Journal of Statistical Planning and Inference 178 (2016) 1–23

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Adaptive deconvolution of linear functionals on the nonnegative real line

Gwennaëlle Mabon *

CREST - ENSAE, 3 avenue Pierre Larousse, 92245 Malakoff, France
MAP5, Université Paris Descartes, 45 ruedes Saints-Pères, 75006 Paris, France

ARTICLE INFO

Article history:
Received 22 April 2016
Received in revised form 22 April 2016
Accepted 30 April 2016
Available online 13 May 2016

MSC:
primary 62G07
secondary 62G99
62J99

Keywords:
Deconvolution
Nonparametric density estimation
Adaptive estimation
Linear functionals
Laguerre basis
Mean squared risk

ABSTRACT

In this paper we consider the convolution model $Z = X + Y$ with X of unknown density f, independent of Y, when both random variables are nonnegative. Our goal is to estimate linear functionals of f such as $\langle \psi, f \rangle$ for a known function ψ assuming that the distribution of Y is known and only Z is observed. We propose an estimator of $\langle \psi, f \rangle$ based on a projection estimator of f on Laguerre spaces, present upper bounds on the quadratic risk and derive the rate of convergence in function of the smoothness of f, g and ψ. Then we propose a nonparametric data driven strategy, inspired Goldenshluger and Lepski (2011) method to select a relevant projection space. This methodology permits to estimate the cumulative distribution function of X for instance. In addition it is adapted to the pointwise estimation of f. We illustrate the good performance of the new method through simulations. We also test a new approach for choosing the tuning parameter in Goldenshluger–Lepski data driven estimators following ideas developed in Lacour and Massart (2015).

1. Introduction

In many experiments statisticians do not observe directly the variable of interest X; instead they have at hand observations of Z, equal to the sum of X and another random variable Y. In various situations, Y can modelize a measurement error, and as such, is assumed to be symmetric or centered. But we can also, in reliability fields, observe the sum of the lifetimes of two components, the second one being well known. In survival analysis, X can be the time of infection of a disease and Y the incubation time, and this happens in the so called back calculation problems in AIDS research. In these last two cases, distributions of X and Y are $\mathbb{R}^+$-supported. In this situation, Y is not considered as a noise but as an additional nuisance process. Indeed a noise distribution is assumed to be centered, which is not the case anymore. Thus we consider the following model

$$Z_i = X_i + Y_i, \quad i = 1, \dots, n, \tag{1}$$

where the X_i's are independent identically distributed (i.i.d.) nonnegative random variables (r.v.) with unknown density f. The Y_i's are also i.i.d. nonnegative variables with known density g. We denote by h the density of the Z_i's. Moreover the X_i's and the Y_i's are assumed to be independent, they are not observed. Our goal is to estimate linear functionals of f defined by $\vartheta(f) = \langle \psi, f \rangle = \mathbb{E}[\psi(X_1)]$ with ψ a known function, from observations $Z_1, \dots, Z_n$.

* Correspondence to: CREST - ENSAE, 3 avenue Pierre Larousse, 92245 Malakoff, France.
E-mail address: gwennaelle.mabon@ensae.fr.

http://dx.doi.org/10.1016/j.jspi.2016.04.006
0378-3758/

Those assumptions imply that, in Model (1), $h(x) = (f \star g)(x)$ where $(\varphi \star \psi)(x) = \int \varphi(x-u)\psi(u)\,du$ denotes the convolution product. This setting matches convolution models which is a classical topic in nonparametric statistics. The problem of recovering the signal distribution f when it is observed with an additive noise with known error distribution on the real line, has been extensively studied, see for rates of convergence and their optimality for kernel estimators (Carroll and Hall, 1988; Fan, 1991), for wavelets strategy (Pensky and Vidakovic, 1999), for projection strategies with penalization (Comte et al., 2006; Butucea and Tsybakov, 2008a,b) for the study of sharp asymptotic optimality; with an unknown error density, see Neumann (1997) for kernel estimator and minimax optimality, Johannes (2009) for minimax optimality under various regularity conditions on f, Comte and Lacour (2011) and Kappus and Mabon (2014) for projection strategies with penalization.

The problem of one-sided error in the convolution model has been first introduced by Groeneboom and Wellner (1992) under a constraint of monotonicity of the cumulative distribution function (c.d.f.). They concentrate on deriving nonparametric maximum likelihood estimators (NPMLE) of the c.d.f. Some particular cases have been tackled as uniform or exponential deconvolution by Groeneboom and Jongbloed (2003) and Jongbloed (1998) who propose NPMLE of the c.d.f. of the X_i's, which have explicit expressions. For other cases van Es et al. (1998) circumvent the lack of explicit expression for the NPMLE by proposing an isotonic inverse estimator. In this paper, we do not use the approach of the NPLME.

Moreover Model (1) is also related to the field of mixture models, see Roueff and Rydén (2005) and Rebafka and Roueff (2010) who study in particular mixtures of Exponential and Gamma distributions, which are contained in our framework. These models play a major role in natural sciences phenomena of discharge or disexcitation as in radioactive decays, the electric discharge of a capacitor or the temperature difference between two objects.

On one hand the problem of estimating linear functionals in linear models has been widely studied especially in the setting of the white noise model, see Cai and Low (2003, 2005) and Laurent et al. (2008) for instance. On the other hand adaptive estimation of linear functionals has not been much studied in the convolution model. Butucea and Comte (2009), in the context of Model (1) when the variables X_i's and Y_i's are $\mathbb{R}$-supported, propose a general estimator of $\vartheta(f)$ using a spectral cut-off in the Fourier domain when the random variables are distributed on the real line. They apply it to the pointwise estimation of the density on the real line and prove that their losses in their adaptive procedure is optimal in the minimax sense. They do not prove it for their general estimator. They also apply their adaptive procedure to the pointwise Laplace transform estimation and the stochastic volatility model. Recently Pensky (2015) has improved their results by deriving minimax lower bounds for estimators of a general linear functional of the deconvolution density. The author even extends the techniques when ψ is not integrable or square integrable and considers the possibility that the vector of observations is sparse (i.e. has a lot of zeros).

In this paper, our goal is to establish a specific method for the estimation of linear functional when we know that the random variables are $\mathbb{R}^+$-supported. In that way we can cite the work of Mabon (2014) who proposes a specific estimation for nonnegative variables in Model (1). The methodology, based on a penalized projection strategy in a Laguerre basis, allows to estimate the density and survival function of X. Moreover this work has showed that for certain classes of density, the Laguerre estimation in the convolution model (1) gives faster rates of convergence of the estimators than with the classical method based on Fourier inversion as obtained in Comte et al. (2006) for instance. In particular, it is verified for mixed Gamma distributions. The contribution of this paper is to extend the particular methodology developed for nonnegative variables to estimate linear functionals of f. Thanks to this new methodology, we can derive the pointwise estimation of the probability density function (p.d.f.) f, of the c.d.f. of X and also its Laplace transform.

In Section 2, we explain how projection coefficient estimators in the Laguerre basis can be used to define estimators of linear functionals. Next we lead a theoretical study of estimators of $\vartheta(f)$ and derive upper bound on the quadratic risk and rates of convergence in function of the smoothness of f, g and ψ when these three functions belong to $\mathbb{L}^2(\mathbb{R}^+)$. We show that under some assumptions on g or ψ, the parametric rate of estimation can be achieved. In Section 3 we propose a nonparametric data driven strategy, following Goldenshluger and Lepski (2011) method, for selecting a relevant projection space. In Section 4, when f and g belong to $\mathbb{L}^2(\mathbb{R}^+)$ we adapt the procedure to pointwise estimation of f in the setting of Model (1) in this case ψ is specified and do not belong to $\mathbb{L}^2(\mathbb{R}^+)$ anymore. The method is then illustrated through simulations. Following Lacour and Massart (2015), we apply a new procedure to choose the tuning parameter appearing in the penalization term of the data-driven estimator. This procedure is promising and shows good results.

To sum up the paper is organized as follows. In Section 2, we give the notations, specify the statistical model and estimation procedures for projection estimators of f and $\vartheta(f)$, upper bound on the pointwise mean squared error. In Section 3, we propose a new data-driven procedure for choosing the tuning parameter linear functionals of f which allows in particular to estimate the cumulative distribution function of X. In Section 4, we derive an adaptive procedure for the pointwise estimation of f and provide an empirical study on simulations. All the proofs are postponed to Section 5.

2. Statistical model and estimation procedure

2.1. Notations

For two real numbers a and b, we denote $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. For two functions $\varphi, \psi : \mathbb{R} \to \mathbb{R}$ belonging to $\mathbb{L}^2(\mathbb{R})$, we denote $\|\varphi\|$ the $\mathbb{L}^2$ norm of φ defined by $\|\varphi\|^2 = \int_{\mathbb{R}} |\varphi(x)|^2 dx$, $\langle \varphi, \psi \rangle$ the scalar product between φ and ψ defined by $\langle \varphi, \psi \rangle = \int_{\mathbb{R}} \varphi(x)\psi(x)dx$.