

# Vector Autoregression

# Overview

- Vector Autoregression (VAR) model is an extension of univariate autoregression model to multivariate time series data
- VAR model is a multi-equation system where all the variables are treated as endogenous (dependent)
- There is one equation for each variable as dependent variable. In its reduced form, the right-hand side of each equation includes lagged values of all dependent variables in the system, no contemporaneous variables

# VAR Model

VAR(p) model:

$$Y_t = a + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t$$

where:

- $Y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$  : an  $(n \times 1)$  vector of time series variables
- $a$ : an  $(n \times 1)$  vector of intercepts
- $A_i$  ( $i=1, 2, \dots, p$ ):  $(n \times n)$  coefficient matrices
- $\varepsilon_t$ : an  $(n \times 1)$  vector of unobservable i.i.d. zero mean error term  
(white noise)

# VAR Model

Example: Bivariate VAR(2) Model

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_{11}^1 & b_{12}^1 \\ b_{21}^1 & b_{22}^1 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} b_{11}^2 & b_{12}^2 \\ b_{21}^2 & b_{22}^2 \end{bmatrix} \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Or:

$$y_{1t} = a_1 + b_{11}^1 y_{1t-1} + b_{12}^1 y_{2t-1} + b_{11}^2 y_{1t-2} + b_{12}^2 y_{2t-2} + \varepsilon_{1t}$$

$$y_{2t} = a_2 + b_{21}^1 y_{1t-1} + b_{22}^1 y_{2t-1} + b_{21}^2 y_{1t-2} + b_{22}^2 y_{2t-2} + \varepsilon_{2t}$$

# The Simplest VAR Model

Example: Bivariate VAR(1) Model with no intercepts

The simplest example is a vector autoregressive process (VAR) with two variables and first order dynamics as in:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{pmatrix}, \quad (1)$$

where  $\epsilon_{y,t}$  and  $\epsilon_{x,t}$  are two white-noise variables (correlated or uncorrelated).

# The Marginal Models

# The Simplest VAR Model

Example: Bivariate VAR(1) Model with no intercepts

We now express  $y_t$  as a function of  $y_{t-1}$  and  $\epsilon_{y,t}$  and  $\epsilon_{x,t}$ .  
Solve the first equation for  $x_{t-1}$ :

$$x_{t-1} = (1/a_{12})y_t - (a_{11}/a_{12})y_{t-1} - (1/a_{12})\epsilon_{y,t} \quad (2)$$

Substitution in the second equation of the VAR gives

$$x_t = \frac{a_{22}}{a_{12}}y_t + \left(a_{21} - a_{22}\frac{a_{11}}{a_{12}}\right)y_{t-1} - \frac{a_{22}}{a_{12}}\epsilon_{y,t} + \epsilon_{x,t}, \quad a_{12} \neq 0 \quad (3)$$

# The Final Equation

Finally: Replace  $t$  by  $t + 1$  in the first equation in (1), and replace  $x_t$  by the left hand side of (3):

$$y_{t+1} = a_{11}y_t + a_{12}\left\{\frac{a_{22}}{a_{12}}y_t + \left(a_{21} - a_{22}\frac{a_{11}}{a_{12}}\right)y_{t-1} - \frac{a_{22}}{a_{12}}\epsilon_{y,t} + \epsilon_{x,t}\right\} + \epsilon_{y,t+1} \quad (4)$$

(4) is called a *final equation*. It shows that  $y_t$  given by the VAR (1) follows the 2 order stochastic difference equation:

$$y_t = \underbrace{(a_{11} + a_{22})}_{\phi_1}y_{t-1} + \underbrace{(a_{12}a_{21} - a_{22}a_{11})}_{\phi_2}y_{t-2} + \epsilon_t. \quad (5)$$

The disturbance in the final equation is:

$$\varepsilon_t = \epsilon_{y,t} - a_{22}\epsilon_{y,t-1} + a_{12}\epsilon_{x,t-1}$$

Write the final equation more compactly

$$\phi(L)y_t = \varepsilon_t, \text{ where} \quad (6)$$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2,$$

$$\phi_1 = (a_{11} + a_{22}) \text{ and } \phi_2 = (a_{12}a_{21} - a_{22}a_{11}).$$

The associated characteristic polynomial:

$$p(\lambda) = \lambda^2 - \phi_1 \lambda - \phi_2. \quad (7)$$

# VAR Stationarity

## Based on the Final Equation

- The **LEFT-HAND SIDE** of the univariate series  $\{Y_t\}$  is a stationary AR(2) if both roots of  $\phi(L)=0$  are outside the unit circle.
- This is equivalent to: both roots of  $P(\lambda)=0$  are inside the unit circle.
- (What process does the error term on the **right-hand side** follow? White noise? MA(1)? Other model?)
- Question: What about the univariate series  $\{X_t\}$ ?

# Check for Stationarity

# The Simplest VAR Model

Example: Bivariate VAR(1) Model with no intercepts

The simplest example is a vector autoregressive process (VAR) with two variables and first order dynamics as in:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{pmatrix}, \quad (1)$$

where  $\epsilon_{y,t}$  and  $\epsilon_{x,t}$  are two white-noise variables (correlated or uncorrelated).

# The Companion Form

Let  $\mathbf{z}_t$  denote the vector time series made up of  $y_t$  and  $x_t$ . Define

$$\mathbf{z}_t = [y_t, x_t]' \text{ and } \boldsymbol{\epsilon}_t = [\epsilon_{y,t}, \epsilon_{x,t}]',$$

(1) can be expressed as

$$\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \quad (8)$$

where  $\mathbf{F}$  is the matrix with VAR coefficients, i.e.,

$$\mathbf{F} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

- ▶ Since equation (8) is a companion form, the associated characteristic polynomial of  $\mathbf{F}$  is

$$|\mathbf{F} - \lambda \mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix}$$

where  $\mathbf{I}$  is the identity matrix.

- ▶  $\lambda$  is an eigenvalue (root) of  $\mathbf{F}$  if

$$|\mathbf{F} - \lambda \mathbf{I}| = 0, \quad (9)$$

If we write out (9) we have that if  $\lambda$  is a root in (9), it is also a root in the characteristic equation

$$\lambda^2 - \phi_1 \lambda - \phi_2 = 0$$

with  $\phi_1 = (a_{11} + a_{22})$  and  $\phi_2 = a_{12}a_{21} - a_{11}a_{22}$ .

# The Companion Form

- $F$  is an example of a companion matrix.
- Since (9) is the same equation as the characteristic equation for the final equation for  $y_t$ , we have that  $[y_t, x_t]$  is a stationary process if and only if the eigenvalues of  $F$  have moduli less than 1 (inside the unit circle).

# The Companion Form

The companion form is well suited for generalizations.  
Let  $\mathbf{y}_t$  be the  $n \times 1$  vector.

$$\mathbf{y}_t = [y_{1t}, y_{2t}, \dots, y_{nt}]'$$

The VAR of order  $p$  is:

$$\mathbf{y}_t = \boldsymbol{\phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\phi}_2 \mathbf{y}_{t-2} + \dots + \boldsymbol{\phi}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad (10)$$

where  $\boldsymbol{\phi}_i$  is a  $n \times n$  matrix with coefficients and  $\boldsymbol{\epsilon}_t$  is a vector with white-noise disturbances that may be correlated

# The Companion Form

Write (10) in companion form:

$$\underbrace{\begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix}}_{\boldsymbol{\zeta}_t} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix}}_{\mathbf{F}_{np \times np}} \underbrace{\begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{y}_{t-2} \\ \vdots \\ \mathbf{y}_{t-p} \end{bmatrix}}_{\boldsymbol{\zeta}_{t-1}} + \underbrace{\begin{bmatrix} \boldsymbol{\epsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\mathbf{v}_t} \quad (11)$$

$$\boldsymbol{\zeta}_t = \mathbf{F}\boldsymbol{\zeta}_{t-1} + \mathbf{v}_t, \quad (12)$$

# The Companion Form

- The VAR(p) is a stationary process if and only if all the eigenvalues of  $F$  have moduli less than 1 (inside the unit circle).
- Recall, the eigenvalues are the solutions of

$$|\mathbf{F} - \lambda \mathbf{I}| = 0$$

# Why do we need VAR?

- Time-series data with autoregressive in nature (serially correlated)
- VAR model is one of the most successful and flexible models for the analysis of multivariate time series
- Especially useful for describing the dynamic behavior of economic and financial time series
- Useful for forecasting

# Applications of VAR

- Analysis of system response to different shocks/impacts
- Model-based forecast. In general VAR encompasses correlation information of the observed data and use this correlation information to forecast future movements or changes of the variable of interest

# Applications of VAR

- In economics, VAR is used to forecast macroeconomic variables, such as GDP, money supply, and unemployment
- In finance, predict spot prices and future prices of securities; foreign exchange rates across markets

# Applications of VAR

- In accounting, predict different accounting variables such as sales, earnings, and accruals
- In marketing, VAR can be used to evaluate the impact of different factors on consumer behavior and forecast its future change.

# Forecasting

# Applications of VAR: Forecasting

- 1-step forecast based on information available at time T:

$$Y_{T+1|T} = a + A_1 Y_T + A_2 Y_{T-1} + \dots + A_p Y_{T-p+1}$$

- h-step forecast:

$$Y_{T+h|T} = a + A_1 Y_{T+h-1|T} + A_2 Y_{T+h-2|T} + \dots + A_p Y_{T+h-p|T}$$

# Implementation

- All data have to have same frequency
  - Data with mixed frequency need to be converted to the same frequency
    - Convert higher-frequency data to the frequency of the lowest-frequency data). For example: if we have daily, weekly and monthly data then we will need to convert everything to monthly frequency
    - Interpolate lower-frequency data into high frequency

# SAS & R for VAR

# Example of VAR usage

- Testable hypothesis: there has to be a dependence of DJIA index on its own lag and on lag of total market capitalization and vice versa
- Use return on DJIA index and return on market capitalization
- Monthly observation

# Dataset

Obs	year	month	ret_dji	ret_totval
1	1961	1	.	.
2	1961	2	0.02619	0.032994
3	1961	3	0.02575	0.033013
4	1961	4	0.01442	0.006127
5	1961	5	0.02067	0.022185
6	1961	6	-0.02172	-0.030350
7	1961	7	0.01913	0.033709
8	1961	8	0.03243	0.022871
9	1961	9	-0.03279	-0.020618

**SAS**

# SAS Implementation

- PROC VARMAX

```
proc varmax data=comb;
```

```
model ret_dji ret_totval / p=1;
```

```
run;
```

- $ret\_dji_t = a_1 + b_{11} ret\_dji_{t-1} + b_{12} ret\_totval_{t-1} + \epsilon_{1t}$
- $ret\_totval_t = a_2 + b_{21} ret\_dji_{t-1} + b_{22} ret\_totval_{t-1} + \epsilon_{2t}$

# SAS Output

<i>Model Parameter Estimates</i>						
<i>Equation</i>	<i>Parameter</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>	<i>Variable</i>
<i>ret_dji</i>	<i>CONST1</i>	0.00542	0.00195	2.78	0.0057	1
	<i>AR1_1_1</i>	-0.29801	0.06973	-4.27	0.0001	<i>ret_dji(t-1)</i>
	<i>AR1_1_2</i>	0.37788	0.07499	5.04	0.0001	<i>ret_totval(t-1)</i>
<i>ret_totval</i>	<i>CONST2</i>	0.00802	0.00185	4.32	0.0001	1
	<i>AR1_2_1</i>	0.06674	0.06627	1.01	0.3144	<i>ret_dji(t-1)</i>
	<i>AR1_2_2</i>	-0.04813	0.07127	-0.68	0.4998	<i>ret_totval(t-1)</i>

$$\text{Ret\_dji}_t = 0.005 - 0.298 \text{ret\_dji}_{t-1} + 0.378 \text{ret\_totval}_{t-1} + e_{1t}$$

$$\text{Ret\_totval}_t = 0.008 + 0.067 \text{ret\_dji}_{t-1} - 0.048 \text{ret\_totval}_{t-1} + e_{2t}$$

# SAS Output

---

## *Covariances of Innovations*

<i>Variable</i>	<i>ret_dji</i>	<i>ret_totval</i>
<i>ret_dji</i>	0.00199	0.00156
<i>ret_totval</i>	0.00156	0.00180

---

---

## *Information Criteria*

<i>AICC</i>	-13.663
<i>HQC</i>	-13.6445
<i>AIC</i>	-13.6631
<i>SBC</i>	-13.6156
<i>FPEC</i>	1.165E-6

---

R

# R for VAR

- VARs are implemented in the **vars** package in R. It contains a function `VARselect` to choose the number of lags  $p$  using four different information criteria: AIC, HQ, SC and FPE. We have met the AIC before, and SC is simply another name for the BIC (SC stands for Schwarz Criterion after Gideon Schwarz who proposed it). HQ is the Hannan-Quinn criterion and FPE is the “Final Prediction Error” criterion. Care should be taken using the AIC as it tends to choose large numbers of lags. Instead, for VAR models, we prefer to use the BIC.

# Concerns

- Assuming all variables are endogenous
- If time-series data are nonstationary (containing stochastic trends), while it is possible to estimate VAR in levels, it is preferable to estimate VAR in first differences
- Uncertainty about number of lags (using LR test, Information criteria: AIC, BIC etc.)

# Concerns

- Data requirements (long time series)
- Imprecise estimated coefficients (overfitting the model).  
Solution – restrict or weight coefficients
- Computationally intensive

# References

- **[courses.ttu.edu/isqs6348-westfall/VARPresentation.ppt](https://courses.ttu.edu/isqs6348-westfall/VARPresentation.ppt)**
- Chapter 1: Vector Autoregressions.  
[https://www2.bc.edu/~iacoviel/teach/0809/EC751\\_files/var.pdf](https://www2.bc.edu/~iacoviel/teach/0809/EC751_files/var.pdf)
- Chapter 6: Multivariate time series models.  
[www.nek.lu.se/.../Ch6%20Multivariate%20time%20series%20models](http://www.nek.lu.se/.../Ch6%20Multivariate%20time%20series%20models)

# References

- <http://www.uio.no/studier/emner/sv/oekonomi/ECON5101/v11/undervisningsmateriale/Lect7v11.pdf>
- Dwyer, Gerald P., Jr. Why Are Vector Autoregressions Useful in Finance? <http://jerrydwyer.com/pdf/lectvar.pdf>
- Vector Autoregressions: Forecasting and Reality. <http://www.frbatlanta.org/filelegacydocs/robtallman.pdf>

# References

- Chapter 11: Vector Autoregressive Models for Multivariate Time Series.  
<http://faculty.washington.edu/ezivot/econ584/notes/varModels.pdf>
- <https://www.otexts.org/fpp/9/2>