Class 4: VAR

Macroeconometrics - Spring 2015

Jacek Suda

April 17, 2015

Outline

Outline:

1. Dynamic Structural Models and VAR
2. Identification
3. Multivariate Wold Form and Forecasting
4. Impulse Response Functions
5. Variance Decomposition
6. Identification
   - Short-run restrictions
   - Long-run restrictions
7. Granger Causality

Example: Money Demand

Let

$$\begin{aligned} y_{1t} &= \text{real money balance} = \frac{M}{P}, \\ y_{2t} &= \text{real GNP}. \end{aligned}$$

Money Demand

$$y_{1t} = \gamma_{10} + \beta_{12} y_{2t} + \gamma_{11} y_{1,t-1} + \gamma_{12} y_{2,t-1} + \varepsilon_{1t}.$$

- ε_{1t} encompasses all other factors,
- β_{12} is a short-run elasticity of real money balances, $\left(\frac{M}{P}\right)^d$, with respect to real income,
- lagged terms allow for different long-run elasticity.

Example: Money Supply

Money Supply

$$y_{2t} = \gamma_{20} + \beta_{21} y_{1t} + \gamma_{21} y_{1,t-1} + \gamma_{22} y_{2,t-1} + \varepsilon_{2t}.$$

- β_{21} is a short-run impact of money on output,

Estimating Money Demand or Money Supply by OLS:

- $\hat{\beta}_{12}^{OLS}$ and $\hat{\beta}_{21}^{OLS}$ are inconsistent due to endogeneity (simultaneity).

$\hat{\beta}_{ij}^{OLS}$

- picks up correlation between income and money.

VAR

We want to make inference on causality from correlation.

- Using VAR, we
  1. get correlation structure of data,
  2. can do "macro modelling without pretending to have too much a priori theory" (Sims, 1980),
  3. given the structure, make identification assumptions what shocks are (timing as identification tool).

Structural VAR(1): System

Write down supply and demand equation as a system of equation:

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

Or, in matrix notation

$$B_{2\times 2} y_{t_{2\times 1}} = \Gamma_{0_{2\times 1}} + \Gamma_{1_{2\times 2}} y_{t-1_{2\times 1}} + \varepsilon_{t_{2\times 1}}.$$

- Similar to AR(1) but in a vector form.

Structural VAR(1): Shocks

Shocks

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

Exogenous shocks to each variable $\Rightarrow$ *diagonal variance covariance matrix,* $D = Diag$*.*

Reduced-form VAR(1)

- Solve structural model.
- If $\beta_{12} \neq \beta_{21} \neq |1| \Rightarrow B^{-1}$ exists and

$$\begin{array}{ccccccc} y_t & = & B^{-1}\Gamma_0 & + & B^{-1}\Gamma_1 y_{t-1} & + & B^{-1}\varepsilon_t \\ y_t & = & \underset{2\times 1}{C} & + & \underset{2\times 2}{\Phi} y_{t-1} & + & e_t \end{array},$$

where

$$e_t = B^{-1}\varepsilon_t = \frac{1}{1-\beta_{12}\beta_{21}} \begin{bmatrix} \varepsilon_{1t} + \beta_{12}\varepsilon_{2t} \\ \varepsilon_{2t} + \beta_{21}\varepsilon_{1t} \end{bmatrix} = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \longrightarrow \text{forecast errors}$$

Forecast errors

$$e_t = B^{-1}\varepsilon_t = \frac{1}{1-\beta_{12}\beta_{21}}\begin{bmatrix} \varepsilon_{1t} + \beta_{12}\varepsilon_{2t} \\ \varepsilon_{2t} + \beta_{21}\varepsilon_{1t} \end{bmatrix} = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

- $\frac{1}{1-\beta_{12}\beta_{21}}$ *overall feedback effect.*
- *Total effect of shock ε_{2t} on money: shock to income $\Rightarrow$ money $\Rightarrow$ income $\Rightarrow$ money.*
- *Any forecast error has the form of the linear combination of structural shocks.*

Forecast errors

Forecast errors:

$$
\begin{aligned}
E[e_t] &= 0 \\
E[e_t e_t'] &= E[B^{-1}\varepsilon_t \varepsilon_t' (B^{-1})'] \\
&= B^{-1} E[\varepsilon_t \varepsilon_t'] (B^{-1})' \\
&= B^{-1} D (B^{-1})' \equiv \Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix}.
\end{aligned}
$$

- Ω is not diagonal as both forecast error are affected by both shocks.
- If $|eigenvalue(\Phi)| < 1$ then reduced form VAR can be consistently estimated by OLS (equation by equation).
- Similar to SUR (*Seemingly Unrelated Regressions*): a special case where all xs are the same for each equation.
- Estimation of system of equations with OLS is equivalent to CMLE to SUR.

Problem with reduced form: identification.

Identification

- Structural VAR(1) has 10 parameters:

$$(\gamma_{10}, \gamma_{20}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, \sigma_1^2, \sigma_2^2, \beta_{12}, \beta_{21})$$

- Reduced form VAR(1) has 9 parameters:

$$(c_1, c_2, \phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}, \omega_{11}, \omega_{12}, \omega_{22})$$

- Cannot identify β_{12}, β_{21} from ω_{12}.
  All we have is correlation of income with forecast and money with forecast.
- Infinite number of structural VARs that are consistent with reduced from VAR.
- Need additional restrictions to identify the model (e.g. short-run, long-run, sign restrictions).

Multivariate Wold Form

- Take covariance stationary process, $\{Y_t\}_{-\infty}^{\infty}$.
- The multivariate Wold Form

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \qquad \{\varepsilon_t\} \sim WN.$$

- For VAR(1)

$$y_{t_{2\times 1}} = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}.$$

and

$$y_t = C + \Phi y_{t-1} + e_t, \qquad e_t \sim iid(0, \Omega),$$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}.$$

Multivariate Wold Form

In lag notation,

$$
\begin{aligned}
(I - \Phi L)y_t &= C + e_t \\
y_t &= \mu + (I - \Phi L)^{-1} e_t, \qquad \mu = (I - \Phi)^{-1} C, \text{ as } LC = C \\
y_t &= \mu + \Psi(L) e_t,
\end{aligned}
$$

where

$$
\begin{aligned}
\Psi(L) &= \sum_{k=0}^{\infty} \Psi_k L^k, \quad \Psi_0 = I, \Psi_k = \Phi^k, \\
\Psi(L) &= I + \Phi L + \Phi^2 L^2 + \ldots
\end{aligned}
$$

Forecast for VAR(1)

From Wold Form:

$$y_{t+s} = \mu + e_{t+s} + \Phi e_{t+s-1} + \ldots + \Phi^s e_t + \Phi^{s+1} e_{t-1} + \ldots$$

Then,

$$\begin{aligned} y_{t+s|t} &= \mu + \Phi^s e_t + \Phi^{s+1} e_{t-1} + \ldots \\ &= \mu + \Phi^s (e_t + \Phi e_{t-1} + \ldots) \\ &= \mu + \Phi^s (y_t - \mu) \end{aligned}$$

- So the forecast is just the deviation of the series from their long-run unconditional means.
- Note:

$$\Psi_s[1,1] \neq \frac{\partial y_{1,t+s}}{\partial e_{1t}},$$

because $E[e_{1t}e_{2t}] = \omega_{12} \neq 0$.
- $\frac{\partial y_{1,t+s}}{\partial e_{1t}}$ = univariate effect e_t on $y_{1,t+s}$, ceteris paribus.
- In the univariate case $\Psi_s = \frac{\partial y_{t+s}}{\partial e_t}$.

MSE

Mean square error

$$\begin{aligned} MSE(y_{t+s|s}, y_{t+s}) &= E[(e_{t+s} + \Phi e_{t+s-1} + \ldots + \Phi^{s-1} e_{t+1}) \\ &\quad \cdot (e_{t+s} + \Phi e_{t+s-1} + \ldots + \Phi^{s-1} e_{t+1})'] \\ &= E[e_{t+s} e'_{t+s}] + \Phi E[e_{t+s} e'_{t+s}] \Phi' + \ldots + \Phi^{s-1} E[e_{t+s} e'_{t+s}] (\Phi^{s-1} \\ &= \Omega + \Phi \Omega \Phi' + \ldots + \Phi^{s-1} \Omega (\Phi^{s-1})' \\ &= \sum_{k=0}^{s-1} \Phi^k \Omega (\Phi^k)'. \end{aligned}$$

$$\begin{aligned} \lim_{s \to \infty} MSE &= \sum_{k=0}^{\infty} \Phi^k \Omega (\Phi^k)' \\ &= \sum_{k=0}^{\infty} \Psi_k \Omega \Psi_k' = var(Y_t) \end{aligned}$$

Impulse Response Functions

Given a reduced-form VAR and an identification assumption for B, solve the structural VAR

$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + \varepsilon_t,$$

where

$$\begin{aligned} \Gamma_0 &= Bc, \\ \Gamma_1 &= B\Phi, \\ \varepsilon_t &= Be_t. \end{aligned}$$

Impulse Response Functions

Solve for vector MA in terms of structural shocks:

$$\begin{aligned}(B - \Gamma_1 L)y_t &= \Gamma_0 + \varepsilon_t \\ y_t &= (B - \Gamma_1 L)^{-1}(\Gamma_0 + \varepsilon_t) \\ &= \mu + \theta(L)\varepsilon_t,\end{aligned}$$

where $\mu = (B - \Gamma_1 L)^{-1}\Gamma_0$.

From Wold Form:

$$\begin{aligned}\theta(L)\varepsilon_t &= \Psi(L)e_t \\ &= \Psi(L)B^{-1}\varepsilon_t \\ \Rightarrow \theta(L) &= \Psi(L)B^{-1} \\ &= B^{-1} + \Psi_1 B^{-1}L + \Psi_2 B^{-1}L^2 + \ldots\end{aligned}$$

Impulse Response Functions

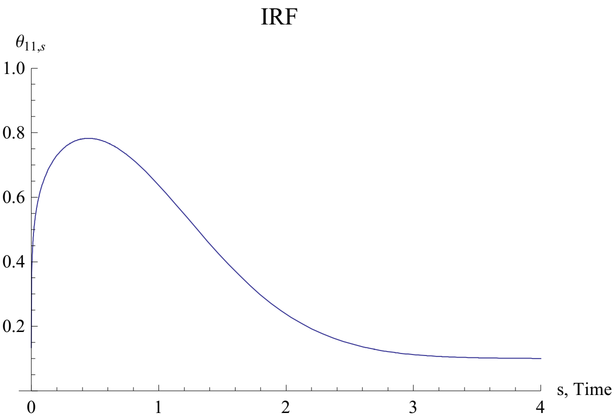
- Therefore,

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

- Note $\theta_0 = B^{-1} \neq I$.
- Then, $\theta_{11,s} = \frac{\partial y_{1,t+s}}{\partial \varepsilon_{1t}}$.
- For n variable system we have n^2 impulse response functions.

Impulse Response Functions

$$\theta_{11,s} = \frac{\partial y_{1,t+s}}{\partial \varepsilon_{1t}}$$



Cumulative Response

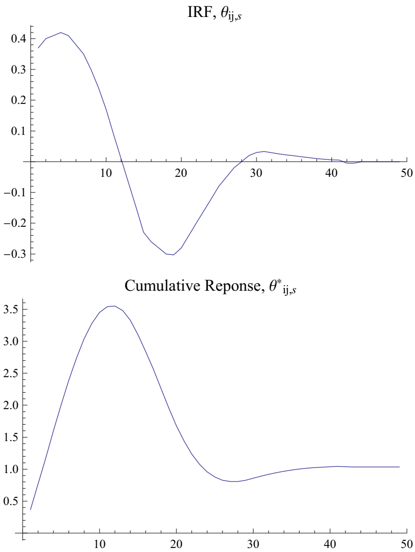
Define Cumulative Response Function

$$\theta^*_{ij,s} = \sum_{k=0}^{s} \theta_{ij,k},$$

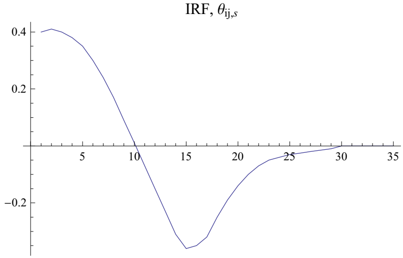
$\theta_{ij}(1) =$ long-run cumulative impact of shock j on variable i.

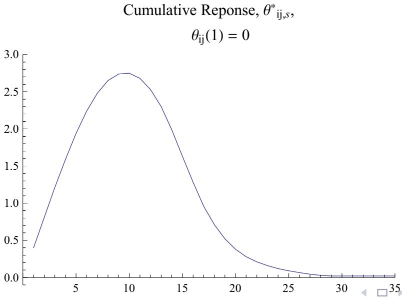
$$\begin{aligned} \theta_{ij}(1) &= \theta_{ij,1} + \theta_{ij,2} + \theta_{ij,3} + \dots \\ \theta_{ij}(1) &= \lim_{s\to\infty} \theta^*_{ij,s} \end{aligned}$$

Cumulative Response



Cumulative Response





Variance decomposition

For MSE at different horizons, what share is due to each structural shocks?

Variance decomposition:

$$p_{i,j}(s) = \frac{\sigma_j^2(\theta_{ij,0}^2 + \ldots + \theta_{ij,s-1}^2)}{MSE(y_{i,t+s|t}, y_{i,t+s})},$$

with i denoting series, and j denoting shocks.

$$MSE^i = MSE(y_{i,t+s|t}, y_{i,t+s}) = \sum_{j=1}^{n} \sigma_j^2(\theta_{ij,0}^2 + \ldots + \theta_{ij,s-1}^2),$$

n variables, n shocks.

VAR(2)

Consider the two equation two-lag system

$$\begin{aligned} x_{1t} &= a_{11}x_{1,t-1} + a_{12}x_{1,t-2} + a_{13}x_{2,t-1} + a_{14}x_{2,t-2} + e_{1,t} \\ x_{2t} &= a_{21}x_{1,t-1} + a_{22}x_{1,t-2} + a_{23}x_{2,t-1} + a_{24}x_{2,t-2} + e_{2,t}. \end{aligned}$$

Defining the vector

$$y_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{1,t-1} \\ x_{2,t-1} \end{bmatrix},$$

the system can be represented in a VAR(1) matrix form

$$y_t = Ay_{t-1} + e_t,$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad e_t = \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ 0 \\ 0 \end{bmatrix}$$

VAR(p)

More generally, the VAR(p) system

$$x_t \quad = \quad c + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \Phi_3 x_{t-3} + \ldots + \Phi_p x_{t-p} + e_t, \qquad e_t \sim iid(0, \Omega),$$

can be written as a VAR(1)

$$y_t = A y_{t-1} + v_t$$

IDENTIFICATION

SHORT-RUN RESTRICTION

VAR(1)

Recall:

- Structural VAR(1)

$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + \varepsilon_t,$$

- Reduced-form VAR(1)

$$\begin{aligned} y_t &= B^{-1}\Gamma_0 + B^{-1}\Gamma_1 y_{t-1} + B^{-1}\varepsilon_t, \\ &= C + \Phi y_{t-1} + e_t. \end{aligned}$$

- Need $\frac{n(n-1)}{2}$ restrictions to identify structural VAR.

Error term

Recall that

$$e_t = B^{-1}\varepsilon_t,$$

where e_t is a forecast error and ε_t is a structural shock. e_t are correlated, ε_t are not correlated.

It is linear relationship.

Short-run Restrictions

Assumptions:

$$\varepsilon_t \sim (0, D), \quad D = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix},$$
$$e_t \sim (0, \Omega), \quad \Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix}.$$

- Suppose B^{-1} is "lower triangular", then B^{-1} and D can be identified from Cholesky decomposition of Ω.

Cholesky decomposition:
For any positive definite symmetric matrix there exist unique decomposed, triangular factorization

$$\Omega = PP' = T\Lambda T',$$

where Λ is a diagonal matrix with positive elements and T is lower diagonal matrix with 1s on diagonal,
P is a lower diagonal matrix.

- Cholesky decomposition is unique.

Short-run Restrictions

Therefore, if B^{-1} is lower triangular then take

$$T = B^{-1},$$

and

$$\Lambda = D,$$

so that

$$var(e_t) = \Omega = B^{-1}D(B^{-1})'$$

Reduced form VAR and Cholesky decomposition $\Rightarrow$ Structural VAR if B^{-1} lower triangular.

- *What does it mean that B^{-1} is lower triangular ?*

Example: Sims, 1980

Consider VAR for $\Delta y_t, \pi_t, i_t$ (output growth, inflation, and interest rate).

- Policy rule:

$$i_t = \beta_{31}\Delta y_t + \beta_{32}\pi_t + \varepsilon_{3t},$$

  where $\beta_{31}\Delta y_t + \beta_{32}\pi_t$ is a reaction function and ε_{3t} is a policy shock.

- Inflation:

$$\pi_t = \beta_{21}\Delta y_t + \gamma_{23} i_{t-1} + \varepsilon_{2t},$$

  with ε_{2t} being, for example, oil price shock.
  Policy variable affects inflation with a lag – an assumption.

- Output growth:

$$\Delta y_t = \gamma_{13} i_{t-1} + \varepsilon_{1t},$$

  with ε_{1t} denoting a productivity/supply shock.
  Timing assumption: it may take a while to have change in i affecting output

Example: Structural VAR

VAR:

$$\begin{bmatrix} 1 & 0 & 0 \\ -\beta_{21} & 1 & 0 \\ -\beta_{31} & -\beta_{32} & 1 \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \pi_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}.$$

- We say that Δy_t does not respond to shocks in i_t and π_t.
- *Sims on γ's:* Why put 0 restrictions if they are not obvious from the model – we can estimate it and see if they are really zero.

Example: Structural VAR(1)

VAR(1):

$$Bx_t = \Gamma_1 x_{t-1} + \varepsilon_t,$$

- $\beta_{13} = \beta_{23} = 0 \Rightarrow$ output and inflation respond to policy shocks with a lag,
- $\beta_{12} = 0 \Rightarrow$ output responds to oil price shock with a lag,
- $\beta_{13} = \beta_{23} = 0 \Rightarrow$ forecast error for output growth is a productivity shock.

Example: Reduced-form VAR(1)

Structural $Bx_t = \Gamma_1 x_{t-1} + \varepsilon_t$ implies a reduced form VAR:

$$\begin{aligned} x_t &= B^{-1}\Gamma_1 x_{t-1} + B^{-1}\varepsilon_t \\ &= \Phi x_{t-1} + e_t. \end{aligned}$$

Since B and B^{-1} are lower triangular,

$$e_t = B^{-1}\varepsilon_t$$

implies

$$\begin{aligned} e_{1t} &= \varepsilon_{1t} \\ e_{2t} &= \varepsilon_{2t} + \beta_{21}\varepsilon_{1t} \\ e_{3t} &= \varepsilon_{3t} + \beta_{22}\varepsilon_{2t} + \beta_{32}\varepsilon_{1t}. \end{aligned}$$

It is recursive identification. We can identify ε_{2t} knowing ε_{1t}.

Wold-causal ordering

Wold-causal ordering:
All variables can be endogenous but

- Δy_t is causally prior to π_t and i_t,
- π_t is causally prior to i_t.

Ordering is what defines the impact.

- We put Δy_t first, π_t second and i_t last.
- *If we put different order, like π_t last, we say that interest rate affect inflation.*
- *Sims: Fed doesn't observe GDP, it has only a lagged value.*

Example: Kilian(2009, AER)

- Lutz Killian (2009, *American Economic Review*):
  Not All Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market.
- Oil shocks: large increases and declines in the price of oil, receive a lot of attention.
- Many recent recessions were preceded by an increase in the price of oil but oil usage is actually a relatively small input compared to GDP.
- Kilian asks "what is an oil price shock and are there different kinds of oil price shocks?" (instead of "what are the effects of an oil price shock?")
- Paper uses VAR analysis to distinguish between shocks to oil prices due to global demand, shocks due to oil supply, and shocks due to speculation in the oil price market.

Kilian(2009): Model

- Three variable monthly VAR in the growth rate of oil production, real global economic activity, and the real price of oil: $zt = (\Delta prod_t, rea_t, rpo_t)$.
- VAR structure

$$A_0 z_t = \alpha + \sum_{i=1}^{24} A_i z_{t-i} + \varepsilon_t,$$

where ε_t are structural shocks and A_0 is lower triangular

$$A_0 = \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$$

- Identifying assumptions
  1. Oil production does not respond within the month to world demand and oil prices.
  2. World demand is affected within the month by oil production, but not by oil prices.
  3. Oil prices respond immediately to oil production and world demand.

Kilian(2009): Shocks

- Since A_0 is lower-triangular, so is A_0^{-1}.
- Reduced-form VAR

$$z_t = A_0^{-1}\alpha + \sum_{i=1}^{24} A_0^{-1} A_i z_{t-i} + A_0^{-1}\varepsilon_t$$

- Reduced form shocks

$$A_0^{-1}\varepsilon_t = e_t = \begin{bmatrix} e_t^{\Delta prod} \\ e_t^{rea} \\ e_t^{rpo} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\Delta prod} \\ \varepsilon_t^{rea} \\ \varepsilon_t^{rpo} \end{bmatrix}$$

Kilian(2009): Findings

- Main Lesson: How the economy reacts to an “oil price shock” will depend on the origins of that shock.
- Shocks to oil supply have limited effects on oil prices and have been of negligible importance in driving oil prices over time.
- Both global demand and speculative oil price shocks can have significant effects on oil prices, but speculative oil price shocks have limited effects on global economic activity.
- Speculative oil-market shocks have accounted for most of the month-to-month movements in oil prices.
- The steady increase in oil prices from 2000 onwards was almost solely due to strong global demand.

Short-run restrictions

Short-run restrictions:

- Use the recursive identification method.
- Construct a set of uncorrelated structural shocks directly from the reduced-form shocks.
- Assumes that certain shocks having effects on only some variables at time t or, alternatively, that some variables only having effects on some variables at time t.
- Corresponds to assuming that B is lower triangular in VAR(1)

$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + \varepsilon_t.$$

- Causal ordering of variables in y_t defines the impact.

Long-run restriction

Too strong assumptions?

- We need a story to tell why B is lower triangular.
- Ordering

$$x_t = \begin{bmatrix} \Delta y_t \\ \pi_t \\ i_t \end{bmatrix}$$

  assumes that output is causally prior to inflation and interest rate: if there is a shock to interest rate, it will take time to be reflected in π_t and Δy_t.
- We want simultaneity in out model (hence VAR) but we assume it away in the first period for identification purposes.
- Additionally: Δp_t^i: commodity prices changes are connected to/directly reflected in interest rate, and both i_t and p_t^i are determined simultaneously.

Too strong assumptions?

- Problems:
  1. short-run identifications may have some limitations–can't be done it in some cases,
  2. we want assumptions on identification that will not just assume answer – we want data to be decisive, not model/identification selected.
- Economic theory gives very little guidance.
- Need method of identification that allows for general B matrix, not only lower triangular.
- Long run identification: impose more plausible restrictions, does not assume Keynesian or classical approach.

Long run identification

- Long-run identification approach: use these theoretically-inspired long-run restrictions to identify shocks and impulse responses.
- Economic theory usually tells us a lot more about what will happen in the longer-run, rather than exactly what will happen today.
- For instance, theory tells us that whatever positive aggregate demand shocks do in the short-run, in the long-run while they should have no effect on output, they have a positive effect on the price level.

Structural VAR

Structural model:

$$By_t = \Gamma_0 + \Gamma_1 y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim iid(0, D).$$

Vector moving average representation:

$$\begin{aligned}
(B - \Gamma_1 L)\, y_t &= \Gamma_0 + \varepsilon_t \\
y_t &= (B - \Gamma_1 L)^{-1}\, \Gamma_0 + (B - \Gamma_1 L)^{-1}\, \varepsilon_t, \\
&= \mu + \theta(L)\varepsilon_t, \qquad \theta_0 = B^{-1}, \theta_1 = \Gamma_1 B^{-1}, \ldots \\
&= \mu + \theta_0 \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots
\end{aligned}$$

- Elements in θ_i tell us what the structure of shocks is.
- We can identify Γ's but not B.

Cumulative impact of the shock

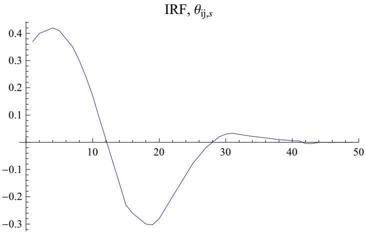
Impulse Response Function:

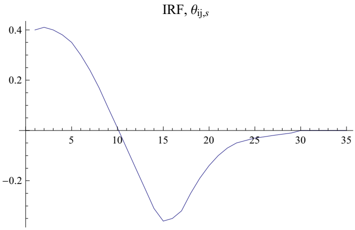
$$\theta_{s,11} = \frac{\partial y_{1,t+s}}{\partial \varepsilon_{1t}}$$

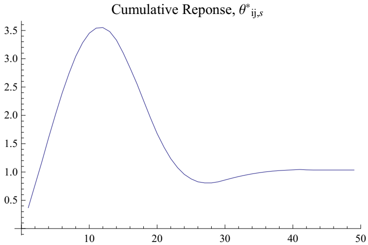
Cumulative impact:

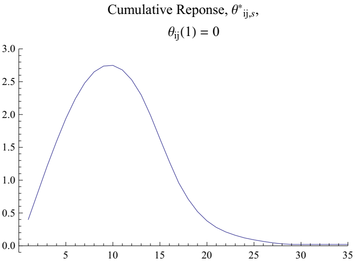
$$\theta^*_{s,11} = \frac{\partial \sum_{j=0}^{s} y_{1,t+j}}{\partial \varepsilon_{1t}} = \sum_{j=0}^{s} \theta_{j,11}.$$

Cumulative impact of the shock









VAR

- When we think about long-run restriction we think about long-run effects of the shock.
- For stationary VAR model $IRF(s) \to 0$ as $s \to \infty$ but it does not have to be true for the long run cumulative response.

Blanchard and Quah, AER 1989

Example: Blanchard and Quah

- Let Δy_t denotes GDP growth *(stationary)* and u_t unemployment

$$x_t = \begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix}.$$

- Then the structural VAR model is

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t, \qquad \varepsilon = \begin{bmatrix} \varepsilon_t^{AS} \\ \varepsilon_t^{AD} \end{bmatrix} \sim iid(0, D).$$

Assumptions

- Assumption: AS and AD shocks are exogenous , uncorrelated.
- They drive fluctuations in each series.

$$D = \begin{bmatrix} \sigma^2_{AS} & 0 \\ 0 & \sigma^2_{AD} \end{bmatrix}.$$

- Assumption: All dynamics are in the VAR, not in the structure of shocks.
- A vector MA version of the model:

$$x_t = \mu + \theta(L)\varepsilon_t \qquad \text{vector MA.}$$

- Long-run variance:

$$\Lambda = var(x_t) = var(\theta(1)\varepsilon_t) = \theta(1)D\theta(1)' \ ,$$

which reflects the cumulative effects of shocks.
- But shocks to Δy_t do not have same periodic persistence as shocks to u_t — they have long run effect.
- We can identify long-run variance from reduced form model.

Reduced form MA

Reduced form:

- Wold form:

$$
\begin{aligned}
x_t &= \mu + \Psi(L)e_t \\
\Lambda &= \Psi(1)\Omega\Psi(1)',
\end{aligned}
$$

- Estimates of each of these variances:

$$
\begin{aligned}
y_t &= \hat{c} + \hat{\Phi} y_{t-1} + e_t, \qquad e_t \sim iid(0, \hat{\Omega}), \\
y_t &= (I - \hat{\Phi})^{-1}\hat{c} + (I - \hat{\Phi}L)^{-1} e_t, \\
&= \hat{\mu} + e_t + \hat{\Psi} e_{t-1} + \hat{\Psi}_2 e_{t-2} + \ldots, \\
& \quad \Psi_1 = \Phi,\ \Psi_2 = \Phi^2,\ \Psi_j = \Phi^j, \\
\hat{\Psi}(1) &= I + \hat{\Psi}_1 + \hat{\Psi}_2 + \ldots
\end{aligned}
$$

- Blanchard and Quah: Aggregate shocks have no long-run effect on level of output.

Structural MA

Structural MA

$$\begin{aligned}
\begin{bmatrix} \Delta y_t \\ u_t \end{bmatrix} &= \begin{bmatrix} \mu_1 \\ \mu_1 \end{bmatrix} + \begin{bmatrix} \theta_{0,11} & \theta_{0,12} \\ \theta_{0,21} & \theta_{0,22} \end{bmatrix} \begin{bmatrix} \varepsilon_t^{AS} \\ \varepsilon_t^{AD} \end{bmatrix} \\
&+ \begin{bmatrix} \theta_{1,11} & \theta_{1,12} \\ \theta_{1,21} & \theta_{1,22} \end{bmatrix} \begin{bmatrix} \varepsilon_{t-1}^{AS} \\ \varepsilon_{t-1}^{AD} \end{bmatrix} \\
&+ \ldots + \begin{bmatrix} \theta_{j,11} & \theta_{j,12} \\ \theta_{j,21} & \theta_{j,22} \end{bmatrix} \begin{bmatrix} \varepsilon_{t-j}^{AS} \\ \varepsilon_{t-j}^{AD} \end{bmatrix} + \ldots
\end{aligned}$$

- Average of past shocks.
- Supply shocks and demand shocks affect both output growth and unemployment...

  ... if short-run restriction: $\theta_{0,12} = 0$.

Long-run restriction

- Long-run assumption: shocks in 1902 does not have any effect on output growth today – accumulation of the shocks conveyed only to some level.
- Both shocks to unemployment and output growth die out.

$$\theta_{j,xy} \to 0 \text{ as } j \to \infty.$$

- Cumulative impacts

$$\begin{aligned} \lim_{s\to\infty} \sum_{j=1}^{\infty} \theta_{j,11} &= \theta_{11}(1), && \textit{cumulative impact of AS shocks on output} \\ \lim_{s\to\infty} \sum_{j=1}^{\infty} \theta_{j,12} &= \theta_{12}(1), && \textit{cumulative impact of AD shocks on output}\ . \end{aligned}$$

- Assumption: $\theta_{12}(1) = 0$.
- AD shock has no long-run impact on the level of output, y_t.

Cholesky decomposition

- Estimate $\hat{\Lambda}$:

$$\hat{\Lambda} = \hat{\Psi}(1)\hat{\Omega}\hat{\Psi}(1)',$$

- Do a Cholesky decomposition of $\hat{\Lambda}$:

$$\hat{\Lambda} = \hat{\theta}(1)\hat{D}\hat{\theta}(1)',$$

with $\hat{\theta}(1)$ lower triangular and $\hat{D}$ diagonal matrices.

VAR

- From the structural form:

$$\begin{aligned} x_t &= \mu + \Psi(L)e_t, \\ x_t &= \mu + \Theta(L)\varepsilon_t. \end{aligned}$$

- Since

$$e_t = B^{-1}\varepsilon_t,$$

we have

$$\Psi(L)B^{-1} = \theta(L).$$

It is true for every VAR model, any L.

- As this hold in general, then it also holds

$$\begin{aligned} \Psi(1)B^{-1} &= \theta(1) \\ \hat{B}^{-1} &= \hat{\Psi}(1)^{-1}\hat{\theta}(1). \end{aligned}$$

Identification

- We have $\hat{\theta}(1)$ identified because of Cholesky decomposition
- Recall $\hat{\Lambda} = \hat{A}\hat{B}\hat{A}'$ is unique with $\hat{A}$ being lower triangular.
- Imposing long-run restriction makes $\hat{\theta}(1)$ lower triangular.
- Since $\hat{\theta}(1)$ is lower triangular we get a $\hat{\theta}(1)$ from the unique decomposition of $\Lambda = \theta(1)D\theta(1)'$.
- In both long-run and and short-run identifications, we construct a lower triangular matrix so we can use Cholesky decomposition.

More on identification

- See Galí (1999, *American Economic Review*), "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" for long-run restrictions.
- Other identification restrictions:
  1. Identification by sign restrictions
  2. Identification from heteroskedasticity
  3. DSGE priors
  4. Identification through regional/multicountry restrictions
  5. Natural experiment approach

  See “Recent Developments in Structural VAR Modeling” NBER Summer Institute lecture by Stock and Watson

Granger Causality

Granger Causality

- Two things behind the notion of Granger causality:
  1. The cause occurs before the effects.
  2. The cause contains information about the effect that is unique and is in no other variable.
- In some case if we have one direction, Granger causality allows us to do inference about causality.
- In some cases, however, it may seem we have Granger causality but it may be driven by the lack of relevant variables in the regression.
- It doesn't require structural assumptions.

Granger(AER)

Granger(AER) causality:

Definition

Process $\{y_{2t}\}$ *"Granger causes"* $\{y_{1t}\}$ *if Mean Square Error of (linear prediction)* $\hat{E}[y_{1,t+s}|\tilde{y}_{1t}] \neq$ *Mean Square Error of* $\hat{E}[y_{1,t+s}|\tilde{y}_{1t}, \tilde{y}_{2t}]$, $(\tilde{y}_{1t} = \{y_{1,t}, y_{1,t-1}, y_{1,t-2}, \ldots\})$.

- y_{2t} contains marginal predictive power above and beyond of what can be observed in y_{1t} alone.
- $\{y_{2t}\}$ provides marginal predictive power for $\{y_{1t}\}$
- Can you improve on MSE by adding $\tilde{y}_{2t}$?
- If y_{1t} can be predicted more efficiently when the information in the y_{2t} process is taken into account then y_{2t} is Granger-causal for y_{1t}.

St. Louis Regression (60s)

- St. Louis Regression (60s): Regress output growth on lagged money growth

$$\Delta Y_t = \alpha + \beta \Delta M_{t-1} + \varepsilon_t,$$

- They find

$$\beta > 0$$

— when money growth is high today, output growth will be high tomorrow.

- St. Louis Fed said it is causal relationship.

Tobin (1970, AER)

Tobin (1970, AER)

- $\beta > 0$ can reflect:
  1. M causes Y
  2. output, Y, causes money, M (*monetary authority just respond to economic conditions*) and output, Y_{t-1}, forecasts future output, Y_t, i.e.

$$cov(\Delta Y_t, \Delta Y_{t-1}) > 0,$$
$$cov(\Delta Y_{t-1}, \Delta M_{t-1}) > 0.$$

- Purely passive rule with real causality of income causing money.

Sims (1972)

Sims (1972)

1. If Y causes M, and Y predicts Y, then output growth Granger causes M money.
2. If Y does not causes M, then Y does not Granger causes M.

Estimate VAR:

$$\begin{aligned} \Delta Y_t &= c_Y + \phi_{11}\Delta Y_{t-1} + \phi_{12}\Delta M_{t-1} + e_{1t}, \\ \Delta M_t &= c_M + \phi_{21}\Delta Y_{t-1} + \phi_{22}\Delta M_{t-1} + e_{2t}. \end{aligned}$$

Sims finds:

- $\phi_{12} \neq 0 \Rightarrow M$ Granger causes Y (Note $\phi_{12} \neq \beta$.)
- $\phi_{21} = 0 \Rightarrow Y$ does not Granger causes M.
- Money is not predicted by income.

Catch

Catch:

- Sims' paper rejects Tobin's story.
- With updated data you don't get the second result: money Granger cause income and income Granger cause money.
- If both GC each other it might be that both causes each other–if we have this simultaneity we have to do the identification restriction.
- Hamilton: there are pitfalls in expectations: stock market can be found to GC a lot of variables: does it mean they cause them? stock market prices can reflect expectations.
- You have to find unidirectional causality.
- Have to convince that this unidirectional causality is not because of expectations.
- Also, failure to reject GC might be due to low power.
- If data are not stationary and you perform Granger causality, sizes of the tests are different than usual. It is relevant to know if there is a unit root.

VAR

A VAR is an n-equation, n-variable model in which each variable is in turn explained by its own lagged values, plus current and past values of the remaining $n - 1$ variables.
In data description and forecasting, VARs have proven to be powerful and reliable tools. Structural inference and policy analysis are, however, inherently more difficult because they require differentiating between correlation and causation; “the identification problem.

Reduced-form VAR

A *reduced form VAR* expresses each variables as a linear function of its own past values, the past values of all other variables being considered and a serially uncorrelated forecast error term. Each equations is estimated by ordinary least squares regression. The error terms are the "surprise" movements in the variables. If the different variables are correlated with each other then the eror terms in the reduced form model will also be correlated across equations.

Recursive VAR

A *recursive VAR* constructs the error terms in each regression equation to be uncorrelated with the error in the preceding equations. This is done by judiciously including some contemporaneous values as regressors. Estimation of each equation by ordinary least squares produces residuals that are uncorrelated across equations. The results depend on the order of the variables: changing the order changes the VAR equations, coefficients, and residuals.

Structural VAR

A *structural VAR* uses economic theory to sort out the contemporaneous links among variables. Structural VARs require "identyfing assumptions" that allow correlations to be interpreted causally.