Statistics and Data Science Seminar

Subspace learning for high dimensional tensor data

Yuefeng Han (Rutgers University)

Abstract: Motivated by modern scientific research, analysis of tensors (multi-dimensional arrays) has emerged as one of the most important and active areas in modern statistics and data science. High-dimensional tensor data routinely arise in a wide range of applications, such as economics, genetics, microbiome studies, brain imaging, and hyperspectral imaging, due to modern data collection capabilities. In many of these settings, the observed tensors are of high dimension and high order, but the important information may lie in dimension-reduced subspaces induced by various structural conditions. This talk aims to develop new methodologies and theories from a perspective of subspace learning. The talk is divided into two parts. In the first part, we introduce a factor approach for analyzing high dimensional dynamic tensors, in a form similar to Tucker tensor decomposition. We propose two estimation methods that are based on the tensor unfolding of lagged cross-product and iterative orthogonal projections of the original dynamic tensors. We also establish computational and statistical guarantees of the proposed methods. In the second part, we investigate a tensor factor model with a CP type low-rank tensor structure. We develop a new computationally efficient estimation procedure, which includes a warm-start initialization and an iterative concurrent orthogonalization scheme. We show that the iterative algorithm achieves ϵ-accuracy guarantee within $\log\log(1/\epsilon)$ number of iterations.

Wednesday, November 17 at 4:00 PM in Zoom