SCHOOL OF DATA SCIENCE

SEMINAR SERIES

Learn Policy Optimally via Efficiently Utilizing Data

Date:	28 January 2019 (Monday)
Time:	2:30pm to 3:30pm
Venue:	P7510, 7/F, Yeung Kin Man Academic Building (YEUNG), City University of Hong Kong

Dr. Yang, Lin
Postdoctoral Researcher
Princeton University

Guest Speaker's profile

Lin Yang is currently a postdoctoral researcher at Princeton University working with Prof. Mengdi Wang. He obtained two Ph.D. degrees simultaneously in Computer Science and in Physics & Astronomy from Johns Hopkins University. Prior to that, he obtained a bachelor's degree from Tsinghua University. His research focuses on developing and applying fast algorithms for large-scale optimization and machine learning. This includes reinforcement learning and steaming methods for optimization and function approximations. His algorithms have been applied to real-world applications including accelerating astrophysical discoveries and improving network security. He published numerous papers in top Computer Science conferences including NeurIPS, ICML, STOC, and PODS. At Johns Hopkins, he was a recipient of the Dean Robert H. Roy Fellowship.

Abstract

Recent years have witnessed increasing empirical successes in reinforcement learning. In contrast, many theoretical problems about reinforcement learning are not well understood even in the most basic setting. For instance, what is the best way to learn an optimal policy for a finite-state Markov decision process from sample transitions and what is its sample complexity? Traditional methods usually use un-necessarily many samples and may be inefficient.
Suppose there is a generative model that allows one to sample state-transitions. We develop a novel algorithm that learns an approximate-optimal policy in near-optimal time and using a minimal number of samples. In particular, our result resolves the long-standing open problem about the sample complexity of Markov decision process. The algorithm integrates the value iteration and variance reduction techniques, and its analysis takes full advantages of monotonicity, contractiveness of the Bellman operator as well as the law of total variance of Markov processes. It makes updates by processing samples in a "streaming" fashion, which requires small memory and is naturally amenable to problems with large-scale data.

The algorithm and analysis can be extended to solve two-person Markov games and contextual Markov decision problems while achieving near-optimal sample complexity. It provides new insights about how to use data efficiently in learning and optimization. I further illustrate several other examples of learning and optimization over streaming data, with applications in accelerating Astrophysical discoveries and improving network securities.

Enquiries: 3442 7887

All are welcome