Linear Bandit Algorithms with Sublinear Time Complexity

Shuo Yang[1] Tongzheng Ren[1] Sanjay Shakkottai[2]
Eric Price[1] Inderjit S. Dhillon[1] Sujay Sanghavi[2]

[1]Department of Computer Science, UT Austin
[2]Department of Electrical and Computer Engineering, UT Austin

ICML 2022

Motivation

- We have minimax-optimal (up to logarithmic factors) linear bandit algorithms [e.g. AyPS11].
- However, all of the existing algorithms require solving the following problem:

$$a = \arg\max_{a \in \mathcal{A}} f(\theta, x_a),$$

where $\mathcal{A}$ is a finite arm set with $|\mathcal{A}| = K$, x_a is the feature corresponding to a. θ is the linear coefficients and f is a function depends on θ and x.

- Typically, this can be solved via an enumeration over $\mathcal{A}$, which leads to a per-step time complexity of $\Theta(K)$. However, it is not computationally efficient when K is large.

Contribution

- Key question: Can we obtain $o(K)$ per-step time complexity?
- We provide an affirmative answer with a novel Maximum Inner Product Search (MIPS) solver for adapted queries.
- We propose an elimination based algorithm and a Thompson Sampling based algorithm to achieve this target with complete characterization on the regret.
- We make discussions on the scenarios when we can achieve $o(K)$ complexity without sacrificing the regret.

MIPS Solver for Adapted Queries

Definition ((c, r, ε)-MIPS problem)

Let $P \subseteq \mathbb{R}^d$ be a finite set of points with $\|p\|_2 \leq 1, \forall p \in P$. Let $q \in \mathbb{R}^d$ be the query with $\|q\|_2 \leq 1$. The (c, r, ε)-approximated max inner product search ((c, r, ε)-MIPS) aims to find $p \in P$ such that $\langle q, p \rangle \geq cr - \varepsilon$ if there exists $p^* \in P$ with $\langle q, p^* \rangle \geq r + \varepsilon$.

Several high probability MIPS solvers with sublinear time complexity have been proposed for non-adaptive queries (with some additional pre-processing steps in $K^{1+o(1)}$ time). We can also adapt [ALRW17] to provide a $(c, r, 0)$ MIPS solver that success with probability at least 0.9. But we need to deal with multiple adaptive queries.

MIPS Solver for Adapted Queries

Our construction to deal with the adapted queries:

- Make an ε-cover of ℓ_2 ball whose size is $\Theta(\exp(d))$ (here ε corresponds to the addictive error in the definition of (c, r, ε)-MIPS problem).
- Initialize $\kappa = \widetilde{\Theta}(d)$ $(c, r, 0)$ non-adaptive MIPS solver (to boost the success probability).
- For each query q, round q to the nearest point in the cover (denoted as $\widehat{q}$), and return the results of any MIPS solver with query $\widehat{q}$.

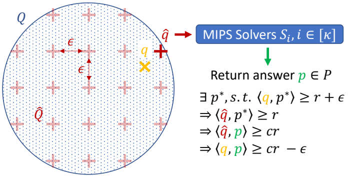

Figure: An illustration of over non-adaptive MIPS solver.

An Elimination-Based Algorithm

- Note that, the standard elliptical uncertainty for linear bandits can be written as:

$$\|x_a\|_{V^{-1}}^2 = \langle x_a x_a^\top, V^{-1} \rangle.$$

  Hence, we can easily query the arm with the highest uncertainty with MIPS solver.
- To get rid of the bad arm with high uncertainty, we can just eliminate those obviously bad arm during the interaction with the environments.
- Can be also adapted to the settings with slowly changed arm set.
- For the details, see our paper.

An Elimination-Based Algorithm

Theorem (Regret and time complexity of the Elimination-Based Algorithm)

For any $\delta \in (0,1)$, with probability at least $1-\delta$, the regret is bounded by

$$R(T) = \widetilde{O}\left(d\sqrt{T} + \eta(T) \cdot T\right)$$

with $\eta(T)$ controlling the approximate MIPS accuracy.

The per-step time complexity is $K^{1-\Theta(\frac{\eta(T)^4}{\log^2 T})+o(\log^{-0.45} K)}$. The overall time complexity overhead (e.g., pre-processing) is $K^{1+o(1)}$.

If T does not scale with K, then we can select $\eta(T) = \Theta(T^{-1/2})$, that can obtain per-step sublinear time complexity with almost no additional cost.

A TS-Based Algorithm

- For TS-based algorithm, we can directly query the arm with the sampled $\widetilde{\theta}$ from the posterior.

Theorem (Regret and time complexity of the TS-Based algorithm)

For any $\delta \in (0,1)$, with probability at least $1-\delta$, the regret is bounded by

$$R(T) = \widetilde{O}\left(d^{3/2}\sqrt{T} + \eta(T)\cdot T\right),$$

with $\eta(T)$ controlling the approximate MIPS accuracy.
The per-step time complexity is $K^{1-\Theta(\eta(T)^2)+o\left(\log^{-0.45} K\right)}$. The overall time complexity overhead is $K^{1+o(1)}$.

If T does not scale with K, then we can select $\eta(T) = \Theta(T^{-1/2})$, that can obtain per-step sublinear time complexity with almost no additional cost.

Experiments

Algorithm	Linear Elim	Sub-Elim, shortlist 30	Sub-Elim, shortlist 100
Regret	3847 ± 212	3795 ± 206	3806 ± 206
Time(s)	29.55	4.22 (3.47)	4.83 (4.09)
Speedup	$\times 1$	$\times$**7.00** ($\times$**8.52**)	$\times$**6.12** ($\times$**7.22**)

Algorithm	Linear TS	Sub-TS, shortlist 30	Sub-TS, shortlist 100
Regret	1193 ± 66	1177 ± 66	1202 ± 68
Time(s)	29.83	19.59 (19.38)	20.63 (20.41)
Speedup	$\times 1$	$\times$**1.52** ($\times$**1.54**)	$\times$**1.45** ($\times$**1.46**)

Table: Experiments on movie recommendation with Movielens-1M. Compared with the standard elimination-based algorithm and TS algorithm, our algorithm can have significant acceleration without sacrificing the cumulative regret. Number in the brackets denote the time and speedup without taking the pre-processing time into account.

Conclusion

- We propose a novel method to achieve sublinear per-step time complexity in linear bandits.
- An adaptive MIPS solver is introduced, which may be of independent interest.

Reference I

[ALRW17] Alexandr Andoni, Thijs Laarhoven, Ilya Razenshteyn, and Erik Waingarten. Optimal hashing-based time-space trade-offs for approximate near neighbors. In Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, pages 47–66. SIAM, 2017.

[AyPS11] Yasin Abbasi-yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K.Q. Weinberger, editors, Advances in Neural Information Processing Systems, volume 24. Curran Associates, Inc., 2011.