

Analysis of Multiple Time Series

Kevin Sheppard

<http://www.kevinsheppard.com>

Oxford MFE

This version: February 14, 2010

February 18, 2010

Today's Material

- Vector Autoregressions
 - ▶ Basic examples
 - ▶ Properties
 - ▷ Stationarity
 - ▶ Revisiting univariate ARMA processes
 - ▶ Forecasting
 - ▶ Granger Causality
 - ▶ Impulse Response functions
- Cointegration
 - ▶ Examining long-run relationships
 - ▶ Determining whether a VAR is cointegrated
 - ▶ Error Correction Models
 - ▶ Testing for Cointegration
 - ▷ Engle-Granger

Lots of revisiting univariate time series.

Why VAR analysis?

- Stationary VARs
 - ▶ Determine whether variable feedback into one another
 - ▶ Improve forecasts
 - ▶ Model the effect of a shock in one series to another
 - ▶ Differentiate between short run and long run dynamics
- Cointegration
 - ▶ Link random walks
 - ▶ Uncover long run relationships
 - ▶ Improve medium to long term forecasting **a lot**

VAR Defined

- Pth order autoregression, AR(P):

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_P y_{t-p} + \epsilon_t$$

- Pth order vector autoregression, VAR(P):

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_P \mathbf{y}_{t-p} + \epsilon_t$$

where \mathbf{y}_t and ϵ_t are k by 1 vectors

- Bivariate VAR(1):

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Compactly expresses two linked models:

$$y_{1,t} = \phi_{01} + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \epsilon_{1,t}$$

$$y_{2,t} = \phi_{02} + \phi_{21} y_{1,t-1} + \phi_{22} y_{2,t-1} + \epsilon_{2,t}$$

Stationarity Revisited

- Stationarity is a statistically meaningful form of regularity.

A stochastic process $\{y_t\}$ is covariance stationary if

$$E[y_t] = \mu \quad \forall t$$

$$V[y_t] = \sigma^2 \quad \sigma^2 < \infty \forall t$$

$$E[(y_t - \mu)(y_{t-s} - \mu)] = \gamma_s \quad \forall t, s$$

- AR(1) stationarity: $y_t = \phi y_{t-1} + \epsilon_t$
 - ▶ $|\phi| < 1$
 - ▶ ϵ_t is white noise
- AR(P) stationarity: $y_t = \phi_1 y_{t-1} + \dots + \phi_P y_{t-P} + \epsilon_t$
 - ▶ Roots of $(z^P - \phi_1 z^{P-1} - \phi_2 z^{P-2} - \dots - \phi_{P-1} z - \phi_P)$ less than 1
 - ▶ ϵ_t is white noise
- **No dependence on t**

- AR(1)

$$\begin{aligned}y_t &= \phi_0 + \phi_1 y_{t-1} + \epsilon_t \\&= \phi_0 + \phi_1(\phi_0 + \phi_1 y_{t-2} + \epsilon_{t-1}) + \epsilon_t \\&= \phi_0 + \phi_1 \phi_0 + \phi_1^2 y_{t-2} + \phi_1 \epsilon_{t-1} + \epsilon_t \\&= \phi_0 + \phi_1 \phi_0 + \phi_1^2(\phi_0 + \phi_1 y_{t-3} + \epsilon_{t-2}) + \phi_1 \epsilon_{t-1} + \epsilon_t \\&= \phi_0 \sum_{i=0}^{\infty} \phi_1^i + \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i} \\&= (1 - \phi_1)^{-1} \phi_0 + \sum_{i=0}^{\infty} \phi_1^i \epsilon_{t-i}\end{aligned}$$

- VAR(1)

$$\begin{aligned} \mathbf{y}_t &= \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t \\ &= \Phi_0 + \Phi_1 (\Phi_0 + \Phi_1 \mathbf{y}_{t-2} + \epsilon_{t-1}) + \epsilon_t \\ &= \Phi_0 + \Phi_1 \Phi_0 + \Phi_1^2 \mathbf{y}_{t-2} + \Phi_1 \epsilon_{t-1} + \epsilon_t \\ &= \Phi_0 + \Phi_1 \Phi_0 + \Phi_1^2 (\Phi_0 + \Phi_1 \mathbf{y}_{t-3} + \epsilon_{t-2}) + \Phi_1 \epsilon_{t-1} + \epsilon_t \\ &= \sum_{i=0}^{\infty} \Phi_1^i \Phi_0 + \sum_{i=0}^{\infty} \Phi_1^i \epsilon_{t-i} \\ &= (\mathbf{I}_k - \Phi_1)^{-1} \Phi_0 + \sum_{i=0}^{\infty} \Phi_1^i \epsilon_{t-i} \end{aligned}$$

Properties of a VAR(1) and AR(1)

$$\text{AR}(1) : y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

$$\text{VAR}(1) : \mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t$$

	AR(1)	VAR(1)
Mean	$\phi_0 / (1 - \phi_1)$	$(\mathbf{I}_k - \Phi_1)^{-1} \Phi_0$
Variance	$\sigma^2 / (1 - \phi_1^2)$	$(\mathbf{I} - \Phi_1 \otimes \Phi_1)^{-1} \text{vec}(\Sigma)$
s^{th} Autocovariance	$\gamma_s = \phi_1^s V[y_t]$	$\Gamma_s = \Phi_1^s V[\mathbf{y}_t]$
$-s^{\text{th}}$ Autocovariance	$\gamma_{-s} = \phi_1^s V[y_t]$	$\Gamma_{-s} = V[\mathbf{y}_t] \Phi_1^{s'}$

Autocovariances of vector processes are not symmetric, but $\Gamma_s = \Gamma_{-s}'$

- Stationarity

- ▶ AR(1): $|\phi_1| < 1$
- ▶ VAR(1): $|\lambda_i| < 1$ where λ_i are the eigenvalues of Φ_1

Stock and Bond VAR

- VWM from CRSP
- 10 Year bond *returns* from FRED (H.15)
- May 1953 until December 2009 (680 months)

$$\begin{bmatrix} VWM_t \\ 10YR_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11,1} & \phi_{12,1} \\ \phi_{21,1} & \phi_{22,1} \end{bmatrix} \begin{bmatrix} VWM_{t-1} \\ 10YR_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Market model:

$$VWM_t = \phi_{01} + \phi_{11,1} VWM_{t-1} + \phi_{12,1} 10YR_{t-1} + \epsilon_{1,t}$$

- Long bond model

$$10YR_t = \phi_{01} + \phi_{21,1} VWM_{t-1} + \phi_{22,1} 10YR_{t-1} + \epsilon_{2,t}$$

- Estimates

$$\begin{bmatrix} VWM_t \\ 10YR_t \end{bmatrix} = \begin{bmatrix} 9.733 \\ (0.000) \\ 1.058 \\ (0.000) \end{bmatrix} + \begin{bmatrix} 0.097 & 0.301 \\ (0.104) & (0.000) \\ -0.095 & 0.299 \\ (0.000) & (0.000) \end{bmatrix} \begin{bmatrix} VWM_{t-1} \\ 10YR_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

Campbell's VAR

- Campbell's (1996) asset pricing paper
- State variables include:
 - ▶ VWM: Value weighted market from CRSP
 - ▶ LBR: Real labor income growth from FRED
 - ▶ RTB: Relative T-bill from French's site
 - ▶ TERM: 10 year minus 3 month from FRED
 - ▶ DIV: Dividend yield from CRSP

	Raw Data				
	VWM_{t-1}	LBR_{t-1}	RTB_{t-1}	$TERM_{t-1}$	DIV_{t-1}
VWM_t	0.073 (0.937)	0.668 (0.000)	-0.050 (0.260)	-0.000 (1.000)	0.183 (0.000)
LBR_t	0.002 (0.001)	-0.164 (0.292)	0.002 (0.518)	0.000 (0.999)	-0.060 (0.000)
RTB_t	0.130 (0.000)	0.010 (0.121)	0.703 (0.688)	-0.010 (0.534)	0.137 (0.000)
$TERM_t$	-0.824 (0.000)	-2.888 (0.000)	0.069 (0.062)	0.960 (0.916)	4.028 (0.000)
DIV_t	0.001 (0.811)	-0.000 (1.000)	-0.001 (0.999)	-0.000 (0.737)	-0.045 (0.108)

Interpreting Estimates

	Standardized Series				
	VWM_{t-1}	LBR_{t-1}	RTB_{t-1}	$TERM_{t-1}$	DIV_{t-1}
VWM_t	0.073 (0.037)	0.112 (0.011)	-0.113 (0.010)	-0.011 (0.337)	0.007 (0.886)
LBR_t	0.012 (0.882)	-0.164 (0.000)	0.027 (0.472)	0.065 (0.004)	-0.013 (0.717)
RTB_t	0.057 (0.086)	0.001 (0.991)	0.703 (0.000)	-0.119 (0.000)	0.002 (0.964)
$TERM_t$	-0.029 (0.778)	-0.017 (0.624)	0.006 (0.851)	0.960 (0.000)	0.005 (0.911)
DIV_t	0.024 (0.640)	-0.000 (0.982)	-0.043 (0.012)	-0.043 (0.679)	-0.045 (0.108)

VAR(P) is really a VAR(1)

- Companion form:

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \epsilon_t$$

- Reform into a single VAR(1) where $\boldsymbol{\mu} = E[\mathbf{y}_t]$

$$\mathbf{z}_t = \boldsymbol{\Upsilon} \mathbf{z}_{t-1} + \boldsymbol{\xi}_t$$

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{y}_t - \boldsymbol{\mu} \\ \mathbf{y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{y}_{t-P+1} - \boldsymbol{\mu} \end{bmatrix}, \quad \boldsymbol{\Upsilon} = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \dots & \Phi_{P-1} & \Phi_P \\ \mathbf{I}_k & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_k & \mathbf{0} \end{bmatrix}$$

- ▶ All results can be directly applied to the companion form.
- ▶ Can also be used to transform AR(P) into VAR(1)

Revisiting Univariate Forecasting

- Consider standard AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

- Optimal 1-step ahead forecast:

$$\begin{aligned} E_t[y_{t+1}] &= E_t[\phi_0] + E_t[\phi_1 y_t] + E_t[\epsilon_{t+1}] \\ &= \phi_0 + \phi_1 y_t + 0 \end{aligned}$$

- Optimal 2-step ahead forecast:

$$\begin{aligned} E_t[y_{t+2}] &= E_t[\phi_0] + E_t[\phi_1 y_{t+1}] + E_t[\epsilon_{t+2}] \\ &= \phi_0 + \phi_1 E_t[y_{t+1}] + 0 \\ &= \phi_0 + \phi_1(\phi_0 + \phi_1 y_t) \\ &= \phi_0 + \phi_1 \phi_0 + \phi_1^2 y_t \end{aligned}$$

- Optimal h -step ahead forecast:

$$E_t[y_{t+h}] = \sum_{i=0}^{h-1} \phi_1^i \phi_0 + \phi_1^h y_t$$

Forecasting with VARs

- Identical to univariate case

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t$$

- Optimal 1-step ahead forecast:

$$\begin{aligned} E_t[\mathbf{y}_{t+1}] &= E_t[\Phi_0] + E_t[\Phi_1 \mathbf{y}_t] + E_t[\epsilon_{t+1}] \\ &= \Phi_0 + \Phi_1 \mathbf{y}_t + \mathbf{0} \end{aligned}$$

- Optimal h-step ahead forecast:

$$\begin{aligned} E_t[\mathbf{y}_{t+h}] &= \Phi_0 + \Phi_1 \Phi_0 + \dots + \Phi_1^{h-1} \Phi_0 + \Phi_1^h \mathbf{y}_t \\ &= \sum_{i=0}^{h-1} \Phi_1^i \Phi_0 + \Phi_1^h \mathbf{y}_t \end{aligned}$$

- Higher order forecast can be recursively computed

$$\begin{aligned} \mathbf{y}_t &= \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \dots + \Phi_P \mathbf{y}_{t-P} + \epsilon_t \\ E_t[\mathbf{y}_{t+h}] &= \Phi_0 + \Phi_1 E_t[\mathbf{y}_{t+h-1}] + \dots + \Phi_P E_t[\mathbf{y}_{t+h-P}] \end{aligned}$$

What makes a good forecast?

- Forecast residuals

$$\hat{e}_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$$

- Residuals are *not* white noise
- Can contain an MA($h - 1$) component
 - ▶ Forecast error for $y_{t+1} - \hat{y}_{t+1|t-h+1}$ was not known at time t .
- Plot your residuals
- Residual ACF
- Mincer-Zarnowitz regressions
- Three period procedure
 - ▶ Training sample: Used to build model
 - ▶ Validation sample: Used to refine model
 - ▶ Evaluation sample: Ultimate test, ideally 1 shot

Multi-step Forecasting

- Two methods
- Iterative method
 - ▶ Build model for 1-step ahead forecasts

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \epsilon_t$$

- ▶ Iterate forecast out to period h

$$\hat{\mathbf{y}}_{t+h|t} = \sum_{i=0}^{h-1} \Phi_1^i \Phi_0 + \Phi_1^h \mathbf{y}_t$$

- ▶ Makes efficient use of information
 - ▶ Imposes a lot of structure on the problem

- Direct Method

- ▶ Build model for h -step ahead forecasts

$$\mathbf{y}_t = \Phi_0 + \Phi_h \mathbf{y}_{t-h} + \epsilon_t$$

- ▶ Directly forecast using a pseudo 1-step ahead method

$$\hat{\mathbf{y}}_{t+h|t} = \Phi_0 + \Phi_h \mathbf{y}_t$$

- ▶ Robust to some nonlinearities

Multi-step Forecast Evaluation

- Multistep forecast evaluation is identical to one-step ahead forecast evaluation with one caveat
- h -step ahead forecast errors may be correlated with any forecast error not known at time t

$$\hat{e}_{t+1|t-h+1}, \hat{e}_{t+2|t-h+2}, \dots, \hat{e}_{t+h-1|t-1}$$

- Leads to a MA($h - 1$) structure in the forecast errors
- Solutions:
 - ▶ Use regular GMZ regression with a Newey-West covariance estimator

$$y_{t+h} - \hat{y}_{t+h|t} = \beta_1 + \beta_2 \hat{y}_{t+h|t} + \boldsymbol{\gamma} \mathbf{x}_t + \eta_t$$

$$H_0 : \beta_1 = \beta_2 = \boldsymbol{\gamma} = 0, H_1 : \beta_1 \neq 0 \cup \beta_2 \neq 0 \cup \boldsymbol{\gamma}_j \neq 0 \exists j$$

- ▶ Explicitly model the MA($h - 1$) and use a standard covariance estimator

$$y_{t+h} - \hat{y}_{t+h|t} = \beta_1 + \beta_2 \hat{y}_{t+h|t} + \boldsymbol{\gamma} \mathbf{x}_t + \eta_t + \sum_{i=1}^{h-1} \theta_i \eta_{t-i}$$

Note: Null is the same; does not impose a restriction on $\boldsymbol{\theta}$

Stock and Bond VAR

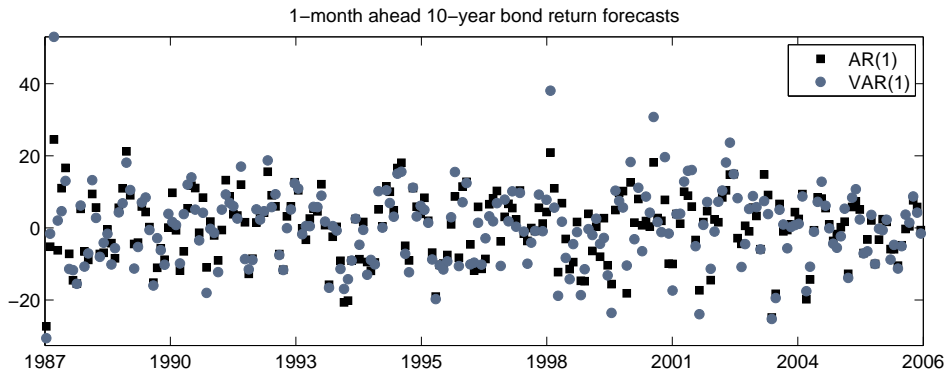
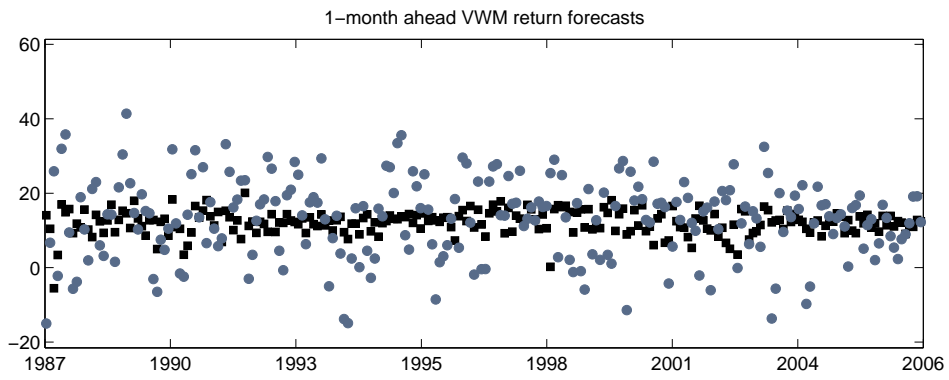
- Estimates from VAR

$$\begin{aligned} VWM_t &= 9.733 & +0.097 & VWM_{t-1} & +0.301 & 10YR_{t-1} \\ & (0.00) & (0.10) & & (0.00) & \\ 10YR_t &= 1.058 & -0.095 & VWM_{t-1} & +0.229 & 10YR_{t-1} \\ & (0.00) & (0.00) & & (0.00) & \end{aligned}$$

- Estimates from AR

$$\begin{aligned} VWM_t &= 9.92 & +0.088 & VWM_{t-1} \\ & (0.00) & (0.06) & \\ 10YR_t &= .073 & & +0.306 & 10YR_{t-1} \\ & (0.83) & & (0.00) & \end{aligned}$$

Comparing AR and VAR forecasts



Estimation and Identification

- Univariate Identification: Box-Jenkins
 - ▶ Use ACF and PACF to determine AR and MA lag order
 - ▶ Examine residuals
 - ▶ Parsimony principle
- The autocorrelation of a scalar process is defined

$$\rho_s = \frac{\gamma_s}{\gamma_0}$$

where γ_s is s^{th} the autocovariance

- ▶ Regression coefficient:

$$y_t = \mu + \rho_s y_{t-s} + \epsilon_t$$

- Partial autocorrelation ψ_s
 - ▶ Regression interpretation of s^{th} partial autocorrelation:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{s-1} y_{t-s+1} + \psi_s y_{t-s} + \epsilon_t$$

- ▶ ψ is the s^{th} partial autocorrelation

ACF and CCF

- Multivariate equivalents
 - ▶ ACF and PACF have same regression definitions
 - ▶ Cross-correlation function

$$\rho_{xy,s} = \frac{E[(x_t - \mu_x)(y_{t-s} - \mu_y)]}{\sqrt{V[x_t]V[y_t]}}$$

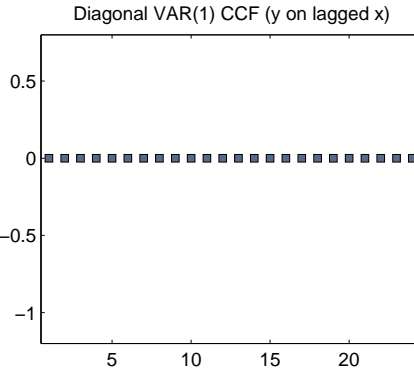
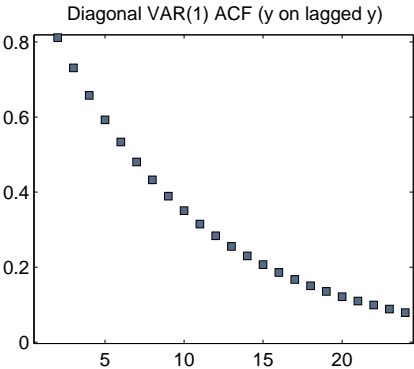
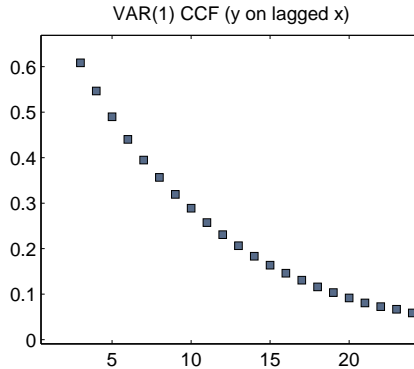
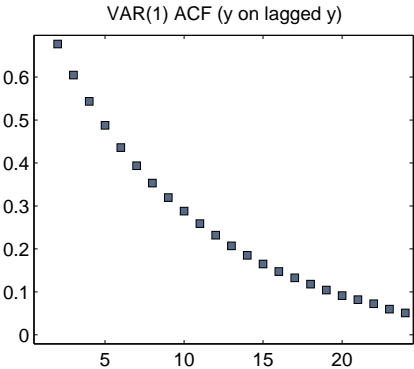
$$\rho_{yx,s} = \frac{E[(y_t - \mu_y)(x_{t-s} - \mu_x)]}{\sqrt{V[x_t]V[y_t]}}$$

- ▶ Generally different
- ▶ Cross-partial-correlation function $\psi_{xy,s}$

$$x_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{s-1} y_{t-s+1} + \psi_{xy,s} y_{t-s} + \epsilon_t$$

- Deeper issue: too many and too complicated
- Simple solution: Model selection

ACFs and CCFs



Model Selection

- Step 1: Pick maximum lag length

- ▶ Information criteria

$$\text{AIC: } \ln |\Sigma(P)| + k^2 P \frac{2}{T}$$

$$\text{Hannan-Quinn IC (HQIC): } \ln |\Sigma(P)| + k^2 P \frac{\ln \ln T}{T}$$

$$\text{SIC: } \ln |\Sigma(P)| + k^2 P \frac{\ln T}{T}$$

- ▷ $\Sigma(P)$ is the covariance of the residuals using P lags

- ▷ $|\cdot|$ is the determinant

- ▶ Hypothesis testing based

- ▷ General to Specific

- ▷ Specific to General

- ▶ Likelihood Ratio

$$(T - P_2 k^2) (\ln |\Sigma(P_1)| - \ln |\Sigma(P_2)|) \stackrel{A}{\sim} \chi_{(P_2 - P_1)k^2}^2$$

Identification in Campbell's VAR

- Maximum lag: 12 (1 year)

Lag Length	AIC	HQC	BIC	LR	P-val
0	6.30	5.84	5.50	1841	0.00
1	3.18	2.79	2.57	210.6	0.00
2	2.89	2.57	2.45	1428	0.00
3	0.25	0.00	0.00	43.68	0.01
4	0.24	0.07	0.18	33.11	0.13
5	0.26	0.15	0.38	133.2	0.00
6	0.05	0.01	0.35	47.45	0.00
7	0.02	0.06	0.50	20.08	0.74
8	0.05	0.16	0.72	28.22	0.30
9	0.06	0.24	0.91	22.84	0.59
10	0.07	0.32	1.11	18.91	0.80
11	0.10	0.42	1.31	54.06	0.00
12	0.00	0.39	1.39	N/A	N/A

Granger Causality

- **First fundamentally new concept**
- Examines whether lags of one variable are helpful in predicting another

Definition (Granger Causality)

A scalar random variable $\{x_t\}$ is said to **not** Granger cause $\{y_t\}$ if $E[y_t | x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, \dots] = E[y_t | y_{t-1}, y_{t-2}, \dots]$. That is, $\{x_t\}$ does not Granger cause if the forecast of y_t is the same whether conditioned on past values of x_t or not.

Granger Causality

- Translates directly into a restriction in a VAR
- Unrestricted

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Restricted so that x_t does not GC y_t

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

$$x_t = \phi_{01} + \phi_{11}x_{t-1} + \phi_{12}y_{t-1} + \epsilon_{1,t}$$

$$y_t = \phi_{02} + \phi_{22}y_{t-1} + \epsilon_{2,t} \Leftarrow \text{No } x_t!$$

More Granger Causality

- In P lag model

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \dots + \Phi_P \mathbf{y}_{t-P} + \epsilon_t$$

the null hypothesis is

$$H_0 : \phi_{ij,1} = \phi_{ij,2} = \dots = \phi_{ij,P} = 0$$

- Alternative is

$$H_0 : \phi_{ij,1} \neq 0 \text{ or } \phi_{ij,2} \neq 0 \text{ or } \dots \text{ or } \phi_{ij,P} \neq 0$$

- Likelihood Ratio test

$$(T - Pk^2) (\ln |\Sigma_r| - \ln |\Sigma_u|) \stackrel{A}{\sim} \chi_P^2$$

- Σ_u is the covariance of the errors from unrestricted model
- Σ_r is the covariance of the errors from restricted model
- $T - Pk^2$ is number of observations minus number of free parameters in unrestricted model
 - ▶ Why χ_P^2 ?

Campbell's VAR

- Campbell (1996) asset pricing paper
- State vector evolves as a VAR
- State variables include:
 - ▶ VWM: Value weighted market from CRSP
 - ▶ LBR: Real labor income growth from FRED
 - ▶ RTB: Relative T-bill from French's site
 - ▶ TERM: 10 year minus 3 month from FRED
 - ▶ DIV: Dividend yield from CRSP

Granger Causality in Campbell's VAR

- Using model with lags 1, 3 and 12 (SIC Search)
- $H_0 : \phi_{ij,1} = \phi_{ij,3} = \phi_{ij,12} = 0$
- $H_1 : \phi_{ij,1} \neq 0$ or $\phi_{ij,3} \neq 0$ or $\phi_{ij,12} \neq 0$
- i represent series being affected by lags of series j

Exclusion	VWM		LBR		RTB		TERM		DIV	
	Stat	P-val	Stat	P-val	Stat	P-val	Stat	P-val	Stat	P-val
<i>VWM</i>	–	–	21.07	0.00	11.55	0.02	7.93	0.09	11.26	0.02
<i>LBR</i>	4.98	0.29	–	–	4.07	0.40	4.62	0.33	3.68	0.45
<i>RTB</i>	3.97	0.41	1.99	0.74	–	–	210.66	0.00	40.03	0.00
<i>TERM</i>	7.13	0.13	16.40	0.00	4.25	0.37	–	–	2.84	0.59
<i>DIV</i>	181.86	0.00	2.09	0.72	40.48	0.00	12.98	0.01	–	–
All	46.79	0.00	21.01	0.18	276.24	0.00	28.65	0.03	287.59	0.00

Impulse Response Functions

- **Second fundamentally new concept**
- Complicated dynamics of a VAR make direct interpretation of coefficients difficult
- Solution is to examine impulse responses
- The impulse response function of y_i with respect to a shock in ϵ_j , for any j and i , is defined as the change in y_{it+s} , $s \geq 0$ for a unit shock in ϵ_{jt}
 - ▶ Hard to decipher
- As long as \mathbf{y}_t is covariance stationarity it must have a VMA representation,

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t + \boldsymbol{\Xi}_1 \boldsymbol{\epsilon}_{t-1} + \boldsymbol{\Xi}_2 \boldsymbol{\epsilon}_{t-2} + \dots$$

- $\boldsymbol{\Xi}_j$ are the impulse responses!
- Why?
 - ▶ Directly measure the effect in period j of any shock

AR(P) and MA(∞)

- Any stationary AR(P)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_P y_{t-P} + \epsilon_t$$

can be represented as an MA(∞)

$$y_t = \phi_0 / (1 - \phi_1 - \phi_2 - \dots - \phi_P) + \epsilon_t + \sum_{i=1}^{\infty} \theta_i \epsilon_{t-i}$$

- AR(1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$$

becomes

$$y_t = \phi_0 / (1 - \phi_1) + \epsilon_t + \sum_{i=1}^{\infty} \phi_1^i \epsilon_{t-i}$$

- Stationary VAR(P) have the same relationship to VMA(∞)

$$\mathbf{y}_t = \mathbf{\Phi}_0 + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \dots + \mathbf{\Phi}_P \mathbf{y}_{t-P} + \boldsymbol{\epsilon}_t$$

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\epsilon}_t + \boldsymbol{\Xi}_1 \boldsymbol{\epsilon}_{t-1} + \boldsymbol{\Xi}_2 \boldsymbol{\epsilon}_{t-2} + \dots$$

Solving IR

- Painful to directly solve for Ξ_j
- Easy in VAR(1)

$$\mathbf{y}_t = (\mathbf{I}_K - \Phi_1)^{-1} \Phi_0 + \epsilon_t + \Phi_1 \epsilon_{t-1} + \Phi_1^2 \epsilon_{t-2} + \dots$$

- $\Xi_j = \Phi_1^j$
- Just let MATLAB do it for you
- Confidence intervals are also somewhat painful
 - ▶ Explained in notes

Considerations for Shocks

- Simple bivariate VAR(1)

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \phi_{01} \\ \phi_{02} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- How you 'shock' matters
- Depends on correlation between $\epsilon_{1,t}$ and $\epsilon_{2,t}$
- 3 methods
 - ▶ Ignore correlation and just shock $\epsilon_{j,t}$ with a 1 standard deviation shock
 - ▶ Use Choleski to factor Σ and use $\Sigma^{1/2}\mathbf{e}_j$ where \mathbf{e}_j is a vector of zeros with 1 in the j^{th} position
 - ▶ Use spectral decomposition to factor Σ and use $\Sigma^{1/2}\mathbf{e}_j$

$$\Sigma = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix} \quad \Sigma_C^{1/2} = \begin{bmatrix} 1 & 0 \\ .5 & .866 \end{bmatrix} \quad \Sigma_S^{1/2} = \begin{bmatrix} 0.96 & .26 \\ .26 & .96 \end{bmatrix}$$

Example of the different shocks

- Define the error covariance

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_x\sigma_y\rho \\ \sigma_x\sigma_y\rho & \sigma_y^2 \end{bmatrix}$$

- ▶ Nonscaled

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- ▶ Standardized

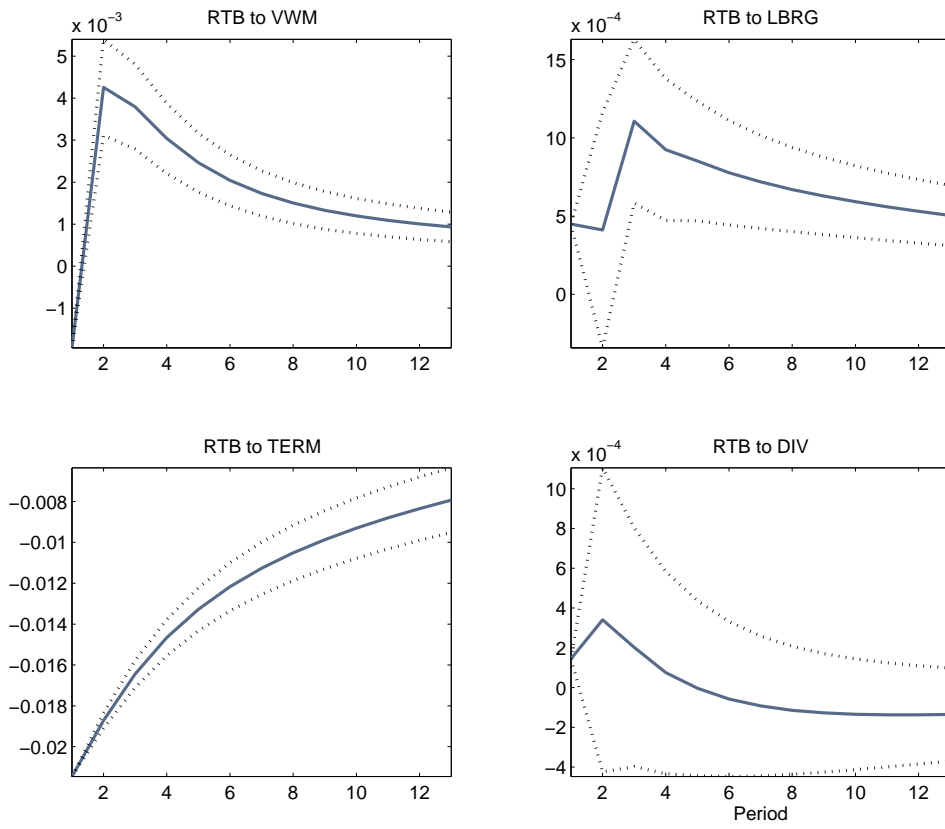
$$\begin{bmatrix} 0 \\ \sigma_x \end{bmatrix}$$

- ▶ Choleski

$$\Sigma_C^{1/2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma_x & 0 \\ \sigma_y\rho & \sigma_y\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sigma_y\sqrt{1-\rho^2} \end{bmatrix}$$

$$\begin{bmatrix} \sigma_x & 0 \\ \sigma_y\rho & \sigma_y\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_x \\ \sigma_y\rho \end{bmatrix}$$

Impulse Responses



Cointegration

- Cointegration is the VAR version of unit roots
- Establishes long run relationships between two unit root variables
 - ▶ Consumption: Unit root
 - ▶ Income: Unit root
 - ▶ Consumption - Income : ????
- I(1): Unit root, $\Delta I(1) \sim I(0)$
- I(0): Stationary
- Definition:
 - ▶ x_t and y_t are I(1) They are **cointegrated** if there exists a vector β with both elements non-zero such that

$$\beta_1 x_t - \beta_2 y_t \sim I(0)$$

- ▶ Strong link between x_t and y_t
- ▶ Both are random walks but difference is mean reverting
- ▶ Mean reversion to the trend (stochastic trend)

What does cointegration look like?

$$\mathbf{y}_t = \Phi_{ij} \mathbf{y}_{t-1} + \epsilon_t$$

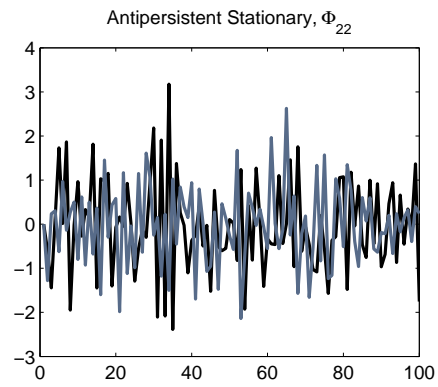
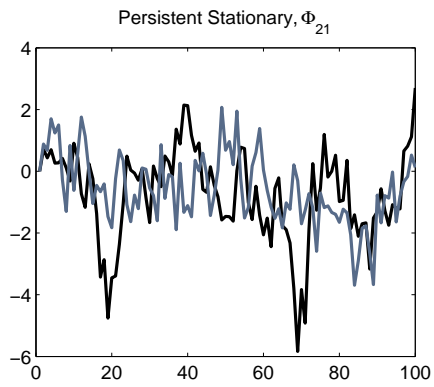
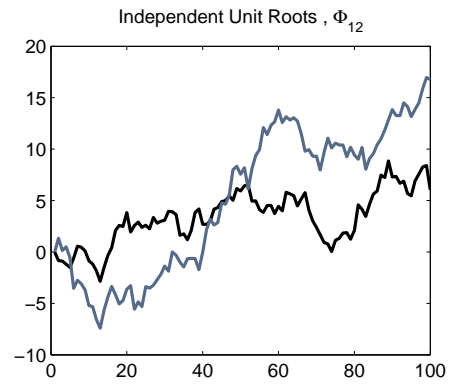
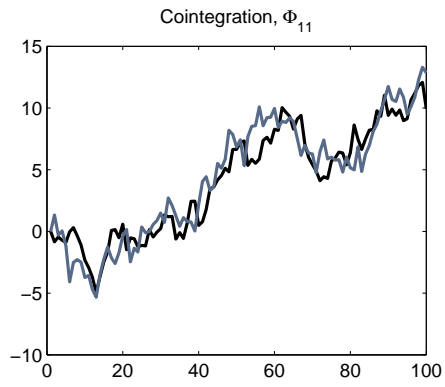
$$\Phi_{11} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix}$$
$$\lambda_i = 1, 0.6$$

$$\Phi_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$\lambda_i = 1, 1$$

$$\Phi_{21} = \begin{bmatrix} .7 & .2 \\ .2 & .7 \end{bmatrix}$$
$$\lambda_i = 0.9, 0.5$$

$$\Phi_{22} = \begin{bmatrix} -.3 & .3 \\ .1 & -.2 \end{bmatrix}$$
$$\lambda_i = -0.43, -0.06$$

Persistence, Anti-persistence and Cointegration



How do we know when a VAR is cointegrated?

- Eigenvalue condition determines whether a VAR(1) is cointegrated

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Cointegrated if only 1 eigenvalue is unity.
- If all less than 1: ?
- If both 1: two independent unit roots

$$\Phi_{11} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix}$$
$$\lambda_i = 1, 0.6$$

$$\Phi_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$\lambda_i = 1, 1$$

$$\Phi_{21} = \begin{bmatrix} .7 & .2 \\ .2 & .7 \end{bmatrix}$$
$$\lambda_i = 0.9, 0.5$$

$$\Phi_{22} = \begin{bmatrix} -.3 & .3 \\ .1 & -.2 \end{bmatrix}$$
$$\lambda_i = -0.43, -0.06$$

Error Correction Models

- Major point of cointegration
 - ▶ Cointegrated \Leftrightarrow Error correction model
- What is an error correction model?
 - ▶ Cointegrated VAR:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- ▶ Error correction model:

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 & .2 \\ .2 & -.2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- ▶ Normalized form

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 \\ .2 \end{bmatrix} [1 \quad -1] \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- $[1 \quad -1]$ is cointegrating vector
- $[-.2 \quad .2]'$ measures the speed of adjustment

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

Subtracting $[y_{t-1} \ x_{t-1}]'$ from both sides

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} - \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} - \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \left(\begin{bmatrix} .8 & .2 \\ .2 & .8 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 & .2 \\ .2 & -.2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 \\ .2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

Cointegrating vectors

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 & .2 \\ .2 & -.2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$
$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} -.2 \\ .2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}$$

- Cointegrating relationship can always be decomposed

$$\Delta \mathbf{y}_t = \boldsymbol{\pi} \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t$$

$$\boldsymbol{\pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$$

- $\boldsymbol{\alpha}$ measures the speed of convergence
- $\boldsymbol{\beta}$ contain the cointegrating vectors
- Number of cointegrating vectors is $\text{rank}(\boldsymbol{\alpha} \boldsymbol{\beta}')$

$$\boldsymbol{\alpha} \boldsymbol{\beta}' = \begin{bmatrix} 0.3 & 0.2 & -0.36 \\ 0.2 & 0.5 & -0.35 \\ -0.3 & -0.3 & 0.39 \end{bmatrix}$$

- How many?

Determining the cointegrating vectors

$$\Delta \mathbf{y}_t = \boldsymbol{\pi} \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t$$

$$\boldsymbol{\pi} = \begin{bmatrix} 0.3 & 0.2 & -0.36 \\ 0.2 & 0.5 & -0.35 \\ -0.3 & -0.3 & 0.39 \end{bmatrix}$$

- Put $\boldsymbol{\pi}$ in row echelon form

$$\boldsymbol{\pi} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.3 \\ 0 & 0 & 0 \end{bmatrix}$$

- Recall $\boldsymbol{\pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$

$$\boldsymbol{\beta} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -0.3 \end{bmatrix} \quad \boldsymbol{\alpha} = \begin{bmatrix} .3 & .2 \\ .2 & .2 \\ -.3 & -.3 \end{bmatrix}$$

Solving for the cointegrating vectors

$$\alpha\beta' = \begin{bmatrix} 0.3 & 0.2 & -0.36 \\ 0.2 & 0.5 & -0.35 \\ -0.3 & -0.3 & 0.39 \end{bmatrix}$$

$$\text{Row-Echelon Form} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \beta_1 & \beta_2 \end{bmatrix}$$

and α has 6 unknown parameters. $\alpha\beta'$ can be combined to produce

$$\pi = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{11}\beta_1 + \alpha_{12}\beta_2 \\ \alpha_{21} & \alpha_{22} & \alpha_{21}\beta_1 + \alpha_{22}\beta_2 \\ \alpha_{31} & \alpha_{32} & \alpha_{31}\beta_1 + \alpha_{32}\beta_2 \end{bmatrix}$$

