先端科学技術研究科　修士論文要旨

所属研究室 (主指導教員)	知能システム制御 (杉本　謙二　教授)		
学籍番号	2011088	提出日	令和 4年 1月 7日
学生氏名	北村　俊徳		
論文題目	Dynamic KL Regularization in Reinforcement Learning: Theoretical Error Propagation Analysis and an Algorithm 強化学習における動的なKL正則化の誤差伝搬解析とアルゴリズムへの応用		

要旨

The recent boom in the literature on entropy-regularized value iteration approaches reveals that Kullback-Leibler (KL) regularization brings advantages to reinforcement learning (RL) algorithms by canceling out errors under mild assumptions. However, existing analyses focus on fixed regularization with a constant weighting coefficient and do not consider cases where the coefficient is allowed to change dynamically. In this paper, we study the dynamic coefficient scheme and present the first asymptotic error bound. Based on the dynamic coefficient error bound, we propose an effective scheme to tune the coefficient according to the magnitude of error in favor of more robust learning. Complementing this development, we propose a novel algorithm, Geometric Value Iteration (GVI), that features a dynamic error-aware KL coefficient design with the aim of mitigating the impact of errors on performance. Our experiments demonstrate that GVI can effectively exploit the trade-off between learning speed and robustness over uniform averaging of a constant KL coefficient. The combination of GVI and deep networks shows stable learning behavior even in the absence of a target network, where algorithms with a constant KL coefficient would greatly oscillate or even fail to converge.