

Slope heuristics and optimal excess risks bounds in heteroscedastic least-squares regression

Adrien Saumard
Universidad de Valparaíso, CIMFAV

May, 7th 2014

The systematical study of model selection procedures, especially since the early nineties, has led to the design of penalties that often allow to achieve minimax rates of convergence and adaptivity for the selected model, in the general setting of risk minimization (Koltchinskii [3], Massart [4]).

However, the proposed penalties often suffer from their dependencies on unknown or unrealistic constants. As a matter of fact, under-penalization has generally disastrous effects in terms of efficiency. Indeed, the model selection procedure then loses any bias-variance trade-off and so, tends to select one of the biggest models in the collection.

Birgé and Massart ([2]) proposed quite recently a method that empirically adjusts the level of penalization in a linear Gaussian setting. This method of calibration is called "slope heuristics" by the authors, and is proved to be optimal in their setting. It is based on the existence of a minimal penalty, which is shown to be half the optimal one.

Arlot and Massart ([1]) have then extended the slope heuristics to the more general framework of empirical risk minimization. They succeeded in proving the optimality of the method in heteroscedastic least-squares regression, a case where the ideal penalty is no longer linear in the dimension of the models, not even a function of it. However, they restricted their analysis to histograms for technical reasons. They conjectured a wide range of applicability for the method.

We will present some results that prove the validity of the slope heuristics in heteroscedastic least-squares regression for more general linear models than histograms. The models considered here are equipped with a localized orthonormal basis, among other things. We show that some piecewise polynomials and Haar expansions satisfy the prescribed conditions.

We will insist on the analysis when the model is fixed. In particular, we will focus on deviations bounds for the true and empirical excess risks of the estimator. Empirical process theory and concentration inequalities are central tools here, and the results at a fixed model may be of independent interest.

References

- [1] S. Arlot and P. Massart. Data-driven calibration of penalties for least-squares regression. *J. Mach. Learn. Res.*, 10:245–279 (electronic), 2009.
- [2] L. Birgé and P. Massart. Minimal penalties for Gaussian model selection. *Probab. Theory Related Fields*, 138(1-2):33–73, 2007.
- [3] Vladimir Koltchinskii. *Oracle inequalities in empirical risk minimization and sparse recovery problems*, volume 2033 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 38th Probability Summer School held in Saint-Flour, 2008, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].
- [4] Pascal Massart. *Concentration inequalities and model selection*, volume 1896 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard.