# Vector Autoregressive (VARs) models

Interdependent and Dynamic System Models

Process of using VARs models

1. Set of variables
2. Optimal number of lags

- Evaluation of the model

3. Stability Test – Unit Roots test & Unit Circle
4. Exogeneity Test – Granger Test
5. Implications of the model
   - Impulse Response Function
   - Forecast Error Variance Decomposition



By Tatre Jantarakolica

# Model Specification

Consider simple bivariate system:

$$x_t = \gamma_{10} - a_{12} y_t + \gamma_{11} x_{t-1} + \gamma_{12} y_{t-1} + \varepsilon_{xt}$$

$$y_t = \gamma_{20} - a_{21} x_t + \gamma_{21} x_{t-1} + \gamma_{22} y_{t-1} + \varepsilon_{yt}$$

Assume $x_t$ and $y_t$ are stationary.

$\varepsilon_{xt}$ *and* $\varepsilon_{yt}$ are uncorrelated white-noise disturbances. $\begin{pmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{pmatrix} \sim IID\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma^2_{\varepsilon_x} & 0 \\ 0 & \sigma^2_{\varepsilon_y} \end{pmatrix} \right)$

Both $x_t$ and $y_t$ have contemporaneous effect on each other, thus, endogeneity problem occurs.

# Model Specification

Structural-form can be rewritten in matrix:

$$\begin{pmatrix} 1 & a_{12} \\ a_{21} & 1 \end{pmatrix}\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ \gamma_{20} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}\begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{pmatrix}$$

$$AY_t = \Gamma_0 + \Gamma_1 Y_{t-1} + \boldsymbol{\varepsilon}_t$$

Reduced-form model can be derived by multiply $A^{-1}$ and obtain VAR model in standard form:

$$A^{-1}AY_t = A^{-1}\Gamma_0 + A^{-1}\Gamma_1 Y_{t-1} + A^{-1}\varepsilon_t$$

$$Y_t = \Delta_0 + \Delta_1 Y_{t-1} + \in_t$$

# Model Specification

VAR bivariate system in standard form:

$$Y_t = \Delta_0 + \Delta_1 Y_{t-1} + \in_t$$

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} + \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

$$x_t = \delta_{10} + \delta_{11} x_{t-1} + \delta_{12} y_{t-1} + e_{1t}$$

$$y_t = \delta_{20} + \delta_{21} x_{t-1} + \delta_{22} y_{t-1} + e_{2t}$$

where

$$\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} \sim IID\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, A^{-1}\Sigma A^{-1\prime} = \Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right)$$

# Example

$$Y_t = \Delta_0 + \Delta_1 Y_{t-1} + \in_t$$

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} + \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

Let

$$\Delta_0 = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}, \ \Delta_1 = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix}$$

and Var-Cov $$\Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 16 & -4 \\ -4 & 25 \end{pmatrix}$$

# Model Selection

## 1. Information Criteria – Best fit the data

In order to determine the lag order of VAR(p), Akaike Information Criteria (AIC), Schwarz (SIC), or Hannan & Quinn (HQIC) can be used:

$$AIC(p) = \log\left(\det\left(\hat{\Omega}_p\right)\right) + 2\frac{pm^2}{n}$$

$$SIC(p) = \log\left(\det\left(\hat{\Omega}_p\right)\right) + \log(n)\frac{pm^2}{n}$$

$$HQIC(p) = \log\left(\det\left(\hat{\Omega}_p\right)\right) + 2\log\left(\log(n)\right)\frac{pm^2}{n}$$

Choose model with lowest value of AIC, SIC, or HQIC.

# Model Selection

## 2. Final Prediction Error – Accurate Forecast

To choose the best model in term of the most accurate forecast, FPE can be employed.

Choose model with lowest value of FPE.

## 3. LR Test – Significant Coefficients

To choose the best model in term of the significant coefficients, LR test can be employed.

Perform LR tests until lag coefficients turn out insignificant.

# Stability

From
$$Y_t = \Delta_0 + \Delta_1 Y_{t-1} + \in_t$$
$$Y_t = \Delta_0 + \Delta_1 \left( \Delta_0 + \Delta_1 Y_{t-2} + \in_{t-1} \right) + \in_t$$
$$= \left( I + \Delta_1 \right) \Delta_0 + \Delta_1^2 Y_{t-2} + \Delta_1 \in_{t-1} + \in_t$$

For $n$ iterations
$$Y_t = \left( I + \Delta_1 + \cdots + \Delta_1^n \right) \Delta_0 + \sum_{i=0}^{n} \Delta_1^i \in_{t-i} + \Delta_1^{n+1} Y_{t-n-1}$$

For stability property, convergence is required.
$$Y_t = \mu + \sum_{i=0}^{\infty} \Delta_1^i \in_{t-i} \qquad where \quad \mu = [\bar{x} \quad \bar{y}]'$$

All eigen values of matrix $\Delta_1$ should be less than $1$ or lie inside unit circle.

# Stationarity

For Stationary property, means, variances, and covariances are needed to be constant.

$$\bar{x} = \left[\delta_{10}\left(1-\delta_{22}\right)+\delta_{12}\delta_{20}\right]/\Phi$$

$$\bar{y} = \left[\delta_{20}\left(1-\delta_{11}\right)+\delta_{21}\delta_{10}\right]/\Phi$$

$$where \quad \Phi = \left(1-\delta_{11}\right)\left(1-\delta_{22}\right)-\delta_{12}\delta_{21}$$

Variance-Covariance Matrix:

$$E\left(Y_t-\mu\right)^2 = E\left[\sum_{i=0}^{\infty}\Delta_1^i \in_{t-i}\right]^2$$

$$= \left(I+\Delta_1^2+\Delta_1^4+\Delta_1^6+\cdots\right)\Sigma$$

$$= \left(I-\Delta_1^2\right)^{-1}\Sigma$$

# Exogeneity Test

## Granger exogeneity test

$$\begin{pmatrix} \Delta_{11}(L) & \Delta_{12}(L) \\ \Delta_{21}(L) & \Delta_{22}(L) \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

To test whether $X_t$ is exogenous variables, the hypotheses is to test

$$H_0 : \Delta_{21}(L) = 0 \ and \ \Delta_{12}(L) = 0$$

# Impulse Response Function

From $$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} + \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

Stability property:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix}^i \begin{pmatrix} e_{1t-i} \\ e_{2t-i} \end{pmatrix}$$

Derive to be impulse response function

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{pmatrix} \begin{pmatrix} e_{1t-i} \\ e_{2t-i} \end{pmatrix}$$

*where* $\phi_{11}(i),\ \phi_{12}(i),\ \phi_{21}(i),\ \phi_{22}(i)$ *are impulse response functions*

# Impulse Response Function

Example: $\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} + \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$

Let $\Delta_0 = \begin{pmatrix} \delta_{10} \\ \delta_{20} \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}, \ \Delta_1 = \begin{pmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{pmatrix} = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix}$

and Var-Cov $\Omega = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 16 & -4 \\ -4 & 25 \end{pmatrix}$

Set $Y_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\in_1 = \begin{pmatrix} e_{11} \\ e_{21} \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$

Then, $Y_1 = \Delta_1 Y_0 + \in_1 = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$

# Impulse Response Function

Example:

Then,

$$Y_2 = \Delta_1 Y_1 + \in_2 = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.6 \\ 0.8 \end{pmatrix}$$

and

$$Y_3 = \Delta_1 Y_2 + \in_3 = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix} \begin{pmatrix} 1.6 \\ 0.8 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.72 \\ 0.72 \end{pmatrix}$$

Impulse of $x$ on

| Period | $x$ | $y$ |
|---|---|---|
| 1 | 4.00 | 0.00 |
| 2 | 1.60 | 0.80 |
| 3 | 0.72 | 0.72 |
| 4 | 0.36 | 0.50 |
| 5 | 0.19 | 0.32 |

Impulse of $y$ on

| Period | $x$ | $y$ |
|---|---|---|
| 1 | 0.00 | 5.00 |
| 2 | 0.50 | 2.50 |
| 3 | 0.45 | 1.35 |
| 4 | 0.32 | 0.77 |

# Orthogonal Innovation

In general, the innovations in VAR are not contemporaneously independent of one another.

Thus, set of orthogonal innovations is used to analyze impulse response function analysis.

$$\mathbf{u}_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$ where: $$u_{1t} = b_{11} e_{1t} \qquad b_{11} = \frac{1}{s_1}$$ and $$u_{2t}^* = e_{2t} - b_{21} e_{1t} \qquad u_{2t} = \frac{u_{2t}^*}{s_{2.1}}$$

$s_1$ is standard deviation of $e_{1t}$ and $s_{2.1}$ is standard error of regression $u_{2t}^*$

# Orthogonal Innovation

5

Transformation of innovation matrix to be orthogonal innovations can be stated as:

$$\mathbf{u}_t = P \in_t \quad \text{or} \quad \in_t = P^{-1}\mathbf{u}_t$$

where: $P = \begin{pmatrix} \frac{1}{s_1} & 0 \\ \frac{-b_{21}}{s_{2.1}} & \frac{1}{s_{2.1}} \end{pmatrix}$ then, $P^{-1} = \begin{pmatrix} s_1 & 0 \\ b_{21}s_1 & s_{2.1} \end{pmatrix}$

Sample covariance matrix for $\mathbf{u}_t$ is:

$$E\left(\mathbf{u}_t\mathbf{u}_t'\right) = P\left(E\left(\in_t\in_t'\right)\right)P' = P\hat{\Omega}P' = I$$

Then, Choleski Factorization of matrix $\hat{\Omega}$ is:

$$\hat{\Omega} = P^{-1}\left(P^{-1}\right)'$$

# Choleski Factorization

In analyzing VARs, identification of VARs system using Choleski factorization can be applied by setting order for contemporaneous effects of each variables.

Theoretical framework should be used for setting the order.

# IRF – Orthogonal Innovation

From previous example

$$s_1 = \sqrt{16} = 4 \qquad b_{21} = -4/16 = -0.25$$

$$s_{2.1} = \sqrt{s_2^2\left(1 - r_{12}^2\right)} = \sqrt{25\left[1 - \left((-4)^2/(16)(25)\right)\right]} = 4.8990$$

Thus, $P^{-1} = \begin{pmatrix} 4 & 0 \\ -1 & 4.8990 \end{pmatrix}$

Let $\mathbf{u}_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Then, $\in_1 = P^{-1}\mathbf{u}_1 = \begin{pmatrix} 4 & 0 \\ -1 & 4.8990 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$

# IRF – Orthogonal Innovation

From previous example

Impulse Responses from $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

| Period | $x$ | $y$ |
|---|---|---|
| 1 | 4.00 | -1.00 |
| 2 | 1.50 | 0.30 |
| 3 | 0.63 | 0.45 |
| 4 | 0.30 | 0.35 |
| 5 | 0.15 | 0.23 |
| 6 | 0.09 | 0.14 |
| 7 | 0.05 | 0.09 |
| 8 | 0.03 | 0.06 |

# Forecast Error Variance Decomposition

- Forecast error variance decomposition reveals the proportion of the movements in a sequence due to its own shocks versus shocks to the other variable.
- However, since variance decomposition also contains identification problem, the Choleski factorization is also applied by setting up the order of variables.

# Forecast Error Variance Decomposition

## Forecast Error

From VAR(1): $Y_t = \Delta_1 Y_{t-1} + \in_t$

1-period forecast $\hat{Y}_{t+1} = E\left(Y_{t+1} \middle| Y_t, \ldots, Y_1\right) = \Delta_1 Y_t$

2-period forecast $Y_{t+2} = \Delta_1^2 Y_t + \Delta_1 \in_{t+1} + \in_{t+2}$

Then $\hat{Y}_{t+2} = \Delta_1^2 Y_t$

In general: $Y_{t+s} = \Delta_1^s Y_t + \Delta_1^{s-1} \in_{t+1} + \cdots + \Delta_1 \in_{t+s-1} + \in_{t+s}$

Then $\hat{Y}_{t+s} = \Delta_1^s Y_t$

Forecast Error:

$$Y_{t+s} - \hat{Y}_{t+s} = \in_{t+s} + \Delta_1 \in_{t+s-1} + \cdots + \Delta_1^{s-1} \in_{t+1}$$

# 5 Forecast Error Variance Decomposition

## Forecast Error

Var-Cov matrix for forecast errors $s$ periods ahead.

$$\sum(s) = \Omega + \Delta_1 \Omega \Delta_1' + \Delta_1^2 \Omega (\Delta_1')^2 + \cdots + \Delta_1^{s-1} \Omega (\Delta_1')^{s-1}$$

From Choleski factorization $\hat{\Omega} = P^{-1}(P^{-1})'$

Then,

$$\sum(s) = P^{-1}(P^{-1})' + (\Delta_1 P^{-1})(\Delta_1 P^{-1})' + \cdots + (\Delta_1^{s-1} P^{-1})(\Delta_1^{s-1} P^{-1})'$$

# Forecast Error Variance Decomposition

## Variance Decomposition

1-period ahead forecast error variance:

$$\mathrm{var}(\in_t) = \Omega$$

Then, $\Omega = P^{-1}\,\mathrm{var}(\mathbf{u}_t)\left(P^{-1}\right)'$

$$= \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}$$

$$= \begin{pmatrix} c_{11}v_1 & c_{12}v_2 \\ c_{21}v_1 & c_{22}v_2 \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}$$

$$= \begin{pmatrix} c_{11}^2 v_1 + c_{12}^2 v_2 & c_{11}c_{21}v_1 + c_{12}c_{22}v_2 \\ c_{11}c_{21}v_1 + c_{12}c_{22}v_2 & c_{21}^2 v_1 + c_{22}^2 v_2 \end{pmatrix}$$

# Forecast Error Variance Decomposition

## Variance Decomposition

According to Orthogonal innovations,

$\text{var}(u_1) = v_1 = 1 \,\&\, \text{var}(u_2) = v_2 = 1$ and $c_{12} = 0$

Then, $$\Omega = \begin{pmatrix} c_{11}^2 v_1 + c_{12}^2 v_2 & c_{11}c_{21}v_1 + c_{12}c_{22}v_2 \\ c_{11}c_{21}v_1 + c_{12}c_{22}v_2 & c_{21}^2 v_1 + c_{22}^2 v_2 \end{pmatrix}$$

$$= \begin{pmatrix} c_{11}^2 v_1 & c_{11}c_{21}v_1 \\ c_{11}c_{21}v_1 & c_{21}^2 v_1 + c_{22}^2 v_2 \end{pmatrix} = \begin{pmatrix} c_{11}^2 & c_{11}c_{21} \\ c_{11}c_{21} & c_{21}^2 + c_{22}^2 \end{pmatrix}$$

Variances of 1-period forecast error are:

$\text{var}(\hat{x}_1) = c_{11}^2$ and $\text{var}(\hat{y}_1) = c_{21}^2 + c_{22}^2$

# Forecast Error Variance Decomposition

## Variance Decomposition

Variance of $\hat{x}_1$ can be decomposed as $1$ component $c_{11}^2$ caused by innovation from $x$.

Variance of $\hat{y}_1$ can be decomposed as 2 components $c_{21}^2$ and $c_{22}^2$, where,

$\dfrac{c_{21}^2}{c_{21}^2 + c_{22}^2}$ represents % of variance of $\hat{y}_1$ caused by orthogonal innovation from $x$.

$\dfrac{c_{22}^2}{c_{21}^2 + c_{22}^2}$ represents % of variance of $\hat{y}_1$ caused by orthogonal innovation from $y$.

# FEVD

From previous example: $\Delta_1 = \begin{pmatrix} 0.4 & 0.1 \\ 0.2 & 0.5 \end{pmatrix}$

$$P^{-1} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ -1 & 4.8990 \end{pmatrix} \quad P^{-1}\left(P^{-1}\right)' = \Omega = \begin{pmatrix} 16 & -4 \\ -4 & 25 \end{pmatrix}$$

$$\Delta_1 P^{-1} = \begin{pmatrix} 1.50 & 0.4899 \\ 0.30 & 2.4495 \end{pmatrix} \quad \Delta_1 P^{-1}\left(\Delta_1 P^{-1}\right)' = \begin{pmatrix} 2.49 & 1.65 \\ 1.65 & 6.09 \end{pmatrix}$$

$$\Delta_1^2 P^{-1} = \begin{pmatrix} 0.63 & 0.4409 \\ 0.45 & 1.3227 \end{pmatrix} \quad \Delta_1^2 P^{-1}\left(\Delta_1^2 P^{-1}\right)' = \begin{pmatrix} 0.5913 & 0.8667 \\ 0.8667 & 1.9521 \end{pmatrix}$$

## FEVD of $x$ on $\mathbf{u}_1$

| Period | $x$ | $y$ |
|---|---|---|
| 1 | $\frac{4^2}{16} =$ 100.00% | $\frac{0}{16} =$ 0.00% |
| 2 | $\frac{\left(4^2+(1.50)^2\right)}{(16+2.49)} =$ 98.70% | $\frac{0+\left(0.4899^2\right)}{(16+2.49)} =$ 1.30% |
| 3 | $\frac{\left(4^2+(1.50)^2+(0.63)^2\right)}{(16+2.49+0.5913)} =$ 97.72% | $\frac{\left(0+(0.4899)^2+(0.4409)^2\right)}{(16+2.49+0.5913)} =$ 2.28% |