



HAL
open science

Inégalités d'oracle exactes pour la prédiction d'une matrice en grande dimension

Stéphane Gaïffas, Guillaume Lecué, Alexandre B. Tsybakov

► **To cite this version:**

Stéphane Gaïffas, Guillaume Lecué, Alexandre B. Tsybakov. Inégalités d'oracle exactes pour la prédiction d'une matrice en grande dimension. 42èmes Journées de Statistique, 2010, Marseille, France, France. inria-00494750

HAL Id: inria-00494750

<https://hal.inria.fr/inria-00494750>

Submitted on 24 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INÉGALITÉS D'ORACLE EXACTES POUR LA PRÉDICTION D'UNE MATRICE EN GRANDE DIMENSION

Stéphane Gaïffas, Guillaume Lecué, Alexandre Tsybakov

*Université Pierre et Marie Curie - Paris 6,
CNRS, Université Paris-Est, Marne la Vallée,
CREST and Université Pierre et Marie Curie - Paris 6*

1. Abstract. We consider the problem of prediction of a high dimensional noisy matrix of size $m \times T$, meaning that mT is much larger than the sample size n . We focus on penalization methods, using the so-called trace or nuclear norm and the Frobenius norm. Such algorithms are now popular, since they are very efficient for matrix completion problems, as well as other problems, such as multi-task learning, see [2, 4, 3, 5, 7, 1]. In this work, we propose sharp oracle inequalities for these algorithms in the noisy case, using a statistical learning theory point of view.

1. Résumé. Nous considérons le problème de prédiction d'une matrice de taille $m \times T$ en "grande dimension", c'est-à-dire où mT est bien plus grand que la taille de l'échantillon n . Pour cela, nous considérons l'algorithme de minimisation de la norme trace, ou des versions "élastiques" de la norme. Cet algorithme est maintenant bien connu et utilisé pour les problèmes de complétion de matrice ou d'apprentissage multi-tâche, voir notamment [2, 4, 3, 5, 7, 1], parmi d'autres. Dans ce travail, nous proposons des inégalités d'oracle "exactes" pour ces algorithmes.

References

- [1] Francis R. Bach. Consistency of trace norm minimization. *J. Mach. Learn. Res.*, 9:1019–1048, 2008.
- [2] E. J. Candès and Y. Plan. Tight oracle bounds for low-rank matrix recovery from a minimal number of random measurements. Technical report, Department of Statistics, Stanford University, 2009.
- [3] E. J. Candès and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, to appear.
- [4] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Found. of Comput. Math.*, 9:717–772, 2008.
- [5] E. J. Candès and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Trans. Inform. Theory*, to appear.

- [6] M. Fazel, H. Hindi, and S. Boyd. Rank minimization and applications in system theory. In *In American Control Conference*, pages 3273–3278. AACC, 2004.
- [7] Angelika Rohde and Alexandre B. Tsybakov. Estimation of high-dimensional low rank matrices. Technical report, Universität Hamburg and Université Paris 6, 2009.