Università di Pavia

Vector AutoRegression Model

Eduardo Rossi

Vector autoregressions (VARs) were introduced into empirical economics by C.Sims (1980), who demonstrated that VARs provide a flexible and tractable framework for analyzing economic time series.

Identification issue: since these models don't dichotomize variables into "endogenous" and "exogenous", the exclusion restrictions used to identify traditional simultaneous equations models make little sense.

A Vector Autoregression model (VAR) of order p is written as:

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t$$

$$\mathbf{y}_t : (N \times 1) \quad \mathbf{\Phi}_i : (N \times N) \ \forall i, \quad \boldsymbol{\epsilon}_t : (N \times 1)$$

$$E(\boldsymbol{\epsilon}_t) = 0 \quad E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}'_\tau) = \begin{cases} \mathbf{\Omega} & t = \tau \\ \mathbf{0} & t \neq \tau \end{cases}$$

$\mathbf{\Omega}$ positive definite matrix.

A VAR is a vector generalization of a scalar autoregression.
The VAR is a system in which each variable is regressed on a constant and p of its own lags as well as on p lags of each of the other variables in the VAR.

$$\begin{aligned} [\mathbf{I}_N - \mathbf{\Phi}_1 L - \cdots - \mathbf{\Phi}_p L^p]\, \mathbf{y}_t &= \mathbf{c} + \boldsymbol{\epsilon}_t \\ \mathbf{\Phi}(L)\mathbf{y}_t &= \mathbf{c} + \boldsymbol{\epsilon}_t \end{aligned}$$

with

$$\mathbf{\Phi}(L) = [\mathbf{I}_N - \mathbf{\Phi}_1 L - \cdots - \mathbf{\Phi}_p L^p]$$

$\mathbf{\Phi}(L)$ $(N \times N)$ matrix polynomial in L.

The element (i, j) in $\mathbf{\Phi}(L)$ is a scalar polynomial in L

$$\delta_{ij} - \phi_{ij}^{(1)} L - \phi_{ij}^{(2)} L^2 - \ldots \phi_{ij}^{(p)} L^p$$

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Stationarity: A vector process is said *covariance stationary* if its first and second moments, $E[\mathbf{y}_t]$ and $E[\mathbf{y}_t \mathbf{y}'_{t-j}]$ respectively, are independent of the date t.

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t$$

First equation:

$$y_{1t} = c_1 + \phi_{11}^{(1)} y_{1t-1} + \phi_{12}^{(1)} y_{2t-1} + \ldots + \phi_{1N}^{(1)} y_{Nt-1} + \epsilon_{1t}$$

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{c} + \mathbf{\Phi}_1 \left[\mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-2} + \boldsymbol{\epsilon}_{t-1}\right] + \boldsymbol{\epsilon}_t \\
&= \mathbf{c} + \mathbf{\Phi}_1 \mathbf{c} + \mathbf{\Phi}_1^2 \mathbf{y}_{t-2} + \boldsymbol{\epsilon}_t + \mathbf{\Phi}_1 \boldsymbol{\epsilon}_{t-1} \\
\mathbf{y}_t &= \ldots \\
\mathbf{y}_t &= \mathbf{c} + \mathbf{\Phi}_1 \mathbf{c} + \ldots + \mathbf{\Phi}_1^{k-1} \mathbf{c} + \mathbf{\Phi}_1^k \mathbf{y}_{t-k} + \boldsymbol{\epsilon}_t + \mathbf{\Phi}_1 \boldsymbol{\epsilon}_{t-1} + \ldots + \mathbf{\Phi}_1^{k-1} \boldsymbol{\epsilon}_{t-k+1}
\end{aligned}$$

$$E[\mathbf{y}_t] = \sum_{j=0}^{k-1} \mathbf{\Phi}_1^j \mathbf{c} + \mathbf{\Phi}_1^k E[\mathbf{y}_{t-k}]$$

The value of this sum depends on the behavior of $\mathbf{\Phi}_1^j$ as j increases.

Let $\lambda_1, \lambda_2, \ldots, \lambda_N$ be the eigenvalues of $\mathbf{\Phi}_1$, the solutions to the characteristic equation

$$|\mathbf{\Phi}_1 - \lambda \mathbf{I}_N| = 0$$

then, if the eigenvalues are all distinct

$$\mathbf{\Phi}_1 = \mathbf{QMQ}^{-1}$$

$$\mathbf{\Phi}_1^k = \mathbf{QM}^k\mathbf{Q}^{-1}$$

$$\mathbf{M}^k = diag(\lambda_1^k, \lambda_2^k, \ldots, \lambda_N^k)$$

If $|\lambda_i| < 1, \;\; i = 1, \ldots, N$

$$\mathbf{\Phi}^k \to 0, \;\; k \to \infty$$

If $|\lambda_i| \geq 1, \;\; \forall i$ then one or more elements of $\mathbf{M}^k$ are not vanishing, and may be tending to ∞.

VAR(p):

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t$$

as a VAR(1) (Companion Form):

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t$$

$$\boldsymbol{\xi}_t = \begin{bmatrix} \mathbf{y}_t \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix} \quad (Np \times 1) \quad \boldsymbol{\xi}_{t-1} = \begin{bmatrix} \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p} \end{bmatrix} \quad (Np \times 1)$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \dots & \mathbf{\Phi}_{p-1} & \mathbf{\Phi}_p \\ \mathbf{I}_N & 0 & \dots & 0 & 0 \\ 0 & \mathbf{I}_N & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \mathbf{I}_N & 0 \end{bmatrix} \quad (Np \times Np) \qquad \mathbf{v}_t = \begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (Np \times 1)$$

$$E[\mathbf{v}_t \mathbf{v}_\tau'] = \begin{cases} \mathbf{Q} & t = \tau \\ \mathbf{0} & t \neq \tau \end{cases} \qquad \mathbf{Q} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \quad (Np \times Np)$$

$$\mathbf{y}_t = \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad t = 1, \ldots, T$$

If the process $\mathbf{y}_t$ has finite variance and an autocovariance sequence that converges to zero at an exponential rate, then $\boldsymbol{\xi}_t$ must share these properties. This is ensured by having the Np eigenvalues of $\mathbf{F}$ lie *inside the unit circle.*

The determinant defining the characteristic equation is

$$|\mathbf{F} - \lambda \mathbf{I}_{Np}| = (-1)^{Np} |\lambda^p \mathbf{I}_N - \lambda^{p-1} \mathbf{\Phi}_1 - \ldots - \mathbf{\Phi}_p| = 0$$

The required condition is that the roots of the equation

$$|\lambda^p \mathbf{I}_N - \lambda^{p-1}\mathbf{\Phi}_1 - \ldots - \mathbf{\Phi}_p| = 0$$

a polynomial of order Np must lie inside the unit circle.

Stability condition can also be expressed as the roots of

$$|\mathbf{\Phi}(z)| = 0$$

lie outside the unit circle, where $\mathbf{\Phi}(z)$ is a $(N \times N)$ matrix polynomial in the lag operator of order p.

When $p = 1$, the roots of

$$|\mathbf{I}_N - \mathbf{\Phi}_1 z| = 0$$

outside the unit circle, i.e. $|z| > 1$, implies that the eigenvalues of $\mathbf{\Phi}_1$ be inside the unit circle. Note that the eigenvalues, roots of $|\mathbf{\Phi}_1 - \lambda \mathbf{I}_N| = 0$, are the reciprocal of the roots of $|\mathbf{I}_N - \mathbf{\Phi}_1 z| = 0$.

Three conditions are necessary for stationarity of the VAR(p) model:

- Absence of mean shifts;
- The vectors $\{\epsilon_t\}$ are identically distributed, $\forall t$;
- Stability condition on $\mathbf{F}$.

If the process is covariance stationary we can take the expectations of both sides of

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t$$

$$\begin{aligned}
\boldsymbol{\mu} &= \mathbf{c} + \mathbf{\Phi}_1 \boldsymbol{\mu} + \ldots + \mathbf{\Phi}_p \boldsymbol{\mu} \\
\boldsymbol{\mu} &= [\mathbf{I}_N - \mathbf{\Phi}_1 \ldots - \mathbf{\Phi}_p]^{-1} \mathbf{c} \\
&= \mathbf{\Phi}(1)^{-1} \mathbf{c}
\end{aligned}$$

If the VAR(p) is stationary then it has a VMA(∞) representation:

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t + \boldsymbol{\Psi}_1 \boldsymbol{\epsilon}_{t-1} + \boldsymbol{\Psi}_2 \boldsymbol{\epsilon}_{t-2} + \ldots \equiv \boldsymbol{\mu} + \boldsymbol{\Psi}(L)\boldsymbol{\epsilon}_t$$

$\mathbf{y}_{t-j}$ is a linear function of $\boldsymbol{\epsilon}_{t-j}, \boldsymbol{\epsilon}_{t-j-1}, \ldots$ each of which is uncorrelated with $\boldsymbol{\epsilon}_{t+1}$ for $j = 0, 1, \ldots$.

It follows that

- $\boldsymbol{\epsilon}_{t+1}$ is uncorrelated with $\mathbf{y}_{t-j}$ for any $j \geq 0$.
- Linear forecast of $\mathbf{y}_{t+1}$ based on $\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots$ is given by

$$\widehat{\mathbf{y}}_{t+1|t} = \boldsymbol{\mu} + \boldsymbol{\Phi}_1(\mathbf{y}_t - \boldsymbol{\mu}) + \boldsymbol{\Phi}_2(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \ldots + (\mathbf{y}_{t-p+1} - \boldsymbol{\mu})$$

$\boldsymbol{\epsilon}_{t+1}$ can be interpreted as the fundamental innovation in $\mathbf{y}_{t+1}$, that is the error in forecasting $\mathbf{y}_{t+1}$ on the basis of a linear function of a constant and $\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots$.

A forecast of $\mathbf{y}_{t+s}$ on the basis of $\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots$ will take the form

$$\widehat{\mathbf{y}}_{t+s|t} = \boldsymbol{\mu} + \mathbf{F}_{11}^{(s)}(\mathbf{y}_t - \boldsymbol{\mu}) + \mathbf{F}_{12}^{(s)}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \ldots + \mathbf{F}_{1p}^{(s)}(\mathbf{y}_{t-p+1} - \boldsymbol{\mu})$$

The moving average matrices $\boldsymbol{\Psi}_j$ can be calculated as:

$$\boldsymbol{\Psi}(L) = [\boldsymbol{\Phi}(L)]^{-1}$$

$$\boldsymbol{\Psi}(L)\boldsymbol{\Phi}(L) = \mathbf{I}_N$$

$$[\mathbf{I}_N + \mathbf{\Psi}_1 L + \mathbf{\Psi}_2 L^2 + \ldots][\mathrm{I}_N - \mathbf{\Phi}_1 L - \mathbf{\Phi}_2 L^2 + \ldots - \mathbf{\Phi}_p L^p] = \mathbf{I}_N$$

Setting the coefficient on L^1 equal to the zero matrix,

$$\mathbf{\Psi}_1 - \mathbf{\Phi}_1 = 0$$

on L^2

$$\mathbf{\Psi}_2 = \mathbf{\Psi}_1 \mathbf{\Phi}_1 + \mathbf{\Phi}_2$$

In general for L^s

$$\mathbf{\Psi}_s = \mathbf{\Phi}_1 \mathbf{\Psi}_{s-1} + \ldots + \mathbf{\Phi}_p \mathbf{\Psi}_{s-p} \quad s = 1, 2, \ldots$$

$$\mathbf{\Psi}_0 = I_N, \quad \mathbf{\Psi}_s = 0 \quad s < 0$$

The innovation in the MA(∞) representation is $\boldsymbol{\epsilon}_t$, *the fundamental innovation* for $\mathbf{y}$.
There are alternative MA representation based on VWN processes other than $\boldsymbol{\epsilon}_t$.
Let $\mathbf{H}$ be a nonsingular $(N \times N)$ matrix

$$\mathbf{u}_t \equiv \mathbf{H}\boldsymbol{\epsilon}_t$$

$\mathbf{u}_t \sim VWN$.

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu} + \mathbf{H}^{-1}\mathbf{H}\boldsymbol{\epsilon}_t + \boldsymbol{\Psi}_1\mathbf{H}^{-1}\mathbf{H}\boldsymbol{\epsilon}_{t-1} + \ldots \\ &= \boldsymbol{\mu} + \mathbf{J}_0\mathbf{u}_t + \mathbf{J}_1\mathbf{u}_{t-1} + \mathbf{J}_2\mathbf{u}_{t-2} + \ldots \end{aligned}$$

$$\mathbf{J}_s \equiv \boldsymbol{\Psi}_s\mathbf{H}^{-1}$$

For example $\mathbf{H}$ can be any matrix that diagonalizes $\mathbf{\Omega}$, the var-cov of $\boldsymbol{\epsilon}_t$,

$$\mathbf{H}\mathbf{\Omega}\mathbf{H}' = \mathbf{D}$$

the elements of $\mathbf{u}_t$ are uncorrelated with one another.
It is always possible to write a stationary VAR(p) process as a convergent infinite MA of a VWN whose elements are mutually uncorrelated.

To obtain the MA representation for the fundamental innovations, we must impose $\mathbf{\Psi}_0 = \mathbf{I}_N$ (while $\mathbf{J}_0$ is not the identity matrix).

- For a covariance stationary process, the parameters $\mathbf{c}, \mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p$ could be defined as the coefficients of the projections of $\mathbf{y}_t$ on $1, \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}$.
- $\boldsymbol{\epsilon}_t$ is uncorrelated with $\mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}$ by the definition of $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p$.
- The parameters of a VAR can be estimated consistently with n OLS regressions.
- $\boldsymbol{\epsilon}_t$ defined by this projection is uncorrelated with $\mathbf{y}_{t-p-1}, \mathbf{y}_{t-p-2}, \cdots$
- The assumption of $\mathbf{y}_t \sim VAR(p)$ is basically the assumption that p lags are sufficient to summarize the dynamic correlations between elements of $\mathbf{y}$.

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t + \boldsymbol{\Theta}_1 \boldsymbol{\epsilon}_{t-1} + \ldots + \boldsymbol{\Theta}_q \boldsymbol{\epsilon}_{t-q}$$

$\boldsymbol{\epsilon}_t \sim VWN$, $\boldsymbol{\Theta}_j$ $(N \times N)$ matrix of MA coefficients $j = 1, 2, \ldots, q$.

$$E(\mathbf{y}_t) = \boldsymbol{\mu}$$

$$\begin{aligned}
\boldsymbol{\Gamma}_0 &= E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})'] \\
&= E[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'] + \boldsymbol{\Theta}_1 E[\boldsymbol{\epsilon}_{t-1} \boldsymbol{\epsilon}_{t-1}'] \boldsymbol{\Theta}_1' + \ldots + \boldsymbol{\Theta}_q E[\boldsymbol{\epsilon}_{t-q} \boldsymbol{\epsilon}_{t-q}'] \boldsymbol{\Theta}_q' \\
&= \boldsymbol{\Omega} + \boldsymbol{\Theta}_1 \boldsymbol{\Omega} \boldsymbol{\Theta}_1' + \ldots + \boldsymbol{\Theta}_q \boldsymbol{\Omega} \boldsymbol{\Theta}_q'
\end{aligned}$$

$$\mathbf{\Gamma}_j = E[(\boldsymbol{\epsilon}_t + \mathbf{\Theta}_1\boldsymbol{\epsilon}_{t-1} + \ldots + \mathbf{\Theta}_q\boldsymbol{\epsilon}_{t-q})(\boldsymbol{\epsilon}_{t-j} + \mathbf{\Theta}_1\boldsymbol{\epsilon}_{t-j-1} + \ldots + \mathbf{\Theta}_q\boldsymbol{\epsilon}_{t-j-q})']$$

$$\mathbf{\Gamma}_j = \begin{cases} \Theta_j\mathbf{\Omega} + \mathbf{\Theta}_{j+1}\mathbf{\Omega}\mathbf{\Theta}'_1 + \mathbf{\Theta}_{j+2}\mathbf{\Omega}\mathbf{\Theta}'_2 + \ldots + \mathbf{\Theta}_q\mathbf{\Omega}\mathbf{\Theta}'_{q-1} & j = 1, \ldots, q \\ \mathbf{\Omega}\mathbf{\Theta}'_{-j} + \mathbf{\Theta}_1\mathbf{\Omega}\mathbf{\Theta}'_{-j+1} + \ldots + \mathbf{\Theta}_{q+j}\mathbf{\Omega}\mathbf{\Theta}'_q & j = -1, \ldots, -q \\ 0 & |j| > q \end{cases}$$

$\mathbf{\Theta}_0 = \mathbf{I}_N$. Any VMA($\infty$) is covariance stationary.

Given:

$$\boldsymbol{\Gamma}_j = E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-j} - \boldsymbol{\mu})']$$

$$\boldsymbol{\Gamma}_{-j} = E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t+j} - \boldsymbol{\mu})']$$

then

$$\begin{aligned} \boldsymbol{\Gamma}_j &\neq \boldsymbol{\Gamma}_{-j} \\ \boldsymbol{\Gamma}'_j &= \boldsymbol{\Gamma}_{-j} \end{aligned}$$

$$\begin{aligned} \{\boldsymbol{\Gamma}_j\}_{1,2} &= cov(y_{1t}, y_{2t-j}) \\ \{\boldsymbol{\Gamma}_j\}_{2,1} &= cov(y_{2t}, y_{1t-j}) \\ \{\boldsymbol{\Gamma}_{-j}\}_{1,2} &= cov(y_{1t}, y_{2t+j}) \end{aligned}$$

$$\begin{aligned} \boldsymbol{\Gamma}_j &= E[(\mathbf{y}_{t+j} - \boldsymbol{\mu})(\mathbf{y}_{(t+j)-j} - \boldsymbol{\mu})'] \\ &= E[(\mathbf{y}_{t+j} - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})'] \\ \boldsymbol{\Gamma}'_j &= E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t+j} - \boldsymbol{\mu})'] = \boldsymbol{\Gamma}_{-j} \end{aligned}$$

Companion form:

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t$$

$$\boldsymbol{\Sigma} = E[\boldsymbol{\xi}_t \boldsymbol{\xi}_t']$$

$$\begin{aligned}
\boldsymbol{\Sigma} &= E\left\{ \begin{bmatrix} \mathbf{y}_t - \boldsymbol{\mu} \\ \mathbf{y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{y}_{t-p+1} - \boldsymbol{\mu} \end{bmatrix} [(\mathbf{y}_t - \boldsymbol{\mu})' \ldots (\mathbf{y}_{t-p+1} - \boldsymbol{\mu})'] \right\} \\
&= \begin{bmatrix} \boldsymbol{\Gamma}_0 & \boldsymbol{\Gamma}_1 & \ldots & \boldsymbol{\Gamma}_{p-1} \\ \boldsymbol{\Gamma}_1' & \boldsymbol{\Gamma}_0 & \ldots & \boldsymbol{\Gamma}_{p-2} \\ \vdots & & & \vdots \\ \boldsymbol{\Gamma}_{p-1}' & \boldsymbol{\Gamma}_{p-2}' & \ldots & \boldsymbol{\Gamma}_0 \end{bmatrix} \quad (Np \times Np)
\end{aligned}$$

$$\begin{aligned} E[\boldsymbol{\xi}_t \boldsymbol{\xi}_t'] &= E[(\mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t)(\mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t)'] \\ &= \mathbf{F}E(\boldsymbol{\xi}_{t-1}\boldsymbol{\xi}_{t-1}')\mathbf{F}' + E(\mathbf{v}_t\mathbf{v}_t') \end{aligned}$$

where $\mathbf{F}E(\boldsymbol{\xi}_{t-1}v_t') = 0$.

$$\boldsymbol{\Sigma} = \mathbf{F}\boldsymbol{\Sigma}\mathbf{F}' + \mathbf{Q}$$

$$vec(\boldsymbol{\Sigma}) = vec(\mathbf{F}\boldsymbol{\Sigma}\mathbf{F}') + vec(\mathbf{Q})$$

$$vec(\boldsymbol{\Sigma}) = vec(\mathbf{F}\boldsymbol{\Sigma}\mathbf{F}') + vec(\mathbf{Q})$$

$$\begin{aligned} vec(\boldsymbol{\Sigma}) &= (\mathbf{F} \otimes \mathbf{F})vec(\boldsymbol{\Sigma}) + vec(\mathbf{Q}) \\ &= [\mathbf{I}_{(Np)^2} - (\mathbf{F} \otimes \mathbf{F})]^{-1} vec(\mathbf{Q}) \end{aligned}$$

The eigenvalues of $(\mathbf{F} \otimes \mathbf{F})$ are all of the form $\lambda_j \lambda_j$ where λ_i and λ_j are eigenvalues of $\mathbf{F}$. Since $|\lambda_i| < 1, \ \forall i$, it follows that all eigenvalues of $(\mathbf{F} \otimes \mathbf{F})$ are inside the unit circle $|\mathbf{I}_{(Np)^2} - (\mathbf{F} \otimes \mathbf{F})| \neq 0$.

The j-th autocovariance of $\boldsymbol{\xi}_t$ is

$$E[\boldsymbol{\xi}_t \boldsymbol{\xi}'_{t-j}] = \mathbf{F} E[\boldsymbol{\xi}_{t-1} \boldsymbol{\xi}'_{t-j}] + E[\mathbf{v}_t \boldsymbol{\xi}'_{t-j}]$$

$$\boldsymbol{\Sigma}_j = \mathbf{F} \boldsymbol{\Sigma}_{j-1}$$

$$\boldsymbol{\Sigma}_j = \mathbf{F}^j \boldsymbol{\Sigma} \quad j = 1, 2, \ldots$$

The j-th autocovariance of $\mathbf{y}_t$ is given by the first rows and n columns of $\boldsymbol{\Sigma}_j$

$$\boldsymbol{\Gamma}_j = \boldsymbol{\Phi}_1 \boldsymbol{\Gamma}_{j-1} + \ldots + \boldsymbol{\Phi}_p \boldsymbol{\Gamma}_{j-p}$$

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \tag{1}$$

$$\boldsymbol{\epsilon}_t \sim i.i.d.N(\mathbf{0}, \mathbf{\Omega})$$

$(T + p)$ observations. Conditioning on the first p observations we estimate using the last T observations.
Conditional likelihood:

$$f_{Y_T, Y_{T-1}, \ldots, Y_1 | Y_0, \ldots, Y_{1-p}}(\mathbf{y}_T, \ldots, \mathbf{y}_1 | \mathbf{y}_0, \ldots, \mathbf{y}_{1-p}; \theta)$$

$$\boldsymbol{\theta} = (\mathbf{c}', vec(\boldsymbol{\Phi}_1)', vec(\boldsymbol{\Phi}_2)', \ldots, vec(\boldsymbol{\Phi}_p)', vech(\boldsymbol{\Omega})')'$$

$$\mathbf{y}_t | \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{-p+1} \sim N(\mathbf{c} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \ldots + \boldsymbol{\Phi}_p \mathbf{y}_{t-p}, \boldsymbol{\Omega})$$

$$\mathbf{x}_t = \begin{bmatrix} 1 \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p} \end{bmatrix} \quad (Np+1) \times 1$$

$$\boldsymbol{\Pi}' \equiv \begin{bmatrix} \mathbf{c} & \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \ldots & \boldsymbol{\Phi}_p \end{bmatrix} \quad (N \times (Np+1))$$

$$E(\mathbf{y}_t | \mathbf{y}_{t-1}, \ldots, \mathbf{y}_{-p+1}) = \boldsymbol{\Pi}' \mathbf{x}_t$$

$$f_{Y_t|Y_{t-1},\ldots,Y_{1-p}}(\mathbf{y}_t|\mathbf{y}_{t-1},\ldots,\mathbf{y}_{1-p};\boldsymbol{\theta}) =$$
$$(2\pi)^{-\frac{n}{2}}|\boldsymbol{\Omega}^{-1}|^{\frac{1}{2}}\exp\left(-\frac{1}{2}(\mathbf{y}_t-\boldsymbol{\Pi}'\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}(\mathbf{y}_t-\boldsymbol{\Pi}'\mathbf{x}_t)\right)$$

The joint density, conditional on the $\mathbf{y}_0,\ldots,\mathbf{y}_{1-p}$

$$\begin{aligned} f_{Y_t,\ldots,Y_1|Y_0,\ldots,Y_{1-p}}(\mathbf{y}_t,\ldots,\mathbf{y}_1|\mathbf{y}_0,\ldots,\mathbf{y}_{1-p};\boldsymbol{\theta}) &= f_{Y_t|Y_{t-1},\ldots,Y_{1-p}}(\theta)\times \\ & f_{Y_{t-1},\ldots,Y_1|Y_0,\ldots,Y_{1-p}}(\theta) \end{aligned}$$

Recursively, the likelihood function for the full sample, conditioning on $\mathbf{y}_0, \ldots, \mathbf{y}_{1-p}$, is the product of the single conditional densities:

$$f_{Y_T,\ldots,Y_1|Y_0,\ldots,Y_{1-p}} = \prod_{t=1}^{T} f_{Y_t|Y_{t-1},\ldots,Y_{1-p}}(\mathbf{y}_t|\mathbf{y}_{t-1},\ldots,\mathbf{y}_{1-p};\boldsymbol{\theta})$$

The log likelihood function:

$$\mathcal{L}(\theta) = -\frac{TN}{2}\log(2\pi) + \frac{T}{2}\log|\boldsymbol{\Omega}^{-1}| - \frac{1}{2}\sum_{t=1}^{T}\left[(\mathbf{y}_t - \boldsymbol{\Pi}'\mathbf{x}_t)'\boldsymbol{\Omega}^{-1}(\mathbf{y}_t - \boldsymbol{\Pi}'\mathbf{x}_t)\right]$$

The ML estimate of $\mathbf{\Pi}$:

$$\widehat{\mathbf{\Pi}}' = \left[\sum_{t=1}^{T} y_t \mathbf{x}_t'\right] \left[\sum_{t}^{T} \mathbf{x}_t \mathbf{x}_{t=1}'\right]^{-1}$$

The j-th row of $\widehat{\mathbf{\Pi}}'$ is:

$$\mathbf{u}_j' \widehat{\mathbf{\Pi}}' = \mathbf{u}_j' \left[\sum_{t=1}^{T} y_t \mathbf{x}_t'\right] \left[\sum_{t}^{T} \mathbf{x}_t \mathbf{x}_t'\right]^{-1}$$

This is the estimated coefficient vector from an OLS regression of y_{jt} on $\mathbf{x}_t$. ML estimates are found by an OLS regression of y_{jt} on a constant and p lags of all the variables in the system.
The ML estimate of $\mathbf{\Omega}$ is given by:

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

Asymptotic distribution of $\widehat{\mathbf{\Pi}}$

The ML estimates $\widehat{\mathbf{\Pi}}$ and $\widehat{\mathbf{\Omega}}$ will give consistent estimates of the population parameters even if the true innovations are non-gaussian:

$$\begin{aligned} \Phi(L)\mathbf{y}_t &= \boldsymbol{\epsilon}_t \\ \boldsymbol{\epsilon}_t &\sim i.i.d.(\mathbf{0}, \mathbf{\Omega}) \\ E[\epsilon_{it}\epsilon_{jt}\epsilon_{lt}\epsilon_{mt}] &< \infty \quad \forall i, j, l, m \end{aligned}$$

the roots of

$$|\mathbf{I}_N - \mathbf{\Phi}_1 z - \ldots - \mathbf{\Phi}_p z^p| = 0$$

Let $K \equiv N \cdot p + 1$, $\mathbf{x}_t'$ $(1 \times K)$:

$$\mathbf{x}_t' = [1\,, \mathbf{y}_t'\,, \mathbf{y}_{t-1}'\,, \ldots\,, \mathbf{y}_{t-p}']$$

$$\widehat{\boldsymbol{\pi}}_T = vec(\widehat{\mathbf{\Pi}}_T)$$

$$\widehat{\boldsymbol{\pi}}_T = \begin{bmatrix} \widehat{\boldsymbol{\pi}}_{1,T} \\ \vdots \\ \widehat{\boldsymbol{\pi}}_{N,T} \end{bmatrix}$$

$$\widehat{\boldsymbol{\pi}}_{i,t} = \left(\sum_t \mathbf{x}_t \mathbf{x}_t' \right)^{-1} \left(\sum_t \mathbf{x}_t y_{it} \right)$$

$$\widehat{\boldsymbol{\Omega}}_T = \frac{1}{T} \sum_t \widehat{\boldsymbol{\epsilon}}_t \widehat{\boldsymbol{\epsilon}}_t'$$

$$\widehat{\epsilon}_{it} = y_{it} - \mathbf{x}_t' \widehat{\boldsymbol{\pi}}_{i,T}$$

Then,

$$\frac{1}{T}\sum_t \mathbf{x}_t\mathbf{x}_t' \xrightarrow{p} \mathbf{Q} = E\left(\mathbf{x}_t\mathbf{x}_t'\right)$$

$$\widehat{\boldsymbol{\pi}}_T \xrightarrow{p} \pi$$

$$\widehat{\boldsymbol{\Omega}}_T \xrightarrow{p} \boldsymbol{\Omega}$$

$$\sqrt{T}(\widehat{\boldsymbol{\pi}}_T - \boldsymbol{\pi}) \xrightarrow{d} N(0, (\boldsymbol{\Omega} \otimes \mathbf{Q}^{-1}))$$

$$\sqrt{T}(\widehat{\boldsymbol{\pi}}_{i,T} - \boldsymbol{\pi}) \xrightarrow{d} N(0, (\sigma_i^2 \mathbf{Q}^{-1}))$$

$$\sigma_i^2 = E(\epsilon_{it}^2)$$

$$\widehat{\sigma}_i^2 = \frac{1}{T}\sum_t \widehat{\epsilon}_{it}^2 \xrightarrow{p} \sigma_i^2$$

$$\widehat{\boldsymbol{\pi}}_i \approx N\left(\boldsymbol{\pi}_i, \widehat{\sigma}_i^2 \left(\sum_t \mathbf{x}_t \mathbf{x}_t'\right)^{-1}\right)$$

OLS t and F statistics applied to the coefficients of any single equation in the VAR are asymptotically valid. A more general hypothesis

$$\mathbf{R}\boldsymbol{\pi} = \mathbf{d}$$

can be tested using a generalization of the Wald form of the OLS χ^2 test

$$\sqrt{T}(\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d}) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{R}\left(\boldsymbol{\Omega} \otimes \mathbf{Q}^{-1}\right)\mathbf{R}'\right)$$

$$
\begin{aligned}
\xi^2(m) &= T(\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d})' \left[\mathbf{R}\left(\widehat{\boldsymbol{\Omega}}_T \otimes \widehat{\mathbf{Q}}_T^{-1}\right)\mathbf{R}'\right]^{-1} (\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d}) \\
&= (\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d})' \left[\mathbf{R}\left(\widehat{\boldsymbol{\Omega}}_T \otimes (T\widehat{\mathbf{Q}}_T)^{-1}\right)\mathbf{R}'\right]^{-1} (\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d}) \\
&= (\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d})' \left[\mathbf{R}\left(\widehat{\boldsymbol{\Omega}}_T \otimes \left(\sum_t \mathbf{x}_t\mathbf{x}_t'\right)^{-1}\right)\mathbf{R}'\right]^{-1} (\mathbf{R}\widehat{\boldsymbol{\pi}}_T - \mathbf{d})
\end{aligned}
$$