

MINIMOS CUADRADOS RECURSIVOS*

JOSE LUIS ARRUEAT**

I. Introducción

Los mínimos cuadrados recursivos surgen en la década del 70 en la literatura econométrica, a partir de contribuciones como las de Harvey y Phillips (1974) y Brown, Durbin y Evans (1975). En fechas más recientes se han incorporado a libros de texto, tales como Harvey (1981a., 1981b.) y Johnston (1984).

El propósito de este trabajo es presentar los elementos básicos de este enfoque y la utilidad que presta para la realización de dóxicas de cambio estructural y autocorrelación. Resulta útil, asimismo, para contrastar la presencia de heteroscedasticidad y para inferir si tanto los parámetros de un modelo lineal general como su varianza son constantes a través del tiempo. Sin embargo, estos últimos enfoques no se discutirán en detalle aunque se darán las referencias apropiadas.

El presente trabajo está organizado como sigue: en la segunda sección se presenta un análisis comparativo de los mínimos cuadrados ordinarios y los mínimos cuadrados recursivos (MCO y MCR, de ahora en más), prestando particular atención a las propiedades de los residuos obtenidos por uno y otro método. El tema de la tercera sección es la dóxima de estabilidad estructural o dóxima de Chow, como comúnmente se la llama, a partir del esquema de MCR. Este contraste de hipóte-

(*) Universidad Nacional de Córdoba e Instituto Di Tella

(**) Agradezco especialmente a los Dres. Fernando Ferrero, Aldo A. Aranda, Raúl E. García y Eduardo Aime, por comentarios y sugerencias. Ninguno de ellos es, por supuesto responsable de los errores que pudieran subsistir en este trabajo.

sis forma una parte central de los cursos de Econometría en los cuales se plantea a partir de estimaciones por MCO, cuya demostración es tediosa, mientras que el enfoque de MCR produce demostraciones muy sencillas.

En la cuarta sección se plantea el uso de los residuos obtenidos por MCR para docimar la presencia de autocorrelación de primer orden y se discuten las ventajas y desventajas relativas que comporta este método con relación al tradicional debido a Durbin y Watson.

Los dócimas de Brown, Durbin y Evans (1975), conocidas como Cusum y Cusum de cuadrados se mencionan brevemente en la quinta sección y seguidamente se considera una forma de cálculo alternativa para los MCR que está dada por el filtro de Kalman. La ventaja de este último enfoque es que resulta de aplicación mucho más general para modelos de mínimos cuadrados generalizados. Dicha ventaja para la estimación econométrica ha sido frecuentemente mencionada en la literatura.

Por último, la sexta sección contiene un breve resumen de los temas abordados en el presente trabajo, de las conclusiones obtenidas y presenta además algunas sugerencias para investigaciones futuras.

II. Mínimos Cuadrados Ordinarios (MCO) y Mínimos Cuadrados Recursivos (MCR)

La representación de esta sección sigue, en lo fundamental, la brindada por Johnston (1984, Cap. 10, págs. 384-387).

Dado el modelo lineal general

$$Y = X\beta + u$$

$$u \sim N(0, \sigma^2 I) \quad (2.1)$$

en el que Y es un vector de $(n \times 1)$ de variables dependientes, X una matriz de $(n \times k)$ de variables explicativas, β un vector de $(k \times 1)$ de parámetros a estimar y u es un vector de errores aleatorios de $(n \times 1)$ con distribución normal multivariante de media cero y matriz de varianzas y covarianzas $\sigma^2 I$.

Para estimar β por MCO se utiliza la conocida fórmula

$$b_{MCO} = (X'X)^{-1} X'Y \quad (2.2)$$

que bajo los supuestos aquí considerados es también el estimador máximo-verosímil de β .

Si consideramos a continuación los residuos obtenidos por MCO, vale decir

$$e = Y - Xb_{MCO} \quad (2.3)$$

resulta sencillo demostrar la siguiente relación entre e y u ,

$$e = Mu \quad (2.4)$$

en que M es la siguiente matriz,

$$M = I - X(X'X)^{-1} X' \quad (2.5)$$

Sobre la base de estos resultados se puede calcular la matriz de varianzas y covarianzas de e , la que resulta igual a

$$E(ee') = \sigma^2 M \quad (2.6)$$

dado que la matriz M de (2.5) es idempotente y simétrica y el vector e tiene esperanza igual al vector nulo. Por lo tanto los residuos calculados por MCO exhiben, en general, heteroscedasticidad y autocorrelación, aún cuando los errores u en (2.1) sean homoscedásticos y no presenten autocorrelación y esto causa dificultades para los contrastes de hipótesis referidas a heteroscedasticidad y autocorrelación.

Consideremos a continuación el esquema de MCR. Si simbolizamos por X_{r-1} la matriz de dimensiones $(r-1) \times k$ que surge de incluir las $r-1$ primeras filas de la matriz X definida en (2.1), podemos estimar el vector por MCO como sigue:

$$b_{r-1} = (X'_{r-1} X_{r-1})^{-1} X'_{r-1} Y_{r-1} \quad (2.7)$$

En la expresión (2.7) b_{r-1} simboliza el estimador por MCO basado en las $(r-1)$ primeras observaciones. En forma análoga, Y_{r-1} es el vector de valores de la variable dependiente que contiene las $(r-1)$ primeras

observaciones.

Valiéndose de b_{r-1} se puede predecir el valor de la variable dependiente para el momento r , vale decir, y_r , a partir del conocimiento de los valores de las variables explicativas para igual período r , es decir, x_r . El error de predicción está dado por la siguiente fórmula

$$y_r - x_r' b_{r-1} \quad (2.8)$$

mientras que su varianza es igual a

$$\sigma^2 (1 + x_r' (X_{r-1}' X_{r-1})^{-1} x_r) \quad (2.9)$$

Conviene aclarar en este punto que el vector x_r definido anteriormente es un vector de k filas y una columna. Por ese motivo se considera su transpuesta en (2.8) y (2.9).

Definiremos el residuo recursivo w_r como sigue

$$w_r = \frac{y_r - x_r' b_{r-1}}{f_r^{1/2}} \quad (2.10)$$

en que f_r está dado por la siguiente expresión

$$f_r = 1 + x_r' (X_{r-1}' X_{r-1})^{-1} x_r \quad (2.11)$$

Resulta fácil demostrar que

$$w_r \sim N(0, \sigma^2) \quad (2.12)$$

bajo el supuesto de normalidad usado en (2.1) puesto que es una función lineal de variables normales y la predicción obtenida por MCO es insesgada.

Se puede obtener una secuencia de residuos recursivos aplicando el siguiente procedimiento.

1. Elíjase una base de k observaciones, digamos las k primeras observaciones de la muestra. Calcúlese el vector b_k y el residuo recursivo

$$w_{k+1} = \frac{y_{k+1} - x'_{k+1} b_k}{f_{k+1}^{1/2}} \quad (2.13)$$

en que

$$f_{k+1} = 1 + x'_{k+1} (X'_k X_k)^{-1} x_{k+1} \quad (2.14)$$

2. Extiéndase la base para incluir las primeras $k+1$ observaciones y calcúlese b_{k+1} y w_{k+2} .

3. Repítase el paso 2, agregando una nueva observación cada vez, hasta incluir los n observaciones de que consta la muestra.

De esta manera se ha generado una secuencia de $(n-k)$ residuos recursivos, como el definido en la fórmula (2.10) para $r=k+1, \dots, n$.

Dichos residuos recursivos revisten gran importancia práctica puesto que bajo los supuestos de normalidad (2.1) y por extensión de (2.12), el vector de residuos w_r para $r=k+1, \dots, n$, que simbolizaremos por w tiene distribución normal multivariante con vector de media cero y matriz de varianzas y covarianzas escalar, lo que simbolizaremos como sigue

$$w \sim N(0, \sigma^2 I_{n-k}) \quad (2.15)$$

Nótese que w tiene sólo $n-k$ componentes. La demostración de (2.15) se encuentra en Johnston (1984) y Phillips y Harvey (1974).

El cálculo de los residuos recursivos se puede llevar a cabo de dos formas alternativas. La primera consiste en aplicar MCO repetidamente, comenzando con las k primeras observaciones y agregando una observación por vez hasta agotar las n observaciones de la muestra. Existe, sin embargo, un algoritmo recursivo alternativo que involucra los siguientes cálculos:

$$(X'_r X_r)^{-1} = (X'_{r-1} X_{r-1})^{-1} - \frac{(X'_{r-1} X_{r-1})^{-1} x'_r (X'_{r-1} X_{r-1})^{-1}}{f_r} \quad (2.16)$$

en que f_r se definió en (2.11) y además,

$$b_r = b_{r-1} + (X_r' X_r)^{-1} x_r (y_r - x_r' b_{r-1}) \quad (2.17)$$

La fórmula (2.16) presenta la importante ventaja de que es preciso calcular la inversa de una matriz una sola vez, vale decir cuando $r = k$. A partir de allí, se recurre a (2.16) para $r = k+1, \dots, n$. El cálculo de la fórmula (2.17) no entraña ninguna dificultad adicional.

Existe una alternativa a esta forma de proceder que ha sido sugerida por Harvey (1981b., págs. 192-193) que plantea dos aspectos fundamentales. El primero consiste en reformular la ecuación (2.17) como sigue:

$$b_r = b_{r-1} + (X_{r-1}' X_{r-1})^{-1} x_r \frac{1}{f_r} (y_r - x_r' b_{r-1}) \quad (2.17')$$

Todos los términos que aparecen en esta fórmula han sido definidos previamente. A primera vista no proporciona ninguna ventaja el uso de (2.17') en lugar de (2.17). Con referencia al filtro de Kalman al que aludiremos posteriormente, la expresión

$$(X_{r-1}' X_{r-1})^{-1} x_r \frac{1}{f_r} \quad (2.18)$$

se llama "ganancia de Kalman".

El segundo ingrediente de la propuesta de Harvey, que puede usarse independientemente del primero, consiste en comenzar las iteraciones para el cálculo de los b_i a partir de:

$$b_0 = 0 \quad (2.19)$$

y

$$(X_0' X_0)^{-1} = KI \quad (2.20)$$

en que K es un número finito pero grande, usualmente del orden de 10^6 ó 10^9 e I es la matriz identidad de $k \times k$. Aunque (2.19) se trate de un artificio arbitrario, la fórmula (2.20) determina que el efecto de este procedimiento sobre los coeficientes b_i se torne despreciable des-

pués de k iteraciones. No he tenido acceso a ninguna demostración formal de esta proposición pero sí he realizado cálculos numéricos que la avalan.

Usando las ecuaciones (2.16) y (2.17) se puede demostrar que:

$$RSS_r = RSS_{r-1} + w_r^2 \quad r = k + 1, \dots, n \quad (2.21)$$

en la que RSS_r simboliza la suma de los cuadrados de los residuos basada en las r primeras observaciones y se define como sigue:

$$RSS_r = (Y_r - X_r b_r)' (Y_r - X_r b_r) \quad (2.22)$$

En la tercera sección aplicaremos los residuos recursivos para efectuar dócimas de cambio estructural. Antes de finalizar esta sección parece oportuno puntualizar lo siguiente:

Primero: Una vez procesadas las n observaciones de la muestra, el estimador b_n es, salvo errores de redondeo, exactamente igual al estimador b_{MCO} de la fórmula (2.2).

Segundo: La suma de cuadrados de residuos obtenidos a partir de las n observaciones mediante (2.22) es igual a:

$$RSS_n = e'e \quad (2.23)$$

en que e está definido por la fórmula (2.3). Esto, obviamente, no es novedoso puesto que se trata de la forma habitual de calcular la suma de los cuadrados de los residuos basados en el método de MCO. Veremos a continuación como realizar un cálculo equivalente a partir de los residuos recursivos. Si utilizamos (2.21) con $r = k+1$, se obtiene

$$RSS_{k+1} = RSS_k + w_{k+1}^2 \quad (2.24)$$

pero obviamente $RSS_k = 0$, por lo que

$$RSS_{k+1} = w_{k+1}^2 \quad (2.25)$$

Procediendo de esta manera, se demuestra que

$$RSS_n = \sum_{j=k+1}^n w_j^2 \quad (2.26)$$

ó, en notación matricial,

$$\text{RSS}_n = w' w \quad (2.27)$$

y por lo tanto, igualando los segundos miembros de las expresiones (2.23) y (2.27) resulta que

$$e' e = w' w \quad (2.28)$$

III. Dócima de Chow a partir de los Residuos Recursivos

La dócima de Chow de cambio estructural se plantea como sigue: se dispone de n observaciones referidas a un vector de variables dependientes (Y) y de k variables explicativas (X). Y es por lo tanto un vector de $n \times 1$ mientras que X es una matriz de $n \times k$, tal como en el modelo presentado en (2.1).

Sin embargo, se quiere docimar la hipótesis de que las n_1 primeras observaciones de Y , que simbolizaremos por Y_1 , y las $n - n_1 = n_2$ restantes, que simbolizaremos por Y_2 están generadas por el mismo proceso generador de datos. Si simbolizamos por X_1 las n_1 primeras observaciones de X y por X_2 las $n - n_1 = n_2$ restantes, podemos expresar el modelo general de la siguiente forma:

$$Y_1 = X_1 \beta_1 + u_1 \quad (3.1)$$

$$Y_2 = X_2 \beta_2 + u_2$$

$$u_1 \sim N(0, \sigma_1^2 I_{n_1}) \quad u_2 \sim N(0, \sigma_2^2 I_{n_2})$$

en que β_1 y β_2 son vectores ($k \times 1$) y (u_1, u_2) son vectores de errores aleatorios de dimensiones ($n_1 \times 1$) y ($n_2 \times 1$) respectivamente.

Existen dos procedimientos diferentes para docimar la hipótesis nula de que $\beta_1 = \beta_2$. Ambos suponen que $\sigma_1^2 = \sigma_2^2$. El primero supone además que tanto n_1 como n_2 son mayores que k , vale decir, el número de observaciones pertenecientes a cada uno de los regímenes permite su estimación por MCO por separado.

El segundo procedimiento supone que $n_1 > k$ y $n_2 \leq k$. Johns-

ton (1984, pág. 220) lo trata de la siguiente manera, usando residuos obtenidos por MCO:

1) Hacer la regresión de Y sobre X usando las $n = n_1 + n_2$ observaciones y obtener la suma de los cuadrados de los residuos $e_*' e_*$.

2) Hacer la regresión de Y sobre X usando solamente las n_1 primeras observaciones y obtener la suma de los cuadrados de los residuos $e_1' e_1$.

3) Calcular el estadístico F tal como se indica a continuación:

$$F = \frac{(e_*' e_* - e_1' e_1)/n_2}{e_1' e_1 / (n_1 - k)} \quad (3.2)$$

Dicho estadístico F tiene n_2 y $(n_1 - k)$ grados de libertad en el numerador y denominador respectivamente. La regla de decisión consiste en rechazar $H_0: \beta_1 = \beta_2$ en caso de que el valor obtenido mediante el empleo de la fórmula (3.2) exceda el valor crítico de la distribución F para un nivel de la significación fijado de antemano.

Johnston (1984) plantea una demostración heurística de la fórmula (3.2) y refiere a Fisher (1970) para una demostración rigurosa.

Si en lugar de basar la dócima de Chow a partir de las estimaciones por MCO, se parte de los MCR, tal como lo propone Harvey (1976), la presentación rigurosa de este contraste de hipótesis se vuelve muy sencilla.

En efecto, usando los resultados sobre los residuos recursivos presentados en la segunda sección, se puede ver de inmediato que las sumas de cuadrados de residuos $e_*' e_*$ y $e_1' e_1$ necesarias para usar (3.2) se pueden calcular como sigue:

$$e_1' e_1 = \sum_{r=k+1}^{n_1} w_r^2 \quad (3.3)$$

$$e_*' e_* = \sum_{r=k+1}^{n_1+n_2} w_r^2 \quad (3.4)$$

en que w_r ha sido definido mediante la fórmula (2.13).

Utilizando (3.3) y (3.4) resulta inmediato comprobar que la fórmula (3.2) se puede expresar como

$$F = \frac{\sum_{r=n_1+1}^{n_1+n_2} w_r^2/n_2}{\sum_{r=k+1}^{n_1} w_r^2/(n_1-k)} \quad (3.5)$$

La ventaja que resulta del empleo de (3.5) en lugar de (3.2) radica simplemente en que se puede recurrir a teoremas de estadística elemental para afirmar que (3.5) tiene una distribución F con n_2 y $(n_1 - k)$ grados de libertad en el numerador y denominador respectivamente por la propiedad de los residuos recursivos establecida en la fórmula (2.15).

IV. Dócima para detectar autocorrelación de errores usando Residuos Recursivos.

La presentación de esta sección se basa en Judge y otros (1982, págs. 449-455) con relación a la dócima de Durbin y Watson. Esta dócima es la más popular de las dócimas para detectar la presencia de autocorrelación de los errores en el contexto del modelo lineal general con gresiones fijas en muestras repetidas, debido a que: a) es aplicable a muestras finitas, b) posee mayor potencia que otras dócimas alternativas, c) se basa en los residuos obtenidos por MCO y d) la mayoría de los programas de cómputos calculan el estadístico apropiado en forma rutinaria.

A partir del estimador de por MCO,

$$b_{MCO} = (X'X)^{-1} X'Y \quad (4.1)$$

se calculan los residuos e , dados por la siguiente fórmula

$$e = Y - X b_{MCO} \quad (4.2)$$

El estadístico d de Durbin y Watson se define como sigue:

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (4.3)$$

bajo el supuesto de que la muestra contiene n observaciones. La fórmula (4.3) puede reescribirse de la siguiente manera

$$d = \frac{e'Ac}{e'e} \quad (4.4)$$

en que la matriz A está dada por la fórmula (4.5)

$$A = \begin{vmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 1 \end{vmatrix} \quad (4.5)$$

Dado que e en la fórmula (4.4) se puede expresar de acuerdo a la siguiente fórmula

$$e = M u \quad (4.6)$$

en que u es el vector de errores aleatorios y M está dada por

$$M = I - X (X'X)^{-1} X' \quad (4.7)$$

se llega a la siguiente expresión para d :

$$d = \frac{u' M A M u}{u' M u} \quad (4.8)$$

puesto que la matriz M es idempotente y simétrica.

A partir de (4.8) se desprende que d es un cociente de dos formas cuadráticas del vector u , pero si el numerador tiene distribución chi-cuadrado ni se distribuye independientemente del denominador.

La distribución de d depende de los valores particulares que asuma la matriz M y por lo tanto de X y es por eso que Durbin y Watson no lograron tabularla para todas las situaciones posibles. Por el contrario, sus tablas contienen valores comúnmente simbolizados d_L

y d_U , siendo d_L la cota superior y d_L la cota inferior respectivamente. El valor crítico de d , que simbolizaremos por d^* para un nivel de significación determinado, está comprendido entre d_L y d_U , es decir,

$$d_L \leq d^* \leq d_U \quad (4.9)$$

El cálculo del valor d^* se puede hacer siguiendo los métodos de Imhof o Pan Jie-Jan (véanse los trabajos de Durbin y Watson y Koerts y Abrahamse citados por Harvey (1981a), (págs. 200-201). Sin embargo existe consenso en que dicho cálculo es laborioso e insume mucho tiempo aún para una computadora.

Los residuos recursivos proveen una alternativa de cálculo sencillo. En efecto, una vez obtenido los w_t se puede calcular el siguiente estadístico, conocido como razón de von Neumann modificada

$$\left(\frac{\delta^2}{s^2}\right)^* = \frac{\sum_{t=k+2}^n (w_t - w_{t-1})^2 / (n-k-1)}{\sum_{t=k+1}^n w_t^2 / (n-k)} \quad (4.10)$$

El resultado obtenido mediante la aplicación de (4.10) debe confrontarse con los valores críticos tabulados por Press y Brooks y que se encuentran reproducidos en los libros de Johnston (1984) y Harvey (1981 a).

Phillips y Harvey (1974) realizaron experimentos de Monte Carlo para evaluar las bondades relativas de este enfoque. Sus conclusiones apuntan a que el método exacto de Durbin y Watson, vale decir, el que utiliza el valor d^* es el que brinda mayor potencia. Sin embargo, la dódima basada en los residuos recursivos tiene mayor potencia que el uso del estadístico de Durbin y Watson cuando se usa como criterio para rechazar la hipótesis nula que el valor de d observado caiga por debajo de d_L .

Resumiendo, la ventaja que comporta el uso de los residuos recursivos consiste en la eliminación de la zona de incertidumbre comprendida entre d_L y d_U . Ello se consigue, sin embargo, a costa de una menor potencia. El mejor curso de acción es calcular el valor d^* por los métodos ya mencionados siempre que ello no sea prohibitivo.

V. MCR y el filtro de Kalman (FK)

Además de las d6cimas presentadas en las secciones tercera y cuarta de este trabajo, existen otros usos importantes de los residuos recursivos con fines de inferencia estadística. Dichos usos se refieren fundamentalmente a la detección de una mala especificación funcional, estudiada por Harvey y Collier (1977), para docimar la presencia de heteroscedasticidad en los residuos, tema que ha sido analizado por Harvey y Phillips (1974) y las d6cimas llamadas Cusum y Cusum de cuadrados propuestas por Brown, Durbin y Evans (1975) que tienen como objetivo determinar si ha habido cambios en los parámetros de un modelo lineal y/o en la varianza del término de error aleatorio. Tal como se anticipara en la introducción, no se hara una discusión detallada de estas d6cimas.

Varios comentaristas de Brown, Durbin y Evans (1975), fundamentalmente Priestley y Young, puntualizaron la íntima conexión existente entre la técnica de MCR y FK. Este tema fue retomado por Harvey (1981b) y origina la fórmula (2.17') de la segunda sección. Resulta muy sencillo obtener estimaciones idénticas a las brindadas por MCR a partir del FK y en lo que resta de esta sección trataré de ilustrar la conveniencia de adoptar este último enfoque porque resulta de aplicación a modelos más generales.

Considérese el siguiente modelo

$$y_t = x_t' \beta_t + \epsilon_t \quad t = 1, 2, \dots, n \quad (5.1)$$

$$\beta_t = M\beta_{t-1} + \eta_t \quad t = 1, 2, \dots, n \quad (5.2)$$

Considerado por Chow (1981). Los símbolos utilizados representan lo siguiente: y_t es la observación de la variable dependiente en el momento t , x_t es el vector de $(k \times 1)$ de valores de las variables explicativas del modelo (5.1), β_t es un vector de $(k \times 1)$ de parámetros y ϵ_t es el término de error, también referido al momento t .

La diferencia entre (5.1) y los modelos de regresión convencionales estriba en que el vector de parámetros no se supone aquí constante sino que evoluciona de acuerdo a la fórmula (5.2), que analizaremos seguidamente.

M es una matriz de $k \times k$ elementos constantes e invariantes res-

pecto del tiempo y η_t es un error aleatorio.

La descripción del sistema dado por las ecuaciones (5.1) y 5.2) no es completa sin especificar las propiedades de los términos de error ϵ_t y η_t . Al respecto es habitual postular las siguientes

$$\epsilon_t \sim N(0, \sigma^2 I) \quad (5.3)$$

$$\eta_t \sim N(0, \sigma^2 P) \quad (5.4)$$

$$\text{cov}(\epsilon_t, \eta_s) = 0 \quad \forall t, S \quad (5.5)$$

Resulta inmediato comprobar que el modelo de regresión convencional, es decir con β constante, se obtiene a partir del modelo dado por (5.1) y (5.5) bajo los siguientes supuestos:

$$M = I \quad (5.6)$$

$$P = 0 \quad (5.7)$$

(5.6) postula que M es igual a la matriz identidad y (5.7) que P es igual a la matriz nula. Como el FK es una técnica diseñada para estimar modelos del tipo del aquí considerado, sin imponer necesariamente (5.6) y (5.7), resulta una técnica de aplicación más general que MCR. Gran número de autores señalan la conveniencia del FK para la estimación econométrica, entre ellos Harvey (1981a, 1981b), Chow (1981), Judge y otros (1985), y Garbade (1977). Las ventajas que brinda este enfoque son dos:

Primero, permite evaluar las funciones de verosimilitud con mayor rapidez. Segundo, al no necesitar inversiones de matrices de $n \times n$ sino sólo de $k \times k$, siendo k normalmente mucho menor que n, posibilita la estimación por métodos maximo-verosímiles de modelos inabordable por los métodos convencionales.

Ilustraré estos aspectos a partir de varios experimentos realizados con un modelo del tipo de descrito por las ecuaciones (5.1) a (5.5), con M dada por (5.6) y P dada por la siguiente expresión

$$P = \begin{vmatrix} 0,01 & 0 \\ 0 & 0,01 \end{vmatrix} \quad (5.8)$$

β_0 , por su parte, es igual a

$$\beta_0 = \begin{bmatrix} 3,0 \\ 3,0 \end{bmatrix} \quad (5.9)$$

por lo que se puede advertir que $k=2$. Los modelos fueron generados por el programa GAUSS fijando un valor de sigma cuadrado igual a uno, y las variables x_t consisten en una ordenada al origen y una extracción de un generador aleatorio uniforme entre 2 y 3. Obviamente, los valores de ϵ_t y η_t fueron generados por generadores aleatorios normales con las matrices de covarianzas ya especificadas.

Chow (1981) demuestra cómo se puede estimar un modelo como el planteado por Mínimos Cuadrados Generalizados (MCG) y mediante la aplicación del FK. Aplicando su método se obtuvieron los siguientes resultados:

Primero, para un tamaño muestral de 30, $n=30$, el tiempo de cálculo insumido por MCG fue de 36,09 segundos contra 8,74 segundos del FK.

Segundo, para $n=60$, los tiempos respectivos fueron de 152,47 y 16,64 segundos.

Tercero, el método de MCG no puede implementarse en la versión de GAUSS disponible, aunque sí en las más nuevas, para $n > 90$. Pero aún las nuevas versiones enfrentan seguramente un límite.

Cuarto, hasta ahora se ha supuesto que los elementos de P son conocidos. Cuando esto no sucede, que es la situación más realista de la práctica econométrica, deben estimarse. Los cálculos preliminares realizados revelan que cada iteración del método máximo-verosímil requiere 8 minutos aproximadamente para $n=30$ cuando se evalúa la función de verosimilitud por los métodos convencionales contra aproximadamente 50 segundos si se usa el enfoque del FK.

VI. Conclusiones

Los MCR representan una técnica valiosa para docimar hipótesis en el contexto del modelo lineal general con regresores fijos en muestras repetidas. En la segunda sección se describió la técnica básica, así como las propiedades de los residuos recursivos y sus relaciones con los resi-

duos habitualmente calculados por MCO.

En la tercera sección se ilustró el uso de los residuos recursivos para aplicar la d6cimo de Chow. Debe destacarse la gran simplificación conceptual que es posible obtener siguiendo este enfoque.

Se describi6 el uso de los residuos recursivos para detectar la presencia de autocorrelaci6n de primer orden en la cuarta secci6n. Nuevamente, esta t6cnica permite simplificaciones importantes que en este caso evitan la zona de incertidumbre del estadístico de Durbin y Watson.

Por 6ltimo en la quinta secci6n se mencionaron brevemente otros usos de los residuos recursivos para hacer inferencias. Se enfatiz6 que para fines de c6culo el filtro de Kalman ofrece ventajas significativas dado que es una t6cnica de uso m6s general aplicable, por ejemplo, a modelos lineales con coeficientes aleatorios.

Estos 6ltimos modelos han adquirido popularidad en la literatura econom6trica a partir de 1973 y deberían, en mi opini6n, incorporarse a los cursos de Econometría para lo cual los m6todos aqu6 descritos deben jugar un papel importante. Asimismo se debe refinar el marco conceptual de dichos modelos. Para este 6ltimo objetivo son muy valiosos los art6culos de Cooley y Prescott (1976) y Garbade (1977).

REFERENCIAS

- BROWN, R.L., DURBIN, J. and EVANS, J.M. (1975), *Techniques for Testing the Constancy of Regression Relationships over Time (with Discussion)*, Journal of the Royal Statistical Society, Series B (Methodological), Vol. 37, N6 2, p6gs. 149-192.
- COOLEY, T.F. and PRESCOTT, E.C. (1976), *Estimation in the presence of stochastic parameter variation*, Econometrica, Vol. 44, N6 1, January, p6gs. 167-184.
- CHOW, G.C. (1960), *A test of equality between sets of observations in two linear regressions*, Econometrica, Vol. 28, N6 3, July, p6gs. 591-605.
- CHOW, G.C. (1981), *Econometric analysis by control methods*, New York, Wiley, Capítulo 6: Applications of the Kalman filter to the estimation of econometric models, p6gs. 85-107.
- DUNCAN, D.B. and HORN, S.D. (1972), *Linear dynamic recursive estimation from the viewpoint of regression analysis*, Journal of the American Statistical Association, Vol. 67, N6 340, December, p6gs. 815-621.
- DURBIN, J. (1969), *Tests for serial correlation in regression analysis based on the periodogram of least-squares residuals*, Biometrika, Vol. 56, N6 1, March, p6gs. 1-15.

- FISHER, F.M. (1970), *Tests of equality between sets of coefficients in two linear regressions. an expository note*, *Econometrica*, Vol. 38, N^o 2, March, págs. 361-366.
- GARBADE, K. (1977), *Two methods for examining the stability of regression coefficients*, *Journal of the American Statistical Association*, Vol. 72, N^o 357, March, págs. 54-63.
- HARVEY, A.C. (1976), *An alternative proof and a generalization of a test for structural change*, *The American Statistician*, Vol. 30, N^o 3, August, págs. 122-123.
- HARVEY, A.C. (1981a), *The Econometric Analysis of Time Series*, Philip Allan, Oxford.
- HARVEY, A.C. (1981b), *Time Series Models*, Philip Allan, Oxford.
- HARVEY, A.C. and COLLIER, P. (1977), *Testing for functional misspecification in regression analysis*, *Journal of Econometrics*, Vol. 6, N^o 1, July, págs. 103-119.
- HARVEY, A.C. and PHILLIPS, G.D.A. (1974), *A comparison of the power of some tests for heteroscedasticity in the general linear model*, *Journal of Econometrics*, Vol. 2, N^o 4, December, págs. 307-316.
- HARVEY, A.C. and PHILLIPS, G.D.A. (1979), *Maximum likelihood estimation of regression models with autoregressive - moving average disturbances*, *Biometrika*, Vol. 68, N^o 1, April, págs. 49-58.
- JOHNSTON, J. (1984), *Econometric Methods*, Third edition, McGraw Hill, New York.
- JUDGE, G.G., CARTER HILL, R., GRIFFITHS, W.E., LUETKEPOHL, H. and LEE, T.C. (1982), *Introduction to the theory and practice of econometrics*, Wiley, New York.
- JUDGE, G.G., CARTER HILL, R., GRIFFITHS, W.E., LUETKEPOHL, H. and LEE, T.C. (1985), *The theory and practice of econometrics*, Wiley, New York, Appendix C. the Kalman filter, págs. 980-984.
- MONTERO, H.E. (1982), *Modelización de series temporales no estacionarias utilizando filtro de Kalman*. Tesis doctoral no publicada, Facultad de Ciencias Económicas, Universidad Nacional de Córdoba, marzo, mimeo.
- PHILLIPS, G.D.A. and HARVEY, A.C. (1974), *A simple test for serial correlation in regression analysis*, *Journal of the American Statistical Association*, Vol. 69, N^o 348, December, págs. 935-939.

MINIMOS CUADRADOS RECURSIVOS

RESUMEN

El trabajo presenta los elementos fundamentales del método de mínimos cuadrados recursivos enfatizando su utilidad para las d6cimas de cambio estructural y autocorrelaci6n.

Con relaci6n al cambio estructural, el m6todo permite una presentaci6n m6s did6ctica que el enfoque habitual asociado a la F. de Chow. Por lo que respecta a la autocorrelaci6n, esta t6cnica tiene ventajas y desventajas relativas al uso de los procedimientos de Durbin y Watson que se discuten en el trabajo.

Por 6ltimo se relacionan los m6nimos cuadrados recursivos con el filtro de Kalman y se ilustran las ventajas de 6ste 6ltimo para la estimaci6n por m6nimos cuadrados generalizados.

RECURSIVE LEAST SQUARES

SUMMARY

The paper presents the key features of the recursive least squares method emphasizing its usefulness for tests of structural break and autocorrelation.

With regard to structural break, this method lends itself to a simpler classroom presentation than the more traditional Chow's F test. In connection with autocorrelation, the method's advantages and disadvantages vis-a-vis the Durbin Watson procedures are discussed.

The paper also relates recursive least squares to the Kalman filter and provides some evidence to establish the superiority of this latter approach for generalized least squares estimation.