Numerical Analysis and Scientific Computing Seminar

Finite-Time Performance of Distributed Temporal Difference Learning on Multi-Agent Reinforcement Learning

Thinh T. Doan

Georgia Institute of Technology

Abstract: The rapid development of low-cost sensors, smart devices, communication networks, and learning algorithms has enabled data driven decision making in large-scale multi-agent systems. Prominent examples include mobile robotic networks and autonomous systems. The key challenge in these systems is in handling the vast quantities of information shared between the agents in order to find an optimal policy that maximizes an objective function. Among potential approaches, distributed reinforcement learning, which is not only amenable to low-cost implementation but can also be implemented in real time, has been recognized as an important approach to address this challenge.

The focus of this talk is to consider the policy evaluation problem in multi-agent reinforcement learning, one of the most fundamental problems in this area. In this problem, a group of agents operate in an unknown environment, where their goal is to cooperatively evaluate the global discounted accumulative reward composed of local rewards observed by the agents. For solving this problem, I consider a distributed variant of the popular temporal difference learning, often referred to as TD(λ) for some constant $\lambda \in [0,1]$. My main contribution is to provide a finite-analysis on the performance of this distributed TD(λ) for both constant and time-varying step sizes. The key techniques are to utilize tools from distributed optimization and stochastic approximation in analyzing the underlying algorithm. In particular, I derive an explicit formula for the upper bound on the rates of the proposed method as a function of the constant λ and the network topology characterized the communication between the agents. In addition, my results theoretically address an important question of TD learning from numerical observations, that is, $\lambda = 1$ gives the best approximation of the function values while $\lambda = 0$ leads to better performance when there is a large variance in the algorithm. Finally, I conclude my talk with some discussion about my future research in the context of distributed decision making on multi-agent systems..

Friday, November 8, 2019, 2:00 pm

Mathematics and Science Center: MSC W303

Mathematics
Emory University