

Sean Meyn

The Projected Bellman Equation in Reinforcement Learning

A topic of discussion throughout the 2020 Simons program on reinforcement learning: is the Q-learning algorithm convergent outside of the tabular setting? It is now known that stability can be assured using a matrix gain algorithm, but this requires assumptions, which begs the next question: does a solution to the projected Bellman equation exist? This is the most minimal requirement for convergence of any algorithm.

The question was resolved in very recent work. A solution does exist, subject to two assumptions: the function class is linear, and (far more crucial) the input used for training is a form of epsilon-greedy policy with sufficiently small epsilon. Moreover, under these conditions it is shown that the Q-learning algorithm is stable, in terms of bounded parameter estimates. Convergence remains one of many open topics for research.

In short, sufficient optimism is not only valuable for algorithmic efficiency, but is a means to algorithmic stability.