Waseda International Symposium

Topological Data Science, Causality, Analysis of Variance & Time Series

dedicated to Professor Taniguchi's retirement

Date: March 7 – March 9, 2022

Venue: Conference Room (大会議室), 1st Floor, Bldg. 55N,
Nishi-Waseda Campus, Waseda University
(Access map: https://www.waseda.jp/fsci/en/access/)

Organizers: Yan Liu & Yuichi Goto

Supported by:

- JSPS KAKENHI Kiban (S) Grand-in-Aid No. 18H05290 (M. Taniguchi)
- Waseda Research Institute for Science & Engineering
  Institute for Mathematical Science

Waseda International Symposium

"Topological Data Science, Causality, Analysis of Variance & Time Series Analysis"

dedicated to Professor Taniguchi's retirement

Date: March 7 – March 9, 2022

Venue: Conference Room (大会議室), 1st Floor, Bldg. 55N,
Nishi-Waseda Campus, Waseda Univ.
(Access map: https://www.waseda.jp/fsci/en/access/)

Organizers: Yan Liu & Yuichi Goto

Waseda International Symposium is supported by:

- **JSPS KAKENHI Kiban (S) Grand-in-Aid No. 18H05290 (M. Taniguchi)**
- **Waseda Research Institute for Science & Engineering, Institute for Mathematical Science**

Program

March 7

9:00–9:10: Liu, Yan (Waseda Univ.)
Opening

Session I (9:10–10:30): chaired by Ing, Ching-Kang

9:10–9:30: Chung, Moo K. (Univ. of Wisconsin-Madison) ZOOM
Dynamic topological data analysis on time varying graphs

9:30–9:50: De Oliveira, Victor (Univ. of Texas) ZOOM
On Information About Covariance Parameters in Gaussian Matérn Random Fields

9:50–10:10: Fujimori, Kou (Shinshu Univ.)
Sparse principal component analysis for high-dimensional stationary time series

10:10–10:30: Ombao, Hernando (KAUST) ZOOM
Spectral Non-Linear Granger Causality for Multivariate Time Series

10:30–10:40: Coffee Break

Session II (10:40–11:40): chaired by Xue, Yujie

10:40–11:00: Amano, Tomoyuki (Univ. of Electro-Communications) ZOOM
Analysis of CL and Estimating Function Estimators for Financial Time Series Models

11:00–11:20: Dou, Xiaoling (Waseda Univ.)
A smoothed empirical beta copula

11:20–11:40: Lin, Liang-Ching (National Cheng Kung Univ.) ZOOM
Monitoring Photochemical Pollutants for Anomaly Detection based on Symbolic Interval-Valued Data Analysis

11:40–13:00: Lunch Time

Session III (13:00–14:20): chaired by Shiraishi, Hiroshi

13:00–13:20: Chan, Ngai Hang (Chinese Univ. of Hong-Kong) ZOOM
Optimal Change-Point Estimation in Time Series

13:20–13:40: Chen, Cathy W.S. (Feng Chia Univ.) ZOOM
Bayesian modelling of integer-valued transfer function models

13:40–14:00: Chen, Ying (National Univ. of Singapore) ZOOM
Probabilistic Forecasting for Daily Electricity Loads and Quantiles for Curve-to-Curve Regression

14:00–14:20: Koul, Hira (MSU) ZOOM
Minimum Distance Estimation in Linear Errors-in-Variables Regression Model

14:20–14:40: Coffee Break

Session IV (14:40–15:40): chaired by Shiohama, Takayuki

14:40–15:00: Goto, Yuichi (Waseda Univ.)
Homogeneity tests for one-way models with dependent errors under correlated groups

15:00–15:20: Guo, Mei-Hui (National Sun Yat-sen Univ.) ZOOM
Change point detection of nonlinear heteroscedastic time series models

15:20–15:40: Hirukawa, Junichi (Niigata Univ.)
Innovation algorithm of fractionally integrated ($I(d)$) process and applications on the estimation of parameters

15:40–16:00: Coffee Break

Session V (16:00–17:00): chaired by Nishiyama, Yoichi

16:00–16:20: Dette, Holger (Ruhr-Univ. Bochum) ZOOM
Testing relevant hypotheses in functional time series via self-normalization

16:20–16:40: Negri, Ilia (Univ. of Bergamo) ZOOM
Z-process method for change point problems in time series

16:40–17:00: Luati, Alessandra (Univ. of Bologna) ZOOM
Semiparametric modeling of multiple quantiles

17:00–17:10: Coffee Break

Session VI (17:10–18:30): chaired by Fujimori, Kou

17:10–17:30: Giraitis, Liudas (Queen Mary Univ. of London) ZOOM
Choosing between persistent and stationary volatility

17:30–17:50: Francq, Christian (ENSAE) ZOOM
Inference on multiplicative component GARCH without any small-order moment

17:50–18:10: Patilea, Valentin (ENSAI) ZOOM
On Wilks' Theorem for Conditional Moment Equations with Weakly Dependent Data

18:10–18:30: Preinerstorfer, David (Univ. of St.Gallen) ZOOM
Consistency of p-norm based tests in high dimensions

March 8

Session VII (9:00–10:20): chaired by Chan, Ngai Hang

9:00–9:20: Dallakyan, Aramayis and Pourahmadi, Mohsen (Texas A&M Univ.) ZOOM
Learning Bayesian Networks through Birkhoff Polytope: A Relaxation Method

9:20–9:40: Stoffer, David S (Univ. of Pittsburgh) ZOOM
Efficient Fitting of Stochastic Volatility Models

9:40–10:00: Tsay, Ruey S (Univ. of Chicago Booth School of Business) ZOOM
Divide-and-Conquer: A Distributed Hierarchical Factor Approach to Modeling Large-Scale Time Series Data

10:00–10:20: Wang, Yuan (Univ. of South Carolina) ZOOM
Topological Inference with Spectral Permutation Test

10:20–10:30: Coffee Break

Session VIII (10:30–11:50): chaired by Dou, Xiaoling

10:30–10:50: Huang, Shih-Feng (National Univ. Kauhsiung) ZOOM
A network autoregressive model with GARCH effects and its applications

10:50–11:10: Ing, Ching-Kang (National Tsing Hua Univ.) ZOOM
Model selection for unit-root time series with many predictors

11:10–11:30: Whang, Yoon-Jae (Seoul National Univ.) ZOOM
Testing for Time Stochastic Dominance

11:30–11:50: Nishiyama, Yoichi (Waseda Univ.)
On rank statistics of PageRank and MarkovRank

11:50–12:10: Coffee Break

Art Session I (12:10–12:50): chaired by Ogata, Hiroaki

Kurita, Mari (New England Conservatory SCE, Boston) ZOOM
Serious Opera and Comic Opera in Italy in the 18th century – Musical Styles and their Impact

12:50–14:00: Lunch Time

Session IX (14:00–15:00): chaired by Akashi, Fumiya

14:00–14:20: Ogata, Hiroaki (Tokyo Metropolitan Univ.)
Pair circulas modelling for circular multivariate time series

14:20–14:40: Shiohama, Takayuki (Nanzan Univ.)
Topological data analysis based classification in time series

14:40–15:00: Shiraishi, Hiroshi (Keio Univ.)
Time Series Quantile Regression by using Random Forests

15:00–15:10: Coffee Break

Session X (15:10–16:10): chaired by Xu, Xiaofei

15:10–15:30: Shimizu, Yasutaka (Waseda Univ.)
A quite new approach to cohort-wise mortality prediction under survival energy hypothesis

15:30–15:50: Monti, Anna (Univ. of Sannio) ZOOM
Tango: dance and statistical thinking

15:50–16:10: Hallin, Marc (Université libre de Bruxelles) ZOOM
Fully Distribution-Free Center-Outward Rank Tests for Multiple-Output Regression and MANOVA

16:10–16:20: Coffee Break

Celebration Session (16:20–17:00): chaired by Kakizawa, Yoshihide

Taniguchi, Masanobu (Waseda Univ.)
The Long and Curved Road

17:00–17:20: Coffee Break

Art Session II (17:20–18:00): chaired by Negri, Ilia

Harada, Yuya (Tokyo Univ. of the Arts)

Pianist 🎹: Hirao, Yui (Tokyo Univ. of Arts)

Performance expression and statistical consideration in opera (Classic Music)

18:00–18:15: Taniguchi, Masanobu (Waseda Univ.)
End Credits

March 9

Session XI (10:00–11:00): chaired by Amano, Tomoyuki

10:00–10:20: Xu, Xiaofei (Waseda Univ.)
Long-memory Log-linear Zero-inflated Generalized Poisson Autoregression for COVID-19 Pandemic Modeling

10:20–10:40: Xue, Yujie (Waseda Univ.)
Hellinger distance estimation for non-regular spectra

10:40–11:00: Akashi, Fumiya (Univ. of Tokyo)
Self-weighted GEL method based on spatial median

11:00–11:10: Coffee Break

Session XII (11:10–11:50): chaired by Guo, Mei-Hui

11:10–11:30: Honda, Toshio (Hitotsubashi Univ.) ZOOM
Forward variable selection for ultra-high dimensional quantile regression models

11:30–11:50: Yamashita, Satoshi (ISM) (Video Letter)
Statistical matching method for data profiling and fusion

11:50–14:20: Lunch Time

Session XIII (14:20–15:40): chaired by Chen, Cathy W.S.

14:20–14:40: Matsuda, Yasumasa (Tohoku Univ.)
Functional regression models for spatio-temporal data

14:40–15:00: Minami, Mihoko (Keio Univ.)
Regression Tree and Clustering for Distributions, and Homogeneous Structure of Population Characteristics

15:00–15:20: So, Mike KP (Hong Kong Univ. of Science and Technology) ZOOM
Topological properties of pandemic and financial networks: prediction and risk management

15:20–15:40: Liu, Yan (Waseda Univ.)
Statistical and Topological Inference for Local Granger Causality

15:40–16:00: Coffee Break

Session XIV (16:00–17:00): chaired by Chen, Ying

16:00–16:20: Solvang, Hiroko Kato (Institute of Marine Research) ZOOM
Categorical data analysis to infer directional associations among rorquals and their prey species

16:20–16:40: Verdebout, Thomas (Université libre de Bruxelles) ZOOM
Asymptotic power of Sobolev tests for uniformity on hyperspheres

16:40–17:00: Proietti, Tommaso (Univ. of Rome Tor Vergata) ZOOM
Peaks, Gaps, and Time Reversibility of Economic Time Series

17:00–17:10: Coffee Break

Session XV (17:10–18:10): chaired by Hirukawa, Junichi

17:10–17:30: Zakoian, Jean-Michel (ENSAE) ZOOM
Testing hypotheses on the innovations distribution in semi-parametric conditional volatility models

17:30–17:50: Taylor, Robert (Univ. of Essex) ZOOM
Extensions to IVX Methods of Inference for Return Predictability

17:50–18:10: Ronchetti, Elvezio (Univ. of Geneva) ZOOM
Optimal transportation through saddlepoints

18:10–18:20: Coffee Break

Session XVI (18:20–18:40): chaired by Goto, Yuichi

18:20–18:40: Hayamizu, Momoko (Waseda Univ.)
Treefit: measuring the tree-likeness of point clouds and application to the analysis of single-cell gene expression data

18:40–18:50: Goto, Yuichi (Waseda Univ.)
Closing

Abstract

Chung, Moo K.

Title: Dynamic topological data analysis on time varying graphs

Abstract: Topological data analysis (TDA) has been used in analyzing static graphs and networks. Instead of analyzing graphs at one fixed scale that may not be optimal, TDA builds a collection of nested subgraphs over every possible threshold. Such a multiscale approach often enables to quantify subtle topological differences in networks. However, it is still not clear how TDA can be adapted for time varying graphs. In this talk, we present a novel dynamic-TDA based on the Wasserstein distance on graph filtrations. The method utilizes the scalable nature of graph filtrations in speeding up computation. The method is applied in modeling the state changes of functional brain networks. We show how to solve various regression, inference and clustering problems using the proposed dynamic-TDA.

De Oliveira, Victor

Title: On Information About Covariance Parameters in Gaussian Matérn Random Fields

Abstract: The Matérn family of covariance functions is currently the most commonly used for the analysis of geostatistical data due to its ability to describe different smoothness behaviors. Yet, in many applications the smoothness parameter is set at an arbitrary value. This practice is due to unqualified claims in the literature stating that geostatistical data have little or no information about the smoothness parameter. This work critically investigates this claim and shows it is not true in general. Specifically, it is shown that the information the data have about the correlation parameters varies substantially depending on the true model and sampling design and, in particular, the information about the smoothness parameter can be large, in some cases larger than the information about the range parameter. In light of these findings, we suggest to abandon the aforementioned practice and instead establish inferences from data — based estimates of both range and smoothness parameters, especially for strongly dependent non--smooth process observed on an irregular sampling design. A data set of daily rainfall totals is used to judge the claim as well as to motivate and illustrate the proposed methods.

Fujimori, Kou

Title: Sparse principal component analysis for high-dimensional stationary time series

Abstract: In this talk, we discuss the sparse principal component analysis for high-dimensional stationary processes. The standard principal component analysis performs poorly when the dimension of the process is large. We establish the oracle inequalities for penalized principal component estimators for the processes including heavy-tailed time series. The rate of convergence of the estimators is established. We also elucidate the theoretical rate for choosing the tuning parameter in penalized estimators. The

performance of the sparse principal component analysis is demonstrated by numerical simulations. (Joint work with Yan Liu, Yuichi Goto and Masanobu Taniguchi)

Ombao, Hernando

Title: Spectral Non-Linear Granger Causality for Multivariate Time Series

Abstract: One of the key goals in analyzing multivariate time series is to characterize and estimate the cross-dependence structure among the components. Traditional approaches (e.g., coherence and correlation) capture only linear dependence. This serious limitation could lead to false conclusions under non-linearity. Keeping this as our motivation, we propose a procedure for identifying non-linear and frequency-band specific Granger causality (Spec NLGC) connections. The advantages of the Spec NLGC approach over traditionally used VAR-based models will be demonstrated using simulations and in the analysis of EEG data. It uncovers non-linear dynamics and yielded novel and insightful findings. The time-evolving Spec NLGC connections gives more meaningful insights regarding the frequency specific connectivity changes at the onset of epileptic seizure as compared to VAR based PDC connections. These confirm the viability of the proposed algorithm as tool for exploring brain connectivity. This is joint work with Archishman Biswas and Malik Sultan of the Biostatistics Group at KAUST.

Amano, Tomoyuki

Title: Analysis of CL and Estimating Function Estimators for Financial Time Series Models

Abstract: There has been proposed many financial time series models in order to represent behaviors of data in the finance and many researchers have investigated these models. One of the most fundamental estimators for financial time series models is the conditional least squares estimator (CL estimator). CL estimator has two advantages; it can be calculated easily and it does not need the knowledge of the innovation. Hence CL estimator has been widely used. However Amano and Taniguchi (2008) showed CL estimator is not asymptotically optimal in general for ARCH model. On the other hand Chandra and Taniguchi (2001) constructed the optimal estimating function estimator (G estimator) for ARCH model based on Godambe's optimal estimating function and showed G estimator is better than CL estimator in the sense of the sample mean squared error by simulation. In this talk we apply CL and G estimators to famous financial time series models (ARCH, GARCH, CHARN) and show G estimator is better than CL estimator in the sense of the efficiency theoretically. Furthermore we derive the condition of the asymptotical optimality of G estimator based on LAN.

Dou, Xiaoling

Title: A smoothed empirical beta copula

Abstract: Empirical beta copula is a non-parametric method of describing dependence structure of multivariate data. In this method, no parameter is needed to be estimate. This advantage makes the empirical beta copula easy to use. However, when the sample size

is large, the computation of this method fails as the Gamma functions in the copula formula become infinity. To make the calculation possible, we can use Stirling's approximation formula to approximate the Gamma functions, then take logarithm and exponential. For practical utility of this copula, we propose a new method to smooth the empirical beta copula by making grids, then estimate the empirical beta copula for the means of data in grids, and multiply the number of the data in grids as weights.

Lin, Liang-Ching

Title: Monitoring Photochemical Pollutants for Anomaly Detection based on Symbolic Interval-Valued Data Analysis

Abstract: This study considers monitoring photochemical pollutants for anomaly detection based on symbolic interval-valued data analysis. For this task, we construct control charts based on the principal component scores of symbolic interval-valued data. Herein, the symbolic interval-valued data are assumed to follow a normal distribution, and an approximate expectation formula of order statistics from the normal distribution is used in the univariate case to estimate the mean and variance via the method of moments. Moreover, we consider the bivariate case wherein we use the maximum likelihood estimator calculated from the likelihood function derived under a bivariate copula. In addition, we establish the procedures for the statistical control chart based on the univariate and bivariate interval-valued variables, and the procedures are potentially extendable to higher dimensional cases. Monte Carlo simulations and real data analysis using photochemical pollutants confirm the validity of the proposed method, and the results particularly show the superiority over the conventional method using the averages in identifying the date on which the abnormal maximum occurred. (Joint work with Meihui Guo and Sangyeol Lee)

Chan, Ngai Hang

Title: Optimal Change-Point Estimation in Time Series

Abstract: TBA

Chen, Cathy W. S.

Title: Bayesian modelling of integer-valued transfer function models

Abstract: External events are commonly known as interventions that often affect times series of counts. This research introduces a class of transfer function models that include four different types of interventions on integer-valued time series: abrupt start and abrupt decay (additive outlier), abrupt start and gradual decay (transient shift), abrupt start and permanent effect (level shift), and gradual start and permanent effect. We propose integer-valued transfer function models incorporating a generalized Poisson, log-linear generalized Poisson, or negative binomial to estimate and detect these four types of

interventions in a time series of counts. Utilizing Bayesian methods, which are adaptive Markov chain Monte Carlo (MCMC) algorithms to obtain the estimation, we further employ deviance information criterion (DIC), posterior odd ratios, and mean squared standardized residual for model comparisons. As an illustration, this study evaluates the effectiveness of our methods through a simulation study and application upon crime data in Albury City, New South Wales Australia. Simulation results show that the MCMC procedure is reasonably effective. The empirical outcome also reveals that the proposed models are able to successfully detect the locations and type of interventions.

Chen, Ying

Title: Probabilistic Forecasting for Daily Electricity Loads and Quantiles for Curve-to-Curve Regression

Abstract: Probabilistic forecasting of electricity load curves is of fundamental importance for effective scheduling and decision making in the increasingly volatile and competitive energy markets. We propose a novel approach to construct probabilistic predictors for curves (PPC), which leads to a natural and new definition of quantiles in the context of curve-to-curve linear regression. There are three types of PPC: a predict set, a predictive band and a predictive quantile, and all of them are defined at a pre-specified nominal probability level. In the simulation study, the PPC achieve promising coverage probabilities under a variety of data generating mechanisms. When applying to one day ahead forecasting for the French daily electricity load curves, PPC outperform several state of the art predictive methods in terms of forecasting accuracy, coverage rate and average length of the predictive bands. For example, PPC achieve up to 2.8-fold of the coverage rate with much smaller average length of the predictive bands. The predictive quantile curves provide insightful information which is highly relevant to hedging risks in electricity supply management. (Joint work with Xiuqin Xu, Yannig Goude and Qiwei Yao. Available at https://arxiv.org/ abs/2009.01595)

Koul, Hira

Title: Minimum Distance Estimation in Linear Errors-in-Variables Regression Model

Abstract: We develop analogs of a class of weighted empirical minimum distance estimators of the underlying parameters in errors-in-variables linear regression models, when the regression error distribution and the conditional distribution of conditionally centered measurement error, given the surrogate, are symmetric around the origin. This class of estimators is defined as the minimizers of integrals of the square of a certain symmetrized weighted empirical process of the residuals. It includes the least absolute deviation and an analog of the Hodges- Lehmann-Sen estimators. We first develop this class of estimators when the distributions of the true covariates and measurement errors are known, and then extend them to the case when these distributions are unknown but validation data is available. An example of the distributions of the errors and covariate is given where the Pitman's asymptotic relative efficiency of some m.d. estimators, relative to the bias corrected LSE, increases to infinity as the ME variance increases to infinity.

Findings of a simulation study show significant superiority of some members of the proposed class of estimators over the bias corrected least squares estimator, in finite samples. In particular, the analog of the Hodges-Lehmann-Sen estimator is seen to be much more robust against the increasing measurement error variance compared to the bias corrected least squares estimator, when the regression error distribution is t2. (Joint work with Pei Geng)

Goto, Yuichi

Title: Homogeneity tests for one-way models with dependent errors under correlated groups

Abstract: In this talk, we introduce homogeneity tests for one-way models with dependent errors. Existing tests are constructed under the assumption of independence between groups. However, this assumption is quite restrictive and impractical. Hence, we propose a test that allows us to deal with correlated groups. A proposed test statistic can be used for both fixed effects model and random effects model. First, we show that, under the null hypothesis, the proposed test statistic converges to the chi-square distribution. From this result, an asymptotically size alpha test can be constructed. Then, we prove the consistency of the test, that is, the power of the test converges to one under the alternative hypothesis. Furthermore, we show our test has a non-trivial power under the local alternative hypothesis. In real data analysis, we test whether or not there exist random effects in three industries. (Joint work with Koichi Arakaki, Yan Liu, and Masanobu Taniguchi)

Guo, Mei-Hui

Title: Change point detection of nonlinear heteroscedastic time series models

Abstract: The SVR-ARMA-GARCH models provide flexible model fitting and good predictive power for nonlinear heteroscedastic time series data. This study explores change point detection of the SVR-ARMA-GARCH model, considering the detection method of Lee et al. (2020) based on cumulative sum of residuals. Firstly, we propose an alternating recursive estimation (ARE) method for the SVR-ARMA-GARCH model to improve the estimation of residuals. Secondly, we propose a new test statistic Ts based on the Brownian Bridge standardized by its time point to improve the detection power in the late stage of the studying time period. Our numerical analysis show that the proposed ARE method of the SVR-ARMA-GARCH models, compared with Chen et. al. (2010), improve the prediction performance and residual estimation for nonlinear heteroscedastic time series models generated from the Lorenz system. And the proposed test statistic Ts has high detection power in the late stage of the time period which makes it more applicable for dynamic monitoring and model adjustment of time series. In the empirical study, we use BDI data to illustrate the performance and advantages of the proposed methods. (Joint work with Hsing-Kai Wang)

Hirukawa, Junichi

Title: Innovation algorithm of fractionally integrated ($I(d)$) process and applications on the estimation of parameters

Abstract: The long memory phenomena frequently occur in the empirical studies of various fields. The fractionally integrated process is the one of the suitable candidates which appropriately represents the long memory property. There are two recursive algorithms for determining the one-step predictors of time series, that is, the Durbin-Levinson algorithm and the innovation algorithm. The Durbin-Levinson algorithm for the fractionally integrated process is well-known and widely used, which naturally derives the Cholesky factorization of the inverse matrix of the covariance matrix of the process. In this talk, we derive the innovation algorithm for the fractionally integrated process. The result is also applied to the derivation of the Cholesky factorization of the covariance matrix of the process in the explicit form. Moreover, the asymptotic theory of Gaussian maximum likelihood estimator (GMLE) is derived in terms of the innovation algorithm. (Joint work with Kou Fujimori)

Dette, Holger

Title: Testing relevant hypotheses in functional time series via self-normalization

Abstract: In this talk we develop methodology for testing relevant hypotheses in a tuning-free way. Our main focus is on functional time series, but extensions to other settings are also discussed. Instead of testing for exact equality, for example for the equality of two mean functions from two independent time series, we propose to test a {\it relevant} deviation under the null hypothesis. In the two sample problem this means that an L^2-distance between the two mean functions is smaller than a pre-specified threshold. For such hypotheses self-normalization, which was introduced by Shao (2010) and is commonly used to avoid the estimation of nuisance parameters, is not directly applicable. We develop new self-normalized procedures for testing relevant hypotheses and demonstrate the particular advantages of this approach in the comparisons of eigenvalues and eigenfunctions.

Negri, Ilia

Title: Z-process method for change point problems in time series

Abstract: Recently a Z-process method were introduced as a general unified approach based on some partial estimation functions to construct a statistical test in some change point problems for ergodic models but also for some models where the Fisher information matrix is random. In this paper, we consider the problem of testing for parameter changes in time series models based on this Z-process method. As examples, we consider the change of the parameter in some linear and nonlinear time series models. Simulation results are also reported for illustration.

Luati, Alessandra

Title: Semiparametric modeling of multiple quantiles

Abstract: We develop a semiparametric model to track a large number of quantiles of a time series. The model satisfies the condition of non-crossing quantiles and the defining property of fixed quantiles. A key feature of the specification is that the updating scheme for time varying quantiles at each probability level is based on the gradient of the check loss function, that forms a martingale difference sequence. Theoretical properties of the proposed model are derived, such as weak stationarity of the quantile process and consistency and asymptotic normality of the estimators of the fixed parameters. The model can be applied for filtering and prediction. We also illustrate a number of possible applications such as: i) semiparametric estimation of dynamic moments of the observables, ii) density prediction, and iii) quantile predictions. (Joint work with Leopoldo Catania)

Giraitis, Liudas

Title: Choosing between persistent and stationary volatility

Abstract: This paper suggests a multiplicative volatility model where volatility is decomposed into a stationary and a non-stationary persistent part. We provide a testing procedure to determine which type of volatility is prevalent in the data. The persistent part of volatility is associated with a nonstationary persistent process satisfying some smoothness and moment conditions. The stationary part is related to stationary conditional heteroskedasticity. We outline theory and conditions that allow the extraction of the persistent part from the data and enable standard conditional heteroskedasticity tests to detect stationary volatility after persistent volatility is taken into account. Monte Carlo results support the testing strategy in small samples. The empirical application of the theory supports the persistent volatility paradigm, suggesting that stationary conditional heteroskedasticity is considerably less pronounced than previously thought. (Joint work with Ilias Chronopoulos and George Kapetanios)

Francq, Christian

Title: Inference on multiplicative component GARCH without any small-order moment

Abstract: In multiplicative component GARCH models, the volatility is decomposed into the product of two factors which often received interpretations in terms of "short run" (high frequency) and "long run" (low frequency) components. While two-component volatility models are widely used in applied works, some of their theoretical properties remain unexplored. We show that the strictly stationary solutions of such models do not admit any small-order finite moment, contrary to classical GARCH. It is shown that the strong consistency and asymptotic normality of the Quasi-Maximum Likelihood estimator hold despite the absence of moments. Tests for the presence of a long-run volatility relying on the asymptotic theory and a bootstrap procedure are proposed. Our results are illustrated

via Monte Carlo experiments and real financial data. (Joint work with Baye Matar Kandji and Jean-Michel Zakoian)

Patilea, Valentin

Title: On Wilks' Theorem for Conditional Moment Equations with Weakly Dependent Data

Abstract: Many statistical and econometric predictive models, in particular time series models, are defined by conditional moment equations. The common attitude for proceeding with inference in such models is to derive an alternative finite set of unconditional moment equations from the original model and to build inference based on them. However, in general, the finite set of unconditional moment equations define a different model, that is a different set of probability distributions, in which the identification of the parameter of interest, and hence the validity of the inference, can be lost. We here define a finite set of estimating equations which guarantee the identification, and we build a new empirical likelihood inference approach using these estimating equations, with weakly dependent data. We provide sufficient conditions under which the empirical log-likelihood ratio statistics converges to a chi-square distribution. The general methodology is applied to nonlinear and semiparametric time series models. The talk is based on joint work with Marie du Roy de Chaumaray and Matthieu Marbac (CREST).

Preinerstorfer, David

Title: Consistency of p-norm based tests in high dimensions

Abstract: Many commonly used test statistics are based on a norm measuring the evidence against the null hypothesis. To understand how the choice of a norm affects power properties of tests in high dimensions, we study the consistency sets of p-norm based tests in the prototypical framework of sequence models with unrestricted parameter spaces, the null hypothesis being that all observations have zero mean. The consistency set of a test is here defined as the set of all arrays of alternatives the test is consistent against as the dimension of the parameter space diverges. We characterize the consistency sets of p-norm based tests. This characterization reveals an unexpected monotonicity result that allows us to construct novel tests that dominate, with respect to their consistency behavior, all p-norm based tests without sacrificing size.

Dallakyan, Aramayis and Pourahmadi, Mohsen

Title: Learning Bayesian Networks through Birkhoff Polytope: A Relaxation Method

Abstract: We establish a novel framework for learning a directed acyclic graph (DAG) when data are generated from a Gaussian, linear structural equation model. It consists of two parts: (1) introduce a permutation matrix as a new parameter within a regularized Gaussian log-likelihood to represent variable ordering; and (2) given the ordering, estimate the DAG structure through sparse Cholesky factor of the inverse covariance matrix. For permutation

matrix estimation, we propose a relaxation technique that avoids the NP-hard combinatorial problem of order estimation. Given an ordering, a sparse Cholesky factor is estimated using a cyclic coordinate wise descent algorithm which decouples row-wise. Our framework recovers DAGs without the need for an expensive verification of the acyclicity constraint or enumeration of possible parent sets. We establish numerical convergence of the algorithm, and consistency of the Cholesky factor estimator when the order of variables is known. Through several simulated and macro-economic datasets, we study the scope and performance of the proposed methodology.

Stoffer, David S

Title: Efficient Fitting of Stochastic Volatility Models

Abstract: The stochastic volatility model is a popular tool for modeling the volatility of assets. The model is a nonlinear and non-Gaussian state space model and presents some challenges not seen in general. Many approaches have been developed for Bayesian analysis that rely on numerically intensive techniques such as Markov chain Monte Carlo (MCMC). Convergence and mixing problems still plague MCMC algorithms used for the model. We present an approach that ameliorates the slow convergence and mixing problems when fitting stochastic volatility models. The approach accelerates the convergence by exploiting the geometry of one of the targets. We demonstrate the method on various numerical examples.

Tsay, Ruey S

Title: Divide-and-Conquer: A Distributed Hierarchical Factor Approach to Modeling Large-Scale Time Series Data

Abstract: This paper proposes a hierarchical approximate-factor approach to analyzing high-dimensional, large-scale heterogeneous time series data using distributed computing. The new method employs a multiple-fold dimension reduction procedure using Principal Component Analysis (PCA) and shows great promises for modeling large-scale data that cannot be stored nor analyzed by a single machine. Each computer at the basic level performs a PCA to extract common factors among the time series assigned to it and transfers those factors to one and only one node of the second level. Each 2nd-level computer collects the common factors from its subordinates and performs another PCA to select the 2nd-level common factors. This process is repeated until the central server is reached, which collects factors from its direct subordinates and performs a _nal PCA to select the global common factors. The noise terms of the 2nd-level approximate factor model are the unique common factors of the 1st-level clusters. We focus on the case of 2 levels in our theoretical derivations, but the idea can easily be generalized to any finite number of hierarchies, and the proposed method is also feasible for the data stored and to be analyzed by a single machine with heterogeneous and multilevel subcluster structures. We discuss some clustering methods when the group memberships are unknown and introduce a new diffusion index approach to forecasting. We further extend the analysis to

unit-root nonstationary time series. Asymptotic properties of the proposed method are derived for the diverging dimension of the data in each computing unit and the sample size T. We use both simulated and real examples to assess the performance of the proposed method in finite samples, and compare our method with the commonly used ones in the literature concerning the forecasting ability of extracted factors. (Joint work with Zhaoxing Gao)

Wang, Yuan

Title: Topological Inference with Spectral Permutation Test

Abstract: Topological data analysis (TDA) can decode multiscale patterns in electroencephalographic (EEG) signals not captured by standard temporal and spectral features. A challenge for applying TDA to groups of long EEG recordings is the ambiguity of performing statistical inference and computational efficiency. To address this problem, we advance a unified inference framework based on a fast permutation test for comparing the TDA descriptor persistence landscape (PL) between two groups of multi-trial EEG signals. The topological inference framework is applied to investigate the EEG correlates of speech sensorimotor impairment in post-stroke aphasia patients under a speech altered auditory feedback (AAF) paradigm. Our analysis reveals a significant difference between the PL features extracted from the event- related potential (ERP) response in aphasia vs. control groups over the parietal-occipital and occipital regions when there is no pitch shift in the auditory feedback and over the parietal region when there an upward pitch shift. The findings validate the application of TDA analysis as a robust tool for investigating the neural correlates of speech sensorimotor impairment in neurological patients suffering from speech-language disorders.

Huang, Shih-Feng

Title: A network autoregressive model with GARCH effects and its applications

Abstract: In this study, a network autoregressive model with GARCH effects, denoted by NAR-GARCH, is proposed to depict the return dynamics of stock market indices. A GARCH filter is employed to marginally remove the GARCH effects of each index, and the NAR model with the Granger causality test and Pearson's correlation test with sharp price movements is used to capture the joint effects caused by other indices with the most updated market information. The NAR-GARCH model is designed to depict the joint effects of nonsynchronous multiple time series in an easy-to-implement and effective way. The returns of 20 global stock indices from 2006 to 2020 are employed for our empirical investigation. The numerical results reveal that the NAR-GARCH model has satisfactory performance in both fitting and prediction for the 20 stock indices, especially when a market index has strong upward or downward movements.

Ing, Ching-Kang

Title: Model selection for unit-root time series with many predictors

Abstract: We study model selection for unit-root time series when a large number of predictors is present. We propose a new model selection algorithm, called FSR+HDIC+Trim, that leverages the advantages of the Forward Stepwise Regression (FSR) and the high-dimensional information criterion (HDIC). Under a set of mild assumptions that allow unknown, flexible multiplicities and locations of the roots on the unit circle as well as conditionally heteroskedastic errors, we establish the selection consistency of FSR+HDIC+Trim. We also demonstrate the outstanding finite-sample performance of the proposed method by simulation studies and an example using U.S. monthly macroeconomic data.

Whang, Yoon-Jae

Title: Testing for Time Stochastic Dominance

Abstract: We propose nonparametric tests for the null hypothesis of time stochastic dominance. Time stochastic dominance makes a partial order of different prospects over time based on the net present value criteria for general utility and time discount function classes. For example, time stochastic dominance can be used for ranking investment strategies or environmental policies based on the expected net present value of the future benefits. We consider an L_p-integrated test statistic and derive its large sample distribution. We suggest bootstrap procedures that allow for time dependence in a panel data structure. We describe two approaches, the contact-set approach and the numerical delta method, that may lead to enhanced power compared to the conventional least-favorable-case based approach. We prove the asymptotic validity of our testing procedures. We investigate the finite sample performance of the tests in simulation studies. As an illustration, we apply the proposed tests to evaluate the Million Baht Village Fund Program in Thailand and carbon emission trading scheme in China. (Joint work with Kyungho Lee and Oliver Linton)

Nishiyama, Yoichi

Title: On rank statistics of PageRank and MarkovRank

Abstract: An important statistic in analyzing some (finite) network data, called *PageRank*, and a related new statistic, which we call *MarkovRank*, are studied in this paper. The PageRank was originally developed by the cofounders of *Google*, Sergey Brin and Larry Page, to optimize the ranking of websites for their search engine outcomes, and it is computed using an iterative algorithm, based on the idea that nodes with a larger number of incoming edges are more important.

The aim of the paper is to analyze the common features and some significant differences between PageRank, which is well established, and MarkovRank, which is newly introduced. A common merit of the two Ranks is that both statistics can be easily

computed by either the mathematical computation or the iterative algorithm. According to the analysis of some examples, these two statistics seem to return somewhat different values, but the resulting rank statistics of both statistics are not far away from each other. One of the differences is that only MarkovRank has the property that its rank statistic does not depend on any tuning parameter, and it is determined only through the adjacency matrix for given network data. Moreover, it is also shown that the rank statistic of MarkovRank is identical to or "finer than"' that of the stationary distribution vector of the corresponding Markov chain (with finite state space) whenever the Markov chain is regular. Thus, MarkovRank may have a potential to play a similar role to PageRank from a practical point of view, not only commonly with light computational tasks, but also with some new theoretical validity.

Kurita, Mari

Title: Serious Opera and Comic Opera in Italy in the 18th century – Musical Styles and their Impact

Abstract: The word Opera describes various artistic realities such as poetry, music, scenography, acting and dance. The foundation of Opera was laid down by a group of composers from the Neapolitan School, who were active in Naples at the end of the 17th and 18th centuries. Serious Opera elaborated historical or Greek mythological topics with heroic or aristocratic characters. The stage of Serious Opera was the place where castrati sang beautiful Bel canto arias with great skill. On the other hand, Comic Opera reflected the popular and local world, often with political and social satire. It was in contrast to the aristocratic and cosmopolitan Serious Opera. Intermezzo was a short Comic Opera that was performed during the intervals of a Serious Opera. Opposite to the traditional French Opera, which reflected the ideals of the Enlightenment, Intermezzo used simple and free expression. Serious Opera influenced many subsequent composers such as Gluck who achieved the unification of traditional Italian Opera and traditional French Opera. Very interestingly, the greatest masterpiece of Opera that combines the characteristics of Serious Opera and Comic Opera is Mozart's Don Giovanni. The 18th century Italian Opera, which laid the foundation for Opera, not only influenced Gluck and Mozart, but also played a major role in the history of Opera.

Ogata, Hiroaki

Title: Pair circulas modelling for circular multivariate time series

Abstract: Modelling multivariate circular time series is considered. The cross-sectional and serial dependence is described by circulas, which are analogs of copulas for circular distributions. Due to a simple expression of the dependence structure, we decompose a multivariate circula density to a product of several pair circula densities. Moreover, in order to reduce the number of pair circula densities, we consider strictly stationary multi-order Markov processes. Some simulation studies are provided to see the behavior of the proposed model.

Shiohama, Takayuki

Title: Topological data analysis based classification in time series

Abstract: Time series often contain outliers and level shifts or structural changes. These unexpected events are of the utmost importance in anomaly detection. The presence of such unusual events can easily mislead conventional time series analysis and yield erroneous conclusions. Anomaly detection methods for time series have been studied for decades and demonstrated to be useful in many applications. There exist many notable methods in machine learning, which include clustering analysis, isolation forests, and classifiers using artificial neural networks. Most of these techniques often are most effective when there are many additional features. In this study, we use topological data analysis (TDA) in order to provide more accurate classifier that can also detect unusual events in time series.

Shiraishi, Hiroshi

Title: Time Series Quantile Regression by using Random Forests

Abstract: We consider the estimation of conditional quantiles for dependent data by using random forests. The quantile regression introduced by Koenker and Bassett (1978) has been discussed under time series setting by some paper such as QAR (Koenker and Xiao, 2004) or CAViaR (Engle and Manganeli, 2004). On the other hand, Quantile Regression Forests (QRF) is introduced under regression model by Meinshausen (2006). We discuss the QRF under time series setting.

Shimizu, Yasutaka

Title: A quite new approach to cohort-wise mortality prediction under survival energy hypothesis

Abstract: We propose a new approach to mortality prediction by “Survival Energy Model (SEM)”. We assume that a human is born with initial energy, which changes stochastically in time and the human dies when the energy vanishes. Then, the time of death is represented by the first hitting time of the survival energy (SE) process to zero. This study assumes that SE follows a (time-inhomogeneous) diffusion process or an inverse Gaussian process, and defines the “mortality function”, which is the first hitting time distribution function of a SE process. Although SEM is a fictitious construct, we illustrate that this assumption has a high potential to yield a good parametric family of the cumulative distribution of death, and the parametric family yields surprisingly good predictions for future mortality rates. This work is published by Shimizu, et al. (2020). ``Why does a human die? A structural approach to cohort-wise mortality prediction under survival energy hypothesis", ASTIN Bulletin, vol.51 (1), 191-219.

Monti, Anna Clara

Title: Tango: dance and statistical thinking

Abstract: The talk briefly recalls the origins of Tango, reviews the main styles and types of music and illustrates the fundamental steps and figures. Tango does not rely on well established sequences of steps but is heavily based on improvisation. At any moment, the dance depends on the connections among the dancers, but it is also influenced by the music, the style, the expertise of the dancers and their previous interactions. The harmony of a tango relies on how these information are processed by the dancers. During a tango event, dancers alternate active and idling times, and typically they change their partner. A model is outlined to describe the activity of a female dancer. Some videos end the talk.

Hallin, Marc

Title: Fully Distribution-Free Center-Outward Rank Tests for Multiple-Output Regression and MANOVA

Abstract: Extending rank-based inference to a multivariate setting such as multiple-output regression or MANOVA with unspecified d-dimensional error density has remained an open problem for more than half a century. None of the many solutions proposed so far is enjoying the combination of distribution-freeness and efficiency that makes rank-based inference a successful tool in the univariate setting. A concept of *center-outward* multivariate ranks and signs based on measure transportation ideas has been introduced recently. Center-outward ranks and signs are not only distribution-free but achieve in dimension $d>1$ the (essential) maximal ancillarity property of traditional univariate ranks, hence carry all the "distribution-free information" available in the sample. We derive here the Hájek representation and asymptotic normality results required in the construction of center-outward rank tests for multiple-output regression and MANOVA. When based on appropriate spherical scores, these fully distribution-free tests achieve parametric efficiency in the corresponding models. (Based on arxiv:2007.15496, joint with Daniel Hlubinka and Šárka Hudecová)

Taniguchi, Masanobu

Title: The Long and Curved Road

Abstract: This talk surveys my life research. The following topics will be delivered.
(1) Introduction of spectral divergence and discussion on efficiency and robustness.
(2) Development in high-order asymptotic theory of time series analysis. Beyond the simultaneous equation analysis.
(3) Statistical analysis of "curved" stochastic models.
(4) Foundation of time series discriminant analysis
(5) Statistical theory based on integral functional of nonparametric spectral estimators. Semiparametric estimation for spectra. Introduction of high-order asymptotic theory for semiparametric time series estimators.
(6) LAN based asymptotic theory for time series, including long memory ones.

(7) Systematic approach for portmanteau tests.
(8) Empirical likelihood approach for time series.
(9) Non-regular estimation for time series, and Bartlett adjustment for non-standard settings.
(10) Asymptotic theory of shrinkage estimation for time series.
(11) Asymptotic theory for portfolio estimation.
(12) New look at circular distributions in view of high-order spectral distribution of stationary processes.
(13) Analysis of variance for time series

Harada, Yuya

Title: Performance expression and statistical consideration in opera (Classic Music)

Abstract: (i) Play multiple operas with different age of different performance expressions.
(ii) Listen to the performance and consider whether statistical consideration can be done.

Xu, Xiaofei

Title: Long-memory Log-linear Zero-inflated Generalized Poisson Autoregression for COVID-19 Pandemic Modeling

Abstract: This paper describes the dynamics of count time series of daily new cases arising from the COVID-19 pandemic using a long-range dependent model. A new long memory model, LFIG model (Log-linear zero-inflated generalized Poisson integer-valued Fractionally Integrated GARCH process), is proposed to provide a novel integer-valued long memory process enabling to account for the multiple features including serial dependence (positive or negative), over-dispersion, zero-inflation, nonlinearity in a unified framework. We adopt an adaptive Bayesian Markov Chain Monte Carlo (MCMC) sampling scheme for parameters estimation. This new modeling is applied to the daily new confirmed cases of COVID-19 pandemic in six countries including Japan, Vietnam, Italy, the United Kingdom, Brazil and the United States. The numerical study demonstrates good interpretation and forecasting performance of long-memory modeling to COVID-19 pandemic. (Joint work with Ying Chen, Yan Liu, Yuichi Goto and Masanobu Taniguchi)

Xue, Yujie

Title: Hellinger distance estimation for non-regular spectra

Abstract: For a Gaussian stationary process, in this talk, we will derive a time series Hellinger distance for spectra f and g: $T(f,g) = \int \log\left\{ 1/2\sqrt{f/g} + 1/2\sqrt{g/f} \right\} \mathrm{d}\lambda$. By evaluating $T(f_\theta, f_{\theta+h})$ of the form $\mathrm{O}(h^\alpha)$, we elucidate the $n^{1/\alpha}$-consistent asymptotics of

the maximum likelihood estimator of θ for non-regular spectra. For regular spectra, we introduce the minimum Hellinger distance estimator $\theta^T\left(\hat{g}_n\right) = \mathrm{argmin}_\theta T(f_\theta, \hat{g}_n)$ where $\hat{g}_n$ is a nonparametric spectral density estimator, and as a benchmark, we introduce the Whittle divergence estimator $\theta^W\left(\hat{g}_n\right)$. It can be shown that both $\theta^T\left(\hat{g}_n\right)$ and $\theta^W\left(\hat{g}_n\right)$ are asymptotically efficient, and that the former is more robust than the latter. Besides, small numerical studies will be provided. (Joint work with Masanobu Taniguchi)

Akashi, Fumiya

Title: Self-weighted GEL method based on spatial median

Abstract: In this talk, we consider estimation and testing problems of multivariate autoregressive processes with possibly heavy-tailed errors. Based on the idea of spatial medians, we extended the self-weighting approach (Ling, 2005) to the multivariate possibly infinite variance processes. We propose a self-weighted generalized empirical likelihood (GEL) estimator and show its asymptotic distribution. The GEL statistics for the testing problem of linear hypothesis of the model coefficients is also proposed, and we construct a feasible testing procedure regardless of whether the model has a finite or infinite variance. Some simulation experiments are given.

Honda, Toshio

Title: Forward variable selection for ultra-high dimensional quantile regression models

Abstract: We propose forward variable selection procedures with a stopping rule for feature screening in ultra-high dimensional quantile regression models. For such very large models, penalized methods do not work and some preliminary feature screening is necessary. We demonstrate the desirable theoretical properties of our forward procedures by taking care of uniformity w.r.t. subsets of covariates properly. The necessity of such uniformity is often overlooked in the literature. Our stopping rule suitably incorporates the model size at each stage. We also present the results of simulation studies and a real data application to show their good finite sample performances. (Joint work with Chien-Tong Lin in National Tsing Hua University)

Yamashita, Satoshi

Title: Statistical matching method for data profiling and fusion

Abstract: Statistical matching techniques aim to build a dataset by combining different data sources. In recent years, matching techniques have been employed in various fields. However, because of some difficulties, there are only a few applications to company data. In this study, we proposed a new statistical matching methodology for company datasets by employing multinomial logistic regression model. The weighted distance was used to

compute the probability of true match pairs through the model. The probability helps classify the record pairs as truly matched or not. We applied these techniques to a commercial company dataset and the official economic census micro data. The results showed that our method performs better than the nearest neighbor method used in the previous study in terms of true match rate.

Matsuda, Yasumasa

Title: Functional regression models for spatio-temporal data

Abstract: Functional regression is an extension of regression when both dependent and independent variables are function-valued. In this talk, we regard spatial data as square integrable function-valued random variables, and construct a regression model by a convolution operator. We propose a frequency domain approach to estimate parameters that can overcome typical difficulties in spatial data analysis, including irregularly spaced observation locations with huge sample sizes, lots of NAs and so on. We clarify the asymptotic regime under which the estimator is consistent and asymptotic normal. We demonstrate our functional regression analysis with an application to the spatial dataset of NTT Docomo human mobility survey.

Minami, Mihoko

Title: Regression Tree and Clustering for Distributions, and Homogeneous Structure of Population Characteristics

Abstract: We often collect samples on characteristics of different observation units and wonder whether the characteristics of the observation units have similar distributional structure. In this talk, we consider methods to find homogeneous subpopulations using regression tree and clustering for distributions approaches based on a modified Jensen-Shannon divergence. We present a statistic to test cluster homogeneity and a hierarchical testing procedure to find the minimal homogeneous tree structure of the distribution of a population characteristic. As a motivational example, we introduce yellowfin tuna fork length data collected from the tuna catch of purse-seine vessels that operated in the eastern Pacific Ocean during 2003 - 2007. We apply the regression tree and clustering for distributions methods to these data to study the spatial structure of the length composition of the yellowfin tuna catch, and discuss the insights the results of the different methods provide about spatial structure of the size composition of yellowfin tuna catch in the eastern Pacific Ocean. We also apply the hierarchical testing procedure to these data to find the minimal homogeneous tree structure of yellowfin tuna fork length distributions. This is a joint work with Dr. Cleridy E. Lennert-Cody of Inter-American Tropical Tuna Commission.

So, Mike K P

Title: Topological properties of pandemic and financial networks: prediction and risk management

Abstract: In this paper, we consider network analysis as a tool to link multiple, possibly high-dimensional, variables together. In epidemiological research, forming a network to relate different locations or people together can provide information on the propagation of infectious diseases and the severity of disease outbreak. In financial research, it is possible to assess how different financial institutions are related, through which we can quantify financial network connectedness. We demonstrate how we may use pandemic and financial network topological properties for predicting pandemic and financial risk. Our examples using COVID-19 data and stock market data showcase that network analysis can be a valuable modeling tool for risk management.

Liu, Yan

Title: Statistical and Topological Inference for Local Granger Causality

Abstract: Granger causality has been employed to investigate causality relations between components of stationary multiple time series. We generalize this concept by developing statistical inference for local Granger causality for multivariate locally stationary processes. Our proposed local Granger causality approach captures time- evolving causality relationships in nonstationary processes. The proposed local Granger causality is well represented in the frequency domain and estimated based on the parametric time-varying spectral density matrix using the local Whittle likelihood. Under regularity conditions, we demonstrate that the estimators converge to multivariate normal in distribution. Additionally, the test statistic for the local Granger causality is shown to be asymptotically distributed as a quadratic form of a multivariate normal distribution. For practical demonstration, the proposed local Granger causality method uncovered new functional connectivity relationships between channels in brain signals. Moreover, the method was able to identify structural changes in financial data. (Joint work with Masanobu Taniguchi and Hernando Ombao)

Solvang, Hiroko Kato

Title: Categorical data analysis to infer directional associations among rorquals and their prey species

Abstract: Several rorqual species migrate annually to take advantage of high summer productivity and feeding opportunities at high latitudes. Recent climate change and fluctuations in environmental conditions in ocean have affected spatial distribution of both zooplankton and fish species known to be potential rorqual prey. Investigating the directional association among rorquals and their prey is important to explore causality in their trophic interactions in the ecosystem, especially in light of recent ecosystem change. For this purpose, a relevant and practical statistical method is required to link whale observations with prey abundance. To meet this challenge, we consider an approach based on a categorical data analysis provided by Sakamoto and Akaike 1998. The basic procedure in the analysis is to find the best fit model among model candidates, including

combinations of several conditional probabilities to investigate possible causal relationships among variables. In this study, we involve additional procedure for verifying optimum threshold for the categorization of continuous data. We examine the proposed approach by simulation study. The procedures were applied to two data sets, one collected during eco system surveys to the west and north of Svalbard in 2014-2017, the other one collected during an international synoptic krill survey in the Scotia Sea west of the Antarctic Peninsula in 2019.

Verdebout, Thomas

Title: Asymptotic power of Sobolev tests for uniformity on hyperspheres

Abstract: One of the most classical problems in multivariate statistics is considered, namely, the problem of testing isotropy, or equivalently, the problem of testing uniformity on the unit hypersphere. Rather than restricting to tests that can detect specific types of alternatives only, we consider the broad class of Sobolev tests. While these tests are known to allow for omnibus testing of uniformity, their non-null behavior and consistency rates, unexpectedly, remain largely unexplored. To improve on this, we thoroughly study the local asymptotic powers of Sobolev tests under the most classical alternatives to uniformity, namely, under rotationally symmetric alternatives. We show in particular that the consistency rate of Sobolev tests does not only depend on the coefficients defining these tests but also on the derivatives of the underlying angular function at zero. (Joint work with Eduardo García-Portugués and Davy Paindaveine)

Proietti, Tommaso

Title: Peaks, Gaps, and Time Reversibility of Economic Time Series

Abstract: By locating the running maxima and minima of a time series, and measuring the current deviation from them, it is possible to generate processes that are analytically relevant for the analysis of the business cycle and for characterizing bull and bear phases in financial markets. First, the measurement of the time distance from the running peak originates a first order Markov chain, whose characteristics can be used for testing time reversibility of economic dynamics and specific types of asymmetries in financial markets. Secondly, the gap processes can be combined to provide a nonparametric measure of the growth cycle. The paper derives the time series properties of the gap process and other related processes that arise from the same measurement context, and proposes new nonparametric tests of time reversibility and new measures of the output gap.

Zakoian, Jean-Michel

Title: Testing hypotheses on the innovations distribution in semi-parametric conditional volatility models

Abstract: The paper considers the problem of testing assumptions on the innovations in GARCH-type models. We propose tests of different hypotheses: adequacy of a parametric quantile, mean-median equality, symmetry of extreme quantiles and zero-median in presence of a conditional mean. The tests rely on the asymptotic distribution of the empirical distribution function of the residuals. They are generally model-free (though not estimation-free) and thus are simple to implement. Efficiency comparisons are made and a numerical study based on simulated and real data is provided. (Joint work with Christian Francq)

Taylor, Robert

Title: Extensions to IVX Methods of Inference for Return Predictability

Abstract: Predictive regressions are widely employed in empirical finance in scenarios where we wish to predict returns series using predictors (lagged macroeconomic and financial variables) which may be strongly persistent and endogenous. Standard methods of statistical inference are invalid in such scenarios. Extended IV [IVX] estimation and inference has proved particularly valuable in this endeavour as it is asymptotically valid for both strongly and weakly persistent regressors, and regardless of any endogeneity present. The contribution of this paper is threefold. First we demonstrate that, provided either a suitable bootstrap implementation is employed or heteroskedasticity consistent standard errors are used, the IVX-based predictability tests of Kostakis et al. (2015) retain asymptotically valid inference under the null hypothesis under considerably weaker assumptions on the innovations than are required by Kostakis et al. (2015) in their analysis. In particular, we allow for quite general forms of conditional and unconditional heteroskedasticity, neither of which are tied to a parametric model. Second, and associatedly, we develop asymptotically valid bootstrap implementations of the IVX tests under these conditions. Monte Carlo simulations show that the bootstrap methods we propose can deliver considerably more accurate finite sample inference than the asymptotic implementation of these tests used in Kostakis et al. (2015) under certain problematic parameter constellations, most notably for their implementation against one-sided alternatives, and where multiple predictors are included. Third, under the same conditions as we consider for the full-sample tests, we show how sub-sample implementations of the IVX approach, coupled with a suitable bootstrap, can be used to develop asymptotically valid one-sided and two-sided tests for the presence of temporary windows of predictability. (Joint work with Matei Demetrescu, Iliyan Georgiev, and Paulo Rodrigues)

Ronchetti, Elvezio

Title: Optimal transportation through saddlepoints

Abstract: We showcase some unexplored connections between saddlepoint approximations, measure transportation, and some key topics in information theory by reviewing selectively some fundamental results available in the literature. We start with the

link between Esscher's tilting (a result rooted in information theory that lies at the heart of saddlepoint approximations) and the solution of the dual Kantorovich problem (which lies at the heart of measure transportation theory) via the Legendre transform of the cumulant generating function. We then investigate these links in the framework of M-estimators and quantile regression. The unveiled connections offer the possibility to view saddlepoint approximations from different angles, putting under the spotlight the links to e.g. convex analysis (via the notion of duality) or differential geometry (via the notion of geodesic). (Joint work with Davide La Vecchia and Andrej Ilievski)

Hayamizu, Momoko

Title: Treefit: measuring the tree-likeness of point clouds and application to the analysis of single-cell gene expression data

Abstract: Single-cell gene expression data that comprehensively capture the state of individual cells is one of the breakthroughs in life sciences in recent years and is regarded as a treasure trove of novel insights into cells and cell differentiation processes. However, the methodology for extracting information from this new type of data is not yet established and there are many mathematical challenges that still need to be resolved. For example, 'trajectory inference' (TI), a task to estimate a tree topology representing a cell differentiation trajectory from high-dimensional point cloud data, is a hot topic in computational biology, and various bioinformatics methods for TI have been proposed and implemented to date. However, it remains difficult to quantify the validity of estimated trees because of the fundamental problem that there is no standard method for measuring the goodness-of-fit between point cloud data and tree models. In this talk, I will present a mathematical method and software to address this problem (https://hayamizu-lab.github.io/treefit/).