THE UNIVERSITY OF CHICAGO

Department of Statistics
STATISTICS COLLOQUIUM

JOINT SEMINAR WITH THE DEPARTMENT OF COMPUTER SCIENCE

ZHUORAN YANG

Department of Operations Research and Financial Engineering
Princeton University

Demystifying (Deep) Reinforcement Learning: The Pessimist, The Optimist, and Their Provable Efficiency

MONDAY, March 8, 2021 at 2:00 PM
Via Zoom (session information will be e-mailed to subscribers)

ABSTRACT

Coupled with powerful function approximators such as deep neural networks, reinforcement learning (RL) achieves tremendous empirical successes. However, its theoretical understandings lag behind. In particular, it remains unclear how to provably attain the optimal policy with a finite regret or sample complexity. In this talk, we will present the two sides of the same coin, which demonstrates an intriguing duality between pessimism and optimism.

In the offline setting, we aim to learn the optimal policy based on a dataset collected a priori. Due to a lack of active interactions with the environment, we suffer from the insufficient coverage of the dataset. To maximally exploit the dataset, we propose a pessimistic least-squares value iteration algorithm, which achieves a minimax-optimal sample complexity.

In the online setting, we aim to learn the optimal policy by actively interacting with an environment. To strike a balance between exploration and exploitation, we propose an

---

optimistic least-squares value iteration algorithm, which achieves a \sqrt{T} regret in the presence of linear, kernel, and neural function approximators.