

Yun Wei

Postdoctoral Fellow The Statistical and Applied Mathematical Sciences Institute Duke University

Parameter Estimations in Finite Mixture Models

Abstract. Finite mixture models are among the most widely used models to address data heterogeneity. Recent work [Henrich & Kahn 2018] establishes the optimal uniform convergence rate (in the minimax sense) using minimum Kolmogorov-Smirnov distance estimators. Subsequently, the work Wu & Yang 2020 proves that the method of moments also achieves the optimal uniform convergence rate for Gaussian mixture models. We propose a general framework and estimator that includes the two previous methods as special cases and establish the convergence rate under the general framework. The general framework can generate novel methods, and one instance is based on the maximum mean discrepancy, which is also shown to achieve optimal uniform convergence rate. Pointwise convergence rates are also established under the general framework. In the second part of the talk, we consider the finite mixture of product distribution with the special structure that the product distributions in each mixture component are also identically distributed. In this setup, each mixture component consists of samples from repeated measurements, making such data exchangeable sequences. Applications of the model include psychological studies and topic modeling. We show that with sufficient repeated measurements, a model that is not originally identifiable becomes identifiable. The posterior contraction rate for parameter estimation is also obtained, showing that repeated measurements are beneficial for estimating parameters in each mixture component. These results hold for general probability kernels, including all regular exponential families, and can be applied to hierarchical models. Based on joint work with Xuanlong Nguyen and Sayan Mukherjee.

Biographical Sketch. Dr. Yun Wei received the B.S. from Nankai University and the Ph.D from University of Michigan in applied and interdisciplinary mathematics (with concentration in statistics). He is currently a postdoc at the Department of Statistical Science at Duke University. His primary research interests are latent variable models, high dimensional statistics, machine learning, information theory, and optimization.

¹Zoom meeting ID: 96753797217 password: mathstat