

Plans d'expériences "space filling" pour le Dantzig selector

Le Dantzig selector permet de faire des choix de sous-modèles dans des modèles linéaires même quand le nombre de paramètres est du même ordre, voire beaucoup plus grand que le nombre de données. Il est basé sur l'idée que le vecteur des paramètres est sporadique : la plupart de coefficients sont nuls. Comme dans la méthodes LASSO on minimise la norme L^1 du paramètre estimé mais sous une contrainte de type L^∞ sur les erreurs.

La méthode fonctionne si les régresseurs, pris par paquets de dimension inférieure au nombre de données, ont une matrice d'information dont les plus grandes valeur propres sont majorées et dont les plus petites sont minorées. Sous ces conditions on obtient des inégalités oracle.

Ces conditions sont vérifiées, avec une grande probabilité, si les régresseurs sont des directions prises au hasard sur la sphère unité de \mathbb{R}^n (n est le nombre de données). Le but de la thèse est de choisir des directions de manière (plus) déterministe, en utilisant les critères et les idées des plans "space-filling". On utilisera pour cela des hypercubes latins, des plans "orthogonal array" des quantifications de la sphere, par exemple par des "lattice rules" ainsi que des algorithmes stochastiques de type recuit simulé pour améliorer éventuellement les plans construits par les méthodes précédentes.

References

Travaux de Emmanuel Candes, California Institute of Technology et Terence Tao (UCLA).

Box, Hunter and Hunter (2005). Statistics for experimenters second edition. Wiley

Kochol M., Constructive approximation of a ball by polytopes. Math. Slovaca 44(1) :99-150, 1994.

Koehler et Owen (1996). Computer experiments. In : Ghosh and Rao Eds, Handbook of Statistics, Vol. 13, pp.261-307. Elsevier Science.

Nuyens D. and Cools R. Fast algorithms for comonent-by-component construction of rank-1 Lattice rule in shift invariant reproducing kernel Hilbert spaces. *Mathematics of computation*, 75, 254, 903-920.