

Abstract:

Different types of dynamics have been studied in repeated game play, and one of them which has received much attention recently consists of those based on "no-regret" algorithms from the area of machine learning. Plenty of work has been done showing that when using a specific no-regret algorithm, convergence of plays to equilibria can be shown and better quality of outcomes in terms of the price of anarchy can be reached in some classes of games. We would like to cover two main types of algorithms, namely gradient-based and best-response-based dynamics, played in their corresponding classes of games and their convergences to equilibria in this talk.

We show that when employing the mirror-descent algorithm, a well-known generic no-regret algorithm, the actual plays converge quickly to equilibria in convex potential games. The bandit model considers a probably more realistic and prevalent setting with only partial information, in which at each time step each player only knows the cost of her own currently played strategy, but not any costs of unplayed strategies. We propose a family of bandit algorithms based on the mirror-descent algorithms previously presented, and show that when each agent individually adopts such a bandit algorithm, their joint mixed strategy profile quickly converges with implications.

In a competitive pricing game for multiple agents over a planning horizon of periods, at each period agents set their selling prices and receive stochastic demand from buyers. Agents do not know their underlying demand curves, but they wish to determine the selling prices to maximize total revenue under competition. We design a learning algorithm to facilitate coordinated dynamic pricing, in which agents estimate their demand functions based on observations and adjust their best-response pricing decisions. We show that the joint pricing decisions, determined by estimated demand functions, converge to underlying equilibria. We also show the no-regret property of our algorithm to provide long-term incentives of adopting the algorithm for agents.