Efficient and Robust Black-box Integral-approximation and Optimization

Yueming Lyu

Faculty of Engineering and Information Technology

University of Technology Sydney

A thesis submitted for the degree of

Doctor of Philosophy

September 2021

Certificate of Original Authorship

I hereby declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. The contents of this dissertation are original and have not been submitted in whole or in part for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: 01-October-2021

I would like to dedicate this thesis to my loving parents, who support me all the time.

Acknowledgements

First, I would like to express my deepest gratitude to my supervisor, Prof. Ivor W. Tsang, who gave me the research opportunity in his group. My research journey was not easy. I was depressed and suffered a lot. It was Prof. Ivor W. Tsang who provided me an opportunity to continue research on machine learning. Prof. Ivor W. Tsang always supported me to be patient to do in-depth analysis instead of shallow work. His insightful guidance helped me to focus on the key reasons to solve the problems. Prof. Ivor W. Tsang encouraged me to explore the research area that I am interested. He guided me to know how to think in both theoretical and empirical views of research. Prof. Ivor W. Tsang taught me how to divide and conquer problems and hierarchically solve them. His critical thinking and broad view gave me insightful guidance and let me know how a good researcher should be. Prof. Ivor W. Tsang's advice and support improved my skills and confidence to be an independent researcher. It is a great fortune for me to be a Ph.D. student under his supervision.

Second, I would like to thank many friends who have supported me. I want to thank Dr. Jiangchao Yao and Dr. Yuangang Pan for discussing both research and life. It was a memorable time for waiting for the subway and discussing research together. I appreciate the time with them and Xiaowei Zhou, Xingrui Yu, Xu Chen, Yan Zhang, Jin Li, Yaxin shi, Jinliang Deng, Yinghua Yao for discussion and lunch together. I am so grateful for such a relaxed life and beautiful experience in Sydney. I would also like to thank Dr. Yuan Yuan, Dr. Yanbin Liu, Dr. Xin Yu, Dr. Xiaolin Zhang, Fan Ma, Tianqi Tang, Guangrui Li, and Guang Li for their support and discussion. I would thank Dr. Peng Sun and Dr. Li Shen for the discussion and for working together. I may miss mentioning many others, but I thank them for their help.

Finally, I would like to dedicate this thesis to my dearest parents and family. Thank you so much for being there, supporting me all the time.

Abstract

Black-box optimization and black-box integral approximation are important techniques for machine learning, industrial design, and simulation in science. This thesis investigates black-box integral approximation and black-box optimization by considering the closed relationship between them. For integral approximation, we develop a simple closed-form rank-1 lattice construction method based on group theory. Our method reduces the number of distinct pairwise distance values to generate a more regular lattice. Furthermore, we investigate structured points set for integral approximation on hyper-sphere. Our structured point sets can serve as a good initialization for black-box optimization. Moreover, we propose stochastic black-box optimization with implicit natural gradients for black-box optimization. Our method is very simple and has only the step-size hyper-parameter. Furthermore, we develop a batch Bayesian optimization algorithm from the perspective of frequentist kernel methods, which is powerful for low-dimensional black-box optimization problems. We further apply our structured integral approximation techniques for kernel approximation. In addition, we develop structured approximation for robust deep neural network architecture, which results in an elegant and simple architecture that preserves optimization properties. Moreover, we develop adaptive loss as a tighter upper bound approximation for expected 0-1 risk, robust and trainable with SGD.

Contents

List of Figures

List of Tables

List of Algorithms