View Reviews

Paper ID
2775

Paper Title
Average-case Complexity of Teaching Convex Polytopes via Halfspace Queries

Reviewer #1

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).
This paper discusses the problem of teaching convex polytopes, i.e., regions induced by intersecting arbitrary n-halfspaces in R^d, assuming the presence of a perfect teaching oracle. It shows the average teaching complexity of this problem is \Theta(d) instead of \Theta(n) in the worst case analysis. Notably, the average teaching complexity has no dependence on n. When learning, the learning complexity is \Theta(n) for iid queries and \Theta(d log n) when the learner chooses the points to query.

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.
The paper is well-written and well-positioned to other related works. The results are theoretical with practical implications. They certainly add to the current knowledge on teaching and learning-by-query complex hypothesis classes. I believe providing tight bounds on the average teaching complexity of polytopes is of general interest to the wide community of ICML, machine learning, and learning theory.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)
Solid contribution to relevant problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)
Worthy contributions, but not surprising

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)
Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)
N/A. Experimental evaluation not needed, not applicable, or not standard for this work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)
Mostly clear, but improvements needed, as recommended in the detailed comments.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)
Yes

9. [Questions for authors] Please provide questions for authors to address during the author feedback period. (Optional, to help authors focus their response to your review.)
This work assumes version space learners and honest responses to each query, I would appreciate if the authors comment on how to tackle the work with faulty oracles.

10. Please provide an "overall score" for this submission.
Weak Accept: Borderline, tending to accept

11. [Confidence] Please provide your confidence in your assessment of this submission.
I am knowledgeable and willing to defend my evaluation, but there's a chance I missed something.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.
Agreement accepted

Reviewer #2

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).
The paper studies the minimum number of examples (half-spaces) to locate a randomly chosen region determined by the arrangement of n d-dimensional hyperplanes. In particular, the introduction of a generalization of general position allows to give sharper bounds for the case of hyperplanes in this "relaxed general position"-parameterization. The compaison to the case of active learning and passive learning is also provided

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.
The introduction of the relaxed general position appears to give better resolution on the bound achievable. The paper builds on computational geometry perspective and on the possibility to count the regions in the case of the "relaxed general position" structure. I guess some of the structural results used here could be of independent interest for other learning settings.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)
Reasonable contribution to a minor problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)
One idea that surprised me by its originality, solid contributions otherwise

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)
Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)
N/A. Experimental evaluation not needed, not applicable, or not standard for this work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)

Very clear, only minor flaws.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)
Yes

9. [Questions for authors] Please provide questions for authors to address during the author feedback period. (Optional, to help authors focus their response to your review.)
A minor point: the subroutine of algo1 is called FindLabels in the pseudocode and FindTS in the text.

10. Please provide an "overall score" for this submission.
Weak Accept: Borderline, tending to accept

11. [Confidence] Please provide your confidence in your assessment of this submission.
I am knowledgeable about this area, but I am not certain about, and am willing to change my evaluation.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.
Agreement accepted

Reviewer #5

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).
This paper considers the average complexity of teaching a target region among the intersections of n fixed hyperplanes in R^d. It shows that if these hyperplanes are in d'-"relaxed general position", then the average teaching dimension is d'. This is done by bounding the number of faces and regions of the intersections. They further extend their results to efficient active learning algorithm to find the target region, teaching phi-separable dichotomy, and teaching ranking through pairwise comparisons.

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.
Strengths:
- it considers an interesting problem of average teaching complexity, and shows the average complexity has a much better dependency on n (the number of hyperplanes).
- The results look reasonable and technically non-trivial though I didn't carefully check their proofs due to time constraints.

Weaknesses:
- I'm not very sure how relevant/significant this problem is. This paper assumes the hyperplanes are fixed, while in a typical learning setting, theses hyperplanes need to be learned.
- the main contribution of this paper is showing that the average teaching complexity is small for regions induced by "relaxed general position of hyperplanes". However, it is not clear to me how interesting/reasonable this definition is. It seems there are arrangements of hyperplanes that don't belong to any d'-relaxed general position. If so, Can the authors comment on why this definition is interesting other than allowing for tighter teaching complexity bounds, and is there any interesting positions that can't be covered by this definition?
- I'm quite confused by the phi-dichotomy section (Section 6). It seems it is the same as learning/teaching a single linear classifier in the kernel space induced by phi. Isn't there prior work for this linear problem?
+ relatedly, Definition 3 and 4 also look quite confusing. By this definition, d'+1 general position implies d' general position, but by corollary 2 the latter is easier to teach..
- This paper is not very easy to read for me besides above problem on section 6. For several sentences I have to read many times to understand them. For example:

+ in line 159~160, the meaning of "arrangement" was not quite clearly me. The paper use "induce" frequently and it caused some trouble for me. It would be better to give a more formal definition of A(H), faces, R(A(H)), etc.
+ line 162: what does it mean by R^d- \cup_h\in H h?
+ In Definition 1&2, it took me some time to realize "otherwise" means k>d.
+ Algorithm 1: the text description above refers to FindTS but in Algorithm 1 it is FindLabels. Also, it is not entirely clear to me how this subroutine works through the description. A formal statement would be helpful.

—
Response to author feedback:

The author feedback did address some of my concern. I'm not familiar with computational geometry literature, but I'm still not sure if this paper is interesting from the machine learning perspective. Some detailed comments:

§6 and relevance of the problem of learning/teaching ϕ-dichotomies
Thanks for clarifying. It would be more clearer if this could be stated in the paper. However, it seems the duality map is quite straightforward, and I'm not quite sure if there is any new/interesting insights we can get from this connection.

Relevance of the problem to ML community
I'm still not convinced that learning from fixed set of hyperplanes is an interesting problem. For related questions you mentioned, I agree the ranking one is interesting, but I don't see how this work provides new insights to perceptron or ϕ-separable dichotomy.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)
Reasonable contribution to a minor problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)
One idea that surprised me by its originality, solid contributions otherwise

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)
Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)
N/A. Experimental evaluation not needed, not applicable, or not standard for this work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)
This paper is difficult to understand in places because of typos, lack of organization, or another flaws. The paper would need significant/major improvement in writing.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)
Yes

10. Please provide an "overall score" for this submission.
Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.
I am knowledgeable and willing to defend my evaluation, but there's a chance I missed something.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.
Agreement accepted

Reviewer #6

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).
This paper investigates the average-case teaching complexity of point location problem using halfspace queries. Unlike previous works, which consider the worst case complexity, this work shows that the average case teaching complexity is Theta(d) and learning complexity is Theta(d log n).

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.
1. Locating a target region among those induced by intersections of n halfspaces (also known as point location problem) is a fundamental problem in computational geometry. This work shows an important average case teaching complexity result.
2. The paper is very clearly written and the proof ideas are very nicely distilled for readers.
3. I verified most proofs in the paper and they look correct to me.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)
Solid contribution to relevant problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)
One idea that surprised me by its originality, solid contributions otherwise

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)
Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)
N/A. Experimental evaluation not needed, not applicable, or not standard for this work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)
Very clear, only minor flaws.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)
Yes

10. Please provide an "overall score" for this submission.
Accept: Good paper

11. [Confidence] Please provide your confidence in your assessment of this submission.
I am knowledgeable and willing to defend my evaluation, but there's a chance I missed something.

Reviewer #7

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).
The paper discusses the complexity of identifying a connected region among those that are defined by an arrangement of n hyperplanes in R^d. The identification procedure consists of testing a sequence of hyperplanes: each test splits the space into two halfspaces, with exactly one of them including the target region.

The authors introduce the concept of d'-relaxed general position arrangement, a natural generalization of general position arrangements, and provide tight bounds on both the average (uniform distribution over regions) and worst-case number of tests required for different learning models, namely, passive learning, active learning and teacher-based learning.

-

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.
STRENGTHS
The paper is well written and the authors present a thorough study for a fairly interesting topic. The techniques employed are also interesting.

WEAKNESS
- While the problem considered by the author is interesting, its relevance to the ML community is not very clear to me. Why this new kind of hyperplane arrangement proposed and studied in the paper is indeed relevant?

2. The "novelty factor" is not particularly strong:

-- The paper by [Fukuda 91] gives an O(d) bound on the average teaching for general arrangements. The contribution here is providing a better bound for a particular case that, while natural, does not have a clear relevance.

-- The bounds for the active learning problem are obtained with an approach very similar to that employed by [Jamieson and Novak 2001]. I do not see it as a major issue when the problem under consideration is relevant but, here, I think that it counts.

SPECIFIC COMMENTS
First bullet in page 2. The O(d) bound for general position arrangements follows from [Fukuda 91]. Thus, I do not see the first bullet as a contribution, at least not the way it is presented.

NOTE
None of the options for items 3 and 4 is a good fit to my opinion.

********** POST REBUTALL COMMENTS **********
Taking into account the discussion and the new review so far, I'm leaning towards rejection due to the following reasons:
(i) lack of a strong motivation for ML
(ii) Doubts about the value of the technical contribution. Maybe people from Computational Geometry community are in a better position to evaluate and appreciate the main results of this paper.

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)
Reasonable contribution to a minor problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)
Worthy contributions, but not surprising

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)
Technically adequate for its area, solid results

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)
N/A. Experimental evaluation not needed, not applicable, or not standard for this work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)
Very clear, only minor flaws.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)
Yes

10. Please provide an "overall score" for this submission.
Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.
I am knowledgeable and willing to defend my evaluation, but there's a chance I missed something.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.
Agreement accepted

Reviewer #8

Questions

1. [Summary] Please summarize the main claims/contributions of the paper in your own words (1-2 sentences or paragraphs).
This paper studies the average case complexity of teaching and learning convex polytopes and \phi separable dichotomy. In the context of teaching, we are faced with the following problem: we have a set of hyperplanes that generates a partition of \R^d; given a target cell in the partition, the teacher would like to send the learner all faces

of the cell. In the context of query learning, the learner can query for every hyperplane, which side the target cell is with respect to that hyperplane, with the goal of identifying the cell. The paper focuses on the average case analysis, in that it measures the teaching / learning complexity in the case when the target cell is chosen uniformly at random from the partition.

The paper shows that, for hyperplanes with d'-relaxed general position, the average teaching complexity is \Theta(d') (Thm 2) and the average query learning complexity is \Theta(d' log n) (Thm 3). In contrast, the worst case teaching and query learning complexity can be as large as \Theta(n). The result also gives another proof of an upper bound on the average number of extreme points in (Cover, 1965).

2. [Detailed comments] Describe the strengths and weaknesses of the work, with respect to the following criteria: soundness of the claims (theoretical grounding, empirical evaluation), significance and novelty of the contribution, relation with prior work, clarity of writing, and relevance to the ICML community.
Strengths:
- the results under the assumption of d'-relaxed general position, and the average query learning complexity result appears to be new.

Weaknesses:
- I had a hard time to appreciate the notion of d'-relaxed general position, especially its relevance to the machine learning community. The authors also acknowledged that if we are only shooting for a O(d) teaching complexity, the problem is basically solved in (Fukuda, 1991). Also, from a cursory read, it seems that the d'-relaxed general position is equivalent to the existence of a d'-dimensional subspace V, such that after projecting all hyperplanes onto V, the hyperplanes are in general position? Then why isn't the O(d') teaching complexity upper bound in Theorem 2 an immediate consequence of Fukuda's result?

- Theorem 3, although interesting, seem to be a direct adaptation of (Jamieson and Nowak, 2011).

- It is hard for me to appreciate the significance of Section 6. After the \phi transformation of all points and using the point-line duality, aren't the results in this section a straightforward consequence of Theorem 2?

Some minor comments:

1. It would be better to replace the text "otherwise has null intersection" in Definitions 1 and 2 to a more explicit description, something like "any subset of k > d hyperplanes has empty intersection" - otherwise it is easy for the reader (at least me) to mistakenly think that Definition 2 is weaker than Definition 1.
2. The notation in Section 6 is too abstract, and I am not sure if it is necessary to introduce heavy notation to convey the ideas.
3. For Definition 6, if hyperplane h is a face of r^* (indicating its informativeness), then why does there always exist some point in the boundary of h such that it is fully contained in r^*?

3. [Relevance and Significance] (Is the subject matter important? Does the problem it tries to address have broad interests to the ICML audience or has impact in a certain special area? Is the proposed technique important, and will this work influence future development?)
Reasonable contribution to a minor problem

4. [Novelty] (Is relation to prior work well-explained, does it present a new concept or idea, does it improve the existing methods, or extend the applications of existing practice?)
Worthy contributions, but not surprising

5. [Technical quality] (Is the approach technically sound. The claims and conclusions are supported by flawless arguments. Proofs are correct, formulas are correct, there are no hidden assumptions.)
A paper that may be strong in other respects, but not technically

6. [Experimental evaluation] (Are the experiments well designed, sufficient, clearly described? The experiments should demonstrate that the method works under the assumed conditions, probe a variety of aspects of the novel methods or ideas, not just the output performance, present comparisons with prior work, test the limits and check the robustness of the novel methods or ideas, and demonstrate their practical relevance.)
N/A. Experimental evaluation not needed, not applicable, or not standard for this work

7. [Clarity] (Is the paper well-organized and clearly written, should there be additional explanations or illustrations?)
Mostly clear, but improvements needed, as recommended in the detailed comments.

8. [Reproducibility] (are there enough details to reproduce the major results of this work?)
Yes

10. Please provide an "overall score" for this submission.
Weak Reject: Borderline, tending to reject

11. [Confidence] Please provide your confidence in your assessment of this submission.
I am knowledgeable and willing to defend my evaluation, but there's a chance I missed something.

16. Please acknowledge that you have read the author rebuttal. If your opinion has changed, please summarize the main reasons in the Detailed comments sections.
Agreement accepted