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

parameter-free approachability algorithms

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 $\beta > 0$.

1) Let h be a regularizer on $\mathscr{C}^\circ$ such that for all $x \in \mathbb{R}^d$ and $\lambda \geqslant 0$,

$$h(\lambda x) - \min h = \lambda^\beta \left(h(x) - \min h \right).$$

Consider the DA algorithm for approachability associated with regularizer h, constant parameter 1, and oracle α:

$$a_t = \alpha \left(\nabla h^* \left(\sum_{s=0}^{t-1} r_s \right) \right), \qquad t \geqslant 0.$$

Let $\mathscr{X}_0 \subset \mathscr{C}^\circ$ be a nonempty closed set. Let $T \geqslant 0$.

a) Prove that for all $\lambda > 0$,

$$\max_{x \in \lambda \mathscr{X}_0} \left\langle \sum_{t=0}^{\mathrm{T}} r_t, x \right\rangle \leqslant \lambda^\beta \left(\max_{\mathscr{X}_0} h - \min h \right) + \sum_{t=0}^{\mathrm{T}} \mathrm{D}_{h^*}(y_t + r_t, y_t),$$

where $y_t = \sum_{s=0}^{t-1} r_s$ for all $t \geqslant 0$.

b) Deduce that

$$\max_{x \in \mathscr{X}_0} \left\langle \sum_{t=0}^{\mathrm{T}} r_t, x \right\rangle \leqslant 2 \left(\max_{\mathscr{X}_0} h - \min h \right)^{1/\beta} \left(\sum_{t=0}^{\mathrm{T}} \mathrm{D}_{h^*}(y_t + r_t, y_t) \right)^{1-1/\beta}.$$

2) Let H be a mirror map compatible with $\mathscr{C}^\circ$. We assume that H admits a minimum on $\mathbb{R}^d$, that the minimizer x_0 belongs to $\mathscr{C}^\circ$, and that for all $x \in \mathbb{R}^d$ and $\lambda \geqslant 0$,

$$\mathrm{H}(\lambda x) - \min \mathrm{H} = \lambda^\beta \left(\mathrm{H}(x) - \min \mathrm{H} \right).$$

Consider the OMD algorithm for approachability associated with regularizer H, constant step-size 1, oracle α, and initial action $a_0 = \alpha(x_0)$:

$$x_{t+1} = \operatorname*{arg\,max}_{x \in \mathscr{C}^\circ} \left\{ \langle \nabla \mathrm{H}(x_t) + r_t, x \rangle - \mathrm{H}(x) \right\} \quad \text{and} \quad a_{t+1} = \alpha\left(x_{t+1}\right), \quad t \geqslant 0.$$

Prove that for all $\mathrm{T} \geqslant 0$,

$$\max_{x \in \mathscr{X}_0} \left\langle \sum_{t=0}^{\mathrm{T}} r_t, x \right\rangle \leqslant 2 \left(\max_{\mathscr{X}_0} \mathrm{H} - \min \mathrm{H} \right)^{1/\beta} \left(\sum_{t=0}^{\mathrm{T}} \mathrm{D}_{\mathrm{H}^*}(\nabla \mathrm{H}_t(x_t) + r_t, \nabla \mathrm{H}(x_t)) \right)^{1-1/\beta}.$$

3) Let $1 < p \leqslant 2$. Consider algorithms from the above families associated with ℓ_p regularizer on $\mathscr{C}^\circ$ and ℓ_p mirror map on $\mathbb{R}^d$ respectively:

$$h_p = \frac{1}{2} \|\cdot\|_p^2 + \mathrm{I}_{\mathscr{C}^\circ} \quad \text{and} \quad \mathrm{H}_p = \frac{1}{2} \|\cdot\|_p^2.$$

Using (without proof) the fact that h_p and H_p are $(p-1)$-strongly convex for $\|\cdot\|_p$, derive corresponding guarantees.

4) Let $\mathrm{L} > 0$. In the context of regret minimization on the simplex, assume that payoff vectors $(u_t)_{t \geqslant 0}$ are bounded as $\|u_t\|_\infty \leqslant \mathrm{L}$ for all $t \geqslant 0$. Then derive guarantees for the above algorithms corresponding to ℓ_p regularizer and mirror map. Which value of p minimizes the regret bounds thus obtained?

5) In the context of regret learning in finite two-player zero sum games, conduct numerical experiments to compare the performance of the above ℓ_p algorithms, the exponential weights algorithm, RM and RM+.

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