

In [1]: `pacman::p_load(tidyverse,ggplot2)`

```
e1=rnorm(n=250,mean=10,sd=5)
e2=rnorm(n=250,mean=13,sd=4)

y1_0=5
y2_0=10

df=data.frame(t=1:250)

df$y1=0
df$y2=0
###univariate ar(1)

n=250
z0=2
z=vector()
for(i in 1:n){
  if(i==1){
    z[i]=z0+e1[i]
  }else{
    z[i]=10+0.8*z[i-1]+e1[i]
  }
}
```

11.2 The Stationary Vector Autoregression Model

Let $\mathbf{Y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ denote an $(n \times 1)$ vector of time series variables. The basic p -lag *vector autoregressive* (VAR(p)) model has the form

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{\Pi}_1 \mathbf{Y}_{t-1} + \mathbf{\Pi}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{\Pi}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T \quad (11.1)$$

where $\mathbf{\Pi}_i$ are $(n \times n)$ coefficient matrices and $\boldsymbol{\varepsilon}_t$ is an $(n \times 1)$ unobservable zero mean white noise vector process (serially uncorrelated or independent) with time invariant covariance matrix $\boldsymbol{\Sigma}$. For example, a bivariate VAR(2) model equation by equation has the form

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^1 & \pi_{12}^1 \\ \pi_{21}^1 & \pi_{22}^1 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} \quad (11.2)$$

$$+ \begin{pmatrix} \pi_{11}^2 & \pi_{12}^2 \\ \pi_{21}^2 & \pi_{22}^2 \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (11.3)$$

or

$$\begin{aligned} y_{1t} &= c_1 + \pi_{11}^1 y_{1t-1} + \pi_{12}^1 y_{2t-1} + \pi_{11}^2 y_{1t-2} + \pi_{12}^2 y_{2t-2} + \varepsilon_{1t} \\ y_{2t} &= c_2 + \pi_{21}^1 y_{1t-1} + \pi_{22}^1 y_{2t-1} + \pi_{21}^2 y_{1t-2} + \pi_{22}^2 y_{2t-2} + \varepsilon_{2t} \end{aligned}$$

where $\text{cov}(\varepsilon_{1t}, \varepsilon_{2s}) = \sigma_{12}$ for $t = s$; 0 otherwise. Notice that each equation has the same regressors – lagged values of y_{1t} and y_{2t} . Hence, the VAR(p) model is just a *seemingly unrelated regression* (SUR) model with lagged variables and deterministic terms as common regressors.

In lag operator notation, the VAR(p) is written as

$$\mathbf{\Pi}(L)\mathbf{Y}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t$$

where $\mathbf{\Pi}(L) = \mathbf{I}_n - \mathbf{\Pi}_1 L - \dots - \mathbf{\Pi}_p L^p$. The VAR(p) is stable if the roots of

$$\det(\mathbf{I}_n - \mathbf{\Pi}_1 z - \dots - \mathbf{\Pi}_p z^p) = 0$$

lie outside the complex unit circle (have modulus greater than one), or, equivalently, if the eigenvalues of the companion matrix

$$\mathbf{F} = \begin{pmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 & \dots & \mathbf{\Pi}_p \\ \mathbf{I}_n & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_n & \mathbf{0} \end{pmatrix}$$

have modulus less than one. Assuming that the process has been initialized in the infinite past, then a stable VAR(p) process is stationary and ergodic with time invariant means, variances, and autocovariances.

If \mathbf{Y}_t in (11.1) is covariance stationary, then the unconditional mean is given by

$$\boldsymbol{\mu} = (\mathbf{I}_n - \mathbf{\Pi}_1 - \dots - \mathbf{\Pi}_p)^{-1} \mathbf{c}$$

The *mean-adjusted* form of the VAR(p) is then

$$\mathbf{Y}_t - \boldsymbol{\mu} = \mathbf{\Pi}_1 (\mathbf{Y}_{t-1} - \boldsymbol{\mu}) + \mathbf{\Pi}_2 (\mathbf{Y}_{t-2} - \boldsymbol{\mu}) + \dots + \mathbf{\Pi}_p (\mathbf{Y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t$$

The basic VAR(p) model may be too restrictive to represent sufficiently the main characteristics of the data. In particular, other deterministic terms such as a linear time trend or seasonal dummy variables may be required to represent the data properly. Additionally, stochastic exogenous variables may be required as well. The general form of the VAR(p) model with deterministic terms and exogenous variables is given by

$$\mathbf{Y}_t = \Pi_1 \mathbf{Y}_{t-1} + \Pi_2 \mathbf{Y}_{t-2} + \cdots + \Pi_p \mathbf{Y}_{t-p} + \Phi \mathbf{D}_t + \mathbf{G} \mathbf{X}_t + \varepsilon_t \quad (11.4)$$

where \mathbf{D}_t represents an $(l \times 1)$ matrix of deterministic components, \mathbf{X}_t represents an $(m \times 1)$ matrix of exogenous variables, and Φ and \mathbf{G} are parameter matrices.

Example 64 *Simulating a stationary VAR(1) model using S-PLUS*

A stationary VAR model may be easily simulated in S-PLUS using the **S+FinMetrics** function **simulate.VAR**. The commands to simulate $T = 250$ observations from a bivariate VAR(1) model

$$\begin{aligned} y_{1t} &= -0.7 + 0.7y_{1t-1} + 0.2y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= 1.3 + 0.2y_{1t-1} + 0.7y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

```
In [3]: #Multivariate

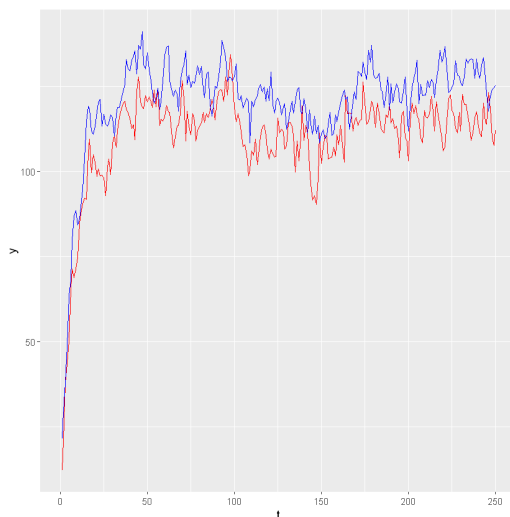
t=250

constants=c(-0.7,1.3)
coefs=c(0.7,0.2,0.2,0.7)
y0=c(5,10)

y=vector()
x=vector()
for (i in 1:t){
  if (i==1){
    y[i]=y0[1]+e1[i]
    x[i]=y0[2]+e2[i]
  }else {
    y[i]=constants[1]+coefs[1]*y[i-1]+coefs[2]*x[i-1]+e1[i]
    x[i]=constants[2]+coefs[3]*y[i-1]+coefs[4]*x[i-1]+e2[i]
  }
}

df=data.frame(t=1:250,x=x,y=y)
```

```
In [4]: df %>% ggplot(aes(x=t))+
  geom_line(aes(y=y),color="red")+
  geom_line(aes(y=x),color="blue")
```



Matrix representation

The matrix representation is row wise as shown

$$\mathbf{Y}(\text{rowwise}) = \text{constants}(\text{rowwise}) + \text{coefficients} * \mathbf{Y}(-1)(\text{rowwise}) + \text{error}(\text{rowwise})$$

```
In [5]: YMAT1=df[,c(2,3)] %>% t %>% as.matrix

constmat= data.frame(
  cx=rep(constants[1],250),cy=rep(constants[2],250)
) %>% t %>% as.matrix

errorformat=data.frame(ex=e1,ey=e2) %>% t %>% as.matrix

coefmat=matrix(data=coefs,nrow=2,ncol=2,byrow=TRUE)

YMAT2=constmat+coefmat%*%YMAT1+errorformat
```

11.2.1 Estimation

Consider the basic VAR(p) model (11.1). Assume that the VAR(p) model is covariance stationary, and there are no restrictions on the parameters of the model. In SUR notation, each equation in the VAR(p) may be written as

$$\mathbf{y}_i = \mathbf{Z}\boldsymbol{\pi}_i + \mathbf{e}_i, \quad i = 1, \dots, n$$

where \mathbf{y}_i is a $(T \times 1)$ vector of observations on the i^{th} equation, \mathbf{Z} is a $(T \times k)$ matrix with t^{th} row given by $\mathbf{Z}'_t = (1, \mathbf{Y}'_{t-1}, \dots, \mathbf{Y}'_{t-p})$, $k = np + 1$, $\boldsymbol{\pi}_i$ is a $(k \times 1)$ vector of parameters and \mathbf{e}_i is a $(T \times 1)$ error with covariance matrix $\sigma_i^2 \mathbf{I}_T$. Since the VAR(p) is in the form of a SUR model

where each equation has the same explanatory variables, each equation may be estimated separately by ordinary least squares without losing efficiency relative to generalized least squares. Let $\hat{\boldsymbol{\Pi}} = [\hat{\boldsymbol{\pi}}_1, \dots, \hat{\boldsymbol{\pi}}_n]$ denote the $(k \times n)$ matrix of least squares coefficients for the n equations.

Let $vec(\hat{\boldsymbol{\Pi}})$ denote the operator that stacks the columns of the $(n \times k)$ matrix $\hat{\boldsymbol{\Pi}}$ into a long $(nk \times 1)$ vector. That is,

$$vec(\hat{\boldsymbol{\Pi}}) = \begin{pmatrix} \hat{\boldsymbol{\pi}}_1 \\ \vdots \\ \hat{\boldsymbol{\pi}}_n \end{pmatrix}$$

Under standard assumptions regarding the behavior of stationary and ergodic VAR models (see Hamilton (1994) or Lütkepohl (1991)) $vec(\hat{\boldsymbol{\Pi}})$ is consistent and asymptotically normally distributed with asymptotic covariance matrix

$$\widehat{avar}(vec(\hat{\boldsymbol{\Pi}})) = \hat{\boldsymbol{\Sigma}} \otimes (\mathbf{Z}'\mathbf{Z})^{-1}$$

where

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T-k} \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'$$

and $\hat{\boldsymbol{\varepsilon}}_t = \mathbf{Y}_t - \hat{\boldsymbol{\Pi}}' \mathbf{Z}_t$ is the multivariate least squares residual from (11.1) at time t .

11.2.2 Inference on Coefficients

The i^{th} element of $vec(\hat{\Pi})$, $\hat{\pi}_i$, is asymptotically normally distributed with standard error given by the square root of i^{th} diagonal element of $\hat{\Sigma} \otimes (\mathbf{Z}'\mathbf{Z})^{-1}$. Hence, asymptotically valid t-tests on individual coefficients may be constructed in the usual way. More general linear hypotheses of the form $\mathbf{R} \cdot vec(\Pi) = \mathbf{r}$ involving coefficients across different equations of the VAR may be tested using the Wald statistic

$$Wald = (\mathbf{R} \cdot vec(\hat{\Pi}) - \mathbf{r})' \left\{ \mathbf{R} \left[\widehat{avar}(vec(\hat{\Pi})) \right] \mathbf{R}' \right\}^{-1} (\mathbf{R} \cdot vec(\hat{\Pi}) - \mathbf{r}) \quad (11.5)$$

Under the null, (11.5) has a limiting $\chi^2(q)$ distribution where $q = rank(\mathbf{R})$ gives the number of linear restrictions.

11.2.3 Lag Length Selection

The lag length for the VAR(p) model may be determined using model selection criteria. The general approach is to fit VAR(p) models with orders $p = 0, \dots, p_{max}$ and choose the value of p which minimizes some model selection criteria. Model selection criteria for VAR(p) models have the form

$$IC(p) = \ln |\tilde{\Sigma}(p)| + c_T \cdot \varphi(n, p)$$

where $\tilde{\Sigma}(p) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t'$ is the residual covariance matrix *without a degrees of freedom correction* from a VAR(p) model, c_T is a sequence indexed by the sample size T , and $\varphi(n, p)$ is a penalty function which penalizes large VAR(p) models. The three most common information criteria are the Akaike (AIC), Schwarz-Bayesian (BIC) and Hannan-Quinn (HQ):

Testing Linear Hypotheses

Now, consider testing the hypothesis that $\Pi_1 = \mathbf{0}$ (i.e., \mathbf{Y}_{t-1} does not help to explain \mathbf{Y}_t) using the Wald statistic (11.5). In terms of the columns of $vec(\Pi)$ the restrictions are $\pi_1 = (c_1, 0, 0)'$ and $\pi_2 = (c_2, 0, 0)$ and may be expressed as $\mathbf{R}vec(\Pi) = \mathbf{r}$ with

$$\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

In []:

```
In [6]: model=vars::VAR(df[,2:3], p = 1, type = c("const"))
model
#notice that intercept values are different than mean
```

VAR Estimation Results:
=====

Estimated coefficients for equation x:
=====

Call:

x = x.l1 + y.l1 + const

x.l1	y.l1	const
0.6826432	0.2086953	15.7562743

Estimated coefficients for equation y:
=====

Call:

y = x.l1 + y.l1 + const

x.l1	y.l1	const
0.191004	0.673937	13.221108

In [7]: model %>% summary

```

VAR Estimation Results:
=====
Endogenous variables: x, y
Deterministic variables: const
Sample size: 249
Log Likelihood: -1435.795
Roots of the characteristic polynomial:
0.878 0.4786
Call:
vars::VAR(y = df[, 2:3], p = 1, type = c("const"))

Estimation results for equation x:
=====
x = x.l1 + y.l1 + const

      Estimate Std. Error t value Pr(>|t|)
x.l1    0.68264    0.03824  17.852 < 2e-16 ***
y.l1    0.20870    0.03898   5.354 1.97e-07 ***
const  15.75627    2.04940   7.688 3.57e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.03 on 246 degrees of freedom
Multiple R-Squared: 0.9161,    Adjusted R-squared: 0.9154
F-statistic: 1343 on 2 and 246 DF,  p-value: < 2.2e-16

Estimation results for equation y:
=====
y = x.l1 + y.l1 + const

      Estimate Std. Error t value Pr(>|t|)
x.l1    0.19100    0.04467   4.276 2.73e-05 ***
y.l1    0.67394    0.04554  14.800 < 2e-16 ***
const  13.22111    2.39409   5.522 8.49e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.707 on 246 degrees of freedom
Multiple R-Squared: 0.8807,    Adjusted R-squared: 0.8798
F-statistic: 908.4 on 2 and 246 DF,  p-value: < 2.2e-16

```

Covariance matrix of residuals:

```

      x      y
x 16.24  1.30
y  1.30 22.16

```

Correlation matrix of residuals:

```

      x      y
x 1.00000 0.06853
y 0.06853 1.00000

```

In [8]: resids=vars::normality.test((model))
resids\$resid %>% cov

A matrix: 2 × 2 of type dbl

```

      x      y
x 16.107595  1.289528
y  1.289528 21.981430

```

In []:

Stationarity Condition

1. (y_t, x_t) are both stationary when the eigenvalues of ϕ are less than one in absolute value.
2. (y_t, x_t) are both integrated of order one, and y_t and x_t are cointegrated when one eigenvalue is unity, and the other eigenvalue is less than one in absolute value.
3. (y_t, x_t) are both integrated of order two if both eigenvalues of ϕ are unity.

```
In [9]: ## όπου φ η μήτρα συντελεστών
model %>% summary -> summr
summr$roots
```

Your code contains a unicode char which cannot be displayed in your current locale and R will silently convert it to an escaped form when the R kernel executes this code. This can lead to subtle errors if you use such chars to do comparisons. For more information, please see <https://github.com/IRkernel/repr/wiki/Problems-with-unicode-on-windows>

```
0.877991319268583 · 0.478588852123629
```

```
In [10]: coefmat=vars::Bcoef(model)
coefmat[, -3]
eigen(coefmat[, -3])
```

A matrix: 2 × 2 of type dbl

	x.l1	y.l1
x	0.6826432	0.2086953
y	0.1910040	0.6739370

```
eigen() decomposition
$values
[1] 0.8779913 0.4785889
```

```
$vectors
      [,1]      [,2]
[1,] 0.7300671 -0.7150123
[2,] 0.6833755 0.6991119
```


OLS Estimation

1. Each equation in the VAR can be estimated by OLS.
2. Then the variance covariance matrix for the error vector

$$\Omega = E w_t w_t' = \begin{pmatrix} \sigma_u^2 & \sigma_{u,v} \\ \sigma_{u,v} & \sigma_v^2 \end{pmatrix}$$

can be estimated by

$$\hat{\Omega} = T^{-1} \begin{pmatrix} \sum \hat{u}_t^2 & \sum \hat{u}_t \hat{v}_t \\ \sum \hat{u}_t \hat{v}_t & \sum \hat{v}_t^2 \end{pmatrix}$$

where \hat{u}_t and \hat{v}_t are residuals.

```
In [15]: datay=data.frame(y=y,y1=lag(y),x1=lag(x))
```

```
modely=lm(data=datay,y~y1+x1)
```

```
resy=modely$residuals
```

```
modely %>% summary
```

Call:

```
lm(formula = y ~ y1 + x1, data = datay)
```

Residuals:

Min	1Q	Median	3Q	Max
-11.1215	-3.2550	-0.4494	2.9876	16.0776

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	13.22111	2.39409	5.522	8.49e-08 ***
y1	0.67394	0.04554	14.800	< 2e-16 ***
x1	0.19100	0.04467	4.276	2.73e-05 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.707 on 246 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 0.8807, Adjusted R-squared: 0.8798

F-statistic: 908.4 on 2 and 246 DF, p-value: < 2.2e-16

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

In [16]: `datax=data.frame(x=x,y1=lag(y),x1=lag(x))`

```
modelx=lm(data=datay,x~y1+x1)
resx=modelx$residuals
modelx %>% summary
```

Call:

```
lm(formula = x ~ y1 + x1, data = datay)
```

Residuals:

Min	1Q	Median	3Q	Max
-10.9306	-2.4150	0.1156	2.6742	9.8363

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	15.75627	2.04940	7.688	3.57e-13 ***
y1	0.20870	0.03898	5.354	1.97e-07 ***
x1	0.68264	0.03824	17.852	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.03 on 246 degrees of freedom
(1 observation deleted due to missingness)

Multiple R-squared: 0.9161, Adjusted R-squared: 0.9154

F-statistic: 1343 on 2 and 246 DF, p-value: < 2.2e-16

In [17]: `cov(cbind(resy,resx))`

A matrix: 2 × 2 of type dbl

	resy	resx
resy	21.981430	1.289528
resx	1.289528	16.107595

In []:

Impulse response

Forecasts for longer horizons h (h -step forecasts) may be obtained using the *chain-rule of forecasting* as

$$\mathbf{Y}_{T+h|T} = \mathbf{c} + \mathbf{\Pi}_1 \mathbf{Y}_{T+h-1|T} + \cdots + \mathbf{\Pi}_p \mathbf{Y}_{T+h-p|T}$$

where $\mathbf{Y}_{T+j|T} = \mathbf{Y}_{T+j}$ for $j \leq 0$. The h -step forecast errors may be expressed as

$$\mathbf{Y}_{T+h} - \mathbf{Y}_{T+h|T} = \sum_{s=0}^{h-1} \mathbf{\Psi}_s \boldsymbol{\varepsilon}_{T+h-s}$$

where the matrices $\mathbf{\Psi}_s$ are determined by recursive substitution

$$\mathbf{\Psi}_s = \sum_{j=1}^{p-1} \mathbf{\Psi}_{s-j} \mathbf{\Pi}_j \quad (11.6)$$

with $\mathbf{\Psi}_0 = \mathbf{I}_n$ and $\mathbf{\Pi}_j = 0$ for $j > p$.¹ The forecasts are unbiased since all of

Wold representation

11.4.2 Impulse Response Functions

Any covariance stationary VAR(p) process has a Wold representation of the form

$$\mathbf{Y}_t = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \mathbf{\Psi}_1 \boldsymbol{\varepsilon}_{t-1} + \mathbf{\Psi}_2 \boldsymbol{\varepsilon}_{t-2} + \cdots \quad (11.11)$$

where the $(n \times n)$ moving average matrices $\mathbf{\Psi}_s$ are determined recursively using (11.6). It is tempting to interpret the (i, j) -th element, ψ_{ij}^s , of the matrix $\mathbf{\Psi}_s$ as the dynamic multiplier or impulse response

$$\frac{\partial y_{i,t+s}}{\partial \varepsilon_{j,t}} = \frac{\partial y_{i,t}}{\partial \varepsilon_{j,t-s}} = \psi_{ij}^s, \quad i, j = 1, \dots, n$$

However, this interpretation is only possible if $\text{var}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma}$ is a diagonal matrix so that the elements of $\boldsymbol{\varepsilon}_t$ are uncorrelated. One way to make the errors uncorrelated is to follow Sims (1980) and estimate the *triangular structural* VAR(p) model

$$\begin{aligned} y_{1t} &= c_1 + \gamma'_{11} \mathbf{Y}_{t-1} + \cdots + \gamma'_{1p} \mathbf{Y}_{t-p} + \eta_{1t} \\ y_{2t} &= c_1 + \beta_{21} y_{1t} + \gamma'_{21} \mathbf{Y}_{t-1} + \cdots + \gamma'_{2p} \mathbf{Y}_{t-p} + \eta_{2t} \\ y_{3t} &= c_1 + \beta_{31} y_{1t} + \beta_{32} y_{2t} + \gamma'_{31} \mathbf{Y}_{t-1} + \cdots + \gamma'_{3p} \mathbf{Y}_{t-p} + \eta_{3t} \\ &\vdots \\ y_{nt} &= c_1 + \beta_{n1} y_{1t} + \cdots + \beta_{n,n-1} y_{n-1,t} + \gamma'_{n1} \mathbf{Y}_{t-1} + \cdots + \gamma'_{np} \mathbf{Y}_{t-p} + \eta_{nt} \end{aligned} \quad (11.12)$$

In matrix form, the triangular structural VAR(p) model is

$$\mathbf{B} \mathbf{Y}_t = \mathbf{c} + \mathbf{\Gamma}_1 \mathbf{Y}_{t-1} + \mathbf{\Gamma}_2 \mathbf{Y}_{t-2} + \cdots + \mathbf{\Gamma}_p \mathbf{Y}_{t-p} + \boldsymbol{\eta}_t \quad (11.13)$$

where

where

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\beta_{21} & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\beta_{n1} & -\beta_{n2} & \cdots & 1 \end{pmatrix} \quad (11.14)$$

is a lower triangular matrix with 1's along the diagonal. The algebra of least squares will ensure that the estimated covariance matrix of the error vector $\boldsymbol{\eta}_t$ is diagonal. The uncorrelated/orthogonal errors $\boldsymbol{\eta}_t$ are referred to as *structural* errors.

The triangular structural model (11.12) imposes the *recursive causal ordering*

$$y_1 \rightarrow y_2 \rightarrow \cdots \rightarrow y_n \quad (11.15)$$

The ordering (11.15) means that the contemporaneous values of the variables to the left of the arrow \rightarrow affect the contemporaneous values of the variables to the right of the arrow but not vice-versa. These contemporaneous effects are captured by the coefficients β_{ij} in (11.12). For example, the ordering $y_1 \rightarrow y_2 \rightarrow y_3$ imposes the restrictions: y_{1t} affects y_{2t} and y_{3t} but y_{2t} and y_{3t} do not affect y_{1t} ; y_{2t} affects y_{3t} but y_{3t} does not affect y_{2t} . Similarly, the ordering $y_2 \rightarrow y_3 \rightarrow y_1$ imposes the restrictions: y_{2t} affects y_{3t} and y_{1t} but y_{3t} and y_{1t} do not affect y_{2t} ; y_{3t} affects y_{1t} but y_{1t} does not affect y_{3t} . For a VAR(p) with n variables there are $n!$ possible recursive causal orderings. Which ordering to use in practice depends on the context and whether prior theory can be used to justify a particular ordering. Results from alternative orderings can always be compared to determine the sensitivity of results to the imposed ordering.

Once a recursive ordering has been established, the Wold representation of \mathbf{Y}_t based on the orthogonal errors $\boldsymbol{\eta}_t$ is given by

$$\mathbf{Y}_t = \boldsymbol{\mu} + \boldsymbol{\Theta}_0 \boldsymbol{\eta}_t + \boldsymbol{\Theta}_1 \boldsymbol{\eta}_{t-1} + \boldsymbol{\Theta}_2 \boldsymbol{\eta}_{t-2} + \cdots \quad (11.16)$$

where $\boldsymbol{\Theta}_0 = \mathbf{B}^{-1}$ is a lower triangular matrix. The impulse responses to the orthogonal shocks η_{jt} are

$$\frac{\partial y_{i,t+s}}{\partial \eta_{j,t}} = \frac{\partial y_{i,t}}{\partial \eta_{j,t-s}} = \theta_{ij}^s, \quad i, j = 1, \dots, n; s > 0 \quad (11.17)$$

where θ_{ij}^s is the (i, j) th element of $\boldsymbol{\Theta}_s$. A plot of θ_{ij}^s against s is called the *orthogonal impulse response function* (IRF) of y_i with respect to η_j . With n variables there are n^2 possible impulse response functions.

where θ_{ij}^s is the (i, j) th element of Θ_s . A plot of θ_{ij}^s against s is called the *orthogonal impulse response function* (IRF) of y_i with respect to η_j . With n variables there are n^2 possible impulse response functions.

In practice, the orthogonal IRF (11.17) based on the triangular VAR(p) (11.12) may be computed directly from the parameters of the non triangular VAR(p) (11.1) as follows. First, decompose the residual covariance matrix Σ as

$$\Sigma = \mathbf{A}\mathbf{D}\mathbf{A}'$$

where \mathbf{A} is an invertible lower triangular matrix with 1's along the diagonal and \mathbf{D} is a diagonal matrix with positive diagonal elements. Next, define the structural errors as

$$\boldsymbol{\eta}_t = \mathbf{A}^{-1}\boldsymbol{\varepsilon}_t$$

These structural errors are orthogonal by construction since $\text{var}(\boldsymbol{\eta}_t) = \mathbf{A}^{-1}\Sigma\mathbf{A}^{-1'} = \mathbf{A}^{-1}\mathbf{A}\mathbf{D}\mathbf{A}'\mathbf{A}^{-1'} = \mathbf{D}$. Finally, re-express the Wold representation (11.11) as

$$\begin{aligned}\mathbf{Y}_t &= \boldsymbol{\mu} + \mathbf{A}\mathbf{A}^{-1}\boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_1\mathbf{A}\mathbf{A}^{-1}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_2\mathbf{A}\mathbf{A}^{-1}\boldsymbol{\varepsilon}_{t-2} + \cdots \\ &= \boldsymbol{\mu} + \boldsymbol{\Theta}_0\boldsymbol{\eta}_t + \boldsymbol{\Theta}_1\boldsymbol{\eta}_{t-1} + \boldsymbol{\Theta}_2\boldsymbol{\eta}_{t-2} + \cdots\end{aligned}$$

where $\boldsymbol{\Theta}_j = \boldsymbol{\Psi}_j\mathbf{A}$. Notice that the structural B matrix in (11.13) is equal to \mathbf{A}^{-1} .

In [18]:

Άλλες σημειώσεις

https://www.fsb.miamioh.edu/lij14/672_s7.pdf (https://www.fsb.miamioh.edu/lij14/672_s7.pdf)

Impulse Response

- Using the lag operator we can show the $MA(\infty)$ representation for the VAR(1) is

$$z_t = w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \dots + \phi^j w_{t-j} + \dots \quad (9)$$

- The coefficient in the MA representation measures the impulse response:

$$\phi^j = \frac{dz_t}{dw_{t-j}} \quad (10)$$

Note ϕ^j is a 2×2 matrix for a bivariate system.

Impulse Response III

- In general u_t and v_t are contemporaneously correlated (not-orthogonal), i.e., $\sigma_{u,v} \neq 0$
- Therefore we can not, say, hold v constant and let only u vary.
- However, we can always find a lower triangular matrix A so that

$$\Omega = AA' \quad (\text{Cholesky Decomposition}) \quad (11)$$

- Then define a new error vector \tilde{w}_t as (linear transformation of old error vector w_t)

$$\tilde{w}_t = A^{-1}w_t \quad (12)$$

- By construction the new error is orthogonal because its variance-covariance matrix is diagonal:

$$\text{var}(\tilde{w}_t) = A^{-1}\text{var}(w_t)A^{-1'} = A^{-1}\Omega A^{-1'} = A^{-1}AA'A^{-1'} = I$$

Cholesky Decomposition

Let $A = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$. The Cholesky Decomposition tries to solve

$$\begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} \sigma_u^2 & \sigma_{u,v} \\ \sigma_{u,v} & \sigma_v^2 \end{pmatrix}$$

The solutions for a, b, c always exist and they are

$$a = \sqrt{\sigma_u^2} \quad (13)$$

$$b = \frac{\sigma_{u,v}}{\sqrt{\sigma_u^2}} \quad (14)$$

$$c = \sqrt{\sigma_v^2 - \frac{\sigma_{u,v}^2}{\sigma_u^2}} \quad (15)$$

Positive Definite Matrix

Note c is always a real number since Ω is a variance-covariance matrix, and so is positive definite (i.e., $\sigma_v^2 - \frac{\sigma_{u,v}^2}{\sigma_u^2}$ is always positive because the determinant of Ω , or the second leading principal minor, is positive).

Impulse Response to Orthogonal Errors

Rewrite the $MA(\infty)$ representation as

$$z_t = w_t + \phi w_{t-1} + \dots + \phi^j w_{t-j} + \dots \quad (16)$$

$$= AA^{-1}w_t + \phi AA^{-1}w_{t-1} + \dots + \phi^j AA^{-1}w_{t-j} + \dots \quad (17)$$

$$= A\tilde{w}_t + \phi A\tilde{w}_{t-1} + \dots + \phi^j A\tilde{w}_{t-j} + \dots \quad (18)$$

This implies that the impulse response to the orthogonal error \tilde{w}_t after j periods ($j = 0, 1, 2, \dots$) is

$$j\text{-th orthogonal impulse response} = \phi^j A \quad (19)$$

where A satisfies (11).

Reduced and structural form

4

Matrix Form

The matrix form for VAR(1) is

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix} \quad (3)$$

or

$$z_t = \phi z_{t-1} + w_t \quad (\text{Reduced Form}) \quad (4)$$

where $z_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix}$, $\phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}$, and $w_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix}$

Reduced Form and Structural Form

1. (4) is called reduced form and the reduced form error w_t is not orthogonal.
2. The so called structural form is a linear transformation of the reduced form:

$$z_t = \phi z_{t-1} + w_t \quad (20)$$

$$\Rightarrow A^{-1} z_t = A^{-1} \phi z_{t-1} + A^{-1} w_t \quad (21)$$

$$\Rightarrow A^{-1} z_t = A^{-1} \phi z_{t-1} + \tilde{w}_t \quad (22)$$

where the structural form error \tilde{w}_t is orthogonal.

Structural Form

1. The Structural Form VAR is

$$A^{-1}z_t = A^{-1}\phi z_{t-1} + \tilde{w}_t \quad (\text{Structural Form}) \quad (23)$$

2. Note A^{-1} is lower triangular. So we have following recursive form (suggested by Sims (1980))

$$A^{-1}z_t = \begin{pmatrix} 1/a & 0 \\ -b/(ac) & 1/c \end{pmatrix} \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} (1/a)y_t \\ (-b/(ac))y_t + (1/c)x_t \end{pmatrix}$$

3. That means x_t will not appear in the regression for y_t , whereas y_t will appear in the regression for x_t .
4. In words, x does not contemporaneously affect y , but y contemporaneously affects x .
Example is that x is Hong Kong's interest rate, while y is US interest rate.
5. Sims method is limited in the sense that the result is sensitive to choice or rank of y and x .

Structural Form VAR I

1. In general, the reduced form VAR(1) is

$$z_t = \phi z_{t-1} + B\tilde{w}_t \quad (24)$$

and the corresponding structural form is

$$B^{-1}z_t = B^{-1}\phi z_{t-1} + \tilde{w}_t \quad (25)$$

2. Because B^{-1} is generally not diagonal, y_t and x_t are related contemporaneously in the structural form.
3. Because we use B^{-1} to account for the contemporaneous correlation, we can always (safely) assume the structural error is orthogonal, i.e.,

$$\text{var}(\tilde{w}_t) = \mathbf{I}.$$

4. The structural form error vector and reduced form error vector are related since

$$B\tilde{w}_t = w_t$$

Structural form

Why is it called structural form?

1. Because in the structural VAR there is instantaneous interaction between y_t and x_t .
2. Both y_t and x_t are endogenous, and the regressors include the current value of endogenous variables in the structural form.
3. The structural VAR is one example of the simultaneous equation model (SEM)
4. We cannot estimate the structural VAR using per-equation OLS, due to the bias of simultaneity.
5. We can estimate the reduced form using per-equation OLS. Then we recover the structural form from the reduced form, with (identification) restriction imposed.

Structural Form VAR II

1. Let $\Omega = E(w_t w_t')$ be the observed variance covariance matrix. It follows that

$$BB' = \Omega$$

2. The goal of structural VAR analysis is to obtain B , which is not unique (for a bivariate system Ω has 3 unique elements, while B has 4 elements to be determined).
3. The Sims (1980) structural VAR, which is of the recursive form, imposes the restriction that B is lower triangular.
4. The Blanchard Quah structural VAR obtains B by looking at the long run effect of the \tilde{w}_t .

MA representation

MA Representation

Consider a univariate AR(p)

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t$$

Using the lag operator we can write

$$(I - \phi_1 L - \dots - \phi_p L^p) y_t = e_t$$

We can obtain the MA representation by inverting $(I - \phi_1 L - \dots - \phi_p L^p)$:

$$\begin{aligned} y_t &= (I - \phi_1 L - \dots - \phi_p L^p)^{-1} e_t \\ &= e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots \end{aligned}$$

22

Long Run Effect

The long run effect (LRE) of e_t is

$$\text{LRE} \equiv \frac{dy_t}{de_t} + \frac{dy_{t+1}}{de_t} + \frac{dy_{t+2}}{de_t} + \dots \quad (26)$$

$$= I + \theta_1 + \theta_2 + \dots \quad (27)$$

$$= (I - \phi_1 - \dots - \phi_p)^{-1} \quad (28)$$

In words, we need to invert the lag polynomial $(I - \phi_1 L - \dots - \phi_p L^p)$ and replace L with identity matrix I .

Simulated from

https://www.fsb.miamioh.edu/li/14/672_s7.pdf (https://www.fsb.miamioh.edu/li/14/672_s7.pdf)

```
In [19]: set.seed(1000) # so the simulation result can be duplicated
n = 200 # sample size = 200
z = as.matrix(cbind(rep(0, n), rep(0, n)))
w = as.matrix(cbind(rnorm(n), rnorm(n)))
phi = as.matrix(cbind(c(0.3, 0.5), c(0, 0.6)))
for (i in 2:n) {
  z[i,] = phi %*% z[i-1,] + w[i,]
}

In [20]: library(vars)

Warning message:
"package 'vars' was built under R version 4.1.3"
Loading required package: MASS

Attaching package: 'MASS'

The following object is masked from 'package:dplyr':

  select

Loading required package: strucchange

Warning message:
"package 'strucchange' was built under R version 4.1.1"
Loading required package: zoo

Attaching package: 'zoo'

The following objects are masked from 'package:base':

  as.Date, as.Date.numeric

Loading required package: sandwich

Warning message:
"package 'sandwich' was built under R version 4.1.1"

Attaching package: 'strucchange'

The following object is masked from 'package:stringr':

  boundary

Loading required package: urca

Warning message:
"package 'urca' was built under R version 4.1.1"
Loading required package: lmtest

Warning message:
"package 'lmtest' was built under R version 4.1.1"

In [21]: VARselect(z)
```

\$selection

AIC(n): 1 HQ(n): 1 SC(n): 1 FPE(n): 1

\$criteria

A matrix: 4 × 10 of type dbl

	1	2	3	4	5	6	7	8	9	10
AIC(n)	-0.06176601	-0.04826772	-0.05658535	-0.03983166	-0.02112784	0.002446965	0.02916935	0.05837397	0.09514411	0.1069014
HQ(n)	-0.02022953	0.02095977	0.04033312	0.08477780	0.13117261	0.182438416	0.23685179	0.29374741	0.35820854	0.3976568
SC(n)	0.04077159	0.12262829	0.18266905	0.26778114	0.35484336	0.446776575	0.54185736	0.63942039	0.74454892	0.8246646
FPE(n)	0.94010777	0.95290182	0.94504889	0.96108756	0.97934766	1.002879385	1.03027719	1.06112921	1.10129361	1.1148446

```
In [22]: var.1c = VAR(z, p=1, type = "both")  
summary(var.1c)  
var.1c = VAR(z, p=1, type = "none")  
summary(var.1c)
```

Warning message in VAR(z, p = 1, type = "both"):
 "No column names supplied in y, using: y1, y2 , instead."
 "

VAR Estimation Results:

=====

Endogenous variables: y1, y2

Deterministic variables: both

Sample size: 199

Log Likelihood: -546.002

Roots of the characteristic polynomial:

0.4243 0.4243

Call:

VAR(y = z, p = 1, type = "both")

Estimation results for equation y1:

=====

y1 = y1.l1 + y2.l1 + const + trend

	Estimate	Std. Error	t value	Pr(> t)
y1.l1	0.227132	0.069460	3.270	0.00127 **
y2.l1	-0.083555	0.049185	-1.699	0.09095 .
const	-0.179921	0.140330	-1.282	0.20132
trend	0.002479	0.001232	2.013	0.04553 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9492 on 195 degrees of freedom
 Multiple R-Squared: 0.0859, Adjusted R-squared: 0.07183
 F-statistic: 6.108 on 3 and 195 DF, p-value: 0.0005426

Estimation results for equation y2:

=====

y2 = y1.l1 + y2.l1 + const + trend

	Estimate	Std. Error	t value	Pr(> t)
y1.l1	0.576700	0.071714	8.042	8.4e-14 ***
y2.l1	0.580314	0.050781	11.428	< 2e-16 ***
const	-0.186417	0.144883	-1.287	0.200
trend	0.001542	0.001272	1.212	0.227

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.98 on 195 degrees of freedom
 Multiple R-Squared: 0.5378, Adjusted R-squared: 0.5307
 F-statistic: 75.64 on 3 and 195 DF, p-value: < 2.2e-16

Covariance matrix of residuals:

	y1	y2
y1	0.90099	-0.05112
y2	-0.05112	0.96041

Correlation matrix of residuals:

	y1	y2
y1	1.00000	-0.05496
y2	-0.05496	1.00000

Warning message in VAR(z, p = 1, type = "none"):
 "No column names supplied in y, using: y1, y2 , instead."
 "

```

VAR Estimation Results:
=====
Endogenous variables: y1, y2
Deterministic variables: none
Sample size: 199
Log Likelihood: -549.51
Roots of the characteristic polynomial:
  0.43  0.43
Call:
VAR(y = z, p = 1, type = "none")

Estimation results for equation y1:
=====
y1 = y1.l1 + y2.l1

      Estimate Std. Error t value Pr(>|t|)
y1.l1  0.25493    0.06887   3.702 0.000278 ***
y2.l1 -0.05589    0.04771  -1.171 0.242875
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9565 on 197 degrees of freedom
Multiple R-Squared: 0.06981,    Adjusted R-squared: 0.06037
F-statistic: 7.392 on 2 and 197 DF,  p-value: 0.0008023

Estimation results for equation y2:
=====
y2 = y1.l1 + y2.l1

      Estimate Std. Error t value Pr(>|t|)
y1.l1  0.58693    0.07050   8.325 1.4e-14 ***
y2.l1  0.59648    0.04885  12.211 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9792 on 197 degrees of freedom
Multiple R-Squared: 0.5342,    Adjusted R-squared: 0.5295
F-statistic: 113 on 2 and 197 DF,  p-value: < 2.2e-16

Covariance matrix of residuals:
      y1      y2
y1  0.9104 -0.0391
y2 -0.0391  0.9578

Correlation matrix of residuals:
      y1      y2
y1  1.00000 -0.04187
y2 -0.04187  1.00000

```

In []:

Original simulation

```

In [26]: #with df adjustment
n=250
sigmau2 = var(resy)*(n-1)/(n-2)
sigmav2 = var(resx)*(n-1)/(n-2)
sigmauv = cov(resy,resx)*(n-1)/(n-2)
omegahat = as.matrix(cbind(c(sigmau2, sigmauv), c(sigmauv, sigmav2)))

```

In [33]:

```

A matrix: 2 × 2 of type dbl

22.070065  1.294728
1.294728  16.172545

```

```

In [27]: ### Cholesky decomposition of omega
Aprime = chol(omegahat)
A = t(Aprime)
A
#The matrix A will be used to construct impulse response to orthogonal (or structural form) errors

```

```

A matrix: 2 × 2 of type
dbl

4.6978787  0.000000
0.2755984  4.012056

```


Impulse Response to Orthogonal Errors

The responses of y and x to the one-unit impulse of the orthogonal (structural-form) the error for y can be obtained as

43

Simulate Impulse Response II

Denote the impulse response after j periods by $IR(j)$, which is a 2×1 vector. It can be computed recursively as

$$IR(1) = A(1c) \quad (31)$$

$$IR(2) = \phi_1 IR(1) \quad (32)$$

$$IR(3) = \phi_1 IR(2) + \phi_2 IR(1) \quad (33)$$

$$\dots = \dots \quad (34)$$

$$IR(j) = \phi_1 IR(j-1) + \phi_2 IR(j-2) \quad (35)$$

To get the impulse response to the second structural form error, just let $\tilde{w}_t = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $A(1c)$ should change to $A(2c)$. Everything else remains unchanged.

```
In [30]: phihat=Bcoef(model)[-3]
         wtilde = as.matrix(c(1,0))
         A1c = A%*%wtilde # or A1c = A[,1]
         if1 = A1c
         if2 = phihat%*%if1
         if3 = phihat%*%if2
         if4 = phihat%*%if3
```

```
In [32]: #You can get the same result using the R command
var1c.11 = irf(model, impulse = "y", response="x", boot=TRUE)
var1c.11
```

Impulse response coefficients

```
$y
      x
[1,] 0.0000000
[2,] 0.9801142
[3,] 1.3296035
[4,] 1.3918728
[5,] 1.3294918
[6,] 1.2187017
[7,] 1.0946183
[8,] 0.9728428
[9,] 0.8597841
[10,] 0.7575806
[11,] 0.6664402
```

Lower Band, CI= 0.95

```
$y
      x
[1,] 0.0000000
[2,] 0.6760214
[3,] 0.9360768
[4,] 0.9607720
[5,] 0.8990895
[6,] 0.8049446
[7,] 0.6886456
[8,] 0.5851518
[9,] 0.5023533
[10,] 0.4305505
[11,] 0.3687119
```

Upper Band, CI= 0.95

```
$y
      x
[1,] 0.0000000
[2,] 1.3849531
[3,] 1.7342365
[4,] 1.7160405
[5,] 1.6468387
[6,] 1.5229915
[7,] 1.3772664
[8,] 1.2438140
[9,] 1.1183418
[10,] 1.0024658
[11,] 0.8881378
```

```
In [37]: causality(model, cause = c("y"))
```

\$Granger

Granger causality H0: y do not Granger-cause x

data: VAR object model

F-Test = 28.663, df1 = 1, df2 = 492, p-value = 1.324e-07

\$Instant

H0: No instantaneous causality between: y and x

data: VAR object model

Chi-squared = 1.164, df = 1, p-value = 0.2806

```
In [38]: causality(model, cause = c("x"))
```

\$Granger

Granger causality H0: x do not Granger-cause y

data: VAR object model

F-Test = 18.282, df1 = 1, df2 = 492, p-value = 2.289e-05

\$Instant

H0: No instantaneous causality between: x and y

data: VAR object model

Chi-squared = 1.164, df = 1, p-value = 0.2806

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []:

In []: