

LA PREDICCIÓN ECONOMICA

Antoni Espasa

**Banco de España. Servicios de Estudios
Estudios Económicos, n.º 18, 1980**

Estoy muy agradecido a Pedro Martínez Méndez por sus comentarios a una primera versión de este trabajo.

INDICE

	<u>Páginas</u>
RESUMEN	7
I. INTRODUCCION	9
II. METODOS BASADOS EN EL ALISADO DE LAS SERIES	13
II.1. Método del alisado exponencial simple	14
II.2. El método Holt-Winters para series no estacionales.	16
II.3. El método Holt-Winters para series con estaciona- lidad multiplicativa	16
II.4. Comentarios a los métodos basados en el alisado exponencial	17
III. EL METODO BOX-JENKINS	19
III.1. El proceso de identificación y estimación	21
III.2. La validación de los resultados	25
III.3. La predicción	26
IV. EL METODO AUTORREGRESIVO POR ETAPAS	27
V. LA PREDICCIÓN ECONOMETRICA	29
APENDICE. Notas sobre los modelos ARIMA	35
BIBLIOGRAFIA	41

RESUMEN

En este trabajo se comentan distintos métodos de predicción, de menor a mayor complejidad, que pueden ser utilizados para la predicción de series económicas, especialmente de aquéllas de las que se dispone, al menos, de información mensual, como son las series monetarias y la mayoría de los indicadores de coyuntura. A pesar de esta conveniencia de información estadística mensual, los procedimientos y técnicas que se presentan en este trabajo son también adecuados para la predicción de series trimestrales e incluso anuales.

Los métodos del alisado exponencial y de la autorregresión por etapas, así como la combinación de ambos, son automáticos, en el sentido de que se pueden programar en un ordenador y luego pueden ser aplicados a un número considerable de series sin necesitar para ello de personal especializado que tenga que intervenir en el proceso de predicción. Este procedimiento puede ser de interés cuando hay que predecir un gran número de series. Los otros métodos basados en modelos ARIMA y modelos econométricos necesitan de personal especializado para construir dichos modelos y su aplicación se referirá, generalmente, a un número limitado de series especialmente relevantes para la toma de decisiones. No obstante hoy en día programar la predicción de un número grande de series económicas por estos últimos procedimientos más sofisticados tiene un costo que es asumible en bastantes ocasiones.

LA PREDICCIÓN ECONOMICA (*)

INTRODUCCION

En el mundo económico se plantea con frecuencia el problema de elegir entre alternativas diversas sin conocer con exactitud cuáles serán sus efectos, ya que éstos dependen de hechos que son desconocidos a la hora de tomar la decisión. A menudo la forma de actuar ante tal situación consiste en que el agente económico se forma una idea de cómo ocurrirán los hechos que desconoce y decide en consecuencia. Los inconvenientes de este proceder son:

1. Esta forma o técnica de apreciar o predecir el futuro no es transmisible ni intercambiable con otros agentes económicos.

2. Es difícil evaluar los éxitos, y

3. Es difícil aprender de los fracasos.

La razón de estos inconvenientes radica en que no se dispone de un sistema explícito para la predicción.

El desarrollo de métodos de predicción transmisibles, intercambiables y evaluables es un área de la investigación de gran importancia actual.

En este trabajo los métodos de predicción que se comentan se basan en la experiencia anterior observada en el sistema. La idea básica consiste en que estudiando la realización pasada de los

(*) Este trabajo se basa principalmente en los estudios de los profesores Granger y Newbold, publicados después de Granger (1966). Una sistematización de los resultados y experiencia obtenidos con tales trabajos se encuentra en el libro de Granger y Newbold (1977) que recomendamos al lector interesado en estos temas.

fenómenos económicos se puede identificar, al menos en parte, la pauta de evolución de los mismos.

No obstante, con frecuencia se dan cambios que no son detectables por el simple estudio del pasado. Por ello al realizar una predicción es de *suma importancia resaltar las hipótesis que se han hecho sobre el comportamiento del sistema.*

En toda predicción, no sólo la económica, hay un problema de costo-calidad que suele ser específico no sólo del fenómeno a predecir sino de la finalidad pretendida con la predicción y, por tanto, del agente o institución que la realiza o encarga. Dado que a mayor calidad de la predicción corresponde un costo mayor para su obtención, la decisión dependerá de qué infraestructura dispone el agente económico que necesita la predicción y de la sensibilidad de los beneficios, que espera con ella, respecto de la calidad de la predicción. En el mundo real, el examen profundo de estos aspectos es esencial y previo a la realización de la predicción. No obstante, dado que el problema de costo-calidad está muy ligado a situaciones concretas, no permite apenas un estudio teórico del mismo con abstracción de la realidad específica en que se pueda plantear. No obstante, el mencionado problema sí que tiene un efecto importante en la orientación que ha tenido y tiene la investigación sobre métodos de predicción. Este efecto ha consistido en que dicha investigación no se ha limitado, ni mucho menos, a métodos que teóricamente darían los mejores resultados, pero cuyo coste en determinadas situaciones los haría prohibitivos, sino que ha cubierto un amplio espectro de métodos con costes de aplicación distintos y por tanto de utilidad variable según las circunstancias. En parte debido a este tipo de razones nuestra exposición se concentrará principalmente en métodos de predicción univariante.

La predicción que tratamos en los primeros epígrafes es univariante en el sentido de que se basará en las realizaciones pasadas y presentes del fenómeno a predecir, pero que ignorará cualquier otra información existente sobre el universo en que se genera el fenómeno.

En el universo económico, un fenómeno particular no se genera independientemente de otros fenómenos, sino que aparece interrelacionado con ellos. Así, pues, la predicción univariante no es la óptima, en el sentido que ignora una información que puede ser útil. Sin embargo, hay razones poderosas para considerar tal predicción con detalle. En efecto:

1. La construcción de un modelo univariante en base al cual obtengamos predicciones es extraordinariamente más simple e inmensamente más barata que la construcción de un modelo multivariante;

2. Por otra parte, las mejoras en la predicción que se consigan con modelos multivariantes respecto a los modelos univariantes pueden ser sólo marginales. Por ello, es importante estudiar cuánta variación de una serie se puede explicar por su propio pasado y para qué tipo de comportamiento de la misma se necesita información sobre factores ajenos;

3. Así pues, la predicción univariante servirá como punto de comparación;

4. Además, con frecuencia se da el caso de que disponemos de información sobre el fenómeno que queremos predecir, pero no sobre otros fenómenos relacionados con él;

5. Por último, como veremos más adelante, las predicciones univariantes se pueden combinar con las predicciones por otros métodos.

Dado que la investigación sobre el tema se ha concentrado principalmente en predictores lineales en las observaciones pasadas de la serie a predecir, nos limitaremos a ellos en este trabajo.

II. METODOS BASADOS EN EL ALISADO DE LAS SERIES

Las ventajas de estos métodos radican en:

1. Que su aplicación tiene un coste más barato que la de otros métodos,
2. Dan predicciones con rapidez,
3. Son automáticos, y
4. No necesitan de personal especializado.

El inconveniente principal de estos métodos consiste en que, aun dentro de los métodos univariantes, sólo son óptimos¹ si la serie a la que se aplican viene generada por un tipo de estructura determinado, es decir, por un proceso estocástico específico². Por ello al exponer estos métodos será conveniente estudiar bajo qué hipótesis son óptimos.

La versión más simple de estos métodos concebiría la serie económica como formada localmente por un nivel (M) y un componente aleatorio (ϵ_t) con esperanza matemática cero y varianza σ^2 .

Dicha hipótesis sugiere que, promediando observaciones consecutivas, los elementos aleatorios de las distintas observaciones tenderán a cancelarse entre sí y el promedio reflejará más claramente

¹ Un predictor es óptimo dentro de una clase de predictores (por ejemplo los predictores lineales) si de entre ellos es él que minimiza el error cuadrático medio de las predicciones a un período por delante, es decir, las predicciones hechas en el momento t para el valor de la serie en $t + 1$.

² Para los conceptos de proceso estocástico, estacionariedad, procesos autorregresivos, de medias móviles y procesos ARMA y ARIMA, en general, véase el apéndice.

el nivel de la serie. Para tener en cuenta que éste cambia con el tiempo, el promedio a aplicar sería del tipo:

$$\bar{X}_t = \frac{1}{n+1} \sum_{j=0}^{j=n} X_{t-j}. \quad (1)$$

Ahora bien, esta estimación pondera por igual todas las observaciones, cuando la naturaleza de las series económicas que se dan en la realidad sugiere que sería conveniente dar más peso al pasado más reciente. Esta idea nos lleva a considerar ponderaciones decrecientes a medida que la observación se aleja de t .

II.1. Método del alisado exponencial simple

Este método sugiere que el fenómeno económico a predecir viene generado por el siguiente modelo:

$$X_t = M_t + \varepsilon_t \quad (2)$$

en el que M_t es el nivel local de la serie, que a su vez viene determinado por:

$$M_t = aX_t + a(1-a)X_{t-1} + a(1-a)^2X_{t-2} + \dots \quad (3)$$

o expresado de otra forma por:

$$M_t = M_{t-1} + a(X_t - M_{t-1}) = aX_t + (1-a)M_{t-1}. \quad (4)$$

Es decir, el nivel se estima en función del nivel estimado para el período anterior más una corrección en función de la nueva observación. En (4) se ve que cuanto mayor es a mayor es el impacto de la nueva observación en la estimación del nivel y cuanto menor es a (véase (3)) mayor es la memoria del sistema para estimar el nivel o, dicho de otra forma, mayor es el número de valores retardados de X que entran en la estimación de M_t .

Para hacer operativo el modelo se necesita fijar un valor inicial de M en base al cual se pueda aplicar recurrentemente (4). Bajo la hipótesis (2) el valor inicial lógico es:

$$M_1 = X_1$$

Dado que el componente aleatorio no es predecible, la predicción que hagamos en el tiempo n será el nivel de la serie tal como se estima en el momento de la predicción, es decir M_n . Esta es la predicción que en el momento n se hace, con este método, para

cualquier momento del tiempo posterior a n . Es decir, si llamamos $f_{n,h}$ a la predicción hecha en el período n , para el período $n+h$ tenemos que,

$$f_{n,h} = M_n \quad \forall h > 0.$$

Esta característica de que todas las predicciones futuras hechas en n sean iguales es el mayor inconveniente de este método. Por el contrario, su gran ventaja radica en que para realizar la predicción no necesitamos guardar todo el pasado de la serie sino solamente M_{n-1} y X_n , con lo que el costo de aplicación es bajísimo.

Es fácil demostrar que el método es óptimo cuando la serie viene generada por un modelo integrado y de medias móviles³ de orden uno, IMA (1, 1). Es decir, cuando:

$$X_t - X_{t-1} = \varepsilon_t - (1 - a) \varepsilon_{t-1} \quad (5)$$

o, si utilizamos el operador de retardos L , que es tal que

$$L^j X_t = X_{t-j}, \quad (6)$$

(5) se puede expresar así:

$$(1 - L) X_t = (1 - \theta_1 L) \varepsilon_t, \quad \theta_1 = (1 - a) \quad (7)$$

Obsérvese que para valores de $a=1$, la serie viene generada por el proceso denominado sendero aleatorio. Este proceso parece bastante adecuado por ejemplo para explicar la formación de precios en mercados altamente especulativos. Por otra parte, los estudios posteriores a Granger (1966), sobre la forma típica del espectro de las series económicas, parecen indicar (véase Granger y Newbold (1977)) que el proceso IMA (1, 1) —quizá con la inclusión de una constante— puede tomarse como una aproximación inicial para una amplia gama de series económicas que hayan sido previamente corregidas de variación estacional.

Con todo esto tenemos que el método de esta subsección es demasiado simplista, pero la experiencia sobre la estructura de las series económicas, principalmente aquellas que carecen de variación estacional, apunta que la utilización del método como la aproximación más burda al problema de predicción tiene cierto fundamento.

³ Para la explicación de este tipo de conceptos véase el apéndice.

II.2. El método Holt-Winters para series no estacionales

El método anterior se utiliza raramente en la práctica. En la literatura se han propuesto desarrollos más sofisticados del mismo que son utilizados principalmente a nivel microeconómico. De entre éstos nos fijaremos en los métodos Holt-Winters, que en su versión más sencilla consideran a la serie formada localmente por un nivel, una tendencia y una perturbación aleatoria, de acuerdo con el modelo

$$X_t = M_t + T_t + \varepsilon_t \quad (8)$$

en el que

$$M_t = aX_t + (1 - a)(M_{t-1} + T_{t-1}) \quad y$$

$$T_t = b(M_t - M_{t-1}) + (1 - b)T_{t-1}. \quad (9)$$

Para aplicar el método se toman como valores iniciales:

$$T_2 = X_2 - X_1$$

$$M_2 = X_2$$

y las predicciones se hacen de acuerdo con la fórmula:

$$F_{n,h} = M_n + h T_n,$$

donde vemos que a las predicciones futuras se les permite que tengan una evolución tendencial. El método es óptimo si la serie viene generada por un modelo IMA (2, 2). Este método no incorpora un componente estacional en el modelo para X_t por lo que no es adecuado para series que muestren variación estacional.

II.3. Método de Holt-Winters para series con estacionalidad multiplicativa

Una característica importante de gran número de series económicas consiste en que tienen un marcado comportamiento estacional. Para tener esto en cuenta la versión del método Holt-Winters más conocida propone el siguiente modelo:

$$X_t = (M_t + T_t) S_t + \varepsilon_t, \quad (11)$$

donde el factor estacional S_t tiene un efecto multiplicativo sobre el nivel y la tendencia. El modelo se complementa con:

$$M_t = a(X_t/S_{t-s}) + (1 - a)(M_{t-1} + T_{t-1}) \quad (12)$$

$$T_t = b(M_t - M_{t-1}) + (1 - b)T_{t-1} \quad (13)$$

$$S_t = c(X_t/M_t) + (1 - c)S_{t-s}, \quad (14)$$

donde s es el retardo estacional que normalmente será de 4 en series trimestrales y de 12 en series mensuales.

Los valores iniciales para poder aplicar de forma recurrente (12), (13) y (14) son:

$$S_t = X_t / ((1/s) \sum_{j=1}^{j=s} X_j), \quad t = 1, 2, 3, \dots, s$$

$$M_s = \frac{1}{s} \sum_{j=1}^{j=s} X_j,$$

$$T_s = 0.$$

Las predicciones se hacen de acuerdo con las fórmulas:

$$\begin{aligned} f_{n,h} &= (M_n + h T_n) S_{n+h-s} & h &= 1, 2, \dots, s \\ &= (M_n + h T_n) S_{n+h-2s} & h &= s + 1, \dots, 2s \\ &\vdots & & \\ &\vdots & & \\ &\vdots & & \end{aligned} \quad (15)$$

En este caso $f_{n,h}$ no es un predictor lineal, por lo que no es posible encontrar un proceso estocástico lineal para el que dicho predictor sea óptimo. No obstante es uno de los métodos basados en el alisado exponencial más utilizado.

Existe también una versión del método con estacionalidad aditiva que se puede demostrar que es óptima para modelos del tipo:

$$\begin{aligned} (1 - L)^2 (1 - L^s) X_t &= (1 - \theta_1 L - \theta_2 L^2 - \theta_s L^s - \\ &- \theta_{s+1} L^{s+1} - \theta_{s+2} L^{s+2}) \epsilon_t \end{aligned} \quad (16)$$

II.4. Comentarios a los métodos basados en el alisado exponencial

Además de los métodos citados existen otros varios que se basan también en el alisamiento exponencial. (Véase Granger y Newbold (1977) Capítulo 5.) Entre ellos podemos citar: el método

de Harrison (1965), el de Harrison y Stevens (1971), que utiliza el análisis bayesiano para incorporar información sobre si el sistema se encuentra en situación normal o de cambio en el nivel o en la tendencia o sometido a una perturbación momentánea fuerte, el de Brown (1962), etc. La descripción de todos ellos alargaría excesivamente este documento y por otra parte, como veremos más adelante, el método de Holt-Winters descrito en el epígrafe anterior es el que parece dar mejores resultados en la predicción.

En las fórmulas para el cálculo del nivel, tendencia y factor estacional de los epígrafes anteriores aparecen las constantes a , b y c a las que se ha de asignar un valor para hacer operativas dichas fórmulas. Cuanto más bajas son esas constantes más memoria tiene el proceso para efectuar la estimación de M_t , T_t y S_t y menos pesa en ella las observaciones recientes. En consecuencia, en una serie muy aleatoria convendrá dar poco peso al presente eligiendo valores bajos para las constantes. Tenemos pues que los valores a asignar a a , b y c dependen de la naturaleza de la serie en cuestión. No obstante, Reid (1969) sugiere los valores $a=0,8$; $b=0,7$ y $c=0,9$ y apunta que en la práctica las predicciones no son muy sensibles a variaciones en valor absoluto menor que 0,1 sobre los valores óptimos de las constantes.

En cualquier caso el ideal es utilizar la muestra de observaciones disponibles para estimar a , b y c eligiendo, por ejemplo, para estos parámetros aquellos valores que minimizan la suma de los cuadrados de los errores en la predicción de un período hacia delante. Incluso con esta estimación los métodos de esta sección continúan siendo automáticos, aunque ahora el costo de aplicación es mayor.

Como hemos visto, los métodos basados en el alisamiento exponencial son muy sencillos y son particularmente interesantes —debido a su bajo costo de aplicación— para situaciones en las que el número de series a predecir es de varios cientos. Por ejemplo, al hacer previsiones sobre la coyuntura económica basada en la información de los indicadores económicos disponibles; al hacer previsiones sobre el nivel de existencias de un gran número de productos, etc. Por otra parte, el método es óptimo si la serie a la que se aplica viene generada por un proceso estocástico determinado. En la medida en que tal hipótesis sea aproximadamente válida las predicciones serán casi óptimas. Pero ciertamente la hipótesis puede distar mucho de la realidad de una serie concreta, en cuyo caso las predicciones serán normalmente malas, es decir, con grandes errores. Por ello al aplicar los métodos de esta sección es conveniente incorporarles un sistema que advierta para qué series se están realizando predicciones inaceptables al objeto de aplicar a dichas series métodos más sofisticados, como los que tratamos en las secciones siguientes.

III. EL METODO BOX-JENKINS

Los métodos anteriores son automáticos ya que cualquiera que sea la serie a predecir se la considera generada por el mismo modelo. Box y Jenkins (1970) proponen una *clase de modelos* y una *forma de proceder* para que los datos elijan el modelo adecuado para cada serie en cuestión. Esta metodología es mucho más flexible que las anteriores, pero en la estrategia a seguir en su aplicación el investigador tiene que interpretar los resultados de las etapas intermedias y tomar decisiones sobre la validez de diversos modelos al caso en estudio, para llegar a aquél que es más adecuado a los datos. La gran virtud del método Box-Jenkins es dicha flexibilidad, que a su vez ha sido la fuente principal de las críticas al método. Así se puede señalar que:

1. El procedimiento Box-Jenkins no es tan rápido como los basados en el alisado exponencial,
2. Necesita de personal con cierta especialización para su aplicación, y
3. Requiere series temporales bastante largas.

Con esto, el costo de las predicciones por este procedimiento es superior al de los métodos anteriores. La contrapartida está en que con él se obtienen, en general, predicciones mejores.

La clase de modelos que proponen Box-Jenkins son los denominados modelos «integrados, autorregresivos y de medias móviles» (ARIMA)⁴ de la forma:

$$\begin{aligned} (1 - \phi_1 L - \dots - \phi_p L^p) (1 - \phi'_1 L^s - \dots - \phi'_{p_1} L^{s p_1}) (1 - L^s)^{d_1} (1 - L)^d Z_t = \\ = (1 - \theta_1 L - \dots - \theta_q L^q) (1 - \theta'_1 L^s - \dots - \theta'_{q_1} L^{s q_1}) a_t, \end{aligned} \quad (17)$$

⁴ Una introducción genérica y simple a los modelos ARIMA se da en el Apéndice.

donde Z_t es la variable económica y a_t es una perturbación aleatoria del tipo ruido blanco, es decir, con una distribución normal con media cero y varianza σ^2 , y que es independiente de cualquier perturbación aleatoria correspondiente a otro momento de tiempo (t') distinto de t . El modelo puede incluir también una constante e incluso un polinomio de tiempo. La justificación de este tipo de modelos podría verse de la siguiente forma. Wold en 1938 demostró (véase por ejemplo Anderson (1971), sección 7.6.3.) que los procesos estocásticos estacionarios⁵ se pueden representar por medio de un proceso de medias móviles (MA) infinito. Ahora bien, un proceso de ARMA del tipo.

$$(1 - \phi_1 L - \dots - \phi_p L^p) W_t = (1 - \theta_1 L - \dots - \theta_q L^q) a_t, \quad (18)$$

puede aproximar la representación de Wold con un nivel de aproximación establecido de antemano dando a P y Q los valores necesarios. Por otra parte, la estacionalidad de las series económicas parece que puede captarse adecuadamente y consumiendo menos grados de libertad si la estructura AR y MA de (18) se descompone de forma multiplicativa como aparece en los dos primeros paréntesis de los términos de la izquierda y de la derecha de (17). Por último, la estructura de (18) supone que W_t viene generada por un proceso estacionario. Obviamente las series económicas no son, en general, estacionarias pero se observa que sucesivas diferenciaciones regulares y estacionales de una serie económica (Z_t) de la forma:

$$(1 - L^s)^{d_1} (1 - L)^d Z_t = W_t, \quad (19)$$

convierten a la serie en estacionaria. Con ello está justificado (17), que teóricamente se presenta como un esquema susceptible de aproximar con suficiente adecuación una amplia gama de procesos estocásticos. La práctica en la aplicación de estos modelos parece confirmar tales impresiones teóricas.

Establecida la clase de modelos ARIMA, en cada caso concreto en base a una serie histórica los datos han de:

1. *Identificar*, los valores de d , d_1 , p , p_1 , q y q_1 así como los parámetros ϕ , ϕ' , θ y θ' que son distintos de cero.
2. *Estimar*, por medio de la máxima verosimilitud el valor de los parámetros ϕ , ϕ' , θ y θ' identificados con valores distintos de cero y la varianza (σ^2) del elemento residual a_t .
3. Estimado el modelo *se examinarán los residuos* (es decir, los valores estimados para a_t) para comprobar que el modelo obtenido es adecuado. Si la comprobación resulta afirmativa se utilizará el modelo para predecir. Si resulta negativa se iniciará de

⁵ Sobre el concepto de estacionariedad véase el apéndice pág. 36.

nuevo el proceso de identificación, estimación y validación de los resultados.

III.1. El proceso de identificación y estimación

La identificación viene dirigida por lo que se ha denominado el «*principio de parsimonia*», consistente en obtener una formulación del modelo con el menor número posible de parámetros desconocidos que posteriormente habrá que estimar. Como dicen Box-Jenkins (1970), este principio no plantea un problema matemático, sino que es una cuestión de encontrar cómo tienden a comportarse las series económicas, probando las ideas existentes sobre series temporales y tratando de desarrollar aquellos conceptos que se muestran fructíferos.

Los instrumentos más utilizados en la etapa de identificación son la *función de autocorrelación* y la *función de autocorrelación parcial*. Si W_t es un proceso estacionario con media δ , la autocorrelación de orden k es la correlación entre W_t y W_{t+k} , es decir,

$$\rho_k = \frac{E\{(W_t - \delta)(W_{t+k} - \delta)\}}{[E\{(W_t - \delta)^2\}]^{1/2} [E\{(W_{t+k} - \delta)^2\}]^{1/2}} \quad (20)$$

donde E es el operador de valores esperados. Estos valores son desconocidos y en la práctica han de estimarse a partir de una muestra, W_1, W_2, \dots, W_N , de N observaciones. A los valores estimados ($\hat{\rho}_k$) de la función de autocorrelación se les denomina *correlograma*. Un estimador adecuado de ρ_k es:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0},$$

donde:

$$\hat{\gamma}_k = \frac{1}{N} \sum_{t=1}^{N-k} (W_t - \bar{W})(W_{t+k} - \bar{W}), \quad k=0,1,2,\dots$$

y \bar{W} es la media aritmética de la muestra.

Un examen cuidadoso del correlograma nos conduce, generalmente, a particularizar en unos pocos modelos específicos (normalmente no más de dos o tres) dentro de la clase ARIMA, las estructuras aptas para explicar la serie en cuestión. Para ello se utilizan un conjunto de resultados cuya demostración puede encontrarse en Box-Jenkins (1970). Estos resultados se refieren normalmente a la función de autocorrelación teórica que es desconocida. El investigador dispone únicamente del correlograma estimado en

base a una muestra y por tanto sus valores están sometidos a errores muestrales. En consecuencia, es importante disponer, además del correlograma, de la estimación de las desviaciones típicas de sus valores para ver en qué medida éstos son significativamente distintos de cero o no. Es decir, al considerar el correlograma estimado no hay que tomarlo al pie de la letra, sino interpretarlo teniendo en cuenta las bandas de confianza en las que, para un nivel de significación dado, se pueden mover sus valores.

Hemos mencionado antes que para gran parte de las series económicas la formulación general del modelo contenida en (18) puede particularizarse en modelos del tipo:

$$\begin{aligned} (1-\phi_1L-\dots-\phi_pL^p)(1-\phi'_1L^s-\dots-\phi'_{p_1}L^{sp_1})W_t = \\ = (1-\theta_1L-\dots-\theta_qL^q)(1-\theta'_1L^s-\dots-\theta'_{q_1}L^{sq_1})a_t. \end{aligned} \quad (21)$$

Esto supone que tanto la parte autorregresiva como la de medias móviles de (18) admite una descomposición (multiplicativa) en una *parte regular* y en una *parte estacional*. La primera engloba retardos de orden inferior al desfase estacional (s) y la segunda a retardos que son múltiplos de dicho desfase. Esta simplificación (en cuanto a grados de libertad) de (18) contenida en (21) parece revelarse como bastante adecuada en muchos casos, por lo que en el resto de esta sección nos limitaremos a ella. Sin embargo, conviene advertir que en determinados casos es posible que se necesiten formulaciones ARIMA más genéricas que (21).

La formulación (21) permite considerar el correlograma en dos partes. Una la compuesta por los retardos 1 a $s-1$ que denominaremos regular y otra la compuesta por los retardos múltiplos de s , que denominaremos estacional. Los resultados que mencionamos a continuación como guía en la fase de identificación, los referimos al correlograma regular, ignorando la parte estacional del modelo. Es decir, suponiendo el siguiente modelo:

$$(1-\phi_1L-\dots-\phi_pL^p)W_t = (1-\theta_1L-\dots-\theta_qL^q)a_t \quad (22)$$

en el que $p_1 = q_1 = 0$. En la práctica los resultados que siguen los aplicaremos primero a la parte regular del correlograma para identificar p y q y posteriormente realizaremos con la parte estacional una aplicación análoga.

La identificación de d y d_1

R.1) (Resultado 1). Si la serie Z_t no es estacionaria, en el sentido de que d y/o d_1 en (17) son distintos de cero, las autocorrelaciones no tenderán a cero incluso para retardos relativamente

altos del correlograma⁶. En consecuencia conviene examinar también los correlogramas de los distintos W_t obtenidos de acuerdo con (19) para los distintos pares de valores d y d_1 . Normalmente, la consideración de valores hasta un máximo de dos para d y de uno para d_1 es suficiente. A la vista de todos estos correlogramas, escogeremos para d y d_1 los valores más pequeños que conduzcan a una serie W_t que según su correlograma aparezca como estacionaria.

Los restantes resultados se refieren a series estacionarias.

La identificación de p y q

Para ello tendremos en cuenta los siguientes resultados:

R.2) Si una serie es ruido blanco, en nuestro ejemplo

$$W_t = a_t,$$

los valores de su correlograma son, por definición, cero.

R.3) Si $q = 0$ el proceso se denomina *autorregresivo* y el correlograma para todos sus retardos *toma valores no nulos que van decreciendo* de acuerdo con una función formada por exponenciales decrecientes y oscilaciones sinusoidales también decrecientes.

R.4) Si $p = 0$ el proceso se denomina de *medias móviles* y en nuestro ejemplo será *de orden q* . En tal caso la función de autocorrelación *toma valores cero para retardos mayores que q* , es decir para $k > q$.

Esta característica de que para un proceso MA(q) la función de autocorrelación es cero más allá de $k=q$, mientras que para un proceso AR(q) continúa siendo no nula pero decreciente, es de especial importancia para diferenciar ambos procesos.

R.5) Si el proceso es ARMA de orden (p, q), y $p > q$ la función de autocorrelación tiene la forma de exponenciales decrecientes y/o oscilaciones sinusoidales también decrecientes, pero si $q > p$ las autocorrelaciones a partir del retardo $q - p + 1$ siguen dichas pautas pero no lo hacen las correlaciones iniciales de cero a $q - p$. Esta última característica es particularmente útil para diferenciar un proceso AR de uno ARMA.

Tenemos pues que tras la identificación de d y d_1 el examen del correlograma de W_t nos indicará si tal serie viene generada por un proceso:

- I) Ruido blanco
- II) De medias móviles (MA)

⁶ Conviene advertir que esta característica se observará en el correlograma estimado, ya que sus valores teóricos tal como han sido definidos antes no existen en este caso.

- III) Autorregresivo (AR) o
- IV) Autorregresivo y de medias móviles (ARMA).

En los dos primeros casos la identificación ha concluido pues en I) $p = q = 0$ y en II) el valor de q se determina verificando que para retardos superiores a q y menores que s , los valores del correlograma no son significativamente distintos de cero.

Si el correlograma estimado indica que el proceso generado puede ser del tipo AR, sabemos que su función de autocorrelación teórica no tiene un punto de corte, más allá del cual valga cero, sino que colea hasta el infinito. En consecuencia, el correlograma es de poca utilidad para distinguir si el proceso es $AR(p)$ o $AR(p + 1)$. Para ello utilizamos (véase Box-Jenkins (1970), páginas 64 y siguientes) un artificio denominado *función de autocorrelación parcial*, que explota el hecho de que todo proceso $AR(p)$ puede describirse por medio de p funciones no nulas de los coeficientes de autocorrelación.

La autocorrelación parcial de orden k , ϕ_{kk} , en un proceso estacionario de W_t , es la correlación parcial entre W_t y W_{t-k} una vez eliminadas en ambas la influencia lineal de los retardos intermedios⁷. Dicha autocorrelación parcial viene dada por el conjunto de ecuaciones (Yule-Walker)

$$\varrho_j = \sum_{i=1}^k \Phi_{ki} \varrho_{j-i}, \quad j=1,2,\dots,k. \quad (22)$$

Para estimar ϕ_{kk} se pueden utilizar sistemas de ecuaciones del tipo (22) sustituyendo los coeficientes de autocorrelación ϱ_j por sus valores muestrales. No obstante, hay métodos más eficientes para estimar ϕ_{kk} (sobre ellos véase Box-Jenkins (1970)). Al igual que antes conviene disponer en la etapa de identificación de un estimador de la varianza de ϕ_{kk} que permita discernir si se puede o no considerar el valor estimado como distinto de cero.

Si un proceso es $AR(p)$, $\phi_{(p+1), (p+1)}$ es cero, así la función de autocorrelación parcial nos sirve para identificar p , contrastando si $\phi_{(p+1), (p+1)}$ no es significativamente distinto de cero⁸.

Identificados p y q se aplicarán los resultados R.2 a R.5 al correlograma estacional para identificar de forma análoga p_1 y q_1 .

Identificados d , d_1 , p , p_1 , q y q_1 podemos utilizar la muestra de W_t para, por medio del método de la máxima verosimilitud,

⁷ Obsérvese que en un proceso $AR(k)$: $\phi_{kk} = \Phi_k$.

⁸ Obsérvese que un proceso $MA(q)$ tiene una representación alternativa como un proceso AR infinito y, por tanto, su función de autocorrelación parcial no se corta en el retardo q sino que colea hasta el infinito.

estimar los coeficientes de (17). El método consiste en minimizar la suma de los cuadrados de los residuos, para lo que existen diferentes algoritmos. Los programas de estimación suelen calcular también las desviaciones típicas de los estimadores con las que se puede verificar si los coeficientes son o no significativamente distintos de cero.

Con frecuencia en el proceso de identificación no se concluye con un único conjunto de valores para d , d_1 , p , p_1 , q y q_1 sino posiblemente con dos o tres. En dichos casos estimaremos ambos modelos y si están anidados (es decir uno es una particularización del otro) por medio de un test del cociente de la máxima verosimilitud contrastaremos ambos modelos. Si los modelos no están anidados, construiremos otro más general que tenga a ambos como casos particulares y por medio del test del cociente de la máxima verosimilitud se podrá verificar el modelo general contra los dos modelos iniciales. Bajo el supuesto de que el modelo general es innecesario aceptaremos el particular que lo rechace más fuertemente.

III.2. La validación de los resultados

Estimado el modelo hay que comprobar que es adecuado. Dos tipos de comprobaciones son las más fructíferas.

1. Ajustar un modelo más general

Para ver si el modelo estimado es adecuado formularemos otro más general⁹, es decir con más parámetros, que pasaremos a estimar. Si los parámetros añadidos no se muestran significativos se acepta el modelo original y, en caso contrario, el nuevo modelo. Conviene advertir que no se deben añadir al mismo tiempo coeficientes extra en la parte autorregresiva y su correspondiente de medias móviles, ya que si el modelo original es cierto los nuevos estimadores tendrán grandes desviaciones típicas.

2. Comprobar que los residuos son aleatorios (ruido blanco)

Esto se hará examinando su correlograma y viendo si tal hipótesis parece cierta. Box y Pierce (1970) proponen el siguiente test:

$$Q = n \sum_{k=1}^M r_k^2(\hat{a}),$$

⁹ Esto se hará teniendo en cuenta las dudas surgidas en el proceso de identificación.

donde $r_k^2(\hat{a})$ es el coeficiente de autocorrelación para el retardo k obtenido con los residuos estimados y n el número de residuos. El valor de Q se compara con el tabulado para χ^2_{M-P-Q} ¹⁰ y se rechaza la hipótesis de residuos aleatorios si Q excede el valor de la tabla a un nivel de significación dado. Si se rechaza dicha hipótesis el modelo no es adecuado y necesita revisarse. Los correlogramas de los residuos son una buena ayuda para empezar de nuevo el proceso de identificación-estimación y validación.

III.3. La predicción

Una vez que se dispone de un modelo adecuado que se ha estimado a partir de una serie Z_1, Z_2, \dots, Z_n , podemos utilizarlo para predecir los valores futuros Z_{n+h} , $h = 1, \dots$. Se demuestra (véase Box-Jenkins (1970)) que la predicción óptima en el tiempo n de Z_{n+h} es simplemente su esperanza condicional en el tiempo n . Así las predicciones se harán con (17) sustituyendo t por $n+h$ y tomando expectativas condicionadas en el tiempo n . Tales expectativas para $Z_n, Z_{n-1}, \dots; a_n, a_{n-1}, \dots$ son sus valores conocidos en el tiempo n . Las expectativas de $a_{n+1}, a_{n+2}, \dots, a_{n+h-1}$ son obviamente cero y las de $Z_{n+1}, Z_{n+2}, \dots, Z_{n+h-1}$ son las predicciones que de dichas cantidades se hace en n . Con ello el proceso de generación de predicciones, dado el modelo (17) que ha sido estimado, es un proceso de paso a paso realizando por turno las predicciones para $h = 1, 2, 3, \dots$

¹⁰ $P = p + p_1$ y $Q = q + q_1$.

IV. EL METODO AUTORREGRESIVO POR ETAPAS

Como hemos dicho antes, la ventaja del método Box-Jenkins radica en ser un método que requiere de personal especializado que investigue las propiedades de la serie cuyos valores futuros se quiere predecir, para obtener de esta forma el modelo más adecuado para cada caso concreto. Esto implica que en el proceso de producción de predicciones se necesita la intervención humana en grado considerable. El ideal sería disponer de un método que permitiese a los datos buscar su propio modelo dentro de una clase amplia y que a la vez fuese automático. Con este fin varios autores (véase Coen et.al. (1969), Kendall (1973) y Newbold y Granger (1974)) han sugerido ajustar un modelo autorregresivo del tipo:

$$W_t = \sum_{j=1}^P \phi_j W_{t-j} + a_t \quad (23)$$

por medio del procedimiento de regresión por etapas.

El método debe aplicarse a series estacionarias por lo que normalmente W_t se define como las primeras diferencias de la serie original. En la aplicación del método se requiere fijar de antemano P . Para ello Granger y Newbold (1977) sugieren el valor de 13 para series sin variación estacional, series trimestrales y series mensuales cortas. Para series mensuales largas proponen el valor de 25. El método requiere también especificar qué procedimiento de regresión por etapas se va a utilizar. El procedimiento de Payne (1973) que parte de la especificación más general y va contrastando qué retardos se pueden excluir, al mismo tiempo que verifica si se deben incluir los excluidos previamente, parece el procedimiento más recomendable. En él habrá que especificar qué tipo de test se utilizará para excluir e incluir regresores y a qué nivel de significación se realizará.

Este método es similar al de Box-Jenkins pero para convertirlo en automático ha habido que limitar los modelos ARIMA considerados en la sección anterior a la subclase de modelos autorregresivos integrados, es decir, modelo ARIMA con la parte de medias móviles igual a cero. Además dado que el mismo conjunto de datos se someten a un número considerable de tests, los problemas del agotamiento de los datos («data mining») pueden ser particularmente importantes. Por otra parte, la hipótesis de que a_t en (23) es un ruido blanco es esencial para la validez asintótica de los test que se realizan. Por todo ello, el método de autorregresión por etapas es, teóricamente, considerablemente inferior al método Box-Jenkins, pero respecto a los métodos del alisado exponencial contiene una flexibilidad mayor.

V. LA PREDICCIÓN ECONOMETRICA

Los métodos de predicción expuestos hasta el momento no tienen en cuenta los resultados de la teoría económica a no ser para elegir las variables que hay que predecir. La predicción econométrica se diferencia de la predicción estadística de las secciones anteriores en que utiliza los resultados de la teoría económica para formular sus modelos. En el caso más simple estos modelos son uniecuacionales¹¹, en los que la única variable dependiente, llamémosle Y , viene determinada por un conjunto de variables explicativas, llamémosles X_1, X_2, \dots, X_k . A falta de mayor información teórica la relación funcional entre Y y las X 's se especifica como lineal. Esta formulación referida a un tiempo t es:

$$Y_t = \beta_1 X_{1t} + \dots + \beta_k X_{kt} + \varepsilon_t, \quad (24)$$

donde ε_t es una perturbación estocástica que sufre el sistema en el momento t y que recoge el efecto de todas las variables explicativas omitidas y que de momento supondremos que es del tipo ruido blanco. Si se dispone de series temporales sobre Y, X_1, \dots, X_k podemos estimar unos valores para los parámetros β_1, \dots, β_k , escogiendo aquellos valores que minimizan la suma de los cuadrados de las perturbaciones aleatorias del sistema. Estimado el sistema podemos utilizarlo para predecir. Así, si queremos predecir el valor de Y en el tiempo $n + h$ sustituiremos t por $n + h$ en (24), los coeficientes por sus valores estimados, $\hat{\beta}_j$, ε_{n+h} por su valor

¹¹ Como advertíamos en la introducción, los fenómenos económicos se determinan de forma simultánea por lo que su representación requiere modelos multiecuacionales. No obstante, para simplificar la exposición en este trabajo nos referiremos únicamente a modelos uniecuacionales.

esperado, que de acuerdo con la hipótesis de ruido blanco es cero, con lo que obtenemos:

$$\hat{Y}_{n+h} = \hat{\beta}_1 X_{1,n+h} + \dots + \hat{\beta}_k X_{k,n+h}. \quad (25)$$

Si los valores de las equis en $n + h$ fuesen conocidos la aplicación de (25) nos daría la predicción de Y_{n+h} . En general, los valores de las equis en $n + h$ son desconocidos y hay que sustituirlos por una predicción de los mismos, que se puede obtener aplicando algunos de los métodos de predicción anteriores.

Tenemos, pues, que en principio, el método se muestra muy poderoso para predecir, ya que tiene en cuenta la relación estructural existente entre Y y las variables explicativas. Sin embargo, en su aplicación con frecuencia los modelos ARIMA dan predicciones mejores que las basadas en modelos de regresión del tipo (24). Esto sin embargo tiene su explicación.

En efecto, con frecuencia la teoría económica no es tan explícita como para decirnos cuáles son las k variables que determinan Y y es el investigador empírico quien hace la elección subjetiva de dichas variables, muchas veces tras experimentar con diversos conjuntos de k regresores. Si se hace esta experimentación se suele utilizar con excesiva frecuencia el cuadrado del coeficiente de correlación múltiple (coeficiente de determinación R^2) como criterio para preferir un conjunto de regresores sobre otro. Esto, sin embargo, puede dar lugar a regresiones espúreas. Así, Granger y Newbold (1977) presentan evidencia estadística, en base a series generadas artificialmente, de que si las series Y y equis vienen generadas por ciertos procesos estocásticos como senderos aleatorios, IMA (1, 1), etc., se puede obtener R^2 elevados al estimar modelos como (24) con variables equis que no tienen ninguna correlación teórica con Y . El trabajo de Granger y Newbold (1977) es una seria advertencia sobre las formas de validar un modelo de regresión. Una de las hipótesis de (24) es que las perturbaciones del modelo (ε_t) son estacionarias y además del tipo ruido blanco, así si dicha hipótesis no se cumple los resultados de una estimación mínimo cuadrático ordinaria (MCO) no son válidos. En consecuencia, para contrastar si los datos confirman la mencionada hipótesis es indispensable examinar los residuos y verificar que son estacionarios del tipo ruido blanco. La necesidad de estas contrastaciones ha sido desde hace mucho reconocida en la teoría econométrica, como ejemplo baste citar el trabajo de Sargan (1964). Sin embargo, en gran parte de los trabajos empíricos se ha hecho caso omiso de las recomendaciones teóricas.

Lo expuesto es razón suficiente para no aceptar los resultados de trabajos empíricos sobre modelos de regresión que no vayan

acompañados de un examen de los residuos que los contraste y valide.

No obstante, los problemas de especificar un modelo de regresión a partir de la teoría económica son mayores todavía, en el sentido que las variables explicativas no actúan sobre Y de forma inmediata sino distribuida en el tiempo. Asimismo los impactos de las perturbaciones que sufre el modelo tienen efecto durante más de un período. Sin embargo, la teoría económica dice muy poco sobre cómo especificar dichas estructuras dinámicas de la parte sistemática y de los errores del modelo. En consecuencia, el investigador tiene que utilizar la información contenida en los datos para realizar la especificación dinámica. En la práctica actual se siguen principalmente dos caminos para resolver, en función de la información muestral, el problema de la identificación de las estructuras dinámicas. Así, en el campo de Análisis de Series Temporales los modelos ARIMA se han generalizado dando lugar a lo que se ha denominado modelos ARIMA con funciones de transferencias (modelos ARIMAX) del tipo¹²:

$$Y_t = \delta_1^{-1}(L) \omega_1(L) X_{1,t-b_1} + \dots + \delta_k^{-1}(L) \omega_k(L) X_{k,t-b_k} + \phi^{-1}(L) \theta(L) a_t, \quad (26)$$

donde $\delta_1(L)$..., $\delta_k(L)$, $\omega_1(L)$..., $\omega_k(L)$, $\phi(L)$ y $\theta(L)$ son polinomios de L. La identificación de (26), al igual que en la sección III, se realiza mediante un proceso de experimentación basado en el examen de los residuos y a partir de una identificación inicial en base a las funciones de correlación cruzada. En estos modelos la identificación de las funciones de transferencia de los inputs y la de las perturbaciones se suele hacer en etapas separadas, es decir, ignorando una posible interrelación entre ellas. Es posteriormente en la etapa de estimación cuando se realiza la estimación completa del modelo. No obstante esta etapa de identificación no suele ser muy útil cuando la influencia de los inputs es principalmente en la tendencia o componente permanente del output.

En el departamento de Econometría de la London School of Economics, se está desarrollando otro acercamiento¹³ al problema

¹² Estos son modelos econométricos que permiten gran flexibilidad en la relación dinámica entre las variables. En la literatura econométrica dichos modelos son conocidos por el nombre de modelos con retardos racionales distribuidos.

¹³ Una exposición sobre este acercamiento al tema puede verse en Espasa (1976).

de identificación de las estructuras dinámicas que considera modelos del tipo:

$$\delta(L) Y_t = \omega_1(L) X_{1t} + \dots + \omega_k(L) X_{kt} + a_t^{14} \quad (27)$$

que utilizando la notación matricial,

$$X = (\underline{y} \ \underline{x}_1 \ \dots \ \underline{x}_k), \text{ (matriz)}^{15} \quad (28)$$

$$\underline{\phi}'(L) = [-\delta(L) \ \omega_1(L) \ \dots \ \omega_k(L)], \text{ (vector fila)}, \quad (29)$$

se puede expresar así:

$$X \underline{\phi}^{(m)}(L) + \underline{a} = 0, \quad (30)$$

donde X, definida en (28), es la matriz de observaciones de las variables del modelo; $\underline{\phi}^{(m)}(L)$ es el vector de parámetros (obsérvese que cada componente del vector es un polinomio en L) tal como se define en (29) y el superíndice m indica que los polinomios de los elementos del vector son de orden m; \underline{a} es el vector de perturbaciones aleatorias.

El modelo (30) es la generalización de

$$X \underline{a}^{(l)}(L) + \underline{y} = 0 \quad (31)$$

donde

$$\begin{aligned} \varrho^{(r)}(L) \underline{y} &= \underline{a}, \text{ y} \\ r + l &= m. \end{aligned} \quad (32)$$

En efecto, el modelo formado por (31) y (32) supone que el vector polinomial $\underline{\phi}^{(m)}(L)$ de (30) se puede descomponer de la forma:

$$\underline{\phi}^{(m)}(L) = \varrho^{(r)}(L) \underline{a}^{(l)}(L)^{16}, \quad (33)$$

en donde $\varrho^{(r)}(L)$ representa la estructura dinámica de los errores (EDE) y $\underline{a}^{(l)}(L)$ la estructura dinámica sistemática (EDS). El proceso de identificación se hace en dos etapas: primero se identi-

¹⁴ Este modelo es un caso particular de (26), en el que se aproxima $\Phi^{-1}(L) \Theta(L)$ por $\delta(L)^{-1}$ y en donde $\delta_1(L) = \delta_2(L) = \dots = \delta_k(L) = \delta(L)$.

¹⁵ Es una matriz $(k + 1) \times T$, es decir, es una matriz que recoge T observaciones sobre cada una de las $k + 1$ variables que entran en el modelo.

¹⁶ Obsérvese que $\underline{\phi}^{(m)}(L)$ y $\underline{a}^{(l)}(L)$ son vectores de $(k + 1)$ polinomios de órdenes m y l, respectivamente, mientras que $\varrho^{(r)}(L)$ es un escalar formado por un polinomio de orden r.

fica m y condicional al valor identificado para m se identifica r y t . Sobre los tests empleados y la forma de proceder véase Espasa (1976) y la bibliografía allí citada.

La conclusión de todo esto es la siguiente. La especificación dinámica de los modelos econométricos (EDS y EDE) es muy compleja y en gran parte de los trabajos empíricos no ha sido considerada adecuadamente y en nuestra opinión ésta es la razón principal por lo que con frecuencia el método Box-Jenkins de la sección III produce predicciones mejores que los modelos econométricos. En efecto, se puede demostrar (véase Zellner y Palm (1974)) que si las variables equis de (26) vienen generadas por procesos estocásticos ARIMA el modelo (18) es una forma final de (26). Es decir, los modelos de la sección III dan gran consideración a la estructura dinámica del modelo, que aunque de forma excesivamente genérica —modelo (18)— respecto modelos más específicos como (26) va en la línea adecuada de la especificación de modelo. Sin embargo en muchas aplicaciones los modelos econométricos no han considerado —a menudo debido al coste que ello supone— correctamente las EDS y EDE con lo que se ha concluido con modelos inadecuados para representar el mundo real y por tanto para predecirlo. En la medida en que los modelos econométricos se especifiquen y estimen teniendo en cuenta estructuras del tipo (26) y (30) es de esperar que la predicción econométrica dé mejores resultados, que la simple predicción univariante de las secciones anteriores.

Pero como decíamos en la introducción, el trabajar con modelos del tipo (18) o del tipo (26) o con la generalización multiecuacional de ambas, para obtener predicciones es algo que ha de decidirse en función de los distintos costes de aplicación y de los beneficios esperados con los diferentes modelos. En nuestra opinión los modelos ARIMA son útiles como paso inicial y previo para llegar luego a modelos econométricos. Estos últimos son especialmente importantes en el campo macroeconómico, donde no sólo interesa el predecir, sino también el disponer de un modelo estructural. En el campo microeconómico se pueden dar más a menudo situaciones en las que los métodos de las secciones II a IV pueden ser suficientes para las necesidades que se pretende cubrir. Señalemos, también, que los métodos de las secciones II a IV se pueden combinar entre sí (véase Granger y Newbold (1977) capítulo 8) o con los de ésta, con lo que con frecuencia se obtienen resultados mucho mejores que los proporcionados por cualquiera de los métodos individuales. En particular, Granger y Newbold (1977) presentan evidencia de que la combinación del método Holt-Winters con el de la autorregresión por etapas dan, en general, resultados de una calidad similar a los del método Box-Jenkins (sección III).

APENDICE

NOTAS SOBRE LOS MODELOS ARIMA

Estas notas van orientadas a facilitar al lector no familiarizado con el tema la lectura de la sección III. Con la finalidad de ganar en claridad se ha sacrificado grandemente la rigurosidad en la exposición.

En Economía las condiciones bajo las que se observa un fenómeno económico en un momento dado son únicas, es decir, no se vuelven a repetir y no se pueden reproducir artificialmente. En general, cuando un fenómeno tiene tal comportamiento la forma de poder inferir en base a una muestra (serie temporal), sobre las características del mismo es suponiendo que dicho fenómeno viene generado por un proceso estocástico estacionario.

La característica peculiar de un *proceso estocástico* consiste en que las propiedades estadísticas de la variable aleatoria (fenómeno económico), llamémosle X , son específicas a cada momento del tiempo, es decir, X_1, X_2, \dots, X_T , son variables aleatorias con funciones de distribución distintas. El proceso estocástico viene caracterizado por las funciones de distribución de cada X_i y por funciones de distribución conjuntas de todos y cada uno de los conjuntos que se pueden formar con las X 's.

Obviamente con una única realización temporal no es posible inferir sobre las características del proceso estocástico tal como se ha definido arriba. Para que sea posible la inferencia estadística necesitamos imponer condiciones más restrictivas. Generalmente la hipótesis que se introduce es la de estacionariedad¹.

¹ Las hipótesis evolutivas consumen más grados de libertad, por lo que su aplicación es más difícil en Economía. Sobre hipótesis evolutivas véase: Priestley, «Evolutionary spectra and non-stationary processes». *Journal of the Royal Statistical Association*, B, 1965, págs. 228-40.

Así, un *proceso estocástico es estacionario*:

- si todas las X_i tienen la misma esperanza matemática (media)
- si todas las X_i tienen la misma varianza
- si la correlación entre X_{t_1} y X_{t_2} sólo dependen de la diferencia entre t_1 y t_2 y no de t_1 y t_2 .

Una clase de procesos estocásticos son los llamados *procesos autorregresivos (AR)*,

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + a_t \quad (1)$$

en los que X_t depende de p X 's anteriores y de un ruido blanco a_t . Si utilizamos el operador lineal L , que cumple la condición

$$L^j X_t = X_{t-j}, \quad (2)$$

(1) se puede expresar así:

$$(1 - \phi_1 L - \dots - \phi_p L^p) X_t = a_t. \quad (1')$$

Para que un proceso AR sea estacionario, las raíces del polinomio $\phi_p(L) = (1 - \phi_1 L - \dots - \phi_p L^p)$ han de estar fuera del círculo unitario.

Otro tipo de *procesos estocásticos* son los llamados de *medias móviles (MA)*

$$X = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} = (1 - \theta_1 L - \dots - \theta_q L^q) a_t. \quad (3)$$

El proceso (1'), que llamaremos AR(p), cuando es estacionario se puede representar también por

$$X_t = (1 - \phi_1 L - \dots - \phi_p L^p)^{-1} a_t = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j},$$

es decir AR(p) puede verse como un proceso MA infinito.

Del mismo modo si las raíces del polinomio $\theta_q(L) = (1 - \theta_1 L - \dots - \theta_q L^q)$ están fuera del círculo unitario se puede escribir (3) de la forma:

$$(1 - \theta_1 L - \dots - \theta_q L^q)^{-1} X_t = a_t \quad \text{ó}$$

$$X_t + \sum_{j=1}^{\infty} \pi_j X_{t-j} = a_t,$$

es decir, MA(q) puede verse como un proceso AR infinito. Es decir, el proceso es invertible.

Combinando (1) y (3) obtenemos un proceso autorregresivo y de medias móviles (ARMA).

$$(1-\phi_1 L-\dots-\phi_p L^p) X_t = (1-\theta_1 L-\dots-\theta_q L^q) a_t, \quad (4)$$

que si los polinomios $\Phi_p(L)$ y $\Theta_q(L)$ tienen las raíces fuera del círculo unitario, el proceso es estacionario e invertible.

Wold en 1938 demostró (ver Anderson (1971) sección 7.6.3) que los procesos estocásticos estacionarios se pueden representar por medio de un proceso MA infinito. Al tener infinitos parámetros la descomposición de Wold no es operativa. La ventaja de los procesos ARMA (p, q) representados en (4) es que tienen un número finito de parámetros y sin embargo tienen gran generalidad y sirven para representar la mayor parte de los procesos estacionarios encontrados en la práctica.

La mayoría de las series temporales en Economía no son estacionarias, pues los niveles de las series varían con el tiempo. Sin embargo, sucesivas diferenciaciones de las series económicas (Z_t)

$$(1-L)^d Z_t = W_t, \quad (5)$$

sí que suelen comportarse como estacionarias. Así la representación ARMA (p, q) es bastante adecuada para W_t y para Z_t tenemos el modelo

$$(1-\phi_1 L-\dots-\phi_p L^p) (1-L)^d Z_t = (1-\theta_1 L-\dots-\theta_q L^q) a_t, \quad (6)$$

que se denomina ARIMA (p, d, q), es decir, autorregresivo integrado de medias móviles. La denominación de integrado viene del hecho que Z_t se define como la suma de las W 's. Por ejemplo para $d = 1$

$$\begin{aligned} (1-L) Z_t &= W_t \\ Z_t - Z_{t-1} &= W_t \\ Z_t &= W_t + W_{t+1} + \dots \end{aligned}$$

Cuando la parte autorregresiva de un modelo ARIMA es cero, el modelo se denomina IMA (d, q).

Hasta ahora el modelo representado en (5) no es muy útil para el análisis de series económicas, ya que para series estacionales los valores de p y q tendrían que ser muy altos. Pasemos pues a considerar *modelos más adecuados para series estacionales*. La exposición de estos modelos está basada en la que se hace en Nelson (1973), capítulo tercero.

Primero conviene advertir que la variación estacional que se encuentra en las series económicas no tiene un carácter determinístico, sino que varía con el tiempo. Un proceso estocástico es puramente estacional si para cualquiera de sus variables aleatorias X_{t_1} y X_{t_2} la correlación entre ellas es distinta de cero solamente para valores de γ ($\gamma = t_1 - t_2$) que son múltiplos del retardo estacional (s). Este retardo es de 4 en las series trimestrales y de 12 en las mensuales. Todo esto en una serie mensual implica que las observaciones correspondientes a un mismo mes, e.g. enero, están interrelacionadas pero no lo están con las observaciones referentes a otros meses.

Al igual que antes, este tipo de procesos se puede representar con bastante generalidad mediante un modelo ARIMA estacional, de la forma

$$\begin{aligned} (1 - \phi'_1 L^s - \phi'_2 L^{2s} - \dots - \phi'_{p_1} L^{p_1 s}) (1 - L^s)^{d_1} Z_t = \\ = (1 - \theta'_1 L^s - \theta'_2 L^{2s} - \dots - \theta'_{q_1} L^{q_1 s}) \varepsilon_t. \end{aligned} \quad (7)$$

Procesos puramente estacionales no ocurren en la práctica por lo que Box-Jenkins (1970) proponen considerar que en (7), ε_t viene determinado por un proceso ARIMA normal del tipo

$$(1 - \phi_1 L - \dots - \phi_p L^p) (1 - L)^d \varepsilon_t = (1 - \theta_1 L - \dots - \theta_q L^q) a_t. \quad (8)$$

Combinando (7) y (8) obtenemos el modelo ARIMA estacional

$$\begin{aligned} (1 - \phi_1 L - \dots - \phi_p L^p) (1 - \phi'_1 L^s - \dots - \phi'_{p_1} L^{p_1 s}) (1 - L^s)^{d_1} (1 - L)^d Z_t = \\ = (1 - \theta_1 L - \dots - \theta_q L^q) (1 - \theta'_1 L^s - \dots - \theta'_{q_1} L^{q_1 s}) a_t. \end{aligned} \quad (9)$$

La estructura estacional en (9) entra de forma multiplicativa ya que la observación Z_t se obtiene filtrando primero el ruido blanco a_t con un filtro no estacional (8), y luego con un filtro estacional (7). Así Z_t puede considerarse generada de una de estas dos formas:

1. A través del modelo estacional (7) en el que la serie input no estacional es ε_t . La serie ε_t es una serie ajustada de estacionalidad correspondiente a Z_t .

2. A través del modelo no estacional

$$(1 - \phi_1 L - \dots - \phi_p L^p) (1 - L)^d Z_t = (1 - \theta_1 L - \dots - \theta_q L^q) v_t \quad (10)$$

en el que la serie input estacional es v_t generada por:

$$\begin{aligned} (1-\phi'_1 L^s - \dots - \phi'_{p_1} L^{p_1 s}) (1-L^s)^{d_1} v_t &= \\ = (1-\theta'_1 L^s - \dots - \theta'_{q_1} L^{q_1 s}) a_t. \end{aligned} \tag{11}$$

BIBLIOGRAFIA

- Anderson, T. W.** (1971): *The Statistical Analysis of Time Series*, Wiley, New York.
- Box, G. E. P. y G. M. Jenkins** (1970): *Time Series Analysis Forecasting and control*, Holden Day, San Francisco.
- Box, G. E. P. y D. A. Pierce** (1970): «Distribution of residual autocorrelation in autoregressive integrated moving average time series models», *Journal of the American Statistical Association*, V. 65, páginas 1509-1526.
- Brown, R. G.** (1962): *Smoothing, Forecasting and Prediction of Discrete Time Series*, Englewood Cliffs, New Jersey, Prentice Hall.
- Coen, P. G., E. D. Gomme y M. G. Kendall** (1969): «Lagged relationships in economic forecasting». *Journal of the Royal Statistical Association*, Serie A, vol. 132.
- Espasa, A.** (1976): *Estimación y selección de modelos econométricos dinámico*, Estudios Económicos, n.º 11, Banco de España. Servicio de Estudios.
- Granger, C. W. J.** (1966): «The typical spectral shape of an economic variable». *Econometrica*, vol. 34, págs. 150-161.
- Granger, C. W. J. y P. Newbold** (1977): *Forecasting Economic Series*. Academic Press, New York.
- Harrison, P. J.** (1965): «Short-Term sales forecasting», *Applies Statistics*, vol. 14, págs. 102-139.
- Harrison, P. J. y C. F. Stevens** (1971): «A Bayesian approach to Short-term forecasting» *Oper. Res. Q.*, vol. 32, págs. 341-362.
- Kendall, M. G.** (1973): *Time Series*, Charles Griffin, London.
- Nelson, C. R.** (1973): *Applied Time Series Analysis*, Holden Day.
- Newbold, P. y C. W. J. Granger** (1974): «Experience with forecasting univariate time series and the combination of forecasts», *J. Roy. Stat. Soc.*, A, vol. 137, págs. 131-146.

- Reid, D. J.** (1969): «A comparative study of time series prediction techniques on economic data», Tesis doctoral, Departamento de Matemáticas, Universidad de Nottingham.
- Sargan, J. D.** (1964): «Wages and Prices in the United Kingdom: A Study in Econometric Methodology», *Colston Papers*, 17, Butterworths Scientific Publications.
- Zellner, A. y F. Palm** (1974): «Time Series Analysis and Simultaneous Equation Econometric Models», *Journal of Econometrics*, vol. 2, páginas 17-54.