Accepted Manuscript

Training sparse least squares support vector machines by the QR decomposition

Xiao-Lei Xia

PII: S0893-6080(18)30207-7
DOI: https://doi.org/10.1016/j.neunet.2018.07.008
Reference: NN 3991

To appear in: *Neural Networks*

Received date : 2 November 2017
Revised date : 7 July 2018
Accepted date : 11 July 2018

Please cite this article as: Xia, X.-L., Training sparse least squares support vector machines by the QR decomposition. *Neural Networks* (2018), https://doi.org/10.1016/j.neunet.2018.07.008

Training Sparse Least Squares Support Vector Machines by the QR Decomposition

Xiao-Lei Xia[a,1,*]

[a] *Centre for Cancer Research and Cell Biology, School of Medicine, Dentistry and Biomedical Sciences Queen's University Belfast, BT9 7BL, UK*

Abstract

The solution of an LS-SVM has suffered from the problem of non-sparseness. The paper proposed to apply the KMP algorithm, with the number of support vectors as the regularization parameter, to tackle the non-sparseness problem of LS-SVMs. The idea of the kernel matching pursuit (KMP) algorithm was first revisited from the perspective of the QR decomposition of the kernel matrix on the training set. Strategies are further developed to select those support vectors which minimizes the leave-one-out cross validation error of the resultant sparse LS-SVM model. It is demonstrated that the LOOCV of the sparse LS-SVM can be computed accurately and efficiently. Experimental results on benchmark datasets showed that, compared to the SVM and variants sparse LS-SVM models, the proposed sparse LS-SVM models developed upon KMP algorithms maintained comparable performance in terms of both accuracy and sparsity.

Keywords: least-squares support vector machines, kernel matching pursuit, QR decomposition, sparseness

1. Introduction

Support Vector Machines (SVM) [22] are a family of algorithms for classification and regression and have enjoyed widespread applications to various pattern recognition tasks since its introduction two decades ago. The standard SVM follows the structural risk minimization principle and is formulated as a quadratic programming (QP) problem subject to inequality constraints. As opposed to the SVM, the least-squares SVM (LS-SVM) which is an important variant of the SVM, adopts equality constraints [20]. Gestel et al. gave a Bayesian perspective on the formulation of the LS-SVM and also investigated its connection with Gaussian Process and kernel fisher discriminant analysis [8]. On the other

* Corresponding author
Email address: xxia01@qub.ac.uk (Xiao-Lei Xia)
[1] Tel: 44-2890975828