with Christian B. Hansen

Course Description:

As in many other fields, economists are increasingly making use of high-dimensional models – models with many unknown parameters that need to be inferred from the data. Such models arise naturally in modern data sets that include rich information for each unit of observation (a type of "big data") and in nonparametric applications where researchers wish to learn, rather than impose, functional forms. High-dimensional models provide a vehicle for modeling and analyzing complex phenomena and for incorporating rich sources of confounding information into economic models.

Our goal in this course is two-fold. First, we wish to provide an overview and introduction to several modern methods, largely coming from statistics and machine learning, which are useful for exploring high-dimensional data and for building prediction models in high-dimensional settings. Second, we will present recent proposals that adapt high-dimensional methods to the problem of doing valid inference about model parameters and illustrate applications of these proposals for doing inference about economically interesting parameters.

Outline (Preliminary and Incomplete):

Lecture 1: Traditional Nonparametrics and an Introduction to High-Dimensional Modeling

- Hastie, T., R. Tibshirani, and J. Friedman (2009), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer. [Elements from Chapters 2, 5, 7, 8.7, 10]
- James, G., D. Witten, T. Hastie, and R. Tibshirani (2014), An Introduction to Statistical Learning with Applications in R, Springer. [Elements from Chapters 2, 3, 5, 7, 8.2]
- Li, Q. and J. S. Racine (2007), Nonparametric Econometrics: Theory and Practice, Princeton University Press. [Elements from Chapters 2, 14]

Lecture 2: Penalized Estimation Methods

- Hastie, T., R. Tibshirani, and J. Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer. [Elements from Chapters 3-4]
- James, G., D. Witten, T. Hastie, and R. Tibshirani (2014), *An Introduction to Statistical Learning with Applications in R*, Springer. [Chapters 6]
- Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012), "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 81(2), 608-650.
- Belloni, A. and V. Chernozhukov (2013), "Least squares after model selection in high-dimensional sparse models," *Bernoulli*, 19(2), 521-547.
- Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016), "Inference in high-dimensional panel models with an application to gun control," *Journal of Business and Economic Statistics*, 34(4), 590-605.
- Chetverikov, D., Z. Liao, and V. Chernozhukov (2019), "On cross-validated lasso," arXiv:1605.02214.
- Fan, J. and J. Lv (2008), "Sure independence screening for ultrahigh dimensional feature space," Journal of the Royal Statistical Society, Series B, 70(5), 849-911.

Lecture 3: Tree-based Methods

- Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen (1984), *Classification and Regression Trees*, CRC Press.
- Hastie, T., R. Tibshirani, and J. Friedman (2009), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer. [Chapters 9, 10, 15, 16]
- James, G., D. Witten, T. Hastie, and R. Tibshirani (2014), An Introduction to Statistical Learning with Applications in R, Springer. [Chapter 8]
- Asher, S., D. Nekipelov, P. Novosad, and S. Ryan (2016), "Classification Trees for Heterogeneous Moment-Based Models," NBER Working Paper 22976.
- Athey, S. and G. Imbens (2016), "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 113 (27), 7353–7360.

- Bajari, P., D. Nekipelov, S. P. Ryan, and M. Yang (2015), "Machine Learning Methods for Demand Estimation," *American Economic Review*, 105 (5): 481-85. (See also NBER Working Paper 20955)
- Breiman, L. (1996), "Bagging Predictors," Machine Learning 26: 123-140
- Friedman, J., T. Hastie, and R. Tibshirani (2000), "Additive logistic regression: A statistical view of boosting (with discussion)," Annals of Statistics, 28, 337-407
- Schapire, R. (1990), "The strength of weak learnability," Machine Learning, 5, 197-227
- Taddy, M., C. Chen, J. Yu, and M. Wyle (2015), "Bayesian and empirical Bayesian forests," *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning*, 37, 967-976.
- Wager, S. and S. Athey (2018), "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113(523), 1228-1242.
- Wager, S. and G. Walther (2016), "Adaptive Concentration of Regression Trees, with Application to Random Forests," arxiv:1503.06388
- Wager, S., T. Hastie, and B. Efron (2014), "Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife," Journal of Machine Learning Research, 15, 1625−1651

Lecture 4: An Overview of High-Dimensional Inference

- Ai, C. and X. Chen (2012), "The semiparametric efficiency bound for models of sequential moment restrictions containing unknown functions," *Journal of Econometrics* 170, 442-457.
- Andrews, D. W. K. (1994), "Asymptotics for semiparametric econometric models via stochastic equicontinuity," *Econometrica*, 62, 43{72.
- Belloni, A., D. Chen, V. Chernohukov, and C. Hansen (2012), "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain," Econometrica, 80(6), 2369-2430
- Belloni, A., V. Chernozhukov, and C. Hansen (2014), "Inference on Treatment Effects after Selection amongst High-Dimensional Controls," Review of Economic Studies, 81(2), 608-650
- Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016), "Inference in high-dimensional panel models with an application to gun control," *Journal of Business and Economic Statistics*, 34(4), 590-605.
- Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2017), "Program Evaluation and Causal Inference with High-Dimensional Data," *Econometrica*, 85(1), 233-298.
- Bickel, P. J. (1982), "On adaptive estimation," *Annals of Statistics*, 10, 647-671
- Bickel, P. J., C. A. J. Klaassen, Y. Ritov, and J. A.Wellner (1998), *Efficient and Adaptive Estimation for Semiparametric Models*, Springer.
- Chen, X., O. Linton, and I. van Keilegom (2003), "Estimation of semiparametric models when the criterion function is not smooth," *Econometrica* 71, 1591-1608.
- V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018), "Double/debiased machine learning for treatment and structural parameters," *Econometrics Journal*, 21(1), C1-C68.
- Chernozhukov, V., C. Hansen, and M. Spindler (2015), "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments," American Economic Review, 105(5), 486-490

- Chernozhukov, V., C. Hansen, and M. Spindler (2015), "Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach," Annual Review of Economics, 7, 649-688
- Hansen, C. and D. Kozbur (2014), "Instrumental Variables Estimation with Many Weak Instruments Using Regularized JIVE," *Journal of Econometrics*, 182(2), 290-308
- Hansen, C., D. Kozbur, and S. Misra (2018), "Targeted undersmoothing," arXiv: 1706:07328
- Ibragimov, I. A. and R. Z. Hasminskii (1981), *Statistical Estimation: Asymptotic Theory*, Springer-Verlag, New York.
- Javanmard, A. and A. Montanari (2014), "Confidence Intervals and Hypothesis Testing for High-Dimensional Regression," *Journal of Machine Learning Research*, 15, 2869-2909
- Levit, B. Y. (1975). "On the efficiency of a class of nonparametric estimates," *Theory of Probability and Its Applications* 20, 723-740
- Linton, O. (1996), "Edgeworth approximation for MINPIN estimators in semiparametric regression models," *Econometric Theory* 12, 30-60.
- Newey, W. (1990), "Semiparametric efficiency bounds," *Journal of Applied Econometrics*, 5, 99-135.
- Newey, W. (1994), "The asymptotic variance of semiparametric estimators," *Econometrica*, 62, 1349-1382.
- Newey, W. K., F. Hsieh, and J. Robins (1998), "Undersmoothing and bias corrected functional estimation," Working paper, MIT Economics Dept., http://economics.mit.edu/files/11219.
- Newey, W. K., F. Hsieh, and J. M. Robins (2004), "Twicing kernels and a small bias property of semiparametric estimators," *Econometrica*, 72, 947-962.
- Robins, J. and A. Rotnitzky (1995), "Semiparametric efficiency in multivariate regression models with missing data," *Journal of the American Statistical Association*, 90, 122-129.
- Robinson, P. M. (1988): "Root-N-consistent semiparametric regression," *Econometrica*, 56(4), 931–954.
- van de Geer, S., P. Bühlmann, Y. Ritov, and Dezeure, R. (2014), "On asymptotically optimal confidence regions and tests for high-dimensional models," *Annals of Statistics*, 42, 1166-1202
- van der Laan, M. J. and S. Rose (2011), *Targeted Learning: Causal Inference for Observational and Experimental Data*, Springer.
- Zhang, C.-H. and S. Zhang (2014), "Confidence intervals for low dimensional parameters in high dimensional linear models," *Journal of the Royal Statistical Society, Series B*, 76(1), 217-242

Lecture 5: Further Examples

- Belloni, A., D. Chen, V. Chernohukov, and C. Hansen (2012), "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain," *Econometrica*, 80(6), 2369-2430
- Belloni, A., V. Chernozhukov, and C. Hansen (2014), "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, 28(2), 29-50

- Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016), "Inference in high-dimensional panel models with an application to gun control," *Journal of Business and Economic Statistics*, 34(4), 590-605.
- Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2017), "Program Evaluation and Causal Inference with High-Dimensional Data," *Econometrica*, 85(1), 233-298.
- Chernozhukov, V., C. Hansen, and M. Spindler (2015), "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments," *American Economic Review*, 105(5), 486-490
- Chernozhukov, V., C. Hansen, and M. Spindler (2015), "Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach," *Annual Review of Economics*, 7, 649-688
- Gentzkow, M., J. Shapiro, and M. Taddy (2019), "Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech," *Econometrica*, 87(4), 1307-1340.
- Kleinberg, J., J. Ludwig, S. Mullainathan, and Z. Obermeyer (2015), "Prediction Policy Problems," *American Economic Review: Papers and Proceedings*, 105(5), 491-495

Further References

- Belloni, A., and V. Chernozhukov (2011): "l1-penalized quantile regression in high-dimensional sparse models," *Annals of Statistics*, 39(1), 82–130.
- Belloni, A., V. Chernozhukov, and C. Hansen (2010) "Inference for High-Dimensional Sparse Econometric Models," *Advances in Economics and Econometrics*. 10th World Congress of Econometric Society, Shanghai, 2010.
- Belloni, A., V. Chernozhukov, K. Kato (2013): "Uniform Post Selection Inference for LAD Regression Models," arXiv:1304.0282.
- Belloni, A., V. Chernozhukov, L. Wang (2011): "Square-Root-LASSO: Pivotal Recovery of Sparse Signals via Conic Programming," *Biometrika*, 98(4), 791–806.
- Belloni, A., V. Chernozhukov, Y. Wei (2016): "Post-selection inference for generalized linear models with many controls," *Journal of Business and Economic Statistics,* 34(4), 606-619.
- Bickel, P., Y. Ritov and A. Tsybakov, "Simultaneous analysis of Lasso and Dantzig selector", *Annals of Statistics*, 2009.
- Candes E. and T. Tao, "The Dantzig selector: statistical estimation when p is much larger than n," *Annals of Statistics*, 2007.
- Chen, J., and Z. Chen (2008), "Extended Bayesian Information Criterion for Model Selection with Large Model Spaces," *Biometrika*, 95, 759–771.
- Donald S. and W. Newey, "Series estimation of semilinear models," *Journal of Multivariate Analysis*, 1994.
- Farrell, M. (2015), "Robust inference on average treatment effects with possibly more covariates than observations," *Journal of Econometrics*, 174, 1-23.
- Frank, I. E., J. H. Friedman (1993): "A Statistical View of Some Chemometrics Regression Tools," *Technometrics*, 35(2), 109–135.
- Gautier, E., A. Tsybakov (2011): "High-dimensional Instrumental Variables Regression and Confidence Sets," arXiv:1105.2454v2

- Gilchrist, D. S. and E. G. Sands (2016), "Something to talk about: Social spillovers in movie consumption," Journal of Political Economy, 124(5), 1339-1382.
- Hahn, J. (1998): "On the role of the propensity score in efficient semiparametric estimation of average treatment effects," Econometrica, pp. 315–331.
- Imbens, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," The Review of Economics and Statistics, 86(1), 4–29.
- Jing, B.-Y., Q.-M. Shao, Q. Wang (2003): "Self-normalized Cramer-type large deviations for independent random variables," Ann. Probab., 31(4), 2167–2215.
- D. Kozbur (2016), "Testing-based forward model selection," arXiv:1512.02666
- Leeb, H., and B. M. Potscher (2008): "Can one estimate the unconditional distribution of post-model-selection estimators?" Econometric Theory, 24(2), 338–376.
- Rudelson, M., R. Vershynin (2008): "On sparse reconstruction from Fourier and Gaussian Measurements", Comm Pure Appl Math, 61, 1024-1045.
- Tibshirani, R, "Regression shrinkage and selection via the Lasso," J. Roy. Statist. Soc. Ser. B, 1996.
- H. Wang (2009), "Forward regression for ultra-high dimensional variable screening," Journal of the American Statistical Association, 104, 1512-1524.

Duration of the course: 27 – 30 June (~20 hours, 9:30am - 5pm)

Seminar: 30th June (morning) + some time to talk to participants