

Análisis de cointegración con modelos VARIMA

José L. Gallego¹ Carlos Díaz²

¹Departamento de Economía
Universidad de Cantabria

²Department of Economics
University of Leicester



Seminario 9-4-2013
Departamento de Economía (UC)

CONTENIDOS

INTRODUCCIÓN

- Metodología Box-Jenkins
- Contrastes de raíces unitarias
- Cointegración mutivariante

Modelos VARIMA cointegrados

- Cointegración en un modelo VIMA(1,1)
- Cointegración en un modelo VARIMA($p, 1, q + 1$) general

Ejemplo

- Simulación de series cointegradas

Experimentos Monte-Carlo

Conclusiones

MODELO ARIMA(P,D,Q)

- ▶ Proceso lineal general estacionario (Box y Jenkins 1970):

$$(1 - \pi_1 B - \pi_2 B^2 - \dots) z_t = a_t \quad \Leftrightarrow \quad z_t = (1 + \psi_1 B + \psi_2 B^2 + \dots) a_t$$

- ▶ Condiciones de estacionariedad e invertibilidad

$$\sum_{j=1}^{\infty} |\psi_j| < \infty \quad \text{y} \quad \sum_{j=1}^{\infty} |\pi_j| < \infty$$

- ▶ Proceso estacionario ARMA(p,q):

$$(1 - \phi_1 B - \dots - \phi_p B^p) z_t = (1 - \theta_1 B - \dots - \theta_p B^p) a_t$$

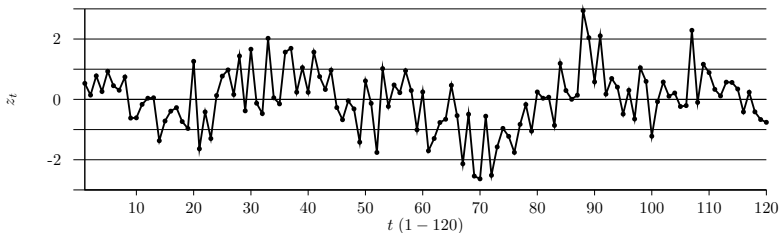
- ▶ Proceso no-estacionario ARIMA(p,d,q):

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d z_t = (1 - \theta_1 B - \dots - \theta_p B^p) a_t$$

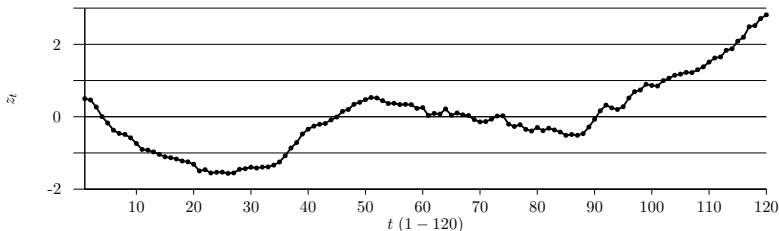
$$\phi(B) \nabla^d z_t = \theta(B) a_t$$

SERIES NO ESTACIONARIAS

► Serie simulada con un modelo IMA(1,1)

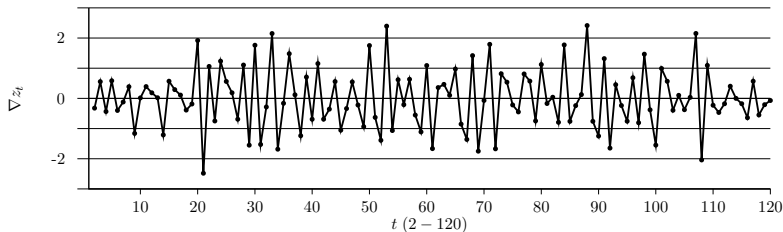


► Serie simulada con un modelo IMA(2,1)

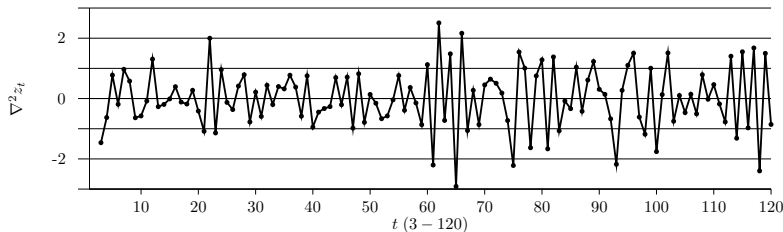


SERIES ESTACIONARIAS

- ▶ Primera diferencia de una serie simulada con IMA(1,1)



- ▶ Segunda diferencia de una serie simulada con un IMA(2,1)



METODOLOGÍA BOX-JENKINS

- ▶ Procedimiento iterativo
 1. Identificación
 2. Estimación
 3. Diagnósis
- ▶ Herramientas gráficas: gráfico temporal, FAS, FAP.
- ▶ Estrategias para identificar el orden de integración
 1. Infradiferenciar

$$(1 - \phi B)z_t = a_t$$

2. Sobrediferenciar

$$(1 - B)z_t = (1 - \theta B)a_t$$

CONTRASTES DE RAÍCES AR UNITARIAS

▶ Proceso AR(1)

$$(1 - B)z_t = (\phi - 1)z_{t-1} + a_t$$

▶ Hipótesis

$$H_0 : \phi = 1 \quad \text{frente} \quad H_1 : \phi < 1$$

▶ Estadístico de Dickey-Fuller (JASA 1979)

$$t_{(\hat{\phi}-1)} \overset{a}{\sim} DF$$

▶ Extensiones

1. Paramétricas: Said y Dickey (Biometrika 1984).
2. No-paramétricas: Phillips y Perron (Biometrika 1988).

CONTRASTE DE RAÍZ MA UNITARIA

- ▶ Modelo IMA(1,1)

$$(1 - B)z_t = (1 - \theta B)a_t$$

- ▶ Hipótesis a contrastar

$$H_0 : \theta = 1 \quad \text{frente} \quad H_0 : \theta \neq 1$$

- ▶ Estadístico LMPI (LBIU) de Nyblom y Makelainen (JASA 1983)

$$NM = \frac{1}{n} \frac{\sum_{t=1}^n (\sum_{\tau=1}^t \tilde{z}_\tau)^2}{\sum_{t=1}^n \tilde{z}_t^2} \stackrel{a}{\sim} CvM(1)$$

- ▶ Extensiones:

1. Paramétricas: Saikkonen y Luukkonen (JASA 1993), Leybourne y McCabe (JBES 1994), Tam y Reinsel (JASA 1997).
2. No paramétricas: Tanaka (ET 1990), Kwiatkowski et al. (JoE 1993).

MODELO VARIMA(p,D,Q)

- ▶ Proceso estacionario VARMA(p,q) (Box y Tiao, JASA 1981):

$$(\mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p) \mathbf{z}_t = (\mathbf{I}_m - \Theta_1 B - \dots - \Theta_p B^p) \mathbf{a}_t$$

- ▶ Proceso no-estacionario ARIMA(p,d,q):

$$(\mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p) \mathbf{z}_t = (\mathbf{I}_m - \Theta_1 B - \dots - \Theta_p B^p) \mathbf{a}_t$$

$$\Phi(B) \nabla^d \mathbf{z}_t = \Theta(B) \mathbf{a}_t$$

- ▶ Problema: combinaciones lineales de series no estacionarias pueden ser estacionarias (Box y Tiao, Biometrika 1977).
- ▶ Cointegración: Granger (JoE 1981), Engle y Grange (JoE 1987).

COINTEGRACIÓN MULTIVARIANTE

1. Extensiones multivariante del contraste Dickey-Fuller: Stock y Watson (JASA 1988), Johansen (Econometrica 1991).
2. Extensiones multivariante del contraste Nyblom-Makelainen: Nyblom y Harvey (ET 2000).
3. Contrastes basados en modelos VARMA: Saikkonen (1992), Yap and Reinsel (1995), Lütkepohl and Claessen (1997).

COINTEGRACIÓN EN UN MODELO VIMA(1,1)

- ▶ Modelo VIMA(1,1) para un vector de m series temporales

$$\nabla \mathbf{z}_t = (\mathbf{I}_m - \Theta B) \mathbf{a}_t, \quad t = 1, \dots, n,$$

- ▶ $\mathbf{a}_t \sim N(\mathbf{0}, \Omega_a)$,
- ▶ $|\mathbf{I}_m - \Theta B| = 0$ for $|B| \geq 1 \Rightarrow \mathbf{z}_t \sim I(1)$.
- ▶ Descomposición de Engle-Granger (1987)

$$\nabla \mathbf{z}_t = (\mathbf{I}_m - \Theta) \mathbf{a}_t + \nabla \Theta \mathbf{a}_t$$

- ▶ Una matriz \mathbf{C} , ($r \times m$), con r relaciones de cointegración

$$\nabla \mathbf{C} \mathbf{z}_t = \mathbf{C} (\mathbf{I}_m - \Theta) \mathbf{a}_t + \nabla \mathbf{C} \Theta \mathbf{a}_t$$

- ▶ Descomp. espectral de $\mathbf{I}_m - \Theta = \mathbf{P}_1 \Lambda_1 \mathbf{Q}_1 + \mathbf{P}_2 \Lambda_2 \mathbf{Q}_2$ ($\mathbf{Q}_1 \mathbf{P}_2 = \mathbf{0}$)

$$\mathbf{C} = \mathbf{Q}_1 \Rightarrow \mathbf{C} (\mathbf{I}_m - \Theta) = \mathbf{0} \Rightarrow \mathbf{C} \Theta = \mathbf{C}$$

CONTRASTE DE COINTEGRACIÓN

- ▶ Proceso VIMA(1,1) escalar

$$\nabla \mathbf{y}_t = (1 - \theta B)\mathbf{e}_t, \quad \mathbf{y}_t = \mathbf{Cz}_t \text{ y } \mathbf{e}_t = \mathbf{Ca}_t$$

- ▶ Hipótesis a contrastar (cointegración frente a no-cointegración)

$$H_0 : \theta = 1 \quad \text{frente a} \quad H_1 : \theta < 1$$

- ▶ Estadístico LMPI de Nyblom y Harvey (ET 2000)
- ▶ Estadístico NH basado en $\hat{\mathbf{e}}_t = E(\mathbf{e}_t | \nabla \mathbf{Y}_2, \dots, \nabla \mathbf{Y}_n)$:

$$L = \frac{1}{rn} \text{tr} \left[\left(\sum_{t=1}^n \hat{\mathbf{e}}_t \hat{\mathbf{e}}_t' \right)^{-1} \sum_{t=1}^n \left(\sum_{\tau=1}^t \hat{\mathbf{e}}_\tau \sum_{\tau=1}^t \hat{\mathbf{e}}_\tau' \right) \right] \stackrel{a}{\sim} \frac{1}{r} CVM(r).$$

- ▶ Valores críticos: Hansen (1990), Gallego and Díaz (2011).

PROCEDIMIENTO DE CONTRASTE

1. Estimar el modelo VIMA(1,1), $\nabla \mathbf{z}_t = (\mathbf{I}_m - \Theta B)\mathbf{a}_t$, por máxima verosimilitud exacta.
2. Calcular los autovalores y autovectores de $(\mathbf{I}_m - \hat{\Theta})$ usando la descomposición $(\mathbf{I}_m - \hat{\Theta})\mathbf{P} = \mathbf{P}\Lambda$.
3. Invertir la matriz \mathbf{P} y construir \mathbf{C} seleccionando las filas asociadas con los r autovalores más pequeños λ_i ($i = \dots, r$).
4. Normalizar los autovectores de manera que $\mathbf{C} = [\mathbf{I}_r | \mathbf{C}_*]$.
5. Generar y transformar los residuos exactos, $\hat{\mathbf{e}}_t = \mathbf{C}\hat{\mathbf{a}}_t$.
6. Calcular el estadístico de Nyblom-Harvey, L .
7. Comparar L con el valor crítico asociado al nivel de significación deseado.

COINTEGRACIÓN EN UN VARIMA($p, 1, q + 1$)

- ▶ Modelo VARIMA($p, 1, q + 1$)

$$\Phi_p(B)\nabla\mathbf{z}_t = \Theta_{q+1}(B)\mathbf{a}_t, \quad t = 1, \dots, n,$$

- ▶ $\mathbf{a}_t \sim N(\mathbf{0}, \Omega_a)$,
- ▶ $\Phi(B) = \mathbf{I}_m - \Phi_1 B - \dots - \Phi_p B^p$, $|\Phi(B)| = 0$ para $|B| > 1$,
- ▶ $\Theta(B) = \mathbf{I}_m - \Theta_1 B - \dots - \Theta_{q+1} B^{q+1}$, $|\Theta(B)| = 0$ para $|B| \geq 1$.

- ▶ Descomposición de Engle and Granger (1987)

$$\Phi_p(B)\nabla\mathbf{z}_t = \Theta_{q+1}(1)\mathbf{a}_t + \Theta_q^*(B)\nabla\mathbf{a}_t.$$

- ▶ Una matriz \mathbf{C} , ($r \times m$), con r relaciones de cointegración

$$\mathbf{C}\Phi_p\nabla\mathbf{z}_t = \mathbf{C}\Theta(1)\mathbf{a}_t + \mathbf{C}\Theta^*(B)\nabla\mathbf{a}_t.$$

- ▶ Modelo VARMA(p, q) para las r relaciones de cointegración

$$\Phi_p^*\nabla\mathbf{Cz}_t = \mathbf{C}\Theta(1)\mathbf{a}_t + \Theta^{**}(B)\nabla\mathbf{Ca}_t.$$

CONTRASTE PARAMÉTRICO DE COINTEGRACIÓN

- ▶ Modelo VARIMA($p, 1, q + 1$) restringido

$$\Phi_p^*(B)\nabla\mathbf{y}_t = \Theta_q^{**}(B)(1 - \theta B)\mathbf{e}_t, \quad \mathbf{y}_t = \mathbf{Cz}_t \text{ y } \mathbf{e}_t = \mathbf{Ca}_t$$

- ▶ Hipótesis a contrastar (cointegración frente a no-cointegración)

$$H_0 : \theta = 1 \quad \text{frente a} \quad H_1 : \theta < 1.$$

- ▶ Extensión de los contrastes paramétricos de no-invertibilidad.
- ▶ Contraste basado en los residuos exactos ($\hat{\mathbf{e}}_t = \mathbf{D}\hat{\mathbf{a}}_t$)

$$L = \frac{1}{rn} \text{tr} \left[\left(\sum_{t=1}^n \hat{\mathbf{e}}_t \hat{\mathbf{e}}_t' \right)^{-1} \sum_{t=1}^n \left(\sum_{\tau=1}^t \hat{\mathbf{e}}_\tau \sum_{\tau=1}^t \hat{\mathbf{e}}_\tau' \right) \right] \stackrel{a}{\sim} \frac{1}{r} CVM(r).$$

SIMULACIÓN DE SERIES COINTEGRADAS

- ▶ Modelo IMA(1,1) trivariante

$$\nabla \mathbf{z}_t = (\mathbf{I}_m - \Theta B) \mathbf{a}_t, \quad t = 1, \dots, 120,$$

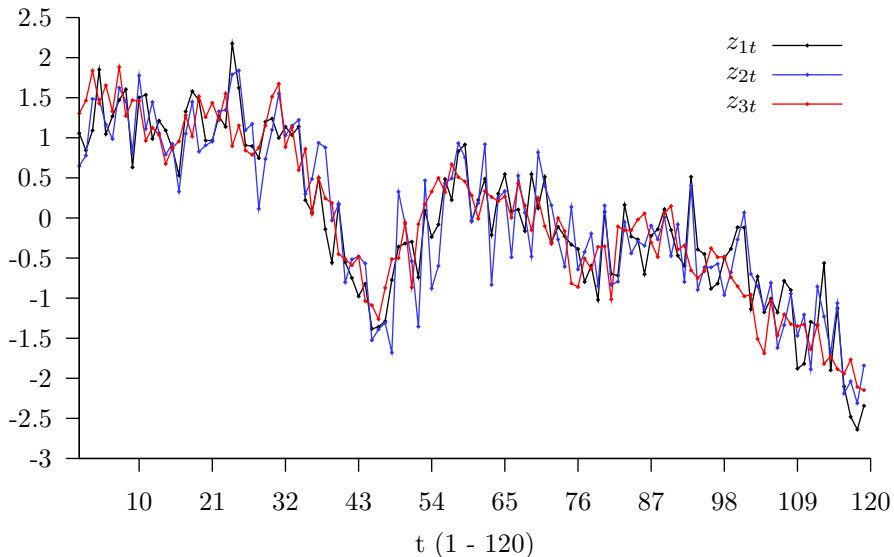
- ▶ Descomposición espectral de $\Theta = \mathbf{P}\Lambda\mathbf{P}^{-1}$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0.25 & -0.5 & 0.75 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.7 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0.25 & -0.5 & 0.75 \end{pmatrix}$$

- ▶ Matrices de parámetros

$$\Theta = \begin{pmatrix} 0.85 & 0.30 & -0.45 \\ -0.15 & 1.30 & -0.45 \\ -0.15 & 0.30 & 0.55 \end{pmatrix}, \quad \Omega_a = 10^{-3} \begin{pmatrix} 2.89 & 1.43 & 0.17 \\ 1.43 & 3.91 & -0.06 \\ 0.17 & -0.06 & 1.46 \end{pmatrix}$$

TRES SERIES COINTEGRADAS



MÁXIMA VEROSIMILITUD EXACTA

- ▶ Preestimaciones: $\hat{\Theta} = 0.8\mathbf{I}_3$ and $\hat{\Omega}_a = \mathbf{I}_3$.
- ▶ Estimaciones finales

$$\hat{\Theta} = \begin{pmatrix} .729 & .321 & -.561 \\ -.271 & 1.3026 & -.567 \\ -.264 & .285 & .442 \end{pmatrix} \text{ se}(\hat{\Theta}) = \begin{pmatrix} .073 & .070 & .078 \\ .072 & .076 & .073 \\ .066 & .061 & .073 \end{pmatrix}$$

$$\hat{\Omega}_a = 10^{-3} \begin{pmatrix} 2.677 & 1.652 & .118 \\ 1.652 & 3.636 & .111 \\ .118 & .111 & 1.615 \end{pmatrix} \quad \log\text{-lik} = 576.62$$

- ▶ Raíces de $|\mathbf{I}_3 - \hat{\Theta}\mathbf{B}| = 0$

Raíz	Real	Imaginaria	Módulo	Frecuencia	Periodo
1	2.1074	0	2.1074	0	Inf
2	0.999996	-0.002803	1.0000	0.000446	2241.77
3	0.999996	0.002803	1.0000	0.000446	2241.77

ANÁLISIS DE COINTEGRACIÓN

▶ Autovalores de $I_3 - \hat{\Theta}$

Real	Imaginaria	Módulo
0.525477	0	0.525477
0.000004	0.002803	0.002803
0.000004	-0.002803	0.002803

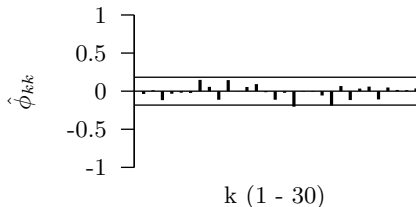
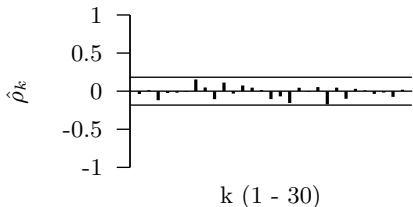
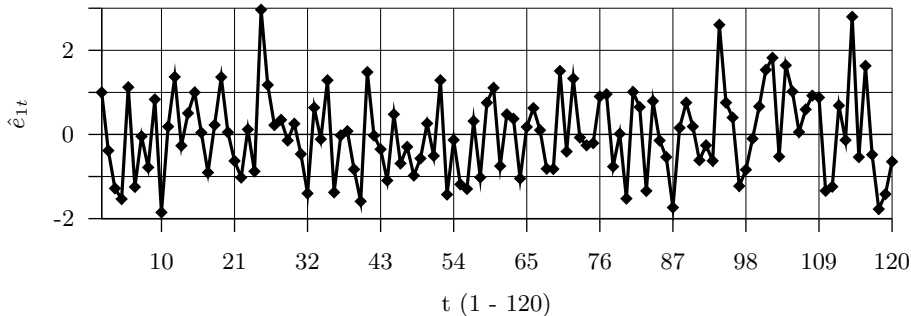
▶ Relaciones de cointegración: $(I_3 - \hat{\Theta})\hat{P} = \hat{P}\hat{\Lambda}$

$$\hat{P}^{-1} = \begin{pmatrix} 0.874 & -0.969 & 1.835 \\ -0.539 & 1.516 & -0.996 \\ 0.397 & -0.811 & 0.430 \end{pmatrix} \Rightarrow \hat{C} = \begin{pmatrix} 1 & 0 & -0.9498 \\ 0 & 1 & -0.9953 \end{pmatrix}$$

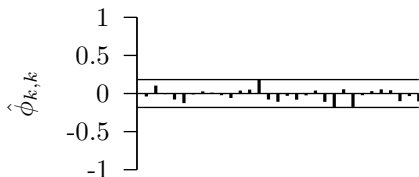
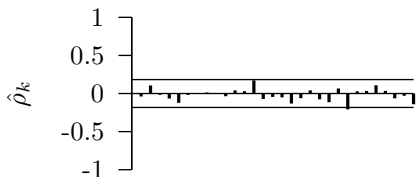
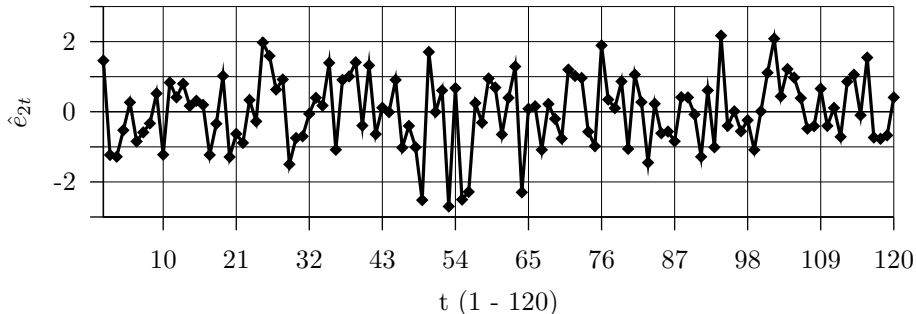
▶ Contrastes de cointegración

L1	p-value	L0	p-value	NH	p-value
0.090512	0.780340	0.084649	0.813968	0.084649	0.813968

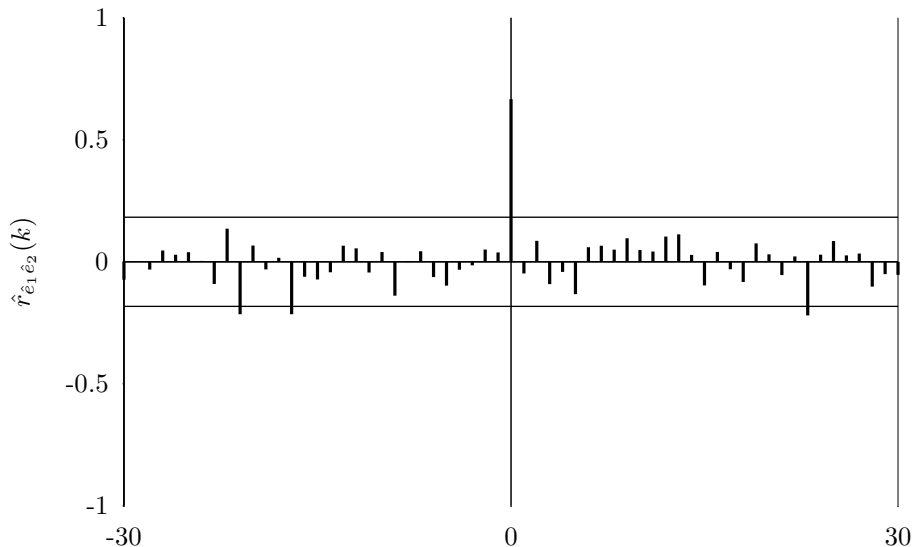
DIAGNOSIS: PRIMERA RELACIÓN DE COINTEGRACIÓN



DIAGNOSIS: SEGUNDA RELACIÓN DE COINTEGRACIÓN



DIAGNOSIS: FUNCIÓN DE CORRELACIÓN CRUZADA



CONTRASTE DE JOHANSEN

Caso 1: sin componentes deterministas ($p=4$)

Rank	Eigenvalue	Trace test	p-value	Lmax test	p-value
0	0.241212	35.7176	0.000983	31.7438	0.000103
1	0.030245	3.9739	0.705353	3.5318	0.702356
2	0.003836	0.442024	0.567474	0.442024	0.568122

Vectores de cointegración: $(1, 0, -1.64)$ y $(0, 1, -0.44)$

Caso 2: término constante ($p=4$)

Rank	Eigenvalue	Trace test	p-value	Lmax test	p-value
0	0.303968	61.6870	2.37e-07	41.6714	0.000006
1	0.157977	20.0156	0.004448	19.7740	0.004940
2	0.002099	0.241598	0.479939	0.241598	0.623067

Vectores de cointegración: $(1, 0, -1.01)$ y $(0, 1, -1.02)$

EXPERIMENTO I

- ▶ Proceso generador de datos

$$\nabla \mathbf{z}_t = (\mathbf{I}_m - \Theta B) \mathbf{a}_t, \quad t = 1, \dots, 120,$$

- ▶ Matrices de parámetros

$$\Theta = \begin{pmatrix} 0.85 & 0.30 & -0.45 \\ -0.15 & 1.30 & -0.45 \\ -0.15 & 0.30 & 0.55 \end{pmatrix}, \quad \Omega_a = 10^{-3} \begin{pmatrix} 2.89 & 1.43 & 0.17 \\ 1.43 & 3.91 & -0.06 \\ 0.17 & -0.06 & 1.46 \end{pmatrix}$$

- ▶ Número de realizaciones: 1000
- ▶ Tamaños empíricos

r	5% v.c.	$\hat{\alpha}$
1	0.473	4.1%
2	0.351	7.1%

RESUMEN Y EXTENSIONES

- ▶ Contraste de cointegración en modelos VARIMA models con el estadístico de Nyblom and Harvey basado en residuos exactos.
- ▶ Cointegración en sistemas triangulares y contraste de tendencias comunes.
- ▶ Extensiones adicionales: modelos VARIMA con deriva o variables de intervención; cointegración estacional.
- ▶ Programa Empiricus

<http://www.empiricus.unican.es>