Goodness-of-fit test for noisy directional data

Thanh Mai Pham Ngoc [1] & Peter T. Kim [2] & Jae Yong Koo [3] & Claire Lacour [4]

[1] *Laboratoire de Mathématiques, UMR 8628, Université Paris Sud, 91405 Orsay Cedex, France, Email: thanh.pham_ngoc@math.u-psud.fr*
[2] *Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada. Email: pkim@uoguelph.ca.*
[3] *Department of Statistics, Korea University, Anam-Dong Sungbuk-Ku, Seoul 136-701 Korea. Email: jykoo@korea.ac.kr*
[4] *Laboratoire de Mathématiques, UMR 8628, Université Paris Sud, 91405 Orsay Cedex, France, Email: claire.lacour@math.u-psud.fr*

Résumé. Nous disposons d'observations bruitées Y_i sur la sphère unité $\mathbb{S}^2$ de $\mathbb{R}^3$: $Y_i = \varepsilon_i X_i,\ i = 1 \cdots n$. Celles-ci sont obtenues à partir de directions X_i corrompues par une rotation aléatoire $\varepsilon_i \in SO(3)$. Les X_i et ε_i sont supposés indépendants. La densité des observations Y_i est la convolée de la densité originelle et de celle du bruit $f_Y = f_\varepsilon \star f$. Nous supposerons que la densité des ε_i est connue. Nous cherchons à tester si la densité des X_i provient de la densité uniforme sur la sphère ou non à partir des données bruitées Y_i. Le test d'adéquation se fait dans un cadre non-paramétrique, les alternatives étant exprimées sur des classes de Sobolev ou celles des fonctions analytiques. Nous considérons deux types de régularité sur le bruit, un bruit ordinairement régulier et très régulier. Les vitesses obtenues par notre procédure sont optimales. Il est à noter qu'avec un bruit très régulier les vitesses de séparation ne sont que logarithmiques pour des classes de Sobolev mais deviennent polynomiales pour des densités analytiques. Nous illustrerons nos résultats théoriques par des simulations et des données réelles provenant de problématiques en astrophysique.

Mots-clés. Déconvolution sphérique, alternatives non-paramétriques, classe de Sobolev, classes analytiques, bruit régulier et très régulier, harmoniques sphériques.

Abstract. We consider the nonparametric goodness-of-fit test of the uniform density on the sphere when we have observations whose density is the convolution of an error density and the true underlying density. We will deal specifically with the smooth and supersmooth error case, this latter includes the Gaussian distribution. Similar to deconvolution density estimation, the smoother the error density the harder is the rate recovery of the test problem. When considering nonparametric alternatives expressed over Sobolev and analytic classes, we show that it is possible to obtain original separation rates. Furthermore, we show that our adaptive statistical procedure attains these optimal rates.

Keywords. Spherical deconvolution, nonparametric alternatives, Sobolev classes, Analytic classes, smooth and supersmooth noise, spherical harmonics.

1 Introduction

We consider the spherical convolution model. We observe: $Y_i = \varepsilon_i X_i, \quad i = 1, \ldots, N$ where the ε_i are i.i.d. random variables of SO(3) the rotation group in $\mathbb{R}^3$ and the X_i's are i.i.d. random variables on $\mathbb{S}^2$, the unit sphere of $\mathbb{R}^3$. We suppose that X_i and ε_i are independent. We also assume that the distributions of Y_i and X_i are absolutely continuous with respect to the uniform measure on $\mathbb{S}^2$ and we set f_Y and f the densities of Y_i and X_i respectively. The distribution of ε_i is absolutely continuous with respect to the probability Haar measure on SO(3) and we will denote the density of the ε_i's by f_ε. We suppose that f_ε is known. Then we have $f_Y = f_\varepsilon \star f$, where $*$ denotes the convolution product which is defined by:

$$f_\varepsilon \star f(\omega) = \int_{\mathrm{SO}(3)} f_\varepsilon(u) f(u^{-1}\omega) du. \tag{1}$$

Roughly speaking, the spherical convolution model provides a setup where each genuine observation X_i is contaminated by a small random rotation. The aim of the present paper is to provide a nonparametric adaptive minimax goodness-of-fit testing procedure on f from the noisy observations Y_i. More precisely, let f^0 being the uniform density on $\mathbb{S}^2$, we consider the problem of testing the null hypothesis $f = f^0$ against alternatives expressed in L^2 norm over Sobolev classes and analytic classes.

Convolution models have been extensively studied in the Euclidean setting. However, so far, only estimation has been treated in the spherical setup. The pioneer works of [3], [6], [7] introduced a minimax estimation procedure based on the Fourier basis of $L^2(\mathbb{S}^2)$. Recently, [5] proposed an optimal and adaptive hard thresholding estimation procedure based on needlets. Goodness-of-fit testing has mainly focused on the case of direct observations. Indeed, very few works have been devoted to the case of indirect observations. Let us cite the works of [1] for the inverse regression problem and [4] for the multivariate convolution density model. [2] built minimax nonparametric goodness-of-fit testing for convolution models based on kernels methods made a step forward by building an adaptive testing procedure in the noisy setup.

In this work, we establish several results for both smooth and supersmooth noises. We exhibit the optimal rates for adaptive cases. We prove that our statistical procedure attains those optimal rates.

The plan of the submission is as follows. In Section 2, we give a brief overview about harmonic analysis on SO(3) and $\mathbb{S}^2$ which will be necessary throughout the paper. In Section 3 we define the test hypotheses and the smoothness assumptions about the unknown density f and the noise ε_i. We also introduce the adaptive goodness-of-fit testing procedure. In Sections 4 we compute upper bounds for testing rates for the ordinary smooth noise case and the super smooth noise case. Note that for sake of clarity and concision we decide note to present the lower bounds but they can be found in [1] and [2].

2 Some preliminaries

This section provides the basic tools of harmonic analysis on the group $SO(3)$ and the unit sphere $\mathbb{S}^2$.

Let $L^2(SO(3))$ denote the space of square integrable functions on $SO(3)$.

Consider the *rotational harmonics*

$$D_{klm}(w) = D_{klm}(\tilde{a}, \tilde{b}, \tilde{c}) = e^{-il\tilde{a}} d^k_{lm}(\cos\tilde{b}) e^{-im\tilde{c}} \quad \text{for} \quad w \in SO(3),$$

where $(\tilde{a}, \tilde{b}, \tilde{c})$ are the *Euler angles* for w with $\tilde{a} \in [0, 2\pi)$, $\tilde{b} \in [0, \pi)$, $\tilde{c} \in [0, 2\pi)$, d^k_{lm} for $-k \le l, m \le k$, $k \in \mathbb{N} := \{0, 1, 2, \ldots\}$ are related to the Jacobi polynomials.

Let $h \in L^2(SO(3))$. We define the *rotational Fourier transform* on $SO(3)$ by

$$\hat{h}_{klm} = \int_{SO(3)} h(w) D_{klm}(w) dw, \tag{2}$$

with dw the probability Haar measure on $SO(3)$. We think of (2) as the matrix entries of the $d_k \times d_k$ matrix $\hat{h}_k = \left[\hat{h}_{klm}\right]$, where $d_k := 2k+1$, $-k \le l, m \le k$, $k \in \mathbb{N}$.

We shall do an analogue analysis for $\mathbb{S}^2$. Any point $x \in \mathbb{S}^2$ can be represented by

$$x = (\cos b \sin a, \sin b \sin a, \cos a)^\top,$$

where $a \in [0, \pi)$, $b \in [0, 2\pi)$ and superscript "$\top$" denotes transpose.

Let the *spherical harmonics* be

$$\phi_{kl}(x) = \phi_{kl}(a, b) = (-1)^l \sqrt{\frac{(2k+1)(k-l)!}{4\pi(k+l)!}} P^k_l(\cos a) e^{ilb} \quad \text{for} \quad x \in \mathbb{S}^2, \tag{3}$$

where $a \in [0, \pi)$, $b \in [0, 2\pi)$, $-k \le l \le k$, $k \in \mathbb{N}$ and P^k_l are the Legendre functions.

Let $L^2(\mathbb{S}^2)$ denote the space of square integrable functions on $\mathbb{S}^2$ and $f \in L^2(\mathbb{S}^2)$. We define the *spherical Fourier transform* on $\mathbb{S}^2$ by

$$\hat{f}_{kl} = \int_{\mathbb{S}^2} f(x) \overline{\phi_{kl}(x)} dx, \tag{4}$$

where dx is the spherical measure on $\mathbb{S}^2$ and the overbar denotes complex conjugation. Again we think of (4) as the vector entries of the d_k vector $\hat{f}_k = \left[\hat{f}_{kl}\right]$, $-k \le l \le k$, $k \in \mathbb{N}$.

We have the following *convolution property* which will be useful for the construction of the test procedure.

$$(\widehat{h \star f})_k = \hat{h}_k \hat{f}_k \quad \text{for} \quad k \in \mathbb{N}. \tag{5}$$

3 Testing uniformity on the sphere

Model and assumptions

The uniform density f^0 on $\mathbb{S}^2$ is given by $f^0 = \frac{1}{4\pi}\mathbb{1}_{\mathbb{S}^2}$. Define a separation measure between f^0 and f by $\Delta_f := \|f - f^0\|_2$. The uniform testing problem is

$$H_0 : f = f^0 \quad \text{versus} \quad H_a : f \in \mathcal{H}_a(\mathcal{F}, M, \psi_N)$$

based on the random sample $Y_1, \ldots, Y_N$, where the alternative is

$$\mathcal{H}_a(\mathcal{F}, M, \psi_N) = \{f : f \in \mathcal{F} \quad \text{and} \quad \Delta_f \geq M\psi_N\}$$

for a class $\mathcal{F}$ of densities, a testing rate ψ_N and M a constant.

For any test T, the maximal misclassification error rate (or the risk for the zero-one loss) is defined by

$$R_N(T, \mathcal{F}, M, \psi_N) = \mathbb{P}_{f^0}(T = 1) + \sup\Big\{\mathbb{P}_f(T = 0) : f \in \mathcal{H}_a(\mathcal{F}, M, \psi_N)\Big\}. \tag{6}$$

In this paper, we will focus on alternative based on Sobolev classes and analytic functions. Let

$$\|f\|^2_{W_\alpha} = \sum_{k\geq 0}(k + k(k+1))^\alpha \|\hat{f}_k\|^2.$$

We denote the Sobolev space by $W_\alpha(\mathbb{S}^2, R) = \left\{f : \mathbb{S}^2 \to \mathbb{R}^+,\ \int_{\mathbb{S}^2} f = 1 \text{ and } \|f\|^2_{W_\alpha} \leq \frac{1}{4\pi} + R^2\right\}$, for some fixed constant $R > 0$.

As for the analytic classes, the analytic norm is defined as

$$\|f\|^2_{\mathcal{A}_{p,r}} = \sum_{k\geq 0} \exp\left(2pk^r\right) \left\|\hat{f}_k\right\|^2.$$

An analytic class of densities is given by $\mathcal{A}_{p,r}(Q) = \left\{f : \mathbb{S}^2 \to \mathbb{R}^+,\ \int_{\mathbb{S}^2} f = 1 \text{ and } \|f\|^2_{\mathcal{A}_{p,r}} \leq \frac{1}{4\pi} + Q^2\right\}$, where $p > 0$, $r > 0$, $Q > 0$ are finite constants.

Noise assumptions

Let $\mathcal{E}_k$ be the d_k-dimensional vector space spanned by $\{\phi_{kl} : -k \leq l \leq k\}$ for each $k \in \mathbb{N}$. Thus any $v \in \mathcal{E}_k$ can be written as $v = \hat{v}_k^\top \phi_k$ and through Parseval's identity, the usual L^2-norm is $\|v\|_2^2 = \|\hat{v}_k\|^2$. Now according to (5), $\hat{h}_k$ and $\hat{h}_k^{-1}$ are also understood as maps $\hat{h}_k : \mathcal{E}_k \to \mathcal{E}_k$ and $\hat{h}_k^{-1} : \mathcal{E}_k \to \mathcal{E}_k$ defined by

$$\hat{h}_k v = (\hat{h}_k \hat{v}_k)^\top \phi_k \quad \text{and} \quad \hat{h}_k^{-1} v = (\hat{h}_k^{-1}\hat{v}_k)^\top \phi_k \quad \text{for} \quad v = \hat{v}_k^\top \phi_k \in \mathcal{E}_k.$$

Again by Parseval's identity, $\|\hat{h}_k v\|_2^2 = \|\hat{h}_k \hat{v}_k\|$ for all $k \geq 0$. Consequently, we have the operator inequality,

$$\left\|\hat{h}_k\right\|_{\text{op}} = \sup_{v\neq 0, v\in\mathcal{E}_k} \frac{\left\|\hat{h}_k v\right\|_2}{\|v\|_2} \quad \text{and} \quad \left\|\hat{h}_k^{-1}\right\|_{\text{op}} = \sup_{v\neq 0, v\in\mathcal{E}_k} \frac{\left\|\hat{h}_k^{-1} v\right\|_2}{\|v\|_2}. \tag{7}$$

For all $k \in \mathbb{N}$, the matrix $\hat{h}_k$ is invertible and there exist constants $q > 0$, $s > 0$, $b_0 > 0$, $b_1 > 0$, $\nu_0, \nu_1 \in \mathbb{R}$ such that

$$\left\|\hat{h}_k^{-1}\right\|_{\text{op}} \leq b_0 k^{\nu_0} \exp\left(qk^s\right) \tag{8}$$

and

$$\left\|\hat{h}_k\right\|_{\text{op}} \leq b_1 k^{-\nu_1} \exp\left(-qk^s\right). \tag{9}$$

The error density h is referred to ordinary smooth if $s = 0$ and to supersmooth otherwise.

Test procedure

The test procedure relies on the construction of an estimator of the separation measure Δ_f.

By (5) we can write,

$$\hat{f}_k = \hat{f}_{\varepsilon k}^{-1} \hat{f}_{Yk},$$

provided of course that the matrices $\hat{f}_{\varepsilon k}$ are invertible for all $k \in \mathbb{N}$ in a range of interest.

We construct an empirical version $\hat{f}_{Yk}$ of as $\hat{f}_{Yk}^N = \frac{1}{N}\sum_{j=1}^N \phi_k(Y_j)$.

Define

$$G(y_1, y_2) := \sum_{k=1}^{K} \langle \Phi_k(y_1), \Phi_k(y_2)\rangle \quad \text{where} \quad \Phi_k = \hat{f}_{\varepsilon k}^{-1}\phi_k, \quad \text{for} \quad y_1, y_2 \in \mathbb{S}^2. \tag{10}$$

Then $U_K := \frac{1}{N(N-1)} \sum_{1\leq i_1 \neq i_2 \leq N} G(Y_{i1}, Y_{i2})$ is a natural estimator of Δ_f^2.

4 Asymptotic results for the supersmooth noise case

Sobolev alternatives

Theorem 1. *Let* $\psi_N = (\log N)^{-2\alpha/s}$ *and* $K_0 > 0$. *We consider* $K^* = \left\lfloor \left(\frac{1}{q}\log(N)/8\right)^{1/s}\right\rfloor$ *and the test statistic*

$$D_N = \mathbb{1}_{\{|U_{K^*}|/t_{K^*}^2 > K_0\}}$$

with $t_K^2 = K^{2\nu_0+1}\exp(2qK^\beta)/N$. *Then, if* $M > K_0 + ((4\pi)^{-1} + R^2)(1/(16q))^{-2\alpha/s}$,

$$\lim_{N\to\infty} R_N(D_N, W_\alpha(\mathbb{S}^2, R), M, \psi_N) = 0$$

Analytic alternatives

Theorem 2. *Suppose (8) holds. Let* $p \geq p_0 > 0$*,* $0 < s \leq \frac{r}{2}$ *and* $Q > 0$*. Let* $L = \left\lceil \left(\frac{2}{M_2 s}\log\log N\right)^{\mathbb{1}_{(0\leq s\leq 1)}\frac{1}{2-s}+\mathbb{1}_{(s>1)}\frac{1}{s}} \right\rceil$ *and* $U = \left\lceil 2\left(\frac{\log N}{2p_0}\right)^{\frac{1}{s}}\right\rceil$*. Let the test* D_N *be defined by* $D_N := \mathbb{1}_{(\max_{K\in\mathcal{K}}|U_K|/\tau_K>1)}$ *with* $\tau_K = \xi\frac{K^{2\nu_0+1-s/2}\exp(2qK^s)}{N}$ *and* $\xi = \mathbb{1}_{(0\leq s\leq 1)}K^{(2-s)/2} + \mathbb{1}_{(s>1)}K^{s/2}$ *, where* $\mathcal{K} = \left\{K : L \leq K \leq U\right\}$*. Let*

$$\psi_N = \exp\left[-\frac{1}{2}\log N + q\left(\frac{\log N}{2p}\right)^{s/r} + \frac{2\nu_0+1+(1-s)\mathbb{1}(0\leq s\leq 1)}{2r}\log\frac{\log N}{2p}\right].$$

Then if $M > (Q^2+2)^{\frac{1}{2}}$*,*

$$\lim_{N\to\infty} R_N(D_N, \mathcal{A}_{p,r}(Q), M, \psi_N) = 0.$$

References

[1] Bissantz, N. Claesken, G., Holzmann, H., and Munk, A. (2009). *Testing for lack of fit in inverse regression-with applications to biophotonic imaging*, J.R. Stat. Soc. Ser. B Stat. Methodol., 71(1):25-48.

[2] Butucea, C., Matias, C. and Pouet, C. (2009). *Adaptive goodness-of-fit testing from indirect observations.*, Ann. Inst. Henri Poincaré Probab. Stat., 45(2): 352-372.

[3] Healy, Jr. D. M., Hendriks, H. and Kim, P. T. (1998), *Spherical deconvolution*, J. Multivariate Anal., 67(1): 1-22.

[4] Holzmann, H., BIssantz, N. and Munk, A. (2007). *Density testing in contaminated sample.* J Multivariate Anal. 98(1): 57-75.

[5] Kerkyacharian, G., Pham Ngoc, T. M., and Picard, D. (2011). *Localized spherical deconvolution.* Ann. Statist., 39(2): 1042-1068.

[6] Kim, P. T. and Koo, J.-Y. (2002). *Optimal spherical deconvolution* J. Multivariate Anal., 80(1): 21-42.

[7] Kim, P. T., Koo, J.-Y., and Park, H. J. (2004). *Sharp minimaxity and spherical deconvolution for super smooth error distributions.* J. Multivariate Anal., 90(2): 384-392.