# COINTEGRACION: UNA PANORAMICA (*)

Juan J. Dolado

* Este artículo ha sido preparado para *Estadística Española*. Su contenido se ha visto beneficiado por comentarios de A. Banerjee, J. Galbraith, D. Hendry, T. Jenkinson y participantes en seminarios en las universidades de Londres, París y Valencia, en el Banco de España y en el Instituto Nacional de Estadística.

Banco de España. Servicio de Estudios
Documento de Trabajo n.º 8902

El Banco de España al publicar esta serie pretende facilitar la difusión de estudios de interés que contribuyan al mejor conocimiento de la economía española.

Los análisis, opiniones y conclusiones de estas investigaciones representan las ideas de los autores, con las que no necesariamente coincide el Banco de España.

ISSN: 0213-2702

ISBN: 84/7793- 026 - 0

Depósito legal: M. 28943 - 1989

Imprenta del Banco de España

## 1. INTRODUCCION

La mayoría del trabajo econométrico está asentado sobre el supuesto de estacionariedad. Hasta hace poco tiempo, este supuesto apenas se cuestionaba y el análisis económico procedía como si todas las series temporales económicas fuesen estacionarias, al menos alrededor de tendencias determinísticas. Las series estacionarias deben tener, al menos, medias y varianzas incondicionales constantes, un supuesto que raramente se verifica en economía. La importancia de dicho supuesto ha sido reconocida durante mucho tiempo, sobre todo desde que los importantes artículos de Granger y Newbold (1974) y Nelson y Plosser (1982) pusieron en alerta a la profesión sobre las consecuencias de la no estacionariedad. Las variables integradas constituyen una clase específica de variables no estacionarias con importantes propiedades tanto desde el punto de vista estadístico como económico. Estas propiedades se derivan de la presencia de tendencias estocásticas (raíces unitarias), en vez de únicamente determinísticas, teniendo las innovaciones de un proceso integrado carácter permanente en vez de transitorio. Por ejemplo, en términos de análisis del bienestar, ello implica que los costes de errores de expectativas producidos, por ejemplo, por variaciones en la política económica tienen un carácter mucho más relevante que en el caso en que las perturbaciones fuesen puramente transitorias.

En particular, la presencia de raíces unitarias en muchas series económicas resulta ser una implicación teórica del uso racional de la información disponible por parte de los agentes económicos. Ejemplos típicos entre variables financieras son los contratos a futuros, los precios de las acciones, los tipos de interés nominales y reales y el tipo de cambio, mientras que entre las variables reales encontramos el paro (histéresis), la inversión y el consumo. Quizás sea este último el ejemplo más conocido, derivado de añadir expectativas racionales a la teoría de la renta permanente. A la vista de esta epidemia de raíces unitarias en economía, ha aparecido una literatura voluminosa en estos últimos años sobre la contrastación, estimación y especificación de modelos que contienen variables integradas, por lo que el

propósito de esta panorámica es proporcionar una guía a través de esta literatura crecientemente compleja. Con el fin de limitar en lo posible el contenido de este artículo, nos circunscribiremos al análisis de la cointegración, es decir, al análisis de modelos multivariantes con series integradas.

Resulta comprensible que en estos últimos años se haya producido un importante progreso asociado con el tema de la cointegración[1]. La razón principal de este interés creciente reside en el hecho de que esta rama de investigación abre la posibilidad de estimar y contrastar la existencia de relaciones de equilibrio a largo plazo entre diversas variables, tal como sugiere la teoría económica. Tal como comentábamos previamente, la evidencia empírica sugiere que la mayoría de las series económicas son integradas, requiriendo un cierto grado de diferenciación a fin de convertirlas en estacionarias. Sin embargo, la teoría económica raramente sugiere relaciones de equilibrio que no sean funciones estacionarias de las variables originales, esto es, sin transformar. Ello implica que, de acuerdo con la teoría económica, existen mecanismos económicos que hacen que, a lo largo del tiempo, las variables se comporten estocásticamente del mismo modo. En otras palabras, mientras que las variables que conforman el modelo pueden ser todas ellas integradas, las desviaciones respecto a determinadas relaciones de equilibrio, sugeridas por la teoría resultan ser estacionarias

Se reconoce desde hace mucho tiempo que las técnicas usuales de análisis de regresión, en presencia de variables integrada, puede llevarnos a conclusiones erróneas, incluso cuando el análisis formal de este hecho sea reciente (veáse Phillips (1986))[2]. El problema de las regresiones "espúreas", tal como fue puesto de manifiesto por Granger y Newbold (1974), produjo un grado creciente de desconfianza en el análisis econométrico. Ello condujo a muchos economistas a adoptar el enfoque de Box y Jenkins (1976) a través del cual todas las series se convertían en series estacionarias, previamente al análisis de regresión, con lo que, en la mayoría de los casos, sólo se consideraban variables diferenciadas (por ejemplo, en el enfoque de "funciones de transferencia"). Ello implicaba que, en general, este tipo de

metodología ignoraba las propiedades a largo plazo del modelo ignorando, por tanto, las relaciones de equilibrio sugeridas por la teoría económica. Todo ello dificultaba la interpretación de los modelos.

Una importante línea de investigación que intentaba ofrecer una respuesta a dichos problemas, se basó en el uso de modelos de mecanismo de corrección del error (ECM) en la modelización econométrica. La idea de modelar ecuaciones de comportamiento económico en forma de "servo-mecanismos" que traten de corregir de forma sistemática errores cometidos en el pasado, se remonta al trabajo de Sargan (1964), basado en ideas de Phillips (1954), y tiene la ventaja de retener información sobre el nivel de las variables y, consecuentemente, sobre las relaciones de equilibrio entre dichas variables. Numerosos ejemplos de aplicaciones satisfactorias de este tipo de metodología han aparecido posteriormente en la literatura (véase, por ejemplo, inter alia. Davidson et al. (1978), Hendry y Mizon (1978), Davidson y Hendry (1981) Currie (1981) y Salmon (1982)), mientras que el análisis estadístico subyacente ha sido desarrollado por Hendry y Richard (1983, 1984). Aún así, existía una laguna en la comprensión de cómo los modelos de regresión con variables integradas podían conducir a términos de error estacionarios e incluso ir más allá y suponer perturbaciones i.i.d. Esta laguna ha sido cubierta mediante un importante artículo de Granger (1981), donde se establece, aunque de manera informal, la equivalencia entre los conceptos de cointegración y modelos ECM. Posteriormente, Granger y Weiss (1983) y Engle y Granger (1987) han desarrollado formalmente las ideas anteriores a través de un conjunto de teoremas de representación que establecen las conexiones existentes entre dichos conceptos. En resumen, se demuestra que un modelo ECM produce un conjunto de variables cointegradas y que la existencia de dicho conjunto implica una representación de las variables en forma de un modelo ECM. En gran medida, el concepto de cointegración provee un soporte estadístico formal al uso de modelos ECM y proporciona adicionalmente un procedimiento genérico para contrastar la validez de las predicciones, en términos de relaciones de equilibrio, derivadas de diversas teorías económicas.

El artículo está organizado del siguiente modo. La Sección 2 ilustra brevemente el marco estadístico necesario para analizar series

integradas, concentrándose, en aras de simplicidad, en el caso más simple de integración, i.e. el paseo aleatorio. La Sección 3 analiza las restricciones que la existencia de cointegración impone sobre las diversas representaciones de un sistema multivariante, así como los métodos de descomposición de dicho sistema en sus varios componentes integrados. Se presta especial atención a sistemas integrados de primer orden, aunque otros casos más generales también son brevemente analizados. La Sección 4 trata de estimación, contrastación y predicción en sistemas cointegrados. Finalmente, la sección 5 extiende parte del análisis previo a un marco de maxima verosimilitud, un enfoque desarrollado muy recientemente y que parece ser una de las líneas de investigación más fructíferas para el futuro.

## 2. PROPIEDADES ESTADISTICAS DE VARIABLES INTEGRADAS

Una serie temporal debilmente estacionaria debería tener media y varianza incondicionales invariantes con el tiempo. Sin embarog, la mayoría de las series económicas no satisfacen esta propiedad, al tener primer y segundos momentos que aumentar con el tamaño muestral (véase Escribano (1987) para la definición precisa de integración en el momento i-smo de un proceso estocástico)[3]. Tales series son no-estacionarias y pueden requerir un cierto grado de diferenciación para lograr estacionariedad. Una variable que requiere d diferencias para inducir estacionariedad se denomina integrada de orden d y se denota como I(d) (véase Granger (1982)). Un ejemplo de serie I(1) es el paseo aleatorio con deriva, cuyo proceso generador de los datos (PGD) es

$$\Delta y_t = -\alpha\, y_{t-1} + \mu + \varepsilon_t ; \quad \alpha = 0,\ y_0 = 0,\ \varepsilon_t \sim niid(0,\ \sigma_\varepsilon^2) \qquad (1)$$

o

$$y_t = \mu t + S_t \ ; \quad S_t = \sum_1^t \varepsilon_i \qquad (2)$$

Nos concentraremos en esta sección en las propiedades estadísticas que se derivan de la presencia de una única raíz unitaria como en (1), y para simplificar suponderemos que $\mu=0$. En general, las series integradas, como

$y_t$, resultan ser funciones lineales de una tendencia (si $\mu \neq 0$). Las desviaciones de dicha tendencia, $S_t$, son no-estacionarias al resultar de la acumulación de perturbaciones aleatorias pasadas, dando lugar al concepto de series integradas.

Con el fin de generalizar la especificación del PGD, abandonamos el supuesto simplificador de perturbaciones iid y pasamos a imponer ciertas condiciones más generales sobre la secuencia $\{\varepsilon_t\}$. Estas condiciones son necesarias si se quiere obtener distribuciones no degeneradas para el conjunto de estadísticos descritas posteriormente. El conjunto más débil de condiciones que puede imponerse se encuentra definido detalladamente en Phillips (1987) y puede resumirse del siguiente modo

(a) $E(\varepsilon_t) = 0 \quad \forall t$

(b) $\sup E|\varepsilon_t|^{2\beta} < \infty$ para algún $\beta > 2$

(c) $\sigma^2 = \lim E(T^{-1} S_T^2)$ existe y $\sigma^2 > 0$

(d) $\varepsilon_t$ tiene una mixtura fuerte con coeficientes de mixtura $\alpha_m$, tales que $\Sigma \alpha_m^{(1-2/\beta)} < \infty$.

La condición (b) restringe el grado de heterogeneidad del proceso, mientras que (c) controla la normalización en un grado que permite la existencia de distribuciones asintóticas no degeneradas. La varianza asintótica $\sigma^2$ es igual a $2\pi f(0)$, siendo $f(0)$ el espectro de $\varepsilon_t$ evaluado en la frecuencia cero. La condición (d) modera el grado de dependencia temporal en relación con la probabilidad de observaciones atípicas (veáse White (1984)).

La generalidad del anterior conjunto de condiciones implica que el modelo (1) engloba una gran variedad de PGD's. Entre estos se encuentra virtualmente cualquier modelo ARIMA con una raiz unitaria e incluso modelos ARMAX con una raíz unitaria y procesos exógenos no integrados.

Con el fin de derivar las distribuciones asintóticas que nos interesan, es necesario, tal como ocurre en el marco estacionario, utilizar una sucesón de variables aleatorias cuya convergencia se asegura a través de determinadas transformaciones. En términos más precisos, en el marco no estacionario necesitamos construir la sucesión de sumas parciales $\{S_t\}$ que ha de ser transformada de manera que sus elementos pertenezcan al espacio D(0,1) de todas las funciones de valor real en el intervalo [0,1] que son contínuas por la derecha y tienen límites finitos por la izquierda. Esto se logra mediante la sustitución de la serie original $\{S_t\}$ por la sucesión concentrada

$$S_T(r) = T^{-1/2} S_{[Tr]}, \quad r\varepsilon[0,1] \tag{3}$$

donde [z] representa la parte entera de z. De esta manera se concentra el dominio de $\{y_t\}$ en el intervalo [0,1], indiciando las observaciones mediante r. Así, por ejemplo, si T=100, la observación original $y_{50}$ estará indiciada por $r\varepsilon(0,50,\ 0,51)$. La elección de $T^{-1/2}$ como grado de normalización asegura que la serie transformada no sea ni explosiva ni converja a cero. Por ejemplo si $\varepsilon_t$ es iid, entonces var $(y_T)=T\sigma^2_\varepsilon$, con lo que su desviación típica será de orden $Op(T^{1/2})$ y, por tanto, este será el orden elegido para modificar el rango de $\{y_t\}$.

Bajo el conjunto de condiciones (a)-(d), tendremos que cuando $T\to\infty$, $S_T(r) \to \sigma W(r)$, donde "->" denota convergencia débil en probabilidad. W(r) es un proceso de Wiener (o movimiento Browniano) con varianza unitaria, que pertenece al espacio C[0,1] de funciones de variable real en el intervalo [0,1]. Este resultado de convergencia se conoce como Teorema de Donsker (véase Billingsley (1968)). Notese que W(r) se comporta como un paseo aleatorio en tiempo contínuo, tal que para r fijo se distribuye como N(0, r) y tiene incrementos independientes. Además, una extensión del teorema de Slutsky para variables aleatorias estacionarias, también se verifica en este contexto, en el sentido de que si g(.) es cualquier función contínua en C[0,1], entonces $S_T(r)\to\sigma W(r)$ implica que $g[S_T(r)]\to g[\sigma W(r)]$. Este resultado se conoce como Teorema de la Correspondencia

Contínua (véase Billingsley (1968)). La diferencia más importante entre la teoría asintótica aplicable a procesos estacionarios y la aplicable a procesos integrados, reside en el hecho de que mientras en la primera los momentos muestrales convergen a constantes, en la segunda convergen a variables aleatorias. Igualmente, como resultado de la ausencia de estacionariedad y ergodicidad, los Teoremas Centrales del límite tradicionales se sustituyen por Teoremas Funcionales del Límite (véase Billingsley (1968))[4].

Como ilustración de los resultados anteriores, los siguientes momentos muestrales estadarizados convergen a funcionales de Wiener

$$(i) \quad T^{-2} \Sigma y_t^2 \rightarrow \sigma^2 \int_0^1 W(r)^2 \, dr$$

$$(ii) \quad T^{-3/2} \Sigma y_t \rightarrow \sigma \int_0^1 W(r) \, dr$$

$$(iii) \quad T^{-1} \Sigma y_{t-1} \varepsilon_t \rightarrow \sigma^2/2[W(1)^2 - \sigma_\varepsilon^2/\sigma^2]$$

Nótense las divergencias entre las órdenes de magnitud de estos procesos límite con los correspondientes a los proceso estacionarios convencionales, i.e. $Op(T^2)$ en vez de $Op(T)$ en (i), $Op(T^{3/2})$ en vez de $Op(T)$ en (ii) y $Op(T)$ en vez de $Op(T^{1/2})$ en (iii). Estas diferencias iluminan las características no convencionales de la consistencia y distribución de los coeficientes que sirven como base para contrastar raíces unitarias y que serán de gran importancia en la discusión posterior del concepto de cointegración.

Si, por ejemplo, se aplican mínimos cuadrados ordinarios (MCO) a (1), resulta fácil derivar en base a los proceso límite (i)–(iii), que el estimador de la pendiente, $\hat{\alpha}$, y su t–ratio convergen a las siguientes distribuciones límite (cuando $\mu$=0)

$$T \hat{\alpha} \rightarrow \frac{\frac{1}{2}[W(1)^2 - \sigma_\varepsilon^2/\sigma^2]}{\int_0^1 W(r)^2 \, dr} \tag{4}$$

$$t_\alpha = 0 \rightarrow \frac{\frac{1}{2} \sigma[W(1)^2 - \sigma_\varepsilon^2/\sigma^2]}{\sigma_\varepsilon [\int_0^1 W(r)^2 \, dr]^{1/2}} \tag{5}$$

En (4) se observa que $\hat{\alpha}$ converge a su verdadero valor, $\alpha=0$, a velocidad $Op(T^{-1})$ en vez del orden convencional $Op(T^{-1/2})$. Igualmente, en (5), se observa que el t-ratio correspondiente tiene una distribución límite no degenerada que difiere de la distribución normal estandarizada que se utiliza en teoría asintótica convencional. Las distribuciones empíricas de (4) y (5) han sido obtenidas por simulación y constituyen la base de los contrastes de raíces unitarias, tanto en sus versiones paramétricas (véase, inter alia, Fuller (1976) Dickey y Fuller (1979, 1981), Evans y Savin (1981), Dickey et al. (1986) y Dickey y Pantula (1987)), como en las no paramétricas (véase Phillips y Ouliaris (1988) y Perron (1988)). Tal como se hizo notar en la Introducción este es un tema que no se trata en la presente panorámica.

## 3. COINTEGRACION Y SUS DIFERENTES REPRESENTACIONES

En esta sección se extiende el concepto de proceso integrado, examinado en la sección anterior, a un contexto multivariante, analizando el concepto de cointegración, es decir la posibilidad de las existencia de relaciones lineales entre variables integradas que resultan tener un orden menor de integración. El caso más interesante es aquel en que la relación existente es estacionaria, I(0), ya que en este caso puede aplicarse el análisis de regresión convencional. Además, se discuten las diferentes representaciones de procesos cointegrados, entre los cuales la más útil es aquella en términos de modelos ECM, donde los agentes económicos utilizan "servo-mecanismos" correctores de los errores cometidos previamente, de acuerdo con una relación de equilibrio preestablecido.

### 3.1. Descomposición de una serie Temporal en Componentes Integrados

Comenzamos por extender la definición de proceso integrado a un vector de series temporales. Sea $\{x_t\}$ un proceso n-dimensional tal que, después de sustraer los componentes determinísticos, tiene la siguiente representación ARIMA multivariante

$$\Phi(L)\, x_t = \Theta(L)\varepsilon_t; \quad \varepsilon_t \sim NI(0,\Omega) \tag{6}$$

donde $\Phi(L)$ y $\Theta(L)$ son (nxn) matrices de polinomios en el operador de desfases L, de orden p y q respectivamente. El polinomio autorregresivo es tal que det $\Phi(L)$ tiene todas sus raíces fuera del círculo unitario, excepto un número d de raíces iguales a la unidad, y $\Phi(0)=\Theta(0)=I$

De acuerdo con (6) definiremos a $\{x_t\}$ como un proceso integrado de orden d, I(d), si tiene la siguiente representación de Wold

$$(1-L)^d x_t = C(L)\,\varepsilon_t \tag{7}$$

donde $C(L) = \Phi^*(L)\,\Theta(L)/[\det \Phi(L)/(1-L)^d]$, siendo $\Phi^*(L)$ la matriz adjunta de $\Phi(L)$. C(L) es una fracción racional que admite un desarrollo en serie con coeficientes decrecientes en forma exponencial. Nótese que el álgebra de los procesos integrados es tal que si $d_1$ y $d_2$ representan los respectivos órdenes de integración de dos series temporales, entonces

$$I(d_1) + I(d_2) = I\{\max(d_1, d_2)\} \tag{8}$$

A partir de la representación de Wold (7), el proceso $\{x_t\}$ puede descomponerse en sus componentes integrados de orden decreciente de la siguiente forma

$$x_t = C(1)(1-L)^{-d}\varepsilon_t - C^{(1)}(1)(1-L)^{-(d-1)}\varepsilon_t + \ldots + \frac{C^{(d-1)}(1)(-1)^{d-1}}{(d-1)!}(1-L)^{-1}\varepsilon_t + C^*(L)\varepsilon_t \tag{9}$$

Esta representación se obtiene mediante un desarrollo de Taylor de la matriz C(L) hasta el orden (d-1), dividiendo posteriormente por $(1-L)^d$. Resulta fácilmente demostrable que la matriz residual $C^*(L)$ tiene un determinante cuyas raíces se encuentran fuera del círculo unitario. De esta manera se logra una descomposición de $x_t$ en términos de sus varios componentes integrados, que puede reescribirse como

$$x_t = x_t^{(d)} + x_t^{(d-1)} + \ldots + x_t^{(1)} + x_t^{(0)} \tag{10}$$

donde $x_t^{(k)} = [(-1)^{d-k}\, C^{(d-k)}(1)\, (1-L)^{-k}/(d-k)!]\varepsilon_t$

Resulta evidente de (10), que la descomposición no es única, ya que siempre pueden reordenarse los diferentes componentes, por ejemplo $(x_t^{(d-k)}+x_t^{(d-k-1)}/2)$ de forma que tenga el mismo orden de integración que $x_t^{(d-k)}$. Entre las infinitas posibles descomposiciones de $x_t$, se ha escogido aquella en que cada uno de los componentes tiene la representación ARIMA más sencilla. Resulta claro que (10) verifica esta propiedad y, en este sentido, se denominará descomposición canónica

Seguidamente analizamos la existencia de combinaciones lineales de los elementos de $x_t$, tales que algunos (o todos) de los componentes integrados desaparezcan. En este sentido, se define a $x_t$ como cointegrada de orden k, C(d,k) si existe un vector (1xn), $\alpha$, tal que $\alpha' x_t$ sea I(d-k). Esto es, a partir de (10)

$$\alpha' x_t = \alpha' x_t^{(d-k)} + \ldots + \alpha' x_t^{(0)} \sim I(d-k) \tag{11}$$

La descomposición (11), de acuerdo con (10), implica que

$$\alpha' C(1) = \alpha' C^{(1)}(1) = \ldots = \alpha' C^{(d-k+1)}(1) = 0 \tag{12}$$

es decir Ker $C^{(1)}(1)' \neq \{0\}$ (i=1,..., d-k+1). Por tanto si definimos los espacios vectoriales

$$C_s = \bigcap_{k=0}^{s} \text{Ker } [C^{(k)}(1)'] \tag{13}$$

podemos alternativamente definir a $x_t$ como cointegrada C(d,k) si y solo si existe un vector $\alpha$ tal que $\alpha \in C_{d-k+1}$ y $\alpha \notin C_{d-k}$

Ejemplo 1: Si $x_t$ es C(1,1), ello implica que existe, al menos, un vector $\alpha$ tal que $\alpha' C(1)=0$.

La relación existente entre los conceptos de cointegración y equilibrio a largo plazo aparece ahora más nítida. Una manera natural de definir el

equilibrio entre un conjunto de variables es postular una relación del tipo $\alpha' x_t = 0$. Por ejemplo, si se supone que una cierta proporción $\gamma$ de los aumentos en la productividad laboral (q en logaritmos) se traslada a aumentos en el salario real ((w-p) en logaritmos), entonces, en equilibrio (aparte de constantes), se verifica la relación $w - p = \gamma q$. Por tanto si $w - p - \gamma q = 0$ en cualquier período t, el mercado de trabajo estará en equilibrio (nótese que si $\gamma = 1$, la relación previa representa la condición de primer orden de una empresa competitiva con una función de producción Cobb-Douglas). Es claro que los salarios reales tenderán a responder con cierto retraso a los cambios en la productividad y la forma en que se restaura el equilibrio puede ser un proceso complejo. En este caso, la perturbación $e_t = w_t - p_t - \gamma q_t$ mide la desviación de la relación de equilibrio, o desequilibrio, en el período t. Si $(w-p)_t$ y $q_t$ están cointegradas, entonces, de acuerdo con las definiciones anteriores, la desviación $e_t$ ha de ser I(0). Una forma obvia de contrastar esta teoría es determinar el orden de integración de $e_t$. Si no es posible rechazar la hipótesis nula de una raiz unitaria en $e_t$, no existirá un mecanismo de ajuste de los salarios reales a su valor de equilibrio, en cuyo caso la relación postulada constituye un concepto irrelevante. La teoría económica sugiere muchas relaciones de equilibrio similares a la anterior, por ejemplo, consumo y renta disponible (Campbell (1987) y Banerjee y Dolado (1988)); tipo de cambio nominal y nivel de precios relativos (Edison y Kloveland (1987)); renta nominal, tipo de interés y saldos nominales de dinero (Hendry y Ericsson (1986) y Dolado (1988)), tipos de interés a corto y largo plazo (Campbell y Shiller (1987)), etc. La contrapartida estadística a las anteriores relaciones de equilibrio es la existencia de cointegración entre las variables previamente mencionadas.

El ejemplo anterior tiene también la virtud de ilustrar varias características del concepto de cointegración. Así, por ejemplo, si $w_t$ y $p_t$ están cointegradas C(d,b), el vector de cointegración ha de ser único ya que si existiese cualquier otra combinación lineal independiente de la anterior, siempre se podría expresar una variable en términos de la otra añadiendo o sustrayendo $w_t$ (o $p_t$) que es I(d). Sin embargo, cuando se trata de más de dos variables, por ejemplo $w_t$, $p_t$ y $q_t$, entonces puede existir

más de un vector de cointegración. Así $p_t - w_t = \beta q_t$ representaría una ecuación de precios de "mark-up", que coexiste con la relación de salarios. Nótese que si existiesen, en este caso, tres vectores independientes, entonces las tres variables serían estacionarias y cualquier relación entre ellas también sería estacionaria, con lo que el concepto de cointegración está vacio de contenido.

La existencia de cointegración también posibilita la identificación de modelos que en un contexto estacionario no estarían identificados. Para ilustrar este hecho, considérese el siguiente ejemplo tomado de Engle y Granger (1987). Supóngase que $x_t = (y_t, z_t)'$ está generado de acuerdo con el siguiente PGD

Ejemplo 2

$$y_t + \beta z_t = v_t \quad ; \quad v_t = \rho_1 v_{t-1} + \varepsilon_{1t}$$

$$y_t + \alpha z_t = u_t \quad ; \quad u_t = \rho_2 u_{t-1} + \varepsilon_{2t} \qquad (14)$$

$$\varepsilon_{it} \sim \text{niid}(0, \sigma_i^2) \ (i=1,2) \quad ; \quad E(\varepsilon_{it} \varepsilon_{js}) = 0$$

Resulta claro que si $y_t$ y $z_t$, fuesen estacionarios, entonces $\alpha$ y $\beta$ no estarían identificados. Sin embargo, dependiendo del valor tomado por los parámetros $\rho_i$'s, surgen los siguientes cuatro casos

(i) $\rho_1 = 1$, $\rho_2 < 1$; implica que $y_t$ y $z_t$ son I(1) y el vector de cointegración es $(1,\alpha)$. Nótese que para un tamaño muestral T suficientemente grande, la perturbación $v_t$ tendrá una varianza que tiende a infinito. Por tanto, cualquier método de minimice la suma de cuadrados de los residuos, ponderada o no ponderada, identificará la segunda ecuación.

(ii) $\rho_1 < 1$, $\rho_2 = 1$; implica, simétricamente al caso anterior, que $y_t$ y $z_t$ son I(1) y el vector identificado de cointegración es $(1,\beta)$

(iii) $\rho_1=\rho_2=1$; implica que $y_t$ y $z_t$ son I(1) pero no existe ningún vector de cointegración

(iv) $\rho_1<1$, $\rho_2<1$; implica que $y_t$ y $z_t$ son I(0), lo que provoca falta de identificación.

Finalmente, nótese que aunque la discusión previa ha sido desarrollada en términos de un marco de relaciones estáticas entre variables integradas, no existe razón alguna para ignorar la existencia de relaciones dinámicas de cointegración. Por ejemplo, supóngase el siguiente PGD

Ejemplo 3

$$y_t = \beta z_t + \gamma \Delta z_t + \varepsilon_{1t}$$

$$\Delta^2 z_t = \varepsilon_{2t} \qquad (15)$$

$$\varepsilon_{it} \sim \text{niid}(0,\sigma_i^2)(i=1,2),\ E(\varepsilon_{it}\ \varepsilon_{js}) = 0$$

Se deriva de (15) que $y_t$ y $z_t$ son I(2) y de la primera ecuación se deriva el vector de cointegración $(1,-\beta-\gamma)$ que produce una combinación lineal de $y_t$, $z_t$ y $\Delta z_t$ que es I(0). Nótese que $y_t$ y $z_t$ son C(2,1) y, por tanto, necesitan asociarse con otra variable I(1), $\Delta z_t$ en este caso, para lograr una condición necesaria de cointegración.

## 3.2. Teoremas de Representación

Una vez que el concepto de cointegración ha sido analizado, es importante examinar las restricciones que impone sobre las diferentes representaciones AR o MA que se derivan de la representación general ARIMA expresada en (6). En línea con el artículo de Engle y Granger (1987), comenzaremos por analizar el caso más sencillo de la clase de Teoremas de Representación, que corresponde a que $x_t$ sea C(1,1). Por tanto cada elemento del vector $x_t$, $x_{it}$, es I(1) y existen combinaciones lineales de dichos elementos que son I(0). El siguiente conjunto de resultados se conoce en la literatura como Teorema de Representación de Granger[5].

Sea $\{x_t\}$ un proceso estocástico n-variante C(1,1), tal que existen r vectores de cointegración (r<n-1) agrupados en la matriz (nxr) $\Gamma$. Entonces existen las siguientes representaciones

i) Representación MA

A partir de (7) con d=1, $x_t$ sigue el proceso MA

$$(1-L)x_t = C(L)\varepsilon_t \tag{16}$$

y de acuerdo con (11)

$$\alpha' C(1)=0 \rightarrow \Gamma' C(1)=0 \tag{17}$$

Por tanto, las filas de $\Gamma$ pertenecen al espacio ker de C(1)'. Como el número de filas en C(1) es n, igual al rango más la dimensión del espacio ker (véase, por ejemplo Dhrymes (1978)), se sigue que

$$\text{rango } C(1) = n-r \tag{18}$$

por lo que C(1) no tiene rango completo

ii) Representación AR

La representación (7) puede reescribirse en forma AR

$$A(L)x_t = \varepsilon_t \tag{19}$$

donde

$$A(L)\,C(L) = 1-L \rightarrow A(1)\,C(1) = 0 \tag{20}$$

Por tanto, las filas de A(1) pertenecen al espacio ker de C(1)' y a partir de (17) puede escribirse del siguiente modo:

$$A(1) = B\Gamma' \tag{21}$$

donde B es una matriz (nxr). Puesto que A(1) ε ker C(1)' se sigue que

$$\text{rango } A(1) = r \tag{22}$$

Dado que A(1) es singular, la representación VAR será tal que los procedimientos numéricos utilizados para calcular la representación MA resultarán altamente inestables, signo de multicolinealidad, lo que dificulta el análisis de multiplicadores en las funciones impulso-respuesta.

iii) Representación ECM

Utilizando un desarrollo de primer orden de Taylor de A(L) alrededor de L=1, se obtiene

$$A(L) = A(1) + A^*(L)\,(1-L) \tag{23}$$

donde el determinante de $A^*(L)$ tiene sus raíces fuera del círculo unitario. Sumando y restando A(1) L a (19) y utilizando (21), se obtiene

$$A(L)x_t = [B\Gamma'L+A^{**}(L)(1-L)]x_t = \varepsilon_t \tag{24}$$

donde $A^{**}(L) = A(1)+A^*(L)$. Por tanto $A^{**}(0) = A(1)+A^*(0)$ y, dado (23), $A^{**}(0) = I$. Consecuentemente, (24) puede reescribirse en términos de un modelo multivariante ECM

$$A^{**}(L)\,\Delta\, x_t = -B\Gamma'\, x_{t-1} + \varepsilon_t \tag{25}$$

Nótese que en la representación (25) la existencia de cointegración se deriva de la presencia de los niveles de las variables, de forma que la estimación de modelos VAR en diferencias dará lugar a errores de especificación si las variables están cointegradas. Igualmente la estimación de modelos VAR en niveles que ignoren las restricciones (21) presentan graves

problemas al obtener la representación MA, tal como se comentó previamente. Sin embargo, Sims et al (1986) han demostrado recientemente que en un modelo VAR donde existen relaciones de cointegración entre las variables, siempre que cualquier variable pueda ser escrita como variable estacionaria con media cero, por ejemplo $\Delta y_t$, $\Delta x_t$ y $(y_{t-1} - \alpha x_{t-1})$ y sus desfases, en un modelo bivariante, sus coeficientes se distribuirán normalmente de acuerdo con la teoría asintótica convencional. Como veremos posteriormente, esto no es completamente cierto cuando se trate de los coeficientes del propio vector de cointegración.

La representación (25) también ilustra la conexión existente entre los conceptos de cointegración y "causalidad de Granger". Por ejemplo, si se trata de un sistema bivariante deberá haber causalidad en la direción z --> y si el segundo elemento de la matriz B, que es (2x1), es nulo, o en la dirección y --> z si es el primer elemento el que es nulo. Si ambos elementos de B difieren de cero, entonces existe causalidad en ambas direcciones y ninguna de las variables puede considerarse como débilmente exógenea con respecto a los parámetros de la otra ecuación (véase Engle et al. (1983)). En este último caso existen restricciones no lineales a través de las ecuaciones, y para su estimación eficiente se necesita un procedimiento de información completa

iv) Representación TC (Tendencias Comunes)

La idea subyacente a esta representación, tal como ha sido sugerida por Aoki (1987) y Stock y Watson (1989), es que si un grupo de variables están cointegradas, deberá existir un conjunto de tendencias estocástics comunes que guien las representaciónes ARIMA de las variables individualmente. Estas tendencias comunes desaparecerán mediante el uso de determinanadas combinaciones lineales entre las variables. Con el fin de ilustrar este método de descomposición, utilicemos una versión de primer orden de (10) cuando d=1, obteniendo

$$x_t = C(1)\, x_t^{(1)} + C^*(L)\varepsilon_t = C(1)x_t^{(1)} + \sum_{1}^{t} \left[ \sum_{0}^{t-i} C_j - C(1) \right] \varepsilon_i \qquad (26)$$

o bien

$$x_t = C(1)x_t^{(1)} - \sum_{1}^{i} [\sum_{t-i+1}^{\infty} C_j]\varepsilon_i \qquad (27)$$

Las tendencias estocásticas vienen dadas por $x_t^{(1)} = \sum_{1}^{t} e_i$ , mientras que el segundo término en (27) es estacionario, para t suficientemente grande, dado que las matrices $C_j$ decaen exponencialmente para j suficientemente grande. A partir de (18) sabemos que el rango C(1) es n−r y por tanto existen (n−r) tendencias comunes en en el sistema, de ahí la denominación TC para esta representación.

Con el fin de ilustrar las diferentes representaciones examinadas en esta sección, considérese el siguiente ejemplo

Ejemplo 4

Supongamos que $x_t=(y_t,z_t)'$ tiene la siguiente representación MA

$$(1-L)\begin{pmatrix} y_t \\ z_t \end{pmatrix} = (1-0,2L)^{-1}\begin{pmatrix} 1-0,6L & 0,8L \\ 0,2L & 1-0,6L \end{pmatrix}\varepsilon_t$$

Así pues

$$C(1) = \begin{pmatrix} 0,5 & 1,0 \\ 0,25 & 0,5 \end{pmatrix} \rightarrow \text{rango } C(1) = 1$$

Por tanto, $y_t$ y $z_t$ son C(1,1). El vector de cointegración, normalizando el coeficiente de $y_t$, es $\alpha'=(1,-2)$

Invirtiendo la matriz MA, se obtiene la repesentación AR

$$\begin{pmatrix} 1-0,6L & -0,8L \\ -0,2L & 1-0,6L \end{pmatrix}\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \varepsilon_t$$

donde

$$A(1) = \begin{pmatrix} 0,4 & -0,8 \\ -0,2 & 0,4 \end{pmatrix} \rightarrow \text{rango } A(1) = 1$$

Nótese a partir de (21) que

$$A(1) = B\alpha' = \begin{pmatrix} 0,4 \\ -0,2 \end{pmatrix}(1,-2)$$

Por tanto la representación ECM viene dada por

$$(1-L)\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0,4 \\ -0,2 \end{pmatrix}(1,-2)\begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \varepsilon_t$$

En este caso existe causalidad en ambas direcciones. Finalmente la representación TC viene dada por

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0,5 & 1 \\ 0,25 & 0,5 \end{pmatrix}\begin{pmatrix} y_t^{(1)} \\ z_t^{(1)} \end{pmatrix} - \sum_{1}^{i} \left[ \sum_{t-i+1}^{\infty} C_j \right]\varepsilon_i$$

$$\text{con } C_j = (0,2)^j \begin{pmatrix} -0,5 & 1 \\ 0,25 & -0,5 \end{pmatrix}$$

Nótese que en la representación TC, el producto interno $\alpha' x_t$ elimina las tendencias comunes $y_t^{(1)}$ y $z_t^{(1)}$

## 3.3. Extensiones de los Teoremas de Representación

En esta sección se extienden los teoremas de la sección anterior, con especial atención a la representación ECM, a casos en que el vector $x_t$ tiene más de una raíz unitaria o posee raíces unitarias en las frecuencias estacionales. Con el fin de ilustrar la forma de proceder en casos más generales, consideremos la representación ECM en los casos de una doble raíz unitaria $(1-L)^2$ y de una raíz unitaria en la frecuencia semestral, $(1-L^2)$.

Ejemplo 5

Supongamos que $\{x_t\}$ tiene la representación de Wold

$$(1-L)^2 x_t = C(L)\varepsilon_t \tag{28}$$

donde C(L) puede desarrollarse del siguiente modo

$$C(L) = \Upsilon_1 + \Upsilon_2(1-L) + C^*(1-L)^2 \tag{29}$$

con $\Upsilon_1=C(1)$, $\Upsilon_2 = -C^{(1)}(1)$ y $C^*(L)$ es invertible. Supongamos que existe un vector de cointegración $\alpha$ tal que

$$\alpha'\Upsilon_1 = \alpha'\Upsilon_2 = 0 \tag{30}$$

A partir de la representación VAR, se obtiene

$$A(L)\ C(L) = (1-L)^2 \rightarrow A(1)\ C(1) = 0 \tag{31}$$

por lo que A(1) puede escribirse como

$$A(1) = B\alpha' \tag{32}$$

Derivando (31) con respecto a L y evaluando en L=1, se obtiene

$$A'(1)C(1) + A(1)\ C'(1) = 0 \tag{33}$$

Utilizando (30) y (32) se sigue que $A'(1)C(1) = 0$, de forma que $A'(1)$ puede escribirse en la forma

$$A'(1) = D\ \alpha' \tag{34}$$

Utilizando un argumento similar al usado en la obtención de (25), una expansión de $A(L)$ similar a (29) conduce a la siguiente representación ECM

$$A^{**}(L)\Delta^2\ x_t = -B\alpha' x_{t-1} - (B-D)\alpha'\Delta\ x_{t-1} + \varepsilon_t \tag{35}$$

con $A^{**}(0) = I$

Ejemplo 6

Supongamos que $\{x_t\}$ tiene la representación de Wold

$$(1-L^2)\ x_t = C(L)\ \varepsilon_t \tag{36}$$

La matriz $C(L)$ puede escribirse como

$$C(L) = \Upsilon_1(1-L) + \Upsilon_2(1+L) + C^*(L)(1-L^2) \tag{37}$$

donde $\Upsilon_1 = C(-1)/2$, $\Upsilon_2 = C(1)/2$ y $C^*(L)$ es invertible

A partir de la representación VAR

$$A(L)\ C(L) = (1-L)(1+L) \rightarrow A(1)\ C(1) = A(-1)\ C(-1) = 0 \tag{38}$$

Así pues, habrá cointegración en la raíz $(1+L)$ si existe un vector de cointegración $\alpha_1$ tal que $\alpha'_1\Upsilon_1 = \alpha'_1\ C(-1) = 0$. De forma similar habrá cointegración en la raíz $(1-L)$ si existe un vector $\alpha_2$ tal que $\alpha'_2\Upsilon_2 = \alpha'_2\ C(1) = 0$. Por tanto podemos expresar $A(1)/2 = B\alpha'_2$ y

$A(-1)/2=D\alpha'_1$. Utilizando el desarrollo (37) para A(L) se obtiene la representación ECM

$$A^{**}(L)\Delta_2 x_t = B\alpha'_2 \Delta x_{t-1} - D\alpha'_1 (x_{t-1}+x_{t-2})+\varepsilon_t \qquad (39)$$

con $A^{**}(0) = I$

Nótese que si no existe cointegración en (1+L), $\alpha_1=0$ y el segundo término en el lado derecho de (39) desaparece. Igualmente si no hay cointegración en (1-L), $\alpha_2=0$ y el primer término no aparece en (39).

Argumentos similares pueden utilizarse para obtener las representaciones ECM para otras frecuencias estacionales (en el caso de periodicidad trimestral, véase Engle et al. (1988) y Banerjee et al. (1989)) e incluso cuando existen raíces en frecuencias regulares y estacionales (véase Osborn et al. (1989)).

## 4. ESTIMACION, CONTRASTACION Y PREDICCION EN SISTEMAS COINTEGRADOS

En esta sección se discuten aspectos relativos a la estimación, contrastación y predicción en sistemas cointegrados. Con el fin de simplificar el análisis limitaremos la discusión a sistemas C(1,1), ofreciendo generalizaciones cuando sea necesario

### 4.1 Estimación de Vectores de Cointegración

Considérese el problema de estimar el vector de cointegración $\alpha$ en el siguiente modelo estático

$$\alpha' x_t = e_t \qquad (40)$$

Si todos los elementos de $x_t$ son I(1), en general cualquier combinación de dichas variables será I(1), excepto en el caso de la combinación lineal dada por el vector $\alpha$. Por tanto, casi todas las combinaciones

lineales darán lugar a procesos $e_t$ con varianza asintóticamente infinita. Dado que un procedimiento sencillo de estimación, como MCO, minimiza la varianza de $e_t$, el vector estimado $\alpha$, después de normalizar a la unidad el coeficiente de una de las variables, representará una excelente aproximación al verdadero vector de cointegración $\alpha$, si es que éste existe.

La sencillez del resultado anterior explica, en gran proporción, por qué el interés en el concepto de cointegración ha explotado como si se tratase de una serie integrada. Ello implica que para parametrizar una relación de equilibrio a largo plazo entre un conjunto de variables, todo lo que se necesita es estimar, mediante MCO, una regresión estática entre dichas variables. Esta regresión puede incluso estimarse en la primera etapa del proceso de especificación dinámica, tal como sugieren Engle y Granger (1987) en el procedimiento en "dos etapas" que discutirá posteriormente. En cualquier caso, dicha regresión puede utilizarse como un contraste inicial que indique la medida en que las predicciones de equilibrio sugeridas por la teoría económica están en consonancia con los datos disponibles. De esta manera se tendrá evidencia de si es fructífero emplear recursos posteriores en la modelización de la dinámica a corto plazo, alrededor de la relación de equilibrio, tal como ilustra la representación ECM.

De acuerdo con la discusión anterior, el estimador minimocuadrático de cualquier vector de cointegración convergerá rápidamente a su verdadero valor. Para analizar esta propiedad considérese el siguiente PGD

Ejemplo 7

$$y_t = \alpha z_t + e_t \quad ; \quad e_t = (1-\rho L)\xi_t \quad ; \quad |\rho|<1$$

$$\Delta z_t = \varepsilon_t$$

$$\varepsilon_t \sim \text{niid}(0,\sigma_\varepsilon^2),\ \xi_t \sim \text{niid}(0,\sigma_\xi^2);\ E(\varepsilon_t \xi_s) = 0$$

El modelo que se estima es

$$y_t = \mu + \alpha x_t + u_t$$

lo que da lugar al siguiente conjunto de resultados

$$T(\hat{\alpha}-\alpha) \rightarrow f_1(W) \quad ; \quad T(1-R^2) \rightarrow f_3(W) \quad ; \quad t_{\alpha=0} \rightarrow f_4(W)$$

$$T^{1/2}(\hat{\mu}) \rightarrow f_2(W) \quad ; \quad DW \rightarrow 2(1-\rho)$$

donde $f_i(W)$ representan genéricamente funcionales de procesos de Wiener del tipo (i)-(iii) en la Sección 2 (véase Stock (1987)). La interpretación de los resultados ilustra claramente la discusión informal previa. La pendiente estimador en la regresión estática converge a su verdadero valor $\alpha$ a velocidad $Op(T^{-1})$ en vez de $Op(T^{-1/2})$ como ocurre con series I(0). La intuición es clara ya que $\hat{\alpha}$ se computa con el ratio de una covarianza, que de acuerdo con (iii) es $Op(T^{-1})$, respecto a una varianza, que de acuerdo con (i) es $Op(T^{-2})$, dado que tanto $y_t$ como $z_t$ son I(1). Por tanto el sesgo es $Op(T^{-1})$. Sin embargo, el t-ratio de $\hat{\alpha}$ tiene una distribución límite no degenerada que, en general, no coincide con la normal estandarizada por lo que los valores reportados de dicho estadístico carecen usualmente de sentido. Puesto que, en este caso, tanto $y_t$ como $z_t$ carecen de deriva (los componentes determinísticos han sido sustraidos), $\hat{\mu}$ converge consistentemente a cero, pero a una velocidad $Op(T^{-1/2})$. El coeficiente de correlación múltiple $R^2$ también es $Op(T^{-1})$ consistente de la unidad, reflejando el hecho de que en un modelo bivariante cointegrado existe un único vector de cointegración, por lo que el producto de las pendientes en las regresiones de y sobre z y de z sobre y, es la unidad. Esta característica será explotada posteriormente. Finalmente, el estadístico Durbin-Watson (DW) converge al resultado estándar, bajo el supuesto simplificador de una perturbación AR(1).

Un importante tema asociado a los resultados anteriores se refiere a la existencia de sesgos de simultaneidad o errores en las variables. Los sesgos que aparecen en dichas ocasiones, se derivan de las correlaciones existentes entre los regresores y las perturbaciones, que normalmente se

suponen de orden Op(T). Sin embargo, dado que, por ejemplo, en el modelo del Ejemplo 6, $\Sigma z_t e_t$ tiene un orden probabilístico inferior al de $\Sigma z_t^2$, tales sesgos desaparecen asintóticamente. Ello implica que, términos de consistencia de los estimadores, ninguno de los problemas anteriores importa.

Ejemplo 8

Considérese el modelo canónico de errores en las variables

$$y_t = \beta z_t^* + u_t$$
$$z_t = z_t^* + \varepsilon_t$$

donde $E(z^*_t u_t)=E(z^*_t \varepsilon_t)=E(u_t \varepsilon_t)= 0$; plim $T^{-1} \Sigma x_t^2 = \sigma_x^2$; plim $T^{-1} \Sigma x_t^{*2} = \sigma_{x^*}^2$.

El estimador MCO de $\beta$ es tal que

$$\text{plim } \hat{\beta} = \beta \sigma_{x^*}^2/\sigma_x^2 < \beta$$

Por tanto $\beta$ está sesgado a la baja y además no está identificado, puesto que $\sigma_{x^*}^2$ no puede estimarse a partir de los momentos muestrales. Maravall (1979) discute condiciones generales bajo las cuales la identificación de $\beta$ es posible. Estas condiciones explotan ciertos rasgos de la dinámica del modelo. Por ejemplo si se conoce que

$$x^*_t = \rho x^*_{t-1} + v_t; \quad |\rho| < 1; \quad v_t \sim \text{niid}(0, \sigma_v^2)$$

$$E(v_t u_t) = E(v_t \varepsilon_t) = 0$$

entonces

$$\text{plim } \beta = (1+\gamma)^{-1} \beta; \quad \gamma = \sigma_\varepsilon^2(1-\rho^2)/\sigma_v^2$$

Puesto que $x_t$ sigue el siguiente proceso ARIMA

$$x_t = \rho\, x_{t-1} + v_t + \varepsilon_t - \rho\, \varepsilon_{t-1}$$

entonces $\rho$, $\sigma^2_\varepsilon$ y $\sigma^2_v$ pueden estimarse consistentemente a partir de los momentos muestrales $\gamma_{xx}$, $\gamma_{xx-1}$ y $\gamma_{xx-2}$. Sin embargo, si $x^*_t$ es I(1), i.e. $\rho=1$, entonces $\gamma=0$ y $\beta$ puede ser identificado únicamente a partir de $\hat{\beta}$, que es consistente.

Este resultado tiene implicaciones importantes para determinados argumentos que abundan en textos de macroeconomía. Así, por ejemplo, Friedman (1957) explica la divergencia existente entre las estimaciones de las propensiones media y marginal a consumir en base a un argumento basado en "errores en las variables", i.e. la confusión entre renta actual y permanente. Si la renta sigue un proceso integrado este argumento no es válido, tal como se ha visto previamente. Stock (1988), aplicando la teoría asintótica de procesos de Wiener, demuestra que la propensión marginal a consumir presenta un sesgo negativo de 0,15 en la regresión estática de consumo sobre renta. Ajustando dicho sesgo, demuestra que ambas propensiones son aproximadamente igual a la unidad.

El resultado más importante de est sección se refiere a la "super"consistencia" de $\alpha$. Sin embargo, tal como se apuntaba en el párrafo anterior, los sesgos, a pesar de ser $Op(T^{-1})$, pueden ser importantes en muestral finitas. Supóngase que los ratios de convergencia de dos estimadores fuesen $T^{-1/2}$ y $10^{10}\ T^{-1}$, en este caso se necesitarían muestras muy grandes para que segundo estimador dominase al primero. En un amplio estudio de Monte-Carlo, Banerjee et al. (1986) examinan los sesgos derivados de la regresión estática. Dado que el coeficiente de correlación múltiple $R^2$ converge a la unidad también a velocidad $Op(T^{-1})$, proponen $(1-R^2)$ como una aproximación al sesgo. Por ejemplo si se supone un PGD como en el Ejemplo 7, entonces la relación entre sesgo y $R^2$ resulta ser

$$\hat{\alpha} - \alpha = \alpha(1-R^2) + Op(T^{-1}) \qquad (41)$$

Si T es suficientemente grande, $\alpha$=1 y $R^2$=0.90, el valor de $\alpha$ estará alrededor de 0.9, lo que implica un sesgo apreciable. Este resultado sugiere que las relaciones de cointegración estimadas en modelos estáticos con $R^2$ no muy cercanos a la unidad, deben interpretarse con escepticimo (véase, por ejemplo, Campbell y Shiller (1987), donde el $R^2$ de la relación entre tipos de interés a corto y largo plazo es 0.8). Sin embargo, resulta importante destacar que en el contexto de regresión múltiple, el $R^2$ de una ecuación siempre aumenta cuando se introduce una variable adicional, por lo que un $R^2$ alto no es condición necesaria para la obtención de sesgos reducidos.

En el Ejemplo 7, el PGD es esencialmente estático, con lo que el modelo reproduce las características del PGD. En general, cuando la relación entre $y_t$ y $z_t$ sea dinámica, el modelo estático traslada toda la dinámica al residuo $u_t$. Los términos dinámicos pueden reparametrizarse en términos de variable I(0), en la forma $\Delta y_{t-i}$, $\Delta z_{t-i}$ y $(y-\alpha z)_{t-i}$ donde los valores de i dependerán de la naturaleza del proceso ARIMA multivariante que genere $x_t=(y_t, z_t)'$, tal como se muestra en la representación ECM de un modelo C(1,1) dada en (25). Con el fin de ofrecer una simple ilustración de este hecho sustitúyase la primera ecuación en el PGD del Ejemplo 6 por

$$y_t = \alpha_1 y_{t-1}+\alpha_2 z_t+\alpha_3 z_{t-1}+\varepsilon_t;\ \varepsilon_t \sim niid(0,\sigma_\varepsilon^2) \qquad (42)$$

La expresión anterior puede reparametrizarse como

$$y_t = \alpha z_t + e_t \qquad (43)$$

donde $\alpha = (\alpha_2+\alpha_3)(1-\alpha_1)^{-1}$ y $e_t = -(1-\alpha_1)^{-1}[\alpha_1 \Delta y_t+\alpha_3 \Delta z_t-\varepsilon_t]$

Puesto que $\Delta y_t$ y $\Delta z_t$ son I(0), $e_t$ es I(0) y contiene correlación serial. Phillips (1988) demuestra que es precisamente la presencia de correlación serial en $e_t$, el motivo de obtener sesgos en la estimación de $\alpha$ en muestras finitas. Una alternativa obvia es la estimación $\alpha$ directamente a partir de (42) (cuando $z_t$ sea débilmente exógena, como en el

presente caso), obtenido a partir del ratio $\hat{\hat{\alpha}}=(1-\hat{\alpha}_1)^{-1}(\hat{\alpha}_1+\hat{\alpha}_3)$. En este caso la perturbación $e_t$ carece de correlación serial. Banerjee et al. (1986) comparan, mediante simulación, los sesgos que se obtienen por ambos métodos, regresión estática y dinámica, obteniendo una clara ventaja del método dinámico.

Finalmente, otra importante implicación que se deriva de la existencia de cointegración, es que, en contraste con el análisis estándar de regresión donde la multicolinealidad se considera un problema, en el contexto de sistemas cointegrados dinámicos, su existencia es esencial. Por ejemplo, considérese el modelo (42) reparametrizado como

$$\Delta y_t = \alpha_2 \Delta z_t + (\alpha_1 - 1) y_{t-1} + (\alpha_2 + \alpha_3) z_{t-1} \tag{44}$$

Típicamente $y_{t-1}$ y $z_{t-1}$ serán altamente colineales y consecuentemente sus t-ratios serán bajos. Eliminar cualquiera de las dos variables sería una solución desastrosa al probema de multicolinealidad, puesto que en ese caso el regresando sería I(0) y el conjunto de los regresores, I(1). Siendo I(1), los segundos momentos de $y_{t-1}$ y $z_{t-1}$ son muy grandes relativamente a los de $\Delta z_t$ y, por tanto, cuando la matriz (X'X) se invierta, la submatriz que incluye $y_{t-1}$ y $z_{t-1}$ tendrá una "quasi-singularidad" correspondiente al vector de cointegración $(1,-\alpha)$. Esto, sin embargo, no es un problema, pues todo lo que se necesita es raparametrizar (44) en la forma ECM, esto es

$$\Delta y_t = \alpha_2 \Delta z_t + (\alpha_1 - 1)(y_{t-1} - \alpha z_{t-1}) + \varepsilon_t \tag{45}$$

donde los regresores son ahora mucho menos colineales. En este sentido, la presencia de multicolinealidad es un signo de la existencia de cointegración en la relación estimada.

## 4.2. Procedimientos de estimación en dos etapas: Estática y Dinámica

Una vez que el vector de cointegración $\alpha$ ha sido estimado a partir de una regresión estática del tipo (43), Engle y Granger (1987)

sugieren sustituir $\hat{e}_{t-1}$ en la representación ECM, por ejemplo (45), y estimar en una segunda etapa los parámetros $\alpha_1$ y $\alpha_2$ asociados a las variables I(0). La intuición subyacente a este procedimiento es que el estimador $\hat{\alpha}$, en la primera etapa, es $Op(T^{-1})$, mientras que los estimadores $\hat{\alpha}_1$ y $\hat{\alpha}_2$, en la segunda etapa, son $Op(T^{-1/2})$. Engle y Granger demuestran que los estimadores en dos etapas de los coeficientes de una ecuación perteneciente a un sistema ECM, utilizando $\alpha$ como verdadero valor, tienen idénticas distribuciones límite a las obtenidas a partir de estimadores máximoverosímiles que utilicen el verdadero valor de $\alpha$. Aplicaciones recientes de esta metodología de estimación, se encuentra, inter alia, en los trabajos de Hall (1986), Jenkinson (1986), Campbell y Shiller (1987), Campbell (1987) y Andrés et al. (1989).

Tal como se mencionaba previamente, un procedimiento alternativo de estimación en dos etapas, consiste en estimar directamente $\alpha$ a partir de la relación dinámica, del tipo (42), obteniendo $\hat{\hat{\alpha}}$. Banerjee et al. (1988) demuestran los estimadores obtenidos por este procedimiento tienen las mismas propiedades que los anteriores. Además de la posible reducción en sesgo de este segundo procedimiento (véase Banerjee et al. (1986)), cuando $z_t$ es débilmente exógena, es posible demostrar (véase Dolado y Ericson (1989)) que $T(\hat{\hat{\alpha}}-\alpha)$ es asintóticamente normal mientras que $T(\hat{\alpha}-\alpha)$ no es normal, cuando el modelo tiene dinámica. Puesto que muchos modelos de regresión pueden interpretarse como distribuciones condicionales autónomas o, en otras palabras, los regresores en dichos modelos son débilmente exógenos, en estos casos puede llevarse a cabo inferencia sobre los coeficientes de las variables que aparecen en la relación de cointegración.

Como una crítica adicional del primer procedimiento en dos etapas, Gourieroux et al. (1987) examinan las velocidades de convergencia de $\alpha$ y $\alpha$, bajo exogeneidad débil, cuando las variables son I(d). Se analizan tres casos de no estacionariedad, incluyendo diferenciación fraccional, i.e. $0,5<d<1$, $d=1$ y $d>1$, encontrando que los grados de convergencia de $\hat{\alpha}$ son $T^{-(2d-1)}$, $T^{-1}$ y $T^{-1}$, mientras que los correspondientes a $\hat{\hat{\alpha}}$ son $T^{-d}$ en todos los casos[6]. A la vista de estos resultados, resulta claro que $\hat{\hat{\alpha}}$

converge más rápidamente que $\hat{\alpha}$ a su verdadero valor en todos los casos, excepto cuando d=1. Dado que d=1 es un caso tan especial, de ello se deriva cierta desconfianza hacia el procedimiento estático en dos-etapas.

Finalmente es necesario destacar la importancia de imponer las raíces unitarias, que se conoce que existen a fin de llevar a cabo inferencia sobre el vector de cointegración bajo la hipótesis de cointegración. Tal como se mencionaba anteriormente, Sims et al. (1986) han demostrado la posibilidad de aplicar técnicas estandar de contrastación en modelos VAR con regresores integrados y en presencia de cointegración. Sin embargo, este resultado no se verifica cuando se trata de hacer inferencia sobre los parámetros del vector de cointegración (véase Phillips (1988) y Dolado y Ericsson (1989)), tal como se ilustra en el siguiente ejemplo

Ejemplo 9

Considérese el siguiente PGD

$$y_t = \beta x_t + \varepsilon_{1t}$$

$$x_t = \alpha x_{t-1} + \varepsilon_{2t}; \ \alpha = 1; \ x_0 = 0, \ \varepsilon_i \sim niid(0,\sigma_i^2); \ E(\varepsilon_{1t} \ \varepsilon_{2s}) = \rho \ \sigma_1 \ \sigma_2 \ \delta_{ts}$$

Supongamos inicialmente que se impone $\alpha=1$. Como las perturbaciones están correlacionadas $\beta$ ha de estimarse por máxima verosimilitud, que corresponde a estimar $\beta$ por MCO en el modelo ampliado

$$y_t = \beta \ x_t + \gamma \ \Delta \ x_t + w_t$$

donde $\gamma = \rho\sigma_1/\sigma_2$ y $E[w_t \varepsilon_{2t}] = 0$. Puede, entonces, demostrarse que, de acuerdo con (i)-(iii) en la Sección 2

$$T(\hat{\beta}-\beta) \rightarrow \frac{\int_0^1 V(r) \, dW(r)}{\int_0^1 V^2(r) dr} \equiv MN$$

donde $P_t(=\sum_1^t \varepsilon_{2i}) \rightarrow V(r)$ y $S_t(=\sum_1^t w_i) \rightarrow W(r)$. Dada que, por construcción, V y W son procesos de Wiener independientes, es posible demostrar que el funcional de Wiener anterior sigue una mixtura de distribuciones normales (MN) véase Phillips and Park (1988)) que es equivalente a la distribución $N(0,\sigma_2^2\sigma_w^2/2)$.

Sin embargo, si $\alpha=1$ no se impone, sino que $\alpha$ se estima, entonces esposible demostrar que el estimador maximoverosímil de $\beta$ en este caso, es tal que

$$T(\hat{\beta}-\beta) \rightarrow \gamma \frac{\frac{1}{2}[V(1)^2-1]}{\int_0^1 V(r)^2 dr} + \frac{\int_0^1 V(r)\, dW(r)}{\int_0^1 V^2(r)\, dr}$$

Puesto que el primer término de la distribución asintótica corresponde a (4) (con $\sigma_\varepsilon^2=\sigma^2$), se conoce que esta distribución no es estándar. Por tanto, a menos que $\gamma=0$ ($\rho=0$) la inferencia sobre $\beta$ ha de hacerse con distribuciones simuladas.

## 4.3 Contrastación

Contrastar la existencia de cointegración entre un conjunto de variables corresponde a contrastar la existencia de una raíz unitaria. Por tanto, el análisis subyacente a este tipo de contraste, tal como se apunta brevemente en la Sección 2, se aplica también a este caso. La diferencia estriba en que ahora se contrasta la presencia de una raíz unitaria en una serie construida, i.e. residuos minimocuadráticos, en vez de series originales. Por este motivo, los valores críticos de este tipo de contrastes han de ser ajustados al alza, ya que de otra forma el error de tipo 1 estaría exagerado. Si no se rechaza la existencia de raíces unitarias en dichos residuos, las "desviaciones del equilibrio" serían integradas y por tanto no existiría un mecanismo de ajuste sitemático del sistema hacia la solución de equilibrio.

En las aplicaciones empíricas los contrastes de mayor uso son el contraste de cointegración basado en el estadístico Durbin–Watson (CRDW) y los contrastes de Dickey y Fuller (1979) en sus versiones simples (DF) y aumentado (ADF). El contraste CRDW, sugerido por Sargan y Bhargava (1983) está basado en la definición usual del estadístico DW

$$CRDW = (\Sigma\ \hat{e}_t^2)^{-1}\ [\Sigma(\hat{e}_t - \hat{e}_{t-1})^2] \qquad (46)$$

donde $\hat{e}_t$ son los residuos minimocuadráticos de (40). Existen diversos problemas con este contraste. En efecto, mientras que cuando se trata de contrastar el grado de integración de una serie (con una constante como único regresor), los valores críticos del contraste, computados por Sargan y Bhargava, son exactos, el mismo contraste aplicado al caso de cointegración depende del número de regresores en la relación de cointregación, existiendo sólo cotas de la región crítica. Esto ocurre porque, como en el caso estandar de aplicación del DW (basado en la hipótesis nula de perturbaciones incorrelaciondas), los valores críticos exactos del CRDW (basado en la hipótesis de una raíz unitaria en las perturbaciones) son una función del PGD. Las cotas de este contraste proporcionan un intervalo que puede utilizarse para no rechazar la hipótesis nula de ausencia de cointegración, pero la amplitud de dicho intervalo aumenta conforme lo hace el número de regresores. Sin el uso, por ejemplo, de la Rutina de Imhof (1961), es difícil hacer inferencia cuando el valor del CRDW cae dentro del intervalo. Por supuesto, resulta factible computar valores críticos exactos para un PGD dado. Así, Engle y Granger (1987) presentan valores críticos para un paseo aleatorio bivariante, pero dichos valores no son generalmente aplicables a otros experimentos cuyos PGP sean más complejos. En este sentido el CRDW resulta ser asintóticamente similar (véase Cox y Hinkley (1974)) cuando el verdadero proceso seguido por las perturbaciones en un AR(1).

Problemas similares afectan al estadístico DF, pero no afectan al ADF si el número de desfases $\Delta e_{t-i}$ que aparecen en el contraste coincide con el del PGD. En la práctica el número de desfases que se utiliza es, en gran medida, ad–hoc y ello afecta a la potencia del contraste. Todo ello

sugiere que quizá se deba prestar mayor atención a las versiones no-paramétricas (véase Ouliaris et al. (1988)) o recursivas (véase Dolado et al. (1989)) de dichos contrastes.

Con fines ilustrativos se presenta en el Cuadro 1, los valores críticos (5%) de los tres contrastes anteriores, computados por Engle y Yoo (1987), donde el PGD es un paseo aleatorio n-variante, con n≤5

[Vease Cuadro 1]

Tal como ocurre con los contrastes de raíces unitarias en series temporales individuales es importante destacar la falta de potencia de estos contrastes cuando la verdadera raíz se encuentra en un entorno de la unidad. Así, por ejemplo, Engle y Granger muestra que cuando el PGD de las perturbaciones de la relación de cointegración es un AR(1) con parámetro igual a 0,9, las potencias de los contrastes CRDW, DF y ADF (al 5%) son 20%, 15% y 11% respectivamente. Cuando se altera el PGD a un proceso AR más general la potencia del ADF es del 60%, dominando a los contrastes CRDW y DF. A la vista de estos resultados quizá resulte conveniente aconsejar el uso del CRDW en una primera instancia, dada su simplicidad, y, en caso de incertidumbre sobre la posibilidad de rechazo de la hipótesis nula, utilizar el ADF[7].

Un enfoque alternativo, discutido previamente, se basa en contrastar la presencia de cointegración utilizando la solución de equilibrio a largo plazo del modelo dinamico, como en (42), especialmente cuando los regresores sean débil o fuertemente exógenos. En la representación ECM, como en (45), el t-ratio del término de corrección del error, $(\hat{\alpha}_1-1)$, es un estadístico de interés, ya que si no existiese cointegración debería converger a cero a velocidad $Op(T^{-1})$. Banerjee et al. (1986) muestran, por simulación como dicho t-ratio tiene el tamaño adecuado al 5%, aunque los resultados de Dolado y Ericsson (1989) muestran que puede haber divergencias para otros PGD's. Sin embargo, de acuerdo con el estudio de Banerjee et al., la potencia de este contraste es muy superior a la de los contrastes basados en la regresión estática. Una posible explicación de esta ventaja comparativa reside en los sesgos en muestras finitas que presenta el enfoque estático.

Finalmente, resulta importante comentar que, dada la fragilidad de los contrastes, otros procedimientos auxiliares menos refinados estadísticamente puede resultar útiles. La observación del correlograma de los residuos puede ayudar en gran medida. Granger y Weiss (1983) también sugieren aumentar los coeficientes estimados del vector de cointegración en una cierta proporción, digamos el 10%, y examinar, entonces, si la suma de cuadrados de las desviaciones del nuevo vector es bastante superior a la obtenida con el vector estimado. La intuición de este procedimiento informa es clara, puesto que sólo el uso del verdadero vector de cointegración hará que dicha varianza sea finita y, por tanto, será fácilmente distinguible de otros casos.

## 4.4. Predicción con Sistemas Cointegrados

Tal como han demostrado Engle y Yoo (1987), las predicciones derivadas de sistemas cointegrados tienen una propiedad que no comparte los sistemas de variables integradas. Esta propiedad implica la existencia de combinaciones lineales de dichas predicciones que son idénticamente cero para horizontes temporales largos, independientemente del punto de origen y, además, la varianza del error de predicción de dichas combinaciones es finita, mientras que, para cualquier otra combinación, diverge conforme se amplie el horizonte de predicción. El siguiente ejemplo ilustra esta propiedad.

### Ejemplo 10

Considerése el proceso $\{x_t\}$ cuyo PGD es

$$\Delta x_t = \varepsilon_t - C_1 \varepsilon_{t-1};\ \varepsilon_0 = x_0 = 0;\ \varepsilon_t \sim (0,\Omega)$$

tal que $x_t$ es C(1,1). Por tanto existe un vector $\alpha$, tal que

$$\alpha' C(1) = \alpha'(I - C_1) = 0$$

El valor $x_{t+h}$ puede expresarse recursivamente como

$$x_{t+h} = \varepsilon_{t+h} + (I - C_1) \sum_{1}^{t+h-1} \varepsilon_j$$

Por tanto

$$E(x_{t+h}/t) = x_{t+h/t} = (I - C_1) \sum_{1}^{t} \varepsilon_j$$

y

$$e_{t+h/t} = \varepsilon_{t+h} + (I - C_1) \sum_{1}^{h} \varepsilon_{t+j}$$

Se sigue que

$$\alpha' \, x_{t+h/t} = 0$$

y

$$\text{var}\,(\alpha' \, x_{t+h/t}) = \alpha' \, \Omega \, \alpha < \infty$$

Con el fin de examinar el comportamiento predictivo de métodos de estimación que ignoran la cointegración existente y aquellos que la tienen en cuenta. Engle y Yoo analizan el comportamiento predictivo relativo a un sistema cointegrado estimado como un VAR irrestringido (equivalente a aplicar MCO ecuación a ecuación) y reparametrizado como un modelo ECM, aplicando el procedimiento estático en dos etapas El PGD viene dado por

$$\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0,8 & 0,1 \\ 0,8 & 0,6 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \varepsilon_t \qquad (47)$$

donde ambas varibles son I(1) y el vector de cointegración es (1,−0,5).

Los resultados, medidos a través del Error Cuadrático Medio de la Predicción (ECMP) muestran una ligera ventaja del método VAR hasta un horizonte de 4 períodos, problablemente debido al sesgo en la estimación del vector cointegrador, pero posteriormente el método ECM se muestra muy superior. Así, después de 20 períodos, el ECMP del modelo ECM es un 40% inferior al del VAR, y probablemente esa mejora en la predicción sería muy superior si el vector de cointegración se hubiese estimado con menor sesgo.

## 5. ESTIMACION Y CONTRASTACION EN UN CONTEXTO DE MAXIMA VEROSIMILITUD

El propósito de esta sección es el de ofrecer un resumen de ciertos desarrollos recientes (véase Johansen (1988) y Johansen y Juselius (1988)) que tratan los problemas de estimación y contrastación de sistemas cointegrados en un contexto de máxima verosimilitud[8]. El enfoque está limitado, por el momento, a sistemas C(1,1) que admiten la representación VAR dada por

$$x_t = \sum_{1}^{p} A_i \, x_{t-i} + \varepsilon_t ; \; \varepsilon_t \sim NI(0,\Omega) \qquad (48)$$

La hipótesis nula de interés es que la dimensión del subespacio de vectores de cointegración es r, una hipótesis que, en este contexto, viene expresada mediante una simple formulación paramétrica en términos de los vectores de cointegración y sus ponderaciones respectivas. Se utiliza un procedimiento de máxima verosimilitud para contrastar dicha hipótesis.

El procedimiento se basa en la reparametrización de (48) en términos de

$$\Delta x_t = -A(1) \, x_{t-1} + \sum_{2}^{p} \Pi_i \, \Delta x_{t-p+1} + \varepsilon_t \qquad (49)$$

donde $A(1) = I - \Sigma A_i$ y $\Pi_i = - \sum_{1}^{p} A_i$

Si el sistema (49) está cointegrado, se conoce, a partir de (21), que $A(1)=B\Gamma'$. La hipótesis es, por tanto, $H_0$: rango $A(1)=r$. Existen tres casos de interés. En primer lugar, si $r=n$ entonces la matriz $A(1)$ tiene rango completo por lo que $x_t$ es estacionario. En segundo lugar, si $r=0$ entonces no existe cointegración y el modelo debería ser estimado en diferencias, eliminando el término en niveles en (49). En tercer lugar si $0<r<n$, existen r vectores de cointegración

La función de verosimilitud (en logs) de (49) puede escribirse del siguiente modo

$$L(B,\Gamma,\Pi_i;\ \Omega) = -\frac{T}{2}\log|\Omega| - \frac{1}{2}\sum_1^T[\Delta x_t+B\Gamma' x_{t-1}-\Sigma\ \Pi_i\Delta x_{t-i+1}]'\Omega^{-1}$$

$$[\Delta x_t+ B\Gamma'\ x_{t-1}- \Sigma\ \Pi_i\ \Delta x_{t-i+1}] + \text{const} \qquad (50)$$

A continuación se concentra la función L respecto a las matrices $\Pi_i$'s. Como es bien conocido del análisis de regresión multivariante, ello se logra mediante la obtención de las matrices de residuos de las regresiones de $\Delta x_t$ y $x_{t-1}$ sobre $\Delta x_{t-1}\ldots\ \Delta x_{t-p+1}$, denotando dichas matrices para $R_{0t}$ y $R_{1t}$ respectivamente. Seguidamente se formula el siguiente modelo de regresión concentrado.

$$R_{ot} = -B\Gamma'\ R_{1t} + \varepsilon_t \qquad (51)$$

Dado que (51) es no lineal en las matrices B y $\Gamma$, sólo es posible concentrar respecto a B si $\Gamma$ se toma como dada. Por tanto, los regresores en (51) habrán de ser interpretados como $\Gamma'R_{1t}$. Las soluciones obtenidas para B y $\Omega$ vienen dadas por

$$B(\Gamma) = -S_{01}\Gamma[\Gamma'\ S_{11}\ \Gamma]^{-1} \qquad (52)$$

y

$$\Omega(\Gamma) = S_{00} - S_{01}\,\Gamma[\Gamma'\,S_{11}\Gamma]^{-1}\Gamma'S_{10} \qquad 53$$

tal que $S_{ij} = T^{-1}\sum_{1}^{T} R_{it}\,R'_{jt}$ $(i,j=0,1)$

Por tanto, la función de verosimilitud concentrada viene dada por

$$L^*(\Gamma) = \text{const} - T/2\,\log\,|\Omega(r)| \qquad (54)$$

Maximizando (54) respecto a r es equivalente a

$$\min_{\Gamma}\ |S_{00}-S_{01}\Gamma(\Gamma'S_{11}\Gamma)^{-1}\Gamma'S_{10}| \qquad (55)$$

A partir de la siguiente conocida relación entre determinantes de matrices particionadas (véase Dhrymes (1978))

$$\begin{vmatrix} S_{00} & S_{01}\Gamma \\ \Gamma'S_{10} & \Gamma'S_{11}\Gamma \end{vmatrix} = |S_{00}|\,|\Gamma'S_{11}\,\Gamma-\Gamma\,S_{10}\,S_{00}^{-1}S_{01}\Gamma| =$$

$$= |\Gamma'S_{11}\Gamma|\,|S_{00}-S_{01}\Gamma(\Gamma'S_{11}\Gamma)^{-1}\Gamma'S_{10}| \qquad (56)$$

se sigue que minimizar (55) es equivalente a

$$\min_{\Gamma}\ |\Gamma'S_{11}\Gamma|^{-1}\,|\Gamma'S_{11}\Gamma-\Gamma'S_{10}S_{00}^{-1}S_{01}\Gamma| \qquad (57)$$

Puesto que la solución al problema (57) carece claramente de unicidad, se impone la normalización $\Gamma'S_{11}\Gamma = I$, convirtiendo el problema en

$$\min_{\Gamma}\ |I-\Gamma'S_{10}S_{00}^{-1}\Gamma| \ \text{ s.a. } \Gamma'S_{11}\Gamma = I \qquad (58)$$

La solución al anterior problema es bien conocida en la literatura (véase, por ejemplo, Anderson (1984)) y se obtiene calculando los valores propios de la matriz $S_{10}S_{00}^{-1}S_{01}$ con respecto a la matriz $S_{11}$, resolviendo

$$|\hat{\lambda}\, S_{11} - S_{10}\, S_{00}^{-1}\, S_{01}| = 0 \tag{59}$$

Los vectores propios correspondientes vienen dados por

$$S_{10}\, S_{00}^{-1}\, S_{01}\, \hat{v}_i = \hat{\lambda}_i\, S_{11}\, \hat{v}_i \quad (i=1...n) \tag{60}$$

$$\text{s.a. } \hat{v}_i'\, S_{11}\, \hat{v}_j = S_{ij}$$

donde $\delta$ es la delta de Kroenecker.

Resulta posible demostrar que los valores propios $\hat{\lambda}_i \varepsilon (0,1)$ y que los vectores de cointegración agrupados en la matriz $\Gamma$, corresponden a los r vectores propios, $\hat{v}_i (i=1...r)$ asociados con los r valores propios mayores. Nótese que los estimadores maximoverosímiles de $\Pi_1 (i=2...p)$ y de B y $\Omega$ pueden obtenerse a partir de (49), (52) y (53) sustituyendo $\Gamma$ por su estimador $\hat{\Gamma}$.

Con el fin de constuir un contraste basado en el ratio de verosimilitud (LR), se sustituye (56) en (54), obteniendo

$$L^*(\Gamma) = \text{cnst} - T/2 \log \left| |\Gamma' S_{11} \Gamma|^{-1} (|S_{00}||\Gamma' S_{11} \Gamma - \Gamma' S_{10} S_{00}^{-1} S_{01} \Gamma|) \right| \tag{61}$$

A continuación, utilizando (59) y la propiedad que establece que el determinante de una matriz es igual al producto de sus valores propios, bajo la hipótesis nula de que existen r vectores de cointegración, la función $L^*$ toma el valor

$$\tilde{L}^*(\Gamma) = -\frac{T}{2} \log |S_{00}| - \frac{T}{2} \sum_{1}^{r} \log (1-\hat{\lambda}_i) + \text{const} \tag{62}$$

donde los valores propios $\hat{\lambda}_i$ han sido ordenados de menor a mayor ($\hat{\lambda}_1 < \ldots < \hat{\lambda}_r$). El máximo valor de la función, $\hat{L}^*$, correspondiente a la hipótesis no restringida, es igual a (62) con el sumatorio de 1 a n. por tanto el contraste LR de $H_0$: rango $\Gamma$ = r viene dado por

$$LR = 2\,(\hat{L}^* - \tilde{L}^*) = -T \sum_{r+1}^{n} \log\,(1-\hat{\lambda}_i) \qquad (63)$$

Johansen (1987) demuestra que la distribución de LR no es estandar, correspondiendo a la versión multivariante del cuadrado de (i) en la Sección 2, dada por la siguiente traza (tr)

$$tr[\int_0^1 W(r)dW(r)'(\int_0^1 W(r)W'(r)dr)^{-1}\int_0^1 dW(r)W'(r)] \qquad (64)$$

Los valores críticos al 5% han sido tabulados por simulación y se presentan en el Cuadro 2 para n=5

[ Vease Cuadro 2]

También es posible considerar hipótesis acerca de restricciones lineales en B y $\Gamma$ bien separada o conjuntamente. Las hipótesis nulas acerca de B y $\Gamma$ tiene la forma $H_0: B = H\Upsilon$ y $H_0'\Gamma = G\Psi$ mientras que la hipótesis conjunta será $H''_0: B\Gamma' = H\Upsilon\Psi'G'$. Las matrices H y G son (nxs) y (nxq) respectivamente, reduciendo el número de parámetros en B a $\Upsilon$(sxr) y en $\Gamma$ a $\Psi$(qxr) donde r<s<n y r<q<n. Todas estas hipótesis suponen restricciones sobre la matriz A(1) que contiene $n^2$ parámetros (nr+(n–r)r, bajo la hipótesis de cointegración), mientras que bajo $H_o, H'_0$ y $H'_0$ contiene sr+(n–r)r, qr+(n–r) y qr+(s–r)r respectivamente Johansen y Juselius (1988) demuestran que la formulación de Wald de estos contrastes se distribuyen asintóticamente como distribuciones Ji–cuadrado con (n–s)r, (n–q)r y (2n–q–s)r grados de libertad respectivamente. La intuición subyacente a estos resultados es similar a la presentada en el Ejemplo 9. A este respecto, es posible demostrar que los resultados de Dolado y Ericsson (1989) acerca de

la distribución asintótica normal de $\Gamma$, son fácilmente generalizables en este contexto, suponiendo que $\Omega$ es diagonal y que si $b_{ij} \neq 0$, entonces $b_{ik}$ y $b_{kj}$ (k=1...n) son cero. En otras palabras, se elimina la posibilidad de que aparezca una misma relación de cointegración en más de una ecuación, lo que unido a la diagonalidad de $\Omega$, implica la existencia de regresores débilmente exógenos respecto a los elementos de $\Gamma$.

Finalmente, resulta interesante hacer notar la similitud del anterior procedimiento con el procedimiento de "tendencias comunes" desarrollado por Stock y Watson (1988b). Este último procedimiento permite determinar el número de vectores de cointegración por medio del contraste de la hipótesis nula de que el vector $x_t$ presenta (n-r) raíces unitarias distintas, frente a la alternativa de que tiene (n-m) (m>r) raíces unitarias. El contraste se construye transformando $x_t$ tal que los primeros r componentes correspondan a los componentes estacionarios, mientras qu los restantes (n-r) correspondan a componentes integrados. Bajo la hipótesis nula, los valores propios de la autorregresión de primer orden de los (n-r) componentes deberán ser igual a la unidad. Bajo la hipotesis alternativa, solamente (n-m) valores propios habían de ser igual a la unidad, de forma que el valor propio (n-m+1)-smo deberá ser inferior a la unidad. El estadístico del contraste viene dado por $T(\hat{\lambda}_{n-m+1}-1)$ y sus valores críticos han sido obtenidos por simulación, encontrándose en el artículo de Stock y Watson.

Cuadro 1

Valores Críticos del Contraste de Cointegración (5%)

| n | T | CRDW | DF | ADF |
|---|---|---|---|---|
| 2 | 50 | 0,78 | -3,67 | -3,29 |
| | 100 | 0,39 | -3,37 | -3,17 |
| | 200 | 0,20 | -3,37 | -3,25 |
| 3 | 50 | 0,89 | -4,11 | -3,75 |
| | 100 | 0,45 | -3,93 | -3,62 |
| | 200 | 0,30 | -3,78 | -3,78 |
| 4 | 50 | 1,10 | -4,35 | -3,98 |
| | 100 | 0,55 | -4,22 | -4,02 |
| | 200 | 0,40 | -4,18 | -4,13 |
| 5 | 50 | 1,10 | -4,76 | -4,15 |
| | 100 | 0,66 | -4,58 | -4,36 |
| | 200 | 0,47 | -4,48 | -4,43 |

Nota: Los valores críticos del contraste CRDW han sido tomados de Engle y Granger (1987) para n=2. Los restantes valores han sido interpolados utilizando los resultados en Sargan y Bhargava (1983). Los valores críticos de los contrastes DF y ADF han sido tomados de Engle y Yoo (1987).

Cuadro 2

Valores Críticos del Contraste LR (5%)

| r | LR |
|---|---|
| 0 | 57,2 |
| 1 | 38,6 |
| 2 | 23,8 |
| 3 | 12,0 |
| 4 | 4,2 |

Nota: Los valores críticos han sido tomados de Johansen y Juselius (1988).

## Notas

1. Como muestra del interés que ha despertado el tema podemos citar que, desde 1986, han aparecido en la literatura dos números especiales de revistas dedicadas al tema (véase Oxford Bulletin of Economics and Statistics (1986) y Journal of Economic Dynamics and Control (1988)) y dos panorámicas (véase Dolado y Jenkinson (1987) y Stock y Watson (1988a).

2. El interés en el tema se remonta a nombres tan ilustres como Jevons (1884), Working (1934) y principalmente Yule (1926).

3. A lo largo de este artículo se restringe la discusión al concepto de integración en varianza. Un análisis excelente de dicho concepto aplicado a otros momentos se encuentra en Escribano (1987).

4. Una discusión extensa de las propiedades asintóticas de procesos I(d) con d>1 puede encontrarse en Gourieroux et al. (1987).

5. Una ampliación de los Teoremas de Representación a procesos no lineales puede encontrarse en Escribano (1986).

6. Para un análisis del concepto de diferenciación fraccional véase Granger y Joyeux (1980).

7. Bajo una interpretación similar puede llevarse a cabo contrastes de cointegración en diferentes frecuencias (véase Engle et al. (1988) y Osborn et al. (1989)).

8. Este enfoque está claramente relacionado con el enfoque de análisis canónico propuesto por Box y Tiao (1981) y Peña y Box (1987).

## Referencias

Anderson, T. (1984). "An Introduction to Multivariate Statistical Analysis", New York, Wiley.

Andres, J., Escribano, A. Molinas, C. y Taguas, D. (1988). "Los Determinantes de la Inversión Productiva Privada en España" (de próxima aparición en Moneda y Crédito (nueva época)).

Aoki, M. (1987). "An Alternative Measure of Random Walk Components in Time Series" Economics Letters, 24, 227-230.

Banerjee, A., Dolado, J., Hendry, D. and Smith, G. (1986). "Exploring Equilibrium Relationshisps in Econometrics through Static Models: Some Monte Carlo Evidence" Oxford Bulletin of Economics and Statistics 48, 253-277.

Banerjee, A. and Dolado, J. (1988). "Tests of the Life Cycle Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small-Sample Interpretations" Oxford Economic Papers, 40, 610-633.

Banerjee, A., Galbraith, J. and Dolado, J. (1988) "Dynamic Specification with the General Error-Correction Form" (de próxima aparición en Oxford Bulletin of Economics and Statistics).

Banerjee, A., Dolado, J., Hendry, D. and Galbraith, J. (1989) "Equilibrium and Error Correction in Econometrics" (de próxima aparición en Oxford University Press).

Billingsley, P. (1986). "Convergence of Probability Measures" New York, Wiley

Box, G. and Jenkins, G. (1976). "Time Series Analysis: Forecasting and Control; San Francisco, Holden-Day.

Box, G. and Tiao, G. (1981). "A Canonical Analysis of Multiple Time Series with Applications" Biometrika, 64, 355-365.

Campbell, J. (1987). "Does Saving Anticipate Declining Labour Income? An Alternative Test of the Permanent Income Hypothesis" Econometrica, 55, 1249–1273.

Campbell, J. And Shiller, R. (1987). "Cointegration and Tests of Present Value Models" Journal of Political Economy, 95, 1062–1088.

Cox, D. and Hinckley D. (1974). "Theoretical Statistics" London, Chapman and Hall.

Currie, D. (1981). "Some Long–Run Features of Dynamic Time–Series Models" Economic Journal, 363, 704–715.

Davidson, J., Hendry, D. Srba, F. and Yeo, S. (1978) "Econometric Modelling of the Aggregate Time Series Relationship between Consumers Expenditure and Income in the U.K." Economic Journal, 88, 661–632.

Davidson, J. and Hendry, D. (1981). "Interpreting Evidence: The Behaviour of Consumers 'Expenditure in the U.K." European Economic Review, 16, 177–192.

Dickey, D., Bell, W. and Miller, R. (1986). "Unit Roots in Time–series Models: Tests and Implications" American Statistician, 40, 12–26.

Dickey, D. and Fuller, W. (1979. "Distribution of the Estimators for Autoregressive Time–Series with a Unit Root". Journal of the American Satistical Association, 74, 427–431.

Dickey, D. and Fuller, W. (1981) "Likelihood Ratio Statistics for Autoregresive Time–Series with a Unit Root" Econometrica, 49, 1057–1072.

Dickey, D. and Pantula, S. (1987). "Determining the Order of Differencing in Univariate Time–Series" Journal of Business and Economic Statistics, 5, 455–462.

Dolado, J. and Jenkinson (1987) "Cointegration: A Survey of Recent Developments" (de próxima aparición en Economic Perspectives).

Dolado, J., Banerjee, A. and Galbraith, J. (1988). "Using Recursive Methods to look for a Unit Root in GNP" (de próxima aparición en Review of Economics and Statistics).

Dolado, J. and Ericsson, J. (1989) "Inference in Conditional dynamic Models with Integrated Variables" (mimeo).

Dolado, J. (1988). "Innovación Financiera, Inflación y Estabilidad de la Demanda de ALP en España" Boletín Económico, Abril, 19–35.

Dhrymes, P. (1978). "Mathematics for Econometrics" New York Springer–Verlag.

Edison, H. and Kloveland, J. (1987). "A Quantitative Reassessment of the Purchasing Power Parity: Evidence from Norway and the U.K." Journal of Applied Econometrics, 2. 209–334.

Engle, R. Hendry, D. and J. Richard (1983). "Exogeneity" Econometrica, 51, 277–304.

Engle, R. and Granger, C. (1987). "Cointegration and Error Correction: Representation, Estimation and Testing" Econometrica, 49, 1057–1072.

Engle, R. and Yoo, S. (1987). "Forecasting and Testing in Cointegrated Systems" Journal of Econometrics, 35, 143–159.

Engle, R., Granger, C., Hylleberg, S. and Yoo, S. (1988). "Seasonal Integration and Cointegration" University of S. Diego Discussion Paper 88–32.

Escribano, A. (1986). "Identification and Modelling of Economic Relationships in a Crowing Economy" Ph. D. Dissertation University of California, S. Diego.

Escribano, A. (1987). "Cointegration, Time Co-trends and Error Corretion Systems: An Alternative Approach" CORE Discussion Paper 8715.

Evans, G. and Savin, N. (1981). "Testing for Unit Roots I'Econometrica 49, 753-777.

Friedman, M. (1957). "A Theory of the Consumption Function" Princeton N.J. Pricenton University Press.

Fuller, W. (1976). "Introduction to Statistical Time-Series". New York. Wiley.

Gourieroux. C., Maurel, F. and Monfort, A. (1987). "Regression and Non Stationarity" ENSAE Discussion Paper 8708, Institut National de la Statistique et des Etudes Economiques.

Granger, C. and Newbold, P. (1974). "Spurious Regressions in Econometrics" Journal of Econometrics, 9, 111-120.

Granger, C. (1981). "Some Properties of Time-Series Data and their Use in Econometric Model Specification" Journal of Econometrics, 16, 121-130.

Granger, C. and Weiss, A. (1983). "Time-series Analysis of Error Correcting Models" in Karlin, S., Amemiya, T. and Goodman, L. (eds). Studies in Econometric Time-Series and Multivariate Statistics. New York. Academic Press.

Hall, S. (1986). "An Application of the Granger and Engle Two-Step Estimation Procedure to U.K. Aggregate Wage DAta" Oxford Bulletin of Economics and Statistics, 48, 229-240.

Hendry, D. and Mizon, G. (1978). "Serial Correlation as a Convenient Simplification, not a Nuissance: A Comment on a Study of the Demand for Money by the Bank of England" Economic Journal, 88, 549-563.

Hendry, D. and Richard, J.F. (1982). "On the Formulation of Empirical Models in Dynamic Econometrics" Journal of Econometrics, 20, 3–33.

Hendry, D. and Richard, J.F. (1983). "The Econometric Analysis of Economic Time Series" International Statistical Review, 51, 111–163.

Hendry, D. and Ericsson, N. (1987). "Assertion Without Empirical Basis: An Econometric Appraisal of 'Monetary Trends... in the U.K.' by M. Friedman and A. Schwartz 'Applied Economics Discussion Paper 25, University of Oxford.

Imhof, P. (1961). "Computing the Distribution of Quadratic Forms in Normal Variates" Biometrika, 48, 419–426.

Jenkinson, T. (1986) "Testing Neo–Classical Theories of Labour Demand: An Application of Cointegration Techniques" Oxford Bulletin of Economics and Statistics, 48, 241–251.

Journal of Economic Dynamics and Control (1988). "Economic Time–Series with Random Walks and other Non Stationary Components" ed. by M. Aoki.

Johansen, S. (1988). "Statistical Analysis of Cointegration Vectors" Journal of Economic Dynamics and Control, 12, 231–254.

Johansen, S. and Juselius, K. (1988) "Hypothesis Testing for Cointegration Vectors with an Application to the Demand for Money in Denmark and Finland" Institute of Economics Discussion Paper 88–05. University of Copenhagen.

Jevons, W. (1884). "Investigation in Currency and Finance" London. Mc. Millan.

Maravall, A. (1979). "Identification in Dynamic Shock–Error Models" New York, Springer–Verlag.

Nelson, C. and Plosser, c. (1982). "Trends and Random Walks in Macroeconomic Time-Series" Journal of Monetary Economics, 10, 139-162.

Osborn, D. Chui, A., Smith, J. and Birchenhall, C. (1989). "Seasonality and the Order of Integration for Consumption" (de próxima aparición en Oxford Bulletin of Economics and Statistics)

Oxford Bulletin of Economics and Statistics (1986). "Economic Modelling with Cointegrated Variables" ed by D. Hendry.

Peña, D. and Box, G. (1987). "Identifying a Simplifying Structure in Time Series" Journal of the American Statistical Association, 82, 836-843.

Perron, P. (1988). "Trends and Random Walks in Macroeconomic Time-Series. Further Evidence from a New Approach" Journal of Economic Dynamics and Control, 12, 297-332.

Phillips, A. (1954). "Stabilization Policy in a Closed Economy" Economic Journal, 64, 290-323.

Phillips, P. (1986). "Understanding Spurious Regressions In Econometrics" Journal of Econometrics, 33, 311-340.

Phillips, P. (1987). "Time Series Regression with a Unit Root" Econometrics, 55, 277-301.

Phillips, P. and Ouliaris, S. (1988). "Testing for Cointegration Using Principal Components Methods" Journal of Economic Dynamics and Control, 12, 205-230.

Phillips, P. (1988). "Optimal Inference in Cointegrated Systems" Cowles Foundation Discussion Paper 866.

Phillips, P. adn Park, J. (1988). "Asymptotic Equivalence of Ordinary Leasat Squases and Generalized Least Square in Regressions with Integrated Variables" Journal of the American Statistical Association, 83, 111–115.

Salmon, M. (1982). "Error Correction Mechanisms" Economic Journal, 92, 615–629.

Sargan, J. D. (1964). "Wages and Prices in the U.K.. A Study in Econometric Methodology" in Hart, P., Mills, G. and Whittaker, J. (eds) Econometric Analysis for National Economic Planning, London, Butterworths.

Sargan, J.D. and Bhargava, A. (1983). "Testing Residuals from Least Square Regression for Being Generated by the Gaussian Random Walk" Econometrica, 51, 153–174.

Sims, C., Stock, J. and Watson, M. (1986). "Inference in Linear Time–Series Models with Some Unit Roots" (mimeo).

Stock, J. (1987). "Asymptotic Properties of Least Squares Estimates of Cointegration Vectors" Econometrica, 1035–1056.

Stock, J. and Watson, M. (1988a). "Variable Trends in Economic Time Series" Journal of Economic Perspectives, 2, 147–174.

Stock, J. and Watson, M. (1988b) "Testing for Common Trends" Journal of the American Statistical Association, 83, 1097–1107.

White, H. (1984). "Asymptotic Theory for Econometricians" London, Academic Press.

Working, H. (1934). "A Random Difference Series for Use in the Analysis of Time–Series" Journal of the American Statistical Association, 29, 11–24.

Yule, G. (1926). "Why Do We Sometiemes get Nonsense-Correlations Between Time-Series? A Study in Sampling and the Nature of Time-Series" Journal of the Royal Statistical Society, 89, 1-64.

**DOCUMENTOS DE TRABAJO** (1):

*8501* **Agustín Maravall**: Predicción con modelos de series temporales.
*8502* **Agustín Maravall**: On structural time series models and the characterization of components.
*8503* **Ignacio Mauleón**: Predicción multivariante de los tipos interbancarios.
*8504* **José Viñals**: El déficit público y sus efectos macroeconómicos: algunas reconsideraciones.
*8505* **José Luis Malo de Molina y Eloísa Ortega**: Estructuras de ponderación y de precios relativos entre los deflactores de la Contabilidad Nacional.
*8506* **José Viñals**: Gasto público, estructura impositiva y actividad macroeconómica en una economía abierta.
*8507* **Ignacio Mauleón**: Una función de exportaciones para la economía española.
*8508* **J. J. Dolado, J. L. Malo de Molina y A. Zabalza**: Spanish industrial unemployment: some explanatory factors *(versión inglés)*. El desempleo en el sector industrial español: algunos factores explicativos *(versión español)*.
*8509* **Ignacio Mauleón**: Stability testing in regression models.
*8510* **Ascensión Molina y Ricardo Sanz**: Un indicador mensual del consumo de energía eléctrica para usos industriales, 1976-1984.
*8511* **J. J. Dolado y J. L. Malo de Molina**: An expectational model of labour demand in Spanish industry.
*8512* **J. Albarracín y A. Yago**: Agregación de la Encuesta Industrial en los 15 sectores de la Contabilidad Nacional de 1970.
*8513* **Juan J. Dolado, José Luis Malo de Molina y Eloísa Ortega**: Respuestas en el deflactor del valor añadido en la industria ante variaciones en los costes laborales unitarios.
*8514* **Ricardo Sanz**: Trimestralización del PIB por ramas de actividad, 1964-1984.
*8515* **Ignacio Mauleón**: La inversión en bienes de equipo: determinantes y estabilidad.
*8516* **A. Espasa y R. Galián**: Parsimony and omitted factors: The airline model and the census X-11 assumptions *(versión inglés)*. Parquedad en la parametrización y omisiones de factores: el modelo de las líneas aéreas y las hipótesis del census X-11 *(versión español)*.
*8517* **Ignacio Mauleón**: A stability test for simultaneous equation models.
*8518* **José Viñals**: ¿Aumenta la apertura financiera exterior las fluctuaciones del tipo de cambio? *(versión español)*. Does financial openness increase exchange rate fluctuations? *(versión inglés)*.
*8519* **José Viñals**: Deuda exterior y objetivos de balanza de pagos en España: Un análisis de largo plazo.
*8520* **José Marín Arcas**: Algunos índices de progresividad de la imposición estatal sobre la renta en España y otros países de la OCDE.
*8601* **Agustín Maravall**: Revisions in ARIMA signal extraction.
*8602* **Agustín Maravall y David A. Pierce**: A prototypical seasonal adjustment model.
*8603* **Agustín Maravall**: On minimum mean squared error estimation of the noise in unobserved component models.
*8604* **Ignacio Mauleón**: Testing the rational expectations model.
*8605* **Ricardo Sanz**: Efectos de variaciones en los precios energéticos sobre los precios sectoriales y de la demanda final de nuestra economía.
*8606* **F. Martín Bourgón**: Indices anuales de valor unitario de las exportaciones: 1972-1980.
*8607* **José Viñals**: La política fiscal y la restricción exterior. (Publicada una edición en inglés con el mismo número).
*8608* **José Viñals y John Cuddington**: Fiscal policy and the current account: what do capital controls do?
*8609* **Gonzalo Gil**: Política agrícola de la Comunidad Económica Europea y montantes compensatorios monetarios.
*8610* **José Viñals**: ¿Hacia una menor flexibilidad de los tipos de cambio en el sistema monetario internacional?
*8701* **Agustín Maravall**: The use of ARIMA models in unobserved components estimation: an application to spanish monetary control.

*8702* **Agustín Maravall**: Descomposición de series temporales: especificación, estimación e inferencia (Con una aplicación a la oferta monetaria en España).
*8703* **José Viñals y Lorenzo Domingo**: La peseta y el sistema monetario europeo: un modelo de tipo de cambio peseta-marco.
*8704* **Gonzalo Gil**: The functions of the Bank of Spain.
*8705* **Agustín Maravall**: Descomposición de series temporales, con una aplicación a la oferta monetaria en España: Comentarios y contestación.
*8706* **P. L'Hotellerie y J. Viñals**: Tendencias del comercio exterior español. Apéndice estadístico.
*8707* **Anindya Banerjee y Juan Dolado**: Tests of the Life Cycle-Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small-Sample Interpretations.
*8708* **Juan J. Dolado y Tim Jenkinson**: Cointegration: A survey of recent developments.
*8709* **Ignacio Mauleón**: La demanda de dinero reconsiderada.
*8801* **Agustín Maravall**: Two papers on arima signal extraction.
*8802* **Juan José Camio y José Rodríguez de Pablo**: El consumo de alimentos no elaborados en España: Análisis de la información de Mercasa.
*8803* **Agustín Maravall y Daniel Peña**: Missing observations in time series and the «dual» autocorrelation function.
*8804* **José Viñals**: El Sistema Monetario Europeo. España y la política macroeconómica. (Publicada una edición en inglés con el mismo número).
*8805* **Antoni Espasa**: Métodos cuantitativos y análisis de la coyuntura económica.
*8806* **Antoni Espasa**: El perfil de crecimiento de un fenómeno económico.
*8807* **Pablo Martín Aceña**: Una estimación de los principales agregados monetarios en España: 1940-1962.
*8808* **Rafael Repullo**: Los efectos económicos de los coeficientes bancarios: un análisis teórico.
*8901* **M.ª de los Llanos Matea Rosa**: Funciones de transferencia simultáneas del índice de precios al consumo de bienes elaborados no energéticos.
*8902* **Juan J. Dolado**: Cointegración: una panorámica.

(1) Los Documentos de Trabajo anteriores a 1985 figuran en el catálogo de publicaciones del Banco de España.

**Información**: Banco de España
Sección de Publicaciones. Negociado de Distribución y Gestión
Teléfono: 446 90 55, ext. 2180
Alcalá, 50. 28014 Madrid