A Prediction Interval Estimation Method for KMSE

Changha Hwang[1], Kyung Ha Seok[2], and Daehyeon Cho[3]

[1] Division of Information and Computer Sciences,
Dankook University, Seoul 140-714, Korea
chwang@dankook.ac.kr
[2] Corresponding Author, Department of Data Science,
Inje University, Kyungnam 621-749, Korea
skh@stat.inje.ac.kr
[3] Department of Data Science,
Inje University, 621-749, Kyungnam, Korea
cho@stat.inje.ac.kr

Abstract. The kernel minimum squared error estimation(KMSE) model can be viewed as a general framework that includes kernel Fisher discriminant analysis(KFDA), least squares support vector machine(LS-SVM), and kernel ridge regression(KRR) as its particular cases. For continuous real output the equivalence of KMSE and LS-SVM is shown in this paper. We apply standard methods for computing prediction intervals in nonlinear regression to KMSE model. The simulation results show that LS-SVM has better performance in terms of the prediction intervals and mean squared error(MSE). The experiment on a real date set indicates that KMSE compares favorably with other method.

1 Introduction

In forecasting tasks the prediction interval gives the range in which you could have a certain level of confidence of finding an individual value of the predicted output for a given input value. Truly reliable prediction systems require the prediction to be qualified by a confidence measure such as prediction or confidence interval. This important issue has received little systematic study. However, this has been paid attention in neural information processing and chemical engineering communities.

Chryssolouris [1] has derived a technique to quantify the confidence intervals for the prediction of neural network models by adopting a variant of the linearisation methodology. Shao et al. [7] have proposed a novel method of computing confidence bounds on predictions from a neural network with determined structure. De Veaux et al. [3] also have proposed a method of computing prediction intervals for neural networks and compared them with prediction intervals based on multivariate adaptive regression splines using generalized additive model (MARS/GAM). Yang et al. [13] have suggested a method of estimating confidence bound for neural networks for the purpose of the prediction of rock

L. Wang, K. Chen, and Y.S. Ong (Eds.): ICNC 2005, LNCS 3610, pp. 536–545, 2005.

porosity values from seismic data for oil reservoir characterization. Seok et al. [6] have presented a Bayesian approach to computing the prediction intervals for support vector machine(SVM) regression and shown SVM regression achieves better performances than the neural networks and MARS in predicting intervals. There are some other literatures related to this issue.

SVM, originally introduced by Vapnik, solves the weak point of neural network such as the existence of local minima in the area of statistical learning theory and structural risk minimization(Vapnik [11]). One of its prominent advantages is the idea of using kernels to realize the nonlinear transforms without knowing the detailed transforms. According to this idea, other authors proposed a class of kernel-based algorithms, such as the kernel Fisher discriminant analysis(KFDA)(Mika et al. [4]), the least squares support vector machine(LS-SVM)(Suykens and Vandewalle [9] , Suykens et al. [10]), and the kernel ridge regression(KRR)(Saunders et al. [5]).

Xu et al. [12] have generalized the conventional minimum squared error method to yield a new type of nonlinear learning machine, by using the kernel idea and adding different regularization terms. They have named the proposed learning machines as KMSE algorithm. KMSE algorithm adopts the idea of kernel function of SVM which is one of the most influential developments in the machine learning.

KMSE model can be viewed as a general framework that includes KFDA, LS-SVM, and KRR as its particular cases. Suykens et al. [8] have proposed a large scale algorithm for LS-SVM by implementing a Hestenes-Stiefel conjugate gradient algorithm for solving the linear equation system. A large scale algorithm for KMSE can be derived without any difficulty by using this idea.

In this paper we discuss a method to compute prediction intervals by applying standard methods for computing prediction intervals in nonlinear regression to the KMSE for regression tasks. The simulation results show that LS-SVM has better performance in terms of the prediction intervals and MSE. The experiment on a real date set indicates that LS-SVM compares favorably with MARS/GAM.

The rest of this paper is organized as follows. Section 2 gives an overview of LS-SVM and KMSE. Section 3 discusses briefly how to compute prediction intervals for KMSE model. Section 4 illustrates the method with a computer generated data and a real data from a polymer process.

2 LS-SVM and KMSE

Let the training data set D be denoted by $\{(\mathbf{x}_k, y_k), k = 1, \ldots, n\}$, with each input $\mathbf{x}_k \in R^d$ and the output $y_k \in R$. It is commonly assumed that

$$y = f(\mathbf{x}, \boldsymbol{\alpha}^*) + \epsilon, \tag{1}$$

where ϵ is independently and identically distributed with zero mean and $\boldsymbol{\alpha}^*$ is the true value of parameters.

Assume a nonlinear function $\boldsymbol{\varphi}(\mathbf{x}) : R^d \rightarrow R^h$ maps the input space to a so-called higher dimensional feature space. It is important to note that the dimension h of this space is defined only in an implicit way.