Econ 227
Nonparametric and Semiparametric Methods

Xinwei Ma
Department of Economics
University of California San Diego

Winter 2023

Course Description

The primary goal of this course is to introduce modern nonparametric and semiparametric techniques in econometrics. The course contains three parts. In the first part, we provide a rigorous introduction to **kernel methods**, which are routinely used to estimate probability density functions and regression curves. In the second part, we briefly discuss an alternative class of nonparametric estimators, the class of so-called **series estimators**. Finally, in the third part, we examine econometric models that involve both finite and infinite dimensional components. The focus will be on estimation and inference techniques for the finite dimensional component in such **semiparametric models**. We provide a unified framework to analyze asymptotic properties of semiparametric estimators, and then we will study in detail the problem of estimating average treatment effects under selection on observable. If time permits, we will also introduce the semiparametric efficiency theory.

The course will be taught at an advanced level. It is designed as a course for students who plan to choose econometrics as their primary field or secondary field, and for students who want to understand and use sophisticated empirical techniques in their research. **Prerequisites are Econ 220A, B and C**.

Useful Information

The information below reflects the current schedule. Please check the syllabus and Canvas announcements regularly for updates.

Instructor:	Xinwei Ma (x1ma@ucsd.edu).
Class Website:	Canvas.
Lectures:	Monday/Wednesday, 3:30-5:00pm @ Economics Building 300.
Office hours:	Tuesday, 11:00am–12:00pm, or by appointment @ Economics Building 225.

Assessment

The final grade will depend on one **take-home midterm exam** (40%), and one **take-home final exam** (60%). The midterm exam will be posted on February 6, and the final exam will be posted on March 15. You will have **one week to work on each of the exams**.

Midterm and final exams should be submitted electronically through the course website.

Schedule

The following is a tentative schedule of this course, and is likely to change depending on our pace through the quarter. I reserve the right to modify this schedule, and add/replace some topics as needed.

Week 1

- Topics: overview of the course · kernel density estimation · bias and variance expansion · asymptotic normality · bandwidth selection · multi-variate density estimation · density derivative estimation
- References: Li and Racine (2007) · Pagan and Ullah (1999)

Week 2

- Topics: uniform convergence rate · optimal (minimax) rate
- References: Silverman (1978), Stone (1980), Pagan and Ullah (1999), Li and Racine (2007)

Week 3

- Topics: local constant estimation of conditional mean · local polynomial regression · regression discontinuity designs
- References: Fan and Gijbels (1996), Pagan and Ullah (1999), Li and Racine (2007), Calonico, Cattaneo and Titiunik (2014), Cattaneo, Jansson and Ma (2019a)

Week 4

- Topics: characteristic function · classical measurement error · kernel deconvolution
- References: Zhang (1990), Fan (1991), Feller (1991)

Week 5

- Topics: series estimation of conditional mean · pointwise, uniform and L^2 convergence rates · asymptotic normality
- References: Newey (1997), Pagan and Ullah (1999), Chen (2007), Li and Racine (2007), Cattaneo and Farrell (2013), Belloni, Chernozhukov, Chetverikov and Kato (2015)

Week 6

- Topics: introduction to semiparametric problems · examples of semiparametric estimators · average density estimation
- References: Bickel and Ritov (1988), Ritov and Bickel (1990), Newey and McFadden (1994), Pagan and Ullah (1999), Chen, Linton and van Keilegom (2003), Tsiatis (2006), Li and Racine (2007), Cattaneo and Jansson (2019)

Week 7

- Topics: selection on observable · treatment effect · inverse probability weighting · asymptotic expansion and asymptotic normality
- References: Hirano, Imbens and Ridder (2003), Tsiatis (2006), Imbens and Wooldridge (2009), Cattaneo (2010), Abadie and Cattaneo (2018)

Week 8

- Topics: bias fo semiparametric estimators · small bias property · small bandwidth and many covariates asymptotics
- References: Newey, Hsieh and Robins (2004), Cattaneo and Jansson (2018), Chernozhukov, Escanciano, Ichimura and Newey (2018), Cattaneo, Jansson and Ma (2019b)

Week 9

– Topics: efficiency of parametric maximum likelihood estimators · introduction to semiparametric efficiency theory
– References: Newey (1990), Bickel, Klaassen, Bickel, Ritov and Wellner (1998), van der Vaart (2000), Tsiatis (2006), Kosorok (2008)

Week 10

– Topics: parametric submodel · tangent space · regular estimator · irregular estimator · efficient influence function and efficiency bound · efficiency bound of the average treatment effect
– References: Newey (1990), Bickel, Klaassen, Bickel, Ritov and Wellner (1998), Hahn (1998), van der Vaart (2000), Hirano, Imbens and Ridder (2003), Tsiatis (2006), Kosorok (2008), Cattaneo (2010)

References

Abadie, A. and Cattaneo, M. D. (2018). "Econometric Methods for Program Evaluation," *Annual Review of Economics*, *10*, 465–503.

Belloni, A., Chernozhukov, V., Chetverikov, D., and Kato, K. (2015). "Some New Asymptotic Theory for Least Squares Series: Pointwise and Uniform Results," *Journal of Econometrics*, *186*(2), 345–366.

Bickel, P. J., Klaassen, C. A., Bickel, P. J., Ritov, Y., and Wellner, J. A. (1998). *Efficient and Adaptive Estimation for Semiparametric Models*, New York: Springer.

Bickel, P. J. and Ritov, Y. (1988). "Estimating Integrated Squared Density Derivatives: Sharp Best Order of Convergence Estimates," *Sankhyā: The Indian Journal of Statistics, Series A*, *50*, 381–393.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, *82*(6), 2295–2326.

Cattaneo, M. D. (2010). "Efficient Semiparametric Estimation of Multi-valued Treatment Effects Under Ignorability," *Journal of Econometrics*, *155*(2), 138–154.

Cattaneo, M. D. and Farrell, M. H. (2013). "Optimal convergence rates, Bahadur representation, and asymptotic normality of partitioning estimators," *Journal of Econometrics*, *174*(2), 127–143.

Cattaneo, M. D. and Jansson, M. (2018). "Kernel-Based Semiparametric Estimators: Small Bandwidth Asymptotics and Bootstrap Consistency," *Econometrica*, *86*(3), 955–995.

Cattaneo, M. D. and Jansson, M. (2019). "Average Density Estimators: E¢ ciency and Bootstrap Consistency," *Working paper*.

Cattaneo, M. D., Jansson, M., and Ma, X. (2019a). "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association, forthcoming*.

Cattaneo, M. D., Jansson, M., and Ma, X. (2019b). "Two-step Estimation and Inference with Possibly Many Included Covariates," *Review of Economic Studies*, *86*(3), 1095–1122.

Chen, X. (2007). "Large Sample Sieve Estimation of Semi-nonparametric Models," In J. J. Heckman and E. Leamer (eds.) *Handbook of Econometrics, Volume VI*, New York: Elsevier Science B.V. 5549–5632.

Chen, X., Linton, O., and van Keilegom, I. (2003). "Estimation of Semiparametric Models When the Criterion Function is Not Smooth," *Econometrica*, *71*(5), 1591–1608.

Chernozhukov, V., Escanciano, J. C., Ichimura, H., and Newey, J. M., Whitney K. Robins (2018). "Locally Robust Semiparametric Estimation," *Working paper*.

Fan, J. (1991). "On the Optimal Rates of Convergence for Nonparametric Deconvolution Problems," *The Annals of Statistics*, *19*(3), 1257–1272.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*, New York: Chapman and Hall.

Feller, W. (1991). *An Introduction to Probability Theory and Its Applications, Volume II*, New York: John Wiley, 2nd edition.

Hahn, J. (1998). "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, *66*(2), 315–331.

Hirano, K., Imbens, G. W., and Ridder, G. (2003). "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, *71*(4), 1161–1189.

Imbens, G. W. and Wooldridge, J. M. (2009). "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, *47*(1), 5–86.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*, New York: Springer.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*, Princeton: Princeton University Press.

Newey, W. K. (1990). "Semiparametric Efficiency Bounds," *Journal of Applied Econometrics*, *5*(2), 99–135.

Newey, W. K. (1997). "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, *79*(1), 147–168.

Newey, W. K., Hsieh, F., and Robins, J. M. (2004). "Twicing Kernels and a Small Bias Property of Semiparametric Estimators," *Econometrica*, *72*(3), 947–962.

Newey, W. K. and McFadden, D. L. (1994). "Large Sample Estimation and Hypothesis Testing," In R. F. Engle and D. L. McFadden (eds.) *Handbook of Econometrics, Volume IV*: Elsevier Science B.V. 2111–2245.

Pagan, A. and Ullah, A. (1999). *Nonparametric Econometrics*, New York: Cambridge University Press.

Ritov, Y. and Bickel, P. J. (1990). "Achieving Information Bounds in Non and Semiparametric Models," *Annals of Statistics*, *18*(2), 925–938.

Silverman, B. W. (1978). "Weak and Strong Uniform Consistency of the Kernel Estimate of a Density and its Derivatives," *Annals of Statistics*, *6*(1), 177–184.

Stone, C. J. (1980). "Optimal Rates of Convergence for Nonparametric Estimators," *Annals of Statistics*, *8*(6), 1348–1360.

Tsiatis, A. A. (2006). *Semiparametric Theory and Missing Data*, New York: Springer.

van der Vaart, A. W. (2000). *Asymptotic Statistics*, New York: Cambridge University Press.

Zhang, C.-H. (1990). "Fourier Methods for Estimating Mixing Densities and Distributions," *The Annals of Statistics*, *18*(2), 806–831.