Journal of Multivariate Analysis 100 (2009) 2364–2375

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Asymptotic normality and confidence intervals for inverse regression models with convolution-type operators

Nicolai Bissantz *, Melanie Birke

Fakultät für Mathematik, Ruhr-Universität Bochum, Germany

ARTICLE INFO

Article history:
Received 3 July 2008
Available online 21 April 2009

AMS subject classifications:
primary 62G08
secondary 62G15
62G20

Keywords:
Bootstrap
Inverse problems
Model selection
Testing

ABSTRACT

We consider inverse regression models with convolution-type operators which mediate convolution on $\mathbb{R}^d$ ($d \geq 1$) and prove a pointwise central limit theorem for spectral regularisation estimators which can be applied to construct pointwise confidence regions. Here, we cope with the unknown bias of such estimators by undersmoothing. Moreover, we prove consistency of the residual bootstrap in this setting and demonstrate the feasibility of the bootstrap confidence bands at moderate sample sizes in a simulation study.

1. Introduction

Suppose that we have observations $(\mathbf{z_k}, Y_\mathbf{k})$, $\mathbf{k} = (k_1, \ldots, k_d) \in \{-n, \ldots, n\}^d$, from the model

$$Y_\mathbf{k} = g(\mathbf{z_k}) + \epsilon_\mathbf{k}, \tag{1}$$

where $g = \Psi * \theta$ is a one-to-one convolution operator with a function Ψ, the $\mathbf{z_k} = \left(\frac{k_1}{na_n}, \ldots, \frac{k_d}{na_n}\right)$ are fixed design points, the $\epsilon_\mathbf{k}$'s are i.i.d. errors with $E\epsilon_\mathbf{k} = 0$, $E\epsilon_\mathbf{k}^2 = \sigma^2$ ($\mathbf{k} = (k_1, \ldots, k_d)$), and a_n is a sequence which converges asymptotically to zero. The observable signal g can be represented as the image of the signal θ under the operator

$$(\mathcal{K}\theta)(z) = \int_{\mathbb{R}^d} \Psi(z - \mathbf{t})\theta(\mathbf{t})\mathrm{d}\mathbf{t}.$$

Recovery of the signal θ from the data $(\mathbf{z_k}, Y_k)$ in model (1) is a statistical inverse problem (e.g. [1,2]) which is closely related to density deconvolution (e.g. [3–5]). It is usually assumed in nonparametric deconvolution regression models (e.g. [6]) that the function θ is periodic (say on $[0, 1]$), and that A is thus a convolution operator on $[0, 1]$ with periodic Ψ which is, however, often unrealistic in practice. Examples are the deconvolution of astronomical and biological images from telescopic and microscopic imaging devices which involves deconvolution, but where the signal is usually not periodic. In this paper we will discuss the estimation of the signal θ from model (1), which appears to be more appropriate in this context. A main difficulty in this situation is that the reconstruction of θ from $g = \mathcal{K}\theta$ at any location x on the real line requires (at least asymptotically) information on g on the full real line. We therefore use a design which includes an additional sequence $a_n \to 0$ to ensure that the design points $\mathbf{z_k}$ will asymptotically exhaust $\mathbb{R}^d$.

* Corresponding address: Fakultät für Mathematik, Ruhr-Universität Bochum, NA 3/70, Universitätsstr. 150, D-44780 Bochum, Germany.
E-mail address: nicolai.bissantz@rub.de (N. Bissantz).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2009.04.004

In this paper we discuss pointwise convergence properties of Fourier-based estimators in model (1). The estimator and some useful assumptions are introduced in Section 2. Asymptotic normality and confidence intervals are discussed in Section 3 and a bootstrap version of the confidence intervals in Section 4. Whereas it is known that asymptotic confidence intervals do not perform well for moderate sample sizes (e.g. [7], in the direct density estimation context, and [8], for uniform confidence bands in density deconvolution), we demonstrate a satisfactory performance of the bootstrap confidence intervals in a simulation study in Section 5. Finally, in order to keep the paper more readable, all proofs are deferred to the Appendix.

2. Prerequisites: Estimator, notation and assumptions

Notation. In the following, we consider the $\mathbf{j}$th derivative of a function or estimator $\hat{\theta}_n(\mathbf{x})$, which depends on a d-variate covariable $\mathbf{x}$. By the $\mathbf{j}$th derivative $\mathbf{j} = (j_1, \ldots, j_d)$ we denote the partial derivative $\partial^j / \partial x_1^{j_1} \cdots \partial x_d^{j_d}$, where $j = j_1 + \cdots + j_d$, and we suppose $j_1, \ldots, j_d$ to be such that $j \le p$, where θ has partial derivatives of order p which are all continuous. Moreover, $\boldsymbol{\omega}^{\mathbf{j}}$, where $\boldsymbol{\omega} \in \mathbb{R}^d$, means $\omega_1^{j_1} \cdots \omega_d^{j_d}$.

The estimator. We consider a Fourier estimator which is based on the Fourier transform Φ_k of some kernel function k, which causes the regularisation of the estimator. To this end we make the following regularity assumptions on the Fourier transform Φ_k.

Assumption 1. The Fourier transform Φ_k of k is symmetric and supported on $[-1, 1]^d$ with $\Phi_k(\omega) = 1$ for $\omega \in [-b, b]^d$, $b > 0$, and $|\Phi_k(\omega)| \le 1$ for all $\omega \in [-1, 1]^d$.

The estimator is now defined as

$$\hat{\theta}_n^{(\mathbf{j})}(\mathbf{x}) = \hat{\theta}_{n,h}^{(\mathbf{j})}(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} (-\mathrm{i}\boldsymbol{\omega})^{\mathbf{j}} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \mathbf{x} \rangle} \Phi_k(h\boldsymbol{\omega}) \frac{\widehat{\Phi}_g(\boldsymbol{\omega})}{\Phi_\Psi(\boldsymbol{\omega})} \mathrm{d}\boldsymbol{\omega}, \quad 0 \le j \le p. \tag{2}$$

Here $h > 0$ is a smoothing parameter called the bandwidth, and $\widehat{\Phi}_g$ is the empirical Fourier transform of g defined by

$$\widehat{\Phi}_g(\boldsymbol{\omega}) = \frac{1}{N a_n^d} \sum_{\mathbf{r} \in \{-n,\ldots,n\}^d} Y_{\mathbf{r}} \mathrm{e}^{\mathrm{i}\langle \boldsymbol{\omega}, z_{\mathbf{r}} \rangle},$$

where $N = n^d$.

The estimator $\hat{\theta}_n^{(\mathbf{j})}$ can be written in kernel form as follows:

$$\hat{\theta}_n^{(\mathbf{j})}(\mathbf{x}) = \frac{1}{N h^{j+d} a_n^d} \sum_{\mathbf{r} \in \{-n,\ldots,n\}^d} Y_{\mathbf{r}} K_n^{(j)}\left(\frac{\mathbf{x} - \mathbf{z}_{\mathbf{r}}}{h}\right),$$

where the kernel

$$K_n^{(\mathbf{j})}(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} (-\mathrm{i}\boldsymbol{\omega})^{\mathbf{j}} \mathrm{e}^{-\mathrm{i}\langle \boldsymbol{\omega}, \mathbf{x} \rangle} \frac{\Phi_k(\boldsymbol{\omega})}{\Phi_\Psi(\boldsymbol{\omega}/h)} \mathrm{d}\boldsymbol{\omega}, \quad 0 \le \mathbf{j} \le p,$$

depends on n through h. The effective kernel $K_n(j)$ depends on the kernel function k, which determines the damping of the contribution of the estimated Fourier transform of $\theta^{(\mathbf{j})}$ at large frequencies to the estimator. This results in a regularisation of the estimator which is required since $\hat{\Phi}_g$ is dominated by noise at large frequencies where the Fourier transform Φ_g of g decays to zero. In more detail, $\hat{\theta}_n^{(\mathbf{j})}$ is closely related to the spectral-cut-off estimator, which would result in Φ_k the indicator function on $[-1, 1]$. We mention that in the time domain, the smoothed spectral-cut-off estimator which is used here yields a kernel estimator with "flat-top" kernels [9].

Hence, the estimator $\hat{f}_n(\mathbf{x})$ may be written as

$$\hat{\theta}_n^{(\mathbf{j})}(\mathbf{x}) = \sum_{\mathbf{r} \in \{-n,\ldots,n\}^d} Y_{\mathbf{r}} \frac{1}{N h^{j+d} a_n^d} K_n^{(\mathbf{j})}\left(\frac{\mathbf{x} - \mathbf{z}_{\mathbf{r}}}{h}\right) = \sum_{\mathbf{r} \in \{-n,\ldots,n\}^d} Y_{\mathbf{r}} w_{\mathbf{j},\mathbf{r},n}(\mathbf{x}),$$

with weights

$$w_{\mathbf{j},\mathbf{r},n}(\mathbf{x}) = \frac{1}{N h^{j+d} a_n^d} K_n^{(\mathbf{j})}\left(\frac{\mathbf{x} - \mathbf{z}_{\mathbf{r}}}{h}\right). \tag{3}$$

Further assumptions. We will make the following common assumptions on Φ_k and Ψ. Our first assumption is that Ψ is ordinary smooth, i.e. we consider mildly ill-posed problems in model (1).