

MỘT ƯỚC LƯỢNG CHUỖI TRỰC GIAO CHO BÀI TOÁN GIẢI CHẬP MẬT ĐỘ PHI THAM SỐ

AN ORTHOGONAL SERIES ESTIMATOR FOR NONPARAMETRIC DENSITY ECONVOLUTION PROBLEM

LÊ THỊ HỒNG THUY^(*) và CAO XUÂN PHƯƠNG^(**)

TÓM TẮT: Cho X là một biến ngẫu nhiên liên tục với hàm mật độ f_X bị chặn và có giá com-pắc trong $[-\pi, \pi]$. Trong bài viết, chúng tôi khảo sát bài toán ước lượng f_X trên cơ sở một mẫu ngẫu nhiên từ phân phối của biến ngẫu nhiên Y được sinh ra từ mô hình $Y = X + \varepsilon$. Ở đây, ε là một nhiễu ngẫu nhiên tuân theo một phân phối biết trước. Bằng cách áp dụng phương pháp ước lượng chuỗi trực giao, chúng tôi đề xuất một ước lượng phi tham số cho f_X . Sau đó, chúng tôi thiết lập một số kết quả hội tụ của ước lượng tương ứng theo sai số bình phương tích phân trung bình dưới một số điều kiện nào đó đặt ra trên các phân phối của X và ε .

Từ khóa: bài toán giải chập mật độ; ước lượng chuỗi trực giao; sự bền vững; các cấp độ hội tụ.

ABSTRACT: Let X be a continuous random variable with a density function f_X blocked and has a compact supported on $[-\pi, \pi]$. In this study, we examine the estimation problem f_X on the basis of a random sample from the distribution of random variables Y generated by the model $Y = X + \varepsilon$. Here ε is a random variable compiled with a known distribution. By applying an orthogonal series estimation method, we propose a nonparametric estimator of f_X . We then establish some convergence results of the respective estimates against the mean integral squared error under some certain conditions set on the distributions of X and ε .

Key words: density deconvolution problem; orthogonal series estimator; consistency; levels of convergence.

1. ĐẶT VẤN ĐỀ

Cho X, Y và ε là các biến ngẫu nhiên liên tục, nhận các giá trị thực và liên hệ nhau thông qua đẳng thức $Y = X + \varepsilon$. Giả thiết các biến X và ε là độc lập. Gọi f_Y, f_X và f_ε lần lượt là hàm mật độ của Y, X và ε , trong đó, giả thiết f_X là hàm bị chặn và có giá chứa trong đoạn $[-\pi, \pi]$. Thêm vào đó, giả thiết f_ε được biết chính xác. Cho $(Y_1, \dots, Y_n)$ là một mẫu ngẫu nhiên kích thước n được chọn từ phân phối của

Y . Bài toán đặt ra là ước lượng $f_X(x), x \in [-\pi, \pi]$, trên cơ sở các quan trắc ngẫu nhiên $Y_1, Y_2, \dots, Y_n$ cũng như hàm mật độ f_ε của ε . Bài toán này thuộc loại các bài toán ngược trong thống kê.

Giả thiết về sự độc lập của X và ε dẫn đến phương trình $f_Y = f_X * f_\varepsilon$, trong đó hàm $(f_X * f_\varepsilon)(x) := \int_{-\infty}^{+\infty} f_X(x-u)f_\varepsilon(u)du$ là tích chập của f_X và f_ε . Như vậy, bài toán xác định f_X chính là bài toán giải phương trình tích chập trên, còn được gọi

^(*) ThS. Trường Đại học Văn Lang, thuy.lth@vlu.edu.vn

^(**) TS. Trường Đại học Tôn Đức Thắng, caoxuanphuong@tdtu.edu.vn, Mã số: TCKH23-01-2020

là bài toán giải chấp mật độ. Phạm vi ứng dụng của bài toán này khá rộng, chẳng hạn trong thống kê y học [1, tr.1176-1185], kinh tế lượng [2], phân tích hồi quy [3, tr.145-168], thiên văn học [4, tr.483-506]... Thập niên 80, thế kỷ XX được xem là giai đoạn mở đầu cho việc nghiên cứu bài toán này, các nghiên cứu tiên phong tiêu biểu [5, tr.169-184], [6, tr.1184-1186], [7, tr.3325-3342], [8, tr.1257-1272]. Từ đó đến nay, bài toán này được nghiên cứu sâu rộng trên nhiều khía cạnh khác nhau, như sự bền vững, sự tối ưu về tốc độ hội tụ, sự lựa chọn các tham số chính hóa theo dữ liệu... Những tổng hợp tương đối đầy đủ về bài toán này có thể tìm thấy trong bài viết của A. Meister [9].

Về cơ bản, phương pháp ước lượng nhân giải chấp là cách tiếp cận phổ biến nhất cho ước lượng f_X . Phương pháp này được giới thiệu lần đầu tiên [5, tr.169-184] và được cải tiến trong nhiều nghiên cứu sau đó. Các ước lượng loại nhân có hạn chế trong việc cài đặt trên máy tính vì cần dùng đến một phương pháp số thích hợp để tính xấp xỉ tích phân. Ngoài ra, nó cũng chỉ được sử dụng với điều kiện $\varphi_\varepsilon(t) \neq 0$ với mọi $t \in \mathbf{R}$, trong đó, φ_ε là hàm đặc trưng của ε . Có nhiều phân phối thỏa mãn điều kiện này, chẳng hạn như các phân phối chuẩn, Cauchy, Laplace... Tuy nhiên, cũng có một vài phân phối thông dụng vi phạm điều kiện này, điển hình là các phân phối đều và tam giác. Một vài nghiên cứu [10, 2023-2053], [11, 201-231] đã phát triển phương pháp sóng nhỏ (wavelet method) để xây dựng các ước lượng chuỗi trực giao cho f_X . Tuy nhiên, ứng dụng của những nghiên cứu này bị giới hạn trong điều kiện φ_ε không triệt tiêu trên $\mathbf{R}$. Bài viết [12] giới thiệu một ước lượng chuỗi trực giao cho f_X dưới giả thiết các không điểm (zeros) của hàm φ_ε được biết chính xác và tuần hoàn. Nếu sự tuần hoàn này bị bỏ đi, phương pháp được đề nghị trong bài viết này cũng không thể áp dụng. Cho đến nay, chúng tôi vẫn chưa thấy có nghiên

cứu nào phát triển một phương pháp ước lượng chuỗi trực giao cho f_X mà có thể được áp dụng bất kể là φ_ε có không điểm trong $\mathbf{R}$, thậm chí không nhất thiết phải biết chính xác vị trí các không điểm trong trường hợp φ_ε có không điểm..

2. NỘI DUNG

2.1. Các ký hiệu

Sau đây, chúng tôi giới thiệu một số ký hiệu sẽ được sử dụng trong bài báo này. Các ký hiệu $\mathbf{N}$, $\mathbf{Z}$ và $\mathbf{R}$ tương ứng biểu thị các tập hợp số nguyên dương, số nguyên và số thực. Với một tập con hữu hạn A của $\mathbf{Z}$, ký hiệu $|A|$ biểu thị số phần tử của A kỳ vọng và phương sai của một biến ngẫu nhiên U lần lượt được ký hiệu là EU và $\text{Var } U$. Ngoài ra, ký hiệu φ_U biểu thị hàm đặc trưng của U , tức là $\varphi_U(t) := E(e^{itU})$ với $t \in \mathbf{R}$ và i là đơn vị ảo. Với các tham số dương a_n và b_n phụ thuộc vào cỡ mẫu n , ký hiệu $a_n \leq O(b_n)$ có nghĩa rằng, tồn tại một hằng số dương C không phụ thuộc vào n sao cho $a_n \leq Cb_n$ với mọi n đủ lớn. Ngoài ra, ta viết $a_n \sim b_n$ khi $a_n \leq O(b_n)$ và $b_n \leq O(a_n)$.

2.2. Ước lượng

Trước tiên, nhắc lại rằng, $L^2([-\pi, \pi])$ là một không gian Hilbert với tích vô hướng $\langle u, v \rangle := \int_{-\pi}^{\pi} u(x)\overline{v(x)}dx$ với mọi $u, v \in L^2([-\pi, \pi])$ và được trang bị chuẩn $\|u\| := \sqrt{\langle u, u \rangle}$. Ngoài ra, nếu $\{\zeta_k\}_{k \in \mathbf{Z}}$ là một cơ sở trực chuẩn của $L^2([-\pi, \pi])$ thì với bất kỳ hàm $u \in L^2([-\pi, \pi])$, ta có biểu diễn $u = \sum_{k=-\infty}^{+\infty} \langle u, \zeta_k \rangle \zeta_k$.

Ta trở lại vấn đề ước lượng f_X . Dưới giả thiết rằng, hàm f_X bị chặn trên $[-\pi, \pi]$, ta có $f_X \in L^2([-\pi, \pi])$. Vì họ $\{(2\pi)^{-1/2} e^{ikx} : k \in \mathbf{Z}\}$ là một cơ sở trực chuẩn đầy đủ trong không gian Hilbert $L^2([-\pi, \pi])$ nên ta biểu diễn $f_X(x)$ (với $x \in [-\pi, \pi]$) dưới dạng:

$$f_X(x) = \sum_{k=-\infty}^{+\infty} \left\langle f_X, \frac{1}{\sqrt{2\pi}} e^{ikx} \right\rangle \frac{1}{\sqrt{2\pi}} e^{ikx} \\ = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \varphi_X(-k) e^{ikx} = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \frac{\varphi_Y(-k)}{\varphi_\varepsilon(-k)} e^{ikx} \quad (1)$$

Nếu $Z(\varphi_\varepsilon) = \emptyset$, trong đó $Z(\varphi_\varepsilon) := \{k \in \mathbf{Z} : \varphi_\varepsilon(-k) = 0\}$ sử dụng các đẳng thức

$$\left\langle f_X, e^{ikx} \right\rangle = \int_{-\pi}^{\pi} f_X(x) e^{-ikx} dx = \varphi_X(-k) \quad \text{và}$$

$\varphi_Y = \varphi_X \varphi_\varepsilon$. Với mẫu ngẫu nhiên $(Y_1, \dots, Y_n)$, ta

sử dụng đại lượng $\hat{\varphi}_Y(-k) := n^{-1} \sum_{j=1}^n e^{-ikY_j}$ để ước lượng $\varphi_Y(-k)$, với mỗi $k \in \mathbf{Z}$. Trong (1), bằng cách thay $\varphi_Y(-k)$ bởi $\hat{\varphi}_Y(-k)$, ta thu được một ước lượng hình thức cho $f_X(x)$ dưới dạng sau:

$$\tilde{f}_X(x) := \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \frac{\hat{\varphi}_Y(-k)}{\varphi_\varepsilon(-k)} e^{ikx}$$

Tuy nhiên, ta thấy $\tilde{f}_X$ có thể không xác định vì hai lý do sau và nhất thiết phải đề xuất một phiên bản hiệu chỉnh cho $\tilde{f}_X$. Lý do đầu tiên là vì $\lim_{|k| \rightarrow +\infty} 1/\varphi_\varepsilon(-k) = \infty$. Ta có thể giải quyết vấn đề này bằng cách cắt chuỗi thành một tổng hữu hạn nhằm mục đích loại bỏ các giá trị $|k|$ đủ lớn. Lý do thứ hai là nếu $Z(\varphi_\varepsilon) \neq \emptyset$ thì $1/\varphi_\varepsilon(-k)$ với $k \in Z(\varphi_\varepsilon)$ sẽ không xác định. Ý tưởng để giải quyết vấn đề này là thay thế đại lượng $\varphi_\varepsilon(-k)$ trong mẫu số bởi một tham số dương phụ thuộc vào cỡ mẫu n bất cứ khi nào $\varphi_\varepsilon(-k)$ trở nên đủ gần 0. Tổng hợp từ các phân tích ở trên, chúng tôi đề nghị ước lượng $f_X(x)$ với $x \in [-\pi, \pi]$ bởi đại lượng:

$$\hat{f}_{X;N_n,\delta_n}(x) := \frac{1}{2\pi} \sum_{k=-N_n}^{N_n} \frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k) e^{ikx}}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \quad (2)$$

Trong đó, các tham số $N_n \in \mathbf{N}$ và $\delta_n \in (0, +\infty)$ phụ thuộc vào cỡ mẫu n và sẽ được chọn sau.

2.3. Sự hội tụ của ước lượng

2.3.1. Mệnh đề

Xét ước lượng $\hat{f}_{X;N_n,\delta_n}$ trong (2). Với bất kỳ $N_n \in \mathbf{N}$ và $\delta_n > 0$, ta có

$$\mathbf{E} \left\| \hat{f}_{X;N_n,\delta_n} - f_X \right\|^2 \leq \frac{1}{2\pi} E_{1;N_n,\delta_n} + \frac{1}{2\pi} E_{2;N_n,\delta_n} + \frac{1}{2\pi} E_{3;N_n}$$

Trong đó:

$$E_{1;N_n,\delta_n} := \sum_{k=-N_n}^{N_n} \left| \frac{|\varphi_\varepsilon(k)|^2}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} - 1 \right| |\varphi_X(k)|^2,$$

$$E_{2;N_n,\delta_n} := \frac{1}{n} \sum_{k=-N_n}^{N_n} \frac{|\varphi_\varepsilon(k)|^2}{\max\{|\varphi_\varepsilon(k)|^4; \delta_n^2\}},$$

$$E_{3;N_n} := \sum_{|k| > N_n} |\varphi_X(k)|^2$$

Chứng minh: Ta có:

$$\mathbf{E} \left\| \hat{f}_{X;N_n,\delta_n} - f_X \right\|^2 = \mathbf{E} \int_{-\pi}^{\pi} \left| \frac{1}{2\pi} \sum_{k=-N_n}^{N_n} \frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k) e^{ikx}}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} - \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \varphi_X(-k) e^{ikx} \right|^2 dx.$$

Vì họ $\{e^{ikx} : k \in \mathbf{Z}\}$ là trực giao và vì đẳng thức $\mathbf{E}|U - a|^2 = |\mathbf{E}U - a|^2 + \text{Var} U$, ta suy ra:

$$\begin{aligned} \mathbf{E} \left\| \hat{f}_{X;N_n,\delta_n} - f_X \right\|^2 &= \frac{1}{2\pi} \sum_{k=-N_n}^{N_n} \mathbf{E} \left| \frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} - \varphi_X(-k) \right|^2 + \frac{1}{2\pi} \sum_{|k| > N_n} |\varphi_X(k)|^2 \\ &= \frac{1}{2\pi} \sum_{k=-N_n}^{N_n} \left| \mathbf{E} \left(\frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \right) - \varphi_X(-k) \right|^2 \\ &\quad + \frac{1}{2\pi} \sum_{k=-N_n}^{N_n} \text{Var} \left(\frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \right) + \frac{1}{2\pi} \sum_{|k| > N_n} |\varphi_X(k)|^2. \end{aligned}$$

Kết hợp đẳng thức này với các đánh giá

$$\begin{aligned} \mathbf{E} \left(\frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \right) &= \frac{\varphi_\varepsilon(k) \mathbf{E} \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} = \frac{|\varphi_\varepsilon(k)|^2 \varphi_X(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \\ \text{Var} \left(\frac{\varphi_\varepsilon(k) \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \right) &= \frac{|\varphi_\varepsilon(k)|^2 \text{Var} \hat{\varphi}_Y(-k)}{\max\{|\varphi_\varepsilon(k)|^4; \delta_n^2\}} \leq \frac{|\varphi_\varepsilon(k)|^2}{n \max\{|\varphi_\varepsilon(k)|^4; \delta_n^2\}} \end{aligned}$$

Ta nhận được kết luận của mệnh đề.

Cho trước $\alpha > 0$ và $L > 0$. Ta ký hiệu $S_{\alpha,L}$ là tập hợp tất cả các hàm mật độ f sao cho hàm đặc trưng φ_f tương ứng, xác định bởi $\varphi_f(t) := \int_{-\infty}^{+\infty} f(x) e^{itx} dx$

với $t \in \mathbf{R}$, thỏa mãn $\sum_{k=-\infty}^{+\infty} |\varphi_f(k)|^2 |k|^{2\alpha} \leq L$. Với

lớp $S_{\alpha,L}$, ta ký hiệu $R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] := \sup_{f \in S_{\alpha,L}} \mathbf{E} \left\| \hat{f}_{X;N_n,\delta_n} - f \right\|^2$,

gọi là rủi ro tối đa của ước lượng $\hat{f}_{X;N_n,\delta_n}$ trên lớp $S_{\alpha,L}$. Sau đây, dưới các điều kiện ràng buộc nào đó cho φ_ε , chúng ta sẽ thiết lập một số chặn trên về tốc độ hội tụ của $R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}]$. Thật ra, ta đưa ra hai giả thiết liên quan đến φ_ε như sau:

(A.1): Tồn tại $c_1, \beta > 0$ sao cho
 $|\varphi_\varepsilon(t)| \geq c_1(1+|t|)^{-\beta}$ với $t \in \mathbf{R}$.

(A.2): Tồn tại $c_2, \eta, \mu, \nu, \vartheta > 0$ và $\kappa \geq 0$ sao cho
 $|\varphi_\varepsilon(t)| \geq c_2 |\sin(\eta t)|^\mu (1+|t|)^{-\nu} e^{-\kappa|t|^\vartheta}$ với $t \in \mathbf{R}$.

2.3.2. Định lý

Cho $\alpha > 0$ và $L > 0$.

a) Dưới giả thiết (A.1), chọn $N_n \sim n^{1/(2\alpha+2\beta+1)}$
 và $\delta_n \sim n^{-2\beta/(2\alpha+2\beta+1)}$, ta được

$$R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O\left(n^{-2\alpha/(2\alpha+2\beta+1)}\right).$$

b) Cho giả thiết (A.2) được thỏa mãn. Nếu
 $\kappa = 0$ thì chọn $N_n \sim n^{2/(8\alpha\mu+4\nu+4\mu+6\alpha+3)}$ và
 $\delta_n \sim n^{-(8\alpha\mu+4\nu+4\mu+2\alpha+1)/(8\alpha\mu+4\nu+4\mu+6\alpha+3)}$, ta
 được $R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O\left(n^{-4\alpha/(8\alpha\mu+4\nu+4\mu+6\alpha+3)}\right)$.

Nếu $\kappa > 0$ thì chọn $\delta_n \sim n^{-\xi}$,
 $N_n \sim (\tau \ln n)^{1/\vartheta}$ với $0 < \xi < 1$ và $0 < \tau < \xi/(2\kappa)$,
 ta được $R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O\left((\ln n)^{-2\alpha/\vartheta}\right)$.

Chúng minh: a) Giả sử $f_X \in S_{\alpha,L}$. Đặt
 $A_{N_n,\delta_n} := \{k \in [-N_n, N_n] : |\varphi_\varepsilon(k)|^2 < \delta_n\}$ và xét $0 < \delta_n < c_1^2/4^\beta$.
 Với bất kỳ $k \in A_{N_n,\delta_n}$, ta có $c_1^2(1+|k|)^{-2\beta} < \delta_n$ hay
 $c_1^2/\delta_n < (1+|k|)^{2\beta}$. Vì $0 < \delta_n < c_1^2/4^\beta$, ta suy ra $|k| > 1$,
 do đó $c_1^2/\delta_n < 4^\beta |k|^{2\beta}$, tương đương $|k| > [c_1^2/(4^\beta \delta_n)]^{1/(2\beta)} := M_n$.
 Tóm lại, $A_{N_n,\delta_n} \subset \{k \in \mathbf{Z} : |k| > M_n\}$. Sử dụng quan hệ
 này và giả thiết $f_X \in S_{\alpha,L}$, ta có đánh giá:

$$\begin{aligned} E_{1;N_n,\delta_n} &\leq \sum_{k \in A_{N_n,\delta_n}} |\varphi_X(k)|^2 \leq \sum_{|k| > [c_1^2/(4^\beta \delta_n)]^{1/(2\beta)}} |\varphi_X(k)|^2 \\ &= \sum_{|k| > [c_1^2/(4^\beta \delta_n)]^{1/(2\beta)}} |\varphi_X(k)|^2 |k|^{2\alpha} |k|^{-2\alpha} \leq O\left(\delta_n^{\alpha/\beta}\right). \end{aligned}$$

Tiếp theo, ta có:

$$E_{2;N_n,\delta_n} \leq \frac{1}{n} \sum_{k=-N_n}^{N_n} \frac{1}{|\varphi_\varepsilon(k)|^2} \leq \frac{1}{n} \sum_{k=-N_n}^{N_n} c_1^{-2} (1+|k|)^{2\beta} \leq O\left(\frac{N_n^{2\beta+1}}{n}\right),$$

$$E_{3;N_n} = \sum_{|k| > N_n} |\varphi_X(k)|^2 |k|^{2\alpha} |k|^{-2\alpha} \leq L N_n^{-2\alpha}.$$

Kết hợp Mệnh đề 2.3.1 với các đánh giá
 trên của $E_{1;N_n,\delta_n}$, $E_{2;N_n,\delta_n}$ và $E_{3;N_n}$, ta suy ra

$$R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O\left(\delta_n^{\alpha/\beta} + \frac{N_n^{2\beta+1}}{n} + N_n^{-2\alpha}\right).$$

Chọn $N_n \sim n^{1/(2\alpha+2\beta+1)}$ và $\delta_n = N_n^{-2\beta} \sim n^{-2\beta/(2\alpha+2\beta+1)}$,
 ta được $R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O\left(n^{-2\alpha/(2\alpha+2\beta+1)}\right)$.

b) Giả sử $f_X \in S_{\alpha,L}$. Ta có:

$$E_{2;N_n,\delta_n} \leq \frac{1}{n} \sum_{k=-N_n}^{N_n} \frac{1}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} \leq \frac{2N_n+1}{n\delta_n},$$

$$E_{3;N_n} = \sum_{|k| > N_n} |\varphi_X(k)|^2 |k|^{2\alpha} |k|^{-2\alpha} \leq L N_n^{-2\alpha}.$$

Tiếp theo, ta tập trung vào việc đánh giá
 $E_{1;N_n,\delta_n}$. Với $\rho_n > 0$ mà sẽ được chọn sau, đặt:

$$A_{1;N_n,\rho_n} := \{k \in [-N_n, N_n] \cap \mathbf{Z} : |\varphi_\varepsilon(k)| \leq \rho_n\},$$

$$A_{2;N_n,\rho_n} := \{k \in [-N_n, N_n] \cap \mathbf{Z} : |\varphi_\varepsilon(k)| > \rho_n\}.$$

Với mỗi $k \in A_{1;N_n,\rho_n}$, tồn tại duy nhất số
 nguyên l_k phụ thuộc vào k sao cho
 $(2l_k - 1)\pi/(2\eta) \leq k < (2l_k + 1)\pi/(2\eta)$, và do đó
 $|k - l_k \pi/\eta| \leq \pi/(2\eta)$. Từ đó, ta có đánh giá:

$$\begin{aligned} \rho_n &\geq |\varphi_\varepsilon(k)| \geq c_2 |\sin(\eta k)|^\mu (1+|k|)^{-\nu} e^{-\kappa|k|^\vartheta} \\ &\geq c_2 \left| \sin \eta \left(k - \frac{l_k \pi}{\eta} \right) \right|^\mu (1+N_n)^{-\nu} e^{-\kappa N_n^\vartheta} \\ &\geq 2^{-\nu} c_2 \left| \frac{2\eta}{\pi} \left(k - \frac{l_k \pi}{\eta} \right) \right|^\mu N_n^{-\nu} e^{-\kappa N_n^\vartheta}. \end{aligned}$$

Do vậy, $k \in [l_k \pi/\eta - R_n; l_k \pi/\eta + R_n] =: I_{k,n}$
 với $R_n := 2^{\nu/\mu-1} c_2^{-1/\mu} \pi N_n^{\nu/\mu} e^{\kappa N_n^\vartheta/\mu} 1/\mu$. Với lưu ý

rằng $|l_k| \leq 1/2 + N_n \eta/\pi$, ta kết luận

$$A_{1;N_n,\rho_n} \subset \bigcup_{|l_k| \leq 1/2 + N_n \eta/\pi} (I_{k,n} \cap \mathbf{Z}). \text{ Đánh giá này cho thấy:}$$

$$|A_{1;N_n,\rho_n}| \leq \sum_{|l_k| \leq 1/2 + N_n \eta/\pi} |I_{k,n} \cap \mathbf{Z}|$$

$$\leq O\left(N_n^{\nu/\mu+1} e^{\kappa N_n^\vartheta/\mu} 1/\mu\right).$$

Bây giờ, ta có:

$$\begin{aligned}
E_{1;N_n,\delta_n} &= \sum_{k \in A_{1;N_n,\rho_n}} \left| \frac{|\varphi_\varepsilon(k)|^2}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} - 1 \right|^2 |\varphi_X(k)|^2 \\
&+ \sum_{k \in A_{2;N_n,\rho_n}} \left| \frac{|\varphi_\varepsilon(k)|^2}{\max\{|\varphi_\varepsilon(k)|^2; \delta_n\}} - 1 \right|^2 |\varphi_X(k)|^2 \\
&\leq |A_{1;N_n,\rho_n}| + \frac{(2N_n+1)\delta_n^2}{\rho_n} \leq O\left(N_n^{\nu/\mu+1} e^{\kappa N_n^{\theta/\mu}} \rho_n^{1/\mu} + \frac{N_n \delta_n^2}{\rho_n}\right).
\end{aligned}$$

Kết hợp Mệnh đề 2.3.1 với các đánh giá của

$E_{1;N_n,\delta_n}$, $E_{2;N_n,\delta_n}$ và $E_{3;N_n}$ ở trên, ta suy ra:

$$R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O\left(N_n^{\nu/\mu+1} e^{\kappa N_n^{\theta/\mu}} \rho_n^{1/\mu} + \frac{N_n \delta_n^2}{\rho_n} + \frac{N_n}{n \delta_n} + N_n^{-2\alpha}\right).$$

Khi $\kappa > 0$, ta chọn $\delta_n \sim n^{-\xi}$, $N_n \sim (\tau \ln n)^{1/\theta}$ và $\rho_n \sim n^{-a}$ với $0 < \xi < 1$, $0 < \tau < \xi / (2\kappa)$ và

$k\tau < a < \xi / 2$, và khi đó ta được $R[\hat{f}_{X;N_n,\delta_n}; S_{\alpha,L}] \leq O((\ln n)^{-2\alpha/\theta})$.

Định lý đã được chứng minh.

3. KẾT LUẬN

Bài viết đã khảo sát mô hình sai số đo cộng tính $Y = X + \varepsilon$ và giới thiệu một ước lượng chuỗi trực giao phụ thuộc vào hai tham số chính hóa cho hàm mật độ f_X của X trên cơ sở mẫu ngẫu nhiên $(Y_1, Y_2, \dots, Y_n)$ từ phân phối của Y . Ước lượng này là vững theo trung bình tương ứng với sai số bình phương tích phân trung bình. Dưới một số điều kiện chính quy được giả định cho các hàm đặc trưng của X và ε , một số tốc độ hội tụ theo cỡ mẫu n đã được thiết lập.

TÀI LIỆU THAM KHẢO

- [1] J. D. Tournier, F. Calamante, D. G. Gadian, A. Connelly (2004), Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution, *NeuroImage*, 23.
- [2] J. L. Horowitz (1998), *Semiparametric Methods in Econometrics*, Springer-Berlin Heidelberg, New York.
- [3] J. L. Horowitz, M. Markatou (1996), *Semiparametric estimation of regression models for panel data*, *The review of Economic Studies*, 63(1).
- [4] N. Bissantz, L. Dumbgen, H. Holzmann, A. Munk (2007), *Nonparametric confidence bands in deconvolution density estimation*, *Journal of the Royal Statistical Society-Series B (Statistical Methodology)*, 69(3).
- [5] L. Stefanski, R. Carroll (1990), *Deconvoluting kernel density estimators*, *Statistics*, 21(2).
- [6] R. Carroll, P. Hall (1988), *Optimal rates of convergence for deconvolving a density*, *Journal of American Statistical Association*, 83(404).
- [7] R. L. Taylor, M. H. Zhang (1990), *On a strongly consistent nonparametric density estimator for the deconvolution problem*, *Communications in Statistics-Theory and Methods*, 19(9).
- [8] J. Fan (1991), *On the optimal rates of convergence for nonparametric deconvolution problems*, *The Annals of Statistics*, 19(3).
- [9] A. Meister (2009), *Deconvolution problems in nonparametric statistics*, Springer-Verlag, Berlin.
- [10] M. Pensky, B. Vidakovic (1999), *Adaptive wavelet estimator for nonparametric density deconvolution*, *The Annals of Statistics*, 27(6).
- [11] K. Lounici, R. Nickl (2011), *Global uniform risk bounds for wavelet deconvolution estimators*, *The Annals of Statistics*, 39(1).
- [12] A. Meister (2008), *Deconvolution from Fourier-oscillating error densities under decay and smoothness restrictions*, *Inverse Problems*, 24.

Ngày nhận bài: 05-2-2020. Ngày biên tập xong: 19-8-2020. Duyệt đăng: 24-9-2020