Abstracts

Sharp Adaptive Nonparametric Testing for Sobolev Ellipsoids

MICHAEL NUSBAUM

(joint work with Pengsheng Ji)

Consider the Gaussian white noise model in sequence space

\[ Y_j = f_j + n^{-1/2} \xi_j, \quad j = 1, 2, \ldots \]

with signal \( f = \{f_j\}_{j=1}^{\infty} \) and \( \xi_j \sim N(0, 1) \) independent. For some \( \rho, \beta, M > 0 \), consider hypotheses of "no signal" vs. an ellipsoid with \( l_2 \)-ball removed:

\[ H_0 : f = 0 \quad \text{against} \quad H_a : f \in \Sigma(\beta, M) \cap B_{\rho}, \]

\[ B_{\rho} = \left\{ f \in l_2 : \|f\|_2^2 \geq \rho \right\}, \quad \Sigma(\beta, M) = \left\{ f : \sum_{j=1}^{\infty} j^{2\beta} f_j^2 \leq M \right\}. \]

Consider \( \alpha \)-tests \( \phi \) and their worst case type II error over the alternative:

\[ \Psi_n(\phi, \rho, \beta, M) := \sup_{f \in \Sigma(\beta, M) \cap B_{\rho}} (1 - E_{n,f} \phi). \]

Ingster [8] found the critical rate for \( \rho \to 0 \), the so-called separation rate \( \rho_n \approx n^{-4\beta/(4\beta+1)} \), where a nontrivial type II error behaviour occurs:

\[ 0 < \lim \inf_{\phi \alpha\text{-test}} \Psi_n \quad \text{and} \quad \lim \sup_{\phi \alpha\text{-test}} \inf \Psi_n < 1 - \alpha. \]

This rate is known as the optimal rate for nonparametric testing. As in nonparametric estimation (cf. Pinsker [13]), the step from optimal minimax rate to optimal constant has been made, with the result by Ermakov [3]:

Suppose \( \alpha \in (0, 1) \) and \( \rho_n \sim (cn)^{-4\beta/(4\beta+1)} \) for some \( c > 0 \). Then

\[ \inf_{\phi \alpha\text{-test}} \Psi_n(\phi, \rho, \beta, M) = \Phi(z_\alpha - cM^{-1/4\beta} \eta_{\beta}) + o(1) \]

where \( z_\alpha \) is the upper \( \alpha \)-quantile of \( N(0, 1) \) and \( \eta_{\beta} = (2\beta+1)^{1/2}(4\beta+1)^{-1/2-1/4\beta} \).

We address the question of sharp minimax adaptive testing, that is the question of whether this constant can be attained by tests which do not depend on \( (\beta, M) \). For minimax estimation with \( l_2 \)-loss over ellipsoids \( \Sigma(\beta, M) \), cf. Efroimovich and Pinsker [2], Golubev [6], Tsybakov [16]. Adaptation to Pinsker’s constant is possible there, without a penalty such as rate loss. For testing, Spokoiny [15] showed that for adaptation to \( (\beta, M) \), there is a rate penalty of order \( (\log \log n)^{1/2} \). Essentially this result concerns adaptation to \( \beta \) only; indeed \( M \) is irrelevant for the optimal rate. Moreover, for adaptation to \( \beta \) only, Ingster and Suslina [9] obtained a sharp constant, within the \( (\log \log n)^{1/2} \) rate loss framework. Adaptation to both parameters \( (\beta, M) \) is an open problem.

We first consider the problem of adaptation to \( M \) only, assuming \( \beta \) known.
**Theorem 1.** Suppose \( c > 0, 0 < M_1 < M_2 < \infty \) and \( \rho_n \sim (cn)^{-4\beta/(4\beta+1)} \). Then there is no test \( \phi_n \) satisfying \( E_{n,0}\phi_n \leq \alpha + o(1) \) and both relations:

\[
\Psi_n(\phi_n, \rho_n, \beta, M_1) \leq \Phi(z_\alpha - cM_1^{-1/4\beta} \eta_\beta) + o(1), \quad i = 1, 2.
\]

In view of (2), adaptation to \( M \) only is impossible at the separation rate. We now replace the constant \( c \) in \( \rho_n \sim (cn)^{-4\beta/(4\beta+1)} \) by a sequence \( c_n \to \infty \) arbitrarily slow. Then \( \Phi(z_\alpha - c_n M^{-1/4\beta} \eta_\beta) \to 0 \), and taking the standard log-asymptotics approach, it turns out that adaptation to Ermakov’s constant is possible.

**Theorem 2.** Assume \( c_n \to \infty \) and \( c_n = o(n^K) \) for every \( K > 0 \). If \( \rho_n = (c,n)^{-4\beta/(4\beta+1)} \) then there exists a test \( \phi_n \) fulfilling \( E_{n,0}\phi_n \leq \alpha + o(1) \) and for all \( M > 0 \)

\[
\limsup_n \frac{1}{c_n^2} \log \Psi_n(\phi_n, \rho_n, \beta, M) \leq -\frac{M^{-1/2\beta} \eta_\beta}{2}.
\]

Ermakov [4] showed that the r. h. s. above is also the best achievable for tests possibly depending on \( M \). Hence there is no “penalty for adaptation” here, except that one has to change the optimality criterion. Proofs for this case are in [10].

For the problem of full adaptation to \((\beta, M)\), we first state a lower asymptotic risk bound for known \( M \) and unknown \( \beta \in [\beta_1, \beta_2] \), a variation of a result of Ingster and Suslina [9]. Assume that \( 0 < \beta_1 < \beta_2 \) and that \( M > 0 \) is fixed. Let \( D \) be arbitrary and define a radius sequence \( \rho_{n,\beta,M} \) by

\[
(r_{n,\beta,M})^{-4\beta/(4\beta+1)} = n^{-1/2} M^{1/4\beta} \eta_\beta^{-1} \left( (2 \log \log n)^{1/2} + D \right).
\]

Then for any sequence of tests \( \phi_n \) satisfying \( E_{n,0}\phi_n \leq \alpha + o(1) \)

\[
\sup_{\beta \in [\beta_1, \beta_2]} \Psi_n(\phi_n, \rho_{n,\beta,M}, \beta, M) \geq (1 - \alpha) \Phi(-D) + o(1).
\]

Here the test sequences \( \phi_n \) are assumed not to depend on \( \beta \) (but possibly on \( M \)); the radius \( \rho_{n,\beta,M} \) depends on \( \beta \) and \( M \). The concept of a radius varying with \( \beta \) (inside the risk supremum) has been introduced by Spokoiny [15] in the context of rate adaptivity. In the refinement of [9], Ermakov’s constant \( M^{-1/2\beta} \eta_\beta \) enters the critical radius \( \rho_{n,\beta,M} \) as well.

The attainability of the bound (4) is shown in [9] for tests depending on \( M \). We show it for tests not depending on \( M \), establishing adaptivity in \((\beta, M)\).

**Theorem 3.** Assume that \( 0 < \beta_1 < \beta_2 \) and \( 0 < M_1 < M_2 \) are fixed. Let \( D \) be arbitrary and let \( \rho_{n,\beta,M} \) be the radius sequence in (7). Then there exists a test \( \phi_n \) fulfilling \( E_{n,0}\phi_n \leq \alpha + o(1) \) and

\[
\sup_{\beta \in [\beta_1, \beta_2], M \in [M_1, M_2]} \Psi_n(\phi_n, \rho_{n,\beta,M}, \beta, M) \leq (1 - \alpha) \Phi(-D) + o(1).
\]

It turns out that there is no additional penalty for \( M \) being unknown, and there is no need to consider tail probabilities.

Ingster and Suslina [9] establish their lower bound (4) for \( l_p \) ellipsoids of smoothness \( r \) with shrinking \( l_p \) ellipsoids of smoothness \( s \) removed, and also Besov classes, but not for sup-norm settings. Lepski and Tsybakov [12] prove a sharp minimax result in testing when the alternative is a Hölder class with a sup-norm.
ball removed. This represents a testing analog of the minimax estimation result of Korostelev [11] and also a sup-norm analog of Ermakov [3]; for the regression case cf. [5]. When $\beta$ is given, Dümbgen and Spokoiny [1] establish a sharp adaptivity result with respect to the size parameter $M$ only. The case of unknown $(\beta, M)$ seems to be an open problem for sup-norm testing; for the estimation case cf. [7]. But in [1] a test is given which is adaptive rate optimal without a log log $n$-type penalty. Rohde [14] considers the sup-norm case for regression with nongaussian errors, combining methods of [1] with ideas related to rank tests.

References


Reporter: Axel Munk