Exploration-exploitation trade-off for continuous-time reinforcement learning

Yufei Zhang
(LSE)

21 June 2023, **5:30 - 6:30pm** c.t. Berlin time
IRTG 2544: "Stochastic Analysis in Interaction"
— Berlin Probability Colloquium —

Abstract

Recently, reinforcement learning (RL) has attracted substantial research interests. Much of the attention and success, however, has been for the discrete-time setting. Continuous-time RL, despite its natural analytical connection to stochastic controls, has been largely unexplored and with limited progress. In particular, characterising sample efficiency for continuous-time RL algorithms remains a challenging and open problem.

In this talk, we develop a framework to analyse model-based reinforcement learning in the episodic setting. We then apply it to optimise exploration-exploitation trade-off for linear-convex RL problems, and report sublinear (or even logarithmic) regret bounds for a class of learning algorithms inspired by filtering theory. The approach is probabilistic, involving analysing learning efficiency using concentration inequalities for correlated continuous-time observations, and applying stochastic control theory to quantify the performance gap between applying greedy policies derived from estimated and true models.

TU Berlin
Institut für Mathematik
Straße des 17. Juni 136
10623 Berlin
MA 041