Waseda International Symposium

Topological Data Science, Causality & Time Series Analysis

Date: February 27 – 29, 2020

Venue: Nishi-Waseda Campus, Waseda University
Building 55N, Meeting Room 2, 2nd floor
(Access map: https://www.waseda.jp/fsci/en/access/)

Organizer: Masanobu TANIGUCHI
(Research Institute for Science & Engineering, Waseda University)

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 & Time Series Analysis"

Date: February 27–29, 2020

Venue: Nishi-Waseda Campus, Waseda University
Building 55N, Meeting Room 2, 2nd floor

(Access map: https://www.waseda.jp/fsci/en/access/)

Organizer: Masanobu TANIGUCHI

(Research Institute for Science & Engineering, Waseda University)

This workshop 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

February 27

13:20–13:30: Masanobu Taniguchi (Waseda University)
Opening

Session I (13:30–14:50): chaired by Christian Francq

13:30–14:10: Yan Liu (Waseda University)
Statistical Inference for Persistence Landscapes of the Granger Causality

14:10–14:50: Liang-Ching Lin (National Cheng Kung University)
Symbolic Interval-Valued Data Analysis for Time Series Based on Auto-Interval-Regressive Models

14:50–15:05: Coffee break

Session II (15:05–17:15): chaired by Elvezio Ronchetti

15:05–15:45: Shih-Feng Huang (National University of Kaohsiung)
Multi-Asset Empirical Martingale Price Estimators for Financial Derivatives

15:45–16:30: Moo K. Chung (University of Wisconsin-Madison)
Exact topological inference on resting-state brain networks

16:30–17:15: Rainer von Sachs (Université catholique de Louvain)
Intrinsic wavelet regression for curves of Hermitian positive definite matrices

February 28

Session III (9:30–10:30): chaired by David Preinerstorfer

9:30–10:00: Yuichi Goto (Waseda University)
Estimation of Trigonometric Moments for Circular Distribution of MA(p) Type by Using Binary Series

10:00–10:40: Kou Fujimori (Waseda University)
Moment convergence of the generalized maximum composite likelihood estimators for determinantal point processes

10:40–10:50: Coffee Break

Session IV (10:50–12:10): chaired by Michael Eichler

10:50–11:30: Yuan Wang (University of South Carolina)
Group Analysis of Multi-Trial Brain Signals via Persistent Morse Homology

11:30–12:10: Fumiya Akashi (University of Tokyo)
Robust regression on hyper-spheres with unspecified heteroscedastic errors and smooth approximation of object functions

12:10–13:20: Lunch Time

Session V (13:20–14:40): chaired by Shih-Feng Huang

13:20–14:00: Ying Chen (National University of Singapore)
Regularized Partially Functional Autoregressive Model

14:00–14:40: Michael Eichler (Maastricht University)
Causal Inference from Multivariate Time Series: Principles and Problems

14:40–14:55: Coffee Break

Session VI (14:55–16:25): chaired by Mike K P So

14:55–15:40: Christian Francq (ENSAE)
Count and duration time series with equal conditional stochastic and mean orders

15:40–16:25: Elvezio Ronchetti (University of Geneva)
Saddlepoint approximations for short and long memory time series: A frequency domain approach

Special Art Session VII (16:30–18:10):
chaired by Rainer von Sachs & Moo K. Chung

16:30–17:15: Anna Clara Monti (University of Sannio)
Tango: dance and statistical thinking

17:25–18:10: Yuya Harada (Tokyo University of the Arts)
Pianist 🎹: Sayo Zenyoji (Tokyo University of the Arts)
The vocal performance of OPERA & Japanese artistic songs

18:30– Buffet Party

February 29

Session VIII (9:30–10:10): chaired by Liang-Ching Lin

9:30–10:10: Koji Tsukuda (University of Tokyo)

A change detection procedure for an ergodic diffusion process

10:10–10:25: Coffee Break

Session IX (10:25–11:50): chaired Ying Chen

10:25–11:05: David Preinerstorfer (University Libre Brussels)

Functional sequential treatment allocation

11:05–11:50: Mike K P So (HKUST)

Bayesian Network Analysis of Systemic Risk in Financial Markets

Abstracts

February 27 (13:20–17:15)

Yan Liu

Title: Statistical Inference for Persistence Landscapes of the Granger Causality

Abstract: We propose a topological approach to statistically analyzing the Granger causality. Granger introduced his celebrated new measure of causality in the sense of prediction errors of multivariate time series 50 years ago. We localize his idea and construct a theory based on locally stationary processes for its alternative version, a natural refinement for stationary processes by Hosoya. To construct the theory, we provide a Gaussian approximation of the suprema of empirical spectral processes. Especially, the local extension of the theory serves for the statistical inference for the Granger causality curve. In addition, we provide a bootstrap procedure for the approximation to construct confidence bands. Finally, we discuss the persistence diagrams and persistence landscapes for the causality curves and numerically construct some examples of locally stationary processes for our simulations studies. (Joint work with Akitoshi Kimura, Masanobu Taniguchi and Hernando Ombao)

Liang-Ching Lin

Title: Symbolic Interval-Valued Data Analysis for Time Series Based on Auto-Interval-Regressive Models

Abstract: This study considers interval-valued time series data. To characterize such data, we propose an auto-interval-regressive (AIR) model using the order statistics from normal distributions. Furthermore, to better capture the heteroscedasticity in volatility, we design a heteroscedastic volatility AIR (HVAIR) model. We derive the likelihood functions of the AIR and HVAIR models to obtain the maximum likelihood estimator. Monte Carlo simulations are then conducted to evaluate our methods of estimation and con rm their validity. A real data example from the S&P 500 Index is used to demonstrate our method. (Joint work with Hsiang-Lin Chien and Sangyeol Lee)

Shih-Feng Huang

Title: Multi-Asset Empirical Martingale Price Estimators for Financial Derivatives

Abstract: This study proposes an empirical martingale simulation (EMS) and an empirical P-martingale simulation (EPMS) as price estimators for multi-asset financial derivatives. Under mild assumptions on the payoff functions, strong consistency and asymptotic normality of the proposed estimators are established. Several simulation scenarios are conducted to investigate the performance of the proposed price estimators under multivariate geometric Brownian motion, multivariate GARCH models, multivariate jump-diffusion models, and multivariate stochastic volatility models. Numerical results indicate that the multi-asset EMS and EPMS price estimators are capable of improving the efficiency of their Monte Carlo counterparts. In addition, the asymptotic distribution serves as a persuasive approximation to the finite-sample distribution of the EPMS price estimator, which helps to reduce the computation time of finding confidence intervals for the prices of multi-asset derivatives.

Moo K. Chung

Title: Exact topological inference on resting-state brain networks

Abstract: Advances in functional magnetic resonance imaging (fMRI) enable us to measure spontaneous fluctuations of neural signals in the brain in higher spatial and temporal resolution than before. Many previous studies on resting-state fMRI have mainly focused on the topological characterization of graph theory features that fluctuate over the choice of parameters. Persistent homology provides a more coherent mathematical framework for quantifying brain networks that are robust to changes. Instead of looking at networks at a fixed scale, persistent homology charts the changes in topological features such as Betti numbers over every possible parameter. In doing so, it reveals the most persistent topological features that are robust to parameter and noise. In this talk, the exact probability distribution on the Betti numbers that are used in determining the statistical significance will be discussed (Network Neuroscience 3:674-694). Two open mathematical problems (three-sample test and higher order Betti number) related to Betti numbers will be presented.

Rainer von Sachs

Title: Intrinsic wavelet regression for curves of Hermitian positive definite matrices

Abstract: Intrinsic wavelet transforms and wavelet estimation methods are introduced for curves in the non-Euclidean space of Hermitian positive definite matrices, with in mind the application to Fourier spectral estimation of multivariate stationary time series. The main focus is on intrinsic average-interpolation wavelet transforms in

the space of positive definite matrices equipped with an affine-invariant Riemannian metric, and convergence rates of linear wavelet thresholding are derived for intrinsically smooth curves of Hermitian positive definite matrices. In the context of multivariate Fourier spectral estimation, intrinsic wavelet thresholding is equivariant under a change of basis of the time series, and nonlinear wavelet thresholding is able to capture localized features in the spectral density matrix across frequency, always guaranteeing positive definite estimates. The finite-sample performance of intrinsic wavelet thresholding is assessed by means of simulated data and compared to several benchmark estimators in the Riemannian manifold. Further illustrations are provided by examining the multivariate spectra of trial-replicated brain signal time series recorded during a learning experiment.

February 28 (9:30–18:10)

Yuichi Goto

Title: Estimation of Trigonometric Moments for Circular Distribution of MA(p) Type by Using Binary Series

Abstract: Directional statistics have received a great deal of interest in recent years, and a variety of distributions on the circle have been proposed. In this talk, we propose circular distributions of a moving average model of order p type which includes the cardioid distribution, and discuss estimation of trigonometric moments based on binary series. We give an explicit form of the root n consistent

estimator based on clipped series, which enables us to construct an efficient estimator by the Newton--Raphson iterative method. We also show a robustness of the proposed estimator when the probability density function is contaminated with a noise term.

Kou Fujimori

Title: Moment convergence of the generalized maximum composite likelihood estimators for determinantal point processes

Abstract: The maximum composite likelihood estimator for parametric models of determinantal point processes (DPPs) is discussed. Since the joint intensities of these point processes are given by determinant of positive definite kernels, we have the explicit form of the joint intensities for every order. This fact enables us to consider the generalized maximum composite likelihood estimator for any order. In this talk, we introduce the two-step generalized composite likelihood estimator and shows the moment convergence of the estimator under a stationarity. Moreover, our results can yield information criteria for statistical model selection within DPPs. (Joint work with Sota Sakamoto and Yasutaka Shimizu)

Yuan Wang

Title: Group Analysis of Multi-Trial Brain Signals via Persistent Morse Homology

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.

Fumiya Akashi

Title: Robust regression on hyper-spheres with unspecified heteroscedastic errors and smooth approximation of object functions

Abstract: Statistical treatment for a random vector on a hyper-spheres attracts a lot attention recently, and has various applications such as seismic wave analysis, analysis for orientation of wild fire, etc. In this talk the nonlinear regression model whose predictor is a random vector on a hyper-sphere is considered. It is well known that the classical method in “linear statistic” does not work for spherical random vectors. To construct a robust estimator for the nonlinear regression function, this talk employees L1-regression method and

kernel-type objective function. The proposed local-linear estimator has asymptotic normality even if the error process has infinite variance, dependent structure or heteroscedasticity. The smooth approximation of the L1 objective function is also proposed. Some simulation experiments illustrate desired finite sample properties of the proposed method. (Joint work with Holger Dette)

Ying Chen

Title: Regularized Partially Functional Autoregressive Model

Abstract: We propose a partially functional autoregressive model (pFAR) to describe the dynamic evolution of serially correlated functional data. This model provides a unified framework to depict both the serial dependence on multiple lagged functional covariates and the associated relation with ultrahigh-dimensional exogenous scalar covariates. Estimation is conducted under a two-layer sparsity assumption, where only a small number of groups and elements are supposed to be active, yet their number and location are unknown in advance. We establish the asymptotic properties of the estimator and perform simulation studies to investigate its finite sample performance. We demonstrate the application of the pFAR model using daily natural gas flow curves data in the high pressure pipeline of German gas transmission network. The gas demand and supply are influenced by their historical values and 85 scalar covariates varying from price to temperature. The model provides insightful interpretation and good out-of-sample forecast accuracy compared to several popular alternative models. (Joint work with Thorsten Koch and Xiaofei Xu)

Michael Eichler

Title: Causal Inference from Multivariate Time Series: Principles and Problems

Abstract: In time series analysis, inference about cause-effect relationships among multiple time series is commonly based on the concept of Granger causality, which exploits temporal structure to achieve causal ordering of dependent variables. One major and well-known problem in the application of Granger causality for the identification of causal relationships is the possible presence of latent variables that affect the measured components and thus lead to so-called spurious causalities. This raises the question about whether Granger causality is an appropriate tool for causal learning; indeed, there are many researchers that deny any such claim. To answer the question in more depth, we present a graph-theoretic approach for describing and analysing Granger-causal relationships in multivariate time series that are possibly affected by latent variables. It is based on mixed graphs in which directed edges represent direct influences among the variables while dashed edges - directed or undirected - indicate associations that are induced by latent variables. We show how such representations can be used for inductive causal learning from time series and discuss the underlying assumptions and their implications for causal learning.

Christian Francq

Title: Count and duration time series with equal conditional stochastic and mean orders

Abstract: We consider a positive-valued time series whose conditional distribution has a time-varying mean, which may depend

on exogenous variables. The main applications concern count or duration data. Under a contraction condition on the mean function, it is shown that stationarity and ergodicity hold when the mean and stochastic orders of the conditional distribution are the same. The latter condition holds for the exponential family parametrized by the mean, but also for many other distributions. We also provide conditions for the existence of marginal moments and for the geometric decay of the beta-mixing coefficients. We give conditions for consistency and asymptotic normality of several estimators of the conditional mean parameters which do not require fully specifying the conditional distribution. We compare Quasi-Maximum Likelihood Estimators (QMLEs) (in particular the Poisson QMLE and the Exponential QMLE) and weighted least squares estimators. Simulation experiments and illustrations on series of stock market volumes and of greenhouse gas concentrations show that the multiplicative-error form of usual duration models deserves to be relaxed, as allowed in our approach. (Joint work with Abdelhakim Aknouche)

Elvezio Ronchetti

Title: Saddlepoint approximations for short and long memory time series: A frequency domain approach

Abstract: Saddlepoint techniques provide numerically accurate, small sample approximations to the distribution of estimators and test statistics. Except for a few simple models, these approximations are not available in the framework of stationary time series. We contribute to fill this gap. Under short or long range serial dependence, for

Gaussian and non Gaussian processes, we show how to derive and implement saddlepoint approximations for two relevant classes of frequency domain statistics: ratio statistics and Whittle's estimator. We compare our new approximations to the ones obtained by the standard asymptotic theory and by two widely-applied bootstrap methods. The numerical exercises for Whittle's estimator show that our approximations yield accuracy's improvements, while preserving analytical tractability. A real data example concludes the paper. (Joint work with Davide La Vecchia)

Anna Clara Monti

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.

Yuya Harada

Pianist Sayo Zenyoji

Title: The vocal performance of OPERA & Japanese artistic songs

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.

February 29 (9:30-11:50)

Koji Tsukuda

Title: A change detection procedure for an ergodic diffusion process

Abstract: The change point test for an ergodic diffusion process is considered under the setting of continuous observation. For this problem, there is an approach based on the weak convergence in sup-norm metric of a random process relating to an estimating equation. Asymptotic null distributions of some test statistics are established by using this weak convergence and the continuous mapping theorem. In this presentation, we propose a different test statistic and justify the proposed procedure through the weak convergence theory in $L^2(0,1)$.

David Preinerstorfer

Title: Functional sequential treatment allocation

Abstract: We consider a setting, in which a policy maker assigns subjects to treatments, observing each outcome before the next subject arrives. Initially, it is unknown which treatment is best, but the sequential nature of the problem permits learning about the effectiveness of the treatments. While the multi-armed-bandit literature has shed much light on the situation when the policy maker compares the effectiveness of the treatments through their mean, economic decision making often requires targeting purpose specific characteristics of the outcome distribution, such as its inherent degree of inequality, welfare or poverty. In this talk we introduce and study sequential learning algorithms when the distributional characteristic of interest is a general functional of the outcome distribution.

Mike K P So

Title: Bayesian Network Analysis of Systemic Risk in Financial Markets

Abstract: Analyzing systemic risk in financial markets has been an active research area in financial econometrics, risk management and big data analytics. This paper proposes an approach based on network analysis to study the interrelationship between financial companies. We develop statistical models to understand how the financial network, and thus systemic risk, changes over time. We adopt Bayesian inference methods to estimate the financial network, do network prediction and use listed companies in Hong Kong to illustrate our idea.