

ESTIMACIÓN DE LA DENSIDAD MARGINAL DE PROCESOS DE MEDIA MÓVIL UTILIZANDO ESTIMADORES NO PARAMÉTRICOS DE MÁXIMA VEROSIMILITUD.

Ángeles Saavedra¹, Ricardo Cao².

¹Departamento de Estadística e Investigación Operativa.
Universidade de Vigo.

²Departamento de Matemáticas.
Universidade de A Coruña.

RESUMEN

Este traballo estudia el estimador de tipo convolutivo de la densidad marginal de procesos de media móvil presentado por sus autores en 1997. Mediante un estudio de simulación se compara el error cuadrático medio integrado del estimador de tipo convolutivo para distintos estimadores de los parámetros, con el estimador de Parzen-Rosenblatt y con el propio estimador de tipo convolutivo supuesto que los parámetros son conocidos.

Palabras y frases clave: Estimador tipo núcleo, error cuadrático medio, series de tiempo, procesos de media móvil.

Clasificación AMS: 62G05.

1. INTRODUCCIÓN

Sea $\{z_t\}$ un proceso de media móvil de orden q , $MA(q)$, dado por

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q},$$

con $\{a_t\}$ variables aleatorias iid y $\theta_1, \theta_2, \dots, \theta_q$ constantes tal que las raíces de la ecuación

$$1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q = 0$$

caen fuera del círculo unidad.

Dadas unas estimaciones $\{\hat{\theta}_i\}_{i=1}^q$ de los parámetros y unos valores iniciales $\{\tilde{a}_i\}_{i=0}^{-q+1}$ acotados en probabilidad puede definirse el estimador de la densidad marginal del proceso como

$$\hat{f}_{\hat{\theta}}(x) = \frac{1}{n^{q+1}h} \sum_{j_1, \dots, j_{q+1}=1}^n L_{q, \hat{\theta}} \left(\frac{x + \hat{\theta}_1 \hat{a}_{j_1} + \cdots + \hat{\theta}_q \hat{a}_{j_q} - \hat{a}_{j_{q+1}}}{h} \right),$$

donde

$$\begin{aligned} \hat{a}_t &= z_t + \sum_{i=1}^q \hat{\theta}_i \hat{a}_{t-i}, \\ L_{q, \hat{\theta}}(t) &= \int \cdots \int K(t + \hat{\theta}_1 u_1 + \cdots + \hat{\theta}_q u_q) K(u_1) \cdots K(u_q) du_1 \cdots du_q, \end{aligned}$$

K es una función núcleo y h es el parámetro de suavización y n el tamaño muestral. Este estimador ha sido estudiado por Saavedra y Cao (1999a,b) quienes, bajo ciertas condiciones de regularidad y considerando un h de orden $n^{-2/5}$, han obtenido un orden paramétrico de convergencia en procesos del tipo $MA(1)$:

$$\int \left(\hat{f}_{\hat{\theta}}(x) - f(x) \right)^2 dx = O_P(n^{-1}).$$

Este orden de convergencia paramétrico indica que estamos en un contexto en el cual las dificultades en la estimación de los parámetros del proceso afectarán directamente al error final que comete el estimador de la densidad marginal $\hat{f}_{\hat{\theta}}(x)$. Por esto, será importante obtener buenos estimadores de los parámetros del modelo.

Gran parte de las propiedades de eficiencia de los estimadores de mínimos cuadrados se basa en el hecho de que la densidad de los errores $\{a_t\}$ se asume normal, puesto que en este caso los estimadores mínimo-cuadráticos coinciden con los de máxima verosimilitud. Si la distribución de los errores no es normal, pero se supone conocida, los métodos basados en máxima verosimilitud siguen siendo aplicables, pero en la mayoría de los casos suponen cierta complejidad de cálculo.

Si no se asume ninguna distribución en los errores, puede extenderse el principio de máxima verosimilitud a un contexto no paramétrico, en el cual la densidad de los errores podrá ser estimada mediante el método núcleo. Algunos de estos estimadores no paramétricos de máxima verosimilitud (NPMLE) fueron comparados en un estudio de simulación con los estimadores clásicos en Cao, Hart y Saavedra (2003), sugiriéndose que los NPMLE deben ser utilizados cuando no se dispone de información sobre la distribución de los errores.

2. ESTIMACIÓN DE LA DENSIDAD MARGINAL

El objetivo de este trabajo es comparar el comportamiento, desde el punto de vista del error cuadrático medio, del estimador de la densidad marginal de tipo convolutivo, $\hat{f}_{\hat{\theta}}$, frente al clásico estimador tipo núcleo propuesto por Parzen-Rosenblatt. Queremos además cuantificar el impacto que el tipo de estimador empleado para aproximar θ puede tener en el error final cometido por el estimador $\hat{f}_{\hat{\theta}}$.

Se llevó a cabo un estudio de simulación con el fin de comparar el error cuadrático medio integrado (MISE) cometido por distintos estimadores de la densidad marginal de un proceso $\{z_t\}$ de media móvil de orden 1, $MA(1)$. Los estimadores considerados fueron: a) el estimador tipo núcleo de Parzen-Rosenblatt, $\hat{f}_{PR}$, b) el estimador de tipo convolutivo cuando θ se supone conocido, $\tilde{f}_{\theta}$, c) el estimador de tipo convolutivo cuando se estima θ mediante el método de los momentos, $\hat{f}_{\theta_{MOM}}$ y d) el estimador de tipo convolutivo cuando se estima θ mediante un estimador del tipo NPMLE, $\hat{f}_{\theta_{NPMLE}}$. El estimador no paramétrico de máxima verosimilitud utilizado en esta simulación está basado en el estimador de Hsieh-Manski (ver Cao, Hart y Saavedra (2003)). Como distribución del ruido se consideró una distribución $N(0, 1)$ así como distintas mixturas de densidades normales (ver Marron y Wand (1992)), con el fin de comparar el comportamiento de los estimadores de la densidad marginal ante diferentes tipos de dificultad.

Las tablas que se adjuntan resumen los valores, aproximados mediante el estudio de simulación, del parámetro que minimiza el $MISE$, h_{MISE} , y del valor mínimo del $MISE$, $MISE(h_{MISE})$. Ésta se ha llevado a cabo considerando el tamaño muestral $n = 500$ y diferentes valores para el parámetro θ .

$n = 500$		$\hat{f}_{PR}$	$\hat{f}_\theta$	$\hat{f}_{\hat{\theta}_{MOM}}$	$\hat{f}_{\hat{\theta}_{NPMLE}}$
$\theta = -0.5$	h_{MISE}	0.3193	0.1560	0.1379	0.0282
$\rho = 0.4$	$MISE(h_{MISE})$	1.98E-3	1.05E-3	1.15E-3	3.56E-3
$\theta = 0.25$	h_{MISE}	0.3146	0.2552	0.2413	0.1966
$\rho = -0.2353$	$MISE(h_{MISE})$	1.62E-3	1.2E-3	1.2E-3	1.65E-3
$\theta = 0.5$	h_{MISE}	0.3146	0.1447	0.1379	0.0282
$\rho = -0.4$	$MISE(h_{MISE})$	1.56E-3	5.56E-4	6.74E-4	2.99E-3
$\theta = 0.8$	h_{MISE}	0.3146	0.096	0.0689	0.0198
$\rho = -0.49$	$MISE(h_{MISE})$	1.53E-3	2.94E-4	5.25E-4	1.12E-2

Tabla 1. Valores aproximados de h_{MISE} y $MISE(h_{MISE})$ cuando el ruido sigue una distribución $N(0, 1)$.

$n = 500$		$\hat{f}_{PR}$	$\hat{f}_\theta$	$\hat{f}_{\hat{\theta}_{MOM}}$	$\hat{f}_{\hat{\theta}_{NPMLE}}$
$\theta = -0.5$	h_{MISE}	5.6206	0.6077	3.2413	4.475
$\rho = 0.4$	$MISE(h_{MISE})$	2E-2	4.84E-2	2.39E-1	1.91E-1
$\theta = 0.25$	h_{MISE}	1.8406	0.5827	1.6069	1.7133
$\rho = -0.2353$	$MISE(h_{MISE})$	2.41E-1	7.34E-2	2.41E-1	2.33E-1
$\theta = 0.5$	h_{MISE}	1.4482	0.5965	1.4758	1.3937
$\rho = -0.4$	$MISE(h_{MISE})$	1.31E-1	3.57E-2	1.31E-1	1.24E-1
$\theta = 0.8$	h_{MISE}	1.931	0.5834	0.862	0.8943
$\rho = -0.49$	$MISE(h_{MISE})$	7.72E-2	2.21E-2	4.39E-2	4E-2

Tabla 2. Valores aproximados de h_{MISE} y $MISE(h_{MISE})$ cuando el ruido sigue una distribución fuertemente asimétrica.

$n = 500$		$\hat{f}_{PR}$	$\hat{f}_\theta$	$\hat{f}_{\hat{\theta}_{MOM}}$	$\hat{f}_{\hat{\theta}_{NPMLE}}$
$\theta = -0.5$	h_{MISE}	1.8448	1.6206	1.6242	1.6385
$\rho = 0.4$	$MISE(h_{MISE})$	1.93E-2	3.36E-2	3.38E-2	3.37E-2
$\theta = 0.25$	h_{MISE}	3.4999	3.5725	3.6147	4.6135
$\rho = -0.2353$	$MISE(h_{MISE})$	2.35E-2	2.98E-2	3E-2	3E-2
$\theta = 0.5$	h_{MISE}	1.8448	1.6187	1.6528	1.6242
$\rho = -0.4$	$MISE(h_{MISE})$	1.87E-2	3.32E-2	3.37E-2	3.36E-2
$\theta = 0.8$	h_{MISE}	1.7039	1.2651	1.2758	1.2711
$\rho = -0.49$	$MISE(h_{MISE})$	2E-2	3.87E-2	3.95E-2	3.94E-2

Tabla 3. Valores aproximados de h_{MISE} y $MISE(h_{MISE})$ cuando el ruido sigue una distribución peine discreto.

Comparando las segundas columnas con las terceras y cuartas puede comprobarse como, tal como era de esperar, la estimación del parámetro θ del modelo incrementa el error $MISE$ cometido. La comparación entre las terceras y cuartas columnas nos confirma que el estimador de los momentos es la mejor opción cuando la densidad subyacente es la gaussiana. En el caso de que la distribución de los errores sea distinta de la $N(0, 1)$ los estimadores no paramétricos de máxima verosimilitud se presentan como una buena elección.

3. AGRADECIMIENTOS

La investigación del primer autor llevada a cabo en este trabajo ha sido parcialmente financiada por el proyecto PGIDIT02PXIA30003PR de la Xunta de Galicia y por el Vicerrectorado de Investigación de la Universidad de Vigo. La del segundo autor ha sido parcialmente financiada por el proyecto PGIDIT03PXIC10505PN. Asimismo ambos autores han sido parcialmente financiados en su investigación por el proyecto BFM2002-00265 del Ministerio de Ciencia y Tecnología.

4. REFERENCIAS

- Cao, R., Hart, J.D. y Saavedra, A. (2003). "Nonparametric maximum likelihood estimators for AR and MA time series". *J. Stat. Comput. Simul.* 73(5), 347–360.
- Marron, J.S. y Wand, M.P. (1992). "Exact mean integrated squared error". *Ann. Statist.* 20, 712–736.
- Saavedra, A. y Cao, R. (1999a). "Rate of convergence of a convolution-type estimator of the marginal density of a MA(1) process". *Stoch. Process. Appl.* 80, 129–155.
- Saavedra, A. y Cao, R. (1999b). "A comparative study of two convolution-type estimators of the marginal density of moving average process". *Comput. Stat.* 14, 355–373.