Time series analysis: Inference and prediction
Analyse des séries temporelles : inférence et prédiction
(Org: **Masoud Nasari** (Bank of Canada and Carleton University) and/et **Mohamedou Ould-Haye** (Carleton University))

S. EJAZ AHMED, Brock U
Implicit BIAS Problems in in High Dimensional Predictive Models: "Cut the Bias"

The rapid growth in the size and scope of data sets in a variety of disciplines have naturally led to the usage of the term, Big Data. The analysis of such data is important in multiple research fields such as digital marketing, gene expression arrays, social network modeling, clinical, genetics and phenotypic data, bioinformatics, personalized medicine, environmental, neuroscience, astronomy, nanoscience, among others. In high-dimensional models where number of predicting variables is greater than observations, many penalized regularization strategies were studied for simultaneous submodel selection and post-estimation. Generally speaking, submodel are subject to inherited bias, and the prediction based on a selected submodel may not be preferable. For this reason, we propose a high-dimensional shrinkage strategy to improve the prediction performance of a submodel. Such a high-dimensional shrinkage estimator (HDSE) is constructed by shrinking a overfitted model estimator in the direction of a candidate submodel. We demonstrate that the proposed HDSE performs uniformly better than the overfitted model estimator. Interestingly, it improves the prediction performance of a given candidate submodel. The relative performance of the proposed HDSE strategy is appraised by both simulation studies and the real data analysis.

SUMANTA BASU, Cornell Uniersity
Frequency-domain graphical models for multivariate time series

Graphical models offer a powerful framework to capture intertemporal and contemporaneous relationships among the components of a multivariate time series. For stationary time series, these relationships are encoded in the multivariate spectral density matrix and its inverse. We will present adaptive thresholding and penalization methods for estimation of these objects under suitable sparsity assumptions. We will discuss new optimization algorithms and investigate consistency of estimation under a double-asymptotic regime where the dimension of the time series increases with sample size. If time permits, we will introduce a frequency-domain graphical modeling framework for multivariate nonstationary time series that captures a new property called conditional stationarity.

YOUSSOUPH CISSOKHO, University of Ottawa
Estimation of cluster functionals for regularly varying time series: sliding blocks estimators.

Cluster indices describe extremal behaviour of stationary time series. We consider their sliding blocks estimators. Using a modern theory of multivariate, regularly varying time series, we obtain central limit theorems under conditions that can be easily verified for a large class of models. In particular, we show that in the Peaks-Over-Threshold framework, sliding and disjoint blocks estimators have the same limiting variance.

MARIE-CHRISTINE DUKER, Cornell Univeristy
Detecting fractal connectivity in high-dimensional time series

The long-run variance matrix and its inverse, the so-called precision matrix, give respectively information about correlations and partial correlations between dependent component series of multivariate time series around zero frequency. This talk will present non-asymptotic theory for estimation of the long-run variance and precision matrices for high-dimensional time series under general assumptions on the dependence structure including long-range dependence. The presented results for thresholding and penalizing versions of the classical local Whittle estimator ensure consistent estimation in a possibly high-dimensional regime. The highlight of this talk is a concentration inequality of the local Whittle estimator for the long-run variance matrix around the true model parameters. In particular, it handles simultaneously the estimation of the memory parameters which enter

the underlying model. Finally, we study the temporal and spatial dependence of multiple realized volatilities of global stock indices.

BOUCHRA NASRI, Université de Montréal
Test of serial dependence for multivariate time series with arbitrary distributions

Tests of serial independence are presented for a fixed number of consecutive observations from a stationary time series, first in the univariate case, and then in the multivariate case, where even vectors of large dimensions can be used. The common distribution function of the time series is not assumed to be continuous, and the tests statistics are based on the multilinear copula process. A case study using a time series of images is used to illustrate the usefulness of the methodologies presented.

MOHAMEDOU OULD-HAYE, Carleton University
Spectral analysis of Time series: Random sampling and Stationarity test

We investigate some aspects of time series data spectral analysis. We consider random sampling of continuous processes and stationarity tests.

BRUNO RÉMILLARD, HEC Montreal
Change-point problems for multivariate time series using pseudo-observations

In this talk, it is shown that under weak assumptions, the change-point tests designed for independent random vectors can also be used with pseudo-observations for testing change-point in the joint distribution of non-observable random vectors, the associated copula, or the margins, without modifying the limiting distributions. In particular, change-point tests can be applied to the residuals of stochastic volatility models or conditional distribution functions applied to the observations, which are prime examples of pseudo-observations. Since the limiting distribution of test statistics depends on the unknown joint distribution function or its associated unknown copula when the dimension is greater than one, we also show that iid multipliers and traditional bootstrap can be used with pseudo-observations to approximate P-values for the test statistics. Numerical experiments are performed in order to compare the different statistics and bootstrapping methods. Examples of applications to change-point problems are given. This is joint work with Bouchra R. Nasri and Tarik Bahraoui.

IDRISS SEKKAK, École de santé publique - Université de Montréal
The stochastic epidemic modelling: The influence of incidence rates and perturbations.

Epidemic models provide an insight on how to react to an epidemic outbreak. For this matter, we investigate several aspects of stochastic dynamical systems according to different incidence rates and perturbations. We carry out a thorough analysis to show the existence of the global and positive solutions. We explore the extinction and the persistence of the disease regarding a derived stochastic threshold of the model. Moreover, we use suitable Lyapunov functions in order to explore the impact of the perturbation on the stability around the equilibrium points. Finally, we give some numerical illustrations to support our analytical results.

MAMADOU Y. THIOUB, HEC Montreal
Goodness-of-fit tests and robust regime selection procedure for general hidden Markov models

This work presents powerful goodness-of-fit procedures for general Markov regime-switching models with covariates when the outcomes are continuous, discrete, or zero-inflated. The EM algorithm is used for the estimation method and a randomized Rosenblatt's transform is applied to obtain formal goodness-of-fit tests. The latter then served for selecting the number of regimes. Numerical experiments are used to assess the finite sample performance of the proposed methodologies and to compare with other criteria for the selection of models, including Bayesian methods. Finally, the proposed methodologies are implemented in an R package available on CRAN.

RICARDAS ZITIKIS, Western University

Detecting systematic anomalies when inputs are stationary time series

We shall present an anomaly-detection method when systematic anomalies, possibly statistically very similar to genuine inputs, are affecting control systems. The method allows anomaly-free inputs to originate from a wide class of random sequences. To illustrate how the method works on data, we shall provide a controlled experiment with anomaly-free inputs following an ARMA time series model under various contamination scenarios.