Journal of Multivariate Analysis 124 (2014) 398–417

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Application of second generation wavelets to blind spherical deconvolution

T. Vareschi

Université Denis Diderot Paris 7 and CNRS-UMR 7099, 175 rue du Chevaleret 75013 Paris, France

ARTICLE INFO

Article history:
Received 14 January 2013
Available online 4 December 2013

AMS subject classifications:
62G05
62G99
65J20
65J22

Keywords:
Blind deconvolution
Blockwise SVD
Spherical deconvolution
Second generation wavelets
Nonparametric adaptive estimation
Linear inverse problems

ABSTRACT

We address the problem of spherical deconvolution in a non-parametric statistical framework, where both the signal and the operator kernel are subject to measurement errors. After a preliminary treatment of the kernel, we apply a thresholding procedure to the signal in a second generation wavelet basis. Under standard assumptions on the kernel, we study the minimax performances of the resulting algorithm in terms of $\mathbb{L}^p$ losses ($p \geq 1$) on Besov spaces on the sphere. We hereby extend the application of second generation spherical wavelets to the blind spherical deconvolution framework. It is important to stress that the procedure is adaptive with regard to both the target function sparsity and the kernel blurring effect. We end with the study of a concrete example, putting into evidence the improvement of our procedure on the recent blockwise SVD algorithm of Delattre et al. (2012).

1. Introduction

1.1. Statistical framework

Consider the following problem: we aim at recovering a signal $\boldsymbol{f}$ defined on the 2-dimensional sphere $\mathbb{S}^2$. $\boldsymbol{f}$ is not observed directly, but through the action of a blurring process modeled by a linear operator $\boldsymbol{K}$, and further contaminated by an additive Gaussian white noise. This is resumed in the classic white noise model

$$\boldsymbol{Y}_\varepsilon = \boldsymbol{Kf} + \varepsilon \dot{\boldsymbol{W}} \tag{1.1}$$

where $\dot{\boldsymbol{W}}$ is a white noise on $\mathbb{L}^2(\mathbb{S}^2)$ and $\boldsymbol{K} : \mathbb{L}^2(\mathbb{S}^2) \to \mathbb{L}^2(\mathbb{S}^2)$ is a measurable operator. We shall further restrict the shape of $\boldsymbol{K}$ by assuming that it is a convolution operator on $\mathbb{L}^2(\mathbb{S}^2)$, a classic framework [14,19,18] enjoying convenient mathematical properties (see Section 1.2). Namely, we suppose that there exists $\boldsymbol{h} \in \mathbb{L}^2(\mathrm{SO}(3))$ such that

$$\boldsymbol{Kf}(\omega) = \int_{\mathrm{SO}(3)} \boldsymbol{f}(g^{-1}\omega)\boldsymbol{h}(g)dg \tag{1.2}$$

where dg is the Haar measure on SO(3). So to speak, $\boldsymbol{f}$ is averaged on a neighborhood of ω and weighted according to $\boldsymbol{h}(g)$ for each rotation g^{-1} applied to ω. Alternatively, in a density estimation framework, one observes a random n-sample $(\boldsymbol{\theta}_1 X_1, \ldots, \boldsymbol{\theta}_n X_n)$ of $Z = \boldsymbol{\theta} X$ with density $\boldsymbol{Kf}$, where $\boldsymbol{\theta}$ is a random element in SO(3) (the group of rotations on $\mathbb{R}^3$) with density $\boldsymbol{h}$, and X has density $\boldsymbol{f} \in \mathbb{L}^2(\mathbb{S}^2)$. Formally we have $\varepsilon \sim n^{-1/2}$, and one can show that (1.2) holds as well [14].

E-mail addresses: thomas.vareschi@univ-paris-diderot.fr, vareschi@math.univ-paris-diderot.fr.

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2013.11.012

In practice, the blurring operator $\boldsymbol{K}$ is seldom directly observable and rather subject to measurement errors. For example $\boldsymbol{K}$ can be unknown but approximated via preliminary inference, or it can be the result of an unknown perturbation applied to a known operator. Following Efromovich and Koltchinskii [9] and Hoffmann and Reiß [15], we model the error in the operator as an additive Gaussian operator white noise. The observed result is a noisy version $\boldsymbol{K}_\delta$, satisfying

$$\boldsymbol{K}_\delta = \boldsymbol{K} + \delta \dot{\boldsymbol{B}} \tag{1.3}$$

where $\dot{\boldsymbol{B}}$ is a Gaussian operator white noise on $\mathcal{L}\big(\mathbb{L}^2(\mathbb{S}^2)\big)$ the set of linear endomorphisms of $\mathbb{L}^2(\mathbb{S}^2)$, independent from $\dot{\boldsymbol{W}}$. The meaning of models (1.1) and (1.3) is as follows: for $u, v, w, \in \mathbb{L}^2(\mathbb{S}^2)$, observable quantities take the forms

$$\langle \boldsymbol{Kf}, u\rangle + \varepsilon\alpha(u), \qquad \langle \boldsymbol{K}v, w\rangle + \delta\beta(v, w)$$

where $\alpha(u)$ and $\beta(v, w)$ are both Gaussian centered variables with respective variances $\|u\|_2^2$ and $\|v\|_2^2\|w\|_2^2$. Moreover, if $u', v', w' \in \mathbb{L}^2(\mathbb{S}^2)$ are other candidate functions, we have

$$\mathbb{E}[\alpha(u)\alpha(u')] = \langle u, u'\rangle_{\mathbb{L}^2(\mathbb{S}^2)}$$
$$\mathbb{E}[\beta(v, w)\beta(v', w')] = \langle v, v'\rangle_{\mathbb{L}^2(\mathbb{S}^2)}\langle w, w'\rangle_{\mathbb{L}^2(\mathbb{S}^2)}.$$

Many scientific fields call upon simple and efficient tools for the resolution of (1.1). Spherical deconvolution is for example well illustrated by the study of ultra high energy cosmic rays (UHECR), which are high energy radiations hitting the earth from apparently random directions. They could originate from long-lived relic particles from the Big Bang. Alternatively, they could be generated by the acceleration of standard particles, such as protons, in extremely violent astrophysical phenomena. They could also originate from Active Galactic Nuclei (AGN), or from neutron stars surrounded by extremely high magnetic fields. Discriminating among these different hypotheses involves the precise reconstruction of the probability density generating their observations. One could ask for example whether the latter is uniformly distributed among the sphere, or if it is constituted of superimposed localized spikes. In practice however, observations $(X_1, \dots, X_n)$ of such radiations are often subject to various physical perturbations. We model these by a random rotation $\boldsymbol{\theta}$, which is to say we actually observe $(\boldsymbol{\theta}_1 X_1, \dots \boldsymbol{\theta}_n X_n)$, n realizations of the random variable $Z = \boldsymbol{\theta}X$. The difficulty of the problem is characterized by the spreading of $\boldsymbol{h}$, the density of θ, around the neutral element of SO(3): the less localized, the more difficult the estimation of $\boldsymbol{f}$. Moreover, the law of $\boldsymbol{\theta}$ is not known in general, even if some assumptions can restrict its shape. In this case, preliminary inference is necessary, and leads to an estimator $\boldsymbol{K}_\delta$ of $\boldsymbol{K}$ according to (1.3).

Case of a known operator

We shall consider here the case where $\delta = 0$, and expose the path which finally led to the introduction and use of needlets. Spherical harmonics constitute the most natural set of functions to expand $\boldsymbol{f} \in \mathbb{L}^2(\mathbb{S}^2)$. Their frequency localization furthermore makes them ideally suited to spherical deconvolution, as they realize a blockwise SVD of $\boldsymbol{K}$ (as shown in Proposition 1.1), a property which guarantees the stability of its inversion. It prompted Healy et al. [14] to solve the spherical deconvolution problem with their use, hereby reaching optimal $\mathbb{L}^2$ rates of convergence on Sobolev spaces (Kim and Koo [19]). Unfortunately their performances can prove quite poor when the loss is measured by other $\mathbb{L}^p$ norms, $1 \le p \le \infty$, since they lack localization in the spatial domain (see [13]). The recent development of spherical wavelets [27,22] reversed this compromise, the latter being well localized in the spatial domain but very poorly in the frequential one. This makes them useful when a direct estimation of $\boldsymbol{f}$ is involved (see for example Freeden et al. [10] or Freeden et al. [11] for applications to geophysics and atmospheric sciences), but irrelevant in the setting of spherical deconvolution. The solution to this problem was finally brought by Narcowich et al. [24], who introduced a new set of functions, called needlets, which preserve the frequential localization of spherical harmonics and remedy their lack of spatial localization. Thereby, needlets inherit the stability of spherical harmonics in spherical deconvolution. They were subsequently exploited by Kerkyacharian et al. [18], who designed a procedure involving needlets attaining near-minimax rates of convergence for $\mathbb{L}^p$ losses $(1 \le p \le \infty)$ on Besov spaces (which definition is given in Section 2.3). Needlets also found various applications in the case of a direct estimation of $\boldsymbol{f}$, whether in astrophysics [21,13] or brain shape modeling [28].

Case of an unknown operator and Galerkin projection

The main methods in the context of blind deconvolution involve SVD and Galerkin schemes (see [3,4,15] for example). Galerkin projections were for example successfully applied to blind deconvolution on Hilbert spaces [9] or on Besov spaces on $[0, 1]^d$ [15,4]. They are based upon a discretization of (1.1) and (1.3) through the choice of appropriate test functions. Suppose we want to recover a function f from the observation of $g = Kf$. Let $(V_n)_{n\ge 0}$ and $(W_n)_{n\ge 0}$ be two increasing sequences of finite n-dimensional subspaces in $\mathbb{L}^2(\mathbb{S}^2)$, which admit the respective orthogonal bases $\boldsymbol{\varphi} = (\varphi_k)_{k\le n}$ and $\boldsymbol{\psi} = (\psi_k)_{k\le n}$. The Galerkin approximation $f_G \in V_n$ of f is the solution of the equation

$$\langle Kf_G, v\rangle = \langle g, v\rangle, \quad \forall v \in W_n. \tag{1.4}$$

This equation actually amounts to solving a finite dimensional linear system. Indeed, for $\gamma \in V_n$, note γ^n the vector whose components are $(\langle \gamma, \varphi_k\rangle)_{k\le n}$ and K^n the matrix with entries $(\langle K\varphi_k, \psi_k\rangle)_{k,k'\le n}$. Then $f_G \in V_n$ and we have

$$g^n = K^n f_G^n.$$