Journal of Computational and Applied Mathematics 315 (2017) 17–31

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

On the moment-recovered approximations of regression and derivative functions with applications

Robert M. Mnatsakanov *, Broti Garai

Department of Statistics, P.O. Box 6330, West Virginia University, Morgantown, WV 26506, USA

ARTICLE INFO

Article history:
Received 7 April 2016
Received in revised form 25 October 2016

Keywords:
Moment-determinate functions
Laplace transform inversion
Nonparametric estimation

ABSTRACT

In this paper three formulas for recovering the conditional mean and conditional variance based on product moments are proposed. The upper bounds for the uniform rate of approximations of regression and derivatives of some moment-determinate function are derived. Two cases where the support of underlying functions is bounded and unbounded from above are studied. Based on the proposed approximations, novel nonparametric estimates of the distribution function and its density in multiplicative-censoring model are constructed. Simulation study justifies the consistency of the estimates.

1. Introduction

In many statistical problems the goal is to express the response Y variable as a function of the predictive variable X, i.e., $Y = r(X)$. Suppose instead of observing the variables (X, Y), we only have information about the averages/moments of the form: $m(k, j) = E(Y^k X^j)$, with $k = 0, 1, 2$ and $j = 0, 1, \ldots, \alpha$. The objective of this paper is to recover regression function $r(x) = E(Y|X = x)$ and conditional variance $\sigma^2(x) = \mathrm{Var}(Y|X = x)$ given the information contained in the sequence of product moments $m(k, j)$, $k = 0, 1, 2$ and $j = 0, 1, \ldots, \alpha$. Also we address the problem of recovering the derivative of a density function of some random variable, say $Y \sim g$, given only its moments $E(Y^j)$ for $j = 0, 1, \ldots, \alpha$. Application in the multiplicative-censoring model is outlined as well. In particular, we derived the consistent nonparametric estimates of associated distribution and its density function.

In this paper it is assumed that the regression function that is defined on the positive half line is moment-determinate (M-determinate). The conditions under which a function is M-determinate or M-indeterminate is a classical mathematical problem and has been studied in many works, see, for example, [1–4] among others. Besides, there are several techniques that provide the reconstruction of probability density function from the sequence of its moments. Let us mention here only three of them. The maximum entropy method was used in [5–7,4], to reconstruct the unknown density from its moments. The article [8] showed that under certain conditions the density function can be approximated by $\boldsymbol{f}_d(x) = \psi(x) \sum_{\ell=0}^{d} \boldsymbol{\xi}_\ell\, x^\ell$. Here $\psi(x)$ is the initial density approximation of $f(x)$, the $\boldsymbol{\xi}_i$'s are determined by equating the first d moments obtained from $\boldsymbol{f}_d(x)$ to those of X, i.e. via

$$\begin{bmatrix} \boldsymbol{\xi}_0 \\ \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_d \end{bmatrix} = \begin{bmatrix} m(0) & m(1) & \cdots & m(d) \\ m(1) & m(2) & \cdots & m(d+1) \\ \cdots & \cdots & \cdots & \cdots \\ m(d) & m(d+1) & \cdots & m(2d) \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \boldsymbol{\mu}_X(1) \\ \vdots \\ \boldsymbol{\mu}_X(d) \end{bmatrix}.$$

* Corresponding author.
E-mail address: rmnatsak@stat.wvu.edu (R.M. Mnatsakanov).

http://dx.doi.org/10.1016/j.cam.2016.10.023

In the previous equation $\boldsymbol{\mu}_X(j) = E(\boldsymbol{X}^j)$ and $m(j) = \int_a^b x^j \psi(x)dx$.

In the framework of moment problems (Hausdorff, Stieltjes and Hamburger), (see [9–12]) a method for recovering the distribution function (df) as well as corresponding density is proposed. These reconstructions are based on the sequence of ϕ-transformed moments $\{m_\phi(j)\}_{j=0}^{\alpha}$ of F up to some order α.

The approximated df and density can be written in a unified form: for each $\alpha \in \mathbb{N}$, $F_{\alpha,\phi} := \mathscr{K}_\alpha^{-1} m_\phi$ and $f_{\alpha,\phi} := \mathscr{B}_\alpha^{-1} m_\phi$ where

$$\left(\mathscr{K}_\alpha^{-1} m_\phi\right)(x) = \sum_{k=0}^{[\alpha\phi(x)]} \sum_{j=k}^{\alpha} \binom{\alpha}{j}\binom{j}{k}(-1)^{j-k} m_\phi(j) \tag{1}$$

and

$$\left(\mathscr{B}_\alpha^{-1} m_\phi\right)(x) = \frac{\Gamma(\alpha+2)|\phi'(x)|}{\Gamma([\alpha\phi(x)]+1)} \sum_{j=0}^{\alpha-[\alpha\phi(x)]} \frac{(-1)^j m_\phi(j+[\alpha\phi(x)])}{j!(\alpha-[\alpha\phi(x)]-j)!}, \quad x \in S. \tag{2}$$

Here, $m_\phi(j) = \int_S [\phi(t)]^j dF(t)$, for $j = 0, 1, \ldots, \alpha$, and $S := \text{supp}\{F\}$ is a support of F. It is assumed that the map $\phi : S \mapsto [0, 1]$ is specified according to the form of S. For example, $\phi(x) = x/T$, if $S = [0, T]$, and $\phi(x) = b^{-x}$ if $S = \mathbb{R}_+$, for some $0 < T < \infty$ and $b > 1$. The recovered constructions $F_{\alpha,\phi}$ and $f_{\alpha,\phi}$ provide stable approximates of F and f, respectively (see [12]).

It is worth mentioning that the first two methods were applied only for reconstruction of a density f while the last one can be used for recovering arbitrary M-determinate function (e.g., the derivative) as well (see [12] and Example 5 in Section 3). This fact enables us to apply the construction (2) to approximate the regression function, the conditional variance as well as the derivative function.

Under very smooth conditions on underlying regression function r, [13] approximated the conditional expectation $E(Y|X = x)$ using the sequence of moments $m(1, j)$ for $j = 0, 1, \ldots, 2n$. Namely, when the support of r is finite, say [0,1], the following approximation was introduced:

$$r_n(x) = \frac{1}{f(x)} \sum_{j=0}^{2n} \nu_{[n(M_n+x)/M_n],j} \frac{n_{j+1}}{M_n^{j+1}} m(1, j), \quad 0 \le x \le 1. \tag{3}$$

Here $\nu_{i,j}$ is the (i, j)th entry of the inverted Vandermonde matrix $V(-n, -n+1, \ldots, n)$ [14], $f(x) = F'(x)$ is the marginal density of X and M_n is a sequence of positive real numbers such that $M_n/n \to 0$ as $n \to \infty$. The author showed that the reconstruction of $r(x)$ based on (3) provides an improved approximation if compared to normal approximation.

The article is organized as follows. In Section 2 two cases of recovering regression function are investigated: (a) the distribution of a predictive variable X is known and (b) the distribution of X is unknown. In Section 3 approximation of the derivative function g' for some continuous density function g is provided utilizing the proposed construction. The upper bounds for the uniform rate of convergence for the approximated regression and the derivative functions are derived and asymptotic behavior of the distances between the approximated and the true regression functions is investigated. Furthermore, in Section 3 we address the problem of estimating the distribution function F and corresponding density f in the framework of multiplicative-censoring model. In Sections 2 and 3 the simulation study and a comparison with other constructions are conducted. Graphical illustrations and tables with the values of errors of recovered functions are provided. The mean squared error and integrated mean squared error of associated nonparametric estimators will be the subject of investigation in a separate work.

2. Approximations of $r(x)$

2.1. Distribution of X is known

In this subsection we will approximate regression function $r(x)$ when the distribution of X is known. In addition, we assume that the distribution function F of X possesses a finite support, i.e., $\text{supp}\{F\} = [0, T]$; $0 < T < \infty$. For the sake of simplicity we assume $T = 1$. Two cases are considered in this subsection: (1) X follows the uniform distribution on (0,1), i.e., $X \sim U(0, 1)$ (see Example 1), and (2) $X \sim \text{Beta}(a, b)$ with known parameters a and b (see Example 2). The density of X is denoted by $\beta(x, a, b) = \frac{1}{\text{B}(a,b)} x^a (1-x)^{b-1}$ for $x \in (0, 1),\ a > 0,\ b > 0$. Here $\text{B}(a, b)$ is the classical Euler-beta function.

2.1.1. Model 1: $X \sim U(0, 1)$

Let us assume $X \sim U(0, 1)$ with some joint probability density function of (X, Y) denoted by $h : [0, 1]^2 \to \mathbb{R}_+$. Denote $m_k = \{m(k, j), j = 0, 1, \ldots, \alpha\}$ where $k = 0, 1, 2$, and

$$m(k, j) = E(Y^k X^j) = \int_0^1 \int_0^1 y^k x^j h(y, x)\, dy\, dx.$$