

SCHOOL OF DATA SCIENCE

SEMINAR SERIES

Sorted Concave Penalized Regression

Date: 1 February 2019 (Friday)

Time: 2:30pm to 3:30pm

Venue: P7510, 7/F, Yeung Kin Man Academic Building (YEUNG),
City University of Hong Kong

Dr Feng, Long
Postdoctoral Associate
Department of Biostatistics
Yale University



Guest Speaker's profile

Long Feng is currently a Postdoctoral Associate in the Department of Biostatistics at Yale University. His research interests focus on high-dimensional data, imaging data, and empirical Bayes methods. Before coming to Yale, he obtained his Ph.D in Statistics from Rutgers University.

Abstract

The Lasso is biased. Concave penalized least squares estimation (PLSE) takes advantage of signal strength to reduce this bias, leading to sharper error bounds in prediction, coefficient estimation and variable selection. For prediction and estimation, the bias of the Lasso can be also reduced by taking a smaller penalty level than what selection consistency requires, but such smaller penalty level depends on the sparsity of the true coefficient vector. The sorted L1 penalized estimation (Slope) was proposed for adaptation to such smaller penalty levels. However, the advantages of concave PLSE and Slope do not subsume each other. We propose sorted concave penalized estimation to combine the advantages of concave and sorted penalizations. We prove that sorted concave penalties adaptively choose the smaller penalty level and at the same time benefits from signal strength, especially when a significant proportion of signals are stronger than the corresponding adaptively selected penalty levels. A local convex approximation, which extends the local linear and quadratic approximations to sorted concave penalties, is developed to facilitate the computation of sorted concave PLSE and proven to possess desired prediction and estimation error bounds. We carry out a unified treatment of penalty functions in a general optimization setting, including the penalty levels and concavity of the above mentioned sorted penalties and mixed penalties motivated by Bayesian considerations. Our analysis of prediction and estimation errors requires the restricted eigenvalue condition on the design, not beyond, and provides selection consistency under a required minimum signal strength condition in addition. Thus, our results also sharpens existing results on concave PLSE by removing the upper sparse eigenvalue component of the sparse Riesz condition.