Generalized Linear Bandits revisited

Julian Zimmert
Humboldt University of Berlin

Csaba Szepesvri
University of Alberta

Abstract

Previous work on generalized linear bandit (GLB) suggested that algorithms based on optimisn can achieve a $\sqrt{T}$-regret analogous to the linear bandit setting. We argue however, that these results should be treated as asymptoic results, as they include generally huge constants that might even be a priori unknown. Our novel lower bound for GLB proves that these constants are not mere artifacts of the analysis, but that the regret might be as large as $\Omega(\frac{1}{2} + \frac{d-1}{2d+6})$. We prove that this lower bound has a matching upper bound up to log factors on a finite time setting with unit ball arms.

1 Introduction

The K-armed bandit problem is an extensively studied problem in online learning **??** An agents choses one of the K arms at each timesteps and observes a reward that depends on the selected action. The agent tries to play a strategy, that maximizes the sum of rewards. This setting is extended to problems where the arms exhibit a structure. That means playing one arm might provide information about the reward distribution for arms that were never played. A basic setting of this kind, which received a lot of attention, are linear bandits **??** It is assumed that each arm has a feature vector x_i and the mean reward observed when playing arm i is given by $x_i^T \theta_*$, while the derivation from the mean is governed by sub-gaussian noise. The optimal strategy needs to carefully balance between playing arms to gain information about the reward structure (Exploration) and commiting to arms which are most promising given the available information (Exploitation). Lower bounds show that this

Preliminary work. Under review by AISTATS 2017. Do not distribute.

Linear bandits were further extended to generalized linear bandits (GLB) in

In the classical K-armed bandit problem, an agent selects at each time step one of the K arms and receives a reward that depends on the chosen action. The aim of the agent is to choose the sequence of arms to be played so as to maximize the cumulated reward. There is a fundamental trade-off between gathering experimental data about the reward distribution (exploration) and exploiting the arm which seems to be the most promising. In the basic multi-armed bandit problem, also called the independent bandits problem, the rewards are assumed to be random and distributed independently according to a probability distribution that is specific to each arm see [1, 2, 3, 4] and references therein. Recently, structured bandit problems in which the distributions of the rewards pertaining to each arm are connected by a common unknown parameter have received much attention [5, 6, 7, 8, 9]. This model is motivated by the many practical applications where the number of arms is large, but the payoffs are interrelated. Up to know, two different models were studied in the literature along these lines. In one model, in each times step, a side-information, or context, is given to the agent first. The payoffs of the arms depend both on this side information and the index of the arm. Thus the optimal arm changes with the context [5, 6, 9]. In the second, simpler model, that we are also interested in here, there is no side-information, but the agent is given a model that describes the possible relations

2 Problem setting

3 Lower bounds

We now present the main result of the paper. The general lower bound proofs the claim that the constants are not mere artifacts of the analysis, while the algorithm dependent bound proofs that there is an unsuperable gap when using the pseudo MLE for subgaussian noise.

3.1 Algorithm independent bound

The exponent converges to 1 as the dimension increases, that means for a finite time horizon and large dimensions, the regret will almost be undistinguishable from linear regret. In real world application, we typically deal with relatively large dimensions. Unless we are very careful about the chosen link function, it might be impossible to learn anything meaningfull in a given time. That is why we distinguish in 4 between different families of link functions, to provide safe-to-use applications.

3.2 Pseudo Maximum-Likelihood based lower bound

For simplicity, we assume a simple Explore then Commit algorithm. We show that for any time horizon T, we can construct a problem where with constant probability the pseudo MLE choses an arm not better than random. During the exploration, we suffer an regret even larger than what the general lower bound predicts.

4 Upper bounds

The lower bound given in the previous section is of a relatively odd nature. To our best knowledge, the hasn't been a similar exponent derived for any bandit related lower bound. This naturally raises the question, whether this is even a tight lower bound. While for practical purposes, this is of little relevance because it is already prohibitive with the existing bound, it is still of academic interest. Secondly, it is essential for provide well behaving function that are guaranteed to suffer $\sqrt{T}$ regret instead.

We are not yet able to provide a complete answer to these question. Based on a slightly simplified setting, we can show that

- the lower bound is indeed tight up to log factors
- symmetric convex-concave link functions are safe-to-use

Restrictions We believe the results hold even without further restrictions, but removing them would complicate the proofs to a point that would burst all limits of our appendix. In this section, we assume that

- the arms consists of the d-dimensional unit ball
- $||\theta_*||$ is known.
- The derivative $\dot{\mu}$ is axis symmetric and monotonous

4.1 concave-convex link functions

We will first

4.2 convex-concave link functions

4.2.1 Citations in Text

Citations within the text should include the author's last name and year, e.g., (Cheesman, 1985). References should follow any style that you are used to using, as long as their style is consistent throughout the paper. Be sure that the sentence reads correctly if the citation is deleted: e.g., instead of "As described by (Cheesman, 1985), we first frobulate the widgets," write "As described by Cheesman (1985), we first frobulate the widgets." Be sure to avoid accidentally disclosing author identities through citations.

4.2.2 Footnotes

Indicate footnotes with a number[1] in the text. Use 8 point type for footnotes. Place the footnotes at the bottom of the column in which their markers appear, continuing to the next column if required. Precede the footnote section of a column with a 0.5 point horizontal rule 1 inch (6 picas) long.[2]

4.2.3 Figures

All artwork must be centered, neat, clean, and legible. All lines should be very dark for purposes of reproduction, and art work should not be hand-drawn. Figures may appear at the top of a column, at the top of a page spanning multiple columns, inline within a column, or with text wrapped around them, but the figure number and caption always appear immediately below the figure. Leave 2 line spaces between the figure and the caption. The figure caption is initial caps and each figure should be numbered consecutively.

Make sure that the figure caption does not get separated from the figure. Leave extra white space at the bottom of the page rather than splitting the figure and figure caption.

This figure intentionally left non-blank

Figure 1: Sample Figure Caption

[1] Sample of the first footnote.
[2] Sample of the second footnote.

4.2.4 Tables

All tables must be centered, neat, clean, and legible. Do not use hand-drawn tables. Table number and title always appear above the table. See Table 1.

Use one line space before the table title, one line space after the table title, and one line space after the table. The table title must be initial caps and each table numbered consecutively.

Table 1: Sample Table Title

PART	**DESCRIPTION**
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

5 SUPPLEMENTARY MATERIAL

If you need to include additional appendices during submission, you can include them in the supplementary material file.

6 INSTRUCTIONS FOR CAMERA-READY PAPERS

For the camera-ready paper, if you are using LaTeX, please make sure that you follow these instructions. (If you are not using LaTeX, please make sure to achieve the same effect using your chosen typesetting package.)

1. Download `fancyhdr.sty` – the `aistats2017.sty` file will make use of it.

2. Begin your document with

   ```
   \documentclass[twoside]{article}
   \usepackage[accepted]{aistats2017}
   ```

   The `twoside` option for the class article allows the package `fancyhdr.sty` to include headings for even and odd numbered pages. The option `accepted` for the package `aistats2017.sty` will write a copyright notice at the end of the first column of the first page. This option will also print headings for the paper. For the *even* pages, the title of the paper will be used as heading and for *odd* pages the author names will be used as heading. If the title of the paper is too long or the number of authors is too large, the style will print a warning message as heading. If this happens additional commands can be used to place as headings shorter versions of the title and the author names. This is explained in the next point.

3. If you get warning messages as described above, then immediately after `\begin{document}`, write

   ```
   \runningtitle{Provide here an alternative
   shorter version of the title of your
   paper}
   \runningauthor{Provide here the surnames
   of the authors of your paper, all
   separated by commas}
   ```

   Note that the text that appears as argument in `\runningtitle` will be printed as a heading in the *even* pages. The text that appears as argument in `\runningauthor` will be printed as a heading in the *odd* pages. If even the author surnames do not fit, it is acceptable to give a subset of author names followed by "et al."

4. Use the file sample_paper.tex as an example.

5. Both submitted and camera-ready versions of the paper are 8 pages, plus any additional pages needed for references.

6. If you need to include additional appendices, you can include them in the supplementary material file.

7. Please, don't change the layout given by the above instructions and by the style file.

Acknowledgements

Use unnumbered third level headings for the acknowledgements. All acknowledgements go at the end of the paper. Be sure to omit any identifying information in the initial double-blind submission!

References

References follow the acknowledgements. Use an unnumbered third level heading for the references section. Any choice of citation style is acceptable as long as you are consistent. Please use the same font size for references as for the body of the paper—remember that references do not count against your page length total.

J. Alspector, B. Gupta, and R. B. Allen (1989). Performance of a stochastic learning microchip. In D. S. Touretzky (ed.), *Advances in Neural Information Processing Systems 1*, 748-760. San Mateo, Calif.: Morgan Kaufmann.

F. Rosenblatt (1962). *Principles of Neurodynamics.* Washington, D.C.: Spartan Books.

G. Tesauro (1989). Neurogammon wins computer Olympiad. *Neural Computation* **1**(3):321-323.