Agnostic Reinforcement Learning

The goal of this collaboration was to improve our understanding of what are the fundamental limits of learning to control complex, high dimensional dynamical systems. In particular, we sought to answer questions such as: is it possible to control in domains where the dynamics cannot be learned to uniform precision? What are the fundamental problem quantities which determine whether learning to control is possible?

To this end, we considered these questions in the context of adaptive control of the linear quadratic regulator (LQR) , perhaps one of the most basic and influential optimal control problems which has been studied extensively since the mid 20th century. While this model has recently received significant attention from the learning theory community, we surprisingly did not have a precise understanding of what are the fundamental quantities which determine the hardness of learning to control. In particular, previous work had shown regret bounds which scale polynomially in the ambient dimension of the state and control input. However, it was previously unknown whether these measures of problem complexity were sharp in general.

In our work, we establish the first sublinear regret bounds for LQR based on intrinsic dimension. We replace dependence on the ambient dimension of the system state with more natural problem quantities which are determined by properties of the process noise covariance. These are the first regret bounds for LQR which hold for infinite dimensional systems. Furthermore, while our results establish that it is possible to learn to control linear systems with infinite dimensional states, we also show that it is impossible to achieve sublinear regret if the input dimension is unbounded. In short, our results illustrate how learning to control is possible only for systems that have relatively simple action spaces.

These results are described in our paper, "Towards a Dimension-Free Understanding of Adaptive Linear Control" which was accepted for publication at the 34th Conference of Learning Theory. A talk about our work may be found here. Looking forward, we aim to ask similar questions regarding the fundamental limits of learning to control in other domains.