Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 79 (2009) 2051–2063
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Linear programming support vector regression with wavelet kernel: A new approach to nonlinear dynamical systems identification

Zhao Lu [a], Jing Sun [b,*], Kenneth R. Butts [c]

[a] *Department of Electrical Engineering, Tuskegee University, Tuskegee, AL 36088, USA*
[b] *Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI 48109, USA*
[c] *North America Technical Center, Toyota Motor Corporation, Ann Arbor, MI 48105, USA*

Received 6 September 2007; received in revised form 2 August 2008; accepted 30 October 2008
Available online 18 November 2008

Abstract

Wavelet theory has a profound impact on signal processing as it offers a rigorous mathematical framework to the treatment of multiresolution problems. The combination of soft computing and wavelet theory has led to a number of new techniques. On the other hand, as a new generation of learning algorithms, support vector regression (SVR) was developed by Vapnik et al. recently, in which ε-insensitive loss function was defined as a trade-off between the robust loss function of Huber and one that enables sparsity within the SVs. The use of support vector kernel expansion also provides us a potential avenue to represent nonlinear dynamical systems and underpin advanced analysis. However, for the support vector regression with the standard quadratic programming technique, the implementation is computationally expensive and sufficient model sparsity cannot be guaranteed. In this article, from the perspective of model sparsity, the linear programming support vector regression (LP-SVR) with wavelet kernel was proposed, and the connection between LP-SVR with wavelet kernel and wavelet networks was analyzed. In particular, the potential of the LP-SVR for nonlinear dynamical system identification was investigated.
© 2008 IMACS. Published by Elsevier B.V. All rights reserved.

Keywords: Support vector regression; Wavelet kernel; Nonlinear systems identification; Linear programming

1. Introduction

Mathematical models that capture the behavior of dynamical systems are of great importance in almost all fields of science and engineering, specifically in control, signal processing and information science. Given that most systems encountered in the real world are complex and nonlinear, one challenge in developing a useful model is to achieve a proper trade-off between model simplicity and accuracy. Since a model is always only an approximation of real phenomena, having an approximation theory which allows for the analysis of model quality is of substantial importance. A fundamental principle in system modeling is the well-recognized Occam's razor hypothesis: 'plurality should not be posited without necessity', or in other words, the simpler a solution is, the more reasonable it is. This concept, known as the parsimonious principle, which ensures the simplest possible model that explains the data, is particularly relevant in nonlinear model building because the size of a nonlinear model can easily become explosively large.

* Corresponding author.
E-mail address: jingsun@umich.edu (J. Sun).

0378-4754/$36.00 © 2008 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.matcom.2008.10.011

Forward selection using the orthogonal least squares (OLS) is an effective construction method that is capable of producing parsimonious linear-in-the-weights nonlinear models with excellent generalization performance [3]. Alternatively, the state-of-the-art sparse kernel modeling techniques, such as the support vector machine (SVM) [17,18] and relevant vector machine [9], have been gaining popularity in data modeling applications. SVM algorithms yield prediction functions that are expanded on a subset of training vectors, or support vectors, hence their names. Sparsity, defined as the ratio of the number of support vectors over the number of data points in the training set, is used to measure the model size and simplicity, thereby allowing the evaluation of the model quality against the parsimonious principle. For linear approximation, it has been pointed out in [2] that the solution found by SVM for regression is a trade-off between sparsity of the representation and closeness to the data. SVMs extend this linear interpretation to nonlinear approximation by mappings to a higher-dimensional feature space. This space can be of very high dimension, even infinite, because the parameters of weights are not explicitly calculated. By using certain kernel functions in the approximation function, nonlinear mappings can be made from input space to output, while the training procedure is concerned with linear mappings in an implied feature space.

In the conventional quadratic programming support vector machines (QP-SVMs), the prediction function yielded often contains redundant terms. The economy of an SVM prediction model is dependent on a sparse subset of the training data being selected as support vectors by the optimization technique. In many practical applications, the inefficiency of the conventional SVMs scheme for selecting support vectors could lead to infeasible models. This is particularly apparent in regression applications where the entire training set can be selected as support vectors if error insensitivity is not included [5].

A recent study has compared the standard SVM and uniformly regularized orthogonal least squares (UROLS) algorithms using time series prediction problems, and has found that both methods have similar excellent generalization performance but the resulting model from SVM is not sparse enough [12]. It is explained that the number of support vectors found by quadratic programming algorithm in a SVM is only an upper bound on the number of necessary and sufficient support vectors, and this is due to the linear dependencies between support vectors in feature space. Some efforts have been made attempt to control the sparsity in support vector machines [5].

Among a number of successful applications of SVMs in practice, it has been shown that the use of support vector kernel expansion also provides us a potential avenue to represent nonlinear dynamical systems and underpin advanced analysis [6,7,16]. Although it is believed that the formulation of SVM embodies the structural risk minimization principle, thus combining excellent generalization properties with a sparse model representation, data modeling practicians have begun to realize that the capability for the standard quadratic programming SVR (QP-SVR) method to produce sparse models has perhaps been overstated. For example, it has been shown that the standard SVM technique is not always able to construct parsimonious models in system identification [6].

In this article, for the purpose of developing an innovative and efficient identification algorithm for complex nonlinear dynamical systems, the issue of model sparsity was addressed from two different perspectives. First, the linear programming support vector regression (LP-SVR) is used to capitalize on the advantages of the model sparsity, the flexibility in using more general kernel functions, and the computational efficiency of linear programming [8,10], as compared to OP-SVR. The idea of LP-SVR is to use the kernel expansion as an ansatz for the solution, but to use a different regularizer, namely the ℓ_1 norm of the coefficient vector. In other words, for LP-SVR, the nonlinear regression problem is treated as a linear one in the kernel space, rather than in the feature space as in the case of QP-SVR. Second, considering the transient characteristic of nonlinear dynamical systems, an appropriate kernel function which is capable to capture the underlying nonstationary dynamics accurately might be expected to yield a more compact and sparse representation. Due to the localization feature in both frequency and time domains, wavelets have been successfully used to represent a much larger class of signals than Fourier representation [1]. Unlike Fourier-based analyses that use global sine and cosine functions as bases, wavelet analysis use bases that are localized in time and frequency to represent nonstationary signals more effectively. As a result, a wavelet expansion representation is much more compact and easier to implement.

This paper focuses on developing a new machine learning algorithm by combining the wavelet kernel function with LP-SVR, and particularly exploring their strength in identification of complex nonlinear dynamical systems. Special attention is paid to the sparsity of the generated model and its role in reducing the generalization error. This paper is organized as follows. In the next section, a brief review about wavelet and wavelet networks are given. The algorithm of LP-SVR with wavelet kernel is developed and discussed in Section 3. A case study with application to nonlinear dynamical system identification is conducted in Section 4, with concluding remarks in Section 5.

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات