

Análisis de tendencias comunes y cointegración en espacio de estados

José Mondéjar Jiménez E-mail: Jose.Mondejar@uclm.es

Manuel Vargas Vargas E-mail: Manuel.Vargas@uclm.es

Universidad de Castilla-La Mancha
Facultad de Ciencias Sociales de Cuenca

Resumen: La detección de cointegración o la obtención de componentes comunes de tendencia ha sido abordada principalmente mediante la representación VARMA de procesos estocásticos, mientras que la representación en espacio de estados ha recibido menos atención en la literatura especializada, aunque presenta ventajas computacionales y analíticas que justifican su estudio. El trabajo se centrará en el análisis de series temporales no estacionarias representadas en espacio de estados. Se expondrá la justificación del método, los algoritmos desarrollados y una propuesta de modelización alternativa a los planteamientos ya clásicos como el contraste de cointegración de Johansen o la descomposición de Beveridge y Nelson.

Palabras clave: Cointegración, tendencias comunes, espacio de estados, filtro de Kalman.

Códigos JEL: C32

Abstract: Cointegration relationships or common trends detection has been undertaken mainly by VARMA representation of stochastic processes, while the specialized literature has paid less attention to the state-space modelisation, although it presents computational and analytical advantages that justify its study. This paper will focus on the state-space analysis of nonstationarity series. The justification of the method and its algorithms will be exposed and an alternative approach to the classical Johansen or Beveridge and Nelson methods is proposed.

Key Words: Cointegration, common trends, state-space models, Kalman filter.

Classification JEL: C32

Para citar este artículo puede utilizar el siguiente formato:

Mondéjar Jiménez, J. y Vargas Vargas, M. "Análisis de tendencias comunes y cointegración en espacio de estados" en *Contribuciones a la Economía*, septiembre 2006. Texto completo en <http://www.eumed.net/ce/>

1. Introducción

Con el desarrollo del análisis de series múltiples se han puesto de manifiesto deficiencias en la modelización ARMA, tales como la gran necesidad de parámetros para capturar las relaciones entre las variables o la dificultad de identificar modelos, problemas que aún no se han resuelto de forma satisfactoria. Simultáneamente y en el campo de la ingeniería, la teoría de sistemas ha desarrollado diversos algoritmos que permiten identificar y modelizar procesos estocásticos de forma distinta a como se viene haciendo en Economía Cuantitativa. Esta formulación de modelos en espacio de estados para series temporales múltiples no es desconocida en Economía; sin embargo su utilización ha estado restringida a objetivos concretos tales como el cálculo “fácil” de la función de verosimilitud o el análisis estructural.

Este paralelismo en el tratamiento de series temporales ha propiciado la aparición de consideraciones y técnicas que, aunque desarrolladas en ámbitos distintos, pueden proporcionar resultados fructíferos en cualquiera de ellos. Este hecho se ha concretado en la aparición de diversos trabajos que pretenden salvar las diferencias metodológicas y terminológicas entre ambos campos, resaltando la equivalencia básica de ambas metodologías. Así, obras como las de Moore (1981), Otter (1985) o Aoki (1983, 1990) han introducido en el campo económico conceptos y algoritmos relacionados con los modelos en espacio de estados que no eran utilizados con anterioridad y que resuelven de forma alternativa problemas básicos en el análisis de series. En algunos casos, estas soluciones son equivalentes a las que se obtienen con el tratamiento clásico mientras que en otros presentan ventajas relativas; son estos últimos casos los que justifican un estudio detallado de esta nueva metodología.

Una representación en espacio de estados fue propuesta por Akaike (1975, 1976) mediante un método de correlación canónica, con la desventaja de que puede conducir a representaciones en espacio de estados no equivalentes cuando se cambia el orden de las variables endógenas. Para evitar este inconveniente, Aoki (1983, 1990) sugiere un método

alternativo relacionado con la descomposición en valores singulares de la matriz de autocovarianzas estimadas. Las matrices de los modelos se obtienen a través de una aproximación de la matriz tipo Hankel anterior por otras de rango inferior. Posteriormente, Mittnik (1989) propone un método para la obtención de modelos en espacio de estados para series temporales múltiples, relacionado con el de Aoki pero que utiliza las matrices de autocorrelación. Este método conserva las propiedades del método anterior y además permite usar input como variables exógenas para el sistema. Este método presenta ventajas computacionales cuando se trabaja con procesos autorregresivos. Por último, destacar la propuesta de Bauer y Wagner (2002) consistente en la utilización de una regresión entre los valores futuros y pasados de la serie.

Todo proceso estocástico débilmente estacionario de rango completo y con función de densidad espectral $\Phi_y(e^{i\cdot})$ sin ceros en la circunferencia unidad tiene una realización innovacional en espacio de estados:

$$\begin{aligned} X_{t+1} &= F_t X_t + G_t \varepsilon_t \\ Y_t &= H_t X_t + D_t \varepsilon_t \quad t \in Z \end{aligned}$$

donde H_t , F_t , G_t , D_t son matrices reales de dimensiones $(p \times n)$, $(n \times n)$, $(n \times p)$ y $(p \times p)$ respectivamente. La primera ecuación del modelo recibe el nombre de ecuación de transición o estado y representa a un modelo markoviano de primer orden, donde la renovación del vector de estados consta de dos componentes; la evolución propia de X_t resumida en la matriz F y la corrección producida por la innovación a través de la matriz G . La segunda ecuación recibe el nombre de ecuación de medida u observacional. Además, si el proceso fuese invariante se describiría mediante las ecuaciones:

$$\begin{aligned} X_{t+1} &= F X_t + G \varepsilon_t \\ Y_t &= H X_t + D \varepsilon_t \end{aligned}$$

por lo que un sistema lineal, invariante y discreto está caracterizado por el conjunto de matrices (H, F, G, D) , donde el par (F, G) controla el comportamiento de la relación entre el input y el estado, (H, F) el de la relación entre el estado y el output y la matriz D la relación instantánea entre el input y el output. Esta forma del modelo puede ocasionar que cierta autocorrelación

aparezca recogida en la matriz D y no en el estado, para considerarla en éste es mejor una representación de la forma equivalente;

$$\begin{aligned}X_{t+1} &= FX_t + G\varepsilon_t \\ Y_t &= HX_t + \varepsilon_t\end{aligned}$$

En este modelo ε es un proceso innovacional de Y que no presenta autocorrelación y X_t es el vector de estado que actúa como estadístico suficiente para la dinámica del sistema. De esta forma, la mejor predicción del vector de observaciones viene dada por:

$$\hat{Y}_{t+1|t} = \mu + HX_t$$

La renovación del vector de estado va a constar de una evolución propia de X_t resumida en F y de una corrección producida por la innovación del proceso a través de G. De esta forma se resume el comportamiento dinámico en una ecuación en diferencias de primer orden. Esto permite obviar la determinación del número de retardos que presentan una correlación significativa con el valor actual ya que el vector de estado recoge toda la información relevante, por atrasada que sea. Un desarrollo completo puede verse en Vargas (1999a).

2. Estimación del submodelo de tendencia

Para modelizar una serie temporal múltiple sin componentes determinísticas¹, Y_t , en espacio de estados, el primer paso consiste en la detección y estimación del submodelo de tendencia. La primera elección que debe hacerse consiste en la dimensión del vector de tendencias o, equivalentemente, el orden de cointegración.

Siguiendo a Bauer y Wagner (2002) se puede encontrar una relación entre los valores anteriores de la forma:

¹ Como ya se ha comentado, si existen éstas, el algoritmo propuesto se aplicaría a los residuos de una regresión que elimine su efecto.

$$\begin{pmatrix} Y_t \\ Y_{t+1} \\ Y_{t+2} \\ \vdots \end{pmatrix} = \begin{pmatrix} HG & H\bar{F}G & \dots & H\bar{F}^{T-1}G \\ HFG & HF\bar{F}G & \dots & HF\bar{F}^{T-1}G \\ \vdots & \vdots & \vdots & \vdots \\ HF^J G & HF^J \bar{F}G & \dots & HF^J \bar{F}^{T-1}G \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ Y_{t-3} \\ \vdots \end{pmatrix} + \begin{pmatrix} H \\ HF \\ \vdots \\ HF^j \\ \vdots \end{pmatrix} \bar{F}^t x_0 + \\
+ \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & \dots \\ HG & 0 & 0 & \dots & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ HF^{j-1}G & HF^{j-2}G & \dots & HG & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t+1} \\ \varepsilon_{t+2} \\ \vdots \end{pmatrix} = \beta \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ Y_{t-3} \\ \vdots \end{pmatrix} + \begin{pmatrix} H \\ HF \\ \vdots \\ HF^j \\ \vdots \end{pmatrix} \bar{F}^t x_0 + \xi \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t+1} \\ \varepsilon_{t+2} \\ \vdots \end{pmatrix}$$

donde $\bar{F}^t = (F - GH)^t$ y x_0 es el valor inicial del estado². La ecuación anterior describe el comportamiento futuro del proceso como una suma de tres componentes; el primer término recoge el pasado del proceso, el segundo término recoge el efecto del valor inicial del estado x_0 y el tercero muestra el efecto de los ruidos futuros del proceso. Este último término es ortogonal a los otros dos y el último término desaparece cuando $t \rightarrow \infty$ debido a la suposición de que la matriz \bar{F}^t converge a cero.

La matriz β conecta los valores pasados con los valores futuros utilizando la información estructural del sistema, de tal forma que contiene la información relevante de las matrices del sistema (F, G, H). La idea del algoritmo propuesto por Bauer y Wagner (2002) es utilizar esta información para obtener estimaciones de las matrices del sistema.

Para llevar a cabo la estimación sólo disponemos de una serie de observaciones finita por lo que la ecuación anterior la vamos a utilizar de forma truncada. Por lo tanto, vamos a seleccionar dos índices N_f y N_p para definir los vectores truncados de observaciones futuras y pasadas como:

$$Y_{t, N_f}^+ = (Y_t^|, Y_{t+1}^|, \dots, Y_{t+N_f-1}^|)' \\
Y_{t-1, N_p}^- = (Y_{t-1}^|, Y_{t-2}^|, \dots, Y_{t-N_p}^|)'$$

y

² Si x_0 es desconocido hay que estimarlo mediante el filtro de Kalman.

$$E_{t,N_f}^+ = (\varepsilon'_t, \varepsilon'_{t+1}, \dots, \varepsilon'_{t+N_f-1})'$$

Además vamos a considerar:

$$O_f = \left(H', F'H', \dots, (F^{N_f-1})' H' \right)'$$

$$K_p = [G, (F - GH)G, \dots, (F - GH)^{N_p-1} G]$$

y una matriz ξ_f con una primera fila formada por ceros y el siguiente bloque de filas formado por $[H F^{i-2} G, \dots, H G, 0, 0]$ para $i \geq 2$. De esta forma y utilizando los vectores de valores pasados y futuros truncados podemos definir la ecuación truncada de forma compacta como:

$$Y_{t,N_f}^+ = O_f K_p Y_{t-1,N_p}^- + O_f (F - GH)^{N_p} x_{t-N_p} + \xi_f E_{t,N_f}^+$$

La anterior ecuación permite desarrollar el siguiente procedimiento:

1. En el primer paso calcular la regresión de Y_{t,N_f}^+ sobre Y_{t-1,N_p}^- para obtener una estimación $\hat{\beta}_{N_f,N_p}$ de $O_f K_p$.

2. Normalmente $\hat{\beta}_{N_f,N_p}$ es de rango completo, mientras $O_f K_p$ es de rango n para $N_p, N_f \geq n$, donde n denota el verdadero orden del sistema y N_p y N_f son los valores de truncamiento seleccionados. Así, aproximamos $\hat{\beta}_{N_f,N_p}$ por una matriz de rango n mediante la descomposición de $\hat{O}_f \hat{K}_p$.

3. Usar la estimación de \hat{K}_p para estimar el estado como $\hat{x}_t = \hat{K}_p Y_{t-1,N_p}^-$. Una vez estimado el estado, las ecuaciones del sistema pueden ser usadas para obtener estimaciones de las matrices del sistema (F, G, H) por mínimos cuadrados ordinarios: Primero establecemos la regresión de y_t en \hat{x}_t para obtener una estimación \hat{H} y unos residuos $\hat{\varepsilon}_t$. Entonces $\hat{\Psi} = E[\hat{\varepsilon}_t \hat{\varepsilon}_t']$ es una estimación de la varianza de las innovaciones. Por último, establecemos la regresión de \hat{x}_{t+1} en \hat{x}_t y $\hat{\varepsilon}_t$ para obtener las estimaciones \hat{F} y \hat{G} .

Hay que tener en cuenta que la aproximación desarrollada en el segundo paso no nos permite obtener la estimación de $\hat{\beta}_{N_f, N_p}$ directamente, sino una matriz ponderada $\hat{W}_f^+ \hat{\beta}_{N_f, N_p} \hat{W}_p^-$. Si llevamos a cabo la descomposición en valores singulares de esta matriz vamos a tener que $\hat{W}_f^+ \hat{\beta}_{N_f, N_p} \hat{W}_p^- = \hat{U} \hat{\Sigma}_1 \hat{V}'$ donde $\hat{\Sigma}_1$ ³ es una matriz diagonal que contiene los valores singulares ordenados de forma decreciente y $\hat{U} \hat{U}'$ y $\hat{V} \hat{V}'$ son matrices identidad. Para un sistema de orden n vamos a tener exactamente n valores singulares mayores que cero, el resto de valores singulares van a converger a cero.

El orden de estimación en el trabajo de Bauer (1998) va a estar basado en la consideración del tamaño del primer valor singular rechazado, $\hat{\sigma}_{n+1}$, explotando el comportamiento asintótico de las estimaciones va a definir el siguiente criterio:

$$SVC(\hat{n}) = \hat{\sigma}_{\hat{n}+1}^2 + 2\hat{n}p \frac{H_T}{T}$$

donde $H_T > 0$, $H_T/T \rightarrow 0$ denota una cota⁴ que determina las propiedades asintóticas del orden del sistema estimado. El número de parámetros en un modelo de dimensión n es igual a $2np$. El orden del sistema estimado, \hat{n} , es considerado como aquel valor que hace mínimo el criterio $SVC(n)$.

Un planteamiento alternativo al anterior parte del:

$$(I - L) Y_t = H(L) \varepsilon_t$$

Siendo ε_t el proceso de innovación de la serie, se tiene que⁵:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \varepsilon_t \rightarrow W(r) \quad , \quad r \in [0, 1]$$

donde por $[Tr]$ se indica la parte entera de (Tr) y $W(r)$ es un proceso Browiano. De aquí, se puede probar que:

³ Hemos utilizado la notación $\hat{\Sigma}_1$ para diferenciar esta matriz de valores singulares de la obtenida mediante la descomposición de la matriz de Hankel $\hat{\Sigma}$ que vamos a analizar posteriormente y que constituye el elemento fundamental que vamos a utilizar en la determinación de la dimensión del sistema en espacio de los estados.

⁴ En Bauer y Wagner (2002) se propone utilizar un valor para la cota $H_T = \log(T)$.

⁵ Ver, por ejemplo, Gouriéroux y Monfort (1997) o Phillips y Durlauf (1986).

$$T^{-2} \sum_{t=1}^T Y_t Y_t' \rightarrow H(l) \left(\int_0^l W(r) W'(r) dr \right) H(l)'$$

Así, tomando un vector β fuera del subespacio generador de cointegración, se obtiene que:

$$T^{-2} \sum_{t=1}^T (\beta' Y_t)^2 \rightarrow \beta' H(l) \left(\int_0^l W(r) W'(r) dr \right) H(l)' \beta \neq 0$$

por lo que $T^{-1} \sum_{t=1}^T (\beta' Y_t)^2 \rightarrow E[(\beta' Y_t)^2]$ es, al menos, de orden T.

Por otro lado, considerando un vector α de cointegración, se tiene que:

$$T^{-2} \sum_{t=1}^T (\alpha' Y_t)^2 \rightarrow \alpha' H(l) \left(\int_0^l W(r) W'(r) dr \right) H(l)' \alpha = 0$$

con lo que $T^{-1} \sum_{t=1}^T (\alpha' Y_t)^2 \rightarrow E[(\alpha' Y_t)^2]$ es de orden uno.

Considerando ahora la matriz $T^{-1} \Gamma_0 = \frac{T^{-2} \sum_{t=1}^T Y_t Y_t'}{\text{Var}(Y_t)}$, que no es más que la matriz de autocorrelaciones dividida por T, y los autovalores de ésta:

$$\lambda_{k,T} = \min_{\mathcal{E}_{p-k+1}} \max_{x \in \mathcal{E}_{p-k+1}} \frac{x' T^{-1} \Gamma_0 x}{x' x} = \max_{\mathcal{E}_k} \min_{x \in \mathcal{E}_k} \frac{x' T^{-1} \Gamma_0 x}{x' x}$$

con $\lambda_{1,T} \geq \lambda_{2,T} \geq \dots \geq \lambda_{p,T}$, en la definición anterior se observa que, si existe cointegración de orden r, debe haber exactamente r autovalores convergentes a cero, mientras que los (p-r) primeros convergerán a constantes no nulas, por lo que $T^{-1} \Gamma_0$ tenderá a una matriz no nula (ver, por ejemplo, Phillips (1987) o Aoki (1990)).

Definiendo $\gamma_l = \Gamma_l - \Gamma_0$ donde $\Gamma_l = \frac{T^{-1} \sum_{t=1}^T Y_t Y_{t-l}'}{\text{Var}(Y_t)}$, se tiene probado (Phillips y Durlauf (1986) o Aoki (1990) que γ_l converge débilmente a una matriz aleatoria. Como $\Gamma_l = \Gamma_0 + \gamma_l$, dividiendo por T se tiene que la diferencia entre $T^{-1} \Gamma_l$ y $T^{-1} \Gamma_0$ es de orden $O_p(T^{-1})$, por lo que

también en $T^{-1} \Gamma_I$ debe haber exactamente r autovalores cercanos a cero, permaneciendo el resto en constantes no nulas. Con este resultado, se puede afirmar que la matriz de autocorrelaciones de primer orden, Γ_I , sólo presenta r autovalores (o valores singulares) de orden $O_p(I)^6$ y $(p-r)$ de orden $O_p(T)$, no existiendo valores propios con orden de crecimiento superior a T .

Con estos resultados es posible plantear la determinación del orden de cointegración y, posteriormente, el tratamiento bietápico de modelización esbozado en el epígrafe anterior.

Como ya se comentaba, el primer objetivo es estimar un modelo de la forma dada en:

$$\tau_{t+1} = \Lambda \tau_t + G_\tau Y_t^*$$

$$Y_t = H_\tau \tau_t + Y_t^*$$

que resuma la componente estocástica no estacionaria de \mathbf{Y}_t y que, aparte del interés intrínseco de su estudio, permita un posterior análisis de la serie estacionaria \mathbf{Y}_t^* resultante de la eliminación de la tendencia. Para ello, hay que determinar la dimensión del vector τ_t , es decir, el número de factores responsables de la no estacionariedad; sin embargo, el algoritmo de identificación no puede basarse en el análisis de la magnitud de los valores singulares de la matriz de autocorrelación de primer orden, ya que algunos de éstos no están acotados cuando la serie es no estacionaria. Como se ha probado que el orden de éstos es, a lo sumo, $O_p(T)$, definiendo los cocientes de valores singulares entre el primero (y mayor) de ellos:

$$\delta_i = \frac{\lambda_i}{\lambda_1} \quad , \quad \text{para } i = 1, 2, \dots, p$$

se tiene que $\delta_1 = 1$ y $0 \leq \delta_i < 1$ para el resto de cocientes, por lo que no presentan el problema de no acotación antes comentado. Además, si existe cointegración de orden r , se tendrá que los r últimos cocientes $(\delta_{p-r+1}, \delta_{p-r+2}, \dots, \delta_p)$ convergerán rápidamente a cero en función del tamaño muestral, convergiendo los $(p-r)$ restantes a cantidades no nulas.

⁶ Justo los correspondientes a los r nulos en la matriz $T^{-1} \Gamma_I$.

Concretamente, dada la única realización disponible de la matriz de tipo Hankel dada por $\hat{H} = \hat{\Gamma}_I$ y la serie de cocientes de sus valores singulares $(1, \delta_2, \delta_3, \dots, \delta_p)$, el orden de cointegración coincide con el número de cocientes que se consideren no nulos. La determinación de tal orden se lleva a cabo con un planteamiento de contrastación secuencial del orden de cointegración, semejante al de Johansen. Específicamente, se trata de contrastes estadísticos donde la hipótesis nula establece la positividad de los cocientes de valores singulares, mientras que la hipótesis alternativa consiste en la nulidad de algunos de tales cocientes, es decir, la existencia de un número de relaciones de tendencia común inferior al de series. Al no disponerse de las distribuciones teóricas de estos cocientes, la región de rechazo de los contrastes ha de determinarse utilizando la distribución empírica obtenida mediante simulaciones, que permite disponer de los valores críticos de los cocientes δ_i a los niveles de significación habituales, como se recoge en Vargas (1999a).

Este planteamiento se asemeja en algunos puntos a la utilización de estimadores de componentes principales tal y como se hace en Harris (1997). Sin embargo el uso de cocientes de valores singulares permite resolver en parte las limitaciones de estos estimadores. Específicamente, la determinación de autovalores se ve dificultada por la existencia de errores numéricos y la sensibilidad de los determinantes a éstos, efecto que puede invalidar la aplicación práctica del método. Por ello, dada la matriz de tipo Hankel formada por la función de autocorrelación muestral de primer orden, $\hat{H} = \hat{\Gamma}_I$, y su descomposición en valores singulares, el análisis de los cocientes permite el establecimiento del orden del vector de tendencias. Una vez obtenidas todas las matrices de este modelo, el filtrado de Kalman proporciona la serie de tendencias, τ_t , y la de innovaciones, Y_t^* .

3. Estimación de modelos de tendencia común para procesos no estacionarios

Con los resultados dados en los epígrafes precedentes, se está en disposición de plantear de manera ordenada y completa la estimación de modelos en espacio de estados para procesos no estacionarios. Este trabajo se centrará en la detección y estimación de la componente no

estacionaria del proceso, no abordando la modelización del submodelo estacionario, que puede verse en Vargas (1999a), Bauer y Wagner (2002) o Mondéjar (2006).

Así, sea Y_t un proceso estocástico p -dimensional que se supone no estacionario. El primer paso consiste en la determinación del tipo de tendencia existente en las series, ya que el efecto de las tendencias determinísticas no es idéntico al de las estocásticas. En este sentido, los tests clásicos de Nelson y Plosser permiten diferenciar entre ambos tipos, por lo que las tendencias determinísticas pueden ser eliminadas mediante regresión, utilizando los errores de ésta como inputs para la representación en espacio de estados⁷. Con este planteamiento, se estaría tratando un modelo general de la forma:

$$\begin{aligned} Y_t &= \beta Z_t + \xi_t \\ X_{t+1} &= F X_t + G \varepsilon_t \quad t \in Z \\ \xi_t &= \mu_\xi + H X_t + \varepsilon_t \end{aligned}$$

donde en Z_t se recogen los regresores necesarios y μ_ξ representa un nivel general del proceso⁸. Los residuos de esta regresión están ya libres de tendencias determinísticas, por lo que la única fuente de no estacionariedad es la estocástica. Para estudiar el orden de crecimiento de las componentes de ξ_t o, equivalentemente, del número de raíces unitarias existentes en cada serie individual, se aplican los tests de raíces unitarias de Dickey-Fuller o Phillips y Perron, que proporcionan el orden de integrabilidad de cada serie. Dada la práctica habitual en Economía Cuantitativa, las componentes del proceso serán $I(0)$ o bien $I(1)$. Sin embargo, no se va a suponer que todas las componentes tienen el mismo orden. De existir algunas que sean estacionarias, el algoritmo expuesto debe excluirlas de la primera fase, apareciendo ceros en la fila correspondiente de la matriz H_τ .

Una vez determinada la posibilidad de que existan tendencias comunes, se ha de estimar un modelo para éstas de la forma:

⁷ Aquí se está englobando también el caso de tendencias estocásticas con deriva, ya que el efecto de ésta es asintóticamente equivalente al de una tendencia determinística. Por ello, un tratamiento incorrecto de ésta puede producir deficiencias como las que se señalan más adelante.

⁸ Este nivel general coincidirá muchas veces con la media no condicional del proceso. Sin embargo, existen situaciones donde esta elección produce pérdida de precisión en las predicciones, siendo más interesante su estimación.

$$\tau_{t+1} = \Lambda \tau_t + G_\tau \xi_t^*$$

$$\xi_t = \mu_\tau + H_\tau \tau_t + \xi_t^*$$

donde por ξ_t se está representando a las series bajo estudio tras la eliminación de las posibles tendencias determinísticas, bien sean las asociadas a funciones deterministas del tiempo bien las asociadas a derivas en las tendencias estocásticas⁹.

El primer paso consiste en la determinación del número de tendencias comunes, es decir, la dimensión del vector τ_t . Como se ha descrito, la DVS de la matriz de autocorrelación de primer orden:

$$E[\xi_t \xi_{t-1}] = \hat{\Gamma}_1 = \hat{H} = \hat{U} \hat{\Sigma} \hat{V}'$$

permite determinar dicha dimensión decidiendo cuántos de los valores singulares recogidos en la matriz $\hat{\Sigma}$ son divergentes, es decir, cuántos de los cocientes δ_i son significativamente distintos de cero.

Dada la ordenación decreciente de los cocientes, esta decisión se lleva a cabo con contrastes de hipótesis secuenciales, donde el primero adopta la forma:

$$H_0 \equiv \delta_2 > 0$$

$$H_1 \equiv \delta_2 = 0$$

Si se rechaza la hipótesis nula, se considera que existe una única componente de tendencia común que relaciona a las componentes no estacionarias del proceso ξ_t ¹⁰ ya que sólo el primer valor singular se considera divergente. Si, por contra, no se rechaza la hipótesis nula, se considera que existen, al menos, dos valores singulares no convergentes, por lo que existen como mínimo dos tendencias comunes. En este caso, ante la posibilidad de más relaciones cointegrantes, se deben plantear contrastes consecutivos sobre la nulidad de cocientes de la forma:

⁹ En este último caso, la regresión recogida en la ecuación anterior tiene como única finalidad la supresión del efecto de las derivas por lo que, tras la determinación del número de relaciones de cointegración, el algoritmo ha de utilizar los valores originales.

¹⁰ Si alguna de las componentes fuese estacionaria, no aparecería en dicha combinación lineal.

$$H_0 \equiv \delta_j > 0$$

$$H_1 \equiv \delta_j = 0$$

hasta que se rechace la hipótesis nula o se llegue al caso $j = p$. Se tendrá entonces una estimación del número de relaciones de cointegración o, equivalentemente, del número de tendencias linealmente independientes que existen en la serie ξ_t .

Para la realización de los contrastes anteriores se han propuesto los valores críticos de los cocientes δ_i a los niveles habituales recogidos en el anexo (tablas 1 a 4), por lo que se reiterarán los contrastes anteriores hasta que un cociente sea inferior al correspondiente valor crítico.

Una vez establecida la dimensión, k , del vector de tendencias, para estimar las matrices del sistema se utiliza la descomposición junto a los resultados recogidos en las ecuaciones. Así, definiendo las matrices:

$$\begin{aligned}\hat{O} &= \hat{U}\hat{\Sigma}^{1/2}, \quad \hat{h} = \hat{\Sigma}^{1/2}\hat{V}', \quad H' = \hat{\Gamma}_2 \\ \hat{O}^- &= \hat{\Sigma}^{-1/2}\hat{U}', \quad \hat{h}^- = \hat{V}\hat{\Sigma}^{-1/2}\end{aligned}$$

es posible obtener la estimación de la matriz H_τ como las primeras k filas de la matriz \hat{O} , y la matriz dinámica como:

$$\hat{\Lambda} = \hat{O}^- H^\wedge \hat{h}^-$$

Utilizando ahora la secuencia auxiliar de estados:

$$\hat{X}_t = Y R^{-1} \xi_{t-1}$$

y la matriz \hat{H}_τ se obtiene la secuencia de innovaciones auxiliares:

$$\hat{e}_t = \xi_t - \hat{H}_\tau \hat{X}_t$$

Por último, una adaptación de las ecuaciones proporciona la estimación de la matriz de ganancia del modelo:

$$\hat{G}_\tau = E \left[\left(\hat{X}_{t+1} - \hat{\Lambda} \hat{X}_t \right) \hat{e}_t' \right] E \left[\hat{e}_t \hat{e}_t' \right]^{-1}$$

completando la estimación de las matrices del subsistema de tendencias.

Tal y como ya se comentaba en el epígrafe, la secuencia de vectores de estado \hat{X}_t , no constituye una estimación eficiente de τ_t , por lo que es necesario recurrir al filtrado de Kalman para obtener la secuencia de vectores de tendencias junto a la del proceso estacionario ξ_t^* asociado. Este algoritmo recursivo consta de tres etapas básicas: el establecimiento de las condiciones iniciales; la actualización de la información; y la predicción, que sirve como condición inicial para continuar la recursión.

Un aspecto importante del algoritmo de filtrado de Kalman es la determinación del estado inicial y su varianza junto con la influencia que ésta pueda tener sobre aspectos claves como la estabilidad o convergencia de las iteraciones. Sin embargo, cuando se trabaja con procesos estacionarios, esta cuestión queda en segundo término en la mayoría de las aplicaciones empíricas ya que suele ser habitual considerar que $\hat{X}(1|0)=0$ y obtener su varianza particularizando la ecuación a la expresión:

$$P(1) = F P(1) F' + G \Psi G'$$

Esta práctica está avalada por las propiedades de convergencia y estabilidad del filtrado de Kalman. Sin embargo, cuando la amplitud muestral no es grande, puede ocurrir que la consideración de nulidad para el valor inicial del estado provoque un mal ajuste del algoritmo de Kalman mientras que una buena especificación mejore sensiblemente las estimaciones sucesivas. Es en tales casos donde tiene interés el estudio de fórmulas que determinen el Amejor@ valor inicial para el estado¹¹. Igualmente, puede ser fructífero no especificar un único valor para iniciar el filtrado; en muchas aplicaciones, entre las que cabe destacar los modelos económicos no estacionarios o los modelos estructurales, es más realista asumir que las condiciones iniciales son parcialmente difusas, por lo que se introduce una perturbación aleatoria en el estado inicial. Se origina así el llamado filtrado difuso de Kalman (ver, por ejemplo, Kohn y Ansley (1989) o Kitagawa y Gersch (1984) para un análisis detallado de este algoritmo). En la misma línea, se puede considerar toda una distribución para el estado inicial (usualmente normal) que, unida a la

¹¹ Obviamente, el criterio para elegir el punto inicial del vector de estado es el de minimizar la norma de la matriz de varianza $P(t)$ para los sucesivos valores de t .

normalidad asumida para la innovación, puede convertir el filtrado de Kalman en un algoritmo de actualización de los momentos de la distribución.

4. Bibliografía

Akaike, H. (1975): “Markovian representation of stochastic processes by canonical variables”, *SIAM J. Control Optim*, nº 13, pp. 162-173.

Akaike, H. (1976): *Canonical correlation analysis of time series and the use of an information criterion*. En *System Identification and Case Studies* (Eds. R. Mehra y D. Lainiotis). Academic Press, New York.

Aoki, M. (1983): *Notes on Economic Time Series Analysis: System Theoretic Approach*. Heidelberg, Springer-Verlag.

Aoki, M. (1990): *State Space Modeling of Time Series*. Second, Revised and Enlarged Edition. Nueva York, Springer-Verlag.

Aoki, M. y Havenner, A. M. (1991): “State Space modeling of multiple time series”, *Econometrics Reviews*, 10, pp. 1-99.

Banerjee, A., Dolado, J., Galbraith, J.W. y Hendry, D.F. (1993): *Cointegration, Error-Correction and the Econometric Analysis of Non-stationary Data*. Oxford University Press.

Bauer, D. (1998): *Some asymptotic theory for the estimation of linear systems using maximum likelihood methods or subspace algorithms*, Thesis proposal, Technischen Universität Wien. Disponible en <http://www.eos.tuwien.ac.at/Oeko/Teaching/Theses/dissbauer.pdf>

Bauer, D. (2005): “Comparing the CCA subspace method to pseudo maximum likelihood methods in the case of no exogenous input”, *Journal of Time Series Analysis*, 26, pp: 631-648.

Bauer, D. y Wagner, M. (2002): “Estimating cointegrated systems using subspace algorithms”, *Journal of Econometrics*, 111, pp: 47-84.

Bauer, D. y Wagner, M. (2003): “A canonical form for unit root processes in the state space framework”, *Diskussionsschriften* 03-12. Universität Bern. Disponible en <http://www.vwi.unibe.ch/publikationen/download/dp0312.pdf>

Clar, M., Ramos, R. y Suriñach, J. (1998): “Potencialidad de la modelización state-space y el filtro de kalman para el análisis regional. Una aplicación para el índice de actividad industrial”, *Documents de treball de la divisió de ciències jurídiques econòmiques i socials. Col·lecció d’Economia*. Departament d’Econometria, Estadística i Economia Espanyola. Universitat de Barcelona.

Clements, M. y Hendry, D. (1999): *Forecasting non-stationary economic time series modeling trends and cycles in economic time series*, The Massachusetts Institute of Technology Press, Cambridge.

De Jong, D.N. (1992): “Cointegration and Trend-Stationarity in Macroeconomic Time Series”. *Journal of Econometrics*, 52, pp. 347-370.

De Jong, P. y Chu-Chun-Lin, S. (1994): “Stationary and non-stationary State Space Models”. *Journal of Time Series Analysis*, 15, 2, pp. 151-166.

Dickey, D. A. y Fuller, W. A. (1979): “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”, *Journal of the American Statistical Association*, 74, pp. 427-431.

Dickey, D. A. y Fuller, W. A. (1981): “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root”, *Econometrica*, 49, pp. 1057-1072.

Dickey, D. A. y Fuller, W. A. (1984): “Testing for unit roots in seasonal time series”, *Journal of the American Statistical Association*, 79, pp. 355-367.

Durbin, J. y Koopman, S. J. (2001): *Time series Analysis by state space methods*, Oxford University Press. New York.

Engle, R.F. y Granger, C.W.J. (1987): “Cointegration and error correction: Representation, estimation and testing”. *Econometrica*, 55, pp. 251-276.

Gourieroux, C. y Monfort, A. (1997): *Time Series and Dynamic Models*. Cambridge University Press.

Hamilton, J. D. (1994): “State-space models”, *Handbook of Econometrics*, 4, pp. 3039-3080.

Harris, D. (1997): “Principal components analysis of cointegrated time series”. *Econometric Theory*, 13, pp. 529-557.

Harvey, A. C., Koopman S.J. y Shephard, N. (2004): *State space and unobserved component models*. Cambridge University Press, Cambridge.

Havener, A.M. (1997): “A guide to State Space Modeling of Multiple Time Series”. En M. Aoki y A.M. Havener (Eds.) *Applications of Computer Aided Time Series Modeling*. Lecture Notes in Statistics, 119. New-York, Springer-Verlag.

Havener, A.M. y Aoki, M. (1990): “*Deterministic and Stochastic Trends in State Space Models of Nonstationary Time Series*”. University of California, Working Paper 90-9.

Johansen, S. (1988): “Statistical Analysis of Cointegration Vectors”. *Journal of Economic Dynamics and Control*, 12, pp. 231-254.

Johansen, S. (1991): “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Model”. *Econometrica*, 59, pp. 1551-1580.

Johansen, S. (1992): “Determination of Cointegration Rank in the Presence of a Linear Trend”. *Oxford Bulletin of Economics and Statistics*, 54, pp. 383-397.

Johansen, S. (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, London.

Johansen, S. y Juselius, K. (1990): “Maximum likelihood estimation and inference in cointegration, with applications to the demand for money”. *Oxford Bulletin of Economics and Statistics*, 52(2), pp. 170-210.

Johansen, S. y Juselius, K. (1992): “Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the Purchasing Power Parity and the Uncovered Interest Parity for the UK”. *Journal of Econometrics*, 53, pp. 211-244.

Kitagawa, G. y Gersch, W. (1984): “A smoothness priors-state space modeling of time series with trend and seasonality”. *Journal of American Statistical Association*, 79, pp. 378-389.

Kohn, R. y Ansley, C. F. (1989): “Filtering and smoothing algorithms for state space models”, *Computer Mathematical Applications*, 18, pp. 515-528.

Lütkepohl, H. (2005): *New introduction to multiple time series analysis*. Springer Verlag, New York.

Mondéjar, J. (2006): Análisis cuantitativo de la coyuntura económica: Una aplicación de la representación en espacio de estados de series temporales múltiples. Tesis Doctoral, Universidad de Castilla-La Mancha.

Moore, C.B. (1981): "Principal component analysis in linear systems: controllability, observability and model reduction". *IEEE Automatic Control*, 26, pp. 17-32.

Otter, P.W. (1985): *Dynamic Feature Space Modelling, Filtering and Self-Tuning Control of Stochastic Systems*. Springer, Berlin.

Phillips, P.C.B. (1987): "Time series regressions with a unit root". *Econometrica*, 55, pp. 277-301.

Phillips, P.C.B. (1991): "Optimal inference in cointegrated systems". *Econometrica*, 59, pp. 283-306.

Phillips, P.C.B. y Durlauf, S. (1986): "Multiple time series regression with integrated processes". *Review of Economic Studies*, 53, pp. 473-495.

Phillips, P.C.B. y Perron, P. (1988): "Testing for a unit root in time series regression". *Biometrika*, 75, pp. 335-346.

Stock, J.H. y Watson, M.W. (1993): "A Simple Estimator of Cointegrating Vectors in Higher Order integrated Systems". *Econometrica*, 61, pp. 783-820.

Vargas, M. (1999a): Modelización de series temporales múltiples en espacio de estados. Análisis de procesos no estacionarios y cointegración. Tesis Doctoral, Universidad de Castilla-La Mancha.

Vargas, M. (1999b): "Modelización de series temporales estacionarias en espacio de estados". Documento de Trabajo 2/1999/4. Universidad de Castilla-La Mancha. Disponible en http://www.uclm.es/AB/fcee/D_trabajos/2-99-4.pdf

5. Anexo

Tabla 1: Cuantiles para el cociente de valores singulares en procesos de dimensión dos.

	Nivel de significación					
Tamaño muestral	0,01	0,05	0,1	0,9	0,95	0,99
15	0,00177	0,01108	0,02616	0,63194	0,75581	0,88629
25	0,01686	0,03712	0,06231	0,72416	0,80194	0,91466
50	0,03173	0,06446	0,08705	0,78093	0,85007	0,93457
75	0,03886	0,06681	0,09508	0,81525	0,87323	0,93762
100	0,045	0,08139	0,12129	0,83148	0,88888	0,94299

Tabla 2: Cuantiles para el cociente de valores singulares en procesos de dimensión tres.

		Nivel de significación					
Tamaño muestral	Cociente	0,01	0,05	0,1	0,9	0,95	0,99
15	δ_2	0,02934	0,06543	0,09598	0,66349	0,75785	0,87961
	δ_3	0,00063	0,00277	0,00529	0,16693	0,23735	0,42611
25	δ_2	0,04989	0,09285	0,13438	0,67557	0,75256	0,86131
	δ_3	0,00129	0,00626	0,01441	0,24579	0,31252	0,46558
50	δ_2	0,05776	0,10497	0,15166	0,71707	0,77867	0,89175
	δ_3	0,00956	0,02085	0,03208	0,33674	0,42895	0,56349
75	δ_2	0,06768	0,11602	0,17320	0,71415	0,79022	0,87472
	δ_3	0,01376	0,02582	0,03549	0,35776	0,44032	0,62173
100	δ_2	0,06803	0,11396	0,16423	0,72712	0,79463	0,89858
	δ_3	0,02032	0,03518	0,04572	0,37132	0,46618	0,62533

Tabla 3: Cuantiles para el cociente de valores singulares en procesos de dimensión cuatro.

Tamaño muestral	Cociente	Nivel de significación					
		0,01	0,05	0,1	0,9	0,95	0,99
15	δ_2	0,06317	0,10382	0,13928	0,67841	0,76043	0,90102
	δ_3	0,01154	0,02294	0,03141	0,25516	0,31460	0,43414
	δ_4	0,00027	0,00134	0,00251	0,06524	0,08702	0,15325
25	δ_2	0,06877	0,13244	0,17504	0,68958	0,76551	0,88093
	δ_3	0,01634	0,02925	0,04219	0,31097	0,37875	0,55044
	δ_4	0,00049	0,00189	0,00342	0,09337	0,12899	0,21505
50	δ_2	0,09311	0,14539	0,20644	0,70496	0,79339	0,91693
	δ_3	0,02618	0,04099	0,05814	0,35522	0,42515	0,55739
	δ_4	0,00558	0,01055	0,01493	0,13191	0,16626	0,28233
75	δ_2	0,07899	0,15625	0,20431	0,69909	0,76565	0,89213
	δ_3	0,03012	0,04547	0,06396	0,39386	0,45353	0,55975
	δ_4	0,00794	0,01615	0,02019	0,14575	0,18512	0,29920
100	δ_2	0,09783	0,16462	0,22156	0,71819	0,79163	0,90893
	δ_3	0,03064	0,05306	0,07194	0,37811	0,45165	0,56583
	δ_4	0,00983	0,01631	0,02230	0,16586	0,21511	0,32968

Tabla 4: Cuantiles para el cociente de valores singulares en procesos de dimensión cinco.

Tamaño muestral	Nivel de significación						
	Cociente	0,01	0,05	0,1	0,9	0,95	0,99
15	δ_2	0,09651	0,13052	0,17251	0,66727	0,75271	0,90243
	δ_3	0,02003	0,04173	0,05203	0,28884	0,35720	0,48576
	δ_4	0,00636	0,01174	0,01589	0,12149	0,15205	0,24480
	δ_5	0,00016	0,00083	0,00165	0,03256	0,05005	0,08652
25	δ_2	0,08740	0,15008	0,19010	0,67448	0,75435	0,87352
	δ_3	0,02659	0,04574	0,05736	0,30289	0,35721	0,49778
	δ_4	0,00501	0,01075	0,01566	0,12833	0,16438	0,24834
	δ_5	0,00011	0,00059	0,00109	0,03476	0,04814	0,09305
50	δ_2	0,11839	0,19265	0,22934	0,70564	0,80186	0,90681
	δ_3	0,03838	0,06222	0,07739	0,36978	0,43096	0,55134
	δ_4	0,01549	0,02332	0,02879	0,17641	0,21588	0,30211
	δ_5	0,00212	0,00542	0,00751	0,06212	0,08205	0,12346
75	δ_2	0,11598	0,18624	0,23602	0,70026	0,77433	0,87953
	δ_3	0,03959	0,06257	0,08137	0,39610	0,44810	0,55640
	δ_4	0,01621	0,02482	0,03326	0,18307	0,23144	0,32486
	δ_5	0,00516	0,00838	0,01078	0,07549	0,09541	0,15857
100	δ_2	0,12158	0,19365	0,23837	0,70631	0,77541	0,89723
	δ_3	0,04294	0,07562	0,09277	0,39518	0,45436	0,57502
	δ_4	0,01806	0,02812	0,03678	0,19647	0,23352	0,33582
	δ_5	0,00732	0,01018	0,01341	0,08103	0,10696	0,16523