Waseda Seminar on High-dimensional Statistics

Date: November 16 (Wed.), 2022

Venue: Meeting Room, Building 63-1, Nishi-Waseda Campus, Waseda University

(Access map: https://www.waseda.jp/top/en/access/nishiwaseda-campus)

This seminar is supported by

JSPS KAKENHI Kiban (C): Grand-in-Aid for Scientific Research No.20K11719 (Yan Liu)

Waseda Research Institute for Science & Engineering, Institute for Mathematical Science

Program

14:00 ~ 14:45

Efficient Computational Strategies for Regression Model Selection (I)

Erricos J. Kontoghiorghes (Cyprus Univ. of Technology and Birkbeck & Univ. of London)

15:00 ~ 15:45

Efficient Computational Strategies for Regression Model Selection (II)

Erricos J. Kontoghiorghes (Cyprus Univ. of Technology and Birkbeck & Univ. of London)

16:00 ~ 16:45

The Dantzig Selector for Statistical Models of Stochastic Processes in High-dimensional and Sparse Settings (I)

Kou Fujimori (Shinshu Univ.)

17:00 ~ 17:45

The Dantzig Selector for Statistical Models of Stochastic Processes in High-dimensional and Sparse Settings (II)

Kou Fujimori (Shinshu Univ.)

18:30 ~

Welcome Party

Abstract

Efficient Computational Strategies for Regression Model Selection

Erricos J. Kontoghiorghes (Cyprus Univ. of Technology and Birkbeck & Univ. of London)

Abstract: Computational strategies for computing the best-subset regression models are proposed. The algorithms are based on a regression tree structure that generates all possible subset models. An efficient branch-and-bound algorithm that finds the best submodels without generating the entire tree is described. Specifically, the computational burden is reduced by pruning the non-optimal subtrees. The main numerical tool that has been employed is the QR factorization and its modification. This yields in a numerically stable and efficient sub-model estimation procedure. An algorithm that selects the best variable-subset model according to a pre-determined search criterion is presented. This algorithm performs considerably faster than all-subsets variable selection algorithms that rely on the residual sum of squares only. The case of high-dimensional data where the number of variables exceeds the number of observations is also considered. Within this context a theoretical combinatorial solution is proposed. (This is a joint work with Cristian Gatu, Marios Demosthenous, Ana Colubi.)

The Dantzig Selector for Statistical Models of Stochastic Processes in High-dimensional and Sparse Settings

Kou Fujimori (Shinshu Univ.)

Abstract: The Dantzig selector is an estimation procedure for regression models in high-dimensional and sparse settings. The Dantzig selectors for some statistical models of stochastic processes are discussed in this talk. We apply this procedure to Cox's proportional hazards model and some specific models of diffusion processes and prove the consistency and the variable selection consistency of the estimators. Based on partial likelihood and quasi-likelihood methods which were studied intensively in low-dimensional settings, we study these statistical models of stochastic processes in high-dimensional and sparse settings, which need some mathematically challenging tasks to prove the asymptotic properties of the estimators. The consistency of the estimators is derived from the stochastic maximal inequalities to deal with the curse of dimension and some matrix factors and conditions on Hessian matrices of likelihood functions to deal with the sparsity. We use exponential concentration inequalities for martingales and the maximal inequalities using Orlicz norm for the former problem and matrix conditions using restricted eigenvalue, compatibility factor and weak cone invertibility factor for the latter problem, which are known to be weaker conditions than others. We prove that consistency of the estimator implies the variable selection consistency which enables us to reduce the dimension. Using the dimension reduction, asymptotically normal estimators can be constructed.