ECON 241C: Machine Learning for Economics
Fall 2023

Instructor Vira Semenova

- Class time: Thursday 10-12am
- OH: Thursday afternoon (depending on Econometrics seminar schedule). Evans Hall 667 or Zoom. Please book a slot here.
- Email: `vsemenova@berkeley.edu`. Please put ECON 241C in subject line.
- Zoom: `https://berkeley.zoom.us/my/vsemenova`

Course description The course covers topics both in classic nonparametric and modern approaches. Topics included, but are not limited to, decision theory, Bayesian methods, high-dimensional models, causal inference, and applications of machine learning in economics. The course is targeted both for students interested in specializing in econometrics as well as applied students interesting in expanding their econometric toolbox.

Course objectives The students will be given tools to conduct statistical analysis, understand the role of empirical evidence in evaluating economic problems, as well as hands-on experience working with data sets and applying machine learning algorithms to economic data.

Course outline (subject to changes)

	Topic
1	Decision Theory
2	Shrinkage in normal means model
3	Quasi-Bayesian methods
4	Uniformity
5	Sparse Models and Methods. Lasso
6	Double Lasso. Double Machine Learning
7	Heterogenous Treatment Effects
	Causal ML
8	Policy Learning
9	Weak instruments
10	Bandits
11-12	Reinforcement Learning

Assignments

- Check deadlines and Submit assignments at **Gradescope** . The course code for Gradescope is **J3WZJD** .
- Grading.
  - Homeworks (80 %). There will be $N = 5$ of them. The best N-1 grades count (i.e., each of the best HW is worth 20% of course grade). If you hand-in all N homeworks you are guaranteed to earn a passing grade (this does take away the anxiety and allows you to focus on learning). Late homeworks are not graded. **Homeworks may be completed in groups of size 1 or 2** .
  - Final assessment (20 %).

References Required reading is marked by *.

- Review of decision theory. Robert (2001) (Chapter 2*)
- Statistical Learning Theory. Lectures by Philippe Rigollet* (b-courses),
- Concentration. Rigollet and Hutter (2017) (Chapter 1*)
- Complexity and VC dimension. Shalev-Shwartz and Ben-David (2014) (Chapters 2-6*) or Wasserman (2006)(Chapter 2.5*)
- Shrinkage in normal means model. Wasserman (2006)(Chapter 7*), Efron (2010)(Chapter 1), Robert (2001) (Chapter 10.4-10.5)
- Double Machine Learning. Chernozhukov et al. (2018) (Section 2*), Belloni et al. (2017)
- Weak Instruments. Moreira (2003)*, Andrews et al. (2019), Mikusheva and Sun (2020)
- Bandits Lectures by Shipra Agrawal* (b-courses), Slivkins (2021), Russo et al. (2020),
- Model selection Leeb and Potcher (2005)
- High-Dimensional Models. Belloni and Chernozhukov (2011a), Belloni and Chernozhukov (2011b), Belloni et al. (2011), Belloni et al. (2014), Belloni et al. (2018), Negahban et al. (2012)
- Heterogenous Treatment Effects. Athey and Imbens (2016)*, Semenova and Chernozhukov (2021)*, Oprescu et al. (2018), Chernozhukov et al. (2017), Colangelo and Lee (2020),

Useful resources

- Shipra Agarwal's course on multi-armed bandits
- Shipra Agarwal's course on reinforcement learning
- Max Kasy's course on machine learning in economics
- Philippe Rigollet's course on mathematics of machine learning and references therein
- Alex Slivkins course on bandits

Academic Integrity. Please read the Center for Student Conduct's statement on Academic Integrity at `http://sa.berkeley.edu/conduct/integrity`. I take issues of intellectual honest very seriously.

References

Andrews, I., Stock, J., and Sun, L. (2019). Weak instruments in iv regression: Theory and practice. *Annual Review of Economics*, 11:727–753.

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7460.

Belloni, A. and Chernozhukov, V. (2011a). ℓ_1-penalized quantile regression in high-dimensional sparse models. *The Annals of Statistics*, 39(1):82–130.

Belloni, A. and Chernozhukov, V. (2011b). High dimensional sparse econometric models: An introduction.

Belloni, A., Chernozhukov, V., Chetverikov, D., Hansen, C., and Kato, K. (2018). High-dimensional econometrics and regularized GMM. *arXiv e-prints*, page arXiv:1806.01888.

Belloni, A., Chernozhukov, V., Fernandez-Val, I., and Hansen, C. (2017). Program evaluation and causal inference with high-dimensional data. *Econometrica*, 85:233–298.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection amongst high-dimensional controls. *Journal of Economic Perspectives*, 28(2):608–650.

Belloni, A., Chernozhukov, V., and Wang, L. (2011). Square-root lasso: pivotal recovery of sparse signals via conic programming. *Biometrika*, 98(4):791–806.

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. *Econometrics Journal*, 21:C1–C68.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernández-Val, I. (2017). Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments. *arXiv e-prints*, page arXiv:1712.04802.

Colangelo, K. and Lee, Y.-Y. (2020). Double Debiased Machine Learning Nonparametric Inference with Continuous Treatments. *arXiv e-prints*, page arXiv:2004.03036.

Efron, B. (2010). *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction*. Cambridge University Press.

Leeb, H. and Potcher, B. (2005). Model selection and inference: Facts and fiction. *Econometric Theory*, 21:21–59.

Mikusheva, A. and Sun, L. (2020). Inference with many weak instruments.

Moreira, M. (2003). A conditional likelihood ratio test for structural models. *Econometrica*, 71:1027–1048.

Negahban, S. N., Ravikumar, P., Wainwright, M. J., and Yu, B. (2012). A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. *Statistical Science*, 27(4):538–557.

Oprescu, M., Syrgkanis, V., and Wu, Z. S. (2018). Orthogonal random forest for causal inference. *arXiv e-prints*, page arXiv:1806.03467.

Rigollet, P. and Hutter, J.-C. (2017). High-dimensional statistics.

Robert, C. (2001). *The Bayesian choice*. Springer.

Russo, D., Roy, B. V., Kazerouni, A., Osband, I., and Wen, Z. (2020). A tutorial on thompson sampling.

Semenova, V. and Chernozhukov, V. (2021). Debiased machine learning of conditional average treatment effect and other causal functions. *Econometrics Journal*, (24).

Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, New York, NY, USA.

Slivkins, A. (2021). Introduction to multi-armed bandits.

Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer Texts in Statistics, New York, NY, USA.