

Breaking the Curse of Multiagents in a Large State Space: RL in Markov Games with Independent Linear Function Approximation

Qiwen Cui

University of Washington

QWCUI@CS.WASHINGTON.EDU

Kaiqing Zhang

University of Maryland, College Park

KAIQING@UMD.EDU

Simon S. Du

University of Washington

SSDU@CS.WASHINGTON.EDU

Editors: Gergely Neu and Lorenzo Rosasco

Abstract

We propose a new model, *independent linear Markov game*, for multi-agent reinforcement learning with a large state space and a large number of agents. This is a class of Markov games with *independent* linear function approximation, where each agent has its own function approximation for the state-action value functions that are *marginalized* by other players' policies. We design new algorithms for learning the Markov coarse correlated equilibria (CCE) and Markov correlated equilibria (CE) with sample complexity bounds that only scale polynomially with *each agent's own function class complexity*, thus breaking the *curse of multiagents*. In contrast, existing works for Markov games with function approximation have sample complexity bounds that scale with the size of the *joint action space* when specialized to the canonical tabular Markov game setting, which is exponentially large in the number of agents.

Our algorithms rely on two key technical innovations:

- Utilizing *policy replay* to tackle *non-stationarity* incurred by multiple agents and the use of function approximation;
- Separating *learning* Markov equilibria and *exploration* in the Markov games, which allows us to use the *full-information* no-regret learning oracle instead of the stronger bandit-feedback no-regret learning oracle used in the tabular setting.

Our algorithm achieves $\tilde{O}(\text{poly}(m, d_{\max}, H, 1/\epsilon))$ sample complexity to output an ϵ -approximate Markov coarse correlated equilibrium, where m is the number of the players, d_{\max} is the maximum feature dimension of all the players, H is the planning horizon. This is the first result that breaks the curse of multiagents for Markov games with function approximation.

In the tabular case, by adapting the policy replay mechanism for independent linear Markov games, we propose an algorithm with $\tilde{O}(\epsilon^{-2})$ sample complexity to learn Markov CCE, which improves the state-of-the-art result $\tilde{O}(\epsilon^{-3})$ in (Daskalakis et al., 2022), where ϵ is the desired accuracy, and also significantly improves other problem parameters. In addition, we design the first provably efficient algorithm for learning Markov CE that breaks the curse of multiagents.

Furthermore, we propose an iterative-best-response type algorithm that can learn pure Markov Nash equilibria in independent linear Markov potential games, with applications in learning in congestion games.*

Concurrent to this work, Wang et al. (2023) also considered learning in Markov games with independent function approximation. Please refer to our arXiv version for a detailed comparison.

Keywords: reinforcement learning theory, multi-agent reinforcement learning, game theory

*Extended abstract. Full version appears as [arXiv reference, 2302.03673].

References

Constantinos Daskalakis, Noah Golowich, and Kaiqing Zhang. The complexity of markov equilibrium in stochastic games. *arXiv preprint arXiv:2204.03991*, 2022.

Yuanhao Wang, Qinghua Liu, Yu Bai, and Chi Jin. Breaking the curse of multiagency: Provably efficient decentralized multi-agent rl with function approximation. *arXiv preprint arXiv:2302.06606*, 2023.