Single-step Dimension Reduction for High-dimensional Data Analysis with Application in Reinforcement Learning

	2017-03-23
URL	http://doi.org/10.15083/00076185

論文題目: Single-step Dimension Reduction for High-dimensional
Data Analysis with Application in Reinforcement
Learning
（高次元データ解析のためのシングルステップ次元削減と
その強化学習への応用）

氏　　名　　タンカラット　ブット　(Tangkaratt Voot)

Machine learning has become an important field due to the necessity of analyzing complex data. Machine learning is categorized into three paradigms: supervised learning, unsupervised learning, and reinforcement learning. Supervised and unsupervised learning have a long history and they have been extensively studied in the machine learning literature. However, they still have a weakness when they are applied to high-dimensional problems. Reinforcement learning has also been studied for decades, but its study has been rather limited and it has more room for improvement. In this dissertation, we firstly propose a dimension reduction method to mitigate the high dimensionality limitation in supervised and unsupervised learning. Secondly, we propose a dimension reduction method for the conditional density estimation problem. Then, we propose a model-based reinforcement learning method which does not rely on strong assumptions unlike existing methods, and we subsequently improve its performance by utilizing our dimension reduction method. Finally, we propose a contextual reinforcement learning method that effectively learns from high-dimensional contexts by utilizing the idea of dimension reduction.

In the first part of this dissertation, we focus on developing two single-step dimension reduction methods. Dimension reduction is a standard tool in machine learning and many methods were proposed in literature. However, existing methods are sensitive to outliers which are common phenomena in practice. To solve this problem, we propose a dimension reduction method based on the maximization of quadratic mutual information (QMI) which is a robust statistical dependence measure. Solving this maximization

problem requires an accurate estimate of the derivative of QMI. A common approach first estimates QMI from data and then computes the derivatives of the estimated QMI. However, this multi-step approach is not appropriate since the derivative of an accurate QMI estimator is not necessarily an accurate estimator of the derivative of QMI. Instead, we propose to directly estimate the derivative of QMI in a single-step manner without estimating QMI itself. Experimental evaluations on high-dimensional regression problems show that our single-step QMI-based dimension reduction method works better and is more robust against outliers than existing methods.

While our QMI-based dimension reduction method is useful for both supervised and unsupervised learning problems, it may not be the optimal method for a supervised learning problem called conditional density estimation. Conditional density estimation aims to estimate a conditional probability density of output given input, and is highly useful for analyzing a relationship between input and output. However, accurately estimating a conditional density from high-dimensional input is challenging. Moreover, a multi-step approach which first performs dimension reduction and then performs conditional density estimation is not always appropriate. To solve this problem, we propose the least-squares conditional entropy (LSCE) method which simultaneously performs non-parametric conditional density estimation and dimension reduction in an integrated manner. Experimental evaluations on high-dimensional conditional density estimation problems show that LSCE estimates conditional densities more accurately than methods based on the multi-step approach.

In the second part of this dissertation, we focus on utilizing dimension reduction in high-dimensional reinforcement learning problems. The goal of reinforcement learning is to learn an optimal policy which controls an agent to receive the maximum cumulative rewards. Among reinforcement learning methods, policy gradient methods are widely applicable. However, accurately estimating policy gradients requires a large amount of data. The model-based approach can cope with this issue by first estimating a transition model from data and then using the model to accurately estimate policy gradients. An advantage of the model-based approach over the model-free approach is that once the transition model is accurately estimated, the agent does not need to collect more data to learn an optimal policy. Moreover, when the budget for collecting data is limited, the model-based approach does not need to determine the sampling schedule, unlike the model-free approach. However, the state-of-the-art method for transition model estimation is based on the strong assumption that the transition dynamics can be accurately modeled by the Gaussian distribution. To avoid assuming such a strong assumption, we propose a model-based policy gradient method which uses the least-squares conditional density estimation (LSCDE) method to estimate a transition model. LSCDE is an existing non-parametric conditional density estimation method and it can asymptotically estimate any conditional density. We evaluate our proposed method through experiments on benchmark problems and a simulated humanoid robot control problem, and show that our model-based

policy gradient method gives better performance than existing methods when the amount of data is limited.

However, LSCDE still requires a relatively large amount of data in high-dimensional reinforcement learning problems. To mitigate this limitation, we propose to improve our model-based policy gradient method by estimating a transition model using our LSCE method. Experimental evaluations on benchmark control problems and a real humanoid robot control problem show that our model-based policy gradient method with LSCE gives the best performance among compared methods.

Although our model-based reinforcement learning method with single-step dimension reduction works well for standard reinforcement learning, it may not be an appropriate method for contextual reinforcement learning. In contextual reinforcement learning, an agent requires to learn different optimal policies for different contexts of a problem. Contextual reinforcement learning is challenging especially when contexts have high dimensionality. To overcome this challenge, we propose a contextual reinforcement learning method where our key idea is to learn a low-rank representation of a model of the cumulative rewards. We show that learning the low-rank representation actually corresponds to a single-step approach to simultaneously performing dimension reduction for model learning. We evaluate our method on a benchmark problem and robot ball hitting problems based on camera images. The experimental results show that our method gives the best performance among compared methods.

In this dissertation, we have proposed five methods that overcome the weaknesses of existing machine learning methods in high-dimensional problems. Based on our empirical evaluations on both benchmark and real-world problems, we conclude that our methods successfully overcome the above mentioned weaknesses of existing machine learning methods.