The Hong Kong Polytechnic University
Department of Applied Mathematics

Statistics and Data Science Online Colloquium Series

Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models

By

Dr. Yang FENG
New York University

Abstract

Unsupervised learning has been widely used in many real-world applications. One of the simplest and most important unsupervised learning models is the Gaussian mixture model (GMM). In this work, we study the multi-task learning problem on GMMs, which aims to leverage potentially similar GMM parameter structures among tasks to obtain improved learning performance compared to single-task learning. We propose a multi-task GMM learning procedure based on the EM algorithm that not only can effectively utilize unknown similarities between related tasks but is also robust against a fraction of outlier tasks from arbitrary sources. The proposed procedure is shown to achieve the minimax optimal rate of convergence for both parameter estimation error and the excess mis-clustering error, in a wide range of regimes. Moreover, we generalize our approach to tackle the problem of transfer learning for GMMs, where similar theoretical results are derived. Finally, we implement the algorithms in a new R package mtlgmm and demonstrate the effectiveness of our methods through simulations and a real data analysis. To the best of our knowledge, this is the first work studying multi-task and transfer learning on GMMs with theoretical guarantees.

Biograph

Yang Feng is an Associate Professor of Biostatistics at New York University. He got his Ph.D. in Operations Research from Princeton University in 2010. His research interests include machine learning, high-dimensional statistics, social networks, nonparametric statistics, and bioinformatics. Dr. Feng has published over 50 articles in leading statistics and machine learning journals. He is currently an associate editor of the Journal of the American Statistical Association, Journal of Business and Economic Statistics, Statistica Sinica, and Stat. He is a fellow of the American Statistical Association and an elected member of the International Statistical Institute. His research has been supported by multiple NSF and NIH grants.

Date: 9 December 2022 (Friday)
Time: 10:30-11:30 (Hong Kong Standard Time GMT +8)
Venue: Online Talk via Zoom (Meeting ID: 965 2989 1545)
Speaker: Dr. Yang Feng, New York University
Host: Dr. Ting Li, The Hong Kong Polytechnic University
Click to join:
https://polyu.zoom.us/j/96529891545?pwd=VXdWNXluMlBaNTFuS295Tnp2TkZndz09

Click to join (Zoom)

* * * ALL ARE WELCOME * * *