Studies in Advanced Econometrics (212.664 - 002)
Spring 2023

Course:

Lectures:	Tuesday, 09:30-12:15 (Classroom: M306)
TA Sessions:	TBA

Instructor: Prof. Myung Hwan Seo

Office:	16.644
Email:	myunghseo@snu.ac.kr
Office Hours:	Tuesday 13:30–14:30 or by appointment

Course Description: This course intends to expose the students to the current econometric and statistical approaches to the machine learning methods and theory that deal with high-dimensional models and big data environment. It introduces various high-dimensional estimation and model selection methods for general regression models such as the linear regression, instrumental variable regression, logistic regression, quantile regression with a particular emphasis on the least absolute shrinkage and selection operator (aka lasso). It develops various distribution theory to cover related issues on prediction/forecasting and statistical inference. And it also discusses numerical algorithms for the methods and demonstrates the execution of the algorithms. The lectures are based mainly on the first three books in the reference list below.

Evaluation: In-class Written Examination, , Assignment, Presentations.

- Presentation can be given on the following:
    1. Application on real data.
    2. Monte Carlo simulation (for both a new method and an existing one).
    3. Give a short lecture on a related method in machine learning literature that is not covered in the class.
    4. Theoretical extension of a recent paper.

Outline

1. Review of the (linear) regression
    (a) Variable selection and post-selection Inference (Hansen Ch.28)

(b) Series Estimation (Hansen Ch.20)

(c) Logit/Probit (Hansen Ch.25)

(d) Linear Regression with Change-Points (Structural Break/Threshold/Tipping)

2. Introduction to the Lasso ([BG])

   (a) Numerical property of a lasso solution

   (b) Theoretical property

   (c) Optimization Algorithm

3. Variants of Lasso

   (a) Adaptive lasso

   (b) Weighted lasso

   (c) Fussed lasso

   (d) SCAD

4. Inference

   (a) Double selection

   (b) Residual Regression

5. Other Prediction Methods (Machine Learning)

   - Some generic big data environment and applications ([5, 6, 11])
   - Supervised vs Un-supervised Learning
   - Factor Analysis
   - Regression Trees
   - Bagging:
     - Review of Bootstrap
   - Random forest
   - Neural Networks
   - Boosting
   - Support Vector Machines

6. Stochastic Grandient Descent

7. Inference on potentially many parameters

References

[BG] Bühlmann, P. and S. van de Geer (2011). Statistics for High-Dimensional Data. DE, Springer.

[1] Hastie, T., Tibshrani, R., Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, Springer.

[2] Hansen, B. (2022). Econometrics. Princeton University Press.

[3] James, G., D. Witten, T. Hastie, and R. Tibshirani (2014): An Introduction to Statistical Learning: With Applications in R. Springer New York.

[4] van der Vaart, A. and J. Wellner (1996). Weak Convergence and Empirical Processes. Springer.

[5] Athey, S. (2017). Beyond prediction: Using big data for policy problems. Science, 355(6324), 483-485.

[6] Athey, S. (2019). 21. The Impact of Machine Learning on Economics. The economics of artificial intelligence, 507-552.

[7] Athey, S., and Imbens, G. W. (2019). Machine learning methods that economists should know about. Annual Review of Economics, 11, 685-725.

[8] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters.

[9] Andrews, D. W. K. (1994): Empirical Process Methods in Econometrics,in Handbook of Econometrics, Vol.\ Iv. Amsterdam: North-Holland, 2247–2294.

[10] Chen, Xiaohong (2007): "Large Sample Sieve Estimation of Semi-Nonparametric Models," In: James J. Heckman and Edward E. Leamer, Editor(s), Handbook of Econometrics, Elsevier, Volume 6, Part B, 5549-5632,

[11] Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. Journal of Economic Perspectives, 31(2), 87-106.