Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 77 (2007) 693–703
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

A note on uniform consistency of monotone function estimators

Natalie Neumeyer

Department Mathematik, Schwerpunkt Stochastik, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany

Received 4 October 2005; received in revised form 25 October 2006; accepted 26 November 2006
Available online 4 January 2007

Abstract

Recently, Dette et al. [A simple nonparametric estimator of a strictly increasing regression function. Bernoulli 12, 469–490] proposed a new monotone estimator for strictly increasing nonparametric regression functions and proved asymptotic normality. We explain two modifications of their method that can be used to obtain monotone versions of any nonparametric function estimators, for instance estimators of densities, variance functions or hazard rates. The method is appealing to practitioners because they can use their favorite method of function estimation (kernel smoothing, wavelets, orthogonal series, etc.) and obtain a monotone estimator that inherits desirable properties of the original estimator. In particular, we show that both monotone estimators share the same rates of uniform convergence (almost sure or in probability) as the original estimator.

MSC: 62G05

Keywords: Function estimator; Kernel method; Monotonicity; Uniform convergence

1. Introduction

During the last decades much effort has been devoted to the problem of estimating monotone functions. Estimating a monotone density function was considered by Grenander (1956), Groeneboom (1985), Groeneboom and Lopuhaä (1993), Datta (1995), Cheng et al. (1999), and van der Vaart and van der Laan (2003), among others. Even more literature can be found about estimating increasing regression functions, starting with Brunk (1958), Barlow et al. (1972), Mukerjee (1988), Mammen (1991), Ramsay (1988), and Hall and Huang (2001), among many others. For censored data Huang and Zhang (1994) and Huang and Wellner (1995) consider estimators for a monotone density and monotone hazard rate. For monotone estimators of a hazard rate see also Mukerjee and Wang (1993) and Hall et al. (2001).

Appealing to users of common kernel methods is a new method proposed by Dette et al. (2006) for nonparametric regression functions and by Dette and Pilz (2007) for variance functions in nonparametric regression models. The considered estimator is easy to implement, is based on kernel estimators, and, in contrast to many other procedures, does not require any optimization over function spaces. To obtain a monotone estimator for a strictly increasing function g (here $g:[0,1]\to\mathbb{R}$ denotes the regression or variance

E-mail address: neumeyer@math.uni-hamburg.de.

0167-7152/$ - see front matter
doi:10.1016/j.spl.2006.11.004

function), the method consists of first monotonically estimating the distribution function of $g(U)$, i.e. $h(t) = P(g(U) \leqslant t)$, by a kernel method, where U is uniformly distributed in $[0, 1]$. The first step uses a (not necessarily increasing) kernel estimator $\hat{g}$ for g. More precisely, the estimator for h is an integrated kernel density estimator,

$$\hat{h}(t) = \int_{-\infty}^{t} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{a} k\left(\frac{x - \hat{g}(i/N)}{a}\right) \mathrm{d}x, \tag{1.1}$$

where k denotes a density function, $a = a_N = o(1)$ a sequence of bandwidths and N converges to infinity. Noting that $h(t) = g^{-1}(t)$, an increasing estimator for g is then obtained by inversion of $\hat{h}$. Asymptotic normality of the constrained estimator is shown in Dette et al. (2006). An alternative method to obtain the estimator for g^{-1} is mentioned but not further developed in the aforementioned references, namely using

$$\hat{h}(t) = \int_{0}^{1} I\{\hat{g}(x) \leqslant t\} \, \mathrm{d}x \tag{1.2}$$

(where I denotes the indicator function) as an estimator for $\int_0^1 I\{g(x) \leqslant t\} \, \mathrm{d}x = g^{-1}(t)$ (where g is increasing). Note that Dette et al.'s (2006) proof for the asymptotic distribution of $\hat{h}$ defined in (1.1) and its inverse is not easily generalized to obtain asymptotic results about the estimator based on (1.2). The approach to use the inverse $\hat{h}^{-1}$ as an estimator for g, where $\hat{h}$ is defined in (1.2) is related to nondecreasing rearrangements of data considered by Ryff (1965, 1970), and is in principle similar to Polonik's (1995, 1998) work, who constructs estimators for a density f from the identity $f(x) = \int_0^\infty I\{f(x) \geqslant t\} \, \mathrm{d}t$. Here, the density contour clusters $\{x : f(x) \geqslant t\}$ are estimated by the so-called excess mass approach. By choosing the class of sets appropriately, for example, monotone density estimators are obtained. In this case the estimator coincides with Grenander's (1956) estimator. In a more general context, Polonik (1995) shows L^1-consistency of the obtained estimators. The approach is related to the estimation of density level sets, see Tsybakov (1997), among others.

In the paper at hand properties of the two methods (using the inverse of (1.1) or (1.2), respectively, as a monotone estimator of g) will be compared. Both methods are not restricted to monotone estimation of regression or variance functions, neither to the case of kernel or local linear estimators used in the first step. These restricted cases were considered in Dette et al. (2006) to prove asymptotic normality of the new estimators and first order equivalence to the unconstrained estimator. In these references it was also crucial to assume the function g to be strictly increasing with positive derivative. Here we consider the general case to modify any function estimator (using kernels, local polynomials, nearest neighbors, wavelets, splines, orthogonal series, etc.) for any function (density, regression function, variance function, hazard function, etc.) with compact support (or support bounded on one side) to obtain a monotone (either nondecreasing or strictly increasing) estimator. The estimators do not need to be based on an independent and identically distributed sample but can be based on dependent observations such as time series, or on censored observations. Also the original estimators are not supposed to be nonparametric but can be either non-, semi- or parametric. We only assume knowledge about uniform consistency of the original estimator used in the first step.

Both procedures (based on (1.1) and (1.2)) to obtain monotone versions of any function estimator are explained in detail in Section 2. We will show that the monotone modifications of the estimator share the same rates of uniform convergence (almost sure or in probability) as the original unconstrained estimator, see Section 3. Some examples of applications are also given in Section 3 and the details of the proofs are deferred to Section 4.

Throughout the text we call a function g *nondecreasing* provided that $x < y$ implies that $g(x) \leqslant g(y)$ and *increasing* provided that $x < y$ implies that $g(x) < g(y)$. Further, $f|_A$ denotes the function f with domain restricted to the set A.

2. Monotone modifications of function estimators

We explain in the following the method to obtain a monotone modification of any function estimator $\hat{g}$ of an unknown function g, where g is nondecreasing. We restrict ourselves first to the case of a compact support