

View Meta-Reviews

Paper ID

1654

Paper Title

Near-Optimal Policy Optimization for Correlated Equilibrium in General-Sum Markov Games

META-REVIEWER #1

META-REVIEW QUESTIONS

3. Please provide a meta-review for this submission. Your meta-review should explain your decision to the authors. It should augment the reviews and communicate how the reviews, author response, and discussion were used to arrive at a decision. Dismissing or ignoring a review is not acceptable unless you have a good reason for doing so. If you want to make a decision that is not clearly supported by the reviews, perhaps because the reviewers did not come to a consensus, please justify your decision appropriately, including, but not limited to, reading the submission in full and writing a detailed meta-review that explains your decision.

This paper provides a new policy optimization algorithm for learning Correlated Equilibria (CE) in multi-player general-sum Markov Games under full-information feedback, with a fast convergence rate T^{-1} . The result improves on existing results which only give such a convergence rate for the easier problem of learning Coarse Correlated Equilibria (CCE). The algorithm works by combining the OFTRL algorithm with log-barrier regularizer and the Blum-Mansour reduction for swap regret.

All reviewers find the results solid and interesting, and the paper to be well-executed. Therefore, I recommend acceptance with oral presentation, and congratulate the authors for the nice work.

View Reviews

Paper ID

1654

Paper Title

Near-Optimal Policy Optimization for Correlated Equilibrium in General-Sum Markov Games

Reviewer #2

Questions

1. Summary and contributions: Briefly summarize the paper and its contributions. (max. 1000 characters)

This paper studies the general-sum Markov games and proposes an algorithm that achieves a near-optimal $O(T^{-1})$ convergence rate for correlated equilibrium. The convergence rate improves over the previous $O(T^{-1/2})$ rate significantly. The improvement comes from the no swap regret algorithm used in the policy update.

4. Soundness of theory and theorems, if applicable:

correct

5. Justification of your provided judgment in Q4, if not applicable, put in “N/A”. If your choice is “minor errors” or “major errors”, please specify the details of your concerns. (max. 2000 characters)

The theorems and proofs look good.

6. Significance of empirical results, if applicable:

not applicable

7. Justification of your provided judgment in Q6, if not applicable, put in “N/A”. If your choice is “mixed results” or “generally not significant”, please specify the details of your concerns. (max. 2000 characters)

No empirical results.

8. Novelty of the theoretical results and/or empirical methodology:

New results, although incremental and/or combinatorial

9. Justification of your provided judgment in Q8. If your choice is “known results” or “incremental/combinatorial” please specify the previous publication(s) that support your choice. (max. 2000 characters)

The proposed algorithm achieves near-optimal $O(T^{-1})$ convergence rate, improves over the previous $O(T^{-1/2})$ rate significantly. The algorithm is under the widely-used V-learning framework for computing equilibrium in Markov games. The main new ingredient is the no swap regret algorithm used in the policy update, which is shown to

have near-optimal convergence rate in normal-form games. Therefore, the theoretical results are great but the methodology is not very new.

10. Non-conventional contributions: does the submission contain non-conventional research contributions? Examples: novel ideas with not widely accepted assumptions, new problems and/or tasks, “bridging fields” contributions. (max. 2000 characters)

This work is the first to achieve a near-optimal convergence rate in computing correlated equilibrium for Markov games. While no-swap-regret algorithms have demonstrated near-optimal results in normal-form games, extending these findings to more intricate Markov games has remained unclear until this paper, effectively filling this critical gap.

11. Clarity: Is the paper clearly written? Consider whether the paper has clearly (1) stated its contributions, notation and results, (2) explained the meaning of the theoretical assumptions, and/or (3) motivated the proposed methodology well with e.g., examples. (max. 2000 characters)

The paper is well-written and has a clear and understandable presentation of both the proposed algorithm and the theoretical results.

12. Relation to prior work: Is it clearly discussed how this work differs from or relates to prior work in the literature? Any related work missing (if yes provide details in Q13)?

All related works are clearly discussed

13. Additional Comments: add your additional comments, feedback and suggestions for improvement, as well as any further questions for the authors. what would you do differently if you were given the chance to improve the paper? (Optional) (max. 8000 characters)

It will be better if empirical results of the proposed algorithm in synthetic/real games are provided.

14. Reproducibility. Are there enough details to reproduce the main experimental results of this work?

Sufficient amount of details available for reproducing the main results.

15. Paper score:

4: Accept

16. Confidence score:

4: You are confident in your assessment but not absolutely certain. (It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.)

17. Does the submission raise potential ethical concerns? Does this submission raise potential ethical concerns? These include methods, applications or data that create or reinforce unfair biases and/or that have a primary purpose of harm or injury.

No concern

18. Justification of your provided judgment in Q17 about ethical concerns, if not applicable, put in “N/A”. (max. 2000 characters)

N/A

19. Code of conduct. Agree to abide by the AISTATS code of conduct The AISTATS code of conduct can be found here: <https://aistats.org/aistats2024/code-of-conduct.html>

Agreement accepted

20. Confidentiality agreement. Reviewers must keep the paper and supplementary materials (including code submissions and LaTeX source), as well as the reviews, confidential. This includes deleting any submitted code at the end of the review cycle to comply with confidentiality requirements.

Agreement accepted

Reviewer #4

Questions

1. Summary and contributions: Briefly summarize the paper and its contributions. (max. 1000 characters)

The authors propose a novel analysis and algorithm for learning (coarse) correlated equilibria in Markov games. The proposed methods allow for achieving improved regret bounds and therefore improved convergence speeds. The claims are supported primarily by theory.

4. Soundness of theory and theorems, if applicable:

correct

5. Justification of your provided judgment in Q4, if not applicable, put in “N/A”. If your choice is “minor errors” or “major errors”, please specify the details of your concerns. (max. 2000 characters)

As the current paper falls outside the immediate scope of my research, I am unable to identify any immediate, significant weaknesses in the paper.

6. Significance of empirical results, if applicable:

not applicable

7. Justification of your provided judgment in Q6, if not applicable, put in “N/A”. If your choice is “mixed results” or “generally not significant”, please specify the details of your concerns. (max. 2000 characters)

N/A

8. Novelty of the theoretical results and/or empirical methodology:

New results that are significantly different from existing ones

9. Justification of your provided judgment in Q8. If your choice is “known results” or “incremental/combinatorial” please specify the previous publication(s) that support your choice. (max. 2000 characters)

The novelty and relation to existing works is discussed extensively, and is positioned as consisting primarily of combining and extending techniques from recent works, including OFTRL and self-concordant barriers, in order to achieve significant improvements of theoretical regret bounds or convergence rates in Markov games.

10. Non-conventional contributions: does the submission contain non-conventional research contributions? Examples: novel ideas with not widely accepted assumptions, new problems and/or tasks, “bridging fields” contributions. (max. 2000 characters)

N/A

11. Clarity: Is the paper clearly written? Consider whether the paper has clearly (1) stated its contributions, notation and results, (2) explained the meaning of the theoretical assumptions, and/or (3) motivated the proposed methodology well with e.g., examples. (max. 2000 characters)

The presentation is clear. Assumptions are explained, and the problem is motivated in the introduction.

12. Relation to prior work: Is it clearly discussed how this work differs from or relates to prior work in the literature? Any related work missing (if yes provide details in Q13)?

All related works are clearly discussed

13. Additional Comments: add your additional comments, feedback and suggestions for improvement, as well as any further questions for the authors. what would you do differently if you were given the chance to improve the paper? (Optional) (max. 8000 characters)

Most of the important related work appears to have been discussed extensively. The significance of strongly improving regret bounds appears sound to me, given the lengthy and clear exposition, and the importance of general Markov games.

However, given the formulation of novel algorithms with improved asymptotic regret versus existing baselines, the manuscript could be improved by providing empirical support of this improved regret minimization or convergence.

=== POST REBUTTAL ===

I thank the authors for their response. My concerns have been addressed, and I am inclined to increase my score by one. I now understand the difficulty in evaluating

equilibrium gaps due to their intractability, and the special cases from existing literature seem sufficient, if discussed in the manuscript.

14. Reproducibility. Are there enough details to reproduce the main experimental results of this work?

Not applicable: the paper has no experiments

15. Paper score:

4: Accept

16. Confidence score:

2: Somewhat confident (You are willing to defend your assessment, but it is quite likely that you did not understand central parts of the submission or that you are unfamiliar with some pieces of related work. Math/other details were not carefully checked.)

17. Does the submission raise potential ethical concerns? Does this submission raise potential ethical concerns? These include methods, applications or data that create or reinforce unfair biases and/or that have a primary purpose of harm or injury.

No concern

18. Justification of your provided judgment in Q17 about ethical concerns, if not applicable, put in "N/A". (max. 2000 characters)

N/A

19. Code of conduct. Agree to abide by the AISTATS code of conduct The AISTATS code of conduct can be found here: <https://aistats.org/aistats2024/code-of-conduct.html>

Agreement accepted

20. Confidentiality agreement. Reviewers must keep the paper and supplementary materials (including code submissions and LaTeX source), as well as the reviews, confidential. This includes deleting any submitted code at the end of the review cycle to comply with confidentiality requirements.

Agreement accepted

Reviewer #11

Questions

1. Summary and contributions: Briefly summarize the paper and its contributions. (max. 1000 characters)

This paper studies policy optimization algorithms for approaching correlated equilibria in general-sum Markov games. By proposing a new algorithm based on the optimistic-follow-the-regularized-leader (OFTRL) update, the convergence rate improves from $\mathcal{O}(T^{-1/2})$ to $\mathcal{O}(T^{-1})$, which is near-optimal.

4. Soundness of theory and theorems, if applicable:

correct

5. Justification of your provided judgment in Q4, if not applicable, put in “N/A”. If your choice is “minor errors” or “major errors”, please specify the details of your concerns. (max. 2000 characters)

The proof is clear. Because of the proof sketch, the intuition and the logic flow of the proof is easy to follow and understand.

6. Significance of empirical results, if applicable:

not applicable

7. Justification of your provided judgment in Q6, if not applicable, put in “N/A”. If your choice is “mixed results” or “generally not significant”, please specify the details of your concerns. (max. 2000 characters)

N/A

8. Novelty of the theoretical results and/or empirical methodology:

New results that are significantly different from existing ones

9. Justification of your provided judgment in Q8. If your choice is “known results” or “incremental/combinatorial” please specify the previous publication(s) that support your choice. (max. 2000 characters)

The theoretical contribution of this paper is quite exciting. Particularly, by proposing a new algorithm based on OFTRL, it shows that the computational time of converging to correlated equilibria in finite-horizon general-sum Markov games can be improved to $\tilde{O}(T^{-1})$ which matches the result of normal-form games and is near-optimal.

10. Non-conventional contributions: does the submission contain non-conventional research contributions? Examples: novel ideas with not widely accepted assumptions, new problems and/or tasks, “bridging fields” contributions. (max. 2000 characters)

N/A

11. Clarity: Is the paper clearly written? Consider whether the paper has clearly (1) stated its contributions, notation and results, (2) explained the meaning of the theoretical assumptions, and/or (3) motivated the proposed methodology well with e.g., examples. (max. 2000 characters)

The paper is well-written and the proof is clear and easy to follow.

12. Relation to prior work: Is it clearly discussed how this work differs from or relates to prior work in the literature? Any related work missing (if yes provide details in Q13)?

All related works are clearly discussed

14. Reproducibility. Are there enough details to reproduce the main experimental results of this work?

Not applicable: the paper has no experiments

15. Paper score:

5: Top 10% of the accepted papers

16. Confidence score:

4: You are confident in your assessment but not absolutely certain. (It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.)

17. Does the submission raise potential ethical concerns? Does this submission raise potential ethical concerns? These include methods, applications or data that create or reinforce unfair biases and/or that have a primary purpose of harm or injury.

No concern

18. Justification of your provided judgment in Q17 about ethical concerns, if not applicable, put in “N/A”. (max. 2000 characters)

N/A

19. Code of conduct. Agree to abide by the AISTATS code of conduct The AISTATS code of conduct can be found here: <https://aistats.org/aistats2024/code-of-conduct.html>

Agreement accepted

20. Confidentiality agreement. Reviewers must keep the paper and supplementary materials (including code submissions and LaTeX source), as well as the reviews, confidential. This includes deleting any submitted code at the end of the review cycle to comply with confidentiality requirements.

Agreement accepted