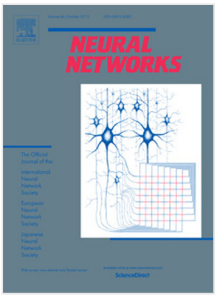
Accepted Manuscript



Model-based reinforcement learning with dimension reduction

Voot Tangkaratt, Jun Morimoto, Masashi Sugiyama

PII: S0893-6080(16)30109-5
DOI: http://dx.doi.org/10.1016/j.neunet.2016.08.005
Reference: NN 3653

To appear in: *Neural Networks*

Received date: 3 February 2016
Revised date: 2 August 2016
Accepted date: 16 August 2016

Please cite this article as: Tangkaratt, V., Morimoto, J., & Sugiyama, M. Model-based reinforcement learning with dimension reduction. *Neural Networks* (2016), http://dx.doi.org/10.1016/j.neunet.2016.08.005

Model-based Reinforcement Learning with Dimension Reduction

Voot Tangkaratt[1], Jun Morimoto[2], and Masashi Sugiyama[3]

[1]Department of Computer Science, The University of Tokyo, Japan
voot@ms.k.u-tokyo.ac.jp
[2]Department of Brain Robot Interface, ATR Computational Neuroscience Laboratory, Japan
xmorimo@atr.jp
[3]Department of Complexity Science and Engineering, The University of Tokyo, Japan
sugi@k.u-tokyo.ac.jp

Abstract

The goal of reinforcement learning is to learn an optimal policy which controls an agent to acquire the maximum cumulative reward. The *model-based* reinforcement learning approach learns a transition model of the environment from data, and then derives the optimal policy using the transition model. However, learning an accurate transition model in high-dimensional environments requires a large amount of data which is difficult to obtain. To overcome this difficulty, in this paper, we propose to combine model-based reinforcement learning with the recently developed *least-squares conditional entropy* (LSCE) method, which simultaneously performs transition model estimation and dimension reduction. We also further extend the proposed method to imitation learning scenarios. The experimental results show that policy search combined with LSCE performs well for high-dimensional control tasks including real humanoid robot control.

Keywords: Model-based reinforcement learning, Transition model estimation, Sufficient dimension reduction

1 Introduction

The goal of reinforcement learning is to learn an optimal policy which controls an agent in an unknown environment to acquire the maximum cumulative reward. Reinforcement learning is appealing because it does not assume strong prior knowledge about the environment and only requires that the agent can collect data through interaction