A fast and more accurate Approximate Bayesian Inference approach for Bayesian Lasso regression

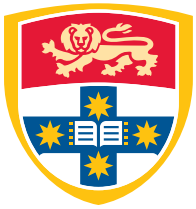
Yuhao Li

A thesis submitted in partial fulfillment of

the requirements for the degree of

Bachelor of Science(Honour)(Data Science)

Mathematics and Statistics



THE UNIVERSITY OF SYDNEY

June 2023

Statement of originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Yuhao Li

Abstract

Variational Approximation: as a deterministic approximation algorithm for Approximate Bayesian Inference(ABI), has been used prevalently among the Bayesian Statistical community for fast Approximate Bayesian Inference(ABI). It is also a faster alternative to Monte Carlo methods such as Markov Chain and Monte Carlo(MCMC). Nevertheless, the Variational Approximation algorithm suffers from the loss of approximation accuracy under some special circumstances, such as underestimating the variance when the correlation of variables becomes large.

In addition, the Lasso penalized regression could be estimated in the Bayesian paradigm to facilitate standard error estimates of coefficients and credible intervals.

In this thesis, we develop two fast and more accurate variational approximation algorithms for the Bayesian Lasso regression problem. The main idea behind these methods involves using the information of local parameter estimates by accommodating the univariate and multivariate lasso distribution, which results in more accurate global parameter approximation by Gaussian Variational Approximation. Our experimental results on numerous real-world datasets suggest their high Variational approximation accuracy with a descent time efficiency, compared with the traditional Monte Carlo methods and Mean-Field Variational Bayes(MFVB).

Acknowledgements

Thanks to Supervisor and family and myself

Finally, I would like to thank my friends my family: my mother, father, brother, grandparents, and cousins. The last couple years has been full of adversity and I could not have overcome it without your support behind the scenes.

Contents

Chapter 1

Introduction

1.1 Motivation

Frequentist Bayesian

Stochastic Sampling based Determinstic Approach

Variational Approximation: an optimimization based technique for approximate bayesian inference, obtain interval estimate and error variance.

1.2 Contribution

Design of a new posterior parameter correction approach based on the posterior estimate of Mean-Field Variation Bayes parameter

1.3 Thesis Organization

This paper will provide a

Chapter 2

Definition and Literature Review

2.1 Variational Inference

2.1.1 Mean Field Variational Bayes

2.2 Expectation Maximization

2.2.1 Bayesian Expectation Maximization

2.3 Markov Chain Monte Carlo(MCMC)

2.3.1 Gibbs Sampler

2.4 Bayesian Paradigm

2.5 Least Absolute Shrinkage and Selection Operator(LASSO) penalized regression

2.5.1 Lasso penalty formulation

2.5.2 Bayesian Lasso regression

2.5.3 Lasso distribution

Univariate Lasso Distribution

Multivariate Lasso Distribution

Chapter 3

Methodlogy

Chapter 4

Experiment Result and Analysis

4.1 Experimental Setting

4.1.1 Parameter selection

4.1.2 Evaluation metric

4.1.3 Experimental datasets

4.2 Experimental Result

Chapter 5

Conclusion

Chapter 6

Limitation and Future Work

Bibliography