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

sparse bandits

Let $1 \leqslant s \leqslant d$ be integers. Consider the multi-armed bandit problem ($\mathcal{X} = \Delta_d$) and assume that each payoff vector from sequence $(u_t)_{t \geqslant 0}$ has at most s nonzero components.

- If payoff vectors are assumed to be in $[-1, 0]^d$, [KP16, Theorem 11] establish a regret bound of order $\sqrt{Ts \log(d/s)}$.
- If payoff vectors are assumed to be in $[0, 1]^d$, [BCL18, Theorem 1] establish a regret bound of order $\sqrt{Ts \log d}$

Rewrite the statements and proofs of those results[1] *using the notation and tools from the course.*

[1] Papers are available at

- https://www.jmlr.org/papers/volume17/15-503/15-503.pdf
- https://arxiv.org/pdf/1711.01037.pdf

References

[BCL18] Sébastien Bubeck, Michael Cohen, and Yuanzhi Li. Sparsity, variance and curvature in multi-armed bandits. In *Algorithmic Learning Theory*, pages 111–127. PMLR, 2018.

[KP16] Joon Kwon and Vianney Perchet. Gains and losses are fundamentally different in regret minimization: The sparse case. *Journal of Machine Learning Research*, 17(229):1–32, 2016.