

Sparsity oracle inequalities for the Lasso

Florentina Bunea, *FSU*
Alexandre Tsybakov, *Paris 6*
Marten H. Wegkamp, *FSU*

Abstract

This paper studies oracle properties of ℓ_1 -penalized least squares in nonparametric regression setting with random design. We show that the penalized least squares estimator satisfies sparsity oracle inequalities, i.e., bounds in terms of the number of non-zero components of the oracle vector. The results are valid even when the dimension of the model is (much) larger than the sample size and the regression matrix is not positive definite. They can be applied to high-dimensional linear regression, to nonparametric adaptive regression estimation and to the problem of aggregation of arbitrary estimators.

AMS 2000 subject classifications: Primary 62G08; secondary 62C20, 62G05, 62G20.

Keywords: sparsity, oracle inequalities, Lasso, penalized least squares, nonparametric regression, dimension reduction, aggregation, mutual coherence, adaptive estimation.



Full Text: [PDF](#)

Bunea, Florentina, Tsybakov, Alexandre, Wegkamp, Marten H., Sparsity oracle inequalities for the Lasso, *Electronic Journal of Statistics*, 1, (2007), 169-194 (electronic). DOI: 10.1214/07-EJS008.

References

- [1] Abramovich, F., Benjamini, Y., Donoho, D.L. and Johnstone, I.M.. (2006). Adapting to unknown sparsity by controlling the False Discovery Rate. *Annals of Statistics* 34(2) 584-653. [MR2281879](#)
- [2] Antoniadis, A. and Fan, J. (2001). Regularized wavelet approximations (with discussion). *Journal of American Statistical Association* 96 939 – 967. [MR1946364](#)
- [3] Baraud, Y. (2002). Model selection for regression on a random design. *ESAIM Probability & Statistics* 7 127–146. [MR1918295](#)
- [4] Birgé, L. (2004). Model selection for Gaussian regression with random design. *Bernoulli* 10 1039–1051. [MR2108042](#)
- [5] Bunea, F., Tsybakov, A.B. and Wegkamp, M.H. (2005) Aggregation for Gaussian regression. Preprint Department of Statistics, Florida state University.
- [6] Bunea, F., Tsybakov, A.B. and Wegkamp, M.H. (2006). Aggregation and sparsity via ℓ_1 -penalized least squares. *Proceedings of 19th Annual Conference on Learning Theory, COLT 2006. Lecture Notes in Artificial Intelligence* 4005 379–391. Springer-Verlag, Heidelberg. [MR2280619](#)
- [7] Candès, E. and Tao, T. (2005). The Dantzig selector: statistical estimation when p is much larger than n . Manuscript.

- [8] Donoho, D.L., Elad, M. and Temlyakov, V. (2004). Stable Recovery of Sparse Overcomplete Representations in the Presence of Noise. Manuscript.
- [9] Donoho, D.L. and Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81 425–455. [MR1311089](#)
- [10] Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression. *Annals of Statistics* 32(2) 407–451. [MR2060166](#)
- [11] Golubev, G.K. (2002). Reconstruction of sparse vectors in white Gaussian noise. *Problems of Information Transmission* 38(1) 65–79. [MR2101314](#)
- [12] Greenshtein, E. and Ritov, Y. (2004). Persistency in high dimensional linear predictor-selection and the virtue of over-parametrization. *Bernoulli* 10 971–988. [MR2108039](#)
- [13] Györfi, L., Kohler, M., Krzyżak, A. and Walk, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. Springer: New York. [MR1987657](#)
- [14] Juditsky, A. and Nemirovski, A. (2000). Functional aggregation for nonparametric estimation. *Annals of Statistics* 28 681–712. [MR1792783](#)
- [15] Kerkycharian, G. and Picard, D. (2005). Thresholding in learning theory. Prépublication n.1017, Laboratoire de Probabilités et Modèles Aléatoires, Universités Paris 6 - Paris 7.
- [16] Knight, K. and Fu, W. (2000). Asymptotics for lasso-type estimators. *Annals of Statistics* 28(5) 1356–1378. [MR1805787](#)
- [17] Koltchinskii, V. (2005). Model selection and aggregation in sparse classification problems. Oberwolfach Reports: Meeting on Statistical and Probabilistic Methods of Model Selection, October 2005 (to appear). [MR2238841](#)
- [18] Koltchinskii, V. (2006). Sparsity in penalized empirical risk minimization. Manuscript.
- [19] Loubes, J.-M. and van de Geer, S.A. (2002). Adaptive estimation in regression, using soft thresholding type penalties. *Statistica Neerlandica* 56 453 – 478.
- [20] Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *Annals of Statistics* 34 (3) 1436–1462. [MR2278363](#)
- [21] Nemirovski, A. (2000). *Topics in Non-parametric Statistics*. Ecole d’Eté de Probabilités de Saint-Flour XXVIII - 1998, Lecture Notes in Mathematics, v. 1738, Springer: New York. [MR1775640](#)
- [22] Osborne, M.R., Presnell, B. and Turlach, B.A (2000a). On the lasso and its dual. *Journal of Computational and Graphical Statistics* 9 319 – 337. [MR1822089](#)
- [23] Osborne, M.R., Presnell, B. and Turlach, B.A (2000b). A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis* 20(3) 389 – 404.
- [24] Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B.* 58 267–288. [MR1379242](#)
- [25] Tsybakov, A.B. (2003). Optimal rates of aggregation. *Proceedings of 16th Annual Conference on Learning Theory (COLT) and 7th Annual Workshop on Kernel Machines*. Lecture Notes in Artificial Intelligence 2777 303–313. Springer-Verlag, Heidelberg.
- [26] Turlach, B.A. (2005). On algorithms for solving least squares problems under an L1 penalty or an L1 constraint. 2004 Proceedings of the American Statistical Association, Statistical Computing Section [CD-ROM], American Statistical Association, Alexandria, VA, pp. 2572-2577.
- [27] van de Geer, S.A. (2006). High dimensional generalized linear models and the Lasso. Research report No.133. Seminar für Statistik, ETH, Zürich.

[28] Wainwright, M.J. (2006). Sharp thresholds for noisy and high-dimensional recovery of sparsity using ℓ_1 -constrained quadratic programming. Technical report 709, Department of Statistics, UC Berkeley.

[29] Zhao, P. and Yu, B. (2005). On model selection consistency of Lasso. Manuscript.

[Home](#) | [Current](#) | [Past volumes](#) | [About](#) | [Login](#) | [Notify](#) | [Contact](#) | [Search](#)

Electronic Journal of Statistics. ISSN: 1935-7524