Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Economic Theory 148 (2013) 825–842
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Notes

No-regret dynamics and fictitious play ☆

Yannick Viossat [a], Andriy Zapechelnyuk [b,*]

[a] *CEREMADE, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, F-75775 Paris, France*
[b] *School of Economics and Finance, Queen Mary, University of London, Mile End Road, London E1 4NS, UK*

Received 1 October 2009; final version received 27 July 2012; accepted 31 July 2012

Available online 9 January 2013

Abstract

Potential based no-regret dynamics are shown to be related to fictitious play. Roughly, these are ε-best reply dynamics where ε is the maximal regret, which vanishes with time. This allows for alternative and sometimes much shorter proofs of known results on convergence of no-regret dynamics to the set of Nash equilibria.

JEL classification: C73; D81; D83

Keywords: Regret minimization; No-regret strategy; Fictitious play; Best-reply dynamics; Nash equilibrium; Hannan set; Curb set

1. Introduction

No-regret strategies are simple adaptive learning rules that recently received a lot of attention in the literature.[1] In a repeated game, a player has a *regret* for an action if, loosely speaking, she

☆ The authors thank William Sandholm whose comments led to a substantial improvement of the paper, as well as Mathieu Faure, Sergiu Hart, Josef Hofbauer, Alexander Matros, Karl Schlag, Eilon Solan and Sylvain Sorin for helpful comments and suggestions. Y. Viossat gratefully acknowledges the support of ANR BLAN RISK.

* Corresponding author.
E-mail addresses: viossat@ceremade.dauphine.fr (Y. Viossat), a.zapechelnyuk@qmul.ac.uk (A. Zapechelnyuk).

[1] These rules have been used to investigate convergence to equilibria in the context of learning in games [19,20,24–26], for combining different forecasts [17,18] (for an overview of the forecast combination literature see [14,43]) and for combining opinions, which is also of interest to management science [33]. In finance this method has been used to derive bounds on the prices of financial instruments [13,15]. This method can be applied to various tasks in computer science, such as job scheduling [36] and routing [9] (for a survey of applicable problems in computer science see [31]).

http://dx.doi.org/10.1016/j.jet.2012.07.003

could have obtained a greater average payoff had she played that action more often in the past. In the course of the game, the player reinforces actions that she regrets not having played enough, for instance, by choosing next action with probability proportional to the regret for that action, as in Hart and Mas-Colell's [24] *regret-matching* rule. Existence of *no-regret strategies* (i.e., strategies that guarantee no regrets almost surely in the long run) is known since Hannan [23]; wide classes of no-regret strategies are identified by Hart and Mas-Colell [25] and Cesa-Bianchi and Lugosi [11].[2]

A *no-regret dynamics* is a stochastic process that describes trajectories of the average correlated play of players and that emerges when every player follows a no-regret strategy (different players may play different strategies). By definition, it converges to the Hannan set (the set of all correlated actions that satisfy the no-regret condition first stated by Hannan [23]).[3] This set is typically large. It contains the set of correlated equilibria of the game and we show that it may even contain correlated actions that put positive weight *only* on strictly dominated actions. Thus convergence of the average play to the Hannan set often provides very little information about what the players will actually play, as it does not even imply exclusion of strictly dominated actions.

In this paper we show that no-regret dynamics are intimately linked to the classical fictitious play process [10]. Drawing on Monderer et al. [38], we first show that contrary to the standard, discrete-time version, continuous fictitious play leads to no regret. We then show that, for a large class of no-regret dynamics, if a player's maximal regret is $\varepsilon > 0$, then she plays an ε-best reply to the average correlated play of the others. Since in this class the maximal regret vanishes (see Corollary 1 below), it follows that, for a good choice of behavior when all regrets are negative, the dynamics is a vanishingly perturbed version of fictitious play.

For two-player finite games, this observation and the theory of perturbed differential inclusions [4,5] allow us to relate formally the asymptotic behavior of no-regret dynamics and of continuous fictitious play (or its time-rescaled version, the best-reply dynamics [22]). In classes of games in which the behavior of continuous fictitious play is well known, this provides substantial information on the asymptotic behavior of no-regret dynamics. In particular, we recover most known convergence properties of no-regret dynamics. Our results do not just allow us to find new and sometimes much shorter proofs of convergence of no-regret dynamics towards the set of Nash equilibria in some classes of games, such as dominance solvable games or potential games. They also allow us to relate the asymptotic behavior of no-regret dynamics and continuous fictitious play in case of divergence, as in the famous Shapley game [41].

These results extend only partially to n-player games (though they fully extend to n-player games with linear incentives [40]). The issue is that in n-player games no-regret dynamics turn out to be related to the correlated version of continuous fictitious play, in which the players play a best reply to the *correlated* past play of the others. This version of fictitious play is defined through a correspondence which is not convex valued. This creates technical difficulties, because the theory of perturbed differential inclusions is not developed for nonconvex-valued correspondences.

A different way to analyze no-regret dynamics is to show that some sets attract nearby solution trajectories. We show that strict Nash equilibria and, more generally, the intersection of

[2] This paper deals with the simplest notion of regret known as *unconditional* (or *external*) regret [20,25,26]. For more sophisticated regret notions, see Hart and Mas-Colell [24], Lehrer [34], and Cesa-Bianchi and Lugosi [12].

[3] The Hannan set of a game is also known as the set of *weak correlated equilibria* [39] or *coarse correlated equilibria* [48, Chapter 3].