Evaluation

Online learning

links with optimization and games

Université Paris–Saclay

UMD-based extension of AdaGrad-Norm and application to games

Let $d \geqslant 1$, $\mathcal{X} \subset \mathbb{R}^d$ a nonempty closed convex set, $K > 0$, $\|\cdot\|$ a norm on $\mathbb{R}^d$, h a regularizer with domain $\mathcal{X}$ that is K-strongly convex for $\|\cdot\|$.

Let $\gamma > 0$. For $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$, let $((x_t, y_t))_{t\geqslant 1}$ be a sequence of strict UMD iterates associated with regularizer h and dual increments $(\gamma_t u_t)_{t\geqslant 1}$, where

$$\gamma_t = \frac{\gamma}{\sqrt{\sum_{s=1}^t \|u_s\|_*^2}}, \qquad t \geqslant 1,$$

with convention $0/0 = 0$.

1) For $x \in \operatorname{dom} h$ and $T \geqslant 1$, derive a guarante on the regret

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle .$$

2) a) In the special cases of dual averaging (with a constant regularizer and dual increments $(\gamma_t u_t)_{t\geqslant 1}$) and online mirror descent (with a constant mirror map and dual increments $(\gamma_t u_t)_{t\geqslant 1}$), derive corresponding algorithms and guarantees.

b) Write corollaries for dual averaging with Euclidean regularizer and mirror descent with Euclidean mirror map.

c) For the entropic regularizer on the simplex, derive the corresponding algorithm and guarantee.

3) Apply to regret learning for finite two-player zero-sum games and derive guarantees. Perform numerical experiments to compare the convergence of the above algorithms (Euclidean DA, Euclidean MD, entropic regularizer) with RM, RM+, classical exponential weights, as well as *optimistic* counterparts of all previous algorithms.

4) Bonus. — Apply to various optimization problems and derive guarantees.

5) Bonus. — Also perform numerical experiments for extensive-form games using the counterfactual regret minimization approach.