14 Vector Autoregressions

Text: pp 312-331.

We are going to talk about **unrestricted vector autoregressions.** This is the **multivariate** (or vector) generalization of the univariate autoregressive model we covered earlier. Here, we can look at the interaction between variables and response of one variable (e.g., energy stock returns) to shocks in other variables (e.g., the Federal funds rate).

14.1 Specification

- Consider two zero-mean (or in deviations from the mean) covariance-stationary time series, $y_{1,t}$ and $y_{2,t}$

  Example: $y_{1,t}$ GDP growth, $y_{2,t}$ the market excess return. Same explanatory variables in each equation. For notational simplicity, suppress the constant.

$$y_{1,t} = a_{11,1}y_{1,t-1} + a_{12,1}y_{2,t-1} + \epsilon_{1,t}$$
$$y_{2,t} = a_{21,1}y_{1,t-1} + a_{22,1}y_{2,t-1} + \epsilon_{2,t}$$

- Write this system of equations in vector/matrix form. Since there is only one lag of each variable, the system is called a VAR(1).

$$y_t = Ay_{t-1} + \epsilon_t$$

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \ A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \ \epsilon_t = \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \overset{iid}{\sim} N\left(\underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{0}, \underbrace{\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}}_{\Sigma} \right)$$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

- This is called the **reduced form** model.

14.2 Estimation

- Estimate each equation separately by least squares.

- Estimate error-covariance matrix Σ with sample counterparts from the regression residuals.

- Select lag length with information criteria (AIC, BIC, etc).
- k is total number of regression coefficients (the $a_{ij,r}$ coefficients in system. In bivariate VAR(1) $k = 6$ including constants.
- For VAR(p),

$$\text{AIC} = 2 \ln |\hat{\Sigma}_p| + \frac{2k}{T}.$$

$$\text{BIC} = 2 \ln |\hat{\Sigma}_p| + \frac{k \ln T}{T}.$$

- $|\Sigma|$ is the determinant of the covariance matrix.

14.3 Impulse Response Analysis

- Remember MA(∞) representation of AR(1) and impulse response?

$$\begin{aligned} y_t &= \rho y_{t-1} + \epsilon_t \\ &= \epsilon_t + \rho\epsilon_{t-1} + \rho^2\epsilon_{t-2} + \rho^3\epsilon_{t-3} + \cdots \end{aligned}$$

Impulse responses

$$\begin{aligned} y_0 &= \epsilon_0 = 1 \\ y_1 &= \rho \\ y_2 &= \rho^2 \\ y_3 &= \rho^3 \end{aligned}$$

Do the same repeated substitution for VAR(1) to get the VMA(∞) (vector moving average)

$$\begin{aligned} y_t &= Ay_{t-1} + \epsilon_t \\ &= \epsilon_t + A\epsilon_{t-1} + A^2\epsilon_{t-2} + A^3\epsilon_{t-3} + \cdots \end{aligned}$$

$A^2 = AA, A^3 = AAA$ etc.

  - We have moving-average representation. Next, employ impulse response analysis to evaluate the dynamic effect of shocks in each variable on (y_{1t}, y_{2t}).
  - Two new issues (decisions). We want to simulate dynamic response of y_{1t} and y_{2t} to a shock to ϵ_{1t}

1. How big should the shock be? This is an issue because you want to compare the response of y_{1t} across different shocks. We must **normalize** the size of the shocks. Usually, people set size of shock to be one standard deviation in size.

   Divide each shock by its standard deviation. (Eviews does this automatically)

2. Need shocks that are unambiguously attributed to y_{1t} and to y_{2t}. If ϵ_{1t} and ϵ_{2t} are correlated, you can't just shock ϵ_{1t} and hold ϵ_{2t} constant. **We need to make the shocks uncorrelated. (Orthogonalizing the shocks).**

- **Orthogonalizing Correlated Variables**. Here is the idea behind orthogonalizing (decorrelating) correlated variables. Not covering the actual way VARs are orthogonalized, just the concepts.

  - Show how to build up correlated random variables from independent random variables.
  - Run the process in reverse to orthogonalize
  - Creating Bi-variate Normal Random Variables
    * Let z_1 and z_2 be independent standard normal random variables. Build the random variables ϵ_1 and ϵ_2 as linear combinations of z_1 and z_2.

$$\begin{aligned} \epsilon_1 &= \sigma_1 z_1 + \mu_1 \\ \epsilon_2 &= \sigma_2 \left(\rho z_1 + \sqrt{(1-\rho^2)} \right) z_2 + \mu_2 \end{aligned}$$

    * ϵ_1 and ϵ_2 are normally distributed. That's because they are linear combinations of normals.
    * See the overlap of z_1 in both ϵ_1 and ϵ_2? That means they are correlated.

$$\begin{aligned} E(\epsilon_1) &= \mu_1, \; E(\epsilon_2) = \mu_2 \\ \text{Var}(\epsilon_1) &= \sigma_1^2, \; \text{Var}(\epsilon_2) = \sigma_2^2 \left(\rho^2 + \left(1-\rho^2\right)\right) = \sigma_2^2 \\ \text{Cov}(\epsilon_1, \epsilon_2) &= \text{E}\left(\sigma_1 z_1 \left(\sigma_2 \left(\rho z_1 + \sqrt{(1-\rho^2)} \right) z_2 \right) \right) = \sigma_1 \sigma_2 \rho \\ \text{Corr}(\epsilon_1, \epsilon_2) &= \rho \end{aligned}$$

    * We built the $\epsilon' s$ from the $z' s$, so given the $\epsilon' s$, we should be able to unpack the $z' s$.
    * The ϵ_1 and ϵ_2 are like the reduced form errors in the VAR. The $z' s$ are like what we call **structural** shocks.

– Reverse Engineer–Recover the $z's$

$$z_1 = \frac{1}{\sigma_1}(\epsilon_1 - \mu_1)$$
$$z_2 = \frac{\sigma_1(\epsilon_2 - \mu_2) - \rho\sigma_2(\epsilon_1 - \mu_1)}{\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

VAR method uses something called the **Choleski (**or Choleski) decomposition of the error covariance matrix, Σ, to do this.

– Orthogonalized (uncorrelated) shocks. Write in matrix form

$$\underbrace{\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}}_{\epsilon_t} = \underbrace{\begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{pmatrix}}_{\Lambda} \underbrace{\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix}}_{z_t}$$
$$\epsilon_t = \Lambda z_t$$

substitute into vector MA(∞) representation.

$$y_t = \epsilon_t + A\epsilon_{t-1} + A^2\epsilon_{t-2} + \cdots$$
$$= \Lambda z_t + A\Lambda z_{t-1} + A^2\Lambda z_{t-2} + \cdots$$

Now we can shock z_2 without disturbing z_1.

– The ordering of the variables can matter, because of the triangular structure of Λ. A time-t shock z_{1t} affects y_{1t} and y_{2t} today. A time-t shock z_{2t} affects y_{2t} today, but y_{1t} with a one-period lag. Let's write out a couple terms:

$$\Lambda z_t = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} = \begin{pmatrix} z_{1t} \\ \rho z_{1t} + z_{2t}\sqrt{-\rho^2+1} \end{pmatrix}$$
$$A\Lambda z_{t-1} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} z_{1t-1} \\ z_{2t-1} \end{pmatrix} = \begin{pmatrix} (a_{11}+\rho a_{12}) z_{1t-1} + a_{12} z_{2t-1}\sqrt{-\rho^2+1} \\ (a_{21}+\rho a_{22}) z_{1t-1} + a_{22} z_{2t-1}\sqrt{-\rho^2+1} \end{pmatrix}$$

- **Some Examples in Eviews. VAR_example.wf1 in Eviews folder**

15 Local Projections

15.1 Pitfalls of VARs

Some pitfalls of VARs have been discovered.

- VAR is optimally designed for one-period ahead forecasting.
- An impulse response, is a function of forecasts at increasingly distant horizons. Therefore misspecification errors are compounded with the forecast horizon.
- It might be better to use a collection of projections local to each forecast horizon instead. This is called a **local projection**.

15.2 Local Projection instead of VAR(1)

The first equation in the VAR is

$$y_{1t+1} = a_1 y_{1t} + b_1 y_{2t} + \epsilon_{t+1,1}$$

Run these regressions. (Look at the dependent variable–is shifted ahead k periods). These are predictive regressions.

$$\begin{aligned}
y_{1t+2} &= a_2 y_{1t} + b_2 y_{2t} + \epsilon_{t+2,2} \\
y_{1,t+3} &= a_3 y_{1t} + b_3 y_{2t} + \epsilon_{t+3,3} \\
&\vdots \\
y_{1,t+k} &= a_k y_{1t} + b_k y_{2t} + \epsilon_{t+k,k}
\end{aligned}$$

- The impulse response of y_1 to a shock to iself is $a_{1,}a_2, ..., a_k$.
- The impulse response of y_1 to a shock to y_2 is $b_1, b_2, ..., b_k$.
- Òscar Jordà worked out the math to prove, if the true DGP is the VAR, the impulse responses from Local Projections and the VAR are identical (asymptotically).
- Construct confidence bands with Newey-West standard errors (the estimate divided by the t-ratio).

15.3 The Bottom Line

Generalizing to beyond a VAR(1). If we want to do local projection to get impulse responses instead of a VAR(p), you regress $y_{1,t+k}$ on a constant, the current value and p lags of $y_{1,t}$ and $y_{2,t}$.

The impulse response of y_1 to a shock in itself is again, the sequence of slope estimates for $y_{1,t}$ and the response to a shock in $y_{2,t}$ is the sequence of slope estimates on $y_{2,t}$.