

ANÁLISIS DE AUTOCORRELACIÓN

1. DEFINICIÓN Y CAUSAS DE AUTOCORRELACIÓN

En este tema se cuestionar, para los modelos que trabajan con datos de series de tiempo, una de las hipótesis que definen el Modelo de Regresión Lineal Normal Clásico (MRLNC). En concreto se analiza la hipótesis que establece que el vector de perturbaciones sigue una distribución según un vector normal esférico.

$$E(u_t) = 0 \qquad E(u_t^2) = \sigma^2 = g_0 \qquad E(u_t u_{t+s}) = 0 \quad \forall s \neq 0$$

La hipótesis de covarianzas nulas es muy interesante desde el punto de vista de las propiedades deseables para los estimadores mínimo cuadráticos ordinarios, pero con frecuencia esta hipótesis es difícil de aceptar en la práctica, en especial cuando las observaciones se suceden en el tiempo. Este problema lo reflejó Malinvaud¹ (1964) señalando que:

“... existe a menudo una correlación positiva entre los términos de perturbación separados s periodos debido al hecho de que los factores no identificados del fenómeno actúan con una cierta continuidad y afectan frecuentemente de análoga manera dos valores sucesivos de la variable endógena.”

Entre los factores no identificados señalados por Malinvaud podría encontrarse un error en la especificación de la forma funcional del modelo y la omisión de variables relevantes que puede dar lugar a un comportamiento sistemático de los residuos que podría interpretarse como autocorrelación cuando en realidad se corrige al especificar correctamente el modelo.

En los casos de incumplimiento de la hipótesis de no autocorrelación es necesario formular el modelo de regresión de un modo más general prescindiendo de esta hipótesis; este modelo recibe el nombre de *modelo de regresión lineal generalizado* y su estimación se realizará aplicando métodos distintos al de mínimos cuadrados ordinarios.

2. MODELIZACIÓN DE LA VARIABLE DE PERTURBACIÓN ALEATORIA

Matemáticamente este supuesto de autocorrelación se expresa a partir de la hipótesis que hace referencia a la covarianza de la perturbación que, como se ha señalado es no nula.

$$E(u_t u_{t+s}) \neq 0 \quad s = 0, \pm 1, \pm 2, \dots$$

se está considerando que el término de perturbación de una observación está relacionado con el término de perturbación de otras observaciones y por lo tanto la covarianza entre ellos es distinta de cero y se define como, $E(u_t u_{t+s}) = g_s \quad s = 0, \pm 1, \pm 2, \dots$

Esto es las covarianzas —o autocovarianzas— son simétricas en el retardo s e independientes del tiempo².

A partir de estas autocovarianzas se pueden definir los coeficientes de autocorrelación; así, el coeficiente de autocorrelación del retardo s es,

$$r_s = \frac{Cov(u_t, u_{t+s})}{\sqrt{Var(u_t)Var(u_{t+s})}}$$

Bajo el supuesto de homocedasticidad —varianzas de la perturbación constantes en el tiempo— el coeficiente de autocorrelación se puede expresar como:

$$r_s = \frac{g_s}{g_0} \quad s = 0, \pm 1, \pm 2, \dots$$

El modelo que ahora se estudia es por tanto un *Modelo de Regresión Lineal Generalizado con autocorrelación* y verifica todas las hipótesis del modelo de regresión lineal normal clásico excepto la que hace referencia a la nulidad de las covarianzas de la perturbación. Este nuevo modelo, en el que se supone que no hay problemas de heterocedasticidad, queda especificado como,

$$Y = X\mathbf{b} + u \quad \text{donde } u \longrightarrow N(0, \mathbf{S}^2 \Omega)$$

Dado que se admite la existencia de autocorrelación pero no de heterocedasticidad la matriz de varianzas y covarianzas de la perturbación - $\sigma^2 \Omega$ - presenta los elementos de la diagonal principal constantes. Es decir, la matriz de varianzas y covarianzas de la perturbación es de la forma,

¹ Malinvaud, E. (1964) "Méthodes Statistiques de l' Econométrie", p. 83, Dunod, Paris.

² Cuando $s=0$, se obtiene la varianza que se define como, $g_0 = E(u_t^2) = \mathbf{S}^2$

$$E(uu') = \begin{pmatrix} \mathbf{s}^2 & \mathbf{g}_1 & \dots & \mathbf{g}_{n-1} \\ \mathbf{g}_1 & \mathbf{s}^2 & \dots & \mathbf{g}_{n-2} \\ \dots & \dots & \dots & \dots \\ \mathbf{g}_{n-1} & \mathbf{g}_{n-2} & \dots & \mathbf{s}^2 \end{pmatrix} = \mathbf{s}^2 \mathbf{\Omega} = \mathbf{s}^2 \begin{pmatrix} 1 & \mathbf{r}_1 & \dots & \mathbf{r}_{n-1} \\ \mathbf{r}_1 & 1 & \dots & \mathbf{r}_{n-2} \\ \dots & \dots & \dots & \dots \\ \mathbf{r}_{n-1} & \mathbf{r}_{n-2} & \dots & 1 \end{pmatrix}$$

donde $\mathbf{\Omega}$ es una matriz simétrica que representa la matriz de correlaciones de las perturbaciones del modelo.

Así, en un modelo de regresión homocedástico si existe autocorrelación aumenta de modo considerable el número de parámetros a estimar; a los k parámetros β hay que añadir ahora la estimación de la varianza de la perturbación σ^2 , y las $n-1$ covarianzas; Esto es, para un modelo con n observaciones hay que estimar $k+n$ parámetros, lo que obliga a suponer alguna estructura de autocorrelación para el término de perturbación, de modo que resulte más sencilla la estimación de estos nuevos parámetros.

Para poder estimar este modelo, obteniendo estimadores con buenas propiedades, se aplican mínimos cuadrados generalizados —o la transformación de Aitken— por lo que es necesario estimar estas correlaciones entre los términos de perturbación.

La forma concreta de estas correlaciones dependerá del proceso generador de las perturbaciones; aunque para su estimación, y puesto que las perturbaciones no se observan se utilizarán los residuos mínimo cuadrático ordinarios (MCO).

3. FORMAS DE AUTOCORRELACIÓN. ESQUEMAS AUTORREGRESIVOS Y DE MEDIAS MÓVILES

El procedimiento práctico consiste en estimar estas correlaciones desconocidas suponiendo alguna estructura de las perturbaciones. Las estructuras más sencillas, y que por otro lado especifican bien el comportamiento de la perturbación, son las que se presentan a continuación como procesos autorregresivos (AR), procesos de medias móviles (MA) o procesos mixtos (ARMA).

- Procesos autorregresivos

Los procesos o filtros autorregresivos están diseñados de modo que el comportamiento de una variable en un instante de tiempo depende de valores pasados de la propia variable. Así, si el valor de la variable u en el momento t depende de su valor en el periodo anterior más un término aleatorio se dice que el proceso es autorregresivo de primer orden (AR(1)). Si la relación de dependencia se establece con los p valores anteriores el proceso será autorregresivo de orden p . Matemáticamente estos procesos se expresan del siguiente modo,

$$\text{AR}(1) \quad u_t = \mathbf{j} u_{t-1} + \mathbf{e}_t$$

$$\text{AR}(2) \quad u_t = \mathbf{j}_1 u_{t-1} + \mathbf{j}_2 u_{t-2} + \mathbf{e}_t$$

...

$$\text{AR}(p) \quad u_t = \mathbf{j}_1 u_{t-1} + \mathbf{j}_2 u_{t-2} + \dots + \mathbf{j}_p u_{t-p} + \mathbf{e}_t$$

Donde \mathbf{e}_t es un proceso de ruido blanco y por tanto con esperanza nula, varianza constante y covarianza nula.

- Procesos de medias móviles

Los procesos de medias móviles, por su parte, se estructuran estableciendo una relación de dependencia entre la variable que se modeliza y un conjunto de retardos de la variable de ruido blanco \mathbf{e}_t . Si sólo existe un retardo de la variable de ruido blanco el proceso será de orden 1, mientras que si existe una combinación lineal de q términos de error de ruido blanco el proceso se denomina MA(q).

$$\text{MA}(1) \quad u_t = \mathbf{e}_t + \mathbf{q} \mathbf{e}_{t-1}$$

$$\text{MA}(2) \quad u_t = \mathbf{e}_t + \mathbf{q}_1 \mathbf{e}_{t-1} + \mathbf{q}_2 \mathbf{e}_{t-2}$$

...

$$\text{MA}(q) \quad u_t = \mathbf{e}_t + \mathbf{q}_1 \mathbf{e}_{t-1} + \mathbf{q}_2 \mathbf{e}_{t-2} + \dots + \mathbf{q}_q \mathbf{e}_{t-q}$$

- Procesos mixtos

Si el comportamiento de una variable se modeliza combinando procesos autorregresivos y de medias móviles se denomina ARMA. Así un modelo ARMA(p,q) se caracteriza por una sucesión de p términos autorregresivos y q términos de medias móviles; esto es,

$$u_t = \mathbf{j}_1 u_{t-1} + \mathbf{j}_2 u_{t-2} + \dots + \mathbf{j}_p u_{t-p} + \mathbf{e}_t + \mathbf{q}_1 \mathbf{e}_{t-1} + \mathbf{q}_2 \mathbf{e}_{t-2} + \dots + \mathbf{q}_q \mathbf{e}_{t-q}$$

El objetivo que se plantea será entonces conocer qué esquema sigue la perturbación y cuál es la mejor estructura para su modelización.

4. DETECCIÓN DE LA AUTOCORRELACIÓN

Para detectar la presencia de autocorrelación se pueden utilizar métodos gráficos y contrastes de hipótesis. A través de los contrastes gráficos se intuirá si existe autocorrelación cuando existan comportamientos sistemáticos para los residuos.

Los contrastes de hipótesis, por su parte, permiten, a través de una regla de decisión, considerar si con los datos de la muestra y con un nivel de significación (α) concreto se debe o no rechazar la hipótesis nula.

Todos los contrastes numéricos de autocorrelación se plantean con idénticas hipótesis; así, podemos señalar que la forma general del contraste es:

H_0 : No existe autocorrelación

H_1 : Existe autocorrelación

Esto es, en la hipótesis nula se considera que el término de perturbación correspondiente a una observación es independiente del correspondiente a cualquier otra observación. En la hipótesis alternativa se señala que el término de error de un modelo econométrico está autocorrelacionado a través del tiempo. Esta hipótesis alternativa, al considerar la existencia de un patrón de comportamiento para los residuos, se puede especificar con procesos autorregresivos —AR(p)—, de medias móviles —MA(q)— o mixtos —ARMA(p,q)— dependiendo del contraste que se vaya a utilizar.

Se presentan a continuación distintos contrastes que permiten detectar si las perturbaciones están o no autocorrelacionadas y, en caso de estarlo, bajo qué esquema.

4.1. Contraste d de Durbin-Watson (1951)

El contraste desarrollado por Durbin y Watson³ es la prueba más frecuentemente empleada para detectar la presencia de autocorrelación en los modelos de regresión. Este contraste permite verificar la hipótesis de no autocorrelación frente a la alternativa de autocorrelación de primer orden bajo un esquema autorregresivo —AR(1): $u_t = \mathbf{r}u_{t-1} + \mathbf{e}_t$ —

Analíticamente el contraste se especifica del siguiente modo:

Formulación de las hipótesis:

$$\begin{aligned} H_0 : \mathbf{r} &= 0 && \text{No existe autocorrelación AR(1)} \\ H_1 : 0 &< |\mathbf{r}| < 1 && \text{Existe autocorrelación AR(1)} \end{aligned}$$

La forma concreta de la hipótesis alternativa establece unas cotas para el coeficiente de correlación; éstas son necesarias para garantizar algunas características del modelo, en concreto que la varianza es finita y se trata por tanto de un proceso no explosivo.

Estadístico de prueba:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

A partir de este estadístico se puede interpretar que,

- Si hay autocorrelación positiva las diferencias entre residuos que distan un periodo es muy pequeña por lo que el valor del estadístico d será próximo a cero.
- Si hay autocorrelación negativa los residuos serán prácticamente iguales pero de signo contrario, su diferencia será por tanto grande y el estadístico será más próximo al límite superior que, como se verá, se establece en cuatro.
- Si no hay autocorrelación, la relación entre los residuos será intermedia y por tanto, el valor del estadístico experimental también alcanzará un valor intermedio.

Para establecer los límites de variación del estadístico d la fórmula anterior se puede desarrollar obteniéndose una expresión en función del coeficiente de autocorrelación muestral de primer orden para los residuos $\hat{\mathbf{r}}$,

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} = \frac{\sum_{t=2}^n (e_t^2 + e_{t-1}^2 - 2e_t e_{t-1})}{\sum_{t=1}^n e_t^2} = \frac{\sum_{t=2}^n e_t^2 + \sum_{t=2}^n e_{t-1}^2 - 2\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}$$

dado que, cuando el tamaño de la muestra es grande, se puede considerar que

$\sum_{t=2}^n e_t^2 \approx \sum_{t=2}^n e_{t-1}^2 \approx \sum_{t=1}^n e_t^2$ entonces el estadístico d se puede expresar como,

$$d \approx \frac{2\sum_{t=2}^n e_t^2 + 2\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}$$

y dado que el coeficiente de correlación empírico de primer orden se calcula, $\hat{r} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}$

entonces el estadístico experimental se puede expresar

$$d \approx 2(1 - \hat{r})$$

Teniendo en cuenta los límites de variación del coeficiente de correlación empírico, $-1 \leq \hat{r} \leq 1$, se puede deducir el rango de variación del estadístico de Durbin-Watson y el signo de la autocorrelación,

$\hat{r} = -1 \rightarrow d \approx 4$ se considera que existe autocorrelación negativa

$\hat{r} = 0 \rightarrow d \approx 2$ indica ausencia de autocorrelación

$\hat{r} = 1 \rightarrow d \approx 0$ se puede admitir que existe autocorrelación positiva

Así, se aprecia que el estadístico experimental tomará valores entre 0 y 4 de tal modo que cuánto más próximo a cero (a cuatro) sea el valor del estadístico d mayor es la evidencia de autocorrelación positiva (negativa). Si el valor del estadístico experimental d es dos, entonces la correlación muestral será nula y por tanto no se detectará un problema de autocorrelación entre las perturbaciones.

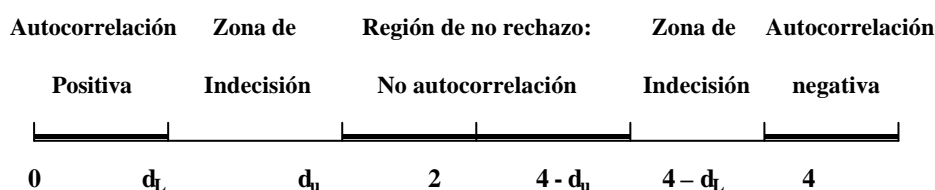
³ J. Durbin y G.S. Watson "Testing for Serial Correlation in Least-Squares Regression" Biometrika, Vol. 38, 1951, pp. 159-171.

No obstante, estos valores (0, 2 y 4) son límites extremos que deben matizarse estableciendo regiones más amplias en las que pueda considerarse si existe o no autocorrelación y, en caso de detectarse, si ésta es positiva o negativa. En este sentido es necesario precisar que la distribución teórica de este estadístico no es sencilla y depende de los valores concretos de la matriz de regresores; por tanto, no existe un valor crítico único que permita establecer una regla de decisión. Para solucionar esta dificultad Durbin y Watson hallaron unos límites superior (d_u) e inferior (d_L) que permiten tomar decisiones acerca de la presencia o ausencia de autocorrelación.

Estos valores señalan el límite superior (d_u) para considerar autocorrelación positiva —esto es, para valores del estadístico experimental superiores a este límite no se rechaza la hipótesis de ausencia de autocorrelación— y el límite inferior (d_L) para no rechazar la hipótesis nula y suponer que las covarianzas de las perturbaciones del modelo son nulas y, por tanto, no están autocorrelacionadas.

Si el valor del estadístico d es superior a dos se puede contrastar la hipótesis nula de no autocorrelación frente a la alternativa de autocorrelación negativa. El análisis es similar pero considerando el valor máximo de 4 como límite para la autocorrelación negativa por tanto los límites anteriores se establecen en los puntos $4-d_u$ y $4-d_L$.

Gráficamente se pueden señalar las regiones del contraste en el siguiente segmento:



Esto es,

$0 < d < d_L$	se rechaza H_0 , existe entonces autocorrelación positiva con un esquema AR(1)
$4-d_L < d < 4$	se rechaza H_0 , existe autocorrelación negativa con un esquema AR(1)
$d_u < d < 4-d_u$	no se rechaza H_0 , no existe autocorrelación
$d_L < d < d_u$	el contraste no es concluyente
$4-d_u < d < 4-d_L$	el contraste no es concluyente

Estos límites dependen del tamaño de la muestra (n) y del número de regresores del modelo (k). Las tablas originales sirven para muestras entre 15 y 100 observaciones y un máximo de 5 regresores. Años más tarde, Savin y White⁴ (1977) publicaron unas tablas más completas que incluyen tamaños de muestra superiores — $5 < n < 200$ — y hasta 20 regresores.

El tratamiento empírico de este contraste requiere de las siguientes fases:

- 1) Estimación por mínimos cuadrados ordinarios (MCO) del modelo de regresión
- 2) Cálculo de los residuos MCO
- 3) Obtención del estadístico d (experimental) de Durbin-Watson
- 4) Búsqueda de los *niveles críticos* del contraste
- 5) Aplicación de la regla de decisión

Un inconveniente que presenta este contraste es que a veces puede no ser concluyente, por lo que hay que considerar, utilizando otros criterios, si existe o no autocorrelación. En este sentido una solución clásica consiste en *ampliar*⁵ las regiones de rechazo considerando así que existe autocorrelación positiva para valores de d inferiores a d_u y autocorrelación negativa si los valores del estadístico experimental son superiores a $4-d_u$.

Este estadístico de uso frecuente, y también generalmente implementado en los programas y aplicaciones informáticas de Econometría, se basa en un conjunto de supuestos acerca de los cuales es necesario reflexionar.

- 1) En primer lugar hay que señalar que el diseño original del contraste se basó en el análisis de un modelo de regresión que incluía término independiente. No obstante, este requisito exigible al modelo fue posteriormente resuelto. En 1980, Farebrother⁶ calculó los valores críticos del contraste para los modelos en los que no existe término independiente.

⁴ N.E. Savin y K.J. White, "The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes of Many Regressors", *Econometrica* Vol., 45, 1977, pp. 1989-1996.

⁵ Un procedimiento práctico de tipo conservador consiste en tomar la decisión de rechazar la hipótesis nula para valores del estadístico experimental inferiores al límite d_u . En este sentido se ha comprobado que las consecuencias de considerar autocorrelación cuando no existe son *menos graves* que las del supuesto contrario; estos resultados se encuentran detallados en H. Theil, A.L. Nagar "Testing the Independence of Regression Disturbances", *Journal of the American Statistical Association*, *Econometrica* Vol. 56, 1961, pp. 793-806 y E.J. Hannan y R.D. Terrel "Testing for Serial Correlation after Least Squares Regression", *Econometrica*, Vol. 36, 1966, pp. 646-660.

2) Junto con la necesidad de término independiente en el modelo, es también un requisito que la matriz de variables explicativas sea no aleatoria, esto es determinista y fija en un muestreo repetido. Por tanto, este contraste no es válido en modelos dinámicos que consideren como regresor retardos de la variable dependiente.

3) La hipótesis alternativa considera que, si las perturbaciones están autocorrelacionadas, el proceso que las genera es autorregresivo de orden 1; esto es, $u_t = \rho u_{t-1} + \epsilon_t$. Sin embargo, se ha comprobado que este estadístico es robusto frente a otras especificaciones de la hipótesis alternativa y, además permite detectar errores de especificación derivados de falta de especificación dinámica y/o de la omisión de variables que estén correlacionadas.

Una prueba más del gran interés que demostró este contraste es el número de económetras que tratan de tabular los niveles críticos para distintos supuestos en los que se basa el contraste de Durbin-Watson. En este sentido podemos citar los siguientes desarrollos:

- 1) Wallis (1972) desarrolló un estadístico similar para modelos de regresión con datos trimestrales en los que se trata de probar la autocorrelación de orden 4 tanto en modelos que incluyen únicamente término independiente como en los modelos que incluyen variables estacionales trimestrales para especificar el nivel autónomo.
- 2) Savin y White (1977) ampliaron las tablas inicialmente propuestas por Durbin y Watson considerando muestras de 6 a 200 observaciones y hasta 20 regresores.
- 3) Ding y Giles (1978) amplían las tablas de Wallis considerando más niveles de significación.
- 4) Farebrother (1980) tabuló los niveles críticos para el supuesto de modelo de regresión sin término independiente.
- 5) King (1981) proporciona los niveles críticos al nivel de significación del 5% para datos de series de tiempo trimestrales y con variables de tendencia y/o variables indicadoras de estacionalidad considerando autocorrelación de cuarto orden.
- 6) King (1983) obtiene los valores d_L y d_U para datos mensuales en los que es conveniente probar una autocorrelación de duodécimo orden.

⁶ Farebrother "The Durbin-Watson Test for Serial Correlation when there is No Intercept in the Regression" *Econometrica*, Vol. 48, 1980, pp. 1553-1563.

4.2. Contraste d_4 de Wallis⁷ (1972)

Este contraste presenta una modificación del estadístico de Durbin-Watson para los modelos que utilizan datos trimestrales en los que, dado el carácter estacional de estas series, se espera que la perturbación de una observación concreta no esté relacionada con la perturbación del periodo inmediatamente anterior sino que dependa de la perturbación del mismo trimestre pero del año anterior, es decir, que la estructura de autocorrelación sea $u_t = \mathbf{j} u_{t-4} + \mathbf{e}_t$

El contraste plantea en la hipótesis nula la ausencia de autocorrelación,

$$H_0 : \mathbf{j} = 0 \quad \text{No existe autocorrelación}$$

$$H_1 : 0 < |\mathbf{j}| < 1 \quad \text{Existe autocorrelación}$$

Para verificar si esta estructura de autocorrelación es o no cierta Wallis propone una

modificación del estadístico de Durbin-Watson que denomina d_4 :
$$d_4 = \frac{\sum_{t=5}^n (e_t - e_{t-4})^2}{\sum_{t=1}^n e_t^2}$$

Este estadístico también fue tabulado por Wallis bajo el supuesto de modelo de regresión con un único término independiente y también para el caso de regresiones que incluyan término independiente y variables ficticias estacionales⁸ (trimestrales). Al igual que el contraste de Durbin-Watson, el estadístico d_4 se ha tabulado suponiendo que la matriz de regresores es no aleatoria y suponiendo también que el modelo tiene término independiente.

Además de este contraste, King⁹ (1983) desarrolló otra modificación del estadístico de Durbin-Watson. En este caso se obtuvieron los valores de los límites superiores (d_U) e inferiores (d_L) de autocorrelación para definir las regiones de rechazo, indecisión y no rechazo cuando se trabaja con datos mensuales.

⁷ K.F. Wallis "Testing for Fourth Order Autocorrelation in Quaterly Regression Equations" *Econometrica*, Vol. 40, 1972, pp. 617-636.

⁸ Para otros niveles de significación pueden consultarse las tablas de D.E.A Giles y M.L. King "Fourth-Order Autocorrelation: Further Significance Points for the Wallis Test" *Journal of Econometrics*, Vol. 8, 1978, pp. 255-259.

⁹ M.L. King "The Durbin Watson Test for Serial Correlation: Bounds for Regressions with Trends and/or Seasonal Dummy Variables", *Econometrica*, Vol. 49, 1981, pp. 1571-1581.

4.3. Contraste h de Durbin¹⁰ (1970)

El contraste de Durbin-Watson, como ya se ha especificado anteriormente, impone como condición para su correcta interpretación que los modelos contengan regresores exclusivamente no aleatorios; con lo cual no se puede aplicar en modelos dinámicos en los que se considere como regresor algún retardo de la variable dependiente. Para corregir esta deficiencia, Durbin¹¹ desarrolló un estadístico que sí puede aplicarse en estos modelos que incluyan retardos de la variable dependiente. Para este caso se ha obtenido un test asintótico para muestras grandes.

La formulación de las hipótesis nula continúa siendo la misma ya que sigue siendo un contraste para la autocorrelación de orden uno bajo un esquema autorregresivo AR(1), $u_t = \rho u_{t-1} + e_t$

$$H_0 : \rho = 0 \quad \text{No existe autocorrelación AR(1)}$$

La hipótesis alternativa, por su parte, se especifica ahora de modo que el contraste se configure como un contraste unilateral; esto es, se van a establecer dos posibles hipótesis alternativas según se considere que la autocorrelación puede ser positiva o negativa. Así, el contraste quedaría especificado

$$H_0 : \rho = 0$$

$$H_1 : \rho < 0$$

$$H_0 : \rho = 0$$

$$H_1 : \rho > 0$$

El estadístico de prueba es $h = \hat{\rho} \sqrt{\frac{n}{1 - n\text{Var}(b_i)}}$ que se distribuye asintóticamente según una distribución $N(0,1)$ lo que, con un nivel de significación del 5%, supone no rechazar la hipótesis nula para los valores de h pertenecientes al intervalo $(-1.645; 1.645)$ ya que se trabaja con un contraste de una sola cola.

Para el cálculo de este estadístico se necesitan conocer los siguientes datos:

- 1) Tamaño de la muestra (n)
- 2) Varianza muestral estimada del coeficiente del regresor aleatorio (Y_{t-1}) en la regresión MCO del modelo a estimar; es decir, obtenida bajo el supuesto de MRLNC [$\text{Var}(b_i)$].

¹⁰ Este contraste fue desarrollado por Durbin a partir de “..una nota de Nerlove y Wallis que afirmaba que la estadística de Durbin y Watson no es aplicable en presencia de variables dependientes rezagadas. Véase M. Nerlove y K.F. Wallis, “Use of Durbin-Watson Statistic in Inappropriate Situations”, *Econometrica*, Vol. 34, 1966, pp. 235-238 (citado en Maddala, 1996).

¹¹ J. Durbin “Testing for Serial Correlation in Least Square Regression when some of the Regressors are Lagged Dependent Variables”, *Econometrica*, Vol. 38, 1970, pp. 410-421.

3) Coeficiente de correlación estimado (\hat{r})

Este coeficiente de correlación estimado se puede calcular a partir de la estimación de una estructura autorregresiva de orden 1 para los residuos —una regresión MCO de los residuos

frente a un retardo de los mismos ($e_t = \rho e_{t-1} + \epsilon_t$); esto es, $\hat{r} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_{t-1}^2}$

Otra posibilidad consiste en calcular esta correlación muestral a partir del valor del estadístico de

prueba del contraste de Durbin-Watson, $\hat{r} \approx 1 - \frac{d}{2}$

El procedimiento de contrastación requiere de la realización de las siguientes fases:

- 1) Estimación MCO del modelo de regresión y obtención de la varianza estimada del coeficiente del regresor aleatorio, $\text{Var}(b_i)$
- 2) Cálculo del coeficiente de correlación estimado
- 3) Cálculo del valor del estadístico experimental h
- 4) Aplicación de la regla de decisión. Si $h > 1,645$ se rechaza la hipótesis nula al nivel de significación del 5% considerando entonces que existe autocorrelación positiva de primer orden. Para el caso de autocorrelación negativa de primer orden, el valor del estadístico experimental h debe ser inferior a $-1,645$.

El principal inconveniente que tiene este contraste es que si el radicando es negativo, esto es $[\text{Var}(b_i) > 1]$, entonces el test falla. Para estos casos Durbin propuso un procedimiento asintótico equivalente¹² y que consiste en lo siguiente:

- 1) Estimar por MCO el modelo de regresión y obtener la serie de residuos MCO
- 2) Estimar una regresión auxiliar en la que los residuos MCO se especifiquen como función de todos los regresores del modelo y también se incluya como regresor adicional un retardo de los residuos.
- 3) Analizar, utilizando el estadístico t habitual, la significación individual del retardo de los residuos de la regresión auxiliar. Si el coeficiente del retardo del residuo es

¹² La utilización de esta *alternativa* no siempre es adecuada; en este sentido se puede señalar que un estudio de Monte Carlo de Maddala y Rao sugiere que esta prueba no es muy potente cuando no puede aplicarse el contraste del estadístico h . (Maddala y Rao "Test for Serial Correlation").

significativamente distinto de cero entonces se considera que existe autocorrelación de primer orden.

Este procedimiento sirvió de base para el contraste, más general, de Breusch-Godfrey (1978) que como se verá a continuación permite contrastar la existencia de otras estructuras de autocorrelación distintas a las autorregresivas de primer orden.

4.4. Contraste de Breusch-Godfrey¹³ (1978)

Los contrastes anteriores, a pesar de su validez y *robustez* para detectar autocorrelaciones de órdenes superiores, se diseñaron inicialmente para contrastar la presencia de procesos autorregresivos de primer orden por lo que el procedimiento adecuado, una vez detectado un problema de autocorrelación, consistirá en el análisis de otros procesos de autocorrelación, ya sean autorregresivos de orden superior, procesos de medias móviles o procesos mixtos.

En este sentido, el contraste de Breusch-Godfrey se especifica con la finalidad de analizar si existe o no autocorrelación de orden superior a uno; para ello, en la hipótesis alternativa se incluyen especificaciones más generales que la del modelo autorregresivo de primer orden y que se pueden generalizar a cualquier especificación ARMA(p,q).

En la hipótesis nula se considera ahora que no existe autocorrelación; la hipótesis alternativa especificará un esquema concreto de autocorrelación.

Por ejemplo, en un modelo autorregresivo de orden p .
$$u_t = \mathbf{j}_1 u_{t-1} + \mathbf{j}_2 u_{t-2} + \dots + \mathbf{j}_p u_{t-p} + \mathbf{e}_t$$

la hipótesis nula se formularía con el supuesto de ausencia de autocorrelación, es decir, nulidad de todos los coeficientes autorregresivos, $H_0 = \mathbf{j}_1 = \mathbf{j}_2 = \dots = \mathbf{j}_p = 0$

Este contraste, al igual que los estudiados hasta el momento, se basa en los residuos MCO y se define como una prueba de significación conjunta de las primeras p autocorrelaciones de los residuos. Para su aplicación empírica es necesario desarrollar las siguientes etapas:

- 1) Estimación por MCO del modelo de regresión y obtención de los residuos MCO (\mathbf{e}_t)
- 2) Estimación de una regresión auxiliar de los residuos \mathbf{e}_t sobre p retardos de los mismos, $\mathbf{e}_{t-1}, \mathbf{e}_{t-2}, \dots, \mathbf{e}_{t-p}$, así como sobre las variables explicativas del modelo original.
- 3) Obtención del coeficiente de determinación (R^2) de la regresión auxiliar.
- 4) Forma del estadístico experimental, $\mathbf{c}_{\text{exp}}^2 = n \cdot R^2$ que se distribuye, bajo la hipótesis nula de no autocorrelación como \mathbf{c}_p^2 , donde p es el número de retardos de los residuos

¹³ T.S. Breusch y L.G. Godfrey "Testing for Autocorrelation in Dynamic Linear Models", Australian Economic Papers, Vol. 17, 1978, pp. 334-355. Y L. G. Godfrey "Testing against General Autorregresive and Moving Average Error Models when the regressors include lagged dependent variables", Econometrica, Vol. 46, 1978, pp. 1293-1302.

incluidos en la regresión auxiliar; esto es, el orden de autocorrelación que se está contrastando; n es el número de observaciones del modelo.

- 5) Regla de decisión: si el valor del estadístico experimental excede del estadístico teórico entonces hay evidencia suficiente para rechazar la hipótesis nula y admitir que existe autocorrelación; en caso contrario no sería correcto rechazar la ausencia de autocorrelación.

Este contraste presenta algunas ventajas frente al estadístico de Durbin-Watson; se puede considerar que el contraste de Breusch-Godfrey puede utilizarse en modelos que incluyan como regresores algunos retardos de la variable endógena, sin que por ello cambien las propiedades del contraste.

En segundo lugar se puede señalar que este contraste permite especificar en la hipótesis alternativa cualquier esquema de autocorrelación ya sea a través de un proceso autorregresivo, de medias móviles o mixto.

A pesar de estas ventajas que lo pueden hacer preferible al contraste de Durbin Watson, no hay que olvidar que para la aplicación de este contraste es necesario especificar una longitud del retardo y que ésta se determinará por un procedimiento de experimentación basado en el análisis de significación individual de los retardos de los residuos, lo que en principio podría dificultar la tarea de selección del orden de autocorrelación.

4.5. Contraste de Sargan (1964)

Cuando el resultado del contraste de Durbin-Watson indica que debe rechazarse la hipótesis nula el origen de esta posible autocorrelación puede deberse a la existencia de errores en la especificación del modelo. Por ejemplo, la omisión de variables relevantes llevaría a incluir esas variables omitidas en el término de perturbación; si estas variables estuvieran correlacionadas podría detectarse un problema de autocorrelación en la perturbación cuando el verdadero origen de ésta se debe a la falta de especificación de aquéllas.

En este sentido Sargan¹⁴ (1964) y posteriormente Hendry y Mizon¹⁵ (1978) buscaron una forma de distinguir, en los casos en que el estadístico de Durbin-Watson detecta la presencia de autocorrelación, si ésta se debe a un error de especificación dinámica o si es realmente un problema de las perturbaciones del modelo.

Así, puede apreciarse cómo un modelo de regresión con perturbaciones AR(1); esto es $Y_t = \mathbf{b} X_t + u_t$ con perturbación $u_t = \mathbf{r} u_{t-1} + \mathbf{e}_t$, puede escribirse como un modelo dinámico autorregresivo, $Y_t = \mathbf{r} Y_{t-1} + \mathbf{b} X_t - \mathbf{r} \mathbf{b} X_{t-1} + \mathbf{e}_t$

Reparametrizando el modelo anterior se puede expresar alternativamente como, $Y_t = \mathbf{b}_1 Y_{t-1} + \mathbf{b}_2 X_t + \mathbf{b}_3 X_{t-1} + e_t$, donde debería verificarse la siguiente restricción $H_0 : \mathbf{b}_1 \mathbf{b}_2 + \mathbf{b}_3 = 0$

Sargan sugiere comenzar con una especificación dinámica y contrastar la restricción anterior antes de probar cualquier análisis de autocorrelación. Si como resultado del contraste $H_0 : \mathbf{b}_1 \mathbf{b}_2 + \mathbf{b}_3 = 0$ se obtiene que esa restricción es cierta, esto es que no se rechaza la hipótesis nula, entonces se puede considerar que los dos modelos son idénticos y por tanto no existe error de especificación con lo que debería pasar a contrastarse si existe o no autocorrelación. En el caso en que se rechaza la hipótesis nula entonces los posibles problemas de autocorrelación detectados se pueden referir a un error de especificación al haber omitido en la especificación inicial los regresores dinámicos X_{t-1} e Y_{t-1} .

¹⁴ J.D. Sargan , "Wages and Prices in the United Kingdom: A Study in Econometric Methodology" en P.E. Hart, G. Mills y J.K. Whitaker (editores), *Econometric Analysis for National Economic Planning*, Colston Papers 16 (London: Butterworth, 1964), pp. 25-54.

¹⁵ D.F. Hendry y G.E. Mizon "Serial Correlation as a Convenient Simplification, Nor a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England" *The Economic Journal*, Vol. 88, sept. 1978, pp. 549-563.

En este mismo sentido se manifiestan Johnston y Dinardo (2001) quienes señalan,

“... en la práctica lo normal es que exista escasa información relativa a la especificación; por esta razón aconsejamos desarrollar una especificación rica en variables a partir de la relación original, con objeto de evitar la necesidad de realizar especificaciones complejas del término de perturbación...”

4.6. Análisis de la Función de Autocorrelación y Función de Autocorrelación Parcial

La función de autocorrelación y la función de autocorrelación parcial constituyen una de las herramientas principales en la identificación de las estructuras autorregresivas y de medias móviles.

Los coeficientes de autocorrelación proporcionan información sobre la relación lineal entre los residuos del modelo separados por k unidades temporales. Es decir, indican el grado de correlación entre cada valor del residuo y los desplazados $1, 2, \dots, h$ periodos. Suponiendo que las varianzas de los residuos son constantes a lo largo del tiempo, el coeficiente de autocorrelación teórico, como ya se ha señalado, se expresa como, $r_s = \frac{g_s}{g_0}$

Esto es,
$$r_0 = \frac{g_0}{g_0} = 1 \qquad r_1 = \frac{g_1}{g_0}$$

donde se verifica que $g_0 > 0 \qquad g_s = g_{-s} \qquad r_s = r_{-s}$

La representación gráfica de estos coeficientes de correlación para un conjunto de retardos sucesivos dará lugar a la función de autocorrelación simple (FAS).

Junto con estos coeficientes de autocorrelación se definen los de autocorrelación parcial que mide la correlación entre dos momentos de tiempo después de eliminar el efecto de los momentos intermedios. La representación gráfica de estos coeficientes recibe el nombre de función de autocorrelación parcial (FAP).

El concepto de autocorrelación parcial es análogo al concepto de coeficiente de regresión parcial. En el MRLNC el k ésimo coeficiente de regresión mide el cambio en el valor medio de la variable dependiente ante un cambio unitario en el regresor k , manteniendo constante la influencia de los demás regresores (o aislando los efectos de los posibles restantes variables explicativas). De modo similar la autocorrelación parcial mide la correlación entre observaciones que están separadas k periodos de tiempo manteniendo constantes las correlaciones en los retardos intermedios. Luego estos coeficientes pueden obtenerse a partir de la siguiente expresión,

$$u_t = \mathbf{j}_1 u_{t-1} + \mathbf{j}_2 u_{t-2} + \mathbf{j}_3 u_{t-3} + \dots + \mathbf{j}_p u_{t-p} + a_t$$

donde \mathbf{j}_p es el coeficiente de autocorrelación parcial ya que mide el efecto adicional de la variable u_{t-p} sobre u_t . Más concretamente recoge el efecto que sobre la variable u_t tiene un retardo de esta misma variable, aislados los efectos de las posibles restantes retardos o considerando éstos como constantes.

El cálculo empírico de estos coeficientes no es materia objeto de estudio de este capítulo; no obstante, podemos señalar que a través de las ecuaciones de Yule-Walker (1930) se pueden establecer las relaciones entre los coeficientes de autocorrelación (\mathbf{r}) y los coeficientes de autocorrelación parcial (\mathbf{j}).

Los patrones que siguen las FAS y FAP son distintos para los procesos AR y MA, por lo que pueden utilizarse para la identificación de la serie de residuos del modelo. En el caso concreto de los procesos autorregresivos AR(p) la FAS decrece geométrica o exponencialmente y la FAP se corta después de p -retardos, esto es presenta p coeficientes significativos. Para el caso de procesos de medias móviles MA(q) ocurre al contrario; la función de autocorrelación parcial decrece de forma geométrica o exponencial y la función de autocorrelación simple se anula para retardos de orden superior a q .

En las estimaciones prácticas estos comportamientos no son tan precisos por lo que a través de contrastes de hipótesis deberá determinarse cuándo un coeficiente estimado (de autocorrelación o de autocorrelación parcial) es considerado nulo a pesar del valor empírico que presente. Para ello se realizan contrastes de significatividad estadística de los coeficientes estableciendo unas bandas de confianza por encima de las cuáles los coeficientes resultan significativos.

Estas bandas pueden calcularse a partir del coeficiente de correlación empírico que presenta la siguiente distribución de probabilidad:

$$\mathbf{r}_k \rightarrow AN \left(0, \sqrt{\frac{1}{n} \left(1 + 2 \sum_{s=1}^{k-1} \hat{\mathbf{r}}_s^2 \right)} \right)$$

Aunque la varianza no es constante, el programa EViews considera que $\sum_{s=1}^{k-1} \hat{r}_s^2 = 0$ y dibuja unas bandas de fluctuación paralelas. Si todos los coeficientes de correlación se sitúan dentro de estos límites el proceso se considera de Ruido Blanco. Cuando existen coeficientes que no se sitúan dentro de las bandas habrá que buscar el patrón de comportamiento según un esquema autorregresivo (AR), de medias móviles (MA) o mixto (ARMA).

Este método gráfico se complementa con otros métodos numéricos como son los contrastes Q de Box-Pierce (1970) y Q de Ljung-Box (1978) que se basan en un análisis de significación de un conjunto de retardos de los residuos. Esto es, la hipótesis nula se formula considerando la ausencia de autocorrelación lo que equivale a considerar la nulidad de un conjunto de estos residuos; $H_0 : r_1 = r_2 = \dots = r_m = 0$

Estos contrastes son muy similares y se definen en ambos casos a partir de la suma acumulada de los coeficientes de correlación empíricos con las siguientes especificaciones.

4.7. Contraste de Box-Pierce-Ljung

Box y Pierce¹⁶ desarrollaron un estadístico que, basado en los cuadrados de los primeros coeficientes de autocorrelación de los residuos, permite analizar si existe o no autocorrelación.

El estadístico se define como una suma acumulada de estos cuadrados de los coeficientes de correlación empíricos; esto es,

$$Q = n \sum_{j=1}^p \hat{\mathbf{r}}_j^2 \quad \text{siendo } \hat{\mathbf{r}}_j = \frac{\sum_{t=j+1}^n e_t e_{t-j}}{\sum_{t=1}^n e_t^2}$$

Bajo la hipótesis nula de no autocorrelación el estadístico Q se distribuye asintóticamente según una χ^2 con grados de libertad igual a la diferencia entre el número de coeficientes acumulados (p) y el número de parámetros estimados al ajustar el proceso ARMA que se considere.

Posteriormente este estadístico fue revisado por Ljung-Box obteniéndose mejores resultados para muestras pequeñas si se utiliza esta otra expresión alternativa.

$$Q' = n(n+2) \sum_{j=1}^p \frac{\hat{\mathbf{r}}_j^2}{(n+j)}$$

Estos estadísticos se definieron inicialmente para el análisis de Series Temporales pero a veces también se utilizan para verificar la hipótesis de autocorrelación en los modelos de regresión. No obstante, esta aplicación en modelos estructurales debe realizarse con cautela ya que la inclusión de variables exógenas en el modelo tiene un efecto desconocido sobre el estadístico experimental¹⁷.

Sin embargo, puesto que se trata de un procedimiento implementado en EViews se presenta aquí como una primera aproximación al análisis de autocorrelación que deberá estudiarse con

¹⁶ G.E.P. Box y D. A. Pierce "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models", Journal of the American Statisticak Association, vol. 65, 1970, pp. 1509-1526.

más detalle con los contrastes presentados anteriormente y que se diseñaban específicamente para modelos de regresión estructurales.

¹⁷ H. Dezhbaksh "The Inapropriate Use of Serial Correlation Test in Dynamic Linears Models", Review of Economics and Statistic, Vol. LXXII, 1990, pp. 126-132.

5. ESTIMACIÓN, INFERENCIA Y PREDICCIÓN

En el capítulo dedicado al análisis del modelo de regresión generalizado ya se ha estudiado que, en este modelo, los estimadores MCO son lineales, insesgados pero no óptimos, por lo que para conseguir una adecuada estimación de los parámetros del modelo y poder realizar los procesos de inferencia se utilizarán los estimadores de mínimos cuadrados generalizados (MCG) también estudiados en el tema once.

En la estimación e inferencia de los parámetros del modelo la autocorrelación origina problemas similares a los ya analizados en el modelo con heterocedasticidad: la estimación por MCO continúa siendo insesgada pero ahora es ineficiente con lo que los procesos de inferencia quedan invalidados.

Dado que los estimadores de MCO no son eficientes —su matriz de varianzas y covarianzas está mal calculada— los procedimientos de inferencia habituales no son válidos ya que es probable que se obtengan resultados que lleven a considerar que los coeficientes β no son estadísticamente significativos —es decir, coeficientes iguales a cero— cuando en realidad sí lo sean. En este caso es aconsejable utilizar los estimadores de Mínimos Cuadrados Generalizados (MCG) o Mínimos Cuadrados Generalizados Factibles (MCGF) ya que presentan mejores propiedades. La forma específica de estos estimadores dependerá del proceso subyacente para la perturbación.

En este sentido Gujarati (1997) concluye que,

“Para establecer intervalos de confianza y probar hipótesis debe utilizarse MCG y no MCO, aún cuando los estimadores derivados de este último sean insesgados y consistentes” y “... las pruebas de significación t y F usuales dejan de ser válidas y, de ser éstas aplicadas, es probable que conduzcan a conclusiones erróneas sobre la significación estadística de los coeficientes de regresión estimados.”

En este epígrafe se aborda la estimación de un modelo de regresión lineal, homocedástico y con perturbaciones autocorrelacionadas. Los distintos procedimientos de estimación que aquí se presentan parten del supuesto de modelos de regresión con matriz de regresores no estocástica; el supuesto de modelos con variables dependientes retardadas —modelos autorregresivos— no es objeto de estudio en este tema.

Se va a estudiar por tanto la estimación en un modelo que como ya se ha señalado cumple las siguientes hipótesis

$$Y = X\mathbf{b} + u$$

$$E(u) = 0$$

$$E(uu') = \mathbf{S}^2 \Omega \quad \Omega \text{ matriz con varianzas constantes y covarianzas no nulas}$$

Para este modelo que presentamos existen distintos métodos de estimación que, básicamente difieren en la información que se tenga acerca de la matriz Ω según que esta sea conocida o desconocida y deba por tanto estimarse.

Si la matriz Ω es conocida se aplican, al igual que se señaló para el problema de heterocedasticidad, MCG ya que el modelo con el que se está trabajando no es un MRLNC sino un MRLG.

Si la matriz Ω es desconocida ésta deberá de estimarse en función del proceso de autocorrelación del modelo. Una vez estimada esta matriz se aplicará lo que se ha dado en llamar MCGF. Estos estimadores de MCGF conservan las propiedades asintóticas de los MCG siempre que los estimadores de la primera etapa —esto es, la estimación de la matriz Ω —sean consistentes.

- Estimación práctica del modelo: aplicación con el programa Eviews

En la práctica la estimación que se va a utilizar consiste en la aplicación directa de alguno de los procedimientos que tiene implementados Eviews. Este programa permite estimar el modelo de regresión por MCO incorporando la estructura de autocorrelación sin más que añadir, a los regresores del modelo, la especificación concreta de la perturbación; esto es el tipo de proceso y el orden del mismo. Así, para el supuesto de un proceso autorregresivo de segundo orden la estimación en Eviews se realizará a partir de los siguientes comandos,

Quick/ Estimate Equation/ Y C X2 X3 AR(1) AR(2)

Si el proceso fuera de medias móviles los términos que se añaden son MA(1), MA(2), ... MA(q) y en caso de un proceso mixto se incluirían tanto los términos autorregresivos como los de medias móviles.

Otra posibilidad que ofrece Eviews consiste en la estimación del modelo por MCO pero con una estimación consistente de la matriz de varianzas y covarianzas de los coeficientes de modo que sean válidos los procesos de inferencia. Para ello, desde la ventana en que se especifica la estimación de MCO se activa el botón Options y se selecciona Heterokedaticiy y la estimación con matriz de varianzas y covarianzas de Newey-West (recuérdese que para el caso de heterocedasticidad la opción a seleccionar era la matriz de White).

El análisis de predicción en los modelos con autocorrelación necesita incluir la estructura de la perturbación; este análisis se va a desarrollar unicamente para el caso de procesos autorregresivos de primer orden. La aplicación para otras estructuras se realizará automáticamente con las opciones de Eviews.

6. PROCESO AUTORREGRESIVO DE PRIMER ORDEN —AR(1)—

El proceso autorregresivo de primer orden —AR(1)— es el proceso más frecuentemente analizado tanto desde un punto de vista teórico como empírico; en este sentido se puede señalar que *“la literatura empírica está aplastantemente dominada por el modelo AR(1) ... [este modelo] ha resistido los contrastes de tiempo y experimentación mostrándose como un modelo razonable para el proceso subyacente que probablemente en realidad, es complejamente opaco.”* (Greene, 1999)

Es necesario señalar también que los procesos de órdenes superiores son, con frecuencia, extremadamente difíciles de analizar y que gran parte de los modelos económicos con problemas de autocorrelación presentan un proceso autorregresivo de primer orden por lo que éste es el supuesto que se considera ahora para analizar la estimación de los modelos de regresión homocedástico y con problemas de autocorrelación.

Definimos ahora el proceso autorregresivo de orden 1 a partir de la estructura,

$$u_t = \mathbf{j} u_{t-1} + \mathbf{e}_t$$

donde \mathbf{e}_t es una variable que se denomina de ruido blanco y que verifica las siguientes propiedades,

$$E(\mathbf{e}_t) = 0 \quad \forall t \quad \text{var}(\mathbf{e}_t) = \mathbf{S}_e^2 \quad \forall t$$

$$E(\mathbf{e}_t \mathbf{e}_{t-1}) = 0 \quad t \neq s \quad \Sigma_{ee} = E(\mathbf{e}\mathbf{e}') = \mathbf{S}_e^2 I$$

La expresión del proceso AR(1) se puede formular también como un proceso de medias móviles de orden infinito sin más que sustituir de forma continuada la variable retardada; esto es,

$$u_t = \mathbf{e}_t + \mathbf{j} \mathbf{e}_{t-1} + \mathbf{j}^2 \mathbf{e}_{t-2} + \dots$$

$$u_t = \sum_{i=0}^{\infty} \mathbf{j}^i \mathbf{e}_{t-i}$$

Las propiedades que verifica el proceso autorregresivo AR(1) son,

$$E(u_t) = 0$$

$$\text{var}(u_t) = E(u_t^2) = \frac{\mathbf{s}_e^2}{1-\mathbf{j}^2} \quad |\mathbf{j}| < 1 \text{ condición necesaria para que la varianza sea finita}$$

Covarianzas

$$E(u_t u_{t-1}) = \mathbf{j} \mathbf{s}_u^2$$

$$E(u_t u_{t-2}) = \mathbf{j}^2 \mathbf{s}_u^2$$

$$E(u_t u_{t-3}) = \mathbf{j}^3 \mathbf{s}_u^2$$

.....

$$E(u_t u_{t-S}) = \mathbf{j}^S \mathbf{s}_u^2$$

Correlaciones

$$\mathbf{r}_1 = \mathbf{j}$$

$$\mathbf{r}_2 = \mathbf{j}^2$$

$$\mathbf{r}_3 = \mathbf{j}^3$$

.....

$$\mathbf{r}_S = \mathbf{j}^S$$

Luego la matriz de varianzas y covarianzas de la perturbación se puede expresar,

$$\Sigma_{uu} = E(uu') = \mathbf{s}_u^2 \begin{bmatrix} 1 & \mathbf{j} & \mathbf{j}^2 & \mathbf{j}^{n-1} \\ & 1 & \mathbf{j} & \mathbf{j}^{n-2} \\ & & 1 & \mathbf{j}^{n-3} \\ & & & \dots \\ & & & & 1 \end{bmatrix} = \frac{\mathbf{s}_e^2}{1-\mathbf{j}^2} \begin{bmatrix} 1 & \mathbf{j} & \mathbf{j}^2 & \mathbf{j}^{n-1} \\ & 1 & \mathbf{j} & \mathbf{j}^{n-2} \\ & & 1 & \mathbf{j}^{n-3} \\ & & & \dots \\ & & & & 1 \end{bmatrix}$$

siendo Ω la matriz,

$$\Omega = \frac{1}{1-\mathbf{j}^2} \begin{bmatrix} 1 & \mathbf{j} & \mathbf{j}^2 & \mathbf{j}^{n-1} \\ & 1 & \mathbf{j} & \mathbf{j}^{n-2} \\ & & 1 & \mathbf{j}^{n-3} \\ & & & \dots \\ & & & & 1 \end{bmatrix}$$

Para la estimación del modelo de regresión aplicando MCG se necesitaría conocer la inversa de esta matriz que es,

$$\Omega^{-1} = \begin{bmatrix} 1 & -\mathbf{j} & 0 & \dots & 0 & 0 & 0 \\ -\mathbf{j} & 1+\mathbf{j}^2 & -\mathbf{j} & \dots & 0 & 0 & 0 \\ 0 & -\mathbf{j} & 1+\mathbf{j}^2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\mathbf{j} & 1+\mathbf{j}^2 & -\mathbf{j} \\ 0 & 0 & 0 & \dots & 0 & -\mathbf{j} & 1 \end{bmatrix}$$

O alternativamente, si el modelo generalizado se quisiera transformar a un modelo clásico utilizando la transformación de Aitken se necesitaría expresar la matriz de paso P que para el supuesto concreto de un modelo autorregresivo de primer orden se formula,

$$P = \begin{bmatrix} \sqrt{1-\mathbf{j}^2} & 0 & 0 & \dots & 0 & 0 \\ -\mathbf{j} & 1 & 0 & \dots & 0 & 0 \\ 0 & -\mathbf{j} & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\mathbf{j} & 1 \end{bmatrix}$$

A partir de esta matriz de paso —también llamada matriz de transformación— se transformarían las variables del modelo del siguiente modo $Y^*=P \cdot Y$, $X^*=P \cdot X$.

Para simplificar la notación vamos a considerar el modelo de regresión simple¹⁸, $Y_t = \mathbf{a} + \mathbf{b}X_t + u_t$ en el que las perturbaciones siguen un esquema autorregresivo de primer orden, $u_t = \mathbf{j} u_{t-1} + \mathbf{e}_t$. Las variables transformadas se calcularían a partir de las siguientes expresiones,

$$Y^* = \begin{bmatrix} (\sqrt{1-\mathbf{j}^2})Y_1 \\ Y_2 - \mathbf{j}Y_1 \\ \dots \\ Y_n - \mathbf{j}Y_{n-1} \end{bmatrix} \quad X^* = \begin{bmatrix} \sqrt{1-\mathbf{j}^2} & (\sqrt{1-\mathbf{j}^2})X_1 \\ 1-\mathbf{j} & X_2 - \mathbf{j}X_1 \\ \dots & \dots \\ 1-\mathbf{j} & X_n - \mathbf{j}X_{n-1} \end{bmatrix}$$

Teniendo en cuenta estas matrices Ω , Ω^{-1} , P se señalan a continuación los distintos procedimientos de estimación para el proceso autorregresivo de primer orden que se está analizando.

1) Estimación con W conocida: Mínimos Cuadrados Generalizados

Para el caso concreto de perturbaciones autocorrelacionadas según un esquema autorregresivo de orden 1 —AR(1)— es relativamente sencillo utilizar este método que se reduciría a la aplicación directa de la fórmula

$$b_* = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} Y$$

Con la forma concreta de la matriz Ω definida anteriormente o transformando las variables y estimando por MCO una regresión de Y^* sobre X^* y teniendo en cuenta que ahora no se debe incluir término constante en el modelo.

Modelo de regresión $Y_t = \mathbf{a} + \mathbf{b} X_t + u_t$ con perturbaciones AR(1) $u_t = \mathbf{j} u_{t-1} + \mathbf{e}_t$

Modelo de regresión expresado en el periodo $t-1$ y multiplicando por ϕ

$$\mathbf{j} Y_{t-1} = \mathbf{j} \mathbf{a} + \mathbf{j} \mathbf{b} X_{t-1} + \mathbf{j} u_{t-1}$$

Se realiza la diferencia entre uno y otro

$$(Y_t - \mathbf{j} Y_{t-1}) = \mathbf{a}(1 - \mathbf{j}) + \mathbf{b} X_t - \mathbf{j} \mathbf{b} X_{t-1} + (u_t - \mathbf{j} u_{t-1})$$

$$(Y_t - \mathbf{j} Y_{t-1}) = \mathbf{a}(1 - \mathbf{j}) + \mathbf{b} (X_t - \mathbf{j} X_{t-1}) + \mathbf{e}_t$$

este modelo se puede expresar como

$$Y_t^* = \mathbf{a}^* + \mathbf{b} X_t^* + \mathbf{e}_t$$

y puesto que ε_t verifica las hipótesis clásicas este modelo —con variables transformadas—se podría estimar por MCO y los estimadores obtenidos serían óptimos; este procedimiento es

¹⁸ Los procedimientos que se van a desarrollar en este epígrafe consideran, por simplicidad, un modelo de regresión simple pero es necesario señalar que la inclusión de otros regresores no cambiaría el procedimiento ya que la corrección se hace para el problema de autocorrelación que es una característica de la perturbación y no de los regresores.

equivalente a la aplicación de MCG con la única diferencia que se prescinde de la primera observación.

Nótese que este procedimiento es similar a “*prescindir*” de la información de la primera fila de la matriz P y estimar un modelo con variable dependiente $(Y_t - \mathbf{j} Y_{t-1})$ sobre una constante y $(X_t - \mathbf{j} X_{t-1})$ utilizando esta transformación desde la segunda a la última observación. Este método, aunque no coincide exactamente con MCG, sí es asintóticamente equivalente.

No obstante, el principal problema que se presenta para aplicar mínimos cuadrados generalizado por cualquiera de estos procedimientos es que el valor de ϕ es desconocido por lo que deberá estimarse como un parámetro más del modelo de regresión aplicándose entonces el método de Mínimos Cuadrados Generalizados Factibles.

2) Estimación con W desconocida: Mínimos Cuadrados Generalizados Factibles

El supuesto analizado en el caso anterior es poco frecuente siendo lo habitual trabajar con modelos que presentan autocorrelación en los que se desconocen los elementos que configuran la matriz Ω . En estos casos es necesario realizar en primer lugar una estimación de dicha matriz que para el caso que estamos analizando —proceso autorregresivo de primer orden— se reduce a una estimación del coeficiente ϕ .

A continuación se señalan algunos procedimientos para estimar el coeficiente de autocorrelación ϕ pero esta lista no es exhaustiva; existen otros métodos, como el de máxima verosimilitud, que aquí no se presentan. Los que aquí se muestran son procedimientos que, básicamente, se desarrollan en dos etapas. En la primera etapa se obtiene una estimación de ϕ que se utiliza, en la segunda etapa, para transformar las variables con las que estimar una ecuación en diferencias generalizada —método este que como ya se ha señalado coincide básicamente con la aplicación de MCG—. Todos estos procedimientos se conocen con el nombre de MCGF ya que, en lugar del verdadero valor del coeficiente ϕ utilizan una estimación del mismo.

- Procedimiento iterativo de Cochrane-Orcutt¹⁹

La estimación planteada por Cochrane-Orcutt es un proceso iterativo que permite estimar el valor del parámetro de autocorrelación desconocido (ϕ).

En una primera etapa se estima el modelo de regresión por MCO y se calcula la serie de residuos MCO; a partir de éstos se realiza una regresión auxiliar de los residuos sobre los residuos del periodo anterior sin incluir término constante. De este modo se obtiene una primera estimación del coeficiente de autocorrelación de primer orden²⁰. A partir de este valor estimado se transforman las variables del modelo de regresión que se utilizan para repetir la estimación del modelo (etapa 1) y continuar con el procedimiento descrito.

Este proceso finaliza cuando el estadístico d de Durbin-Watson indique que los residuos MCO de la etapa 1 son de ruido blanco o, alternatively, se puede finalizar el proceso cuando las estimaciones sucesivas del parámetro ϕ difieran en menos de una cantidad prefijada por ejemplo 0,01 ó 0,005.

Es necesario señalar que la aplicación de este método reduce el número de observaciones de la muestra ya que se omite la información relativa a la primera observación.

- *Procedimiento o Modificación de Prais-Winsten*²¹ (con información completa)

Una modificación del procedimiento de Cochrane-Orcutt fue propuesta por Prais y Winsten (1954); estos autores sugieren *ampliar* el tamaño de muestra incluyendo una transformación para la primera observación que, como consecuencia de la utilización de primeras diferencias, desaparece. En lugar de utilizar el modelo que surge directamente de la transformación se incorpora, para las primeras observaciones de las variables el factor de corrección $\sqrt{(1 - \mathbf{r}^2)}$, con lo que la primera observación para la variable dependiente será $\sqrt{(1 - \mathbf{r}^2)} Y_1$ y

¹⁹ D. Cochrane y G.H. Orcutt, "Application of Least Squares Regressions to Relationships Containing Autocorrelated Error Terms" Journal of the American Statistical Association, Vol. 44, 1949, pp.32-61.

²⁰ La estimación del coeficiente de autocorrelación también se podría calcular a partir del estadístico de Durbin-Watson que

como ya se ha señalado permite obtener, $r \approx 1 - \frac{d}{2}$

²¹ S.J. Prais y C.B. Winsten "Trend Estimation and Serial Correlation" Cowles Commission Discussion Paper, nº 383, Chicago, 1954.

análogamente para la matriz de regresores. Esta modificación permite mejorar la eficiencia en la estimación de muestras pequeñas.

- *Procedimiento en dos etapas de Cochrane-Orcutt*

Este procedimiento se muestra como una versión abreviada del proceso iterativo; la primera etapa sería equivalente y, a continuación, en la segunda y última etapa se estimaría la ecuación en diferencias a partir del coeficiente de autocorrelación estimado.

- *Método de Durbin de dos pasos*

Durbin²² (1960) propone una primera estimación del modelo en cuasi diferencias expresado a partir de la igualdad,

$$Y_t = \mathbf{a}(1-\mathbf{j}) + \mathbf{b}X_t - \mathbf{j} \mathbf{b}X_{t-1} + \mathbf{j} Y_{t-1} + \mathbf{e}_t$$

Como primera estimación del coeficiente de autocorrelación se va a considerar el coeficiente de la variable endógena desplazada, ya que, aunque sesgada, se trata de una estimación consistente de ϕ

Esta primera estimación del coeficiente de autocorrelación puede servir de base para la aplicación de cualquiera de los otros dos métodos presentados con anterioridad. En concreto, Griliches y Rao²³ muestran, a partir de un estudio de Monte Carlo que el estimador obtenido a partir de una primera estimación con el método de Durbin y seguido del procedimiento de Prais-Winsten para las variables transformadas es mejor que otras alternativas. En este sentido Greene (1998) —citando a Harvey y McAvinchey (1981)— afirma que “es bastante peor omitir la primera observación que mantenerla”.

²² J. Durbin “Estimation of Parameters in Time Series Regression Models” Journal of the Royal Statistical Society, ser. B, Vol. 22, 1960, pp. 139-153.

²³ Z. Griliches y P. Rao “Small Sample Properties of Several Two Stage Regression Methods in The Context of Autocorrelated Errors”, Journal of the American Statistical Association, Vol. 64, 1969, pp. 253-272.

Predicción con perturbaciones AR(1)²⁴

Una vez que se ha detecta la presencia de autocorrelación en un modelo de regresión esta información debe considerarse para realizar la predicción; así por ejemplo un modelo de regresión simple con perturbaciones AR(1) la mejor predicción para un periodo extramuestral $n+1$ debe incluir una corrección por autocorrelación; por tanto,

$\hat{Y}_{n+1} = a + b X_{n+1}$ no es la mejor predicción para el periodo $n+1$

ya que $E(u_{n+1}) = \rho u_n$ que podría estimarse a través de $\hat{\rho} u_n = \hat{\rho}(Y_n - a - bX_n)$

Por tanto, la predicción de Y para el periodo $n+1$ debe incorporar esta información obteniéndose el siguiente predictor

$$\hat{Y}_{n+1} = a + b X_{n+1} + \hat{\rho} u_n$$

$$\hat{Y}_{n+1} = a + b X_{n+1} + \hat{\rho}(Y_n - a - bX_n)$$

y este predictor coincide con el que se obtendría habiendo transformado el modelo con cuasidiferencias.

²⁴ Este tema se puede estudiar con más profundidad en a.S. Goldberger "Best Linear Unbiased Prediction in the Generalized Linear Regression Model" Journal of the American Sotstistical Association, Vol. 57, 1962, pp. 369-375.