BANCO CENTRAL DE COSTA RICA

NOTA TÉCNICA
N.º 001 | 2003

Aspectos teóricos sobre algunos temas econométricos

Ana Cecilia Kikut Valverde
Jorge León Murillo
Mario Rojas Sánchez
Carlos Torres Gutiérrez
Bernal Laverde Molina
Evelyn Muñoz Salas
Álvaro Solera Ramírez

Fotografía de portada: “Presentes”, conjunto escultórico en bronce, año 1983, del artista costarricense Fernando Calvo Sánchez. Colección del Banco Central de Costa Rica.

Aspectos teóricos sobre algunos temas econométricos

Ana Cecilia Kikut Valverde[*], Bernal Laverde Molina[†], Jorge León Murillo[‡], Evelyn Muñoz Salas[§], Mario Rojas Sánchez[**], Álvaro Solera Ramírez[††], Carlos Torres Gutiérrez[‡‡]

Resumen

Este documento constituye una recopilación de las presentaciones realizadas durante el taller para el uso del paquete econométrico EViews, impartido por funcionarios del Departamento de Investigaciones Económicas de la División Económica a funcionarios de este mismo departamento, del 19 de mayo al 6 de junio del 2003.

Se estimó importante hacer tal recopilación porque se considera que es un material valioso de consulta para los mismos asistentes a la capacitación y para otras personas interesadas en temas econométricos. Lo anterior con el fin de mejorar los conocimientos en esta área y perfeccionar paulatinamente la aplicación de las herramientas disponibles.

El documento consiste en un conjunto de diapositivas que tratan algunos de los problemas más comunes del método de los mínimos cuadrados ordinarios como son autocorrelación, multicolinealidad, y heterocedasticidad y temas especiales como integración, cointegración, modelos de corrección de errores, cointegración según el enfoque de Johansen, vectores autorregresivos y el filtro de Kalman.

Este campo queda sujeto para ser ampliado con temas más avanzados en un futuro cercano. Por último, es importante mencionar que este material es complemento de otros documentos que se han publicado como parte del curso, tales como Compendio de programas para EViews: I parte (DIE-040-2003-IT) y Principales indicadores para el diagnóstico del análisis de regresión lineal (DIE-026-2003-IT).

Palabras clave: Modelación económica, Mínimos cuadrados ordinarios, Regresión lineal.

Clasificación JEL: C10.

[*] Departamento de Investigación Económica. División Económica, BCCR. kikutva@bccr.fi.cr
[†] Departamento de Investigación Económica. División Económica, BCCR. laverdema@bccr.fi.cr
[‡] Departamento de Investigación Económica. División Económica, BCCR. leonmj@bccr.fi.cr
[§] Departamento de Investigación Económica. División Económica, BCCR. munozse@bccr.fi.cr
[**] Departamento de Investigación Económica. División Económica, BCCR. rojassm@bccr.fi.cr
[††] Departamento de Investigación Económica. División Económica, BCCR. solerara@bccr.fi.cr
[‡‡] Departamento de Investigación Económica. División Económica, BCCR. torresgc@bccr.fi.cr

Theoretical Issues on Some Econometric Topics

Ana Cecilia Kikut Valverde[§§], Bernal Laverde Molina[***], Jorge León Murillo[†††], Evelyn Muñoz Salas[‡‡‡], Mario Rojas Sánchez[§§§], Álvaro Solera Ramírez[****], Carlos Torres Gutiérrez [††††]

Key words: Economic modeling, Ordinary least squares, Linear regression.
JEL codes: C10.

[§§] Department of Economic Research. Email address kikutva@bccr.fi.cr
[***] Department of Economic Research. Email address laverdema@bccr.fi.cr
[†††] Department of Economic Research. Email address leonmj@bccr.fi.cr
[‡‡‡] Department of Economic Research. Email address munozse@bccr.fi.cr
[§§§] Department of Economic Research. Email address rojassm@bccr.fi.cr
[****] Department of Economic Research. Email address solerara@bccr.fi.cr
[††††] Department of Economic Research. Email address torresgc@bccr.fi.cr

TABLA DE CONTENIDO

TALLER PARA EL USO DE E VIEWS

Departamento de Investigaciones Económicas

AUTOCORRELACIÓN Y MULTICOLINEALIDAD EN LOS MODELOS DE REGRESIÓN

DEPARTAMENTO INVESTIGACIONES ECONÓMICAS

Mayo, 2003

I. AUTOCORRELACIÓN

NOTA: Cargar archivo de trabajo:
c:\tallereviews\auto.wf1

DIE 2

I. AUTOCORRELACION

- NATURALEZA DEL PROBLEMA
  - Violación del supuesto de regresión lineal clásico de errores no autocorrelacionados:
    - E(ui uj)=0 para todo i diferente de j
- CONSECUENCIAS
  - β_{MCO} lineales, insesgados y consistentes, pero ineficientes (no tendrán varianza mínima)
  - →R^2 sobreestimado
  - → pruebas t y F pierden validez

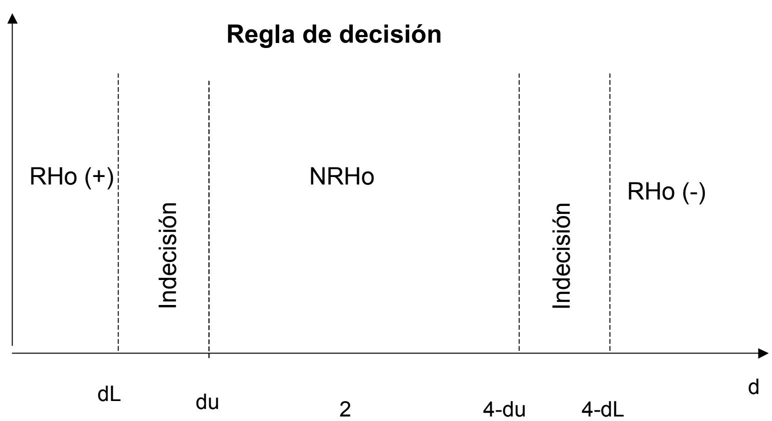
DIE 3

I. AUTOCORRELACION

- DETECCIÓN (7 pruebas)
- **1. Gráfico de los residuos de regresión**
    - *resid* vs tiempo (doble clic sobre *resid*, ***View/Graph/line***)
    - $resid_t$ vs $resid_{t-1}$ (doble clic sobre *resid*, ***Quick/Graph/Scatter***; agregar *resid(-1)* en la caja de diálogo)
  - Para identificar patrones de autocorrelación, comparar con gráficos estilizados de Gujarati (pág. 395):
- **2. Durbin Watson (DW)**
  - Válido para autocorrelación serial de 1° orden en los residuos
  - No aplica para modelos con variable dependiente rezagada (VDR) como variable explicativa
  - Ho: $\rho=0$ (ausencia autocorrelación: $u_t = \rho u_{t-1} + e_t$)
  - H1: $\rho<0$ (aut. –) ó $\rho>0$ (aut. +)
  - Valor crítico (automático en EViews)
  - Valor tabular (tabla DW de Gujarati, pág 800)

I. AUTOCORRELACION

Regla de decisión



I. AUTOCORRELACION

- **3. Durbin h**
  - Válida para modelos con VDR como variable explicativa
  - Detecta autocorrelación serial de 1° orden en los residuos
  - $h=(1-0.5d)\{[n/[1-n(var(\beta_i))]\}^{0.5}$
  - En muestras "grandes", h ~asint N(0,1)
  - EViews no incluye esta prueba (programa JLM)
  - Regla de decisión

    h>1.96 RHo (aut. + de primer orden)

    h<-1.96 RHo (aut. - de primer orden)

    -1.96<h<1.96 NRHo
  - Críticas (*test* General)

Yt en 1, X1, X2, Yt-1 → ut
ut en 1, X1, X2, Yt-1, ut-1 → R^2
$(n-1)R^2$~Chi^2 con 1g° l

DIE 6

I. AUTOCORRELACION

- **4. Correlogramas**
  - Es uno de los mejores *test* cuando "n" es bajo
  - Los correlogramas de la función de autocorrelación simple y parcial se generan en EViews desde la ventana de la ecuación estimada: ***View/Residual Test/Correlogram-Q-statistics***
  - Deben especificarse los rezagos en la caja de diálogo
  - Regla de decisión
    - Si las barras están dentro de las bandas de confianza → NRHo (5% signif.) → no hay autocorrelación
    - Alternativamente, si las probabilidades son <0.05 → RHo (5% signif.) → hay autocorrelación

DIE 7

I. AUTOCORRELACION

- **5. Q de Lung-Box**
  - Prueba más general para autocorrelación tipo ARMA
  - Q~asint Chi^2 con r-p-q grados de libertad, donde:
    - r: g°l de la distribución Chi^2 (arbitrario)
    - p y q: órdenes del proceso de autocorrelación que se quiere probar
  - Ho: ausencia de autocorrelación
  - H1: autocorrelación ARMA(p,q)
  - Se genera en EViews desde la ventana de la ecuación estimada:***View/Residual Test/Correlogram-Q-statistics***
  - Deben especificarse los rezagos en la caja de diálogo (se computa un estadístico Q y su probabilidad asociada para cada rezago).

DIE 8

I. AUTOCORRELACION

- Regla de decisión
  - Probabilidad de cada Q < 0.05 → RHo → hay autocorrelación
  - Este resultado debe concordar con barras del correlograma fuera de las bandas de confianza.

DIE 9

I. AUTOCORRELACION

- **6. Prueba de las "rachas" (*runs tests*)**
  - Def: "racha" es una secuencia ininterrumpida de un mismo signo (+ ó -) en los valores de la serie de los errores.
  - Prueba si el número de "rachas" (*k*) es "alto" o "bajo" respecto a una secuencia aleatoria ("ruido blanco") de *k*.
  - Ho: ausencia de autocorrelación ("rachas" aleatorias)
  - H1: autocorrelación
  - EViews no calcula la prueba (programa OKC)
  - Sean n_1 numero de errores + y n_2 número de errores -
  - Bajo Ho y suponiendo $n_1 > 10$ y $n_2 > 10$, k ~asint. N( ), con:
    - Media: $E(k)=(2n_1n_2/(n_1+n_2))+1$
    - Varianza: $\sigma^2_k=[2n_1n_2(2n_1n_2-n_1-n_2)]/[(n_1+n_2)^2\,(n_1+n_2-1)]$
  - Regla de decisión:
    - $E(k)-1.96\,\sigma^2_k < k < E(k)+1.96\,\sigma^2_k \rightarrow$ NRHo
    - Si *k* está fuera de esos límites RHo (hay autocorrelación)

I. AUTOCORRELACION

- **7. Prueba de Breusch-Godfrey (BG)**
  - Prueba asintótica de Multiplicadores de Lagrange (LM).
  - Válida para modelos con VDR como variable explicativa
  - Se generan en EViews desde la ecuación estimada: ***View/Residual Test/Serial Correlation LM Test***
  - Debe especificarse el orden de la correlación serial a probar en la caja de dialogo.
  - Regla de decisión:
    - El estadístico calculado ***Obs*R-squared*** se compara con el valor tabular de la tabla Chi^2, según los rezagos seleccionados
    - Alternativamente, si la probabilidad del estadístico es <0.05 se RHo $\rightarrow$ hay autocorrelación

I. AUTOCORRELACION

- SOLUCIÓN DEL PROBLEMA
  - ρ conocido: Ej. AR(1) $u_t=\rho u_{t-1}+e_t$

$$
\begin{aligned}
\text{Modelo: } Y_t &= \beta_1 + \beta_2 X_t + u_t \\
Y_{t-1} &= \beta_1 + \beta_2 X_{t-1} + u_{t-1} \\
\rho Y_{t-1} &= \rho\,\beta_1 + \rho\,\beta_2 X_{t-1} + \rho\, u_{t-1} \\
(Y_t-\rho Y_{t-1}) &= (\beta_1-\rho\,\beta_1)+(\beta_2 X_t-\rho\,\beta_2 X_{t-1})+(u_t-\rho\, u_{t-1}) \\
Y_t^* &= \beta_1^* + \beta_2^* X_t^* + e_t
\end{aligned}
$$

Donde: $Y_1(1-\rho^2)^{0.5}$; $X_1(1-\rho^2)^{0.5}$

  - Alternativamente, una vez identificado el proceso AR, se incorpora como una variable explicativa del modelo
  - ρ desconocido: estimar ρ mediante método iterativo de Cochrane-Orcutt, ver Kikut A.C. (1993).

DIE 12

I. AUTOCORRELACION

- **Programa para detectar autocorrelación (auto.prg)**
  - **1. Gráfico de residuos (res3 y res7)**
  - **2. DW (res1)**
  - **3. DH (res5)**
  - **4. Correlogramas (res13)**
  - **5. Q Lung-Box (res13)**
  - **6. Rachas (res8)**
  - **7. Breusch-Godfrey (res9 a res12)**

DIE 13

- II. MULTICOLINEALIDAD

II. MULTICOLINEALIDAD

- NATURALEZA DEL PROBLEMA
  - Relación lineal perfecta o imperfecta entre algunas o todas las variables explicativas del modelo de regresión
  - La multicolinealidad es una característica de la muestra utilizada y no de los errores de estimación

II.1 Multicolinealidad perfecta

- NATURALEZA DEL PROBLEMA
  - Viola el supuesto de ausencia de colinealidad perfecta entre los regresores del modelo
  - La matriz X'X será singular →no se puede calcular $(X'X)^{-1}$ porque su determinate es 0 → no se puede invertir → impide el cálculo de los β_{MCO}
- DETECCION
  - EViews detecta automáticamente el problema:
    - "*near singular matrix*"
- SOLUCIÓN
  - La solución práctica es eliminar la variable explicativa que causa este problema

DIE 16

II.2 Multicolinealidad imperfecta

- NATURALEZA DEL PROBLEMA
  - No viola ningún supuesto del modelo clásico
    - Los β_{MCO} seguirán siendo consistentes e insesgados (y por tanto MELI), pero tendrán asociados "grandes" errores estándar
  - Es un problema grave si el objetivo es estimar ecuaciones de comportamiento (estimación confiable de los β's del modelo)
  - Pero no es necesariamente mala si el objetivo es la predicción
- DETECCION (4 pruebas)
  - EViews no detecta automáticamente el problema, pero se pueden hacer las siguientes pruebas:

DIE 17

II.2 Multicolinealidad imperfecta

- **1. t' s poco significativas y R^2 alto**
  - Con alta colinealidad entre variables, los errores estándar estimados aumentarán mucho, reduciendo el valor de las pruebas *t* → *t*'s poco significativos junto con R^2 elevado
  - Analizar salida estándar de regresión de EViews
  - Pero para distinguir variables colineales de variables superfluas, debe analizarse su capacidad explicativa individual
  - → Se corre una regresión que utilice solo la variable explicativa previamente no significativa . Si ésta continúa siendo no significativa será superflua; en caso contrario será colineal
  - PERO, si se elimina → sesgo de especificación.

DIE 18

II.2 Multicolinealidad imperfecta

- **2. Alta sensibilidad de los parámetros**
  - El valor de los β_{MCO} y sus errores estándar asociados se vuelven extremadamente sensibles al agregar u omitir una observación (parámetros estimados poco robustos)
  - Si se reestima el modelo excluyendo una observación y se notan grandes cambios en el valor y en el signo de los β's estimados, será un indicio de multicolinealidad
  - En EViews se ajusta el tamaño muestral desde la ventana de la ecuación estimada con el botón ***Estimate***

DIE 19

II.2 Multicolinealidad imperfecta

- **3. Matriz de correlación**
  - Se calcula en EViews desde la ventana de la ecuación estimada: ***Procs/Make Regressor Group***, lo cual crea un objeto grupo con todas las variables del modelo (incluida la variable endógena). Luego se pulsa ***View/Correlations/Common Sample***.
  - Regla de decisión: correlaciones cercanas a 1 o -1 indican multicolin.
  - Conviene también calcular el determinante de la matriz: se pulsa ***View/Group Members***; se borra la variable dependiente y se pulsa ***UpdateGroup***. Se asigna un nombre al grupo pulsando ***Name*** y se escribe en la línea de comandos ***sym mcorrel=@cor(<***nombre del grupo***>)*** y se pulsa la tecla *Enter*. Finalmente se calcula el determinante escribiendo en la línea de comandos ***scalar detmcorrel=@det(mcorrel)*** y se pulsa la tecla *Enter*. El valor del determinante se mostrará en la parte inferior de la pantalla (línea de estado) al dar doble clic sobre el objeto *detmcorrel*. Si su valor es cercano a cero habrá indicios de multicolinealidad.

II.2 Multicolinealidad imperfecta

- **4. Factor de inflación de la varianza (FIV)**
  - El FIV muestra como la varianza de un estimador se *infla* por la presencia de la multicolinealidad.
  - A medida que la colinealidad aumenta, la varianza del estimador también aumenta, pudiendo llegar a infinito en el límite.
  - EViews no calcula la prueba (programa de OKC)
  - Regla de decisión:
    - En la práctica, un FIV>10 se considera como indicio de multicolinealidad.
    - Si no hay multicolinealidad el FIV=1.

II.2 Multicolinealidad imperfecta

- SOLUCIÓN
  - Tentación de eliminar la variable colineal, PERO: sesgo de especificación del modelo
  - Datos nuevos o adicionales (ampliar “n”)
  - Transformación de variables (estimación en 1eras diferencias)
  - Información a priori (de trabajos previos o de la Teoría)
  - Otros (combinar corte transversal y series de tiempo; componentes principales; regresión cresta (*ridge*), etc.)

DIE 22

II.2 Multicolinealidad imperfecta

- **Programa para detectar multicolinealidad (mult.prg)**
  - **1. t’s poco significativos y R2 alto (res6)**
  - **2. Alta sensibilidad de los β’s estimados**
  - **3. Matriz de correlación**
  - **4. FIV (mult_fiv)**

DIE 23

III. HETEROCEDASTICIDAD

DIE

1

III.1 Heterocedasticidad: naturaleza

- Un supuesto importante del modelo clásico de regresión lineal es que la varianza de cada término de perturbacion (u_i) es un número constante igual a σ^2. Este es el supuesto de homocedasticidad (igual varianza).
- Cuando no se cumple se supuesto anterior se dice que existe heterocedasticidad, esto es, la varianza de cada término de perturbacion (u_i) no es un número constante ($\sigma_i^{2)}$.
- El problema de heterocedasticidad tiende a ser más común en series de corte transversal que en series de tiempo.

DIE

2

III.2 Heterocedasticidad y la estimación de MCO

- En presencia de heterocesasticidad los estimadores de MCO siguen siendo lineales e insesgados pero dejan de tener la varianza mínima en la clase de estimadores lineales insesgados, es decir, dejan de ser estimadores MELI.

- Lo anterior por cuanto MCO no utiliza la información contenida en la variabilidad desigual de la variable dependiente y asigna una ponderación igual a cada observación. Sin embargo, MCG toma en cuenta esta información y asigna ponderaciones diferentes a cada observación produciendo estimadores MELI.

- MCG es MCO aplicados sobre las variables transformadas que satisfagan los supuestos tradicionales de los mínimos cuadrados.

DIE 3

III.2 Heterocedasticidad y la estimación de MCO

- Cuando se utiliza MCO asumiendo o no heterocedasticidad no se pueden establer intervalos de confianza y pruebas de hipótesis utlizando las pruebas t y F. Lleva a intervalos mayores con lo cual las pruebas t y F producen resultados inexactos, lo que a primera vista parece ser un coeficiente estadísticamente no significativo podría ser significativo.

DIE 4

III.3 Heterocedasticidad : identificación

- No existen reglas fijas y seguras para detectarla, sino solamente unas cuantas normas muy generales.
- Por la naturaleza del problema, método gráfico, prueba de Park, prueba de Glejser, prueba de Goldfeld-Quandt y prueba de White.
- La mayoría de métodos están basados en el exámen de los residuos e_i de MCO, puesto que éstos son los observables y no las perturbaciones u_i de la población.
- El test general de White consiste en el estadístico $n*R^2$, donde R^2 viene de regresar e_i^2 sobre una constante y todas las variables explicativas. El estadístico se distribuye asintóticamente como una chi-cuadrado con P-1 grados de libertad, donde P es el número de regresores en la regresión, no incluyendo la constante.

DIE 5

III.4 Heterocedasticidad : ARGH y GARCH

- En análisis de datos macroeconómicos Engle (1982,1983) y Cragg (1982) encontraron evidencia de que para algunos fenómenos la varianza de los disturbios en modelos de series de tiempo son menos estables de lo que usualmente se asume.
- Los resultados de Engle sugieren que en el análisis de modelos de inflación, los grandes y pequeños errores de predicción parecen ocurrir por grupos, sugiriendo una forma de heterocedasticidad en la que la varianza del error de proyección depende del tamaño de la perturbación precedene.
- Engle sugiere la heterocedasticidad condicional autorregresiva o modelos ARCH como una alternativa para los procesos usuales de series de tiempo.

DIE 6

III.4 Heterocedasticidad : ARGH y GARCH

- Las autocorrelaciones de los residuos al cuadrado proveen evidencia de comportamientos ARCH. Un test LM de ARCH(q) contra la hipótesis de efectos no ARCH puede calcularse a través de T veces el R^2 en la regresión de los e_t^2 sobre una constante y q valores rezagados de los e_t^2. Este estadístico se distribuye como una chi-cuadrado con q grados de libertad. Valores mayores que el valor crítico tabular dan evidencia de presencia de efectos ARCH o GARCH.

DIE 7

III.4 Heterocedasticidad : ARGH y GARCH

- Un proceso ARCH se traduce en que la varianza del error cambia en el tiempo, dependiendo en cada periodo de la magnitud de los errores en periodos anteriores, medidos como el cuadrado del error.

- El proceso GARCH generaliza el proceso, permitiendo que la varianza dependa del rezago de la propia varianza.

DIE 8

III.4 Heterocedasticidad : ARGH y GARCH

- La forma genérica de un modelo GARCH para la varianza del error se puede escribir así:

$$y_t = X_t \beta + \varepsilon_t$$

$$\varepsilon_t \sim (0, \sigma^2_t)$$

$$\sigma^2_t = \omega + \gamma(L)\, \varepsilon^2_t + \lambda(L)\, \sigma^2_t$$

DIE 9

III.5 Heterocedasticidad : medidas remediales

- El método más directo es utilizar mínimos cuadrados ponderados cuando se conoce la varianza heterocedástica.
- Sin embargo, es muy raro que se conozcan los σ^2_i por lo que generalmente se recomienda hacer algunos supuestos sobre la naturaleza de los σ^2_i , para transformar los datos y hacer que las perturbaciones de los datos transformados sean homocedásticas.
- La transformación logarítmica de los datos frecuentemente reduce el problema de heterocedasticidad, dado que comprime las escalas en que se miden las variables. Tiene la ventaja de que la pendiente de la regresión se puede interpretar como elasticidad, esto hace que los modelos logarítmicos sean tan populares en econometría empírica.

DIE 10

Análisis de integración

- Econometría tradicional supone variables estacionarias
- Es decir que sus valores fluctúan aleatoriamente alrededor de una media constante y tienen variancia finita y constante.
- En una serie no estacionaria, cualquier choque que se presente en la serie tendrá carácter permanente.

DIE TALLER PARA EL USO DE EVIEWS 2

Análisis de integración

- Pruebas convencionales basadas en los estadísticos F y t pueden indicar la presencia de una relación entre variables cuando ésta en realidad no existe.
- Los estimadores de mínimos cuadrados no son consistentes.
- Existe la posibilidad de caer en el problema de regresiones espúreas.

DIE TALLER PARA EL USO DE EVIEWS 3

Análisis de integración

- Se ha recurrido a la diferenciación de las variables no estacionarias para convertirlas en estacionarias, lo cual ocasiona una pérdida de información en cuanto a los grados de libertad y al comportamiento de largo plazo de la serie.

DIE TALLER PARA EL USO DE EVIEWS 4

Análisis de integración

- Surge un desarrollo orientado al manejo de series no estacionarias:
  - Raíces unitarias
  - Vectores autorregresivos
  - Cointegración
  - Corrección de errores
  - Vectores de Johansen

DIE TALLER PARA EL USO DE EVIEWS 5

Análisis de cointegración

- La cointegración es una metodología del campo de la econometría dinámica en la que se combina el análisis de series de tiempo con el análisis econométrico tradicional.

Estadísticamente...

- La teoría de la cointegración dice que si las series consideradas en un modelo de regresión ***cointegran***, la inferencia y estimación son válidas aún cuando las variables que lo conforman no son estacionarias.

Económicamente….

- La interrelación existente entre algunas variables económicas, provoca que éstas no tiendan a disociarse mucho entre sí en el largo plazo, es decir existe entre ellas una relación de largo plazo y que ésta es estable.

¿Por qué?

- Aunque dos series se muevan cada una sin una tendencia a volver al promedio, existe una combinación lineal de ellas que si lo hace.

Análisis de integración

- Para poder efectuar un análisis de cointegración, es necesario analizar previamente el ***orden de integración*** de las series, puesto que se requiere que tengan el mismo orden de integración.

Análisis de Integración

- El orden de integración de una serie se refiere al número de veces que ésta debe ser diferenciada para obtener una serie estacionaria.
- Una variable es integrada de orden d, I(d), cuando se obtiene una variable estacionaria después de ‘d’ diferenciaciones.

Análisis de integración

- Algebraicamente una serie estacionaria, I(0), se escribiría de esta forma:

$$X_t = \alpha + \mu_t$$

- No depende de sus valores pasados, solamente de un término aleatorio “U”
- Todo choque es transitorio y tiende a desaparecer en el tiempo.

Análisis de integración

- En Eviews podemos generar artificialmente una serie I(0):

  En línea de comandos:

  Genr I_cero = constante + nrnd

  Si la constante es 0 tenemos una serie ruido blanco

  VIEW/GRAPH/LINE

Análisis de integración

- Una serie I(1) se generará por alguno de los siguientes procesos:

$$X_t = X_{t-1} + \mu_t \qquad SCST$$

$$X_t = \alpha + X_{t-1} + \mu_t \qquad CCST$$

$$X_t = \alpha + X_{t-1} + \beta T + \mu_t \qquad CCCT$$

$$X_t = X_{T-1} + \beta T + \mu_t \qquad SCCT$$

Análisis de integración

- De igual forma en Eviews podemos generar artificialmente una serie I(1):

  En línea de comandos:

  Genr X = constante + X(-1) + nrnd

  VIEW/GRAPH/LINE

Análisis de integración

- La forma de realizar un análisis de grado de integración en una variables es a través de las "Pruebas de Raíz Unitaria", como la prueba de Dickey-Fuller aumentada.

Análisis de integración

- La prueba parte de que la serie de interés es generada por un camino aleatorio:

$$X_t = \alpha + \rho X_{t-1} + e_t$$

- Si $\rho = 1 \Rightarrow I(1)$
- Entonces $H_0 : \rho = 1$
  $H_1 : \rho \langle 1$

Análisis de integración

- Para que esta serie sea I(1), su primera diferencia debe ser I(0)

$$X_t - X_{t-1} = \alpha - (1-\rho)X_{t-1} + \varepsilon_t$$

$$\Delta X_t = \alpha + \gamma\ X_{t-1} + e_t$$

- Efectuamos la prueba: $H_0 : \rho = 1$

$$H_0 : \gamma = 0$$

Procedimiento en E Views

- Se ajusta la ecuación en diferencias por MCO y se analiza la significancia del coeficiente para Xt-1.
- Esta prueba requiere el uso de tablas especiales, como la de MacKinnon.

Procedimiento en E Views

- El procedimiento a seguir se encuentra en la hoja adjunta.
- Para efectos prácticos debe tenerse presente que para el análisis tanto en niveles como en primeras o segundas diferencias:

Ho= la serie no es estacionaria.

Técnicas de Cointegración

- Un conjunto de variables está cointegrado cuando una combinación lineal de ellas es estacionaria I(0).
- Una relación de cointegración puede ser interpretada como una relación de equilibrio de largo plazo, ya que es posible que las variables cointegradas se puedan desviar de esta relación en el corto plazo, pero que su asociación se reanude en el largo plazo (estado estacionario).

Modelos de corrección de errores

- Trabajar con variables I(1) en primeras diferencias, I(0), es válido estadísticamente cuando se construyen modelos univariables (ARMA); pero no cuando es importante analizar si existe alguna relación entre variables.
- Los modelos en primeras diferencias no tienen una solución de largo plazo, debido a que en el estado estacionario: $\Delta y=0$ y $\Delta x=0$, dado que $y_t=y_{t-i}$ y $x_t=x_{t-i}$, y por tanto, no se puede saber si existe una relación de equilibrio entre x e y.

Modelos de corrección de errores

- Existe una clase de modelos que supera este problema, en los cuales se combinan primeras diferencias y rezagos de variables cointegradas con información del equilibrio de largo plazo representado en el término de error rezagado.
- Estos modelos son conocidos como "Modelos de corrección de errores o de corrección de equilibrios" (MCE).
- En estos casos es válido aplicar MCO para estimar las funciones de largo y corto plazo.

Modelos de corrección de errores

- Interpretación de MCE:
  - La variable *y* puede cambiar entre *t-1* y *t*, en respuesta a variaciones en la(s) variable(s) explicativa(s), *x*, entre *t-1* y *t*, y también para corregir cualquier desequilibrio que exista en el periodo previo.
- El coeficiente asociado al término de error rezagado indica la velocidad del ajuste hacia el equilibrio y se interpreta como la proporción del desequilibrio del periodo *t-1* que se corrige en *t*.

Métodos de estimación MCE

- Existen al menos tres métodos:
  - Engle - Granger (método 2 etapas)
  - Engle – Yoo (método 3 etapas)
  - Johansen (sistema de cointegración)
- Los dos primeros se utilizan para estimar un vector de cointegración entre dos variables (y,x).
- El tercero se utiliza para estimar múltiples vectores de cointegración (r), en modelos de más de dos variables (g>2). En estos casos r<g.

Método EG de 2 etapas

- Etapa 1:
  - Verificar que las variables (y, x) son I(1).
  - Estimar con MCO el modelo con los niveles de las variables I(1).
  - Salve los residuos (***resid***) con otro nombre y realice el test de raíz unitaria.
  - Si los residuos son I(0), entonces (y, x) cointegran, es decir, tienen una relación de largo plazo.

Método EG de 2 etapas

- Etapa 2:
  - Utilice los residuos de la primera etapa como una variable en el MCE.

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 (y_{t-1} - \hat{\varphi} x_{t-1}) + \varepsilon_t$$

  - Donde: $[1, \hat{\varphi}]$, es el vector de cointegración.

  - Estime con MCO el MCE.
  - Realice la inferencia estadística a los parámetros.

Método EG de 2 etapas

- MG tiene algunos inconvenientes:
  - Los test de raíz unitaria y cointegración pierden validez en muestras pequeñas.
  - Puede presentarse el sesgo de ecuaciones simultáneas si la causalidad entre las variables (y, x) va en ambas direcciones.
  - Si se comete un error de especificación el modelo de E1, este se traslada al test de cointegración en E2.
  - No es posible realizar pruebas de hipótesis al VC de E1.

Método EY de 3 etapas

- Este método permite realizar inferencia estadísticas sobre el vector de cointegración (VC) estimado en E1.
- Parte de E1 y E2 del método EG y adiciona un E3, en el que se actualiza la estimación del vector de cointegración y su error estándar.
- E3 es un paso algebraicamente técnico y adicional, pero no logra superar las otras desventajas del EG.

Método de Johansen

- Este método está basado en modelos VAR.
- Suponiendo que hay un grupo de g variables (g>2) que son I(1) y que podrían estar cointegradas. Un VAR con k rezagos con tales variables sería:

$$y_t = \beta_1 y_{t-1} + \ldots + \beta_k y_{t-k} + \xi_t$$

Método de Johansen

- Para usar el método de Johansen el modelo anterior debe convertirse en un VECM de la siguiente forma:

$$\Delta y_t = \Pi y_{t-k} + \Gamma_1 \Delta y_{t-1} \ldots + \Gamma_{k-1} y_{t-(k-1)} + \xi_t$$

$$donde: \Pi = (\sum_{j=1}^{k} \beta_i) - I_g \qquad \Gamma = (\sum_{j=1}^{i} \beta_j) - I_g$$

Método de Johansen

- Este VAR contiene g variables en primeras diferencias del lado izquierdo y k-1 de las diferencias de las variables dependientes del lado derecho.
- El método de Johansen tiene las siguientes particularidades:
  - Es un test de máxima verosimilitud que requiere muestras grandes (100 o más datos).
  - Es sensible al número de rezagos del VECM, por lo que es importante elegir el criterio óptimo de selección.
  - El test se centra en el examen de la matriz Π que es la que contiene los coeficientes de largo plazo.

Método de Johansen

- La prueba de cointegración entre las variables (*y*) se calcula a partir del rango de la matriz Π con base en las raíces características que resulten diferentes de cero (eigenvalues).
- Los test estadísticos se refieren a la traza de la matriz y los eigenvalues que resulten significativos.
- H0: VC = r vs H1: VC = r + 1
- El número de VC li (r < g) para obtener solución.

Método de Johansen

- Para elegir el número de rezagos se puede usar el método de Enders (1996):
  - Estimar un VAR con las variables que se desea analizar la cointegración.
  - Chequear que los residuos del VAR estimado no estén correlacionados y sean normales.
  - Tomar nota de los criterios de Akaike y Schwart.
  - Estimar sucesivamente otros VAR con menor número de rezagos y chequear.
  - Elegir el que no tenga correlación serial, normal y menores AK y Sc.

VECTORES AUTO REGRESIVOS VAR

Departamento de Investigaciones Económicas

VAR

- Fueron popularizados por Sims (1980), como una generalización de modelos auto regresivos (AR).
- Se considera como un híbrido entre modelos de series de tiempo AR, y los modelos estructurales de ecuaciones simultaneas.
- Son una alternativa a grandes modelos estructurales de ecuaciones simultaneas.
- Matemáticamente se pueden expresar así:

$$X_{i,t} = c_i + \sum_{s=1}^{n} \alpha_{i,s} X_{i,t-s} + \sum_{s=1}^{n} \alpha_{j\neq i,s} X_{j\neq i,t-s} + \mu_i$$

Ventajas

- No se necesita especificar cual variable es exógeno o endógena (todas son endógenas). Es decir, es ateórico.
- Permite que la variable dependa de más que sus propios rezagos, por lo que es más flexible que un modelo ARMA.
- Al no existir variables contemporaneas en el lado derecho de la ecuación, se puede utilizar MCO.
- Las proyecciónes son por lo general mejores que las proyecciones de modelos estructurales.

DIE 3

Desventajas

- Son ateóricos, por lo que dice poco de decisiones de política.
- Se puede dar el caso de relaciones espurias.
- Los coeficientes no se pueden interpretar económicamente.
- Es dificil determinar la cantidad de rezagos
- Se estiman muchos parámetros
- Todos los componentes del VAR deben ser estacionarios

DIE 4

VII. REPRESENTACIÓN ESTADO-ESPACIO Y EL FILTRO DE KALMAN

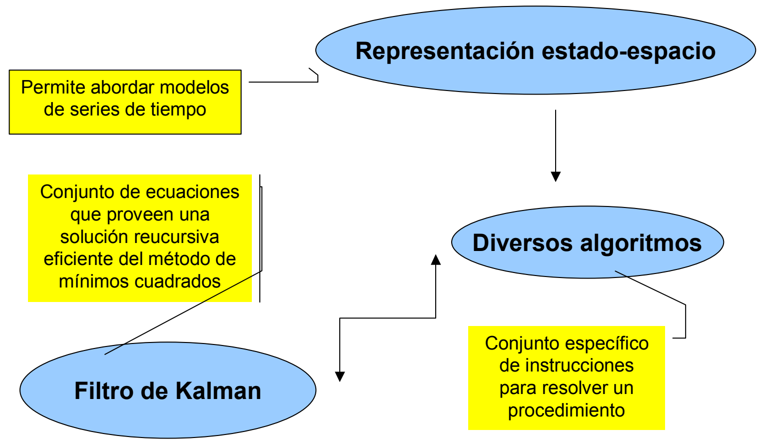
DIE 1

El filtro de Kalman: esquema

- Visualización gráfica
- Generalidades
- Proceso a estimar
- El algoritmo
- El filtro de Kalman y la notación estado-espacio
- Modelos estado-espacio y el filtro de Kalman en EViews

DIE 2

Estado-espacio y el filtro de Kalman: relación



DIE 3

El filtro de Kalman: autor

- Rudolf Emil Kalman
- Nació en 1930 en Hungría
- BS y MS en MIT
- PhD en Columbia 1957
- Desarrolló el filtro entre 1960-61

DIE 4

El filtro de Kalman: generalidades

- Tiene su origen en el documento de R. E. Kalman de 1960 donde describe una solución recursiva al problema del filtrado lineal de datos discretos.
- El filtro es un conjuto de ecuaciones que proveen una solución recursiva eficiente del método de mínimos cuadrados. Permite calcular el estimador óptimo del estado de un sistema en cada momento del tiempo con base en la información disponible en el momento k-1 y actualizar, con la información adicional disponible en el momento k, dichas estimaciones.
- El filtro estima la solución por medio de un proceso de retroalimentación : pronostica y corrige.
- El ciclo del filtro consta de cinco ecuaciones.

El filtro de Kalman: generalidades

- El filtro de Kalman es el principal algoritmo para estimar sistemas dinámicos especificados en la forma estado-espacio.
- En la notación estado-espacio el sistema se formula en términos de una ecuación de estado o proceso y de una ecuación de observación o medida.
- Los modelos estado-espacio tienen muchas aplicaciones econométricas. Entre las que se pueden mencionar: modelación de componentes no observables y parámetros que cambian en el tiempo. También pueden representar modelos Arima.

El filtro de Kalman: proceso a estimar

- El filtro de Kalman resuelve el problema general de estimar el estado X de un proceso controlado en tiempo discreto, el cual es dirigido por una ecuación en diferencia lineal estocástica que asume la siguiente forma:

✓ $X(k) = AX(k-1) + w(k-1)$

Con una medición Z que es:

✓ $Z(k) = HX(k) + v(k)$

- La matriz A relaciona el estado en el periodo previo k-1 con el estado en el momento k. La matriz H relaciona el estado con la medición Z(k). Las variables aleatorias w y v se asumen independientes, ruido blanco y con distribucion normal.

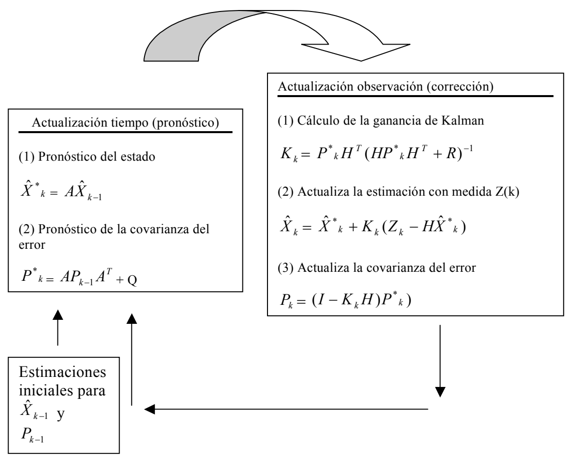
El filtro de Kalman: el algoritmo

- Estima el proceso anterior por medio de un mecanismo de predicción y corrección: el filtro estima el estado del proceso a algún punto en el tiempo y entonces obtiene la retroalimentación por medio de los datos observados. Se debe suponer un valor inicial para la media y varianza del estado para iniciar la estimación.

- El algoritmo se puede derivar a partir de cinco ecuaciones divididas en dos grupos: el primero proyecta el estado (predicción previa) y estima la covarianza del mismo, mientras que el segundo grupo proporciona una estimación posterior mejorada al tomar en cuenta la nueva observación (correción de la estimación previa).

El filtro de Kalman: el algoritmo

Figura 3. Una visión completa del filtro de Kalman



DIE 9

El filtro de Kalman: el algoritmo

- En la derivación de las ecuaciones del filtro el objetivo es encontrar una ecuación que calcule una estimación posterior del estado como una combinación lineal de la estimación previa del estado y la ponderación de la diferencia entre la medida actual y la estimación previa del estado.

DIE 10

El filtro de Kalman y la notación estado-espacio

- La representación estado-espacio se conforma como una herramienta que permite abordar el manejo de un amplio rango de modelos de series de tiempo.
- El filtro de Kalman es el principal algoritmo para estimar sistemas dinámicos especificados en la forma de estado-espacio.
- La representación estado-espacio de un sistema lineal captura la dinámica de un vector $\mathbf{y}_t$ de orden **nx1** en términos de un posible vector no observado α_t, de orden **mx1** conocido como vector de estado.
- Los modelos estado-espacio tienen muchas aplicaciones econométricas.
- En la formalización de la notación estado-espacio el modelo puede reescribirse en términos de dos ecuaciones: una ecuación de estado o proceso y una ecuación de observación o medida.

DIE 11

El filtro de Kalman y la notación estado-espacio

- La notación estado-espacio asume la siguiente forma para la ecuación de estado o proceso, también denominada de transición:

✓ $X_{k+1} = \Phi_k X_k + w_k$

X_{k+1} vector estado al momento k+1

Φ_k matriz transición estado

X_k vector estado al momento k

w_k error de la ecuación de proceso, independiente y con distribución normal

DIE 12

El filtro de Kalman y la notación estado-espacio

- La notación estado-espacio asume la siguiente forma para la ecuación de observación o medida:

✓ $Z_k = M_k X_k + v_k$

Z_k medida de observación al momento k

M_k matriz que relaciona los estados con la medida

X_k vector estado al momento k

v_k error de la ecuación de medida, independiente y con distribución normal

El filtro de Kalman y la notación estado-espacio

- Un ejemplo de representación de un modelo dinámico en la forma de estado-espacio es la siguiente:

✓ $y_t = 2y_{t-1} - y_{t-2} + \varepsilon_t$ (modelo dinámico)

$$X_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}, \Phi_t = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, W_t = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}, M = \begin{bmatrix} 1 & 0 \end{bmatrix}, V_t = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}$$

✓ $X_t = \Phi_t X_{t-1} + w_t$ (ecuación de proceso)

✓ $Z_t = M_t X_t + v_t$ (ecuación de medida)

Modelos estado-espacio y el filtro de Kalman en EViews

- EViews 4.1 tiene dos formas para especificar el modelo a partir de la creación del objeto estado-espacio. Se selecciona **objects/new objects/sspace** desde la barra principal de herramientas. Entonces se crea un objeto estado-espacio y abre una ventana vacía para la especificación del modelo.

- La especificación se puede realizar de dos formas a partir objeto estado-espacio:
- ✓ Utilizando autoespecificación : utiliza ventanas para la creación de formas estándar de estos modelos.
- ✓ Ulizando palabras claves y texto para describir la ecuación de estado, de observación, la estructura de los errores, las condiciones inciales y valores iniciales de los parámetros.

DIE 15

Modelos estado-espacio y el filtro de Kalman en EViews

- Sintáxis de la ecuación de estado : contiene la palabra clave **"@state"** seguida de una especificación que toma en cuenta:

- ➢ Cada ecuación debe tener sólo el nombre de una variable dependiente, expresiones algebraicas no son permitidas.
- ➢ No pueden contener variables dependientes de la ecuación de observación ni adelantos ni rezagos de esas variables.
- ➢ Cada ecuación debe ser lineal en los rezagos de un periodo de los estados. La no linealidad en los estados o la presencia de estados contemporáneos, adelantados o rezagados en varios periodos genera mensajes de error. La restricción de rezagos de un periodo sobre los estados no es restrictiva desde que un mayor orden de rezagos pueden ser escritos como nuevas variables estado.
- ➢ Pueden contener variables exógenas o coeficientes desconocidos y pueden ser no lineales en esos elementos.

DIE 16

Modelos estado-espacio y el filtro de Kalman en EViews

- Las siguientes son ecuaciones de estado válidas:

  @state sv1 = sv1(-1) + [var = exp (c(3))]
  @state sv2 = c(1) + c(2) * sv2(-1) + [var = exp(c(3))]
  @state sv3 = c(1) + exp(c(3)* x/z) + c(2) *sv3(-1) + [var = exp(c(3))]

- Las siguientes no son ecuaciones de estado válidas:
  @state exp(sv1) = sv1(-1) + [var = exp (c(3))]
  @state sv2 = log(sv2(-1)) + [var = exp(c(3))]
  @state sv3 = c(1) + c(2) *sv3(-2) + [var = exp(c(3))]

Modelos estado-espacio y el filtro de Kalman en EViews

- Sintáxis de la ecuación de medida : contiene la palabra clave **“@signal”** seguida de una especificación que toma en cuenta:

- Las variables dependientes pueden envolver expresiones algebraicas.

- No pueden contener valores actuales o adelantos de las variables observables.

- Deben ser lineales en los estados comtemporáneos. La no linealidad en los estados o la presencia de adelantos o rezagos de los estados genera un mensaje de error.

- Pueden tener variables exógenas o coeficientes desconocidos y pueden ser no lineales en esos elementos.

Modelos estado-espacio y el filtro de Kalman en EViews

- Las siguientes son ecuaciones de observación válidas:

  @signal log(y) = c(1) + c(2)*X + C(3)*sv1
  @signal y = sv1 +sv2*X + sv3*Y(-1) + [var = exp(c(1))]
  Z = sv1+ sv2*x + c(1) + [var = exp(c(2))]

- Las siguientes no son ecuaciones de observación válidas:
  @signal log(y) = c(1) + c(2)*X +sv1(-1)
  @signal y = sv1*sv2*x + [var = exp(c(1))]
  Z = sv1 + sv2*x + Z(1) + c(1) + [var = exp(c(2))]

DIE 19

Modelos estado-espacio y el filtro de Kalman en EViews

- En las ecuaciones de estado o medida el término de error se debe especificar explícitamente.
- La forma más fácil es especificando la varianza del término de error por medio de la expresión **"VAR"** seguida de una expresión que difine la varianza. Esta expresión puede ser un valor constante conocido o puede ser una expresión conteniendo parámetros desconocidos que deben ser estimados.
- Por defecto Eviews inicializa los parámetros con los valores en el correspondiene vector. Esto se puede cambiar especificando explicitamente los valores deseados de los parámtros usando la expresión **"PARAM" o "@PARAM".**

DIE 20

Modelos estado-espacio y el filtro de Kalman en EViews

- Por defecto Eviews maneja las condiciones inciales de los estados de una manera difusa: supone una media de cero y la varianza como un número arbitrariamente alto que refleje la incertidumbre acerda del valor.

- Sin embargo si se cuenta con información acerca de los valores de la media y la varianza, se puede crear un vector o matriz que los contenga por medio de las claves **"@mprior"** y **"@vprior"** para la media y la varianza, respectivamente. Estas palabras claves son seguidas del nombre del vector y del nombre de la matriz simétrica definidas previamente.

DIE 21

Modelos estado-espacio y el filtro de Kalman en EViews

- Una manera de ilustrar los modelos estado-espacio y su estimación por medio del algoritmo de Kalman, es a partir de un modelo autorregresivo de serires de tiempo.

- Suponga un modelo AR(2) para la tasa de crecimiento interanual del IMAE

```
vector(2) svec0
svec0.fill 1,0
matrix(2,2) svar0
svar0.fill(b=c) 1, 0.5, 0.5, 2
@mprior svec0
@vprior svar0
param c(1) -1.0 c(2) 0.0 c(3) 0.0 c(4) 0.0
@signal imaein = sv1
@state sv1 = c(4)+c(2)*sv1(-1) + c(3)*sv2(-1) + [var = exp(c(1))]
@state sv2 = sv1(-1)
```

DIE 22

kikutva@bccr.fi.cr
laverdema@bccr.fi.cr
leonmj@bccr.fi.cr
munozse@bccr.fi.cr
sanchezrm@bccr.fi.cr
torresgc@bccr.fi.cr
solerara@bccr.fi.cr