

# CDRodeo : Sélection gloutonne de fenêtres multivariées en estimation de densité conditionnelle

Minh-Lien Jeanne NGUYEN, Université Paris-Sud

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

Dans cet exposé, on s'intéresse à l'estimation de densités conditionnelles en modérément grandes dimensions. La densité conditionnelle, moins restrictive que la fonction de régression, permet d'en dériver de nombreuses autres quantités d'intérêt (quantiles, intervalles de confiance, modes...). À partir d'une famille d'estimateurs à noyau adaptée pour l'estimation de densité conditionnelle, on revisite l'algorithme RODEO [2, 3] pour sélectionner une fenêtre locale multivariée. La méthode répond à plusieurs enjeux : éviter le fléau de la dimension en combinant une exécution rapide et la détection de variables non pertinentes (en cas de fonction parcimonieuse) et converger à vitesse optimale au sens minimax (à un facteur logarithmique près).

## Références

- [1] Nguyen, M.-L.J. (2018, pré-publication) *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.