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

minimizing regret on treeplexes with approachability

Let $\mathcal{T}$ be a treeplex defined as in the course. The aim of the project is to study new algorithms for online linear optimization on $\mathcal{T}$, and to apply them to extensive-form games. Consider

$$
\begin{aligned}
\Sigma_{\max} &= \{\sigma \in \Sigma,\ \nexists \theta' \in \Theta,\ \sigma = p(\theta')\}\,, \\
\nu(\theta) &= \operatorname{Card}\{\theta' \in \Theta,\ p(\theta) = p(\theta')\}\,, \\
\mathfrak{w} &= \left(\mathbb{1}_{\{(\theta,a)\in\Sigma_{\max}\}} \cdot \nu(\theta)^{-1} \cdot \prod_{\substack{(\theta',a')\in\Sigma \\ \theta\in\Theta^{[\theta' a' \downarrow]}}} \nu(\theta')^{-1} \right)_{(\theta,a)\in\Sigma}.
\end{aligned}
$$

1) Prove that $\mathbb{R}_+\mathcal{T} = \{\lambda x,\ \lambda \in \mathbb{R}_+,\ x \in \mathcal{T}\}$ is a closed convex cone.

2) Prove that for all $x \in \mathcal{T}$, the value $\langle \mathfrak{w}, x\rangle$ is the same.

3) Consider the approachability problem with $\mathcal{A} = \mathcal{T}$ as the action set of the Decision Maker, $\mathcal{B} = \mathbb{R}^\Sigma$ as the action set of Nature, outcome function

$$g(a, u) = u - \langle u, a\rangle\, \mathfrak{w}, \qquad a \in \mathcal{T},\ u \in \mathbb{R}^\Sigma,$$

and $\mathcal{C} = (\mathbb{R}_+\mathcal{T})^\circ$ as target set. Prove that $\mathcal{C}$ satisfies Blackwell's condition.

4) Write the corresponding Blackwell (resp. Greedy Blackwell) algorithm and derive guarantees for online linear optimization on $\mathcal{T}$.

5) Implement the above algorithms for solving Kuhn Poker. To do so, write the Euclidean projection onto $\mathbb{R}_+\mathcal{T}$ as a quadratic program (with linear constraints) and use a package such as `cvxopt` to solve it. Compare the performance with CFR and CFR+.

6) Bonus. — Give alternative outcome functions so that $\mathcal{C}$ still satisfies Blackwell's condition and compare the practical performance.