Lectures on principles, ideas and theory in econometric time series:

With examples from cointegration, bootstrap, ARCH, state space and big data models.

Lectures by:
Søren Johansen and Anders Rahbek, University of Copenhagen

February 26, 2019

Course outline

The course will be in two main parts: The first part discusses econometric methods and theory, which are applied in the second part, where selected topics from cointegration, state space models, the bootstrap and multivariate ARCH models, as well as big data modelling will be discussed in detail from recent research.

In part I, we give an introduction, aimed for graduate/Ph.D. level students in econometrics, to *(i)* asymptotic theory for stationary, i.i.d. as well as non-stationary (integrated of order one) variables; *(ii)* theory for the bootstrap; *(iii)* theory for cointegration and for (multivariate) ARCH models; and, *(iv)* theory for the Kalman filter. All theory presented will be in terms of examples where details are explained, rather than providing a general introduction to the field(s).

In part II, we discuss recent research with reference to the theory and methodology introduced in Part I. The topics include:

(i) *Cointegration and adjustment in a common trends causal model and the role of weak exogeneity.*

(ii) *Optimal hedging and cointegration in the presence of heteroscedastic errors.*

(iii) *Bootstrap based inference in stationary and non-stationary (conditionally heteroscedastic) autoregressive models.*

(iv) *Models, Methods and Big Data*

Part I | Introduction to the theory of:

Asymptotic theory for i.i.d., stationary and non-stationary univariate variables

We consider some simple statistical models and discuss a general methodology for conducting likelihood inference.

Example 1. The univariate AR model,

$$x_t = \varrho x_{t-1} + \varepsilon_t$$

for $\varrho = 1$ and $|\varrho| < 1$.

Example 2. The univariate ARCH model,

$$x_t = \sqrt{1 + \varrho x_{t-1}^2} z_t$$

for ϱ such that x_t is (non-)stationary.

Example 3. The common trends model for observation y_t and the unobserved state variable α_t, is given by

$$y_t = \beta\alpha_{t-1} + \varepsilon_t,$$
$$\alpha_t = \varrho\alpha_{t-1} + \eta_t.$$

Literature:

Jensen, S.T. and A. Rahbek (2004), Asymptotic Inference for Nonstationary GARCH, *Econometric Theory,* 20:1203–1226.

Johansen, S. and A. Rahbek (2019) *Lecture notes, unpublished.*

Kristensen, D. and A. Rahbek (2005) Asymptotics of the QMLE for a Class of ARCH(q) Models, *Econometric Theory,* 21:946–961.

Kristensen, D. and A. Rahbek (2010), Likelihood-based Inference for Cointegration with Nonlinear Error-Correction, *Journal of Econometrics,* 158:78–94.

Theory for the bootstrap

We give an introduction to the bootstrap and illustrate with the AR and ARCH models from above.

Example 4. The AR(1) bootstrap,

$$x_t^* = \varrho^* x_{t-1}^* + \varepsilon_t^*$$

where the bootstrap process x_t^* is resampled as a function of the bootstrap parameter, ϱ^*, and the bootstrap innovations, ε_t^*, where in general ϱ^* and ε_t^* are functions of the original data, $x_1, ..., x_T$.

Example 5. The ARCH bootstrap,

$$x_t^* = \sqrt{1 + \varrho^* x_{t-1}^{*2}} z_t^*$$

for ϱ^* and z_t^* functions of the original data $x_1, ..., x_T$.

Literature:

Cavaliere, G. and A. Rahbek (2012), Bootstrap Determination of the Co-Integration Rank in Vector Autoregressive Models, *Econometrica,* 80:1721-1740.

Cavaliere, G., H.B. Nielsen and A. Rahbek (2017), On the Consisteny of the Bootstrap Testing for a Parameter on the Boundary of the Parameter Space, *Journal of Time Series Analysis,* 38:513-534.

Theory for the CVAR and multivariate ARCH

We consider again some examples where here ϱ is a $(p \times p)$-dimensional matrix.

Example 6. The cointegrated vector autoregressive model (CVAR) model for multivariate cointegration,

$$x_t = \varrho x_{t-1} + \varepsilon_t, \quad \varrho = \alpha\beta'.$$

Example 7. The multivariate autoregressive conditional heteroscedastic (ARCH) model,

$$x_t = \Omega_t z_t, \ \Omega_t = I + \varrho x_{t-1} x'_{t-1} \varrho.$$

Literature:

Johansen and Rahbek, (2019) *Lecture notes, unpublished.*

Probabilistic and statistical analysis of the common trends model

The multivariate common trends model for observation $x_t \in \mathbb{R}^p$ and unobserved state variable $\alpha_t \in \mathbb{R}^m$ is given by

$$\begin{aligned} x_t &= \beta\alpha_{t-1} + \varepsilon_t, \\ \alpha_t &= \varrho\alpha_{t-1} + \eta_t. \end{aligned}$$

The lecture will discuss identification of the parameters, and simple inference for β based on a regression estimator for β.

The Gaussian likelihood can be calculated using the Kalman filter, and we discuss the prediction error formulation of the model, and the diffuse and conditional likelihood.

Based on this, we discuss existence, consistency and asymptotic distribution of the maximum likelihood estimator, using score and information.

Literature:

Johansen, S. (2018) Inference in a simple nonstationary state space model. Unpublished

Chang, Y., J. I. Miller, and J. Y. Park (2009) Extracting a common stochastic trend: Theory with some applications. *Journal of Econometrics*, 150, 231–247.

Part II | Research Topics

Cointegration and adjustment in a common trends causal model and the role of weak exogeneity.

The lectures will contain a discussion of causal model for stationary variable and a new causal model for nonstationary variables.

A simple CVAR(1) model for some observed variables, x_t, and some unobserved variables, τ_t, is defined and the question of weak exogeneity in the derived model for the observations is discussed.

The techniques used in the discussion are: Unobserved components models and their CVAR(∞) representation. The Kalman filter technique for deriving a random walk representation of the conditional mean of the unobserved component, $E(\tau_t|x_0, \ldots, x_t)$ and some results from control theory are used to show the existence of the limiting conditional variance of the unobserved component as the solution of a matrix Riccatti equation. A few examples will be used for illustration.

Literature:

Johansen, S. (2019) Cointegration and Adjustment in the infinite order CVAR representation of some partially observed CVAR(1) models, *Econometrics,* 7:2.

Optimal hedging and cointegration in the presence of heteroscedastic errors

The role of cointegration is analysed for optimal hedging of an h-period portfolio. Prices are assumed to be generated by a cointegrated vector autoregressive model allowing for stationary martingale errors, satisfying a mixing condition and hence some heteroscedasticity. The risk of a portfolio is measured by the conditional variance of the h-period return given information at time t. If the price of an asset is nonstationary, the risk of keeping the asset for h periods diverges for large h. The $h-$period minimum variance hedging portfolio is derived, and it is shown that it approaches a cointegrating vector for large h, thereby giving a bounded risk. Taking the expected return into account, the portfolio that maximizes the Sharpe ratio is found, and it is shown that it also approaches a cointegration portfolio.

For constant conditional volatility, the conditional variance can be estimated, using regression methods or the reduced rank regression method of cointegration. In case of conditional heteroscedasticity, however, only the expected conditional variance can be estimated without modelling the heteroscedasticity. The findings are illustrated with a data set of prices of two year forward contracts for electricity, which are hedged by forward contracts for fuel prices. The main conclusion of the paper is that for optimal hedging, one should exploit the cointegrating properties for long horizons, but for short horizons more weight should be put on the remaining dynamics.

Literature:

Gatarek, L. and Johansen, S. (2019) The role of cointegration for optimal hedging with heteroscedastic error term. Unpublished.

Bootstrap based inference in stationary and non-stationary (conditionally heteroscedastic) autoregressive models: Hybrid and shrinking bootstrap.

In this lecture we discuss the general application of the bootstrap for statistical inference in econometric time series models.

We do this by considering in detail the implementation of bootstrap inference in the class of double-autoregressive [DAR] models as well as ARCH models.

DAR models are particularly interesting to illustrate implementation of the bootstrap to time series: first, standard asymptotic inference is usually difficult to implement due to the presence of nuisance parameters under the null hypothesis; second, inference involves testing whether one or more parameters are on the boundary of the parameter space; third, under the alternative hypothesis, fourth or even second order moments may not exist. In most of these cases, the bootstrap is not considered an appropriate tool for inference. Conversely, and taking testing (non-) stationarity to illustrate, we show that although a standard bootstrap based on unrestricted parameter estimation is invalid, a correct implementation of a bootstrap based on *restricted* parameter estimation (restricted bootstrap) is first-order valid; that is, it is able to replicate, under the null hypothesis, the correct limiting null distribution. Importantly, we also show that the behaviour of this bootstrap under the alternative hypothesis may be different because of possible lack of finite second-order moments of the bootstrap innovations. This features makes – for some parameter configurations – the restricted bootstrap unable to replicate the null asymptotic distribution when the null is false. We will see that this drawback can be fixed by using a new 'hybrid' bootstrap, where the parameter estimates used to construct the bootstrap data are obtained with the null imposed, while the bootstrap innovations are sampled with replacement from the unrestricted residuals. We will discuss that this bootstrap, novel in this framework, mimics the correct asymptotic null distribution, respectively of the null to be true or false. Throughout, we use a number of examples from the bootstrap time series literature to illustrate the importance of properly defining and analyzing the bootstrap generating process and associated bootstrap statistics.

Literature:

Cavaliere and Rahbek, *Econometric Theory Lecture 2019*, *Lecce*, *Italy*, unpublished.

Model, methods and big data

"All models are wrong but some are useful".(George Box)

"All models are wrong, and increasingly you can succeed without them." (Peter Norvig, Google's research director)

The lecture is about methods and models. "Methods" means algorithms, and "Models" we know about. What is the interplay between the two? Are models obsolete? What is the role of models in our work? Do models come before methods or the other way around. Does the Big Data revolution try to solve the same problems as before, or do they attack new problems, that we could not even dream of?

To me a model is a way of expressing my understanding of what goes on, so we can communicate with others and construct thought experiments and real experiments, that can further our understanding of what is going on. The lecture will illustrate with a few historical examples of the interplay between models and methods, and will give a brief introduction to some new (model based) results on cointegration and big data.

Literature:

Onatski, A. and C. Wang. (2018) Alternative asymptotics for in cointegration tests in large VARs, *Econometrica*, Vol. 86, No. 4, 1465–1478.

Chris Anderson (2008) The end of theory: The data deluge makes the scientific method obsolete. *Wired Magazine.*

Chang, Y., C. Kim, and J. Park (2016). Nonstationarity in time series of state densities. *Journal of Econometrics,* 192:152–167.

Beare, B. and W. Seo (2018). Representation of I(1) and I(2) autoregressive Hilbertian processes. In press.

Franchi, M. and Paruolo, P. (2018) Cointegration in functional autoregressive processes. In press.