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

solving two-player zero-sum games with blackwell's approachability

The approach from the course for solving two-player zero-sum games was to use two regret minimizing algorithms against each other. Those regret minimizers could be derived from Blackwell's approachability. In this project, we explore a possibly different approach for solving two-player zero-sum games directly with Blackwell's approachability, without defining regret minimizing algorithms as an intermediate step.

Let $m, n \geqslant 1$ be two integers and A a real-valued matrix of size $m \times n$. We consider the following approachability problem.

Let $\mathcal{A} = \Delta_m \times \Delta_n$ the set of actions of the Decision Maker and $\mathcal{B} = \{\varnothing\}$ the set of actions of Nature[1][2]. Consider the following outcome function $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^m \times \mathbb{R}^n$ defined as

$$g((a, b), \varnothing) = \left(\mathrm{A}b - \langle a, \mathrm{A}b\rangle \, \mathbb{1}, -\mathrm{A}^\top a + \langle a, \mathrm{A}b\rangle \, \mathbb{1}\right), \quad (a, b) \in \mathcal{A},$$

and target set $\mathcal{C} = \mathbb{R}^m_- \times \mathbb{R}^n_-$.

[1] not the empty set, but the singleton containing the empty set.

[2] Note that in this problem, the elements of $\mathcal{A}$ will be denoted (a, b), hence b will not be an element of $\mathcal{B}$ but part of an element of $\mathcal{A}$.

1) Prove that $\mathscr{C}$ satisfies Blackwell's condition for outcome function g and give an associated oracle α.

2) Give explicit expressions for Blackwell's algorithm and the Greedy Blackwell algorithm in this approachability problem. Are those algorithms different from RM and RM+ in the context of solving two-player zero-sum games? If so, write the corresponding guarantees, and derive as corollaries upper bounds on the duality gap.

3) Let $p, q > 1$ such that $1/p + 1/q = 1$. Consider regularizer

$$h_p(x) = \frac{1}{2}\|x\|_p^2 + \mathrm{I}_{\mathscr{C}^\circ}(x), \qquad x \in \mathbb{R}^m \times \mathbb{R}^n.$$

   a) Prove that for all $y \in \mathbb{R}^m \times \mathbb{R}^n$,

$$\nabla h_p^*(y) = \left(\|y_+\|_q^{2-q} (y_+)_i^{q-1} \right)_{1 \leqslant i \leqslant m+n},$$

   where $y_+ = (\max(0, y_i))_{1 \leqslant i \leqslant m+n}$.

   b) Give an explicit expression of the action (a_t, b_t) given at time $t \geqslant 1$ by the associated DA algorithm (with constant parameters 1), in terms of A, $(a_s)_{1 \leqslant s \leqslant t}$ and $(b_s)_{1 \leqslant s \leqslant t}$.

   c) Derive guarantees on the duality gap.

4) Define and analyse similar algorithms from the MD family (with constant step-size 1).

5) Perform numerical experiments. Compare the performance of the above algorithms to RM and RM+.

6) Bonus. — Extend this approach to the context of extensive-form games and perform numerical experiments.