

Factored Value Iteration Converges

István Szita* and András Lőrincz†

Abstract

In this paper we propose a novel algorithm, factored value iteration (FVI), for the approximate solution of factored Markov decision processes (fMDPs). The traditional approximate value iteration algorithm is modified in two ways. For one, the least-squares projection operator is modified so that it does not increase max-norm, and thus preserves convergence. The other modification is that we uniformly sample polynomially many samples from the (exponentially large) state space. This way, the complexity of our algorithm becomes polynomial in the size of the fMDP description length. We prove that the algorithm is convergent. We also derive an upper bound on the difference between our approximate solution and the optimal one, and also on the error introduced by sampling. We analyse various projection operators with respect to their computation complexity and their convergence when combined with approximate value iteration.

Keywords: factored Markov decision process, value iteration, reinforcement learning

1 Introduction

Markov decision processes (MDPs) are extremely useful for formalising and solving sequential decision problems, with a wide repertoire of algorithms to choose from [4, 26]. Unfortunately, MDPs are subject to the ‘curse of dimensionality’ [3]: for a problem with m state variables, the size of the MDP grows exponentially with m , even though many practical problems have polynomial-size descriptions. Factored MDPs (fMDPs) may rescue us from this explosion, because they offer a more compact representation [17, 5, 6]. In the fMDP framework, one assumes that dependencies can be factored to several easy-to-handle components.

For MDPs with known parameters, there are three basic solution methods (and, naturally, countless variants of them): value iteration, policy iteration and linear programming (see the books of Sutton & Barto [26] or Bertsekas & Tsitsiklis [4] for an excellent overview). Out of these methods, linear programming is generally

*Eötvös Loránd University, Hungary, Department of Information Systems, E-mail: szityu@gmail.com

†Eötvös Loránd University, Hungary, Department of Information Systems, E-mail: andras.lorincz@inf.elte.hu. Please send correspondence to András Lőrincz.