NSYSU-Waseda International Symposium

Time Series, Machine Learning and Causality Analysis

Date: 2-3 September 2019
Venue: International Conference Room SC 1005,
College of Science Department of Applied Mathematics,
National Sun Yat-sen University, Kaohsiung.

This workshop is supported by:
Japan Scoiety for the Promotion of Science Kiban (S) No.18 H05290 (Prof.Taniguchi, M),
Waseda Research Institute for Science & Engineering, Institute for Mathematical Science,
and The Institute of Marine Research
National Center for Theoretical Sciences

Program

September 2

09:50 – 10:00 Meihui Guo (National Sun Yat-sen University), Opening

Session (I): 10:00~12:10 chaired by Ching-Kang Ing

10:00 – 10:50 Masanobu Taniguchi (Waseda University)
Time Series Analysis under Non-Standard Settings

10:50 – 11:30 Shih-Feng Huang (National University of Kaohsiung)
Stock Market Trend Prediction Using Functional Time Series Approach

11:30 – 12:10 Hai-Tang Chiou (National Tsing Hua University)
Variable Selection for High-Dimensional Regression Models with Time Series and Heteroscedastic Errors

12:10 – 13:30 Lunch break

Session (II): 13:30~15:00 chaired by Akitoshi Kimura

13:30 – 14:20 Takayuki Shiohama (Tokyo University of Science)
Risk-Based Asset Allocation Strategies for Factor Investing

14:20 – 15:00 ShengLi Tzeng (National Sun Yat-sen University)
Parsimonious Imaging Biomarkers for Lung Tumor Classification

15:00 – 15:30 Coffee break

Session (III): 15:30~17:00 chaired by Mong-Na Lo Huang

15:30 – 16:10 Hsiang-Ling Hsu (National University of Kaohsiung)
On model selection from a finite family of possibly misspecified time series models

16:10 – 17:00 Kou Fujimori (Waseda University)
The Dantzig Selector for Statistical Models of Stochastic Processes in High Dimensional and Sparse Settings

18:00 – Dinner

September 3

Session (IV): 10:00~17:00 Free Discussion

Abstract

Taniguchi, Masanobu Waseda University, Japan

Title: Joint circular distributions in view of higher order spectra of time series

Abstract: This talk delivers a series of recent developments for time series analysis under non-standard settings. Concretely the following subjects are addressed:

(i) Non-regular estimation for discontinuous spectra.
(ii) Jackknifed Whittle estimation.
(iii) Higher order asymptotic theory for semi-parametric spectral estimation.
(iv) Asymptotics of realized volatility with ARCH(∞) microstructure noise.
(v) Model selection for contiguously specified spectral family.

The results show unusual aspects of the asymptotic theory, and open a new horizon for time series.

Huang, Shih-Feng National University of Kaohsiung, Taiwan

Title: Stock Market Trend Prediction Using Functional Time Series Approach

Abstract: Thanks to advanced technologies, ultra-high frequency limit order book (LOB) data are now available to data analysts. An LOB contains comprehensive information on all transactions in a market. We use LOB data to investigate the high frequency dynamics of market supply and demand (S-D) and inspect their impacts on intra-daily market trends. The intra-daily S-D curves are fitted with B-spline basis functions. Technique of multiresolution is introduced to capture inhomogeneous curvature of the S-D curves and a lasso-type criterion is employed to select a common basis set. Based on empirical evidence, we model the time varying coefficients in the B-spline interpolation by vector autoregressive models of order p ($\geqq 1$). The Xgboost algorithm is employed to extract information from the areas under the S-D curves to predict the intra-daily market trends. In the empirical study, we analyze the LOB data from LOBSTER (https://lobsterdata.com/). The results show that the proposed approach is able to recover the S-D curves and has satisfactory performance on both curve and market trend predictions.

Chiou, Hai-Tang National Tsing Hua University, Taiwan
(co-author: Ching-Kang Ing, Meihui Guo**)**

Title: Variable Selection for High-Dimensional Regression Models with Time Series and Heteroscedastic Errors

Abstract: Although existing literature on high-dimensional regression models is rich, the vast majority of studies have focused on independent and homogeneous error terms. In this article, we consider the problem of selecting high-dimensional regression models with heteroscedastic and time series errors, which have broad applications in economics, quantitative finance, environmental science, and many other fields. The error term in our model is not only allowed to be short- or long-range dependent, but also contains a high-dimensional dispersion function accounting for heteroscedasticity. By making use of the orthogonal greedy algorithm and the high-dimensional information criterion, we propose a new model selection procedure that can consistently choose the relevant variables in both the regression and the dispersion functions. The finite sample performance of the proposed procedure is also illustrated via simulations and real data analysis.

Shiohama, Takayuki **Tokyo University of Science, Japan**
Title: Risk-Based Asset Allocation Strategies for Factor Investing
Abstract: Factor investing or smart beta strategy has become an established investment style among institutional and individual investors. Since factor returns have been cyclical, and are weakly correlated, the diversification benefit can be obtained by investing in multiple factors. As a result, investors are increasingly turning their attention from single-factor investing to multi-factor strategies, with growing interest in and discussion on how to construct multi-factor portfolios. In this study, we contribute to the existing risk-based asset allocation approaches for factor investing strategies by constructing a multi-factor portfolio based on the inverse weighting method. We also investigate the effectiveness of the factor risk parity (FRP) strategy by constructing active equity FRP portfolios for global stock markets. The results obtained reveal that the IFV portfolios as well as the FRP strategy outperformed market capitalization weighted portfolios significantly by acquiring factor risk premiums successfully.

Tzeng, ShengLi **National Sun Yat-sen University, Taiwan**
Title: Parsimonious Imaging Biomarkers for Lung Tumor Classification
Abstract: Objective: Most existing quantitative imaging biomarkers (QIBs) in medical image classification describing local spatial behaviors based on rectangular images. This study develops a novel method to extract QIBs accounting for more global spatial characteristics of imaging, which can also be applied directly to non-rectangular regions of interest. If there exists strong large-scale spatial structure, our proposal is expected to extract much fewer QIBs. These data-driven may lead to a better understanding of the spatial patterns they capture.
Methods: A spatial imputation method utilizes image intensities within each region of interest from patients. It represents image intensities as linear combinations of basis functions. The algorithm learns the spatial covariance structure for intensities and the spatial covariance can be decomposed into several orthogonal basis functions. The decomposition is analogous to popular principal component analysis. We then use the "principal component scores" as data-driven QIBs.
Results: We applied the proposed method to a study of solitary pulmonary nodules from 85 patients' [F-18]FDG dual time point PET images (DTPI). Each patient was scanned twice for first and delayed imaging, and subsequently classified as benign or malignant with confirmation (22 benign and 63 malignant). We compared the classification performances of support vector machines using (1) the proposed QIBs to (2) that using 120 common QIBs. We found our proposal describing large spatial characteristics. Only 4 proposed QIBs were extracted from the DTPI data. Out-of-sample AUCs in cross validation for the two feature sets were (1) 0.820 and (2) 0.838, while using the proposed QIBs had a higher specificity.
Conclusion: The few proposed QIBs performed almost as well as a far larger number of common QIBs, which implies the parsimonious explanatory ability of this novel data-driven method.

Hsu, Hsiang-Ling National University of Kaohsiung, Taiwan
(co-author: Ching-Kang Ing, Howell Tong**)**
Title: On model selection from a finite family of possibly misspecified time series models
Abstract: Consider finite parametric time series models. `I have *n* observations and *k* models, which model should I choose on the basis of the data alone?' is a frequently asked question in many practical situations. This poses the key problem of selecting a model from a collection of candidate models, none of which is necessarily the true data generating process (DGP). Although existing literature on model selection is vast, there is a serious lacuna in that the above problem does not seem to have received much attention. In fact, existing model selection criteria have avoided addressing the above problem directly, either by assuming that the true DGP is included among the candidate models and aiming at choosing this DGP, or by assuming that the true DGP can be asymptotically approximated by an increasing sequence of candidate models and aiming at choosing the candidate having the best predictive capability in some asymptotic sense. In this article, we propose a misspecification-resistant information criterion (MRIC) to address the key problem directly. We first prove the asymptotic efficiency of MRIC whether the true DGP is among the candidates or not, within the fixed-dimensional framework. We then extend this result to the high-dimensional case in which the number of candidate variables is much larger than the sample size. In particular, we show that MRIC can be used in conjunction with a high-dimensional model selection method to select the (asymptotically) best predictive model across several high-dimensional misspecified time series models.

Fujimori, Kou Waseda University, Japan
Title: The Dantzig selector for statistical models of stochastic processes in high dimensional and sparse settings
Abstract: The Dantzig selector, which was proposed by Cand¥'es and Tao in 2007, is an estimation procedure for regression models in high-dimensional and sparse settings. In this talk, the Dantzig selectors for some statistical models of stochastic processes will be discussed. We apply this procedure to Cox's proportional hazards model and some specific models of diffusion processes and prove the consistencies and the variable selection consistencies of the estimators which enables us to reduce the dimension.

Speakers

Fujimori, Kou, Waseda University, Japan kfujimori@aoni.waseda.jp

Huang, Shih-Feng, National University of Kaohsiung, Taiwan huangsf@nuk.edu.tw

Hsu, Hsiang-Ling, National University of Kaohsiung, Taiwan hsuhl@nuk.edu.tw

Pan, Guangming, Nanyang Technological University, Singapore gmpan@ntu.edu.sg

Shiohama, Takayuki, Tokyo University of Science, Japan shiohama@ms.kagu.tus.ac.jp

Taniguchi, Masanobu, Waseda University, Japan taniguchi@waseda.jp

Tzeng, ShengLi, National Sun Yat-sen University, Taiwan sltzeng@math.nsysu.edu.tw