Sparsity-Agnostic Linear Bandits with Adaptive Adversaries

Tianyuan Jin Kyoungseok Jang Nicolò Cesa-Bianchi

NUS National University of Singapore

UNIVERSITAS STUDIORUM MEDIOLANENSIS

POLITECNICO MILANO

Sparse Linear Bandits

- In each round $t = 1, \cdots, T$

  1. The adversary chooses an arm set $\mathscr{A}_t \subset \mathbb{R}^d$

     - The adversary is allowed to be adaptive, i.e., $\mathscr{A}_t$ can depend in an arbitrary way on the past choices of the learner

  2. The learner choose an arm $A_t \in \mathscr{A}_t$

  3. The learner obtains a reward X_t

     - $X_t = \langle A_t, \theta_* \rangle + \epsilon_t$, where $\{\epsilon_t\}_{t \in [T]}$ are independent conditionally 1-subgaussian random variables

Sparse Linear Bandits

- Model parameter $\theta_* \in \mathbb{R}^d$
  - $S = \|\theta_*\|_0 = \sum_{i=1}^{d} \mathbb{I}\{\theta_i \neq 0\}$, i.e., the number of nonzero components of θ_*
  - S is unknown
- **Goal**: maximize the total reward $\Leftrightarrow$ minimize the regret
  - $R_T = \sum_{t=1}^{T} \max_{a \in \mathcal{A}_t} \langle a, \theta_* \rangle - \sum_{t=1}^{T} \langle A_t, \theta_* \rangle$

Previous Work

Reference	Sparsity-Agnostic	Adaptive Adversary	Regret	Assumptions
[Abbasi-Yadkori et al. 2011]	**Yes**	**Yes**	$\tilde{O}(\min\{d\sqrt{T}, d^2/\Delta\})$	-
[Abbasi-Yadkori et al. 2012]	**No**	**Yes**	$\tilde{O}(\min\{\sqrt{SdT}, dS/\Delta\})$	-
[Pacchiano et al. 2022]	**Yes**	**Yes**	$\tilde{O}(S^2\sqrt{T})$	Nested
			$\tilde{O}(S^2 d^2/\Delta)$	Nested, i.i.d actions
SparseLinUCB	Yes	Yes	$\tilde{O}(S\sqrt{dT}, \max\{d^2, S^2 d\}/\Delta)$	-

- **No bounds** improving on $\tilde{O}(d\sqrt{T})$ regret of the OFUL algorithm without *additional assumptions on the sparsity structure* or on the *action set generation*

SparseLinUCB

- Use the online to confidence set conversion technique
  - Given a sparsity parameter S, we can construct the confidence set of θ $\left\{\theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_{V_{t-1}^2}\| \leq \gamma(S)\right\}$, where radius of the confidence set is $\sqrt{\gamma(S)} = \Theta(\sqrt{S \log T})$.
- We run the confidence set conversion technique over hierarchy of nested confidence sets with radii $\alpha_i = 2^i \log T$ for $i = 1,\ldots,n = \Theta(\log d)$.
- For SparseLinUCB, we choose the larger confidence sets with exponentially small (but fixed) probability

AdaLinUCB

Motivation

- Practical considerations
  - The smaller the confidence radius, the better the experiments obtained
    - However, a small confidence radius does not provide any technical guarantee
- The existing solution uses the Exp3 algorithm to adaptively set the probability of choosing confidence radius at each time step
  - It does have good empirical performance
  - However, the regret bound is of order $T^{2/3}$

AdaLinUCB

- With probability $1-q$, it uses the Exp3 to set the probability for choosing the confidence radius at each time step. With probability q, it choose the confidence set with radii α_n

- The regret of AdaLinUCB is of order $\sqrt{T}$

- In our experiments, we show that AdaLinUCB outperform OFUL in a model selection task

Thank You

For more details, please check our full paper.