The second equation is an AR(1) model for y_{2t}

$$y_{2t} = c + \rho y_{2t-1} + \varepsilon_{2t}$$

Stationarity now only requires $|\phi| < 1$ and $|\rho| < 1$.

The first equation may then be solved for y_{1t} as a function of y_{2t} and ε_{1t}

$$
\begin{aligned}
y_{1t} &= \frac{\alpha}{\phi(1)} + \phi(L)^{-1}\beta(L)y_{2t} + \phi(L)^{-1}\varepsilon_{1t} \\
&= \mu + \psi_\beta(L)y_{2t} + \psi(L)\varepsilon_t \\
\mu &= \frac{\alpha}{\phi(1)} \\
\psi_\beta(L) &= \phi(L)^{-1}\beta(L),\ \psi(L) = \phi(L)^{-1}
\end{aligned}
$$

Since y_{2t} is exogenous, we have two sources of shocks. Note: there can be four types of dynamic multipliers :

$$
\frac{\partial y_{1t+s}}{\partial y_{2t}},\ \frac{\partial y_{1t+s}}{\partial \varepsilon_{2t}},\ \frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}},\ \frac{\partial y_{2t+s}}{\partial \varepsilon_{2t}}
$$

The short-run dyamic multipliers with respect to y_{2t} and ε_{1t} are

$$\begin{aligned}
\frac{\partial y_{1t+s}}{\partial y_{2t}} &= \frac{\partial y_{1t}}{\partial y_{2t-s}} = \psi_{\beta,s} \\
\frac{\partial y_{1t+s}}{\partial \varepsilon_{1t}} &= \frac{\partial y_{1t}}{\partial \varepsilon_{1t-s}} = \psi_{s}
\end{aligned}$$

In the *steady state* or *long-run equilibrium* all variables are constant

$$\begin{aligned}
y_1^* &= \mu + \psi_\beta(L) y_2^* = \mu + \psi_\beta(1) y_2^* \\
y_2^* &= \frac{c}{1-\rho} \\
\psi_\beta(1) &= \phi(1)^{-1}\beta(1) = \frac{\beta_0 + \beta_1}{1-\phi}
\end{aligned}$$

The long-run impact of a change in y_2 on y_1 is then

$$\frac{\partial y_1^*}{\partial y_2^*} = \psi_\beta(1) = \frac{\beta_0 + \beta_1}{1-\phi} = \sum_{s=0}^{\infty} \frac{\partial y_{1t+s}}{\partial y_{2t}}$$

Identification issues

In some applications, identification of the parameters of the SVAR is achieved through restrictions on the parameters of the SMA representation.

Identification through contemporaneous restrictions

Suppose that ε_{2t} has no contemporaneous impact on y_{1t}. Then $\theta_{12}^{(0)} = 0$ and

$$\mathbf{\Theta}_0 = \begin{bmatrix} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix}.$$

Since $\mathbf{\Theta}_0 = \mathbf{B}^{-1}$ then

$$\begin{aligned} \begin{bmatrix} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} &= \frac{1}{\Delta}\begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \\ &\Rightarrow \; b_{12} = 0 \end{aligned}$$

Hence, assuming $\theta_{12}^{(0)} = 0$ in the SMA representation is equivalent to assuming $b_{12} = 0$ in the SVAR representation.

Identification through long-run restrictions

Suppose ε_{2t} has no long-run cumulative impact on y_{1t}. Then

$$\begin{aligned}\theta_{12}(1) &= \sum_{s=0}^{\infty} \theta_{12}^{(s)} = 0 \\ \boldsymbol{\Theta}(1) &= \begin{bmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix}.\end{aligned}$$

This type of long-run restriction places nonlinear restrictions on the coefficients of the SVAR since

$$\begin{aligned}\boldsymbol{\Theta}(1) &= \boldsymbol{\Psi}(1)\mathbf{B}^{-1} = \mathbf{A}(1)^{-1}\mathbf{B}^{-1} \\ &= (\mathbf{I}_2 - \mathbf{B}^{-1}\boldsymbol{\Gamma}_1)^{-1}\mathbf{B}^{-1}\end{aligned}$$

Estimation Issues

In order to compute the structural IRFs, the parameters of the SMA representation need to be estimated. Since

$$\begin{aligned}\boldsymbol{\Theta}(L) &= \boldsymbol{\Psi}(L)\mathbf{B}^{-1} \\ \boldsymbol{\Psi}(L) &= \mathbf{A}(L)^{-1} = (\mathbf{I}_2 - \mathbf{A}_1 L)^{-1}\end{aligned}$$

the estimation of the elements in $\boldsymbol{\Theta}(L)$ can often be broken down into steps:

- $\mathbf{A}_1$ is estimated from the reduced form VAR.

- Given $\widehat{\mathbf{A}_1}$, the matrices in $\boldsymbol{\Psi}(L)$ can be estimated using $\widehat{\boldsymbol{\Psi}}_k = \widehat{\mathbf{A}}_1^k$.

- $\mathbf{B}$ is estimated from the identified SVAR.

- Given $\hat{\mathbf{B}}$ and $\hat{\boldsymbol{\Psi}}_k$, the estimates of $\boldsymbol{\Theta}_k$, $k = 0, 1, \ldots,$ are given by $\hat{\boldsymbol{\Theta}}_k = \widehat{\boldsymbol{\Psi}}_k \hat{\mathbf{B}}^{-1}$.

Forecast Error Variance Decompositions

Idea: determine the proportion of the variability of the errors in forecasting y_1 and y_2 at time $t+s$ based on information available at time t that is due to variability in the structural shocks ε_1 and ε_2 between times t and $t+s$.

To derive the FEVD, start with the Wold representation for $\mathbf{y}_{t+s}$

$$\begin{aligned}\mathbf{y}_{t+s} &= \boldsymbol{\mu} + \mathbf{u}_{t+s} + \boldsymbol{\Psi}_1 \mathbf{u}_{t+s-1} + \cdots \\ &\quad + \boldsymbol{\Psi}_{s-1}\mathbf{u}_{t+1} + \boldsymbol{\Psi}_s \mathbf{u}_t + \boldsymbol{\Psi}_{s+1}\mathbf{u}_{t-1} + \cdots .\end{aligned}$$

The best linear forecast of $\mathbf{y}_{t+s}$ based on information available at time t is

$$\mathbf{y}_{t+s|t} = \boldsymbol{\mu} + \boldsymbol{\Psi}_s \mathbf{u}_t + \boldsymbol{\Psi}_{s+1}\mathbf{u}_{t-1} + \cdots$$

and the forecast error is

$$\mathbf{y}_{t+s} - \mathbf{y}_{t+s|t} = \mathbf{u}_{t+s} + \boldsymbol{\Psi}_1 \mathbf{u}_{t+s-1} + \cdots + \boldsymbol{\Psi}_{s-1}\mathbf{u}_{t+1}.$$

Using

$$\boldsymbol{\varepsilon}_t = \mathbf{B}^{-1}\mathbf{u}_t, \; \boldsymbol{\Theta}_k = \boldsymbol{\Psi}_k \mathbf{B}^{-1}$$

The forecast error in terms of the structural shocks is

$$\begin{aligned}
\mathbf{y}_{t+s} - \mathbf{y}_{t+s|t} &= \mathbf{B}^{-1}\boldsymbol{\varepsilon}_{t+s} + \boldsymbol{\Psi}_1 \mathbf{B}^{-1}\boldsymbol{\varepsilon}_{t+s-1} + \\
&\quad \cdots + \boldsymbol{\Psi}_{s-1}\mathbf{B}^{-1}\boldsymbol{\varepsilon}_{t+1} \\
&= \boldsymbol{\Theta}_0 \boldsymbol{\varepsilon}_{t+s} + \boldsymbol{\Theta}_1 \boldsymbol{\varepsilon}_{t+s-1} + \cdots + \boldsymbol{\Theta}_{s-1}\boldsymbol{\varepsilon}_{t+1}
\end{aligned}$$

The forecast errors equation by equation are

$$\begin{bmatrix} y_{1t+s} - y_{1t+s|t} \\ y_{2t+s} - y_{2t+s|t} \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+s} \\ \varepsilon_{2t+s} \end{bmatrix} + \\
\cdots + \begin{bmatrix} \theta_{11}^{(s-1)} & \theta_{12}^{(s-1)} \\ \theta_{21}^{(s-1)} & \theta_{22}^{(s-1)} \end{bmatrix} \begin{bmatrix} \varepsilon_{1t+1} \\ \varepsilon_{2t+1} \end{bmatrix}$$

For the first equation

$$
\begin{aligned}
y_{1t+s} - y_{1t+s|t} &= \theta_{11}^{(0)}\varepsilon_{1t+s} + \cdots + \theta_{11}^{(s-1)}\varepsilon_{1t+1} \\
&\quad + \theta_{12}^{(0)}\varepsilon_{2t+s} + \cdots + \theta_{12}^{(s-1)}\varepsilon_{2t+1}
\end{aligned}
$$

Since it is assumed that $\boldsymbol{\varepsilon}_t \sim i.i.d.\ (0, \mathbf{D})$ where $\mathbf{D}$ is diagonal, the variance of the forecast error in may be decomposed as

$$
\begin{gathered}
var(y_{1t+s} - y_{1t+s|t}) = \sigma_1^2(s) \\
= \sigma_1^2\left(\left(\theta_{11}^{(0)}\right)^2 + \cdots + \left(\theta_{11}^{(s-1)}\right)^2\right) \\
+ \sigma_2^2\left(\left(\theta_{12}^{(0)}\right)^2 + \cdots + \left(\theta_{12}^{(s-1)}\right)^2\right).
\end{gathered}
$$

The proportion of $\sigma_1^2(s)$ due to shocks in ε_1 is then

$$
\rho_{1,1}(s) = \frac{\sigma_1^2\left(\left(\theta_{11}^{(0)}\right)^2 + \cdots + \left(\theta_{11}^{(s-1)}\right)^2\right)}{\sigma_1^2(s)}
$$

the proportion of $\sigma_1^2(s)$ due to shocks in ε_2 is

$$\rho_{1,2}(s) = \frac{\sigma_2^2 \left(\left(\theta_{12}^{(0)}\right)^2 + \cdots + \left(\theta_{12}^{(s-1)}\right)^2 \right)}{\sigma_1^2(s)}.$$

The forecast error variance decompositions (FEVDs) for y_{2t+s} are

$$\rho_{2,1}(s) = \frac{\sigma_1^2\left(\left(\theta_{21}^{(0)}\right)^2 + \cdots + \left(\theta_{21}^{(s-1)}\right)^2\right)}{\sigma_2^2(s)},$$

$$\rho_{2,2}(s) = \frac{\sigma_2^2\left(\left(\theta_{22}^{(0)}\right)^2 + \cdots + \left(\theta_{22}^{(s-1)}\right)^2\right)}{\sigma_2^2(s)},$$

where

$$var(y_{2t+s} - y_{2t+s|t}) = \sigma_2^2(s)$$
$$= \sigma_1^2\left(\left(\theta_{21}^{(0)}\right)^2 + \cdots + \left(\theta_{21}^{(s-1)}\right)^2\right)$$
$$+\sigma_2^2\left(\left(\theta_{22}^{(0)}\right)^2 + \cdots + \left(\theta_{22}^{(s-1)}\right)^2\right).$$

Identification Using Recursive Causal Orderings

Consider the bivariate SVAR. We need at least one restriction on the parameters for identification. Suppose $b_{12} = 0$ so that $\mathbf{B}$ is lower triangular. That is,

$$\begin{aligned} \mathbf{B} &= \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix} \\ \mathbf{B}^{-1} &= \mathbf{\Theta}_0 = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix} \end{aligned}$$

The SVAR model becomes the recursive model

$$\begin{aligned} y_{1t} &= \gamma_{10} + \gamma_{11} y_{1t-1} + \gamma_{12} y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= \gamma_{20} - b_{21} y_{1t} + \gamma_{21} y_{1t-1} + \gamma_{22} y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

The recursive model imposes the restriction that the value y_{2t} does not have a contemporaneous effect on y_{1t}. Since $b_{21} \neq 0$ a priori we allow for the possibility that y_{1t} has a contemporaneous effect on y_{2t}.

The reduced form VAR errors $\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t$ become

$$\begin{aligned}\mathbf{u}_t &= \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \\ &= \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} - b_{21}\varepsilon_{1t} \end{bmatrix}.\end{aligned}$$

Claim: The restriction $b_{12} = 0$ is sufficient to just identify b_{21} and, hence, just identify $\mathbf{B}$.

To establish this result, we show how b_{21} can be uniquely identified from the elements of the reduced form covariance matrix $\mathbf{\Omega}$. Note

$$\begin{bmatrix} \omega_1^2 & \omega_{12} \\ \omega_{12} & \omega_2^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 & -b_{21} \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} \sigma_1^2 & -b_{21}\sigma_1^2 \\ -b_{21}\sigma_1^2 & \sigma_2^2 + b_{21}^2\sigma_1^2 \end{bmatrix}.$$

Then, we can solve for b_{21} via

$$b_{21} = -\frac{\omega_{12}}{\omega_1^2} = -\rho\frac{\omega_2}{\omega_1},$$

where $\rho = \omega_{12}/\omega_1\omega_2$ is the correlation between u_1 and u_2. Notice that $b_{21} \neq 0$ provided $\rho \neq 0$.

Estimation Procedure

1. Estimate the reduced form VAR by OLS equation by equation:

$$
\begin{aligned}
\mathbf{y}_t &= \hat{\mathbf{a}}_0 + \widehat{\mathbf{A}}_1 \mathbf{y}_{t-1} + \hat{\mathbf{u}}_t \\
\widehat{\mathbf{\Omega}} &= \frac{1}{T}\sum_{t=1}^{T} \hat{\mathbf{u}}_t \hat{\mathbf{u}}_t'
\end{aligned}
$$

2. Estimate b_{21} and $\mathbf{B}$ from $\widehat{\mathbf{\Omega}}$:

$$
\begin{aligned}
\hat{b}_{21} &= -\frac{\widehat{\omega}_{12}}{\widehat{\omega}_1^2}, \\
\widehat{\mathbf{B}} &= \begin{bmatrix} 1 & 0 \\ \hat{b}_{21} & 1 \end{bmatrix}.
\end{aligned}
$$

3. Estimate SMA from estimates of $\mathbf{a}_0, \mathbf{A}_1$ and $\mathbf{B}$:

$$
\begin{aligned}
\mathbf{y}_t &= \widehat{\boldsymbol{\mu}} + \widehat{\mathbf{\Theta}}(L)\widehat{\boldsymbol{\varepsilon}}_t \\
\widehat{\boldsymbol{\mu}} &= \hat{\mathbf{a}}_0(\mathbf{I}_2 - \widehat{\mathbf{A}}_1)^{-1} \\
\widehat{\mathbf{\Theta}}_k &= \widehat{\mathbf{A}}_1^k \widehat{\mathbf{B}}^{-1}, k = 0, 1, \ldots \\
\widehat{\mathbf{D}} &= \widehat{\mathbf{B}}\widehat{\mathbf{\Omega}}\widehat{\mathbf{B}}'.
\end{aligned}
$$

Remark:

Above procedure is numerically equivalent to estimating the triangular system by OLS equation by equation:

$$\begin{aligned} y_{1t} &= \gamma_{10} + \gamma_{11} y_{1t-1} + \gamma_{12} y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= \gamma_{20} - b_{21} y_{1t} + \gamma_{21} y_{1t-1} + \gamma_{22} y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

Why? Since $\text{cov}(\varepsilon_{1t}, \varepsilon_2) = 0$ by assumption, $\text{cov}(y_{1t}, \varepsilon_{2t}) = 0$

Recovering the SMA representation using the Choleski Factorization of Ω.

Claim: The SVAR representation based on a recursive causal ordering may be computed using the Choleski factorization of the reduced form covariance matrix $\mathbf{\Omega}$.

Recall, the *Choleski factorization* of the positive semi-definite matrix Ω is given by

$$\begin{aligned} \mathbf{\Omega} &= \mathbf{P}\mathbf{P}' \\ \mathbf{P} &= \begin{bmatrix} p_{11} & 0 \\ p_{21} & p_{22} \end{bmatrix} \end{aligned}$$

A closely related factorization obtained from the Choleski factorization is the *triangular factorization*

$$\mathbf{\Omega} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}'$$

$$\mathbf{T} = \begin{bmatrix} 1 & 0 \\ t_{21} & 1 \end{bmatrix}, \ \mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

$$\lambda_i \geq 0, i = 1, 2.$$

Consider the reduced form VAR

$$\mathbf{y}_t = \mathbf{a}_0 + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{u}_t,$$
$$\mathbf{\Omega} = E[\mathbf{u}_t\mathbf{u}_t']$$
$$\mathbf{\Omega} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}'$$

Construct a pseudo SVAR model by premultiplying by $\mathbf{T}^{-1}$:

$$\mathbf{T}^{-1}\mathbf{y}_t = \mathbf{T}^{-1}\mathbf{a}_0 + \mathbf{T}^{-1}\mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{T}^{-1}\mathbf{u}_t$$

or

$$\mathbf{B}\mathbf{y}_t = \gamma_0 + \mathbf{\Gamma}_1\mathbf{y}_{t-1} + \varepsilon_t$$

where

$$\mathbf{B} = \mathbf{T}^{-1},\ \gamma_0 = \mathbf{T}^{-1}\mathbf{a}_0,$$
$$\mathbf{\Gamma}_1 = \mathbf{T}^{-1}\mathbf{A}_1, \varepsilon_t = \mathbf{T}^{-1}\mathbf{u}_t.$$

The pseudo structural errors $\boldsymbol{\varepsilon}_t$ have a diagonal covariance matrix $\mathbf{\Lambda}$

$$\begin{aligned} E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] &= \mathbf{T}^{-1}E[\mathbf{u}_t\mathbf{u}_t']\mathbf{T}^{-1\prime} \\ &= \mathbf{T}^{-1}\mathbf{\Omega}\mathbf{T}^{-1\prime} \\ &= \mathbf{T}^{-1}\mathbf{T}\mathbf{\Lambda}\mathbf{T}'\mathbf{T}^{-1\prime} \\ &= \mathbf{\Lambda}. \end{aligned}$$

In the pseudo SVAR,

$$\begin{aligned} \mathbf{B} &= \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix} = \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 \\ -t_{21} & 1 \end{bmatrix} \\ b_{12} &= 0,\ b_{21} = -t_{21} \end{aligned}$$

Ordering of Variables

The identification of the SVAR using the triangular factorization depends on the ordering of the variables in $\mathbf{y}_t$. In the above analysis, it is assumed that $\mathbf{y}_t = (y_{1t}, y_{2t})'$ so that y_{1t} comes first in the ordering of the variables. When the triangular factorization is conducted and the pseudo SVAR is computed the structural $\mathbf{B}$ matrix is

$$\begin{aligned} \mathbf{B} &= \mathbf{T}^{-1} = \left[\begin{array}{cc} 1 & 0 \\ b_{21} & 1 \end{array} \right] \\ &\Rightarrow b_{12} = 0 \end{aligned}$$

If the ordering of the variables is reversed, $\mathbf{y}_t = (y_{2t}, y_{1t})'$, then the recursive causal ordering of the SVAR is reversed and the structural $\mathbf{B}$ matrix becomes

$$\begin{aligned} \mathbf{B} &= \mathbf{T}^{-1} = \left[\begin{array}{cc} 1 & 0 \\ b_{12} & 1 \end{array} \right] \\ &\Rightarrow b_{21} = 0 \end{aligned}$$

Sensitivity Analysis

- Ordering of the variables in $\mathbf{y}_t$ determines the recursive causal structure of the SVAR,

- This identification assumption is not testable

- Sensitivity analysis is often performed to determine how the structural analysis based on the IRFs and FEVDs are influenced by the assumed causal ordering.

- This sensitivity analysis is based on estimating the SVAR for different orderings of the variables.

- If the IRFs and FEVDs change considerably for different orderings of the variables in y_t then it is clear that the assumed recursive causal structure heavily influences the structural inference.

Residual Analysis

One way to determine if the assumed causal ordering influences the structural inferences is to look at the residual covariance matrix $\hat{\mathbf{\Omega}}$ from the estimated reduced form VAR. If this covariance matrix is close to being diagonal then the estimated value of $\mathbf{B}$ will be close to diagonal and so the ordering of the variables will not influence the structural inference.