Constant Step-Size Multiplicative Weight Update in Potential Games

• **Maurizio D'Andrea** (Toulouse School of Economics, UT capitole and ANITI)[1]

Keywords: Repeated Games,MWU Algorithms, Markov Chain, Nash Equilibrium.

Abstract: It is generally recognized that computing Nash equilibium may be very complex, so over the years we wondered about the possibility of obtaining these outcomes through dynamic processes where players adapt their actions over time. In this paper we focus on n-player potential games repeated over time where each agent competes against the others following the multiplicative weights update (MWU) with constant step size. The MWU algorithms are a class of algorithms largely studied in literature in particular in online learning and decision making. In the MWU version that we consider, at each step all players select an action based on their mixed strategies, then they receive information about their vector payoff and they update their mixed strategies.

The goal is to study the behavior of the players induced strategies in the long run. In their paper [1] Mertikopoulos et al. study this problem when the step size vanishes. Indeed in this case the mixed strategy sequence generated by the MWU algorithm can be viewed as a "Robbins–Monro approximation" (an asymptotic pseudo trajectory to be precise) of replicator dynamics. Hence it is possible to use ordinary differential equation theory and in particular the fact that the potential function is a Lyapunov function of the replicator dynamic to prove that in generic potential games the actual sequence of play converges almost surely to a Nash equlibrium. In the constant step size version is not possible to bring the sequence of strategies back to the replicator dynamics. However we can prove that if the sequence of strategies converges almost surely then it converges to a mixed Nash equilibrium. Moreover this equilibrium is weakly stable in the sense that if fixing one of the players to choosing a pure strategy in the support of her strategy with probability one, leaves the other players indifferent between the strategies in their support. This result already allows us to see in which classes of games it is not possible to obtain a convergence of the sequence of strategies. For example we can not obtain the convergence in the class of zero-sum games because there exist zero-sum games without weekly stable Nash equilibrium (Matching Pennies). precisely in this sense Bailey et al. in a recent paper [2] have studied the behavior of the strategies induced by the MWU in the case of zero-sum games proving that they converge to the boundary of the strategies set. The first step of this work is to prove that the same result is true also for generic potential games. In order to do this we note that the dynamic induced by the MWU is an homogeneous Markov chain, then using standard Markov chains techniques we prove that under certain hypotheses the sequence of the strategies converges to the boundary. This result in the simple case of 2-player games with two actions for each player guarantees convergence to the pure Nash equilibrium. Therefore the final goal is to extend these result to obtain convergence in a generic two-player games.

References:

[1] J. Cohen, A. Héliou, P Mertikopoulos. Learning with Bandit Feedback in Potential Games. *31st Conference on Neural Information Processing Systems, Long Beach, CA, USA.*, 2017.

[2] J. P. Bailey, S. G. Nagarajan, G. Piliouras. Stochastic Multiplicative Weights Updates in Zero-Sum Games. *https://arxiv.org/abs/2110.02134*, 2021.

[3] J. R. Norris. Markov Chain. *Cambridge University Press*, 1997.

[1] Phd work supervised by Jérôme Renault and Fabien Gensbittel.