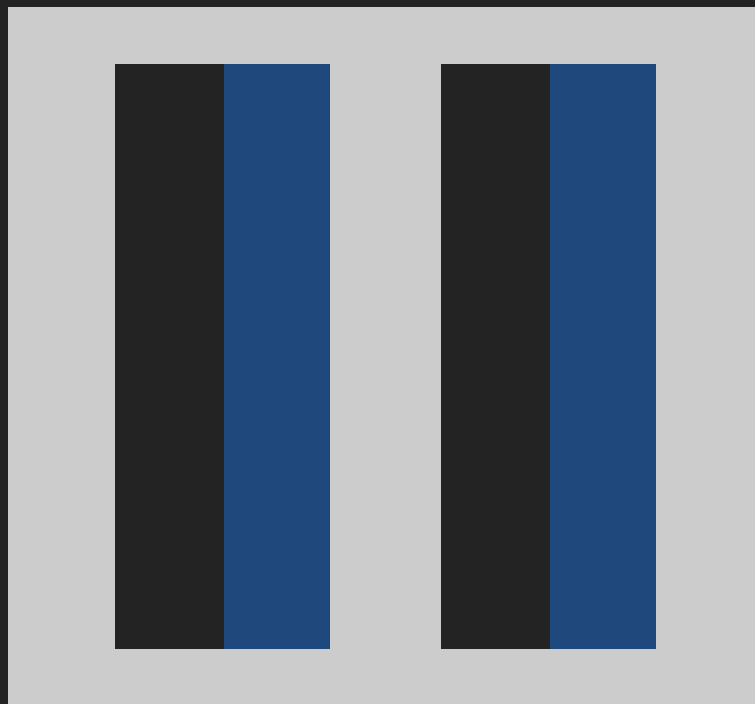


ECONOMETRÍA

JOSÉ ALBERTO MAURICIO



4

REGRESIÓN CON SERIES TEMPORALES

4.2 REGRESIÓN CON SERIES ESTACIONARIAS

EctrGr-JAM-4-2.pdf

Copyright © 2022 - 2024 J.A.M.

ucm.randomshock.com/ectrgr

Versión 2.4 - Enero 2024

REQUISITOS

En esta Sección 4.2 se suponen bien conocidos [1] los aspectos teóricos y prácticos del Análisis de Regresión Lineal cubiertos en los Temas 1 - 3 y en la Sección 4.1 del Tema 4, y [2] todos los procedimientos descritos en las Secciones 1 - 17 de la guía *Introducción al Uso de EViews 4.1*.

BIBLIOGRAFÍA PARA LA SECCIÓN 4.2



Hill, Griffiths, Lim (2018): Capítulo 9.

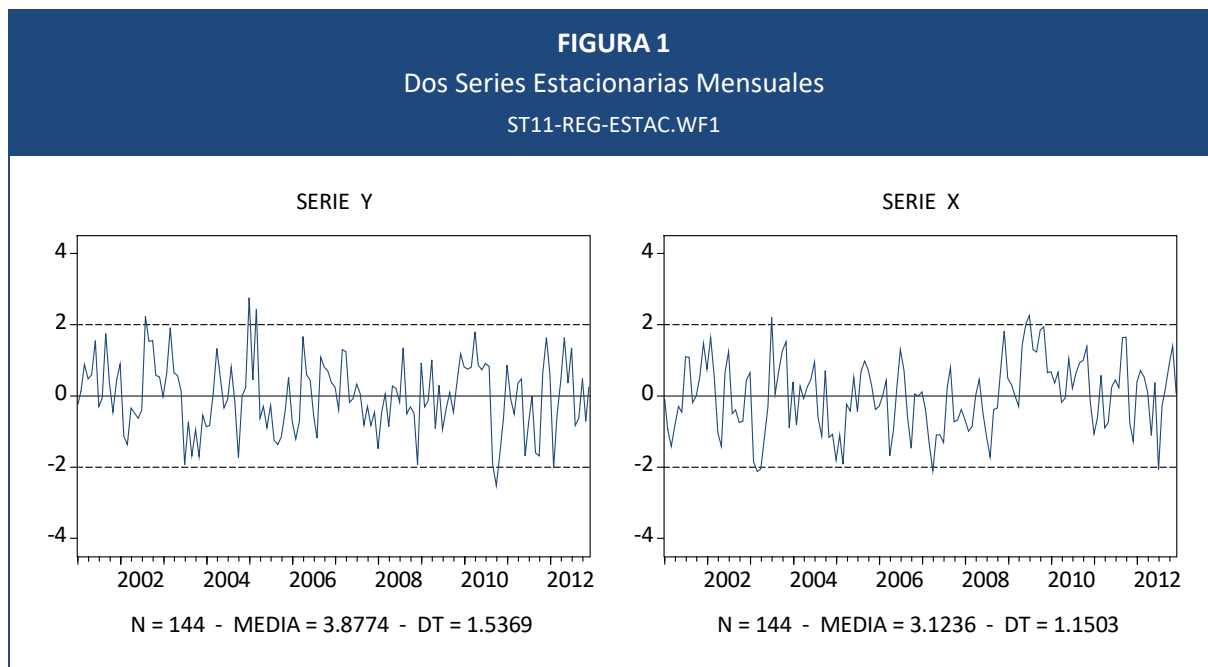
Wooldridge (2020): Capítulos 10 - 12.

Sección 18 de la guía *Introducción al Uso de EViews 4.1*.

En adelante IEV41 ...

PARTE 1 - AUTOCORRELACIÓN

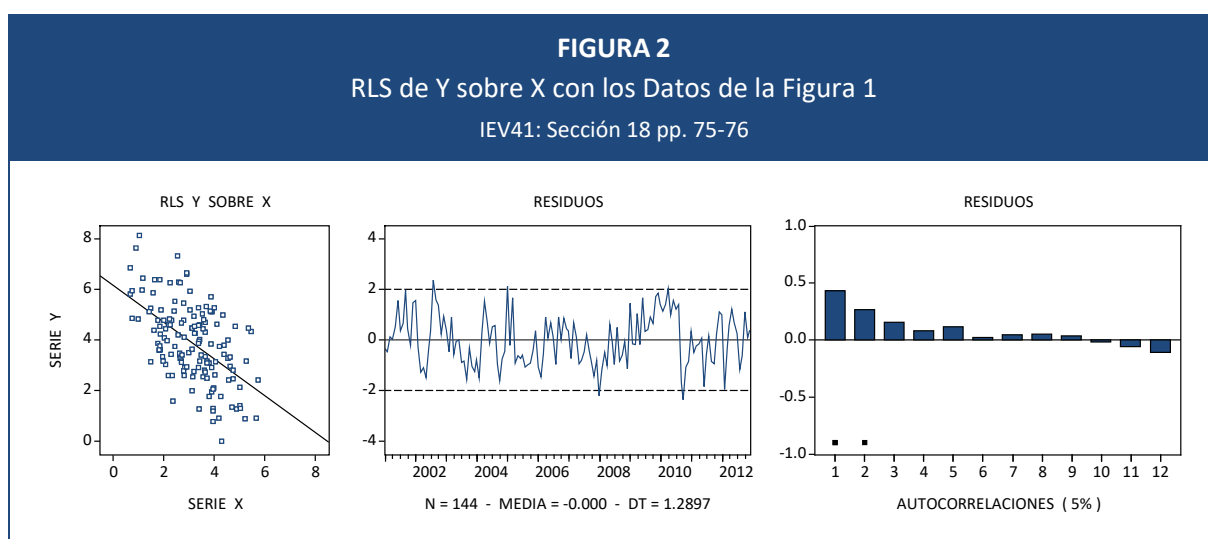
Con frecuencia, los **residuos** de un modelo de regresión estimado por MCO con series estacionarias presentan algún tipo de **dinámica** o **autocorrelación** que sugiere la existencia de **covarianzas distintas de cero** entre algunas **perturbaciones** del modelo.



$$y_t = \begin{matrix} 6.1475 \\ (0.3131) \\ [0.0000] \end{matrix} - \begin{matrix} 0.7267x_t \\ (0.0941) \\ [0.0000] \end{matrix} + \hat{u}_t,$$

[1]

$$N = 144, \hat{\sigma} = 1.2943, \bar{R}^2 = 0.2909.$$



La **autocorrelación** muestral de **orden** $k \geq 1$ de la serie de residuos \hat{u}_t en el último diagrama de la Figura 2 es la estimación MCO de la pendiente en la RLS de \hat{u}_t sobre \hat{u}_{t-k} . En un contexto temporal, la secuencia de autocorrelaciones (barras) en un diagrama de ese

tipo informa sobre el grado medio de dependencia (lineal) de cualquier residuo con respecto a su propio pasado (la dinámica, la inercia, o la memoria, que está presente en la evolución histórica de la serie de residuos).

Como en el caso de la heteroscedasticidad (cuando la dispersión no constante en los residuos se interpretaba como un síntoma de que la varianza de las perturbaciones podía no ser constante), la presencia de autocorrelaciones residuales significativas se puede interpretar como un síntoma de que las correlaciones entre las perturbaciones de un modelo no son todas iguales a cero.

⇒ Modelo RLS basado en las Figuras 1-2 con **autocorrelación de orden 2**:

$$Y_t = \beta_1 + \beta_2 X_t + U_t, \text{ con } E[U_t | \mathbf{X}] = 0, \text{ Var}[U_t | \mathbf{X}] = \sigma^2 \text{ (constante)}, \quad [2]$$

$$\text{Corr}[U_{t_1}, U_{t_2} | \mathbf{X}] = \text{Corr}[U_{t_2}, U_{t_1} | \mathbf{X}] = \begin{cases} \rho_1 \neq 0, & \text{si } |t_1 - t_2| = 1, \\ \rho_2 \neq 0, & \text{si } |t_1 - t_2| = 2, \\ 0, & \text{si } |t_1 - t_2| > 2, \end{cases} \quad [3]$$

con $t, t_1, t_2 = 1, \dots, N$. En este caso, la matriz de varianzas de $\mathbf{U} = [U_1, U_2, \dots, U_N]'$ es

$$\text{Var}[\mathbf{U} | \mathbf{X}] = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & 0 & \dots & 0 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \dots & 0 \\ \rho_2 & \rho_1 & 1 & \ddots & \ddots & \vdots \\ 0 & \rho_2 & \ddots & \ddots & \rho_1 & \rho_2 \\ \vdots & \vdots & \ddots & \rho_1 & 1 & \rho_1 \\ 0 & 0 & \dots & \rho_2 & \rho_1 & 1 \end{bmatrix} = \sigma^2 \mathbf{V} \neq \sigma^2 \mathbf{I},$$

donde ρ_1 y ρ_2 son las **autocorrelaciones** de órdenes 1 y 2 de \mathbf{U} .

$$[2]-[3] \Rightarrow E[Y_t | \mathbf{X}] = \beta_1 + \beta_2 X_t, \text{ Var}[Y_t | \mathbf{X}] = \sigma^2 \text{ (constante)}, \quad [4]$$

$$\text{Corr}[Y_{t_1}, Y_{t_2} | \mathbf{X}] = \text{Corr}[Y_{t_2}, Y_{t_1} | \mathbf{X}] \neq 0 \text{ si } |t_1 - t_2| = 1, 2. \quad [5]$$

En un contexto temporal, la causa de la autocorrelación es la existencia de algún grado de asociación (lineal) entre lo que ocurre con la variable dependiente en momentos distintos de su historia, es decir, la existencia de algún tipo de dinámica (inercia, memoria, dependencia con respecto a su propio pasado) en la evolución histórica de la variable dependiente. Por este motivo, la presencia de autocorrelación siempre se debe considerar cuando se pretende elaborar un modelo con datos de series temporales.

⇒ Modelo RLM con **autocorrelación general**:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{U}, \text{ con } E[\mathbf{U} | \mathbf{X}] = \mathbf{0}, \quad [6]$$

$$\text{Var}[\mathbf{U} | \mathbf{X}] (= E[\mathbf{U}\mathbf{U}' | \mathbf{X}]) = \sigma^2 \mathbf{V}, \quad [7]$$

donde \mathbf{V} es una matriz simétrica **no diagonal** del siguiente estilo (en la que se impone la misma varianza para todas las perturbaciones del modelo y se permite la existencia de autocorrelación entre ellas hasta el máximo orden posible):

$$\mathbf{V} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \cdots & \rho_{N-1} \\ \rho_1 & 1 & \rho_1 & \rho_2 & \cdots & \rho_{N-2} \\ \rho_2 & \rho_1 & 1 & \ddots & \ddots & \vdots \\ \rho_3 & \rho_2 & \ddots & \ddots & \rho_1 & \rho_2 \\ \vdots & \vdots & \ddots & \rho_1 & 1 & \rho_1 \\ \rho_{N-1} & \rho_{N-2} & \cdots & \rho_2 & \rho_1 & 1 \end{bmatrix} \neq \mathbf{I}. \quad [8]$$

Como en el caso de la heteroscedasticidad, reemplazando en el modelo RLM clásico la hipótesis de que $\text{Var}[\mathbf{U} | \mathbf{X}] = \sigma^2 \mathbf{I}$ por [7]-[8], se puede comprobar que

$$E[\hat{\beta}_{\mathbf{W}}] = \beta,$$

$$\text{Var}[\hat{\beta}_{\mathbf{W}} | \mathbf{X}] = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \neq \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}. \quad [9]$$

De nuevo, ningún resultado basado en la expresión $\sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$ para $\text{Var}[\hat{\beta}_{\mathbf{W}} | \mathbf{X}]$ es aplicable ahora (en particular, el Teorema de Gauss-Markov). Además, cualquier cálculo basado en la estimación $\hat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1}$ de $\text{Var}[\hat{\beta}_{\mathbf{W}} | \mathbf{X}]$ es **incorrecto** (en particular, los errores estándar habituales de los estimadores MCO, como los que figuran en [1]).

PARTE 2 - DETECCIÓN

⇒ Autocorrelaciones muestrales de los residuos (como en la Figura 2).

⇒ **Contraste de Breusch-Godfrey:**

1. Hipótesis nula (H_0): ausencia de autocorrelación en las perturbaciones.
2. Hipótesis alternativa (H_1): autocorrelación de orden $P \geq 1$ (con frecuencia, $P = 1$).
3. Estadístico: $BG_P = N \times R_P^2$, donde si \hat{u}_t es la serie de residuos MCO de [6], entonces R_P^2 es el coeficiente de determinación en la **regresión auxiliar** de \hat{u}_t sobre todas las variables explicativas de [6] y las series $\hat{u}_{t-1}, \dots, \hat{u}_{t-P}$.
4. Nivel de significación marginal: $\alpha^* = \Pr[\chi^2(P) \geq BG_P]$. Si α^* es suficientemente pequeño (menor que el nivel de significación escogido), rechazar H_0 en favor de H_1 .

CONTRASTE DE BREUSCH-GODFREY EN EL MODELO [1]

IEV41: Sección 18 pp. 76-77

Con $P = 1$ en H_1 , $BG_1 = 144 \times 0.18925$, donde 0.18925 es el coeficiente de determinación en la regresión con término constante de \hat{u}_t sobre x_t y \hat{u}_{t-1} . El nivel de significación marginal del contraste de Breusch-Godfrey es

$$\alpha^* = \Pr[\chi^2(1) \geq 27.252] = 0.0000.$$

PARTE 3 - UTILIZACIÓN ADECUADA DE MCO

El **estimador de Newey-West** es un estimador adecuado (robusto, consistente) de [9] en presencia de cualesquiera tipos de heteroscedasticidad y autocorrelación:

$$\text{Vár}[\hat{\beta}_{\mathbf{W}} | \mathbf{X}]_{\text{NW}} = \frac{1}{N} \hat{\mathbf{Q}}^{-1} \hat{\mathbf{S}}_{\text{NW}} \hat{\mathbf{Q}}^{-1}, \text{ con}$$

$$\hat{\mathbf{Q}} = \frac{1}{N} \mathbf{X}'\mathbf{X} = \frac{1}{N} \sum_{t=1}^N \mathbf{x}_t \mathbf{x}_t',$$

$$\hat{\mathbf{S}}_{\text{NW}} = \frac{1}{N} \sum_{t=1}^N \hat{u}_t^2 \mathbf{x}_t \mathbf{x}_t' + \sum_{h=1}^H \left(1 - \frac{h}{H+1}\right) \left(\frac{1}{N} \sum_{t=1}^{N-h} [\hat{u}_t \hat{u}_{t+h} (\mathbf{x}_t \mathbf{x}_{t+h}' + \mathbf{x}_{t+h} \mathbf{x}_t')]\right).$$

donde \mathbf{x}_i' es la fila i -ésima de la matriz \mathbf{X} .

En la expresión anterior para $\hat{\mathbf{S}}_{\text{NW}}$, el número $H \leq N - 1$ es tal que se espera o se supone que $\text{Cov}[U_{t_1}, U_{t_2}] = 0$ para $|t_1 - t_2| > H$. En la práctica, H se suele escoger como la parte de entera de $N^{1/4}$, o bien como la parte entera de $4 \times (N/100)^{2/9}$.

Un ajuste muy común del estimador $\hat{\mathbf{S}}_{\text{NW}}$ cuando se emplea con muestras cortas, consiste en multiplicar $\hat{\mathbf{S}}_{\text{NW}}$ por $\frac{N}{N-K}$ (este ajuste está implementado, por ejemplo, en EViews, así como la segunda de las dos sugerencias de párrafo anterior para escoger H en $\hat{\mathbf{S}}_{\text{NW}}$).

ESTIMACIONES DE NEWEY-WEST EN EL MODELO [1]	
IEV41: Sección 18 pp. 77-78	
y_t	$= 6.1475 - 0.7267x_t + \hat{u}_t,$
(0.3495)*	*(0.1160)
[0.0000]	[0.0000]
$N = 144, \hat{\sigma} = 1.2943, \bar{R}^2 = 0.2909.$	

PARTE 4 - MODELOS ADL

⇒ Modelo ADL(1,1):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \gamma_0 X_t + \gamma_1 X_{t-1} + U_t. \quad [10]$$

⇒ Modelo ADL(1,0):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \gamma_0 X_t + U_t. \quad [11]$$

⇒ Modelo ADL(2,1):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \gamma_0 X_t + \gamma_1 X_{t-1} + U_t. \quad [12]$$

⇒ Modelo ADL(2,2):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \gamma_0 X_t + \gamma_1 X_{t-1} + \gamma_2 X_{t-2} + U_t. \quad [13]$$

⇒ Modelo ADL(0,0) (RLS):

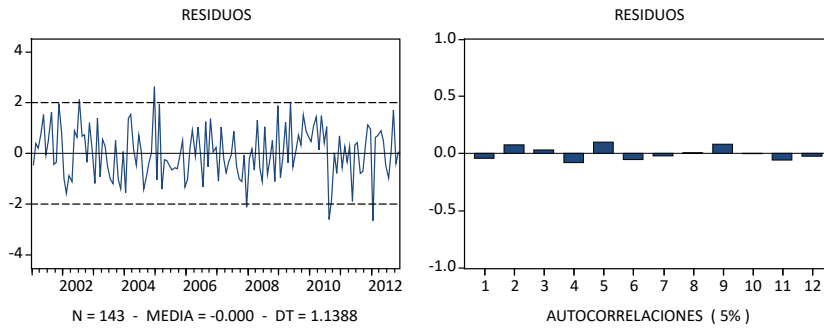
$$Y_t = \beta_0 + \gamma_0 X_t + U_t. \quad [14]$$

IEV41: Sección 18 pp. 78-79

$$y_t = 3.1744 + 0.4495y_{t-1} - 0.9022x_t + 0.5708x_{t-1} + \hat{u}_t,$$

(0.5452)	(0.0752)	(0.0971)	(0.1139)	
[0.0000]	[0.0000]	[0.0000]	[0.0000]	

$$N = 143, \hat{\sigma} = 1.1511, \bar{R}^2 = 0.4428, BG_1 = 11.1430 [0.2850].$$



PARTE 5 - ESTRATEGIA GENERAL CON SERIES ESTACIONARIAS

Estimar varios modelos ADL y seleccionar, entre aquellos cuyos residuos no presenten autocorrelación, el que mejor ajuste proporcione (contrastes de significación y criterios de información); con frecuencia, un modelo ADL(1,1) es suficiente.