EVALUATION
ONLINE LEARNING
LINKS WITH OPTIMIZATION AND GAMES
UNIVERSITÉ PARIS–SACLAY

ADAGRAD FOR APPROACHABILITY AND APPLICATION TO GAMES

Let $d \geqslant 1$ and consider an approachability problem given by some actions sets $\mathcal{A}$, $\mathcal{B}$ and an outcome function $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^d$. Let $\mathcal{C} \subset \mathbb{R}^d$ be a closed convex cone satisfying Blackwell's condition with respect to g, and let $\alpha : \mathcal{C}^\circ \to \mathcal{A}$ be an associated oracle satisfying:

$$x' = \lambda x \quad \text{for some } \lambda > 0 \quad \Longrightarrow \quad \alpha(x') = \alpha(x).$$

We define an adaptation of AdaGrad-Norm for this approachability problem as follows. Let $(b_t)_{t \geqslant 0}$ a sequence in $\mathcal{B}$, $x_0 = 0$, $a_0 = \alpha(x_0)$ and for $t \geqslant 0$,

$$x_{t+1} = \Pi_{\mathcal{C}^\circ}\left(x_t + \frac{\gamma}{\sqrt{\sum_{s=0}^t \|r_s\|_2^2}} r_t \right), \qquad a_{t+1} = \alpha(x_{t+1}),$$

where $\gamma > 0$, $\Pi_{\mathcal{C}^\circ}$ denotes the Euclidean projection onto $\mathcal{C}^\circ$ and for $t \geqslant 0$, $r_t = g_t(a_t, b_t)$.

1) Is this algorithm parameter-free?

2) Derive guarantees.

3) Write the algorithm in the special case of regret minimization on the simplex, and derive guarantees.

4) Apply to regret learning for finite two-player zero-sum games and derive guarantees. Perform numerical experiments to compare the convergence of the algorithm with RM, RM+ and exponential weights.

5) Bonus. — In numerical experiments, also consider optimistic variants of all algorithms.

6) Bonus. — Also consider an adaptation of AdaGrad-Diagonal.