

Identifiability in Inverse reinforcement learning

Samuel Cohen
(University of Oxford)

16 February 2022, 5:15 - 6:15 pm Berlin time
IRTG 2544: "Stochastic Analysis in Interaction"
— Berlin Probability Colloquium —

Abstract

Inverse reinforcement learning attempts to reconstruct the reward function in a Markov decision problem, using observations of agent actions. This problem is ill-posed, and the reward function is not identifiable, even under the presence of perfect information about optimal behavior. We provide a resolution to this non-identifiability for problems with entropy regularization. For a given discrete-time finite-state MDP, we fully characterize the reward functions leading to a given policy and demonstrate that, given demonstrations of actions for the same reward under two distinct discount factors, or under sufficiently different environments, the unobserved reward can be recovered up to a constant. We also give general necessary and sufficient conditions for reconstruction of time-homogeneous rewards on finite horizons, and for action-independent rewards, generalizing recent results of Kim et al. [2021] and Fu et al. [2018].