

Tema 4

Autocorrelación

4.1. El concepto de autocorrelación y su modelización

4.1.1. Introducción

Este tema aborda el problema de autocorrelación, una característica que puede presentar el término de error de un modelo econométrico y que se manifiesta con mayor frecuencia al tratar con datos observados a lo largo del tiempo. Este tipo de datos, llamados datos de series temporales, son observaciones de una variable recogidas a lo largo del tiempo (días, meses, trimestres, años etc.) de cierta unidad económica (empresa, consumidor, región, país etc.). Estos datos presentan un ordenamiento natural, ya que las observaciones están ordenadas de acuerdo al momento del tiempo en que han sido observadas.

Es probable que a la hora de explicar el comportamiento de una variable observada en el tiempo, se tenga que tener en cuenta que sus observaciones puedan estar correlacionadas a lo largo del mismo. Por ejemplo, el consumo de una familia en un año es probable que esté correlacionada con el consumo del año anterior ya que en media, *ceteris paribus*, esperaríamos que una familia no cambiara mucho su comportamiento en su consumo de un año a otro. Esto implicaría posiblemente, introducir en el modelo para explicar el consumo de un año, por ejemplo, el consumo del año anterior. También podemos considerar que un cambio en el nivel de una variable explicativa, por ejemplo la renta de esa familia en un año, tenga efectos sobre el consumo en años sucesivos. Una forma de tener en cuenta esto sería introducir como variable explicativa del consumo en un año no sólo la renta de ese año, sino la renta de años anteriores.

Finalmente, además de factores observables cuya influencia en el consumo se distribuye a lo largo del tiempo, podemos tener factores no observables recogidos en el término de perturbación del modelo cuya influencia se mantiene durante varios periodos presentando lo que se llama autocorrelación o correlación serial en el tiempo.

Como vemos, hay distintas formas de introducir en un modelo efectos que se distribuyen a lo largo del tiempo. Podemos introducir retardos tanto de la variable endógena, por ejemplo Y_{t-1} , como de ciertas variables explicativas, $X_{t-j}, j = 1, \dots, J$, o modelar el término de error u_t en función por ejemplo de su propio pasado $u_{t-j}, j = 1, \dots, p$. Esto es, lo que se conoce como especificación

dinámica del modelo. En muchos casos es necesario reflexionar sobre esta especificación y saber asignar a qué parte, sistemática o de perturbación, corresponde. Una mala especificación dinámica o simplemente funcional de la parte sistemática de un modelo, puede llevar a detectar autocorrelación en el término de perturbación.

En este tema estudiaremos cómo detectar la presencia de correlación serial en el término de error, analizar cuál puede ser la fuente de este problema y si no es un problema de mala especificación, cómo recoger en el modelo esta característica dinámica en el término de error. También estudiaremos las implicaciones de la existencia de autocorrelación en los métodos de estimación alternativos y en los contrastes de hipótesis sobre los parámetros de interés del modelo. En un tema posterior nos ocuparemos de la dinámica en la parte sistemática.

4.1.2. Procesos Autorregresivos y de Medias Móviles

Normalmente se considera como autocorrelación en el término de error de un modelo correctamente especificado en media, al incumplimiento de la hipótesis básica $E(u_t u_s) = 0 \quad t \neq s$. Esto implica que puede existir relación entre pares de perturbaciones u_t, u_s , de forma que no todas las covarianzas entre perturbaciones de distintos periodos de tiempo van a ser cero.

Una forma de recoger este comportamiento es proponer que las perturbaciones siguen algún tipo de proceso estocástico, tal que sus realizaciones están correlacionadas en el tiempo. Vamos a ver dos tipos de procesos, los llamados procesos autorregresivos (AR) y los de medias móviles (MA).

En general, un proceso o modelo Autorregresivo de orden p , que denotamos como $AR(p)$, se define como un modelo de regresión lineal donde la variable dependiente es el término de perturbación y los regresores son sus propias realizaciones pasadas esto es:

$$u_t = \rho_1 u_{t-1} + \dots + \rho_p u_{t-p} + \varepsilon_t$$

donde $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$, y se conoce como un ruido blanco. Un proceso o modelo de medias móviles de orden q , que se denota como $MA(q)$, también se representa como un modelo de regresión lineal donde la variable dependiente es el término de perturbación, pero los regresores son las realizaciones presente y pasadas de la innovación o ruido blanco ε_t .

$$u_t = \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

La combinación de ambos procesos sería un modelo $ARMA(p,q)$.

La matriz de varianzas y covarianzas de las perturbaciones, Ω , es distinta para cada proceso y bajo ciertas condiciones depende de pocos parámetros, que serán los coeficientes del modelo AR o MA y la varianza de ε_t . Lo veremos con dos casos particulares: los procesos $AR(1)$ y $MA(1)$.

$$\textbf{Proceso AR(1):} \quad u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \stackrel{iid}{\sim} (0, \sigma_\varepsilon^2) \quad |\rho| < 1 \quad (4.1)$$

Es el proceso más habitual y simple para modelizar autocorrelación en el término de error de un modelo de regresión adecuadamente especificado, en especial cuando los datos son de frecuencia anual. Sustituyendo recursivamente podemos expresar u_t como una combinación lineal de los valores pasados y presente de ε :

$$u_t = \rho u_{t-1} + \varepsilon_t = \rho^2 u_{t-2} + \rho \varepsilon_{t-1} + \varepsilon_t = \dots = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}$$

Las propiedades de este proceso son las siguientes:

$$E(u_t) = E\left(\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}\right) = \sum_{j=0}^{\infty} \rho^j E(\varepsilon_{t-j}) = 0$$

$$E(u_t^2) = E\left(\sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}\right)^2 = \sum_{j=0}^{\infty} \rho^{2j} E(\varepsilon_{t-j}^2) = \sigma_\varepsilon^2(1 + \rho^2 + \rho^4 + \rho^6 + \dots)$$

por lo que si $|\rho| < 1$:

$$Var(u_t) = \sigma_u^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \quad \forall t$$

Además,

$$E(u_t u_{t-1}) = E[(\rho u_{t-1} + \varepsilon_t) u_{t-1}] = \rho E(u_{t-1}^2) + E(u_{t-1} \varepsilon_t) = \rho \sigma_u^2$$

$$E(u_t u_{t-2}) = E[(\rho u_{t-1} + \varepsilon_t) u_{t-2}] = \rho E(u_{t-1} u_{t-2}) + E(u_{t-2} \varepsilon_t) = \rho^2 \sigma_u^2$$

En general, $E(u_t u_{t-s}) = \rho^s \sigma_u^2$ y por lo tanto:

$$E(uu') = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{bmatrix} \quad (4.2)$$

Es decir, bajo el modelo AR(1) con el parámetro $|\rho| < 1$, el número de parámetros contenidos en $E(uu')$ es solamente dos, σ_ε^2 y ρ . Esto es conveniente a la hora de tener que utilizar un método de estimación que requiera estimar esta matriz de varianzas y covarianzas.

Como se puede apreciar, la correlación entre perturbaciones no es cero para ningún retardo, aunque a medida que están más alejadas entre sí, dado que $|\rho| < 1$, esta autocorrelación va disminuyendo.

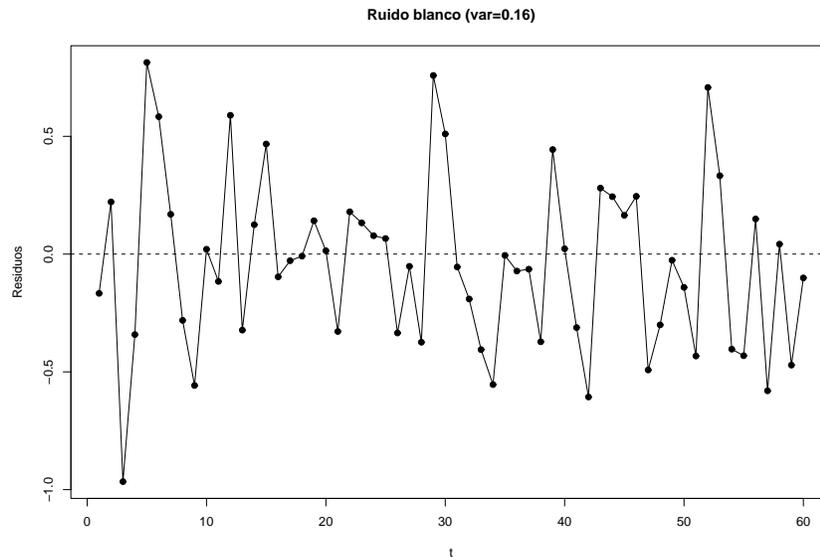
Si $0 < \rho < 1$, todas las covarianzas, y por tanto las correlaciones, son positivas. Esto indica que el efecto de un shock positivo se mantiene en el tiempo con el mismo signo, aunque cada vez será de menor magnitud. Si este signo cambia por un nuevo shock ε_t entonces también se mantendrá en el tiempo el signo negativo. Por lo tanto, las realizaciones de un proceso AR(1) con coeficiente positivo, mostrará una serie temporal con observaciones consecutivas agrupadas del mismo signo. Cuanto más cercano a uno sea el valor de ρ más tardará en general en cambiar el signo.

Si $-1 < \rho < 0$, las covarianzas, y por tanto las correlaciones, alternan el signo siendo positivas entre perturbaciones separadas un número de retardos par y negativas entre aquellas separadas un número de retardos impar. Esto indica que el efecto de un shock positivo tenderá a asociarse con un siguiente error de signo negativo y este con uno positivo, aunque cada vez de menor magnitud. Por lo tanto, las realizaciones de un proceso AR(1) con coeficiente negativo mostrarán una serie temporal con observaciones consecutivas que alternan el signo. Cuanto más cercano a uno en valor absoluto sea el valor de ρ , más se mantendrá la alternancia del signo y su magnitud.

Ejercicio 4.1 Escribe de forma general, sin restricciones sobre sus parámetros, la matriz de varianzas y covarianzas del vector de perturbaciones $E(uu')$ para $T = 10$. ¿Es una matriz simétrica? ¿Cuántos parámetros desconocidos tendrías que estimar? ¿Y si supones que la perturbación sigue un AR(1) con $|\rho| < 1$?

Los siguientes gráficos ilustran la diferencia entre las realizaciones en el tiempo de un proceso que es ruido blanco es decir, sin autocorrelación ($u_t = \varepsilon_t$), versus otras que siguen un AR(1) con coeficiente positivo ($u_t = 0,95u_{t-1} + \varepsilon_t$) o negativo ($u_t = -0,95u_{t-1} + \varepsilon_t$).

Figura 4.1: Ruido blanco $\rho = 0$



$$\text{Proceso MA(1):} \quad u_t = \theta\varepsilon_{t-1} + \varepsilon_t \quad \varepsilon_t \stackrel{iid}{\sim} (0, \sigma_\varepsilon^2) \quad (4.3)$$

Las propiedades de este proceso son las siguientes:

$$E(u_t) = \theta E(\varepsilon_{t-1}) + E(\varepsilon_t) = 0$$

$$E(u_t^2) = E(\theta\varepsilon_{t-1} + \varepsilon_t)^2 = \theta^2 E(\varepsilon_{t-1}^2) + E(\varepsilon_t^2) + 2\theta E(\varepsilon_t \varepsilon_{t-1}) = \sigma_\varepsilon^2(1 + \theta^2) = \sigma_u^2$$

Es decir, las perturbaciones que siguen un MA(1) tienen media cero y varianza constante. Además:

$$E(u_t u_{t-1}) = E(\theta\varepsilon_{t-1} + \varepsilon_t)(\theta\varepsilon_{t-2} + \varepsilon_{t-1}) = \theta\sigma_\varepsilon^2$$

$$E(u_t u_{t-2}) = E(\theta\varepsilon_{t-1} + \varepsilon_t)(\theta\varepsilon_{t-3} + \varepsilon_{t-2}) = 0$$

$$E(u_t u_{t-s}) = 0 \quad \text{si } s \geq 2 \quad \text{y por lo tanto:}$$

Figura 4.2: *Proceso AR(1) con $\rho = 0,95$*

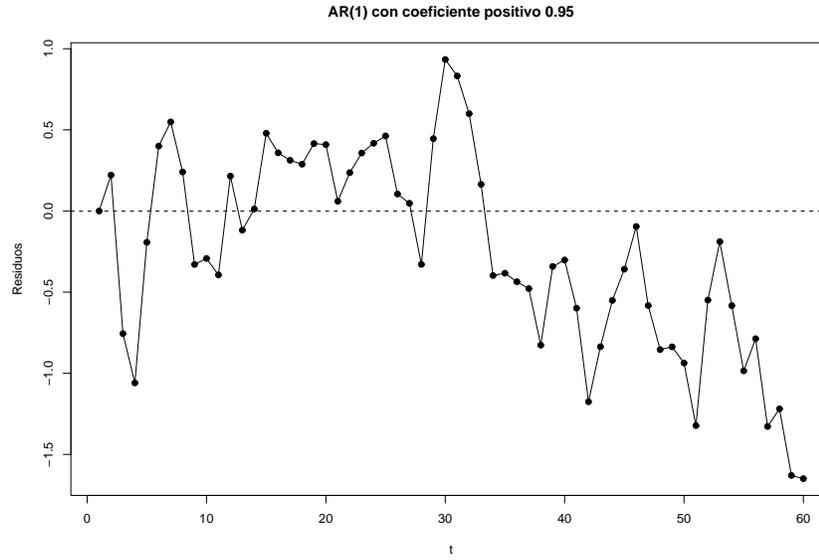
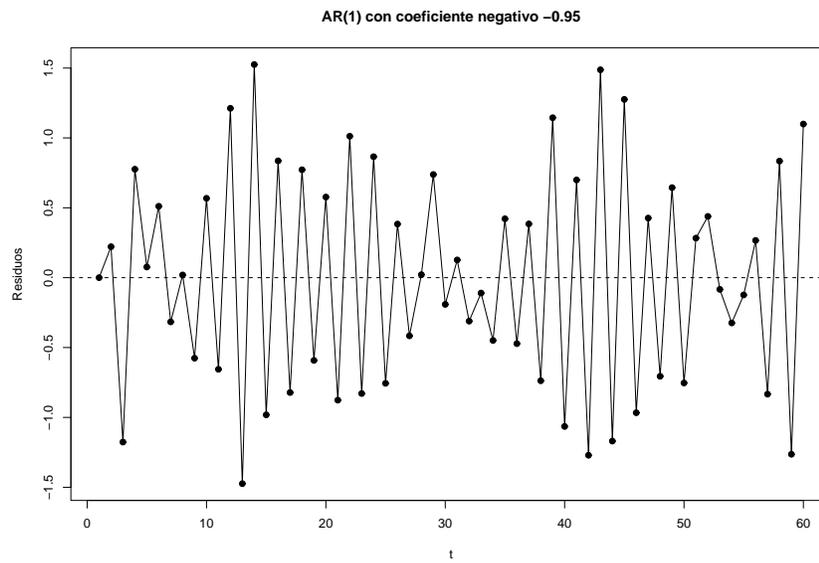


Figura 4.3: *Proceso AR(1) con $\rho = -0,95$*



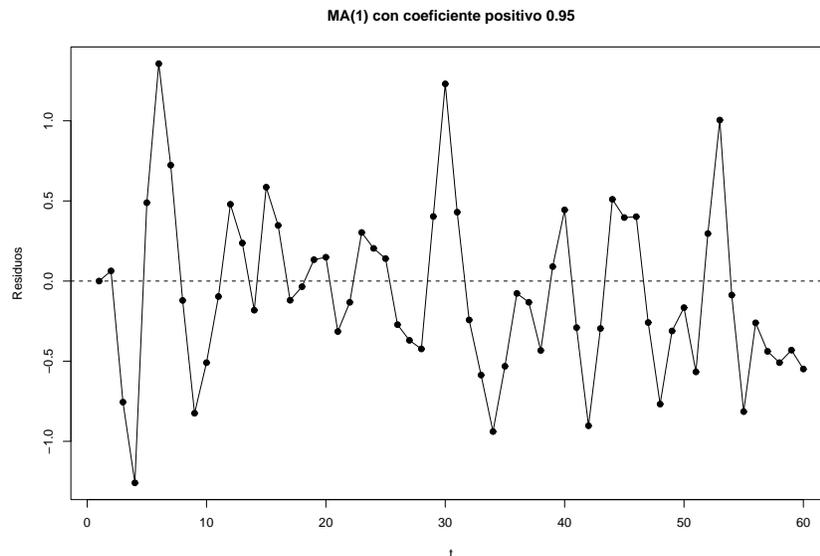
$$E(wu') = \sigma_\varepsilon^2 \begin{bmatrix} 1 + \theta^2 & \theta & 0 & \cdots & 0 \\ \theta & 1 + \theta^2 & \theta & \cdots & 0 \\ 0 & \theta & 1 + \theta^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 + \theta^2 \end{bmatrix} \quad (4.4)$$

Ejercicio 4.2 Calcula la matriz de varianzas y covarianzas si consideramos escribir el proceso MA(1) de la forma $u_t = \varepsilon_t - \delta\varepsilon_{t-1}$ donde $\varepsilon_t \stackrel{iid}{\sim} (0, \sigma_\varepsilon^2)$. ¿Si llamamos $\theta = -\delta$, obtenes lo mismo que antes?

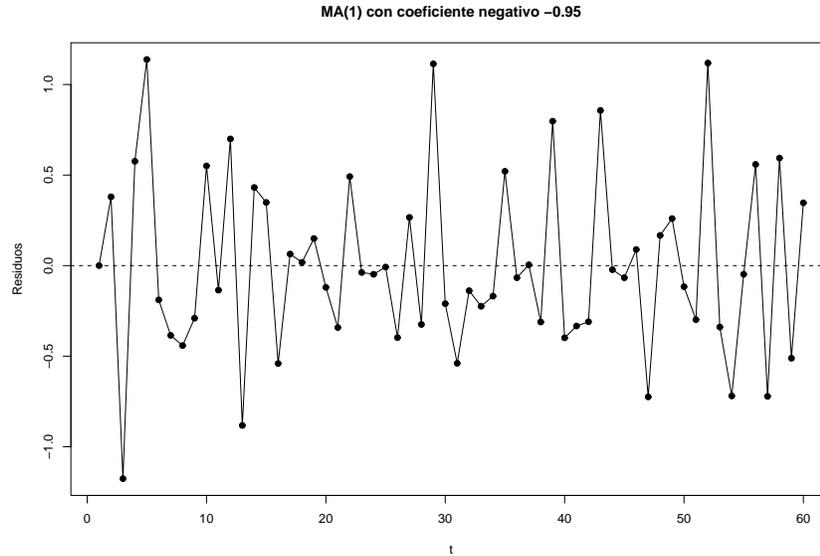
Bajo el proceso MA(1) la matriz $E(uu')$ depende también de dos parámetros, σ_ε^2 y θ . A diferencia del proceso AR(1), las covarianzas entre perturbaciones solamente se mantienen distintas de cero entre aquellas separadas un periodo, siendo el resto igual a cero. Esto indica que si θ es mayor que cero, el efecto de un shock positivo en un periodo tiende a mantenerse en el tiempo con el mismo signo solamente por un periodo, siendo luego igual de factible que la realización sea positiva o negativa dependiendo del shock ε de ese periodo. Si el valor de θ es negativo la alternancia será de positivo a negativo en el siguiente periodo o de negativo a positivo.

Los siguientes gráficos ilustran las realizaciones en el tiempo de un proceso MA(1) positivo ($u_t = 0,95\varepsilon_{t-1} + \varepsilon_t$) y de un MA(1) negativo ($u_t = -0,95\varepsilon_{t-1} + \varepsilon_t$).

Figura 4.4: *Proceso MA(1) con $\theta = 0,95$*



En los gráficos podemos observar que las series generadas por procesos con autocorrelación de orden uno positiva tienden a ser más suaves que las series temporales generadas por procesos con autocorrelación de orden uno negativa. Los procesos autorregresivos tienden a acentuar esta característica relativamente a los de medias móviles. Procesos de mayor orden o combinaciones de estos dos procesos son más difíciles de detectar con un simple gráfico de la serie temporal. Otro tipo de técnicas de identificación de estos procesos son los llamados gráficos de autocorrelación o

Figura 4.5: *Proceso MA(1) con $\theta = -0,95$* 

correlogramas simple y parcial.

4.2. Contrastes de autocorrelación y análisis de residuos

Las perturbaciones no son observables, por lo que los residuos serán nuestra aproximación a estas variables tanto en el análisis gráfico como en los contrastes estadísticos. Al igual que en el tema de heterocedasticidad podemos utilizar como preliminar a un contraste estadístico de autocorrelación, el gráfico de la serie temporal de los residuos obtenidos de estimar por MCO la especificación del modelo de interés. Vamos a estudiar dos contrastes diseñados para contrastar la hipótesis nula de ausencia de autocorrelación frente a diversas alternativas de autocorrelación en el error.

4.2.1. Contraste de Durbin y Watson

Este contraste está diseñado para contrastar la hipótesis nula de no autocorrelación

$$H_0 : \rho = 0, \quad u_t = \varepsilon_t \stackrel{iid}{\sim} (0, \sigma_\varepsilon^2)$$

frente a la alternativa de un proceso AR(1), donde la hipótesis alternativa puede ser

$$H_a : u_t = \rho u_{t-1} + \varepsilon_t \quad 0 < \rho < 1$$

o bien

$$H_a : u_t = \rho u_{t-1} + \varepsilon_t \quad -1 < \rho < 0$$

El procedimiento es el siguiente:

Se estima el modelo de interés por MCO y se calculan los residuos \hat{u}_t para $t = 1, \dots, T$. El estadístico de Durbin-Watson se calcula como:

$$DW = \frac{\sum_{t=2}^T (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^T \hat{u}_t^2}$$

Se puede obtener una relación aproximada entre este estadístico y el coeficiente de autocorrelación de orden uno en los residuos MCO definido como:

$$\hat{\rho} = \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2}$$

siendo esta relación $DW \simeq 2(1 - \hat{\rho})$, por lo tanto, se puede aproximar los valores que tomará el estadístico para ese coeficiente de autocorrelación de orden uno en los residuos.

Así, el valor del estadístico DW está entre $(0, 4)$. Si el valor de DW está cercano a cero indica que el valor de $\hat{\rho}$ está próximo a 1. Si $\hat{\rho} \in (0, 1)$ entonces $DW \simeq (2, 0)$, por lo que valores de DW de 2 hacia 0 indicaran evidencia hacia autocorrelación de orden uno positiva. Si $\hat{\rho} \in (-1, 0)$ entonces $DW \simeq (4, 2)$, por lo que valores de DW de 2 hacia 4 indicaran evidencia hacia autocorrelación de orden uno negativa.

Formalmente para realizar el contraste necesitamos conocer la distribución del estadístico bajo la hipótesis nula de no autocorrelación, pero la distribución de este estadístico DW no es ninguna de las habituales y se demuestra que depende de la matriz X . Durbin y Watson han tabulado unas cotas, d_i (inferior) y d_s (superior) de valores críticos, que permiten decidir si se rechaza la hipótesis nula a un nivel de significación elegido.

Procedimiento: Dado el tamaño muestral T y el número de regresores exceptuando el término independiente $K' = K - 1$, obtenemos en la tabla de valores críticos tabulados las cotas inferior d_i y superior d_s para un nivel de significación elegido.

- Si $DW \in (d_s, 4 - d_s)$, entonces no se rechaza $H_0 : \rho = 0$ frente a cualquiera de las dos alternativas bien $H_a : 0 < \rho < 1$ o bien $H_a : -1 < \rho < 0$. Por lo tanto, no hay evidencia de autocorrelación de orden uno.
- Si $DW < d_i$ se rechaza $H_0 : \rho = 0$ y hay evidencia de autocorrelación de orden uno positiva $H_a : 0 < \rho < 1$.
- Si $DW > 4 - d_i$ se rechaza $H_0 : \rho = 0$ versus la alternativa de autocorrelación de orden uno negativa $H_a : -1 < \rho < 0$.
- Si $DW \in (d_i, d_s)$ o si $DW \in (4 - d_s, 4 - d_i)$, el estadístico cae en las zonas de indeterminación del contraste, por lo que el contraste no es concluyente.

Comentarios:

- Este contraste puede detectar otros procesos AR o MA de mayor orden siempre y cuando den lugar a residuos con una autocorrelación de orden uno significativa. Por la misma razón puede detectar problemas de mala especificación de la parte sistemática del modelo. Pero el contraste no indica cuál puede ser el caso.

- Si la autocorrelación de orden uno no es significativa puede que este contraste no detecte procesos AR o MA de mayor orden.
- En las zonas de indeterminación el contraste no es concluyente. No obstante, se debería tener en cuenta que tampoco se acepta la hipótesis nula de no autocorrelación siendo cautos con los resultados MCO especialmente en la inferencia utilizando los estadísticos t y F habituales. Esta región de indeterminación es mayor cuanto menor número de grados de libertad se tengan (menor T y mayor K).
- El estadístico de DW no es fiable en modelos que incluyen como variables explicativas la variable dependiente retardada Y_{t-1}, Y_{t-2}, \dots u otro tipo de regresor estocástico.

4.2.2. Contraste de Breusch y Godfrey

Este es un contraste más general que el de Durbin-Watson. Se contrasta la hipótesis nula:

$$H_0 : u_t \sim iid(0, \sigma_\varepsilon^2)$$

frente a

$$H_a : u_t = \rho_1 u_{t-1} + \dots + \rho_p u_{t-p} + \varepsilon_t \quad \text{AR}(p) \quad \text{ó}$$

$$u_t = \theta_1 \varepsilon_{t-1} + \dots + \theta_p \varepsilon_{t-p} + \varepsilon_t \quad \text{MA}(p)$$

donde $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$. Luego bajo H_0 el proceso es un ruido blanco, no hay autocorrelación. Bajo la hipótesis alternativa el término de perturbación puede seguir bien un proceso autorregresivo de orden p o bien un proceso de medias móviles del mismo orden, elegido previamente a la realización del contraste. El procedimiento es el siguiente:

- Se estima el modelo de interés por MCO:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_K X_{Kt} + u_t$$

Se calculan los residuos \hat{u}_t para $t = 1, \dots, T$.

- Se obtiene el coeficiente de determinación R^2 de la siguiente regresión auxiliar:

$$\hat{u}_t = \alpha_1 + \alpha_2 X_{2t} + \dots + \alpha_K X_{Kt} + \delta_1 \hat{u}_{t-1} + \dots + \delta_p \hat{u}_{t-p} + \xi_t$$

donde $t = 1, \dots, T$ y p es el número de retardos elegido para el orden del proceso AR o MA bajo la alternativa. Al realizar la regresión auxiliar usando todas las observaciones, incluidas las primeras p ($t = 1, \dots, p$), se necesita el conocimiento de $\hat{u}_0, \dots, \hat{u}_{1-p}$. Una forma de solventar el problema de no conocer esos residuos es la de igualar $\hat{u}_0, \dots, \hat{u}_{1-p}$ a cero.

- El estadístico del contraste y su distribución asintótica bajo la hipótesis nula son:

$$BG = TR^2 \xrightarrow{d, H_0} \chi_p^2 \quad \text{bajo } H_0$$

donde T y R^2 es el tamaño muestral y el coeficiente de determinación, respectivamente, de la regresión auxiliar.

- **Regla de decisión:** Rechazar la hipótesis nula de no autocorrelación a un nivel de significación α , si el valor muestral del estadístico, denotado por BG, es tal que $BG > \chi_{p|\alpha}^2$.

Otra forma de realizar el contraste consiste en omitir las p observaciones iniciales de la regresión auxiliar

$$\hat{u}_t = \alpha_1 + \alpha_2 X_{2t} + \dots + \alpha_K X_{Kt} + \delta_1 \hat{u}_{t-1} + \dots + \delta_p \hat{u}_{t-p} + \xi_t$$

donde $t = p + 1, \dots, T$ y considerar el R^2 de esta regresión y el tamaño muestral disponible $(T - p)$. En este caso el estadístico del contraste y su distribución asintótica bajo la hipótesis nula son:

$$(T - p)R^2 \xrightarrow{d, H_0} \chi_p^2 \quad \text{bajo } H_0$$

siendo la regla de decisión igual que antes.

Comentarios:

- Este contraste es más general que el de Durbin-Watson. Las alternativas pueden ser $AR(p)$ con $p \geq 1$ y también $MA(p)$.
- A diferencia del contraste de Durbin-Watson, sí se puede utilizar en modelos que incluyen como variables explicativas la variable dependiente retardada Y_{t-1}, Y_{t-2}, \dots u otro tipo de regresor estocástico.
- Es un contraste asintótico esto es, bajo H_0 se conoce la distribución asintótica del estadístico. Por lo tanto, es válido para muestras grandes siendo cautos con los resultados para muestras pequeñas.
- La elección del número de retardos p es importante. Un p pequeño puede no capturar autocorrelaciones significativas de mayor orden, pero un p grande puede influir negativamente en la potencia del contraste.

4.2.3. Análisis de los residuos y contrastes: ejemplos

En esta sección vamos a presentar varios ejemplos que ilustran la utilización de los gráficos de residuos para detectar la existencia de autocorrelación y en algunos casos detectar una posible mala especificación del modelo. También ilustramos la utilización del contraste de Durbin-Watson no solamente para contrastar la existencia de autocorrelación de orden uno, sino también como indicador de mala especificación.

Ejemplo 4.1 Autocorrelación positiva

Se dispone del fichero de datos *inv.dat* del libro *Undergraduate Econometrics* de Hill, Griffiths y Judge disponibles en <http://www.wiley.com/college/econ/hill331848/students.html>. Son datos anuales desde 1974 a 2003 de las siguientes variables:

I: Inversión real en billones de dólares.

GNP: Producto Nacional Bruto real en billones de dólares.

R: Tipo de interés.

Se estima por MCO un modelo que relaciona la Inversión real con el Producto Nacional Bruto real y el tipo de interés incluyendo un término independiente.

$$I_t = \beta_1 + \beta_2 GNP_t + \beta_3 R_t + u_t \quad t = 1, \dots, 30$$

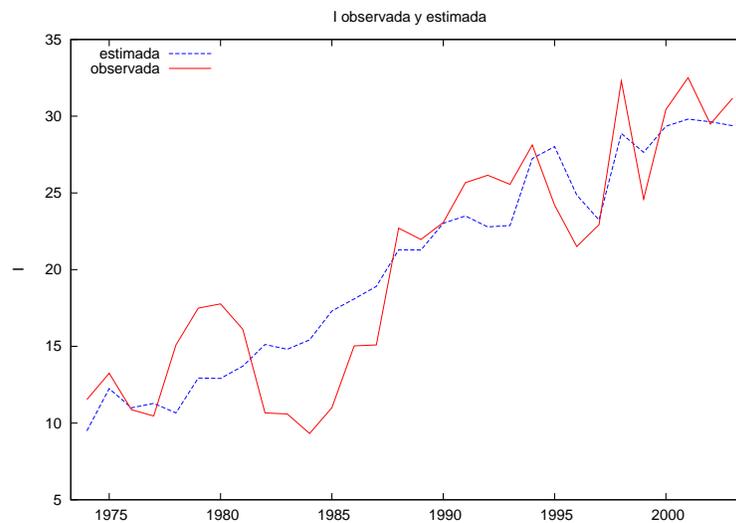
y los resultados de la estimación son:

$$\widehat{I}_t = 6,22494 + 0,769911 GNP_t - 0,184196 R_t$$

$(\widehat{des}(\widehat{\beta}_{MCO})) \quad (2,51089) \quad (0,0717905) \quad (0,126416)$

$$R^2 = 0,816282 \quad SCR = 299,336 \quad DW = 0,852153 \quad \hat{\rho} = 0,567726$$

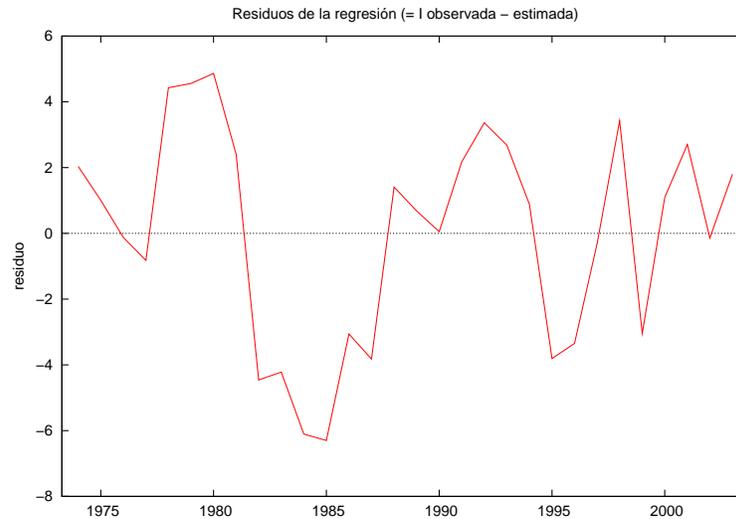
Figura 4.6: Gráfico de la serie de Inversión observada y estimada



La Figura 4.6 muestra el gráfico de las series de Inversión observada y estimada, que representa la evolución en el periodo muestral considerado, de los valores observados de la variable Inversión I_t y de la Inversión estimada \hat{I}_t que predice el modelo ajustado dados los valores observados de las variables explicativas. El modelo estimado no recoge bien los ciclos de expansión y recesión observados en la Inversión, especialmente en los años 1977-1982, y 1982-1988, siendo los observados mucho más pronunciados que los generados por el modelo ajustado. Esto explica los ciclos observados en el gráfico de la serie temporal de los residuos. Este hecho sugiere que el término de perturbación del modelo pueda recoger factores correlacionados en el tiempo tal que presente autocorrelación positiva al menos hasta de orden uno, pudiendo ser también de mayor orden.

La Figura 4.7 muestra la evolución de los residuos $\hat{u}_t = I_t - \hat{I}_t$ a lo largo del periodo muestral considerado $t = 1974, \dots, 2003$. Como es de esperar, ya que hay término constante en la regresión, oscilan alrededor de su media muestral igual a cero. Se puede observar que especialmente de 1973

Figura 4.7: Gráfico de la serie temporal de los residuos MCO



a 1995 se van alternando en ciclos, grupos de residuos seguidos del mismo signo, siendo menos evidente al final de la muestra, donde el ajuste es algo mejor.

Contraste de Durbin-Watson:

$$\begin{array}{llll}
 H_0 : \rho = 0 & u_t \sim iid(0, \sigma_u^2) & & \text{No autocorrelación.} \\
 H_a : 0 < \rho < 1 & u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) & & \text{AR(1) autocorrelación positiva.}
 \end{array}$$

Para $\alpha = 0,05$, $T = 30$, $K' = 2$ las cotas inferior y superior tabuladas respectivamente son $d_i = 1,2837$ y $d_s = 1,5666$. Como $DW = 0,852153 < d_i = 1,2837$, se rechaza H_0 . Hay evidencia de autocorrelación de orden uno positiva en el término de perturbación del modelo .

Ejemplo 4.2 Autocorrelación negativa

Se dispone de observaciones anuales de las variables Consumo (C_t) y Renta (R_t) para un país. Los resultados de la estimación por el método de Mínimos Cuadrados Ordinarios de la función de consumo $C_t = \beta_1 + \beta_2 R_t + u_t$ se muestran a continuación:

Estimación MCO, usando las observaciones 1950-1963 ($T = 14$)
Variable dependiente: C

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	0,3470	1,1371	0,3052	0,7654
R	0,7001	0,0479	14,6101	0,0000
R^2	0,9467	Suma de cuad. residuos	30,6381	
$\hat{\rho}$	-0,7323	Durbin-Watson	3,3158	

Figura 4.8: Gráfico de la serie de Inversión observada y estimada

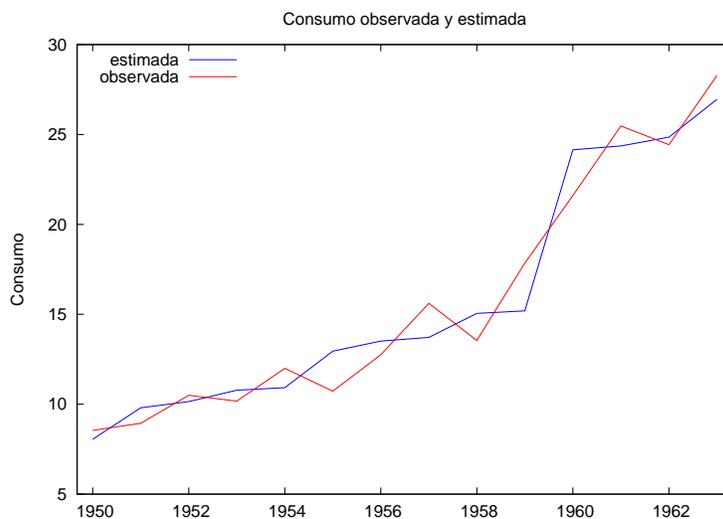
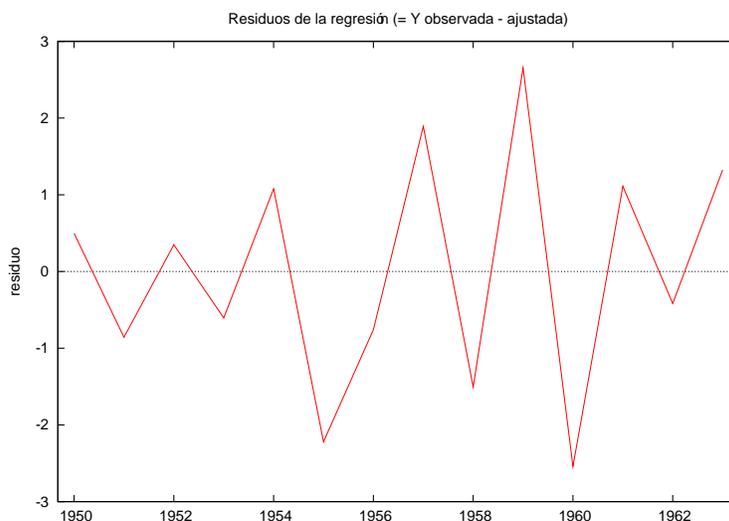


Figura 4.9: Gráfico de la serie temporal de los residuos MCO



La Figura 4.8 representa la evolución en el periodo muestral considerado, de los valores observados de la variable Consumo, C_t y del Consumo estimado, \hat{C}_t que predice el modelo ajustado, dados los valores observados de las variables explicativas. El modelo ajustado parece predecir bien la tendencia creciente del consumo, incluso al final de la muestra donde esta tendencia parece ser más acusada. Pero lo que el modelo no predice bien es la dirección del ciclo, ya que cuando se da una expansión el modelo predice una recesión y cuando se observa una recesión el modelo predice una expansión. Esto hace que en la serie temporal de los residuos de la Figura 4.9 se observe una alternancia de

signos de positivo a negativo y viceversa a lo largo de todo el periodo. Esto sugiere que puede existir autocorrelación de orden uno negativa.

Contraste de Durbin-Watson:

$$\begin{array}{llll} H_0 : \rho = 0 & u_t \sim iid(0, \sigma_u^2) & & \text{No autocorrelación.} \\ H_a : -1 < \rho < 0 & u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) & & \text{AR(1) autocorrelación negativa.} \end{array}$$

Para $\alpha = 0,05$, $T = 14$, $K' = 1$ las cotas inferior y superior tabuladas son respectivamente, $d_i = 1,0450$ y $d_s = 1,3503$. Como $DW = 3,315808 > 4 - d_i = 2,955$ se rechaza H_0 . Hay evidencia de autocorrelación de orden uno negativa en el término de perturbación del modelo.

Los siguientes ejemplos nos servirán para ilustrar que el contraste de Durbin-Watson puede detectar autocorrelación, pero ser una consecuencia de una mala especificación funcional del modelo, o a cambios estructurales que no se han recogido. En estos casos hay que proceder a reespecificar el modelo.

Ejemplo 4.3 Mala especificación. Especificación lineal versus cuadrática

Consideremos los siguientes datos anuales desde 1990 a 2001 observados para las variables Y y X :

Y	6	3	1	1	1	4	6	16	25	36	49	64
X	-4	-3	-2	-1	1	2	3	4	5	6	7	8

Se propone inicialmente la estimación de una relación lineal entre estas dos variables:

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (4.5)$$

Estimación MCO, usando las observaciones 1990-2001 ($T = 12$)

Variable dependiente: Y

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	8,01220	4,05896	1,9740	0,0766
X	4,45591	0,919174	4,8477	0,0007
Suma de cuad. residuos	1501,071	R^2 corregido	0,671647	
$\hat{\rho}$		0,853150	Durbin-Watson	0,321828

El gráfico de la serie observada y ajustada, Figura 4.10, sugiere que el ajuste de una relación lineal no parece ser la más indicada, ya que la serie observada Y presenta una curvatura que no captura el ajuste del modelo. Si analizamos el gráfico de residuos, Figura 4.11, vemos que la serie presenta forma de U, lo que confirma que es mejor una relación cuadrática. La forma que presenta la serie de residuos hace que estos se concentren en grupos de residuos del mismo signo. Esto puede sugerir que el término de error del modelo presenta autocorrelación de orden uno positiva.

El valor del estadístico de Durbin-Watson es $DW = 0,321828$. Las cotas inferior d_i y superior d_s tabuladas para $T = 12$ y $K' = 1$, al nivel de significación del 0,05 son $d_i = 0,971$ y $d_s = 1,331$. Como el valor del estadístico DW es menor que la cota inferior d_i se rechaza la hipótesis nula

Figura 4.10: Gráfico de la serie observada y ajustada con especificación lineal

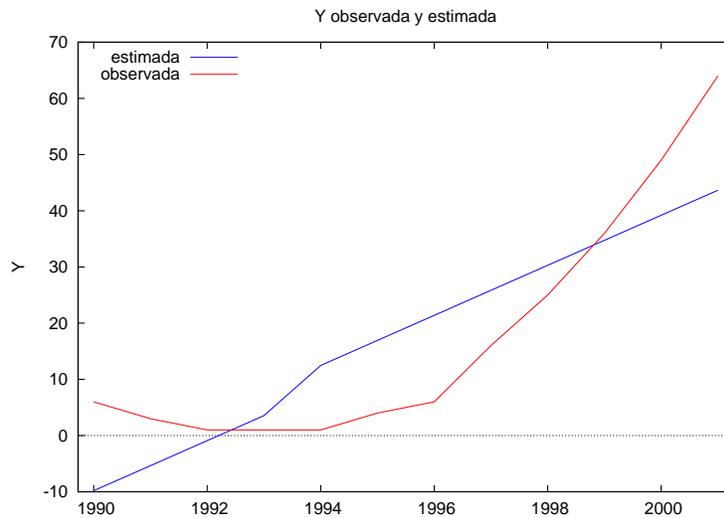
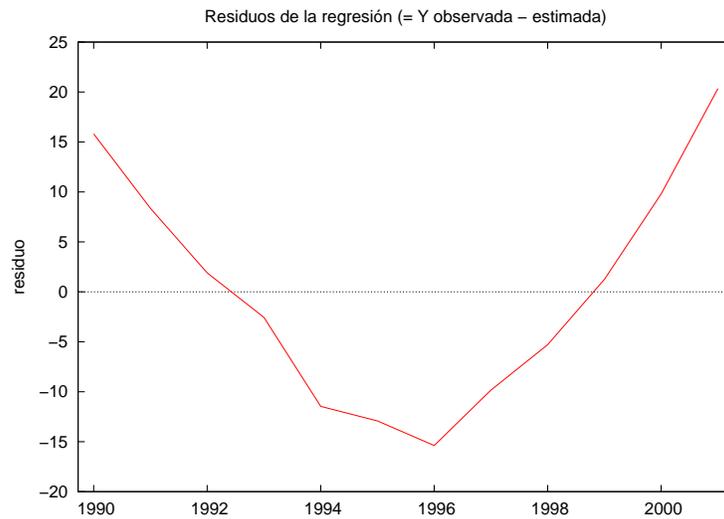


Figura 4.11: Gráfico de residuos MCO de la especificación lineal



de no autocorrelación frente a la alternativa de autocorrelación positiva. El contraste parece estar capturando un problema de mala especificación funcional.

Seguidamente se reespecifica el modelo y se estima por MCO una relación cuadrática entre Y_t y X_t introduciendo como nuevo regresor X_t^2 .

Estimación MCO de la especificación cuadrática.

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 + v_t \quad (4.6)$$

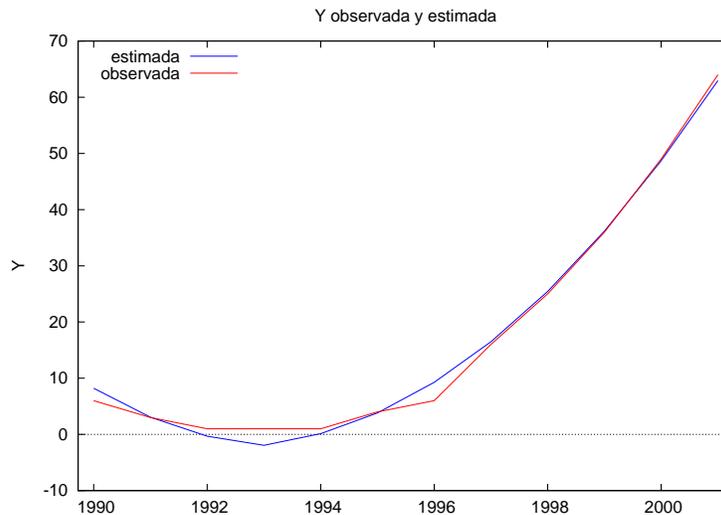
Los resultados, denotando $XSQ = X_t^2$, son los siguientes:

Estimación MCO, usando las observaciones 1990–2001 ($T = 12$)

Variable dependiente: Y

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	-1,7816	0,7408	-2,4050	0,0396
X	1,0334	0,2064	5,0058	0,0007
XSQ	0,8825	0,0407	21,6641	0,0000
Suma de cuad. residuos	28,2431	R^2 corregido	0,9931	
$\hat{\rho}$	0,2984	Durbin-Watson	1,2157	

Figura 4.12: Gráfico de la serie observada y ajustada con especificación cuadrática

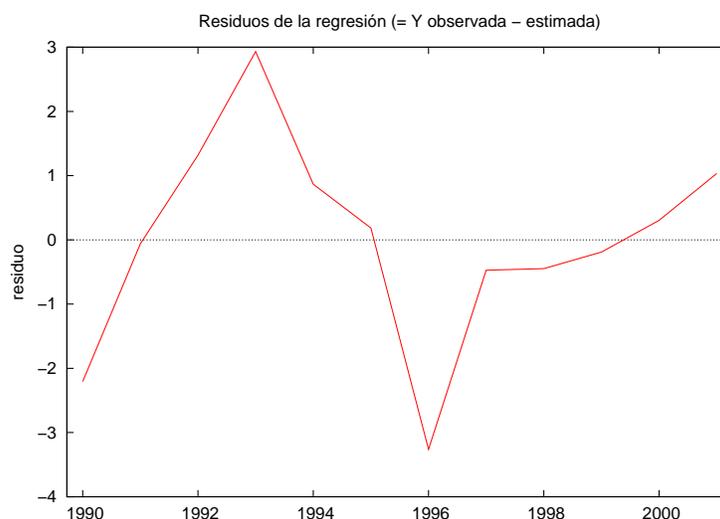


La Figura 4.12 de la relación observada y ajustada muestra claramente que esta especificación se ajusta mucho mejor a los datos. El gráfico de residuos, Figura 4.13, presenta una asociación de residuos positivos y negativos mucho menor que antes. También la magnitud de los residuos, en valor absoluto, es mucho menor.

Por otro lado, el valor del estadístico DW es mucho más cercano a 2 y el valor $\hat{\rho}$ más cercano a cero, por lo que hay una menor evidencia de autocorrelación en el término de perturbación. Mirando a las cotas inferior y superior tabuladas ahora para $T = 12$ y $K' = 2$ son $d_i = 0,812$ y $d_s = 1,579$. El valor muestral del estadístico se encuentra en la zona de indeterminación, $DW = 1,216 \in (0,812; 1,579)$ por lo que el contraste no es concluyente.

Dado que el tamaño muestral es pequeño, la zona de indeterminación del contrastes es bastante amplia y hay que hacer notar que el valor muestral de DW está más cerca de la región de aceptación de la hipótesis nula (hacia cota superior) que de la de rechazo (hacia cota inferior). Aún así, es bastante evidente que el problema de autocorrelación de orden uno positiva detectado anteriormente se debía a una mala especificación funcional o de omisión de la variable X_t^2 . De todas formas, pudiera ser que quede aún cierto grado de autocorrelación en el término de error del modelo reespecificado, aunque tampoco podemos concluir si este es el caso. En cuanto a los resultados de la estimación

Figura 4.13: Gráfico de residuos MCO de la especificación cuadrática



del modelo (4.6), la variable X_t^2 es significativa al 5 %, ya que el valor p asociado a su coeficiente es menor que 0,05. Esto confirma que es más adecuada una relación cuadrática que lineal.

Ejemplo 4.4 Mala especificación. Cambio estructural no recogido

Consideramos los siguientes datos sobre salarios (W) y horas trabajadas (H) en un país durante el periodo 2001-2008. La variable ficticia D recoge un cambio en la política de empleo a partir del año 2005.

Año	2001	2002	2003	2004	2005	2006	2007	2008
W	170	180	165	165	105	95	100	90
H	40	50	30	40	50	35	40	35
D	0	0	0	0	1	1	1	1

En primer lugar, se estima el siguiente modelo para los salarios en función de las horas trabajadas, donde no se tiene en cuenta el cambio en la política de empleo.

$$W_t = \beta_1 + \beta_2 H_t + u_t \tag{4.7}$$

Estimación MCO, usando las observaciones 1-8

Variable dependiente: W

	Coficiente	Desv. Típica	Estadístico t	Valor p
const	102,3210	90,9652	1,1248	0,3036
H	0,7857	2,2436	0,3502	0,7382
Suma de cuad. residuos	10571,43	R^2 corregido	-0,143299	
$F(1,6)$	0,122635	Valor p (de F)	0,738160	
$\hat{\rho}$	0,752888	Durbin-Watson	0,445789	

Figura 4.14: Gráfico de la serie observada y ajustada con el Modelo (4.7)

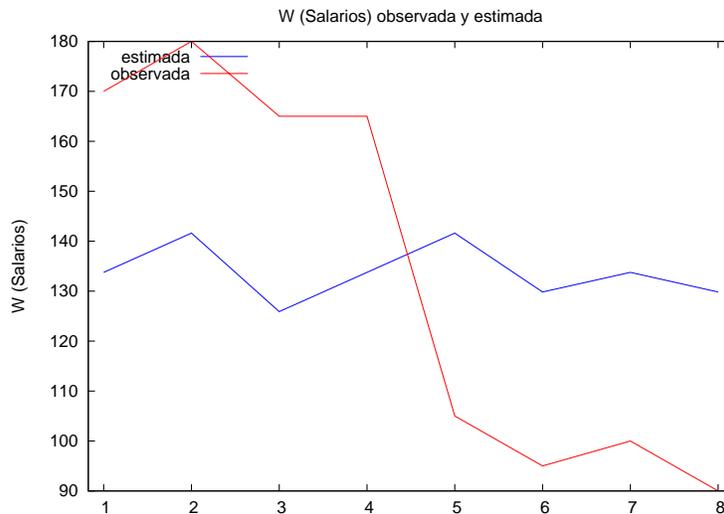
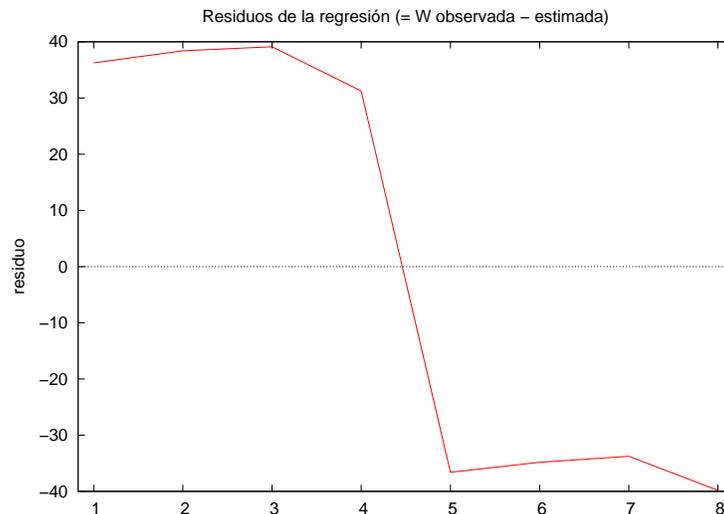


Figura 4.15: Gráfico de residuos MCO del Modelo (4.7)



A la vista del gráfico de la serie observada y ajustada, Figura 4.14, se sospecha que a partir del año 2005 con el cambio en la política de empleo, el nivel medio de los salarios bajó de forma significativa. Este cambio estructural en los salarios no se ha recogido en la especificación, por lo que la serie ajustada de salarios se mantiene mucho más estable que la serie observada. En cuanto al gráfico de residuos, Figura 4.15, se observa un primer grupo de valores positivos, correspondientes a los años previos al cambio y posteriormente un segundo grupo de valores negativos. En consecuencia, se detecta autocorrelación de orden uno positiva, ya que el valor del estadístico de Durbin-Watson es menor que la cota inferior d_i tabulada para el tamaño muestral $T = 8$ y $K' = 1$ esto es, $DW = 0,445789 < d_i = 0,7629$. El valor del coeficiente de autocorrelación de orden uno en los residuos $\hat{\rho} = 0,752888$ también indica autocorrelación positiva.

Tenemos que incluir en el modelo una variable ficticia que recoja ese cambio en el nivel de los salarios mínimos a partir del año 2005. Para ello, consideramos estimar el siguiente modelo ampliado donde además de las horas trabajadas incluimos como regresor la variable ficticia D , para recoger ese cambio estructural:

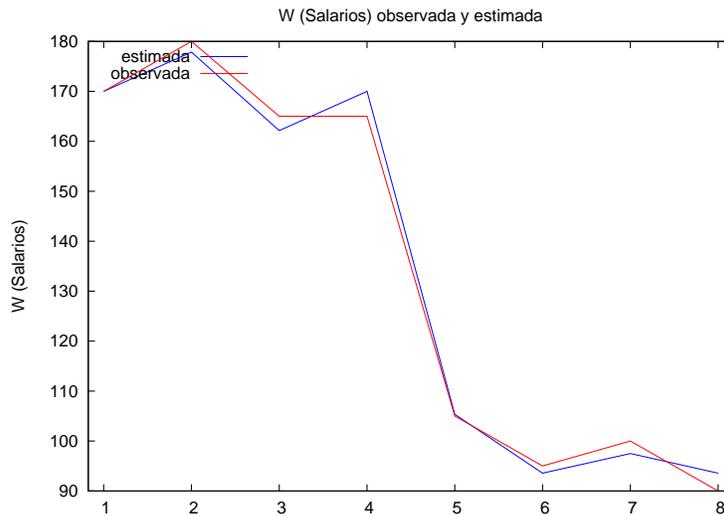
$$W_t = \beta_1 + \beta_2 H_t + \beta_3 D_t + u_t \quad (4.8)$$

Estimación MCO, usando las observaciones 1–8

Variable dependiente: W

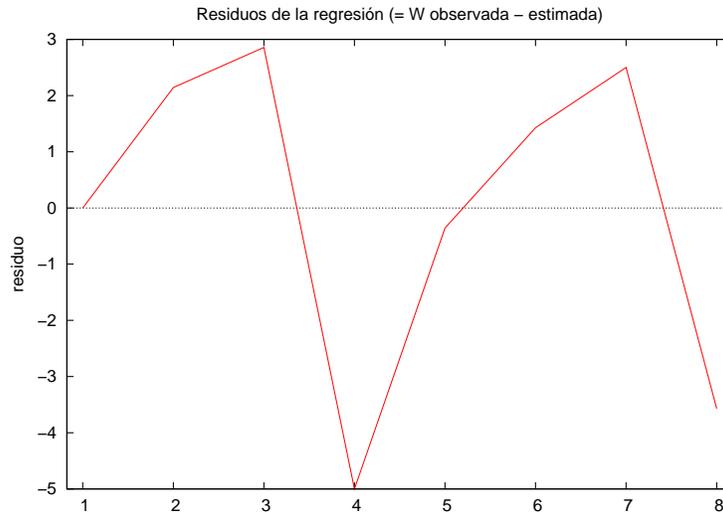
	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	138,5710	7,5381	18,3826	0,0000
H	0,7857	0,1835	4,2817	0,0079
D	-72,5000	2,4275	-29,8659	0,0000
Suma de cuad. residuos		58,92857	R^2 corregido	0,992352
$\hat{\rho}$		-0,265193	Durbin-Watson	2,199134

Figura 4.16: Gráfico de la serie observada y ajustada con el Modelo (4.8)



A la luz de los resultados de estimación y de los gráficos de la serie observada y ajustada y de residuos, Figuras 4.16 y 4.17 respectivamente, es claro que la variable ficticia ha recogido el cambio estructural en el nivel medio de los salarios a partir de 2005. El coeficiente que acompaña a la variable ficticia D es significativamente distinto de cero. El ajuste ha mejorado sustancialmente, tanto en términos del R^2 corregido que ha aumentado considerablemente, como gráficamente. El estadístico de Durbin-Watson toma un valor 2,199 que pertenece a la zona de no rechazo de la hipótesis nula de no autocorrelación ($d_s = 1,7771; 4 - d_s = 2,2229$) dado el nivel de significación 5%, y $T = 8$, $K' = 2$. Por lo tanto, se confirma que la evidencia de autocorrelación en el modelo (4.7) se debía a una mala especificación. Una vez recogido el cambio estructural en la serie de salarios, incluyendo una variable ficticia como variable explicativa adicional a las horas trabajadas, esta evidencia desaparece.

Figura 4.17: Gráfico de residuos MCO del Modelo (4.8)



Ejemplo 4.5 Aplicación de los contrastes de autocorrelación

Utilizamos los datos de la Tabla 6.3 del libro de Greene (1998) *Análisis Económico* que están en el Apéndice: Tablas de datos. Son datos de series temporales con frecuencia anual para la economía americana de los años 1968 a 1982 de las siguientes series:

- INVER: Inversión en términos nominales
- PNB: Producto Nacional Bruto en términos nominales
- IPC: Índice de Precios al Consumo
- TI: Tipo de Interés, promedio anual de la tasa de descuento de la Reserva Federal de NY.

Con estos datos se construyen las series de Inversión y Producto Nacional Bruto en términos reales dividiendo las series nominales por el IPC con año base en 1972 y multiplicando por 10^{-1} con lo cual las series están medidas en trillones de dólares. La Tasa de Inflación se ha calculado como el porcentaje de variación del IPC y el Tipo de Interés es el promedio anual de la tasa de descuento del Banco de la Reserva Federal de Nueva York. El Tipo de Interés Real se obtiene como la diferencia entre el Tipo de Interés Nominal y la Tasa de Inflación.

- $INVERR_t = (INVER_t / IPC_t) \times 0,1$
- $PNBR_t = (PNB_t / IPC_t) \times 0,1$
- $INF_t = ((IPC_t - IPC_{t-1}) / IPC_{t-1}) \times 100$
- $TIREAL_t = TI_t - INF_t$

Partimos de una primera especificación de la función de inversión incluyendo solamente el tipo de interés real. Los resultados obtenidos son los siguientes:

Modelo sin PNBR como regresor:

$$\widehat{INVERR}_t = 0,198273 + 0,006442 TIREAL_t$$

(t-estad) (21,44) (1,42)

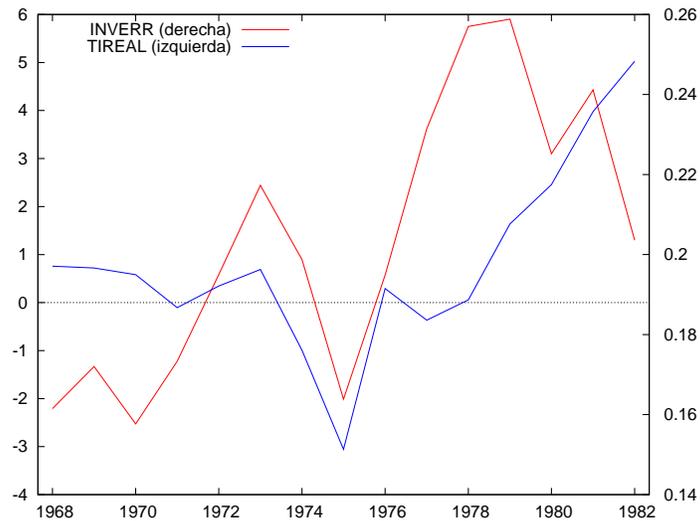
$$T = 15 \quad \bar{R}^2 = 0,068 \quad F(1, 13) = 2,018 \quad DW = 0,556216 \quad \hat{\rho} = 0,669145$$

La variable Tipo de Interés Real, TIREAL, explica muy poco de la Inversión Real, INVERR. El coeficiente de determinación corregido es muy bajo, la variable TIREAL no es significativa al 5 % ya que el valor muestral del estadístico t, en valor absoluto, es menor que el valor crítico 2,16037. A la vista de los resultados del ajuste del modelo sin PNBR como regresor, la especificación propuesta no es muy adecuada.

En la Figura 4.18 se muestra la evolución en el tiempo de la Inversión Real y el Tipo de Interés Real, en el periodo muestral considerado. La serie de Tipo de Interés Real es muy estable en los diez primeros años a excepción del periodo 1973-1976 donde hay una bajada de tipos de interés seguida de una recuperación al mismo nivel anterior. Solamente hay una moderada tendencia creciente a partir de 1978.

La Inversión Real en cambio, presenta una tendencia creciente que se mantiene a lo largo de toda la muestra. Alrededor de ella se muestran ciclos cada vez más pronunciados de expansión y recesión siendo la duración del ciclo más pronunciado de expansión aproximadamente tres años y el de las recesiones sobre dos años.

Figura 4.18: *Evolución temporal de las variables INVERR y TIREAL*



En la Figura 4.19 el gráfico de la serie observada y ajustada por el modelo a la serie INVERR indica

que el modelo no captura nada bien el comportamiento de la Inversión Real. La serie ajustada no recoge la tendencia creciente en la serie ni tampoco los ciclos.

El gráfico de la serie de residuos, Figura 4.20, es casi igual a la serie de la INVERR observada, y dado que no se ha recogido ni siquiera la tendencia creciente, los residuos asociados a la primera mitad de la muestra son negativos y el resto son todos positivos. Esta agrupación de residuos del mismo signo seguidos hace pensar que se pueda detectar un problema de autocorrelación de orden uno positiva, pero el problema es claramente de una mala especificación del modelo.

Figura 4.19: *Serie INVERR observada y ajustada: Modelo sin PNB*

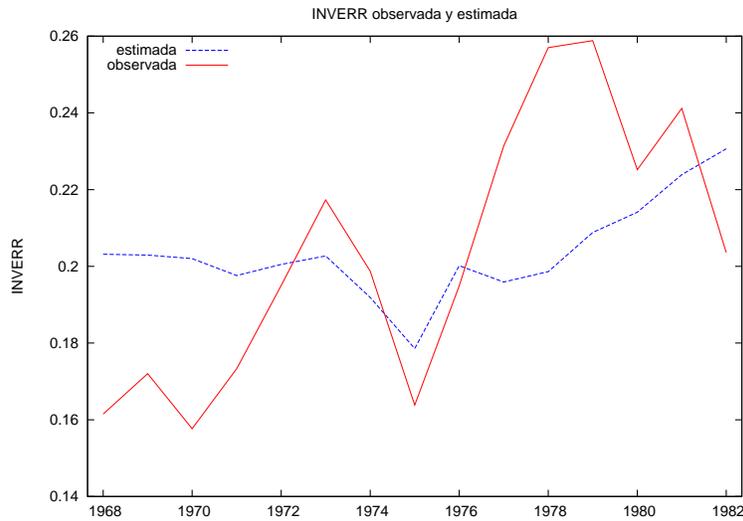
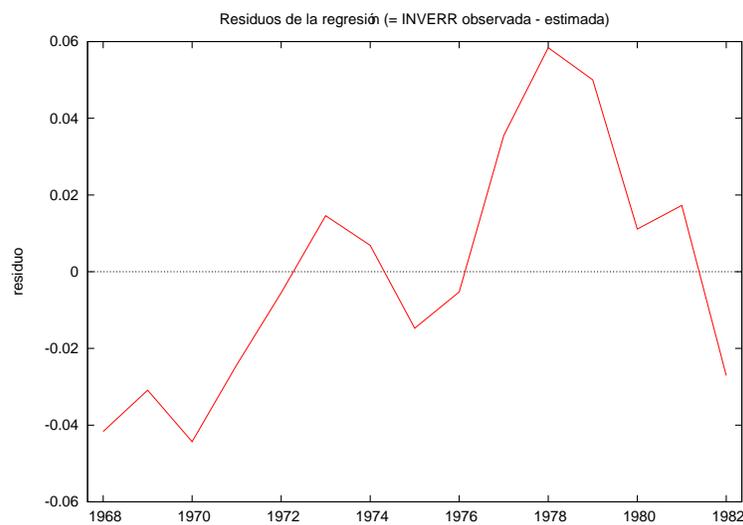


Figura 4.20: *Serie de residuos MCO: Modelo sin PNB*



Contraste de Durbin-Watson:

$$\begin{array}{ll}
 H_0 : \rho = 0 & u_t \sim iid(0, \sigma_u^2) \quad \text{No autocorrelación.} \\
 H_a : 0 < \rho < 1 & u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \quad \text{AR(1) autocorrelación positiva.}
 \end{array}$$

Para $\alpha = 0,05$, el tamaño muestral $T = 15$ y el número de regresores exceptuando el término constante $K' = 1$, las cotas tabuladas inferior y superior respectivamente son $d_i = 1,077$ y $d_s = 1,361$. Como $DW = 0,556216 < d_i = 1,077$, se rechaza H_0 al nivel de significación del 5%. Hay evidencia de autocorrelación de orden uno positiva.

Contraste Breusch-Godfrey:

a) Autocorrelación hasta de orden uno.

$$\begin{array}{ll}
 H_0 : \rho = 0 \quad \text{ó} \quad \theta = 0 & u_t \sim iid(0, \sigma_u^2) \quad \text{No autocorrelación.} \\
 H_a : \rho \neq 0 & u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \quad \text{AR(1) ó} \\
 & \theta \neq 0 \quad u_t = \theta \varepsilon_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \quad \text{MA(1)}
 \end{array}$$

Realizamos la regresión auxiliar correspondiente y calculamos el R^2 para obtener el estadístico. Consideramos $t = 2, \dots, 15$ esto es, 14 observaciones.

$$\begin{array}{ll}
 \hat{u}_t = 0,004 - 0,003 \text{ TIREAL}_t + 0,724 \hat{u}_{t-1} & \\
 \text{t-estad.} \quad (0,63) \quad (-0,99) & (3,49)
 \end{array}$$

El coeficiente de determinación de la regresión auxiliar es $R^2 = 0,524809$ por lo que el valor muestral del estadístico es $BG = 14(0,524809) = 7,347326$. Dado que el valor muestral del estadístico es mayor que el valor crítico $\chi_{1|0,05}^2 = 3,84$, se rechaza H_0 frente a H_a al 0,05 de significación. Hay evidencia de autocorrelación de orden uno. Veamos que ocurre si consideramos una alternativa más general.

b) Autocorrelación hasta de orden tres.

$$\begin{array}{ll}
 H_0 : \rho_1 = \rho_2 = \rho_3 = 0 \quad \text{ó} \quad \theta_1 = \theta_2 = \theta_3 = 0 & u_t \sim iid(0, \sigma_u^2) \quad \text{No autocorrelación.} \\
 H_a : \rho_i \neq 0 \text{ para algún } i & u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \varepsilon_t \quad \text{AR(3) ó} \\
 & \theta_i \neq 0 \text{ para algún } i \quad u_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \varepsilon_t \quad \text{MA(3)}
 \end{array}$$

donde $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$. Realizamos la regresión auxiliar correspondiente y calculamos el R^2 para obtener el estadístico. Consideramos $t = 4, \dots, 15$ esto es, 12 observaciones.

$$\begin{array}{ll}
 \hat{u}_t = 0,005 - 0,001 \text{ TIREAL}_t + 0,838 \hat{u}_{t-1} - 0,308 \hat{u}_{t-2} - 0,069 \hat{u}_{t-3} & \\
 \text{t-estad.} \quad (0,58) \quad (-0,22) & (2,05) \quad (-0,56) \quad (-0,15)
 \end{array}$$

El coeficiente de determinación de la regresión auxiliar es $R^2 = 0,468251$ por lo que el valor muestral del estadístico es $BG = 12(0,468251) = 5,619012$. Dado que el valor muestral del

estadístico es menor que el valor crítico $\chi_{3|0,05}^2 = 7,81$, no se rechaza H_0 frente a H_a al 0,05 de significación. No hay evidencia de autocorrelación hasta de orden tres. Vemos entonces que el incluir más retardos hace que el contraste pierda potencia si estos no son necesarios. De hecho en este caso no parece detectarse la autocorrelación de orden uno que sí era significativa cuando se considera solamente ese retardo.

Los contrastes de autocorrelación sugieren la existencia de autocorrelación positiva de orden uno. Como hemos observado antes, esta evidencia de autocorrelación de orden uno positiva se debe a un pobre ajuste de este modelo. Esto puede estar capturando la omisión de alguna variable relevante que debería de ser incluida para explicar mejor la inversión real. De hecho se debe de reespecificar el modelo incluyendo al menos una variable relevante para explicar la Inversión Real como es el Producto Nacional Bruto en términos reales PNB_R.

Modelo con PNB_R como regresor:

$$\widehat{\text{INVERR}}_t = -0,048683 - 0,002126 \text{TIREAL}_t + 0,197170 \text{PNBR}_t$$

t-estad. (-1,08) (-0,72) (5,51)

$$T = 15 \quad \bar{R}^2 = 0,714 \quad F(2, 12) = 18,496 \quad DW = 1,374911$$

El ajuste ha mejorado sustancialmente en términos del coeficiente de determinación corregido. El valor del estadístico $F(2, 12) = 18,496$ para el contraste de significación conjunta es mayor que el valor crítico al 0,05 de significación (3,89). Luego conjuntamente las variables explicativas TIREAL y PNB_R parecen ser relevantes para explicar la Inversión Real, aunque a nivel de significatividad individual, los estadísticos t muestran que solamente es significativa la variable PNB_R, no siendo TIREAL individualmente significativa (el valor crítico al 5% en las tablas de la t-Student con 12 grados de libertad es 2,179). Por lo tanto, esta especificación del modelo donde además del Tipo de Interés Real se ha incluido el Producto Nacional Bruto en términos reales, parece más adecuada.

La Figura 4.21 muestra el gráfico de la serie INVERR observada y la ajustada por el modelo, indica que esta última especificación captura mejor el comportamiento de la Inversión Real. La serie ajustada recoge la tendencia creciente en la serie, pero no se ajusta del todo bien a los ciclos.

El gráfico de la serie de residuos, Figura 4.22, ya no presenta esa asociación de residuos del mismo signo seguidos, pero muestra los ciclos de expansión y recesión que se daban en la serie de Inversión Real, siendo el ciclo más pronunciado de expansión sobre tres años y las recesiones sobre dos años. Esto sugiere que pueda no ser significativa la autocorrelación de primer orden positiva, pero que pueda existir evidencia de autocorrelación significativa de orden dos o tres.

Contraste de Durbin-Watson:

$$\begin{array}{lll} H_0 : \rho = 0 & u_t \sim iid(0, \sigma_u^2) & \text{No autocorrelación.} \\ H_a : 0 < \rho < 1 & u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) & \text{AR(1) autocorrelación positiva.} \end{array}$$

Para $\alpha = 0,05$, $T = 15$, $K' = 2$ las cotas tabuladas son $d_i = 0,946$ y $d_s = 1,543$.

Como el valor muestral del estadístico está entre los valores inferior y superior, $d_i = 0,946 < DW = 1,347911 < d_s = 1,543$, que es la zona de indeterminación del contraste, el contraste no es concluyente. El valor del estadístico esta más cerca de la cota superior o lo que es lo mismo, más cerca de la zona de aceptación de la hipótesis nula.

Figura 4.21: Serie *INVERR* observada y ajustada: Modelo con *PNBR*

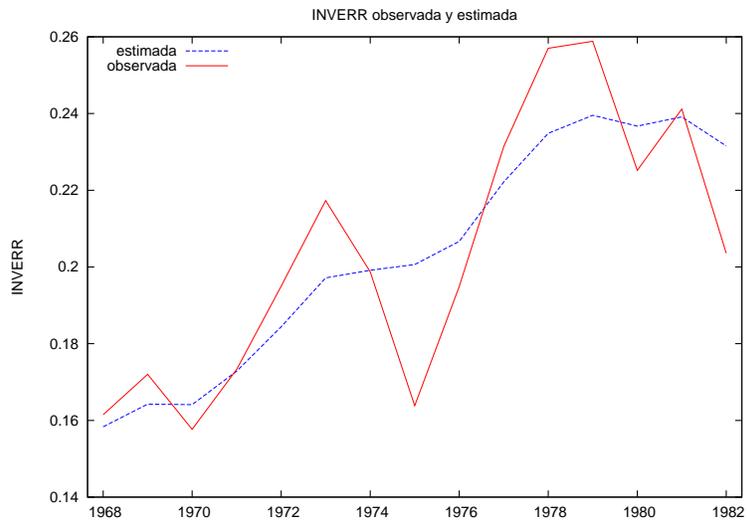
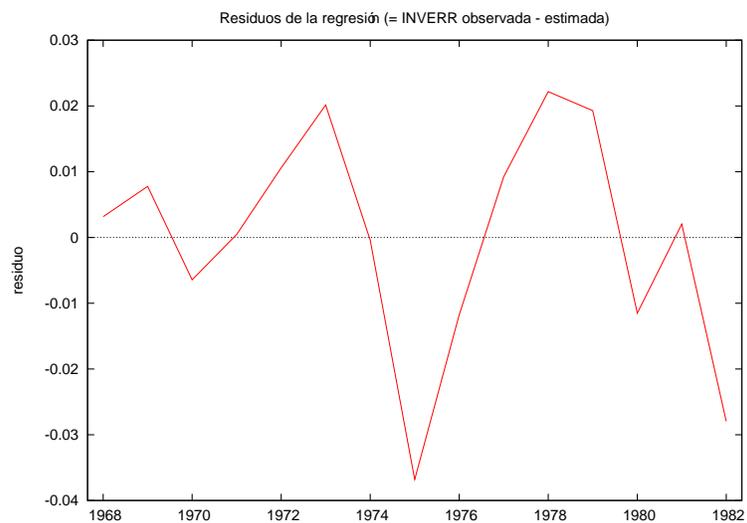


Figura 4.22: Serie de residuos MCO: Modelo con *PNBR*



Contraste Breusch-Godfrey:

a) Autocorrelación hasta de orden uno:

$$H_0 : \rho = 0 \quad \text{ó} \quad \theta = 0 \quad u_t \sim iid(0, \sigma_u^2) \quad \text{No autocorrelación.}$$

$$H_a : \begin{array}{ll} \rho \neq 0 & u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \quad \text{AR}(1) \quad \text{ó} \\ \theta \neq 0 & u_t = \theta \varepsilon_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \quad \text{MA}(1) \end{array}$$

Realizamos la regresión auxiliar correspondiente a este modelo y esta alternativa. Calculamos el R^2 para obtener el estadístico. Consideramos $t = 2, \dots, 15$ esto es, 14 observaciones.

$$\hat{u}_t = -0,001 + 0,0007 \text{PNBR} - 0,0001 \text{TIREAL} + 0,2712 \hat{u}_{t-1}$$

(-0,029)
(0,016)
(-0,034)
(0,785)

El coeficiente de determinación de la regresión auxiliar es $R^2 = 0,058756$ por lo que el valor muestral del estadístico es $BG = 14(0,058756) = 0,822584$. Dado que el valor muestral del estadístico es menor que el valor crítico $\chi_{1|0,05}^2 = 3,84$, no se rechaza H_0 frente a H_a al 5% de significación. No hay evidencia de autocorrelación hasta de orden uno.

Los contrastes de autocorrelación sugieren que no hay problemas de autocorrelación, aunque el de Durbin-Watson no es concluyente. La cuestión es que en estos contrastes las alternativa a contrastar es un proceso con autocorrelación de orden uno. Pudiera ser que la autocorrelación de primer orden no sea muy significativa, pero debido a los ciclos de dos y tres años exista autocorrelación de orden tres. Por ello, hacemos el contraste de Breusch-Godfrey para una alternativa de mayor orden, por ejemplo de orden 3:

b) Autocorrelación hasta de orden 3:

$$H_0 : \rho_1 = \rho_2 = \rho_3 = 0 \quad \text{ó} \quad \theta_1 = \theta_2 = \theta_3 = 0 \quad u_t \sim iid(0, \sigma_u^2) \quad \text{No autocorrelación.}$$

$$H_a : \begin{array}{ll} \rho_i \neq 0 \text{ para algún } i & u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \varepsilon_t \quad \text{AR}(3) \quad \text{ó} \\ \theta_i \neq 0 \text{ para algún } i & u_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \varepsilon_t \quad \text{MA}(3) \end{array}$$

donde $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$. Realizamos la regresión auxiliar correspondiente y calculamos el R^2 para obtener el estadístico. Consideramos $t = 4, \dots, 15$ esto es, 12 observaciones.

$$\hat{u}_t = 0,0674 - 0,051 \text{PNBR} + 0,005 \text{TIREAL} - 0,198 \hat{u}_{t-1}$$

(0,929)
(-0,918)
(1,321)
(-0,490)

$$- 0,011 \hat{u}_{t-2} - 1,106 \hat{u}_{t-3}$$

(-0,032)
(-2,364)

El coeficiente de determinación de la regresión auxiliar es $R^2 = 0,599334$ por lo que el valor muestral del estadístico es $BG = 12(0,599334) = 7,192$. Dado que el valor muestral del estadístico es menor que el valor crítico $\chi_{3|0,05}^2 = 7,81$, no se rechaza H_0 frente a H_a al 5% de significación. No hay evidencia de autocorrelación hasta de orden tres.

¿Qué ocurre si consideramos en el contraste todas las observaciones? gretl realiza de esa forma el contraste de Breusch-Godfrey, donde se igualan a cero el valor de los residuos no disponibles

$\hat{u}_0, \hat{u}_{-1}, \hat{u}_{-2}$. En la ventana de estimación por MCO del modelo de interés, eligiendo Contrastes \rightarrow Autocorrelación se obtiene:

$$\begin{aligned} \hat{u}_t = & 0,021 - 0,016 \text{ PNBR} + 0,003 \text{ TIREAL} - 0,209 \hat{u}_{t-1} \\ & (0,602) \quad (-0,579) \quad (1,251) \quad (-0,591) \\ & - 0,031 \hat{u}_{t-2} - 1,0103 \hat{u}_{t-3} \\ & (-0,099) \quad (-2,587) \end{aligned}$$

El coeficiente de determinación de esta regresión auxiliar es $R^2 = 0,551856$ por lo que el valor muestral del estadístico es $BG = 15(0,551856) = 8,277847$. Dado que el valor muestral del estadístico es mayor que el valor crítico $\chi_{3|0,05}^2 = 7,81$, en este caso se rechaza H_0 frente a H_a al 5% de significación. Habría cierta evidencia de autocorrelación de orden tres, aunque no de orden uno.

Por lo tanto, el problema detectado de autocorrelación de orden uno en el modelo donde no se incluye el PNBR parece haberse resuelto incluyendo esa variable que resulta ser relevante en media para explicar la Inversión real. Todavía se puede considerar cierta evidencia marginal de autocorrelación de orden tres, quizá debido a un componente cíclico residual no explicado por el modelo en la serie de Inversión Real. De todas formas, el número de observaciones es algo limitado y el contraste parece verse afectado por este hecho.

4.3. Consecuencias de la detección de autocorrelación

En los ejemplos anteriores hemos visto que a veces se detecta autocorrelación debida a una mala especificación del modelo. En ese caso, el problema es de omisión de variables relevantes más que de pura autocorrelación. Si las variables omitidas están correlacionadas con las incluidas, entonces habrá sesgos en la estimación de los coeficientes de interés por MCO. Así mismo la inferencia no será válida. En ese caso hay que intentar volver a especificar el modelo incluyendo aquellos factores que son relevantes y que hemos omitido en primera instancia.

Por otro lado, si el problema es puramente de autocorrelación es decir, existen factores no relevantes en media que se recogen en el término de perturbación y presentan correlación entre sí a lo largo del tiempo, las consecuencias son otras:

- Al ser factores no relevantes en media esto es, $E(u) = 0$, si los regresores son no estocásticos, el estimador MCO de los coeficientes β será lineal e insesgado.
- La matriz de varianzas y covarianzas del vector de perturbaciones no esféricas, siga el término de perturbación un proceso AR(p) o MA(q) o una combinación de los dos, es tal que $E(uu') \neq \sigma^2 I$. Esto implica que el estimador MCO ya no es el de menor varianza dentro de la clase de estimadores lineales e insesgados. Por lo tanto, no es el estimador eficiente.
- En muestras grandes, en general el estimador MCO es consistente, pero seguirá siendo no eficiente asintóticamente.
- El estimador usual utilizado para estimar $V(\hat{\beta}_{MCO}) = (X'X)^{-1}X'\Sigma X(X'X)^{-1}$ dado por la expresión $\hat{V}(\hat{\beta}_{MCO}) = \hat{\sigma}^2(X'X)^{-1}$ es un estimador sesgado e inconsistente si $E(uu') \neq \sigma^2 I$.

Por lo tanto, la inferencia utilizando este estimador de $V(\hat{\beta}_{MCO})$ no es válida ni siquiera en muestras grandes.

El estimador de β lineal, insesgado y eficiente en este caso es el obtenido por el método de MCG. Como ya vimos en un tema anterior, el método de estimación de Mínimos Cuadrados Generalizados incorpora en el criterio de estimación la información adicional en la matriz de varianzas y covarianzas de las perturbaciones, $E(uu') = \sigma^2\Omega$. La función objetivo a minimizar viene dada por

$$\text{Min}_{\hat{\beta}} \hat{u}' \Omega^{-1} \hat{u} = \text{Min}_{\hat{\beta}} (Y - X\hat{\beta})' \Omega^{-1} (Y - X\hat{\beta}) \quad (4.9)$$

donde Ω es conocida y σ^2 es un factor de escala al que es invariante el estimador. Para el caso particular de que el término de perturbación siga un AR(1), si el valor de ρ es conocido, la función criterio a minimizar con respecto a β es la siguiente suma de cuadrados:

$$\begin{aligned} SCR(\hat{\beta}) &= (1 - \rho^2) (Y_1 - \hat{\beta}_1 - \hat{\beta}_2 X_{21} - \dots - \hat{\beta}_K X_{K1})^2 \\ &+ \sum_{t=2}^T \left[(Y_t - \rho Y_{t-1}) - \hat{\beta}_1 (1 - \rho) - \sum_{j=2}^K \hat{\beta}_j (X_{jt} - \rho X_{j,t-1}) \right]^2 \end{aligned} \quad (4.10)$$

Minimizar esta función es equivalente a estimar por MCO el siguiente modelo transformado en el que se verifican las hipótesis básicas. Para la primera observación,

$$\sqrt{1 - \rho^2} Y_1 = \beta_1 \sqrt{1 - \rho^2} + \beta_2 \sqrt{1 - \rho^2} X_{21} + \dots + \beta_K \sqrt{1 - \rho^2} X_{K1} + \varepsilon_1 \quad (4.11)$$

y para el resto de observaciones $t = 2, \dots, T$ es,

$$Y_t - \rho Y_{t-1} = \beta_1 (1 - \rho) + \beta_2 (X_{2t} - \rho X_{2,t-1}) + \dots + \beta_K (X_{Kt} - \rho X_{K,t-1}) + \varepsilon_t \quad (4.12)$$

donde $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$. Por tanto, la primera observación sufre una transformación diferente a todas las demás.

Normalmente el valor de los parámetros de los que depende Ω no son conocidos. Por lo general se podrá disponer de un estimador previo de estos parámetros para obtener el estimador MCGF de β . Veamos como obtener este estimador en el caso de que sea un AR(1) el proceso que sigue el término de error del modelo de interés.

4.4. Estimación por MCGF bajo un AR(1)

Suponemos que el modelo de interés es

$$Y_t = \mathbf{X}'_t \beta + u_t \quad u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$$

siendo $\mathbf{X}'_t = [1 \ X_{2t} \ \dots \ X_{Kt}]$ y no se conoce el valor poblacional del parámetro ρ .

El modelo transformado tal que el error es un ruido blanco se puede escribir como:

$$Y_t - \rho Y_{t-1} = (\mathbf{X}'_t - \rho \mathbf{X}'_{t-1}) \beta + \varepsilon_t \quad t = 2, \dots, T$$

donde $\mathbf{X}'_t = [1 \ X_{2t} \ \dots \ X_{Kt}]$.

En general no se considera la primera observación del modelo transformado, porque el estimador así obtenido es asintóticamente equivalente a considerar todas las observaciones y computacionalmente es más sencillo de obtener.

4.4.1. Método de Hildreth y Lu: Red de búsqueda

Dado que $\rho \in (-1, 1)$, se realiza una partición de este intervalo y se estima por MCO el modelo transformado para los diferentes valores del parámetro ρ en esa partición. Se elige aquél valor de ρ y las estimaciones de β por MCO en el modelo transformado correspondientes, que proporcionan una menor suma de cuadrados residual en la estimación.

Los pasos de este método de estimación son:

- Se particiona el intervalo $(-1, 1)$ en una serie de valores equidistantes, $\hat{\rho}_i$, $i = 1, \dots, S$, por ejemplo, $-0,95, -0,9, -0,85, \dots, -0,05, 0, 0,05, \dots, 0,95$. Esta sería la red de búsqueda sobre el parámetro ρ a la hora de minimizar la suma de cuadrados residual al estimar β por MCO en el modelo transformado.
- Para cada ρ_i en la red, se estima β por MCO en el modelo transformado

$$Y_t - \rho_i Y_{t-1} = (\mathbf{X}'_t - \rho_i \mathbf{X}'_{t-1})\beta + w_t \quad t = 2, \dots, T$$

y se obtiene su suma de cuadrados residual

$$SCR_i = \sum_{t=2}^T (\widehat{w}_{it})^2 = \sum_{t=2}^T \left((Y_t - \rho_i Y_{t-1}) - (\mathbf{X}'_t - \rho_i \mathbf{X}'_{t-1})\hat{\beta}_i \right)^2$$

donde \widehat{w}_{it} $t = 2, \dots, T$ son los residuos de estimar β por MCO en el modelo transformado con ese valor ρ_i .

- Los estimadores $\hat{\beta}_{HL}$ y $\hat{\rho}_{HL}$ serán los asociados a la menor suma de cuadrados residual SCR_i del modelo transformado para los valores ρ_i $i = 1, \dots, S$ en la red. Para mayor precisión, se puede repetir el proceso afinando la red de búsqueda en torno al valor $\hat{\rho}_{HL}$ que ha proporcionado la SCR_i mínima.

Ejemplo 4.6 Ejemplo de estimación por Hildreth-Lu.

Seguimos con el Ejemplo 4.1 sobre el modelo que relaciona la Inversión Real con el Producto Nacional Bruto Real y el Tipo de Interés incluyendo un término independiente:

$$I_t = \beta_1 + \beta_2 GNP_t + \beta_3 R_t + u_t \quad t = 1, \dots, 30$$

Consideramos estimar por el método de Hildreth-Lu bajo el supuesto de que u_t sigue un proceso AR(1). El criterio de estimación del método de Hildreth-Lu es elegir el valor de $\hat{\rho}$ en la red tal que la suma de cuadrados residual de estimar el modelo transformado

$$(I_t - \hat{\rho} I_{t-1}) = \beta_1(1 - \hat{\rho}) + \beta_2(GNP_t - \hat{\rho} GNP_{t-1}) + \beta_3(R_t - \hat{\rho} R_{t-1}) + w_t$$

por MCO es la menor.

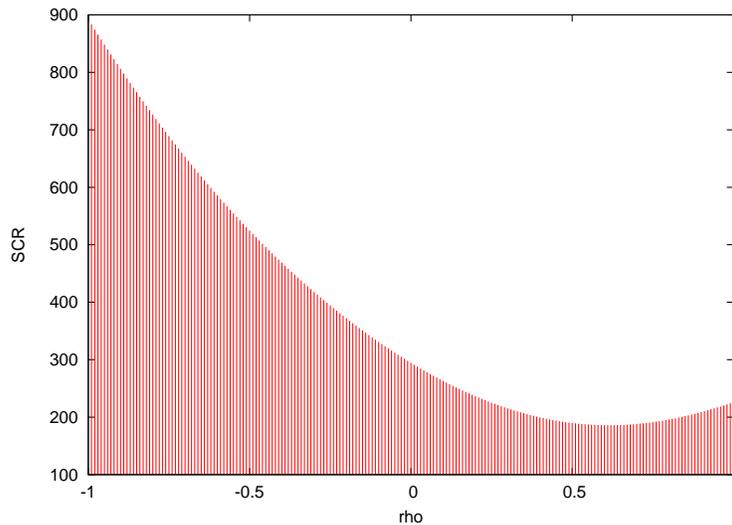
$$\text{Min}_{\hat{\beta}} \text{SCR}(\hat{\beta}) = \text{Min}_{\hat{\beta}} \sum_{t=2}^{t=30} \left(Y_t^* - \hat{\beta}_1 X_{1t}^* - \hat{\beta}_2 X_{2t}^* - \hat{\beta}_3 X_{3t}^* \right)^2$$

donde

$$Y_t^* = I_t - \hat{\rho} I_{t-1}; \quad X_{1t}^* = 1 - \hat{\rho}; \quad X_{2t}^* = \text{GNP}_t - \hat{\rho} \text{GNP}_{t-1}; \quad X_{3t}^* = R_t - \hat{\rho} R_{t-1}$$

En la Figura 4.23 se muestra la Suma de Cuadrados Residual del modelo transformado para cada valor del parámetro ρ en la red de búsqueda.

Figura 4.23: *Función de Suma de Cuadrados Residual del modelo transformado*



El menor valor de la suma de cuadrados residual del modelo transformado es igual a 185,96 y se alcanza para el valor $\hat{\rho} = 0,61$. Ese valor es la estimación elegida para ρ . Dado ese valor obtenemos las estimaciones de los parámetros de interés β de aplicar MCO al modelo transformado:

$$(I_t - 0,61 I_{t-1}) = \beta_1(1 - 0,61) + \beta_2(\text{GNP}_t - 0,61 \text{GNP}_{t-1}) + \beta_3(R_t - 0,61 R_{t-1}) + w_t$$

$$\hat{\beta}_{HL} = \begin{bmatrix} \sum X_{1t}^{*2} & \sum X_{1t}^* X_{2t}^* & \sum X_{1t}^* X_{3t}^* \\ \sum X_{1t}^* X_{2t}^* & \sum X_{2t}^{*2} & \sum X_{2t}^* X_{3t}^* \\ \sum X_{1t}^* X_{3t}^* & \sum X_{2t}^* X_{3t}^* & \sum X_{3t}^{*2} \end{bmatrix}_{\hat{\rho}=0,61}^{-1} \begin{bmatrix} \sum Y_t^* X_{1t}^* \\ \sum Y_t^* X_{2t}^* \\ \sum Y_t^* X_{3t}^* \end{bmatrix}_{\hat{\rho}=0,61}$$

Método de Hildreth–Lu, usando las observaciones 1975–2003 ($T = 29$)

Variable dependiente: I

$$\hat{\rho} = 0,614637$$

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	7,3298	3,6585	2,0035	0,0556
GNP	0,7848	0,1442	5,4418	0,0000
R	-0,2957	0,0786	-3,7596	0,0009

Estadísticos basados en los datos rho-diferenciados:

$$\text{Suma de cuad. residuos} = 185,9631$$

La función de regresión muestral obtenida por tanto es la siguiente:

$$\widehat{(\text{des}(\hat{\beta}_{HL}))} \hat{I}_t = 7,3195 + 0,7851 \text{GNP}_t - 0,2955 \text{R}_t$$

$$(3,6229) \quad (0,1428) \quad (0,0788)$$

4.4.2. Método de Cochrane-Orcutt

Este procedimiento, al igual que el de Hildreth-Lu, consiste en minimizar la suma de cuadrados residual del modelo transformado, pero difiere en la forma de estimar el parámetro ρ . Los pasos a seguir son los siguientes:

1. Se estima el modelo original o de interés $Y_t = \mathbf{X}'_t \beta + u_t$ por MCO y se calculan los residuos \hat{u}_t .
2. Se obtiene $\hat{\rho}$ de la regresión MCO en:

$$\hat{u}_t = \rho \hat{u}_{t-1} + v_t \quad t = 2, \dots, T$$

El estimador así obtenido es el coeficiente de autocorrelación de orden uno en los residuos:

$$\hat{\rho} = \frac{\sum_{t=2}^T \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^T \hat{u}_{t-1}^2}$$

Parece razonable utilizar este estimador como aproximación muestral al coeficiente de autocorrelación de orden uno poblacional entre las perturbaciones ρ . Ahora bien, siempre y cuando los residuos sean función de un estimador de β en el modelo de interés que sea consistente.

3. Estimar β por MCO en el modelo transformado:

$$Y_t - \hat{\rho} Y_{t-1} = (\mathbf{X}'_t - \hat{\rho} \mathbf{X}'_{t-1}) \beta + w_t \quad t = 2, \dots, T$$

Podemos iterar el proceso entre estas dos etapas utilizando en la primera etapa los residuos obtenidos de sustituir en el modelo de interés las estimaciones de los parámetros β , obtenidos en la

segunda etapa. El proceso iterativo finalizará cuando se alcance un criterio de convergencia previamente establecido. Por ejemplo, en términos de la diferencia en valor absoluto de las estimaciones del parámetro ρ , o de la suma de cuadrados residual de la segunda etapa entre dos iteraciones consecutivas.

Ejemplo 4.7 Ejemplo de estimación por Cochrane-Orcutt.

Seguimos con el ejemplo anterior sobre el modelo que relaciona la Inversión Real con el Producto Nacional Bruto real y el tipo de interés incluyendo un término independiente. Estimamos por MCO el modelo de interés

$$I_t = \beta_1 + \beta_2 GNP_t + \beta_3 R_t + u_t \quad t = 1, \dots, 30$$

y guardamos los residuos.

\hat{u}_{1974}	\hat{u}_{1975}	\hat{u}_{1976}	\hat{u}_{1977}	\hat{u}_{1978}	\hat{u}_{1979}	\hat{u}_{1980}	\hat{u}_{1981}
2,036858	1,003142	-0,121917	-0,823617	4,430448	4,556650	4,864468	2,402315
\hat{u}_{1982}	\hat{u}_{1983}	\hat{u}_{1984}	\hat{u}_{1985}	\hat{u}_{1986}	\hat{u}_{1987}	\hat{u}_{1988}	\hat{u}_{1989}
-4,457423	-4,220092	-6,105541	-6,300801	-3,061876	-3,820975	1,406905	0,682641
\hat{u}_{1990}	\hat{u}_{1991}	\hat{u}_{1992}	\hat{u}_{1993}	\hat{u}_{1994}	\hat{u}_{1995}	\hat{u}_{1996}	\hat{u}_{1997}
0,052044	2,172243	3,365111	2,684884	0,883149	-3,807123	-3,346817	-0,307411
\hat{u}_{1998}	\hat{u}_{1999}	\hat{u}_{2000}	\hat{u}_{2001}	\hat{u}_{2002}	\hat{u}_{2003}		
3,423955	-3,045531	1,094203	2,702973	-0,142711	1,799846		

Con estos residuos calculamos el valor de $\hat{\rho}$ de la regresión

$$\hat{u}_t = \rho \hat{u}_{t-1} + v_t \quad t = 2, \dots, 30$$

MCO, usando las observaciones 1975-2003 ($T = 29$)

Variable dependiente: uhat

	Coefficiente	Desv. Típica	Estadístico t	Valor p
uhat_1	0,567726	0,155221	3,6575	0,0010

El estimador así obtenido es el coeficiente de autocorrelación de orden uno en los residuos:

$$\hat{\rho} = \frac{\sum_{t=1975}^{2003} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1975}^{2003} \hat{u}_{t-1}^2}$$

El valor muestral inicial para $\hat{\rho}$ es 0,567726. Con esta primera estimación de ρ se obtiene el modelo transformado:

$$(I_t - 0,56773 I_{t-1}) = \beta_1(1 - 0,56773) + \beta_2(GNP_t - 0,56773 GNP_{t-1}) + \beta_3(R_t - 0,56773 R_{t-1}) + w_t$$

y se estiman por MCO los coeficientes β . De nuevo se calculan los residuos en el modelo de interés utilizando las estimaciones de β en esta etapa y se calcula un nuevo valor de $\hat{\rho}$ en una segunda

la expresión:

$$V(\hat{\beta}_{MCO}) = (X'X)^{-1}X'\Sigma X(X'X)^{-1}$$

El estimador de $V(\hat{\beta}_{MCO})$ dado por

$$\widehat{V}(\hat{\beta}_{MCO}) = \hat{\sigma}^2(X'X)^{-1}$$

donde $\hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{T-k}$, es un estimador sesgado e inconsistente si $\Sigma \neq \sigma^2 I$.

Dado que $\hat{\beta}_{MCO}$ es consistente, si proponemos un estimador alternativo para $V(\hat{\beta}_{MCO})$ que sea consistente podremos definir estadísticos de contraste válidos al menos asintóticamente.

Estimador de Newey-West de $V(\hat{\beta}_{MCO})$:

El estimador consistente de la matriz de varianzas y covarianzas del estimador MCO, robusto a autocorrelación es:

$$\widehat{V}(\hat{\beta}_{MCO})_{NW} = (X'X)^{-1}Q(X'X)^{-1}$$

siendo $Q = \sum_{j=1}^T \hat{u}_t^2 X_t X_t' + \sum_{\ell=1}^L \sum_{t=\ell+1}^T w_\ell \hat{u}_t \hat{u}_{t-\ell} [X_t X_{t-\ell}' + X_{t-\ell} X_t']$ donde $w_\ell = 1 - \frac{\ell}{L+1}$ y $X_t' = [1 \quad X_{2t} \quad \dots \quad X_{Kt}]$

Para realizar inferencia válida al menos asintóticamente, podemos utilizar los estadísticos basados en el estimador de β por MCO y un estimador consistente de su matriz de varianzas y covarianzas, por ejemplo el anterior $\widehat{V}(\hat{\beta}_{MCO})_{NW}$:

Para $q = 1$, el estadístico de contraste y su distribución asintótica bajo H_0 son:

$$\frac{(R\hat{\beta}_{MCO} - r)}{\sqrt{R\widehat{V}(\hat{\beta}_{MCO})_{NW}R'}} \xrightarrow{d, H_0} N(0, 1)$$

Para $q \geq 1$, el estadístico de contraste y su distribución asintótica bajo H_0 son:

$$(R\hat{\beta}_{MCO} - r)' [R\widehat{V}(\hat{\beta}_{MCO})_{NW}R']^{-1} (R\hat{\beta}_{MCO} - r) \xrightarrow{d, H_0} \chi_q^2$$

El utilizar este estimador de $V(\hat{\beta}_{MCO})$ por Newey-West permite seguir utilizando el estimador de los coeficientes MCO para hacer inferencia válida asintóticamente, sin tener que modelizar ni estimar directamente $E(uu') = \Sigma$. Esto permitirá que la inferencia sea válida utilizando los estadísticos anteriores, al menos para muestras grandes. La cuestión es que esto no implica una mejora en la precisión con la que estimamos β por MCO. Para eso tenemos que estimar por *MCG* y esto implica conocer Σ excepto por un factor de escala, o bien por *MCGF* lo que requiere modelizar el proceso de autocorrelación y estimar los parámetros de los que es función Σ .

Ejemplo 4.8 Contraste de significatividad individual de la variable tipo de interés. Utilizamos para realizar el contraste los resultados de estimar los coeficientes por MCO junto a las estimaciones de las desviaciones típicas robustas a la posible existencia de autocorrelación.

$$\widehat{des}(\widehat{\beta}_{MCO})_{NW} \begin{matrix} \hat{I}_t \\ = \\ \end{matrix} \begin{pmatrix} 6,22494 \\ (2,04053) \end{pmatrix} + \begin{pmatrix} 0,769911 \\ (0,0722822) \end{pmatrix} GNP_t - \begin{pmatrix} 0,184196 \\ (0,114826) \end{pmatrix} R_t$$

Contrastamos la significatividad del tipo de interés. Utilizando estos resultados podemos realizar el contraste:

$$\begin{aligned} H_0 : \beta_3 &= 0 \\ H_a : \beta_3 &\neq 0 \end{aligned}$$

con el estadístico

$$t = \frac{\hat{\beta}_{3,MCO}}{\widehat{des}(\hat{\beta}_{3,MCO})_{NW}} \xrightarrow{d,H_0} N(0, 1)$$

Obtenemos el valor muestral del estadístico, que en valor absoluto no supera el valor crítico, $|t| = 1,6041 < 1,96 = N(0, 1)_{|0,025}$, por lo que no rechazamos H_0 para un nivel de significación $\alpha = 5\%$ por lo que no habría evidencia para concluir que el tipo de interés es una variable significativa.

4.6. Inferencia con MCGF

Si consideramos modelizar el proceso de autocorrelación en las perturbaciones, estimar Σ con un estimador consistente $\widehat{\Sigma}$ y utilizar el estimador de β por $MCGF$, que será consistente y asintóticamente eficiente, podemos hacer inferencia utilizando los siguientes estadísticos:

Para $q \geq 1$:

$$(R\hat{\beta}_{MCGF} - r)' [R\widehat{V}(\hat{\beta}_{MCGF})R']^{-1} (R\hat{\beta}_{MCGF} - r) \xrightarrow{d,H_0} \chi_q^2$$

Para $q = 1$:

$$\frac{(R\hat{\beta}_{MCGF} - r)}{\sqrt{R\widehat{V}(\hat{\beta}_{MCGF})R'}} \xrightarrow{d,H_0} N(0, 1)$$

siendo $\widehat{V}(\hat{\beta}_{MCGF})$ un estimador consistente de $V(\hat{\beta}_{MCGF})$. Esto es,

$$\widehat{V}(\hat{\beta}_{MCGF}) = (X'\widehat{\Sigma}^{-1}X)^{-1}$$

Si $\Sigma = \sigma^2\Omega$ entonces un estimador consistente de Σ es $\widehat{\Sigma} = \hat{\sigma}^2\widehat{\Omega}$ donde $\widehat{\Omega}$ es un estimador consistente de Ω utilizado para obtener $\hat{\beta}_{MCGF}$ y

$$\hat{\sigma}^2 = \frac{\hat{u}^*\hat{u}^*}{T - K}$$

donde $\hat{u}^*\hat{u}^* = (Y - X\hat{\beta}_{MCGF})'\widehat{\Omega}^{-1}(Y - X\hat{\beta}_{MCGF}) = (Y^* - X^*\hat{\beta}_{MCGF})'(Y^* - X^*\hat{\beta}_{MCGF})$

Ejemplo 4.9 Contrastamos la significatividad del tipo de interés utilizando los resultados anteriores de la estimación por Hildreth-Lu:

$$\widehat{des}(\hat{\beta}_{HL}) = \begin{matrix} \hat{I}_t & = & 7,31954 & + & 0,785198 & GNP_t & - & 0,295541 & R_t \\ (3,62292) & & (0,142836) & & & & & (0,0788497) \end{matrix}$$

$$\begin{aligned} H_0 : \beta_3 &= 0 \\ H_a : \beta_3 &\neq 0 \end{aligned}$$

con el estadístico

$$t = \frac{\hat{\beta}_{3,HL}}{\widehat{des}(\hat{\beta}_{3,HL})} \xrightarrow{d, H_0} N(0, 1)$$

Como $|t| = 3,7482 > 1,96 = N(0, 1)_{0,025}$ es decir, el valor muestral del estadístico en valor absoluto supera el valor crítico, rechazamos H_0 y concluimos que el tipo de interés es una variable significativa, al 5% de significación.

¿Cuál es la diferencia entre el contraste de significatividad realizado utilizando el estimador de Hildreth-Lu y el que utiliza el estimador MCO del coeficiente que acompaña al tipo de interés utilizando un estimador robusto a autocorrelación de su desviación típica?

La diferencia entre ambos contrastes está en el estimador utilizado para construir el estadístico. El estimador de Hildreth-Lu es un estimador de MCGF y es asintóticamente más eficiente que el estimador MCO. El resultado del contraste es distinto, ya que en el primer caso no se rechaza la hipótesis nula por lo que la variable tipo de interés es una variable significativa, mientras que en el segundo se rechaza la hipótesis nula, pudiendo ser consecuencia de utilizar en el contraste un estimador más eficiente asintóticamente. Por esa razón, aunque ambos contrastes son válidos asintóticamente, puede ser mejor utilizar el contraste basado en MCGF. Pero para ello es necesario estimar consistentemente Σ , lo que no siempre es sencillo.

4.7. Anexo 4.1: Instrucciones básicas de gretl para autocorrelación

En este anexo vamos a indicar las instrucciones básicas para poder realizar la práctica de ordenador del tema. En el anexo del tema anterior ya se han explicado las generales, por lo que nos centraremos solamente en las específicas de autocorrelación.

• **Contrastes de Autocorrelación:** El programa gretl tiene implementados varios contrastes de autocorrelación.

- El estadístico de Durbin-Watson se muestra en los resultados de las estimación por MCO. Para realizar el contraste de Durbin-Watson, hay que obtener los valores de las cotas inferior y superior tabuladas al 5 % para un tamaño muestral dado n , por ejemplo 72 observaciones.

En la ventana principal hay que elegir *Herramientas* → *Tablas estadísticas* → *DW* → se completa el tamaño muestral $n = 72$ y el número de regresores del modelo, excluyendo el término constante, por ejemplo 2.

Se muestran las cotas inferior $d_i \equiv dL$ y superior $d_s \equiv dU$, de la siguiente forma:

Valores críticos al 5% del estadístico de Durbin-Watson, $n = 72$, $k = 2$

$$dL = 1,5611$$

$$dU = 1,6751$$

- El contraste de Breusch-Godfrey se puede obtener en la pantalla de resultados de la estimación MCO. Pulsar:

Contrastes → *Autocorrelación* → *Seleccionar el número de retardos p de la hipótesis alternativa*

Al realizar la regresión auxiliar gretl usa todas las observaciones, incluidas las primeras p igualando aquellos retardos de los residuos no disponibles $\hat{u}_0, \dots, \hat{u}_{1-p}$ a cero. Por supuesto puedes obtener el valor del estadístico de contraste siguiendo todos los pasos aprendidos en clase. Veamos como realizar explícitamente y paso a paso el contraste de Breusch-Godfrey:

1. Se estima el modelo de interés por MCO *Modelo* → *Mínimos Cuadrados Ordinarios*
2. Se guardan los residuos *Guardar* → *residuos* en la ventana de estimación del modelo.
3. Se realiza la regresión auxiliar tal y como se ha explicado en la clase magistral. Si no se consideran todas las observaciones, simplemente se realiza la estimación por MCO eligiendo como variable dependiente la que se ha definido para guardar los residuos y como variables explicativas los regresores del modelo de interés. Además, para incluir como regresores los p retardos de los residuos: en la ventana *gretl:especificar modelo* elegir *retardos* y en la que surge elegir los retardos que se deseen, de 1 a p de la variable dependiente.
4. Se guarda el valor del coeficiente de determinación R^2 de esa regresión auxiliar y se multiplica por el número de observaciones disponibles para el cálculo del estadístico.

- **Estimar de forma consistente la matriz de varianzas y covarianzas de los coeficientes estimados por MCO robusto a autocorrelación.**

La estimación de los coeficientes del modelo es por MCO, pero su matriz de varianzas y covarianzas se estima teniendo en cuenta que hay autocorrelación para que la inferencia usando los estadísticos t y F sea adecuada:

Modelo → *Mínimos Cuadrados* → *elegir la variable dependiente y las variables independientes*

Elegir en esa misma ventana **desviaciones típicas robustas**, y en configurar elegir HAC.

De esta forma podemos hacer los contrastes de significatividad individual con los valores de los estadísticos t que nos muestra el output eligiendo el valor crítico en la distribución $N(0,1)$. Cualquier otro contraste de restricciones lineales los podemos realizar eligiendo en la ventana de estimación *Contrastes* → *Restricciones lineales*. El programa gretl considera para realizar el contraste el estimador robusto a autocorrelación de la matriz de varianzas y covarianzas de los coeficientes estimados por MCO.

- **Estimar la especificación del modelo por el método de Cochrane-Orcutt.**

Modelo → *Series Temporales* → *Cochrane-Orcutt*

En la ventana de resultados se muestra el valor de rho y la SCR del modelo transformado para cada iteración, hasta el valor final alcanzado el criterio de convergencia que por defecto usa gretl. A su vez, se muestran los resultados de la estimación del modelo original por este método. También se muestran una serie de estadísticos basados en los datos rho-diferenciados. Hay que ser cautelosos en la interpretación de estos resultados, ya que se refieren al modelo transformado y no al original. Este modelo tiene regresores estocásticos ya que para la transformación se utiliza la estimación del parámetro ρ . Por lo tanto, el valor del estadístico Durbin-Watson no es comparable con las cotas tabuladas bajo el supuesto de regresores no estocásticos.

Cualquier contraste de restricciones lineales utilizando este estimador lo podemos realizar eligiendo en la ventana de estimación *Contrastes* → *Restricciones lineales*

- **Estimar la especificación del modelo por el método de Hildreth-Lu**

Modelo → *Series Temporales* → *Hildreth-Lu*

En la ventana de resultados se muestra el valor de rho en la red para la cual la SCR es mínima junto con los resultados de la estimación de los parámetros del modelo original. Adicionalmente, se muestra en una nueva ventana la función a minimizar que es la Suma de Cuadrados de Residuos (SCR) del modelo transformado para distintos valores de rho en la red.

- **Tratamiento de los retardos en gretl.**

- Añadir retardos de una variable en concreto a la ventana principal.

Añadir → *Retardos de las variables seleccionadas*

En la ventana que surge poner el número de retardos que se quiere añadir. El problema está en que estos retardos no aparecerán en las ventanas de estimación del modelo por MCO o por otro tipo de métodos.

Otra forma sería *Añadir* → *Definir nueva variable*

En la ventana que surge definir el retardo de la variable deseada, por ejemplo de la variable RD el retardo 4 sería $RD4 = RD(-4)$

Entre paréntesis el número de retardo deseado se acompaña con el signo menos.

Vemos los datos de esas variables

Obs	RD	RD4
1960	57,94	
1961	60,59	
1962	64,44	
1963	70,66	
1964	76,83	57,94
1965	80,00	60,59
1966	84,82	64,44
1967	86,84	70,66
1968	88,81	76,83
1969	88,28	80,00

La primera columna sería RD_t y la segunda RD_{t-4} .

- Añadir retardos de la variable dependiente o de los regresores y usarlos en la estimación. En este caso se añadirán en la misma ventana de estimación:

En la ventana *gretl*: *especificar modelo* elegir *retardos* y en la que surge elegir los retardos que se deseen, bien de forma continua o retardos específicos de forma discontinua, tanto de la variable dependiente como de las variables explicativas que se han definido previamente.

