

A Doubly Robust Approach to Sparse Reinforcement Learning

Tuesday, June 18, 2024 2:30 PM (30 minutes)

We propose a new regret minimization algorithm for episodic sparse linear Markov decision process (SMDP) where the state-transition distribution is a linear function of observed features.

The only previously known algorithm for SMDP requires the knowledge of the sparsity parameter and oracle access to a reference policy.

We overcome these limitations by combining the doubly robust method that allows one to use feature vectors of *all* actions with a novel analysis technique that enables the algorithm to use data from all periods in all episodes. This algorithm is shown to achieve best possible regret (up to log factors)

Primary author: ZEEVI, Assaf (columbia university)

Presenter: ZEEVI, Assaf (columbia university)

Session Classification: Parallel session: Reinforcement learning in MDPs with large state spaces