Sharp asymptotics for isotonic regression

Received: 20 September 1999 / Revised version: 10 May 2001 / Published online: 19 December 2001 – © Springer-Verlag 2001

Abstract. The asymptotic behavior of the isotonic estimator of a monotone regression function (that is the least-squares estimator under monotonicity restriction) is investigated. In particular it is proved that the $L_1$-distance between the isotonic estimator and the true function is of magnitude $n^{-1/3}$. Moreover, it is proved that a centered version of this $L_1$-distance converges at the $n^{1/2}$ rate to a Gaussian variable with fixed variance.

1. Introduction

The problem of estimating a monotone function $f$ in a non-parametric framework has been raised in both cases when $f$ is a density function or a regression function. Grenander (1956) considered the problem of density estimation while Brunk (1958) studied the problem of regression estimation. The estimators they suggested are defined as follows in the case when $f$ is non-increasing.

- Suppose that $f$ is a density function, and suppose we are given i.i.d. random variables $X_1, \ldots, X_n$ with common density function $f$. The so-called “Grenander estimator” of $f$ is defined as the maximum likelihood estimator under the order restriction that $f$ is non-increasing.
- Suppose we are given the regression model $Y_i = f(x_i) + \epsilon_i$, $1 \leq i \leq n$, where the $x_i$’s are fixed real numbers with $x_1 \leq \ldots \leq x_n$ and the $\epsilon_i$’s are i.i.d. random variables with zero mean. Brunk defined the “isotonic estimator” of $f$ as the least-squares estimator under the order restriction that $f(x_1) \geq \ldots \geq f(x_n)$.

It is shown in Grenander and Brunk’s papers that the Grenander and the isotonic estimators are respectively an histogram and a regressogram estimator; they are entirely data-driven and do not require any choice of smoothing parameters. One can find in the literature several algorithms for building the Grenander and the isotonic estimators, see e.g. the so-called “Pool-Adjacent-Violators” algorithm described by Barlow et al (1972). Moreover the Grenander (resp. isotonic) estimator is locally adaptive in the sense that it is as efficient as the best histogram (resp. regressogram) estimator, see Proposition 2.2 of Reboul (1997). These properties make the Grenan-
der and the isotonic estimators attractive in comparison with usual non-parametric estimators such as kernel estimators, regular histogram or regular regressogram estimators. Indeed, kernel, regular histogram and regressogram estimators require the choice of bandwidths that are often chosen in an arbitrary way.

Pointwise convergence of both Grenander estimator and isotonic estimator has been studied: the asymptotic distribution of the difference at a fixed point between the estimator and the function to be estimated is known, see Prakasa Rao (1969) for the density estimation problem, Brunk (1970) and Wright (1981) for the regression estimation problem. But until recently, very few was known about global convergence of these estimators. In the framework of density estimation, Groeneboom (1985) stated a central limit theorem for the $L_1$-distance between the density function to be estimated and its Grenander estimator: a centered version of this distance converges at the $n^{1/2}$ rate to a Gaussian variable with fixed variance. Only a sketch of proof was given then: it is only recently that Groeneboom et al. (1999) established a complete proof of this result. Their new proof essentially does not rely on the sketch suggested in 1985. It has been established simultaneously and independently of our own paper.

The aim of this paper is to study the global convergence in the $L_1$-distance sense, of the isotonic estimator of a monotone regression function $f$. We focus on the non-increasing case. The non-decreasing case can be deduced from that case by changing $f$ into $-f$ since the isotonic estimator of $-f$ is $-\hat{f}_n$, where $\hat{f}_n$ is the isotonic estimator of $f$, see Barlow et al (1972). The following regression model is investigated

$$Y_i = f(x_i) + \varepsilon_i, \ 1 \leq i \leq n$$

where the design points $x_i$ are fixed ($x_i = i/n$) and the $\varepsilon_i$'s are independent and identically distributed random variables with mean zero and known variance $\sigma^2$. The unknown regression function $f$ is known to be decreasing and smooth. Our main result is Theorem 2 below, where it is stated that a centered version of the $L_1$-distance between $f$ and the isotonic estimator of $f$ is asymptotically Gaussian. The rate of convergence of this $L_1$-loss is $n^{1/2}$ and the asymptotic variance does not depend on $f$. Our main motivation is to build a Goodness of fit test for a monotone regression function. Theorem 2 provides such a statistical test even if $\sigma^2$ is unknown (if $\sigma^2$ is unknown, one can replace it by some adequate estimator). The power of this test is studied by Durot and Tocquet (1998).

Here are the main steps of the proof of Theorem 2. A central limit theorem for the $L_1$-loss of the non-parametric least-squares estimator of a decreasing, smooth signal function is stated in the framework of white noise model (see Theorem 3). Thanks to a strong approximation argument (which is a refinement due to Sakhanenko of the Komlós, Major and Tusnády’s construction) a white noise model with signal function $f$, that approaches our regression model is then built. More specifically it is proved that if there exists some large enough $p$ for which $\mathbb{E}|\varepsilon_i|^p$ is finite then the isotonic estimator $\hat{f}_n$ of $f$ obtained in the regression model and the corresponding estimator $f_n$ obtained in the white noise model are asymptotically