pág. 26) se representa gráficamente la función de autocorrelación que Cleveland y Tiao (1976) calculan para el estimador del elemento irregular por el procedimiento X-11, y se propone que, al aplicar el método X-11, se calcule el correlograma de la serie resultante para el componente irregular y se compare con el gráfico indicado, con el fin de determinar si el procedimiento X-11 es adecuado para la aplicación realizada.

4.4.5. Estimación a partir de muestras finitas (*)

Antes de pasar a discutir modelos ARIMA más generales que (4.4.2), es preciso adaptar los resultados de (4.4.6) y (4.4.7) a las condiciones en que normalmente se plantea el estudio, es decir, desconocimiento de los parámetros del proceso ARIMA generador de Y y una muestra finita del mismo.

Cuando los parámetros de (4.4.2) son desconocidos, se han de estimar de forma consistente a partir de los datos muestrales. Se puede demostrar que en este caso se obtienen estimadores consistentes de todos los demás parámetros en (4.4.3), (4.4.7) y (4.4.6), sin más que sustituir los valores desconocidos de $\phi_Y(L)$, $\theta_Y(L)$ y σ_a^2 por sus estimadores consistentes.

A partir de ahora se utilizará la siguiente terminología:

- X_t: representará cualquier señal de interés, T_t, S_t, I_t o cualquier agrupación de dos de ellas. Esta señal vendrá dada por un modelo ARIMA, como el modelo (4.4.9) para el componente estacional.
- X_t^*: es la aproximación de la señal mediante la aplicación de medias móviles simétricas infinitas (filtro) que se derivan de los parámetros teóricos del modelo ARIMA correspondiente a Y_t. Para el componente estacional, su estimador se ha definido en (4.4.6). Tal estimador, pues, supone dados los parámetros del filtro, así como una muestra infinita de Y. El modelo que sigue el estimador es distinto a la señal y en el caso del componente estacional se da en (4.4.10).
- $\hat{X}_t$: es un estimador de la aproximación X_t^*, obtenido sustituyendo los parámetros teóricos del filtro por los que se derivan de una estimación del modelo ARIMA de Y_t y aplicando el filtro a una muestra finita de Y_t[21].

[21] En la práctica el analista no conoce nunca ni la verdadera señal X_t ni su aproximación X_t^*, sólo el estimador $\hat{X}_t$. Por esa razón a partir de la sección séptima y en los demás capítulos del libro se suprimirá el símbolo $\wedge$ con el fin de simplificar la notación y puesto que no hay confusión posible.

En cuanto al hecho de sólo disponer de una muestra finita de Y hay que recordar que los filtros que definen a X_t^* son siempre convergentes; por tanto, y aunque teóricamente se necesite una muestra infinita, en la práctica una muestra suficientemente larga pero finita proporciona una buena aproximación.

Sin embargo, en los extremos de la muestra, y concretamente al final de la misma, no existe un número finito suficiente de observaciones para estimar las señales. En tal caso, centrando el análisis en lo que sucede al final de la muestra, la solución óptima consiste en aproximar los valores futuros desconocidos por predicciones eficientes. Dichas predicciones se obtienen a partir del modelo ARIMA de de la variable observable o, en el caso de que exista y proporcione mejores predicciones, de un modelo econométrico. Se denotará por $\hat{X}_t^{(t+k)}$ al estimador de X_t^* usando información hasta $t+k$[22].

4.4.6. Comentarios finales sobre la aplicación de los métodos de forma reducida

Cuando la serie observada viene generada por un modelo más general que (4.4.2), tal y como se comentó anteriormente no está garantizada la posibilidad de descomponerla de la forma (4.4.1) utilizando este tipo de enfoque.

Al presentar la segunda propiedad de los componentes no observables, se puso un ejemplo de cómo la presencia de un polinomio autorregresivo estacionario en $\phi_Y(L)$ podía inducir a una asignación arbitraria. Pero la imposibilidad de aplicar estos métodos también puede deberse a la presencia de determinadas raíces en $\theta_Y(L)$. Así, incluso en el modelo de las líneas aéreas

$$\Delta\Delta_{12} Y_t = (1-\theta_1 L)(1-\theta_{12} L^{12}) a_t$$

la descomposición sólo existe si $\theta_{12} \geqslant 0$[23]. Afortunadamente, el caso $\theta_{12} < 0$ es muy poco frecuente en series económicas, ya que implica un comportamiento estacional poco habitual.

Por último conviene advertir que la extracción de señales se puede ver muy modificada ante ciertas variaciones en el modelo del agregado. Así, a diferencia de lo que ocurre en la predicción, donde modelos distintos pueden dar predicciones numéricamente muy similares, al

[22] Por lo tanto, la diferencia entre $\hat{X}_t$ y $\hat{X}_t^{(t+k)}$ estriba en que en $\hat{X}_t$ todos los valores de la serie original que entran en su cómputo son valores observados, mientras en $\hat{X}_t^{(t+k)}$ se combinan valores observados y predicciones.

[23] Hillmer y Tiao (1982).

menos a corto plazo, en la extracción de señales modelos aparentemente similares pueden implicar filtros bastante diferentes.

Este punto ha sido puesto de manifiesto, entre otros trabajos, en Maravall (1986b) en relación con la serie mensual de exportaciones españolas para los años 1974 a 1983. Para dicha serie se comprueba que los modelos

$$(1-\phi_1 L)\Delta_{12} Y_t = \mu + (1-\theta_{12} L^{12}) a_t \qquad (4.4.11)$$

$$(1-\phi_1 L)\Delta_{12} Y_t = \mu + (1-\theta_1 L)(1-\theta_{12} L^{12}) a_t \qquad (4.4.12)$$

$$\Delta\Delta_{12} Y_t = (1-\theta_1 L)(1-\theta_{12} L^{12}) a_t \qquad (4.4.13)$$

proporcionan un ajuste y predicciones muy similares, pero las distintas estimaciones de la tendencia que de ellos se derivan difieren significativamente. Esto por otra parte tampoco es de extrañar: por lo dicho en la sección octava del capítulo dos, el lector entenderá muy bien que el modelo (4.4.13), ha de dar resultados diferentes a los obtenidos con (4.4.11) y (4.4.12).

En consecuencia, cuando la extracción de cualquier tipo de señal se vaya a hacer tomando el modelo univariante del agregado como punto de partida, es importante asegurar que ese modelo realmente está validado y, sobre todo, que representaciones alternativas igualmente compatibles con los datos muestrales conducen a estimaciones similares de la señal deseada.

4.5. Métodos basados en modelos estructurales

4.5.1. Fundamento de los métodos estructurales

Otro grupo de métodos de descomposición de una serie temporal en componentes no observables está constituido por los *métodos basados en modelos estructurales*. Estos métodos comparten con los del epígrafe anterior la consideración explícita de los procesos generadores de los componentes; por esa razón para ambos se utiliza la denominación de métodos basados en modelos, y esa característica los distingue claramente de los procedimientos empiricistas.

Lo que separa las dos clases de métodos basados en modelos es el papel que asignan a la información extramuestral a la hora de formular los modelos de los componentes.

En el caso de los métodos en forma reducida, se vio que giran alrededor del modelo del agregado; y a partir de éste se obtienen los modelos de los componentes imponiendo un conjunto mínimo de restricciones, las justas para que la descomposición $Y_t = T_t + S_t + I_t$ esté identificada.

Los métodos de forma estructural imponen más restricciones a priori sobre los modelos de los componentes, y de hecho en ellos los procesos generadores de los componentes están casi totalmente determinados sin necesidad de conocer el modelo ARIMA correspondiente a Y_t.

Así, en el epígrafe anterior la muestra observada del agregado desempeña un papel central, ya que con ella se estima el modelo ARIMA de Y_t y a partir de éste se determinan en gran medida las características de los modelos de los componentes; en cambio la información extramuestral se incorpora de la forma menos restrictiva posible. En el método que se verá en este epígrafe, por el contrario, los procesos generadores de los componentes están especificados en su casi totalidad antes de proceder a observar el agregado y estimar su correspondiente modelo ARIMA, con lo que la información muestral sólo se emplea para estimar un subconjunto reducido del total de parámetros que caracterizan a dichos procesos.

La implicación en términos del filtro a usar para estimar los distintos componentes de una serie temporal concreta es clara: si los procesos generadores de los componentes están casi totalmente determinados de antemano, también lo estarán los filtros a emplear. Ahora bien, esto no significa que estos métodos se puedan equiparar con los empiricistas, ya que:

1. Los filtros siguen siendo específicos para cada serie: se determinan a priori los modelos para los componentes teóricos más adecuados para la serie temporal concreta que se va a descomponer, que no tienen porqué coincidir con los que corresponden a otra serie distinta; y esta individualización también se traslada al filtro a emplear.
2. Puesto que los modelos que se suponen a priori para los componentes implican restringir el proceso generador de datos del agregado, se contrasta si estos modelos supuestos son adecuados comprobando si el modelo del agregado que se obtiene a partir de los datos observados satisface dichas restricciones.

4.5.2. Modelos ARIMA de los componentes no observables (UC-ARIMA): primeros desarrollos (**)

El especificar a priori los procesos generadores de los componentes es una idea relativamente antigua, que ha conocido un nuevo auge desde mediados de los años setenta. De esa época son lo que se dio en llamar *modelos UC-ARIMA* («Unobserved Components ARIMA», ARIMA de componentes no observables), que generaron una primera corriente de trabajos en esta línea, entre los que destacan Engle (1978) y Nerlove et al. (1979).

Por ejemplo, en el citado trabajo de Engle se consideran modelos estructurales de la forma[24]

$$\begin{aligned} Y_t &= T_t + S_t + I_t \\ (1-L)^{d_1}\phi_T(L)\,T_t &= \theta_T(L)\,b_t \\ (1-L^s)^{d_2}\phi_S(L^s)\,S_t &= \theta_S(L^s)\,c_t \\ I_t &= d_t \end{aligned} \tag{4.5.1}$$

donde b_t, c_t y d_t son ruidos blancos normales independientes entre sí. Dados estos modelos de los componentes no observables (es decir, el modelo estructural) entonces Y_t está generado por un modelo (la forma reducida implícita en ese modelo estructural) cuya expresión general es

$$\phi_Y(L)\,\Phi_Y(L^s)(1-L)^d(1-L^s)^D\,Y_t = \theta_Y(L)\,\Theta_Y(L^s)a_t \tag{4.5.2}$$

siendo los órdenes de los polinomios de (4.5.2) los siguientes:

- $\phi_Y(L)$: el de $\phi_T(L)$;
- $\Phi_Y(L^s)$: el de $\phi_S(L^s)$;
- $\theta_Y(L)$: su orden es igual a $\max(d_1+p_1+q_2s,\ d_2s+p_2s+q_1,\ d_1+d_2s+p_1+p_2s)$, siendo p_i el orden de $\phi_i(L)$ y q_i el de $\theta_i(L)$, $i=T, S$; $d_1(d_2)$ el número de diferencias del modelo de la tendencia (componente estacional); y s el período estacional (12 para series mensuales).
- $\Theta_Y(L^s)$: su orden es siempre cero;

y además $d=d_1$, $D=d_2$.

Como es natural la forma estructural (4.5.1) impone un determinado comportamiento para el agregado: dados los coeficientes de (4.5.1) es evidente que los polinomios que forman el modelo ARIMA del agregado tendrán que satisfacer ciertas restricciones.

Cuando se construya un modelo ARIMA para Y_t sin tener en cuenta la estructura (4.5.1) —es decir, cuando se estime la forma reducida sin restringir—, los resultados que se obtienen pueden o no ser muy similares a los que se derivan de imponer (4.5.1) —forma reducida restringida—. En caso afirmativo, la estructura supuesta no está en contradicción con los datos; en caso contrario, existe evidencia en contra de la estructura (4.5.1) como expresión del proceso generador de los valores observados, Y_t.

[24] Se ha cambiado su notación para adaptarla a la de este libro.

Esto último es cierto si la forma reducida no restringida ha sido suficientemente validada. Ahora bien, en tanto en cuanto se postula que Y_t admite una formulación en términos de sumas de componentes en la que cada componente tiene unas características específicas, en la validación del modelo de forma reducida no restringida se debe contemplar también el discriminar entre formulaciones ARIMA alternativas en función de las características de los componentes que de ellas se derivan. En este sentido la formulación de un modelo estructural puede ser un instrumento adicional muy útil en la etapa de validación de un modelo ARIMA. Procediendo de este modo se puede desarrollar un proceso iterativo entre la forma reducida no restringida y el modelo estructural.

En efecto, la proposición de un modelo estructural basado en las características que deben poseer los componentes de Y_t conducirá a un modelo ARIMA denominado de forma reducida restringida, al que denotaremos R^0. A partir de los datos, y empleando la metodología propuesta por Box y Jenkins, se puede construir un modelo ARIMA, forma reducida no restringida, N^0, y en la validación del modelo N^0 se contempla si las propiedades que se derivan para sus componentes, de acuerdo con las hipótesis utilizadas en la sección 4.4, son compatibles con las ideas a priori que el modelo estructural incorpora. De esta confrontación puede surgir la conveniencia de contrastar el modelo ARIMA no restringido, N^0, con otro alternativo, N^1, similar a él, pero que recoja mejor las propiedades estructurales de los componentes.

Al mismo tiempo, la confrontación anterior ha podido sugerir que alguna faceta del modelo estructural sea claramente inaceptable para la serie Y_t en cuestión, y que conviene modificar el modelo estructural, lo que conducirá a un nuevo modelo ARIMA restringido, R^1. El proceso se puede ir repitiendo hasta que se logre una convergencia entre N^h y R^h. Obsérvese que esta forma de integrar la construcción de modelos ARIMA —procedimiento de Box y Jenkins— con la formulación de un marco teórico sobre el que basar la extracción de señales, garantiza que la estimación de componentes que se obtenga por el procedimiento de la sección anterior y el de ésta será la misma.

Es importante destacar que no todo modelo ARIMA de la forma (4.5.2) admite una representación en términos de componentes no observables; en otras palabras, no todo modelo ARIMA es un UC-ARIMA. Por ejemplo, si se reexpresa (4.5.2) como

$$\phi^+(L)\,Y_t = \theta^+(L)\,a_t \tag{4.5.3}$$

donde $\phi^+(L)=\phi_Y(L)\Phi_Y(L^s)(1-L)^d(1-L^s)^D$, y $\theta^+(L)=\theta_Y(L)\Theta_Y(L^s)$, una condición necesaria para que (4.5.3) admita una representación de la forma (4.5.1) es que el orden de $\theta^+(L)$ sea mayor o igual que el de $\phi^+(L)$; de no ser así es imposible obtener el modelo del agregado como implicación de una estructura similar a (4.5.1). Es decir, en los modelos UC-ARIMA necesariamente se ha de cumplir que el orden del polinomio que resume la parte de medias móviles del modelo sea mayor o igual que el del polinomio autorregresivo.

El uso de representaciones de la forma (4.5.1) en el proceso de descomposición de una serie temporal recibió bastante atención en la literatura especializada durante los años setenta y principios de los ochenta. No obstante, la caracterización de los componentes implícita en (4.5.1) es errónea, tal y como demuestra Maravall (1987, 1988a).

Sin pérdida de generalidad supóngase que s es igual a 12 y d_2 es uno, la situación más corriente en el análisis aplicado. Si se admite un proceso generador para el componente estacional dado por

$$(1-L^{12})\,\phi_S(L^{12})\,S_t=\theta_S(L^{12})\,c_t, \tag{4.5.4}$$

se está diciendo que el componente estacional es un proceso no estacionario, ya que requiere el operador $(1\text{-}L^{12})=\Delta_{12}=\Delta U_{11}(L)$. La no estacionariedad se manifiesta en la presencia de ciclos que no se amortiguan de período 2, 2,4, 3, 4, 6, 12 y ∞ meses[25]. Centrando la atención en este último, es evidente que un ciclo de período infinito no es estacional sino que, por definición, es tendencial, y que está inducido por la raíz real y positiva contenida en Δ_{12}.

En consecuencia (4.5.4) no representa un estacional puro, sino que es la mezcla de un componente estacional y una variable que tiene tendencia. Por esa razón cuando en el capítulo 2 se presentó la diferencia estacional, se recalcó que dicha diferencia se descomponía en el producto de una diferencia regular y del operador sumas móviles $U_{11}(L)$, y lo estrictamente estacional correspondía sólo al efecto de este último.

En resumen, en la representación (4.5.1) existe un factor, la diferencia regular implícita en $(1-L^{12})$, que no debe aparecer en un verdadero componente estacional, sino que debe asignarse al proceso generador de la tendencia: compárese la ecuación que aparece en (4.5.1) con la expresión general del modelo del componente estacional usado en la sección anterior.

LLama la atención que este error se repitiera en la literatura especializada[26], cuando trabajos como Cleveland (1972) y Box et al.

[25] Véase el apéndice al capítulo 2 de este libro.

[26] Maravall (1987, pág. 23) señala que los siguientes trabajos, entre otros, cometen

(1978) utilizan el polinomio $U_{11}(L)$ para definir el modelo ARIMA del componente estacional. En todo caso, sirve para recalcar los factores que se reflejan en una diferencia estacional, y que, pese a lo que el nombre pueda hacer pensar, no son todos estrictamente estacionales.

La falta de un acuerdo generalizado en la forma de caracterizar a priori a los componentes ha provocado que la aplicación de planteamientos estructurales al problema de la descomposición haya permanecido, hasta hace poco, en un terreno puramente académico, con escasas incursiones en el análisis aplicado. Sin embargo, últimamente se ha desarrollado un enfoque distinto del planteamiento UC-ARIMA de los setenta, basado en los *modelos estructurales de series temporales*, que sí se ha revelado útil en la práctica[27].

4.5.3. Modelo estructural sobre los componentes de una serie temporal (*)

Este tipo de modelos aparece en Harvey (1981b, 1984), y exposiciones detalladas de su empleo en el problema de la descomposición se encuentran en Harvey y Todd (1983), Maravall (1985) y Fernández (1988). Las formas estructurales que se manejan son muy sencillas, lo que no impide que se puedan aplicar al análisis de un amplio conjunto de fenómenos económicos reales.

El *modelo estructural básico* está definido como:

$$\begin{array}{rlr} Y_t=T_t+S_t+I_t & & \\ \Delta T_t=\beta_{t-1}+f_t & f_t\sim \text{Niid}(0,\sigma_f^2) & \\ \Delta\beta_t=g_t & g_t\sim \text{Niid}(0,\sigma_g^2) & (4.5.5)\\ U_{11}(L)\,S_t=c_t & c_t\sim \text{Niid}(0,\sigma_c^2) & \\ I_t=d_t & d_t\sim \text{Niid}(0,\sigma_d^2) & \end{array}$$

ese error al especificar el componente estacional: Nerlove et al. (1979), Pierce (1978), Pagan (1975), Engle (1978), Cleveland y Tiao (1976), Granger (1978), Harvey (1981b), Ansley (1983), Gourieroux y Monfort (1983), Pierce et al. (1984), Hausman y Watson (1985) y Hylleberg (1986). Curiosamente Cleveland (1972) propone para la tendencia y el componente estacional los siguientes modelos

$$(1-L)^2\,T_t=(1+0{,}26L+0{,}3L^2-0{,}32L^3)\,b_t$$
$$U_{11}(L)\,S_t=(1-0{,}26L^{12})\,c_t,$$

y, sin embargo, en Cleveland y Tiao (1976) se comete el error de asignar una raiz real y positiva al componente estacional.

[27] Estos modelos estructurales también son, por definición, modelos UC-ARIMA, pero los autores vinculados a este enfoque prefieren utilizar la terminología de modelo estructural.

siendo todas las innovaciones independientes entre sí. Para simplificar la exposición se sigue suponiendo que el período estacional es igual a 12.

El modelo estructural (4.5.5) implica un comportamiento para los componentes caracterizado por:

1. La primera ecuación es la definición del agregado como la suma de los tres componentes no observables e independientes entre sí.

2. La segunda ecuación define la tendencia como un sendero aleatorio con deriva, pero esta deriva a su vez —tercera ecuación— es estocástica y sigue otro sendero aleatorio. Por sustitución recursiva, las segunda y tercera ecuaciones dan lugar al siguiente modelo para el componente tendencial:

$$\Delta^2 T_t = (1-\theta_1 L)\, b_t \qquad b_t \sim \text{Niid}(0,\ \sigma_b^2) \tag{4.5.6}$$

donde los parámetros de (4.5.6) están ligados con las varianzas de las innovaciones de las segunda y tercera ecuaciones de (4.5.5) mediante las fórmulas:

$$\sigma_f^2 = \theta_1 \sigma_b^2 \qquad \sigma_g^2 = (1-\theta_1)^2 \sigma_b^2.$$

3. El componente estacional es muy sencillo, ya que su suma móvil es directamente un proceso ruido blanco, lo mismo que el elemento irregular.

Resulta muy ilustrativo comparar el modelo estructural básico (4.5.5)-(4.5.6) con los supuestos que se hicieron para los modelos de los componentes al presentar los métodos de forma reducida. De esa comparación se desprende claramente que el modelo estructural precedente pertenece a la familia de modelos admisibles del epígrafe anterior, con el añadido de que se han incorporado más restricciones que las estrictamente necesarias para identificarlo a partir de un modelo ARIMA correspondiente al agregado.

Así, por ejemplo, en los métodos de forma reducida las únicas restricciones que se imponen sobre los polinomios de la parte de medias móviles de los procesos de la tendencia y del estacional son que sus órdenes sean menores o iguales a los órdenes de sus correspondientes polinomios autorregresivos. Sin embargo, en (4.5.6) el polinomio de medias móviles está más restringido, ya que su orden es igual a uno, y el modelo estructural del componente estacional incluye la restricción de una media móvil de orden cero frente al orden once que normalmente se permite en los procedimientos de forma reducida.

De hecho una característica notable de la estructura (4.5.5) es que está determinada en su casi totalidad con información a priori, salvo por las varianzas de las innovaciones: esas cuatro varianzas son los únicos parámetros desconocidos de toda la estructura, y el papel de la muestra del agregado es posibilitar su estimación.

El modelo estructural básico evidentemente impone un proceso generador de datos muy específico para Y_t: el modelo para dicho agregado ha de ser del tipo

$$\Delta\Delta_{12} Y_t = \theta_{13}(L)\, a_t \tag{4.5.7}$$

un ARIMA (0, 13, 13) con restricciones, pues sus catorce parámetros (los trece de $\theta_{13}(L)$ y σ_a^2) se obtienen a partir de los cuatro parámetros de (4.5.5): véase Maravall (1985).

4.5.4. La estimación de los componentes (*)

La estimación de los parámetros desconocidos se realiza a partir de la función de verosimilitud de (4.5.7) incluyendo las restricciones que se derivan de (4.5.5), procediendo bien en el dominio del tiempo usando el filtro de Kalman, bien en el dominio de las frecuencias: véase Harvey y Todd (1983) para los detalles. Una vez estimados todos los parámetros de (4.5.5)-(4.5.6)-(4.5.7), la estimación de los valores de los componentes correspondientes a la serie observada del agregado se realiza aplicando a ésta los filtros adecuados, de manera similar a como se argumentó para los métodos de forma reducida.

Estos filtros se derivan de (4.5.5) y (4.5.7) por lo que, y en la medida que (4.5.5) recoge las características de las señales —según abogan los defensores de este procedimiento—, el diseño de los filtros se ha realizado teniendo en cuenta las propiedades esenciales de las señales a estimar[28]. Además, y puesto que los filtros se construyen utilizando lo menos posible la información muestral sobre Y, no están distorsionados por los posibles errores que se puedan deslizar al especificar y estimar el modelo ARIMA del agregado a partir de una muestra del mismo.

La forma de proceder en el análisis aplicado está basada en la que se propuso en el subepígrafe 4.5.2: si al estimar un modelo ARIMA para el agregado la especificación resultante es similar al ARIMA restringido (4.5.7), entonces el modelo estructural (4.5.5) es una buena aproximación a la estructura subyacente en el agregado. Si la muestra sobre Y conduce a que existe más de un modelo ARIMA con pro-

[28] Véase Fernández (1990).

piedades predictivas similares se elegirá, si lo hubiere, aquél que esté próximo a (4.5.7). Si los datos no apuntan hacia ningún modelo ARIMA compatible con el modelo estructural básico, no resulta correcto imponer (4.5.5) como mecanismo generador de los valores de Y_t, y el analista debe estudiar si existe otro modelo estructural que sea razonable desde el punto de vista teórico y que sea compatible con los datos.

Vista la forma reducida restringida —(4.5.7)— implicada por la estructura (4.5.5), es de esperar que el modelo estructural sea de aplicación en muchas de las variables económicas, ya que, con las restricciones que se imponen sobre sus parámetros, (4.5.7) resulta muy similar al modelo de las líneas aéreas, que a su vez constituye, en la mayor parte de las veces, una buena aproximación al proceso generador de las series temporales económicas.

Por lo tanto, la descomposición implícita en el modelo estructural básico será similar a la del método X-11 cuando éste sea óptimo; también debe producir resultados similares a los métodos de forma reducida en aquellos casos en que realmente la descomposición está bien definida. De hecho Marawall (1985) demuestra que el método X-11 es óptimo para un modelo estructural básico caracterizado por[29]

$$\sigma_f^2 = 0{,}020 \operatorname{var}(\Delta\Delta_{12} Y_t)$$
$$\sigma_g^2 = 0{,}025 \operatorname{var}(\Delta\Delta_{12} Y_t)$$
$$\sigma_c^2 = 0{,}010 \operatorname{var}(\Delta\Delta_{12} Y_t)$$
$$\sigma_d^2 = 0{,}050 \operatorname{var}(\Delta\Delta_{12} Y_t).$$

La principal aportación de los métodos de este epígrafe es haber mostrado que esquemas sencillos, y al mismo tiempo sensatos, para los componentes son válidos para representar una gran parte de los fenómenos económicos que se observan en la práctica. Además, el hecho de que su forma reducida restringida sea muy parecida al modelo de las líneas aéreas permite soslayar el principal inconveniente de este tipo de métodos, a saber, la imposición a priori de patrones de comportamiento.

[29] Véase también Butter y Mourik (1990) para una comparación de los métodos X-11 y basado en modelos estructurales para series económicas reales.

4.6. Incertidumbre, actualización y variabilidad en los componentes de un fenómeno económico

4.6.1. Planteamiento del problema

Cuando se dispone de un modelo explícito para cada componente, como en los métodos basados en modelos, se pueden realizar inferencias sobre los resultados de la descomposición. En tales casos no sólo es posible estimar los valores puntuales de los componentes en un momento dado, sino también construir intervalos de confianza, aproximar su grado de impredictibilidad a distintos horizontes, etc.

Esta es otra de las ventajas de este tipo de métodos frente a los procedimientos empiricistas, ya que en éstos últimos se carece de una representación del proceso generador de los datos de las variables implicadas, y por lo tanto este tipo de análisis es impracticable.

El interés de realizar inferencias sobre los resultados no es meramente académico: en muchas ocasiones la toma de decisiones, al menos a corto plazo, se basa en la evolución de un componente no observable del indicador. El ejemplo más típico es el diseño de la política monetaria a corto plazo, que se suele realizar en función de la serie ajustada de estacionalidad del indicador de liquidez utilizado (en España, los activos líquidos en manos del público).

Ahora bien, en este caso, y puesto que la señal se aproxima con error, es muy importante acompañar su estimación puntual con un *intervalo de confianza*. Si se desconoce la variabilidad asociada al valor de la señal en un momento concreto, el proceso de toma de decisiones se realiza en parte a ciegas, ya que puede ocurrir que el valor puntual lleve a interpretar la situación presente del fenómeno de forma distinta a como se interpretaría de conocer el intervalo de confianza. Este es un punto que ha sido tratado extensamente en Moore et al. (1981), Pierce et al. (1981) y Maravall y Pierce (1986).

4.6.2. Errores en la estimación de los componentes (*)

En la estimación del valor concreto de un componente en un momento dado se pueden distinguir dos tipos de errores, que combinados forman el llamado *error total*:

1. *Error en el estimador final*: el componente de interés es una magnitud no observable, que se estima de forma óptima condicionándola a los valores pasados, presente y futuros del agregado observado. Este error es precisamente la diferencia entre el valor descono-

cido y su estimación óptima. Para el componente estacional en el momento t, S_t, el error de estimación final FS_t vendrá dado por

$$FS_t = S_t - \hat{E}(S_t/\ldots, Y_{t-1}, Y_t, Y_{t+1}, \ldots) = S_t - \hat{S}_t.$$

2 *Error de revisión*: cuando el último valor observado del agregado corresponde al momento t, los valores de los componentes para $t, t-1, t-2, \ldots, t-k$ se estiman provisionalmente sustituyendo valores desconocidos de la variable observable por predicciones. Así por ejemplo la estimación del componente estacional para t con información hasta $t+k$ vendrá dada por

$$\hat{S}_t^{(t+k)} = \hat{E}(S_t/\ldots, Y_t, \ldots, Y_{t+k}, \hat{Y}_{t+k+1}, \hat{Y}_{t+k+2}, \ldots) = \\ = E(S_t/\ldots, Y_t, \ldots, Y_{t+k});$$

El error de revisión se define como la diferencia entre el estimador óptimo (final) y el estimador provisional, y mide el coste de utilizar predicciones sobre el futuro cuando éste es desconocido.

Para simplificar, el análisis que sigue a continuación se centrará en el error de revisión del estimador concurrente, $RS_t^{(t)} = \hat{S}_t - \hat{S}_t^{(t)}$, pero, en general, un error de revisión vendrá dado por

$$RS_t^{(t+k)} = \hat{S}_t - \hat{S}_t^{(t+k)} = \hat{S}_t - \hat{E}(S_t/\ldots, Y_{t-1}, Y_t, Y_{t+1}, \ldots, Y_{t+k})$$

donde a medida que k aumenta el estimador provisional $\hat{S}_t^{(t+k)}$ es cada vez más parecido al estimador final, y por lo tanto los errores de revisión son cada vez menos importantes.

Tal y como han sido definidos, ambos tipos de errores siguen modelos ARIMA que se obtienen a partir de los modelos de las variables que entran en su definición, y el núcleo de este epígrafe está formado por el análisis de las propiedades estocásticas de estos errores.

Pierce (1980) demostró que el error de estimación final y el error de revisión son independientes, por lo que sus propiedades se pueden estudiar por separado. La presentación de estas propiedades, al igual que otros temas de este epígrafe, está basada en resultados de Maravall (1986a, 1987), y suponen que la serie ha sido descompuesta utilizando un método basado en modelos de forma reducida.

4.6.3. Error en el estimador final (*)

1. El error en el estimador final de cualquier componente es un proceso estocástico *estacionario*.

Por ejemplo, supóngase sin pérdida de generalidad[30] que se pretende estimar el componente estacional, para lo que se plantea la descomposición

$$Y_t = Y_t^a + S_t,$$

donde Y_t^a representa la serie ajustada de estacionalidad ($Y_t^a = T_t + I_t$), y los procesos generadores de los datos de las tres variables implicadas son

$$\begin{aligned} \phi_Y(L)\, Y_t &= \theta_Y(L)\, a_t & a_t &\sim \text{Niid}(0, \sigma_a^2) \\ \phi_S(L)\, S_t &= \theta_S(L)\, c_t & c_t &\sim \text{Niid}(0, \sigma_c^2) \\ \phi^*(L)\, Y_t^a &= \theta^*(L)\, f_t & f_t &\sim \text{Niid}(0, \sigma_f^2) \end{aligned}$$

Se demuestra que el error de estimación final del componente estacional, FS_t, viene generado por el modelo

$$\theta_Y(L)\, FS_t = \theta_S(L)\, \theta^*(L)\, \varepsilon_t \qquad (4.6.1)$$
$$\varepsilon_t \sim \text{Niid}(0, \sigma_\varepsilon^2 = \sigma_c^2 \sigma_f^2 / \sigma_a^2).$$

Como una de las restricciones iniciales del problema es que el modelo univariante de la serie observada sea invertible, es decir, las raíces de $\theta_Y(L)$ estén fuera del círculo unitario, (4.6.1) representa un proceso estocástico estacionario.

2. El error en el estimador final en general *no es ruido blanco*: esto es evidente a partir de (4.6.1), donde se comprueba que sigue un proceso ARMA de orden $(q_Y, q_S + q^*)$, siendo q_Y el orden del polinomio $\theta_Y(L)$, q_S el de $\theta_S(L)$ y q^* el de $\theta^*(L)$.

3. En el caso concreto de la descomposición en componente estacional y serie ajustada, las propiedades estocásticas de ambos errores de estimación final son exactamente *las mismas*. En efecto, de la doble igualdad

$$Y_t = Y_t^a + S_t = \hat{Y}_t^a + \hat{S}_t$$

es inmediato comprobar como el error en el estimador final en el momento t para cualquiera de las señales es igual al de la otra cambiando el signo. En consecuencia, (4.6.1) también representa el proceso generador correspondiente al error de estimación final de la serie ajustada, FY_t^a.

[30] El desarrollo que sigue está basado en descomponer la variable observada en dos componentes, una que es la señal de interés y otra el ruido añadido: véase Maravall (1987), pag 44.

4.6.4. Error de revisión[31] (*)

1. El error de revisión depende únicamente de las innovaciones futuras desconocidas de la serie agregada. Supóngase por ejemplo que se trata de estimar ese valor del componente estacional en el momento t; el estimador final es igual a

$$\hat{S}_t = \hat{\alpha}(L, F)\, Y_t = \sum_{j=-k}^{k} \hat{\alpha}_j Y_{t+j} = \sum_{j=0}^{k} \hat{\alpha}_{-j} Y_{t-j} + \sum_{j=1}^{k} \hat{\alpha}_j Y_{t+j}$$

siendo $2k+1$ la longitud del filtro; el estimador concurrente se obtiene como

$$\hat{S}_t^{(t)} = \sum_{j=0}^{k} \hat{\alpha}_{-j} Y_{t-j} + \sum_{j=1}^{k} \hat{\alpha}_j \hat{Y}_{t+j}$$

siendo $\hat{Y}_{t+j}$ la predicción óptima para $t+j$ con información hasta t. Por lo tanto el error de revisión $RS_t^{(t)}$ será igual a

$$RS_t^{(t)} = \hat{S}_t - \hat{S}_t^{(t)} = \sum_{j=1}^{k} \hat{\alpha}_j Y_{t+j} - \sum_{j=1}^{k} \hat{\alpha}_j \hat{Y}_{t+j} = \sum_{j=1}^{k} \hat{\alpha}_j (Y_{t+j} - \hat{Y}_{t+j}) =$$
$$= \sum_{j=1}^{k} \hat{\alpha}_j e_{t+j} = \sum_{j=1}^{k} \hat{\alpha}_j w_j(F)\, a_t = \hat{\alpha}^*(F)\, a_t \qquad (4.6.2)$$

donde $e_{t+j} = w_j(F)\, a_t$ es el error que se comete con la predicción del valor correspondiente a $t+j$ hecha con j períodos de antelación; puesto que los errores de predicción son una transformación lineal de las innovaciones posteriores a t[32], sigue de forma trivial que el error de revisión asociado al estimador concurrente es a su vez una transformación lineal de innovaciones estrictamente futuras de la variable observada[33].

2. La sucesión de errores de revisión del estimador concurrente $(\ldots, RS_{t-1}^{(t-1)}, RS_t^{(t)}, RS_{t+1}^{(t+1)}, \ldots)$, es un proceso estocástico *estacionario*: se demuestra que en la expresión (4.6.2) $\hat{\alpha}_j^*$ tiende a cero a medida que j tiende a infinito, y a partir de un determinado adelanto k las innovaciones futuras no afectan al error de revisión.

Una importante implicación de este resultado es que la señal y un

[31] Todo el análisis que sigue se limita al error de revisión del estimador concurrente. La generalización a otros errores de revisión es inmediata.

[32] Como se demostró en el capítulo 2, sección novena.

[33] Nótese que $\hat{\alpha}^*(F)$ es tal que el coeficiente correspondiente a F^0 es cero. Esta formulación es distinta a la de Maravall (1987), pág. 47.

estimador inicial suyo, el estimador concurrente, están ya *cointegrados*, pues siendo ambas normalmente no estacionarias su diferencia es un proceso $I(0)$.

3. Lo anterior no significa que el error de revisión del estimador concurrente sea un proceso ruido blanco; al contrario, *existe dependencia dinámica de corto plazo* motivada por el hecho de que las innovaciones correspondientes a distintos momentos del futuro entran en la determinación de su valor presente.

4. El estimador concurrente y su error de revisión son *ortogonales*, como se desprende de la propiedad primera de este error: el estimador está basado en la historia temporal del agregado que se observa hasta t, mientras que el error de revisión sólo depende de las sorpresas posteriores a t.

Por esa razón no se puede utilizar la dependencia dinámica implícita en el proceso generador de datos del error de revisión para mejorar el estimador concurrente.

5. Sean $\hat{S}_t^{(t)}$ y $\hat{S}_t^{(t+1)}$ los estimadores del valor del componente estacional referido a t con información hasta t y $t+1$ respectivamente. De la primera propiedad es inmediato que

$$\hat{S}_t^{(t+1)} - \hat{S}_t^{(t)} = \hat{\alpha}_1 (Y_{t+1} - \hat{Y}_{t+1}) = \hat{\alpha}_1^* a_{t+1}$$

y en general

$$\hat{S}_t^{(t+k)} - \hat{S}_t^{(t)} = \sum_{j=1}^{k} \hat{\alpha}_j (Y_{t+j} - \hat{Y}_{t+j}) = \sum_{j=1}^{k} \hat{\alpha}_j^* a_{t+j}.$$

Por lo tanto la variación en la estimación del valor de la señal en un momento t cuando se dispone de k nuevas observaciones sigue una media *móvil de orden* $k-1$, y en consecuencia es estacionaria. Como

$$RS_t^{(t+k)} - RS_t^{(t)} = \hat{S}_t^{(t)} - \hat{S}_t^{(t+k)},$$

la misma conclusión es válida para los sucesivos errores de revisión que se obtienen al reestimar un valor concreto de la señal a medida que van llegando nuevos valores del agregado.

6. Una de las implicaciones del requisito canónico es la de maximizar las varianzas de los errores de revisión (Maravall, 1988b).

Con la adopción del requisito canónico está implícito el deseo de tener una señal que evolucione de la forma más suave posible: si las variaciones en el tiempo del componente buscado han de ser lo más pequeñas posibles, cualquier nuevo valor de Y exige modificar adecuadamente toda la historia más reciente de las señales sistemáticas para que ese valor sea concordante con ellas. Si por el contrario se admitiese una señal muy irregular que mostrase oscilaciones bruscas, una nueva estimación de la señal para $t+1$ no exigiría tantas modificaciones en las estimaciones correspondientes a t, $t-1$, etc.

4.6.5. Un ejemplo de estimación de los componentes y de las varianzas de los errores de estimación

Un ejemplo del tipo de análisis que se puede realizar utilizando los resultados de 4.4 y de los subepígrafes anteriores se encuentra en Maravall (1987): para la serie mensual de activos líquidos en manos del público, con la muestra 1978 a 1985, se estima el siguiente modelo:

$$\Delta\Delta_{12} Y_t = (1+0{,}19L)(1-0{,}62L^{12})\,a_t$$
$$\sigma_a^2 = 0{,}138 \cdot 10^{-4}$$

Aplicando el método basado en modelos de forma reducida, los modelos teóricos estimados para los componentes resultan ser

$$\Delta^2 T_t = (1+L)(1-0{,}961L)\,b_t \qquad \sigma_b^2 = 0{,}323 \cdot 10^{-5}$$
$$U_{11}(L)\,S_t = (1+2{,}019L+2{,}487L^2+2{,}619L^3+2{,}481L^4+$$
$$+2{,}182L^5+1{,}800L^6+1{,}365L^7+0{,}972L^8+0{,}568L^9+$$
$$+0{,}310L^{10}+0{,}031L^{11})\,c_t \qquad \sigma_c^2 = 0{,}731 \cdot 10^{-6}$$
$$I_t = d_t \qquad \sigma_d^2 = 0{,}149 \cdot 10^{-5}$$

El primer resultado de interés es el que se refiere a la aleatoriedad de cada componente: para la serie original, la desviación típica de la innovación es igual a 0,37%, cifra en que se puede cuantificar la impredictibilidad del agregado; para la tendencia, la misma medida es de apenas 0,18%, del 0,09% para el componente estacional y del 0,12% para el irregular.

Aplicando el mismo análisis para obtener la serie desestacionalizada, su proceso generador de datos se estima como:

$$\Delta^2 Y_t^a = (1+0{,}182L)(1-0{,}961L)\,f_t \qquad \sigma_f^2 = 0{,}925 \cdot 10^{-5},$$

con una desviación típica de la innovación igual al 0,30% del nivel de la serie. Compruébese que esta señal es más errática que la tendencia, cuya desviación estándar innovacional es, necesariamente, menor (0,18%).

Calculando los estadísticos referidos a los errores (total, en el estimador final y de revisión) para la tendencia y la serie desestacionalizada, se tiene que:

1. El error total en el estimador del valor de la señal en t usando información hasta t tiene una varianza de $0{,}618 \cdot 10^{-5}$ (desviación típica de 0,25%) para la tendencia y de $0{,}526 \cdot 10^{-5}$ (0,23%) para la serie desestacionalizada: como consecuencia de ser la tendencia una señal más firme (menos oscilante) que la serie ajustada, tiene siempre, necesariamente, mayor error total en su estimación (en este caso 0,25% frente a 0,23%). No obstante, en este ejemplo al menos, la mayor imprecisión en la estimación de la tendencia no elimina la ventaja de ser una señal menos oscilante.

Estas cifras permiten construir intervalos de confianza para dichas señales, sin las cuales no se puede instrumentar el control monetario. Así, si para el momento t se estima un valor para la tendencia igual a $\hat{T}_t^{(t)}$, hay una confianza del 95% de que el verdadero valor de la señal esté en el intervalo $\hat{T}_t^{(t)} \pm 0{,}0049\hat{T}_t^{(t)}$.

Nótese que los valores son relativamente importantes, ya que las desviaciones típicas de los errores totales de las estimaciones de estas dos señales oscilan entre el 60% y el 70% de la desviación típica de la innovación del agregado.

2. La varianza del error en el estimador final supone algo menos de la mitad de la varianza del error total: para la tendencia es igual a $0{,}299 \cdot 10^{-5}$ (0,17%) y para la serie desestacionalizada a $0{,}254 \cdot 10^{-5}$ (0,16%).

Estas varianzas permiten construir intervalos de confianza para los verdaderos valores de las señales una vez se disponen de suficientes observaciones hacia el pasado y hacia el futuro.

3. Los errores de revisión de los estimadores concurrentes tienen varianzas de $0{,}319 \cdot 10^{-5}$ (0,18%) para la tendencia y $0{,}272 \cdot 10^{-5}$ (0,16%) para la serie desestacionalizada.

Estas cifras son de destacar, ya que la incertidumbre debida al desconocimiento de los valores futuros supone más de la mitad de la varianza del error total.

4. Por lo que se refiere a los errores de revisión con información hasta $t+k$, se observa que cuando se estima el valor de la señal correspondiente a t con información hasta $t+12$, la varianza del error de revisión cae hasta $0{,}842 \cdot 10^{-6}$ (cuya raiz cuadrada —desviación típica— es igual a 0,09%) para la tendencia y a $0{,}106 \cdot 10^{-5}$ (0,09%) para la serie ajustada.

De aquí se extraen las siguientes conclusiones:

4.a) los errores de revisión de los estimadores que se obtienen cuando se dispone de un año más de datos tienen una varianza que supone el 26% de la varianza del error de revisión del estimador concurrente para la tendencia, y el 39% para la serie desestacionalizada;

4.b) la varianza del error de revisión decrece más lentamente para la serie desestacionalizada que para la tendencia, con lo que la sucesión de estimadores de ésta se acerca más rápidamente al estimador final.

5. Al conocer las varianzas y los correlogramas de los distintos tipos de errores se pueden calcular las propiedades estocásticas, y en especial cuantificar el grado de precisión, de las distintas tasas de crecimiento que se estudiarán en el próximo capítulo.

4.7. La evolución subyacente de un fenómeno económico: ¿tendencia o serie desestacionalizada?[34]

Por evolución subyacente o nivel subyacente se ha definido la trayectoria de avance firme y suave de una serie, una vez que a los datos originales se les ha extraído aquellas oscilaciones que dificultan el seguimiento del fenómeno de interés.

Esta trayectoria es la realmente importante para evaluar la evolución del fenómeno, ya que éste oscila alrededor de ella, de forma que las desviaciones sobre la misma se compensan. Precisamente por ello en el nivel subyacente se pueden detectar ciertas peculiaridades básicas del fenómeno, que en cambio pueden ser difícilmente perceptibles en la serie original.

La necesidad de disponer de una señal de este tipo, en la que basar el análisis de corto plazo, ha sido unánimemente reconocida desde hace mucho tiempo. En la mayor parte de los países occidentales, las oficinas de estadística y los organismos oficiales hacen un uso exhaustivo de series ajustadas de estacionalidad, y los analistas económicos suelen utilizar tales series ajustadas como representativas de la evolución subyacente de las correspondientes variables.

La serie desestacionalizada muestra una evolución más suave que la serie original, ya que al suprimir el componente estacional se elimina, en general, la mayor causa de las oscilaciones a corto plazo del fenómeno observado. Por lo tanto la serie desestacionalizada realmente es un candidato válido para aproximar la evolución subyacente.

Sin embargo, dicha serie a su vez es el resultado de combinar dos

[34] Recuérdese que a partir de este epígrafe se emplea la notación T_t, S_t, I_t y Y_t^a para los estimadores de las correspondientes señales, suprimiéndose por tanto el símbolo $\wedge$

componentes: uno todavía más suave, la tendencia, y otro que recoge oscilaciones irregulares.

En ese sentido, la serie desestacionalizada es todavía una serie contaminada, al menos comparada con la tendencia: ésta es una señal pura, ya que no es posible una descomposición adicional de ella, y en consecuencia está más próxima a lo que se entiende por evolución subyacente.

Por eso, aun admitiendo que la serie ajustada de estacionalidad sea una señal válida para aproximar la evolución subyacente, en este libro se defiende la idea de que tal evolución subyacente se aproxima mejor utilizando la tendencia de la serie observada. Si se acepta este resultado, los análisis de corto plazo se deberían basar en esta señal y no, como es la práctica más extendida en la actualidad, en la serie desestacionalizada.

Hay que tener en cuenta que una de las principales causas del componente irregular son los errores de medida. Por sí sola, ésta es ya una razón para preferir la tendencia, pues cuanto más se limpia una señal de otro tipo de oscilaciones, mayor es el peso de estos errores en su evolución. Así, cuando se tiene una señal, como la serie desestacionalizada, donde los movimientos estacionales han sido eliminados, las oscilaciones inducidas por los errores de medida pueden ser notorias en la evolución de dicha señal.

No obstante, es evidente que la serie desestacionalizada sigue siendo la señal preferida por los analistas de la coyuntura para realizar el seguimiento a corto plazo de los fenómenos económicos. A su favor se suelen dar los argumentos que se exponen a continuacion.

1. Al menos en determinadas variables, el componente irregular no está siempre dominado por oscilaciones sin interés económico y resulta ser importante en sí mismo. Si en tales casos se utiliza la tendencia no se incorpora la información que el componente irregular aporta.

Esta afirmación es correcta, pero no representa un argumento en favor de la serie desestacionalizada. Para centrar la discusión, considérese el siguiente ejemplo en términos de agregados monetarios: es cierto que el componente irregular de la serie de activos líquidos en manos del público (ALP) es importante porque va a influir en la evolución a muy corto plazo de los mercados monetarios, y especialmente en los tipos de interés.

En cambio la tendencia está muy relacionada con la evolución a largo plazo de las variables nominales de la economía, sobre todo con las de gasto nominal y precios.

Por lo tanto, existen sobradas razones para extraer el componente irregular de ALP y utilizarlo como un indicador de ciertos factores que causan variaciones a corto plazo en las variables como tipo de interés, índices bursátiles, etc. Así mismo, también es de gran interés obtener la tendencia de la misma serie y usarla en un indicador de la evolución sostenida de variables económicas como precios, gasto, etc.

Sin embargo, no parece existir ningún motivo para combinar en una misma serie la tendencia y el elemento irregular, ya que es difícil imaginar alguna situación en donde tal agregación sea de utilidad: los procesos económicos en los que cada componente interviene son tan dispares, que difícilmente la serie ajustada podrá reflejar el tipo de señal que se requiere. Ambos componentes son importantes, pero por razones distintas, por lo que parece que el usuario desearía disponer de ellos por separado pero no en un solo agregado.

2. La serie ajustada de estacionalidad se estima con más precisión que la tendencia: cuanto más contaminada está una señal —es decir, menos oscilaciones se han eliminando de la misma— mejor se estima.

Es cierto que la serie desestacionalizada se estima mejor que la tendencia, como se ilustró en el epígrafe anterior al discutir la incertidumbre asociada a la estimación de ambas señales para la variable ALP.

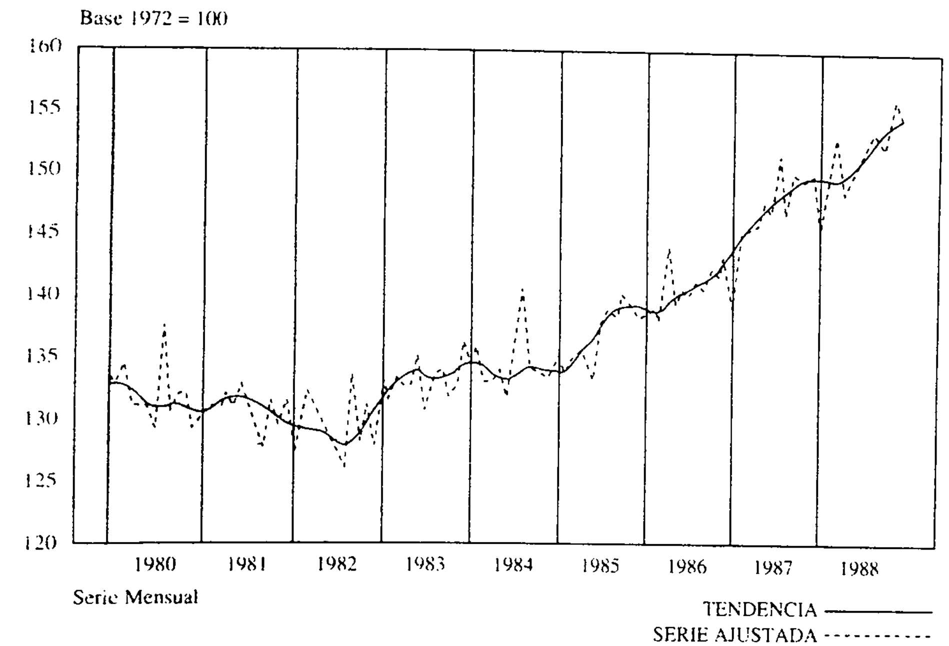
Esto siempre será así, ya que si una serie (en este caso Y_t^a) es igual a otra (T_t) más ruido, por definición siempre se estimará de forma más precisa. Ahora bien, al analizar la evolución subyacente a través de una señal contaminada con oscilaciones irregulares puede ocurrir, y en la mayor parte de los casos ocurre, que el análisis resulte más confuso que si se hiciese sobre una serie más suave, aunque esté peor estimada.

En el gráfico 4.2 se representan ambas señales para el Indice de Producción Industrial, tal y como se estiman en Morales et al. (1989): resulta claro como la contribución del elemento irregular es apreciable, y en consecuencia es ciertamente difícil identificar la serie desestacionalizada con el concepto intuitivo de evolución subyacente.

Hay que tener en cuenta que si la precisión en la estimación realmente fuese un criterio determinante en la selección de la señal, lo correcto sería utilizar siempre la serie original, ya que ésta se estima sin ningún tipo de error a partir de sí misma. Todo proceso de suavización, al ser una operación para extraer un componente no observable, implica pérdida de precisión; la mejor decisión no es escoger la serie que se estima con más precisión, sin tener en cuenta ningún otro tipo de consideración. Al contrario, se ha de procurar establecer una comparación entre precisión en la estimación y menor presencia de oscilaciones para cada una de las señales, y decidir en consecuencia.

En el ejemplo para ALP del epígrafe 4.6 se comprobó como el

GRÁFICO 4.2. *Tendencia y serie ajustada de estacionalidad del índice de producción industrial.*



error total del estimador concurrente de la tendencia tenía una desviación típica igual a 0,25%, frente al 0,23% de la serie desestacionalizada. Con una diferencia tan pequeña, parece difícil de justificar, al menos para este indicador monetario, la decisión a favor de utilizar la serie ajustada de estacionalidad como una señal sobre la que seguir el control monetario.

3. Para poder proceder sin actualizar la señal.

La práctica habitual para estimar la serie desestacionalizada es la siguiente: con información hasta diciembre del año n se avanza una estimación de los valores del componente estacional para los doce meses del año $n+1$; y a medida que se van observando los valores de la serie agregada correspondientes al año $n+1$, se dividen por el valor avanzado del correspondiente coeficiente estacional y se obtienen directamente los correspondientes valores desestacionalizados, sin necesidad de aplicar mes a mes ningún procedimiento de extracción de señales.

Más concretamente, si la observación t se refiere a diciembre del año n, se calcula $S_{t+j}^{(t)}$, $j=1, 2, \ldots, 12$; a partir de aquí

$$\tilde{Y}_{t+j}^a = Y_{t+j}/S_{t+j}^{(t)},$$

donde el símbolo $\sim$ se utiliza porque no es una estimación eficiente dado el conjunto de información disponible; ésta sería

$$Y_{t+j}^{a(t+j)} = Y_{t+j}/S_{t+j}^{(t+j)}$$

es decir, usando el coeficiente estacional que se estima con información hasta el último dato conocido.

Además la eficiencia implica que con cada nuevo dato del agregado se deben actualizar las estimaciones de la señal correspondiente a momentos anteriores: cuando se conozca el dato correspondiente a $t+k$ habría que reestimar los coeficientes estacionales, y con ellos calcular

$$Y_{t+j}^{a(t+k)} = Y_{t+j}/S_{t+j}^{(t+k)}, \qquad j<k.$$

La actualización es importante porque permite valorar la influencia de los nuevos acontecimientos que han incidido en el sistema. Además, al sustituir el futuro desconocido por realizaciones del mismo, la incertidumbre de la señal se va reduciendo, y el estimador provisional se parece cada vez más al definitivo.

A esa operación de actualización no hay que juzgarla como algo negativo. Actualizar implica estimar eficientemente una señal con toda la historia del agregado disponible en ese momento, lo que naturalmente repercute en una información más fiable para el proceso de toma de decisiones. Lo que es difícil de justificar es justamente el tomar decisiones ignorando información disponible, que es lo que se hace al utilizar señales no actualizadas.

En esta cuestión la serie ajustada y la tendencia presentan una diferencia esencial; así como el proceso antes señalado como práctica habitual se puede llevar a cabo con la serie desestacionalizada, con la tendencia no sucede así. Con la tendencia o se incorpora la nueva información contenida en los datos recién observados de forma eficiente, o no se incorpora en absoluto; pero no cabe un término medio, una incorporación parcial como en el caso de la serie ajustada.

Desde ese punto de vista la serie desestacionalizada presenta una «ventaja», ya que se puede tener una estimación ineficiente de esa señal, en el sentido de que puede tener más variabilidad que la estrictamente imprescindible. Pero parece difícil emplear esta propiedad para recomendar su utilización, sobre todo cuando sabemos que las primeras actualizaciones son las que conllevan mayor reducción en el error total de la estimación.

Como conclusión general se puede afirmar que no hay razones, ni teóricas ni prácticas, que justifiquen el uso generalizado de la serie ajustada de estacionalidad como señal en la que basar el análisis de la coyuntura. Por el contrario, se ha visto que sí existen motivos para considerar por separado el componente tendencial y el irregular, así como para actualizar las señales con la llegada de un nuevo dato, con lo que la pretendida «ventaja práctica» de la serie ajustada también desaparece.

En consecuencia en este libro se propone que, en general, el análisis de la evolución a corto plazo de los fenómenos económicos debe basarse en su tendencia, que es el componente no observable que mejor suele aproximar el nivel subyacente de la serie observada.

4.8. Modelización de determinados efectos relevantes en la extracción de señales mediante variables artificiales

4.8.1. Planteamiento del problema

En la sección primera de este capítulo se señaló la necesidad de distinguir la parte determinista de la parte estocástica a la hora de proceder a la descomposición del agregado en componentes no observables. Una vez discutida la descomposición de la parte estocástica, en este epígrafe y en el siguiente se verá el tratamiento de la parte determinista.

Hasta ahora siempre que se han utilizado variables artificiales en un modelo univariante ha sido para captar el efecto de alteraciones, puntuales o permanentes, de carácter exógeno. No obstante, existen ciertos efectos que ocurren de forma recurrente, año tras año, que sólo se pueden modelizar de modo relativamente sencillo de forma determinista. Estos efectos no constituyen un análisis de intervención tal y como lo presentan Box y Tiao (1975), pero como están captados mediante variables artificiales se mantiene la denominación.

En este epígrafe se definirán y estudiarán este tipo de efectos, y en el siguiente se discutirá como proceder a su descomposición en términos de tendencia, estacional e irregular.

4.8.2. Efecto calendario

En variables mensuales que se obtienen por agregación temporal de los correspondientes valores diarios, se tiene que si en estas últimas existe algún tipo de estacionalidad semanal, el valor que toma el agregado mensual depende de la composición del mes.

Un ejemplo típico lo constituyen las series de actividad como el Indice de Producción Industrial (IPI). Esta serie se caracteriza, a efectos de lo que se está tratando, por: *a*) la producción mensual es la suma de las producciones diarias, y *b*) existe una estacionalidad de carácter semanal, que lleva a distinguir, al menos, entre días laborables y no laborables.

La primera característica conlleva que cuantos más días tenga el mes, y a igualdad de los otros factores que influyen en el IPI, mayor será el valor de la serie mensual. La segunda, que si el mes t contiene más días laborables que el mes t-1, el valor de la serie en t tenderá a ser mayor que el de t-1, aunque ambos meses tengan el mismo número total de días y la actividad industrial sea la misma en ambos. Normalmente los días laborales tampoco tienen el mismo nivel de actividad, por lo que en ocasiones habrá que distinguir cuantos días de cada tipo tiene cada mes.

A este tipo de efecto se le conoce en la literatura como *efecto calendario*, y su influencia en el dato mensual es bien conocida por los analistas del corto plazo desde hace décadas. Sin embargo, no fue objeto de un tratamiento estadístico riguroso hasta los trabajos de Eisenpress (1956) y sobre todo Young (1965); Cleveland y Devlin (1980), Liu (1980), Cleveland y Grupe (1982) y Hillmer et al. (1983) proporcionan una perspectiva actual.

Su principal manifestación es inducir en la serie un ciclo de período 2,84 meses[35]. En el análisis aplicado, el efecto calendario se aproxima de forma determinista. Para ello se define un conjunto de variables artificiales $U_{1t}, U_{2t}, \ldots, U_{7t}$, donde U_{1t} es el número de lunes del mes t, U_{2t} el número de martes, ..., y así hasta U_{7t}, que recoge el número de domingos del mes t. El desarrollo que viene a continuación se realiza siguiendo a Cleveland y Grupe (1982).

El efecto calendario total asociado al mes t viene dado por

$$\delta_1 U_{1t} + \delta_2 U_{2t} + \cdots + \delta_7 U_{7t}$$

donde δ_i es un parámetro que recoge el efecto de un día del tipo i sobre la variable de interés. Suponiendo que, con exclusión hecha del efecto de calendario, la variable es estacionaria, estos parámetros se aproximan estimando la relación

$$Y_t = \delta_1 U_{1t} + \delta_2 U_{2t} + \cdots + \delta_7 U_{7t} + n_t \qquad (4.8.1)$$

siendo n_t una perturbación estacionaria.

[35] Este ciclo difícilmente se aprecia en el gráfico de la serie original pero se manifiesta claramente como un máximo local al calcular el (pseudo)espectro de la serie. Sin embargo, y debido a la proximidad de los máximos correspondientes a los ciclos estacionales de periodos 2, 2,4 y 3 meses, puede pasar desapercibido en un primer análisis exploratorio del espectro.

Sin embargo, si se pretende estimar directamente la relación (4.8.1) se obtendrán estimadores muy poco precisos de los coeficientes δ_i, ya que por construcción las variables U_{it} están muy correlacionadas entre sí. Para evitar este problema, es conveniente reparametrizar el modelo mediante el siguiente cambio de variable:

$$V_{it} = U_{it} - U_{7t} \qquad i = 1, 2, \ldots, 6$$

$$V_{7t} = \sum_{i=1}^{7} U_{it}$$

con lo que

$$\sum_{i=1}^{7} \delta_i U_{it} = \sum_{i=1}^{7} \beta_i V_{it},$$

donde los nuevos parámetros β_i están relacionados con los δ_i mediante

$$\delta_i = \beta_i + \beta_7 \qquad i = 1, 2, \ldots, 6$$

$$\delta_7 = \beta_7 - \sum_{i=1}^{7} \beta_i. \qquad (4.8.2)$$

Tal y como se han construido las variables V_{it}, si se estima el modelo

$$Y_t = \beta_1 V_{1t} + \beta_2 V_{2t} + \cdots + \beta_7 V_{7t} + n_t, \qquad (4.8.3)$$

que es simplemente una reparametrización de (4.8.1), se obtienen estimadores con precisión satisfactoria para β_i. Una vez conocidas las estimaciones de β_i es inmediato calcular los valores correspondientes a los estimadores de los parámetros δ_i mediante (4.8.2).

No obstante, nótese como la parametrización (4.8.3) también tiene una interpretación directa:

- V_{7t}: representa el número total de días del mes t;
- $\beta_7 = \Sigma \delta_i / 7$: es el efecto medio diario, o lo que es lo mismo, el aumento en Y_t al añadir un día más al mes t, si se desconoce qué tipo de día (lunes, martes, ...) es, y se aproxima su contribución por un promedio general;
- V_{it}, $i = 1, 2, \ldots, 6$: representa el número de días del tipo i que tiene el mes t con exceso sobre el número de días de un tipo que se toma como referencia, y que en la definición antecedente es el domingo. Por ejemplo, si V_{1t} es igual a cero, entonces en el mes t hay el mismo número de lunes y de domingos;

— β_i, $i=1, 2, \dots, 6$: es el efecto diferencial del tipo de día i con respecto al efecto medio β_7.

Es importante destacar que los coeficientes β_i, $i=1, \dots, 6$, reflejan, por definición, el coeficiente estacional semanal correspondiente al tipo de día i. Dicho coeficiente para el domingo no se puede obtener de forma directa añadiendo un sumando a la ecuación (4.8.3), ya que en este caso se tendría multicolinealidad perfecta; pero llamando β_7^* al coeficiente que se busca, es inmediato comprobar que

$$\beta_7^* = \delta_7 - \beta_7 = -\sum_{i=1}^{6} \beta_i$$

ya que por definición la suma de coeficientes estacionales ha de ser cero.

Como se desprende del análisis anterior, la expresión del efecto calendario y la explotación de la información que contiene son relativamente sencillas. Más complejo es el problema de su estimación: al formular la ecuación (4.8.1) se hizo el supuesto de que Y_t era estacionario, lo que en la mayor parte de las aplicaciones no se cumplirá.

Los estimadores de los parámetros en (4.8.1) o (4.8.3) son consistentes siempre que el residuo sea estacionario, por lo que si Y_t es una variable no estacionaria con transformación estacionaria $\Delta^d \Delta_{12}^D$, (4.8.1) y (4.8.3) se sustituyen por

$$\Delta^d \Delta_{12}^D Y_t = \sum_{i=1}^{7} \delta_i \Delta^d \Delta_{12}^D U_{it} + n_t \qquad (4.8.1')$$

$$\Delta^d \Delta_{12}^D Y_t = \sum_{i=1}^{7} \beta_i \Delta^d \Delta_{12}^D V_{it} + n_t . \qquad (4.8.3')$$

Ahora bien, aunque δ_i o β_i sean consistentes, los estimadores de sus respectivas matrices de varianzas-covarianzas no lo son, pues n_t en general no es ruido blanco. En este caso no se puede realizar el contraste de presencia de efecto calendario, cuya hipótesis nula es

$$H_0: \quad \delta_1 = \delta_2 = \cdots = \delta_7 = 0$$

o alternativamente

$$H_0: \quad \beta_1 = \beta_2 = \cdots = \beta_7 = 0$$

ya que los estadísticos que habitualmente se utilizan no tendrán las distribuciones usuales.

En consecuencia, la mejor recomendación que se puede hacer es la siguiente: cuando se sospecha que una serie presenta efecto calendario, se estima la relación (4.8.3') y se toman las estimaciones de la perturbación, $\hat{n}_t$, como la transformación estacionaria de la parte estrictamente estocástica de la serie. Una vez especificado, estimado y validado un modelo ARMA para estos residuos, se estiman conjuntamente los parámetros del efecto calendario y de la estructura estocástica, y se contrasta si existe efecto calendario. De forma esquemática:

1. Se estima

$$\Delta^d \Delta_{12}^D Y_t = \sum_{i=1}^{7} \beta_i \Delta^d \Delta_{12}^D V_{it} + n_t$$

y se obtienen los residuos $\hat{n}_t$.

2. Se construye un modelo ARMA para los residuos

$$\phi_p(L)\, \hat{n}_t = \theta_q(L)\, a_t$$

3. Se estima conjuntamente el modelo completo

$$\Delta^d \Delta_{12}^D Y_t = \sum_{i=1}^{7} \beta_i \Delta^d \Delta_{12}^D V_{it} + \frac{\theta_q(L)}{\phi_p(L)}\, a_t$$

4. Se contrasta la presencia de efecto calendario

$$H_0: \quad \delta_1 = \delta_2 = \cdots = \delta_7 = 0$$

mediante un estadístico F, que ahora sí seguirá (asintóticamente) la distribución habitual.

5. Si se rechaza la hipótesis nula, hay efecto calendario y las estimaciones de la tercera etapa son las definitivas; si no se rechaza, se reestima el modelo sin efecto calendario

$$\phi_p(L)\, \Delta^d \Delta_{12}^D Y_t = \theta_q(L)\, a_t$$

y éste es el resultado final del proceso de modelización.

4.8.3. Semana Santa o Pascua

Una segunda alteración que se presenta año tras año es la *Semana Santa*, y que al menos en España tiene un efecto cada vez más distorsionador sobre la actividad económica.

La principal característica de la Semana Santa es su *movilidad*: si el Viernes Santo fuese una fecha fija su influencia no presentaría un problema especial, pues aparecería recogida en la estimación del componente estacional por los procedimientos vistos en las secciones anteriores. Sin embargo esa movilidad se traduce en:

a) es necesario diseñar variables artificiales que reflejen en cada año el efecto que ejerce sobre la variable estudiada en los meses de marzo y abril;

b) hace difícil estimar cuál es exactamente la duración del período afectado por la Semana Santa, y

c) existe cierta evidencia en análisis desagregados en el tiempo[36] de que la intensidad del efecto de Semana Santa, así como su duración, depende mucho de la fecha concreta del calendario en que cae. El efecto sobre ciertas variables de una Semana Santa en la que el Viernes Santo ocurre el 25 de marzo puede ser distinto al que tiene si cae el 3 de abril.

En principio, y dada la complejidad de la Semana Santa y su posible evolución en el tiempo, parece recomendable un esquema estocástico para su modelización. No obstante, las series temporales económicas contienen una información limitada sobre esta clase de efecto: únicamente se observa una Semana Santa por año, y además las Semanas Santas de distintos años no son estrictamente comparables. De ahí que en el análisis aplicado se use un esquema determinista para su modelización, con lo que al menos se aproxima el efecto medio.

La modelización del efecto de Semana Santa suele realizarse de la siguiente manera: se define lo que se llama el *período de Semana Santa*, por ejemplo de k días[37]. Si hay información extramuestral lo suficientemente fiable como para ponderar de forma distinta cada día, se puede incorporar al análisis[38]; en caso contrario, se asigna la misma ponderación a cada uno de los k días. En Cancelo y Espasa (1991a) se estima, para cada uno de los días del período comprendido entre el lunes santo a martes de Pascua, la siguiente reducción, en tanto por cien, en el consumo diario de energía eléctrica: 0,65, 1,98, 3,01, 26,40, 34,32, 14,50, 6,46, 39,27 y 6,71, respectivamente. Para ciertas variables, especialmente las relacionadas con la actividad económica, los coeficientes anteriores se pueden utilizar para ponderar los días del período vacacional de Pascua. En lo que sigue se denominará P a la suma de las ponderaciones empleadas.

[36] Véase Loureiro (1986) y Cancelo (1988b, 1990).

[37] Normalmente la muestra no permite estimar con precisión aceptable el valor de k, y el analista debe fijarlo a priori.

[38] Véase Morales et al. (1989).

A continuación se construye una variable artificial SS_t, igual al tanto por uno que la ponderación de los días del período de Semana Santa que caen en el mes t supone sobre P. Si, por ejemplo, k es igual a 10, las ponderaciones son unitarias y en un año determinado ocho días caen en marzo y dos en abril, los valores de SS_t en ese año serán, 0,8 para marzo, 0,2 para abril y 0 para los demás meses.

La definición de estas variables se realiza de forma automática en algunos programas de análisis de series temporales, concretamente en el programa SCA[39]. No obstante, análisis con series diarias[40] han demostrado que para variables muy relacionadas con la actividad económica el efecto de la Semana Santa dura aproximadamente hasta el martes o el miércoles siguientes a la Pascua. Por lo tanto es recomendable no restringir a priori la extensión del efecto limitándolo sólo a lo que sucede hasta el domingo de Pascua, como hace el programa SCA. Ello exige prescindir de la opción automática de construcción de la variable y que el propio analista se encargue de generarla.

Una vez se ha construido la variable SS_t, si Y_t es estacionaria se plantea la expresión

$$Y_t = \alpha SS_t + n_t \qquad (4.8.6)$$

y se estima el coeficiente α. Los problemas asociados a la estimación del efecto de Semana Santa son en todo equivalentes a los vistos al analizar el efecto calendario, con lo que vuelve a ser de aplicación todo lo que allí se dijo.

4.8.4. Festivos intrasemanales

Un tercer tipo de distorsión de carácter estacional está causado por la presencia de *festivos intrasemanales*. El análisis de coyuntura tradicional normalmente no tiene en cuenta su influencia; ello se debe, por un lado, a que considera que su efecto sobre la serie mensual es pequeño; y por otro, que se repite en idénticas condiciones año a año, y en consecuencia queda bien reflejado en la correspondiente media móvil.

Sin embargo, cuando se trasladan resultados referidos a datos

[39] Véase Scientific Computing Associates (1986).

[40] Véase Cancelo (1990) y Cancelo y Espasa (1991a).

diarios al correspondiente agregado mensual se comprueba que el peso de los festivos es más importante de lo que se pensaba[41], ya que:

a) Existen fiestas móviles, como Corpus Christi, que pueden caer en distintos meses y distorsionar el patrón mensual y, por otra parte, en España el calendario oficial de fiestas cambia de un año a otro.

b) La población afectada por una fiesta varía, ya que las Comunidades Autónomas tienen potestad para introducir ciertos cambios en el calendario estatal. Así, el 25 de julio fue festivo en aproximadamente el 90% del territorio nacional en el año 1983, pero sólo en el 30% en 1987[42].

c) Los festivos son móviles dentro de la semana, y si existe estacionalidad de periodo semanal esa movilidad puede distorsionar significativamente el agregado mensual. En Cancelo y Espasa (1991a) se calcula el efecto que una fiesta estándar, según el día de la semana en que caiga, tiene en el consumo de energía eléctrica: en el cuadro 4.1 se reproducen dichas estimaciones. Aplicando este tipo de análisis diarios se estima que en 1988 la fiesta del 1 de mayo tuvo un efecto sobre el consumo total del mes del 0,12%, en el sentido de que el total mensual hubiese sido un 0,12% más alto de no existir dicha fiesta. Sin embargo, en 1987 este efecto fue del 1,2%: la diferencia se debe a que en 1988 el 1 de mayo cayó en domingo y en 1987 en viernes.

d) Relacionado con la movilidad anterior, están los efectos inducidos por los puentes, que pueden reforzar considerablemente el efecto de la fiesta propiamente dicha (véase el cuadro 4.1).

CUADRO 4.1. *Efecto de los festivos de carácter general sobre el consumo de energía eléctrica: (porcentaje de caída en el consumo)*

Día del festivo	*Efecto sobre*					
	Lunes	*Martes*	*Miércoles*	*Jueves*	*Viernes*	*Sábado*
Lunes	29,6	4,2				
Martes	10,7	34,5	3,4			
Miércoles			29,7	4,1		
Jueves				28,7	11,4	2,6
Viernes					29,0	8,8
Sábado						8,3

[41] Véase Cancelo y Espasa (1991b).

[42] De hecho la determinación de los porcentajes no se puede realizar sin considerar el problema concreto que se ha de resolver: estos porcentajes del 90% y del 30% se refieren al consumo de energía eléctrica correspondiente a la zona en fiestas, que era la variable que se analizaba en el trabajo citado en la nota anterior.

Por todo ello, los festivos juegan un papel destacado en el análisis de coyuntura de cierto tipo de variables económicas. Cuando se dispone de un modelo de periodicidad diaria o semanal que contiene información detallada sobre las fiestas, se puede usar dicho modelo para corregir la serie, agregar mensualmente la serie corregida, y basar el análisis de coyuntura en esa nueva serie. Este es el procedimiento seguido por Cancelo y Espasa (1991b) para elaborar un indicador semanal y otro mensual de actividad económica basados en el consumo de energía eléctrica.

Cuando no se dispone de información con periodicidad inferior a la mensual, entonces se puede utilizar el procedimiento previsto en Morales et al (1989)[43]: se define de forma operativa qué se entiende por festivo, se construyen las variables ficticias que recogen el número de días festivos de cada mes[44] y se estima una relación de la forma

$$Y_t = \gamma\ \mathrm{FES}_t + n_t, \qquad (4.8.7)$$

donde para simplificar se supone que Y_t es estacionaria y que una sola variable artificial es suficiente. Los problemas que plantea la estimación de (4.8.7) son idénticos a los de estimar (4.8.1), (4.8.3) o (4.8.6), por lo que no se volverán a discutir.

En el mencionado trabajo de Morales et al (1989), referido al Indice de Producción Industrial, se distinguen dos tipos de variables de festivos, una para las fiestas de carácter nacional y otra para las autonómicas. El efecto de una fiesta nacional se estima en una caída del IPI del 2,46%, coeficiente que tiene asociado un t-ratio de 8,2; en cuanto a las autonómicas[45], su efecto es aproximadamente del 1,57%, con un t-ratio de 3,4.

Estos resultados ponen de manifiesto que, incluso considerando únicamente información de periodicidad mensual, el tener en cuenta los festivos supone mejorar considerablemente la modelización, y con ello la estimación de los componentes no observables.

[43] Véase también el capítulo 7 de este libro.

[44] En el caso más sencillo todos los festivos se tratan por igual, pero es fácil generalizarlo cuando existen razones para distinguir las fiestas por el día de la semana en que caen y/o por la época del año en que ocurren. Además, se les puede dar a las fiestas ponderaciones diferentes en función de la población afectada.

[45] Sólo se consideran aquellas fiestas que afectan a más de un 60% de la población nacional.

4.8.5. Estacionalidad determinista

En muchas ocasiones es corriente encontrar un tratamiento *determinista* del componente estacional; merecen especial mención las siguientes circunstancias:

1. En un *modelo econométrico*: cuando se trata de estudiar la relación dinámica que liga a dos o más variables, el análisis debe hacerse siempre empleando las variables observadas, nunca las señales extraídas de las mismas sobre la base de eliminar una estacionalidad estocástica. Wallis (1974), en un trabajo ya clásico, demostró que la utilización de filtros dinámicos para extraer señales distorsiona la relación dinámica original entre las variables, por lo que es siempre preferible trabajar con las series observadas aunque la relación que se busca sea la existente entre los niveles subyacentes.

Si en estos casos el modelo econométrico incluye como inputs variables económicas que explican parte de la estacionalidad de la variable dependiente, el resto puede explicarse usando una aproximación deterministica basada en variables artificiales. Sin embargo, si los inputs económicos apenas explican la estacionalidad de la variable dependiente puede ser preferible formular un elemento residual, r_t, del tipo $U_{s-1}(L)r_t=\theta(L^s)a_t$, aunque esto complique la estimación.

2. Cuando la muestra es *pequeña*: si sólo se dispone de observaciones para un número reducido de años, los datos son incapaces de discriminar entre estacionalidad determinista y estacionalidad estocástica. En ese caso, y puesto que el coste informativo de la media móvil estacional es muy alto, es preferible aproximar la estacionalidad de forma determinista.

A estos efectos se entiende que una muestra es pequeña cuando se dispone de menos de cinco años de datos. Por ejemplo, el método X-11 ARIMA impone de forma automática que la estacionalidad es determinista siempre que la muestra mensual tenga menos de 60 observaciones: hay que tener en cuenta que la media móvil más corta que dicho método usa para estimar el componente estacional requiere al menos cinco observaciones para cada mes.

3. Cuando *se combinan estacionalidad determinista y estocástica*: ésta es la situación que se analiza detalladamente en Pierce (1978); la referencia Espasa (1983a) puede ser también útil para ilustrar este punto. En general la especificación de este tipo de modelos suele ser bastante compleja, por lo que remitimos al lector interesado a los trabajos mencionados.

La forma más empleada para estimar un componente estacional determinista consiste en definir impulsos estacionales:

$$D_{it}=\begin{cases}1 & \text{si la observación } t \text{ corresponde al mes } i\\ 0 & \text{en otro caso}\end{cases}$$

y plantear expresiones del tipo

$$Y_t=\alpha_0+\sum_{i=1}^{12}\alpha_i D_{it}+n_t \tag{4.8.8}$$

donde α_i se interpreta como el coeficiente estacional correspondiente al mes i, o de forma equivalente, $\alpha_0+\alpha_i$ es el nivel medio correspondiente a dicho mes.

Es bien sabido, véase, por ejemplo, Suits (1957), que los parámetros de (4.8.8) no están identificados, ya que existe multicolinealidad perfecta entre los doce impulsos estacionales y la constante; también es sobradamente conocida la solución, que consiste en eliminar una de las variables, por ejemplo la correspondiente al mes de diciembre, y estimar

$$Y_t=\beta_0+\sum_{i=1}^{11}\beta_i D_{it}+n_t. \tag{4.8.9}$$

Los coeficientes de (4.8.9) y de (4.8.8) están ligados por

$$\begin{aligned}\beta_0&=\alpha_0+\alpha_{12}\\ \beta_i&=\alpha_i-\alpha_{12} \qquad i=1,2,\ldots,11\end{aligned} \tag{4.8.10}$$

y lo que parece la ordenada en el origen en (4.8.9) es de hecho el nivel medio correspondiente al mes que se ha omitido en esa expresión; a su vez, el coeficiente del impulso estacional del mes i, β_i, es la diferencia entre el coeficiente estacional de ese mes y el mes omitido, o lo que es lo mismo, el exceso de estacionalidad (positivo o negativo) del mes i en relación al mes que se toma de referencia.

A partir de (4.8.10), y teniendo en cuenta que forzosamente $\sum_{i=1}^{12}\alpha_i=0$, es inmediato calcular la expresión de los coeficientes α_i en términos de β_i[46].

Una vez estudiada la especificación del modelo, los problemas que surgen al intentar estimar (4.8.9) son idénticos a los que planteaban

[46] El lector familiarizado con el análisis de la varianza habrá observado que si n_t en (4.8.8) es ruido blanco normal, el modelo con estacionalidad determinista se puede plantear como un ANOVA de un factor con doce niveles, correspondientes a los doce meses del año.

los efectos anteriores, y la forma de solucionarlos es la ya vista al tratar el efecto calendario.

La asignación de este tipo de factor determinista a un componente no observable es inmediata. Una vez estimados los coeficientes β_i, el siguiente paso es determinar los estimadores de α_i dando la vuelta al sistema de ecuaciones (4.8.10) —con la ecuación adicional $\sum_{i=1}^{12} \alpha_i = 0$— para tener una estimación de (4.8.8). Como es obvio, α_0 se asigna a la tendencia y $\sum_{i=1}^{12} \alpha_i D_{it}$ al componente estacional.

4.9. Descomposición de la parte determinista de una serie temporal

Para terminar este tema, en este epígrafe se establecen los criterios que rigen la descomposición del efecto de los factores deterministas en tendencia, estacional e irregular. Puesto que el factor a descomponer es por definición determinista, también los componentes lo serán.

4.9.1. Efecto calendario

Si bien el efecto calendario tiene un marcado carácter estacional, tal y como se comprobó en el epígrafe anterior, no se puede asignar sin más en su totalidad al componente estacional, pues su contribución a la variable de interés no suma cero en un año natural.

La descomposición de este efecto en sus componentes no observables se realiza de la siguiente manera[47]: puesto que la contribución total del efecto calendario se puede expresar como

$$\sum_{i=1}^{7} \beta_i V_{it} =$$

$$= \sum_{i=1}^{6} \beta_i V_{it} + \beta_7 LF_t + \beta_7 (V_{7t} - LF_t - 30{,}4375) + \beta_7 \cdot 30{,}4375 \qquad (4.9.1)$$

[47] Para un mayor detalle del proceso véase Hillmer et al. (1983).

donde LF_t es una variable definida como

$$LF_t = \begin{cases} 0{,}75 & \text{si } t \text{ es febrero de un año bisiesto} \\ -0{,}25 & \text{si } t \text{ es febrero de un año no bisiesto} \\ 0 & \text{en otro caso,} \end{cases}$$

de (4.9.1) se tiene:

a) $\beta_7 \cdot 30{,}4375$ es un efecto constante que afecta a todos los meses por igual: obsérvese que es el resultado de multiplicar el efecto promedio de un día por el número medio de días que tiene un mes (contando en este último promedio la distorsión que suponen los bisiestos).

b) $\beta_7(V_{7t} - LF_t - 30{,}4375)$ es el efecto estacional puro: si 30,4375 es el número medio de días del mes, la suma de las diferencias $V_{7t} - 30{,}4375$ para los doce meses de un mismo año tendría que ser cero; sin embargo como esa media se calcula teniendo en cuenta los años bisiestos, no será así salvo que se añada una variable de ajuste para los meses de febrero, que es precisamente lo que hace LF_t[48].

c) $\Sigma \beta_i V_{it} + \beta_7 LF_t$ es el efecto calendario propiamente dicho, y se comprueba que al sumar observaciones correspondientes a varios meses consecutivos es aproximadamente igual a cero. En algunos casos (Morales et al., 1989) se asigna directamente al componente estacional, si bien no es un componente estacional en sentido estricto; otros (Hillmer et al., 1983) prefieren hablar más explícitamente de un componente combinado estacional/de calendario. En cualquier caso, en términos de la descomposición en tendencia, estacional e irregular no hay discusión en asignar esta expresión al estacional.

4.9.2. Semana Santa o Pascua

El efecto de Semana Santa es otro ejemplo de una distorsión de carácter marcadamente estacional que sin embargo no se puede asignar directamente al componente estacional, ya que la suma de sus contribuciones a la variable estudiada, a lo largo de un año natural, no es igual a cero.

Para descomponer la contribución de la Semana Santa, que afecta a los meses de abril y marzo, se parte de la identidad

$$\alpha SS_t = \alpha[SS_t - (1/2)\, MA_t] + \alpha[(1/2)\, MA_t - (1/12)] + (\alpha/12) \qquad (4.9.2)$$

[48] Para un año de 365 días $\Sigma V_{7t} = 365$, $\Sigma(30.4375) = 12 \cdot 30.4375 = 365.25$; pero al ser un año no bisiesto, en febrero de ese año LF_t valdrá $-0{,}25$, por lo que $\Sigma(V_{7t} - LF_t - 30.4375) = \Sigma V_{7t} - (-0.25) - 365.25 = 0$.

donde:

a) $\alpha/12$ es una constante que refleja el efecto medio por mes de la existencia de la Semana Santa para todo el año natural[49].

b) $\alpha[(1/2)MA_t-(1/12)]$, siendo MA_t una variable ficticia que toma el valor uno si t corresponde a marzo o abril y cero en los demás casos, es el componente estacional en sentido estricto, pues la suma para los doce meses del año de la diferencia $(1/2)MA_t-(1/12)$ es igual a cero.

c) $\alpha[SS_t-(1/12)MA_t]$, a diferencia de lo que ocurría con el efecto calendario, se agota en cada Semana Santa, ya que la suma de los valores de esa diferencia para marzo y abril es siempre igual a cero. Si bien éste es el efecto Pascua en sentido estricto, en la descomposición se asigna al componente estacional determinista. Antes se dijo que si la asignación del efecto calendario puro al componente estacional lleva a distorsionar éste, ya que no está garantizado que su suma en doce meses sea cero; esto no ocurre con el efecto Pascua en sentido estricto, $\alpha(SS_t-(1/2)MA_t)$, que si suma cero al tomar los doce valores correspondientes a un año natural.

4.9.3. Festivos

La descomposición del efecto de los festivos se realiza reproduciendo la idea general de los casos anteriores. Sea el caso más sencillo en que sólo hay una variable artificial que refleja el número de festivos que ocurren en el mes considerado. Se define una variable NUMFES_t como el número total de fiestas que ocurre en el año natural al que se refiere la observación t (y que toma el mismo valor para los doce meses correspondientes a un mismo año natural); entonces a partir de la identidad

$$\gamma\ \mathrm{FES}_t=\gamma[\mathrm{FES}_t-(1/12)\,\mathrm{NUMFES}_t]+\gamma(1/12)\,\mathrm{NUMFES}_t \quad (4.9.3)$$

se tiene:

a) $(1/12)\mathrm{NUMFES}_t$ es el número medio de festivos correspondiente al año considerado, y como el último sumando es constante para todos los meses de un mismo año natural, se asigna a la tendencia.

b) Por construcción, la suma de los valores de $\mathrm{FES}_t-(1/12)\,\mathrm{NUMFES}_t$ correspondientes a un mismo año natural es igual a cero, lo que lleva a asignar al componente estacional el primer sumando a la derecha de la igualdad (4.9.3).

En el caso de que la variable FES_t incluya algún tipo de ponderación para tratar a distintos festivos de forma diferente, se aplica el procedimiento anterior modificando la definición de la variable NUMFES_t para tener en cuenta estas ponderaciones.

4.9.4. Variables de tipo escalón

Si la serie experimenta una perturbación exógena de carácter permanente, su efecto se ha de asignar en su totalidad a la tendencia, ya que representa un cambio en la evolución a largo plazo de la misma.

4.9.5. Variables de tipo impulso

Si la perturbación exógena es transitoria, afectando a la variable de interés de forma puntual y luego desapareciendo, lo más razonable es asignarla al elemento irregular, que por definición recoge las anomalías que modifican el corto plazo de la serie.

Es interesante destacar que las intervenciones propiamente dichas no requieren un análisis tan sofisticado como los factores deterministas que se presentaron en el epígrafe anterior. Esto se debe a que, tal y como se comentó en su momento, el efecto calendario, la Semana Santa o los festivos tienen un comportamiento cíclico y recurrente, cosa que no ocurre con una perturbación exógena; ésta afecta a la serie pero es algo ajeno a su evolución normal, y no un aspecto más de la misma que está modelizado de forma determinista.

[49] Un ejemplo aclarará este punto: sea Y_t las ventas de un comercio minorista en el mes t (sin tomar logaritmos), y supóngase que se estima que el efecto de la Semana Santa representa una caída en las ventas de 240 unidades monetarias, repartidas entre marzo y abril. Al cabo del año las ventas anuales estarán 240 unidades por debajo de lo que estarían si no hubiese existido la Semana Santa, por lo que la caída media por mes es igual a $240/12=20$.