Causalidad y Modelos dinámicos

Una aplicación de los modelos dinámicos

La causalidad

- Cuando una variable es causa de otra, los datos manifiestan una alta relación entre ellas
- sin embargo lo contrario no es verdad, una alta relación no implica causalidad
- La causalidad debe estar justificada teoréticamente
- Desde el punto de vista económico se plantean tres problemas alternativos cuando se analizan relaciones entre variables:
  - Causalidad temporal
    - Una variable ocurre cuando otra ocurre previamente: da pie a los modelos ARMAX
  - Endogeneidad
    - se trata de estudiar si un determinado conjunto de variables que intervienen en el modelo se determinan dentro del modelo o vienen dadas por factores externos a él.
  - Simultaneidad
    - Fruto de la endogeneidad, cuando dos variables son a la vez causa y efecto entre ellas. En este caso existen mas de una variable dependiente

Definición de modelos ARMAX

- Las variables independientes pueden afectar a la dependiente en periodos anteriores, pero también, la propia dependiente puede afectarse a si misma. En consecuencia, los modelos dinámicos lineales de tipo económico, pueden presentarse según la siguiente ecuación:

$$A(L)(Y_t - \mu) = B(L)x_t + u_t$$

Donde

$$A(L) = 1 - \alpha_1 L - \ldots - \alpha_s L^s$$

$$B(L) = \beta_0 + \beta_1 L + \ldots + \beta_r L^r$$

suponiendo que u_t es un proceso ARMA(p,q)

- En forma compacta , se escribiría

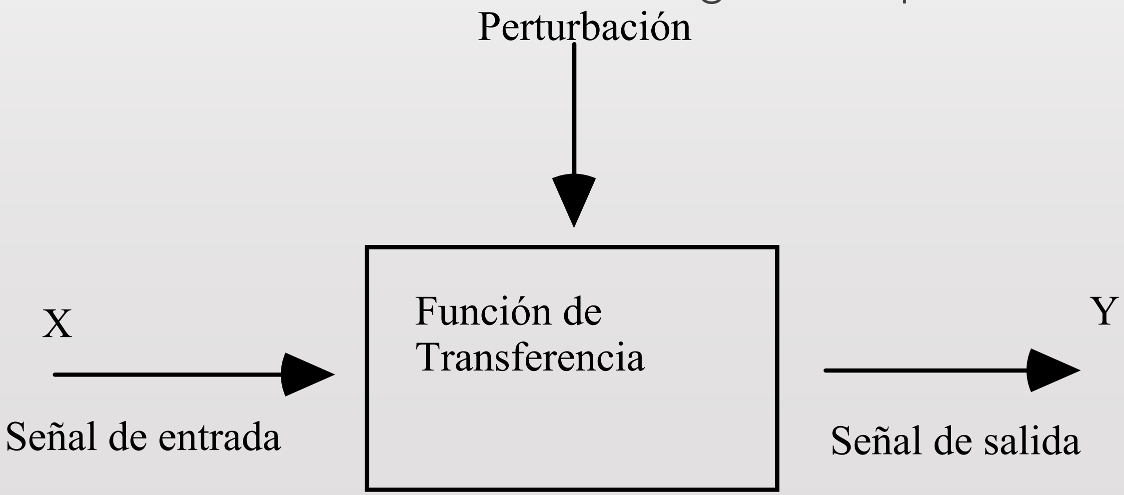
$$Y_t = \mu + \frac{B(L)}{A(L)} x_t + \frac{\phi(L)}{A(L)\theta(L)} \varepsilon_t$$

Función de transferencia

- La ecuación anterior, se puede escribir en forma compacta como

$$Y_t = \mu + \frac{B(L)}{A(L)}x_t + \frac{\theta(L)}{A(L)\phi(L)}\varepsilon_t = \mu + \overbrace{D(L)x_t}^{\text{Función de transferencia}} + \overbrace{\Psi(L)\varepsilon_t}^{\text{perturbación}}$$

que se denomina función de transferencia, según el esquema



Ejemplo: efecto del tipo de interés sobre la masa monetaria

- Supongamos que el tipo de interés (r) condiciona la cantidad de dinero (m) por una función ARMAX del tipo

$$m_t = \beta_0 + \beta_1 r_t + \beta_2 r_{t-1} + \beta_3 r_{t-2} + \alpha_1 m_{t-1} + \epsilon_t$$

- El modelo tendría que

$$A(L) = 1 - \alpha_1 L$$

$$B(L) = \beta_0 + \beta_1 L + \beta_2 L^2$$

$$A(L) = 1 - 0.448L$$

$$B(L) = 0.002 - 0.004L + 0.002L^2$$

Modelo 3: ARMAX, usando las observaciones 1980:3-1999:3 (T = 77)
Estimado usando Mínimos cuadrados (MV condicional)
Variable dependiente: (1-L) m_p

	Coeficiente	Desv. Típica	z	Valor p	
const	0.00505748	0.00177151	2.855	0.0043	***
phi_1	0.448006	0.101040	4.434	9.25e-06	***
rs	0.00173774	0.000894064	1.944	0.0519	*
rs_1	−0.00379740	0.00149118	−2.547	0.0109	**
rs_2	0.00197631	0.000904807	2.184	0.0289	**

Media de la vble. dep. 0.007701 D.T. de la vble. dep. 0.004657
media innovaciones 0.000000 D.T. innovaciones 0.004026
Log-verosimilitud 317.9737 Criterio de Akaike −625.9475
Criterio de Schwarz −614.2285 Crit. de Hannan-Quinn −621.2600

Dinámica sistemática

- Separamos el comportamiento de la serie en dos parte:
  - La dinámica del valor esperado de la dependiente condicionado a la independiente
  - La dinámica de la perturbación.
- La dinámica sistemática hacer relación al comportamiento de la primera.

$$E\left(Y_t / X_t \right) = \mu + \frac{B(L)}{A(L)} X_t = \mu + D(L) X_t$$

donde B(L)=A(L)D(L), y por tanto

$$\delta_j = \sum_{i=1}^{\min(j,r)} \alpha_i \delta_{j-i} + \beta_i \quad si \quad 0 \leq j \leq s$$

$$\delta_j = \sum_{i=1}^{\min(j,r)} \alpha_i \delta_{j-i} \quad si \quad j > s$$

Características de la dinámica sistemática

- El efecto depende de los coeficientes D(L) de retardos implicados, que como se ve son combinaciones de los polinomios A(L) y B(L), afectando a la predicción.
  - La parte correspondiente a la variable independiente B(L) tiene un efecto finito
  - la parte A(L) de la variable dependiente afecta siempre, por ello es necesario conocer sus raíces para asegurar la estabilidad del proceso.

Ejemplo: efecto del tipo de interés sobre la masa monetaria(2)

- En este caso la dinámica sistemática estaría dada por ella función de transferencia

$$\frac{B(L)}{A(L)} = \frac{0.002 - 0.004L + 0.002L^2}{1 - 0.448L}$$

- Un impulso dura infinitamente, pero como la raíz es menor que 1, ese efecto va suavizándose. La serie es estable
- La dinámica de la perturbación, es un ruido blanco inicialmente, pero al introducir la parte de A(L) de la dependiente nos quedará un modelo autorregresivo de orden 1.

$$\frac{1}{A(L)} = \frac{1}{1 - 0.448L}$$

Dinámica sistemática y equilibrio

- Si suponemos un comportamiento estable a largo plazo del modelo, es decir, que existe una tendencia al <u>equilibrio,</u> este vendrá dado por

$$\overline{Y} = \mu + \frac{B(1)}{A(1)}\overline{X} = \mu + \frac{\beta_0 + \beta_1 + \ldots + \beta_s}{1 - \alpha_1 - \ldots - \alpha_r}\overline{X}$$

- El efecto de la variable independiente sobre la dependiente depende del tiempo.
- Ese efecto puede ser diferente en cada periodo de tiempo. Al valor numérico que nos mide el impacto para cada periodo de tiempo se le da el nombre genérico de <u>multiplicador</u>, que por consiguiente se define como el impacto que sobre Y produce una cambio en X de una unidad en ese periodo.

Efecto multiplicador

- Multiplicador a largo plazo o multiplicador total
  - Mide el efecto en el equilibrio

$$D(1) = \delta_{\infty}^{*} = \sum_{j=0}^{\infty} \delta_j$$

- Multiplicador de impacto
  - Mide el efecto instantáneo en el periodo t, vendrá dado por δ_0.
- Multiplicador J-interim
  - Mide el impacto después de J periodos

$$\delta_J^{*} = \sum_{j=0}^{J} \delta_j$$

- Proporción de impacto en el periodo J
  - Mide la parte del impacto total que se ha producido al cabo de J periodos

$$\delta_J^{+} = \delta_J^{*} \big/ \delta_{\infty}^{*}$$

Retardos

- De modo complementario al estudio de los multiplicadores esta el estudio de los retardos, es decir, el tiempo que tarda en producirse un determinado efecto.
  - Retardo mediano
    - El número de periodos necesarios para que se dé la mitad del impacto total
  - Retardo medio
    - cuando todos los dj≥0, se define el como el promedio de los retardos al realizar los impactos, o sea,

$$\bar{\delta}_M = \frac{\sum_{j=0}^{\infty} j\delta_j}{\delta_\infty^*} = \frac{D'(1)}{D(1)} donde \quad D'(L) = \frac{\partial D(L)}{\partial L}$$

Ejemplo: efecto del tipo de interés sobre la masa monetaria(3)

- El valor del logaritmo de la masa monetaria real en el equilibrio sería

$$\overline{Y} = \mu + \frac{B(1)}{A(1)}\overline{X} = \mu + \frac{\beta_0 + \beta_1 + \beta_2}{1-\alpha_1}\overline{X} = 0{,}005057480 + \frac{-0{,}00373131}{1-0.448}8.9682 = 0{,}0037033$$

Multiplicador a largo plazo o multiplicador total

$$D(1) = \delta_\infty^* = \sum_{j=0}^{\infty}\delta_j = -0{,}000151$$

Multiplicador de impacto es 0,00173774
Multiplicador 1-interim

$$\delta_1^* = \sum_{j=0}^{1}\delta_j = -0{,}00373131$$

Caso particular: modelo de Koyck

$$Y_t = \alpha Y_{t-1} + \mu + \beta x_t + u_t \qquad 0 < \alpha < 1$$

donde u seguirá en general un proceso ARMA(p,q).

- Es un modelo de retardos infinitos, donde el polinomio A(L) es de orden 1 y el B(L) de orden 0, entonces los coeficientes de D(L) serán:

$$\delta_i = \beta \alpha^i \qquad i = 0,\ldots\infty$$

- A α se le conoce como *tasa de disminución*
- Como a >0, todos los δ_i tendrán signo el signo de β, que suele ser positivo, y en consecuencia todos los multiplicadores J-interim son crecientes (si β es negativo serian decrecientes).
- Como a<1, el pasado cada vez es menos influyente.
- El multiplicador de impacto será β y el total será $\delta_\infty^* = \beta / 1-\alpha$
- Por eso, a $(1-\alpha)$ se le llama *velocidad de ajuste, puesto que es el cociente entre el impacto a corto plazo y el impacto en el equilibrio.*
- El retardo medio viene dado por $\overline{\delta_M} = \alpha / 1-\alpha$

Interpretación económica del modelo de Koyck

- En la practica este modelo se utiliza para dos casos de modelos económicos:
  - El ajuste parcial
    - Este introduce un nivel deseado y hay dos ecuaciones:
      - La relación entre el nivel deseado de la dependiente y la independiente
      - un ajuste parcial del nivel real hacia el nivel deseado en la dependiente
  - Las exceptivas adaptativas
    - Introduce un modelo con dos ecuaciones:
      - la variable dependiente depende de las expectativas de la variable independiente
      - as expectativas de la independiente son adaptativas en función de su pasado reciente

Modelo de ajuste parcial

- Sea Y* el nivel deseado de la variable Y (por ejemplo el gasto) y X la independiente (por ejemplo, la renta). Inicialmente supondremos que el gasto deseado es una parte β de la renta, luego

$$C_t^* = \beta Y_t + u_t$$

- Sin embargo, el consumidor va a tardar un tiempo en ajustar el gasto real al nivel deseado y sólo consigue un ajuste parcial, aproximado. Denominaremos γ al grado de ajuste, y por tanto, la función de ajuste que denominaremos <u>función de reacción</u> será

$$C_t - C_{t-1} = \gamma(C_t^* - C_{t-1}) + u_t \qquad 0 < \gamma < 1$$

- sustituyendo C* tendremos que

$$C_t = (1-\gamma)C_{t-1} + \beta\gamma Y_t + u_t$$

- Presenta un modelo de Koyck

Modelo de expectativas adaptativas

- Vamos a analizar ahora el modelo de expectativas adaptativas, que en el caso de la renta coincide con el modelo de Friedman de la renta permanente. La variable dependiente depende de las expectativas de la independiente

$$C_t = \beta_0 + \beta_1 Y_t^* + u_t$$

- Y* nos mediría algún concepto de expectativas de la renta a largo plazo, como puede ser el concepto de renta permanente. Normalmente, no es directamente observable, por lo que formulamos una hipótesis sobre como es la formación de sus expectativas.

$$Y_t^* = Y_{t-1}^* + \gamma(Y_t - Y_{t-1}^*) \quad 0 < \gamma < 1$$

- El decisor revisa sus expectativas sobre la base de la discrepancia entre su predicción y el valor real. El parámetro γ mide la reacción a esa discrepancia. A este tipo de expectativas se les conoce como <u>expectativas adaptativas de primer orden.</u> Tomando $\alpha=1-\gamma$, sustituyendo en el modelo

$$C_t = \beta_0 + \beta_1 \frac{1-\alpha}{1-\alpha L} Y_t + u_t$$

- Que es un esquema de Koyck

Modelo de Koyck en Galicia

La constante es significativa peor el efecto del pasado es negativo, no representa expectativas adaptativas

efecto del pasado o es significativo pero podría ser ajuste parcial

Modelo 9: ARMAX, usando las observaciones 1995:3-2011:4 (T = 66)
Variable dependiente: ld_Gastoenconsumofinal
Desviaciones típicas basadas en el Hessiano

	Coeficiente	*Desv. Típica*	*z*	*Valor p*	
const	0.00398481	0.0021658	1.8398	0.0658	*
phi_1	−0.524347	0.147813	−3.5474	0.0004	***
C_ret	−0.242084	0.0888893	−2.7234	0.0065	***
ld_PIBpm	0.815885	0.077232	10.5641	<0.0001	***

Media de la vble. dep. 0.011798 D.T. de la vble. dep. 0.053663
media innovaciones −0.000208 D.T. innovaciones 0.012588
Log-verosimilitud 194.9377 Criterio de Akaike −379.8753
Criterio de Schwarz −368.9270 Crit. de Hannan-Quinn −375.5491

Modelo 8: ARMAX, usando las observaciones 1995:3-2011:4 (T = 66)
Variable dependiente: ld_Gastoenconsumofinal
Desviaciones típicas basadas en el Hessiano

	Coeficiente	*Desv. Típica*	*z*	*Valor p*	
phi_1	−0.895049	0.142374	−6.2866	<0.0001	***
C_ret	0.0895448	0.191224	0.4683	0.6396	
ld_PIBpm	0.766341	0.166053	4.6151	<0.0001	***

Media de la vble. dep. 0.011798 D.T. de la vble. dep. 0.053663
media innovaciones 0.000511 D.T. innovaciones 0.012941
Log-verosimilitud 192.4669 Criterio de Akaike −376.9337
Criterio de Schwarz −368.1751 Crit. de Hannan-Quinn −373.4728

Inferencia en modelos ARMAX

Con variables estacionarias

Estimadores en modelos dinámicos

- Caben varias opciones
  - Estimador de mínimos cuadrados (OLS)
    - Aunque el modelo tenga a la variable retardada como explicativa, si las perturbaciones se comportan como un ruido blanco se puede estimar por OLS, y los estimadores seguirían siendo eficientes asintóticamente, pues coincidirían con los de MLE.
    - Sólo es válido si los residuos son incorrelados
  - Estimador de Variables instrumentales (VI).
    - Consiste en buscar un conjunto de variables incorreladas con las perturbaciones u_t pero correladas con las variables explicativas.
  - Estimador de máxima verosimilitud
    - Consiste en maximizar la función de verosimilitud, es equivalente a minimizar la suma de las perturbaciones al cuadrado, pero trabajando ya sobre perturbaciones que son ruido blanco, por tanto introduce todos los parámetros del modelo ARMA. Tendríamos que tener en cuenta el hecho de los valores iniciales como vimos en las series temporales, que puede afectar a la eficiencia de la estimación.

Estimador de variables instrumentales

- Reescribimos el modelo como

$$Y_t = Z_t \delta + u_t$$

donde Z contiene tanto las variables exógenas como las endógenas retardadas.

- Se debe buscar un conjunto de variables W, del mismo rango que Z, verificando que

1. Debe estar correlada con las variables explicativas
2. Está incorrelada con las perturbaciones

- El estimador IV vendrá dado por $\hat{\delta} = (W'Z)^{-1} W'y$

Propiedades del estimador IV

- Consistente
- Asintóticamente insesgado
- No es eficiente
- Asintóticamente normal
  - Su ley de distribución es aproximadamente la siguiente:

$$T^{-1/2}(\hat{\delta}_i - \delta_i) \text{ es } AN(0, \sigma^2 \Sigma_{ZW}^{-1} \Sigma_{WW} \Sigma_{ZW}^{-1})$$

Test de hipótesis de coeficientes individuales en variables instrumentales

- Hipótesis $\begin{matrix} H_0 : \psi_i = \psi_{0i} \\ H_1 : \psi_i \neq \psi_{0i} \end{matrix}$
- Estadístico

$$t_i = \frac{\sqrt{T}(\hat{\Psi}_i - \Psi_{0i})}{s_{\hat{\Psi}_i}}$$

siendo $s_{\psi i}$ la raíz cuadrada del elemento i-simo de la diagonal de la matriz de varianzas covarianzas estimada

- Ley de distribución:
  - seguirá aproximadamente una ley normal tipificada.

Test de hipótesis de coeficientes conjuntos en variables instrumentales

- Hipótesis $$H_0: \ \psi_i = 0 \ \forall i = 1 \dots r$$ $$H_1: \ \exists i = 1 \dots r \ / \ \psi_i \neq 0$$
- Estadístico $$F = T \sum_{i=1}^{r} \left(\frac{\hat{\psi}_i}{s_{\hat{\psi}_i}} \right)^2$$

siendo s_{yi} la raíz cuadrada del elemento i-simo de la diagonal de la matriz de varianzas covarianzas estimada

- Ley de distribución:
  - seguirá aproximadamente una chi cuadrado con r-p-q grados de libertad.

Inferencia en el efecto tipo de interés masa monetaria

```
Modelo 5: MC2E, usando las observaciones 1980:3-1999:3 (T = 77)
Variable dependiente: d_m_p
Mediante Instrumentos: d_m_p_1
Instrumentos: const y_1 rs rs_1 rs_2

            Coeficiente   Desv. Típica     z       Valor p
  --------------------------------------------------------
  rs          0.00115830   0.00104419      1.109   0.2673
  rs_1       -0.00365633   0.00171477     -2.132   0.0330    **
  rs_2        0.00254538   0.000998645     2.549   0.0108    **
  d_m_p_1     0.943336     0.161694        5.834   5.41e-09 ***

Media de la vble. dep.   0.007701   D.T. de la vble. dep.    0.004657
Suma de cuad. residuos   0.001557   D.T. de la regresión     0.004618
R-cuadrado               0.267516   R-cuadrado corregido     0.237414
F(4, 73)                 55.44188   Valor p (de F)           2.14e-21
rho                     -0.248486   Durbin-Watson            2.485248
```

Test de coeficientes conjunto

Test de coeficientes individuales

Permite ver la endogenidad de las variables

```
Contraste de Hausman -
  Hipótesis nula: Los estimadores de MCO son consistentes
  Estadístico de contraste asintótico: Chi-cuadrado(1) =
8.72651
  con valor p  = 0.00313617

Contraste de sobreidentificación de Sargan -
  Hipótesis nula: todos los instrumentos son válidos
  Estadístico de contraste: LM = 0.00839732
  con valor p  = P(Chi-cuadrado(1) > 0.00839732) =
0.926987

Contraste de Instrumento débil -
  Estadístico F de la primera etapa (2, 72) = 18.4934
  Valores críticos para el tamaño maximal deseado de
MC2E, cuando
   los contrastes se ejecutan a un nivel de significación
nominal del 5%:

  tamaño      10%      15%      20%      25%
    valor    19.93    11.59    8.75     7.25

  El tamaño maximal puede ser superior a 10%
```

Propiedades de los estimadores de máxima verosimilitud

- La estimación consiste en maximizar la función de verosimilitud suponiendo **normales** los ruidos blancos del modelo.
- Sea ψ el vector de parámetros del modelo, entonces el estimador de MV verifica las siguientes propiedades:
  - Asintóticamente insesgado
  - Consistente
  - Asintóticamente eficiente
  - Asintóticamente normal. El estadístico $\sqrt{T}(\hat{\psi} - \psi)$ tiene como distribución asintótica una ley normal $AN\,(0, J^{-1}(\psi))$

siendo J(ψ) la matriz de información de Fisher

Test de hipótesis de coeficientes individuales en estimación de máxima verosimilitud

- Hipótesis
$$H_0 : \psi_i = \psi_{0i}$$
$$H_1 : \psi_i \neq \psi_{0i}$$
- Estadístico
$$t_i = \frac{\sqrt{T}(\hat{\Psi}_i - \Psi_{0i})}{s_{\hat{\Psi}_i}}$$

siendo $s_{\gamma i}$ la raíz cuadrada del elemento i-simo de la diagonal de la matriz de información de Fisher invertida

- Ley de distribución:
  - seguirá aproximadamente una ley normal tipificada.

Test de hipótesis de coeficientes conjuntos en estimación de máxima verosimilitud

- Hipótesis $\begin{array}{l} H_0 : \ \psi_i = 0 \ \forall i = 1 \ldots r \\ H_1 : \ \exists i = 1 \ldots r \ / \ \psi_i \neq 0 \end{array}$
- Estadístico $F = T \sum_{i=1}^{r} \left(\frac{\hat{\psi}_i}{s_{\hat{\psi}_i}} \right)^2$

siendo s_{yi} la raíz cuadrada del elemento i-simo de la diagonal de la matriz de varianzas covarianzas estimada

- Ley de distribución:
  - seguirá aproximadamente una chi cuadrado con r-p-q grados de libertad.

Inferencia en el efecto tipo de interés masa monetaria

Test de coeficientes individuales

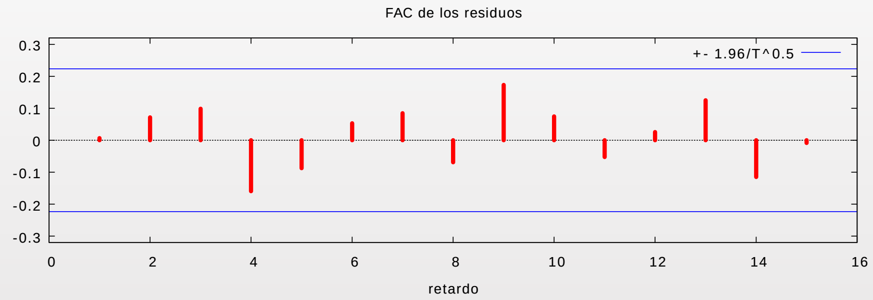
Modelo 3: ARMAX, usando las observaciones 1980:3-1999:3 (T = 77)
Estimado usando Mínimos cuadrados (MV condicional)
Variable dependiente: (1-L) m_p

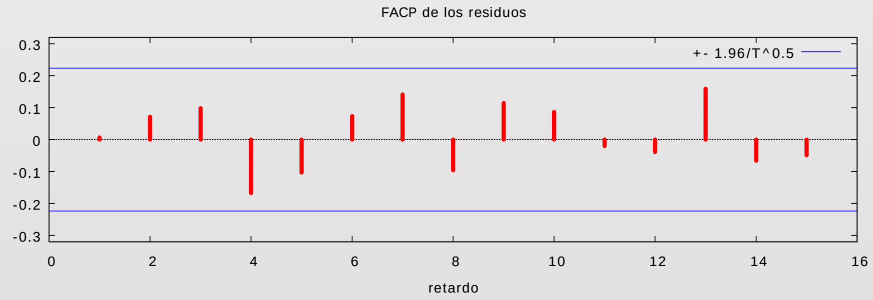
	Coeficiente	Desv. Típica	z	Valor p	
const	0.00505748	0.00177151	2.855	0.0043	***
phi_1	0.448006	0.101040	4.434	9.25e-06	***
rs	0.00173774	0.000894064	1.944	0.0519	*
rs_1	−0.00379740	0.00149118	−2.547	0.0109	**
rs_2	0.00197631	0.000904807	2.184	0.0289	**

Media de la vble. dep.	0.007701	D.T. de la vble. dep.	0.004657
media innovaciones	0.000000	D.T. innovaciones	0.004026
Log-verosimilitud	317.9737	Criterio de Akaike	−625.9475
Criterio de Schwarz	−614.2285	Crit. de Hannan-Quinn	−621.2600

	Real	Imaginaria	Módulo	Frecuencia
AR				
Raíz 1	2.2321	0.0000	2.2321	0.0000

Contraste de normalidad de los residuos -
Hipótesis nula: el error se distribuye normalmente
Estadístico de contraste: Chi-cuadrado(2) = 1.28704
con valor p = 0.525439





Causalidad y exogeneidad

Modelos VAR

- Las series temporales se pueden generalizar al caso multivariante, pero en este caso, normalmente existe interacción entre las diferentes componentes de la serie.
- Sea X un vector de dimensión k, la forma general es

$$A_p(L)X_t = \varepsilon_t, \qquad donde \qquad A_p(L) = \sum_{j=0}^{p} \Phi_j L^j con \qquad \Phi_0 = I_k$$

donde Φ_j es un conjunto de matrices de orden k x k.

- El tratamiento es similar a las series univariantes, y se puede considerar un tipo particular de ecuaciones simultáneas, donde todas las variables (en este caso las componentes de X) son endógenas.
- Esto sugiere la necesidad del estudio de la exogeneidad.

Selección de modelos VAR

- Al igual que ocurría con los modelos de series temporales, es necesario identificar el orden de retardos para estimar el modelo.
- En la practica, la decisión se basa en una serie de criterios estadísticos que tienen en cuenta dos elementos
  - Simplicidad o parsimonia
    - Se busca el mínimo numero de parámetros posible
  - El grado de ajuste a los datos
    - Maximiza la función de verosimilitud

Criterios de selección de modelos VAR

Criterios	Formula	Características
Akaike	$AIC = \log(\hat{\sigma}^{2}) + {}^{2p}/_{n}$	Maximiza la función de verosimilitud restando el numero de parámetros
Schwarz	$BIC = \log(\hat{\sigma}^{2}) + \frac{p\ log\ n}{n}$	Maximiza la probabilidad del modelo
Hannan-Quin	HQ= $\log(\hat{\sigma}^{2}) + \frac{2p \log log\ n}{n}$	Maximiza la información con el mínimo numero de parámetros

Ejemplo de modelo VAR: comportamiento de los tipos de interés: Selección del modelo

```
Sistema VAR, máximo orden de retardos 8

Los asteriscos de abajo indican los mejores (es decir, los mínimos)
valores de cada criterio de información, AIC = criterio de Akaike,
BIC = criterio bayesiano de Schwarz y HQC = criterio de Hannan-Quinn.

retardos  log.veros    p(RV)       AIC          BIC          HQC

   1      -74.28303              2.317832     2.572782     2.419217
   2      -52.37816  0.00000     1.813469*    2.195894*    1.965548*
   3      -51.77988  0.87866     1.909292     2.419192     2.112063
   4      -50.05585  0.48582     1.973404     2.610779     2.226868
   5      -44.62034  0.02805     1.932967     2.697817     2.237124
   6      -44.39503  0.97813     2.039297     2.931621     2.394146
   7      -40.27994  0.08350     2.036055     3.055854     2.441596
   8      -37.18011  0.18473     2.061412     3.208686     2.517646
```

Elegimos el modelo de orden 2

Ejemplo de modelo VAR: comportamiento de los tipos de interés: Estimación

```
Sistema VAR, orden del retardo 2
estimaciones de MCO, observaciones 1980:3-1999:3 (T = 77)
Log-verosimilitud = -65.968407
Determinante de la matriz de covarianzas = 0.019019528
AIC = 2.0252
BIC = 2.3904
HQC = 2.1713
Contraste Portmanteau: LB(19) = 81.3231, gl = 68 [0.1289]

Ecuación 1: rs

      Coeficiente Desv. Típica      Estadístico t     Valor p

const 1.06304     0.65964     1.6115      0.1115
rs_1  1.23028     0.132764    9.2666      <0.0001      ***
rs_2  -0.313196   0.13402     -2.3369     0.0223       **
rl_1  0.327385    0.179292    1.8260      0.0721       *
rl_2  -0.325276   0.175029    -1.8584     0.0673       *
time  -0.00971476 0.00551704  -1.7609     0.0826       *

Media de la vble. dep.   8.867297         D.T. de la vble. dep.
3.211894
Suma de cuad. residuos   18.58063         D.T. de la regresión
0.511565
R-cuadrado   0.976301         R-cuadrado corregido    0.974632
F(5, 71)     584.9892         Valor p (de F)     3.25e-56
rho    0.057706          Durbin-Watson      1.818554
Contrastes F de restricciones cero:
Todos los retardos de rs        F(2, 71) =    138.8 [0.0000]
Todos los retardos de rl        F(2, 71) =   1.7894 [0.1745]
Todas las variables, retardo 2 F(2, 71) =   11.938 [0.0000]
```

```
Ecuación 2: rl

      Coeficiente Desv. Típica      Estadístico t     Valor p
const 1.559 0.455568    3.4221      0.0010      ***
rs_1  -0.0230124  0.091691    -0.2510     0.8026
rs_2  0.0850137   0.0925583   0.9185      0.3615
rl_1  1.5275      0.123825    12.3360     <0.0001     ***
rl_2  -0.701663   0.12088     -5.8046     <0.0001     ***
time  -0.0114293  0.00381024  -2.9996     0.0037      ***

Media de la vble. dep.   9.552747         D.T. de la vble. dep.    2.758632
Suma de cuad. residuos   8.862412         D.T. de la regresión     0.353303
R-cuadrado   0.984677         R-cuadrado corregido     0.983598
F(5, 71)     912.4966         Valor p (de F)     6.23e-63
rho    0.057665          Durbin-Watson      1.852848
Contrastes F de restricciones cero:
Todos los retardos de rs        F(2, 71) =   1.3034 [0.2780]
Todos los retardos de rl        F(2, 71) =   149.21 [0.0000]
Todas las variables, retardo 2 F(2, 71) =   23.076 [0.0000]


Para el sistema en conjunto
Hipótesis nula: el retardo más largo es 1
Hipótesis alternativa: el retardo más largo es 2
Contraste de razón de verosimilitudes: Chi-cuadrado(4) = 50.8237 [0.0000]
```

Sea

$D(x_T^1 | \theta, x_0)$

$\theta \in \Theta$

Exogeneidad

- La exogeneidad estudia si la información se contiene dentro o fuera del modelo
- Hablando estrictamente se diría que una variable Z_t es exógena si, al hacer inferencias sobre los parámetros importantes del modelo, condicionados al conocimiento de Z_t no se pierde información sobre dichos parámetros.
- Para formalizar el concepto hablaremos únicamente de exogeneidad estricta y exogeneidad Dhrymes o D-exogeneidad para el caso de no verificarse la normalidad.
- Para desarrollar estos dos conceptos introducimos el concepto de procesos innovativos

Procesos innovativos

- Sea x_0 el conjunto de condiciones iniciales en el vector X.
- Definimos las expectativas de x a partir de la información obtenida hasta el periodo t-1, como $\varphi_t = E(X_t|\theta, X_{t-1})$
- Llamamos proceso innovativo a $\vartheta_t = X_t - \varphi_t$, es decir, la información sobre x_t obtenida sólo en el periodo t.
- Es evidente que tiene una distribución de media 0, y además, está incorrelado con su pasado, puesto que toda información posible hasta el momento t-1 ya está incluida en φ_t.
- Suponemos que su matriz de covarianzas es constante

Sea

$D(x_T^1 | \theta, x_0)$

$\theta \in \Theta$

Exogeneidad fuerte y debil

- Consideremos un proceso de vectores aleatorios

$$X_t = \begin{pmatrix} y_t \\ z_t \end{pmatrix} \text{ donde } \quad y_t \in \Re^k, z_t \in \Re^G \quad k + G = n$$

- Un conjunto de variables Z se dice que son <u>predeterminadas</u> si Z es independiente de las perturbaciones del modelo, por tanto no depende de los errores de ese periodo y de los periodos siguientes.
- Se dice que es <u>estrictamente exógena</u>, si es independiente de todos los errores, tanto pasados como futuros, y por tanto no hay nada en el modelo que la pueda determinar.
- Se dice que es exogena-Dhrymes, si Z_t está incorrelada con todos los errores, tanto pasados, como presentes, como futuros.
- Es evidente que baja normalidad coincide con la exogeneidad estricta, pero no en otros casos.

Causalidad de Granger

- Un vector aleatorio X se dice que no G-causa Y si la información retardada de la variable X no influye en el valor de Y una vez conocidos los valores retardados de Y.
- A lo largo de lo que sigue supondremos normalidad y en consecuencia el hecho de que X no G-cause Y es equivalente a decir que Y_t no depende de X_{t-j} $j=1..._$, una vez dado Y_{t-j} $j=1..._$
- Cuando X G-causa Y e Y G-causa X entonces se dice que el sistema tiene retroalimentación (feedback), y podemos hablar de simultaneidad de X e Y.

Exogeneidad y G-causalidad

- Consideramos un modelo VAR

$$\begin{pmatrix} A(L) & B(L) \\ C(L) & D(L) \end{pmatrix}\begin{pmatrix} X_t \\ Z_t \end{pmatrix}=\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

donde u tiene matriz de varianzas covarianzas constante Σ y es incorrelada con su pasado.

- Despejando X_t y Z_t , reescribimos el modelo

$$\begin{pmatrix} X_t \\ Z_t \end{pmatrix} = \begin{pmatrix} H_{11}(L) & H_{12}(L) \\ H_{21}(L) & H_{22}(L) \end{pmatrix}\begin{pmatrix} X_t \\ Z_t \end{pmatrix}=\begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} \text{ con } \mathrm{var}(\mathrm{v}) = \Omega = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$$

Exogeneidad y G-causalidad

- Se dice que X <u>no causa</u> Z si $H_{21}(L)=0$
  - el pasado de la variable X no ayuda a mejorar la predicción de la variable Z que se obtendría utilizando sólo el pasado de esta variable
- Se dice que X <u>es predeterminada</u> si $cov(X_t, u_{2t-s}) = 0 \; \forall \, s \; \leq 0$
  - la información del pasado en la variable Z no condiciona la variable X, pasada y contemporánea.
- Se dice que X <u>es exógena, en sentido estricto,</u> si $cov(X_t, u_{2t-s}) = 0 \; \forall \, s$
  - si es independiente de todos los valores de la perturbación aleatoria correspondiente a esa relación.
- Si z_t es predeterminada y X no G-causa Z entonces Z es estrictamente exógena. La implicación inversa se verifica también si el modelo está exactamente identificado.

MODELO DE LOS TEST DE G-CAUSALIDAD

- Para simplificar el planteamiento del test y lo que sigue en el resto de la sección supondremos solo 2 variables **x** e **y,** pero la generalización al caso multivariante es evidente.
- Supongamos dos valores enteros J y K suficientemente altos, tal que

$$Y_t = \alpha + \sum_{i=1}^{k} \beta_i Y_{t-i} + \sum_{j=1}^{J} \gamma_j X_{t-j} + \varepsilon_t$$

de forma que no haya efectos retardados que no se recojan en el modelo.

F-test bajo ruido blanco

- Suposiciones

suponemos inicialmente que las perturbaciones son normales independientes e igualmente distribuidas ε_t es $N(0, \sigma^2)$

- Hipótesis $H_0: \ \gamma_i = 0 \ \forall i = 1 \ldots J$
  $H_1: \ \exists i = 1 \ldots J \ / \ \gamma_i \neq 0$

- Estadístico $F = \dfrac{(SSE_0 - SSE_1)/J}{SSE_1/(T-J-k-1)}$

siendo SSE_0 la suma de cuadrados de los errores en el modelo sin la variable X y SSE_1 la suma de cuadrados de los errores en el modelo incluyendo todas las variables

- Ley de distribución:

seguirá una F de Snedecor con J y T-J-k-1 grados de libertad.

Test bajo errores ARMA: modelo

- Suposiciones

  suponemos inicialmente que las perturbaciones son normales pero siguen un comportamiento ARMA(p,q)

- El procedimiento cambia, puesto que primero conviene ajustar los errores para que se comporten como ruido blanco:

  1º)- Se hace un preblanqueado sobre Y_t y sobre X_t., es decir, se aplica un filtro lineal a cada variable de forma que los residuos sean ruido blanco. Sean e_{xt} y e_{yt} los residuos correspondientes.

  2º)- Se calcula la correlación entre esos residuos $\gamma_k = corr(e_{y_t}, e_{x_{t-k}})$

Test bajo errores ARMA: contraste

- La hipótesis alternativa es de dos tipos según el signo de las correlaciones

$$H_0 : \ \gamma_i = 0 \ \forall k \neq 0$$

$$H_{1A} : \ \exists k > 0 \ / \ \gamma_i > 0$$

$$H_{1B} : \ \exists k < 0 \ / \ \gamma_i > 0$$

Si se verifica H_A se dice que X G-causa Y, si es la otra a<lternativa, entonces Y G-causa X

Si se incluye el retardo 0 se dice que la causalidad es instantánea.

- Estadístico
  - Se utiliza un estadístico tipo Ljung-Box-Pierce

$$Q = T(T+2)\sum_{j=1}^{k}\frac{r_j^2}{T-j}$$

- Ley de distribución:
  - seguirá aproximadamente una chi cuadrado con k-p-q grados de libertad, siendo p y q los ordenes del proceso ARMA.

¿Condicionó el bono US al bono Australiano?

¿Condicionó el bono US al bono Australiano? Selección modelo VAR

```
Sistema VAR, máximo orden de retardos 8

Los asteriscos de abajo indican los mejores (es decir, los mínimos)
valores de cada criterio de información, AIC = criterio de Akaike,
BIC = criterio bayesiano de Schwarz y HQC = criterio de Hannan-Quinn.

retardos  log.veros    p(RV)       AIC          BIC          HQC

   1     -179.04321             5.363571*    5.557842*    5.440645*
   2     -177.83141  0.65837    5.444389     5.768172     5.572845
   3     -172.23670  0.02452    5.398165     5.851462     5.578003
   4     -170.13875  0.38014    5.453297     6.036108     5.684518
   5     -169.22831  0.76866    5.542850     6.255173     5.825452
   6     -168.66458  0.88989    5.642452     6.484289     5.976437
   7     -162.76281  0.01887    5.587328     6.558678     5.972695
   8     -160.86888  0.43548    5.648373     6.749238     6.085123
```

Modelo VAR (1)

¿Condicionó el bono US al bono Australiano? Estimación

IAU no afecta, por tanto IAU no G-Causa IUS

```
Sistema VAR, orden del retardo 1
estimaciones de MCO, observaciones 1972:2-1991:1 (T = 76)
                                         Log-verosimilitud = -193.8778
Determinante de la matriz de covarianzas = 0.56343299
                             AIC = 5.2599
BIC = 5.4439
HQC = 5.3335
Contraste Portmanteau: LB(19) = 89.3597, gl = 72 [0.0809]

Ecuación 1: IAU

             Coeficiente   Desv. Típica   Estadístico
  -----------------------------------------------------
  const       0.676113      0.412725         1.638
  IAU_1       0.871084      0.0417194       20.88         1.66e-032 ***
  IUS_1       0.0962053     0.0470823        2.043        0.0446    **

Media de la vble. dep.  11.49566   D.T. de la vble. dep.   2.630867
Suma de cuad. residuos  44.48981   D.T. de la regresión    0.780672
R-cuadrado              0.914296   R-cuadrado corregido    0.911948
F(2, 73)                389.3840   Valor p (de F)          1.13e-39
rho                     0.021312   Durbin-Watson           1.943841

Contrastes F de restricciones cero:

Todos los retardos de IAU          F(1, 73) =   435.96 [0.0000]
Todos los retardos de IUS          F(1, 73) =   4.1753 [0.0446]

]
```

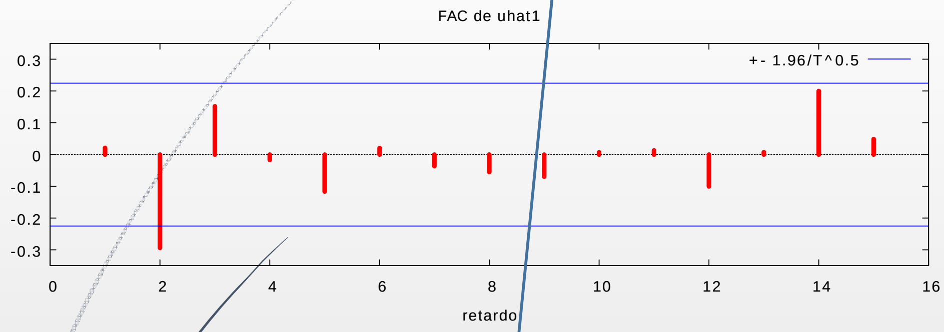
Afectan los das dos variables

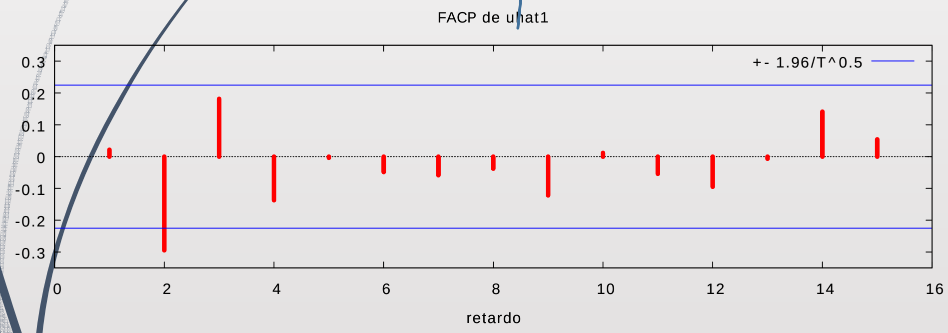
```
Ecuación 2: IUS

             Coeficiente   Desv. Típica   Estadístico t    Valor p
  --------------------------------------------------------------------
              1.10912       0.549005          2.020        0.0470

             -0.0212371     0.0554949        -0.3827       0.7031
        _     0.907350      0.0626286        14.49         3.62e-023
***

Media de la vble. dep.   9.091842   D.T. de la vble. dep.   2.382760
Suma de cuad. residuos   78.72112   D.T. de la regresión    1.038447
R-cuadrado               0.815129   R-cuadrado corregido    0.810064
F(2, 73)                 160.9346   Valor p (de F)          1.74e-27
rho                     -0.146952   Durbin-Watson           2.291397

Contrastes F de restricciones cero:

Todos los retardos de IAU          F(1, 73) =  0.14645 [0.7031]
Todos los retardos de IUS          F(1, 73) =   209.90 [0.0000
```

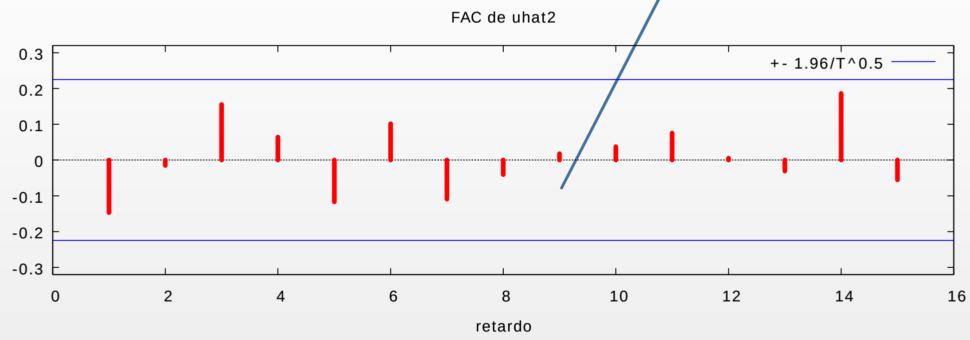
¿Condicionó el bono US al bono Australiano? diagnosis

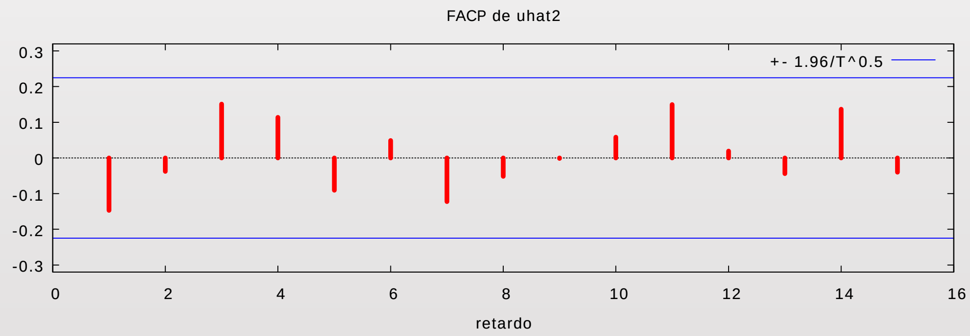
Posible MA(2)

Ruido blanco. Como esta era la ecuación que analizamos puede aceptarse el test realizado









Test de exogeneidad de Hausmann-Wu : idea

- Se estima el modelo por OLS y por IV.
- El estimador OLS
  - Si todas las variables son exógenas, es insesgado, consistente y eficiente
  - Si alguna variable no es exógena, es inconsistente
- El estimador IV es siempre consistente
- Calculamos la diferencia de estimadores $\delta = \hat{\beta}_{OLS} - \hat{\beta}_{IV}$
  - Cuando todas las variables son exógenas, d tiende a cero, puesto que ambos estimadores son consistentes, convergen al mismo parámetro
  - Cuando alguna no es exógena, d se aleja de 0 puesto que el estimador OLS es inconsistente

Test de Hausman-Wu: contraste

- Hipótesis
  H_0 : $\quad Todas\ las\ variables\ son\ exogenas$
  H_1: $existe\ alguna\ variable\ Z, dentro\ de\ las\ independientes, que\ no\ es\ exogena$
- Estadístico

$$HW = \delta' cov(\delta)^{-1}\delta$$

siendo $cov(\delta)$ la matriz de varianzas covarianzas de la diferencia de los estimadores, donde

$$cov(\delta)=\mathrm{cov}(\hat{\beta}_{OLS}) - cov(\hat{\beta}_{IV})$$

- Ley de distribución:
  - seguirá aproximadamente una chi cuadrado con r grados de libertada, siendo r la dimensión de Z, las posibles variables endógenas..

Exogeneidad del bono US

Se rechaza, pero no el de instrumento débil; por lo que hay una sospecha de que no es exógena, Cambiamos de instrumento

```
Modelo 2: MC2E, usando las observaciones 1972:1-1991:1 (T = 77)
Variable dependiente: IAU
Mediante Instrumentos: IUS
Instrumentos: const E

             Coeficiente   Desv. Típica      z       Valor p
  --------------------------------------------------------
  const       -38.2023       34.9097       -1.094    0.2738
  IUS           5.48394       3.85528       1.422    0.1549

Media de la vble. dep.  11.41805   D.T. de la vble. dep.   2.700766
Suma de cuad. residuos  10422.30   D.T. de la regresión   11.78830
R-cuadrado              0.367308   R-cuadrado corregido   0.358873
F(1, 75)                2.023357   Valor p (de F)         0.159042
Log-verosimilitud      -491.5412   Criterio de Akaike     987.0824
Criterio de Schwarz     991.7700   Crit. de Hannan-Quinn  988.9574
rho                     0.879681   Durbin-Watson          0.229714
```

```
Contraste de Hausman -
  Hipótesis nula: Los estimadores de MCO son consistentes
  Estadístico de contraste asintótico: Chi-cuadrado(1) =
129.953
  con valor p  = 4.1951e-030

Contraste de Instrumento débil -
  Estadístico F de la primera etapa (1, 75) = 1.64006
  Valores críticos para el tamaño maximal deseado de MC2E,
cuando los contrastes se ejecutan a un nivel de
significación nominal del 5%:

  tamaño      10%      15%      20%      25%
    valor    16.38     8.96     6.66     5.53

  El tamaño maximal puede ser superior a 25%
```

Exogeneidad del bono US

Se rechazan ambos test, por lo que seguramente no es exógena, hay interacción entre ambos bonos

```
Modelo 3: MC2E, usando las observaciones 1972:2-1991:1 (T = 76)
Variable dependiente: IAU
Mediante Instrumentos: IUS
Instrumentos: const IUS_1

             Coeficiente   Desv. Típica      z       Valor p
  ---------------------------------------------------------
  const       4.44024        1.07999       4.111    3.93e-05   ***
  IUS         0.776017       0.115637      6.711    1.94e-011  ***

Media de la vble. dep.   11.49566   D.T. de la vble. dep.    2.630867
Suma de cuad. residuos   343.3017   D.T. de la regresión     2.153883
R-cuadrado               0.350878   R-cuadrado corregido     0.342107
F(1, 74)                 45.03478   Valor p (de F)           3.39e-09
Log-verosimilitud       -453.6987   Criterio de Akaike       911.3975
Criterio de Schwarz      916.0590   Crit. de Hannan-Quinn    913.2604
rho                      0.873442   Durbin-Watson            0.220589
```

```
Contraste de Hausman -
  Hipótesis nula: Los estimadores de MCO son consistentes
  Estadístico de contraste asintótico: Chi-cuadrado(1) =
6.85274
  con valor p  = 0.00885054

Contraste de Instrumento débil -
  Estadístico F de la primera etapa (1, 74) = 325.477
  Valores críticos para el tamaño maximal deseado de MC2E,
cuando
   los contrastes se ejecutan a un nivel de significación
nominal del 5%:

  tamaño      10%      15%      20%      25%
    valor    16.38     8.96     6.66     5.53

  El tamaño maximal probablemente es menor que 10%
```

Problemas en la G-causalidad

- La G-causalidad presupone que las variables son estacionarias, puesto que si no los test no serian válidos.
- En general cuando tenemos dos procesos aleatorios X_t e Y_t, su relación a lo largo del tiempo podría escribirse como

$$\begin{pmatrix} \alpha(L) & \beta(L) \\ \gamma(L) & \delta(L) \end{pmatrix} \begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

donde a(L), b(L), g(L) y d(L) son polinomios de retardos en L.

- La inexistencia de causalidad implicaría que $\beta(L)=\gamma(L)=0$ y $cov(u_{1t},u_{2t})=0$
- Sin embargo en la estimación práctica existen dificultades para analizar la causalidad si los procesos no son estacionarios.

Dificultades en el estudio de la causalidad

- Supongamos que ambas X_t e Y_t son ARIMA(P_i,d,0), i=X,Y teniendo el mismo orden de diferenciación d.
- Aunque $\beta(L)=\gamma(L)=0$, las series mantienen su tendencia.
- Si consideramos que $Y_t = \mu + \beta X_t + u_t$
- Bajo la suposición de que $\beta(L)=\gamma(L)=0$, el estimador de β debería ser cercano a 0. Sin embargo, al hacer una regresión OLS de Y_t sobre X_t se obtiene que $plim\ \hat{\beta} = \frac{\mu_1}{\mu_2}\frac{\delta(1)}{\alpha(1)} \neq 0\ si\ \mu_1 \neq 0\ y\ \delta(1) \neq 0$
- Por tanto si Y tiene deriva y X tiene algún termo diferente de 0 entonces siempre aparecerá efecto en la regresión, aun siendo espurio.
- De hecho, siempre se verifica que plim $R^2=1$; plim DW=0; plim $\sigma^2=\infty$

Consumo-Renta Galicia

R2 cercano a 1
y DW muy bajo

```
Modelo 2: MCO, usando las observaciones 1995:1-2011:4 (T = 68)
Variable dependiente: Gastoenconsumofinal

             Coeficiente    Desv. Típica   Estadístico t   Valor p
  -----------------------------------------------------------------
  const      813142          56947.8            14.28      7.34e-022 ***
  PIBpm           0.759920       0.00515964     147.3      7.85e-085 ***

Media de la vble. dep.   8875709   D.T. de la vble. dep.    2332208
Suma de cuad. residuos  1.11e+12   D.T. de la regresión    129418.5
R-cuadrado              0.996967   R-cuadrado corregido    0.996921
F(1, 66)                21691.88   Valor p (de F)          7.85e-85
Log-verosimilitud      -895.8877   Criterio de Akaike      1795.775
Criterio de Schwarz     1800.214   Crit. de Hannan-Quinn   1797.534
rho                     0.327597   Durbin-Watson           1.334698
```

Cointegracion

Cointegracion

- El concepto de cointegracion trata de afrontar el problema planteado previamente, buscando dilucidar si realmente existe relación entre las variables.

Si X_t e Y_t son I(1) pero existe una combinación lineal, digamos

$$Z_t = m + aX_t + bY_t$$

tal que Z_t es I(0), entonces X_t, Y_t se dice que están cointegradas.

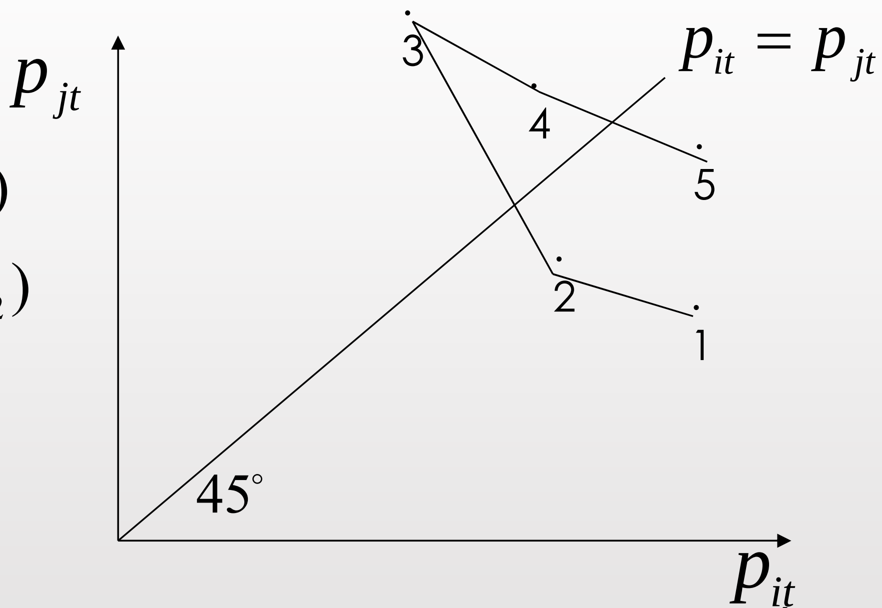
- La idea de la cointegracion es que las dos serie comparten la misma tendencia estocástica y por tanto se pueden integrar conjuntamente

Que es un ATRACTOR?

Interpretación Geometrica de Cointegración

Considere el precio (en el tiempo) de un bien que se comercia en dos localidades diferentes *i y j.*

$$1:(p_{i1}, p_{j1})$$
$$2:(p_{i2}, p_{j2})$$
$$\vdots$$
$$t:(p_{it}, p_{jt})$$



Supongamos $p_{i1} > p_{j1}$ $\longrightarrow$ Demanda irá al lugar *j*

$$\downarrow p_{i1} \text{ y} \uparrow p_{j1}$$

El ajuste no será instantaneo; pero en el largo plazo

$$p_{it} = p_{jt}$$

Interpretación Geométrica de Cointegración(cont)

El concepto de ***atractor*** es el concepto de ***long-run equilibrium*** entre dos procesos estocásticos. Permitimos que las dos variables diverjan en el corto plazo; pero en el largo plazo tienen que converger a una región común. En otras palabras, si no hubiera ningún shock en el sistema de aquí al futuro, los dos procesos estocásticos convergerían a un conjunto atractor común.

Cuestión 1: Escribe en términos intuitivos dos ejemplos económicos donde cointegración está presente. Por qué?

Cuestión 2: Un borracho saliendo de un bar sigue un paseo aleatorio. Su perro sigue otro paseo aleatorio por su cuenta. Llegan a un parque donde no están permitidos los perros sueltos. El borracho le pone la correa al perro y los dos entran en el parque. Estarán sus caminatas o sendas dentro del parque cointegradas? Por qué?

Significado de la cointegración

Descomponemos las series en una parte I(1) y una parte I(0)

$$X_t \equiv AW_t + \tilde{X}_t \qquad I(1) = I(1) + I(0)$$

$$Y_t \equiv W_t + \tilde{Y}_t$$

En ese caso, se puede hacer la siguiente combinación lineal

$$Z_t \equiv X_t - AY_t = AW_t + \tilde{X}_t - AW_t - A\tilde{Y}_t$$

$$Z_t \equiv \tilde{X}_t - A\tilde{Y}_t \rightarrow I(0)$$

por lo que X_t, Y_t tienen un factor I(1) común.

Resultado

Si dos series I(1) tienen un factor común I(1) y un componente I(0) idiosincrásico, entonces ellas están cointegradas.

Por tanto, en ese sentido, dos series cointegradas tienen una tendencia estocástica común

Aplicaciones

- ***Teoría de la paridad del poder adquisitivo (PPP)***
  - Paridad del nivel de vida ajustado a largo plazo
- ***Modelo del Valor Presente Actualizado (PVM)***
  - Intereses a largo plazo y a corto plazo

Test de cointegración

- Vamos a considerar dos tipos de test
  - Para series bivariantes analizamos el test de Engle y Granger, en sus diversas modalidades
    - Sin constante
    - Con constante
    - Con tendencia y constante
  - Para series multivariantes, analizamos el test de Johansen, que permite varias opciones según se introduzca constante o tendencia y según las restricciones que se impongan a la constante o a la tendencia
- Antes de estudiar la cointegracion multivariante, desarrollamos el concepto de corrección de error con la bivariante

Test de Engle-Granger

- Modelo

  Se realiza una regresión en el siguiente modelo $y_t = \alpha + \beta x_t + z_t \; ; \quad \mathrm{y}_t, \mathrm{x}_t \propto \mathrm{I}(1)$

  Se calcula una regresión sobre los residuos

$$\Delta \hat{z}_t = f(t) + \phi \hat{z}_{t-1} + \sum_{i=1}^{p} \lambda_i \Delta \hat{z}_{t-i} + error$$

  Si ϕ es igual a cero entonces los residuos tienen raíces unitarias y las series no están cointegradas

- Hipótesis $H_0: \phi = 0$ $H_1: \phi < 0$
- Estadístico

$$Z_\phi = \frac{\hat{\phi}}{s_\phi}$$

  Siendo sd(ϕ_i) calculado en la regresión sobre la serie, y para T grande se aproxima por 1/T,

- Ley de distribución:
  - seguirá una ley propia tabulada por Mc Kinnon

Regresiones entre variables económicas

- **Regresión por niveles**

$$y_t = \alpha + \beta x_t + z_t \; ; \quad \mathrm{y}_t \, , \mathrm{x}_t \propto \mathrm{I}(1)$$

  - Si las series son cointegradas los residuos son I(0), existe relación entre las variables. En este caso se puede analizar el modelo de corrección de error.
  - β mide el cambio
  - Si las series no son cointegradas los residuos son I(1), la relación es espuria
- **Regresión en diferencias**

$$\Delta y_t = \alpha + \beta \Delta x_t + u_t \; ; \quad \Delta \mathrm{y}_t \, , \Delta \mathrm{x}_t \propto \mathrm{I}(0)$$

  - Siempre existe.
  - β mide la velocidad del cambio

Modelo de corrección de error (MCE): elementos

- El trabajar con series I(1) e I(0) permite analizar separadamente los dos tipos de regresiones anteriores o combinarlos en una sola.
- La idea del MCE es combinar ambas regresiones,
- Para ello, se realiza la regresión por niveles y se introducen los residuos de la regresión por niveles en la regresión por diferencias

$$\Delta y_t = \alpha + \rho z_t + \beta \Delta x_t + u_t \; ; \quad \mathrm{z_t}, \Delta \mathrm{y_t} \, , \Delta \mathrm{x_t} \propto \mathrm{I}(0)$$

- De esa forma, β sigue midiendo la velocidad de cambio del efecto, pero un nuevo parámetro ρ va a medir el efecto de la relación entre las variables consideradas por niveles.

Estimación del MCE

- Metodo de dos etapas de Engle-Granger:
  - (i) Estimar el valor de Z a partir de la regression por niveles
  - (ii) Introducir   el residuo estimado Z en la regression por diferencias

$$\Delta Y_t = \mu_2 + \rho_2(Y_{t-1} - \hat{\alpha}X_{t-1}) + \gamma_{x1}\Delta X_{t-1} + \ldots.. + \gamma_{y1}\Delta Y_{t-1} + \ldots. + \varepsilon_{yt}$$

- Los estimadores OLS en el ECM son consistentes y eficientes.

Consumo renta Galicia

Hay evidencia de una relación cointegrante si:

(a) La hipótesis de existencia de raíz unitaria no se rechaza para las variables individuales y

(b) La hipótesis de existencia de raíz unitaria se rechaza para los residuos (uhat) de la regresión cointegrante.

En este caso se rechazan ambas. NO existe cointegracion

Etapa 1: contrastando la existencia de una raíz unitaria en Gastoenconsumofinal

Contraste aumentado de Dickey-Fuller para Gastoenconsumofinal
incluyendo 4 retardos de (1-L)Gastoenconsumofinal
tamaño muestral 63
hipótesis nula de raíz unitaria: a = 1

contraste con constante
modelo: (1-L)y = b0 + (a-1)*y(-1) + ... + e
valor estimado de (a - 1): -0.00849415
Estadístico de contraste: tau_c(1) = -1.31142
valor p asintótico 0.6266
Coef. de autocorrelación de primer orden de e: -0.028
diferencias retardadas: F(4, 57) = 250.334 [0.0000]

Etapa 2: contrastando la existencia de una raíz unitaria en PIBpm

Contraste aumentado de Dickey-Fuller para PIBpm
incluyendo 4 retardos de (1-L)PIBpm
tamaño muestral 63
hipótesis nula de raíz unitaria: a = 1

contraste con constante
modelo: (1-L)y = b0 + (a-1)*y(-1) + ... + e
valor estimado de (a - 1): -0.0088717
Estadístico de contraste: tau_c(1) = -1.39747
valor p asintótico 0.5852
Coef. de autocorrelación de primer orden de e: 0.165
diferencias retardadas: F(4, 57) = 202.222 [0.0000]

Etapa 3: regresión cointegrante

Regresión cointegrante -
MCO, usando las observaciones 1995:1-2011:4 (T = 68)
Variable dependiente: Gastoenconsumofinal

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	813142	56947.8	14.28	7.34e-022	***
PIBpm	0.759920	0.00515964	147.3	7.85e-085	***

Media de la vble. dep.	8875709	D.T. de la vble. dep.	2332208
Suma de cuad. residuos	1.11e+12	D.T. de la regresión	129418.5
R-cuadrado	0.996967	R-cuadrado corregido	0.996921
Log-verosimilitud	-895.8877	Criterio de Akaike	1795.775
Criterio de Schwarz	1800.214	Crit. de Hannan-Quinn	1797.534
rho	0.327597	Durbin-Watson	1.334698

Etapa 4: contrastando la existencia de una raíz unitaria en uhat

Contraste aumentado de Dickey-Fuller para uhat
incluyendo 4 retardos de (1-L)uhat
tamaño muestral 63
hipótesis nula de raíz unitaria: a = 1

modelo: (1-L)y = (a-1)*y(-1) + ... + e
valor estimado de (a - 1): -0.263054
Estadístico de contraste: tau_c(2) = -2.27483
valor p asintótico 0.3855
Coef. de autocorrelación de primer orden de e: -0.159
diferencias retardadas: F(4, 58) = 16.294 [0.0000]

PIB y Masa monetaria

```
Etapa 1: contrastando la existencia de una raíz unitaria en m_p

Contraste aumentado de Dickey-Fuller para m_p
incluyendo 2 retardos de (1-L)m_p
tamaño muestral 76
hipótesis nula de raíz unitaria: a = 1

  contraste con constante
  modelo: (1-L)y = b0 + (a-1)*y(-1) + ... + e
  valor estimado de (a - 1): 0.000178852
  Estadístico de contraste: tau_c(1) = 0.0657734
  valor p asintótico 0.9632
  Coef. de autocorrelación de primer orden de e: 0.016
  diferencias retardadas: F(2, 72) = 8.897 [0.0004]

Etapa 2: contrastando la existencia de una raíz unitaria en y

Contraste aumentado de Dickey-Fuller para y
incluyendo 2 retardos de (1-L)y
tamaño muestral 76
hipótesis nula de raíz unitaria: a = 1

  contraste con constante
  modelo: (1-L)y = b0 + (a-1)*y(-1) + ... + e
  valor estimado de (a - 1): -1.17457e-006
  Estadístico de contraste: tau_c(1) = -0.000254784
  valor p asintótico 0.9575
  Coef. de autocorrelación de primer orden de e: -0.029
  diferencias retardadas: F(2, 72) = 2.810 [0.0668]
```

Hay evidencia de una relación cointegrante si:
(a) La hipótesis de existencia de raíz unitaria no se rechaza para las variables individuales y
(b) La hipótesis de existencia de raíz unitaria se rechaza para los residuos (uhat) de la regresión cointegrante.
En este caso hay evidencia de cointegracion

```
Etapa 3: regresión cointegrante

Regresión cointegrante -
MCO, usando las observaciones 1980:1-1999:3 (T = 79)
Variable dependiente: m_p

             Coeficiente   Desv. Típica   Estadístico t    Valor p
  ----------------------------------------------------------------
  const       -1.94380      0.0627595        -30.97       3.22e-045 ***
  y            1.42091      0.0140408        101.2        1.16e-083 ***

Media de la vble. dep.   4.404742   D.T. de la vble. dep.    0.185278
Suma de cuad. residuos   0.019981   D.T. de la regresión     0.016109
R-cuadrado               0.992537   R-cuadrado corregido     0.992441
Log-verosimilitud        215.0586   Criterio de Akaike      -426.1172
Criterio de Schwarz     -421.3783   Crit. de Hannan-Quinn   -424.2187
rho                      0.847329   Durbin-Watson            0.271065

Etapa 4: contrastando la existencia de una raíz unitaria en uhat

Contraste aumentado de Dickey-Fuller para uhat
incluyendo 2 retardos de (1-L)uhat
tamaño muestral 76
hipótesis nula de raíz unitaria: a = 1

  modelo: (1-L)y = (a-1)*y(-1) + ... + e
  valor estimado de (a - 1): -0.210015
  Estadístico de contraste: tau_c(2) = -3.40989
  valor p asintótico 0.04132
  Coef. de autocorrelación de primer orden de e: -0.014
  diferencias retardadas: F(2, 73) = 4.588 [0.0133]
```

Modelo de Corrección del Error vectorial (VECM)

Para un VAR bivariante, donde X_t, Y_t son I(1) y estan cointegradas

$$\Delta X_t = c_1 + \rho_1 Z_{t-1} + \beta_1 \Delta X_{t-1} + \ldots.. + \alpha_1 \Delta Y_{t-1} + \ldots. + \varepsilon_{xt}$$
$$\Delta Y_t = c_2 + \rho_2 Z_{t-1} + \gamma_1 \Delta X_{t-1} + \ldots.. + \delta_1 \Delta Y_{t-1} + \ldots. + \varepsilon_{yt}$$

donde $(\varepsilon_{xt}, \varepsilon_{yt})'$ es un ruido blanco bivariante y

$Z_t = X_t - AY_t \rightarrow I(0)$, y

como minimo un $\rho_i \neq 0$

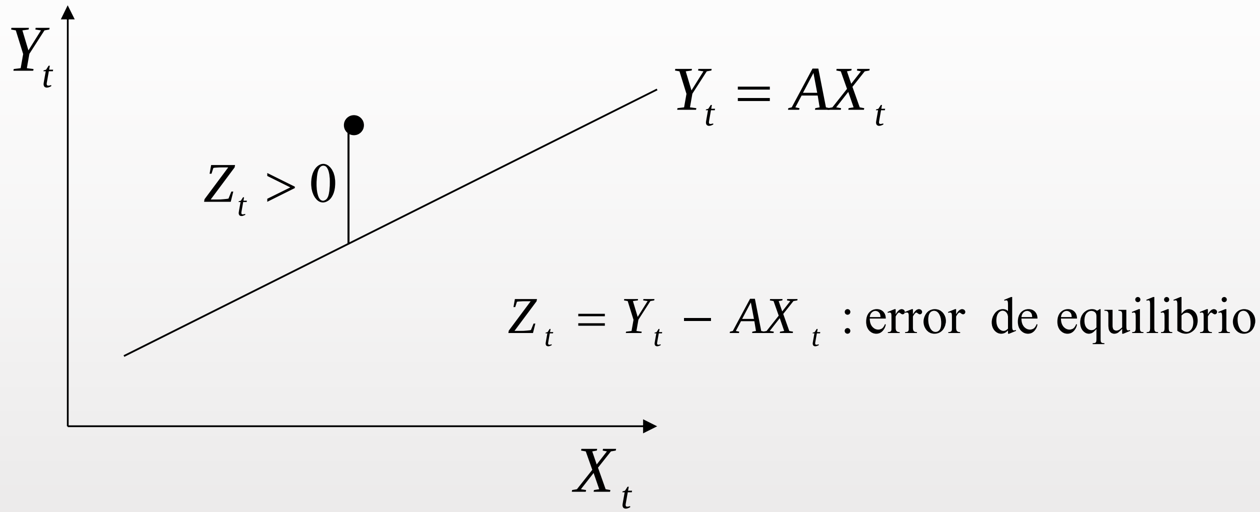
Puesto que si X_t, Y_t no están cointegradas $\Rightarrow Z_t \rightarrow I(1)$, entonces
en el ECM, I(1) no puede explicar I(0), i.e. $\Delta X_t, \Delta Y_t \Rightarrow \rho_1 = \rho_2 = 0$

Resultado

Si X_t, Y_t están cointegradas, entonces existe una representacion ECM y viceversa (Teorema de Representación de Granger).

Intuición Geométrica del Modelo de Corrección del Error

Intuición sobre el ECM



Donde vaya el sistema en el tiempo *t+1* , depende de la magnitud y el signo del error de equilibrio en el periodo anterior.

<u>Dinámica de Corto-Plazo</u>: movimientos en el corto plazo dentro del ECM, que guian a la economía hacia el <u>Equilibrio de Largo Plazo</u> $Y_t = AX_t$

PIB y Masa monetaria
Estimacion

En la regresión por diferencias cointegrada, la demanda monetaria depende solo de su pasado y se corrige en un 7%

Ecuación 1: d_m_p

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	-0.150513	0.0584590	-2.575	0.0121	**
d_m_p_1	0.484411	0.0975595	4.965	4.36e-06	***
d_y_1	-0.0631421	0.0899765	-0.7018	0.4851	
EC1	-0.0787598	0.0297416	-2.648	0.0099	***

Media de la vble. dep.	0.007701	D.T. de la vble. dep.	0.004657
Suma de cuad. residuos	0.001172	D.T. de la regresión	0.004007
R-cuadrado	0.288962	R-cuadrado corregido	0.259741
rho	-0.048926	Durbin-Watson	2.069429

Ecuación 2: d_y

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	0.151562	0.0706132	2.146	0.0352	**
d_m_p_1	0.172513	0.117843	1.464	0.1475	
d_y_1	0.282580	0.108683	2.600	0.0113	**
EC1	0.0758173	0.0359252	2.110	0.0382	**

Media de la vble. dep.	0.005166	D.T. de la vble. dep.	0.005145
Suma de cuad. residuos	0.001710	D.T. de la regresión	0.004840
R-cuadrado	0.150029	R-cuadrado corregido	0.115098
rho	-0.095889	Durbin-Watson	2.184165

El modelo estima las os variables puesto que en cointegracion existe una tendencia común dada por la regresión cointegrante previa que representa la regresión por niveles

Matriz de covarianzas cruzadas entre ecuaciones:

	m_p	y
m_p	1.5221e-005	7.9000e-007
y	7.9000e-007	2.2208e-005

determinante = 3.37405e-010

Regresión cointegrante -
MCO, usando las observaciones 1980:1-1999:3 (T = 79)
Variable dependiente: m_p

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	-1.94380	0.0627595	-30.97	3.22e-045	***
y	1.42091	0.0140408	101.2	1.16e-083	***

The Johansen ML Procedure

- This is based on a VAR approach to cointegration
- All the variables are assumed to be endogenous (although it is possible to include exogenous variables)
- The test relies on the relationship between the rank of a matrix and its eigenvalues or characteristic roots.

You do not need to understand the mechanics of this approach, just how to use it and how to interpret the results

Johansen ML Approach

- The approach to testing for cointegration in a multivariate system is similar to the ADF test, but requires the use of a VAR approach:

$$x_t = A_1 x_{t-1} + u_t$$

$$\Delta x_t = (A_1 - I)x_{t-1} + u_t$$

$$\Delta x_t = \pi x_{t-1} + u_t$$

$$Where : \pi = (A_1 - I)$$

- Where in a system of *g* variables:

x_t and u_t are g x 1 vectors.

A_1 is an g x g matrix of parameters

I is an g x g Identity matrix

Johansen ML Approach

- The rank of π equals the number of cointegrating vectors, indicating common trends
- If π consists of all zeros, as with the ADF test, the rank of the matrix equals zero, all of the *xs* are unit root processes, implying the variables are not cointegrated.
- As with the ADF test, the equation can also include lagged dependent variables, although the number of lags included is important and can affect the result. This requires the use of the Akaike or Schwarz-Bayesian criteria to ensure an optimal lag length.

Main Differences with the Bi-Variate Test for Cointegration

- Using the Johansen Maximum Likelihood (ML) procedure, it is possible to obtain more then a single cointegrating relationship, whereas only one can be obtained with the Engle-Granger test.
- If there is evidence of more than one cointegrating relationship, which one should be used with the Johansen test?
- There are two separate tests for cointegration with the Johansen, but only one with the Engle-Granger which can give different results.
- Given that the Johansen is a maximum likelihood based test (Engle-Granger is OLS based), it requires a large sample.
- The multivariate test is based on a VAR, not a single OLS estimation as with the Engle-Granger approach.

The π Matrix

- As mentioned, *r* is the rank of π and determines the number of cointegrating vectors.
- When r = 0 there are no cointegrating vectors
- If there are *g* variables in the system of equations, there can be a maximum of g-1 cointegrating vectors.
- Π is defined as the product of two matrices: a and β' , of dimension (g x r) and (r x g) respectively.
  - The β gives the long-run coefficients of the cointegrating vectors, the α is known as the adjustment parameter and is similar to an error correction term. The relationship can be expressed as:

$$\Pi = \alpha \beta '$$

Test Statistics

- There are two test statistics produced by the Johansen ML procedure.
- There are the Trace test and maximal Eigenvalue test.
- Both can be used to determine the number of cointegrating vectors present, although they don't always indicate the same number of cointegrating vectors.
- The Trace test is a joint test, the null hypothesis is that the number of cointegrating vectors is less than or equal to r, against a general alternative hypothesis that there are more then r.
- The Maximal Eigenvalue test conducts separate tests on each eigenvalue. The null hypothesis is that there are r cointegrating vectors present against the alternative that there are $(r + 1)$ present.
- The distribution of both test statistics is non-standard.

Multivariate Cointegration and VECMs

- Vector Error Correction Models (VECM) are the basic VAR, with an error correction term incorporated into the model and as with bivariate cointegration, multivariate cointegration implies an appropriate VECM can be formed.
- The reason for the error correction term is the same as with the standard error correction model, it measures any movement away from the long-run equilibrium.
- These are often used as part of a multivariate test for cointegration, such as the Johansen ML test, having found evidence of cointegration of some I(1) variables, we can then assess the short run and potential Granger causality with a VECM.

Precios, tipo de cambio y bonos

```
Contraste de Johansen:
Número de ecuaciones = 5
Orden del retardo = 4
Periodo de estimación: 1973:1 - 1987:2 (T = 58)
Caso 3: constante no restringida

Log-verosimilitud = 1091.96 (Incluyendo un término constante:
927.361)

Rango Valor propio Estad. traza  Valor p  Estad. Lmáx  Valor p
   0     0.58839      123.75 [0.0000]      51.486 [0.0000]
   1     0.38681      72.267 [0.0000]      28.366 [0.0362]
   2     0.27590      43.901 [0.0005]      18.724 [0.1069]
   3     0.25112      25.177 [0.0010]      16.772 [0.0177]
   4     0.13491      8.4053 [0.0037]      8.4053 [0.0037]

Corregido por el tamaño muestral (gl = 37)
Rango Estad. traza Valor p
   0     123.75 [0.0000]
   1     72.267 [0.0002]
   2     43.901 [0.0010]
   3     25.177 [0.0014]
   4     8.4053 [0.0052]

valor propio     0.58839      0.38681      0.27590      0.25112
0.13491
```

beta renormalizado

p1	1.0000	-0.91894	-0.27632	-0.91628	0.087826
p2	-1.2461	1.0000	-0.17758	1.8757	-0.40316
e12	-0.69511	0.20166	1.0000	-0.81291	0.23156
i1	-0.96809	-0.84003	4.4037	1.0000	2.1677
i2	-2.4153	4.1999	-3.1330	-3.2917	1.0000

alfa renormalizado

p1	-0.10872	0.066675	0.022530	-0.012390	-0.0097752
p2	0.071527	0.089729	0.041378	-0.0093785	0.045146
e12	0.099865	0.17795	-0.010100	0.17772	-0.10327
i1	0.11834	0.023797	0.0080935	-0.047319	-0.066851
i2	0.0055687	0.020373	0.072387	0.040304	-0.041610

matriz de largo plazo (alfa * beta')

	p1	p2	e12	i1	i2
p1	-0.16572	0.17886	0.11936	0.11488	0.50305
p2	-0.0098045	-0.042541	0.027832	0.12608	0.15047
e12	-0.23278	0.43028	-0.21202	-0.33676	-0.15048
i1	0.13172	-0.18691	-0.046380	-0.29114	-0.12233
i2	-0.073739	0.092951	0.030226	0.24637	-0.32896

Criticisms of the Johansen Approach

- The result can be sensitive to the number of lags included in the test and the presence of autocorrelation
- If there are more than two cointegrating vectors present, how do we find the most appropriate vector for the subsequent tests.
- If the two test statistics differ, which one gives the correct result?
- This is a large sample test.
- The Wickens critique suggests we often find evidence of cointegration when none exists.

The Approach to Multivariate Cointegration and VECMs

1) Test the variables for stationarity using the usual ADF tests.
2) If all the variables are I(1) include in the cointegrating relationship.
3) Use the AIC or SBIC to determine the number of lags in the cointegration test (order of VAR)
4) Use the trace and maximal eigenvalue tests to determine the number of cointegrating vectors present.
5) Assess the long-run β coefficients and the adjustment a coefficients.
6) Produce the VECM for all the endogenous variables in the model and use it to carry out Granger causality tests over the short and long run.

Conclusion

- When there are more than two variables, we need to use the Johansen ML approach to test for cointegration
- There are two statistics to take into account; the trace and maximum eigenvalue.
- Depending on how many cointegrating vectors are present, we can then test for the short-run using a vector error correction model.