Evaluation

Online learning

links with optimization and games

Université Paris–Saclay

sparse payoff vectors

Let $d \geqslant 0$. Consider online linear optimization on $\mathscr{X} = \Delta_d$. Let $1 \leqslant s \leqslant d$ be an integer, assume that payoff vectors $(u_t)_{t \geqslant 0}$ are in $[0,1]^d$ and that for all $t \geqslant 0$, u_t has at most s nonzero components.

1) Assume that s is known, it the sense that the algorithm may depend on the value of s.
   a) Using a constant ℓ_p regularizer (or mirror map), derive the best possible regret bound.
   b) Using time-dependent regularizers (or mirror maps), derive the best possible *horizon-free* regret bound.
2) Propose an algorithm (e.g. with time-dependent regularizers or mirror maps) which achieves a similar regret bound as above without prior knowledge of s. *If this question is too difficult, just propose an algorithm which somehow tries to progressively adapt to the a priori unknown value of s.*
3) In the context of regret learning in finite two-player zero-sum games, compare the performance of the algorithm proposed in the previous question

with the exponential weights algorithm, regret matching, and regret matching+. *It will be interesting to pay attention to games with sparse matrices (meaning few nonzero entries).*

4) Same question with optimistic variants of considered algorithms.