Stochastic Compositional Optimization for Machine Learning

Abstract: Optimizing functions which could be expressed as nested compositions of T functions in the stochastic setting arise naturally in several modern machine learning problems. Examples include distributionally robust optimization, adversarial training, classification with imbalanced datasets, model agnostic meta-learning and training graph neural networks. In this talk, I will discuss our recent works on developing efficient algorithms for solving the above class of stochastic optimization problems and its variants. First, I will present an algorithm, based on moving-average and linearization techniques, which has an optimal level-independent sample complexity of O_T(e^-4) to obtain an e-approximate first-order stationary solution under appropriate assumptions. The algorithm does not use mini-batches (and is hence applicable to the purely online setting), and is parameter-free. Second, I will present a projection-free version of the algorithm with similar sample complexities, suitable for solving constrained compositional optimization problems. Finally, I will introduce a novel bi-level version of stochastic compositional optimization problem and present an algorithm, along with the corresponding oracle complexity results, for efficiently solving this class of problems. This bi-level formulation will be used to conceptually and numerically illustrate the robustness benefits exhibited by deeper neural networks.

Bio: Krishna Balasubramanian is an assistant professor in the Department of Statistics, University of California, Davis. He is also affiliated with the Graduate Group in Applied Mathematics, the Center for Data Science and Artificial Intelligence Research (CeDAR) and the TETRAPODS Institute of Data Science at UC Davis. He was a visiting scientist at the Simons Institute for the Theory of Computing, UC Berkeley in Fall 2021 and 2022. He completed his PhD in Computer Science from Georgia Institute of Technology, and was a postdoctoral researcher in the Department of Operations Research and Financial Engineering, Princeton University, and the Department of Statistics at UW-Madison. Krishna's research interests include stochastic optimization and sampling, theoretical machine learning, and geometric and topological statistics. His research was/is supported by a Facebook PhD fellowship, and CeDAR and NSF grants. He serves as an editorial board reviewer for the Journal of Machine Learning Research and as a senior area chair for the International Conference on Machine Learning (ICML), Advances in Neural Information Processing Systems (NeurIPS) and Conference on Learning Theory (COLT).

TUESDAY, JANUARY 31, 2023 from 3:30 pm – 4:30 pm, GER 206