

Exploiting Action Impact Regularity and Exogenous State Variables for Offline Reinforcement Learning (Abstract Reprint)

Vincent Liu¹, James R. Wright¹, Mrtha White¹

¹University of Alberta and Alberta Machine Intelligence Institute (Amii), Edmonton, Alberta, Canada

Abstract Reprint. This is an abstract reprint of a journal article by Liu, Wright, and White (2023).

Abstract

Offline reinforcement learning—learning a policy from a batch of data—is known to be hard for general MDPs. These results motivate the need to look at specific classes of MDPs where offline reinforcement learning might be feasible. In this work, we explore a restricted class of MDPs to obtain guarantees for offline reinforcement learning. The key property, which we call Action Impact Regularity (AIR), is that actions primarily impact a part of the state (an endogenous component) and have limited impact on the remaining part of the state (an exogenous component). AIR is a strong assumption, but it nonetheless holds in a number of real-world domains including financial markets. We discuss algorithms that exploit the AIR property, and provide a theoretical analysis for an algorithm based on Fitted-Q Iteration. Finally, we demonstrate that the algorithm outperforms existing offline reinforcement learning algorithms across different data collection policies in simulated and real world environments where the regularity holds.

References

Liu, V.; Wright, J. R.; and White, M. 2023. Exploiting Action Impact Regularity and Exogenous State Variables for Offline Reinforcement Learning. *Journal of Artificial Intelligence Research*, 77: 71–101.