

Análisis del ciclo económico mediante modelos de índices dinámicos. Aplicación a los Índices de la Producción Industrial¹

Alfredo Cristóbal
Enrique M. Quilis

S.G. de Cuentas Nacionales
INSTITUTO NACIONAL DE ESTADÍSTICA

¹ Los autores agradecen las discusiones mantenidas con Ana M^a Abad, Rafael Frutos, Francisco Melis y Julián Pérez acerca de la metodología que se expone en este trabajo. Asimismo, han contado con los comentarios de Hermenegildo Fernández y de participantes en seminarios del Instituto Gallego de Estadística y de las universidades Autónoma de Madrid y de Valencia. Las opiniones expresadas corresponden a los autores y no reflejan necesariamente las del Instituto Nacional de Estadística.

1. Introducción

El objetivo del presente trabajo es exponer una técnica de clasificación cíclica basada en una extensión del modelo factorial, aplicable al caso en que las observaciones provienen de series de tiempo, que se denomina "modelo de índices dinámicos". Esta clase de modelos posibilita tanto la identificación de patrones cíclicos comunes como la clasificación de las series respecto a dichos patrones en función de su desfase. De esta manera se obtienen cuatro categorías de series: adelantadas, coincidentes, retrasadas y acíclicas. Las principales aplicaciones de esta técnica son:

- (a) la identificación y caracterización de patrones cíclicos comunes;
- (b) la clasificación cíclica de las series individuales mediante la estimación de la estructura de desfases subyacente,
- (c) el diseño de índices sintéticos así como la reducción de la dimensionalidad del sistema;
- (d) la simplificación y especificación de modelos ARMA vectoriales y
- (e) el desarrollo de análisis de carácter estructural.

El trabajo se estructura de la siguiente forma. En la segunda sección se presenta de forma sumaria el modelo factorial estático, enfatizando su interpretación estadística. Las extensiones de dicho modelo al caso dinámico son comentadas en la tercera sección. Se presentan dos desarrollos diferentes: el modelo factorial dinámico y el modelo de índices dinámicos, haciéndose explícitas sus semejanzas y diferencias. La cuarta sección está dedicada a la especificación de esta última clase de modelos. Se expone un algoritmo de simplificación así como la interpretación de dichos modelos. Los resultados de la aplicación de esta metodología a 133 Índices de la Producción Industrial (IPI) son expuestos en la sección quinta. El trabajo se cierra con un apartado en el que se exponen algunas extensiones de esta técnica.

2. Modelo factorial estático

El modelo factorial estático considera que la interacción existente entre k variables puede ser explicada como el resultado de su común dependencia respecto de un conjunto más reducido de r factores inobservables que expresan fuentes de variabilidad independientes entre sí.

La relación postulada entre variables y factores se encuentra contenida en la siguiente ecuación:

$$[1] \quad Z_t = \Lambda F_t + \varepsilon_t$$

donde $Z_t:k \times 1$ es el vector de series observadas²; $F_t:r \times 1$ representa a los factores comunes; $\Lambda:k \times r$ es la matriz de cargas que establece la conexión entre variables y factores y $\varepsilon_t:k \times 1$ recoge la variabilidad idiosincrásica de Z_t o, si se prefiere, la parte de Z_t que no está explicada por los factores.

La primera característica notable de la ecuación [1] es que todo lo que se encuentra en su parte derecha es inobservable. Naturalmente, esto plantea un problema de equivalencia observacional que ha de ser resuelto incorporando una serie de hipótesis *a priori* acerca de la naturaleza de los elementos involucrados en [1]. Estas son:

$$[2a] \quad E[F_t] = 0$$

$$[2b] \quad E[F_t F_t'] = I_r$$

Esta hipótesis establece la ortogonalidad de los factores comunes, de manera que la información contenida en cualquiera de ellos no está comprendida en alguno de los restantes factores. Asimismo, en [2a] se establece la nulidad del vector de esperanzas.

$$[3a] \quad E[\varepsilon_t] = 0$$

$$[3b] \quad E[\varepsilon_t \varepsilon_t'] = \Sigma_\varepsilon = \text{diag} (\sigma_{\varepsilon 1} \dots \sigma_{\varepsilon k})$$

Con esta condición se asume que cada componente ε_{it} contiene exclusivamente información referida a la parte de z_{it} no explicada por los factores comunes. Se admite que la variabilidad de este componente "residual" sea también específica. Por estas razones ε_t recibe la denominación de "factor idiosincrásico". Asimismo, en [3a] se establece la nulidad del vector de esperanzas.

$$[4] \quad E[F_t \varepsilon_t'] = 0$$

² Sin pérdida de generalidad se asume $E[Z_t]=0$.

La condición anterior indica que los factores comunes y los idiosincrásicos carecen de elementos informativos comunes.

$$[5] \quad [\Lambda' \Sigma_{\epsilon}^{-1} \Lambda] = \text{diag} (v_1 \dots v_r)$$

Esta condición asegura la identificación del modelo factorial estático así como la ortogonalidad de los factores estimados por máxima verosimilitud.

La interpretación del modelo [1] complementado con las hipótesis [2]..[5] es sencilla. Cada variable z_{it} es expresada como combinación lineal de r patrones comunes ortogonales que recogen los movimientos conjuntos de las k series, más un componente de variabilidad intrínseca. La expresión formal correspondiente es:

$$[6] \quad z_{it} = \lambda_{i1} f_{1t} + \lambda_{i2} f_{2t} + \dots + \lambda_{ir} f_{rt} + \epsilon_{it}$$

La aplicación del modelo factorial estático a un vector de variables Z_t exige resolver, de forma secuencial, los siguientes problemas:

- (a) ¿cuál es el número apropiado de factores, r ?
- (b) ¿cómo estimar la matriz de cargas Λ a partir de la muestra sobre Z_t ?
- (c) ¿cómo recuperar los factores F_t condicionados a la estimación de Λ y a la información muestral sobre Z_t ?

Un tratamiento completo de estas tres cuestiones se encuentra en Mardia *et al.* (1979).

3. Extensión al caso dinámico: modelos factorial dinámico y de índices dinámicos

Un aspecto crucial del modelo expuesto en la sección anterior es su inadecuación para recoger dos de los aspectos básicos de los datos provenientes de series temporales: la dependencia de cada variable de sus propios valores desfasados (autocovariación) y la relación entre ésta y los valores retardados de las demás variables (covariación cruzada). En consecuencia, hay que modificar el modelo [1]..[5] de manera que refleje de forma más eficaz las propiedades dinámicas de los datos.

Existen dos maneras de generalizar el modelo factorial estático para que tenga en cuenta las características estadísticas de las series temporales. La primera, debida a Peña y Box (1987), amplía el modelo incorporando una ecuación que caracteriza la dinámica de los factores y, en consecuencia, la de las series contenidas en Z_t . Dicha ecuación establece que los r factores evolucionan según un modelo ARMA(p,q) multivariante:

$$[7] \quad (I - \Phi_1 B - \dots - \Phi_p B^p) F_t = (I - \Theta_1 B - \dots - \Theta_q B^q) a_t$$

donde Φ_i , $i=1..p$, y Θ_j , $j=1..q$, son matrices $r \times r$ de parámetros fijos pero desconocidos; a_t es una innovación de tipo ruido blanco multivariante con matriz de varianzas y covarianzas Σ_a y B es el operador de desfase ($B^h Z_t = Z_{t-h}$). Las ecuaciones [8a] y [8b] representan el modelo factorial dinámico de Peña y Box:

$$[8a] \quad Z_t = \Lambda F_t + \varepsilon_t$$

$$[8b] \quad \Phi(B) Z_t = \Theta(B) a_t$$

Un análisis detallado de este modelo se encuentra en Peña y Box (1984, 1987), Peña (1990), Escribano y Peña (1994) y Peña y Poncela (1996) y aplicaciones de carácter macroeconómico en Quilis (1994) e INE (1994).

La segunda manera de extender el marco factorial al dominio dinámico está relacionada con la aplicación del análisis de correlaciones canónicas a la especificación de modelos multivariantes de series temporales, bien en el espacio de los estados (Akaike, 1976) bien dentro de la clase de modelos ARMA vectoriales (Cooper y Wood, 1982). La representación que se utiliza en este trabajo se basa en Melis (1985), Cristóbal *et al.* (1996) y Cristóbal y Quilis (1996) y recibe el nombre de "modelo de índices dinámicos".

La idea básica es la siguiente: si el vector observado Z_t no contiene toda la información relevante para especificar la dinámica de las k variables que representa, entonces hay que ampliar su espacio de medida de manera que se incorpore la información dinámica ausente de Z_t . Así, en lugar de trabajar con Z_t se considerará el vector ampliado con sus m valores desfasados. El modelo de índices dinámicos se encuentra en las ecuaciones [9a] y [9b]:

$$[9a] \quad Y_{mt} = \Lambda F_t + \varepsilon_t$$

$$[9b] \quad Y_{mt} = [Z_t, Z_{t-1}, \dots, Z_{t-m}]'$$

Estas expresiones ilustran muy bien las estrategias adoptadas para extender el marco factorial: añadir la ecuación [8b] en el primer caso y extender Z_t en el segundo. La pregunta que surge de forma inmediata es: ¿qué relación guardan ambas representaciones?. La mejor manera de responder consiste en examinar la estructura de la matriz de varianzas y covarianzas: $\Gamma(m,j) = E[Y_{mt} Y_{m,t-j}]$. Esta matriz es la que sustenta el análisis de correlaciones canónicas entre Y_{mt} e $Y_{m,t-j}$ (véase Tsay y Tiao, 1985).

Considerando distintos valores para m y j se obtiene una tabla de doble entrada formada por matrices de dimensión $k(m+1) \times k(m+1)$. Así, en el caso $m=j=0$ se tiene la matriz de varianzas y covarianzas de Z_t . La primera fila de esta tabla, $\Gamma(0,j)$ $j=0,1,2,\dots$, está formada por las matrices de covarianzas cruzadas de Z_t . Estas matrices son de dimensión $k \times k$, generalmente asimétricas y sustentan el proceso de identificación del modelo factorial dinámico³. El modelo de índices dinámicos se deriva del estudio de la primera columna de $\Gamma(m,j)$. Así, se obtiene una secuencia de matrices de tipo Toeplitz, simétricas y de orden creciente $k(m+1) \times k(m+1)$.

Simplificando el argumento, cabría decir que el modelo factorial dinámico analiza la fila $\Gamma(0,j)$, con $j>0$, mientras que el de índices dinámicos considera la primera columna $\Gamma(m,0)$, con $m>0$. Naturalmente, no existe una relación directa necesaria entre ambas especificaciones. Así, por ejemplo, la estructura VARMA(p,q) de Z_t condiciona el rango de $\Gamma(0,j)$ pero no el de $\Gamma(m,0)$ (véase Tsay y Tiao, 1985).

Ambas representaciones ilustran aspectos diferentes del proceso dinámico generador de Z_t . El modelo factorial dinámico posibilita una especificación directa del comportamiento temporal de los factores y es especialmente útil en el marco del análisis de cointegración así como en la especificación de modelos ARMA vectoriales. El modelo de índices dinámicos es interpretable directamente como un modelo factorial estático, por lo que puede ser identificado, estimado y evaluado siguiendo los procedimientos habituales de este modelo. Su mayor utilidad radica en la clasificación dinámica de un vector de series temporales y en el diseño de índices sintéticos de los diferentes grupos así formados.

³ Naturalmente, de forma no exclusiva.

4. Especificación de un modelo de índices dinámicos

En este apartado se expone el procedimiento de especificación, estimación, simplificación e interpretación de un modelo de índices dinámicos, con el objeto de clasificar las señales cíclicas de un vector de series temporales.

Habitualmente, dichas señales han de ser estimadas utilizando procedimientos apropiados de extracción que no son el objeto de este trabajo (véase Quilis, 1995). En consecuencia, se asumirá que Z_t es un vector formado por las señales cíclicas de k series.

Se compone un vector extendido⁴ Y_t de dimensión $k(m+1) \times 1$ añadiendo a Z_t sus valores desplazados hasta el retardo m : $Z_{t-1} \dots Z_{t-m-1}$, Z_{t-m} :

$$[10] \quad Y_t = [Z_t, Z_{t-1}, \dots, Z_{t-m}]'$$

Como se verá, la decisión respecto al máximo retardo m considerado en el análisis tiene consecuencias importantes. A continuación se postula un modelo factorial para Y_t :

$$[11] \quad Y_t = \Lambda F_t + \varepsilon_t$$

donde F_t : $r \times 1$ representa los r factores comunes e inobservables siendo $r \ll k(m+1)$. Λ : $k(m+1) \times r$ es una matriz de parámetros -llamados "cargas"- que relacionan variables observadas con factores. Finalmente, ε_t : $k(m+1) \times 1$ es un vector que recoge los factores específicos de cada variable.

Así, el modelo factorial [11] intenta explicar la evolución conjunta de un vector (extendido) de series temporales en función de un número reducido de fuentes o factores comunes que, dada su ortogonalidad, reflejan aspectos diferentes o dimensiones independientes del proceso generador subyacente.

La metodología propuesta consta, en su forma más general, de cinco etapas:

(1) determinación del número r de factores, (2) estimación del modelo [11], (3) simplificación de la matriz de cargas Λ , (4) determinación de la serie de referencia y (5) clasificación de las $k-1$ series restantes respecto a la determinada en la etapa anterior. A continuación se detalla cada una de estas etapas.

Etapas 1: Determinación del número de factores.

La especificación del número apropiado de factores se realiza examinando los autovalores de la matriz de correlaciones Γ_m de Y_t así como sus aportaciones a la explicación de la suma de varianzas de las series que forman dicho vector (Mardia *et al.*, 1979). Especial relevancia tiene la presencia de efectos de acumulación de los menores autovalores (fenómeno de morrena o *scree*).

Un aspecto que conviene subrayar es la naturaleza dinámica de esta matriz, ya que contiene las matrices de correlaciones cruzadas de Z_t hasta el retardo m :

⁴ Como m es fijo, se omite el subíndice de Y .

$$\Gamma_Y(0) = \begin{bmatrix} \Gamma_Z(0) & \Gamma_Z(1) & \dots & \Gamma_Z(m) \\ \Gamma_Z(1)' & \Gamma_Z(0) & \dots & \Gamma_Z(m-1) \\ \dots & \dots & \dots & \dots \\ \Gamma_Z(m)' & \dots & \dots & \Gamma_Z(0) \end{bmatrix}$$

Esta matriz es de tipo Toeplitz (simétrica por bloques y en su conjunto) y condensa toda la información dinámica del vector original Z_t

Etapa 2: Estimación de la matriz de cargas.

Una vez determinado el número idóneo de factores, se estima el modelo [11]. En este trabajo se ha utilizado el método de los componentes principales porque genera de forma automática factores ortogonales. En otras aplicaciones podría recurrirse al método de la máxima verosimilitud (Mardia *et al.*, 1979).

La relación entre factores y patrones depende de la magnitud del máximo retardo m empleado. Dos series cíclicas puras (sinusoidales) de período de cuatro años y con un desfase de dos años, aparecerán asociadas al mismo factor con cargas altas y positivas si m es igual o superior a 24 meses. Si m es inferior a 24 meses, la primera serie aparecerá asociada al factor con carga alta y positiva y la segunda estará asociada al mismo factor con carga alta y negativa. El parámetro m que determina la resolución de desfases del procedimiento determina también la presencia de cargas negativas.

Etapa 3: Simplificación de la matriz de cargas.

Cuando el análisis aspira a obtener una clasificación de las series consideradas, esto es, una partición de la población analizada en subconjuntos disjuntos, cada serie debe estar asociada con un sólo factor, lo que implica seleccionar en cada fila de la matriz de cargas el elemento máximo en valor absoluto.

Para abreviar la exposición, y dado que en el caso que aquí se analiza se retiene un sólo factor, se considerará que el número apropiado de factores es uno ($r=1$), de forma que la matriz de cargas se reduce a un vector columna. Si r fuera mayor que uno bastaría con repetir r veces el procedimiento que se describe a continuación.

Sea $\lambda_{i,j}$ el elemento típico del vector columna de cargas Λ . El subíndice i ($i=1 \dots k$) hace referencia a la variable i -ésima que compone Z_t y el subíndice j ($j=0 \dots m$) determina el retardo de dicha variable. El procedimiento de simplificación consiste en retener sólo la mayor carga de cada variable, anulando todas las demás. Esto es, definir un nuevo vector columna de cargas Ω cuyo elemento típico $\omega_{i,j}$ se calcula de la siguiente manera:

[12]

$$\omega_{i,j} = \begin{cases} \lambda_{i,j} \Leftrightarrow |\lambda_{i,j}| = \max_j [|\lambda_{i,j}|] \\ 0 \text{ en los demas casos} \end{cases}$$

El vector columna simplificado proporciona los desfases relativos entre las series. Así, si la serie 1 posee su máxima carga en el retardo 0 y la serie 2 presenta la máxima carga en el retardo m, se afirma que la serie 2 adelanta m períodos a la serie 1.

Etapa 4: Determinación de la serie de referencia.

Una de las ventajas de esta técnica es que ofrece una manera puramente empírica de establecer cuál es la serie de referencia. Para ello conviene recordar que el factor común resume toda la información dinámica de las series consideradas, ya que se ha estimado considerando los sucesivos desfases de Z_t (al operar sobre el vector extendido Y_t) y que, en consecuencia, las cargas $\lambda_{i,j}$ especifican la vinculación de cada serie en cada desfase con el factor. De esta manera, podremos identificar como serie de referencia aquella que se encuentra más relacionada con el factor común obtenido. Formalmente:

[13] $z_{h,t}$ es la serie de referencia $\Leftrightarrow h = \arg \max_i [|\omega_{i,j}|]$

Naturalmente, la elección de la serie de referencia puede hacerse *a priori*, en cuyo caso esta etapa se omite.

Etapa 5: Clasificación respecto a la serie de referencia.

La serie de referencia $z_{h,t}$ está máximamente identificada con el factor común F_t en un retardo d ($0 \leq d \leq m$). Para cada variable $i \neq h$ se considera el retardo j correspondiente a su carga simplificada no nula $\omega_{i,j}$ y se calcula su desfase característico ($de_i=d-j$). De esta manera tenemos el siguiente criterio de clasificación:

Tabla 1: Criterio de clasificación cíclica

	Desfase caraterístico: de_i		
	$de_i < -L$	$-L \leq de_i \leq L$	$de_i > L$
Clasificación cíclica	Adelantada: <a>	Coincidente: <c>	Retrasada: <r>

Donde L es un parámetro fijado *a priori*. Naturalmente, la clasificación puede ser más detallada, incluyendo, si es posible, series muy adelantadas y/o muy retrasadas. El caso <i> (aciclicidad) aparece cuando $\omega_{i,j} = 0 \forall j$, esto es, cuando la serie considerada no está identificada de forma alguna con el patrón dominante y, en consecuencia, tampoco lo está con la serie de referencia $z_{h,t}$. Como la anulación exacta de las cargas rara vez se produce, es conveniente fijar de antemano un umbral de significación.

En algunas ocasiones sólo interesa establecer los desfases relativos entre las series asociadas a un mismo patrón y en este caso las etapas 4 y 5 son innecesarias.

5. Aplicación a los índices de producción industrial (IPI)

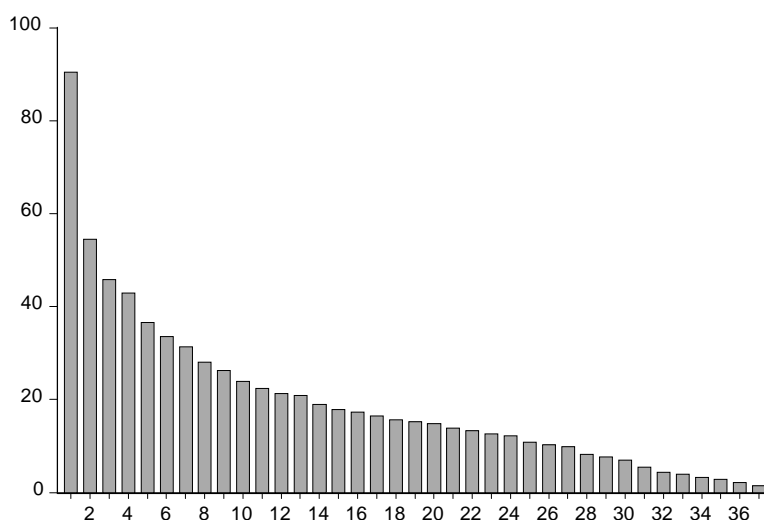
La metodología expuesta en la sección anterior va a ser aplicada a 133 Índices de la Producción Industrial (IPI), elaborados mensualmente por el INE. El período muestral abarca desde enero de 1975 hasta febrero de 1998, con lo que $n=278$.

En general, las series de los IPI están caracterizadas por una amalgama de elementos tendenciales, estacionales e irregulares que dificultan la estimación de su perfil cíclico. Por ello, en este trabajo se han filtrado dichas series por medio de la combinación de un filtro de paso en banda (la tasa interanual) y un suavizador de paso bajo (el filtro AR(2) de tipo Butterworth descrito en Melis, 1991). El resultado es una estimación bastante nítida del ritmo de avance de la serie temporal objeto de estudio.

La resolución de desfases seleccionada es $m=12$, que se ha mostrado como una elección razonable en los diversos ensayos realizados con esta técnica.

En el gráfico 1 se muestran los 37 primeros autovalores de la matriz de correlaciones de Y_t que, como ya se ha comentado, encapsula la dinámica de los 133 IPI considerados. La elevada distancia entre el primer y el segundo autovalor así como la presencia de un fenómeno de morrena (*scree*) a partir del segundo, sugiere que el número apropiado de factores es uno ($r=1$).

Gráfico 1: Autovalores de la matriz Γ_Y



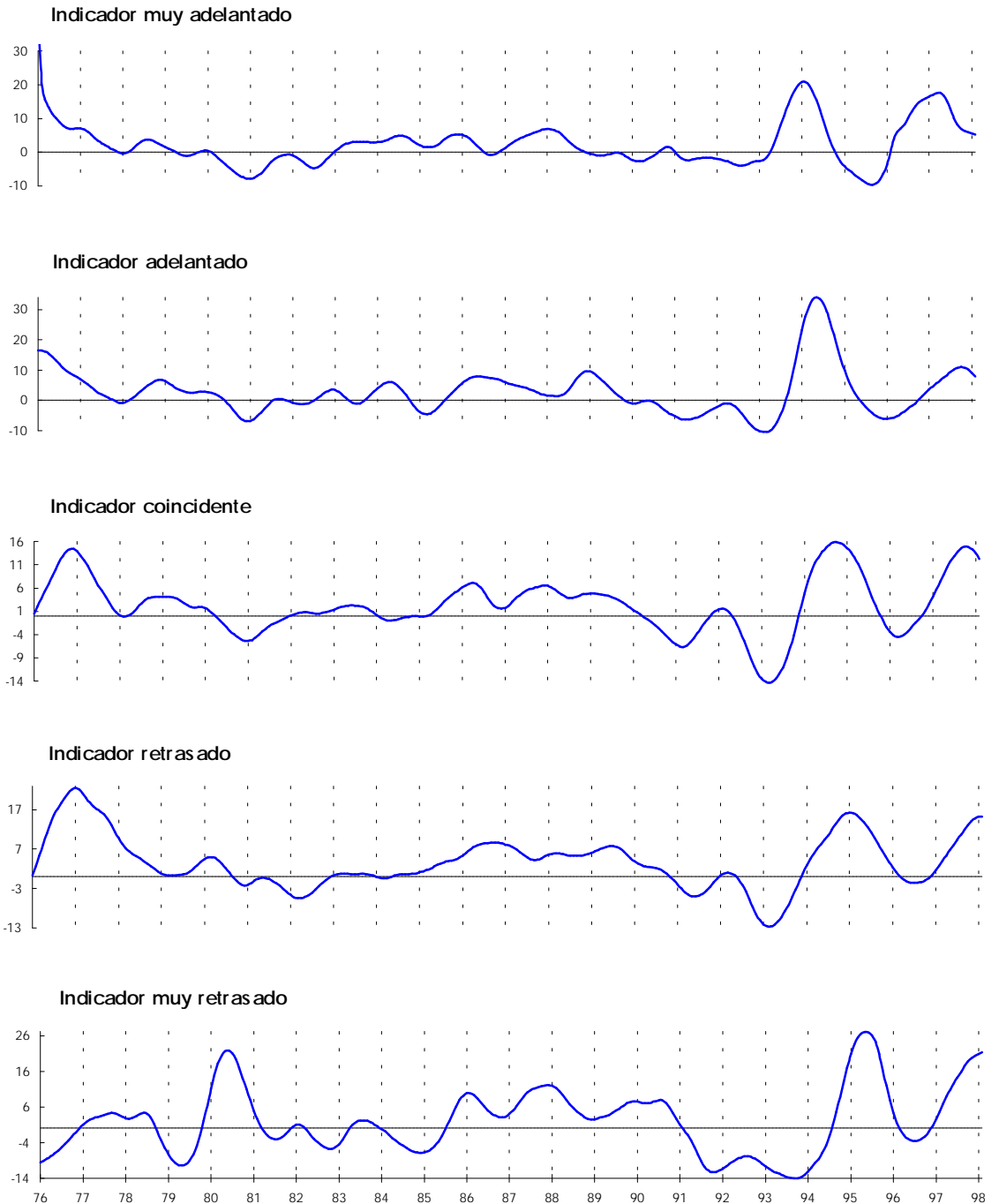
Una vez determinado el número adecuado de factores se puede estimar la matriz de cargas correspondiente y, como ya se ha comentado, especificar la estructura de desfases. No se ha considerado un umbral de significación *a priori* para las cargas individuales habida cuenta del desconocimiento práctico de su distribución. La dimensión de la matriz de tipo Toeplitz Γ_Y es 1729×1729 , lo que proporciona una medida de los requisitos de cómputo de esta clase de técnicas.

Como referencia temporal se ha tomado el punto medio de la retícula de desfases ($d=6$), formándose cinco grupos en función de la distancia de cada desfase característico, de_i . De esta manera se tiene:

- $\langle aa \rangle$: series muy adelantadas: $de_i \in [-6, -5]$
- $\langle a \rangle$: series adelantadas: $de_i \in [-4, -2]$
- $\langle c \rangle$: series coincidentes: $de_i \in [-1, 1]$
- $\langle r \rangle$: series retrasadas: $de_i \in [2, 4]$
- $\langle rr \rangle$: series muy retrasadas: $de_i \in [5, 6]$

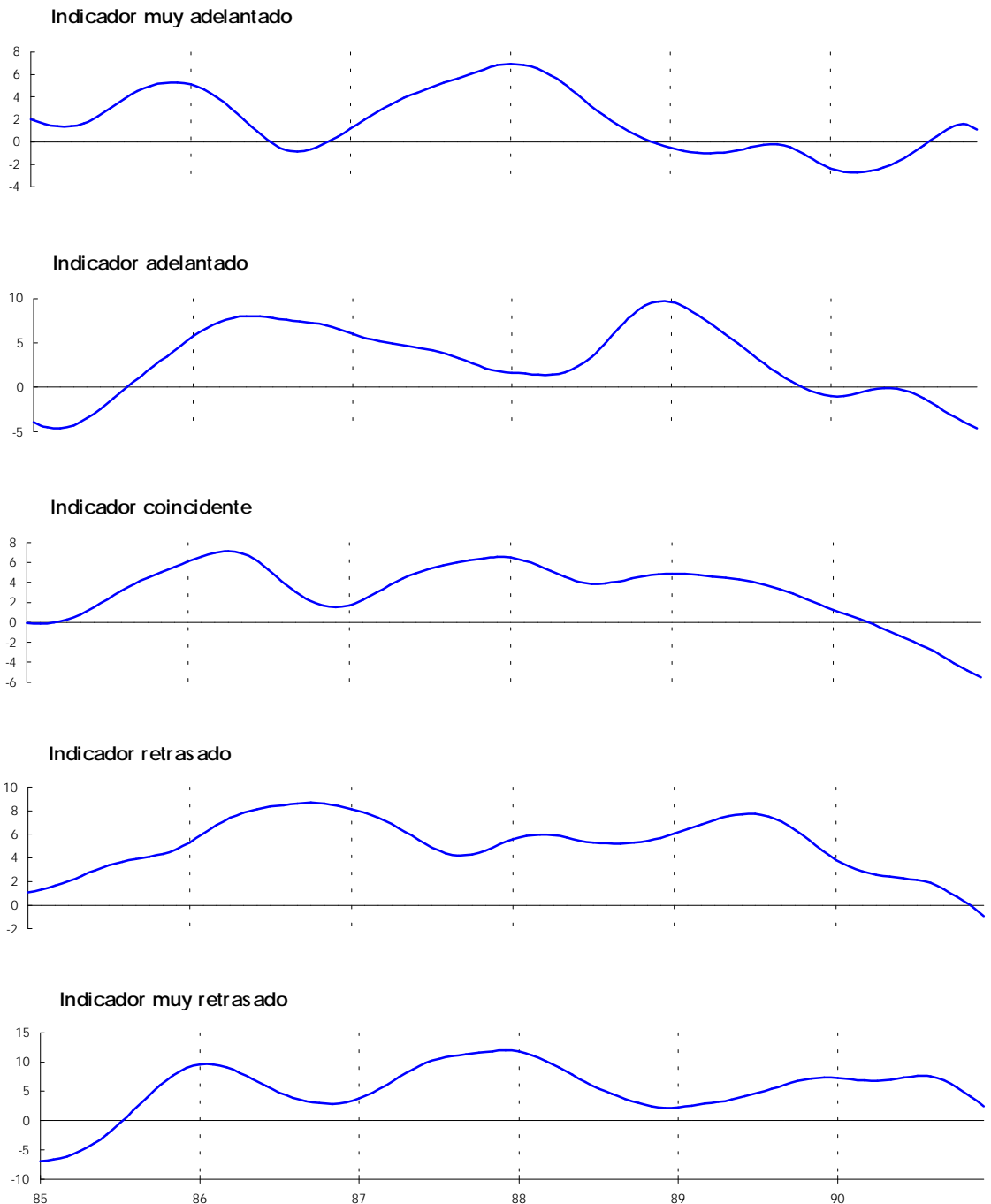
Habida cuenta de la elevada dimensión del sistema aquí modelizado, no se detalla la matriz de cargas correspondiente sino que se ofrecen, en el gráfico 2, los correspondientes cinco índices sintéticos. Dichos índices son una combinación convexa y normalizada de los IPI que integran cada categoría dinámica, tomando como base la matriz de cargas simplificadas. Hay que advertir que los grupos no están, por lo común, ubicados en un punto único sino en una banda temporal. Se ha elegido como centro de cada grupo el desfase que, comenzando en cero, se encuentra separado un trimestre por cada lado del correspondiente al siguiente grupo: 0, 3, 6 y 9. A continuación se colapsa cada banda a dichos puntos reteniendo la estructura interna de desfases. En Cristóbal y Quilis (1996) se encuentra una exposición detallada.

Gráfico 2: Índices sintéticos



Con el fin de apreciar mejor el comportamiento de los índices sintéticos, los gráficos 3 y 4 detallan dos períodos especialmente interesantes. El primero, 1985-1990, abarca buena parte de un ciclo amplio (1985-1991) en el que se encuentran inscritos varios subciclos. El segundo, 1992-1998, recoge la recesión de 1993 y la oscilación correspondiente que termina en 1996.

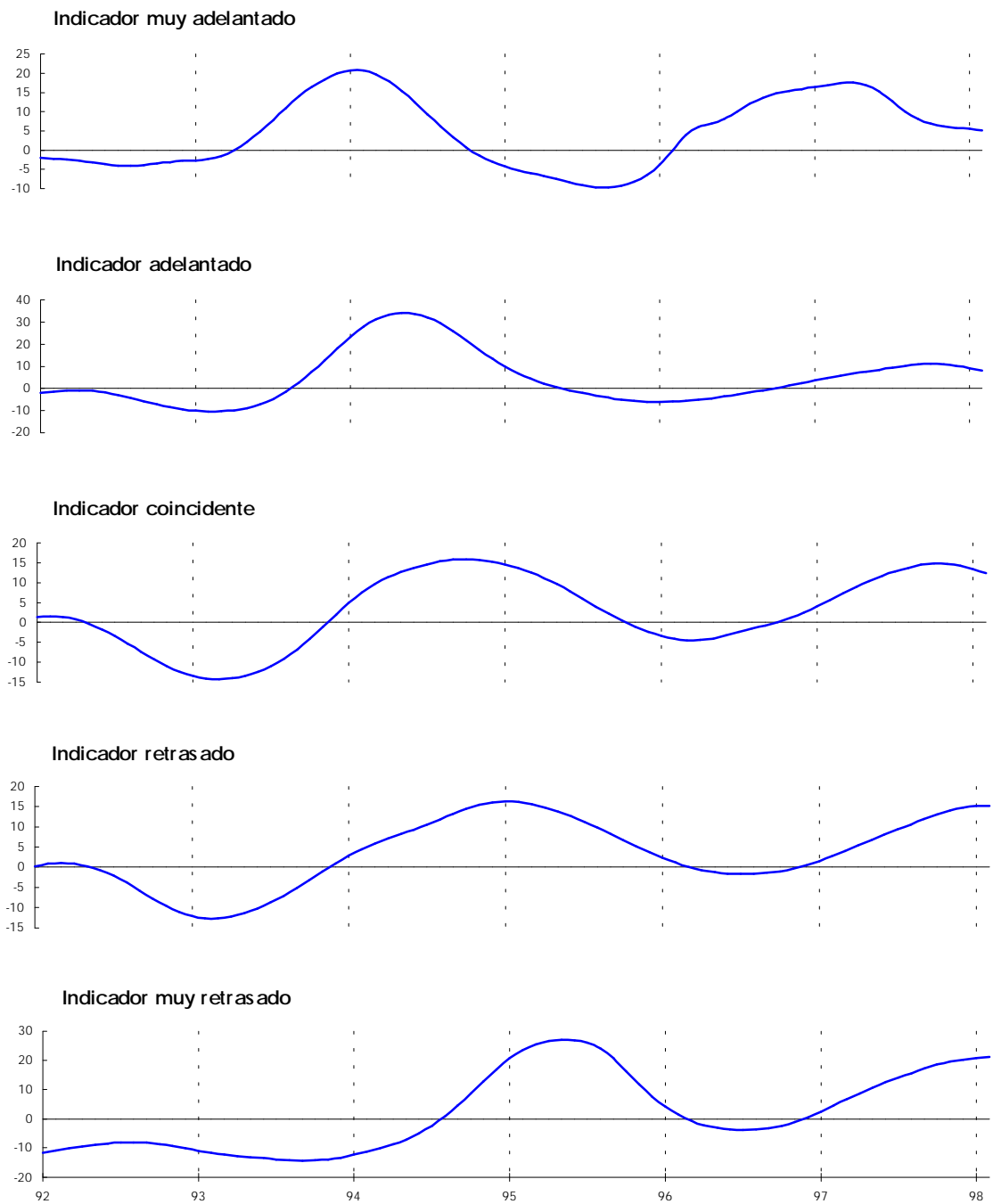
Gráfico 3: Índices sintéticos. Detalle 1985-1990



Como se aprecia en el gráfico anterior, la coherencia temporal entre los indicadores es significativa pero dista de ser perfecta, especialmente en lo que se refiere a los mínimos cíclicos. Por el contrario, como se observa en el gráfico 4, los movimientos cíclicos de los distintos índices guardan una sincronía más rigurosa en la fase 1992-1998, tanto en los mínimos como, sobre todo, en los máximos. Especialmente signi-

ficativa es la presencia de un máximo cíclico en los indicadores muy adelantado, adelantado y coincidente que estarían reflejando un cambio de fase en el conjunto de la producción industrial y que aún no se ha completado, ya que los índices retrasado y muy retrasado no han registrado todavía dicho punto de giro.

Gráfico 4: Índices sintéticos. Detalle 1992-1998



Finalmente, cabe preguntar si estas propiedades dinámicas se encuentran ligadas a alguna característica estructural de los IPI. A modo de ilustración, en la tabla 2 se compara la agrupación cíclica obtenida con la correspondiente a la Clasificación Nacional de Actividades Económicas (CNAE-74).

Tabla 2: Clasificación cíclica y agrupaciones CNAE

División CNAE-74	Clasificación cíclica				
	<aa>	<a>	<c>	<r>	<rr>
1: Energía y agua	1	1	3	1	2
2: Extracción de minerales. Química	6	7	10	8	13
3: Transformación de metales	15	5	13	1	4
4: Otras manufacturas	8	6	8	10	11

En primer lugar, no existe un patrón de concentración claro en la tabla anterior, de manera que los IPI de las distintas ramas de actividad no se agrupan de manera muy señalada en alguna de las cinco categorías dinámicas. No obstante, los índices correspondientes a las industrias básicas y de otras manufacturas tienden a ser retrasados, mientras que los de las transformadoras son predominantemente coincidentes o muy adelantados. Dada la gran incertidumbre existente acerca de los fenómenos dinámicos, parece razonable emplear ambas clasificaciones con el fin de asentar el diagnóstico coyuntural sobre una base más firme, puesto que cada una de ellas aporta información sobre aspectos no contemplados en la otra clasificación.

6. Extensiones

Existen dos líneas de ampliación de este trabajo bastante inmediatas. La primera consiste en profundizar el análisis estructural esbozado al final de la sección anterior, con el objeto de comprender mejor los factores generadores de la dinámica observada. En particular, incorporar elementos relativos a la naturaleza de cada industria como, p.e., el grado en que su producción está orientada a la exportación o a la demanda interna y su especialización en la fabricación de bienes de consumo, intermedios o de capital.

La segunda línea de desarrollo consiste en integrar las conclusiones obtenidas a partir de la clasificación cíclica en un modelo econométrico de tipo VARMA. De esta manera se podrían incluir variables que reflejaran *shocks* globales (p.e., indicadores de demanda agregada) y, al mismo tiempo, examinar la dinámica sectorial interna (p.e., cómo reaccionan unas industrias frente a impulsos específicos de otras).

Así, como ilustración, si x_t representa a una variable de ámbito global y se dispone de tres índices sintéticos desfasados entre sí un período, a_t , c_t y r_t , es posible especificar el siguiente modelo VAR(1) restringido:

[14a]

$$\begin{bmatrix} x \\ a \\ c \\ r \end{bmatrix}_t = \begin{bmatrix} \emptyset_{xx} & 0 & 0 & 0 \\ \emptyset_{ax} & \emptyset_{aa} & 0 & 0 \\ \emptyset_{cx} & \emptyset_{ca} & \emptyset_{cc} & 0 \\ \emptyset_{rx} & \emptyset_{ra} & \emptyset_{rc} & \emptyset_{rr} \end{bmatrix} \begin{bmatrix} x \\ a \\ c \\ r \end{bmatrix}_{t-1} + \begin{bmatrix} \epsilon_x \\ \epsilon_a \\ \epsilon_c \\ \epsilon_r \end{bmatrix}_t$$

En notación matricial:

[14b]
$$W_t = \Phi W_{t-1} + \epsilon_t$$

La estructura triangular de Φ refleja la ordenación recursiva sugerida por la clasificación cíclica. La validación del modelo exige la comparación de las versiones restringida y no restringida por medio de un contraste de razón de verosimilitudes. Asumiendo que el modelo [14] no es rechazado por la muestra, se puede examinar su función de respuesta a los impulsos:

[15]
$$W_t = [I - \Phi B]^{-1} a_t = \Psi(B) a_t,$$

siendo a_t un vector de innovaciones estructurales obtenido a partir de la ortogonalización de los residuos del modelo VAR en forma reducida. El esquema de identificación más apropiado en este contexto es el recursivo, asociado a la descomposición de Cholesky de la matriz de varianzas y covarianzas de los residuos de [14].

Los operadores $\Psi_{i1}(B)$, con $i=2..4$, caracterizan la respuesta cíclica de las industrias de tipo i frente a un *shock* común. Los elementos $\Psi_{ij}(B)$, con $i,j=2..4$ e $i>j$, representan la respuesta de las industrias de tipo i respecto a una innovación en las de tipo j . Naturalmente, $\Psi_{ij}(B) = 0$ si $i<j$, habida cuenta de la naturaleza triangular de Φ .

Finalmente, se puede mejorar el algoritmo de simplificación de la matriz de cargas utilizando técnicas de regresión paso a paso (*stepwise*) en la línea de Bozdogan y Ramírez (1987).

Referencias

Akaike, H. (1976) "Canonical correlation of time series and the use of an information criterion", en Mehra, R.K. y Lainiotis, D.G. (Eds.) *Systems identification: advances and case studies*, Academic Press, New York, U.S.A.

Bozdogan, H. y Ramírez, D.E. (1987) "An expert model selection approach to determine the "best" pattern structure in factor analysis models", en Bozdogan, H. y Gupta, A.K. (Eds.) *Multivariate statistical modeling and data analysis*, D. Reidel Publishing Company, New York, U.S.A.

Cooper, D.M. y Wood, E.F. (1982) "Identifying multivariate time series models", *Journal of Time Series Analysis*, n. 3, p. 153-164.

Cristóbal, A., Quilis, E.M. y Melis, F. (1996) "Las series de cartera de pedidos de la industria y su clasificación dinámica", Instituto de Estudios Fiscales, Documento de Trabajo n. 27/96.

Cristóbal, A. y Quilis, E.M. (1996) "Análisis del ciclo regional mediante modelos de índices dinámicos. Aplicación a los indicadores de clima industrial", Ponencia presentada en las II Jornadas Técnicas de Estadística y Coyuntura Regional, 1-2 febrero, Santiago de Compostela.

Escribano, A. y Peña, D. (1994) "Cointegration and common factors", *Journal of Time Series Analysis*, vol. 15, n. 6, p. 577-586.

INE (1994) "Indicadores cíclicos del entorno internacional de la economía española", *Boletín Trimestral de Coyuntura*, n. 53, p. 61-121.

Mardia, K.V., Kent, J.T. y Bibby, J.M. (1979) *Multivariate analysis*, Chapman and Hall, New York, U.S.A.

Melis, M. (1985) "Caracterización y clasificación de series económicas. Aplicación al sistema de números índices de la producción industrial", Instituto Nacional de Estadística, Documento Interno.

Melis, M. (1991) "La estimación del ritmo de variación en series económicas", *Estadística Española*, n. 126, p. 7-56.

Peña, D. y Box, G.E.P. (1984) "A factor analysis for time series", SCA Corp., Working Paper n. 108.

Peña, D. y Box, G.E.P. (1987) "Identifying a simplifying structure in time series", *Journal of the American Statistical Society*, vol. 82, n. 399, p. 836-842.

Peña, D. (1990) "Cointegración y reducción de dimensionalidad en series temporales multivariantes", *Cuadernos Económicos de ICE*, n. 44, p. 109-126.

Peña, D. y Poncela, P. (1996), "Improving prediction with dynamic factor analysis", Universidad Carlos III, Documento de Trabajo n. 96-63.

Quilis, E.M. (1994) "Medida de la inflación subyacente: un análisis factorial dinámico", *Infomación Comercial Española*, n. 729, p. 101-108.

Quilis, E.M. (1995) "Apuntes de extracción de la señal en series económicas", Instituto Nacional de Estadística, Documento Interno.

Tsay, R.S. y Tiao, G.C. (1985) "Use of canonical analysis in time series model identification", *Biometrika*, vol. 72, n. 2, p. 299-315.