

Gauss Process Regression

Introduction, Comparison and Analysis

Tzu-Heng Lin, 2014011054, W42
Department of Electronic Engineering, Tsinghua University, Beijing, China
lzhbrian@gmail.com

ABSTRACT

¹ Markov Chain Monte Carlo (MCMC) is a technique to make an estimation of a statistic by simulation in a complex model. Restricted Boltzmann Machine(RBM) is a crucial model in the field of Machine Learning. However, training a large RBM model will include intractable computation of the partition functions, i.e. $Z(\theta)$. This problem has aroused interest in the work of estimation using a MCMC methods. In this paper, we first conduct Metropolis-Hastings Algorithm, one of the most prevalent sampling methods, and analyze its correctness & performance, along with the choice of the accepting rate. We then implement three algorithms: TAP, AIS, RTS, to estimate partition functions of an RBM model. Our work not only give an introduction about the available algorithms, but systematically compare the performance & difference between them. We seek to provide an overall view in the field of MCMC.

In this paper, we first give an overall introduction of the gauss process regression.

1. INTRODUCTION

2. CONCLUSION

In this paper, we discuss about the Markov Chain Monte Carlo method which are now undoubtedly one of the most

important sampling methods.

We comprehensively introduce the concept of Metropolis-Hastings Algorithm and conduct an experiment to verify its correctness. We also make some analysis about how accepting rate would interfere the sampling result.

We systematically compare three methods of partition function estimation which are crucial works in training a Restricted Boltzmann Machine or a Deep Belief Network.

As future work, we would like to join more methods to the comparison and if could, propose some improvement to the algorithms available.

3. ACKNOWLEDGEMENT

¹Tzu-Heng Lin is currently an undergraduate student in the Department of Electronic Engineering, Tsinghua University. His research interests include Big Data Mining, Machine Learning, etc. For more information about him, please see <http://lzhbrian.me>. The code in this paper can be found in www.github.com/lzhbrian/gpr. Please feel free to contact him at any time via lzhbrian@gmail.com or linzh14@mails.tsinghua.edu.cn

I would like to thank Yuanxin Zhang, XueChao Wang, for the discussion with me on the algorithms. Without them, I wouldn't have the possibility to accomplish this work in such a short time. This paper is a project of Stochastic Process Course in Tsinghua University, taught by Prof. Zhijian Ou.