REFEREE REPORT ON THE HDR MANUSCRIPT

ONLINE OPTIMIZATION AND LEARNING IN GAMES: THEORY AND APPLICATIONS BY PANAYOTIS MERTIKOPOULOS

This work deals with online optimization algorithms in the framework of multi agents interactions. It is really impressive by the wide range of topics, the variety of approaches and the number of applications. The amount of material presented is much above the usual level of an "habilitation à diriger les recherches" and denotes an amazingly productive research activity.

The manuscript, which is a model of clarity and pedagogical presentation, is divided into 7 parts.

After a short introduction, Part 2 presents the framework of online optimization and of regret minimization. It then describes and analyses the main algorithms (FTL, OGD, MD and DA) and their properties in term of speed of convergence of the time average regret to 0. Furthermore the author provides an extremely clear discussion of their relations and specific characteristics. The last section concerns the game framework where several agents act independently. Equilibria and their link with variational inequalities are described and correlated equilibria, as well as coarse correlated equilibria are introduced.

Part 3 is mainly devoted to the analysis of the continuous time dual averaging dynamics (CDA) in the framework of games. It takes the form:

$$dY_t = V_t\,dt$$
$$X_t = Q(\eta_t Y_t)$$

where $V_t = V(X_t)$ is the evaluation through the map that appears in the variational formulation of the equilibria and Q is the choice map associated to the regularization function in the (DA) algorithm.
Then the author studies successively:
- the link with the replicator dynamics and the projection dynamics in finite games, and more generally Hessian Riemannian gradient systems,
- the constant bound on the regret in the unilateral case,
- the Poincaré recurrence of the trajectory in the 0-sum case,
- for finite games: the elimination of dominated strategies and an extended version of the "Folk Theorem" for evolutionary game dynamics,
- a convergence result for strictly monotone concave games.

The last section deas with learning procedures in the presence of noise. The corresponding (SDA) dynamics is given by:

$$dY_t = V_t dt + \sigma(X_t, t)\,dW_t$$
$$X_t = Q(\eta_t Y_t)$$

where W_t is a Wiener process and η_t a time dependent learning parameter. Precise bounds are given for the regret in the unilateral case. Then asymptotic results are obtained for finite games, extending the properties known in the deterministic case: elimination of dominated strategies, equilibrium convergence and stability properties, time average trajectories and link with (BRD). Finally in the framework of concave games interesting and useful conditions are given to obtain convergence.

Part 4 is the discrete counterpart of Part 3 and consider procedures of the form:

$$Y_{t+1} = Y_t + \gamma_t V_t$$

$$X_{t+1} = Q(Y_{t+1})$$

where the random input is of the form: $V_t = V(X_t) + U_t$, U_t being a noise satisfying usual conditions. The first result is a bound on the regret obtained by discretization of the continuous time dynamics of Part 3 and matching the optimal rate described in Part 2. Then positive conclusions concerning concave games are established under the original concept of "variational stability". The precise analysis covers properties of "variationally stable" points or sets, conditions for global or local convergence and finer results for sharp equilibria. Then several results are obtained for finite games and compared to their analogous versions in Part 3. A last section is concerned with the study of learning with bandit feedback. In the concave games case, the approach follows a simultaneous perturbation stochastic approximation and under a fine tuning of the parameters convergence for strictly monotone games is obtained. As for finite games, under the hypothesis of observed realized payoffs, a exploration and exploitation algorithm with exponential weights allows to achieve convergence in generic potential games.

Part 5 is devoted to the analysis of high-performance computing in a distributed system. The author described the distributed asynchronous stochastic gradient algorithm and some variants aiming at controlling the impact of the delays due to independent inputs. With step sizes adapted to the delays general convergence results are presented with, in the convex case more precise properties relying on the (CDA) dynamics via an asymptotic pseudo trajectory phenomena.

Part 6 deals with wireless networks and signal covariance optimization. This multi-input and multi-output system leads to a semidefinite optimization problem. The author suggests to replace the usual analysis via fixed point by taking advantage of the potential property of the underlying game and then using a learning scheme. Performance guarantees are proved and a deep analysis of complexity is provided.

These two last parts also present numerical experiments and justified discussions.

Part 7 describes a collection of promising research directions including: decomposition of games, memory of the players and acceleration dynamics, evolutionary approach to GAN ...

1. Main comments

Altogether the author shows an impressive and comprehensive knowledge of the field and provides an amazing amount of substantial contributions that cover all the aspects of online learning: continuous and discrete time models, deterministic and stochastic procedures, perfect information and noisy signals, ranging from usual gradient descent in the convex case to complex game theoretical models and from very theoretical results to applied contributions.

More precisely his large collection of results produces significant extensions in various directions allowing for a much deeper understanding of the structure of optimization and learning in games.

Examples would include:

- the link between discrete and continuous time processes,
- a very interesting connection of (CDA) with Hessian dynamics, or "variational stability" and ESS,
- the analysis of elimination of dominated strategies in finite games for large class of processes,
- the proof of versions of 'folk theorem" for (CDA) which corresponds to a wide generalization of properties of the replicator dynamics, even allowing for random perturbation,
- a deep study of the impact of noise and importance of the corresponding step size,
- the comparison between the trajectory and its time average,
- several precise, deep and very interesting discussions of the choice of the modelization and

comparison with previous approaches see e.g. for (SRD): Fudenberg & Harris, Cabrales, Imhof, Imhof & Hofbauer ...,
- a lot of new and important generalizations of results for finite games to games with a continuum of strategies.
One could also mention the introduction of new objects like the 'Fenchel coupling" as a distance or the notion of variational stability that prove to be extremely useful. A further characteristics is the very wide range of tools used that allows a.o. to invoque quite sophisticated arguments (ICT) for very practical problems.

P. Mertikopoulos thus presents a quite unique profile and has already reached a large international recognition which is obviously justified by the qualitative and quantitative amount of his achievements. This is reflected by his signifiquant number of publications (32 + 7 submitted) in the best journals of the field (SIAM Opt., Math Prog., MOR, SIAM Control, JOTA, GEB , JDG, IJGT, IEEE Trans., ...), his very frequent contributions to conferences (50) and his impressive list of collaborators.

Again the work presented is largely exceeding the amount and quality of material traditionally presented for an HDR. It is clear that Panayotis Mertikopoulos masters all the qualifications to conduct a research team and I give an extremely positive advice for his habilitation.

2. Minor remarks

Part 2 presents the basics of online optimization and learning which is now a hot topic in many fields: optimization, statistics, computer science and game theory a.o.
I have the feeling that some tools or ideas introduced early in game theory could have been mentioned: unilateral learning and worst case analysis treated as a game (Blackwell), regret (Hannan, 1956), fictitious play (Brown, 1948 and Robinson, 1951), smooth fictitious play (Fudenberg and Levine, 1995), relation external/internal regret and link to consistency and calibration (Foster, Vohra, Lugosi, Stoltz, Blum, Mansour ...) (where is the notion of correlated equilibria used ?)
Some references could have been added like:
Polyak for OGD,
Nemirovsky for OMD,
Vovk (1990) for exponential weight algorithm,
Nash and Glicksberg (1953) a.o. for continuous games,
link with non atomic games and Wardrop (1953) equilibrium,
coarse equilibria Moulin-Vial, (1978)
Sandholm (2001) on potential games.

p. 15 recall the dependence on $\mathcal{X}$ for H (l. 19) or $\Pi = \Pi_{\mathcal{X}}$ (l. - 5).
p. 26 there is a difference between existence of value and of equilibria in 0-sum games.
p. 35 Example 3.1 (l. 21)
last line [81]wrong reference
p. 44 l. -7 , ??
p. 58, l.-1, ?

Paris, November 17, 2019

Sylvain Sorin
IMJ, Sorbonne Université

Sylvain Sorin