A primer on Structural VARs

Claudia Foroni
Norges Bank

10 November 2014

Refresh: what is a VAR?

- $VAR\left(p\right)$:

$$y_t = \nu + B_1 y_{t-1} + \ldots + B_p y_{t-p} + u_t, \qquad (1)$$

where

$$\begin{aligned}
\underset{K\times 1}{y_t} &= \begin{pmatrix} y_{1t} & \ldots & y_{Kt} \end{pmatrix}' \\
B_i &= K \times K \text{ coefficient matrices} \\
\underset{K\times 1}{\nu_t} &= \begin{pmatrix} \nu_1 & \ldots & \nu_K \end{pmatrix}' \text{ vector of intercepts} \\
\underset{K\times 1}{u_t} &= \begin{pmatrix} u_{1t} & \ldots & u_{Kt} \end{pmatrix}' \text{ white noise} \\
E\left(u_t\right) &= 0,\ E\left(u_t u_t'\right) = \Sigma_u,\ E\left(u_t u_s'\right) = 0, \forall s \neq t
\end{aligned}$$

Example

Example: VAR(1) with three variables: GDP growth (y_t), inflation (π_t), interest rate(r_t)

$$\begin{pmatrix} y_t \\ \pi_t \\ r_t \end{pmatrix} = \begin{pmatrix} \nu_y \\ \nu_\pi \\ \nu_r \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} u_{yt} \\ u_{\pi t} \\ u_{rt} \end{pmatrix} \tag{2}$$

Use of VAR

So far...

- Describe and summarize macroeconomic time series.
- Compute forecasts.

From Now...

- Understand how variables interact.
- Understand the effect of a shock over time on the different variables.
- Understand the contribution of a shock to the behaviour of the different variables.

Reduced-form VAR

$$
\begin{aligned}
y_t &= By_{t-1} + u_t, && (3)\\
u_t &\sim N(0, \Sigma_u) && (4)
\end{aligned}
$$

- Estimation: **OLS.**
- The constant is **not** the mean nor the long-run equilibrium value of the variable.
- The correlation of the residuals reflects the **contemporaneous** relation between the variables.

Reduced-form VAR: interpretation

- Reduced-form VARs do not tell us anything about the structure of the economy.
- We cannot interpret the reduced-form error terms (u_t) as structural shocks.
- In order to perform policy analysis we want to have:
    1. Orthogonal shocks...
    2. ...with economic meaning.
- We need a **structural representation**.

Structural VAR

Ideally we want to know

$$Ay_t = By_{t-1} + e_t, \quad e_t \sim N(0, I) \tag{5}$$

where the ε_t are **serially uncorrelated and independent of each other**.

Estimation of structural VARs

- We cannot cannot estimate the structural form with OLS because it violates one important assumption: **the regressors are correlated with the error term**.
- The A matrix is problematic, since it includes all the contemporaneous relation among the endogenous variables.

How to solve the estimation issue

- Premultiply the SVAR in eq. (5) by A^{-1} :

$$
\begin{aligned}
A^{-1}Ay_t &= A^{-1}By_{t-1} + A^{-1}e_t, \quad e_t \sim N(0, I) \qquad (6) \\
&\Longrightarrow \\
y_t &= Fy_{t-1} + u_t, \quad u_t \sim N(0, \Sigma_u) \qquad (7)
\end{aligned}
$$

- The VAR in eq. (7) is the usual one we are used to estimate: the reduced-form VAR!

From the reduced-form back to the structural form

- From

$$y_t = Fy_{t-1} + u_t, \quad u_t \sim N(0, \Sigma_u) \tag{8}$$

- Back to

$$Ay_t = By_{t-1} + e_t, \quad e_t \sim N(0, I). \tag{9}$$

- We know that:

$$F = A^{-1}B, \tag{10}$$

$$u_t = A^{-1}e_t, \tag{11}$$

$$\Sigma_u = A^{-1}IA^{-1\prime} = A^{-1}A^{-1\prime}. \tag{12}$$

Identification of A and B

- If we know $A^{-1} \implies B = AF$.
- If we know $A^{-1} \implies e_t = Au_t$.
- **Identification**: how to pin down A^{-1}.

Identification problem

$$\underbrace{\Sigma_u}_{\text{symmetric}} = A^{-1}A^{-1\prime} \tag{13}$$

$$\underbrace{\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}}_{\text{6 values}} = \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1}}_{\text{9 unknowns}} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1\prime}$$

- We have 9 unknowns (the elements of A) but only 6 equations (because the variance-covariance matrix is symmetric)
- **The system is not identified!**

Identification schemes

- Zero short-run restrictions (also known as Choleski identification, recursive identification)
- Sign restrictions

and not covered in this primer...

- Zero long-run restrictions (also known as Blanchard-Quah)
- Theory based restrictions
- Via heteroskedasticity

Zero short-run restrictions (Choleski identification)

- Assume A (or equivalently A^{-1}) to be lower triangular:

$$
\begin{aligned}
Ay_t &= By_{t-1} + e_t && (14)\\
y_t &= A^{-1}By_{t-1} + A^{-1}e_t && (15)\\
y_t &= \widetilde{B}y_{t-1} + \widetilde{A}e_t && (16)
\end{aligned}
$$

with

$$
\widetilde{A} = \begin{pmatrix} \widetilde{a}_{11} & 0 & 0 \\ \widetilde{a}_{21} & \widetilde{a}_{22} & 0 \\ \widetilde{a}_{13} & \widetilde{a}_{23} & \widetilde{a}_{33} \end{pmatrix} \quad (17)
$$

Zero short-run restrictions (Choleski identification)

- In our example:

$$\begin{pmatrix} y_t \\ \pi_t \\ r_t \end{pmatrix} = \begin{pmatrix} \widetilde{b}_{11} & \widetilde{b}_{12} & \widetilde{b}_{13} \\ \widetilde{b}_{21} & \widetilde{b}_{22} & \widetilde{b}_{23} \\ \widetilde{b}_{13} & \widetilde{b}_{23} & \widetilde{b}_{33} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} \widetilde{a}_{11} & 0 & 0 \\ \widetilde{a}_{21} & \widetilde{a}_{22} & 0 \\ \widetilde{a}_{13} & \widetilde{a}_{23} & \widetilde{a}_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{\pi t} \\ \varepsilon_{rt} \end{pmatrix} \tag{18}$$

- Remember: identification problem, we had 6 equations and 9 unknowns.
- Now: we set three elements of A equal to 0 $\Longrightarrow$ **6 equations and 6 unknowns: identification**!

Zero short-run restrictions (Choleski identification)

- **Choleski decomposition**:

$$\Sigma_u = P'P \tag{19}$$

with P' lower triangular.

- Since we have .

$$\Sigma_u = A^{-1}A^{-1\prime} \tag{20}$$

with A^{-1} lower triangular

- Then $P' = A^{-1} \implies$ **Choleski allows identification**!

Choleski identification: interpretation

- Choleski identification is also called **recursive** identification. Why?
- Let's rewrite the VAR in eq. (18) one by one:

$$
\begin{aligned}
y_t &= \ldots + \widetilde{a}_{11}\varepsilon_{yt} && (21)\\
\pi_t &= \ldots + \widetilde{a}_{21}\varepsilon_{yt} + \widetilde{a}_{22}\varepsilon_{\pi t} && (22)\\
r_t &= \ldots + \widetilde{a}_{31}\varepsilon_{yt} + \widetilde{a}_{32}\varepsilon_{\pi t} + \widetilde{a}_{33}\varepsilon_{rt} && (23)
\end{aligned}
$$

Choleski identification: interpretation

- Let's look at the shocks:
  - ε_{yt} affects contemporaneously all the variables.
  - $\varepsilon_{\pi t}$ affects contemporaneously π_t and r_t, but not y_t.
  - ε_{rt} affects contemporaneously only r_t, but not y_t and π_t.
- **The order of the variables matters!**
  - The variable placed on top is the most exogenous (it is affected only by a shock to itself).
  - Each variable contemporaneously affects all the variables ordered afterwards, but it is affected with a delay by them.

Sign restrictions

- Remember Choleski:
$$\Sigma_u = P'P, \tag{24}$$
where P is lower triangular.
- This decomposition is **not unique**.
- Take an orthonormal matrix, i.e. any matrix S such that:
$$S'S = I, \tag{25}$$
- Then we can write
$$\Sigma_u = P'P = P'IP = P'S'SP = \mathcal{P}'\mathcal{P}. \tag{26}$$
- $\mathcal{P}$ is generally not lower triangular anymore.

Sign restrictions

- We can draw as many S as we want $\Longrightarrow$ we can have as many $\mathcal{P}$ as we want.
- Question: **is P plausible?**
- We want to check whether the impulse responses implied by $\mathcal{P}$ satisfy a set of sign restrictions, typically theory-driven.

Sign restrictions: a very intuitive example

- **How is a monetary policy shock affecting output?** (Simplified version of Uhlig (JME, 2005)).
- A contractionary monetary policy should (conventional wisdom and theory):
  1. Raise the federal fund rate,
  2. Lower prices.
- What happens to output? Since it is the question we want to answer, **we leave output unrestricted, i.e. we do not make any assumption on it**!
- We keep only the matrices which generate the responses to a monetary policy shock coherent with 1. and 2.

Steps to implement sign restrictions

1. Estimate the resuced-form VAR and obtain F and Σ_u.
2. Compute $P' = chol\left(\Sigma_u\right)$.
3. Draw a random orthonormal matrix S.
4. Compute $A^{-1} = \mathcal{P} = P'S'$.
5. Compute the impulse responses associated with A^{-1}.
6. Are the sign restrictions satisfied?
   - If yes, store the impulse response.
   - If no, discard the impulse response.
7. Repeat 3-6 until you obtain N replications.
8. Report the mean or median impulse response (and its confidence interval).

Impulse response functions

- The impulse response function traces the effect of a one-time shock to one of the structural errors on the current and future values of all the endogenous variables.
- This is possible only when the errors are uncorrelated $\Longrightarrow$ structural form!

How to compute the IRFs

- Remember previous classes on VAR...
- The impulse responses are derived from the MA representation of the VAR.
- Rewrite the VAR(p) in canonical form (i.e. as VAR(1)), as

$$y_t = Fy_{t-1} + A^{-1}e_t \tag{27}$$

$$\begin{aligned} IRF\,(0) &= A^{-1} = P' & (28)\\ IRF\,(1) &= FA^{-1} & (29)\\ IRF\,(2) &= F^2A^{-1} & (30) \end{aligned}$$

Variance decomposition

- The variance decomposition separates the variation in a endogenous variable into the component shocks of the VAR.
- It tells what portion of the variance of the forecast error in predicting $y_{i,T+h}$ is due to the structural shock e_j.
- It provides information about the relative importance of each innovation in affecting the variables.

Historical decomposition

- The historical decomposition tells what portion of the deviation of $y_{i,t}$ from its unconditional mean is due to the shock e_j.
- The structural shocks push the variables away from their equilibrium values.