A Course on Advanced Econometrics

Yongmiao Hong

The Ernest S. Liu Professor of Economics & International Studies
Cornell University

Course Introduction: Modern economies are full of uncertainties and risk. Economics studies resource allocations in an uncertain market environment. As a generally applicable quantitative analytic tool for uncertain events, probability and statistics have been playing an important role in economic research. Econometrics is statistical analysis of economic and financial data. It has become an integral part of training in modern economics and business. This course develops a coherent set of econometric theory and methods for economic models. It is offered as an advanced econometrics course for doctoral students in economics, business, management, statistics, applied mathematics, and related fields.

The course materials are organized in a coherent manner. Chapter 1 is a general introduction to econometrics. It first describes the two most important features of modern economics, namely mathematical modeling and empirical validation, and then discusses the role of econometrics as a methodology in empirical studies. A few motivating economic examples are given to illustrate how econometrics can be used in empirical studies. Finally, it points out the limitations of econometrics and economics due to the fact that an economy is not a repeatedly controlled experiment. Assumptions and careful interpretations are needed when conducting empirical studies in economics and finance.

Chapter 2 introduces a general regression analysis. Regression analysis is modeling, estimation, inference, and specification analysis of the conditional mean of economic variables of interest given a set of explanatory variables. It is most widely applied in economics. Among other things, this chapter interprets the mean squared error and its optimizer, which lays down the probability-theoretic foundation for least squares estimation. In particular, it provides an interpretation for the least squares estimator and its relationship with the true parameter value of a correctly specified regression model.

Chapter 3 introduces the classical linear regression analysis. A set of classical assumptions are given and discussed, and conventional statistical procedures for estimation, inference, and hypothesis testing are introduced. The roles of conditional homoskedasticity, serial uncorrelatedness, and normality of the disturbance of a linear regression model are analyzed in a finite sample econometric theory. We also discuss the generalized least squares estimation as an efficient estimation method of a linear regression model when the variance-covariance matrix is known up to a constant. In particular, the generalized least squares estimation is embedded as an ordinary least squares estimation of a suitably transformed regression model via conditional variance scaling and autocorrelation filtering.

The subsequent chapters 4–7 are the generalizations of classical linear regression analysis when various classical assumptions fail. Chapter 4 first relaxes the normality and conditional homoskedasticity assumptions, two key conditions assumed in the classical linear regression modeling. A large sample theoretic approach is taken. For simplicity, it is assumed that the observed data are generated from an independent and identically distributed random sample. It is shown that while the finite distributional theory is no longer valid, the classical statistical procedures are still approximately applicable when the sample size is large, provided conditional homoskedasticity holds. In contrast, if the data display conditional heteroskedasticity, classical statistical procedures are not applicable even for large samples, and heteroskedasticity-robust procedures will be called for. Tests for existence of conditional heteroskedasticity in a linear regression framework are introduced.

Chapter 5 extends the linear regression theory to time series data. First, it introduces a variety of basic concepts in time series analysis. Then it shows that the large sample theory for i.i.d. random samples carries over to stationary ergodic time series data if the regression error follows a martingale difference sequence. We introduce tests for serial correlation, and tests for conditional heteroskedasticity and autoregressive conditional heteroskedasticity in a time series regression framework. We also discuss the impact of autoregressive conditional heteroskedasticity on inferences of static time series regressions and dynamic time series regressions.

Chapter 6 extends the large sample theory to a very general case where there exist conditional heteroskedasticity and autocorrelation. In this case, the classical regression theory cannot be used, and a long-run variance-covariance matrix estimator is called for to validate statistical inferences in a time series regression framework.

Chapter 7 is the instrumental variable estimation for linear regression models, where the regression error is correlated with the regressors. This can arise due to measurement errors, simultaneous equation biases, and other various reasons. Two- stage least squares estimation method and related statistical inference procedures are fully exploited. We describe tests for endogeneity.

Chapter 8 introduces the generalized method of moments, which is a popular estimation method for possibly nonlinear econometric models characterized as a set of moment conditions. Indeed, most economic theories, such as rational expectations, can be formulated by a moment condition. The generalized method of moments is particularly suitable to estimate model parameters contained in the moment conditions for which the conditional distribution is usually not available.

Chapter 9 introduces the maximum likelihood estimation and the quasi-maximum likelihood estimation methods for conditional probability models and other nonlinear econometric models. We exploit the important implications of correctly specification of a conditional distribution model, especially the analogy between the martingale difference sequence property of the score function and serial uncorrelatedness, and the analogy between the conditional information equality and conditional homoskedasticity. These links can provide a great help in understanding the large sample properties of the maximum likelihood estimator and the quasi-maximum likelihood estimator.

Chapter 10 concludes the book by summarizing the main econometric theory and methods covered in this book, and pointing out directions for further build-up in econometrics.

This course has several important features. It covers, in a progressive manner, various econometrics models and related methods from conditional means to possibly nonlinear conditional moments to the entire conditional distributions, and this is achieved in a unified and coherent framework. We also provide a brief review of asymptotic analytic tools and show how they are used to develop the econometric theory in each chapter. By going through this book progressively, readers will learn how to do asymptotic analysis for econometric models. Such skills are useful not only for those students who intend to work on theoretical econometrics, but also for those who intend to work on applied subjects in economics because with such analytic skills, readers will be able to understand more specialized or more advanced econometrics textbooks.

Textbooks: (1) *Lecture Notes on Advanced Econometrics*; (2) *Advanced Econometrics*, China Higher Education Publisher.

Requirements: The course grade will be based on homework (10%), empirical project (30%), midterm (30%), and final (30%).

Tables of Contents