Statistics and Probability Letters 96 (2015) 346–355

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Nonparametric adaptive density estimation on random fields using wavelet method

Linyuan Li *

Department of Mathematics and Statistics, University of New Hampshire, NH, 03824, USA

ARTICLE INFO

Article history:
Received 19 July 2014
Received in revised form 16 August 2014
Accepted 20 October 2014
Available online 30 October 2014

MSC:
primary 62G05
secondary 62G07
62M30

Keywords:
Random fields
Wavelet estimator
Minimax estimation
Mixing errors

ABSTRACT

We consider non-linear wavelet-based estimators of density functions with stationary random fields, which are indexed by the integer lattice points in the N-dimensional Euclidean space and are assumed to satisfy some mixing conditions. We investigate their asymptotic rates of convergence based on thresholding of empirical wavelet coefficients and show that these estimators achieve nearly optimal convergence rates within a logarithmic term over a large range of Besov function classes $B^s_{p,q}$. Therefore, wavelet estimators still achieve nearly optimal convergence rates for random fields and provide explicitly the extraordinary local adaptability.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample from a density f on the real line, an interesting problem is to estimate f based on the observed data $\{X_i\}$'s. There is a vast literature on density estimation based on independent or weakly dependent random samples. Recently, inference on spatial version of the above data (i.e., spatial data or random fields) receives more and more attention because of its importance for applications. Spatial data arise in many different fields, such as econometrics, epidemiology, astronomy, geophysics, medicine, environmental science, image analysis and oceanography. For a systematic discussion on random fields, see, e.g., Cressie (1991), Anselin and Florax (1995), Guyon (1995), Stein (1999), Banerjee et al. (2004) and among others. Density and regression estimation using kernel method for random fields was investigated by Tran (1990), Tran and Yakowitz (1993), Carbon et al. (1996), Bradley and Tran (1999), Hallin et al. (2001, 2004), Biau (2003) and Biau and Cadre (2004), among others. To the best of our knowledge, the problem of the adaptive nonparametric estimation of the spatial density using the wavelet approach has not been addressed so far. It is the intention of this paper to study density estimation for random variables which show spatial interaction.

Unlike the kernel method, which typically assumes that the underlying curve satisfies certain fixed and known smoothness condition such as two-times continuous differentiability, wavelet-based method assumes that the underlying targeted curve belongs to a large function space which has varying degrees of smoothness. Although the wavelet estimators do not depend on those unknown smoothness parameters, they typically attain so-called optimal convergence rates over that large function space. This optimality demonstrates that the wavelet estimators behave as if one knows the underlying

* Tel.: +1 603 862 4592.
E-mail address: linyuan@cisunix.unh.edu.

http://dx.doi.org/10.1016/j.spl.2014.10.012
0167-7152/

curve in advance. This extraordinary adaptivity property is very useful when one is not sure about the smoothness of the underlying curve. In this paper, we consider the data are observed at regularly spaced lattice points from strictly stationary random fields, their density functions f belong to a large range of Besov function classes $B^s_{p,q}$, $s > 0$, $1 \leq p, q \leq \infty$. We assume that the data possess certain mixing dependence structure, whose formal definition will be presented in the next section. As those in the wavelet literature, we investigate the asymptotic convergence rates of nonlinear wavelet-based estimators. The main contribution of this paper is to show that these estimators with strong mixing spatial data attain nearly optimal convergence rates over the large function classes $B^s_{p,q}$, which extends existing results with independent or weakly dependent errors to random fields.

The rest of this paper is organized as follows. In the next section, we recall briefly the elements of random fields, wavelets, density function spaces $B^s_{p,q}$, and introduce the wavelet-based estimators for the spatial density functions. The main result is described in Section 3, whose proof is provided in Section 4 and Appendix.

2. Preliminaries

This section provides some basic facts about random fields, wavelets, the function spaces $B^s_{p,q}$ for density functions, and the proposed wavelet-based estimators that will be used in the sequel.

2.1. Random fields

We consider that the data are observed at regularly spaced lattice points from a strictly stationary random field. To be specific, let $\mathbb{Z}^N$ denote the integer lattice points in the N-dimensional Euclidean space and $\{X_\mathbf{i}\}$ be a strictly stationary random field indexed by $\mathbb{Z}^N$ and defined on some probability space $(\Omega, \mathcal{F}, P)$. A point $\mathbf{i}$ in $\mathbb{Z}^N$ will be referred to as a site and written as $\mathbf{i} = (i_1, i_2, \ldots, i_N)$. In this paper, we adopt the same dependence structure (mixing conditions) on the random field as in Tran (1990) and Hallin et al. (2004). Let $\mathbf{S}$ and $\mathbf{S}'$ be two collections of sites. The Borel fields $\mathcal{B}(\mathbf{S}) = \mathcal{B}(X_\mathbf{i}, \mathbf{i} \in \mathbf{S})$ and $\mathcal{B}(\mathbf{S}') = \mathcal{B}(X_\mathbf{i}, \mathbf{i} \in \mathbf{S}')$ are the σ-fields generated by the random variables $X_\mathbf{i}$ with sites $\mathbf{i}$ and $\mathbf{i}'$ ranging over $\mathbf{S}$ and $\mathbf{S}'$, respectively. Let $d(\mathbf{S}, \mathbf{S}') := \min\{\|\mathbf{i} - \mathbf{i}'\| \,|\mathbf{i} \in \mathbf{S}, \mathbf{i}' \in \mathbf{S}'\}$ be the Euclidean distance between $\mathbf{S}$ and $\mathbf{S}'$, where $\|\mathbf{i}\| := (i_1^2 + i_2^2 + \cdots + i_N^2)^{1/2}$ stands for the Euclidean norm. We will assume that $X_\mathbf{i}$ satisfies the following mixing condition: There exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that whenever $\mathbf{S}, \mathbf{S}' \subset \mathbb{Z}^N$,

$$\begin{aligned}\alpha\big(\mathcal{B}(\mathbf{S}),\ \mathcal{B}(\mathbf{S}')\big) &= \sup\Big\{\big|P(AB) - P(A)P(B)\big|,\ A \in \mathcal{B}(\mathbf{S}),\ B \in \mathcal{B}(\mathbf{S}')\Big\} \\ &\leq h\big(\mathrm{Card}(\mathbf{S}), \mathrm{Card}(\mathbf{S}')\big)\, \varphi\big(d(\mathbf{S}, \mathbf{S}')\big), \end{aligned} \tag{2.1}$$

where $\mathrm{Card}(\mathbf{S})$ denotes the cardinality of $\mathbf{S}$. Here h is a symmetric positive function nondecreasing in each variable. Throughout the paper, assume that h satisfies either

$$h(n', n'') \leq \min\{n', n''\} \quad \text{or} \tag{2.2}$$

$$h(n', n'') \leq C_0\, (n' + n'' + 1)^{C_1}, \tag{2.3}$$

for some constants $C_0 > 0$ and $C_1 > 1$. If $h \equiv 1$, then $X_\mathbf{i}$ is called strongly mixing.

The function φ in above mixing assumption (2.1) satisfies the following assumption:

A1: There exist constant $0 < \rho < 1$ and some $C_2 > 0$ such that $\varphi(k) \leq C_2 \rho^k$ for all $k \geq 1$.

The exponential decaying rate for the mixing coefficient is relatively strong, but it is used in many places when one deals with exponential inequality for spatial processes. Withers (1981) has shown that autoregressive and more general nonlinear time series models are strongly mixing with exponential mixing rates under certain weak assumptions. Let $I_\mathbf{n}$ be a rectangular region defined by $I_\mathbf{n} = \{\mathbf{i} \in \mathbb{Z}^N, 1 \leq i_l \leq n_N,\ l = 1, 2, \ldots, N\}$ and $\mathbf{n} = (n_1, n_2, \ldots, n_N)$. Assume that we observe $\{X_\mathbf{i}\}$ on $I_\mathbf{n}$. Suppose $\{X_\mathbf{i}\}$ takes values in $\mathbb{R}$ and has density $f(x)$. The joint probability density for $X_{\mathbf{i}_1}$ and $X_{\mathbf{i}_2}$ is denoted with $f_{X_{\mathbf{i}_1}, X_{\mathbf{i}_2}}(x, y)$ satisfies the following assumption:

A2: $|f_{X_{\mathbf{i}_1}, X_{\mathbf{i}_2}}(x, y) - f(x)f(y)| \leq C$ for some constant C and for all $x, y, \mathbf{i}_1$ and $\mathbf{i}_2$.

We write $\mathbf{n} \to \infty$ if $\min\{n_1, n_2, \ldots, n_N\} \to \infty$ and $C_3 < |n_i/n_j| < C_4, i, j = 1, 2, \ldots, N$ for some constants $0 < C_3, C_4 < \infty$. Letter C will be used to denote constants whose values are unimportant and may vary from line to line. All limits are taken as $\mathbf{n} \to \infty$. Also let $\mathbf{n}_\pi = n_1 \cdot n_2 \cdots n_N$.

2.2. Wavelets and function spaces

Our estimators are constructed in terms of wavelets, which form an orthonormal basis of $L^2(\mathbb{R})$. Formally, let $\phi(x)$ and $\psi(x)$ be father and mother wavelets, having the following properties: ϕ and ψ are bounded and compactly supported, and $\int \phi = 1$. We call a wavelet ψ *r-regular* if ψ has r vanishing moments and r continuous derivatives. For the existence of these compactly supported wavelets with high order varnish moments and continuous derivative, see Daubechies (1992). The translations and dilations of above wavelets are defined as

$$\phi_{j_0k}(x) = 2^{j_0/2}\phi(2^{j_0}x - k), \qquad \psi_{jk}(x) = 2^{j/2}\psi(2^jx - k), \quad x \in \mathbb{R},\ j_0,\ j,\ k \in \mathbb{Z}.$$