

# Modelos de Función de Transferencia

Aránzazu de Juan Fernández

Abril 2006

# 1 Introducción

En los modelos de función de transferencia, el objetivo es relacionar dos ó más series temporales elaborando modelos causales de predicción. Se considera la forma de relacionar una serie temporal, denominada *output* en función de una u otras series temporales, que se denominan *inputs*. También se considera a priori que existe una causalidad unidireccional desde los inputs hacia el output, desechando la posibilidad de *feedback*.

Una de las preguntas que surgen habitualmente es ¿por qué utilizar los modelos de función de transferencia si existe el modelo lineal general que relaciona diferentes variables?. Existen diversas razones que explican la necesidad de los modelos de función de transferencia:

1.- En el modelo lineal general, la relación suele ser instantánea y viene establecida a priori.

2.- En el modelo lineal general, la relación va de la variable  $x_t$  a la variable  $y_t$ , pero la variable  $y_t$  no influye sobre la variable  $x_t$ .

3.- En el modelo lineal general, la parte de variable respuesta no explicada por la variable independiente  $x_t$  es un proceso  $u_t$  de variables independientes. No se suele verificar con datos dinámicos.

Estos modelos son muy utilizados en todos los campos científicos para evaluar respuestas dinámicas. Si las variables *inputs* son controlables, estos modelos permiten simular y evaluar políticas alternativas. Si no lo son, ofrecen la posibilidad de estudiar cómo ciertos "escenarios", definidos por posibles evoluciones de la variable explicativa, afectan a la variable respuesta.

Además estos modelos son muy útiles para elaborar predicciones, aunque dependerá del intervalo que se tenga entre observaciones. Si disponemos de

datos anuales (es decir el intervalo entre observaciones es grande), la relación entre la variable explicativa y la respuesta no suele contener retardos y para prever el *output* es necesario conocer el *input*. En el caso más general en que  $x_t$  es estocástico, para preverla es necesario construir su modelo univariante, y es posible que la mejora en la predicción de  $y_{t+k}$  que aporte el modelo dinámico basado en la predicción de  $x_t$  respecto al modelo univariante para  $y_t$  sea pequeña, por el error de predicción de la variable exógena. Sin embargo, el modelo puede ser muy adecuado para la predicción condicional de la respuesta en función de posibles valores de la variable explicativa.

Si el período de observación es corto (diario o minutos), donde es frecuente que existan retardos en la relación entre las variables, de manera que  $x_{t-k}$  afecta a  $y_t$  ( $k > 0$ ). En este caso, la variable  $x_t$  es un indicador avanzado de la respuesta y como el regresor que necesitamos para prever la respuesta en el instante  $t$  es conocido, el modelo de regresión dinámica suele proporcionar mejores predicciones de  $y_t$  que el univariante.

La construcción de los modelos de función de transferencia sigue las mismas etapas que la construcción de los modelos univariantes de series temporales: Identificación, Estimación, Verificación y Predicción.

## 2 Modelos de función de transferencia con un único input

### 2.1 Conceptos generales

El caso más sencillo es el modelo de función de transferencia con un sólo input, que se puede extender sin grandes dificultades a modelos con varios inputs.

En un sistema lineal de un único input y único output, la serie output  $Y_t$  y la serie input  $X_t$  se relacionan a través de un filtro lineal de la siguiente forma:

$$\begin{aligned} Y_t &= v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + N_t \\ &= v(B) X_t + N_t \end{aligned} \quad (1)$$

donde  $v(B) = v_0 + v_1 B + v_2 B^2 + \dots$  se refiere como función de transferencia del filtro de Box y Jenkins (1970), y  $N_t$  es el ruido del sistema que es independiente de la serie input  $X_t$ . Los coeficientes de  $v(B)$  se conocen como la *función de respuesta al impulso*. Para que el sistema (1) sea estable se debe cumplir que una variación finita en el input produzca una variación también finita en el output. Esto es deberá cumplirse que

$$\sum_{j=0}^{\infty} v_j = g \quad (2)$$

siendo  $g$  finito. El valor de  $g$  representa el cambio total en  $Y_t$  motivado por un cambio unitario en  $X_t$  mantenido indefinidamente en el tiempo.

Los propósitos de la modelización de función de transferencia son identificar y estimar la función de transferencia  $v(B)$  y el modelo del ruido  $N_t$

sobre la base de la información que proporcionan la serie input y la serie output.

El modelo (1) es inestimable, debido a que en él aparece un número infinito de parámetros. Este problema se puede aliviar al expresar la función de transferencia como el cociente de dos polinomios (rationales) finitos:

$$v(B) = \frac{\omega(B)B^b}{\delta(B)} \quad (3)$$

donde  $\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s$ ,  $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$  y  $b$  es un parámetro de retardo que representa el tiempo que transcurre antes de que el impulso en la variable input produzca un efecto en la variable output. Para que el sistema sea estable, se asume que las raíces de  $\delta(B) = 0$  caen fuera del círculo unidad.

Sustituyendo (3) en (1), se obtiene:

$$Y_t = \frac{\omega(B)B^b}{\delta(B)} X_t + N_t \quad (4)$$

Por otro lado, el término de error no tiene por qué ser necesariamente un ruido blanco. Se puede suponer, con carácter general, que  $N_t$  sigue un proceso ARIMA, aunque sigue siendo independiente de la variable de input  $X_t$ , esto es:

$$\phi(B)(1 - B)^d N_t = \theta(B)u_t \quad (5)$$

ó

$$N_t = \frac{\theta(B)}{\phi(B)(1 - B)^d} u_t \quad (6)$$

con  $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$  y  $\phi(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$  de manera que todas las raíces de ambos polinomios caen fuera del círculo

unidad siendo  $(1 - B)^d$  el operador diferencias consecutivas utilizado para inducir estacionariedad y  $u_t$  es un ruido blanco.

Sustituyendo (6) en (4), se obtiene:

$$(1 - B)^d Y_t = \frac{\omega(B)B^b}{\delta(B)}(1 - B)^d X_t + \frac{\theta(B)}{\phi(B)}u_t \quad (7)$$

A partir de (7) puede observarse que si el proceso del término de error (5) no es estacionario y, por tanto, debe diferenciarse  $d$  veces para conseguir la estacionariedad, este mismo orden de diferenciación recae tanto sobre la variable dependiente o output como sobre la variable explicativa o input.

Sin embargo, en la práctica, para construir un modelo de función de transferencia, se requiere que tanto la variable dependiente (output) como la variable explicativa (input) sean estacionarias, para lo cual obviamente no tiene por qué cumplirse que ambas variables necesiten el mismo orden de diferenciación para lograr su estacionariedad. Por otra parte, una vez lograda la estacionariedad en ambas variables, el proceso  $N_t$  deberá ser un  $ARMA(p, q)$ , con lo cual el modelo de función de transferencia se puede escribir como:

$$y_t = \frac{\omega(B)B^b}{\delta(B)}x_t + \frac{\theta(B)}{\phi(B)}u_t \quad (8)$$

donde  $y_t = (1 - B)^d Y_t$  y  $x_t = (1 - B)^d X_t$ , siendo  $d'$  el orden de diferenciación de  $Y_t$  y  $d$  el orden de diferenciación de  $X_t$ .

Este modelo de función de transferencia se representa gráficamente como en la figura 1.

a.- *La estructura de la función de transferencia* determina la naturaleza de la influencia de la variable explicativa sobre la variable dependiente.

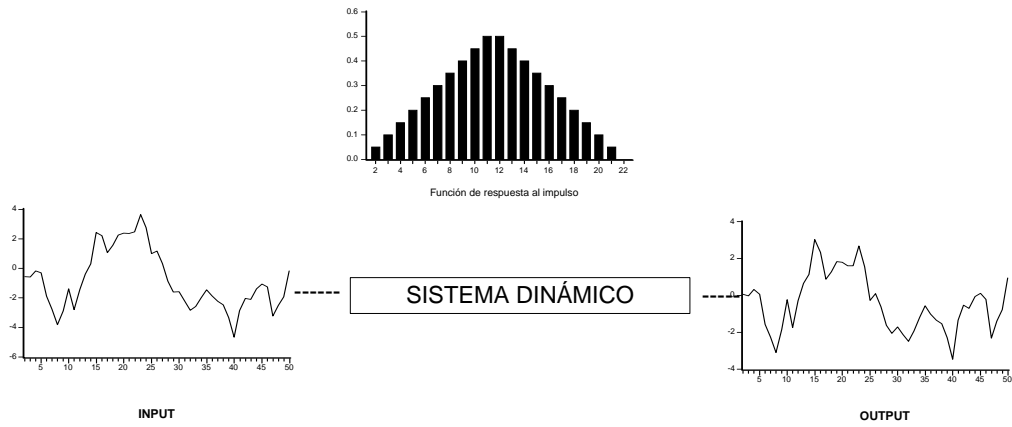


Figure 1: Input y Ouput a partir de un sistema dinámico

b.- *Modelo del ruido* representa un proceso *ARMA* estándar.

La conjunción de ambas partes da lugar al *modelo de función de transferencia completo*. (ver figura 2)

A partir de (3), resulta inmediato que, una vez que conozcamos los valores de  $\omega(B)$ ,  $\delta(B)$  y  $b$ , podremos obtener de inmediato los coeficientes de respuesta del impulso,  $v_j$ , igualando los coeficientes de  $B^j$  de ambos lados de la siguiente ecuación:

$$\delta(B)v(B) = \omega(B)B^b \quad (9)$$

que también puede escribirse como:

$$(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)(v_0 + v_1 B + v_2 B^2 + \dots) = (\omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s) \quad (10)$$

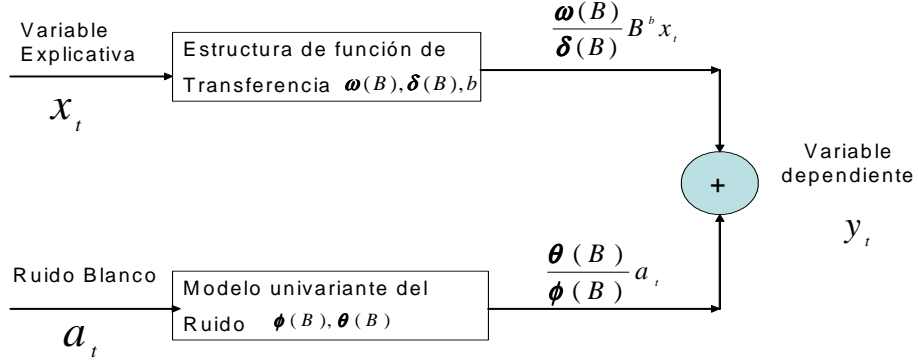


Figure 2:

A partir de esta expresión, se obtiene:

$$\begin{aligned}
 v_j &= 0 && \text{para } j < b \\
 v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} + \omega_0 && \text{para } j = b \\
 v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} + \omega_{j-b} && \text{para } j = b + 1, \dots, b + s \\
 v_j &= \delta_1 v_{j-1} + \delta_2 v_{j-2} + \dots + \delta_r v_{j-r} && \text{para } j > b + s
 \end{aligned} \tag{11}$$

Además, las  $r$  ponderaciones de respuesta al impulso  $v_{b+s}, v_{b+s-1}, \dots, v_{b+s-r+1}$ , sirven como valores iniciales para la ecuación en diferencias:

$$\delta(B)v_j = 0 \quad \text{para } j > b + s \tag{12}$$

De esta forma, las ponderaciones de respuesta al impulso presentan las siguientes características:

- a.- Hay  $b$  coeficientes iguales a 0 ( $v_0, v_1, \dots, v_{b-1}$ )



- b.- Hay  $s - r + 1$  coeficientes  $(v_b, v_{b+1}, \dots, v_{b+s-r})$  que no siguen un patrón de comportamiento fijo.
- c.- Hay  $r$  coeficientes de respuesta al impulso que sirven como valores iniciales  $(v_{b+s-r+1}, v_{b+s-r+2}, \dots, v_{b+s})$
- d.- Para  $j > b + s$ , los coeficientes de respuesta al impulso  $(v_j)$  siguen el patrón autorregresivo de orden  $r$  dado en (12).

Por lo tanto, si conocemos los coeficientes de respuesta al impulso  $(v_j)$ , podemos determinar los valores de  $b, r, s$  como sigue:

- 1.- *El valor de  $b$*  se determinará teniendo en cuenta el hecho de que  $v_j = 0$  para  $j < b$  y  $v_b \neq 0$ .
- 2.- *El valor de  $r$*  se determinará por el patrón de comportamiento de los coeficientes de respuesta al impulso de forma similar a como se identifica el orden  $p$  de un modelo *ARIMA* a través de la función de autocorrelación.
- 3.- En cuanto *al valor de  $s$* , para un valor dado de  $b$ , si  $r = 0$  podrá identificarse fácilmente  $s$  por cuánto se cumplirá que  $v_j = 0$  para  $j > b + s$ . Si  $r \neq 0$ , el valor de  $s$  se encontrará observando cuando comienza a decaer el patrón de la función de respuesta al impulso.

*Algunas funciones de respuesta al impulso típicas*

En la práctica, los valores de  $r$  y  $s$  raramente exceden de 2. Los casos más habituales que suelen encontrarse en variables económicas son:

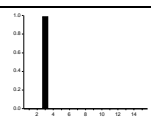
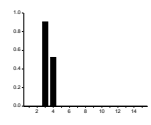
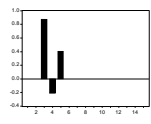
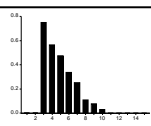
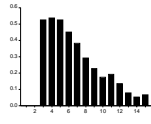
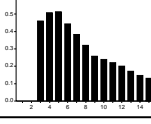
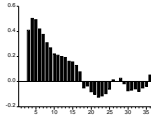
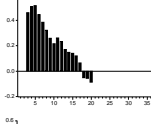
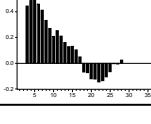
$(b, r, s)$	Función de Transferencia	Ponderaciones de respuesta al impulso
$r = 0$		
$(3, 0, 0)$	$v(B)x_t = \omega_0 x_{t-3}$	
$(3, 0, 1)$	$v(B)x_t = (\omega_0 - \omega_1 B)x_{t-3}$	
$(3, 0, 2)$	$v(B)x_t = (\omega_0 - \omega_1 B - \omega_2 B^2)x_{t-3}$	
$r = 1$		
$(3, 1, 0)$	$v(B)x_t = \frac{\omega_0}{(1-\delta_1 B)} x_{t-3}$	
$(3, 1, 1)$	$v(B)x_t = \frac{(\omega_0 - \omega_1 B)}{(1-\delta_1 B)} x_{t-3}$	
$(3, 1, 2)$	$v(B)x_t = \frac{(\omega_0 - \omega_1 B - \omega_2 B^2)}{(1-\delta_1 B)} x_{t-3}$	
$r = 2$		
$(3, 2, 0)$	$v(B)x_t = \frac{\omega_0}{(1-\delta_1 B - \delta_2 B^2)} x_{t-3}$	
$(3, 2, 1)$	$v(B)x_t = \frac{(\omega_0 - \omega_1 B)}{(1-\delta_1 B - \delta_2 B^2)} x_{t-3}$	
$(3, 2, 2)$	$v(B)x_t = \frac{(\omega_0 - \omega_1 B - \omega_2 B^2)}{(1-\delta_1 B - \delta_2 B^2)} x_{t-3}$	

Tabla 1: Funciones de respuesta al impulso típicas en variables económicas

*Tipo 1:  $r = 0$*

a)  $b = 3, s = 0$

En este caso, la ecuación (10) puede escribirse:

$$v_0 + v_1B + v_2B^2 + \dots = \omega_0B^3 \quad (13)$$

luego

$$\begin{aligned} v_3 &= \omega_0 \\ v_j &= 0 \quad \forall j > 2 \end{aligned} \quad (14)$$

b)  $b = 3; s = 1$

La ecuación (10) en este caso será:

$$v_0 + v_1B + v_2B^2 + \dots = (\omega_0 - \omega_1B)B^3 \quad (15)$$

luego

$$\begin{aligned} v_3 &= \omega_0 \\ v_4 &= -\omega_1 \\ v_j &= 0 \quad \forall j > 4 \end{aligned} \quad (16)$$

c)  $b = 3; s = 2$

La ecuación (10) en este caso será:

$$v_0 + v_1B + v_2B^2 + \dots = (\omega_0 - \omega_1B - \omega_2B^2)B^3 \quad (17)$$

luego

$$\begin{aligned}
v_3 &= \omega_0 \\
v_4 &= -\omega_1 \\
v_5 &= -\omega_2 \\
v_j &= 0 \quad \forall j > 5
\end{aligned} \tag{18}$$

*Tipo 2:  $r = 1$*

a)  $b = 3; s = 0$ : Los coeficientes de la respuesta al impulso decrecen exponencialmente a partir de  $v_{b+s} = v_3$ . En efecto, con  $b = 3, r = 1$  y  $s = 0$ , la ecuación (10) será igual a:

$$(1 - \delta_1 B)(v_0 + v_1 B + v_2 B^2 + \dots) = \omega_0 B^3 \tag{19}$$

obteniéndose:

$$\begin{aligned}
v_3 &= \delta_1 v_2 + \omega_0 = \omega_0 \\
v_4 &= \delta_1 v_3 = \delta_1 \omega_0 \\
v_5 &= \delta_1 v_4 = \delta_1^2 \omega_0 \\
v_j &= \delta_1^{j-3} \omega_0 \quad \forall j > 3
\end{aligned} \tag{20}$$

Además, como tiene que cumplirse que  $|\delta_1| < 1$ , se obtendrá claramente un decrecimiento exponencial de la función de respuesta al impulso a partir de  $v_3$ .

b)  $b = 3; s = 1$ : Los coeficientes de la respuesta al impulso decrecen exponencialmente a partir de  $v_{b+1} = v_4$ . La ecuación (10) será igual a:

$$(1 - \delta_1 B)(v_0 + v_1 B + v_2 B^2 + \dots) = (\omega_0 - \omega_1 B) B^3 \tag{21}$$

obteniéndose:

$$\begin{aligned}
v_3 &= \delta_1 v_2 + \omega_0 = \omega_0 \\
v_4 &= \delta_1 v_3 - \omega_1 = \delta_1 \omega_0 - \omega_1 \\
v_5 &= \delta_1 v_4 \\
v_j &= \delta_1 v_{j-1} \quad \forall j > 4
\end{aligned} \tag{22}$$

Ahora, el decrecimiento de la función de respuesta al impulso se producirá a partir de  $v_4$ .

- c)  $b = 3$ ;  $s = 2$ : El decrecimiento exponencial de la función de respuesta al impulso se producirá ahora a partir de  $v_{b+s} = v_5$ . La ecuación (10) será:

$$(1 - \delta_1 B)(v_0 + v_1 B + v_2 B^2 + \dots) = (\omega_0 - \omega_1 B - \omega_2 B^2) B^3 \tag{23}$$

A partir de esta ecuación, se obtiene:

$$\begin{aligned}
v_3 &= \delta_1 v_2 + \omega_0 = \omega_0 \\
v_4 &= \delta_1 v_3 - \omega_1 = \delta_1 \omega_0 - \omega_1 \\
v_5 &= \delta_1 v_4 - \omega_2 \\
v_6 &= \delta_1 v_5 \\
v_j &= \delta_1 v_{j-1} \quad \forall j > 5
\end{aligned} \tag{24}$$

Dentro de este tipo, el valor de  $b$  se identifica observando cuál es el primer coeficiente de la función de respuesta al impulso que es no nulo; el valor de  $s$  se identifica observando el número de coeficientes de la función de respuesta al impulso que no decrecen a partir de  $v_b$ .  $r = 1$  se identifica por cuanto a partir de  $v_{b+s}$  su comportamiento es como el de un proceso AR(1), es decir, la función de respuesta al impulso

presenta un comportamiento como el de la función de autocorrelación de un proceso AR(1).

*Tipo 3:  $r = 2$*

En este caso, los coeficientes de la función de respuesta al impulso exhiben a partir de  $v_{b+s}$ , bien un comportamiento de decrecimiento exponencial o bien un comportamiento sinusoidal decreciente, dependiendo de si las raíces de  $\delta(B) = 1 - \delta_1 B - \delta_2 B^2$  son reales ( $\delta_1^2 + 4\delta_2 \geq 0$ ) ó complejas ( $\delta_1^2 + 4\delta_2 < 0$ ).

## 2.2 Modelos de función de transferencia con múltiples inputs y estacionales

El modelo se puede generalizar a:

- 1.- La existencia de más de un input
- 2.- que el ruido viene generado por un proceso estacional multiplicativo

### 2.2.1 Generalización a k inputs

En este caso el modelo de función de transferencia sería:

$$\begin{aligned}
 y_t &= \frac{\omega_1(B)}{\delta_1(B)} x_{1,t-b} + \frac{\omega_2(B)}{\delta_2(B)} x_{2,t-b} + \dots + \frac{\omega_k(B)}{\delta_k(B)} x_{k,t-b} + N_t \quad (25) \\
 &= \frac{\omega_1(B)}{\delta_1(B)} B^{b_1} x_{1t} + \frac{\omega_2(B)}{\delta_2(B)} B^{b_2} x_{2t} + \dots + \frac{\omega_k(B)}{\delta_k(B)} B^{b_k} x_{kt} + N_t \\
 &= \sum_{i=1}^k \frac{\omega_i(B)}{\delta_i(B)} B^{b_i} x_{it} + N_t
 \end{aligned}$$

donde  $y_t = (1 - B)^d Y_t$  y  $x_t = (1 - B)^d X_t$ .

## 2.2.2 Generalización para un proceso estacional multiplicativo

El proceso del ruido en este caso sigue el modelo

$$N_t = \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D} u_t \quad (26)$$

con  $\Phi(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps})$  y  $\Theta(B^s) = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs})$ .

Reemplazando (26) en (25), se obtiene:

$$(1-B)^d(1-B^s)^D Y_t = \sum_{i=1}^k \frac{\omega_i(B)}{\delta_i(B)} B^{b_i} (1-B)^d(1-B^s)^D X_{it} + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)} u_t \quad (27)$$

Si el término de error de (26) no es estacionario y por tanto, debe diferenciarse mediante  $(1-B)^d(1-B^s)^D$ , este mismo orden de diferenciación recae tanto sobre  $Y_t$  como las  $k$  variables explicativas. Pero, dado que no es de esperar que todas las variables involucradas en el modelo requieran el mismo número de diferencias, la forma habitual de escribir el modelo de función de transferencia general es:

$$y_t = \sum_{i=1}^k \frac{\omega_i(B)}{\delta_i(B)} B^{b_i} x_{it} + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)} u_t \quad (28)$$

donde  $y_t = (1-B)^{d^0}(1-B^s)^{D^0} Y_t$  y  $x_{it} = (1-B)^{d_i}(1-B^s)^{D_i} X_{it}$ .

## 2.2.3 Generalización del modelo de función de transferencia con parámetros estacionales

El modelo de función de transferencia más general vendría dado por la expresión

$$y_t = \sum_{i=1}^k \frac{\omega_i(B)\Omega_i(B^{s_i})}{\delta_i(B)\Delta_i(B^{s_i})} B^{b_i} x_{it} + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)} u_t \quad (29)$$

donde  $\Omega_i(B^{s_i}) = (\Omega_{i0} - \Omega_{i1}B^{s_i} - \Omega_{i2}B^{2s_i} - \dots - \Omega_{Hi}B^{Hs_i})$  y  $\Delta_i(B^{s_i}) = (1 - \Delta_{1i}B^{s_i} - \Delta_{2i}B^{2s_i} - \dots - \Delta_{Ri}B^{Rs_i})$ .

### 3 Identificación de la función de transferencia

#### 3.1 La función de covarianza y correlación cruzada

La función de correlación cruzada es una medida muy útil de la fuerza y la dirección de la relación entre dos variables aleatorias. Dados dos procesos estocásticos  $x_t$  e  $y_t$  para  $t = 0, \pm 1, \pm 2, \dots$ , decimos que  $x_t$  e  $y_t$  son conjuntamente estacionarios si  $x_t$  e  $y_t$  son ambos procesos univariantes estacionarios y la covarianza cruzada entre  $x_t$  e  $y_t$  es una función de la diferencia temporal.

La covarianza cruzada se define:

$$\gamma_{yx}(j) = E[(y_t - \mu_y)(x_{t-j} - \mu_x)] \quad j = 0, \pm 1, \pm 2, \dots \quad (30)$$

ó

$$\gamma_{xy}(j) = E[(x_t - \mu_x)(y_{t-j} - \mu_y)] \quad j = 0, \pm 1, \pm 2, \dots \quad (31)$$

Dando valores a  $j$  se obtiene la función de covarianzas cruzadas.

Las covarianzas cruzadas no necesitan ser simétricas alrededor de 0. Por el contrario, es muy posible que  $x_t$  sea una variable que cause a  $y_t$  y por lo tanto, esté fuertemente correlacionada con valores futuros de  $y_t$ , de manera que  $\gamma_{yx}(j) \neq 0$ ,  $j > 0$  y que sin embargo no esté correlacionada con valores retardados de  $y_t$ , es decir,  $\gamma_{yx}(-j) = 0$ ,  $j > 0$ . Por lo tanto, en principio,

$$\gamma_{yx}(j) \neq \gamma_{yx}(-j) \quad (32)$$



y sin embargo, se cumple que

$$\gamma_{yx}(j) = \gamma_{xy}(-j) \quad (33)$$

Se deben usar series estacionarias ya que las covarianzas cruzadas sólo podrán interpretarse cuando se efectúen sobre series estacionarias.

*Interpretación de  $\gamma_{yx}(j) = E[(y_t - \mu_y)(x_{t-j} - \mu_x)]$ :*

- a) Un valor positivo de  $j$  en  $\gamma_{yx}(j)$ , denota retardo mientras que un valor negativo de  $j$  denota adelanto.
- b) En  $\gamma_{yx}(j)$ , el  $j$  escrito a la derecha de  $x$  se refiere a los desfases (si  $j$  es positiva) o adelantos (si  $j$  es negativa) de la variable  $x$  respecto al período presente  $t$  de la variable  $y$ . Así pues, si el valor presente de  $y$  viene influenciado por dos valores retardados de la variable  $x$ , deberá cumplirse que  $\gamma_{yx}(2) \neq 0$ . Esto obviamente también implica que  $\gamma_{xy}(-2) \neq 0$  dado que se cumple que  $\gamma_{yx}(2) = \gamma_{xy}(-2)$ .

Dado que suponemos que la variable  $x$  puede influir en  $y$  pero que la inversa no puede darse (esto es, existe causalidad unidireccional de  $x$  hacia  $y$ , pero no existe *feedback*) estamos suponiendo de partida que sólo nos interesarán valores no negativos de  $j$  en  $\gamma_{yx}(j)$ .

La función de correlación cruzada se define estandarizando la función de covarianzas cruzada, es decir,

$$\rho_{yx}(j) = \frac{\gamma_{yx}(j)}{\sigma_y \sigma_x} \quad (34)$$

siendo  $\sigma_y$  y  $\sigma_x$  las desviación típicas de las variables  $y$  y  $x$  respectivamente. Dando valores a  $j$ , se obtiene la función de correlaciones cruzadas. Estas no son simétricas alrededor de 0,  $\rho_{yx}(j) \neq \rho_{xy}(j)$ .

La función de correlación cruzada es útil para determinar:

- a) Si las series  $y$  y  $x$  son estacionarias.
- b) Si  $y$  y  $x$  muestran alguna relación entre ellas.
- c) El tipo de relación que existe entre ambas series.

### 3.2 Relación entre la función de correlación cruzada y la función de transferencia

Dado el modelo de función de transferencia:

$$y_t = v_0x_t + v_1x_{t-1} + v_2x_{t-2} + \dots + N_t \quad (35)$$

donde se supone que  $y_t$  y  $x_t$  son estacionarias y que  $\mu_x = \mu_y = 0$ . Si multiplicamos este modelo por  $x_{t-j}$  a ambos lados

$$y_t x_{t-j} = v_0x_t x_{t-j} + v_1x_{t-1} x_{t-j} + v_2x_{t-2} x_{t-j} + \dots + N_t x_{t-j} \quad (36)$$

y tomamos esperanzas matemáticas:

$$E[y_t x_{t-j}] = v_0E[x_t x_{t-j}] + v_1E[x_{t-1} x_{t-j}] + v_2E[x_{t-2} x_{t-j}] + \dots + E[N_t x_{t-j}] \quad (37)$$

se obtiene:

$$\gamma_{yx}(j) = v_0\gamma_{xx}(j) + v_1\gamma_{xx}(j-1) + v_2\gamma_{xx}(j-2) + \dots + v_j\gamma_{xx}(0) + v_{j+1}\gamma_{xx}(1) + \dots \quad (38)$$

ya que  $E[N_t x_{t-j}] = 0$ . Dividiendo por el producto de las desviaciones típicas de las series,  $\sigma_x \sigma_y$ , se obtiene:

$$\rho_{yx}(j) = \frac{\sigma_x}{\sigma_y} \left[ v_0 \frac{\gamma_{xx}(j)}{\sigma_x^2} + v_1 \frac{\gamma_{xx}(j-1)}{\sigma_x^2} + v_2 \frac{\gamma_{xx}(j-2)}{\sigma_x^2} + \dots + v_j \frac{\gamma_{xx}(0)}{\sigma_x^2} + \dots \right]$$

$$\rho_{yx}(j) = \frac{\sigma_x}{\sigma_y} [v_0 \rho_{xx}(j) + v_1 \rho_{xx}(j-1) + \dots + v_j + v_{j+1} \rho_{xx}(1)] \quad (39)$$

A partir de esta expresión se observa que la relación entre la función de correlación cruzada  $\rho_{yx}(j)$  y la función de respuesta al impulso aparece claramente contaminada por la estructura de autocorrelación de la serie input  $x_t$  ( $\rho_{xx}$  representa la función de autocorrelación de la variable input). Sin embargo, si la serie de input fuese un ruido blanco, todas las autocorrelaciones de  $x_t$  serían nulas y la expresión (39) se simplificaría considerablemente, obteniéndose:

$$\rho_{yx}(j) = \frac{\sigma_x}{\sigma_y} v_j \quad (40)$$

de manera que

$$v_j = \frac{\sigma_y}{\sigma_x} \rho_{yx}(j) \quad (41)$$

Así, si  $x_t$  es un ruido blanco, la función de respuesta al impulso  $v_j$  es directamente proporcional a la función de correlación cruzada  $\rho_{yx}(j)$ .

Para poder escribir una relación como (41), que es la relación básica en la identificación de los modelos de función de transferencia a través de la función de correlación cruzada, se debe considerar:

- 1.- La función de correlación cruzada se define sólo cuando  $x_t$  e  $y_t$  procesos bivariantes conjuntamente estacionarios. Para conseguir la estacionariedad requerida, puede que sean necesarias diferencias y transformaciones de estabilización de varianza. Se asume pues que  $x_t$  e  $y_t$  son conjuntamente estacionarias.
- 2.- Para poder relacionar la función de correlación cruzada y los coeficientes de la función de respuesta del impulso debemos asegurarnos que el

input sea ruido blanco, es decir,  $\rho_{xx}(j) = 0, \forall j \neq 0$ .

### 3.3 El preblanqueo

Para ello, si escribimos el modelo de función de transferencia como:

$$y_t = v(B)x_t + N_t \quad (42)$$

y suponemos que  $x_t$  viene generada por un proceso *ARMA* del tipo:

$$\phi_x(B)x_t = \theta_x(B)\alpha_t \quad (43)$$

donde  $\phi_x(B)$  y  $\theta_x(B)$  son los polinomios autorregresivo y de media móvil del proceso *ARMA* para la serie input que cumplen las condiciones de estacionariedad e invertibilidad usuales y  $\alpha_t$  es un proceso de ruido blanco. Así, si escribimos  $\alpha_t$ :

$$\alpha_t = \frac{\phi_x(B)}{\theta_x(B)}x_t \quad (44)$$

que se denomina serie input *preblanqueada*. Aplicando la misma transformación de preblanqueo a la serie de output, obtenemos:

$$\beta_t = \frac{\phi_x(B)}{\theta_x(B)}y_t \quad (45)$$

De esta forma, si en el modelo de función de transferencia (42) premultiplicamos a ambos lados por  $\frac{\phi_x(B)}{\theta_x(B)}$ , se obtiene

$$\frac{\phi_x(B)}{\theta_x(B)}y_t = v(B)\frac{\phi_x(B)}{\theta_x(B)}x_t + \frac{\phi_x(B)}{\theta_x(B)}N_t \quad (46)$$

Si denominamos  $\varepsilon_t = \frac{\phi_x(B)}{\theta_x(B)}N_t$ , tendríamos

$$\beta_t = v(B)\alpha_t + \varepsilon_t \quad (47)$$

ecuación a partir de la cual se obtiene fácilmente

$$v_j = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\beta\alpha}(j) \quad (48)$$

ya que  $\alpha_t$  es un proceso de ruido blanco. Así pues, si conociésemos los elementos poblacionales  $\rho_{\beta\alpha}(j)$ ,  $\sigma_\beta$  y  $\sigma_\alpha$ , obtendríamos los valores de  $v_j$  y, de acuerdo con las características de los coeficientes de la función de respuesta al impulso, podríamos identificar el modelo de transferencia pertinente. En la práctica, al no conocer los valores poblacionales de  $\rho_{\beta\alpha}(j)$ ,  $\sigma_\beta$  y  $\sigma_\alpha$ , tendremos que estimar estos coeficientes a partir de datos muestrales.

### 3.4 Función de correlación cruzada muestral

Para un conjunto de series temporales,  $x_t$  e  $y_t$ ,  $1 \leq t \leq n$ , la función de correlación cruzada:

$$\rho_{yx}(j) = \frac{\gamma_{yx}(j)}{\sigma_x \sigma_y} \quad j = 0, \pm 1, \pm 2, \dots \quad (49)$$

se estima por la función de correlación cruzada muestral

$$\hat{\rho}_{xy}(j) = \frac{\hat{\gamma}_{xy}(j)}{S_x S_y} \quad j = 0, \pm 1, \pm 2, \dots \quad (50)$$

donde

$$\hat{\gamma}_{xy}(j) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-j} (x_t - \bar{x})(y_{t+j} - \bar{y}) & j \geq 0 \\ \frac{1}{n} \sum_{t=1-j}^n (x_t - \bar{x})(y_{t+j} - \bar{y}) & j < 0 \end{cases} \quad (51)$$

$$S_x = \sqrt{\hat{\gamma}_{xx}(0)}, \quad S_y = \sqrt{\hat{\gamma}_{yy}(0)} \quad (52)$$

y  $\bar{x}$  e  $\bar{y}$  son las medias muestrales de las series  $x_t$  e  $y_t$ .

Para contrastar si algunos de  $\widehat{\rho}_{xy}(j)$  son significativamente distintos de 0 se comparan con sus errores estándar muestrales. Bajo la hipótesis de normalidad, Bartlett (1955) derivó las varianzas y covarianzas aproximadas entre dos correlaciones cruzadas muestrales  $\widehat{\rho}_{xy}(j)$  y  $\widehat{\rho}_{xy}(j+k)$ . La covarianza, bajo la hipótesis de que las dos series están incorreladas y que la serie  $x_t$  es ruido blanco, viene dada por:

$$\text{cov} [\widehat{\rho}_{xy}(j), \widehat{\rho}_{xy}(j+k)] \simeq (n-j)^{-1} \rho_{yy}(k) \quad (53)$$

De lo que se sigue que:

$$\text{Var} (\widehat{\rho}_{xy}(j)) \simeq (n-j)^{-1} \quad (54)$$

Así cuando la serie  $x_t$  es ruido blanco, se puede contrastar la hipótesis de que las dos series  $x_t$  e  $y_t$  no están correlacionadas de forma cruzada comparando los valores de la función de correlación cruzada muestral con sus errores estándar aproximados  $\frac{1}{\sqrt{(n-j)}}$ .

### 3.5 Identificación de la función de respuesta al impulso

Consideremos el modelo de función de transferencia con un sólo input

$$y_t = v(B)x_{t-b} + N_t = \frac{\omega(B)}{\delta(B)}x_t + N_t \quad (55)$$

con  $N_t = \frac{\theta(B)}{\phi(B)}u_t$ . La identificación de la función de respuesta al impulso  $v(B)$  se realizará siguiendo las siguientes etapas:

*Etapas 1:* Considerando que  $y_t$  y  $x_t$  son estacionarias, preblanquear el input de acuerdo con su representación *ARMA*

$$\phi_x(B)x_t = \theta_x(B)\alpha_t \implies \alpha_t = \frac{\phi_x(B)}{\theta_x(B)}x_t \quad (56)$$

donde  $\alpha_t$  es un ruido blanco con media cero y varianza  $\sigma_\alpha^2$ .

*Etapa 2:* Filtrar la serie del output utilizando el mismo filtro que preblanqueó el input, obteniendo:

$$\beta_t = \frac{\phi_x(B)}{\theta_x(B)} y_t \quad (57)$$

*Etapa 3:* Calcular la función de correlación cruzada muestral entre  $\beta_t$  y  $\alpha_t$  para estimar  $v_j$  de acuerdo con la expresión:

$$\hat{v}_j = \frac{\hat{\sigma}_\beta}{\hat{\sigma}_\alpha} \hat{\rho}_{\beta\alpha}(j) \quad (58)$$

Para analizar qué coeficientes de la función de respuesta al impulso son significativos, se puede utilizar el contraste de significación de las correlaciones cruzadas (Bartlett, 1946).

*Etapa 4:* Identificar  $r$  (orden del polinomio  $\delta(B)$ ) y  $s$  (orden del polinomio  $\omega(B)$ ) de acuerdo a las reglas descritas anteriormente. Una vez que se escogen  $b, r$  y  $s$ , se pueden encontrar estimaciones preliminares de los coeficientes de  $\delta(B)$  y  $\omega(B)$  a partir de sus relaciones con  $v_j$ . Así se obtiene una estimación preliminar de la función de transferencia:

$$\hat{v}(B) = \frac{\hat{\omega}(B)}{\hat{\delta}(B)} B^b \quad (59)$$

Las etapas descritas se realizan suponiendo:

- 1.- Se procede a identificar  $v(B)$  considerando que no hay error, esto es  $N_t = 0$ .
- 2.- Se parte de que las series  $y_t$  y  $x_t$  son estacionarias.

### 3.6 Identificación del modelo del ruido

Una vez identificada la función de transferencia, se puede calcular la serie del ruido estimada:

$$\hat{N}_t = y_t - \hat{v}(B)x_t = y_t - \frac{\hat{\omega}(B)}{\hat{\delta}(B)}B^b x_t \quad (60)$$

identificándose el proceso *ARMA* generador de los residuos mediante sus funciones de autorrelación simple y parcial muestrales. Se obtiene así:

$$\phi(B)\hat{N}_t = \theta(B)a_t \quad (61)$$

Conjugando ambas identificaciones, se llega al modelo de función de transferencia:

$$y_t = \frac{\omega(B)}{\delta(B)}B^b x_t + \frac{\theta(B)}{\phi(B)}a_t \quad (62)$$

### 3.7 Comentarios sobre la identificación del modelo de función de transferencia

- 1.- En la construcción del modelo, se asume que todas las variables  $y_t, x_t$  y  $n_t$  son estacionarias. Así, para series no estacionarias se tienen que llevar a cabo transformaciones para la estabilización de la varianza y diferenciación de las series para conseguir primero la estacionariedad.
- 2.- En el proceso de identificación de la función de transferencia  $v(B)$ , se preblanquea la serie input. El modelo preblanqueado se aplica para filtrar la serie de output, pero no necesariamente para preblanquearla. Este es un método normal y simple para construir un modelo de función de transferencia causal. Sin embargo, para construir un posible sistema no causal con fenómeno de *feedback*, donde  $y_t$  viene influida por  $x_t$



y  $x_t$  viene influida por  $y_t$ , ambas series de input y output deberían preblanquearse antes de examinar su función de correlación cruzada. Esto se denomina a menudo *preblanqueo doble* (Granger y Newbold, 1977, cap.7). Sin embargo, generalmente es más ventajoso modelizar un sistema no causal utilizando un proceso vectorial.

## 4 Estimación de modelos de función de transferencia

Después de la identificación tentativa de un modelo de función de transferencia

$$y_t = \frac{\omega(B)}{\delta(B)}x_{t-b} + \frac{\theta(B)}{\phi(B)}a_t \quad (63)$$

necesitamos estimar los parámetros:  $\delta = (\delta_1, \delta_2, \dots, \delta_r)'$ ;  $\omega = (\omega_0, \omega_1, \omega_2, \dots, \omega_s)'$ ;  $\phi = (\phi_1, \phi_2, \dots, \phi_p)'$ ;  $\theta = (\theta_1, \theta_2, \dots, \theta_q)'$  y  $\sigma_a^2$ . Se puede reescribir (63) de la siguiente manera:

$$\delta(B)\phi(B)y_t = \phi(B)\omega(B)x_{t-b} + \delta(B)\theta(B)a_t \quad (64)$$

ó

$$c(B)y_t = d(B)x_{t-b} + e(B)a_t \quad (65)$$

donde

$$\begin{aligned} c(B) &= \delta(B)\phi(B) = (1 - c_1B - c_2B^2 - \dots - c_{r+p}B^{r+p}) \\ d(B) &= \phi(B)\omega(B) = (d_0 - d_1B - d_2B^2 - \dots - d_{s+p}B^{s+p}) \\ e(B) &= \delta(B)\theta(B) = (1 - e_1B - e_2B^2 - \dots - e_{r+q}B^{r+q}) \end{aligned}$$

así:

$$a_t = y_t - c_1 y_{t-q} - \dots - c_{p+r} y_{t-p-r} - d_0 x_{t-b} - d_1 x_{t-b-1} - \quad (66)$$

$$\dots - d_{p+s} x_{t-b-p-s} + e_1 a_{t-1} + e_2 a_{t-2} + \dots + e_{r+p} a_{t-r-p} \quad (67)$$

donde  $c_i, d_j$  y  $e_k$  son funciones de  $\delta_i, \omega_j, \phi_k$  y  $\theta_l$ . Bajo la hipótesis de que los  $a_t$  son  $N(0, \sigma_a^2)$  (ruido blanco gaussiano), tenemos la función de verosimilitud condicional

$$L(\delta, \omega, \phi, \theta, \sigma_a^2/b, x, y, x_0, y_0, a_0) = (2\pi\sigma_a^2)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma_a^2} \sum_{t=1}^n a_t^2 \right] \quad (68)$$

donde  $x_0, y_0$  y  $a_0$  son valores iniciales para calcular los  $a_t$  a partir de (66) similares a los necesarios para la estimación de modelos *ARMA*. En general, se pueden utilizar los métodos de estimación de los modelos *ARIMA* para estimar los parámetros  $\delta, \omega, \phi, \theta$  y  $\sigma_a^2$ . Por ejemplo, haciendo las  $a_t$  desconocidas iguales a sus esperanzas condicionadas de 0, las estimaciones mínimo cuadráticas no lineales de estos parámetros se obtienen minimizando

$$S(\delta, \omega, \phi, \theta/b) = \sum_{t=t_0}^n a_t^2 \quad (69)$$

$a_t^2$  donde  $t_0 = \max\{p+r+1, b+p+s+1\}$ .

Nótese que hasta ahora se ha asumido que  $b$  es conocido. Para valores dados de  $r, s, p$  y  $q$ , si también necesitamos estimar  $b$ , se podría optimizar la ecuación (69) para un rango probable de valores de  $b$ . Entonces,  $b$  se selecciona de manera que sea el valor que proporciona el mínimo total de la suma de cuadrados.

En esta etapa de estimación, se presentan dos problemas (igual que en la estimación de un modelo *ARIMA*):

1.- Determinación de los valores iniciales

2.- En general, estos modelos son no lineales.

Por lo que respecta al problema de la determinación de los valores iniciales, caben dos enfoques en el proceso de estimación (al igual que ocurría en la etapa de estimación de los modelos *ARIMA*): el *enfoque condicional*, que consiste en estimar los parámetros considerando como dados los valores iniciales y el *enfoque exacto* (no condicional) mediante el cual se estiman conjuntamente tanto los parámetros como los valores iniciales.

Adoptando el método de estimación de máxima verosimilitud, en el primer enfoque se trata de maximizar la función de verosimilitud condicional, mientras que en el segundo enfoque deberá maximizarse la función de verosimilitud exacta (Hillmer y Tiao, 1979).

En el enfoque condicional, los problemas son menores por cuánto simplemente se tratará de maximizar la función de verosimilitud condicional:

$$L(\beta/b, x, y, x_0, y_0, a_0) = (2\pi\sigma_a^2)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma_a^2} S_*(\beta) \right] \quad (70)$$

donde  $S_*(\beta) = \sum_{t=1}^n [a_t(\beta/b, x, y, x_0, y_0, a_0)]^2 = \sum_{t=1}^n \hat{a}_t^2$  y  $\beta' = (\delta', \omega', \phi', \theta')$ , siendo  $a_t(\beta/b, x, y, x_0, y_0, a_0) = \hat{a}_t$  los valores que se obtienen (estiman) de  $a_t$  a partir de (67) supuestos conocidos los valores de los parámetros  $\beta$ , el valor de  $b$ , la base informativa  $(y, x)$  y los valores iniciales  $(x_0, y_0, a_0)$ .

Estos estimadores máximo verosimiles condicionales, caso de dejar el suficiente número de observaciones de  $x$  y de  $y$  sin utilizar para tomar dichos datos como valores iniciales y efectuando el supuesto de reemplazar los valores iniciales de la perturbación  $a_0$  por su valor esperado, esto es, por cero,

coinciden con los estimadores MC que se obtendrían de  $\delta, \omega, \phi$  y  $\theta$  tales que minimizan

$$\sum_{t=t_0}^n \hat{a}_t^2$$

siendo  $t_0 = \max \{p + r + 1, b + p + s + 1\}$ .

Dado el problema de la no linealidad, estos estimadores MC serían no lineales debiendo aplicar por ello procedimientos de estimación no lineales.

*Ejemplo de estimación de un modelo de función de transferencia*

Consideremos el siguiente modelo de función de transferencia:

$$y_t = \frac{\omega_0 + \omega_1 B + \omega_2 B^2}{1 - \delta B} x_t + \frac{1}{1 - \phi B} a_t \quad (71)$$

Si denotamos  $n_t = \frac{1}{1 - \phi B} a_t$ , podemos expresar el modelo:

$$y_t = \delta y_{t-1} + (\omega_0 + \omega_1 B + \omega_2 B^2) x_t + (1 - \delta B) n_t \quad (72)$$

en el que  $t_0 = \max \{r + 1, b + s + 1\} = 3$ . De manera que podemos hacer:

$$\begin{aligned} n_3 &= y_3 - \delta y_2 - \omega_0 x_3 - \omega_1 x_2 - \omega_2 x_1 + \delta n_2 \\ n_4 &= y_4 - \delta y_3 - \omega_0 x_4 - \omega_1 x_3 - \omega_2 x_2 + \delta n_3 \end{aligned} \quad (73)$$

.....

comenzando en  $n_2 = 0$  para después obtener las innovaciones  $a_t$  a partir de  $t = 4$ , por medio de

$$\begin{aligned} a_4 &= n_4 - \phi n_3 \\ a_5 &= n_5 - \phi n_4 \end{aligned} \quad (74)$$

.....

y formar por último la suma residual:

$$S^2(b, \delta, \omega, \phi) \equiv \sum_{t=4}^T a_t^2(b, \delta, \omega, \phi/x_0, y_0, a_0) \quad (75)$$

donde las primeras innovaciones se han tomado iguales a su esperanza condicional que es 0.

La expresión genérica de cada  $a_t$  es:

$$\begin{aligned} a_t = & n_t - \phi n_{t-1} = (y_t - \delta y_{t-1} - \omega_0 x_t - \omega_1 x_{t-1} - \omega_2 x_{t-2} + \delta n_{t-1}) - \\ & \phi(y_{t-1} - \delta y_{t-2} - \omega_0 x_{t-1} - \omega_1 x_{t-2} - \omega_2 x_{t-3} + \delta n_{t-2}) \end{aligned} \quad (76)$$

con vector gradiente:

$$\begin{aligned} \frac{\partial a_t}{\partial \delta} &= -y_{t-1} + n_{t-1} + \phi y_{t-2} - \phi n_{t-2} \\ \frac{\partial a_t}{\partial \omega_0} &= -x_t + \phi x_{t-1} \\ \frac{\partial a_t}{\partial \omega_1} &= -x_{t-1} + \phi x_{t-2} \\ \frac{\partial a_t}{\partial \omega_2} &= -x_{t-2} + \phi x_{t-3} \\ \frac{\partial a_t}{\partial \phi} &= -(y_{t-1} - \delta y_{t-2} - \omega_0 x_{t-1} - \omega_1 x_{t-2} - \omega_2 x_{t-3} + \delta n_{t-2}) \end{aligned} \quad (77)$$

por lo que podemos comenzar un algoritmo iterativo del tipo Gauss-Newton a partir de estimaciones iniciales  $\hat{\theta}_0 = (\hat{\delta}_0, \hat{\omega}_{00}, \hat{\omega}_{10}, \hat{\omega}_{20}, \hat{\phi}_0)$  tal y como se realiza en la estimación de modelos no lineales. Un modo de proceder es obtener series temporales de cada uno de los componentes del vector gradiente así como de los propios residuos  $a_t$ , utilizando las estimaciones iniciales de los parámetros para estimar la regresión:

$$\begin{aligned} \hat{a}_t = & (\hat{\delta}_0 - \delta) \frac{\partial a(\hat{\theta})}{\partial \delta} + (\hat{\omega}_{00} - \omega_0) \frac{\partial a(\hat{\theta})}{\partial \omega_0} + (\hat{\omega}_{10} - \omega_1) \frac{\partial a(\hat{\theta})}{\partial \omega_1} + \\ & (\hat{\omega}_{20} - \omega_2) \frac{\partial a(\hat{\theta})}{\partial \omega_2} + (\hat{\phi}_0 - \phi) \frac{\partial a(\hat{\theta})}{\partial \phi} + a_t \end{aligned} \quad (78)$$

que se repite, tomando en cada iteración las últimas estimaciones como iniciales, hasta lograr la convergencia. La aparición del vector  $\hat{\theta}$  hace referencia a que las componentes del vector gradiente de la función  $a_t$  están evaluadas en las preestimaciones. Lograda la convergencia, la matriz de varianzas covarianzas de las estimaciones se aproxima por el producto  $\hat{\sigma}_a^2 \left( \nabla_a \nabla_a^0 \right)^{-1}$ , donde  $\sigma_a^2$  se estima mediante el cociente del valor alcanzado por la suma residual en la última iteración y el número de observaciones ó de grados de libertad.

## 5 Verificación de un modelo de función de transferencia

Después de que el modelo haya sido identificado y sus parámetros estimados, es necesario contrastar la validez del modelo antes de utilizarlo en predicción, control y otros propósitos.

Los requisitos para que un modelo de función de transferencia sea adecuado son:

- 1.- que las estimaciones de sus parámetros sean significativas
- 2.- que los residuos se comporten como ruidos blancos
- 3.- que el modelo no omita parámetros relevantes ni incluya parámetros superfluos ó irrelevantes
- 4.- que el modelo sea estable.

Los requisitos 1 y 4 se comprueban de forma similar a los de los modelos *ARIMA*. Nos centramos en los otros dos grupos de contrastes utilizados para

efectuar la diagnosis: (a) Análisis de los residuos y (b) Omisión e inclusión de parámetros.

## 5.1 Implicaciones teóricas de una incorrecta especificación del modelo de función de transferencia

Si el modelo

$$y_t = \frac{\omega(B)}{\delta(B)}x_{t-b} + \frac{\theta(B)}{\phi(B)}a_t \quad (79)$$

estuviera bien especificado, los residuos deberían comportarse como un ruido blanco y estar incorrelacionados con las variables explicativas del modelo. Por lo tanto, ni las autocorrelaciones de  $\hat{a}_t$  ni las correlaciones cruzadas de  $\hat{a}_t$  y  $x_t$  deberían ser significativas. Sólo cuando el modelo está incorrectamente especificado el término de perturbación incumplirá las condiciones de ruido blanco e incorrelación con el input. Para verlo, escribamos el modelo (79) como:

$$y_t = v(B)x_t + \Psi(B)a_t \quad (80)$$

con  $v(B) = \frac{\omega(B)}{\delta(B)}$ ,  $\Psi(B) = \frac{\theta(B)}{\phi(B)}$ , siendo éste el modelo correcto. Supongamos ahora que hemos identificado incorrectamente el modelo siguiente:

$$y_t = v_0(B)x_t + \Psi_0(B)a_{0t} \quad (81)$$

En este caso, el término de perturbación del modelo incorrectamente especificado ( $a_{0t}$ ) será igual a:

$$a_{0t} = \Psi_0^{-1}(B) [y_t - v_0(B)x_t] \quad (82)$$

Sustituyendo (80) en (82):

$$\begin{aligned} a_{0t} &= \Psi_0^{-1}(B) [v(B)x_t + \Psi(B)a_t - v_0(B)x_t] \\ &= \Psi_0^{-1}(B) [v(B) - v_0(B)] x_t + \Psi_0^{-1}(B)\Psi(B)a_t \end{aligned} \quad (83)$$

A partir de (83) se observa cómo la perturbación del modelo incorrectamente especificado presentará autocorrelaciones y correlaciones cruzadas con el input ( $x_t$ ) y por lo tanto con el input preblanqueado,  $\alpha_t$ .

Ahora bien, la especificación incorrecta de un modelo de función de transferencia puede darse como consecuencia de identificar incorrectamente tan sólo el modelo del ruido ó la función de respuesta al impulso (parte sistemática del modelo de función de transferencia).

Veamos cada una de estas situaciones.

### 5.1.1 Identificación correcta de la función de respuesta al impulso e incorrecta del modelo del ruido

En este caso tendremos  $v(B) = v_0(B)$  y  $\Psi(B) \neq \Psi_0(B)$ , por lo tanto (83) podrá escribirse como:

$$a_{0t} = \Psi_0^{-1}(B)\Psi(B)a_t$$

ó

$$a_{0t} = \frac{\Psi(B)}{\Psi_0^{-1}(B)} a_t \quad (84)$$

Consecuentemente, en este caso, las  $a_{0t}$ 's no presentarán correlaciones cruzadas cruzadas con el input (ó inputs), aunque si presentarán autocorrelación.



### 5.1.2 Identificación incorrecta de la función de respuesta al impulso y correcta del modelo del ruido

En este caso, se cumplirá  $v(B) \neq v_0(B)$  y  $\Psi(B) = \Psi_0(B)$ , siendo (83) igual a:

$$a_{0t} = \Psi^{-1}(B) [v(B) - v_0(B)] x_t + a_t \quad (85)$$

Por lo tanto, en este caso, la perturbación estaría tanto autocorrelacionada como correlacionada cruzadamente con  $x_t$ .

La conclusión fundamental que se extrae del análisis de estas situaciones es que si, una vez estimado el modelo de función de transferencia, sus residuos presentan autocorrelación, pero no hay coeficientes de correlación cruzados con el input significativos, se puede concluir que deberemos actuar en el sentido de corregir la identificación del modelo del ruido.

Mayores dificultades surgen cuando tanto los coeficientes de autocorrelación de los residuos como los coeficientes de correlación cruzada de dichos residuos y el input son significativos. En este caso, puede estar mal identificado ó bien sólo la función de respuesta al impulso o bien tanto dicha función como el modelo del ruido. En este caso, es aconsejable identificar de nuevo la función de respuesta al impulso, incluyendo, por ejemplo, parámetros adicionales en  $\omega(B)$  y/ó  $\delta(B)$  y volver a identificar el modelo del ruido teniendo en cuenta esta nueva especificación en la parte sistemática.

En algún caso, las propias correlaciones cruzadas entre los inputs y los residuos pueden sugerir alguna identificación para reespecificar la función de respuesta al impulso.

Para analizar esta situación, si escribimos en modelo de función de trans-

ferencia en su forma "preblanqueada"

$$\beta_t = v(B)\alpha_t + N_t \quad (86)$$

donde  $\beta_t = \frac{\phi_x(B)}{\theta_x(B)}y_t$  y  $N_t = \frac{\phi_x(B)\theta(B)}{\theta_x(B)\phi(B)}a_t$ , siendo  $\phi_x(B) = 1 - \phi_{x1}L - \dots - \phi_{xp}L^p$   
 $\alpha_t = \frac{\phi_x(B)}{\theta_x(B)}x_t$   $\theta_x(B) = 1 - \theta_{x1}L - \dots - \theta_{xq}L^q$ .

A partir de (86), podemos escribir:

$$v_j = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\beta\alpha}(j) \quad j = 0, 1, 2, \dots \quad (87)$$

Supongamos ahora que especificamos incorrectamente el modelo como sigue:

$$\beta_t = v_0(B)\alpha_t + N_{0t} \quad (88)$$

Así pues, el ruido de este modelo es igual a:

$$N_{0t} = \beta_t - v_0(B)\alpha_t \quad (89)$$

Reemplazando (86) en (89)

$$N_{0t} = v(B)\alpha_t + N_t - v_0(B)\alpha_t = [v(B) - v_0(B)]\alpha_t + N_t \quad (90)$$

De manera análoga a como se obtiene (87), ahora tendríamos:

$$v_j - v_{0j} = \frac{\sigma_{N_0}}{\sigma_\alpha} \rho_{N_0\alpha}(j) \quad j = 0, 1, 2, \dots \quad (91)$$

Por lo tanto, a partir de (91) se desprende que las correlaciones cruzadas entre la variable explicativa preblanqueada y los residuos pueden utilizarse para evaluar qué cambio deberían efectuarse en el polinomio  $v_0(B)$  y consecuentemente en  $\omega_0(B)$  y  $\delta_0(B)$  de manera que se mejore la especificación del modelo.

De todo lo expuesto, cabe concluir que un doble requisito indispensable para que un modelo de función de transferencia pase la etapa de verificación consiste en que los residuos se comporten como un ruido blanco y que no presenten correlaciones cruzadas significativas con el input (inputs) preblanqueado (preblanqueados).

## 5.2 Análisis de los residuos

### 5.2.1 Análisis de las correlaciones cruzadas

Un requisito que debe cumplir un modelo de función de transferencia para que éste supere la etapa de verificación es que no existan correlaciones cruzadas significativas entre los residuos y el input preblanqueado. Se puede verificar el cumplimiento de este requisito:

- 1.- Contrastando la significatividad individual de las correlaciones cruzadas.
- 2.- Contrastando la significatividad conjunta de un grupo de correlaciones cruzadas.

Para analizar la significatividad individual se tiene en cuenta el resultado de Bartlett(1946),

$$Var [r_{\hat{a}\alpha}(j)] \simeq (T - j)^{-1} \quad (92)$$

siendo T el número de residuos calculados ( $T = N - t_0 + 1$ ), donde N es el número de observaciones originales y  $t_0 = \max \{p + r + 1, b + p + s + 1\}$ . Teniendo en cuenta (92), podemos afirmar con un nivel de confianza del 95% que no existe correlación cruzada significativa entre  $\hat{a}_t$  y  $\hat{\alpha}_t$  cuando:

$$|r_{\hat{a}\alpha}| < 1.96 \frac{1}{\sqrt{T - k}}$$

Conjuntamente con el análisis de significatividad, resulta de interés efectuar un contraste de significatividad conjunta. Esto es, contrastar la hipótesis nula de que las primeras  $M + 1$  correlaciones cruzadas entre  $\hat{a}_t$  y  $\hat{\alpha}_t$  son conjuntamente iguales a 0. En este caso, el estadístico tipo Ljung-Box (1978):

$$Q_0 = T(T + 2) \sum_{k=0}^M \frac{1}{T - k} r_{\hat{a}\alpha}^2(j) \quad (93)$$

se distribuye asintóticamente bajo la  $H_0$  como una  $\chi^2$  con  $(M + 1) - (s + 1) - r$  grados de libertad. En (93),  $T = N - t_0 + 1$ ;  $s + 1$  es el número de coeficiente  $\omega_j$  estimados y  $r$  es el número de coeficientes  $\delta_j$  estimados. El número de grados de libertad de  $Q_0$  es independiente del número de parámetros estimados en el modelo del ruido.

Así:

- Si  $Q_0 < \chi_{\varepsilon}^2(M - s - r)$ , no se puede rechazar la hipótesis nula al nivel de significación  $\varepsilon$ , de manera que el modelo supera este requisito de la verificación.

- Si  $Q_0 > \chi_{\varepsilon}^2(M - s - r)$ , se rechaza la hipótesis nula al nivel de significación  $\varepsilon$ , de manera que el modelo no supera este requisito de la verificación por lo que hay que modelizar una nueva función de respuesta al impulso y, en conformidad con ella, una nueva modelización del modelo del ruido de la función de transferencia.

### 5.2.2 Análisis de las autocorrelaciones

El segundo requisito que deben cumplir los residuos es que sean ruido blanco. Para ello, se deberá:

1.- Analizar las funciones de autocorrelación simple y parcial muestrales

de los residuos y comprobar que ambas se comportan como un ruido blanco.

2.- Contrastar de forma conjunta la significatividad de los coeficientes de la función de autocorrelación de los residuos. En este caso, se utiliza el contraste de Ljung-Box (1978)

$$Q_1 = T(T + 2) \sum_{j=1}^M (T - j)^{-1} r_j^2(\hat{a}_t) \quad (94)$$

el cual bajo la hipótesis nula de no significatividad conjunta de que las  $M$  primeras autocorrelaciones son conjuntamente iguales a 0, se distribuye asintóticamente como una  $\chi^2$  con  $M - k$  grados de libertad.

La regla para el contraste sería:

- Si  $Q_1 < \chi_\varepsilon^2(M - k)$ , no se puede rechazar la hipótesis nula para el nivel de significación  $\varepsilon$  y el modelo supera el requisito de la etapa de verificación.

- Si  $Q_1 > \chi_\varepsilon^2(M - k)$ , se rechaza la hipótesis nula para el nivel de significación  $\varepsilon$ . El modelo no supera el requisito de la etapa de verificación y habrá que volver a modelizar un nuevo modelo para el ruido.

## 6 Predicción y modelo óptimo

### 6.1 Introducción

Cuando el modelo estimado supera la etapa de la verificación, se puede utilizar el modelo de función de transferencia para obtener predicciones del output.

Se pueden calcular dos tipos de predicciones: puntual y por intervalos en las cuales se pueden distinguir dos situaciones: a) Se conocen los valores futuros del input (ó inputs) y b) como se desconocen los valores futuros del input (ó inputs) se deben realizar predicciones de los mismos.

## 6.2 Predicción puntual

### 6.2.1 Valores futuros de los inputs conocidos

El modelo de función de transferencia con un sólo input puede representarse como:

$$(1 - B)^{d^0} Y_t = \frac{\omega(B)}{\delta(B)} B^b (1 - B)^d X_t + \frac{\theta(B)}{\phi(B)} a_t \quad (95)$$

ó

$$Y_t = \frac{\omega(B) B^b (1 - B)^d}{\delta(B) (1 - B)^{d^0}} X_t + \frac{\theta(B)}{\phi(B) (1 - B)^{d^0}} a_t \quad (96)$$

Una forma alternativa de escribir el modelo (95) ó (96) es:

$$c(B) Y_t = d(B) B^b X_t + e(B) a_t \quad (97)$$

donde

$$\begin{aligned} c(B) &= \delta(B) \phi(B) (1 - B)^{d^0} = 1 - c_1 B - c_2 B^2 - \dots - c_{p+r+d^0} B^{p+r+d^0} \\ d(B) &= \omega(B) \phi(B) (1 - B)^d = d_0 - d_1 B - d_2 B^2 - \dots - d_{p+s+d} B^{p+s+d} \\ e(B) &= \delta(B) \theta(B) = 1 - e_1 B - e_2 B^2 - \dots - e_{q+r} B^{q+r} \end{aligned} \quad (98)$$

A partir de (97), haciendo  $b = 0$  podemos escribir:

$$\begin{aligned} Y_t &= c_1 Y_{t-1} + c_2 Y_{t-2} + \dots + c_{p+r+d^0} Y_{t-p-r-d^0} + \\ &\quad d_0 X_t - d_1 X_{t-1} - \dots - d_{p+s+d} X_{t-p-s-d} + \\ &\quad a_t - e_1 a_{t-1} - e_2 a_{t-2} - \dots - e_{q+r} a_{t-q-r} \end{aligned} \quad (99)$$

Luego referido al período  $T + l$ , esta expresión será igual a

$$\begin{aligned}
Y_{T+l} = & c_1 Y_{T+l-1} + c_2 Y_{T+l-2} + \dots + c_{p+r+d} Y_{T+l-p-r-d} + \\
& d_0 X_{T+l} - d_1 X_{T+l-1} - \dots - d_{p+s+d} X_{T+l-p-s-d} + \\
& a_{T+l} - e_1 a_{T+l-1} - e_2 a_{T+l-2} - \dots - e_{q+r} a_{T+l-q-r} \quad (100)
\end{aligned}$$

denotando ahora por  $\widehat{Y}_T(l)$  la predicción óptima (en el sentido de tener menor error cuadrático medio de predicción) efectuada para el período  $l$  condicionada a la información disponible hasta el período  $T$  y los valores conocidos de los inputs que son conocidos; esto es,

$$\widehat{Y}_T(l) = E [Y_{T+l}/I_T, X_{T+j}; 0 < j \leq \infty] \quad (101)$$

donde  $I_T$  recoge toda la información disponible en dicho período  $T$ . Este valor esperado se calculará teniendo en cuenta las siguientes reglas:

- 1.- Los valores presentes y pasados de las perturbaciones (esto es,  $a_{T+j}$ , para  $j \leq 0$ ) se reemplazarán por

$$a_{T+j} = Y_{T+j} - \widehat{Y}_{T+j-1}(1) \quad \text{para } j \leq 0$$

mientras que los valores futuros ( $a_{T+j}$ , para  $j > 0$ ) se reemplazarán por su valor esperado, el cual al ser un ruido blanco es igual a 0. Así pues:

$$E [a_{T+j}/I_T] = \begin{cases} 0 & \text{para } j > 0 \\ Y_{T+j} - \widehat{Y}_{T+j-1}(1) & \text{para } j \leq 0 \end{cases} \quad (102)$$

- 2.- Los valores presentes y pasados de las observaciones del input se reemplazarán por sus valores conocidos. Asimismo, como suponemos que los

valores futuros del input son conocidos, también reemplazamos éstos por dichos valores conocidos. Así pues:

$$E[X_{T+j}/I_T, X_{T+j}, 0 < j \leq l] = X_{T+j} \quad \forall j \quad (103)$$

3.- Los valores presentes y pasados del output se reemplazan por los valores observados, mientras que los valores futuros se reemplazan por sus valores predichos apropiados; esto es

$$E[Y_{T+j}/I_T, X_{T+j}, 0 < j \leq l] = \left\{ \begin{array}{l} Y_{T+j} \quad \text{para } j \leq 0 \\ \widehat{Y}_T(l) \quad \text{para } 0 < j < l \end{array} \right\} \quad (104)$$

De acuerdo con estas reglas, la predicción del output  $l$  períodos hacia adelante desde la observación  $T$ -ésima será igual a:

$$\begin{aligned} \widehat{Y}_T(l) = E[Y_{T+l}/I_T, X_{T+j}, 0 < j \leq l] = & c_1 \widehat{Y}_T(l-1) + \dots + c_{l-1} \widehat{Y}_T(1) + c_l Y_T + \\ & c_{l+1} Y_{T-1} + \dots + c_{p+r+d} Y_{T+l-p-r-d} + d_0 X_{T+l} - d_1 X_{T+l-1} - \dots \\ & - d_{p+s+d} X_{T+l-p-s-d} - e_l [Y_T - \widehat{Y}_{T-1}(1)] - e_{l+1} [Y_{T-1} - \widehat{Y}_{T-2}(1)] \\ & - \dots - e_{q+r} [Y_{T+l-q-r} - \widehat{Y}_{T+l-q-r-1}(1)] \end{aligned} \quad (105)$$

### 6.2.2 Valores futuros de los inputs desconocidos

El modelo de partida en este caso (un sólo input) será (95) ó (96) y el modelo *ARIMA* del input será

$$\phi_x(B)(1-B)^d X_t = \theta_x(B) \alpha_t \quad (106)$$

donde  $\phi_x(B) = (1 - \phi_{x1}B - \phi_{x2}B^2 - \dots - \phi_{xp}B^p)$  . El modelo (106) puede escribirse:

$$(1-B)^d X_t = \frac{\theta_x(B)}{\phi_x(B)} \alpha_t \quad (107)$$



A efectos de predicción puntual, la única modificación que existe en este caso consiste en que dado que los valores futuros del input no se conocen, éstos deben predecirse a partir del modelo (107).

## 6.3 Predicción por intervalos

### 6.3.1 Valores futuros de los inputs conocidos

El modelo

$$Y_t = \frac{\omega(B)B^b(1-B)^d}{\delta(B)(1-B)^{d^b}}X_t + \frac{\theta(B)}{\phi(B)(1-B)^{d^b}}a_t \quad (108)$$

puede escribirse como:

$$Y_t = v^*(B)X_t + \Psi^*(B)a_t \quad (109)$$

donde

$$\begin{aligned} v^*(B) &= \frac{\omega(B)B^b(1-B)^d}{\delta(B)(1-B)^{d^b}} = v_0^* + v_1^*B + v_2^*B^2 + \dots \\ \Psi^*(B) &= \frac{\theta(B)}{\phi(B)(1-B)^{d^b}} = 1 + \Psi_1^*B + \Psi_2^*B^2 + \dots \end{aligned} \quad (110)$$

Teniendo en cuenta (110), la expresión (109) puede escribirse:

$$Y_t = v_0^*X_t + v_1^*X_{t-1} + v_2^*X_{t-2} + \dots + a_t + \Psi_1^*a_{t-1} + \Psi_2^*a_{t-2} + \dots \quad (111)$$

Para el período  $T + l$ :

$$\begin{aligned} Y_{T+l} &= v_0^*X_{T+l} + v_1^*X_{T+l-1} + v_2^*X_{T+l-2} + \dots + \\ &v_{l-1}^*X_{T+1} + v_l^*X_T + v_{l+1}^*X_{T-1} + v_{l+2}^*X_{T-2} + \\ &\dots + a_{T+l} + \Psi_1^*a_{T+l-1} + \Psi_2^*a_{T+l-2} + \dots + \Psi_{l-1}^*a_{T+1} + \\ &\Psi_l^*a_T + \Psi_{l+1}^*a_{T-1} + \Psi_{l+2}^*a_{T-2} + \dots \end{aligned} \quad (112)$$

y la predicción puntual óptima desde el período  $T$  para  $l$  períodos hacia adelante, condicionada a que los valores de futuros del input son conocidos,

es igual a

$$\begin{aligned}\widehat{Y}_T(l) &= E[Y_{T+l}/I_T, X_{T+j}, 0 < j \leq l] = \\ &v_0^*X_{T+l} + v_1^*X_{T+l-1} + v_2^*X_{T+l-2} + \\ &+ \dots + v_{l-1}^*X_{T+1} + v_l^*X_T + v_{l+1}^*X_{T-1} + v_{l+2}^*X_{T-2} + \dots + \\ &\Psi_l^*a_T + \Psi_{l+1}^*a_{T-1} + \Psi_{l+2}^*a_{T-2} + \dots\end{aligned}\quad (113)$$

El error de la predicción efectuada en el período  $T$ ,  $l$  períodos hacia adelante  $[e_T(l)]$  será igual a la diferencia entre (112) y (113):

$$e_T(l) = Y_{T+l} - \widehat{Y}_T(l) = a_{T+l} + \Psi_1^*a_{T+l-1} + \Psi_2^*a_{T+l-2} + \dots + \Psi_{l-1}^*a_{T+1}$$

Como  $a_{T+l}, a_{T+l-1}, a_{T+l-2}, \dots, a_{T+1}$  son ruidos blancos, la esperanza matemática del error de predicción es igual a 0, y la varianza es igual a

$$Var [e_T(l)] = E [(e_T(l))^2] = \sigma_a^2(1 + \Psi_1^{*2} + \Psi_2^{*2} + \dots + \Psi_{l-1}^{*2}) \quad (114)$$

Por lo tanto, bajo la hipótesis de normalidad, esto es suponiendo que  $a_t \sim N(0, \sigma_a^2)$ , podemos escribir:

$$Y_{T+l} - \widehat{Y}_T(l) \sim N(0, Var [e_T(l)]) \quad (115)$$

obteniendo el intervalo de confianza para  $Y_{T+l}$ , para un nivel de significación  $\varepsilon$ , mediante la expresión .

$$\widehat{Y}_T(l) \pm N_{\frac{\varepsilon}{2}} DT [e_T(l)] \quad (116)$$

Cuando  $\varepsilon = 0.05$ , la expresión (116) es igual a

$$\widehat{Y}_T(l) \pm 1.96 DT [e_T(l)]$$

siendo  $DT [e_T(l)] = \sigma_a \sqrt{1 + \Psi_1^{*2} + \Psi_2^{*2} + \dots + \Psi_{l-1}^{*2}}$  y donde los valores de  $\Psi_j^*$ 's se calculan a partir de la expresión (110).

### 6.3.2 Valores futuros de los inputs desconocidos

El modelo de partida en este caso:

$$\begin{aligned} Y_t &= \frac{\omega(B)B^b}{\delta(B)(1-B)^{d^b}}X_t + \frac{\theta(B)}{\phi(B)(1-B)^{d^b}}a_t & (117) \\ (1-B)^d X_t &= \frac{\theta_x(B)}{\phi_x(B)}\alpha_t \end{aligned}$$

considerando conjuntamente ambas expresiones

$$Y_t = \frac{\omega(B)\theta_x(B)B^b}{\delta(B)\phi_x(B)(1-B)^{d^b}}\alpha_t + \frac{\theta(B)}{\phi(B)(1-B)^{d^b}}a_t \quad (118)$$

ó bien

$$Y_t = v^{**}(B)\alpha_t + \Psi^*(B)a_t \quad (119)$$

donde  $\Psi^*(B)$  está definido en (110) y

$$v^{**}(B) = \frac{\omega(B)\theta_x(B)B^b}{\delta(B)\phi_x(B)(1-B)^{d^b}} = v_0^{**} + v_1^{**}B + v_2^{**}B^2 + \dots \quad (120)$$

El modelo (119) los podemos escribir en  $T+l$  como sigue:

$$\begin{aligned} Y_{T+l} &= v_0^{**}\alpha_{T+l} + v_1^{**}\alpha_{T+l-1} + v_2^{**}\alpha_{T+l-2} + & (121) \\ &\dots + v_{l-1}^{**}\alpha_{T+1} + v_l^{**}\alpha_T + v_{l+1}^{**}\alpha_{T-1} + \\ &v_{l+2}^{**}\alpha_{T-2} + \dots + a_{T+l} + \Psi_1^*a_{T+l-1} + \Psi_2^*a_{T+l-2} + \dots + \\ &\Psi_{l-1}^*a_{T+1} + \Psi_l^*a_T + \Psi_{l+1}^*a_{T-1} + \Psi_{l-2}^*a_{T-2} + \dots \end{aligned}$$

La predicción puntual óptima desde el período  $T$  para  $l$  períodos hacia adelante, teniendo en cuenta que tanto las  $\alpha$ 's como las  $a$ 's son ruido blanco, es igual a

$$\begin{aligned} \widehat{Y}_T(l) &= E[Y_{T+l}/I_T] = v_l^{**}\alpha_T + v_{l+1}^{**}\alpha_{T-1} + v_{l+2}^{**}\alpha_{T-2} + \dots + & (122) \\ &\Psi_l^*a_T + \Psi_{l+1}^*a_{T-1} + \Psi_{l-2}^*a_{T-2} + \dots \end{aligned}$$

A partir de (121) y (122), calcularemos el error de predicción que, en este caso, será igual a:

$$e_T(l) = Y_{T+l} - \widehat{Y}_T(l) = v_0^{**} \alpha_{T+l} + v_1^{**} \alpha_{T+l-1} + v_2^{**} \alpha_{T+l-2} + \dots + v_{l-1}^{**} \alpha_{T+1} + a_{T+l} + \Psi_1^* a_{T+l-1} + \Psi_2^* a_{T+l-2} + \dots + \Psi_{l-1}^* a_{T+1}$$

Como consecuencia de que tanto las  $\alpha$ 's como las  $a$ 's son ruido blanco,  $E[e_T(l)] = 0$ , y su varianza, teniendo en cuenta que los inputs preblanqueados y las perturbaciones son independientes, será igual a:

$$\begin{aligned} Var(e_T(l)) &= E[(e_T(l))^2] = \sigma_\alpha^2 [v_0^{**2} + v_1^{**2} + \dots + v_{l-1}^{**2}] + \\ &\sigma_a^2 (1 + \Psi_1^{*2} + \Psi_2^{*2} + \dots + \Psi_{l-1}^{*2}) \end{aligned} \quad (123)$$

La predicción por intervalo, suponiendo que tanto  $\alpha_t$  como  $a_t$  se distribuyen como una normal, será en este caso igual a

$$\widehat{Y}_T(l) \pm N_{\frac{\varepsilon}{2}} DT[e_T(l)]$$

siendo  $DT[e_T(l)] = \sqrt{\sigma_\alpha^2 [v_0^{**2} + v_1^{**2} + \dots + v_{l-1}^{**2}] + \sigma_a^2 (1 + \Psi_1^{*2} + \Psi_2^{*2} + \dots + \Psi_{l-1}^{*2})}$ .

## 7 Ejemplo

En esta sección presentaremos un ejemplo de la construcción de un modelo de función de transferencia. Se trata de la relación dinámica existente entre los accidentes de carretera y las variables económicas, extraído del trabajo "*Forecasting traffic accidents using disaggregated data*", aceptado para su publicación en *International Journal of Forecasting*, realizado junto a *A. García-Ferrer y P. Poncela*.

En secciones previas a la construcción del modelo de función de transferencia, se analizan las posibles relaciones existentes entre las variables de accidentes de tráfico en España y variables económicas. De hecho, utilizando la metodología de detección de puntos de cambio y fechado de los ciclos económicos (García-Ferrer y Bujosa (2000)), se encuentran grandes similitudes en el fechado de los ciclos para las variables de accidentes y las variables económicas consideradas en este trabajo: IPI, matriculación de automóviles (NUVE) y consumo de gasolina (CGAS). Esto pone de manifiesto que entre ambos tipos de variables parece existir cierta relación. Parece lógico pensar que una mayor actividad económica, pueda generar una mayor utilización de los automóviles lo que conllevaría una mayor probabilidad de sufrir un accidente de tráfico debido al mayor número de vehículos en las vías de comunicación. Sin embargo, esa relación no tiene por qué ser instantánea, es decir, no quiere decir que una persona sale a la calle con su automóvil e inmediatamente sufra un accidente con una determinada probabilidad. Quiere decir que la mayor actividad económica se irá traduciendo en que progresivamente el número de vehículos en circulación irá aumentando, creciendo así de forma paulatina la probabilidad de accidente. Es decir, se puede esperar que

exista una relación dinámica entre el crecimiento económico y los accidentes de tráfico. Pero deben ser los datos los que nos certifiquen esta creencia y nos indiquen qué tipo de relación se puede encontrar. Por tanto, esperaríamos una relación de causalidad unidireccional de la actividad económica hacia los accidentes, dinámica, con algún retardo a la hora de dejar sentir sus efectos sobre la accidentalidad en las carreteras.

Con objeto de ilustrar todas las etapas de la construcción de un modelo de función de transferencia, presentaremos todos los pasos seguidos hasta encontrar el modelo, etapas que no fueron incluidas en el artículo de referencia por motivos obvios de espacio.

## 7.1 Análisis de los puntos de cambio y del fechado de los ciclos.

Utilizando la metodología desarrollada por García-Ferrer y Bujosa (2000), se obtuvieron los siguientes gráficos que muestran claramente la coincidencia existente entre los ciclos de las variables económicas y las variables de accidentes.<sup>1</sup>

En esta figura, se aprecia cierta anticipación de los ciclos en las variables económicas (IPI y NUVE) sobre las variables de accidentes consideradas (Accidentes totales (ACC), fallecidos por accidente de tráfico (FAT) y heridos en accidentes de tráfico (INJ)).

---

<sup>1</sup>Los datos de este ejemplo son los que figuran en el citado trabajo. Para la obtención de los modelos univariantes y de función de transferencia se utilizó la muestra 1975.01 a 2000.12, dejando las observaciones correspondientes a los años 2001, 2002 y 2003 para realizar la comparación de la capacidad predictiva de los modelos.

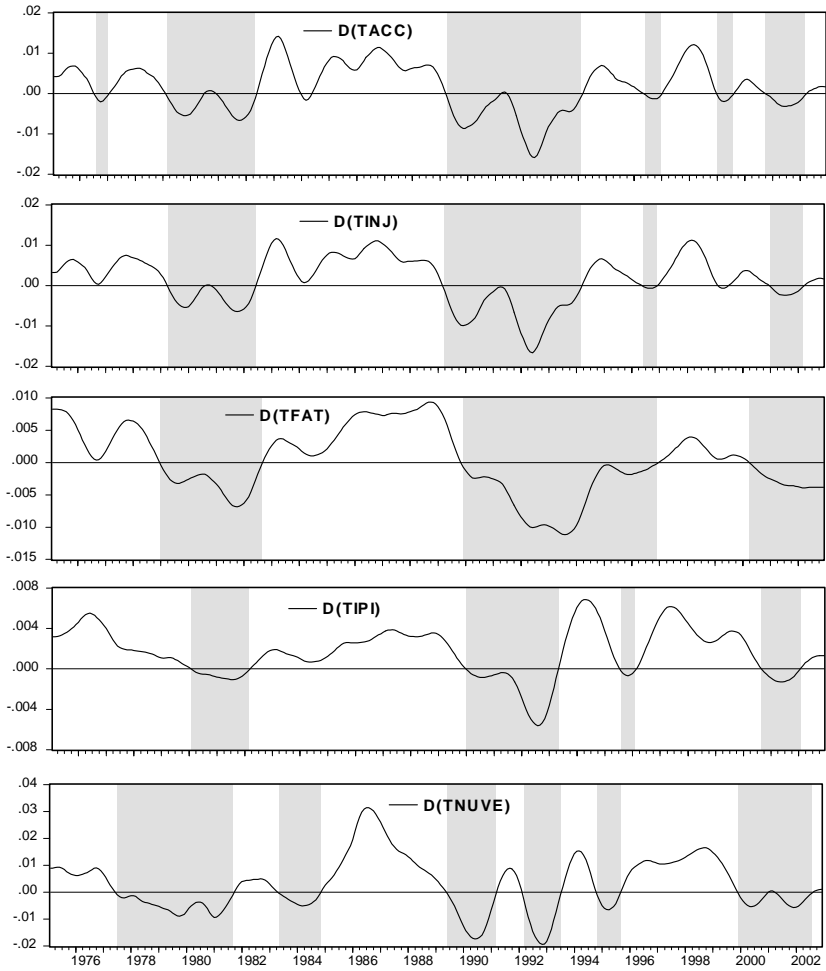


Figure 3: Derivadas de la tendencia de las variables de accidentes y de las variables económicas.

## 7.2 Identificación de los modelos univariantes y análisis de intervención

Utilizando el constraate de raíz unitaria, tanto regular como estacional, desarrollado por Osborn et al (1989), así como las funciones de autocorrelación simple y parcial de las variables en logaritmos, se identifica que las series necesitan una diferencia regular y estacional para conseguir la estacionariedad. A partir de las funciones de autocorrelación simple y parcial de estas variables estacionarias, se identifica para las series de IPI y de matriculaciones un modelo estacional multiplicativo  $MA(1) \times MA(1)_{s=12}$ . Es decir, los logaritmos de las variables objeto del análisis siguen un proceso  $ARIMA(0, 1, 1) \times ARIMA(0, 1, 1)_{s=12}$ .

Los modelos estimados para las variables que consideramos *INPUT* (*IPI* y *NUVE*) fueron los siguientes:

$$\nabla\nabla_{12} \ln(IPI) = (1 - \underset{(0.039)}{0.7404B})(1 - \underset{(0.042)}{0.6945B^{12}})a_t \quad (124)$$

$$\nabla\nabla_{12} \ln(NUVE) = (1 - \underset{(0.044)}{0.6658B})(1 - \underset{(0.031)}{0.8503B^{12}})a_t \quad (125)$$

Estos serán los modelos que se utilizarán para realizar el preblanqueo del *OUPTUT* para luego calcular la función de transferencia. Aunque todo lo que se ha visto sobre la identificación del modelo de función de transferencia se refería a un único input, se puede extrapolar sin dificultad al caso de dos input si, como ponen de manifiesto Liu y Hanssens (1982), los input están incorrelados. En este caso, la correlación entre los input es 0.25 de manera que se puede aplicar esta metodología.



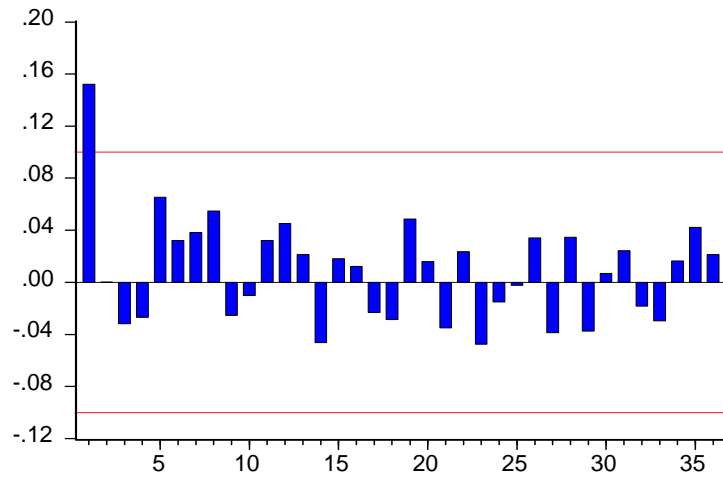


Figure 4: Función de correlación cruzada entre ACC e IPI

### 7.3 Preblanqueo e identificación de la función de transferencia

Utilizando los modelos (124) y (125) se realizó el preblanqueo del *OUTPUT*, se calculó la función de correlación cruzada estimada, obteniéndose las siguientes funciones de respuesta al impulso que aparecen en las figuras 4 y 5.

Se puede observar fácilmente que para el caso del IPI sólo se aprecia el retardo 1 estadísticamente significativo, de manera que la respuesta al impulso será  $\omega_0 B$ , es decir  $(b, s, r) = (1, 0, 0)$ , mientras que para el caso de NUVE se aprecian dos retardos estadísticamente significativos: 1 y 2, siendo entonces la función de respuesta al impulso  $(\omega_0 + \omega_1 B)B$ , es decir

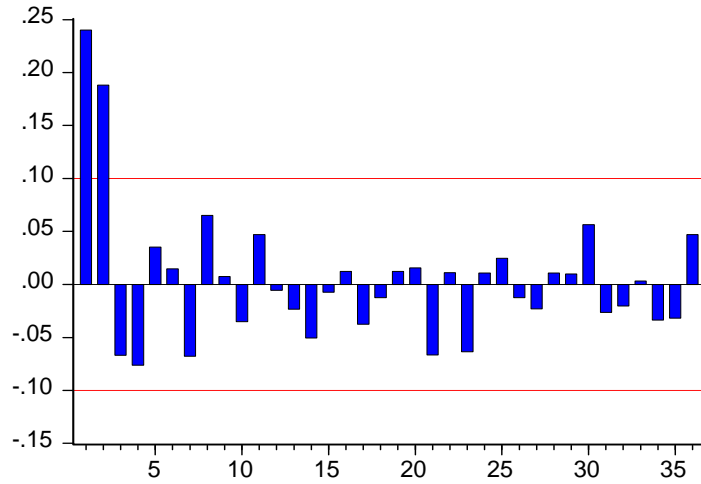


Figure 5: Función de correlación cruzada entre ACC y NUVE

$(b, s, r) = (1, 1, 0)$ . Así, el modelo de función de transferencia será entonces:

$$\nabla\nabla_{12} \ln(ACC) = \omega_0 B \nabla\nabla_{12} \ln(IPI) + (\omega_0 + \omega_1 B) B \nabla\nabla_{12} \ln(NUVE) \quad (126)$$

Al realizar la identificación del modelo del ruido se encontró que los residuos seguían un proceso  $MA(1) \times MA(1)_{s=12}$  y se identificaron diferentes atípicos que se incluyeron en el modelo de función de transferencia completo.

## 7.4 Estimación del modelo de función de transferencia

La estimación de este modelo resultó ser:

$$\begin{aligned}
\nabla\nabla_{12} \ln(ACC) = & -\underset{(.030)}{.100}\nabla\nabla_{12}JUN82 - \underset{(.031)}{.123}\nabla\nabla_{12}NOV93 & (127) \\
& + \underset{(.030)}{.096}\nabla\nabla_{12}APR95 + \underset{(.030)}{.062}\nabla\nabla_{12}MAR97 \\
& + \underset{(.060)}{.119}\nabla\nabla_{12} \ln(IPI) + \left( \underset{(.019)}{.066}B + \underset{(.018)}{.042}B^2 \right) \nabla\nabla_{12} \ln(NUVE) \\
& + \left( 1 - \underset{(.052)}{.454}B \right) \left( 1 - \underset{(.042)}{.728}B^{12} \right) a_t
\end{aligned}$$

La variable JUN92 se corresponde con un cambio de nivel identificado en todas las variables de accidentes analizadas en este artículo producido como consecuencia de un cambio en la legislación en cuanto a la utilización de los cinturones de seguridad. A partir de esa fecha, resultó ser obligatorio llevar puesto el cinturón de seguridad lo que produjo una disminución en el nivel de todas las variables de accidentes. El resto de variables ficticias incluidas se corresponden con comportamientos atípicos tipo AO detectados en el proceso de identificación de residuos atípicos, que están debidamente justificados en el artículo.

## 7.5 Verificación del modelo de función de transferencia

Siguiendo con las etapas del proceso de construcción de un modelo de función de transferencia se realiza la verificación del modelo mediante el análisis, por un lado de las funciones de autocorrelación simple y parcial de los residuos y por otro, mediante la función de correlación cruzada entre los residuos del modelo de función de transferencia y los residuos del modelo del *OUTPUT*.

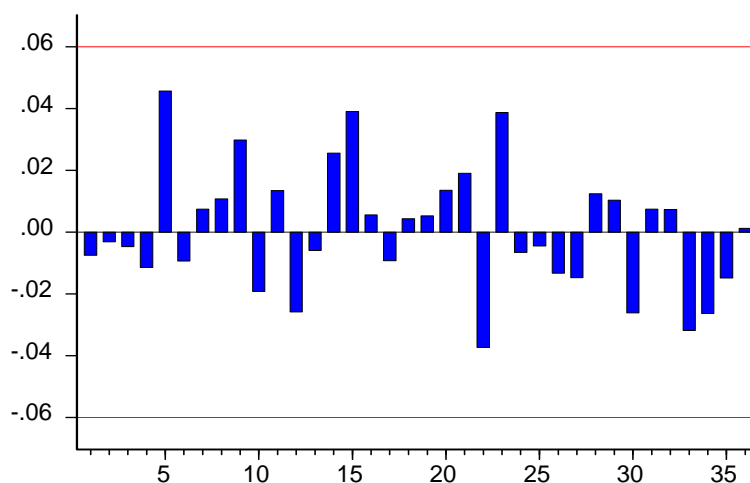


Figure 6: Función de correlación cruzada entre los residuos del modelo del *input IPI* y los residuos del modelo de función de transferencia

Por lo que respecta a los residuos del modelo estimado, los estadísticos de Ljung-Box de este modelo fueron:  $LBQ(12)=14.5$  y  $LBQ(24)=26.2$  lo que nos lleva a no rechazar la hipótesis nula de que los residuos son ruido blanco.

Por lo que respecta a la función de correlación cruzada entre los residuos del modelo de función de transferencia estimado y los residuos de los modelos univariantes de cada uno de los *INPUT* utilizados en la construcción del modelo de función de transferencia (*IPI* y *NUVE*), también aparecen "limpias" en el sentido de que no se puede decir que ninguna correlación cruzada sea significativa. (veáse figura 6 y 7)

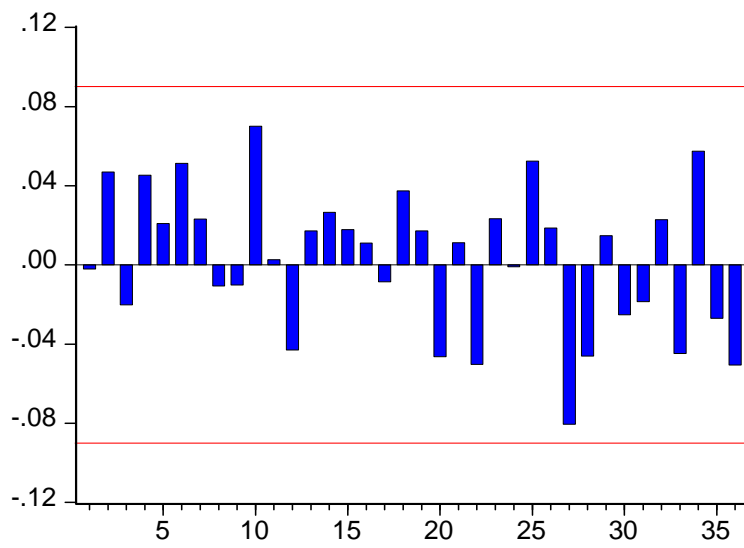


Figure 7: Función de correlación cruzada entre los residuos del modelo del *INPUT NUVE* y los residuos del modelo de función de transferencia

## 7.6 Predicción con el modelo de función de transferencia

Como se comentó anteriormente, en el modelo de función de transferencia, la predicción se realiza una vez que se conocen los valores del *INPUT* por lo que en muchas ocasiones es necesario reservar ciertas observaciones para poder realizar la predicción.

En este caso, los valores de los input (IPI y NUVE) se conocen antes que los valores del output (ACC) debido al retraso de 8 meses con el que se publican las cifras oficiales de accidentes por parte de la Dirección General de Tráfico, por lo que no es necesario reservar observaciones de la muestra para calcular las predicciones. Sin embargo, ya que los objetivos del artículo eran distintos (comparar la bondad predictiva de los modelos de los agregados frente a los modelos de los desagregados), se reservaron las observaciones de los años 2001, 2002 y 2003 para verificar si las predicciones se ajustan a lo que posteriormente sucede.

Se realizan predicciones horizonte 12 meses, para los años 2001, 2002 y 2003. Es decir, se estimó el modelo con las observaciones hasta el año 2000 y se realizaron predicciones para los doce meses del año 2001. Seguidamente, se reestimó el modelo con las observaciones hasta el año 2001 y se realizaron predicciones para el año 2002 y finalmente, se estimó el modelo hasta el año 2002 y se realizaron predicciones para el año 2003. Los resultados predictivos con este modelo fueron bastante satisfactorio, encontrándose valores para las medidas de capacidad predictiva (MAPE, RMSE, APE y FGR) bastante reducidas, por lo que el modelo se puede decir que funciona bien tanto a nivel de estimación como a nivel de predicción.

## 8 Bibliografía

- Anzar, A. y Trávez, F.J. (1993): *Métodos de Predicción en Economía*, vol 2, Ariel Economía.
- Bartlett, M.S. (1946): On the theoretical specification and sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, B*, 8, pp. 27-41.
- Bartlett, M.S. (1955): *Stochastic Processes*, Cambridge University Press.
- Box, G.E.P. y Jenkins, G.M. (1970): *Time Series Analysis: Forecasting and Control*, San Francisco, Holden-Day.
- García-Ferrer, A. y Bujosa, M. (2000): "Forecasting OCDE turning points using unobserved components models with business survey data. *International Journal of Forecasting*, 16(2), pp. 207-227.
- García-Ferrer, A., de Juan, A. y Poncela, P. (2005): Forecasting traffic accidents using disaggregated data. Aceptado para su publicación en *International Journal of Forecasting*.
- Granger, C.W.J. y Newbold, P. (1977): *Forecasting Economic Time Series*, Nueva York, Academic Press.
- Hamilton, J.D. (1994): *Time Series Analysis*, Princeton University Press.
- Hillmer, S.C. y Tiao, G.C. (1979): Likelihood function of stationary multiple autoregressive moving average models. *Journal of the American Statistical Association*, 74, pp. 652-660.

- Liu, L.M. y Hanssens, D.M. (1982): Identification of multiple-input transfer function models. *Communications in Statistics*, 11, pp. 297-314.
- Ljung, G.M. y Box, G.E.P. (1978): On a measure of lack of fit in time series models. *Biometrika*, 65, pp. 297-303.
- Novales, A. (1993): *Econometría*, Madrid, McGraw-Hill.
- Osborn, D.R., Chui, A.P.L., Smith, J.P. y Birchenhall, C.R. (1988): Seasonality and the order of integration for consumption. *Oxford Bulletin of Economics and Statistics*, 50, pp. 361-377.
- Peña, D. (2005): *Análisis de Series Temporales*, Ciencias Sociales, Alianza Editorial.
- Wei, W.W.S. (1990): *Time Series Analysis: Univariate and Multivariate Methods*, Redwood City, Addison Wesley.