View Meta-Reviews

Paper ID

541

Paper Title

Improved Algorithms for Adversarial Linear Contextual Bandits via Reduction

Track Name

COLT2025

META-REVIEWER #1

META-REVIEW QUESTIONS

2. Meta-review

The authors propose an algorithm for adversarial linear contextual bandits that is computationally efficient and can yield meaningful regret bounds in the independent parametrized LCB and Jointly Parametrized LCB settings. This algorithm achieves a rate of d\sqrt{T}, which is a polynomial factor in d better than previous bounds. The algorithm reduces the problem of adversarial contextual bandits where the context is an i.i.d. sample of a distribution D, to adversarial linear contextual bandits.

During the discussion that followed the authors' rebuttal, the reviewers have repeatedly questioned about the significance of the contribution placing this submission in the borderline area. After going over the paper myself (including the appendix) I concluded that the contribution is indeed neat. Work needs to be done in order to clarify some of the points raised by the reviewers.

View Reviews

Paper ID

541

Paper Title

Improved Algorithms for Adversarial Linear Contextual Bandits via Reduction

Track Name

COLT2025

Reviewer #1

Questions

1. Rating (scale of 1-7)

3: Marginally below acceptance threshold

3. Review

The paper aims to provide a reduction from the adversarial linear contextual bandit problem to the adversarial linear bandit problem, leveraging an oracle for the latter that permits some level of model misspecification. (In both problems, loss vectors are adversarial; in the contextual version, the action set is drawn iid from some underlying distribution while in the non-contextual version, the action set is fixed. In both problems, the expected loss of the chosen action is linear, given by the dot product between the loss vector and the chosen action.)

The reduction is applied to get an O(dT^{1/2}) regret bound, tightening the dependence on the dimension of the space, d, with a polynomial-time algorithm. Liu et al. (2023) gave a polynomial-time algorithm with O(d^2T^{1/2}) regret and showed that EXP4 gives a O(dT^{1/2}) regret inefficiently, with complexity that scales as T^d.

Tightening the dependence on d with an efficient algorithm is of some interest. I believe the result is roughly correct but the paper lacks precision and reads more like a working draft.

It would help to outline the structure of the construction (schedule for running Algorithm 5) from the start rather than waiting until page 11.

Can you use the reduction in Foster et al. (Adapting to Misspecification in Contextual Bandits) instead of Algorithm 6 to get the same result?

Edit: Author feedback was reviewed. Thank you for clarifying the focus on computational efficiency. I think there's still an opportunity to significantly improve the readability. The paper's structure could be clearer by outlining how the pieces connect, making it easier for the reader to follow the flow, without having to piece everything together themselves. A few additional comments:
1) I believe Olkhovskaya et al. considers adaptive adversaries, whereas you assume an oblivious adversary but never explicitly discuss the distinction.
2) Terms like 'simulator' and 'free contexts' are not defined--it's clear what these are but they should be defined.
3) I found Definition 1 a bit odd--as far as I can tell, there no need to let c_ts depend on the learner's actions.

Reviewer #2

Questions

1. Rating (scale of 1-7)
3: Marginally below acceptance threshold

3. Review
In this work the authors have proposed an algorithm for adversarial linear contextual bandits that is computationally efficient and can yield meaningful regret bounds in the independent parametrized LCB and Jointly Parametrized LCB settings. This algorithm avoids issues such as assuming log concavity of the context distribution. Moreover this algorithm achieves a rate of d\sqrt{T}, which is a polynomial factor in d better than previous bounds. The algorithm works by reducing the problem of adversarial contextual bandits where the context is an i.i.d. sample of a distribution D, to the setting of adversarial linear contextual bandits problem. The main result in this work appears to be Lemma 4, and Algorithm 6.

The manuscript is fairly well written. A few suggestions to improve readability. The definition of hat{\Omega} is not clear. When reading it, I assumed it is represented in the form of some polytope. It is not exactly discussed how is that representation easy to use. I would like to hear from the authors what do they think is the main innovation, if any, of their work. Is the reduction from contextual to linear via computing the set \hat{\Omega} the main contribution? The misspecified truncated Exp Weights algorithm seems very similar to exponential weights for adversarial linear bandits, since the misspecification level can be explicitly upper bounded, it seems there isn’t substantial challenge in making Algorithm 6 work. Is this correct?
For the small loss results the authors rely on having access to a small loss algorithm that

tolerates misspecification. Is there any such algorithm? And if there is, why do we need Algorithm 6 since a small loss algorithm would subsume the min-max result.

Reviewer #4

Questions

1. Rating (scale of 1-7)

4: Marginally above acceptance threshold

3. Review

Pros) This paper has clear contributions - orderwise improvements on IP and JP, relaxing many assumptions too. I was not able to check the full details of the proofs, but most of the flows seem reasonable, and the techniques they used were quite interesting - by using multiple reduction, authors not only improved JP and IP results, but also devised a unique setting called misspecification-robust (even though known c_t is a bit artificial setting).

Cons) However, I believe that adding a figure will definitely help readers to understand the workflow. If I understood correctly, the flow is IP -> JP -> LinBandit -> Misspecification-robust, but it is not easy to follow the flow and it would be great if there's a diagram that shows the big picture of this paper.

Overall, I believe this paper is above the acceptance threshold.

Typo on page 10: Appendix reference