psd

DISSERTATION PRESENTATION AND DEFENSE

Statistical Methods For Change Point Problems and Bayesian Uncertainty Quantification

WHEN
April 20, 2022
10:00 AM

WHERE
Jones Laboratory, Room 304

Lizhen Nie, PhD candidate

This dissertation considers two topics. Driven by numerous real applications, the first topic is on change point problems, which deals with the testing and localization of change point(s) in a sequence of observations. We focus on the unsupervised, offline, nonparametric setting. We use a new kernel/distance-based framework to propose new statistics for change point problems. This framework has several the benefits including that it applies to any observation lying in general metric spaces, and its lack of parametric assumptions. We study three change-point sub-problems under this framework. First, we start with the classical setting where we analyze the detection and localizing of \emph{abrupt} changes. Theoretical guarantees on false positive rate, power consistency under local alternatives, and the minimax localization rate are derived. The connection between the proposed framework and existing nonparametric methods is also investigated. We then consider a more realistic setting where the change is \emph{gradual} rather than abrupt. We discuss limitations of existing procedures designed for abrupt change points, and propose a new method that is designed to address the challenges in the gradual change point problem. Theoretical guarantees are provided for the proposed method. Finally, we study the much less investigated topic of changes in the relationship of a \emph{pair} of observations. Different from traditional settings, we focus on changes in the \emph{conditional} distribution rather than joint or marginal distributions. Several methods are investigated under this setting, and we illustrate through empirical studies the importance of separating this problem from the unpaired problem, and the superiority of proposed algorithms.

The second topic of this dissertation is on Bayesian uncertainty quantification, where we focus on improving the scalability of traditional methods. We explore the possibility of sampling from an approximate posterior by performing maximum a posteriori optimization on many independently perturbed datasets. This method turns fast optimization into approximate posterior sampling. Beyond its scalability, theoretical support is also provided. We compare our algorithm to the traditional Markov Chain Monte Carlo as well as many state-of-the-art methods. We show, both in simulations and on real data, that our method fares very well in these comparisons, often providing substantial computational gains. The second project puts the previous idea in a general framework of learning approximate loss-driven posteriors from bootstrap approaches, where we investigate the possibility of utilizing generative networks to learn a mapping from bootstrap weights to optimized sample points, which is free from solving optimization problems and further improves the scalability.

stat.uchicago.edu

THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS