UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI INFORMATICA
VIA CELORIA, 18 – 20133 MILANO – ITALY

Nicolò Cesa-Bianchi • nicolo.cesa-bianchi@unimi.it • phone +39-02-503-16280

November 20, 2019

Prof. David Monniaux
Président de la Commission HDR Mathématiques Appliquées et Informatique de l'Université de Grenoble Alpes

Report on the Habilitation thesis of Panayotis Mertikopoulos

I have read the Habilitation thesis of Dr. Mertikopoulos with title "Online Optimization and Learning in Games: Theory and Applications". The thesis is organized to provide an accessible overview of the candidate's research on multi-agent online learning. Results are stated and explained in the context of the relevant definitions. Proofs are mostly not included, but suitable references to the original articles are provided. The thesis itself is beautifully written, very well structured, and carefully formatted. It was for me a real pleasure to read.

The thesis investigates the application of the online convex optimization framework to the problem of learning in games. Chapter 2 provides a clear introduction to regret minimization in online convex optimization according to different feedback assumptions (especially full information, bandit feedback, and gradient feedback). The main algorithimical tools —Follow The Regularized Leader, Online Mirror Descent (OMD), and Dual Averaging— used in the subsequent chapters are introduced, and their connections explained. Although there is not a lot of technical or conceptual novelty here, this is an important preliminary chapter, as the rest of the thesis heavily depends on it. I found the explanation of the interplay between lazy and eager variants of OMD particuarly insightful. The second part of Chapter 2 introduces the basic game-theoretic notions, which the online learning application are built on. The main equilibria (Nash, correlated, and coarse correlated) are defined for continuous and finite games. Again, the clarity of exposition is remarkable, attesting a profound familiarity with the subject. The motivations for the introduction of the different notions of equilibria are compelling, and offer new insights even to the informed reader.

Chapters 3 and 4 contain the main conceptual and technical contributions. The leading theme is the study of the connection between online convex optimization and convergence to equilibria in games. Following previous works along the same lines, this connection is studied in a continuous time setting, which allows a clearer analytical picture of the resulting dynamics. The continuous time analogue of OMD, called CDA, is shown to have constant regret, thus significantly extending previous results for the Replicator Dynamics (a special case of CDA with entropic regularization). The behaviour of CDA in this special case is further investigated with respect to the two types of updates (lazy or eager), which translate to different dynamics (Replicator and Projected) with different properties. Next, the time evolution of CDA's joint action X_t is studied in finite games. A surprising result is that CDA may not lead to convergence even in simple finite 2-player zero-sum games admitting a unique interior Nash equilibrium, in which case almost every trajectory is shown to be Poincaré recurrent. This is stark contrast with the time-averaged orbit of CDA, which is known to lead to Nash equilibria in zero-sum games. Properties connecting CDA to its stable points and to the convergence to Nash equilibria are derived. No similarly strong results were known for CDA. Extensions of these results are also proven for concave games. Chapter 3 is concluded with a series of results concerning the CDA dynamics in the presence of stochastic perturbations of the gradient.

In Chapter 4 the continuous-time results are translated to their discrete-time counterparts. This discretization reveals a trade-off between speed of convergence (obtained with a coarse discretization) and discretization error. The main consequence of this trade-off is a $\sqrt{T}$ dependence on time in the regret. As expected, the lack of convergence of CDA in specifica cases is not affected by the discretization process. Indeed, just like CDA, OMD's joint action X_t does not converge in case of finite zero-sum games with a unique interior Nash equilibrium. The convergence properties of OMD are then studied in the two settings of concave games and finite games (also considered in the continuous case). For monotone games, the property of variational stability is used to prove that in all strictly monotone or strictly coherent games, OMD converges to the unique Nash equilibrium with probability 1. For finite games with perturbed gradients, local convergence to strict Nash equilibria is established. The chapter is closed by a section on learning in games with bandit feedback, that is when the only information available to the players is the payoffs they receive at each stage. In order to compensate this lack of information, OMD is modified using a single-shot perturbed gradient approximation, which is sufficient for regret minization (although it delivers suboptimal convergence rates). In the case of strictly monotone games, bandit OMD is shown to converge to the Nash equilibrium. In the case of strongly monotone games, a (possibly suboptimal) explicit rate of convergence is derived. The concluding result for this section concerns convergence to the strict Nash equilibrium in potential games.

Chapter 5 and 6 investigate applications of the multi-agent framework to a number of problems. These include distributed optimization using asynchronous stochastic gradient descent in master-slave models and multi-core with shared memory models. The online learning techniques are also used to derive and analyze the matrix exponential learning algorithm for throughput maximization in multiple-input and multiple-output (MIMO) systems.

The Habilitation thesis of Dr. Mertikopoulos develops an original and well focused research agenda at the interface of online learning, stochastic optimization, and game theory. This is a very interesting area of research, with a good potential impact on the field of machine learning (especially through applications to GANs and related models, which have also been explored by the candidate). Dr. Mertikopoulos' contributions have significantly advanced the state of the art by investigating the fundamental question of what type of convergence properties we can expect from online learning when run in a multi-agent setting. In addition to that, the thesis shows Dr. Mertikopoulos' deep understanding of the subject, including its historical roots and its most recent developments. In general, I was honestly impressed by the quality and the quantity of the thesis' content.

Dr. Mertikopoulos' publication record is vast and diversified in terms of venues, which include top journals and/or conferences in game theory (Games and Economic Behavior), optimization (Mathematical Programming, Mathematics of Operations Research), machine learning (NeurIPS), and information theory (IEEE Transactions on Information Theory). His scientific output in the last 5-6 years is impressive for a theoretical researcher (20 journal papers, 26 conference publications, and many in submission). His student supervision record is also excellent, with 5 PhD students and 5 post-docs in the last five years.

In summary, the Habilitation thesis of Dr. Mertikopoulos is the product of a mature and independent researcher, able to carry out an original research agenda at the intersection of three important fields, and showing great potential for further high-quality research contributions. For all these reasons, I recommend that the candidate should be allowed to defend his Habilitation.

Sincerely,

Nicolò Cesa-Bianchi