TRABAJOS DE ESTADISTICA
Vol. 3. Núm. 2. 1988 pp 157 a 166

REGRESION ESPECTRAL SESGADA

TUSELL, FERNANDO
Departamento de Estadística Teórica
Facultad de CC.EE. y Empresariales
Universidad del País Vasco

RESUMEN

Se propone un método de regresión espectral adecuado cuando los regresores tienen potencia espectral despreciable sobre bandas de frecuencia estrechas. Se investiga la relación entre el método propuesto y los procedimientos de regresión sesgada habituales en el dominio del tiempo.

Palabras clave: Regresión espectral, suavizado espectral, regresión sesgada, desestacionalización.
Clasificación AMS: Primaria, 62M. Secundaria, 60G.

SUMMARY

A spectral regresion method for dealing with regressors having negligible spectral density over narrow frequency bands is proposed. Its relationship with biased regression methods in the time domain is investigated.

Key words: Spectral regression, biased regression, spectral smoothing, seasonal adjustment.
AMS Clasiffication: Primary, 62M. Secundary, 60G.
Title: Biased spectral regression.

Aceptado, junio 1988.

1. INTRODUCCION

La aproximación espectral a la estimación de modelos de regresión lineal dista de ser nueva. Aunque escasamente empleada en la práctica econométrica habitual, tiene su origen en artículos aparecidos ya en la década de los sesenta, entre los que cabe citar Hannan (1963), Hannan (1965) y Hannan (1967). Con posterioridad han aparecido trabajos que clarifican la relación entre regresión espectral y regresión mínimo cuadrática ordinaria o mínimo cuadrática generalizada, como Engle (1974) y textos como Fishman (1969) o Dhrymes (1971).

El objetivo del presente trabajo es doble. Por una parte, se muestra que el mero hecho de emplear estimadores espectrales consistentes (como periodogramas suavizados) en las ecuaciones normales de la regresión espectral supone la realización implícita de una operación similar a la que da origen al estimador «ridge» (véase Hoerl-Kennard (1970), por ejemplo) en el dominio del tiempo.

Por otra parte, proponemos emplear en regresión espectral un estimador similar en espíritu al de raíces latentes, ideado en el dominio del tiempo por Webster et al. (1973).

2. NOTACION

Consideramos en lo que sigue el caso particular en que se estima una ecuación de regresión con retardos de $Y(t)$ sobre $X(t)$, de la forma:

$$Y(t) = \sum_{j=0}^{L} a_j X(t-j) + n(t) \tag{1}$$

siendo $n(t)$, $X(t)$ e $Y(t)$ sucesiones aleatorias estacionarias, ergódicas y sin componente singular. La sucesión $n(t)$ de ruido no es necesariamente blanca, pero sí suponemos que $X(t)$ y $n(t)$ están incorreladas. Sin pérdida de generalidad y en aras de la simplicidad notacional admitimos también que todas las sucesiones intervinientes tienen media nula.

Sea U una matriz unitaria cuyo elemento genérico hk-ésimo es:

$$u_{hk} = T^{-1/2} \exp\{i\lambda_h(k - T/2)\} \tag{2}$$

donde

$$\begin{aligned} \lambda_h &= 2\pi T^{-1}(h - T/2) \\ T &= \text{longitud de la realización disponible} \\ p &= T/2 \\ h, k &= -p + 1, ..., -1, 0, 1, ..., p \end{aligned}$$

Sea C_{nn} la matriz de covarianzas de $n(t)' = (n(1) ... n(T))$. Entonces, los elementos de la matriz

$$V_{nn} = UC_{nn}U^* \tag{3}$$

donde U^* denota la traspuesta complejo-conjugada de U, son:

$$v_{hl} = T^{-1} \sum_{j=-p+1}^{p} \sum_{k=-p+1}^{p} C_{nn}(k - j) \exp\{i(\lambda_h j - \lambda_l k)\} \tag{4}$$

donde $C_{nn}(k - j)$ denota la autocorrelación de retardo $(k - j)$ del proceso de ruido (y dada la estructura de Toeplitz de la matriz C_{nn} es también su elemento kj-ésimo).

Cuando $T \to \infty$, puede demostrarse (Fishman (1969), pág. 153) que:

$$\lim_{T \to \infty} v_{jk} = \begin{cases} 2\pi f_{nn}(\lambda_j) & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases} \tag{5}$$

Es decir, para T grande $f_{nn}(\lambda_h)$, $h = -p + 1, ..., p$ son proporcionales a los valores propios de la matriz de covarianzas C_{nn}. Un razonamiento de tipo intuitivo en apoyo de (5) puede verse en Whittle (1952) y, en forma modificada, en Hannan (1960).

Como indica el primero de los dos autores citados, la aproximación espectral al problema de la regresión lineal se reduce a explotar la diagonalización aproximada en (3)(5).

3. REGRESION ESPECTRAL Y REGRESION RIDGE GENERALIZADA

Teniendo en cuenta la igualdad aproximada (5) resulta que la minimización de $n(t)'C_{nn}^{-1}n(t)$ (suma de cuadrados generalizados de los residuos) es asintóticamente equivalente a la minimización de:

$$n(t)'U^*V^{-1}Un(t) = \sum_{j=-p+1}^{p} \frac{I_{nn}(\lambda_j)}{f_{nn}(\lambda_j)}$$

Por otra parte, tenemos:

$$n(t) = Y(t) - \sum_{j=0}^{L} a_j X(t-j)$$

y salvo «efectos final» (que son de orden $O(1/T)$ y, por tanto, despreciables cuando $T \to \infty$; véase Hannan (1960), pág. 55), se verifica:

$$I_{nn}(\lambda_j) \approx I_{yy}(\lambda j) + \sum_k \sum_s a_k a_s I_{xx}(\lambda_j) \exp\{i\lambda_j(k-s)\} -$$
$$- 2\sum_s a_s I_{xy}(\lambda_j) \exp\{(i\lambda_j s)\}$$

En consecuencia, se trata de minimizar:

$$\sum_{j=-p+1}^{p} \frac{I_{yy}(\lambda_j)}{f_{nn}(\lambda_j)} + \sum_{k=1}^{L} \sum_{l=1}^{L} a_k a_l \frac{I_{xx}(\lambda_j)}{f_{nn}(\lambda_j)} \exp\{i\lambda_j(k-l)\} -$$
$$-2\sum_{s=0}^{L} a_s \frac{I_{xy}(\lambda_j)}{f_{nn}(\lambda_j)} \exp\{i\lambda_j s\} \tag{6}$$

Derivando respecto a a_s ($s = 1, 2, ..., L$) e igualando las derivadas a cero se obtiene:

$$\sum_{s=0}^{L} \hat{a}_s \sum_{j=-p+1}^{p} \frac{I_{xx}(\lambda_j)}{f_{nn}(\lambda_j)} \exp\{i\lambda_j(l-s)\} =$$
$$= \sum_{j=-p+1}^{p} \frac{I_{xy}(\lambda_j)}{f_{nn}(\lambda_j)} \exp\{i\lambda_j s\} \quad (l = 1, 2, ..., L) \tag{7}$$

Un caso particularmente simple y que estudiaremos primero es aquel en que la perturbación es ruido blanco. Entonces (7) se transforma en:

$$\sum_{s=0}^{L} \hat{a}_s \sum_{j=-p+1}^{p} I_{xx}(\lambda_j) \exp\{i\lambda_j(l-s)\} =$$
$$= \sum_{j=-p+1}^{p} I_{xy}(\lambda_j) \exp\{i\lambda_j^{l}\} \quad (l = 1, 2, ..., L) \tag{8}$$

Salvo por el hecho de que (5) y (6) son igualdades asintóticas —la igualdad estricta no se verifica por «efectos final» derivados de la longitud finita de las muestras—, (8) recoge igualdades en todo análogas a las que en el dominio del tiempo constituyen las ecuaciones normales de la regresión ordinaria.

Frecuentemente se introducen dos modificaciones, tanto en (7) como en (8). Por un lado, se sustituyen $I_{xx}(\lambda_j)$ e $I_{xy}(\lambda_j)$ por estimadores consistentes de los respectivos espectros. Por otro, se reduce el número de bandas de frecuencia sobre las que se suma, habida cuenta de que, tras el suavizado, $\hat{f}_{xx}(\lambda), \hat{f}_{xy}(\lambda)$ varían poco en intervalos de frecuencia de amplitud similar al ancho de la banda del estimador empleado.

La primera de estas modificaciones reemplaza los periodogramas por:

$$\begin{aligned} \hat{f}_{xx}(\lambda_j) &= \sum_{k=-m}^{m} W(k) I_{xx}(\lambda_{j+k}) \\ \hat{f}_{xy}(\lambda_j) &= \sum_{k=-m}^{m} W(k) I_{xy}(\lambda_{j+k}) \end{aligned} \tag{9}$$

siendo $W(.)$ una ventana espectral (desde luego, no tienen por qué ser iguales las ventanas espectrales empleadas en el suavizado de periodograma y periodograma cruzado; prescindimos además, por simplicidad notacional, de considerar separadamente el caso de los puntos extremos, cercanos a 0 y 2π, en que el suavizado no puede ser bilateral).

Si reemplazamos (9) en (8) y tenemos en cuenta que la transformada de Fourier de una convolución es un producto de transformadas de Fourier, vemos que (8) es equivalente a:

$$\sum_{s=0}^{L} \hat{a}_s \sum_{\tau=-T+1}^{T-1} w(\tau/T)\hat{C}_{xx}(\tau)\left[T^{-1} \sum_{j=-p+1}^{p} \exp\{i\lambda_j(k-s-\tau)\}\right] =$$

$$= \sum_{\tau=-T}^{T-1} w(\tau/T)\hat{C}_{xy}(\tau)\left[T^{-1} \sum_{j=-p+1}^{p} \exp\{i\lambda_j(k-\tau)\}\right] \quad (k = 1, 2, ..., L)$$

siendo $w(\tau/T)$ una «ventana de datos», transformada inversa de Fourier de $W(.)$. Como las expresiones en los corchetes son 1/2 cuando los exponentes son cero y cero en caso contrario,

$$\sum_{s=0}^{L} \hat{a}_s w(k-s/T)\hat{C}_{xx}(k-s) = w(k/T)\hat{C}_{xy}(k) \quad (k=1,2,...,L) \tag{10}$$

Reescribiendo (10) en forma matricial, tenemos:

$$\begin{bmatrix} \hat{C}_{xx}(0) & w(1/T)\hat{C}_{xx} & \cdots & w(L/T)\hat{C}_{xx}(L) \\ w(1/T)\hat{C}_{xx}(1) & \hat{C}_{xx}(0) & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ w(LT)\hat{C}_{xx}(L) & \vdots & \cdots & \hat{C}_{xx}(0) \end{bmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_2 \end{bmatrix} =$$

$$= \begin{bmatrix} \hat{C}_{xx}(0) \\ \vdots \\ \vdots \\ w(L/T)\hat{C}_{xy}(L) \end{bmatrix}$$

Ello muestra que, al estimar una ecuación con retardos, el efecto equivalente del suavizado espectral en el dominio del tiempo es realzar la diagonal principal de la matriz de momentos de los regresores, es decir, el mismo que se obtendría al emplear un estimador ridge generalizado:

$$\hat{a} = (\hat{C}_{xx} + G)^{-1}(\hat{C}_{xy} + d) \tag{11}$$

en que G es una matriz simétrica cuyo elemento genérico viene dado por:

$$g_{jk} = -[1 - w((j-k)/T)]\hat{C}_{xx}(j-k)$$

y $\mathbf{d}$ es un vector cuya ordenada j-ésima sería:

$$d_j = -[1 - w((j-1)/T)]\hat{C}_{xy}(j-1)$$

Bajo ciertas circunstancias un estimador como (11) puede interpretarse como el resultado de estimar $\mathbf{a}$ utilizando una distribución a priori $N(\mathbf{d}, G^{-1})$.

4. REGRESION ESPECTRAL EN RAICES LATENTES

La igualdad (7) puede reescribirse así:

$$\sum_{j=-p+1}^{p} f_{nn}^{-1}(\lambda_j) \sum_{s=0}^{L} [a_s I_{xx}(\lambda_j) \exp\{i\lambda_j(l-s)\} -$$

$$I_{xy}(\lambda_j) \exp\{i\lambda_j s\}] = 0 \quad (l = 1, 2, ..., L) \tag{12}$$

Matricialmente:

$$\sum_{-p+1}^{p} f_{nn}^{-1}(\lambda_j)[H(\lambda_j)\hat{\mathbf{a}} - \mathbf{d}(\lambda_j)] = 0 \tag{13}$$

siendo $H(\lambda_j)$ una matriz cuyo elemento genérico h_{mn} es $I_{xx}(\lambda_j) \exp\{i\lambda_j(m-n)\}$. Para cada frecuencia λ_j, igualando el corchete a cero se obtendría el filtro que mejor ajuste proporciona entre las bandas de frecuencia en torno a λ_j de $X(t)$ e $Y(t)$. La expresión (13) pondera las T igualdades —para las frecuencias de Fourier entre 0 y 2_π— de modo inversamente proporcional a la varianza del ruido en cada banda.

Observemos, sin embargo, que si tanto $X(t)$ como $n(t)$ tienen varianza muy pequeña en una determinada banda de frecuencia, nos encontramos ante un caso similar al que en regresión en raíces latentes se denomina «multicolinealidad no predictiva»: podría en efecto ocurrir que $H(\lambda_j) \simeq 0$ lo que haría el correspondiente sumando no informativo a efectos de estimación de $\mathbf{a}$.

Para evitar este inconveniente, el estimador que proponemos consiste en resolver:

$$\sum_{\lambda_j \varepsilon F} f_{nn}^{-1}(\lambda_j)[H(\lambda_j)\hat{\mathbf{a}} - \mathbf{d}(\lambda_j)] = 0 \tag{14}$$

siendo F el conjunto formado por las frecuencias de Fourier para las cuales $I_{xx}(\lambda_j) \simeq 0$.

La sustitución de $I_{xx}(\lambda_j)$ por $f_{xx}(\lambda_j)$ en (14) tendría el efecto indicado en la sección anterior —imposición de una distribución a priori sobre los valores de los parámetros $\mathbf{a}$—, dejando por lo demás subsistente el fundamento del método propuesto.

5. RELACION CON TRABAJOS Y DESARROLLOS EXISTENTES

El estimador que proponemos, aunque motivado por el trabajo realizado en el dominio del tiempo por Webster et al. (1974), presenta similaridades con trabajos de otros autores. Examinarlas resulta clarificador.

Engle (1974) sugirió la conveniencia de emplear regresión espectral con exclusión de bandas de frecuencia cuando existe la sospecha de que la relación lineal existente entre regresando y regresores no es la misma para todas las frecuencias.

El argumento en la sección anterior es en cierto modo similar; lo que proponemos es la exclusión de aquellas bandas de frecuencia en que los regresores tienen potencia espectral despreciable (bandas que, por tanto, no dan lugar a relaciones predictivas entre regresando y regresores).

Es también iluminante considerar el estimador que proponemos en relación a Sims (1974). Implícita en su trabajo está la preocupación por el efecto distorsionante sobre la estimación de los parámetros de fuertes picos estacionales en los espectros de los regresores, dando en ocasiones lugar a lo que llama «sesgo estacional». Muestra que, en el caso del modelo simple que estamos considerando, la minimización de la suma de cuadrados de los residuos equivale a la minimización de:

$$\int_{-\pi}^{\pi} f_{xx}(\lambda)|a(\lambda) - \hat{a}(\lambda)|^2 \, d\lambda \tag{15}$$

donde $a(\lambda)$ y $\hat{a}(\lambda)$ son, respectivamente, las funciones de transferencia (transformadas de Fourier del filtro lineal en (1)) verdadera y estimada. Sims (1974) concluye que la estimación de $\mathbf{a}$ resulta decisivamente influenciada por los valores de $a(\lambda)$ a las frecuencias en que $f_{xx}(\lambda)$ presenta crestas prominentes —puesto que en tal caso las discrepancias entre $a(\lambda)$ y $\hat{a}(\lambda)$ resultan más fuertemente ponderadas en (15)—

Obsérvese que en la minimización de la expresión anterior intervienen con ponderación muy baja aquellas bandas de frecuencia en que $f_{xx}(\lambda) = 0$. Ello hace que sustanciales discrepancias puedan presentarse en la estimación entre $a(\lambda)$ y $\hat{a}(\lambda)$, discrepancias que luego aflorarán al tomar la transformada inversa de Fourier para obtener estimaciones del vector de parámetros $\mathbf{a}$.

El método de regresión espectral en raíces latentes propuesto más arriba, prescinde de las bandas de frecuencia problemáticas en este aspecto: el razonamiento seguido podría ser el mismo empleado por Sims, aunque la motivación de nuestro trabajo es diferente.

Finalmente, es de interés observar que la no consideración de bandas de frecuencia para las que $f_{xx}(\lambda) \simeq 0$ supone eliminar en la diagonalización aproximada de la matriz C_{xx} aludida inmediatamente antes de la fórmula (4) los valores propios aproximadamente nulos. Ello muestra el parentesco del método propuesto con la regresión en componentes principales.

6. CONCLUSIONES

El mismo problema que Sims detectó en el caso de regresión con series temporales fuertemente estacionales se presenta en la regresión con pronunciadas simas en el espectro de los regresores. La alternativa de emplear regresión espectral con exclusión de bandas es válida en un caso como en otro, aunque se racionaliza de modo diferente en ambos.

El método que sugerimos está indicado en aquellos casos en que una o más bandas de frecuencia están ausentes o han sido acusadamente atenuadas. Esto es habitual en Econometría, en que sólo se dispone en ocasiones de series que han sido desestacionalizadas de manera drastica.

Desde un punto de vista conceptual, tanto la exclusión de bandas de frecuencia como el suavizado espectral pueden verse como manipulaciones similares a las que en el dominio del tiempo suponen los denominados métodos de regresión sesgada (fundamentalmente «ridge regression» y regresión en componentes principales). La toma en consideración de esta similitud permite un uso provechoso de la regresión espectral, a la vez que facilita la interpretación de los resultados.

BIBLIOGRAFIA

DHRYMES, P. (1971): *Econometrics. Statistical Foundations and Applications*, Harper and Row, New York, 1970.

ENGLE, R. F. (1974): «Band Spectrum Regression», *Int. Econ. Rev.*, vol. 15, págs. 1-11.

FISHMAN, G. S. (1969): *Spectral Methods in Econometrics*, Harvard University Press, Cambridge (Mass.), 1969.

HANNAN, E. J. (1963) «Regression for Time Series», en *Time Series Analysis*, ed. por M. Ronsenblatt, Wiley, New York, 1963.

HANNAN, E. J. (1965): «The Estimation of Relationships involving Distributed Lags», *Econometrica*, vol. 33, págs. 206-224.

HANNAN, E. J. (1967): «The estimation of a lagged regression relations», *Biometrika*, vol. 54, págs. 409-418.

HOERL, A. E., y KENNARD, R. W. (1970): «Ridge regression. Biased estimation for non-orthogonal problems», *Technometrics*, vol. 12, págs. 55-67.

SIMS, C. A. (1974): «Seasonality in Regressión», *J. Amer. Statist. Assoc.*, vol. 69, págs. 618-626.

WEBSTER, J. T., et al. (1974): «Latent Root Regression Analysis», *Technometrics*, vol. 16, págs. 513-522.

WHITTLE, P. (1952): *Prediction and Regulation by Linear Least Squares Methods*, D. Van Nostrand Co. Inc., Princeton, New Jersey, 1963.