CCIT Report No. 328

The Empirical Bayes Envelope and Regret Minimization in Stochastic Games

S. Manor and N. Shimkin

ABSTRACT

In repeated matrix games, classical results establish the existence of *regret minimizing* strategies that secure an average payoff as high as the best-response payoff (the *Bayes Envelope*) relative to the empirical frequencies of the opponent's actions. In this paper we consider the question of regret minimization in the framework of stochastic (Markov) games, under appropriate recurrence properties. Our approach relies on the observed state-action frequencies. The *empirical Bayes envelope* is naturally defined in this space as the best-response to the stationary strategy of the opponent which agrees with the observed frequencies. This envelope is shown to be non-attainable in general, and we therefore put forward its convexification (the CBE) as a feasible goal. The CBE is shown to be attainable, it guarantees a payoff higher than the minimax value when the opponents play is suboptimal, and reduces to the standard Bayes envelope for repeated matrix games. We further show that if the opponent alone affects the state transitions, then the empirical Bayes envelope is itself convex and attainable. Some concrete examples from information and learning theory are shown to fit our framework, as well as a k-th order extension of the Bayes envelope for repeated matrix games.

The analysis is based on a generalization of Blackwell's approachability theory to stochastic games, which holds under the recurrence assumption related to the existence of a single state which is recurrent under all strategies. Extension of the results to games that satisfy communicating-type conditions are outlined.