Journal of Multivariate Analysis 151 (2016) 110–132

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Adaptive global thresholding on the sphere

Claudio Durastanti

Fakultät für Matematik, Ruhr Universität, Bochum, Germany

ARTICLE INFO

Article history:
Received 12 January 2016
Available online 3 August 2016

AMS 2010 subject classifications:
62G08
62G20
65T60

Keywords:
Global thresholding
Needlets
Spherical data
Nonparametric regression
U-statistics
Besov spaces
Adaptivity

ABSTRACT

This work is concerned with the study of the adaptivity properties of nonparametric regression estimators over the d-dimensional sphere within the global thresholding framework. The estimators are constructed by means of a form of spherical wavelets, the so-called needlets, which enjoy strong concentration properties in both harmonic and real domains. The author establishes the convergence rates of the L^p-risks of these estimators, focusing on their minimax properties and proving their optimality over a scale of nonparametric regularity function spaces, namely, the Besov spaces.

1. Introduction

The purpose of this paper is to establish adaptivity for the L^p-risk of regression function estimators in the nonparametric setting over the d-dimensional sphere $\mathbb{S}^d$. The optimality of the L^p risk is established by means of global thresholding techniques and spherical wavelets known as needlets.

Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be independent pairs of random variables such that, for each $i \in \{1, \dots, n\}$, $X_i \in \mathbb{S}^d$ and $Y_i \in \mathbb{R}$. The random variables $X_1, \dots, X_n$ are assumed to be mutually independent and uniformly distributed locations on the sphere. It is further assumed that, for each $i \in \{1, \dots, n\}$,

$$Y_i = f(X_i) + \varepsilon_i, \qquad (1)$$

where $f : \mathbb{S}^d \mapsto \mathbb{R}$ is an unknown bounded function, i.e., there exists $M > 0$ such that

$$\sup_{x\in\mathbb{S}^d} |f(x)| \le M < \infty. \qquad (2)$$

Moreover, the random variables $\epsilon_1, \dots, \epsilon_n$ in Eq. (1) are assumed to be mutually independent and identically distributed with zero mean. Roughly speaking, they can be viewed as the observational errors and in what follows, they will be assumed to be sub-Gaussian.

In this paper, we study the properties of nonlinear global hard thresholding estimators, in order to establish the optimal rates of convergence of L^p-risks for functions belonging to the so-called Besov spaces.

E-mail address: claudio.durastanti@gmail.com.

http://dx.doi.org/10.1016/j.jmva.2016.07.009

1.1. An overview of the literature

In recent years, the issue of minimax estimation in nonparametric settings has received considerable attention in the statistical inference literature. The seminal contribution in this area is due to Donoho et al. [7]. In this paper, the authors provide nonlinear wavelet estimators for density functions on $\mathbb{R}$, lying over a wide nonparametric regularity function class, which attain optimal rates of convergence up to a logarithmic factor. Following this work, the interaction between wavelet systems and nonparametric function estimation has led to a considerable amount of developments, mainly in the standard Euclidean framework; see, e.g., [3,5,24,26–28,30] and the textbooks [22,44] for further details and discussions.

More recently, thresholding methods have been applied to broader settings. In particular, nonparametric estimation results have been achieved on $\mathbb{S}^d$ by using a second generation wavelet system, namely, the spherical needlets. Needlets were introduced by Narcowich et al. [39,40], while their stochastic properties dealing with various applications to spherical random fields were examined in [2,6,34–36]. Needlet-like constructions were also established over more general manifolds by Geller and Mayeli [18–21], Kerkyacharian et al. [25] and Pesenson [41] among others, and over spin fiber bundles by Geller and Marinucci [16,17].

In the nonparametric setting, needlets have found various applications on directional statistics. Baldi et al. [1] established minimax rates of convergence for the L^p-risk of nonlinear needlet density estimators within the hard local thresholding paradigm, while analogous results concerning regression function estimation were established by Monnier [38]. The block thresholding framework was investigated in Durastanti [9]. Furthermore, the adaptivity of nonparametric regression estimators of spin function was studied in Durastanti et al. [10]. In this case, the regression function takes as its values algebraical curves lying on the tangent plane for each point on $\mathbb{S}^2$ and the wavelets used are the so-called spin (pure and mixed) needlets; see Geller and Marinucci [16,17].

The asymptotic properties of other estimators for spherical data, not concerning the needlet framework, were investigated by Kim and Koo [31–33], while needlet-like nearly-tight frames were used in Durastanti [8] to establish the asymptotic properties of density function estimators on the circle. Finally, in Gautier and Le Pennec [15], the adaptive estimation by needlet thresholding was introduced in the nonparametric random coefficients binary choice model. Regarding the applications of these methods in practical scenarios, see, e.g., [13,14,23], where they were fruitfully applied to some astrophysical problems, concerning, for instance, high-energy cosmic rays and Gamma rays.

1.2. Main results

Consider the regression model given in Eq. (1) and let $\{\psi_{j,k} : j \geq 0, k = 1, \ldots, K_j\}$ be the set of d-dimensional spherical needlets. Roughly speaking, j and K_j denote the resolution level j and the cardinality of needlets at the resolution level j, respectively. The regression function f can be rewritten in terms of its needlet expansion. Namely, for all $x \in \mathbb{S}^d$, one has

$$f(x) = \sum_{j\geq 0} \sum_{k=1}^{K_j} \beta_{j,k} \psi_{j,k}(x),$$

where $\{\beta_{j,k} : j \geq 0, k = 1, \ldots, K_j\}$ is the set of needlet coefficients.

For each $j \geq 0$ and $k \in \{1, \ldots, K_j\}$, a natural unbiased estimator for $\beta_{j,k}$ is given by the corresponding empirical needlet coefficient, viz.

$$\widehat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} Y_i \psi_{j,k}(X_i); \tag{3}$$

see, e.g., Baldi et al. [1] and Härdle et al. [22]. Therefore, the global thresholding needlet estimator of f is given, for each $x \in \mathbb{S}^d$, by

$$\hat{f}_n(x) = \sum_{j=0}^{J_n} \tau_j \sum_{k=1}^{K_{J_n}} \widehat{\beta}_{j,k} \psi_{j,k}(x), \tag{4}$$

where τ_j is a nonlinear threshold function comparing the given j-dependent statistic $\widehat{\Theta}_j(p)$, built on a subsample of $p < n$ observations, to a threshold based on the observational sample size. If $\widehat{\Theta}_j(p)$ is above the threshold, the whole j-level is kept; otherwise it is discarded.

Loosely speaking, this procedure allows one to delete the coefficients corresponding to a resolution level j whose contribution to the reconstruction of the regression function f is not clearly distinguishable from the noise. Following Kerkyacharian et al. [30], we consider the so-called hard thresholding framework, defined as

$$\tau_j = \tau_j(p) = \mathbb{1}\{\hat{\Theta}_j(p) \geq B^{dj} n^{-p/2}\},$$

where $p \in \mathbb{N}$ is even. Further details regarding the statistic $\hat{\Theta}_j(p)$ will be discussed in Section 3.4, where the choice of the threshold $B^{dj} n^{-p/2}$ will also be motivated.