

# Detección de Raíces Unitarias y Cointegración mediante Métodos de Subespacios

Alfredo G<sup>a</sup> Hiernaux\*      Miguel Jerez      José Casals

22 de junio de 2005

---

\*Departamento de Fundamentos del Análisis Económico II. Facultad de Ciencias Económicas. Campus de Somosaguas. 28223 Madrid. email: agarciah@ccee.ucm.es, tel: (+34) 91 394 26 26, fax: 91 394 25 91.

# Detección de raíces unitarias y cointegración mediante métodos de subespacios

## Resumen

En este trabajo se propone un nuevo procedimiento para detectar raíces unitarias basado en métodos de subespacios. Nuestra propuesta tiene tres aspectos originales principales. Primero, la misma metodología puede aplicarse a series individuales o a vectores de series temporales. Segundo, utiliza una familia flexible de criterios de información, cuyas funciones de pérdida pueden adaptarse a las propiedades estadísticas de los datos. Finalmente, no requiere especificar un proceso estocástico para las series analizadas. Un ejercicio de simulación muestra que el método tiene buenas propiedades en muestras finitas y su aplicación práctica se ilustra mediante el análisis de varias series reales.

**Palabras Clave:** Espacio de los Estados, métodos de subespacios, raíces unitarias, cointegración

**Clasificación AMS:** 62M10 - 62H20

# 1. INTRODUCCIÓN

Este trabajo desarrolla un método para detectar raíces unitarias en series temporales que tiene su origen en los métodos aplicados habitualmente a sistemas físicos. En ingeniería aeronáutica y de telecomunicaciones existe una importante tradición de análisis de este tipo de series, utilizando instrumentos que no se han extendido completamente a otros campos. Estas técnicas, que se conocen como de *Identificación de Sistemas*, se ocupan de “la determinación, a partir de un input y un output observados y dentro de una clase específica de modelos, de un modelo que optimiza una función de pérdida que depende del input y del output del proceso en estudio” (Aström y Eykhoff, 1971). El trabajo pionero de Ho y Kalman (1966) sentó las bases de la estrategia de modelización en este campo. Ésta se centra en obtener, a partir de los datos, un modelo en forma de espacio de los estados. Este problema se aborda generalmente con un enfoque geométrico-algebraico, aunque también cabe destacar aproximaciones estadísticas como las de Akaike (1975, 1976) o Larimore (1983, 1990).

Desde los años 90, la identificación de sistemas ha estado dominada por los *Métodos de Subespacios*. Muchos de estos métodos parten de la idea de que, si se observaran los estados de un sistema, las matrices que lo caracterizan podrían estimarse por mínimos cuadrados ordinarios. Puesto que en general no ocurre así, el problema de identificación se reduce a aproximar los estados a partir de los datos, sin tener información *a priori* de la estructura del sistema. El término “subespacio” en este contexto se relaciona con la forma en que se obtienen estos estados, los datos observados se organizan en matrices generadoras de los subespacios de información pasada y futura, a partir de las cuales se resuelve un problema de mínimos cuadrados generalizados con restricción de rango que proporciona las secuencias de estados. Algunos de estos métodos difieren entre sí sólo en las matrices de ponderación usadas en el problema de mínimos cuadrados generalizados.

En comparación con las técnicas convencionales de análisis estadístico de series temporales, como pueden ser Box y Jenkins (1976) o Jenkins y Alavi (1981), estos métodos tienen tres ventajas fundamentales:

1. Permiten identificar un modelo lineal general a partir de los datos, sin información *a*

*priori* sobre la estructura del proceso.

2. Tratan de la misma forma los casos univariante y multivariante.
3. Son computacionalmente eficientes y se apoyan en técnicas numéricas robustas del álgebra lineal.
4. No requieren iteraciones, por lo que carecen de problemas de convergencia.

Por otra parte, la aplicación de estos métodos a series temporales que no proceden de sistemas físicos presenta importantes desafíos. Por ejemplo, las series temporales en ingeniería suelen ser estacionarias y típicamente carecen de componentes estacionales, por lo que los métodos de subespacios disponibles carecen de instrumentos adecuados para acomodar estas características habituales en series económicas.

En este trabajo se aborda un tema no tratado por la literatura de subespacios en ingeniería: la detección de raíces unitarias. Para ello partimos del algoritmo CCA (Canonical Correlation Analysis) de Larimore (1983). Este método resulta especialmente conveniente ya que, debido a las matrices de ponderación que utiliza, el algoritmo calcula eficientemente los coeficientes de correlación canónica (CCC) entre los subespacios de información pasada y futura. Esta propiedad, además de facilitar la interpretación estadística del procedimiento, constituye el fundamento del método que se propone. La idea básica es que los CCC correspondientes a raíces unitarias convergen a sus verdaderos valores exponencialmente, mientras que los demás CCC exhiben una convergencia mucho más lenta. Esta propiedad de los CCC de raíces unitarias, conocida como *superconsistencia*, permite distinguir entre ambos tipos de correlaciones.

La superconsistencia de los CCC fue inicialmente aprovechada por Poskitt (2000) para el análisis de raíces unitarias y relaciones de cointegración. Posteriormente, Bauer y Wagner (2002), de aquí en adelante BW, refinan este planteamiento aprovechando la relación con los métodos de subespacios. El punto de partida de nuestro trabajo es, por tanto, similar al de BW, con las siguientes mejoras:

1. Los criterios de información en los que se basa el método de BW no tienen en cuenta una variable relevante en el cálculo de las matrices que definen el problema de regresión. Los resultados muestran que incluir esta dimensión en el criterio mejora apreciablemente su capacidad discriminante.
2. BW describen la forma matemática de su criterio de información como “un tanto heurística”. Con ello quieren decir que el peso de los distintos términos que se tienen en cuenta se fija arbitrariamente. Aquí, éstos se ajustan empíricamente usando métodos de simulación.

La estructura del artículo es la siguiente. La Sección 2 define la notación básica y resume los resultados fundamentales que se usarán posteriormente. En la Sección 3 se presenta la metodología propuesta y se analizan sus propiedades en muestras finitas mediante ejercicios de simulación. La Sección 4 ilustra la aplicación práctica del método mediante el análisis de dos casos reales. Finalmente, la Sección 5 resume las principales conclusiones.

## 2. MÉTODOS DE SUBESPACIOS Y CORRELACIONES CANÓNICAS

Consideramos un sistema lineal de coeficientes fijos, cuya dinámica puede describirse mediante el siguiente modelo en Espacio de los Estados en forma de innovaciones,

$$\mathbf{x}_{t+1} = \Phi \mathbf{x}_t + E \psi_t \tag{1}$$

$$\mathbf{z}_t = H \mathbf{x}_t + \psi_t \tag{2}$$

donde  $\mathbf{x}_t \in \mathbb{R}^n$  es el vector de estado,  $\mathbf{z}_t \in \mathbb{R}^m$  es el vector de outputs observables y  $\psi_t \in \mathbb{R}^m$  es una secuencia de perturbaciones  $iidN(\mathbf{0}, Q)$ . En este sistema se ha omitido, sin pérdida de generalidad, un vector de variables exógenas. Esta formulación tiene la ventaja de que numerosos modelos de series temporales, incluyendo estructuras VAR y VARMA, pueden representarse directamente en esta forma (ver Aoki, 1990; Terceiro, 1990).

Los métodos de subespacios parten de una representación matricial de las ecuaciones

de estado (1) y de observación (2). Así, de (1) recursivamente, se obtiene que,

$$\mathbf{x}_t = \mathbf{\Phi}^t \mathbf{x}_0 + \sum_{i=0}^{t-1} \mathbf{\Phi}^i \mathbf{E} \psi_{t-i-1} \quad (3)$$

y sustituyendo a su vez (3) en la ecuación de observación (2) se obtiene,

$$\mathbf{z}_t = \mathbf{H} \mathbf{\Phi}^t \mathbf{x}_0 + \mathbf{H} \sum_{i=0}^{t-1} \mathbf{\Phi}^i \mathbf{E} \psi_{t-i-1} + \psi_t \quad (4)$$

que es una representación del vector de variables endógenas,  $\mathbf{z}_t$ , que depende exclusivamente del vector de estado inicial,  $\mathbf{x}_0$ , y de la historia presente y pasada de las innovaciones,  $\psi_t$ .

La ecuación (4) puede escribirse en forma matricial como,

$$\mathbf{Z}_k = \mathbf{O}_i \mathbf{X}_k + \mathbf{H}_i \mathbf{\Psi}_k \quad (5)$$

donde los datos se organizan en  $\mathbf{Z}_k$ , denominada matriz Output Block Hankel (ver Van Overschee y De Moor, 1994). Los elementos de la ecuación (5) se estructuran de la forma siguiente:

1) Matrices Output e Input Block Hankel:

$$\begin{pmatrix} \mathbf{Z}_p \\ \mathbf{Z}_f \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_{0:i-1} \\ \mathbf{Z}_{i:2i-1} \end{pmatrix} = \begin{pmatrix} z_0 & z_1 & \dots & z_{T-2i} \\ z_1 & z_2 & \dots & z_{T-2i+1} \\ \vdots & \vdots & & \vdots \\ z_{i-1} & z_i & \dots & z_{T-i-1} \\ \hline z_i & z_{i+1} & \dots & z_{T-i} \\ z_{i+1} & z_{i+2} & \dots & z_{T-i+1} \\ \vdots & \vdots & & \vdots \\ z_{2i-1} & z_{2i} & \dots & z_{T-1} \end{pmatrix} \quad (6)$$

donde  $\mathbf{z}_0$  es un vector de  $m$  variables en el instante 0 y  $k = p, f$ . La matriz  $\mathbf{\Psi}_k$  tiene la misma estructura que  $\mathbf{Z}_k$ , pero agrupando las secuencias de las perturbaciones  $\psi_t$ . Además,  $i$  es la dimensión del subespacio de información o espacio de filas de  $\mathbf{Z}_k$ . La elección de la dimensión de los subespacios de información pasada y futura,  $p$  y  $f$  respectivamente, se

estudia, entre otros, en Viberg (1995), Chui (1997) o Bauer (1998), siendo habitual tomar  $p = f = i$ . En los algoritmos utilizados en este trabajo (ver García Hiernaux, 2005) se define  $i = \text{máx}(4, h_t)$ , siendo  $h_t$  el entero más cercano a  $\log(T)$ . Esta alternativa se basa en una regla heurística fundamentada en la experiencia empírica, para una discusión más profunda sobre los posibles valores de  $i$  (ver Casals, 1997).

2) Matriz de Observabilidad.

$$\mathbf{O}_i = \begin{pmatrix} \mathbf{H} \\ \mathbf{H}\Phi \\ \mathbf{H}\Phi^2 \\ \vdots \\ \mathbf{H}\Phi^{i-1} \end{pmatrix} \in \mathbb{R}^{im \times n} \quad (7)$$

3) Secuencia de estados.

$$\mathbf{X}_t = (\mathbf{x}_t \quad \mathbf{x}_{t+1} \quad \mathbf{x}_{t+2} \quad \dots \quad \mathbf{x}_{t+T-2i})$$

A partir de esta definición, las secuencias de estados pasados y futuros, que se inician en  $t = 0$  y  $t = i$  respectivamente, pueden expresarse como  $\mathbf{X}_p = \mathbf{X}_0$  y  $\mathbf{X}_f = \mathbf{X}_i$ .

4) Matriz Toeplitz triangular inferior por bloques.

$$\mathbf{H}_i^s = \begin{pmatrix} \mathbf{I}_m & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{H}\mathbf{E} & \mathbf{I}_m & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{H}\Phi\mathbf{E} & \mathbf{H}\mathbf{E} & \mathbf{I}_m & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}\Phi^{i-2}\mathbf{E} & \mathbf{H}\Phi^{i-3}\mathbf{E} & \mathbf{H}\Phi^{i-4}\mathbf{E} & \dots & \mathbf{I}_m \end{pmatrix} \in \mathbb{R}^{im \times im} \quad (8)$$

Ya que se parte de un sistema dinámico lineal, la secuencia de estados futuros puede definirse como  $\mathbf{X}_f = \mathbf{A}_n \mathbf{Z}_p$ , donde  $\mathbf{A}_n$  es una matriz de rango  $n$ . Por tanto, seleccionando adecuadamente los subíndices temporales en (5) y sustituyendo se obtiene,

$$\mathbf{Z}_f = \mathbf{O}_i \mathbf{A}_n \mathbf{Z}_p + \mathbf{H}_i \Psi_f \quad (9)$$

Los algoritmos de subespacios resuelven estos modelos de regresión con restricción de rango dados en (9), estimando las matrices  $\mathbf{O}_i$ ,  $\mathbf{A}_n$  y  $\mathbf{H}_i$ , para después recuperar las matrices del sistema formado por (1) y (2). Para ello, usan la Descomposición en Valores Singulares, SVD, (Eckart y Young, 1936) de  $\mathbf{W}_1 \mathbf{Z}_f \mathbf{W}_2$ , donde  $\mathbf{W}_1$  y  $\mathbf{W}_2$  son dos matrices de ponderación. Cuando  $\mathbf{W}_1 = (\mathbf{Z}_f \mathbf{Z}'_f)^{-\frac{1}{2}}$  y  $\mathbf{W}_2 = \mathbf{Z}'_p (\mathbf{Z}_p \mathbf{Z}'_p)^{-1} \mathbf{Z}_p$ , los Valores Singulares (VS) resultantes de la descomposición SVD coinciden con los CCC entre el subespacio de información futura,  $\mathbf{Z}_f$ , y el de información pasada,  $\mathbf{Z}_p$ .

### 3. DETECCIÓN DE RAÍCES UNITARIAS

BW proponen un procedimiento rápido para detectar raíces unitarias en el que, en lugar de calcular el CCC entre  $z_t$  y  $z_{t-1}$  como sugiere Poskitt (2000), usan los CCC entre los bloques de información pasada  $\mathbf{Z}_p$  y de información futura  $\mathbf{Z}_f$ . Estos autores evalúan una transformación de los  $j$  primeros VS calculados,  $\hat{\sigma}_j$  (donde  $j \leq i$ ), comparándola con una función de penalización  $C(T)$ , que depende del tamaño muestral:

$$1 - \hat{\sigma}_j^2 - C(T) < 0 \quad (10)$$

El VS que cumple esta desigualdad puede considerarse igual a la unidad, correspondiendo a una raíz unitaria del sistema. La ecuación (10) compara la distancia a cero de  $1 - \hat{\sigma}_j^2$  con una función de penalización,  $C(T)$ . Ésta representa una cota superior que implica que el valor  $1 - \hat{\sigma}_j^2$  que se sitúe por debajo, será considerado igual a cero, con lo que el criterio devolverá al menos  $j$  raíces unitarias ya que los VS cumplen  $\hat{\sigma}_1 \leq \hat{\sigma}_2 \leq \dots \leq \hat{\sigma}_i \leq 0$ . Cabe destacar que el análisis de correlaciones canónicas no puede detectar directamente procesos I(2), pues en estos casos tan sólo  $\sigma_j = 1$ . No obstante, esto no causa un problema en la práctica, puesto que al diferenciar la serie y analizar los CCC de la variable transformada se obtendrá la segunda raíz igual a la unidad. La función de pérdida que especifican BW para la ecuación (10) es,

$$C(T) = \frac{\log(T)^2}{T} \quad (11)$$

que supone un decaimiento de la penalización a medida que la muestra crece.

Existen dos problemas fundamentales en la utilización de  $C(T)$ : 1) su uso sólo es correcto si el número de raíces unitarias del sistema ( $d$ ), es igual a la unidad y, 2) no depende de la dimensión del subespacio de información pasada,  $i$ . La primera cuestión es muy relevante, ya que no proporciona buenos resultados en procesos multivariantes con dos o más raíces unitarias. La segunda cuestión supone que cambios en la dimensión del subespacio de información pasada pueden producir cambios importantes en los valores calculados de los CCC. Básicamente, cuanto mayor es la dimensión de los subespacios de información más elevados son los CCC, al igual que un coeficiente de determinación crece cuando aumentan las variables explicativas.

Para resolver estas dos limitaciones, en este trabajo se formula un criterio de decisión de la familia del propuesto por BW que, 1) está basado tanto en la idea de Poskitt (2000), como en un estudio de simulación y, 2) es aplicable a sistemas con múltiples raíces unitarias.

La idea clave es obtener una función de penalización que refleje el límite entre los  $\hat{\sigma}_j$  que se consideren iguales a la unidad y los inferiores, a partir de la propiedad de convergencia a su verdadero valor a tasas diferentes. Para ello, se definen dos condiciones que debe cumplir esta función.

- *Condición 3.1:* Debe depender del número de observaciones ( $T$ ), de la dimensión del subespacio de información pasada ( $i$ ) y del número de raíces unitarias ( $d$ ) que se desee contrastar.
- *Condición 3.2:* Debe aproximarse a cero cuando  $T$  tienda a infinito, para asegurar la consistencia del estimador del número de raíces unitarias,  $\hat{d}$ . Formalmente,  $\lim_{T \rightarrow \infty} G_l(T, i, d) = 0$ , donde  $G_l$  representa las posibles funciones de pérdida.

La forma más general de la familia de criterios propuestos puede formularse como:

$$f(\hat{\sigma}_j) - G_l(T, i, d) \leq 0 \quad (12)$$

donde  $f(\hat{\sigma}_j)$  es una transformación que mide la distancia entre el  $j$ -ésimo VS y la unidad. Cada una de las funciones de pérdida depende del número de observaciones, de la dimensión

del subespacio de información y del número de raíces unitarias que se desee evaluar. En este trabajo, se usan como  $f(\hat{\sigma}_j)$  las expresiones  $1 - \hat{\sigma}_j^2$  y  $1 - \hat{\sigma}_j$ , aunque puede haber un continuo de funciones válidas. La igualdad en la expresión (12) se requiere para que cuando  $\lim_{T \rightarrow \infty} f(\hat{\sigma}_j) = \lim_{T \rightarrow \infty} G_l(T, i, d) = 0$ , el criterio devuelva estimaciones consistentes del número de raíces unitarias.

### 3.1. Procesos univariantes

Con el fin de encontrar una forma funcional concreta a  $G_l(T, i, d)$ , se estudia la distribución empírica de los  $\hat{\sigma}_j$  que corresponden a raíces no unitarias. Para ello, se simulan 1000 repeticiones de modelos cercanos a la no estacionariedad para todos los instantes temporales  $t$ , desde  $t = 20, 21, \dots, 500$ . Al primer VS resultante de cada una de estas simulaciones se le aplica la transformación  $f(\cdot)$ , resumiendo toda esta información de cada instante en un determinado percentil, que queda representado por la serie  $y_{1t}$ . De esta forma, se obtiene una función empírica que mide la distancia a cero de  $\hat{\sigma}_1$ . El problema es encontrar una función de penalización que cumpla las condiciones (3.1) y (3.2) y se ajuste correctamente a  $y_{1t}$ .

Para el caso en que  $d = 0$ , se considera el modelo loglineal siguiente:

$$\log(y_{1t}) = \alpha + \beta_1 \log(t) + \beta_2 \log(i) + \varepsilon_t \quad (13)$$

donde  $\varepsilon_t$  corresponde a un término de error e  $i$  hace referencia a la dimensión del subespacio de información pasada y futura, ya que se asume  $p = f$ .

Una vez definida  $y_{1t}$  se tienen todas las variables necesarias para estimar (13). Esta ecuación genera un tipo particular de función de penalización para cada conjunto de valores de los parámetros estimados  $\alpha, \beta_1$  y  $\beta_2$ . Además, éstos variarán dependiendo del valor del CCC generador de las simulaciones ( $\bar{\sigma}$ ) y del percentil elegido ( $\tau$ ), indispensables para definir la variable  $y_{jt}$ . Por ello, se realizan múltiples simulaciones alterando estos dos valores. El Cuadro 1 muestra los parámetros estimados del modelo (13) para diferentes funciones  $f(\cdot)$  y valores de  $\tau$  y de  $\bar{\sigma}$ .

[INSERTAR CUADRO 1]

El Cuadro 1 incluye, a efectos comparativos, la función de penalización propuesta por BW. Cuando se ajusta el modelo (13) a la serie generada por esta función, puede observarse que sus valores estimados no se distancian demasiado del resto. Por otro lado, como es lógico, los resultados muestran que diferencias en los valores que definen la distribución empírica conducen a distintas funciones de penalización. Estas se denotan por  $G_l$ , con  $l = 1, 2, \dots, 6$ .

[INSERTAR FIGURA 1]

Si  $y_{1t}$  representa suficientemente bien la frontera entre los CCC iguales y los inferiores a la unidad, entonces la función  $G_4$ , que se ajusta correctamente a la distribución empírica seleccionada, es una función de pérdida válida para evaluar si  $d > 0$ . El problema es ahora decidir cuál de todas las posibles series  $y_{1t}$  refleja mejor la frontera que discrimina los CCC. Se considera que, para elegir la función de penalización óptima, debe maximizarse tanto el tamaño como la potencia del criterio de decisión asociado, definidos respectivamente como la probabilidad de que el criterio acierte cuando se cumple la hipótesis nula y la probabilidad de que el criterio acierte cuando ésta no se cumple. Desafortunadamente, en este caso, tamaño y potencia se mueven en sentido opuesto. Por ello, decidimos restringir las distribuciones empíricas a aquellas que cumplan al menos una de las dos condiciones siguientes:

- *Condición 3.3:* Maximizar el tamaño, con una potencia del criterio consistente, es decir que tiende a la unidad cuando aumenta la dimensión de la muestra.
- *Condición 3.4:* Maximizar el tamaño, exigiendo una potencia mínima para un número de observaciones determinado a priori.

Cada una de ellas puede ser utilizada para valorar la proximidad de una serie a la no estacionariedad. Para el caso univariante, cuya decisión básica es si  $d > 0$ , es razonable exigir que la potencia mínima en la segunda condición sea igual al 50 % para  $T = 50$ . Para seleccionar las funciones de penalización y con ellas los criterios definitivos, se utiliza como una aproximación al tamaño del criterio en modelos univariantes, la probabilidad de estimar correctamente el número de raíces unitarias en un paseo aleatorio,  $\nabla z_t = a_t$ . Asimismo, se aproxima la potencia como la probabilidad de que  $\hat{d} = d$  en un modelo estacionario

autorregresivo persistente del tipo  $(1 - .9B)z_t = a_t$ . Los resultados de los ejercicios de simulación de estos modelos con 1000 repeticiones se presentan en los Cuadros 2 y 3.

[INSERTAR CUADROS 2 Y 3]

Los Cuadros 2 y 3 muestran la consistencia, tanto en tamaño como en potencia, de todos los criterios presentados. Resulta relevante observar que varios de los criterios batan al test ADF tanto en potencia como en tamaño. A la vista de los resultados, se elige como función de penalización que cumple la condición (3.3) a  $G_5$ , que pasa a denominarse  $G_a(d > 0)$ . Además, se selecciona la función de pérdida  $G_2$ , de aquí en adelante  $G_b(d > 0)$ , que presenta una potencia del 50 % con un tamaño aceptable y consistente.

### 3.2. Procesos multivariantes

La principal ventaja de la metodología propuesta está en el tratamiento de modelos multivariantes. En primer lugar, tan sólo unos pocos autores como Phillips y Durlauf (1986), Abuaf y Jorion (1990), Flôres et al. (1999) o BW tratan la identificación de raíces unitarias en vectores de series. En segundo lugar, la posibilidad de estimar las funciones de penalización en este tipo de modelos abre las puertas al análisis de cointegración de las variables del sistema.

Como se ha visto, las funciones de penalización estimadas no sólo dependen del tamaño muestral y de la dimensión del subespacio de información pasada, sino también del mínimo número de raíces unitarias que se desee contrastar. Esto implica que cada  $\hat{\sigma}_j$  candidato a ser igual a la unidad, debe ser comparado con una función de penalización diferente. El criterio desarrollado para procesos univariantes indica si la serie analizada es estacionaria o si, por el contrario,  $d > 0$ . En este apartado, se desarrollan los criterios para contrastar secuencialmente si  $d > 1$ ,  $d > 2$ ,  $d > 3$  y  $d > 4$  de forma que, si los criterios aplicados a  $\hat{\sigma}_1$  y  $\hat{\sigma}_2$  señalan que  $d > 1$  y  $d > 2$ , pero rechazan  $d > 3$  para  $\hat{\sigma}_3$ , entonces el algoritmo de detección devolverá la solución  $\hat{d} = 2$ . Para estimar las funciones de penalización se simulan vectores de dos, tres, cuatro y cinco series con VS iguales y distintos a la unidad en cada caso. En cada sistema, se genera al menos un CCC cercano a la unidad, pero no unitario. Después, se estiman distintas distribuciones empíricas para cada proceso, como en el caso

univariante, adaptando un modelo que se ajuste adecuadamente.

La Figura 2 compara las distribuciones empíricas  $y_{1t}$  e  $y_{2t}$  correspondientes a los pares  $(\tau = 90, \bar{\sigma} = 1.0)$  y  $(\tau = 90, \bar{\sigma} = .95)$ , de un sistema bivariante. Además, muestra la función de penalización ajustada a  $y_{1t}$ ,  $G_4(d > 1)$ , y la función de pérdida elegida que representaba la cota superior más baja de los dos criterios seleccionados en la sección anterior,  $G_b(d > 0)$ . Obsérvese que el primer VS, que corresponde a una raíz unitaria del proceso bivariante, se mantiene muy por debajo de  $G_b(d > 0)$  utilizado en el criterio con mayor potencia de los elegidos para contrastar  $d > 0$ . En efecto, si comparamos un proceso multivariante  $(\mathbf{A}_t)$  con  $d$  raíces unitarias, con un mismo sistema  $(\mathbf{B}_t)$  que incorpora una nueva serie no estacionaria (el proceso  $\mathbf{B}_t$  tendría  $d + 1$  raíces iguales a la unidad), entonces  $\hat{\sigma}_j(\mathbf{B}_t) \geq \hat{\sigma}_j(\mathbf{A}_t)$  para  $j = 1, \dots, d + 1$ . Por lo tanto, aprovechando la ordenación de mayor a menor de los VS que devuelve la descomposición SVD, es válido usar los criterios expuestos de forma secuencial, para decidir cuántos de ellos son iguales a la unidad.

[INSERTAR FIGURA 2]

El Cuadro 4 presenta los parámetros estimados del modelo (13) correspondientes a cuatro posibles funciones de penalización para contrastar  $d > 1$ . Puede observarse en la Figura 2 que el modelo especificado para representar las distribuciones empíricas que se usan para contrastar  $d > 0$  también se ajusta correctamente a las que se utilizan para decidir si  $d > 1$ , salvo quizá, en las primeras observaciones.

[INSERTAR CUADRO 4]

Para discriminar entre las posibles funciones de penalización propuestas, se utilizan las condiciones (3.3) y (3.4), establecidas anteriormente. Los resultados de las simulaciones realizadas para evaluar el tamaño y la potencia de los criterios formulados para  $d > 1$  se exponen en los Cuadros 5 y 6. En ellos se refleja la probabilidad de que los diferentes criterios acierten el número de raíces unitarias del proceso generador, utilizando la función de penalización  $G_b(d > 0)$  para decidir si  $d > 0$ .

[INSERTAR CUADROS 5 Y 6]

Para el análisis de procesos bivariantes, se escogen las funciones de penalización  $G_4$  y  $G_1$  del Cuadro 4. Estas se denominan, de aquí en adelante,  $G_a(d > 1)$  y  $G_b(d > 1)$ , respectivamente. La primera maximiza el tamaño dada una potencia débil, pero consistente, y la segunda lo hace dada una potencia mínima del 45 % para  $T = 50$ . Se decide disminuir la potencia mínima exigida con respecto de  $d > 0$ , al advertir que las propiedades de los criterios, tanto en tamaño como en potencia, se degradan ligeramente en muestras cortas cuando aumenta  $d$ . De nuevo, los cuadros reflejan que todos los criterios estiman el número de raíces unitarias consistentemente.

El método de obtención de las funciones de pérdida para procesos trivariantes, tetravariantes y pentavariantes es análogo al expuesto hasta ahora. Sin embargo, el modelo especificado para representarlos adecuadamente es distinto. De hecho, para mejorar el ajuste a las distribuciones empíricas generadas con estos procesos, decidimos modelizar dos submuestras, proponiendo una ecuación específica para representar cada una de ellas. Los modelos ajustados son los siguientes:

$$\begin{aligned} t < t_j^* & \quad y_{jt} &= \alpha_{j1} + \delta_{1j}t + \delta_{2j}t^2 + \delta_{3j}t^3 + \delta_{4j}i + \varepsilon_{1jt} \\ t \geq t_j^* & \quad \log(y_{jt}) &= \alpha_{2j} + \beta_{1j} \log(t) + \beta_{2j} \log(i) + \varepsilon_{2jt} \end{aligned} \quad (14)$$

donde  $t_j^*$  se refiere a la primera realización de la segunda submuestra.

Las estimaciones de las dos funciones de penalización que cumplen las condiciones (3.3) y (3.4), para contrastar la existencia de al menos  $j - 1$  raíces unitarias (con  $j = 3, 4, 5$ ) y los parámetros estimados de los modelos propuestos en (14), se ofrecen en el Cuadro 7. De nuevo, se rebaja la potencia mínima exigida para  $T = 50$ , que hace referencia a la condición (3.4): 40 % para  $d > 2, 3, 4$ . Obsérvese, en el Cuadro 7, cómo aumentar la dimensión del subespacio de información pasada  $i$  afecta positivamente a  $\hat{\sigma}_j$  y, por tanto, negativamente a  $f(\hat{\sigma}_j)$ , para todo  $j$ . Además, cuanto mayor es  $j$  mayor es la sensibilidad de  $y_{jt}$  ante cambios en  $i$ .

[INSERTAR CUADRO 7]

Los procesos generadores y las simulaciones realizadas para valorar el tamaño y la potencia de los criterios que usan las funciones de penalización del Cuadro 7 se presentan en los Cuadros 8 y 9. Todos los criterios estiman  $d$  consistentemente, aunque los resultados revelan un deterioro en sus propiedades cuando aumenta el número de raíces unitarias a contrastar. En este punto se decide no estimar funciones de pérdida para contrastar  $d > 5$ , por varias razones: 1) la utilización econométrica de modelos con más de cinco series temporales es realmente escasa, 2) el uso de estos criterios para procesos con una dimensión mayor pero con un número, a priori, no superior a cinco raíces unitarias es también válido y 3) el deterioro que muestran los criterios con el aumento de  $d$  y la complejidad del análisis cuando aumenta la dimensión del sistema no motiva la búsqueda de funciones de penalización para procesos con un mayor número de raíces iguales a la unidad.

[INSERTAR CUADROS 8 Y 9]

## 4. EJERCICIOS CON DATOS REALES

En esta sección, se utilizan dos conjuntos de series muy analizadas en la literatura econométrica. En los estudios del primero de ellos no se encuentran indicios de cointegración, mientras que lo contrario ocurre en el segundo vector de series escogido.

### 4.1. Series de Precios de la Harina

El vector de series analizadas ( $\mathbf{Y}_t$ ) representa los Índices mensuales de Precios de la Harina en las ciudades de Buffalo, Minneapolis y Kansas City entre los meses de Agosto de 1972 y Noviembre de 1980. Fueron presentadas por Tiao y Tsay (1989) para ilustrar su metodología de especificación y más tarde analizadas por Grubb (1992), Lütkepohl y Poskitt (1996) o Casals (1997).

[INSERTAR FIGURA 3]

En primer lugar, se analiza el orden de integración de cada una de las series por separado. Esto se hace con las dos funciones de penalización obtenidas:  $G_a$  y  $G_b$ . Los resultados del Cuadro 10 señalan la existencia de al menos una raíz unitaria en cada una de las series. Además, cuando las series se diferencian una vez, el VS más grande, correspondiente a la

serie de precios de Minneapolis, resulta igual a 0.257, muy inferior a la unidad. Esto hace que el criterio descarte para cualquiera de las series un grado de integración de orden 2. Por tanto, la conclusión es que cada una de las series presenta una, y sólo una, raíz unitaria. Este es el resultado más común de la literatura (Grubb, 1992; Lütkepohl y Poskitt, 1996; Casals, 1997).

[INSERTAR CUADRO 10]

Una vez decidido que cada serie es  $I(1)$ , es posible estudiar la existencia de cointegración determinando el orden de integración del vector de las tres series. Los resultados de este análisis se presentan en el Cuadro 11. Se observa que el criterio propuesto encuentra tres raíces unitarias en el sistema, independientemente de la función de pérdida especificada. Es decir, el proceso tiene tres tendencias, cada una perteneciente a una de las series que conforman el sistema, descartando la existencia de cointegración. Otros autores como Tiao y Tsay (1989), con los llamados modelos de componentes escalares, o Lütkepohl y Poskitt (1996), con un análisis tradicional, también han rechazado la hipótesis de cointegración.

[INSERTAR CUADRO 11]

## 4.2. Series de tipos de interés de EEUU

En este caso se analiza el orden de integración de diversas series de tipos de interés de la economía de Estados Unidos entre agosto de 1985 y enero de 2003. Martín Manjón y Treadway (1997) estudian la relación existente entre estas series, encontrando que los tipos con un vencimiento inferior o igual a dos años presentan una tendencia común a largo plazo con el tipo objetivo de la Reserva Federal (Fed). En este trabajo se analiza el tipo objetivo de la Fed conjuntamente con el tipo de interés efectivo y los tipos a tres y seis meses del mercado secundario, proceso denotado por  $Z_t$ .

[INSERTAR FIGURA 4]

El análisis de cada serie por separado concluye que todas ellas son  $I(1)$ , independientemente de la función de pérdida usada. Al diferenciar las series, ninguna de ellas genera un VS suficientemente elevado como para que el criterio encuentre una segunda raíz unitaria.

El Cuadro 12 resume los resultados del estudio del vector de series. En este caso, el criterio revela la existencia de una única raíz unitaria. Por tanto, la tendencia que presenta cada serie es común. De este modo, se revela la existencia de tres relaciones de cointegración, una por cada raíz unitaria que desaparece en el análisis conjunto de los datos. Este es el mismo resultado que obtienen Martín Manjón y Treadway (1997) con un análisis de cointegración (en modelos bivariantes) basado en contrastes de no invertibilidad sobre la posible combinación lineal estacionaria diferenciada.

[INSERTAR CUADRO 12]

## 5. CONCLUSIONES

En este trabajo se propone una metodología de detección de raíces unitarias basada en métodos de subespacios y en el análisis de correlaciones canónicas. El método se basa en dos nuevos criterios de información que permiten estimar el número de raíces unitarias, tanto en el marco univariante como multivariante.

Las funciones de penalización subyacentes a los criterios propuestos se ajustan mediante técnicas de simulación. Esta idea sugiere de forma natural la posibilidad de adaptar el criterio de información a muestras con propiedades estadísticas especiales. Por ejemplo, se sabe que los rendimientos financieros observados en alta frecuencia suelen presentar distribuciones con colas más gruesas que las de una normal y, ocasionalmente, una asimetría apreciable. En este contexto, nuestra metodología permite plantearse el diseño de un criterio optimizado para este tipo de muestras.

El trabajo incluye ejercicios de simulación que analizan las propiedades de estos criterios en muestras finitas. Este análisis indica que los procedimientos propuestos tienen una capacidad apreciable para estimar el número de raíces unitarias y el rango de cointegración en distintas situaciones. Además a diferencia de otros procedimientos, como ocurre en el método de Johansen (1988, 1991) que requiere una especificación previa tipo VAR, la metodología presentada no precisa ajustar un modelo a los datos, con la consiguiente disminución de riesgo de error.

## Referencias

- Abuaf, N. y Jorion, P. (1990). Purchasing Power Parity in the Long Run. *The Journal of Finance*, vol. 45, 157–174.
- Akaike, H. (1975). Markovian Representation of Stochastic Processes by Canonical Variables. *SIAM Journal of Control*, vol. 13, núm. 1, 162–173.
- Akaike, H. (1976). *Canonical Correlation Analysis of Time Series and the Use of an Information Criterion*. Academic Press.
- Aoki, M. (1990). *State Space Modelling of Time Series*. Springer Verlag, New York.
- Aström, K. J. y Eykhoff, P. (1971). System Identification, A Survey. *Automatica*, vol. 7, 123–162.
- Bauer, D. (1998). *Some Asymptotic Theory for the Estimation of Linear Systems Using Maximum Likelihood Methods or Subspace Algorithms*. Tesis Doctoral, TU Wien.
- Bauer, D. y Wagner, M. (2002). Estimating Cointegrated Systems Using Subspace Algorithms. *Journal of Econometrics*, vol. 111, 47–84.
- Box, G. E. P. y Jenkins, G. M. (1976). *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco, 2<sup>a</sup> ed.
- Casals, J. (1997). *Métodos de Subespacios en Econometría*. Tesis Doctoral, Universidad Complutense de Madrid.
- Chui, N. L. C. (1997). *Subspace Methods and Informative Experiments for System Identification*. Tesis Doctoral, Pembroke College Cambridge.
- Dickey, D. A. y Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with Unit Root. *Journal of the American Statistical Association*, vol. 74, 427–431.
- Eckart, C. y Young, G. (1936). The Approximation of One Matrix by Another of Lower Rank. *Psychometrika*, vol. 1, 211–218.
- Flores, R. G., Jorion, P., Preumont, P. Y., y Szafarz, A. (1999). Multivariate Unit Roots Test of PPP Hypothesis. *Journal of Empirical Finance*, vol. 6, 335–353.
- García Hiernaux, A. (2005). *Identificación de Modelos para Series Temporales mediante Métodos de Subespacios*. Tesis Doctoral, Universidad Complutense de Madrid.
- Grubb, H. (1992). A Multivariate Time Series Analysis of Some Flour Price Data. *Applied Statistics*, vol. 41, 95–107.
- Ho, B. y Kalman, R. (1966). Effective Construction of Linear State-Variable Models from Input-Output Functions. *Regelungstechnik*, vol. 14, 545–548.
- Jenkins, G. M. y Alavi, A. S. (1981). Some Aspects of Modelling and Forecasting Multivariate Time Series. *Journal of Time Series Analysis*, vol. 2, núm. 1, 1–47.
- Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, vol. 12, 231–254.

- Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, vol. 59, 1551–1580.
- Larimore, W. E. (1983). System Identification, Reduced-Order Filtering and Modeling Via Canonical Variate Analysis. *Proceedings of the American Control Conference, ACC, San Francisco*, páginas 445–451.
- Larimore, W. E. (1990). Canonical Variate Analysis in Identification, Filtering and Adaptive Control. *Proceedings of the 29th Conference on Decision and Control, Hawaii*, páginas 596–604.
- Lütkepohl, H. y Poskitt, D. S. (1996). Specification of Echelon Form VARMA Models. *Journal of Business and Economic Statistics*, vol. 14, núm. 1, 69–79.
- Martín Manjón, R. y Treadway, A. (1997). The Fed Controls Only One of the two Interest Rates in the U.S. Economy. *Documento de Trabajo del ICAE*, , núm. 9716.
- Phillips, P. B. C. y Durlauf, N. S. (1986). Multiple Time Series Regression with Integrated Process. *Review of Economic Studies*, vol. 53, 473–495.
- Poskitt, D. S. (2000). Strongly Consistent Determination of Cointegrating Rank Via Canonical Correlations. *Journal of Business and Economic Statistics*, vol. 18, 77–90.
- Terceiro, J. (1990). *Estimation of Dynamics Econometric Models with Errors in Variables*. Springer-Verlag, Berlin.
- Tiao, G. C. y Tsay, R. S. (1989). Model Specification in Multivariate Time Series. *Journal of the Royal Statistical Society, B Series*, vol. 51, núm. 2, 157–213.
- Van Overschee, P. y De Moor, B. (1994). N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems. *Automatica*, vol. 30, núm. 1, 75–93.
- Viberg, M. (1995). Subspace-Based Methods for the Identification of the Linear Time-Invariant Systems. *Automatica*, vol. 31, núm. 12, 1835–1852.

Cuadro 1: Funciones de penalización estimadas para contrastar  $d > 0$

$f(\hat{\sigma}_1) = y_{1t}$	Distribución Empírica		Valores Estimados			Función de Penalización
	$\bar{\sigma}$	$\tau$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	
$1 - \hat{\sigma}_1^2$	0.975	95	1.09	-0.51	-0.17	$G_1(d > 0)$
$1 - \hat{\sigma}_1^2$	0.975	90	0.60	-0.50	-0.10	$G_2(d > 0)$
$1 - \hat{\sigma}_1^2$	0.95	80	0.10	-0.36	-0.10	$G_3(d > 0)$
$1 - \hat{\sigma}_1^2$	0.95	90	0.52	-0.40	-0.10	$G_4(d > 0)$
$1 - \hat{\sigma}_1$	0.95	90	0.10	-0.44	-0.05	$G_5(d > 0)$
$1 - \hat{\sigma}_1$	0.975	95	0.65	-0.56	-0.16	$G_6(d > 0)$
$\log(T)^2/T^*$	-	-	0.98	-0.52	-0.10	$C(T)$

\* Corresponde a la función de penalización de Bauer y Wagner (2002)

Cuadro 2: Tamaño de los criterios para  $d > 0$   
a partir del proceso  $\phi(B)z_t = a_t$

T	ADF <sup>†</sup>	$C^\ddagger$	$G_1$	$G_2$	$G_3$	$G_4$	$G_5$	$G_6$
50	0.950	0.953	0.960	0.879	0.888	0.953	0.970	0.956
100	0.951	0.984	0.986	0.935	0.958	0.989	0.995	0.984
500	0.939	0.996	0.997	0.996	0.997	1.0	1.0	0.997

<sup>†</sup> Test aumentado de Dickey y Fuller (1979) al 95 % con un retardo.

<sup>‡</sup> Corresponde a la función de penalización de Bauer y Wagner (2002).

$\phi(B) = 1 - B$ .

Cuadro 3: Potencia de los criterios para  $d > 0$   
a partir del proceso  $\phi(B)z_t = a_t$

T	ADF <sup>†</sup>	$C^\ddagger$	$G_1$	$G_2$	$G_3$	$G_4$	$G_5$	$G_6$
50	0.114	0.246	0.208	0.505	0.464	0.246	0.188	0.230
100	0.341	0.451	0.420	0.762	0.628	0.376	0.290	0.451
500	1.0	1.0	1.0	1.0	1.0	0.994	0.981	1.0

<sup>†</sup> Test aumentado de Dickey y Fuller (1979) al 95 % con un retardo.

<sup>‡</sup> Corresponde a la función de penalización de Bauer y Wagner (2002).

$\phi(B) = 1 - .9B$ .

Cuadro 4: Funciones de penalización estimadas para  $d > 1$

Distribución Empírica $f(\hat{\sigma}_2) = y_{2t}$	Valores Estimados		Función de Penalización			
	$\bar{\sigma}$	$\tau$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	
$1 - \hat{\sigma}_2^2$	0.95	75	0.43	-0.39	-0.07	$G_1$
$1 - \hat{\sigma}_2^2$	0.95	80	0.50	-0.39	-0.07	$G_2$
$1 - \hat{\sigma}_2^2$	0.95	85	0.57	-0.39	-0.07	$G_3$
$1 - \hat{\sigma}_2^2$	0.95	90	0.67	-0.39	-0.06	$G_4$

Cuadro 5: Tamaño de los criterios para  $d > 1$

a partir del proceso  $\Phi(\mathbf{B})\mathbf{z}_t = \mathbf{a}_t$

T	$G_1$	$G_2$	$G_3$	$G_4$
50	0.708	0.779	0.837	0.901
100	0.854	0.891	0.924	0.953
500	0.997	0.998	0.999	0.999

$\Phi(\mathbf{B}) = \text{diag}(\nabla, \nabla)$

Cuadro 6: Potencia de los criterios para  $d > 1$

a partir del proceso  $\Phi(\mathbf{B})\mathbf{z}_t = \mathbf{a}_t$

T	$G_1$	$G_2$	$G_3$	$G_4$
50	0.452	0.373	0.293	0.204
100	0.516	0.428	0.337	0.239
500	0.993	0.989	0.968	0.947

$\Phi(\mathbf{B}) = \text{diag}(\nabla, (1 - .9B))$

Cuadro 7: Funciones de penalización propuestas para  $d > 2$ ,  $d > 3$  y  $d > 4$

$j$	Distribución Empírica			$t < t^*$				$t \geq t^*$			Función de Penalización	
	$f(\hat{\sigma}_j) = y_{jt}$	$\bar{\sigma}$	$\tau$	$\hat{\alpha}_1$	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{\delta}_4$	$\hat{\alpha}_0$	$\hat{\beta}_1$		$\hat{\beta}_2$
3	$1 - \hat{\sigma}_j^2$	0.95	90	-0.305	0.040	$-6.5 \times 10^{-4}$	$3.3 \times 10^{-6}$	-	0.786	-0.328	-0.226	$G_a(d > 2)$
3	$1 - \hat{\sigma}_j^2$	0.95	60	-0.353	0.036	$-5.9 \times 10^{-4}$	$3.0 \times 10^{-6}$	-	0.188	-0.285	-0.172	$G_b(d > 2)$
4	$1 - \hat{\sigma}_j^2$	0.95	90	-0.635	0.044	$-6.0 \times 10^{-4}$	$2.7 \times 10^{-6}$	-	1.589	-0.437	-0.365	$G_a(d > 3)$
4	$1 - \hat{\sigma}_j^2$	0.95	60	-0.621	0.041	$-5.6 \times 10^{-4}$	$2.5 \times 10^{-6}$	-	1.557	-0.469	-0.418	$G_b(d > 3)$
5	$1 - \hat{\sigma}_j^2$	0.95	90	-0.317	0.032	$-3.2 \times 10^{-4}$	$1.1 \times 10^{-6}$	-0.076	1.313	-0.383	-0.280	$G_a(d > 4)$
5	$1 - \hat{\sigma}_j^2$	0.95	60	-0.366	0.030	$-3.0 \times 10^{-4}$	$1.0 \times 10^{-6}$	-0.063	1.131	-0.361	-0.378	$G_b(d > 4)$

Cuando  $j = 3, 4$ , entonces  $t^* = 88$ ;

Cuando  $j = 5$ , entonces  $t^* = 121$ .

Cuadro 8: Tamaño de los criterios para  $d > 2$ ,  $d > 3$  y  $d > 4$   
a partir de los procesos  $\Phi_k(\mathbf{B})\mathbf{z}_t = \mathbf{a}_t$

T	$d > 2$		$d > 3$		$d > 4$	
	$G_a$	$G_b$	$G_a$	$G_b$	$G_a$	$G_b$
50	0.897	0.601	0.903	0.603	0.882	0.600
100	0.919	0.669	0.921	0.644	0.934	0.626
500	0.998	0.996	0.997	0.926	0.990	0.966

$\Phi_1(\mathbf{B}) = \text{diag}(\nabla, \nabla, \nabla)$ ;  $\Phi_2(\mathbf{B}) = \text{diag}(\nabla, \nabla, \nabla, \nabla)$ ;  
 $\Phi_3(\mathbf{B}) = \text{diag}(\nabla, \nabla, \nabla, \nabla, \nabla)$

Cuadro 9: Potencia de los criterios para  $d > 2$ ,  $d > 3$  y  $d > 4$   
a partir del proceso  $\Phi_k(\mathbf{B})\mathbf{z}_t = \mathbf{a}_t$

T	$d > 2$		$d > 3$		$d > 4$	
	$G_a$	$G_b$	$G_a$	$G_b$	$G_a$	$G_b$
50	0.920	0.402	0.940	0.398	0.110	0.408
100	0.161	0.487	0.166	0.460	0.144	0.438
500	0.741	0.951	0.854	0.994	0.458	0.844

$\Phi_1(\mathbf{B}) = \text{diag}(\nabla, \nabla, (1 - .9B))$ ;  $\Phi_2(\mathbf{B}) = \text{diag}(\nabla, \nabla, \nabla, (1 - .9B))$ ;  
 $\Phi_3(\mathbf{B}) = \text{diag}(\nabla, \nabla, \nabla, \nabla, (1 - .9B))$

Cuadro 10: Estudio del orden de integración de los Precios de la Harina

$j$	$\hat{\sigma}_j$	$f(\hat{\sigma}_j) - G_a(d > j - 1)$	$f(\hat{\sigma}_j) - G_b(d > j - 1)$	$\hat{d}_{G_a}$	$\hat{d}_{G_b}$
<i>Buffalo</i>					
1	0.9901	$0.0099 - 0.1344 < 0$	$0.0197 - 0.1551 < 0$	1	1
2	0.0168	$0.9997 - 0.2945 > 0$	$0.9997 - 0.2279 > 0$	1	1
<i>Minnesota</i>					
1	0.9902	$0.0098 - 0.1344 < 0$	$0.0195 - 0.1551 < 0$	1	1
2	0.0197	$0.9996 - 0.2945 > 0$	$0.9996 - 0.2279 > 0$	1	1
<i>Kansas City</i>					
1	0.9901	$0.0099 - 0.1344 < 0$	$0.0197 - 0.1551 < 0$	1	1
2	0.0219	$0.9995 - 0.2945 > 0$	$0.9995 - 0.2279 > 0$	1	1

$\hat{\sigma}_j$  se corresponde al  $j$ -ésimo valor singular.  $G_a$  y  $G_b$  son las funciones de pérdida propuestas.  $\hat{d}_{G_a}$  y  $\hat{d}_{G_b}$  es el número de raíces unitarias encontrado al analizar los  $j$  primeros valores singulares.

*Cuadro 11: Estudio del orden de integración del proceso  $\mathbf{Y}_t$*

$j$	$\hat{\sigma}_j$	$f(\hat{\sigma}_j) - G_a(d > j - 1)$	$f(\hat{\sigma}_j) - G_b(d > j - 1)$	$\hat{d}_{G_a}$	$\hat{d}_{G_b}$
1	0.9912	$0.0088 - 0.1344 < 0$	$0.0175 - 0.1551 < 0$	1	1
2	0.9420	$0.1126 - 0.2945 < 0$	$0.1126 - 0.2279 < 0$	2	2
3	0.8837	$0.2191 - 0.3370 < 0$	$0.2191 - 0.2460 < 0$	3	3
4	0.5117	$0.7382 - 0.3649 > 0$	$0.7382 - 0.2788 > 0$	3	3

$\hat{\sigma}_j$  se corresponde al  $j$ -ésimo valor singular.  $G_a$  y  $G_b$  son las funciones de pérdida propuestas.  $\hat{d}_{G_a}$  y  $\hat{d}_{G_b}$  es el número de raíces unitarias encontrado al analizar los  $j$  primeros valores singulares.

*Cuadro 12: Estudio del orden de integración del proceso  $\mathbf{Z}_t$*

$j$	$\hat{\sigma}_j$	$f(\hat{\sigma}_j) - G_a(d > j - 1)$	$f(\hat{\sigma}_j) - G_b(d > j - 1)$	$\hat{d}_{G_a}$	$\hat{d}_{G_b}$
1	0.9952	$0.0048 - 0.0970 < 0$	$0.0095 - 0.1070 < 0$	1	1
2	0.8797	$0.2261 - 0.2205 > 0$	$0.2261 - 0.1797 > 0$	1	1
3	0.7324	$0.4636 - 0.2648 > 0$	$0.4636 - 0.1991 > 0$	1	1
4	0.4028	$0.8377 - 0.2642 > 0$	$0.8377 - 0.1968 > 0$	1	1

$\hat{\sigma}_j$  se corresponde al  $j$ -ésimo valor singular.  $G_a$  y  $G_b$  son las funciones de pérdida propuestas.  $\hat{d}_{G_a}$  y  $\hat{d}_{G_b}$  es el número de raíces unitarias encontrado al analizar los  $j$  primeros valores singulares.

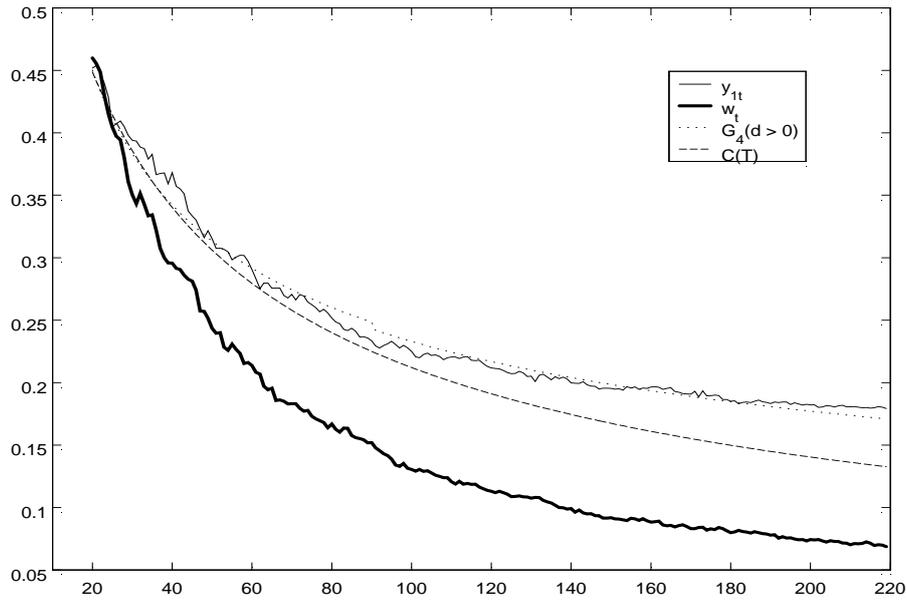


Figura 1: Valores Singulares (VS) simulados y funciones de penalización para procesos univariantes.  $y_{1t}$  corresponde al primer VS de un proceso estacionario,  $w_t$  corresponde al primer VS de un proceso no estacionario.  $G_4(d > 0)$  es una función de penalización propuesta para procesos univariantes y  $C(T)$  es la función de penalización utilizada por Bauer y Wagner (2002).  $y_{1t}$  y  $w_t$  reflejan sensiblemente la diferencia entre las tasas de convergencia de los VS a sus valores asintóticos.

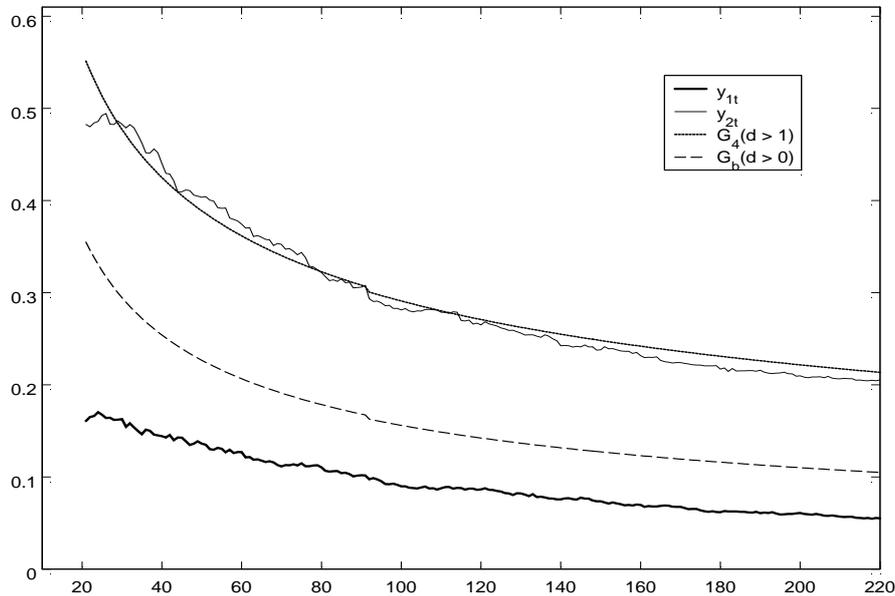


Figura 2: Valores Singulares (VS) simulados y funciones de penalización para procesos bivalentes.  $y_{1t}$  corresponde al primer VS de un sistema bivalente no estacionario (cuyos procesos univariantes son  $I(1)$  e  $I(0)$ ).  $y_{2t}$  corresponde al segundo VS del mismo sistema,  $G_4(d > 1)$  es una de las funciones de penalización propuestas para detectar al menos dos raíces unitarias y  $G_b(d > 0)$  es la función de penalización usada para detectar al menos una raíz unitaria.

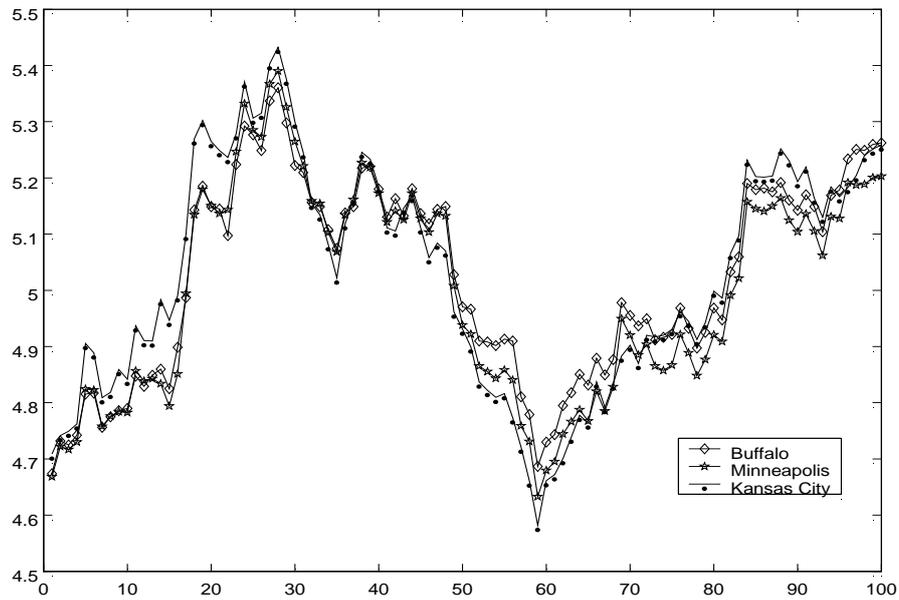


Figura 3: Logaritmo de los Precios de la Harina en las diferentes ciudades

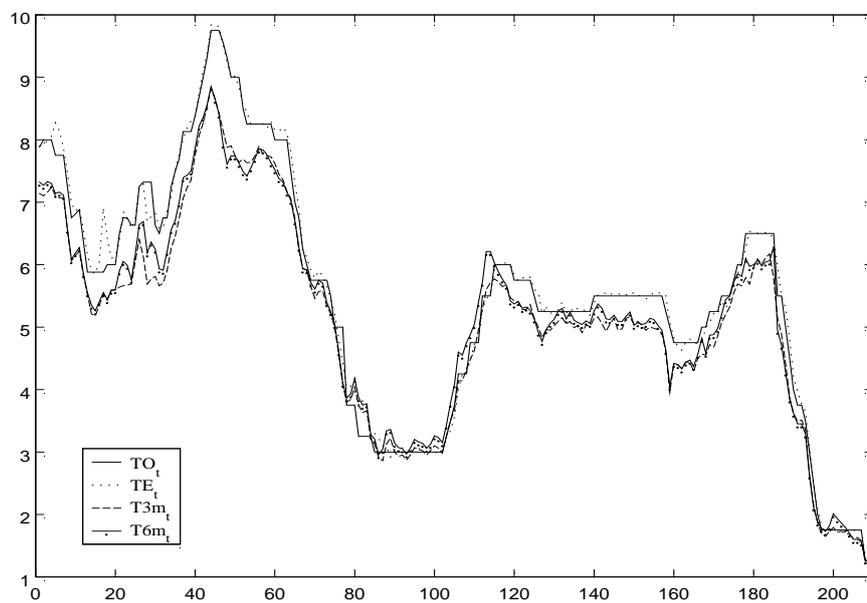


Figura 4: Tipos de interés a corto plazo.  $TO_t$  es el tipo objetivo,  $TE_t$  es el tipo efectivo y  $T3m_t$  y  $T6m_t$  son los tipos a 3 y 6 meses respectivamente. La muestra analizada se extiende desde agosto de 1985 hasta enero de 2003.

# Subspace-Based Methods for the determination of Unit Roots and Cointegrating Ranks

## **Abstract**

In this work we propose a new procedure to detect unit roots based on subspace methods. It has three main original aspects. First, the same method can be applied to single or multiple time series. Second, it uses a flexible family of information criteria, which loss functions can be adapted to the statistical properties of the data. Last, it does not require the specification of a stochastic process for the series analyzed. A simulation exercise shows that the method has good finite sample properties and its application is illustrated with the analysis of several real time series.

**Keywords:** State-space models, subspace methods, unit roots, cointegration