RUTGERS
School of Arts and Sciences

Department of Statistics

Xianyang Zhang
Statistics Department
Texas A&M University

Change-point Detection: Computation and Statistical Inference

Wednesday, November 16, 2022
11:50 AM

110 Frelinghuysen Road, Hill Center, Room 552

Zoom Meeting: Meeting ID: 99075124232
Password: 952486

https://rutgers.zoom.us/j/99075124232?pwd=UDdPVjRncXZFcXpvbFE0OWJyMVdSUT09

Abstract: Change-point analysis is concerned with detecting and locating structure breaks in the underlying model of a data sequence. It finds an abundance of applications in a wide variety of fields, for example, bioinformatics, finance, and engineering. This talk provides an overview of two different change-point detection frameworks in the literature. The first approach is based on minimizing a cost function over possible numbers and locations of change points. Such an approach requires finding the cost value repeatedly over different segments of the data set, which can be time-consuming. To tackle this issue, we introduce a new method based on sequential gradient descent to find the cost value accurately and efficiently. The core idea is to update the cost value using the information from previous steps without re-optimizing the objective function. Numerical studies show that the new approach can be orders of magnitude faster than the Pruned Exact Linear Time method without sacrificing estimation accuracy. The second approach combines two-sample hypothesis testing with segmentation techniques. A particular challenge within this framework is dealing with the high-dimensionality of data and the nonparametric nature of structure break. We develop a new methodology to detect structural breaks in the distributions of a sequence of high-dimensional observations. We show that the new approach is more efficient than the existing methods.

Bio: Xianyang Zhang is an Associate Professor in the statistics department at Texas A&M University. He obtained his Ph.D. in statistics from the University of Illinois at Urbana Champaign in 2013. His research interests include high dimensional/large-scale statistical inference, kernel methods, genomics data analysis, functional data analysis, time series, and econometrics.