

Hybrid Least-Squares Algorithms for Approximate Policy Evaluation^{*}

Jeff Johns, Marek Petrik, and Sridhar Mahadevan

Department of Computer Science
University of Massachusetts Amherst
{johns,petrik,mahadeva}@cs.umass.edu

Abstract. The goal of approximate policy evaluation is to “best” represent a target value function according to a specific criterion. Different algorithms offer different choices of the optimization criterion. Two popular least-squares algorithms for performing this task are the *Bellman residual* method, which minimizes the Bellman residual, and the *fixed point* method, which minimizes the *projection* of the Bellman residual. When used within policy iteration, the fixed point algorithm tends to ultimately find better performing policies whereas the Bellman residual algorithm exhibits more stable behavior between rounds of policy iteration. We propose two *hybrid least-squares algorithms* to try to combine the advantages of these algorithms. We provide an analytical and geometric interpretation of hybrid algorithms and demonstrate their utility on a simple problem. Experimental results on both small and large domains suggest hybrid algorithms may find solutions that lead to better policies when performing policy iteration.

Reference

1. Johns, J., Petrik, M., Mahadevan, S.: Hybrid Least-Squares Algorithms for Approximate Policy Evaluation. Machine Learning (2009) DOI: 10.1007/s10994-009-5128-4

^{*} This is an extended abstract of an article published in the machine learning journal [1].