VAR PROCESSES

1 Introduction to Vector Processes

Reading on matrices:
Chris Orme, *Lecture Notes in Linear Algebra*, cost £1.00, from Room N.4.3, Dover St. building.

Suppose we want to model k related time series $y_{1t}, y_{2t}, \ldots, y_{kt}$. Define the vector

$$\mathbf{y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{kt} \end{bmatrix}, \qquad (k \times 1)$$

For example, to analyse monetary policy, might use rate of interest (r_t), the output growth (g_t) and the rate of inflation (Δp_t):

$$\mathbf{y}_t = \begin{bmatrix} r_t \\ g_t \\ \Delta p_t \end{bmatrix}.$$

Develop a multivariate time series model for $\mathbf{y}_t$.

1.1 Vector white noise

The $k \times 1$ vector white noise process $\boldsymbol{\varepsilon}_t$ satisfies

$$\begin{aligned} E(\boldsymbol{\varepsilon}_t) &= \mathbf{0} \\ E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') &= \boldsymbol{\Sigma} \\ E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s') &= \mathbf{0}, s \neq t \end{aligned}$$

Thus:

- each element has mean zero;
- variance-covariance matrix is constant over time;
- elements have zero autocorrelations and zero cross-correlations over time. For example, $k = 2$ & $s = t - 1$, we require

$$E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_{t-1}') = \begin{bmatrix} E(\varepsilon_{1t}\varepsilon_{1,t-1}) & E(\varepsilon_{1t}\varepsilon_{2,t-1}) \\ E(\varepsilon_{2t}\varepsilon_{1,t-1}) & E(\varepsilon_{2t}\varepsilon_{2,t-1}) \end{bmatrix} = \mathbf{0}.$$

Vector white noise has not only $E(\varepsilon_{it}\varepsilon_{i,t-j}) = 0, j = 1, 2, ...,$
but also cross-correlations over time,

eg. $E(\varepsilon_{1t}\varepsilon_{2,t-j}) = E(\varepsilon_{2t}\varepsilon_{1,t-j}) = 0, j = 1, 2, ...$

Implication:

All past elements of $\boldsymbol{\varepsilon}_{t-j}$ are uncorrelated with current $\boldsymbol{\varepsilon}_t$.

1.2 VAR(P) processes

Vector autoregressive process of order P, or VAR(P), is

$$\mathbf{y}_t = \boldsymbol{\delta} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + ... + \boldsymbol{\Phi}_P \mathbf{y}_{t-P} + \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\varepsilon}_t$ is vector white noise with $E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Sigma}$.

Analogous to AR(p).

Note:
Any interrelations at t between y_{it} captured in $\boldsymbol{\Sigma}$;
All interrelations over time between y_{it} & $y_{k,t-j}$ captured by VAR coefficients.

In lag operator notation:

$$\boldsymbol{\Phi}(L)\mathbf{y}_t = \boldsymbol{\delta} + \boldsymbol{\varepsilon}_t$$

where $\boldsymbol{\Phi}(L)$ is the $k \times k$ matrix polynomial

$$\boldsymbol{\Phi}(L) = \mathbf{I}_k - \boldsymbol{\Phi}_1 L - \boldsymbol{\Phi}_2 L^2 - ... - \boldsymbol{\Phi}_P L^P.$$

Example: VAR(1) with $\boldsymbol{\delta} = \mathbf{0}$ is

$$\mathbf{y}_t = \boldsymbol{\delta} + \mathbf{\Phi}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t.$$

For $k = 2$:

$$\begin{aligned}
\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} &= \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \\
&= \begin{bmatrix} \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \varepsilon_{1t} \\ \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + \varepsilon_{2t} \end{bmatrix}
\end{aligned}$$

with

$$\begin{aligned}
E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') &= \begin{bmatrix} E(\varepsilon_{1t}^2) & E(\varepsilon_{1t}\varepsilon_{2t}) \\ E(\varepsilon_{2t}\varepsilon_{1t}) & E(\varepsilon_{2t}^2) \end{bmatrix} \\
&= \mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}
\end{aligned}$$

and

$$\mathbf{\Phi}(L) = \mathbf{I}_2 - \mathbf{\Phi}L = \begin{bmatrix} 1 - \phi_{11}L & -\phi_{12}L \\ -\phi_{21}L & 1 - \phi_{22}L \end{bmatrix}$$

Points to note:

- For VAR with $P > 1$, an additional subscript (or superscript) is needed for elements of $\mathbf{\Phi}_j$ to indicate lag j, in addition to the two subscripts used in a VAR(1) for the equation and the variable.
- The VAR(P) treats each of the k variables in the same way: there is no distinction between endogenous and exogenous variables.
- The VAR focuses on dynamic interrelationships between the variables.

2 Properties of VAR Processes

2.1 Stationarity

A VAR process is (second-order) stationary if:

$$\begin{aligned} E(\mathbf{y}_t) &= \boldsymbol{\mu} \text{ for all } t \\ E(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})' &= \boldsymbol{\Gamma}_0 \text{ for all } t \\ E(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-j} - \boldsymbol{\mu})' &= \boldsymbol{\Gamma}_j \text{ for all } t \text{ and any } j \end{aligned}$$

$\boldsymbol{\Gamma}_j = E(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-j} - \boldsymbol{\mu})'$ is the autocovariance matrix at lag j.

Example: VAR between interest rates & inflation, at lag $j = 1$:

$$\begin{aligned} \boldsymbol{\Gamma}_1 &= E(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-1} - \boldsymbol{\mu})' \\ &= E\begin{bmatrix} r_t - \mu_r \\ \Delta p_t - \mu_{\Delta p} \end{bmatrix} \begin{bmatrix} r_{t-1} - \mu_r & \Delta p_{t-1} - \mu_{\Delta p} \end{bmatrix} \\ &= \begin{bmatrix} E(r_t - \mu_r)(r_{t-1} - \mu_r) & E(r_t - \mu_r)(\Delta p_{t-1} - \mu_{\Delta p}) \\ E(\Delta p_t - \mu_{\Delta p})(r_{t-1} - \mu_r) & E(\Delta p_t - \mu_{\Delta p})(\Delta p_{t-1} - \mu_{\Delta p}) \end{bmatrix} \end{aligned}$$

Diagonal elements of $\boldsymbol{\Gamma}_j$ are lag j autocovariances; Off-diagonal elements of $\boldsymbol{\Gamma}_j$ are lag j cross-covariances.

Assume $\boldsymbol{\delta} = \mathbf{0}$, and recall that, for nonsingular matrix $\mathbf{A}$,

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} adj[\mathbf{A}]$$

In VAR $\mathbf{\Phi}(L)\mathbf{y}_t = \boldsymbol{\varepsilon}_t \Rightarrow \mathbf{y}_t = \mathbf{\Phi}^{-1}(L)\boldsymbol{\varepsilon}_t$.

Premultiply by $|\mathbf{\Phi}(L)|$:
$\Rightarrow$

$$|\mathbf{\Phi}(L)|\,\mathbf{y}_t \;=\; adj[\mathbf{\Phi}(L)]\boldsymbol{\varepsilon}_t$$

where $adj[\mathbf{\Phi}(L)]$ is adjoint matrix of $\mathbf{\Phi}(L)$.

A determinant is scalar, so same AR $|\mathbf{\Phi}(L)|$ applies to each variable.
$\Rightarrow$

Each variable in $\mathbf{y}_t$ is stationary if all roots of

$$|\mathbf{\Phi}(z)| \;=\; \left|\mathbf{I}_k - \mathbf{\Phi}_1 z - \mathbf{\Phi}_2 z^2 - ... - \mathbf{\Phi}_P z^P\right| = 0$$

lie outside the unit circle.

This is also the stationarity condition for the VAR.

Examples: VAR(1) with $k = 2$

1. $\phi_{11} = 1, \phi_{12} = -.6, \phi_{21} = .5, \phi_{22} = -.7$. Then

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & -.6 \\ .5 & -.7 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

and

$$\mathbf{\Phi}(L) = \mathbf{I}_2 - \mathbf{\Phi} L = \begin{bmatrix} 1 - L & .6L \\ -.5L & 1 + .7L \end{bmatrix}$$

$\Rightarrow$

$$\begin{aligned} |\mathbf{\Phi}(z)| &= \begin{vmatrix} 1 - z & .6z \\ -.5z & 1 + .7z \end{vmatrix} \\ &= (1 - z)(1 + .7z) + .3z^2 \\ &= 1 - .3z - .7z^2 + .3z^2 \\ &= 1 - .3z - .4z^2 = (1 - .8z)(1 + .5z) \end{aligned}$$

Both roots are outside the unit circle and the VAR is stationary, despite $\phi_{11} = 1$.

2. $\phi_{11} = .9, \phi_{12} = .1, \phi_{21} = .2, \phi_{22} = .8$,

$$\begin{aligned} |\mathbf{\Phi}(z)| &= \begin{vmatrix} 1 - \phi_{11}z & -\phi_{12}z \\ -\phi_{21}z & 1 - \phi_{22}z \end{vmatrix} = \begin{vmatrix} 1 - .9z & -.1z \\ -.2z & 1 - .8z \end{vmatrix} \\ &= (1 - .9z)(1 - .8z) - .02z^2 \\ &= 1 - 1.7z + .7z^2 = (1 - z)(1 - .7z). \end{aligned}$$

System is nonstationary. Due to factor $(1 - z)$, both $y_{1t}, y_{2t} \sim I(1)$.

2.2 Mean and Autocovariances

For stationary VAR(P), take expectations in

$$\mathbf{y}_t = \boldsymbol{\delta} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + ... + \boldsymbol{\Phi}_P \mathbf{y}_{t-P} + \boldsymbol{\varepsilon}_t$$

$\Rightarrow$

$$\boldsymbol{\mu} = \boldsymbol{\delta} + \boldsymbol{\Phi}_1 \boldsymbol{\mu} + \boldsymbol{\Phi}_2 \boldsymbol{\mu} + ... + \boldsymbol{\Phi}_P \boldsymbol{\mu}$$

& hence

$$\boldsymbol{\mu} = (\mathbf{I}_k - \boldsymbol{\Phi}_1 - \boldsymbol{\Phi}_2 - ... - \boldsymbol{\Phi}_P)^{-1} \boldsymbol{\delta} = \boldsymbol{\Phi}^{-1}(1) \boldsymbol{\delta}.$$

This is a generalisation of the AR(p) mean result.

Stationarity ensures $|\boldsymbol{\Phi}(1)| \neq 0$, since $|\boldsymbol{\Phi}(z)|$ does not have a unit root.

Can also obtain $\boldsymbol{\Gamma}_j$ as function of stationary VAR coefficient matrices $\boldsymbol{\Phi}_1, ..., \boldsymbol{\Phi}_P$.

2.3 Moving Average Representation

Every stationary VAR process has an infinite order vector moving average (VMA) representation:

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{\Phi}^{-1}(L)[\boldsymbol{\delta} + \boldsymbol{\varepsilon}_t] \\
&= \boldsymbol{\mu} + \mathbf{\Phi}^{-1}(L)\boldsymbol{\varepsilon}_t \\
&= \boldsymbol{\mu} + \sum_{\ell=0}^{\infty} \mathbf{\Psi}_\ell \boldsymbol{\varepsilon}_{t-\ell}.
\end{aligned}$$

as $\boldsymbol{\mu} = \mathbf{\Phi}^{-1}(1)\boldsymbol{\delta}$; note that $\mathbf{\Psi}_0 = \mathbf{I}_k$.

For VAR(1), $\mathbf{\Phi}^{-1}(L) = (\mathbf{I}_k + \mathbf{\Phi}L + \mathbf{\Phi}^2 L^2 + ...)$, so that

$$\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\mu} + (\mathbf{I}_k + \mathbf{\Phi}L + \mathbf{\Phi}^2 L^2 + ...)\boldsymbol{\varepsilon}_t \\
&= \boldsymbol{\mu} + \sum_{i=0}^{\infty} \mathbf{\Phi}^i \boldsymbol{\varepsilon}_{t-i}.
\end{aligned}$$

and $\mathbf{\Psi}_\ell = \mathbf{\Phi}^\ell$.

For general VAR(P), $\mathbf{\Psi}_\ell$ is a function of the VAR coefficient matrices $\mathbf{\Phi}_1, ..., \mathbf{\Phi}_P$.

3 Interpreting VARs

There are many coefficients in a VAR. Say $P = 4$ with $k = 3$; each VAR equation involves $3 \times 4 = 12$ coefficients plus intercept
$\Rightarrow 3 \times 13 = 39$ coefficients in the system.
With, $k = 6$ and $P = 4$, $6 \times (6 \times 4 + 1) = 150$ coefficients!

Dynamic interrelationships in the VAR can be complex.
Say VAR(1) in g_t, Δp_t and r_t:

$$\begin{aligned}\Delta p_t &= \delta_1 + \phi_{11}\Delta p_{t-1} + \phi_{12}g_{t-1} + \phi_{13}r_{t-1} + \varepsilon_{pt}\\ g_t &= \delta_2 + \phi_{21}\Delta p_{t-1} + \phi_{22}g_{t-1} + \phi_{23}r_{t-1} + \varepsilon_{gt}\\ r_t &= \delta_3 + \phi_{31}\Delta p_{t-1} + \phi_{32}g_{t-1} + \phi_{33}r_{t-1} + \varepsilon_{rt}\end{aligned}$$

What is the effect of r on future Δp?

- Rate of interest directly affects future inflation through $\phi_{23}r_{t-1}$.
- Also indirect effect through g_t (r_{t-1} influences g_t and g_{t-1} appears in Δp_t equation).
- Both g_{t-1} & Δp_{t-1} affect r_t, leading to a feedback effect through r_t.

Impulse response functions used for VAR interpretation.

3.1 Impulse Response Functions

Impulse response function is the dynamic effect of a disturbance ε_{jt}.

For monetary policy VAR, an interest rate disturbance ε_{rt} affects each of $g, \Delta p$ & r.
For k variables there are k^2 impulse response functions, one for each of k disturbances on k variables.

VMA representation implies

$$\frac{\partial \mathbf{y}_{t+\ell}}{\partial \boldsymbol{\varepsilon}_t'} = \boldsymbol{\Psi}_\ell$$

so that

$$\frac{\partial y_{i,t+\ell}}{\partial \varepsilon_{jt}} = \psi_{ij}^{\ell}$$

where ψ_{ij}^{ℓ} is $(i,j)th$ element of $\boldsymbol{\Psi}_\ell$.

$\Rightarrow$ impulse response function for y_j on y_i is

$$\psi_{ij}^{\ell}, \qquad \ell = 0, 1, 2, ...$$

As $\boldsymbol{\Psi}_0 = \mathbf{I}_k$

$$\frac{\partial y_{it}}{\partial \varepsilon_{jt}} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} .$$

This is response to $\varepsilon_{jt} = 1.$

Can set $\varepsilon_{jt} = a_j$ and show impulse responses as $a_j \psi_{ij}^{\ell}$ ($\ell = 0, 1, 2, ...$).

Shock a_j often taken as estimated standard deviation of ε_j.

3.2 Orthogonalised Impulse Response Functions

Impulse response functions above do not take account of disturbance covariance matrix $\mathbf{\Sigma}$.

Effectively take $\varepsilon_{jt} = a_j, \varepsilon_{it} = 0, i \neq j$ despite $E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij} \neq 0$.

Therefore, transform ε_{jt} to u_{jt} that are mutually orthogonal;
that is, $E(u_{it}u_{jt}) = 0, i \neq j$.

Define $\mathbf{u}_t$ by

$$\mathbf{u}_t = \mathbf{C}\boldsymbol{\varepsilon}_t$$

where $\mathbf{C}$ is lower triangular and nonsingular ($k \times k$), with diagonal elements of unity and

$$E(\mathbf{u}_t\mathbf{u}_t') = \mathbf{D}, \mathbf{D} \text{ diagonal.}$$

This is the *Cholesky decomposition*.

Due to orthogonality, no disturbance u_{jt} provides information about another $u_{is} (i \neq j)$;
therefore u_{jt} affects no $y_{it} (i \neq j)$.

As $\boldsymbol{\varepsilon}_t = \mathbf{C}^{-1}\mathbf{u}_t$, VMA is

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{\ell=0}^{\infty} \boldsymbol{\Psi}_\ell \mathbf{C}^{-1}\mathbf{u}_{t-\ell}$$

$\Rightarrow$

$$\frac{\partial \mathbf{y}_{t+\ell}}{\partial \mathbf{u}_t'} = \boldsymbol{\Psi}_\ell \mathbf{C}^{-1}.$$

The *orthogonalised impulse response function* is for y_j on y_i is:

$$\frac{\partial y_{i,t+\ell}}{\partial u_{jt}} \qquad \ell = 0, 1, 2, ...$$

which is the (i, j) element of $\boldsymbol{\Psi}_\ell \mathbf{C}^{-1}$.

In practice, computed with $u_{jt} = a_j$
& all other $u_{it} = 0$.

a_j typically equal to estimated standard deviation of u_{jt}.

It is important to understand that lower triangular $\mathbf{C}$ implies a causal order:

- $\varepsilon_{1t} \longrightarrow \varepsilon_{2t}$;
- $\varepsilon_{1t}, \varepsilon_{2t} \longrightarrow \varepsilon_{3t}$;
- $\varepsilon_{1t}, \varepsilon_{2t}, ..., \varepsilon_{k-1,t} \longrightarrow \varepsilon_{kt}$.

That is, within t, variables ordered first may cause later ones, but not vice versa.

Ordering of variables can have a substantial impact on orthogonalised impulse response functions.

Some econometricians and statisticans are sceptical about them.

VAR itself gives no information about causal ordering and hence it depends on *a priori* economic beliefs.

Example: Consider again the monetary policy VAR, with variables are ordered as:

$$\mathbf{y}_t = \begin{bmatrix} \Delta p_t \\ g_t \\ r_t \end{bmatrix}.$$

Orthogonalised impluse response functions here implicitly assume:

- current inflation may influence current growth;
- current inflation & growth may influence current interest rates;
- current interest rates & growth do not affect inflation in period t;
- interest rates at t do not affect current growth.

Estimated VAR(4) model for UK inflation, GDP growth and interest rates.

Variable order 1: Inflation, growth, interest rates
Variable order 2: Growth, inflation, interest rates

- Same VAR coefficients

```
 Orthogonalised Impulse Responses to one SE
             shock to GROWTH
          Effect on INFLATION
*******************************************
```

Horizon	Variable Order 1	Variable Order 2
0	0.00	-.15842
1	.09203	.01082
2	-.00937	-.07926
3	.09854	.05633
4	.16212	.09323
5	.12177	.06113
6	.09203	.03498
7	.07004	.02097
8	.06767	.01070
9	.06448	.01914
10	.05878	.01655
11	.05008	.01104
12	.04518	.00893

```
******************************************
```

4 Modelling Issues

4.1 Estimation and Inference

Estimation of a VAR(P) process is by OLS, separately for each equation.

Although $E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Sigma}$ (not necessarily diagonal), OLS is optimal.

Comment: Assumes unrestricted VAR (ie, same explanatory variables in each equation and no restrictions between coefficients).

Hypothesis tests for individual coefficients use usual t-ratios.

A test of, say $\boldsymbol{\Phi}_P = \mathbf{0}$, requires a system test, as coefficients in all k equations are involved.

These are based on

$$\widehat{\boldsymbol{\Sigma}}_P = \frac{1}{T-P} \sum_{t=1}^{T} \mathbf{e}_t \mathbf{e}_t'.$$

with residuals $\mathbf{e}_t$ ($t = P+1, ..., T$) from VAR(P).

Note: OLS equation by equation can be shown to minimise $\ln \left| \widehat{\boldsymbol{\Sigma}}_P \right|$.

Usual system test is a *likelihood ratio* test,

$$LR = T\left[\ln\left|\widehat{\mathbf{\Sigma}}(R)\right| - \ln\left|\widehat{\mathbf{\Sigma}}(U)\right|\right]$$

where $\widehat{\mathbf{\Sigma}}(R), \widehat{\mathbf{\Sigma}}(U)$ are the restricted and unrestricted estimated disturbance variance matrices.

When the restrictions are valid, asymptotically $LR \sim \chi^2_{df}$ with $df =$ number of restrictions.

For testing

$$H_0 : \mathbf{\Phi}_P = \mathbf{0}$$

unrestricted model is VAR(P);
restricted model is VAR($P-1$).

Under this H_0, $LR \overset{a}{\sim} \chi^2_{k^2}$. [There are k zero restrictions in each of k equations.]

4.2 VAR Order Specification

As for AR(p), two approaches for specification of P :

1. "Testing down": Start with maximum order P^*. Test
$$H_0 : \mathbf{\Phi}_{P^*} = \mathbf{0}.$$
If H_0 not rejected, move to order $P^* - 1$ and test
$$H_0 : \mathbf{\Phi}_{P^*-1} = \mathbf{0}.$$
Continue until H_0 is rejected.

2. Use a model specification criterion to select between orders $0, 1, ..., P^*$. Most common are
$$AIC(P) \; = \; \ln\left|\widehat{\mathbf{\Sigma}}_P\right| + \frac{2Pk^2}{T-P}$$
and
$$SIC(P) \; = \; \ln\left|\widehat{\mathbf{\Sigma}}_P\right| + \frac{Pk^2\ln(T-p)}{T-P}.$$
Select $\overline{P}$ that minimises criterion.

Note: These are generalisations of criteria for univariate AR(P), where $k = 1$.