Journal of Multivariate Analysis 166 (2018) 300–319

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Nonparametric density estimation for spatial data with wavelets

Johannes T.N. Krebs

Department of Statistics, University of California, Davis, One Shields Avenue, 95616, CA, USA

A R T I C L E I N F O

Article history:
Received 4 August 2017
Available online 4 April 2018

AMS 2010 subject classifications:
primary
62G07
62H11
65T60
secondary
65C40
60G60

Keywords:
Besov spaces
Density estimation
Hard thresholding
Rate of convergence
Spatial lattice processes
Strong spatial mixing
Wavelets

A B S T R A C T

Nonparametric density estimators are studied for d-dimensional, strongly spatial mixing data which are defined on a general N-dimensional lattice structure. We consider linear and nonlinear hard thresholded wavelet estimators derived from a d-dimensional multi-resolution analysis. We give sufficient criteria for the consistency of these estimators and derive rates of convergence in $L^{p'}$ for $p' \in [1, \infty)$. For this reason, we study density functions which are elements of a d-dimensional Besov space $B^s_{p,q}(\mathbb{R}^d)$. We also verify the analytic correctness of our results in numerical simulations.

1. Introduction

This article considers methods of nonparametric density estimation for spatially dependent data with wavelets. There is an extensive literature on the density estimation problem for iid data or time series. Recently, inference techniques for spatial data have gained importance because of their relevance in modern applications such as image analysis, forestry, epidemiology or geophysics. See the monographs of [11] and [23] for a systematic introduction on spatial data and random fields.

So far when working with random fields, the kernel method has been a popular tool both in regression and density estimation; see, e.g., [3,8,9,27,28]. More recently, Dabo-Niang and Yao [13] extended the kernel method to functional stationary random fields and estimated the spatial density with respect to a reference measure. Dabo-Niang et al. [12] proposed a kernel method in spatial density estimation which also allows for spatial clustering. Amiri et al. [1] studied asymptotic properties of a recursive version of the Parzen–Rozenblatt estimator.

While the kernel method is efficient if the density has unbounded support, it can sometimes suffer from boundary issues for densities with compact support. Furthermore, the kernel method typically requires the density to satisfy certain smoothness conditions like two-times continuous differentiability. In situations where the density function does not meet these requirements, the wavelet method is an alternative which often performs relatively well because it adapts

E-mail address: jtkrebs@ucdavis.edu.

https://doi.org/10.1016/j.jmva.2018.03.013

automatically to the regularity of the curve to be estimated. Wavelet estimators assume that the underlying curve belongs to a function space with certain degrees of smoothness. These smoothness parameters can be more subtle than the differentiability criterion from above and will be clarified later. The wavelet estimators do not depend on the smoothness parameters; nevertheless, they behave as if the true curve is known in advance and attain the optimal rates of convergence. This is in particular true for the hard thresholding estimator of Donoho et al. [19].

However, estimating the density of spatial data with wavelets has received little attention. Only the special case of time series has been thoroughly investigated: the wavelet method for density and regression estimators for multivariate and stationary time series is studied in [4,5,37,38]. In a recent article, Li [36] studied wavelet estimators for compactly supported one-dimensional Besov densities on stationary and strongly mixing random fields.

In the present article, we continue with these considerations for d-dimensional densities and study the linear and the hard thresholding estimator based on wavelets without assuming that they are isotropic. It is well-known that the hard thresholding estimator performs better than its linear analogue for certain densities in the one-dimensional setting. We will show a similar behavior for multivariate density functions.

The hard thresholding estimator has a linear basic component with respect to a coarse level j_0. Additionally, nonlinear details are added for higher levels $j_0 \le j \le j_1$ if their contribution is significant in the statistical sense. This implies that this estimator can converge faster than the linear estimator in certain parameter settings.

The generalization to arbitrary dimensions is non-trivial, in particular because the definitions of the underlying Besov space $B^s_{p,q}(\mathbb{R}^d)$ have to be generalized to the d-dimensional case. For isotropic wavelets there already exist such generalizations; see, e.g., [30,39]. However, as we also allow for density estimators with nonisotropic wavelets, we need a more general definition. This is one of the unique features of the present work. Moreover, we allow for density functions on $\mathbb{R}^d$ which do not necessarily have compact support.

We assume that $Z = \{Z(s) : s \in \mathbb{Z}^N\}$ is a random field with equal marginal laws on $\mathbb{R}^d$ which admit a square integrable density f with respect to the d-dimensional Lebesgue measure λ^d. Then for an orthonormal basis $\{b_u : u \in \mathbb{N}_+\}$ of $L^2(\lambda^d)$, we have the representation $f = \sum_{u\in\mathbb{N}_+} \langle f, b_u\rangle\, b_u$, where $\langle\cdot,\cdot\rangle$ is the inner product on the function space $L^2(\lambda^d)$. Since f is a density, we have the fundamental relationship between an observed sample $\{Z(s) : s \in I\}$ with $I \subseteq \mathbb{Z}^N$ and a coefficient $\langle f, b_u\rangle$ from this representation, viz.

$$\langle f, b_u\rangle = \mathrm{E}\,[\, b_u\{Z(s)\}\,] \approx \frac{1}{|I|}\sum_{s\in I} b_u\{Z(s)\}.$$

It is well-known that replacing the true coefficient with the empirical approximation yields a consistent density estimate for an iid sample of one-dimensional data under certain conditions, see, e.g., [18,29]. In the particular case of wavelets, Kerkyacharian and Picard [33] derived rates of convergence for the linear wavelet estimator.

In contrast to linear wavelet estimators, nonlinear wavelet estimators are particularly useful if the density curve features high-frequency oscillations or exhibits an erratic behavior. Rates of convergence of the hard thresholded wavelet estimator were studied in [19] and [25]. Since then the wavelet method for the density problem has been studied in various special settings: Hall et al. [24], Cai [7] and Chicken and Cai [10] considered rates of convergence for wavelet block thresholding. Giné and Nickl [22] gave several uniform limit theorems for wavelet density estimators for a compactly supported density and iid sample data. Xue [44] studied wavelet-based density estimation under censorship. Giné and Madych [21] investigated wavelet projection kernels in the density estimation problem. In this article, we continue the analysis for multivariate sample data which are spatially dependent.

This manuscript is organized as follows. We give the fundamental definitions and summarize the main facts of Besov spaces in d dimensions in Section 2. In Section 3 we study in detail the wavelet density estimators. We give criteria which are sufficient for the consistency of the nonparametric estimators and establish rates of convergence. Section 4 is devoted to numerical applications. We use an algorithm proposed by Kaiser et al. [31] for the simulation of the random field and estimate its marginal density with the linear and the hard thresholded wavelet estimator. Section 5 contains the proofs of the results from Section 3. Appendix A contains useful inequalities for dependent sums. As the wavelet estimators are a priori not necessarily a density, we consider in Appendix B the question under which circumstances a normalized estimator is consistent.

2. Notation and definitions

This section is divided in four parts. First, we introduce the concepts for multidimensional wavelets. Second, we define the multidimensional Besov spaces. Third, we explain the data generating process. Finally, we define the wavelet density estimator. In the following, we write $L^2(\lambda^d)$ for $L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda^d)$, where λ^d is the d-dimensional Lebesgue measure and we write

$$\|f\|_{L^p(\lambda^d)} = \left(\int_{\mathbb{R}^d} |f|^p \,\mathrm{d}\lambda^d\right)^{1/p}$$

for the L^p-norm of a function f on $\mathbb{R}^d$.

We begin with well-known results on wavelets in d dimensions; see, e.g., the monograph of [2].