Statistical Inference for Change Points in High-Dimensional Data

Estimation and testing of change points in high-dimensional data have wide applications in many disciplines, such as biological science, economics and finance. In this talk, we introduce a new U-statistic based approach to both problems and show its advantage over several existing methods via theory and simulations. The talk consists of two parts.

In the first part, we will introduce a new test based on U-statistics for testing a mean shift in high-dimensional data. The test aims to detect dense alternatives and is tuning parameter free. At the core of our theory, we show weak convergence of a sequential U-statistic based process, and derive the limiting distribution under both the null and alternatives.

In the second part, we will discuss a change point location estimator which maximizes a new U-statistic based objective function. Under mild and easily interpretable assumptions, we derive its convergence rate and asymptotic distribution after suitable centering and normalization. A comparison with the popular least squares based approach illustrates the theoretical advantage of ours. A bootstrap-based approach is also proposed to construct a confidence interval with accurate coverage, which is corroborated by simulation results.