ABSTRACT
Estimation of mixture models, a fundamental problem itself, is an example of machine learning and statistical problems where nonconvexity arises due to superposition and composition of simpler models. This problem may possess spurious local minimizers that are not globally optimal, both in theory and in practice. In this talk, we will present several positive results for mixture problems in the lens of nonconvex optimization.

For a mixture of two components, we show that all local minima are globally optimal, and EM converges to a global optimum from random initialization. This result applies to a broad class of log-concave mixture problems including mixtures of Gaussians and linear regressions. Our analysis highlights the challenge due non-contraction in Euclidean distance, and the role of angle contraction.

For more than two components, spurious local minima provably exist. One such local minimizer puts multiple centers at one true component, and another center in the middle of multiple true components. We prove that this is essentially the only type of spurious local minima under certain conditions. Consequently, standard algorithms such as gradient descent and EM are guaranteed to converge to a solution that partially identifies the true model.

Leveraging the above structural results, we will discuss several algorithmic ideas for avoiding local minima and finding the global optimum. Over-parametrization, in which one over-specifies the number of components, stands out as a promising approach. We will also briefly discuss the SDP relaxation approach, an extreme form of over-parametrization.

BIO
Yudong Chen is an assistant professor at the School of Operations Research and Information Engineering, Cornell University. Before joining Cornell, he was a postdoctoral scholar at the Department of Electrical Engineering and Computer Sciences at University of California, Berkeley. He obtained his Ph.D. in Electrical and Computer Engineering from the University of Texas at Austin, and his M.S. and B.S. from Tsinghua University. His research interests include machine learning, high-dimensional statistics, convex and nonconvex optimization, and their applications in financial, communication and computer networks. His work has been recognized by an NSF CAREER award and the Second Place of INFORMS Nicholson Student Paper Competition.

DETAILS
Monday, May 10, 2021 at 12 noon CST.