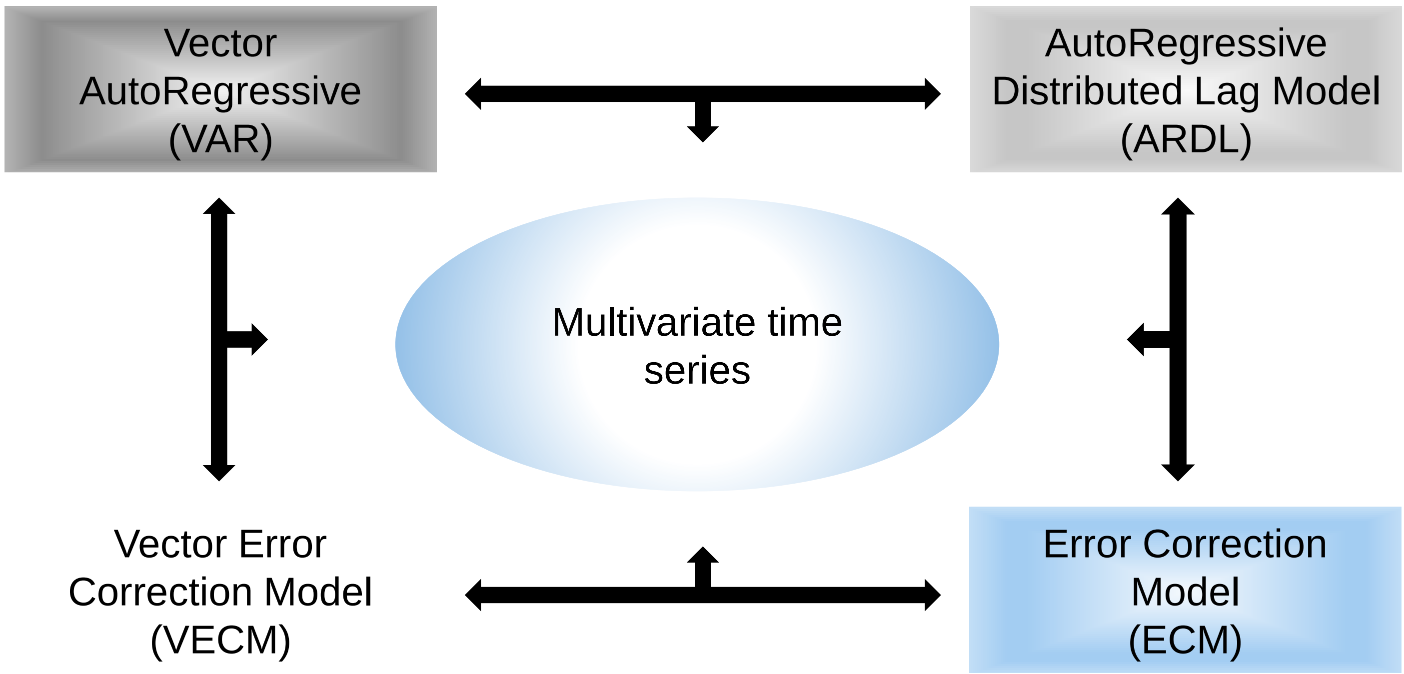
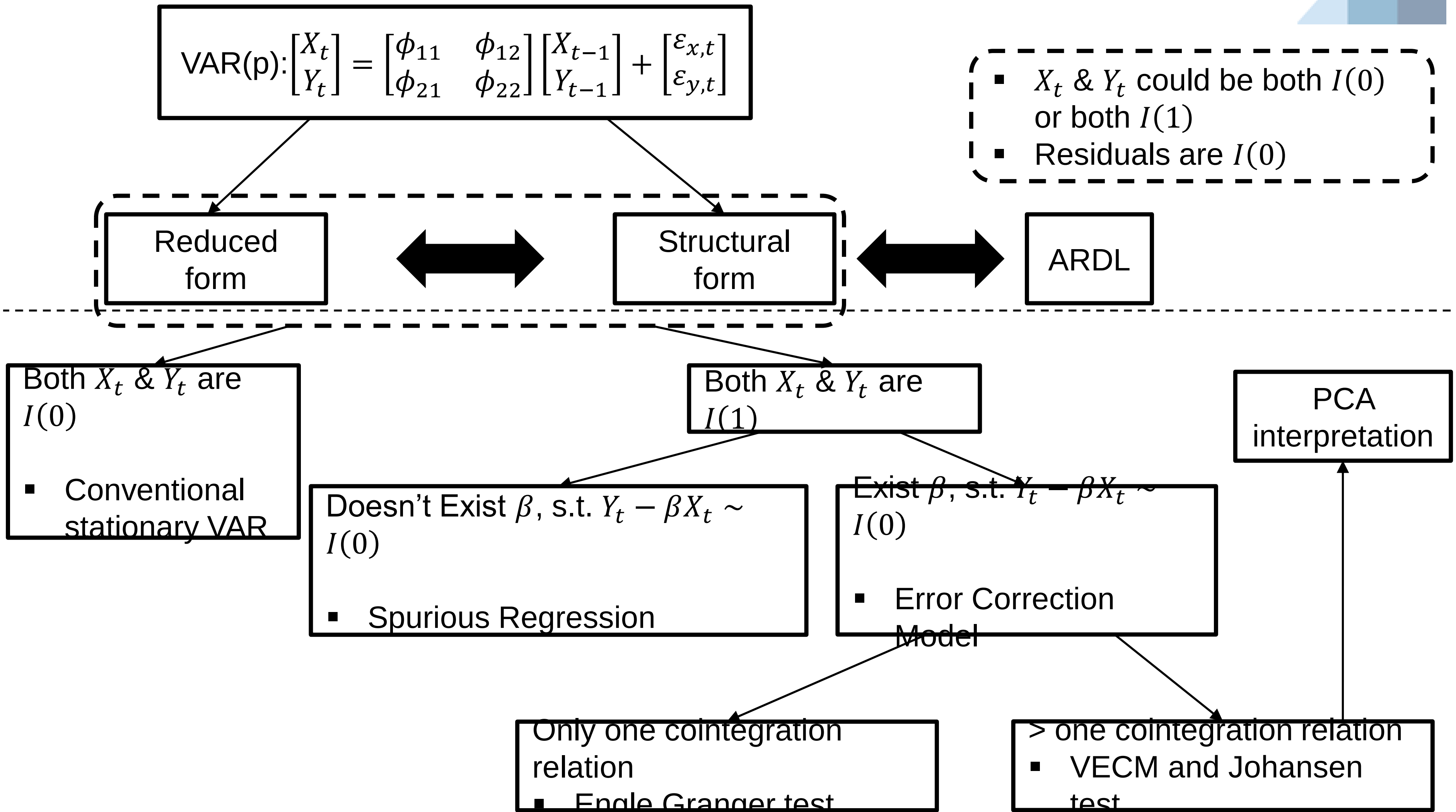
MacroEconometric Forecasting

Topic:

Cointegration and Vector Error Correction Models (VECMs)

Presented by Munisa Yashnarbekova


Vector AutoRegressive (VAR)
AutoRegressive Distributed Lag Model (ARDL)
Multivariate time series
Vector Error Correction Model (VECM)
Error Correction Model (ECM)


VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test

VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$

- X_t & Y_t could be both $I(0)$ or both $I(1)$
- Residuals are $I(0)$

Reduced form

Structural form

ARDL

Both X_t & Y_t are $I(0)$

- Conventional stationary VAR

Both X_t & Y_t are $I(1)$

PCA interpretation

Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$

- Spurious Regression

Exist β, s.t. $Y_t - \beta X_t \sim I(0)$

- Error Correction Model

Only one cointegration relation

- Engle Granger test

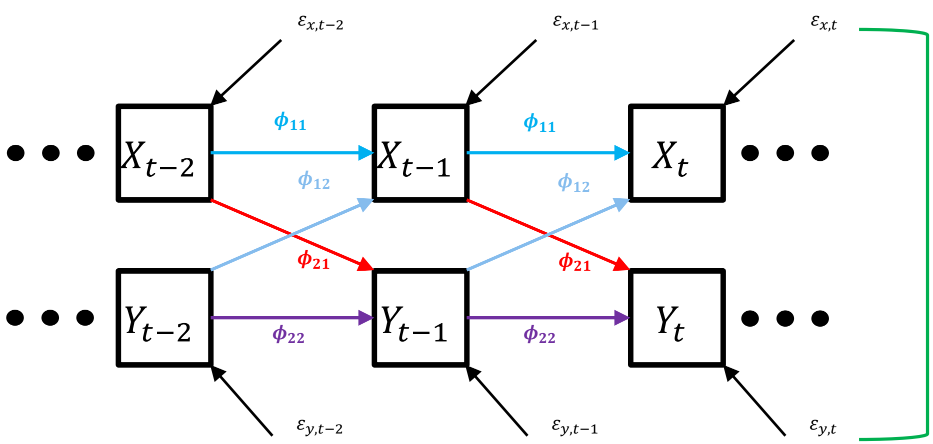
> one cointegration relation

- VECM and Johansen test

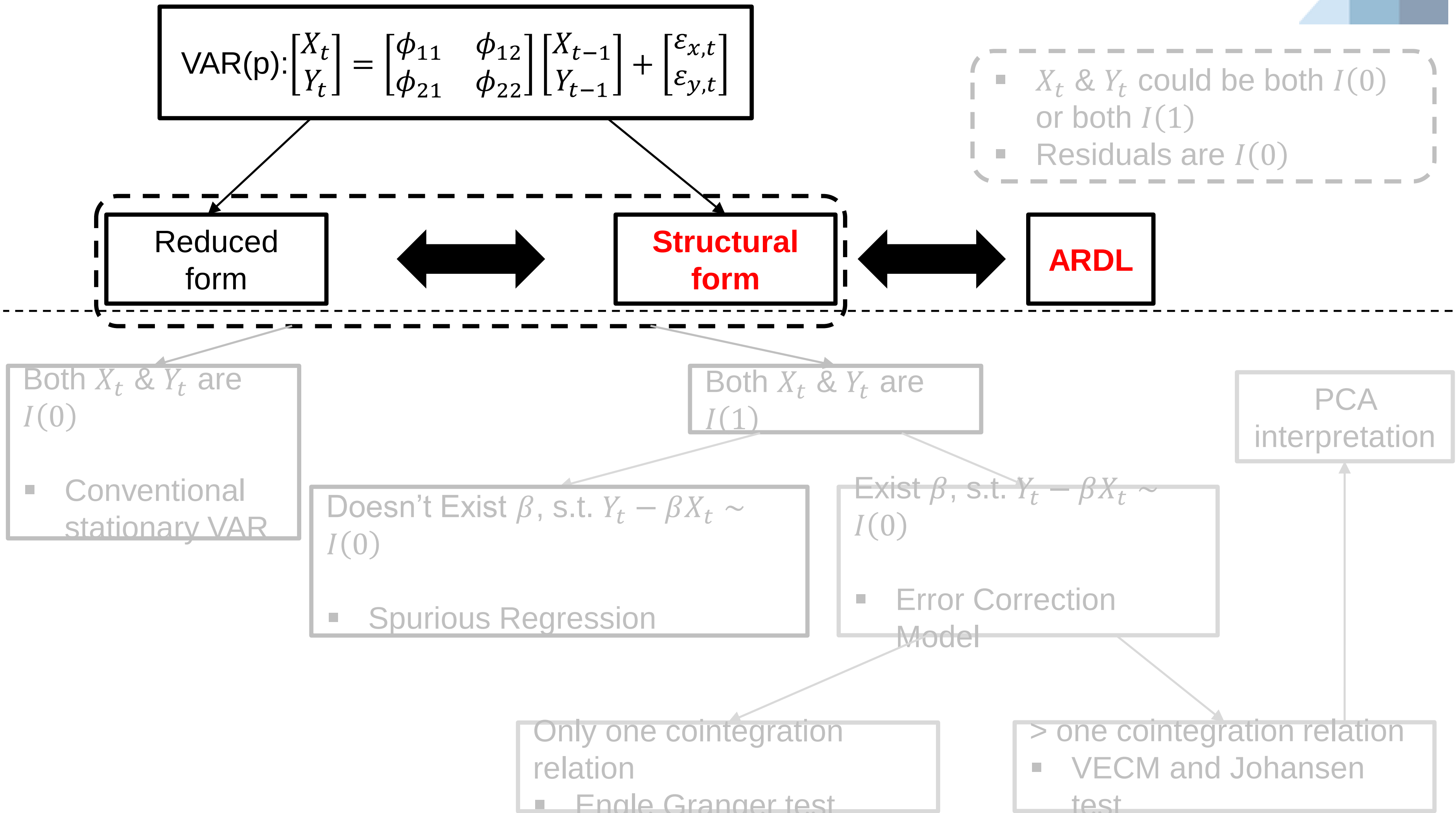
Reduced form of VAR

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix},$$

where $\begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \sim WN(0, \Sigma)$, and $\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$.



- Relation between X_t & Y_t are not directly shown
- Cholesky Decomposition: there exists $L = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$, s.t. $\Sigma = LDL'$, where $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$
- $L^{-1}\Sigma L'^{-1} = D$, where $L^{-1} = \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix}$
- $\begin{bmatrix} \phi_{11}^* & \phi_{12}^* \\ \phi_{21}^* & \phi_{22}^* \end{bmatrix} = L^{-1} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$, $\begin{bmatrix} \varepsilon_{x,t}^* \\ \varepsilon_{y,t}^* \end{bmatrix} = L^{-1} \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$, then
  $$L^{-1} \begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11}^* & \phi_{12}^* \\ \phi_{21}^* & \phi_{22}^* \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t}^* \\ \varepsilon_{y,t}^* \end{bmatrix}$$
- $Cov\left(\begin{bmatrix} \varepsilon_{x,t}^* \\ \varepsilon_{y,t}^* \end{bmatrix}\right) = L^{-1}\Sigma L'^{-1} = D$

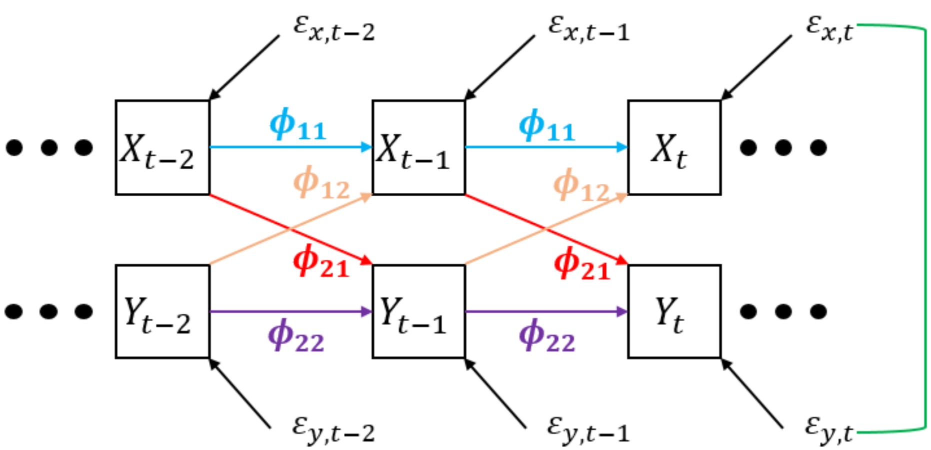


VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test

Structural form of VAR

$$L^{-1}\begin{bmatrix}X_t\\Y_t\end{bmatrix}=\begin{bmatrix}1 & 0\\-c & 1\end{bmatrix}\begin{bmatrix}X_t\\Y_t\end{bmatrix}=\begin{bmatrix}\phi_{11}^* & \phi_{12}^*\\\phi_{21}^* & \phi_{22}^*\end{bmatrix}\begin{bmatrix}X_{t-1}\\Y_{t-1}\end{bmatrix}+\begin{bmatrix}\varepsilon_{x,t}^*\\\varepsilon_{y,t}^*\end{bmatrix}$$

$$\begin{cases}X_t=\phi_{11}^*X_{t-1}+\phi_{12}^*Y_{t-1}+\varepsilon_{x,t}^*\\Y_t=cX_t+\phi_{21}^*X_{t-1}+\phi_{22}^*Y_{t-1}+\varepsilon_{y,t}^*\end{cases}$$



ARDL

- Relation between X_t & Y_t is indicated by c, as a results, $\varepsilon_{x,t}^*$ & $\varepsilon_{y,t}^*$ are no longer correlated
- Reduced form and structural form are equivalent
- $Y_t=cX_t+\phi_{21}^*X_{t-1}+\phi_{22}^*Y_{t-1}+\varepsilon_{y,t}^*$ is **A**uto**R**egressive **D**istributed **L**ag Model –ARDL(1,1)
- $X_t=\phi_{11}^*X_{t-1}+\phi_{12}^*Y_{t-1}+\varepsilon_{x,t}^*$ $=0Y_t+\phi_{11}^*X_{t-1}+\phi_{12}^*Y_{t-1}+\varepsilon_{x,t}^*$ is also ARDL

Fitting function of ARDL/ARX:

- Matlab:

```
Spec = vgxset('n',1,'nAR',1,'nX',1,'Constant',false);
estSpec = vgxvarx(Spec,Y(:,2),num2cell(Y(:,1)),[]);
```

- R Univariate: `fit = arima(y,c(1,0,0),xreg = x)`
- R Multivariate: `fit = VAR(cbind(y,z),p = 1,type = 'none',exogen = x)`

- SAS: `proc VARMAX;model y = x / p=1 noint;`

More than two variables

- Reduced form

$$\begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Kt} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \cdots & \phi_{1K} \\ \vdots & \ddots & \vdots \\ \phi_{K1} & \cdots & \phi_{KK} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \\ \vdots \\ X_{K,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{K,t} \end{bmatrix}, \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{K,t} \end{bmatrix} \sim N(0, \Sigma).$$

- Structural form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21}^* & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{K1}^* & l_{K2}^* & \cdots & 1 \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Kt} \end{bmatrix} = \begin{bmatrix} \phi_{11}^* & \cdots & \phi_{1K}^* \\ \vdots & \ddots & \vdots \\ \phi_{K1}^* & \cdots & \phi_{KK}^* \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \\ \vdots \\ X_{K,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t}^* \\ \varepsilon_{2,t}^* \\ \vdots \\ \varepsilon_{K,t}^* \end{bmatrix}, cov\left(\varepsilon_{i,t}^*, \varepsilon_{j,t}^*\right) = 0$$

- ARDL

$$\begin{cases} X_{1t} = \phi_{11}^* X_{1,t-1} + \cdots + \phi_{1K}^* X_{K,t-1} + \varepsilon_{1,t}^* \\ X_{2t} = \phi_{21}^* X_{1,t-1} + \cdots + \phi_{2K}^* X_{K,t-1} - l_{21}^* X_{1t} + \varepsilon_{2,t}^* \\ X_{3t} = \phi_{31}^* X_{1,t-1} + \cdots + \phi_{3K}^* X_{K,t-1} - l_{31}^* X_{1t} - l_{32}^* X_{2t} + \varepsilon_{3,t}^* \\ \vdots \\ X_{Kt} = \phi_{K1}^* X_{1,t-1} + \cdots + \phi_{KK}^* X_{K,t-1} - \Sigma_{i=1}^{K-1} X_{it} + \varepsilon_{K,t}^* \end{cases}$$

What if VAR(p) instead of VAR(1)?

- VAR(2)

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \phi_{11}^{(2)} & \phi_{12}^{(2)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}, \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \sim WN(0, \Sigma)$$

- Transform to VAR(1)

$$\begin{bmatrix} X_{t-1} \\ Y_{t-1} \\ X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \phi_{11}^{(2)} & \phi_{12}^{(2)} & \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} & \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{bmatrix} \begin{bmatrix} X_{t-2} \\ Y_{t-2} \\ X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$$

Companion matrix

- Any VAR(p) could be transformed to VAR(1)

VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$

- X_t & Y_t could be both $I(0)$ or both $I(1)$
- Residuals are $I(0)$

Reduced form ⟷ Structural form ⟷ ARDL

Both X_t & Y_t are $I(0)$
- Conventional stationary VAR

Both X_t & Y_t are $I(1)$

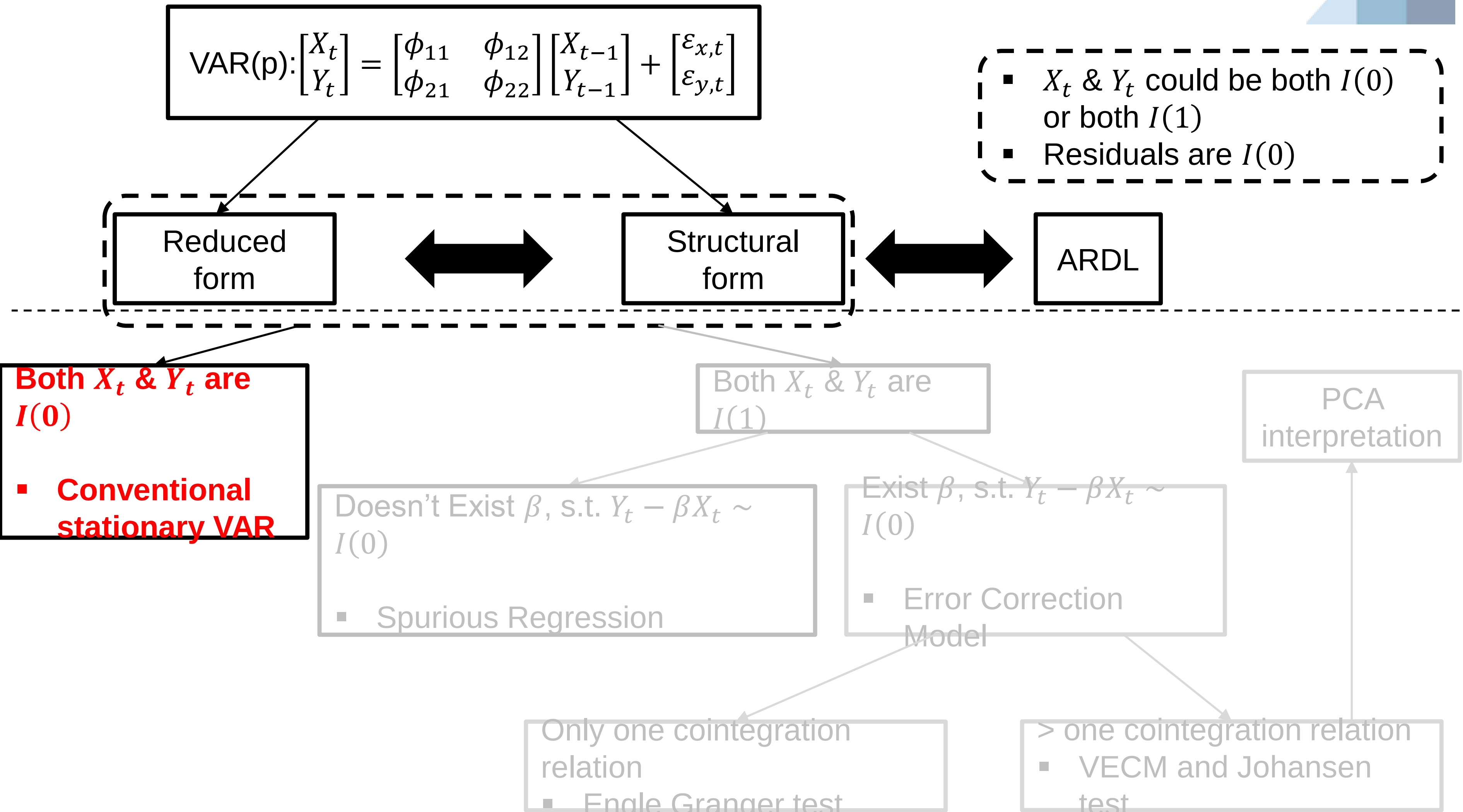
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
- Spurious Regression

Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
- Error Correction Model

Only one cointegration relation
- Engle Granger test

> one cointegration relation
- VECM and Johansen test

PCA interpretation


VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test

- $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
- Roots of $\left| I - \begin{bmatrix} \phi_{11}B & \phi_{12}B \\ \phi_{21}B & \phi_{22}B \end{bmatrix} \right|$ have modulus greater than 1

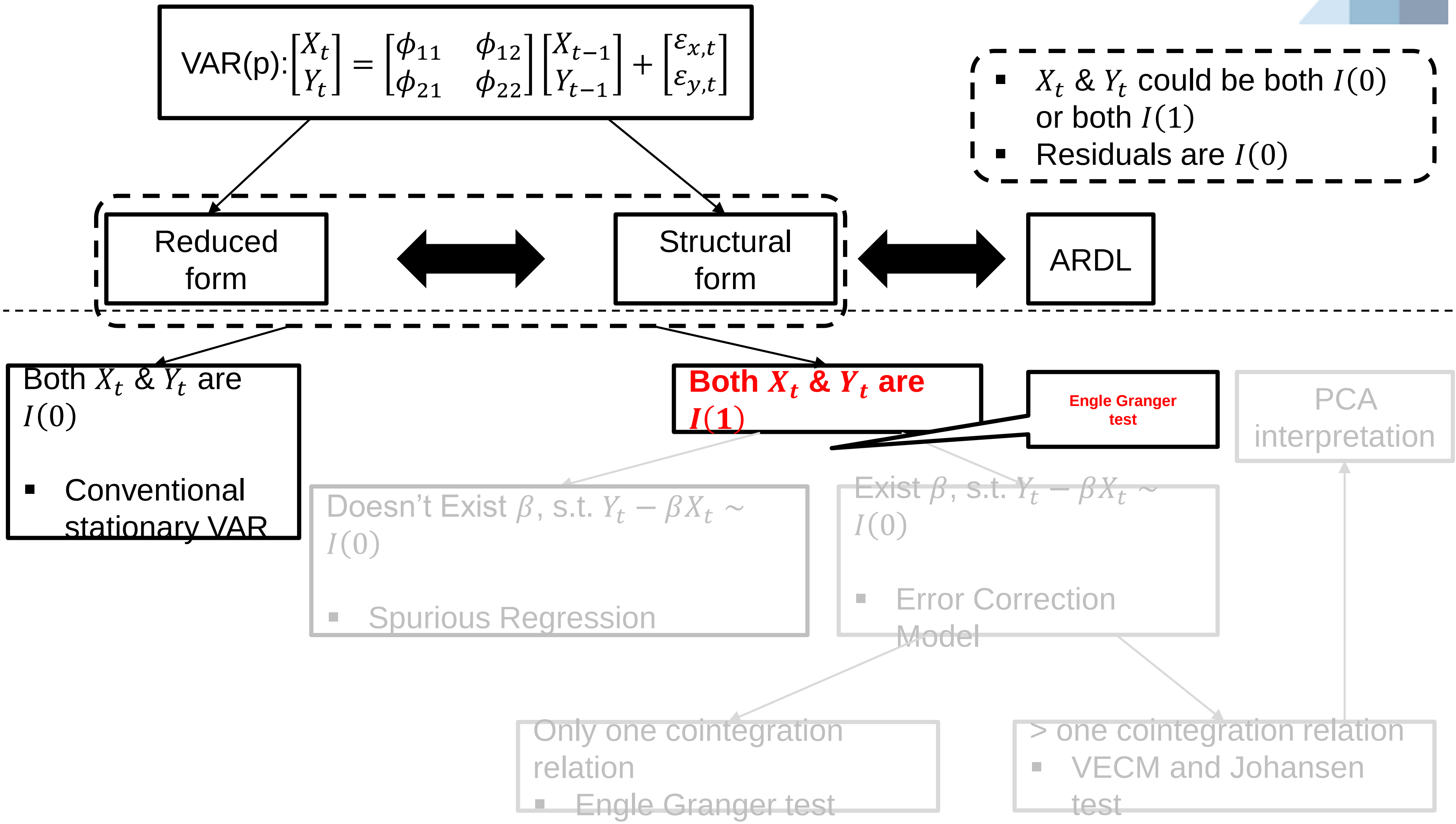
Fitting function of stationary VAR:

- Matlab:

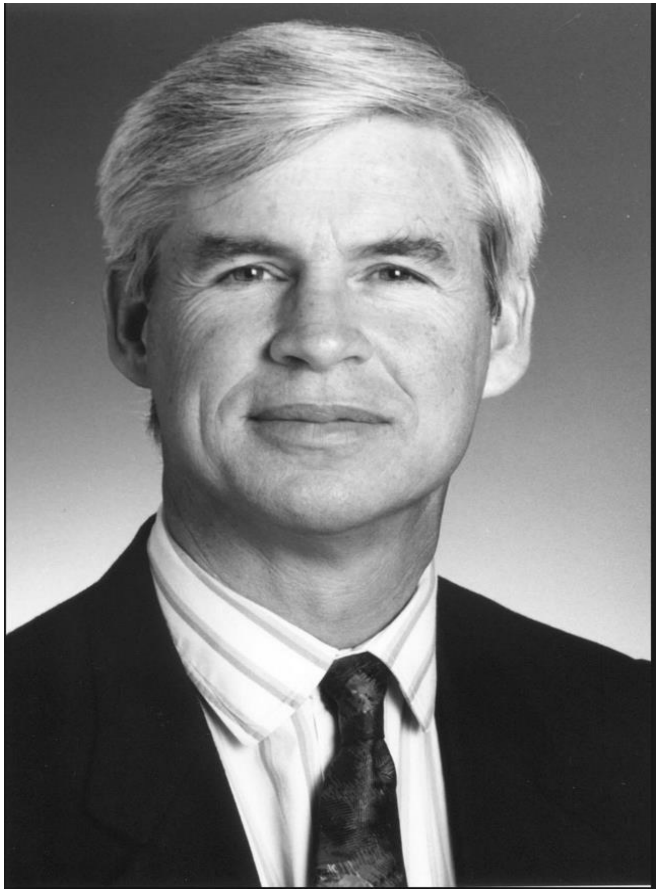
```
   Spec = vgxset('n',2,'nAR',1,'Constant',false);
estSpec = vgxvarx(Spec,Y);
```

- R Reduced form: `fit  = VAR(cbind(x,y), p = 1, type = 'none')`
- R Structural form: `fit2 = SVAR(fit,Amat = matrix(c(1,0,NA,1),2,2))` $\Leftarrow \mathbf{L} = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$

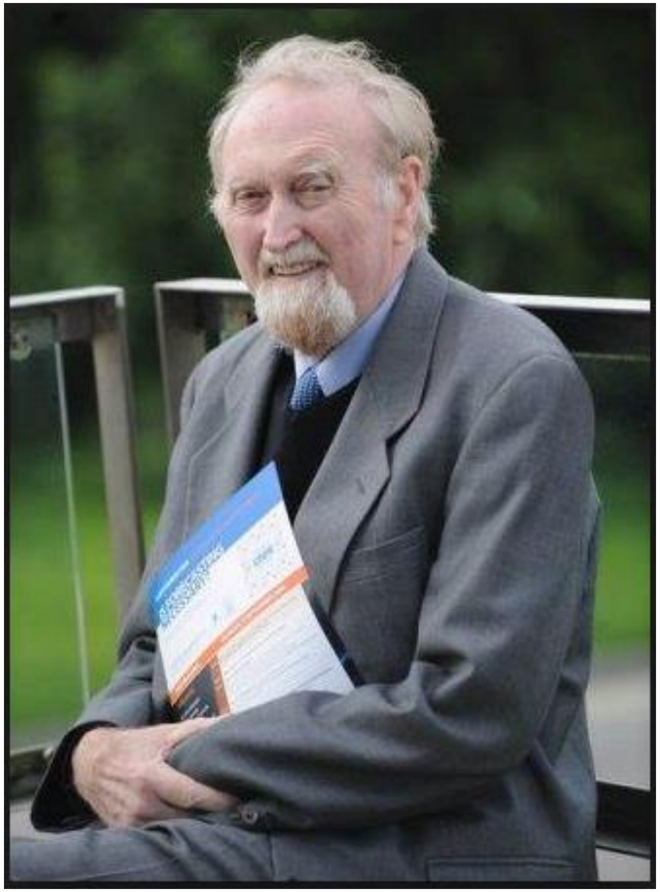
- SAS: `proc VARMAX;model x y / p=1 noint;`



VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
Engle Granger test
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test



Robert F. Engle (born in 1942)
American economist, currently teaches at New York University, Stern School of Business



Clive Granger (1934 – 2009)
British economist, taught at University of Nottinghan in Britain & University of California, San Diego in US

They shared Nobel Memorial Prize in Economic Sciences in 2003 for the methods of ARCH (Engle), and Co-integration (Granger) respectively.

Cointegration:

There are two $I(1)$ series, X_t and Y_t, if there exists β s.t. $Y_t - \beta X_t \sim I(0)$, we say X_t and Y_t are cointegrated, or there is one cointegrating relation between X_t and Y_t, and the state of $Y_t - \beta X_t \sim I(0)$ is called the long-run equilibrium between X_t and Y_t

Test cointegration:

- Engle Granger test
  - Fit OLS – $Y_t = \hat{\alpha} + \hat{\beta} X_t + \hat{Z}_t$
  - Test $\hat{Z}_t$ is I(0) or I(1)

How to test I(0) v.s. I(1):

Dickey-Fuller test & Augmented Dickey-Fuller test

	Dickey Fuller test
▪ If there's no drift or linear trend	$Z_t = \phi Z_{t-1} + u_t$, where u_t is white noise
$\begin{cases} H_0: \phi = 1 \\ H_1: \lvert\phi\rvert < 1 \end{cases}$	Under H_0, Z_t is a random walk, while under H_1, Z_t is stationary AR
	$Z_t = \phi Z_{t-1} + u_t \Leftrightarrow \nabla Z_t = (\phi - 1) Z_{t-1} + u_t = \gamma Z_{t-1} + u_t$
$\begin{cases} H_0: \phi = 1 \\ H_1: \lvert\phi\rvert < 1 \end{cases} \Leftrightarrow \begin{cases} H_0: \gamma = 0 \\ H_1: \gamma < 0 \end{cases}$	Under H_0, $t_\gamma = \frac{\hat{\gamma}_{OLS}}{SE(\hat{\gamma}_{OLS})}$ is not standard normal, and its limiting distribution is called Dickey-Fuller distribution, which doesn't have closed form representation
▪ To capture the nonzero mean	$Z_t = c + \phi Z_{t-1} + u_t$, where u_t is white noise
	$Z_t = c + \phi Z_{t-1} + u_t \Leftrightarrow \nabla Z_t = c + (\phi - 1) Z_{t-1} + u_t = c + \gamma Z_{t-1} + u_t$
▪ To capture the nonzero mean & linear trend	$Z_t = c + \delta t + \phi Z_{t-1} + u_t$, where u_t is white noise
	$Z_t = c + \delta t + \phi Z_{t-1} + u_t \Leftrightarrow \nabla Z_t = c + \delta t + (\phi - 1) Z_{t-1} + u_t = c + \delta t + \gamma Z_{t-1} + u_t$
	○ Nonzero c indicates existence of drift ○ Nonzero δ indicates existence of linear trend ○ $\gamma = 0$ means non-stationarity, and rejection of $\gamma = 0$ indicates stationarity of Z_t

Augmented Dickey Fuller test

- DF test assumes u_t is white noise, by adding enough lags of ∇Z_{t-j} on the right hand side, this property could be guaranteed approximately
- $\nabla Z_t = c + \delta t + \gamma Z_{t-1} + \Sigma_{j=1}^{p} \phi_j \nabla Z_{t-j} + u_t$

How to choose # of lags?

- Ng and Perron (1995): Set a upper bound p_{max}, then decrease p till ϕ_p is significant
- Schwert (1989): $p = \left[12 \cdot \left(\frac{T}{100}\right)^{1/4}\right]$, T is the number of data points
- AIC, BIC, etc.

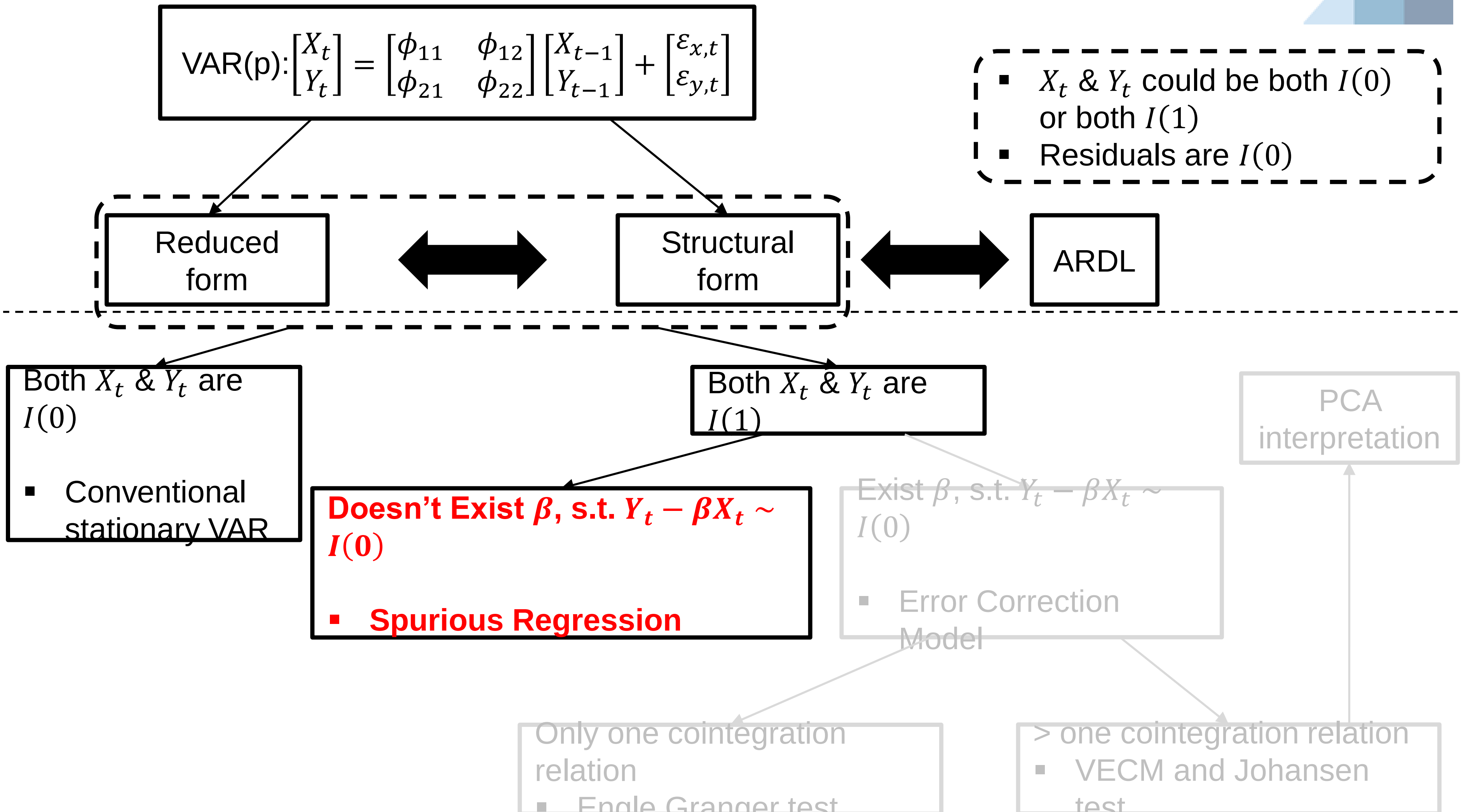
- Nonzero c indicates existence of drift
- Nonzero δ indicates existence of linear trend
- $\gamma = 0$ means non-stationarity, and rejection of $\gamma = 0$ indicates stationarity of Z_t

EG test

- Matlab:
```
[h,pvalue,stat,cValue,reg] = egcitest(Y_new,'test','t2','creg','nc');
```

- R: `fit5 = egcm(x,y)`

- SAS: `proc AUTOREG; model y = x/ nlag=1 stationarity =(adf=1);`


VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test

Spurious Regression

- EG test indicates there's no β s.t. $Y_t - \beta X_t \sim I(0)$

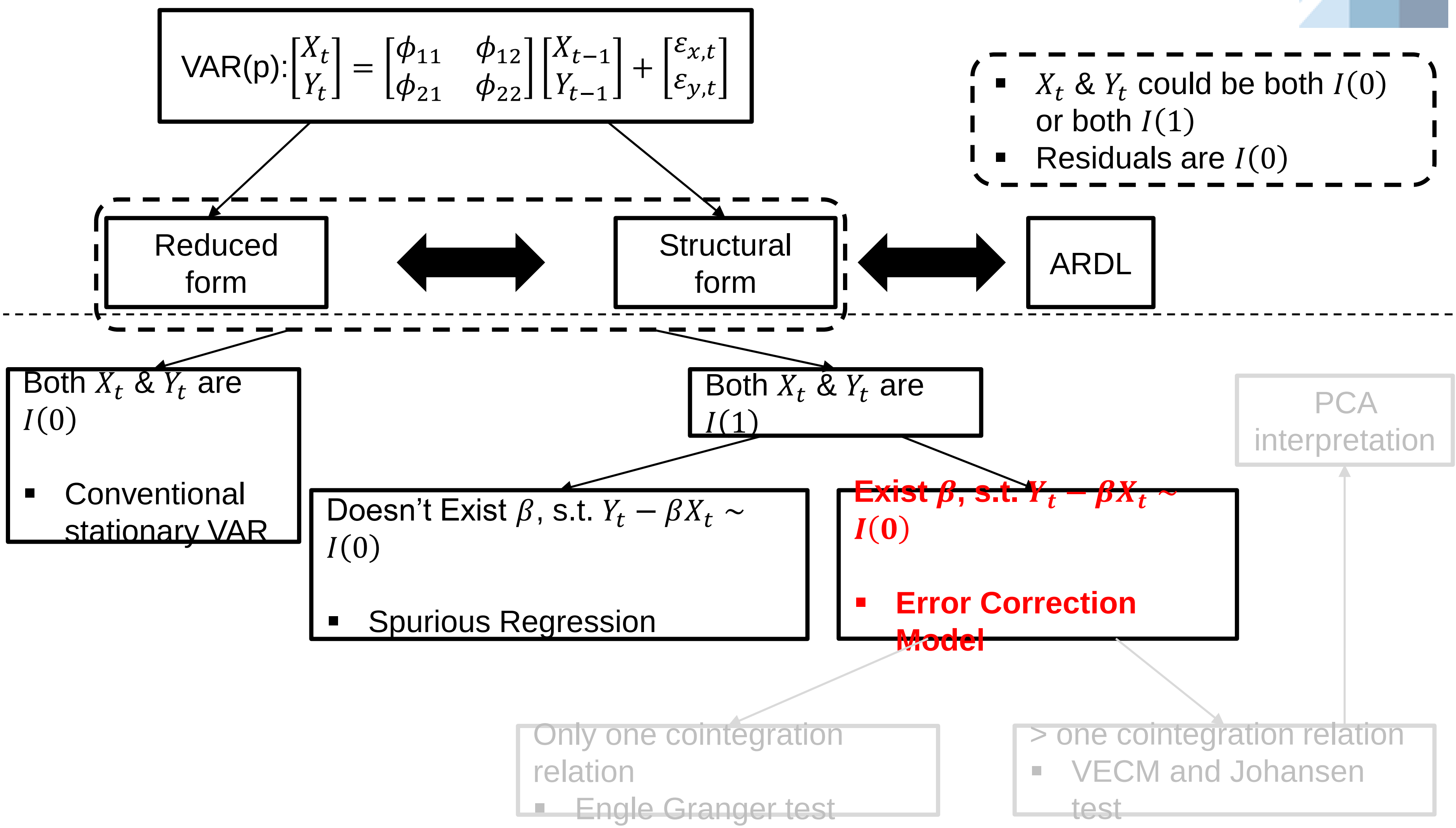
E.g. X_t: Monthly salary at Microsoft; Y_t: Population in an developing African city

- Two series are totally unrelated
- Increase in X_t & Y_t doesn't cause each other
- OLS of Y_t on X_t will generate high R^2 because $corr(X_t, Y_t)$ is high
- OLS is not reliable when spurious regression is present

Matlab:
```
ToEstMd = regARIMA(0,1,0);
ToEstMd.Intercept = 0;
EstMd = estimate(ToEstMd,Y(:,3),'X',Y(:,1),'Display','params');
```

R:
```
fit <- arima(z, c(0,1,0), xreg=x);
fit <- lm(diff(z) ~ diff(x));
```

SAS:
```
proc arima data=simul;
identify var=z(1) crosscorr=x(1);
estimate input=x;
run;
```



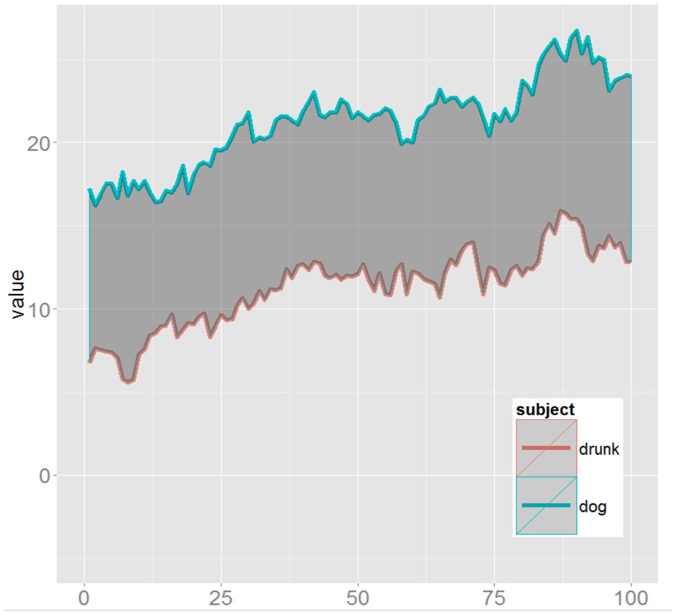
VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test

An example of cointegration

- **Drunk woman and a stray dog**

Drunk's position is a random walk along real line: $X_t = X_{t-1} + u_t$, where u_t is white noise
Her dog also wanders aimlessly as a random walk: $Y_t = Y_{t-1} + w_t$, where w_t is white noise

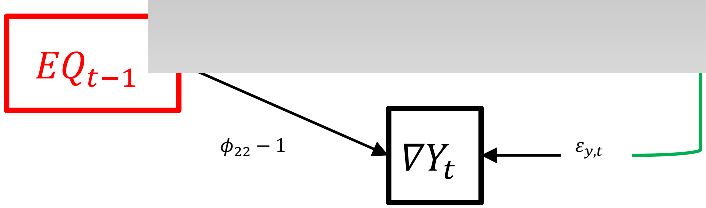
- **What if the dog belongs to the drunk?**
  - They wouldn't be far away from each other
  - Drunk's current position X_t is not only affected by her previous position X_{t-1}, but also affected her distance from her dog previously, i.e. $Y_{t-1} - X_{t-1}$
  - Dog's current position Y_t is not only affected by her previous position Y_{t-1}, but also affected her distance from her dog previously, i.e. $Y_{t-1} - X_{t-1}$

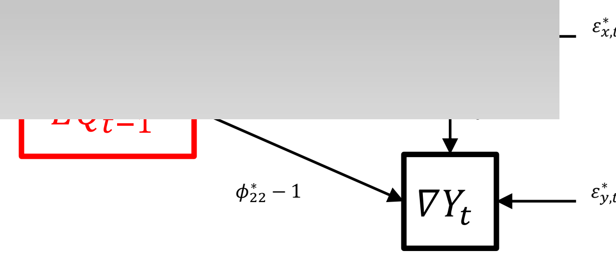


Error Correction Model (ECM)

- Reduced: $\begin{cases} X_t = \phi_{11} X_{t-1} + \phi_{12} Y_{t-1} + \varepsilon_{x,t} \\ Y_t = \phi_{21} X_{t-1} + \phi_{22} Y_{t-1} + \varepsilon_{y,t} \end{cases}$
- Structural: $\begin{cases} X_t = \phi^*_{11} X_{t-1} + \phi^*_{12} Y_{t-1} + \varepsilon^*_{x,t} \\ Y_t = cX_t + \phi^*_{21} X_{t-1} + \phi^*_{22} Y_{t-1} + \varepsilon^*_{y,t} \end{cases}$
- Substitut ... $\nabla X_t + X_{t-1}$
- $\begin{cases} \nabla X_t = \ldots \\ \nabla Y_t = (\ldots \end{cases}$ ... $\Big) + \varepsilon^*_{x,t}$ ... $X_{t-1}\Big) + \varepsilon^*_{y,t}$

- If roots of $\left| I - \begin{bmatrix} \phi_{11} B & \phi_{12} B \\ \phi_{21} B & \phi_{22} B \end{bmatrix} \right| = 0$ are all outside of unit circle – stationary
- $\left| I - \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \right| = 0$, meaning $(1 - \phi_{11})(1 - \phi_{22}) = \phi_{12}\phi_{21}$
- From which it could be proven that $\frac{\phi_{11}-1}{\phi_{12}} = \frac{\phi_{21}}{\phi_{22}-1} = \frac{\phi^*_{11}-1}{\phi^*_{12}} = \frac{\phi^*_{21}}{\phi^*_{22}-1}$



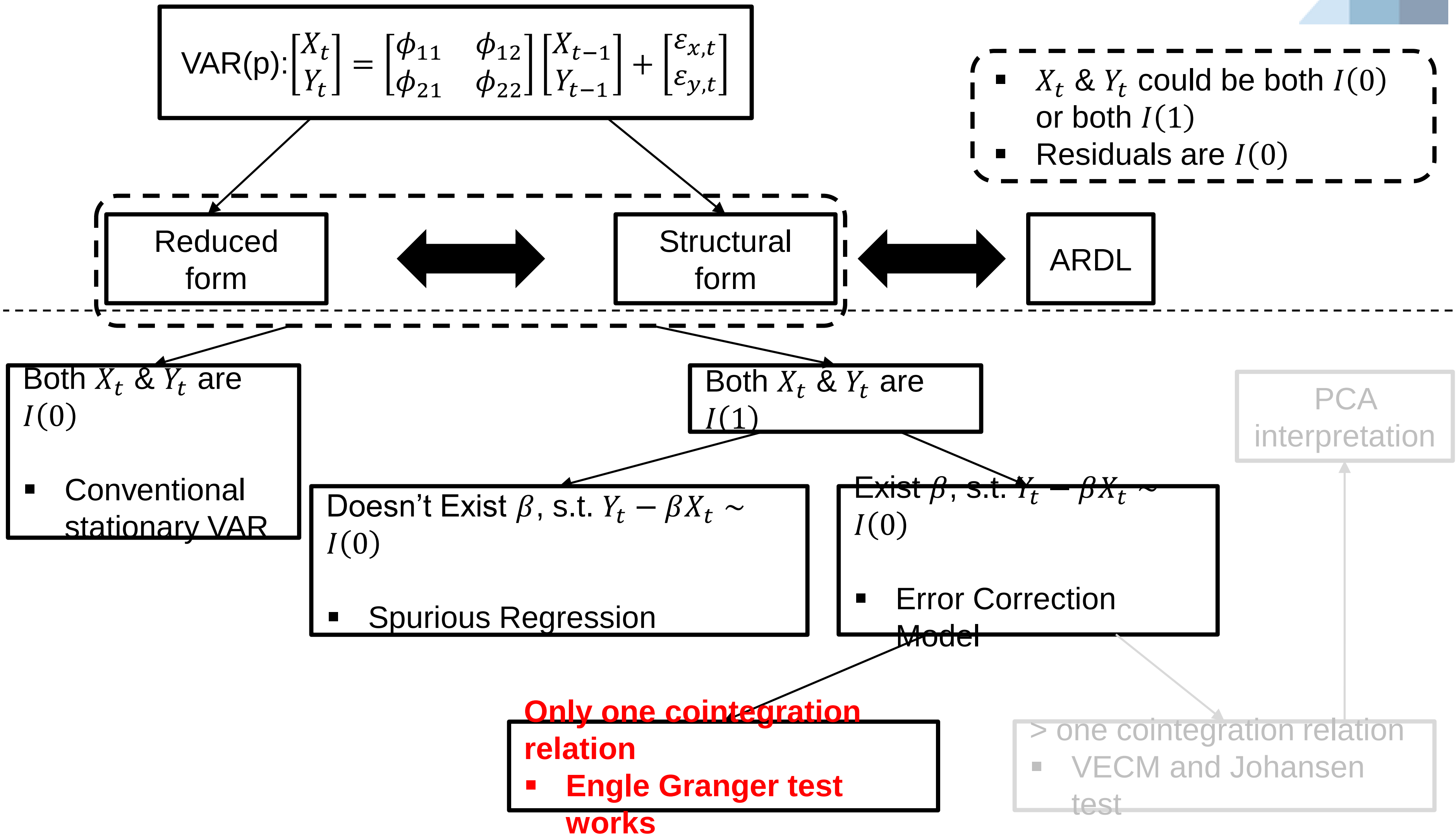


How to fit ECM model?

EG two-step fitting:

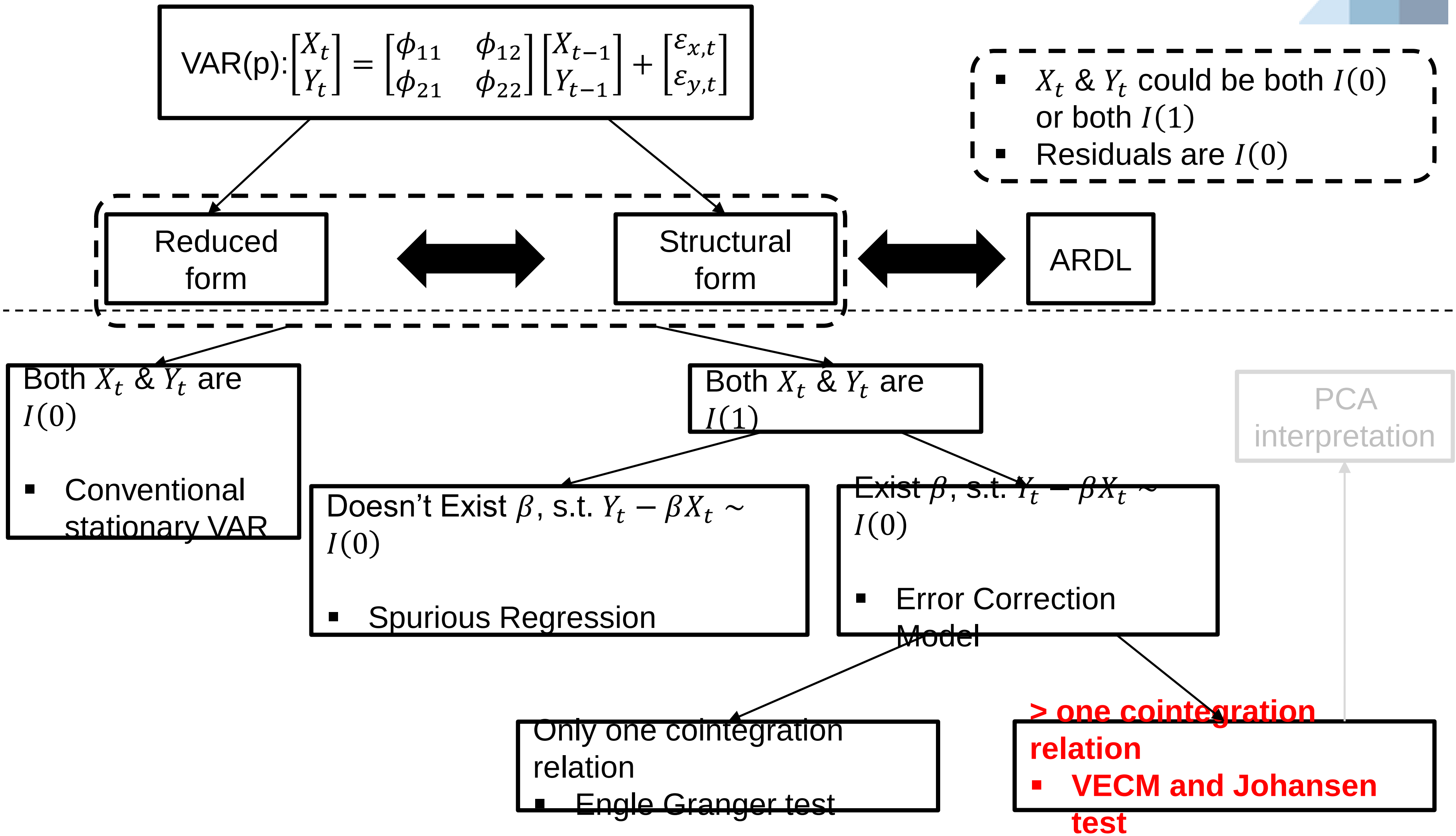
- Fit OLS: $Y_t = \hat{\alpha} + \hat{\beta} X_t + \hat{u}_t$
  (Since X_t & Y_t are cointegrated, $\hat{\alpha}$ & $\hat{\beta}$ are cointegrated)

- Fit OLS: $\nabla Y_t = \hat{\theta} \nabla X_t + \hat{\phi}\left(Y_{t-1} - \hat{\alpha} - \hat{\beta} X_{t-1}\right) + \hat{\varepsilon}_t$


VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
▪ X_t & Y_t could be both $I(0)$ or both $I(1)$
▪ Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
▪ Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
▪ Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
▪ Error Correction Model
Only one cointegration relation
▪ Engle Granger test works
> one cointegration relation
▪ VECM and Johansen test

What if there're more than two variables?

- There're X_{1t}, X_{2t} and X_{3t}
  - If we're confident that there wouldn't be more than one cointegration relation among them, i.e. #cointegration = 0 or 1
  - Engle Granger test:
    Fit OLS: $X_{3t} = \hat{\beta}_0 + \hat{\beta}_1 X_{1t} + \hat{\beta}_2 X_{2t} + \hat{u}_t$
    ADF test on $\hat{u}_t$
  - If there're two cointegration relations: i.e. $X_{3t} - \beta_1 X_{1t} - \beta_2 X_{2t} \sim I(0)$ & $X_{3t} - \beta_1' X_{1t} - \beta_2' X_{2t} \sim I(0)$, it means $(\beta_1 - \beta_1')X_{1t} + (\beta_2 - \beta_2')X_{2t} \sim I(0)$
  - There is strong multicollinearity between X_{1t} & X_{2t}
  - Thus…

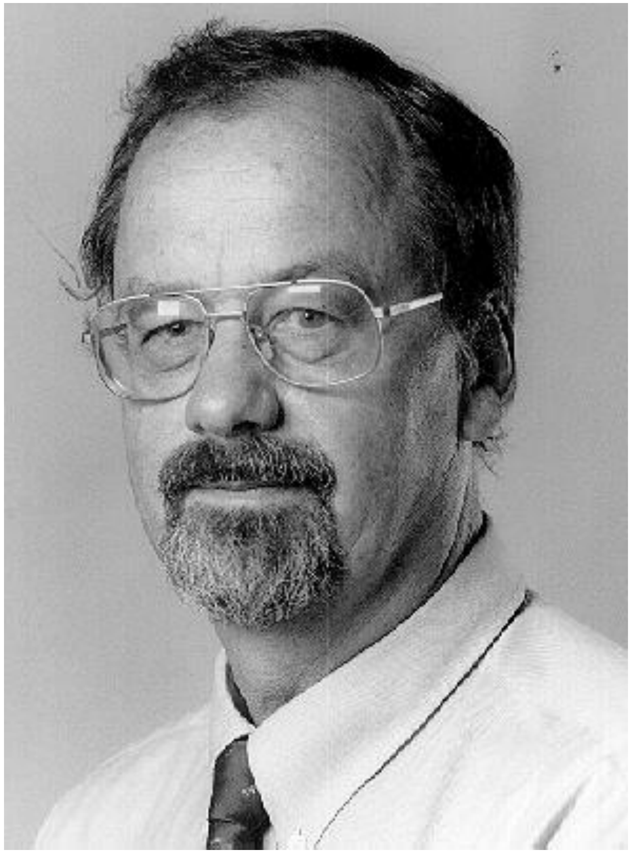

VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
▪ X_t & Y_t could be both $I(0)$ or both $I(1)$
▪ Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
▪ Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
▪ Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
▪ Error Correction Model
Only one cointegration relation
▪ Engle Granger test
> one cointegration relation
▪ VECM and Johansen test

VECM for Cointegration
Johansen Approach

Tingjun Ruan

Johansen Methodology (1988)

- The drawback of Engle Granger approach: it can only identify a ***single equilibrium*** relationship among the variables. If we have more than two variables in the model, then there is a possibility of having more than one cointegration relationships
- Johansen (1988) proposed a framework of estimating and testing of vector error correction model (VECM) *based on vector auto regressive (VAR)* equations with which we can find out how many cointegrating relationships exist among variables
- Johansen VECM is more general for testing ***multiple cointegrating relationships*** when there are more than two variables.
- For k variables, we can have up to k-1 cointegrations



Soren Johansen (born in 1939 Denmark)

Danish Statistician and Econometrician, who is known for his research in time series analysis, in particular the theory of cointegration in collaboration with Katarina Juselius.

He is currently a professor at the Department of Econometrics, University of Copenhagen.

Multiple Cointegration Example - A Drunk, Her Dog and A boyfriend

- Murry (1994) proposed the classic (one) cointegration relationship: A drunk and her dog who adjust their paths to avoid straying too far apart
- Smith and Harrison extended this classic example by including the third party -- the drunken boyfriend, to illustrate ***multiple*** cointegration and error correction

http://www-stat.wharton.upenn.edu/~steele/Courses/434/434Context/Co-integration/DrunkDogBoyfriendConintegration.pdf)

- Both the **drunk** $\boldsymbol{X_t}$ and **her dog** Oliver (unleashed) $\boldsymbol{Y_t}$ can be modeled by the random walk:

$$X_t - X_{t-1} = u_t$$
$$Y_t - Y_{t-1} = w_t$$

The **boyfriend** Kinley $\boldsymbol{Z_t}$ who does not hold drunk's hand:

$$Z_t - Z_{t-1} = v_t$$

$$X_t, Y_t, Z_t \sim I(1)$$

The drunk adjusts her walk to respond to the positions of both Oliver (her dog) and Kinley (her boyfriend)

Error Correction model:

$$\nabla X_t = \alpha(X_{t-1} - Y_{t-1}) + \beta(X_{t-1} - Z_{t-1}) + \gamma \nabla Y_t + \theta \nabla Z_t + \varepsilon_t$$

- The drunk's behavior is affected by two independent cointegrations, corresponding to two long-run relationship $X_t - Y_t$ and $X_t - Z_t$ respectively
- The values of adjustment/speed coefficient α and β reveal the strengths of her adjustments to Oliver and Kinley respectively; if $\alpha > \beta$, she exhibits the stronger attraction to the dog than the boyfriend
- With three variables it is not possible to have more than two cointegration relationships, otherwise they are all stationary or I(0) process which causes contradiction

VECM

- A k-dimentional Vector Auto Regressive (VAR) model with p lags

$$X_t = \Pi_1 X_{t-1} + \cdots + \Pi_p X_{t-p} + \varepsilon_t$$

Where X_t is a vector of k time series at time t and ε_t is a vector white noise with covariance matrix Σ_ε

- If the variables from the VAR are cointegrated, the VAR model presented can be equivalently re-written in the form of vector error correction model as shown below (Engel and Granger 1987):

$$\nabla X_t = \Pi \mathrm{X}_{t-1} + \Gamma_1 \nabla \mathrm{X}_{t-1} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_t$$

$$\Pi = \Pi_1 + \cdots + \Pi_p - I_k$$

$$\Gamma_k = -\sum_{j=k+1}^{p} \Pi_j \,, k = 1, \ldots, p-1$$

VECM

- A k-dimentional ***Vector Error Correction Model*** (VECM) for the VAR(p) process can be written as:

$$\nabla X_t = \Pi_{[k\times k]}\mathrm{X}_{\mathrm{t-1}} + \Gamma_1 \nabla \mathrm{X}_{\mathrm{t-1}} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_{\mathrm{t}}$$

- In the VECM, ∇X_t and its lag are all I(0)
- $\Pi_{[k\times k]}\mathrm{X}_{\mathrm{t-1}}$ is an error correction term;$\Pi_{[k\times k]}$ contains long run relationships
- $Rank\left(\Pi_{[k\times k]}\right)$ = number of cointegration
- If there are r co-integration vectors, $\Pi_{[k\times k]}$ can be expressed as $\Pi_{[k\times k]} = \alpha_{[k\times r]}\beta'_{[r\times k]}$
- $\boldsymbol{\alpha}$ contains the speed of adjustment parameters which interpreted as the weight with which each co-integration vector appears in a given equation
- $\boldsymbol{\beta}$ contains the coefficient of long-run relationship

$Rank(\Pi)$

- $\boldsymbol{rank(\Pi) = 0}$

$\Rightarrow \Pi = 0 \Rightarrow X_t \sim I(1)$ and not cointegrated

The VECM reduces to

$$\nabla X_t = \Gamma_1 \nabla X_{t-1} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_t$$

$\Rightarrow \nabla X_t$ follows a VAR(p-1) model

- $\boldsymbol{rank(\Pi) = k}$

$\Rightarrow \Pi$ has full rank$\Rightarrow X_t$ is stationary / I(0)

In this case we can simply analyze X_t directly

- $\mathbf{0} < \boldsymbol{rank}(\boldsymbol{\Pi}) = \boldsymbol{r} < \boldsymbol{k}$

There are r cointegration relations, Π can be written as $\mathbf{\Pi}_{[\mathbf{k}\times\mathbf{k}]} = \boldsymbol{\alpha}_{\mathbf{k}\times\mathbf{r}}\boldsymbol{\beta}'_{\mathbf{r}\times\mathbf{k}}$, where β has r linearly independent columns representing the cointegrating vectors

$$\nabla X_t = \alpha\beta' \mathrm{X}_{\mathrm{t}-1} + \Gamma_1 \nabla \mathrm{X}_{\mathrm{t}-1} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_{\mathrm{t}}$$

$\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$ is not unique since we have:

$$\alpha\beta' = \alpha H H^{-1}\beta' = (\alpha H)\left(\beta H^{-1'}\right)' = \alpha^*\beta^{*\prime}$$

Where $\alpha^* = \alpha H, \beta^* = \beta H^{-1'}$

It is common to require that $\beta' = [I_r, \beta_1{}']$, where I_r is the $r \times r$ identity matrix and β_1 is a $(k - r) \times r$ matrix

Determine the Number of Cointegration

- The rank of the long-run impact matrix Π equals the number of cointegration, and it is evaluated to determine the number of eigenvalues that are different from zero

- Johansen and Juselius (1990) proposed two methods for determining the number of cointegrating relations.
  - **Trace Statistics** based on a likelihood ratio test about a trace of the matrix
  - **Maximum Eigenvalue Statistics** based on eigenvalues (maximum) obtained from estimation procedure
  - These cointegration tests are formulated in term of the estimated eigenvalues $\hat{\lambda}_i$ of the matrix Π

Johansen's Trace Statistic

- The trace statistic considers whether the trace is increased by adding more eigenvalues beyond the r^{th} eigenvalue
- Consider the hypothesis: $\begin{cases} H_0: Rank(\Pi) = r_0 \\ H_1: Rank(\Pi) > r_0 \end{cases}$
- Likelihood ratio test statistics:

$$LR_{tr}(r_0) = -T \sum_{i=r_o+1}^{k} ln\left(1 - \widehat{\lambda}_i\right)$$

Note:

- If $rank(\Pi) = r_0, \text{then } \hat{\lambda}_{r_0+1}, \dots, \hat{\lambda}_k$ should all be close to zero, or $LR_{tr}(r_0)$ should be small that we cannot reject H_0 as $ln\left(1 - \widehat{\lambda}_i\right) \approx 0 \; as \; i > r_o$
- If $rank(\Pi) > r_0, then \; \hat{\lambda}_{r_0+1}, \dots, \hat{\lambda}_k$ should be non-zero but less than one, and $LR_{tr}(r_0)$ should be large as $ln\left(1 - \widehat{\lambda}_i\right) \ll 0 \; for \; some \; i > r_0$.

Johansen's Maximum Eigenvalue Statistic

- It is a sequential procedure to determine the number of cointegrations
- Consider the hypothesis: $\begin{cases} H_0: Rank(\Pi) = r_0 \\ H_1: Rank(\Pi) = r_0 + 1 \end{cases}$
- Test Statistic:

$$LR_{max}(r_0) = -Tln(1 - \hat{\lambda}_{r_0+1})$$

Note:

- The test consists of ordering the largest eigenvalues in descending order and considering whether they are significantly different from zero.
- Suppose we obtained n eigenvalues denoted $\lambda_1 > \lambda_2 > \cdots > \lambda_k$, if the variables under are not cointegrated, the rank of Π is zero and all the eigenvalues should be equal to zero, therefore, $\ln\left(1 - \widehat{\lambda}_i\right)$ equals to 0.

- The asymptotic null distribution of both tests are not Chi-square but instead a complicated function of Brownian motion and the critical values of the test statistics are non-standard
- Critical values for both statistics are provided by Johansen and Juselius (1990)

Maximum Likelihood Estimation of Cointegrated VECM

- Once we found the $rank(\Pi) = r$ $(0 < \mathrm{r} < \mathrm{n})$, then the cointegrated VECM can be written:

$$\nabla X_t = \alpha\beta' X_{t-1} + \Gamma_1 \nabla X_{t-1} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_t$$

Where ε_t is Gaussian white noise

- Reduced rank multivariate regression
- Johansen derived the maximum likelihood estimation of the parameters under the reduce rank restriction
- $\hat{\beta}_{mle} = (\hat{v}_1, \ldots, \hat{v}_r)$, where $\hat{v}_i$ are the eigenvectors associated with the eigenvalues $\hat{\lambda}_i$
- We first estimate β using MLE, plug in $\hat{\beta}$ then the MLEs of the remaining parameters can be estimated by OLS (since all the variables are I(0) process):

$$\nabla X_t = \alpha \boldsymbol{\hat{\beta}'_{mle}} X_{t-1} + \Gamma_1 \nabla X_{t-1} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_t$$

Note:

- Johansen approach estimates β using MLE instead of OLS which is used in E-G two step as if r (# of cointegration) > 1 may cause multicolinearity

Maximum likelihood estimation

$$\nabla X_t = \alpha\beta' X_{t-1} + \Gamma_1 \nabla X_{t-1} + \cdots + \Gamma_{p-1} \nabla X_{t-p+1} + \varepsilon_t$$

Consider two multivariate regression:

$$\nabla X_t = \Omega_1 \nabla X_{t-1} + \cdots + \Omega_{p-1} \nabla X_{\mathrm{t-p+1}} + \mathrm{u_t}$$

$$X_{t-1} = \Xi_1 \nabla X_{t-1} + \cdots + \Xi \nabla X_{t-p+1} + w_t$$

Let $\hat{u}_t$ and $\widehat{w}_t$ be the residuals of the two equation, Define sample covariance matrix:

$$S_{00} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_t \hat{u}_t{}' \quad S_{01} = \frac{1}{T-p} \sum_{t=p+1}^{T} \hat{u}_t \widehat{w}_t{}' \quad S_{11} = \frac{1}{T-p} \sum_{t=p+1}^{T} \widehat{w}_t \widehat{w}_t{}'$$

Solving the eigenvalue :

$$\left|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = 0$$

Denote the eigenvalue and eigenvector pairs by $(\hat{\lambda}_i, \hat{v}_i), \hat{\lambda}_1 > \cdots > \hat{\lambda}_k$

$$\hat{\beta}_{mle} = (\hat{v}_1, \ldots, \hat{v}_r)$$

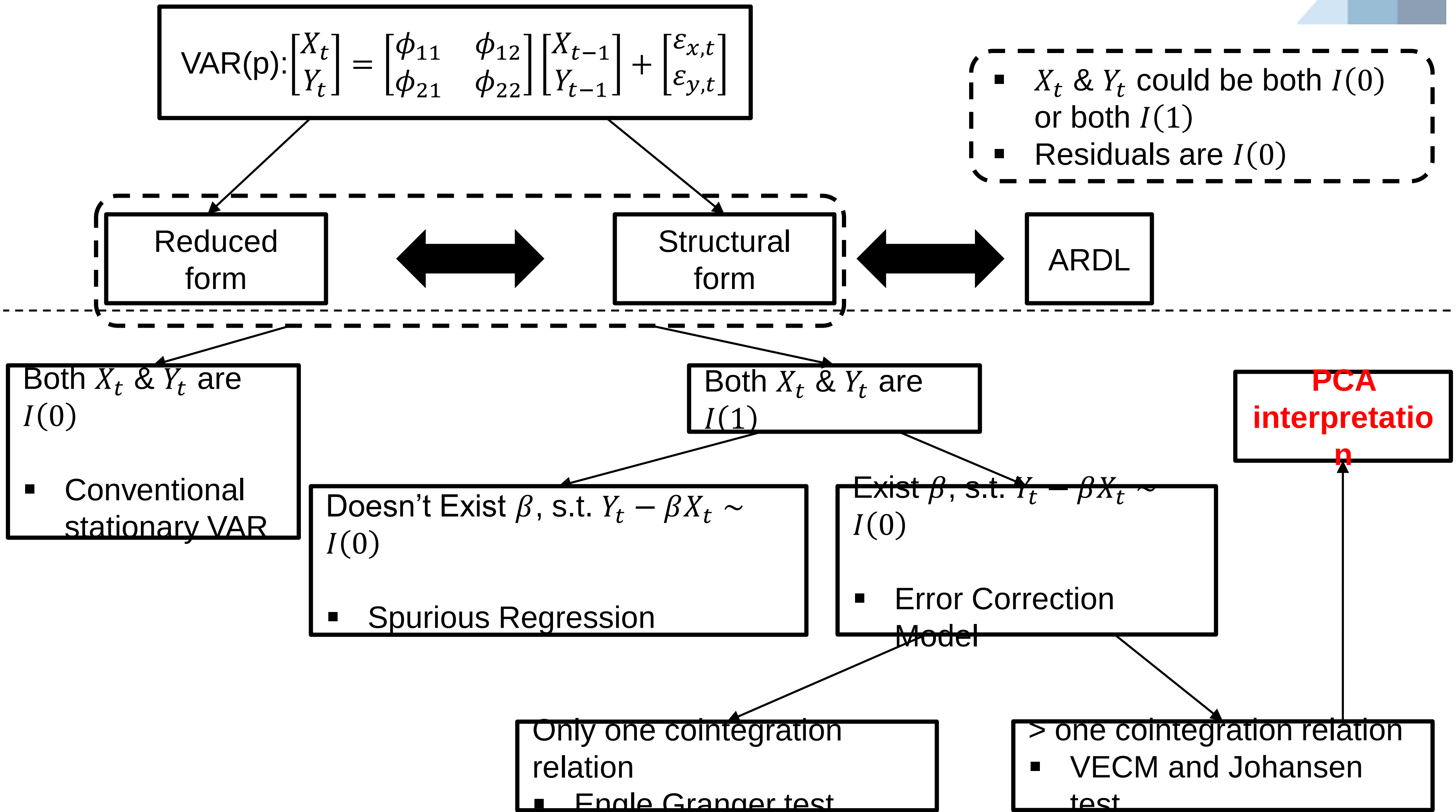
Functions of fitting VECM and Johansen test

- Matlab Johansen Procedure : `[h,pvalue,~,~,mles] = jcitest(Y,'Model','H2','lags',1,'display','params')`
- Matlab VECM to VAR : `VAR = vectovar('VEC',{A0 A1},'C',C)`
- R Johansen Procedure:
- R VECM to VAR:

```
fit = ca.jo(cbind(x,y,z),ecdet = 'none',K = 2,spec = 'transitory')
fit2 = vec2var(fit,r = 1)
```

- SAS Johansen Procedure :

```
proc VARMAX;
    model x y z / p=1 noint
cointtest=(johansen=(TYPE=TRACE))
            ecm = (rank=1)
            print = (iarr estimates);
run;
```



VAR(p): $\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix}$
X_t & Y_t could be both $I(0)$ or both $I(1)$
Residuals are $I(0)$
Reduced form
Structural form
ARDL
Both X_t & Y_t are $I(0)$
Conventional stationary VAR
Both X_t & Y_t are $I(1)$
PCA interpretation
Doesn't Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Spurious Regression
Exist β, s.t. $Y_t - \beta X_t \sim I(0)$
Error Correction Model
Only one cointegration relation
Engle Granger test
> one cointegration relation
VECM and Johansen test

Linear decomposition of cointegration

Ruofeng Wen

The key of time series regression: Trend

- Without $I(1)$, **everything** discussed before can be described by $VAR(1)$
- And $VAR(1)$ can be solved consistently and efficiently by (Generalized) Least Square, and thus have no difference with ordinary regression with possible serial correlation
- Deterministic trend (e.g. $\alpha + \beta t$) and stochastic trend (e.g. random walk) will cause spurious results in regression due to $I(1)$ residual.
- Trend can be eliminated through **differencing**, but **co-integration** analysis provides a better alternative when applicable.

- Stationary Series can be modelled with linear regression directly:

$$Y_t = \alpha + \beta X_t + u_t$$

$I(0)$ $I(0)$ $I(0)$

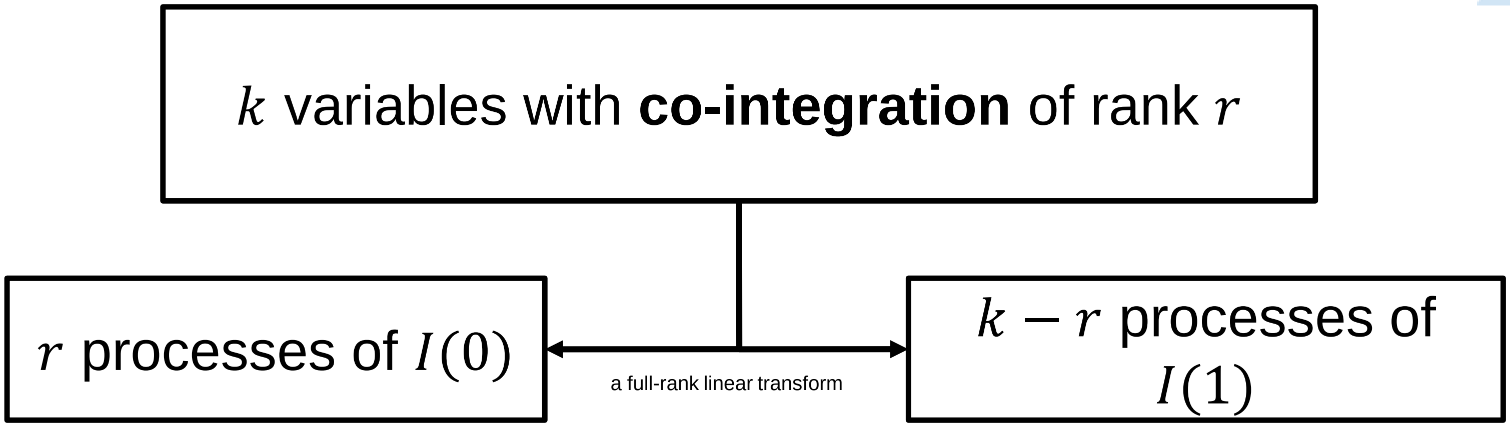
- Series with stochastic trend can be modelled by (order q) differencing

$I(1)$ $I(1)$ $I(?)$

$$Y_t = \alpha + \beta X_t + \epsilon_t$$

$$\Delta Y_t = \eta + \gamma \Delta X_t + u_t$$

$I(0)$ $I(0)$ $I(0)$


k variables with **co-integration** of rank r
r processes of $I(0)$
a full-rank linear transform
$k - r$ processes of $I(1)$

So what *is* co-integration?

- Consider Johansen Procedure of a VAR(1)

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \epsilon_t \Rightarrow \Delta\mathbf{x}_t = (\mathbf{A} - \mathbf{I})\mathbf{x}_{t-1} + \epsilon_t$$

Here one unit root denotes an zero eigenvalue in $(\mathbf{A} - \mathbf{I})$, or a $\lambda_i = 1$ eigenvalue in $\mathbf{A}$

$$\text{*Jordan} = \begin{pmatrix} \lambda_i & 1 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{pmatrix}$$

- For any diagonalizable* matrix $\mathbf{A}$, its eigen-decomposition gives

Does not change $I(0)$ or $I(1)$ status

$$\mathbf{A} = \mathbf{Q}^{-1}\mathbf{D}\mathbf{Q}$$

where $\mathbf{D} = diag(\lambda_1, \dots, \lambda_n)$ contains sorted eigenvalues of $\mathbf{A}$. Then we have

$$\boldsymbol{y}_t = \mathbf{D}\boldsymbol{y}_{t-1} + e_t$$

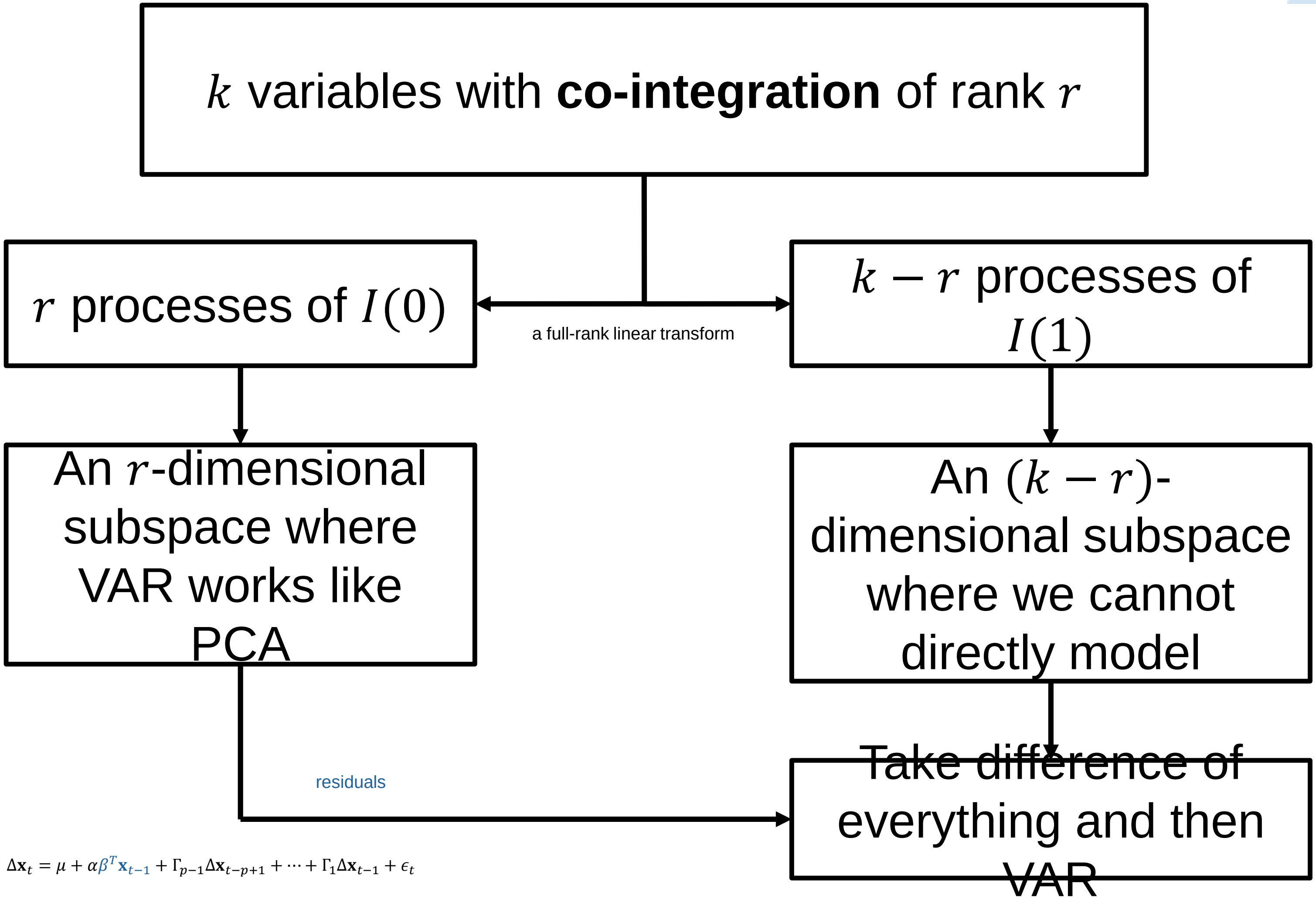
where $\boldsymbol{y}_t = \mathbf{Q}\mathbf{x}_t$. Note in Johansen if the rank of co-integration is r, then $(\mathbf{A} - \mathbf{I})$ is of rank r, namely the number of $\lambda_i = 1$ is actually $k - r$, and others less than 1. Thus:

$$y_{t,1} = y_{t-1,i} + e_{t,1}$$
$$\dots$$
$$y_{t,n-r} = y_{t-1,n-r} + e_{t,n-r}$$

$k - r$ variables $\sim I(1)$

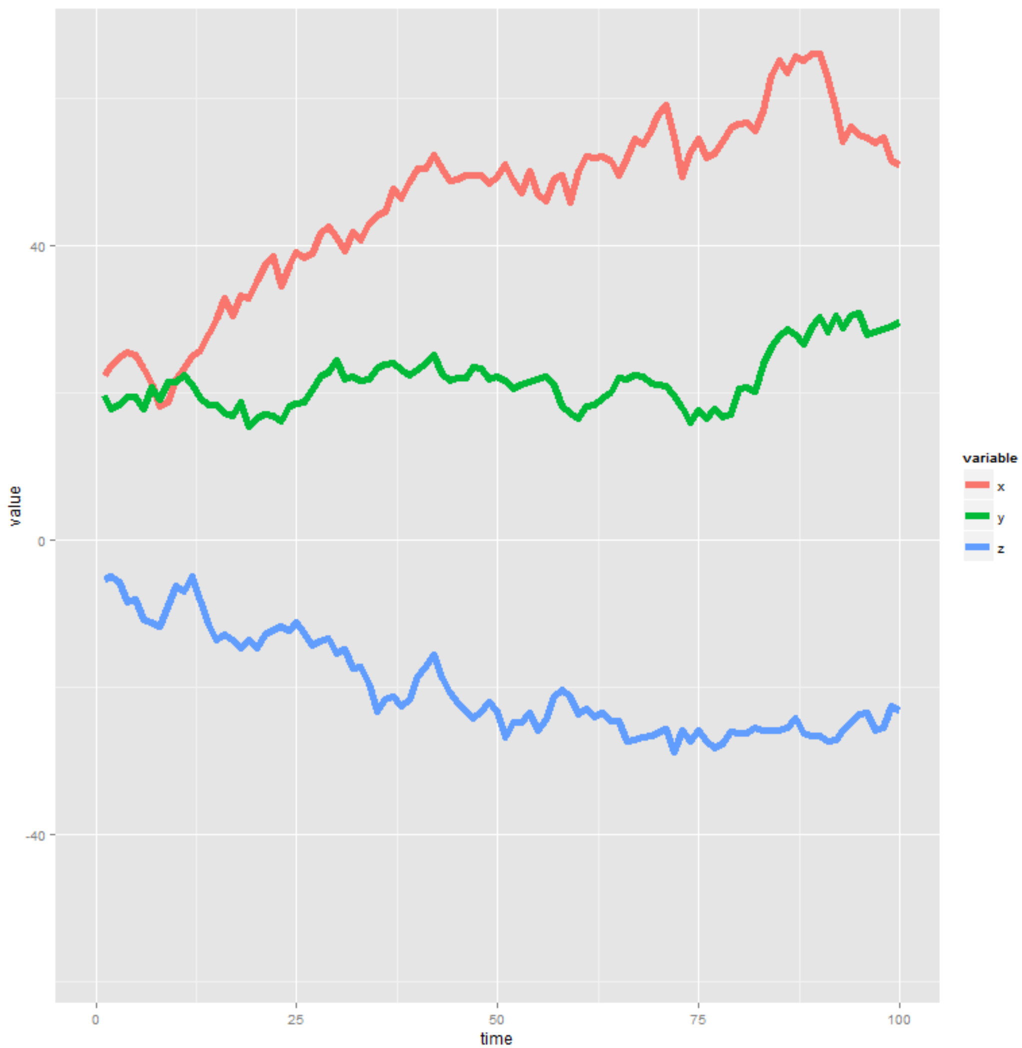
$$y_{t,n-r+1} = \lambda_{n-r+1} y_{t-1,i} + e_{t,n-r+1}$$
$$\dots$$
$$y_{t,n} = \lambda_n y_{t-1,n} + e_{t,n}$$

r variables $\sim I(0)$



$$\Delta \mathbf{x}_t = \mu + \alpha\beta^T\mathbf{x}_{t-1} + \Gamma_{p-1}\Delta\mathbf{x}_{t-p+1} + \cdots + \Gamma_1\Delta\mathbf{x}_{t-1} + \epsilon_t$$

Data Simulation



$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t$$

Data generation

independent

$$z_1(t) = z_1(t-1) + u_1(t) \qquad I(1)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t) \qquad I(0)$$
$$z_3(t) = z_3(t-1) + u_3(t) \qquad I(1)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

$$z_2(t) = \frac{2}{3} y_{t-1} - \frac{1}{3} x_{t-1}$$

ECM (E-G)

$$\nabla y_t = -\frac{1}{2}\left(\frac{2}{3} y_{t-1} - \frac{1}{3} x_{t-1}\right) + 0 \nabla y_{t-1} + \epsilon_{2t}$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

VECM (Johansen)

$$\begin{pmatrix} \nabla x \\ \nabla y \\ \nabla z \end{pmatrix}_t = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

$cor(\epsilon_1, \epsilon_2) \neq 0$ $\qquad$ $cor(\epsilon_1, \epsilon_2 - 0.2\epsilon_1) = 0$

Structural VAR

$$\begin{pmatrix} 1 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 - 0.2\epsilon_1 \\ \epsilon_3 \end{pmatrix}_t$$

ARDL (ARX)

$$y_t = 0.2 x_t + 0.6 y_{t-1} + e_t$$

Spurious

$$z_t = \beta x_t + e'_t \quad \rightarrow \quad \hat{\beta} \neq 0$$
$$\nabla z_t = \beta \nabla x_t + e'_t \quad \rightarrow \quad \hat{\beta} = 0$$

R analysis

Data generation

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

```r
require(reshape2)
require(ggplot2)
require(urca)
require(forecast)
require(vars)
require(egcm)

# Simulation
set.seed(7)
n = 1100
sigma = sqrt(c(1,1,3))

# underlying latent variables
# z1: I(1)
z1 = 0:(n - 1)
for (i in 2:n)
  z1[i] = z1[i - 1] + rnorm(1,sd = sigma[1])

# z2: ARIMA(1,0,0)
z2 = 0:(n - 1)
for (i in 2:n)
  z2[i] = 0.5 * z2[i - 1] + rnorm(1,sd = sigma[2])

# z3: I(1 )
z3 = 0:(n - 1)
for (i in 2:n)
  z3[i] = z3[i - 1] + rnorm(1,sd = sigma[3])

# rotation matrix A
A = matrix(c(2,1,0,-1,1,0,0,0,1),3,3)
temp = A %*% rbind(z1,z2,z3)
x = temp[1, 100:(n-1)]
y = temp[2, 100:(n-1)]
z = temp[3, 100:(n-1)]
```

R analysis

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_{\mathrm{t}}$$

$cor(\epsilon_1, \epsilon_2) \neq 0$

$cor(\epsilon_1, \epsilon_2 - 0.2\epsilon_1) = 0$

Structural VAR

$$\begin{pmatrix} 1 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 - 0.2\epsilon_1 \\ \epsilon_3 \end{pmatrix}_{\mathrm{t}}$$

```r
# VAR directly
fit3 = VAR(cbind(x,y,z), p = 1, type = 'n')
summary(fit3)

# strutural VAR
fit8 = SVAR(fit3,Amat = t(matrix(c(1,0,0,NA,1,0,0,0,1),3,3)))
```

R analysis

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

$cor(\epsilon_1, \epsilon_2) \neq 0$ $\qquad$ $cor(\epsilon_1, \epsilon_2 - 0.2\epsilon_1) = 0$

$$\begin{pmatrix} 1 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 - 0.2\epsilon_1 \\ \epsilon_3 \end{pmatrix}_t$$

```
VAR Estimation Results:
=========================

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

      Estimate Std. Error t value Pr(>|t|)
x.l1  0.844866   0.019799  42.672  < 2e-16 ***
y.l1  0.297653   0.039232   7.587  7.5e-14 ***
z.l1 -0.007070   0.005301  -1.334    0.183
---

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

      Estimate Std. Error t value Pr(>|t|)
x.l1  0.164695   0.012687   12.98   <2e-16 ***
y.l1  0.665679   0.025140   26.48   <2e-16 ***
z.l1 -0.003465   0.003397   -1.02    0.308
---

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

      Estimate Std. Error t value Pr(>|t|)
x.l1 -0.030157   0.014928  -2.020   0.0436 *
y.l1  0.050833   0.029581   1.718   0.0860 .
z.l1  0.992239   0.003997 248.271   <2e-16 ***
---
```

R analysis

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

$cor(\epsilon_1, \epsilon_2) \neq 0$ $\qquad$ $cor(\epsilon_1, \epsilon_2 - 0.2\epsilon_1) = 0$

Structural VAR

$$\begin{pmatrix} 1 & 0 & 0 \\ -0.2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 - 0.2\epsilon_1 \\ \epsilon_3 \end{pmatrix}_t$$

```
Warning message:
In SVAR(fit3, Amat = t(matrix(c(1, 0, 0, NA, 1, 0, 0, 0, 1), 3,  :
  Convergence not achieved after 10000 iterations.

SVAR Estimation Results:
========================
Estimated A matrix:
     x y z
x  1.0 0 0
y -0.9 1 0
z  0.0 0 1
```

SAS analysis

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

```
/* Vector Autoregressive*/
proc VARMAX data=simul;
        model x y z / p=1 noint;
run;
```

- `p=1:` specify the order of VAR to be 1 since a VAR(1) is fitted here ;
- `noint:` it indicates an intercept is not included in the model.

SAS analysis output

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

Model Parameter Estimates

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
x	AR1_1_1	0.84487	0.01980	42.67	0.0001	x(t-1)
	AR1_1_2	0.29765	0.03923	7.59	0.0001	y(t-1)
	AR1_1_3	-0.00707	0.00530	-1.33	0.1826	z(t-1)
y	AR1_2_1	0.16469	0.01269	12.98	0.0001	x(t-1)
	AR1_2_2	0.66568	0.02514	26.48	0.0001	y(t-1)
	AR1_2_3	-0.00347	0.00340	-1.02	0.3078	z(t-1)
z	AR1_3_1	-0.03016	0.01493	-2.02	0.0436	x(t-1)
	AR1_3_2	0.05083	0.02958	1.72	0.0860	y(t-1)
	AR1_3_3	0.99224	0.00400	248.27	0.0001	z(t-1)

Covariances of Innovations

Variable	x	y	z
x	4.90422	0.88950	-0.07509
y	0.88950	2.01376	-0.07400
z	-0.07509	-0.07400	2.78804

Matlab analysis

$z_1(t) = z_1(t-1) + u_1(t)$
$z_2(t)$
$= 0.5z_2(t-1) + u_2(t)$
$z_3(t) = z_3(t-1) + u_3(t)$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

```
Spec = vgxset('n',3,'nAR',1,'Constant',false);
estSpec = vgxvarx(Spec,Y);
vgxdisp(estSpec)
```

- `vgxset:` set VARMAX model specification parameters;
- `('n',3,'nAR',1,'Constant',false):`
  a 3-dimensional VAR(1) with no intercept is fitted;
- `vgxvarx:` estimate the VAR parameters;
- `vgxdisp:` display the fitted model parameters and statistics.

Matlab analysis output

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \mathbf{A} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t$$

VAR

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_t = \begin{pmatrix} \frac{5}{6} & \frac{1}{3} & 0 \\ \frac{1}{6} & \frac{2}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

```
Model  : 3-D VAR(1) with No Additive Constant
         Conditional mean is AR-stable and is MA-invertible
AR(1) Autoregression Matrix:
      0.844866        0.297653     -0.00706968
      0.164695        0.665679     -0.00346546
    -0.0301574       0.0508327        0.992239
Q Innovations Covariance:
       5.64198          1.2861       -0.220827
        1.2861         2.21713       -0.150859
     -0.220827       -0.150859         2.80505
```

R analysis

```
# ARX/ARDL (univariate use arima(y,xreg = x))
fit4 = VAR(cbind(y,z),p = 1,type = 'none',exogen = x)
summary(fit4)

# ARX with all I(1)
fit6 = arima(z,c(0,1,0),xreg = x)
summary(fit6)
fit7 = lm(diff(z) ~ diff(x))
summary(fit7)

# spurious regression
fit9 = lm(z ~ x)
summary(fit9)
```

$$y_t = 0.2x_t + 0.6y_{t-1} + e_t$$ ARDL (ARX)

Spurious

$$z_t = \beta x_t + e'_t \quad \rightarrow \quad \hat{\beta} \neq 0$$
$$\nabla z_t = \beta \nabla x_t + e'_t \quad \rightarrow \quad \hat{\beta} = 0$$

R analysis

```
Call:
lm(formula = diff(z) ~ diff(x))

Residuals:
    Min      1Q  Median      3Q     Max 
-6.1030 -1.0757  0.0537  1.0641  5.0542 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.018583   0.052993   0.351    0.726
diff(x)     -0.002451   0.023256  -0.105    0.916
---

Call:
lm(formula = z ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-33.145 -11.049   3.621  10.396  27.518 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.45272    0.84049   4.108 4.32e-05 ***
x           -0.67996    0.02298 -29.588  < 2e-16 ***
---
```

```
VAR Estimation Results:
========================= 

Estimation results for equation y: 
================================== 
y = y.l1 + z.l1 + exo1 

      Estimate Std. Error t value Pr(>|t|)    
y.l1  0.616969   0.022942  26.893   <2e-16 ***
z.l1 -0.002243   0.003259  -0.688    0.492    
exo1  0.190141   0.011582  16.418   <2e-16 ***
---
```

$$y_t = 0.2x_t + 0.6y_{t-1} + e_t$$

ARDL (ARX)

↕

Spurious

$$z_t = \beta x_t + e'_t \quad \rightarrow \quad \hat{\beta} \neq 0$$

$$\nabla z_t = \beta \nabla x_t + e'_t \quad \rightarrow \quad \hat{\beta} = 0$$

SAS analysis

```
/* ARDL(ARX) - Univariate*/
proc VARMAX data=simul;
                model y z = x / p=1 noint;
run;
```

$$y_t = 0.2x_t + 0.6y_{t-1} + e_t$$ ARDL (ARX)

```
/* ARDL(ARX) - Multivariate */
proc VARMAX data=simul;
                model y = x / p=1 noint;
run;
```

- `y z = x` : a VAR includes time series Y&Z with X as a exogenous variable is fitted;
- `p=1`: order of the VAR model, “xlag= “ if you want to include lags of X in the model.

SAS analysis output

ARDL(ARX)

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
y	XL0_1_1	0.19106	0.01150	16.61	0.0001	x(t)
	AR1_1_1	0.61782	0.02290	26.98	0.0001	y(t-1)

$$y_t = 0.2x_t + 0.6y_{t-1} + e_t$$ ARDL (ARX)

SAS analysis

```
/*Regression of two I(1) series*/
proc reg data=simul;
model dz = dx/noint;
run;
```

OR

```
proc arima data=simul;
identify var=z(1) crosscorr=x(1) ;
estimate input=x noint;
run;
```

```
/*Spurious Regression*/
proc reg data=simul;
model z = x;
run;
```

Spurious

$$z_t = \beta x_t + z_{t-1} + {e'}_t \quad \rightarrow \quad \hat{\beta} \neq 0$$

$$\nabla z_t = \beta \nabla x_t + {e'}_t \quad \rightarrow \quad \hat{\beta} = 0$$

- `var=z(1)`: time series z is modeled as response, the first order diffence of z is taken;
- `crosscorr=x(1)`: x is crosscorrelated with z and the first order difference of x is included in the model;
- `estimate input=x`: specify x is the input time series, and x must be listed in `crosscorr=`;
- Here **`Proc arima`** did the same fit as taking the first order difference of the two I(1) time series and then conduct OLS fit on them.

SAS analysis output

Spurious Regression

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	3.45272	0.84049	4.11	<.0001
x	1	-0.67996	0.02298	-29.59	<.0001

Regression of two I(1) series

Conditional Least Squares Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
NUM1	-0.0025931	0.02324	-0.11	0.9112	0	x	0

Spurious

$$z_t = \beta x_t + z_{t-1} + e'_t \quad \rightarrow \quad \hat{\beta} \neq 0$$

$$\nabla z_t = \beta \nabla x_t + e'_t \quad \rightarrow \quad \hat{\beta} = 0$$

Matlab analysis

```
%% ARDL multivariate
Spec = vgxset('n',2,'nAR',1,'nX',2,'Constant',false);
xx = cellfun(@(x) [x,0;0,x],...
    num2cell(Y(:,1)),'UniformOutput',false);
estSpec = vgxvarx(Spec,Y(:,2:3),xx,[]);
vgxdisp(estSpec)
```

$$y_t = 0.2x_t + 0.6y_{t-1} + e_t$$ ARDL (ARX)

```
%% ARDL univariate
Spec = vgxset('n',1,'nAR',1,'nX',1,'Constant',false);
estSpec = vgxvarx(Spec,Y(:,2),num2cell(Y(:,1)),[]);
vgxdisp(estSpec)
```

- When fit a ARDL with two response variable, we need to specify the number of exogenous variables to be 2 via (`'nX',2`) in `vgxset` although we only have one exogenous variable. Then `cellfun` is used to create a an m-by-1 cell (m is the number of observations) to specify the exogenous variable, and j-th (j=1,…,m) cell contains a 2-by-2 diagonal matrix with j-th element of exogenous variable on the diagonal.
- `vgxvarx`: estimate the VAR parameters;
- `vgxdisp`: display the fitted model parameters and statistics.

Matlab analysis

ARDL(ARX)

```
Model   : 1-D VARMAX(1,0,1) with No Additive Constant
          Conditional mean is AR-stable and is MA-invertible
b Regression Parameter:
      0.208368
AR(1) Autoregression Matrix:
      0.583531
Q Innovations Covariance:
       1.92822
```

$$y_t = 0.2x_t + 0.6y_{t-1} + e_t$$

ARDL (ARX)

Matlab analysis

```
%% Spurious Regression
fitlm(Y(:,1),Y(:,3))
```

Spurious

$$z_t = \beta x_t + z_{t-1} + e'_t \quad \rightarrow \quad \hat{\beta} \neq 0$$

$$\nabla z_t = \beta \nabla x_t + e'_t \quad \rightarrow \quad \hat{\beta} = 0$$

```
%% Regression of two I(1) time series
ToEstMd = regARIMA(0,1,0);
ToEstMd.Intercept = 0;
EstMd = estimate(ToEstMd,Y(:,3),'X',Y(:,1),'Display','params');
```

- `regARIMA:` build regression model with ARIMA error. In practice, it takes first order difference of the two series first and then do the model fit;
- `estimate:` estimate the parameters of the regression model. First argument specifies the type of the ARIMA model, and follows by the response data. `'X',Y(:,1)` specifies the input time series and `'Display','params'` would show the fitted results;
- `fitlm:` fit a linear regression model.

Matlab analysis

Spurious Regression

```
Estimated Coefficients:
                   Estimate       SE         tStat       pValue
                   ________    ________    ________    __________

    (Intercept)    -0.017239   0.048051    -0.35877       0.71984
    x1               0.11008   0.021087      5.2205    2.1718e-07
```

Spurious

$$z_t = \beta x_t + z_{t-1} + e'_t \quad \rightarrow \quad \hat{\beta} \neq 0$$

$$\nabla z_t = \beta \nabla x_t + e'_t \quad \rightarrow \quad \hat{\beta} = 0$$

Regression of two I(1) series

```
Regression with ARIMA(0,1,0) Error Model:
------------------------------------------
Conditional Probability Distribution: Gaussian

                                  Standard          t
 Parameter        Value           Error          Statistic
-----------    -----------    ------------    -----------
Intercept                0           Fixed          Fixed
    Beta1      -0.00259313       0.0241766      -0.107258
 Variance          2.79657        0.124087        22.5372
```

R analysis

Data generation

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t = \mathbf{A}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t$$

The $r = 1$ co-integration is

$$z_2(t) = \frac{2}{3} y_t - \frac{1}{3} x_t \sim I(0)$$

ECM (E-G)

$$\nabla y_t = -\frac{1}{2}\left(\frac{2}{3} y_{t-1} - \frac{1}{3} x_{t-1}\right) + 0 \nabla y_{t-1} + \epsilon_{2t}$$

VECM (Johansen)

$$\begin{pmatrix} \nabla x \\ \nabla y \\ \nabla z \end{pmatrix}_t = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

```r
# Engle-Granger
fit5 = egcm(x,y)
summary(fit5)

# johansen, 'r <= 1' hypothesis significant, indicating cointegration rank = 1
fit = ca.jo(cbind(x,y,z),ecdet = 'none',K = 2,spec = 'transitory')
summary(fit)

# johansen to VAR, agree with VAR?
fit2 = vec2var(fit,r = 1)
fit2
```

R analysis

Data generation

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5 z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t = \mathbf{A}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t$$

The $r = 1$ co-integration is

$$z_2(t) = \frac{2}{3} y_t - \frac{1}{3} x_t \sim I(0)$$

ECM (E-G)

$$\nabla y_t = -\frac{1}{2}\left(\frac{2}{3} y_{t-1} - \frac{1}{3} x_{t-1}\right) + 0 \nabla y_{t-1} + \epsilon_{2t}$$

VECM (Johansen)

$$\begin{pmatrix} \nabla x \\ \nabla y \\ \nabla z \end{pmatrix}_t = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

```
Y[i] =   0.4918 X[i] +   0.3336 + R[i], R[i] =   0.5221
R[i-1] + eps[i], eps ~ N(0,  1.5242^2)
        (0.0031)        (0.7706)
(0.0271)

R[1000] = -1.4247 (t = -0.800)

######################
# Johansen-Procedure #
######################

Eigenvalues (lambda):
[1] 1.855745e-01 8.725927e-03 3.742696e-05
```

```
Values of test statistic and critical value of
test:

           test 10pct  5pct  1pct
r <= 2 |   0.04  6.50  8.18 11.65
r <= 1 |   8.75 12.91 14.90 19.19
r = 0  | 204.86 18.90 21.07 25.75

Eigenvectors, normalised to first column:
              x.l1      y.l1      z.l1
x.l1  1.000000000 1.0000000  1.000000
y.l1 -2.008998832 0.8679812  1.066943
z.l1 -0.001280153 1.4539880 -1.417780
```

SAS analysis

Data generation

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t = \mathbf{A}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t$$

The $r = 1$ co-integration is

$$z_2(t) = \frac{2}{3}y_t - \frac{1}{3}x_t \sim I(0)$$

ECM (E-G)

$$\nabla y_t = -\frac{1}{2}\left(\frac{2}{3}y_{t-1} - \frac{1}{3}x_{t-1}\right) + 0\nabla y_{t-1} + \epsilon_{2t}$$

```
proc AUTOREG data = simul;
   model y = x/ nlag=1 stationarity =(adf);
run;
```

- `nlag=1` : specifies the order of the autoregressive error process or the subset of autoregressive error lags to be fitted;
- `stationarity =(adf)`: produces the augmented Dickey-Fuller unit root test. And if a regressor is presented, E-G test will be carried out to test the cointegration. The adf option is only available in recent SAS versions (after SAS 9.2).

SAS analysis

Data generation

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t = \mathbf{A}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t$$

The $r = 1$ co-integration is

$$z_2(t) = \frac{2}{3}y_t - \frac{1}{3}x_t \sim I(0)$$

ECM (E-G)

$$\nabla y_t = -\frac{1}{2}\left(\frac{2}{3}y_{t-1} - \frac{1}{3}x_{t-1}\right) + 0\nabla y_{t-1} + \epsilon_{2t}$$

Estimates of Autoregressive Parameters			
Lag	Coefficient	Standard Error	t Value
1	-0.517473	0.027100	-19.09

Engle-Granger Cointegration Test			
Type	Lags	Tau	Pr < Tau
Single Mean	1	-14.9642	<.0010
Trend	1	-14.9100	<.0010

Rejection of null hypothesis indicates existence of cointegration.

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	0.3336	0.1140	2.93	0.0035
x	1	0.4918	0.003118	157.73	<.0001

SAS analysis

```
proc VARMAX data=simul;
   model x y z / p=1 noint lagmax=1
                 cointtest=(johansen=(TYPE=TRACE))
                                   ecm=(rank=1)
                         print=(iarr estimates);
run;
```

VECM (Johansen)

$$\begin{pmatrix} \nabla x \\ \nabla y \\ \nabla z \end{pmatrix}_t = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

- `p=1` specifies the order of the vector autoregressive process;
- `lagmax=1` specifies the maximum number of lags for which results are computed and displayed by the PRINT= option;
- `cointtest=(johansen=(TYPE=TRACE))` specifies johansen test to test for cointegration, here `TYPE=TRACE` prints the cointegration trace test which is the default and `TYPE=MAX` prints the cointegration maximum eigenvalue test;
- `ecm=(rank=1)` is used to fit the VECM model after we know the number of cointegration from the johansen test, `rank=1` specifies the number of cointegration is 1 in our case;
- `print=(iarr estimates)` is used to print the reparameterized VAR coefficient matrices for the VCEM model via `iarr`. And `estimates` will print the coefficients estimates and the significance of the esitmeates.

SAS Output

Cointegration Rank Test Using Trace						
H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process
0	0	0.2434	285.7178	24.08	NOINT	Constant
1	1	0.0066	7.1336	12.21		
2	2	0.0005	0.5354	4.14		

Johansen test indicates a cointegration of order 1.

Infinite Order AR Representation				
Lag	Variable	x	y	z
1	x	0.84862	0.30159	-0.00033
	y	0.16673	0.66782	0.00037
	z	-0.02715	0.05409	0.99994

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
D_x	AR1_1_1	-0.15138	0.01964			x(t-1)
	AR1_1_2	0.30159	0.03914			y(t-1)
	AR1_1_3	-0.00033	0.00004			z(t-1)
D_y	AR1_2_1	0.16673	0.01258			x(t-1)
	AR1_2_2	-0.33218	0.02507			y(t-1)
	AR1_2_3	0.00037	0.00003			z(t-1)
D_z	AR1_3_1	-0.02715	0.01482			x(t-1)
	AR1_3_2	0.05409	0.02953			y(t-1)
	AR1_3_3	-0.00006	0.00003			z(t-1)

VECM in VAR form

VECM (Johansen)

$$\begin{pmatrix} \nabla x \\ \nabla y \\ \nabla z \end{pmatrix}_t = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{pmatrix} \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{t-1} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}_t$$

Matlab analysis

Data generation

$$z_1(t) = z_1(t-1) + u_1(t)$$
$$z_2(t) = 0.5z_2(t-1) + u_2(t)$$
$$z_3(t) = z_3(t-1) + u_3(t)$$

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}_t = \mathbf{A}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t$$

The $r = 1$ co-integration is

$$z_2(t) = \frac{2}{3}y_t - \frac{1}{3}x_t \sim I(0)$$

ECM (E-G)

$$\nabla y_t = -\frac{1}{2}\left(\frac{2}{3}y_{t-1} - \frac{1}{3}x_{t-1}\right) + 0\nabla y_{t-1} + \epsilon_{2t}$$

```
Y_new = [Y(:,2) Y(:,1)];
[h,pvalue,stat,cValue,reg] = ...  egcitest(Y_new,'test','t2','creg','nc');

table(h, pvalue, reg.coeff,...
'VariableName',{'h','pvalue','coefficient'})
```

```
ans =

    h       pvalue     coefficient
  _____    _______    ___________

  true      0.001       0.49972
```

- `'test','t2'`: indicates z-test (t1) or t-test (t2);
- `'creg','nc'`: indicates no constant or trend in the EG test.
- `h = true`: indicates the existence of cointegration.

Matlab analysis

VECM (Johansen)

$$\begin{pmatrix}\nabla x\\ \nabla y\\ \nabla z\end{pmatrix}_t = \begin{pmatrix}\frac{1}{2}\\ -\frac{1}{2}\\ 0\end{pmatrix}\begin{pmatrix}-\frac{1}{3} & \frac{2}{3} & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}_{t-1} + \begin{pmatrix}\epsilon_1\\ \epsilon_2\\ \epsilon_3\end{pmatrix}_t$$

```
[h,pvalue,~,~,mles] = ... jcitest(Y,'Model','H2','lags',0,...
'display','full')
```

r	h	stat	cValue	pValue	eigVal
0	1	212.1266	24.2747	0.0010	0.1853
1	0	7.6481	12.3206	0.3002	0.0071
2	0	0.5129	4.1302	0.6598	0.0005

Johansen test indicates a cointegration of order 1.

```
r = 1

A =
   -0.5398
    0.5945
   -0.0968

B =
    0.2804
   -0.5587
    0.0006
```

```
>> A*B'

ans =

   -0.1514    0.3016   -0.0003
    0.1667   -0.3321    0.0004
   -0.0271    0.0541   -0.0001
```

- `'Model','H2'`: no intercepts or trends in the cointegrating relations and there are no trends in the data. This model is appropriate if all series have zero mean;
- `'lags',0` : indicates the number of lagged differences in VECM. Here the original series is VAR(1), thus the number of lagged difference in VECM is 0. It's different from the `p=` option when fit VECM using SAS.

Reference and source

- Conceptual Econometrics Using R (ISSN Book 41) 1st Edition, by Hrishikesh D. Vinod (Editor)
- Principles of Macroeconometric Modeling (Volume 36) (Advanced Textbooks in Economics, Volume 36) by L.R. Klein, W. Welfe, et al. | Oct 5, 1999
- Macroeconomic Modeling and Macroeconometric Simulation: Illustrated with a developing economy Model (Macroeconometric model Book 1) Book 1 of 1: Macroeconometric model | by Kannapiran Arjunan | Jun 9, 2020
- Global and National Macroeconometric Modelling: A Long-Run Structural Approach by Anthony Garratt, Kevin Lee, et al. | May 4, 2012
- Simulation of a macroeconometric model with multiple time series considerations (Wayne economic papers) by Rosemary Rossiter | Jan 1, 1982