Time Series Analysis

Vector Autoregression (VAR) Models

Jude C. Hays

February 9, 2021

Outline

Introduction to VAR Models

- Box-Jenkins intervention models can be used to analyze the effect of a deterministic event (or treatment) on the mean of a time series y_t at the point when the intervention occurs.
- Transfer function models allow us to analyze the effect of a stochastic intervention z_t on y_t, assuming there is no feedback from y_t to z_t.
- Vector autoregression (VAR) models can be used to analyze the properties of a system of equations in which all the variables are jointly endogenous.

Introduction to VAR Models

- A VAR is an n-equation, n-variable (linear) model in which each variable is explained by its own lagged values and lagged values of the remaining n-1 variables.
- Sims (1980) developed the VAR model to address problems with structural equation models, which require many assumptions for identification.

"Because existing large models contain too many *incredible restrictions*, empirical research aimed at testing competing macroeconomic theories too often proceeds in a single- or few-equation framework. For this reason alone, it appears worthwhile to investigate the possibility of building large models in a style which does not tend to accumulate restrictions so haphazardly...It should be feasible to estimate large-scale macromodels as unrestricted reduced forms, treating all variables as endogenous" (Sims 1980, emphasis added).

Structural and Standard (Reduced) Forms

- As with an AR process, the order of the VAR (p) determines the number of lags included for each variable.
- Thus, each equation in a VAR contains $np + 1$ parameters and the systems as a whole has $n^2 p + n$ parameters.
- In the simple two-variable case, the *structural version* of the *first-order* VAR model is

$$\begin{aligned} y_t &= b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt} \\ z_t &= b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \end{aligned}$$

where y_t and z_t are assumed stationary and ε_{yt} and ε_{zt}, the *structural disturbances*, are uncorrelated white-noise disturbances with standard deviations σ_y and σ_z respectively.

- Note that we can rewrite this system as

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

Structural and Standard (Reduced) Forms

- Or, more compactly, using matrix notation, as

$$\mathbf{B}\mathbf{x_t} = \mathbf{\Gamma_0} + \mathbf{\Gamma_1}\mathbf{x_{t-1}} + \varepsilon_\mathbf{t}$$

- Premultiplying both sides by $\mathbf{B}^{-1}$ gives the VAR in its *standard* or *reduced* form

$$\mathbf{x_t} = \mathbf{A_0} + \mathbf{A_1}\mathbf{x_{t-1}} + \mathbf{e_t}$$

where $\mathbf{A_0} = \mathbf{B}^{-1}\mathbf{\Gamma_0}$, $\mathbf{A}_1 = \mathbf{B}^{-1}\mathbf{\Gamma}_1$ and $\mathbf{e_t} = \mathbf{B}^{-1}\varepsilon_\mathbf{t}$.

- Given $\mathbf{B}^{-1} = \begin{bmatrix} 1/1 - b_{12}b_{21} & -b_{12}/1 - b_{12}b_{21} \\ -b_{21}/1 - b_{12}b_{21} & 1/1 - b_{12}b_{21} \end{bmatrix}$, the reduced-form disturbances are

$$\begin{aligned} e_{1t} &= (\varepsilon_{yt} - b_{12}\varepsilon_{zt})/(1 - b_{12}b_{21}) \\ e_{2t} &= (\varepsilon_{zt} - b_{21}\varepsilon_{yt})/(1 - b_{12}b_{21}) \end{aligned}$$

Reduced-form Disturbances

- Given our assumption that the structural disturbances, ε_{yt} and ε_{zt}, are white noise, the reduced-form disturbances, e_{1t} and e_{2t}, have zero means, constant variances, and are individually serially uncorrelated.
- Unlike the structural disturbances, however, the reduced-form disturbances will be correlated

$$Ee_{1t}e_{2t} = -(b_{21}\sigma_y^2 + b_{12}\sigma_z^2)/(1 - b_{12}b_{21})^2$$

  if either b_{12} or b_{21} is nonzero.
- We will refer to the variance-covariance matrix of the reduced-form disturbances as

$$\mathbf{\Sigma} = \begin{bmatrix} \mathrm{var}(e_{1t}) & \mathrm{cov}(e_{1t}, e_{2t}) \\ \mathrm{cov}(e_{1t}, e_{2t}) & \mathrm{var}(e_{2t}) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

Identifying Structural VAR Models

- Note that the structural version of the VAR cannot be estimated because of the contemporaneous (zero-order) relationships between y_t and z_t, which imply $\mathrm{cov}(y_t, \varepsilon_{zt}) \neq 0$ and $\mathrm{cov}(z_t, \varepsilon_{yt}) \neq 0$.
- The parameters of the structural model are not identified, but the parameters of the reduced-form model are.
- We can, however, identify the parameters of the structural model, if we are willing to make some restrictions.
- To understand how we can gain identification by imposing restrictions, think about the parameters we can estimate from the reduced-form first-order VAR model. There are nine: six coefficients (a_{10}, a_{20}, a_{11}, a_{12}, a_{21} and a_{22}) and the three elements of the variance-covariance matrix for the reduced-form residuals ($\mathrm{var}(e_{1t})$, $\mathrm{var}(e_{2t})$ and $\mathrm{cov}(e_{1t}, e_{2t})$).

Identifying Structural VAR Models

- The structural version of the VAR has ten parameters: eight coefficients (b_{10}, b_{20}, b_{12}, b_{21}, γ_{11}, γ_{12}, γ_{21} and γ_{22}) and two unknown elements of the variance-covariance matrix for the structural disturbances (σ_y^2 and σ_z^2).
- Hence, we need one restriction, leaving nine unknown parameters of the structural model. In this case, with the reduced-form parameter estimates, we have nine quantities that we can plug into nine equations to solve for the nine unknown parameters of the structural model.
- One very common approach is to impose zero-order restrictions that makes the system of equations recursive.
- For example, if we restrict b_{21} to be zero, then z_t has a contemporaneous effect on y_t, but y_t affects z_t with a one-period lag.

Identifying Structural VAR Models

- Again, it is the contemporaneous relationships in the **B** matrix that make the structural model unidentified.
- In an n-variable VAR model, where the **B** matrix is $n \times n$, we need to make $(n^2 - n)/2$ restrictions to be identified.
- Importantly, identification allows us to distinguish the reduced form residuals (e_{1t} and e_{2t}) from the structural innovations (ε_{yt} and ε_{zt}).
- In our simple case

$$\mathbf{B} = \begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}, \mathbf{B}^{-1} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix}, \text{ which implies}$$

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & -b_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

Identifying Structural VAR Models

- To create a recursive system of equations, we make **B** upper triangular.
- This forces $(n^2 - n)/2$ elements of **B** to be zero.
- When we impose this kind of recursive structure on the relationship between the reduced-form and structural disturbances (via $\mathbf{B}^{-1}$), we call it a **Choleski decomposition**.
- The Choleski decomposition creates an **ordering** of the variables where some are *causally prior* to others.
- Once we are identified, we can do *innovation accounting* analysis.

The Impulse Response Function

- Just as a univariate autoregressive (AR) process has a moving average (MA) representation, a vector autoregression (VAR) has a vector moving average (VMA) representation.
- Our reduced-form (first-order) VAR is

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

- Without an initial condition, the solution to this model is

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

The Impulse Response Function

- We can write this in terms of the structural disturbances by substituting $\mathbf{e_t} = \mathbf{B}^{-1}\varepsilon_\mathbf{t}$

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

- It is useful to simplify this expression by defining

$$\phi_i = \frac{A^i}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

which gives

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

The Impulse Response Function

- The four sets of coefficients–$\phi_{11}(i)$, $\phi_{12}(i)$, $\phi_{21}(i)$, $\phi_{22}(i)$–are the impulse response functions.
- The coefficient $\phi_{jk}(i)$ gives the effect of a one-unit innovation in variable k at time t on variable j at time $t+i$.
- We typically plot these coefficients to display how the variables in the system respond to various shocks.
- It is best to bootstrap the confidence intervals for the $\phi_{jk}(i)$.

Forecast Error Variance Decomposition

- The *forecast error variance decomposition* tells us the proportion of the movement in the time series for a particular variable that is attributable to its own innovations/shocks versus the innovations/shocks in other variables.
- This gives us a nice quantitative summary of how important one variable is for explaining the dynamics of another.
- More specifically, in the case of our two-equation VAR for $\{y_t\}$ and $\{z_t\}$, this decomposition tells us how much of our uncertainty about future values of y_t is attributable to the (unpredictable) white noise process underlying y_t versus the (unpredictable) white noise process underlying z_t.

Forecast Error Variance Decomposition

- To better understand *forecast error variance decomposition* it is useful to compare forecast error variance in the univariate and multivariate cases.
- The time t forecast error for y_{t+j} is $e_t(j) = y_{t+j} - E_t y_{t+j}$.
- In the univariate AR(1) case, for example, the time t forecast error for y_{t+2} is

$$\begin{aligned} e_t(2) &= a_1(y_{t+1} - E_t y_{t+1}) + \varepsilon_{t+2} \\ &= \varepsilon_{t+2} + a_1 \varepsilon_{t+1} \end{aligned}$$

Forecast Error Variance Decomposition

- This generalizes to j periods into the future as

$$e_t(j) = \varepsilon_{t+j} + a_1\varepsilon_{t+j-1} + \ldots + a_1^{j-1}\varepsilon_{t+1}$$

- The variance of this forecast error is

$$\mathrm{var}[e_t(j)] = \sigma^2[1 + a_1^2 + \ldots + a_1^{2(j-1)}]$$

- In the case of a p^{th}-order VAR for $\{y_t\}$ and $\{z_t\}$, the equivalent quantities for forecasts of y_t are

$$\begin{gathered} e_{yt}(j) = \phi_{11}(0)\varepsilon_{yt+j} + \phi_{11}(1)\varepsilon_{yt+j-1} + \ldots + \phi_{11}(j-1)\varepsilon_{yt+1} \\ +\phi_{12}(0)\varepsilon_{zt+j} + \phi_{21}(1)\varepsilon_{zt+j-1} + \ldots + \phi_{21}(j-1)\varepsilon_{zt+1} \end{gathered}$$

$$\begin{gathered} \mathrm{var}[e_{yt}(j)] = \sigma_y^2[\phi_{11}(0)^2 + \phi_{11}(1)^2 + \ldots + \phi_{11}(j-1)^2] \\ +\sigma_z^2[\phi_{12}(0)^2 + \phi_{21}(1)^2 + \ldots + \phi_{21}(j-1)^2] \end{gathered}$$

Determining Lag Length and Selecting Variables

- To this point we have worked with a two-equation first-order VAR. In practice, we will not know the values of p and n beforehand, and so one of the first steps in our analysis is to specify the reduced-form VAR.
- To this end, we use a Likelihood Ratio test of cross-equation restrictions and the multivariate generalizations of the AIC and SBC.
- The Likelihood Ratio Statistic is

$$LR = (T - c)(\ln|\Sigma_r| - \ln|\Sigma_{ur}|)$$

where T is the number of usable observations, c is is the number of parameters estimated in each equation of the unrestricted model, and $|\Sigma|$ is the determinant of the variance covariance matrix for the restricted (r) and unrestricted (ur) models.

Determining Lag Length and Selecting Variables

- Under the null hypothesis, LR is distributed as a χ^2 with degrees of freedom equal to the overall number of restrictions *in the system*.
- If we are evaluating lag length, the overall number of restrictions in the system is $n^2(\Delta p)$. For example, in a four-equation VAR, if we are testing 8 versus 4 lags, the overall number of restrictions is 64.
- If we are evaluating whether or not to add another variable to an n-equation system, the overall number of restrictions is np. For example, in a two-equation fourth-order VAR, if we are testing whether to add a third variable, the number of restrictions is 8. The unrestricted model has two equations with 13 parameters each, and the restricted model has two equations with 9 parameters each.

Determining Lag Length and Selecting Variables

- The multivariate generalizations of the AIC and SBC are

$$\begin{aligned} AIC &= T \ln |\Sigma| + 2N \\ SBC &= T \ln |\Sigma| + N \ln(T) \end{aligned}$$

where N is the total number of parameters in the system of equations.

Granger Causality Tests

- Granger causality implies that current and past values of one time series are useful for forecasting future values of another time series.
- In the context of a VAR, this amounts to a F-test of the null hypothesis that all of the coefficients on the lags of a particular variable are jointly zero.
- For example, in a p^{th}-order VAR for $\{y_t\}$ and $\{z_t\}$, y_t does not Granger cause z_t if

$$a_{21}(1) = a_{21}(2) = ... = a_{21}(p) = 0$$

- Note that Granger causality tests do not evaluate the contemporaneous (zero-order) relationship between y_t and z_t, which could be causal.

How to Conduct VAR Analysis

1. Select the variables for the VAR model based on your theory.
2. Choose the appropriate lag length for the model and evaluate your variable selection using the likelihood ratio test and the multivariate generalizations of AIC and SBC.
3. Conduct Granger causality tests for all the variables in the model.
4. Conduct Innovation Accounting analysis.