Opuscula Mathematica • Vol. 30 • No. 2 • 2010

http://dx.doi.org/10.7494/OpMath.2010.30.2.203

Zbigniew Szkutnik

A NOTE ON MINIMAX RATES OF CONVERGENCE IN THE SPEKTOR-LORD-WILLIS PROBLEM

Abstract. In this note, attainable lower bounds are constructed for the convergence rates in a stereological problem of unfolding spheres size distribution from linear sections, which shows that a spectral type estimator is strictly rate minimax over some Sobolev-type classes of functions.

Keywords: Poisson inverse problem, rate minimaxity, singular value decomposition, stereology.

Mathematics Subject Classification: 62G05, 65J22.

1. INTRODUCTION

Consider a Poisson inverse problem of estimating a function $f \in L^2([0,1],\mu)$, with $d\mu(x) = xdx$, based on an observation of an inhomogeneous Poisson process on [0,1] with intensity function ng with respect to the measure $d\lambda(y) = ydy$, where

$$g(y) = (\mathcal{K}f)(y) = 2\int_y^1 f(x)d\mu(x), \tag{1.1}$$

and n is the "size of the experiment" that will tend to infinity in the asymptotic setup. This may serve as a model of a stereological problem, known as the Spektor-Lord-Willis (SLW) problem, and defined as follows. A population of spheres of random radii is randomly placed in an opaque medium. An experimenter is interested in estimating the distribution of the radii, but the only available data are the lengths of the line segments that are intersections of the spheres with a random linear probe through the medium. A practical motivation for studying such problems may come, e.g., from metallurgy, where linear intercepts are measured on polished metallographic sections (cf., [3] or [4], p.117), from geology, where drilling data are analysed, or from medicine, because of biopsy data. The formulation of the problem dates back to Spektor ([8]) and Lord and Willis ([6]). Various approaches to the problem are studied in [9], pp. 296–299. More recently, the SLW problem was discussed in [2] and [10]. Because of mathematical tractability reasons, the intensities of the Poisson processes were taken with respect to $d\mu$ and $d\lambda$, thus leading to (1.1). The minimax risk was considered over some Sobolev ellipsoids defined in terms of the singular functions of the operator $\mathcal{K} : L^2([0,1],\mu) \to L^2([0,1],\lambda)$. More specificaly, one has (see, [2])

Proposition 1.1. *The singular values of the operator* $\mathcal{K}$ *in* (1.1) *are* $b_\nu = 2/[\pi(2\nu+1)]$, $\nu = 0,1,\ldots,$ *with the right singular functions* $\phi_\nu(x) = 2\sin[(2\nu+1)\pi x^2/2]$ *and the left singular functions* $\psi_\nu(y) = 2\cos[(2\nu+1)\pi y^2/2]$.

The estimated function f is assumed to belong to the class

$$\mathcal{F}_{a,C} = \left\{ \sum_{\nu=0}^{\infty} c_\nu \phi_\nu \colon c_0 = 1, \sum_{\nu=1}^{\infty} (2\nu+1)^{2a} c_\nu^2 \le C^2 \right\},$$

with some $a > 1/2$ and for some C. Regularity of the functions from $\mathcal{F}_{a,C}$ is described by the following proposition, proved in [2].

Proposition 1.2. *Let k be a natural number.*

(a) *If* $f \in \mathcal{F}_{a,C}$ *with* $a > k + 1/2$, *then f is k times continuously differentiable in* $[0,1]$.
(b) *If* $f \in \mathcal{F}_{k,C}$, *then f has k weak derivatives that are square integrable in* $[0,1]$ *with respect to* $dm(x) := x^{1/2}dx$.

Define the risk of an estimator $\tilde{f}_n$ as the mean integrated square error

$$M(\tilde{f}_n, f) = \mathrm{E}_f \|\tilde{f}_n - f\|^2,$$

where $\|\cdot\|$ denotes the $L^2([0,1],\mu)$ norm. With $f \in \mathcal{F}_{a,C}$ one would expect the minimax convergence rates $n^{-2a/(2a+3)}$ (cf., e.g., [5], or [7]). Indeed, it was proved in [2] that $n^{-2a/(2a+3)}$ is an upper bound for the convergence rate. The lower bounds obtained in [10] and in [2] were, however, faster by some logarithmic factors. In this note, we obtain $n^{-2a/(2a+3)}$ as a lower bound thus proving strict minimaxity of the estimator developed in [2].

2. THE RESULT

Denote by $\rho(P,Q)$ the Hellinger affinity between probability measures P, Q and by $\Delta(\omega,\omega')$ the Hamming distance between two finite, binary sequences ω, ω' of the same length. The following version of the Assouad Lemma will be used (cf., [1]).

Lemma 2.1. *Let* $\{P_\omega,\, \omega \in \mathcal{D}\}$ *be a family of distributions indexed by* $\mathcal{D} = \{0,1\}^m$ *and* $X_1,\ldots,X_n$ *an i.i.d. sample from a distribution in the family. Assume that*

$\rho(P_\omega, P_{\omega'}) \geq \bar{\rho}$ *for each pair* $(\omega, \omega') \in \mathcal{D}^2$ *such that* $\Delta(\omega, \omega') = 1$. *Then, for any estimator* $\hat{\omega}(X_1, \ldots, X_n)$ *with values in* $\mathcal{D}$,

$$\sup_{\omega \in \mathcal{D}} \mathrm{E}_\omega \left[\Delta\left(\hat{\omega}, \omega\right)\right] \geq m \bar{\rho}^{2n}/4,$$

where E_ω *denotes the expectation when the* X_i *have the distribution* P_ω.

A good lower bound for the risk can be obtained with a possibly large number of well separated functions in $\mathcal{F}_{a,C}$ for which the corresponding data distributions are close to each other. In order to describe the action of $\mathcal{K}$ in a tractable way, the functions will be defined in terms of the singular functions.

Theorem 2.2. *For the class of estimators*

$$\mathcal{T} = \{\tilde{f}_n \,:\, \mathrm{E}_f \|\tilde{f}_n\|^2 < \infty, f \in \mathcal{F}_{a,C}\},$$

there exists a constant c *such that*

$$\inf_{\tilde{f}_n \in \mathcal{T}} \sup_{f \in \mathcal{F}_{a,C}} M(\tilde{f}_n, f) \geq c\, n^{-2a/(2a+3)}.$$

Proof. For an integer $m = m(n)$, let $\omega = (\omega_1, \ldots, \omega_m)$ with $\omega_i \in \{0, 1\}$ and let b_k, ϕ_k and ψ_k be as in Proposition 1. Define

$$f_\omega = \phi_0 + \delta_m \sum_{i=1}^{m} \omega_i (\phi_{m+2i-2} + \phi_{m+2i-1})$$

with some positive δ_m. In order to have $f_\omega \in \mathcal{F}_{a,C}$ for all ω, it suffices that $\delta_m^2 \sum_{\nu=m}^{3m-1} (2\nu+1)^{2a} \leq C^2$, or that $(6m)^{2a+1} \leq 2C^2 \delta_m^{-2}(2a+1)$, and we can take

$$\delta_m^2 \asymp m^{-(2a+1)} \tag{2.1}$$

to satisfy the condition. Set $g_\omega = \mathcal{K} f_\omega$, $f_0 = \phi_0$ and $g_0 = \mathcal{K} f_0$. To each f_ω there corresponds an observable Poisson process $\mathcal{N}_{ng_\omega}$ with intensity function ng_ω or, equivalently, n i.i.d. copies of a Poisson process $\mathcal{N}_{g_\omega}$. Denote by $\mathcal{L}(\mathcal{N}_g)$ the distribution of $\mathcal{N}_g$. As in [2], one has

$$\rho\left(\mathcal{L}(\mathcal{N}_{g_\omega}), \mathcal{L}(\mathcal{N}_{g_{\omega'}})\right) = \int \sqrt{\frac{d\mathcal{L}(\mathcal{N}_{g_\omega})}{d\mathcal{L}(\mathcal{N}_{g_0})} \frac{d\mathcal{L}(\mathcal{N}_{g_{\omega'}})}{d\mathcal{L}(\mathcal{N}_{g_0})}} d\mathcal{L}(\mathcal{N}_{g_0}) = \exp\left[-H^2(g_\omega, g_{\omega'})\right],$$

where $H^2(g_\omega, g_{\omega'}) = \int_0^1 \left(\sqrt{g_\omega} - \sqrt{g_{\omega'}}\right)^2 d\lambda/2$. With $\Delta(\omega, \omega') = 1$, one has $g_{\omega'} = g_\omega \pm \delta_m(b_k\psi_k + b_{k+1}\psi_{k+1})$, for some k between m and $3m-2$. Standard calculation gives

$$H^2(g_\omega, g_{\omega'}) = \frac{\delta_m^2}{2b_0} \int_0^1 \frac{(b_k\psi_k + b_{k+1}\psi_{k+1})^2}{\psi_0} \left(\sqrt{\frac{g_{\omega'}}{b_0\psi_0}} + \sqrt{\frac{g_\omega}{b_0\psi_0}}\right)^{-2} d\lambda.$$

The second factor under the integral is bounded and cut away from zero (cf., [2]). Hence,

$$H^2(g_\omega, g_{\omega'}) \asymp \delta_m^2 b_k^2 \left[\int_0^1 \frac{(\psi_k + \psi_{k+1})^2}{\psi_0} d\lambda + \left(1 - \frac{b_{k+1}}{b_k}\right)^2 \int_0^1 \frac{\psi_{k+1}^2}{\psi_0} d\lambda - \right.$$
$$\left. -2\left(1 - \frac{b_{k+1}}{b_k}\right) \int_0^1 \frac{(\psi_k + \psi_{k+1})\psi_{k+1}}{\psi_0} d\lambda \right]. \tag{2.2}$$

Since $\psi_k(y) + \psi_{k+1}(y) = 4\cos[(k+1)\pi y^2]\psi_0(y)$, one easily obtains $\int_0^1 (\psi_k + \psi_{k+1})^2/\psi_0 d\lambda = O(1)$. Further, $\int \psi_{k+1}^2/\psi_0 d\lambda \asymp \log(2k+3)$ (cf., [2]) and, because $1 - b_{k+1}/b_k = 2/(2k+3)$, the second term in (2.2) is $o(1)$. The same holds true for the third term, because

$$\left| \int_0^1 \frac{\psi_k \psi_{k+1}}{\psi_0} d\lambda \right| \le \left[\int_0^1 \frac{\psi_k^2}{\psi_0} d\lambda \int_0^1 \frac{\psi_{k+1}^2}{\psi_0} d\lambda \right]^{1/2} \asymp \log(2k+3).$$

Consequently $H^2(g_\omega, g_{\omega'}) = O(\delta_m^2 b_m^2) = O(\delta_m^2 m^{-2}) = O(m^{-(2a+3)})$. Now, for any estimator $\tilde{f}_n$ of f, take $\tilde{\omega} \in \mathcal{D} = \{0,1\}^m$ such that $\|f_{\tilde{\omega}} - \tilde{f}_n\| = \min_{\omega \in \mathcal{D}} \|f_\omega - \tilde{f}_n\|$. Then $\|f_{\tilde{\omega}} - f_\omega\| \le \|f_{\tilde{\omega}} - \tilde{f}_n\| + \|f_\omega - \tilde{f}_n\|$ and

$$\sup_{f \in \mathcal{F}_{a,C}} \mathrm{E}_f \|\tilde{f}_n - f\|^2 \ge \max_{\omega \in \mathcal{D}} \mathrm{E}_{f_\omega} \|\tilde{f}_n - f_\omega\|^2 \ge \frac{1}{4} \max_{\omega \in \mathcal{D}} \mathrm{E}_{f_\omega} \|f_{\tilde{\omega}} - f_\omega\|^2 =$$
$$= \frac{2\delta_m^2}{4} \max_{\omega \in \mathcal{D}} \mathrm{E}_{f_\omega} [\Delta(\tilde{\omega}, \omega)] \ge \frac{\delta_m^2 m \bar{\rho}^{2n}}{8} \asymp m^{-2a} \bar{\rho}^{2n},$$

because of the Assouad Lemma and because of (2.1). Take $m \asymp n^{1/(2a+3)}$. Then $H^2(g_\omega, g_{\omega'}) = O(n^{-1})$, which implies that $\bar{\rho}^{2n} \asymp 1$, and $\sup_{f \in \mathcal{F}_{a,C}} \mathrm{E}_f \|\tilde{f}_n - f\|^2 \ge cn^{-2a/(2a+3)}$. This completes the proof. □

Although the idea of the proof in [2] was quite similar, the functions were defined there as

$$f_\omega = \phi_0 + \delta_m \sum_{i=m}^{2m-1} \omega_{i-m+1} \phi_i,$$

which only gave $H^2(g_\omega, g_{\omega'}) \asymp m^{-(2a+3)} \log m$ and, consequently, the disturbing logarithmic factor in the lower bound. On the other hand, our choice of f_ω produced distributions $\mathcal{L}(\mathcal{N}_{g_\omega})$ slightly closer to each other, namely $H^2(g_\omega, g_{\omega'}) \asymp m^{-(2a+3)}$, which proved sufficient to obtain sharp, attainable bounds for the convergence rates.

Acknowledgments

Work partially supported by AGH local grant 10.420.03.

REFERENCES

[1] L. Birgé, *The Brouwer Lecture 2005. Statistical estimation with model selection*, 2006. (available at arXiv:math.ST/0605187 v2)

[2] A. Dudek, Z. Szkutnik, *Minimax unfolding spheres size distribution from linear sections*, Statist. Sinica **18** (2008), 1063–1080.

[3] J.-H. Han, D.-Y. Kim, *Determination of three-dimensional grain size distribution by linear intercept measurement*, Acta Mater. **46** (1998), 2021–2028.

[4] J.E. Hilliard, L.R. Lawson, *Stereology and Stochastic Geometry*, Kluwer, Dordrecht, 2003.

[5] I.M. Johnstone, B.W. Silverman, *Speed of estimation in positron emission tomography and related inverse problems*, Ann. Statist. **18** (1990), 251–280.

[6] G.W. Lord, T.F. Willis, *Calculation of air bubble distribution from results of a Rosiwal traverse of aerated concrete*, A.S.T.M. Bull. **56** (1951), 177–187.

[7] A.C.M. van Rooij, F.H. Ruymgaart, *Asymptotic minimax rates for abstract linear estimators*, J. Statist. Plann. Inference **53** (1996), 389-402.

[8] A.G. Spektor, *Analysis of distribution of spherical particles in non-transparent structures*, Zavodskaja Laboratorija **16** (1950), 173–177.

[9] D. Stoyan, W.S. Kendall, J. Mecke, *Stochastic Geometry and its Applications*, Akademie-Verlag, Berlin, 1987.

[10] Z. Szkutnik, *A new solution to an old stereological problem*, Bulletin of the International Statistical Institute **LXII** (2007), 4577–4580.

Zbigniew Szkutnik
szkutnik@agh.edu.pl

AGH University of Science and Technology
Faculty of Applied Mathematics
al. Mickiewicza 30, 30-059 Kraków, Poland

Received: November 12, 2009.
Accepted: November 30, 2009.