Adaptive estimation of linear functionals by model selection

B. Laurent*, C. Ludena‡, C. Prieur*
*INSA de Toulouse, LSP; ‡IVIC, Venezuela

Let T be a linear functional defined over a certain Hilbert space $\mathbb{H}$. We consider the problem of estimating $T(s)$, $s \in \mathbb{H}$ based on the Gaussian observation

$$Y(t) = \langle s, t \rangle + \frac{\sigma}{\sqrt{n}} L(t), \text{ for all } t \in \mathbb{H},$$

where L is some centered Gaussian isonormal process, which means that L maps isometrically $\mathbb{H}$ onto some Gaussian subspace of $\mathbb{L}_2(\Omega)$. For all $t, u \in \mathbb{H}$, $\text{Cov}(L(t), L(u)) = \langle t, u \rangle$.

Our main goal will be developing procedures which adapt to the smoothness of the underlying function $s \in \mathbb{H}$ in the framework of model selection as proposed by Barron et al. [1] and Birgé and Massart [3].

Following Birgé [2] we take a model selection point of view at adaptive estimation via Lepski's method. In order to construct the adaptive estimator we consider a family of nested finite dimensional linear subspaces of $\mathbb{H}$. Over each subspace S of the family, we consider an estimator of $T(s)$ based on the estimation of the orthogonal projection of s onto S. The problem is thus establishing a procedure to select an estimator among the collection of estimators that we have built. We present a general procedure based on a penalized criterion. We obtain an oracle inequality, which shows that the quadratic risk of the selected estimator is close to the infimum of the quadratic risks of the estimators that we have considered in the collection.

As an the example, we consider the estimation of the r^{th} derivative of a function at a point using a multiresolution analysis framework. Minimax rates are shown to hold uniformly over Besov spaces. We present numerical examples and we compare our method to a global (not pointwise) model selection estimator.

References

[1] A.R. Barron, L. Birgé, P. Massart (1999), *Risk bounds for model selection via penalization*, Probab. Theory Related Fields, **113**, 301-415.

[2] L. Birgé (2001). *An alternative point of view on Lepski's method.* State of the art in probability and statistics (Leiden, 1999), 113–133, IMS Lecture Notes Monogr. Ser., 36, Inst. Math. Statist., Beachwood, OH.

[3] L. Birgé, P. Massart (2001). *Gaussian model selection.* J. Eur. Math. Soc. (JEMS) **3**, no. 3, 203–268.