Supervised Linear Dimension Reduction

Wei Bian

Faculty of Engineering and Information Technology

University of Technology, Sydney

A thesis submitted for the degree of

Doctor of Philosophy

2012

Certificate of Authorship/Originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature

Acknowledgements

On having completed this thesis, I am especially thankful to my advisor Prof. Dacheng Tao, who had led me to an at one time unfamiliar area of academic research, and trusted me and gave me as much as possible freedom to purse my own research interests. Dacheng has taught me how to think and study independently and how to solve a difficult scientific problem in flexible but rigorous ways. He has sacrificed much of his precious time for developing my academic research skills. He has also given me great help and support in life.

I am thankful to the group members I met in the Hong Kong Polytechnic University, Nanyang Technological University, and University of Technology, Sydney, including Tianyi Zhou, Bo Geng, Chao Zhang, Bo Xie, Yang Mu, and many others. I learned a lot from these smart people, and I was always inspired by the interesting and in-depth discussions with them. I enjoyed the wonderful atmosphere,being with them, of both academic research and daily life.

I am incredibly grateful to my mother for her generosity and encouragement. This thesis is definitely impossible to be completed without her constant support and understanding. I am also thankful to my friends who have companied me, though not always at my side, through the arduous journey of four and a half years.

Abstract

Supervised linear dimension reduction (SLDR) is one of the most effective methods for complexity reduction, which has been widely applied in pattern recognition, computer vision, information retrieval, and multimedia data processing. This thesis explores SLDR by enriching the theory of existing methods and by proposing new methods.

In the first part of this thesis, we present theoretical analysis of Fisher's linear discriminant analysis (LDA), one of the most representative methods for SLDR. 1) Classical asymptotic analysis of LDA is based on a fixed dimensionality, and thus does not apply in the case where the dimensionality and the training sample number are proportionally large. Besides, the classical result does not provide quantitative information on the performance of LDA. To address these limitations, we present an asymptotic generalization analysis of LDA, allowing both the dimensionality and the training sample number to be proportionally large, from which we principally obtain an asymptotic generalization bound that quantitatively describes the performance of LDA in terms of the dimensionality and the training sample number. 2) We study a new regularization method for LDA, termed the block-diagonal regularization. By partitioning variables into small groups and treating them independently, block-diagonal regularization effectively reduces the dimensionality to training sample number ratio and thus improves the generalization ability of LDA. We present a theoretical justification of the block-diagonally regularized LDA by investigating its approximation and sample errors. We show that the block-diagonally regularized LDA performs competitively compared to other types of regularized LDA, e.g., with the Tikhonov regularization and the banded regularization.

In the second part of this thesis, we propose two new methods for SLDR. 1) The first method is for parametric SLDR, termed max-min distance analysis (MMDA). MMDA optimizes the projection matrix by maximizing the minimum pairwise distance of all class pairs in the dimension reduced space. Thus, it duly considers the separation of all classes and overcomes the "class separation" problem of existing parametric SLDR methods that close class pairs tend to merge in the dimension reduced space. 2) The second method is for nonparametric SLDR, which uses minimizing the asymptotic nearest neighbor classification error (MNNE) as the criterion for optimizing the projection matrix. Theoretically, we compare MNNE with other criteria, e.g., maximizing mutual information (MMI) and minimizing Bhattacharyya bound. We show that MMNE is superior to these two criteria in terms of the closeness to the Bayes optimal criterion. Empirical studies show that the proposed methods, MMDA and MNNE, achieve state-of-the-art performance for parametric and nonparametric SLDR, respectively.

Contents

List of Figures

List of Tables