EC402: Vector Autoregressions (II)

Christian Julliard

Department of Economics and FMG
London School of Economics

Outline

Outline

Structural VAR's

- Let X_t be a $n \times 1$ vector of time series. A S-VAR takes the form

$$\Gamma_0 X_t + \Gamma_1 X_{t-1} + \ldots .\Gamma_p X_{t-p} = \underbrace{c}_{n\times 1} + \underbrace{\varepsilon_t}_{n\times 1} \text{ where } \varepsilon_t \sim N(0, \Sigma) \tag{1}$$

  where Σ is $n \times n$ as well as each Γ_i matrix, c is a vector of constants and ε_t spans the space of *innovations* to X_t.
- This implies that each variables can potentially depend on past *and* current values of all the other variables.
- Two commonly used normalizations are: $\Sigma = I$ (the identity matrix) or each variable has coefficient 1 in one of the $\Gamma(L)$. (We will use the former in what follows)

- We assume Γ_0 is full rank i.e. the system can be solved to determine X_t from past X and ε, that is we can obtain the "reduced form"

$$X_t = \gamma + B(L) X_{t-1} + v_t \tag{2}$$

where $v_t = \Gamma_0^{-1}\varepsilon_t$, $\gamma = \Gamma_0^{-1} c$ and $B(L) = -\Gamma_0^{-1}\left(\Gamma_1 + \Gamma_2 L.... + \Gamma_p L^{p-1}\right)$.

Note: we can estimate equation (2) as discussed before.

Identification

- As in the simultaneous equation case, a key question is whether we can recover the parameters of the structural form (1) from the parameters of the reduced form (2).
- Let's consider the normalization $\Sigma = I$.
- The reduced form has in γ and $B(L)$ as many parameters as in c and $\Gamma_1, ...\Gamma_p$.
- Moreover, we have that

$$v_t \sim N\left(0; \Gamma_0^{-1}\left(\Gamma_0^{-1}\right)'\right)$$

- so we could hope to recover Γ_0 from the covariance matrix of v_t.

Problem: there are many $n \times n$ matrixes G such that $GG' = \Gamma_0^{-1}\left(\Gamma_0^{-1}\right)'$.

Why? There are $(n+1)\,n/2$ free elements in $\Gamma_0^{-1}\left(\Gamma_0^{-1}\right)'$ while Γ_0 has n^2 free elements.

$\Rightarrow$ If we want identification we need *at least* $(n-1)\,n/2$ restrictions.

Note: this is only a necessary condition, not a sufficient one.

- The most commonly used approach to obtain identification is to *impose restrictions in the* Γ_0 *matrix* alone. There are at least two good reasons why this is appealing:

First these restrictions have a natural economic interpretation as assumptions about delays in the reactions of particular variables.

Example

It might be natural to assume that, in setting the interest rate, the Fed reacts contemporaneously to inflation, but that the level of inflation in the economy is not influenced immediately by the actions of the central bank. In writing a S-VAR for i and π, this assumption would be formulated as

$$\underbrace{\begin{bmatrix} \times & \times \\ 0 & \times \end{bmatrix}}_{\Gamma_0} \begin{bmatrix} i_t \\ \pi_t \end{bmatrix} + \Gamma_1 \begin{bmatrix} i_{t-1} \\ \pi_{t-1} \end{bmatrix} + \ldots.\Gamma_p \begin{bmatrix} i_{t-p} \\ \pi_{t-p} \end{bmatrix} = c + \varepsilon_t$$

where $\times$ denotes non-zero elements.

Second the structural estimation can be performed in a relatively simple way by:

1. estimating γ and $B(L)$ performing OLS equation by equation in the reduced form,
2. and then maximizing

$$-\frac{Tn}{2}\log 2\pi + \frac{T}{2}\log\left|\left(\Gamma_0^{-1}\left(\Gamma_0^{-1}\right)'\right)^{-1}\right| +$$
$$-\frac{1}{2}trace\left[\left(\Gamma_0^{-1}\left(\Gamma_0^{-1}\right)'\right)^{-1}\left(\sum_{t=0}^{T}\hat{v}_t'\hat{v}_t\right)\right]$$

with respect to Γ_0, where $\hat{v}_t$ are the reduced form OLS residuals.

Note: this last expression is simply the likelihood maximized with respect to γ and $B(L)$.

- The last step is straightforward if we have exactly $(n-1)\,n/2$ restrictions, but it requires numerical optimization if we have more than $(n-1)\,n/2$ restrictions (i.e. if the system is over-identified).

Note: imposing at least $(n-1)\,n/2$ restrictions in the Γ_0 matrix is only a necessary condition, but it is not sufficient since we also need the restrictions to be *linearly independent* otherwise Γ_0 wouldn't be invertible.

Example: 2 variables system

we need at least $(n-1)\,n/2 = (2-1)\,2/2 = 1$ zero restrictions. Examples that works are Γ_0 of the form

$$\begin{bmatrix} \times & 0 \\ \times & \times \end{bmatrix} \text{ or equivalently } \begin{bmatrix} 0 & \times \\ \times & \times \end{bmatrix}.$$

A useless Γ_0 is

$$\begin{bmatrix} \times & 0 \\ \times & 0 \end{bmatrix}$$

since it is not invertible

Example: 3 variables system

we need at least $(n-1)\,n/2 = (3-1)\,3/2 = 2$ zero restrictions. An exactly identified Γ_0 is then of the form

$$\begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}.$$

An interesting one is

$$\begin{bmatrix} \times & 0 & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}.$$

Here we have enough zeros but we don't have identification since it would not be invertible (the last two rows are not linearly independent). Note: we can identify the first equation with respect to the other two, but we cannot identify the last two equations separately.

Outline

Impulse-Response Functions

- Given the S-VAR (where for simplicity I'm disregarding the vector of constant terms)

$$\Gamma_0 X_t = A(L) X_{t-1} + \varepsilon_t \text{ where } \varepsilon_t \sim N(0, I) \quad (3)$$

we might be interest in how a shock in one of the equations influences the other variables in the system.

- That is, we might want to know the *response* of a variable to a shock (*impulse*) in another variable.

Example: we might want to know how a shock to the current interest rate would affect inflation in the future.

- The IRF's are a very useful data summary since they concisely report the link between variables over time at different horizons.

- The system (3) can be rewritten as

$$X_t = \Gamma_0^{-1} A(L) X_{t-1} + \Gamma_0^{-1} \varepsilon_t$$

and this implies that

$$\begin{aligned} X_{t+1} &= \Gamma_0^{-1} A(L) X_t + \Gamma_0^{-1} \varepsilon_{t+1} \\ &= \left[\Gamma_0^{-1} A(L)\right]^2 X_{t-1} + \Gamma_0^{-1} A(L) \Gamma_0^{-1} \varepsilon_t + \Gamma_0^{-1} \varepsilon_{t+1} \\ X_{t+2} &= \Gamma_0^{-1} A(L) X_{t+1} + \Gamma_0^{-1} \varepsilon_{t+2} \\ &= \left[\Gamma_0^{-1} A(L)\right]^3 X_{t-1} + \left[\Gamma_0^{-1} A(L)\right]^2 \Gamma_0^{-1} \varepsilon_t + \Gamma_0^{-1} A(L) \Gamma_0^{-1} \varepsilon_{t+1} + \Gamma_0^{-1} \varepsilon_{t+2} \end{aligned}$$

- We can therefore define the Impulse Response Function (IRF) of the $j-th$ variable in X to a shock in the $i-th$ equation as

$$\frac{\partial E_t \left[X_{j,t+T}\right]}{\partial \varepsilon_{i,t}} = \left\{ \left[\Gamma_0^{-1} A(L)\right]^T \Gamma_0^{-1} \right\}_{ji} \qquad (4)$$

where $\{\}_{ji}$ denotes the (j, i) element.

Note: we can obtain correct IRF's for some of the shocks even if the system is only partially identified.

Example

Suppose we want a S-VAR for inflation, π, output gap, x and the interest rate, i and believe that the CB reacts contemporaneously to news about π and x but that π and x reacts with a lag to i. We then have:

$$\underbrace{\begin{bmatrix} \times & \times & 0 \\ \times & \times & 0 \\ \times & \times & \times \end{bmatrix}}_{\Gamma_0} \begin{bmatrix} \pi_t \\ x_t \\ i_t \end{bmatrix} = \underbrace{A(L)}_{3\times 3} \begin{bmatrix} \pi_{t-1} \\ x_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^{\pi} \\ \varepsilon_t^{x} \\ \varepsilon_t^{i} \end{bmatrix}; \ \varepsilon_t \sim N(0, I).$$

This system is not identified unless we add a zero in one of the first two equations. But this would imply that either inflation reacts with a lag to output gap, or that output gap react with a lag to inflation.

Note: the third row of Γ_0 is identified with respect to the first two ones.

So, if we are only interested in the IRF's of an ε^i shock, we can simply add an arbitrary zero restriction in one of the first two rows of Γ_0 and equation (4) will give the appropriate answer for the ε^i shocks (but not for the others).

Partial Identification

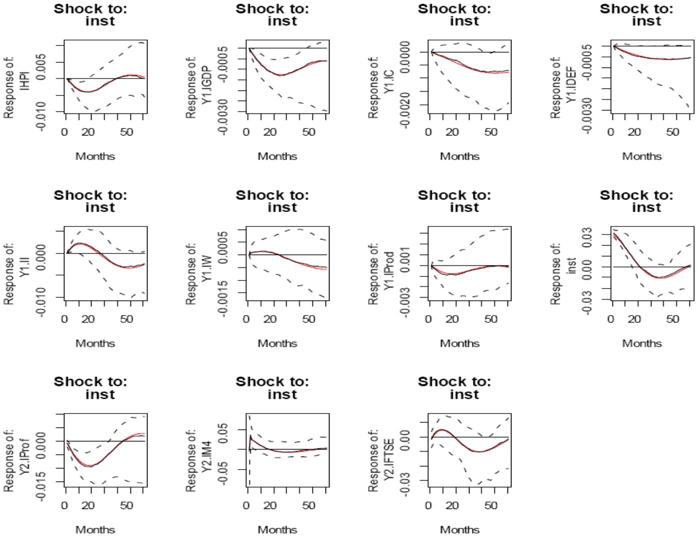

Figure 2. Impulse Response Functions to a Monetary Policy Shock (UK monthly data).

Cholesky Indetification

- This is a very popular approach to identification of the IRF's from a reduced form VAR of the form

$$X_t = B(L)\,X_{t-1} + v_t \text{ where } v_t \sim N(0, \Omega) \tag{5}$$

- It is based on fact that, for any real symmetric positive definite matrix Ω, there exist a unique lower triangular matrix A, with 1s along the main diagonal, and a unique diagonal matrix D with positive entries along the main diagonal such that

$$\Omega = ADA' = AD^{1/2}D^{1/2}A' = PP' \text{ where } P := AD^{1/2}$$

Note: P is also lower triangular.

- Define the $n \times 1$ vector $u_t := A^{-1} v_t$ and note that its covariance matrix is diagonal since

$$\begin{aligned} E\left(u_t u_t'\right) &= \left(A^{-1}\right) E\left(v_t v_t'\right)\left(A^{-1}\right)' = \left(A^{-1}\right) \Omega \left(A'\right)^{-1} \\ &= \left(A^{-1}\right) A D^{1/2} D^{1/2} A' \left(A'\right)^{-1} = D \end{aligned}$$

- That is, the elements of u are *orthogonalized errors* with variance given by D
- Similarly $\eta_t := P^{-1} v_t = D^{-1/2} u_t$ is also a vector of orthogonalized errors.
- Based on these observation, researchers often reports IRF's given by either

$$\frac{\partial E_t\left[X_{j,t+T}\right]}{\partial u_{i,t}} \text{ or } \frac{\partial E_t\left[X_{j,t+T}\right]}{\partial \eta_{i,t}} \tag{6}$$

- The Cholesky decomposition in not a magic wand. It is simply one particular set of restrictions on the Γ_0^{-1} matrix of the structural form, and it will make sense iff this set of restrictions does.

Note: if we think of the Cholesky decomposition as identifying the "true" structural shocks we should have

$$\Gamma_0^{-1} = AD^{1/2} := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 & 0 \\ a_{31} & a_{32} & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 \end{bmatrix} D^{1/2} \tag{7}$$

Note: this approach imposes an *order* among the variables: the first variable in the system will respond contemporaneously only to its own structural shock, the second variable in the system will respond contemporaneously to its own structural shock and to the structural shock of the first variable, and so on.

$\Rightarrow$ this approach will make sense iff this zero restrictions and ordering of the variables are economically meaningful.

Outline

Lucas (1976), Sims(1987), Leeper and Zha(2001)

LUCAS' CRITIQUE. Models that assume that agents' expectations are linear function of the past are likely to be wrong in projecting the effects of systematic policy changes (since policy changes change agents' forecasting rules)

PROBLEM FOR SVAR? (Sims(1987), Leeper and Zha(2001))

- VAR are linear approximations, and can do quite well (especially in the short run) as long as long as the "true" model non-linearity is not too sever.

$\therefore$ If the model fits the data well, and there is little sign of nonlinearity in the sample period, policy disturbances similar to the ones observed in the past will be projected accurately even if they have been generate by "regime shifts."

Note: we can also model regime shifts in the SVAR setting