Journal of Multivariate Analysis 160 (2017) 93–104

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Nonparametric estimation of a function from noiseless observations at random points

Benedikt Bauer [a], Luc Devroye [b], Michael Kohler [a], Adam Krzyżak [c,*], Harro Walk [d]

[a] Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstraße 7, 64289 Darmstadt, Germany
[b] School of Computer Science, McGill University, 3450, rue University, Montréal (QC), Canada H3A 2K6
[c] Department of Computer Science and Software Engineering, Concordia University, 1455, boul. de Maisonneuve ouest, Montréal (QC), Canada H3G 1M8,
[d] Fachbereich Mathematik, Universität Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 12 August 2016
Available online 19 June 2017

AMS subject classifications:
primary 62G05
secondary 62G20

Keywords:
Multivariate scattered data approximation
Rate of convergence
Supremum norm error

ABSTRACT

In this paper we study the problem of estimating a function from n noiseless observations of function values at randomly chosen points. These points are independent copies of a random variable whose density is bounded away from zero on the unit cube and vanishes outside. The function to be estimated is assumed to be (p, C)-smooth, i.e., (roughly speaking) it is p times continuously differentiable. Our main results are that the supremum norm error of a suitably defined spline estimate is bounded in probability by $\{\ln(n)/n\}^{p/d}$ for arbitrary p and d and that this rate of convergence is optimal in minimax sense.

1. Introduction

1.1. Multivariate scattered data approximation

Approximation problems in which the input data is a set of deterministic distinct points are so-called scattered data approximation problems which have been extensively studied in the literature. In a typical setting we are given a set of deterministic points $(x_1, y_1), \ldots, (x_n, y_n) \in [0, 1]^d \times \mathbb{R}$ and try to find a function m from a given function space, e.g., a Sobolev space, that fits the data closely. In scattered data approximation the points are not assumed to occupy a regular grid but rather are scattered around the space making the reconstruction problem difficult. The most popular approaches include the moving least squares approximation [6,10,15,18,32,33], schemes based on radial basis functions or constant functions on spheres [9,16,24–26], multiquadric interpolants [21] and the smoothing spline approach. The latter one can be posed as the regularized least squares problem where one minimizes the criterion

$$\sum_{i=1}^{n}\{m(x_i) - y_i\}^2 + \lambda\, \|m\|_H^2$$

* Corresponding author.
E-mail addresses: bbauer@mathematik.tu-darmstadt.de (B. Bauer), lucdevroye@gmail.com (L. Devroye), kohler@mathematik.tu-darmstadt.de (M. Kohler), krzyzak@cs.concordia.ca (A. Krzyżak), walk@mathematik.uni-stuttgart.de (H. Walk).

http://dx.doi.org/10.1016/j.jmva.2017.05.010
0047-259X/

over a class of functions H. The classes of functions include Beppo-Levi Space [9] and Reproducing Kernel Hilbert Space [17]. In the moving least squares approach we seek function m^* which is a solution of the minimization problem

$$\min_{m\in P}\left[\sum_{i=1}^{n}\{m(x_i)-y_i\}^2 w(x,x_i)\right], \tag{1}$$

where P is a finite-dimensional subspace (usually spanned by polynomials) of a space of continuous functions on a compact set Ω. Weight functions w are typically local, radial functions. It can be shown under mild conditions that the solution of problem (1) exists and is unique [32]. For the rate of approximation define the separation distance q_X and the mesh norm $h_{X,\Omega}$ as follows:

$$q_X=\frac{1}{2}\min_{1\le j<k\le n}\|x_j-x_k\| \quad \text{and} \quad h_{X,\Omega}=\sup_{x\in\Omega}\min_{j\in\{1,\ldots,n\}}\|x-x_j\|,$$

where $\|x\|$ denotes the Euclidean norm of $x\in\mathbb{R}^d$. Assume that a global constant c_1 exists such that the data separation condition

$$q_X\le h_{X,\Omega}\le c_1\, q_X \tag{2}$$

holds on the data set. Then under the condition that Ω is compact and satisfies the so-called cone condition we get for $f\in C^p(\Omega)$ the approximation bound $\|m-m^*\|_{\infty,\Omega}\le c_2\, h^p_{X,\Omega}$; see, e.g., [32,33]. Hence if $x_1,\ldots,x_n$ are scattered approximately evenly in $[0,1]^d$, we get

$$\|m-m^*\|_{\infty,[0,1]^d}\le c_3\, n^{-p/d}. \tag{3}$$

The approximation error bounds for the radial basis function interpolations may be found in [33] and [20].

1.2. The problem studied in this paper

In practice it is not clear, especially in high dimensions, at which locations a function should be sampled. A simple but effective way is to generate sampling points randomly from the uniform distribution on a ball or cube. The rest of the paper will be devoted to estimation of an unknown function m observed at such random scattered data. Our main question is how the error bound in (3) changes in this case. Obviously the result in (3) is not applicable in this case since condition (2) does not hold. Nevertheless it is natural to conjecture that a bound similar to (3) should hold for suitably defined estimates, even if the data points are randomly and not deterministically distributed. However, it is not clear how the definition of the estimates should be changed in order to be able to show such a result.

To formulate our problem precisely, let $X,X_1,\ldots,X_n$ be independent and identically distributed random variables with values in $[0,1]^d$ and let $m:[0,1]^d\to\mathbb{R}$ be a (measurable) function. Given the data $\mathcal{D}_n=\{(X_1,m(X_1)),\ldots,(X_n,m(X_n))\}$, we are interested in constructing an estimate $m_n=m_n(\cdot,\mathcal{D}_n):\mathbb{R}^d\to\mathbb{R}$ such that the supremum norm error

$$\|m_n-m\|_{\infty,[0,1]^d}=\sup_{x\in[0,1]^d}|m_n(x)-m(x)|$$

is small.

1.3. Main results

It is well-known that we need smoothness assumptions on m in order to derive nontrivial results on the rate of convergence of the global error of a function estimate (see, e.g., [8], Theorem 3.1). In the sequel we assume that m is (p,C)-smooth for some $p=k+s$ for some $k\in\mathbb{N}_0$, $s\in(0,1]$ and $C>0$, i.e., (roughly speaking, see below for the exact definition) it is p-times continuously differentiable. Furthermore we will assume throughout this paper that there exists a constant $c_4>0$ such that

$$\Pr\{X\in S_r(x)\}>c_4\, r^d$$

does hold for all $x\in[0,1]^d$ and all $0<r\le 1$, where $S_r(x)$ denotes the (closed) ball of radius r around x. (This condition is in particular satisfied if X has a density with respect to the Lebesgue–Borel measure which is bounded away from zero on $[0,1]^d$.) We will show that in this case we can construct a spline estimate $m_n=m_n(\cdot,\mathcal{D}_n)$ such that

$$\|m_n-m\|_{\infty,[0,1]^d}=O_{\mathbf{P}}\left[\left\{\ln(n)/n\right\}^{p/d}\right], \tag{4}$$

where we write $Z_n=O_{\mathbf{P}}(Y_n)$ if the nonnegative random variables Z_n and Y_n satisfy $\lim_{c\to\infty}\limsup_{n\to\infty}\Pr(Z_n>c\,Y_n)=0$. Furthermore we show that the above rate of convergence is optimal in some minimax sense.