

# CDRodeo : Sélection gloutonne de fenêtres multivariées pour l'estimation de densité conditionnelle

Minh-Lien Jeanne NGUYEN, Universités Paris-Sud et Paris-Dauphine

**Mots-clés** : estimation non paramétrique, grande dimension, parcimonie, densité conditionnelle, algorithmes gloutons, estimateurs à noyau,

Dans cet exposé, on s'intéresse à la sélection de la fenêtre pour l'estimation à noyau de densités conditionnelles. On s'attaque plus spécifiquement aux cas de dimensions modérément grandes, et l'enjeu est de contourner le fléau de la dimension. La méthode CDRodeo [1], inspirée de [2, 3], offre plusieurs avantages : une vitesse de convergence optimale au sens minimax (à un facteur logarithmique près), des temps d'exécution compétitifs grâce au caractère itératif et glouton de la procédure, ainsi que la détection de variables pertinentes en cas de parcimonie. En particulier, CDRodeo est doublement adaptatif : en la régularité et en la structure de parcimonie.

## Références

- [1] Nguyen, M.-L.J. (2018, en révision) *Nonparametric method for sparse conditional density estimation in moderately large dimensions*. hal-01688664.
- [2] Lafferty J.D., Wasserman L.A. (2008) *Rodeo : Sparse, greedy nonparametric regression*. Annals of Statistics, Vol. 36, No. 1, 28-63.
- [3] Liu H., Lafferty J.D., Wasserman L.A. (2007) *Sparse Nonparametric Density Estimation in High Dimensions Using the Rodeo*. AISTATS, 283-290.