Evaluation
Online learning
links with optimization and games
Université Paris–Saclay

approachability algorithms based on time-dependent norms

Consider the approachability framework from the course and corresponding notation. Let $\mathscr{C} \subset \mathbb{R}^d$ be a closed convex set satisfying Blackwell's condition and $\alpha : \mathscr{C}^\circ \to \mathscr{A}$ an associated oracle such that

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

The goal of this project is to define two families of parameter-free algorithms for approachability and apply them to regret minimization on the simplex, and then study their practical behavior in the context of solving games.

Let $\beta > 0$ and $(\|\cdot\|_{(t)})_{t\geqslant 1}$ be a sequence of norms in $\mathbb{R}^d$ such that $\|\cdot\|_{(t+1)} \geqslant \|\cdot\|_{(t)}$ for all $t \geqslant 1$.

1) Consider regularizers

$$h_t(x) = \frac{1}{\beta}\|x\|_{(t)}^{\beta} + \mathrm{I}_{\mathscr{C}^\circ}(x), \qquad x \in \mathbb{R}^d,\ t \geqslant 1.$$

Consider the DA algorithm for approachability associated with regularizers $(h_t)_{t\geqslant 1}$, constant parameter 1, and oracle α:

$$a_t = \alpha\left(\nabla h_t^*\left(\sum_{s=1}^{t-1} r_s\right)\right), \qquad t \geqslant 1.$$

Let $\mathscr{X}_0 \subset \mathscr{C}^\circ$ be a nonempty closed set. Let $T \geqslant 1$.

a) For all $\lambda > 0$, establish an upper bound on

$$\max_{x \in \lambda \mathscr{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle ,$$

where $y_t = \sum_{s=1}^{t-1} r_s$ for all $t \geqslant 1$.

b) By dividing the above by λ, using the properties of regularizers $(h_t)_{t \geqslant 1}$, and considering a judicious value for λ, deduce an upper bound on

$$\max_{x \in \mathscr{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle .$$

c) Let $L > 0$. Study the special case where $\beta = 2$ and for each $t \geqslant 1$,

$$A_t = \sqrt{L I_d + \sum_{s=1}^{t-1} \operatorname{diag}(r_s r_s^\top)} \quad \text{and} \quad \|\cdot\|_{(t)} = \|\cdot\|_{A_t} ,$$

where the square root is to be understood component-wise.

2) Consider mirror maps

$$H_t(x) = \frac{1}{\beta} \|x\|_{(t)}^{\beta} , \qquad x \in \mathbb{R}^d,\ t \geqslant 1.$$

a) Consider the OMD algorithm for approachability associated with regularizers $(H_t)_{t \geqslant 1}$, constant step-size 1, oracle α, and initial action $a_1 = \alpha(0)$. Let $T \geqslant 1$. With a similar analysis as in the previous question, establish an upper bound on

$$\max_{x \in \mathscr{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle .$$

b) Study the special case of AdaGrad-Diagonal.

3) Let $M > 0$. In the context of regret minimization on the simplex, assume that payoff vectors $(u_t)_{t \geqslant 1}$ are bounded as $\|u_t\|_\infty \leqslant M$ for all $t \geqslant 1$. Then derive guarantees for the above algorithms.

4) In the context of regret learning in finite two-player zero sum games, conduct numerical experiments to compare the performance of the above algorithms studied in questions 1c) and 2b), the exponential weights algorithm, RM and RM+.

5) Bonus. — Include in numerical experiments the *optimistic* counterpart of each algorithm.

6) Bonus. — Conduct numerical experiments in the context of extensive-form games.