

# MODELOS CON PARÁMETROS VARIABLES EN EL ANÁLISIS DE SERIES TEMPORALES

D. PEÑA SÁNCHEZ DE RIVERA

*Este trabajo compara la representación de modelos de previsión univariante en el espacio de los estados con la formulación ARIMA. Se expone como ambos enfoques conducen a expresiones de previsión similares y se discuten las ventajas alternativas de ambas metodologías. La conclusión principal de este trabajo es que la formulación ARIMA es más general y permite diferenciar claramente los "superparámetros", que son fijos, de los parámetros de la ecuación de previsión, que son adaptativos y por lo tanto variables en el tiempo. Los "superparámetros" determinan la evolución y estructura de los parámetros adaptativos. Este aspecto ofrece importantes ventajas teóricas y prácticas que se ilustran mediante un ejemplo concreto de previsión univariante en que se comparan ambos métodos.*

## 1. INTRODUCCION

Desde mediados de los años 60, el filtro de Kalman /13/ y la formulación de modelos económicos en términos de la clásica representación utilizada en ingeniería, en el espacio de los estados, ha tenido una popularidad creciente. Una ventaja clara de esta representación es la existencia de un procedimiento recursivo eficiente de estimación: El filtro de Kalman.

En Economía, el filtro de Kalman se ha utilizado, entre otras áreas, en la estimación de modelos econométricos con parámetros que se suponen variables en el tiempo (véase la referencia /3/ o la /22) en la estimación de modelos input-output estocásticos /23/ y /22/ y en la previsión con modelos econométricos /14/. Una breve discusión de estas aplicaciones se encuentra en /7/.

Un área reciente de aplicación de este enfoque es para la previsión univariante de series temporales con parámetros que varían con el tiempo. En esta línea se encuentra el trabajo de Harrinson y Stevens /11/ y Morrinson y Pike /16/.

Este trabajo está dirigido a estudiar la aplicación de esta metodología en el contexto de la previsión de series temporales, y está

estructurado como sigue: En la sección 2 presentamos la formulación del filtro de Kalman y en la sección 3 comentamos brevemente su utilización en Econometría. En la sección 4, estudiamos la formulación de modelos de previsión univariante con este enfoque, comparándola con la representación ARIMA. La sección 5, presenta una comparación empírica de ambos métodos utilizando una serie publicada en la literatura. Las conclusiones del trabajo se resumen en la sección 6.

## 2. EL FILTRO DE KALMAN

Presentaremos en primer lugar la formulación discreta del problema de estimación, que puede resolverse mediante el filtro de Kalman, en el contexto de un problema de ingeniería. A continuación, expondremos su relación con un procedimiento iterativo de mínimos cuadrados, en el marco de la estimación de un modelo de regresión. Finalmente, comentaremos su aplicación para construir un modelo econométrico con parámetros que, se supone, varían con el tiempo.

En Ingeniería, suponemos que el estado de un sistema físico dinámico, que observamos en instantes discretos, puede caracterizarse por un vector de estado  $X_t$  que describe completamente la situación del sistema en el momento  $t$ .

- D. Peña. Escuela de Ingenieros Industriales. Depto. de Estadística. José Gutiérrez Abascal, 2. Madrid - 6.  
- Article rebut el Maig del 1.980.

Por ejemplo,  $\underline{X}_t$  puede representar la posición de un móvil que evoluciona en el tiempo con determinadas leyes físicas.

El vector de estado es un vector de variables que, conjuntamente con el resto de los inputs al sistema, determina su evolución posterior. Es decir, si conocemos el vector de estado  $\underline{X}_{t-1}$  y el efecto de la perturbación aleatoria que afecta a la evolución del sistema en ese instante,  $\underline{\omega}_{t-1}$ , el estado en  $t$  queda determinado por la "ecuación de estado":

$$\underline{X}_t = \underline{\Phi}_t \underline{X}_{t-1} + \underline{\Gamma}_t \underline{\omega}_{t-1} \quad (1)$$

Donde  $\underline{\Phi}_t$  es una matriz conocida de coeficientes variables en el tiempo que gobierna la transición sistemática del estado  $t-1$  al estado  $t$ ,  $\underline{\omega}_t$  es un vector aleatorio de ruido blanco<sup>1</sup> y  $\underline{\Gamma}_t$  es una matriz de coeficientes variables que determinan el efecto de las perturbaciones exteriores al sistema,  $\underline{\omega}_{t-1}$ , sobre la evolución de cada variable de estado. Las matrices  $\underline{\Phi}_t$  y  $\underline{\Gamma}_t$  se suponen conocidas para todo  $t$ .

Se supone también un sistema de mediación de un vector de observaciones empíricas  $\underline{Z}_t$  que informa acerca del sistema en el estado  $t$ , a través de la relación:

$$\underline{Z}_t = \underline{H}_t \underline{X}_t + \underline{V}_t \quad (2)$$

Donde  $\underline{H}_t$  es una matriz conocida que explicita la relación sistemática en cada instante entre las observaciones  $\underline{Z}_t$  y el estado del sistema  $\underline{X}_t$ , que no es observable directamente. El error de medida  $\underline{V}_t$ , suponemos es ruido blanco<sup>1</sup>.

El problema planteado es: dada una secuencia de vectores de medición ( $\underline{Z}_1 \dots \underline{Z}_t$ ), se trata de estimar el estado del sistema, así como de prever su evolución en los instantes ( $t+1 \dots t+k$ ). En la hipótesis de que el criterio de estimación es mínimos cuadrados, este problema fue resuelto originalmente por Kalman (1960), utilizando las propiedades de las proyecciones ortogonales sobre un espacio de Hilbert. Posteriormente el filtro ha sido justificado por otros métodos. Destacaremos la deducción del filtro como un algoritmo recursivo mínimocuadrático /24/ y el enfoque Bayesiano de Ho y Lee o Aoku /2/.

Una buena exposición del algoritmo en un contexto de ingeniería se encuentra en Medith

/15/.

### 3. APLICACION A MODELOS ECONOMETRICOS

La conversión del problema descrito por (1) y (2) al campo de los modelos económicos se ha abordado tradicionalmente suponiendo, que el vector de estado  $\underline{X}_t$ , es el vector de parámetros del modelo. Entonces, la ecuación (2) representa la forma estructural del modelo y la (1) la ecuación que describe la evolución de los parámetros -que dejan de ser fijos- a lo largo del tiempo.

Para concretar, supongamos el modelo de regresión uniecuacional:

$$y_t = \underline{x}'_t \underline{\beta}_t + \varepsilon_t \quad (3)$$

$$\underline{\beta}_t = \underline{\Phi}_t \underline{\beta}_{t-1} + \underline{\Gamma}_t \underline{\omega}_{t-1} \quad (4)$$

Donde, en la ecuación (3)  $y_t$  es un escalar, igual al valor observado de la variable dependiente en el momento  $t$ ,  $\underline{x}_t$  es el vector<sup>2</sup> de observaciones de las variables explicativas, en el momento  $t$ ,  $\underline{\beta}_t$  es el vector de parámetros y  $\varepsilon_t$  la perturbación. Esta ecuación es similar a la (2), pero ahora el vector de medición  $\underline{Z}_t$  es el escalar  $y_t$ , la matriz  $\underline{H}_t$  que relaciona los valores del vector de estado con el vector de observaciones es ahora simplemente el vector  $\underline{x}'_t$  y el vector de estado del sistema es el vector de parámetros  $\underline{\beta}_t$ . La ecuación (4) describe la evolución temporal del vector de estado  $\underline{\beta}_t$  y es análoga a (1). El modelo de regresión clásico supone que la matriz  $\underline{\Phi}_t$  es unitaria, para todo  $t$ , y que la matriz  $\underline{\Gamma}_t$  es idénticamente nula (esto es equivalente a suponer que los parámetros son cantidades fijas que no varían con el tiempo). La formulación de (3) y (4) elimina esa restricción permitiendo que los parámetros del modelo sigan un proceso estocástico especificado por (4).

La aplicación del filtro de Kalman a la formulación (3) y (4) procede en dos etapas. En la primera, se utiliza (4) para hacer una previsión de  $\hat{\underline{\beta}}_{t/t-1}$ , es decir del valor del vector de parámetros dada la información hasta el instante  $t-1$ , que en la segunda etapa es revisado al recibir la información del instante  $t$  ( $\underline{x}'_t, y_t$ ) para obtener la estimación óptima  $\hat{\underline{\beta}}_t$  en ese instante. En el caso en que los parámetros  $\underline{\beta}_t$  se supongan fijos, no se a

plica la primera etapa de "previsión" y la aplicación del filtro de Kalman equivale a los mínimos cuadrados recursivos de Plackett --- /18/ y conduce a un algoritmo que permite obtener iterativamente los estimadores mínimo-cuadráticos dada la información hasta el instante  $t$ , sin necesidad de recalcularlos desde el principio cuando recibimos nuevas observaciones. El algoritmo es:

$$\hat{\beta}_t = \hat{\beta}_{t-1} + P_t X_t (y_t - X_t' \hat{\beta}_{t-1}) \quad (5)$$

$$P_t = P_{t-1} - P_{t-1} X_t' a^{-1} X_t P_{t-1} \quad (6)$$

$$a = 1 + X_t' P_{t-1} X_t \quad (7)$$

$$\hat{\sigma}_t^2 = \frac{1}{t} \sum_{i=1}^t (y_i - X_i' \hat{\beta}_i)^2 \quad (8)$$

Donde  $\hat{\beta}_t$  es el vector de parámetros utilizando las observaciones hasta el instante  $t$  inclusive,  $P_t$  es, salvo por una constante, una estimación de la matriz de varianzas y covarianzas de las variables  $x$  en el instante  $t$  y  $a$  es un escalar. Prescindiendo del efecto debido a los valores iniciales,  $P_t$  es la matriz  $(X'X)_t^{-1}$ , siendo  $X_t$  la matriz de observaciones de las variables explicativas hasta el instante  $t$ .

La ecuación (5) admite una interpretación intuitiva simple. Si tenemos en cuenta que la matriz estimada de varianzas y covarianzas de las  $t$  observaciones disponibles es:

$$\hat{\Sigma}_t = \hat{\sigma}_t^2 (X'X)_t^{-1} = \hat{\sigma}_{t=t}^2 P_t = \{s_{ij,t}^2\}$$

donde  $s_{ij,t}^2$  son las varianzas y covarianzas estimadas hasta el momento  $t$ . Si llamamos  $e_t^*$  al error que cometeríamos al predecir el valor obtenido  $y_t$  con los parámetros estimados en  $t-1$ :

$$e_t^* = y_t - X_t' \hat{\beta}_{t-1}$$

La modificación de un parámetro cualquiera  $\beta_t^j$  en el instante  $t$ , al observar los valores de las  $k$  variables explicativas es, según (5):

$$\hat{\beta}_t^j = \hat{\beta}_{t-1}^j + \frac{e_t^*}{\sigma} \left[ \sum_{i=1}^k X_{it} s_{ij,t-1}^2 \right] \quad (9)$$

que indica como la modificación del coeficiente  $\beta_t^j$  depende: a) del error relativo  $e_t^*$ ; b) de los valores de las  $X$  observadas que tengan covariación no nula con la variable  $X_j$ .

En este contexto, este algoritmo puede demos-

trarse<sup>5</sup> que es equivalente a la utilización de un algoritmo de Newton no lineal para obtener el valor óptimo de  $\hat{\beta}$  que minimiza la suma de los cuadrados de los residuos de las  $t$  observaciones con valor inicial  $\hat{\beta}_{t-1}$ . El óptimo se alcanza en un solo paso debido a que la función es cuadrática.

Este análisis presupone que los parámetros son constantes en el tiempo. Rubin /2/, sugirió por primera vez la hipótesis de que los parámetros de un modelo de regresión pueden ser variables que varían temporalmente. Hildreth y Houck /12/, Cooley y Prescott /5/, Rosenberg /20/ y Cooper /6/, entre otros, han presentado modelos econométricos en que se especifica una estructura de variación de parámetros<sup>6</sup> del tipo reflejado por (4). La obtención de los parámetros en cada instante puede estimarse eficientemente mediante el filtro de Kalman<sup>7</sup>.

La parametrización de modelos en esta línea presenta problemas importantes. En primer lugar, puede ocultar una mala formulación de una relación simple entre las variables o una errónea formulación de la perturbación aleatoria. En segundo lugar, contrastar rigurosamente esta hipótesis no es fácil con los tamaños muestrales habituales en Econometría.

Respecto al primer punto, consideremos que, en realidad, la relación entre  $y_t$  y  $x_t$  viene dada por la función de transparencia<sup>8</sup>:

$$y_t = v(B) x_t + \rho(B) a_t \quad (10)$$

Donde  $\rho(B)$  es un filtro no necesariamente estacionario y  $a_t$  un proceso de ruido blanco. Supongamos que, en su lugar, especificamos el modelo de regresión con parámetros variables:

$$y_t = x_t \beta_t + \alpha_t \quad (11)$$

Entonces, despejando  $\beta_t$  en (11) y utilizando (10):

$$\beta_t = \frac{y_t}{x_t} + \frac{\alpha_t}{x_t} = \frac{1}{x_t} v(B) x_t + \frac{1}{x_t} \rho(B) a_t + \frac{1}{x_t} \alpha_t$$

y aparentemente el parámetro  $\beta_t$  tiene una estructura estocástica que depende de los filtros  $v(B)$ ,  $\rho(B)$  y del modelo univariante de  $x_t$ . Si este modelo es

$$x_t = \psi_x(B) u_t \quad (12)$$

tendremos que:

$$\beta_t = \frac{1}{x_t} \{ \psi(B) \psi_x(B) v_t + \rho(B) a_t + \alpha_t \} = \frac{\psi^*(B) \epsilon_t}{\psi_x(B) v_t}$$

y será el cociente de dos procesos ARIMA generales. Es interesante señalar que los modelos propuestos en la literatura conducen a casos especiales de la formulación (10). Por ejemplo, Hildreth y Houck /12/ proponen el modelo:

$$y_t = x_t \beta_t + \alpha_t$$

$$\beta_t = \omega_0 + v_t$$

Donde  $v_t$  y  $\alpha_t$  son ruidos blancos incorrelados. Entonces:

$$y_t = x_t \omega_0 + x_t v_t + \alpha_t$$

y si  $x_t$  sigue el proceso (12):

$$y_t = \omega_0 x_t + \rho(B) a_t$$

donde:

$$\rho(B) a_t = v_t \psi_x(B) v_t + \alpha_t$$

El mismo razonamiento simple sirve para mostrar como los modelos anteriormente citados con parámetros variables, son casos particulares de (10), que pueden surgir como consecuencia de una mala especificación de la función de transferencia y/o el modelo del ruido.

El segundo aspecto que queremos comentar es el problema de contrastar, la hipótesis de que los parámetros varían con el tiempo. Si tenemos una muestra grande, el procedimiento obvio es dividirla en tramos y estimar el modelo en cada uno de los tramos. Si hay evidencia estadística de que los parámetros estimados son distintos, parece razonable preguntarse: a) Si han existido intervenciones o factores puntuales anómalos en los distintos períodos que puedan explicar los cambios observados; b) Si posibles variables causales no incluidas en el modelo pueden ser responsables de los cambios. El primer problema puede tratarse con el Análisis de Intervención de Box y Tiao /25/ y el segundo introduciendo explícitamente las variables apropiadas en el modelo. Pensamos que el acudir a un modelo con parámetros que varían en el tiempo sólo debe hacerse cuando exista una teoría con-

sistente que explique por qué y cómo varían estos parámetros.

## 4. APLICACION A MODELOS UNIVARIANTES DE SERIES TEMPORALES

### 4.1. Introducción

El enfoque metodológico anterior ha sido recientemente extendido al análisis univariante de series temporales. La formulación del modelo en este caso es:

$$y_t = f_t' \beta_t + \epsilon_t \quad (13)$$

$$\beta_t = \phi_t \beta_{t-1} + \Gamma_t \omega_{t-1} \quad (14)$$

La especificación de los vectores  $f_t$  y  $\beta_t$  conduce a distintas clases de modelos. Pagan (1975) ha mostrado cómo utilizar esta representación para formular un modelo de descomposición clásico en tendencia, estacionalidad y parte irregular. Restringiéndonos a la formulación de Harrinson y Stevens (1976) y Morrison y Pike (1977), escribiremos:

$$f_t' = (1, 0, \dots, 0)$$

$$\beta_t' = (\mu_t, a_{1t}, \dots, a_{kt})$$

$$E[\epsilon_t] = 0$$

$$E[\epsilon_t \epsilon_j] = R(t) \delta_{tj}$$

$$E[\omega_t] = 0$$

$$E[\omega_t \omega_j] = Q(t) \delta_{tj}$$

siendo  $\delta_{ij}$  la delta de Kronecker y  $\phi_t$  y  $\Gamma_t$  matrices que varían con el tiempo, que se suponen conocidas.

La formulación del modelo univariante con esta estructura permite la estimación recursiva de los parámetros  $\beta_t$ , el "estado" del sistema, y la predicción de  $y_t$  utilizando el filtro de Kalman.

Este enfoque se denomina "previsión bayesiana" por Harrison y Stevens /11/, que utilizan la interpretación bayesiana del filtro de Kalman para justificar la parametrización -- (13) y (14). Según estos autores una ventaja importante de este procedimiento es que

puede aplicarse cuando no disponemos de datos históricos que pueden ser sustituidos por información a priori para comenzar el algoritmo recursivo. La crítica básica a este enfoque es que supone conocidas las matrices  $\phi_t$  y  $\Gamma_t$  y dado que esta metodología no incluye procedimientos de identificación, esto conduce a que las matrices se especifiquen arbitrariamente. El problema es especialmente grave, porque al no disponer de procedimientos diagnósticos o de crítica del modelo, una especificación provisional a priori, que puede ser quizás inevitable al disponer de poca información muestral, se convierte en permanente por la imposibilidad de que los datos muestrales nos conduzcan a rechazar el modelo.

#### 4.2. Modelos ARIMA

Un enfoque alternativo, propugnado por Box y Jenkins /4/ es tomar como formulación general para la representación de una serie la clase de modelos ARIMA dada por:

$$\phi_p(B) \nabla^d z_t = \theta_q(B) a_t$$

Donde  $\phi_p(B)$  y  $\theta_q(B)$  son operadores polinómicos de órdenes  $p$  y  $q$  respectivamente en el operador de retardo  $B$ , con ceros fuera del círculo unidad y  $\nabla$  es el operador diferencia definido por  $\nabla = 1 - B$ . Un resumen de los procedimientos operativos para construir un modelo ARIMA para una serie dada puede verse en /19/.

La determinación de la ecuación de previsión  $\hat{z}_t(1)$  para obtener una estimación del valor de la serie  $z_t$  en el instante  $t+1$  dada la información disponible hasta el instante  $t$ , requiere: a) Establecer la ecuación de previsión (recta, parábola, senoide, etc...); b) Establecer cómo vamos a modificar los parámetros de la ecuación de previsión, es decir cómo obtener  $\hat{z}_{t+1}(1)$  cuando observamos el nuevo valor  $z_{t+1}$  de la serie; c) Determinar la distribución de probabilidad de la previsión para obtener intervalos de confianza de las previsiones puntuales, para distintos instantes de tiempo.

Al construir un modelo ARIMA para una serie, estos tres factores quedan automáticamente definidos. En términos muy generales podemos decir que el operador general autorregresivo

$\rho(B) = \phi(B) \nabla^d$  determina la ecuación de previsión, el operador media móvil va a fijar la adaptación de los parámetros, mientras que el operador,  $\rho(B)^{-1} \theta(B)$ , va a determinar los errores de previsión para los distintos retardos.

Si llamamos  $\hat{z}_t(1)$  a la función de previsión calculada en el instante  $t$  para la previsión de horizonte 1, que será por lo tanto la previsión de  $z_{t+1}$  conocidos  $z_t$  y  $z_{t-1}, \dots$ ; esta función viene determinada por la solución de la ecuación homogénea en diferencias: (en la hipótesis  $p+d > q$ )

$$\rho(B) \hat{z}_t(1) = 0 \quad (15)$$

donde el operador  $B$  actúa sobre 1, siendo ahora  $t$  fijo. La solución general de esta ecuación es:

$$\hat{z}_t(1) = \underline{f}'_1 \underline{\beta}_t \quad (16)$$

Donde  $\underline{\beta}'_t = (\beta_1(t), \dots, \beta_{p+d}(t))$  es un vector de coeficientes variables que van a adaptarse al cambiar el origen y  $\underline{f}'_1$  es un vector:

$$\underline{f}'_1 = (f_1(1), \dots, f_{p+d}(1))$$

con componentes  $f_i(1)$  que son funciones exponenciales, polinómicas o sinusoidales. La ecuación de adaptación de los parámetros  $\underline{\beta}_t$  cuando recibimos la observación  $z_{t+1}$  es:

$$\underline{\beta}_{t+1} = \underline{L} \underline{\beta}_t + \underline{h} e_t(1) \quad (17)$$

donde  $\underline{L}$  es una matriz de constantes fijas que acopla el vector  $\underline{\beta}_{t+1}$  al nuevo origen  $z_{t+1}$ ,  $e_t(1)$  es el error de previsión definido por:

$$e_t(1) = z_{t+1} - \hat{z}_t(1)$$

y  $\underline{h}$  es un vector de constantes que dependen básicamente de la parte media móvil del proceso ARIMA.

La estructura del modelo determina también únicamente las varianzas de los errores de previsión para cualquier número de períodos en adelante.

Por lo tanto, la ecuación del modelo explicita sin ambigüedad los tres componentes necesarios para la ecuación de previsión.

Observemos que: a) La estructura autorregresiva del modelo  $\rho(B)$  va a determinar la forma de las funciones  $f_i(1)$  y los parámetros  $\beta_t$  de la ecuación de previsión. b) la adaptación de estos parámetros va a ser determinada por la parte media móvil.

Por ejemplo, para el modelo:

$$\nabla^2 Z_t = (1 - \theta_1 B - \theta_2 B^2) a_t$$

La previsión  $Z_t(\ell)$  realizada en  $t$  con horizonte  $\ell$  es, (véase Box y Jenkins /4/ pp. 146-147), una línea recta con parámetros  $\beta_1(t)$  y  $\beta_2(t)$ :

$$\hat{Z}_t(\ell) = \beta_1(t) + \beta_2(t) \ell$$

que aunque son fijos en  $t$ , van adaptándose al recibir nueva información de acuerdo con:

$$\beta_1(t+1) = \beta_1(t) + \lambda_0 e_t(1)$$

$$\beta_2(t+1) = \beta_2(t) + \lambda_1 e_t(1)$$

siendo  $\lambda_0$  y  $\lambda_1$  constante determinadas por:

$$\lambda_0 = 1 + \theta_2 \quad \lambda_1 = 1 - \theta_1 - \theta_2$$

y siendo  $e_t(1)$  la diferencia entre el valor previsto en  $t$  para  $t+1$ ,  $\hat{Z}_t(1)$  y el valor observado  $Z_{t+1}$ .

Es importante señalar que este enfoque recoge como caso particular los modelos deterministas utilizados a veces en previsión. En el ejemplo anterior si  $\theta_2 = -1$  y  $\theta_1 = 2$ , los errores  $e_t(1)$  no son tenidos en cuenta en la actualización de los parámetros del modelo -- ( $\lambda_0 = \lambda_1 = 0$ ) y la función de previsión se convierte en una recta determinista. En general, la casi cancelación de operadores entre la parte media móvil y autorregresiva indicará la necesidad de un término determinista. El lector interesado en este punto puede acudir a /1/.

#### 4.3. Comparación entre ambas formulaciones

Si utilizamos (16) y (17), un modelo ARIMA puede escribirse:

$$Z_{t+1} = \underline{f}_1' \underline{\beta}_t + e_t(1) \quad (18)$$

$$\underline{\beta}_{t+1} = \underline{I} \underline{\beta}_t + \underline{h} e_t(1) \quad (19)$$

que es la formulación "de estado" análoga a (13) y (14).

Las diferencias principales entre ambas formulaciones es que en el enfoque de Box y Jenkins los vectores  $\underline{f}_1$  y  $\underline{h}$  y la matriz  $\underline{I}$  se determinan empíricamente a partir de los datos y no se especifican a priori arbitrariamente. Un aspecto fundamental de esta metodología es la distinción entre lo que podríamos llamar "hiperparámetros" del proceso que determinan la estructura de (18) y (19) y la memoria del proceso y que son fijos y los parámetros  $\underline{\beta}_t$  de la ecuación de previsión que son adaptativos y variables a lo largo del proceso. La expresión ARIMA en forma de ecuación en diferencias estocásticas permite la estimación eficiente de estos "hiperparámetros" (los coeficientes autorregresivos y media móvil) que van a determinar inmediatamente la estructura de la representación (18) y (19).

La expresión del modelo ARIMA como (18) y (19) conduce a la sospecha de que los modelos propuestos en la literatura mediante (13) y (14) deben de ser casos particulares de la clase de modelos ARIMA. Por ejemplo, Harrison y Stevens /11/ proponen los dos modelos univariantes siguientes. El "steady" model:

$$Y_t = \mu_t + \epsilon_t$$

$$\mu_t = \mu_{t-1} + v_t$$

Donde  $\epsilon_t$  y  $v_t$  son procesos de ruido blanco - incorrelados. Este modelo equivale a:

$$Y_t = \mu_{t-1} + v_t + \epsilon_t = Y_{t-1} + \epsilon_{t-1} + \epsilon_t + v_t$$

$$\nabla Y_t = (1 - \theta B) a_t$$

Donde  $a_t$  es ruido blanco,  $B$  es el operador de retardo  $BY_t = Y_{t-1}$ ;  $\nabla = 1 - B$ . Hemos aplicado -- que la suma de dos procesos de media móvil -- independientes es igual a un nuevo proceso de media móvil de orden igual al orden máximo -- de los procesos que sumamos /8/.

Un análisis equivalente conduce a demostrar que el "Linear Growth model" de estos autores es simplemente el ARIMA (0,2,2).

Morrison y Pike /16/ proponen dos modelos

adaptativos que, en su opinión: "Aparecen frecuentemente en la práctica e ilustran la capacidad de modelado" de la formulación (13) y (14). El primer modelo es:

$$Y_t = \mu_t + \epsilon_t \quad (20)$$

$$\mu_t = \mu_{t-1} + \alpha_{t-1} \quad (21)$$

$$\alpha_{t-1} = \rho \alpha_{t-2} + \omega_{t-2} \quad (22)$$

Donde  $\epsilon_t$  y  $\omega_t$  son procesos independientes de ruido blanco. Este modelo puede expresarse:

$$Y_t = \mu_{t-1} + \rho \alpha_{t-2} + \omega_{t-2} + \epsilon_t$$

$$Y_t = Y_{t-1} - \epsilon_{t-1} + \rho \alpha_{t-2} + \omega_{t-2} + \epsilon_t$$

y como, según (22):

$$(1 - \rho B) \alpha_{t-1} = \omega_{t-2} + \alpha_{t-1} = (1 - \rho B)^{-1} \omega_{t-2}$$

$$\nabla Y_t = \nabla \epsilon_t + (1 - \rho B) \rho^{-1} \omega_{t-3} + \omega_{t-2}$$

$$(1 - \rho B) \nabla Y_t = (1 - \rho B) \nabla \epsilon_t + \omega_{t-2}$$

$$(1 - \rho B) \nabla Y_t = (1 - \theta_1 B - \theta_2 B^2) a_t$$

Es decir, equivalente al ARIMA (1,1,2), aplicando de nuevo la propiedad de que la suma de procesos media móvil es también media móvil.

El segundo modelo es:

$$Y_t = \mu_t + \epsilon_t \quad (23)$$

$$\mu_t = \mu_{t-1} + \alpha_{t-1} \quad (24)$$

$$\alpha_{t-1} = \rho \mu_{t-2} \quad (25)$$

que conduce, por operaciones análogas, al modelo:

$$(1 - B - \rho B^2) Y_t = (1 - B - \rho B^2) \epsilon_t$$

Dado que una de las raíces de la ecuación característica del modelo es estrictamente mayor que uno, este modelo no será estacionario y tampoco invertible<sup>9</sup>. La media del proceso es la solución de la ecuación en diferencias:

$$\mu_t = \mu_{t-1} + \rho \mu_{t-2}$$

que es  $\mu_t = A_1 G_1^t + A_2 G_2^t$ , siendo  $A_1$  y  $A_2$  constantes que dependen del origen de tiempos.

Como una de las dos raíces es mayor que la unidad, la media seguirá un crecimiento exponencial. Un inconveniente claro de este modelo es que incorpora una tendencia determinista fijada por  $\rho$  y por lo tanto será poco flexible para seguir la evolución de la serie si ésta presenta cambios en su estructura de evolución, como es esperable en el futuro.

En resumen, no parecen existir razones teóricas de peso que indiquen las ventajas de la formulación (13) y (14), ya que puede considerarse equivalente a la representación (18) y (19) de un modelo ARIMA y si existen, sin embargo, problemas importantes. La identificación y crítica del modelo puede hacerse consistentemente en la representación clásica ARIMA como ecuación en diferencias estocásticas y no en la formulación en espacio de los estados, que si es conveniente para interpretar el modelo desde un punto de vista de previsión. Creemos que los modelos ARIMA son más generales y permiten utilizar la información muestral en forma más eficiente, lo que debe conducir a mejores previsiones. Como ilustración de este último punto utilizaremos la serie analizada por Morrison y Pike /16/ para ilustrar la eficacia de su método.

## 5. APLICACIONES

Morrison y Pike efectúan previsiones de la serie de demanda de electricidad en USA en el período (1945-1972). Estiman tres modelos -- utilizando los 14 primeros años de datos y -- general previsiones año a año desde 1958 en adelante. Según el criterio de error cuadrático medio mínimo, escogen el modelo dado por (23), (24), (25). Vamos a comparar sus resultados con los obtenidos por un análisis, mediante modelos ARIMA, utilizando la metodología de Box y Jenkins.

Tomaremos únicamente, como hacen estos autores, los 14 datos iniciales como muestra dada para identificar y estimar el modelo. La Serie es claramente no estacionaria (figura 1) y heterocedástica. El diagrama rango/media tomando las observaciones en grupos de 4 tiene el aspecto de la figura 2. Dado el crecimiento línea del Rango con la media local, el logaritmo parece adecuado, aunque quizás una transformación de la fami-

lia Box-Cox en  $0 < \lambda < 1$  podría ser más eficiente. Si aceptamos el logaritmo, la transformación  $\nabla \ln Z$  produce, prácticamente, estacionariedad. Los coeficientes de autocorrelación simple (cas) y parcial (cap) de la serie  $\nabla \ln Z$  con una y dos diferencias se encuentran en la tabla 1.

La serie es presumiblemente estacionaria con una diferencia, aunque su media es, en este caso, distinta de cero. Esto representa un inconveniente ya que el modelo que obtendríamos con esta transformación:

$$\nabla \ln Z = 0,091 + a_t \quad (.01)$$

$$\hat{\sigma} = 0,030 \quad (26)$$

o equivalentemente:

$$Z_t = Z_{t-1} \cdot e^{0,091 u_t} = 1,10 Z_{t-1} \cdot u_t$$

donde  $u_t$  es log-normal, tiene una tendencia

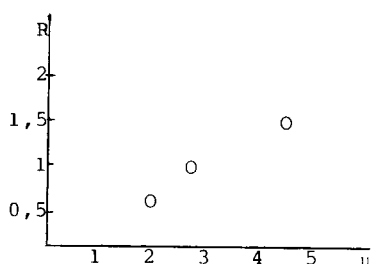


Fig. 2  
Diagrama Rango/Media

determinista que implica un crecimiento del 10% fijo anual. Para evitar este inconveniente y dado el pequeño número de observaciones disponibles, es conveniente trabajar con un modelo más adaptativo que no tenga una media fija. Esto nos conduce a un modelo con dos diferencias que, al estimar el parámetro MA, resulta:

$$\nabla^2 \ln Z = (1 - .86B)a \quad (.08)$$

$$\hat{\sigma} = 0,044 \quad (27)$$

Aunque la desviación típica aumenta en (27) respecto al modelo<sup>10</sup> (26), este modelo es más flexible por lo que, dado el pequeño número de datos de que disponemos, le escogeremos<sup>11</sup>. Una crítica frecuente, y desde nuestro punto de vista infundada, respecto a los modelos ARIMA es su dificultad de interpretación y este argumento se esgrime a favor de la representación en el espacio de los estados. Si desarrollamos el modelo anterior, tendremos:

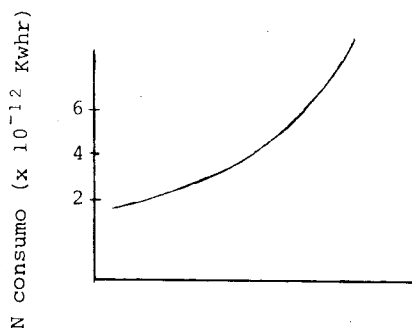


Fig. 1  
Gráfico de la serie

Tabla 1  
fas y fap de  $\nabla \ln Z_t$  y  $\nabla^2 \ln Z_t$

		1	2	3	4	5	6
$\nabla$	cas	-0,07	-0,26	-0,13	0,13	0,10	-0,16
	cap	-0,07	-0,27	-0,19	0,03	0,04	-0,14
$\nabla^2$	cas	-0,39	-0,09	-0,16	0,13	0,17	-0,14
	cap	-0,39	-0,29	-0,41	-0,28	0,00	-0,09



$$(1 - .86)^{-1} \nabla^2 \text{Ln } Z_t = a_t$$

$$(1 + .86B + .86^2 B^2 + \dots)(1 - B) \text{Ln } \frac{Z_t}{Z_{t-1}} = a_t$$

y llamando:

$$\text{Ln } T_{t-1} = .24 \text{Ln } \frac{Z_{t-1}}{Z_{t-2}} + .24 \cdot 86 \text{Ln } \frac{Z_{t-2}}{Z_{t-3}} + .24 \cdot 86^2 \text{Ln } \frac{Z_{t-3}}{Z_{t-4}} + \dots$$

$$T_{t-1} = \left(\frac{Z_{t-1}}{Z_{t-2}}\right)^{.24} \left(\frac{Z_{t-2}}{Z_{t-3}}\right)^{.21} \left(\frac{Z_{t-3}}{Z_{t-4}}\right)^{.18} \left(\frac{Z_{t-4}}{Z_{t-5}}\right)^{.15} \dots$$

$$v_t = e_t^a$$

El modelo puede escribirse:

$$\text{Ln } Z_t = \text{Ln } Z_{t-1} + \text{Ln } T_{t-1} + a_t \quad (23)$$

que indica que el valor (en logaritmos) de la serie en un instante es igual al valor inmediatamente anterior más una ponderación de las tasas de crecimiento en todos los períodos anteriores con pesos decrecientes en el tiempo, más el ruido. En la métrica inicial:

$$Z_t = Z_{t-1} \cdot T_{t-1} \cdot v_t \quad (29)$$

$$\frac{Z_t}{Z_{t-1}} = T_{t-1} \cdot v_t$$

que explicita como la tasa de crecimiento observada en  $t$  es explicada como una media de las tasas de crecimiento en meses anteriores, por el efecto de un ruido log-normal.

Las ecuaciones (28) y (29) describen la estructura empírica de variación deducible de los datos observados. Otra forma de escribir el modelo es estudiando su función de previsión, que será en logaritmos:

$$\text{Ln } \hat{Z}_t(\ell) = b_0(t) + b_1(t) \cdot \ell$$

Donde los coeficientes  $b_0(t)$  y  $b_1(t)$  se revisan con:

$$b_0(t) = b_0(t-1) + b_1(t-1)$$

$$b_1(t) = b_1(t-1) + \lambda \hat{a}_t$$

$$\text{siendo } \hat{a}_t = \text{Ln } Z_t - \text{Ln } \hat{Z}_{t-1}(1) \quad \text{y } \lambda = 1 - \theta =$$

$$= .24$$

En la métrica original:

$$\hat{Z}_t(\ell) = K_0(t) \cdot K_1(t)^\ell$$

y las ecuaciones de revisión de los parámetros son ahora:

$$K_0(t) = K_0(t-1) \cdot K_1(t-1)$$

$$K_1(t) = K_1(t-1) \cdot \hat{v}_t^\lambda$$

Donde  $\hat{v}_t = \frac{Z_t}{Z_{t-1}(1)}$ , es el cociente entre el valor observado y el previsto en el período anterior.

La tabla 2 recoge las previsiones a un período adelante<sup>12</sup> del modelo (27) y las obtenidas, para las mismas condiciones, por Morrison y Pike.

Los valores de la suma cuadrática de los errores de previsión están calculados siempre con dos decimales y mejoran los resultados dados por Morrison y Pike. Los valores entre paréntesis dan el error cuando la previsión se da redondeada, de manera que tenga la misma precisión que las observaciones reales, y son por lo tanto los errores que se calculan directamente de la tabla.

Los signos + y - junto a los valores previstos indican si la previsión ha estado por arriba o por debajo del valor observado. Aunque el pequeño número de rachas no es significativo

Tabla 2

Período	Observación real.	Previsión M/P	Previsión ARIMA
15	6,83	6,84 (+)	6,82 (-)
16	7,21	7,37 (+)	7,43 (+)
17	7,76	7,89 (+)	7,81 (+)
18	8,31	8,46 (+)	8,39 (+)
19	8,90	9,06 (+)	8,98 (+)
20	9,53	9,69 (+)	9,60 (+)
21	10,4	10,4 (=)	10,3 (-)
22	11,1	11,1 (=)	11,2 (+)
23	12,1	11,9 (-)	12,0 (-)
24	13,1	12,8 (-)	13,0 (+)
25	13,9	13,8 (-)	14,1 (+)
26	14,7	14,8 (+)	14,9 (+)
27	15,8	15,9 (+)	15,8 (+)
Suma cuadrática de errores		0,335 (0,28)	0,270 (0,19)

(con  $\alpha = 0,05$ ) para indicar dependencia de -- los errores, mediante un test de rachas, si sugiere serias dudas acerca de la aleatoriedad de los residuos estimados. Para estudiar más detalladamente estos residuos hemos analizado los errores de previsión de ambos modelos. La tabla 3 presenta los coeficientes de autocorrelación simple y parcial para los errores de previsión a un período de adelanto para ambos modelos. (ver Tabla 3)

El alto valor significativo del coeficiente de autocorrelación de orden 1 entre las estimaciones del modelo M/P nos indica que los errores están correlados y conduciría, en todo caso, a serias dudas sobre la bondad de la especificación. Los errores de previsión del modelo ARIMA no muestran indicios significativos de que se aparten de la estructura de ruido blanco.

Los errores de previsión del modelo M/P presentan la estructura:

$$(1 - .48B) \varepsilon_t = (1 + .98B) a_t \quad (26) \quad (.10)$$

Donde  $\varepsilon_t$  es la serie de errores de previsión y  $a_t$  es un proceso de ruido blanco.

## 6. CONCLUSIONES

La principal conclusión de este trabajo es - que la parametrización de modelos dinámicos con parámetros que varían con el tiempo, de manera que "encajen" en la formulación donde el filtro de Kalman es aplicable, presenta problemas importantes no resueltos y no es claro que suponga, hasta el momento, un avance respecto a los modelos paramétricos de series temporales recomendados por Box y Jenkins. En el caso de modelos de regresión una mala especificación de la función de transferencia y/o del ruido puede ser responsable de

Tabla 3

RETARDO:		1	2	3	4	5	6
M/P	Cas	0,62	0,17	-0,00	-0,25	-0,37	-0,31
	Cap	0,62	-0,35	0,13	-0,47	0,17	-0,30
ARIMA	Cas	-0,16	-0,33	+0,20	-0,03	-0,39	0,19
	Cap	-0,16	-0,36	0,08	-0,11	-0,38	-0,00

los problemas observados.

En el análisis y previsión de series univariantes, los modelos recomendados en la literatura son o bien casos particulares de los modelos ARIMA o presentan problemas de invertibilidad. El deseable carácter adaptativo de la función de previsión se consigue de -- forma más clara con un modelo ARIMA que permite diferenciar, metodológicamente, los parámetros fijos del modelo, de las constantes instantáneas que generan la función de previsión, que van adaptándose en el tiempo de acuerdo con la eficacia de las previsiones obtenidas. Una de las ventajas aducidas para la formulación y estimación mediante el filtro de Kalman, la adaptabilidad de las previsiones y la posibilidad de realizar previsiones "on-line", está pues, claramente implícita, en los modelos paramétricos ARIMA.

En resumen, pensamos que es necesario mucha más investigación para que la formulación en el espacio de los estados, en la línea que hemos comentado, pueda convertirse en una herramienta importante para la investigación empírica de sistemas económicos.

## 7. NOTAS

<sup>1</sup>Utilizaremos en este trabajo la expresión: vector ruido blanco, para un proceso vectorial con:

$$E \{ a_t \} = 0$$

$$E \{ a_t a_h \} = \Sigma \delta_{th}$$

Donde  $\delta_{th}$  es la delta de Kronecker.

<sup>2</sup>Todos los vectores que definimos en este trabajo son vectores columna. Utilizaremos la notación convencional  $x_t'$  para representar el vector transpuesto (vector fila).

El vector  $x_t$  puede ser un vector de variables predeterminadas que se fijan arbitrariamente y son por lo tanto conocidas para todo t (como suponemos ocurre con la matriz  $\Phi_t$  en el planteamiento estandar de la ecuación de estado (1)) o puede ser la observación en t de un proceso estocástico multivariante (que clásicamente en econometría se supone estacionario) en cuyo caso tendríamos

una "ecuación de estado" generalizada.

<sup>3</sup>Para iniciar el algoritmo recursivo necesitamos unos valores iniciales  $\hat{\beta}_0$  y  $P_{t=0}$  del vector de parámetros y de su matriz de varianzas y covarianzas. Estos valores pueden especificarse a priori con una justificación Bayesiana. Si queremos que los resultados finales no dependan de los valores iniciales podemos especificar  $\hat{\beta}_0 = 0$  y  $P_{t=0}$  diagonal con valores muy grandes en la diagonal principal, lo que indica que la varianza supuesta de los estimadores  $\hat{\beta}_0$  es muy grande y en muy pocas iteraciones el algoritmo se verá libre de las condiciones iniciales. Si disponemos de una muestra previa de observaciones en la que hemos estimado  $\hat{\beta}$  y  $P$ , estos valores pueden utilizarse para comenzar el cálculo recursivo. Obsérvese que el algoritmo no requiere ninguna inversión matricial lo que le hace muy eficiente para el cálculo automático.

<sup>4</sup>Relativo a su varianza estimada  $\sigma_t^2$ . La ecuación refleja que si  $e_t^*$  es cero, no hay revisión del valor de  $\hat{\beta}_t$ .

<sup>5</sup>El algoritmo de Newton para minimizar una función  $S(\hat{\beta})$  conduce a iteraciones con:

$$\hat{\beta}_t = \hat{\beta}_{t-1} - (H(\hat{\beta}_{t-1}))^{-1} \nabla S(\hat{\beta}_{t-1})$$

Siendo  $H(\hat{\beta}_{t-1})$  la matriz Hessiano y  $\nabla S(\hat{\beta}_{t-1})$  el vector gradiente de la función, ambos evaluados en la iteración  $t-1$ . Supongamos que la función a minimizar es cuadrática y del tipo:

$$S(\hat{\beta}) = (Y - X\hat{\beta})'(Y - X\hat{\beta}) = e'e$$

Entonces, si tenemos una solución inicial  $\hat{\beta}_{t-1}$ , la iteración siguiente para obtener  $\hat{\beta}_t$  vendrá dada por:

$$\begin{aligned} \frac{\partial S(\hat{\beta})}{\partial \hat{\beta}} &= \nabla S(\hat{\beta}_{t-1}) = 2 X'X \hat{\beta}_{t-1} - 2X'Y = \\ &= -2X'e_{t-1} = -2(\underline{x}_1 \dots \underline{x}_t) \begin{pmatrix} e_{1,t-1}^* \\ \vdots \\ e_{t,t-1}^* \end{pmatrix} \end{aligned}$$

donde  $\underline{x}_i$  es la columna correspondiente a las observaciones de las variables explicativas, en el instante  $t = i$ , y  $e_{i,t-1}^*$  es el error de previsión para las observaciones  $i$  cuando tomamos  $\hat{\beta}_{t-1}$  como valor del vector de parámetros. Si  $\hat{\beta}_{t-1}$  es un estimador mínimo-

cuadrático dadas las  $t-1$  primeras observaciones, los residuos para dichas  $t-1$  observaciones son ortogonales a las variables  $X$  observadas, por lo que:

$$\underline{x}_1 e_{1,t-1}^* + \dots + \underline{x}_{t-1} e_{t-1,t-1}^* = 0$$

y el vector gradiente en la iteración  $t$  será:

$$\nabla S(\hat{\beta}_{t-1}) = -2 \underline{x}_t e_{t,t-1}^*$$

el Hessiano vendrá dado por:

$$\frac{\partial^2 S(\hat{\beta})}{\partial \hat{\beta}_{t-1}^2} = H(\hat{\beta}_{t-1}) = 2(X'X)_t$$

y será siempre positivo definido. Por lo tanto, si sustituimos en la expresión general del algoritmo:

$$\hat{\beta}_t = \hat{\beta}_{t-1} + (X'X)_t^{-1} e_{t,t-1}^* \underline{x}_t'$$

que, teniendo en cuenta la definición de  $e_{t,t-1}^*$ , es precisamente (5).

<sup>6</sup>Las especificaciones de (4) suelen ser muy simples, generalmente con  $\phi_t$  y  $\Gamma_t$  iguales a la matriz unidad para todo  $t$ .

<sup>7</sup>Para ello es necesario especificar la forma de las matrices  $\phi_t$  y  $\Gamma_t$  para todo  $t$  y la matriz de varianzas y covarianzas del vector de perturbación  $\omega$ . El algoritmo procede entonces utilizando (4) para prever  $\hat{\beta}_{t/t-1}$  y  $P_{t/t-1}$ , previsión del vector de parámetros en  $t$  y de su matriz de covarianzas dada la información hasta  $t-1$ , que se "revisan" con (5), (6) y (7) al recibir nueva información.

<sup>8</sup> $v(B) = v_0 + v_1 B + v_2 B^2 + \dots$  donde las  $v_i$  son constantes a determinar y  $B$  el operador de retardo definido por  $Bx_t = x_{t-1}$ . El operador (posiblemente de orden infinito)  $v(B)$  se parametriza escuetamente mediante

$$v(B) = \frac{\omega(B) B^b}{\delta(B)}$$

Donde:  $\omega(B) = \omega_0 - \omega_1 B - \dots - \omega_S B^S$

$$\delta(B) = 1 - \delta_1 B - \dots - \delta_Y B^Y$$

$$\rho(B) = \frac{\theta(B)}{\phi(B)} = 1 + \rho_1 B + \rho_2 B^2 + \dots$$

La ecuación (10) descompone  $y_t$  en dos componentes ortogonales: La parte de  $y_t$  que es  $\underline{ex}$

plicada por las variaciones de  $x_t$  ( $v(B)x_t$ ) y la parte de  $y_t$  que se explica por la propia historia de  $y_t$  más un proceso de ruido blanco. La construcción de funciones de transferencia de este tipo ha sido estudiada detalladamente por Box y Jenkins /4/.

<sup>9</sup>Box y Jenkins imponen la restricción de que la parte MA sea invertible, que conduce al resultado intuitivo de que el pasado influirá menos en el presente cuanto más alejado esté. El modelo (15), (16) y (17) no satisface esa restricción y por lo tanto no entre dentro de la clase de modelos "razonables" propugnados por estos autores.

<sup>10</sup>Para precisar la importancia del aumento de la varianza podemos construir un intervalo de confianza para  $\hat{\sigma}_a^2$  utilizando que  $n \hat{\sigma}_a^2 / \sigma^2$  es una  $\chi^2$  con  $n$  grados de libertad. El intervalo del 95% es  $(0 - 0,073)$  que sugiere que el valor estimado  $\hat{\sigma}_a = 0,03$  puede variar apreciablemente en otras muestras si este modelo es correcto. El valor de  $\hat{\sigma}_a$  obtenido con el modelo (22) no es por lo tanto, en términos estadísticos, significativamente grande.

<sup>11</sup>La elección entre ambos modelos puede formalizarse en términos Bayesianos. Un contraste clásico con la  $t$  de Student para comprobar si son distintas las medias de los residuos en distintos tramos en el modelo (21) no descarta, al 90%, la hipótesis de igualdad, por lo que la segunda diferencia no aparecería justificada utilizando sólo la información muestral. Si adoptamos un punto de vista Bayesiano y especificamos una probabilidad a priori baja de que las medias sean iguales, como el número de observaciones disponibles es muy pequeño, la prior dominará la verosimilitud y la distribución posterior vendrá determinada por la prior, lo que justifica la elección del modelo con dos diferencias. En nuestra opinión cuando disponemos de una serie muy breve, como en este caso, es conveniente, en general, trabajar con un modelo que no contenga términos deterministas, lo que garantiza una mayor flexibilidad para la previsión.

<sup>12</sup>Es interesante comentar que habría ocurrido de haber elegido el modelo (21). Los errores de previsión serían más grandes y al -

cabo de 5 ó 6 periodos la evidencia a favor de tomar el modelo con dos diferencias sería muy grande, ya que aparecería significativamente un cambio de nivel en la media para el modelo  $\sqrt{\ln Z_t}$ .

## 8. AGRADECIMIENTOS

Agradezco a Antoni Espasa, del Servicio de Estudios del Banco de España, sus acertados comentarios a una versión preliminar de este trabajo.

## 9. REFERENCIAS

- /1/ ABRAHAM, B y G.E.P. Box.: "Deterministic and Forecast-adaptive time-dependent Models". Applied Statistics, 27, nº 2, pp. 120-13), 1978.
- /2/ AOKI, M.: "Optimization of Stochastic System". Academic-Press, 1967.
- /3/ ATHANS, M.: "The Discrete Time Linear-Quadratic-Gaussian Stochastic control -- problem". Annals of Economic and Social Measurement, 3, pp. 49-64, 1972.
- /4/ BOX, G.E.P. y G.M. JENKINS.: "Times Series Analysis, Forecasting and Control". Holden-Day, 1976.
- /5/ COOLEY, T. y PRESCOTT, E.C.: "Varying Parameter Regression: a Theory and some applications". Annals of Economic and Social Measurement. pp. 463-473, 1973.
- /6/ COOPER, J.P.: "Time Varying Regression - coefficients. A mixed estimation approach and operational limitations of the general markov structure". Annals of Economic and Social Measurement. pp. 525-530, 1973.
- /7/ CHOW, G.: "Analysis and Control of Dynamic Economic Systems. Wiley, 1975.
- /8/ GRANGER, C.W. y MORRIS, M.J.: "Time Series Modelling and Interpretation". Journal of Royal Statistical Society A. 139, 2 pp. 246-257, 1976.

- /9/ GRANGER, C.W. y NEWBOLD.: "Forecasting Economic Time Series". Academic Press, 1977.
- /10/ HANNAN, E.J.: "The identification and parameterization of ARMAX and State space models". *Econometrica*. Vol. 44, n° 4. pp. 713-723, 1976.
- /11/ HARRISON, P.J. y STEVENS, C.F.: "Bayesian Forecasting". *Journal of Royal Statistical Society*, 1976.
- /12/ HILDRETH, C. y HOUCK, J.: "Some estimator for a linear model with random coefficients". *Journal of American Statistical Association*. pp. 584-595, 1968.
- /13/ KALMAN, R.E.: "A new approach to linear Filtering and prediction Problems". *Journal of Basic Engineering*, (ASME), Vol. 82D, pp. 35-45, 1960.
- /14/ MARIANO, R.S. y SCHELICHER, S.: "On the use of Kalman Filters in Economic Forecasting". University of Pennsylvania. Dpto. of Economics Discussion paper 247, Philadelphia, 1972.
- /15/ MEDITH, J.S.: "Stochastic Optimal linear Estimation and Control. Mc Graw Hill, 1969.
- /16/ MORRISON, G.W. y PIKE, D.H.: "Kalman filtering applied to Statistical Forecasting". *Management Science*, Vol. 23, n° 7 pp. 768-774, 1977.
- /17/ PAGAN, A. "A note of the extraction of Components from series". *Econometrica*. Vol. 43, pp. 163-168, 1975.
- /18/ PLACKETT, R.L.: "Some theorems in least squares". *Biometrika*, 37 pp. 147-157, 1950.
- /19/ PEÑA, D.: "La Metodología de Box-Jenkins. Una aplicación a la previsión del consumo de Gasolina". *Información Comercial Española*. 542. pp. 135-142, 1978.
- /20/ ROSENBERG, B.: "The Analysis of a cross-section of Time Series by stochastically convergent parameter regression". *Annals of Economic And Social Measurement*. pp. 399-450, 1973.
- /21/ RUBIN, H.: "Note on Random coefficients". En *Statistical Inference in Dynamic Economic Systems*. Koopman editor, Wiley, 1950.
- /22/ TERCEIRO, J. "Aplicaciones del filtro de Kalman en Economía". IX Reunión Nacional de S.E.I.O.E., Barcelona, 1977.
- /23/ WISKWAKARMA, K.P. y otros.: "Optimal Prediction of Inter-Industry Demand". *Econometric Institute Report 7022*. Netherlands School of Economics. Rotterdam.
- /24/ YOUNG, P. "Recursive approaches to Time Series Analysis". *The Institute of Mathematics and its Applications*. pp. 209-224, 1974.
- /25/ BOX, G.E.P. & TIAO, G.C.: "Intervention analysis with applications to economic and environmental problems". *Journal of American Statistical Association*, v. 70, pp. 70-79. 1975.

