Cointegración y regresiones espúreas

Profesora: Graciela Sanroman
Econometría II, FCCEEA-UDELAR

2016

Objetivo:
Analizar las propiedades de los estimadores en un modelo de regresión lineal que incluye variables no estacionarias.
Consideremos por ejemplo el caso de

$$\begin{aligned} y_t &= y_{t-1} + e_t \text{ con } e_t \sim iid(0, \sigma_e^2) \text{ y } cov(e_t, e_s) = 0 \ \forall t \neq s \\ x_t &= x_{t-1} + u_t \text{ con } u_t \sim iid(0, \sigma_u^2) \text{ y } cov(u_t, u_s) = 0 \forall t \neq s \end{aligned}$$

$$\text{Suponemos } y_0 = 0 \ \text{ y } x_0 = 0$$

Si $cov(e_t, u_s) = 0 \ \forall t, s$ los procesos y_t y x_t serán independientes, por lo cuál al estimar β por MCO una regresión

$$y_t = \alpha + \beta x_t + \varepsilon_t \ \ t = 1, 2, ...T$$

esperaríamos que sólo se rechazara la hipótesis $\beta = 0$ en el 5% de los casos (si se usa un nivel de significación del 5%).

- Sin embargo, en la práctica observamos que rechazaremos $\beta = 0$ más frecuentemente y que la tasa de rechazo aumentará a medida que T aumente. Entonces dicho contraste será:
  - sesgado
  - inconsistente.
- Esto se da así debido al carácter no estacionario de las variables, el cual hace que los supuestos en los que se basa el modelo de regresión multiple no sean válidos y por lo tanto la distribución de probabilidad de los estimadores de los coeficientes y de los estadísticos se ve afectada.

Ejemplo

Simulamos datos que cumplen con:

$y_t = y_{t-1} + e_t$ con $e_t \sim iid(0,\, 0.5)$ y $cov(e_t, e_s) = 0\ \forall t \neq s$

$x_t = x_{t-1} + u_t$ con $u_t \sim iid(0,\, 1)$ y $cov(u_t, u_s) = 0\ \forall t \neq s$

$y_0 = 0$ y $x_0 = 0$

$cov(e_t, u_s) = 0\ \forall t, s$

Estimamos

$$y_t = \alpha + \beta x_t + \varepsilon_t \ \ t = 1, 2, \ldots T$$

Extraemos 100 muestras de distinto tamaño (T) y observamos la cantidad de veces que rechazamos la hipótesis $Ho:\ \beta = 0$ (esperamos rechazar 5 veces)

Tamaño muestra	Cantidad de rechazos en 100 muestras
10	35
20	46
30	54
50	65
100	80
150	83

REGRESION ESPURIA

- Entonces, estamos encontrando un efecto significativo de x en y que en realidad no existe.
- Davidson y MacKinnon (1993): con *e* y *u* independientes e identicamente distribuidos:
  - 10000 muestras:
  - tamaño T=50; se rechaza en torno a 66.2% de las veces
  - tamaño T=250 se rechaza en torno a 84.7% de las veces
- Engel y Newbold (1974) demostraron que aun cuando y y x sean independientes, si ambas son paseos aleatorios una regresión de *y* sobre *x* proporcionarà estadìsticos *t* significativos en un porcentaje muy elevado de los casos y llamaron a esta situación **REGRESION ESPURIA**.

Recordar: en el contexto de datos de corte transversal se denomina regresión espuria al caso en el cual la significación del coeficiente asociado a una variable (digamos $x1$) se origina en un problema de omisión de variable relevante (digamos que $x2$ está omitida) cuando $x1$ y $x2$ están correlacionadas.

Retomemos:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- Para que el estadìstico t de $\widehat{\beta}_1$tenga una distribución aproximada normal estandar en muestras grandes u_t como mínimo debe ser un proceso de media cero y sin correlaciòn serial.
- Bajo

  $$Ho: \ \beta_1 = 0 \qquad\qquad y_t = \beta_0 + u_t$$

  y si suponemos que y_t es un paseo aleatorio que empieza en $y_0 = 0$ entonces la ecuación anterior serà valida sòlo si $\beta_0 = 0$, y lo que es más importante

  $$u_t = y_t = \sum_{j=1}^{t} e_t$$

  es decir u_t es un paseo aleatorio bajo Ho, lo cuál viola los supuestos de Gauss-Markov

Además, si y_t o x_t son paseos aleatorios con deriva el problema de regresiòn espuria se agrava.

Cuando regresamos una variable I(1) sobre otra variable I(1) el estadìstico t :

- no tiene una distribución asintótica normal estándar,
- tiende a infinito cuando T tiende a infinito.

Adicionalmente el comportamiento del R-cuadrado no es el habitual que se encuentra en datos de corte transversal o con series temporales I(0). En estos casos:

$$p \lim R^2 = 1 - \frac{\sigma_u^2}{\sigma_y^2}$$

pero no ocurre lo mismo con regresiones espurias con variables I(1): el límite en probabilidad del R^2 no està bien definido
IMPLICACION: es muy probable que el R^2 sea muy alto si y y x son procesos de series temporales incluso cuando ambas variables sean estocásticamente independientes.

Algunas propiedades de combinaciones lineales de distintos tipos de series temporales:

$$\begin{aligned}
y_t &\sim I(0) \Rightarrow (a + by_t) \sim I(0) \\
y_t &\sim I(1) \Rightarrow (a + by_t) \sim I(1) \\
y_t &\sim I(0) \text{ y } x_t \sim I(0) \Rightarrow (ax_t + by_t) \sim I(0) \\
y_t &\sim I(0) \text{ y } x_t \sim I(1) \Rightarrow (ax_t + by_t) \sim I(1) \\
y_t &\sim I(1) \text{ y } x_t \sim I(1) \Rightarrow (ax_t + by_t) \sim I(1) \text{ EN GENERAL} \\
a, b &\in R
\end{aligned}$$

ENTONCES: SI y es I(1) y alguno de los regresores es también I(1) es muy probable que el modelo resulte significativo utilizando las pautas tradicionales de inferencia y ello es un resultado erróneo: REGRESION ESPURIA.

Una alternativa es diferenciar el modelo:

$$\Delta y_t = \beta_1 \Delta x_t + \Delta u_t$$

Limitaciones en la estimación del modelo en primeras diferencias:

- el estimador es menos eficiente ya que
  - Δx_t tiene menos variabilidad que x_t
  - Δu_t tiene más variabilidad que u_t (de hecho Δu_t es un proceso MA(1) no invertible)
- Podría ser muy útil relacionar los niveles de y y x cuando entre y y x hay una relaciòn de equilibrio en el largo plazo.

Algunos ejemplos de relaciones de equilibrio en el largo plazo:

- tasas de interés reales a 30 y 90 dìas
- tasas de interés nominales e inflación
- salarios nominales e inflación
- productividad del trabajo y salarios reales
- consumo y pbi
- precios domésticos y precios internacionales de bienes transables (paridad de los poderes de compra)

Asociados al concepto de relaciones económicas de largo plazo surge en econometría el concepto de: **COINTEGRACION**

COINTEGRACION

- El concepto de cointegración lo formalizaron Engel y Granger (1987).
- Consideremos dos variables I(1)

$$\begin{aligned} y_t &= y_{t-1} + e_t \text{ con } e_t \sim iid(0, \sigma_e^2) \text{ y } cov(e_t, e_s) = 0 \ \forall t \neq s \\ x_t &= x_{t-1} + u_t \text{ con } u_t \sim iid(0, \sigma_u^2) \text{ y } cov(u_t, u_s) = 0 \forall t \neq s \end{aligned}$$

Consideremos una combinación lineal

$$n_t = \alpha_1 y_t + \alpha_2 x_t$$

n_t será *en general* $I(1)$. Sin embargo, es posible que para algún par de valores particulares (α_1, α_2) la variable n_t sea I(0)

- Si se cumple que existe un par de valores (α_1, α_2) que hacen que n_t sea I(0) decimos que y y x **están cointegradas**

COINTEGRACION

- Notar
  - n_t no será I(0) en general, es decir para cualquier par $(\alpha_1^{'}, \alpha_2^{'}) \neq (\alpha_1, \alpha_2)$
  - pero si $n_t = \alpha_1 y_t + \alpha_2 x_t$ es $I(0)$ para algún par de valores particulares (α_1, α_2) también será válido que $cn_t = c\alpha_1 y_t + c\alpha_2 x_t$ es $I(0)$ para cualquier constante c, es decir también será válido para cualquier par $(c\alpha_1, c\alpha_2)$
- Entonces, sin perder generalidad se fija el coeficiente de la y en el valor 1 (corresponde a $c = \frac{1}{\alpha_1}$) y denominamos β al coeficiente de x_t (que corresponde a $\frac{\alpha_2}{\alpha_1}$) y $\varepsilon_t = \frac{n_t}{\alpha_1}$ escribimos:
$$\varepsilon_t = y_t - \beta x_t$$
- Si se cumple que ε_t es I(0) decimos que y y x están cointegradas. Recordar que si ε_t es I(0) entonces:
  - media constante
  - varianza constante
  - covarianzas $\text{cov}(\varepsilon_t, \varepsilon_{t-j})$ que dependen sólo de j (la cantidad de períodos entre las ε)

COINTEGRACION

Ejemplo: tasas de interés reales a 30 y 90 dìas

- las series unitarias tienden a ser procesos integrados,
- la diferencia entre ambas esperamos que sea estacionaria.

En este ejemplo la cointegraciòn tiene sustento en la teoría económica: **arbitraje**
Las tasas de interés tienen una relaciòn a largo plazo:

$$t90_t = t30_t + \mu + e_t$$

en el corto plazo pueden haber desvìos respecto al equilibrio de largo plazo, pero esos desequilibrios tienden a desaparecer.
En este ejemplo además, esperamos un valor muy especifico para β, esperamos $\beta = 1$.
Esto es muy importante, para proceder a la estimación debemos distinguir las situaciones en donde β es conocido de aquellas en las cuales no se conoce.

Objetivo:

Someter a prueba la hipótesis de existencia de una relación de cointegración entre dos variables y y x, es necesario distinguir:

- Estimación de relaciones de cointegración cuando β es conocido

- Estimación de relaciones de cointegración cuando β es desconocido

Cuando β es conocido la prueba de cointegraciòn es como sigue:

- calculo $\varepsilon_t = y_t - \beta x_t$ (en el caso de tasas de interés a 90 y 30 días $\beta = 1$ entonces calculo $y_t - x_t$)
- realizo una prueba de raiz unitaria tradicional a ε_t resultante

Cuando β es conocido (continuación)
Notar que en este caso esperamos

$$\varepsilon_t = \mu + v_t$$

- por lo cual para hacer la prueba de raíz unitaria formularía el modelo:

$$\Delta\varepsilon_t = \mu + \theta\varepsilon_t + \sum_{j=1}^{p} \gamma_j \Delta\varepsilon_{t-j} + v_t$$

- Los valores críticos a utilizar corresponden al caso de los vistos en los contrastes ADF de una serie con deriva, en este caso (para muestras grandes, los valores críticos son):

Significación	1%	5%	10%
valor critico	-3.43	-2.86	-2.57

Si el valor del estadístico es menor que -2.86 (mayor que 2.86 en valor absoluto) rechazo (al nivel de significación del 5%) la hipótesis de raiz unitaria, si es mayor no la rechazo.

- SI NO RECHAZO RAIZ UNITARIA: LA EVIDENCIA SUGIERE QUE LAS SERIES NO ESTAN COINTEGRADAS
- SI RECHAZO RAIZ UNITARIA (ε ES $I(O)$) TENGO EVIDENCIA DE QUE HAY COINTEGRACION

Cuando β es desconocido
Es más dificil contrastar la existencia de una relación de cointegración, cuando no conocemos el verdadero valor de β (ejemplo ingreso-consumo, salario-inflación)
Si y y x estàn cointegradas, entonces el estimador MCO de β en

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

es consistente, el problema es que bajo la hipótesis nula se supone que las variables no están cointegradas lo que implica que bajo H_0 estamos ante una regresión espuria.

Cuando β es desconocido (continuación)

Pero esto tiene una solución, el contraste para verificar la presencia de cointegración cuando β es desconocida se realiza como sigue

- Se obtiene $\widehat{\beta}$ en una regresión MCO de y sobre x
- Se calcula $\widehat{u}_t = y_t - \widehat{\alpha} - \widehat{\beta} x_t$
- SE REALIZA UNA PRUEBA DE RAIZ UNITARIA A $\widehat{u}_t$ incluyendo constante pero utilizando una tabla especial (Davidson y MacKinnon, 1993) ya que la distribución Dickey-Fuller no es válida en este contexto debido a que para la serie $\widehat{u}_t$ surge de estimaciones $\widehat{\alpha} - \widehat{\beta}$. Los valores críticos a utlizar en muestras grandes es:

Significación	1%	5%	10%
valor crítico	-3.90	-3.34	-3.04

Cuando β es desconocido (continuación)

Si el valor del estadístico es menor que -3.34 rechazo la hipótesis de raiz unitaria (al 5% de significación), si es mayor no la rechazo

- SI NO RECHAZO RAIZ UNITARIA: LA EVIDENCIA SUGIERE QUE LAS SERIES NO ESTAN COINTEGRADAS
- SI RECHAZO RAIZ UNITARIA (u ES $I(O)$) TENGO EVIDENCIA DE QUE HAY COINTEGRACIÒN

Retomemos el ejemplo con variables simuladas. Simulamos datos que cumplen con:

$y_t = y_{t-1} + e_t$ con $e_t \sim iid(0, 0.5)$ y $cov(e_t, e_s) = 0 \ \forall t \neq s$

$x_t = x_{t-1} + u_t$ con $u_t \sim iid(0, 1)$ y $cov(u_t, u_s) = 0 \ \forall t \neq s$

$y_0 = 0$ y $x_0 = 0$

$cov(e_t, u_s) = 0 \ \forall t, s$

Estimamos $y_t = \alpha + \beta x_t + \varepsilon_t \ \ t = 1, 2, \ldots T$

Extraemos 100 muestras de distinto tamaño (T) y observamos la cantidad de veces que rechazamos la hipótesis $Ho: \ \beta = 0$, también observamos la cantidad de rechazos de la hipótesis de raíz unitaria de la serie $\widehat{u}_t = y_t - \widehat{\alpha} - \widehat{\beta} x_t$

		Significación 5%
Tamaño muestra	Prueba t	Contraste raíz unitaria residuos
10	35	17
20	46	14
30	54	11
50	65	7
100	80	4
150	83	4

Si y y x no están cointegradas no podré estimar la relación entre los niveles de estas variables pero aun puedo intentar encontrar relaciòn entre estas variables utilizando las primeras diferencias pero debemos tener claro de que se tratan las relaciones (no incluye relación entre los niveles de estas variables).

$$\Delta y_t = \beta_1 \Delta x_t + \Delta u_t$$

VARIABLES CON DERIVA

- En el análisis anterior no hemos considerado el caso en el que y y x tienen deriva, este supuesto es razonable por ejemplo para las series de tasas de interes pero no para otras (producto, precios) si y y x tienen deriva la esperanza de $E(y_t)$ y $E(x_t)$ son funciones del tiempo (generalmente creciente) y el análisis de cointegración debe tener esto en cuenta.
- La definición estricta de cointegración implica que $y_t - \theta x_t$ sea I(0) sin tendencia.

Consideremos

$$\begin{aligned} y_t &= \delta t + g_t \\ x_t &= \lambda t + h_t \end{aligned}$$

con g y h procesos I(1)
$E(\Delta y_t) = E(\delta(t - t + 1) + \Delta g_t) = \delta$ es la deriva de y
$E(\Delta x_t) = E(\lambda(t - t + 1) + \Delta h_t) = \lambda$ es la deriva de x
Si y y x están cointegradas tiene que existir β tal que $y_t - \beta x_t$ sea $I(0)$

$$y_t - \beta x_t = (\delta - \beta\lambda)t + g_t - \beta h_t$$

que es normalmente un proceso estacionario alrededor de una tendencia, pero la definición estricta de cointegración exige que no haya tendencia y ello implicaría $\delta = \beta\lambda$

Entonces en procesos integrados con deriva es posible que las partes estocásticas g_t, h_t estén cointegradas pero que el parametro β de cointegración no resulte en $\delta = \beta\lambda$, es decir no elimine la tendencia temporal lineal.
Podremos contrastar si existe cointegración entre g y h sin detenernos en la parte de tendencia regresando

$$y_t \text{ sobre una constante, una tendencia lineal y } x_t$$

y aplicamos los contrastes DF o ADF habituales a los residuos $\widehat{u}$, pero ahora los valores críticos son diferentes a los anteriormente vistos

Significación	1%	5%	10%
valor crítico	-4.32	-3.78	-3.50

En este caso la existencia de cointegración no excluye la posibilidad de que $y_t - \beta x_t$ tenga tendencia lineal pero no es un proceso I(1).

NOTACION

Decimos que dos series son CI(1,1) si ambas series son I(1) y los residuos de la regresión de una sobre la otra es I(0), es decir si ambas están cointegradas
En general diremos que dos series son CI(d,q) si ambas variables son integradas de orden d y los residuos de la regresión de una sobre la otra es Integrado de orden d-q.

Más de dos variables:
Supongamos que tenemos 3 variables

$$x1, \ x2 \ y \ x3$$

todas ellas $I(1)$
¿es posible que exista una relaciòn de largo plazo estable entre ellas?
Cointegración: se complica ya que podría haber ninguno, 1 o 2 vectores de cointegración (en general con k variables podrían existir hasta $k-1$ relaciones de cointegración.
En el caso de más de 2 variables se utiliza Johansen (1988) pruebas de la traza y el valor propio (no lo veremos).

Resumen del proceso de estimación de relaciones de cointegración en el caso de dos variables

1. Comprobar que el orden de integración de las variables es 1 en todos los casos.
2. Estimar modelo con variables de igual orden de integración.
3. Analizar el orden de integración de los residuos, si los residuos son $I(0)$ estamos frente a variables cointegradas (recordar que si no conocemos el valor del vector de cointegración es necesario estimarlos y en dicho caso tenemos que utilizar valores críticos diferentes a los utilizados en las pruebas ADF).

Si la evidencia señale que estamos en un caso de variables CI(1,1) para realizar inferencia sobre los coeficientes del modelo podemos recurrir al método de MCO dinámicos propuesto por Stock y Watson, consiste en estimar una regresión

$$y_t = \alpha + \beta x_t + \sum_{j=-p}^{j=+p} \gamma_j \Delta X_{t+j} + u_t$$

Estimando los errores estándar HAC (robustos a heterocedasticidad y autocorrelación) los estadísticos t en muestras grandes se comportan bien, es decir tienen distribución normal aproximada.

Ordenes de integración mayores

Consideremos que tenemos un conjunto de $k+1$ variables

$$y, x_1, x_2, x_3, \ldots, x_k$$

Realizar una regresión de y sobre el vector x sólo puede ser interpretable si todas son integradas de igual orden:

- Si todas son I(0) puedo realizar inferencia con los métodos habituales.
- Si son I(d) con d$\geqslant$ 1 debo PRIMERO proceder a analizar si están cointegradas antes de realizar cualquier análisis de signo, significación y magnitud de los parámetros.

En general se dice que un conjunto de variables son CI(d,q) si todas las variables son integradas de orden d y los residuos de la regresión de unas sobre otras es Integrado de orden d-q.

PASOS:

1. Comprobar el orden de integración de las series
2. Estimar modelos con variables de igual orden de integración: analizar el orden de integración de los residuos.
3. Similar a lo anterior pero con más variables

PREDICCIÓN

Determinar el orden de integración de una serie es fundamental a la hora de REALIZAR PREDICCIONES.

PREDICCIONES

Para realizar predicciones basadas en el análisis univariado de las series recurrimos a la modelización ARIMA (o SARIMA si tenemos series mensuales o trimestrales).
Por ejemplo,

- si estamos frente a una serie $y_t \sim \mathrm{I}(0)$, las predicciones se basarán en la modelización de un modelo ARMA para y_t.
- si estamos frente a una serie $y_t \sim \mathrm{I}(1)$, las predicciones se basarán en la modelización de un modelo ARMA para Δy_t.

Es importante destacar que los errores de predicción por utilizar un modelo ARMA incorrecto serán mucho menos graves que cuando la predicción se basa en un modelo en el que el orden de integración utilizado difiere del verdadero orden de integración de la serie

PREDICCIONES

En los casos en los cuáles tengo evidencia que conocer x ayuda en la predicción de y, distinguiré el caso en el cuál y_t y x_t son CI(1,0) del caso en que y_t y x_t son CI(1,1).
Si y_t y x_t son CI(1,0) (es decir I(1) no cointegradas) utilizaré el modelo

$$\Delta y_t = \beta_0 + \sum_{j=1}^{j=p} \delta_j \Delta x_{t-j} + \sum_{j=1}^{j=p} \rho_j \Delta y_{t-j} + u_t$$

Es posible ver esta ecuación como un modelo de retardos distribuidos (sólo que aquí la ecuación es en primeras diferencias en lugar que en niveles) podemos determinar por ejemplo el multiplicador de impacto y el multiplicador de largo plazo.

PREDICCIONES Mecanismo de Corrección del Error

Si y, x son CI(1,1), Para realizar predicciones de medio y largo plazo se utilizará el MODELO DE CORRECCION DEL ERROR, que se sustenta en la existencia de una relación de equilibrio (COINTEGRACIÓN) entre la variable a predecir (y) y una variable que la determina (x).
Si y_t y x_t son CI(1,1) (es decir I(1) cointegradas) utilizaré el modelo

$$\Delta y_t = \beta_0 + \alpha(y_{t-1} - \beta x_{t-1}) + \sum_{j=1}^{j=p} \delta_j \Delta x_{t-j} + \sum_{j=1}^{j=p} \rho_j \Delta y_{t-j} + u_t$$

A este modelo se lo denomina VECTOR DE CORRECCION DEL ERROR O MODELO DE CORRECCION DEL ERROR

Mecanismo de corrección del error

$$\Delta y_t = \beta_0 + \alpha(y_{t-1} - \beta x_{t-1}) + \sum_{j=1}^{j=p} \delta_j \Delta x_{t-j} + \sum_{j=1}^{j=p} \rho_j \Delta y_{t-j} + u_t$$

Al término

$$\alpha(y_{t-1} - \beta x_{t-1})$$

se le denomina TERMINO DE CORRECCION DEL ERROR

- Si $y_t \sim I(1), x_t \sim I(1)$ entonces $\Delta y_t \sim I(0)$, $\Delta x_t \sim I(0)$ por lo cual el lado izquierdo de la ecuación anterior es $I(0)$ y también lo sería el lado derecho si $(y_t - \beta x_t)$ es $I(0)$
- Si y_t, x_t son CI(1,1) con vector de cointegración $(1, \beta)$ entonces $(y_{t-1} - \beta x_{t-1}) \sim I(0)$, por lo que la ecuación puede ser estimada y realizar inferencia como es tradicional.

Mecanismo de corrección del error

¿Por qué agregar el término $\alpha(y_{t-1} - \beta x_{t-1})$ a la ecuación de predicción?

- Si existe una relación de equilibrio de largo plazo entre y_t y x_t, entonces esperamos que $y_{t-1} - \beta x_{t-1}$ sea 0, pero sabemos:
  - que en un momento dado es improbable que ese equilibrio se satisfaga completamente
  - que el desequilibrio existente influenciará las variaciones Δy_t ya que la tendencia será a re-establecer el equilibrio
- Desvíos respecto al equilibrio de largo plazo:
  - si $y_{t-1} > \beta x_{t-1}$ entonces en el período t existirán presiones a la baja sobre y
  - si $y_{t-1} < \beta x_{t-1}$ entonces en el período t existirán presiones a la alza sobre y
- Corrección del error:
  - el modelo indica que en un período t dado se "corregirá" una magnitud α del desvío que presente y en el período $t-1$ respecto a la relación de largo plazo esperada con respecto a x
  - esperamos $\alpha < 0$
- El mecanismo de corrección del error actúa de forma que y retorne a su valor de equilibrio

Estimación del modelo de corrección del error

Aquí también tenemos que distinguir dos casos

- β es conocido
- β es desconocido

Estimación del modelo de corrección del error: β conocido

En este caso la estimación del modelo es muy sencilla:

1. Obtengo $s_{t-1} = y_{t-1} - \beta x_{t-1}$
2. Regreso (MCO) Δy_t sobre $s_{t-1}, \Delta x_{t-1}, ..., \Delta x_{t-p}, \Delta y_{t-1}, ...\Delta y_{t-q}$

Estimación del modelo de corrección del error: β desconocido

En este caso la estimación del modelo es también sencilla pero requiere una etapa adicional:
1. Regreso (MCO) y_t sobre x_t y obtengo $\widehat{\beta}$
2. Obtengo $\widehat{s}_{t-1} = y_{t-1} - \widehat{\beta} x_{t-1}$
3. Regreso Δy_t sobre $\widehat{s}_{t-1}, \Delta x_{t-1}, ..., \Delta x_{t-p}, \Delta y_{t-1}, ...\Delta y_{t-q}$

Este método se denomina procedimiento en dos etapas de Engle y Granger. Como Engle y Granger (1987) demuestran, en este procedimiento podemos hacer caso omiso (asintóticamente) de que $\widehat{\beta}$ ha sido estimado en una regresión preliminar.

Comentarios finales

1. En la práctica empírica se suele utilizar Johansen (1988) para este problema de estimación. No obstante, en el caso de que sólo se consideren dos variables y, x el procedimiento de Engel y Granger tanto para la estimación del vector de cointegración como para la estimación del MCE son válidas y tienen buenas propiedades (en muestras grandes)
2. Las propiedades de los contrastes presencia de relaciones cointegración dependen del tamaño de la muestra (con menos de 60 observaciones los resultados no son confiables).
3. Al igual que en los contrastes de raíz unitaria, la potencial existencia de cambios estructurales complejiza el problema de decidir si existe o no una relación de cointegración entre las variables.

Bibliografía:
Wooldridge: Introducción a la econometría: un enfoque moderno.
Stock y Watson: Introduction to econometrics

Bibliografía complementaria:
Enders: Applied Econometrics Time Series
Hamilton: Time Series Analysis