Machine Learning in Econometrics

February 2016

The goal of this course is to give an introduction to Machine Learning in Econometrics. In the first part, methods from Machine Learning are presented. In the second part, applications of those methods in Econometrics are discussed. Moreover, the statistical software package R is used for illustration of the methods.

Lectures

1. Introduction

- Definitions
- Basic Concepts
- Challenges in High-Dimensions

Part I. How to make predictions?

1. Linear Regression and Extensions

- Recap: Linear Regression
- Regression Splines
- Smoothing Splines

2. Ridge Regression
3. Lasso Regression

- Basic Principle
- Some Results
- Extensions

4. Of Trees and Forests

- Regression Trees
- Bagging
- Random Forests

5. Boosting

- Basic Idea
- L_2Boosting for Regression

6. Model Selection: How to choose between different models?

Part II. Estimation and Inference of Structural Parameters and Treatment Effects

1. Partialling-out
2. Inference on Selected Target Variables in Regressions in High-Dimensions
3. IV Estimation in High-Dimensions
4. The Orthogonality Principle

Problem Sets

- [PS1](#) Introduction to R /RStudio and some important R Packages
- [Ps2](#) Linear Regression and Extensions
- [Ps3](#) Ridge and Lasso Regression

Literature

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2001. *The Elements of Statistical Learning.* Springer Series in Statistics. New York, NY, USA: Springer New York Inc.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2014. *An Introduction to Statistical Learning: With Applications in R.* Springer Publishing Company, Incorporated.

Shalev-Shwartz, Shai, and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms.* New York, NY, USA: Cambridge University Press. http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf.

Victor Chernozhukov, Martin Spindler, Chris Hansen. 2016. *hdm: High-Dimensional Metrics (Vignette).*