Vector Autoregressive Models

Richard G. Pierse

1 Vector Autoregressive Models

The pth order vector autoregressive model or *VAR* model can be written as

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

with

$$\boldsymbol{\varepsilon}_t \sim iid\ N(\mathbf{0}, \mathbf{\Omega})\,.$$

where y_t is a $n \times 1$ vector of variables at time t and $\mathbf{c}$ is an intercept. Other deterministic components can be added to the model without affecting the analysis. There are pn^2 parameters in the $\mathbf{\Phi}$ matrices. Making use of the lag operator L, defined by $L^k x_t = x_{t-k}$, the equation can be rewritten as

$$\mathbf{\Phi}(L)\mathbf{y}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t \tag{1}$$

where

$$\mathbf{\Phi}(L) = \mathbf{\Phi}_0 L^0 - \mathbf{\Phi}_1 L^1 - \cdots - \mathbf{\Phi}_p L^p\,,$$

$\mathbf{\Phi}_0 = \mathbf{I}$, and, to ensure stationarity, the roots of $|\mathbf{\Phi}(L)|$ lie *outside the unit circle*.

In 'pure' *VAR* models no *a priori* economic restrictions are imposed on $\mathbf{\Phi}$.

1.1 Estimating *VAR* Models

The restriction that $\mathbf{\Phi}_0 = \mathbf{I}$ implies that there are no current endogenous variables in the model. Equations are related solely through the off-diagonal elements in the covariance matrix $\mathbf{\Omega}$. Thus the maximum likelihood estimator for $\mathbf{\Phi}$ is simply the *OLS* estimator. The error covariance matrix $\mathbf{\Omega}$ is consistently estimated by

$$\widehat{\mathbf{\Omega}} = \frac{1}{T}\sum_{t=1}^{T} \widehat{\boldsymbol{\varepsilon}}_t \widehat{\boldsymbol{\varepsilon}}_t'$$

where $\widehat{\boldsymbol{\varepsilon}}_t$ is the $n \times 1$ vector of *OLS* residuals.

1.2 Granger causality and *VARs*

Granger causality asks how useful is a set of variables $\mathbf{y}_2$ for forecasting another set of variables $\mathbf{y}_1$. It has nothing to do with the notion of causality in any philosophical sense.

Definition 1 $\mathbf{y}_2$ *does not Granger cause* $\mathbf{y}_1$ *if for all* $s > 0$

$$\begin{aligned} & \text{MSE}\ [\widehat{E}(\mathbf{y}_{1,t+s} \mid \mathbf{y}_{1,t}, \mathbf{y}_{1,t-1}, \cdots)] \\ = \ & \text{MSE}\ [\widehat{E}(\mathbf{y}_{1,t+s} \mid \mathbf{y}_{1,t}, \mathbf{y}_{1,t-1}, \cdots, \mathbf{y}_{2,t}, \mathbf{y}_{2,t-1}, \cdots)] \end{aligned}$$

Consider the model

$$\begin{bmatrix} \mathbf{\Phi}_{11}(L) & \mathbf{\Phi}_{12}(L) \\ \mathbf{\Phi}_{21}(L) & \mathbf{\Phi}_{22}(L) \end{bmatrix} \begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{y}_{2t} \end{bmatrix} = \mathbf{c} + \boldsymbol{\varepsilon}_t$$

where

$$\boldsymbol{\varepsilon}_t \sim iid\ N(\mathbf{0}, \mathbf{I})$$

If $\mathbf{\Phi}_{12}(L) = 0$ in this model *then* $\mathbf{y}_2$ does not Granger cause $\mathbf{y}_1$.

1.3 Impulse Response Functions

The *VAR* model can always be transformed into the infinite Moving Average (*Wold*) representation:

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \mathbf{\Psi}_1 \boldsymbol{\varepsilon}_{t-1} + \mathbf{\Psi}_2 \boldsymbol{\varepsilon}_{t-2} + \cdots +$$

where

$$\frac{\partial\ \mathbf{y}_{t+s}}{\partial\ \boldsymbol{\varepsilon}_t} = \mathbf{\Psi}_s$$

is the response of $\mathbf{y}$ in period $t+s$ to a shock in period s. Considered as a function of s, $\mathbf{\Psi}_s$ is known as the *impulse response function.*

The matrices $\mathbf{\Psi}_s$ can be computed *recursively* using the formula

$$\mathbf{\Psi}_s = \sum_{j=1}^{\min(p,s)} \mathbf{\Psi}_{s-j} \mathbf{\Phi}_j \quad , \quad s = 1, 2, 3, \cdots$$

where $\mathbf{\Psi}_0 = \mathbf{I}$. Alternatively, they can be generated by simulation of the *VAR* model.

1.4 Orthogonal shocks

Note that the off-diagonal terms in $\mathbf{\Omega}$ imply that the shocks $\boldsymbol{\varepsilon}_t$ are *correlated* between equations. However, $\mathbf{\Omega}^{-1}$ can always be factorised as

$$\mathbf{\Omega}^{-1} = \mathbf{HH}'$$

where $\mathbf{H}$ is a lower triangular matrix having zeros above the diagonal (This is known as a *Cholesky decomposition*). It then follows that

$$\boldsymbol{\varepsilon}_t^* = \mathbf{H}'\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \mathbf{H}'\mathbf{\Omega}\mathbf{H}) = N(\mathbf{0}, \mathbf{I})$$

and the transformed shocks $\boldsymbol{\varepsilon}_t^*$ are *orthogonal*. This orthogonalisation *is not unique* and depends on the ordering of the variables in the *VAR*. Sometimes there may be a natural *recursive* ordering of the equations that justifies the Cholesky factorisation.

Pesaran (1996) proposes a generalised orthogonalisation that does not depend on the ordering of variables in the *VAR*. This is implemented in version 4 of the *MicroFit* computer package.

2 Structural *VAR* Models

2.1 'Pure' *VAR* Model (Atheoretic)

First consider the *atheoretic VAR* model in $MA(\infty)$ form

$$\mathbf{y}_t = \mathbf{\Psi}(L)\boldsymbol{\varepsilon}_t = \mathbf{\Psi}^*(L)\boldsymbol{\varepsilon}_t^*$$

where $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \mathbf{\Omega}$ and $\boldsymbol{\varepsilon}_t^* = \mathbf{H}'\boldsymbol{\varepsilon}_t$ so that

$$\mathbf{\Psi}^*(L) = \mathbf{\Psi}(L)\mathbf{H}'$$

and the shocks $\boldsymbol{\varepsilon}_t^*$ are *orthogonal* with covariance matrix

$$E(\boldsymbol{\varepsilon}_t^*\boldsymbol{\varepsilon}_t^{*\prime}) = \mathbf{H}'\mathbf{\Omega}\mathbf{H} = \mathbf{I}\,.$$

2.2 Structural Model (Economic)

Compare this with the model

$$\mathbf{y}_t = \mathbf{\Phi}(L)^{-1}\boldsymbol{\eta}_t$$

where the errors $\boldsymbol{\eta}_t$ are assumed to be structural, being derived from economic theory. We make the *assumption* that the structural errors $\boldsymbol{\eta}_t$ are *orthogonal*:

$$\boldsymbol{\eta}_t \sim iid\ N(\mathbf{0}, \mathbf{I})\,.$$

We then have the relationship between the two models that

$$\boldsymbol{\eta}_t = \mathbf{S}\boldsymbol{\varepsilon}_t^*$$

so that

$$\boldsymbol{\Phi}(L)^{-1}\mathbf{S} = \boldsymbol{\Psi}^*(L)$$

where the matrix $\mathbf{S}$ is *orthogonal* since

$$\begin{aligned} \mathbf{S}\boldsymbol{\varepsilon}_t^* &\sim iid\ N(\mathbf{0}, \mathbf{S}\mathbf{S}') = \boldsymbol{\eta}_t \sim iid\ N(\mathbf{0}, \mathbf{I}) \\ &\Rightarrow \mathbf{S}\mathbf{S}' = \mathbf{I} \quad , \quad \mathbf{S}^{-1} = \mathbf{S}' \quad , \quad \mathbf{S}'\mathbf{S} = \mathbf{I}\,. \end{aligned}$$

The economic restrictions on $\boldsymbol{\varepsilon}_t^*$ are expressed in terms of the matrix $\mathbf{S}$ alone. Imposing structure is equivalent to identifying $\mathbf{S}$.

2.3 *VAR* models and Structural models

2.3.1 Simultaneous Equations Models

There is a close relationship between *VAR* models and simultaneous equations models. Consider the standard simultaneous equations model

$$\mathbf{B}_0\,\mathbf{y}_t = \boldsymbol{\Gamma}\,\mathbf{z}_t + \mathbf{u}_t \quad , \quad \mathbf{B}_0 \neq \mathbf{I} \tag{2}$$

$$\mathbf{u}_t \sim iid\ N(\mathbf{0}, \boldsymbol{\Sigma})$$

where $\mathbf{z}_t$ is a set of *predetermined* variables comprising:

a) lags of endogenous variables $\mathbf{y}_t$

b) current and lagged *exogenous* variables $\mathbf{x}_t$ that are *not* explained within the system

Let the two components of $\mathbf{z}_t$ be separated out into

$$\boldsymbol{\Gamma}\mathbf{z}_t = \mathbf{B}^*(L)\mathbf{y}_{t-1} + \mathbf{C}(L)\mathbf{x}_t\,.$$

Then the system can be completed by adding a *VAR* model for $\mathbf{x}_t$:

$$\mathbf{D}(L)\,\mathbf{x}_t = \mathbf{v}_t \quad , \quad \mathbf{D}_0 = \mathbf{I}$$

so that, stacking,

$$\begin{bmatrix} \mathbf{B}(L) & -\mathbf{C}(L) \\ \mathbf{0} & \mathbf{D}(L) \end{bmatrix} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{x}_t \end{bmatrix} = \begin{matrix} \mathbf{u}_t \\ \mathbf{v}_t \end{matrix}$$

where $\mathbf{B}(L) = \mathbf{B}_0 - \mathbf{B}^*(L)L$. This can be re-expressed as

$$\mathbf{\Phi}(L)\,\mathbf{w}_t = \boldsymbol{\eta}_t \qquad (3)$$

where $\mathbf{w}_t = (\mathbf{y}_t' : \mathbf{x}_t')'$ and $\boldsymbol{\eta}_t = (\mathbf{u}_t' : \mathbf{v}_t')'$ and the first term in $\mathbf{\Phi}(L)$ is given by

$$\mathbf{\Phi}_0 = \begin{bmatrix} \mathbf{B}_0 & -\mathbf{C}_0 \\ \mathbf{0} & \mathbf{I} \end{bmatrix}.$$

The matrices $\mathbf{B}(L)$ and $\mathbf{C}(L)$ will typically have *overidentifying* restrictions (usually zero restrictions) deriving from economic theory.

2.3.2 The Reduced and Final Forms

Now define

$$\mathbf{\Phi}(L) = \mathbf{\Phi}_0\mathbf{\Phi}^*(L)$$

where $\mathbf{\Phi}_0^* = \mathbf{I}$. Then

$$\mathbf{\Phi}^*(L)\ \mathbf{w}_t = \mathbf{\Phi}_0^{-1}\boldsymbol{\eta}_t = \boldsymbol{\eta}_t^* \qquad (4)$$

is the *reduced form* of the system. In form this is a *VAR* model but the structure imposes *restrictions* on $\mathbf{\Phi}^*(L)$.

Inverting the lag polynomial $\mathbf{\Phi}^*(L)$ gives the *final form* representation

$$\mathbf{w}_t = \mathbf{\Phi}^*(L)^{-1}\,\mathbf{\Phi}_0^{-1}\boldsymbol{\eta}_t = \mathbf{\Phi}(L)^{-1}\boldsymbol{\eta}_t$$

which is an infinite moving average representation in terms of the *structural* errors $\boldsymbol{\eta}_t$.

Thus the simultaneous equations system (2) can be seen to be a *VAR* system like (4) that is *subject to a set of (over)identifying restrictions.* The real difference between structural *VARs* and conventional structural equation systems is the nature of the identification restrictions that are imposed.

3 Identification restrictions

Identification in the model (3) requires the imposition of n^2 restrictions. In conventional simultaneous equations models, identification is generally achieved by imposing zero restrictions on the coefficients on the predetermined variables in the matrices $\mathbf{\Phi}_1, \cdots, \mathbf{\Phi}_p$. Sims (1980) argues against this type of identifying restriction on the dynamics. Instead, structural VAR modellers have sought to impose identifying restrictions either on the matrix of contemporaneous coefficients $\mathbf{\Phi}_0$, on the covariance matrix $\mathbf{\Omega}$, or on the *long run* coefficients

$$\mathbf{\Phi}(1) = \mathbf{\Phi}_0 - \mathbf{\Phi}_1 \cdots - \mathbf{\Phi}_p = \mathbf{\Psi}(1)^{-1}.$$

For example Sims (1980) suggests imposing the $n \times n - 1$ restrictions that $\mathbf{\Omega}$ is diagonal, plus the $n \times n + 1$ restrictions that $\mathbf{\Phi}_0$ is lower triangular with ones on the diagonal. These restrictions together make the system *recursive* and exactly identified. By contrast, Blanchard and Quah (1989) and King *et al.* (1991) identify a structural VAR through restrictions on $\mathbf{\Phi}(1)$.

4 Further Reading

Watson (1994) in Volume IV of the *Handbook of Econometrics* treats both pure and structural VAR models and has a very comprehensive bibliography. Hamilton (1994) and Lütkepohl (1993) also contain good treatments of 'pure' VAR models.

References

[1] Blanchard, O.J. and D. Quah (1989), 'The dynamic effects of aggregate demand and supply disturbances', American Economic Review, 79, 655–73.

[2] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, NJ.

[3] King, R.G., C.I. Plosser, J.H. Stock and M.W. Watson (1991), 'Stochastic trends and economic fluctuations', *American Economic Review*, 81, 819–840.

[4] Lütkepohl, H. (1993), *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

[5] Pesaran, M.H. (1996), 'MicroFit Version 4 Manual', *mimeo.*

[6] Sims, C.A. (1980), 'Macroeconomics and reality', *Econometrica*, 48, 1–48.

[7] Watson, M.W. (1994), 'Vector autoregressions and cointegration', Chapter 47 in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics, Volume IV*, North-Holland, Amsterdam.