



Regularized least squares support vector regression for the simultaneous learning of a function and its derivatives

Jayadeva^a, Reshma Khemchandani^b, Suresh Chandra^{b,*}

^a IBM India Research Lab, Block – C, Institutional Area Vasant Kunj, New Delhi 110 070, India

^b Department of Mathematics, Indian Institute of Technology, Hauz Khas, New Delhi 110 016, India

ARTICLE INFO

Article history:

Received 17 March 2007

Received in revised form 7 April 2008

Accepted 17 April 2008

Keywords:

Support vector machines

Regularized least squares

Machine learning

Function approximation

ABSTRACT

In this paper, we propose a regularized least squares approach based support vector machine for simultaneously approximating a function and its derivatives. The proposed algorithm is simple and fast as no quadratic programming solver needs to be employed. Effectively, only the solution of a structured system of linear equations is needed.

© 2008 Published by Elsevier Inc.

1. Introduction

The last decade has witnessed the evolution of Support Vector Machines (SVMs) as a powerful paradigm for pattern classification and regression [2,4,23]. Support vector machines techniques have been successfully applied to a wide variety of areas, e.g. computer vision [7], image retrieval [26], etc. Support Vector Regression (SVR) attempts to fit a regressor through a given set of data samples, where the points may be in the pattern space or in a higher-dimensional feature space. SVR involves the solution of a quadratic minimization problem with linear inequality constraints which also minimizes the structural complexity.

Estimation of a function along with its derivatives, or partial derivatives, is an important problem with diverse applications [9,15]. Derivatives of estimated static relations are often used for linearization in control and in extended Kalman filtering [3]. A wide variety of real problems, especially in linear circuits are concerned with nonparametric estimation of a function and its derivative [18]. The automatic estimation of derivatives can be employed for the construction of asymptotic local confidence intervals for the nonparametric estimate of the regression function and its first derivatives [16]. Other applications include Optimal control, optimization processes, computer graphics and image processing. For instance, in volumetric modelling, or when modelling shapes, a technique that can estimate the shape from measurements of the height as well as gradient or curvature at a few points would be able to find a better representation of the shape. In many cases, the curvature or gradient may be available at a different set of points than at which the height of the surface is available. An example is modelling a terrain from a remote sensing data as well as ground data. Although function approximation methods have been widely applied to modelling device characteristics, as in [14], such modelling is limited unless it can also estimate the derivatives or partial derivatives, which carry important information about nonlinearity in the device's behaviour. Information about

* Corresponding author. Tel.: +91 11 26591479; fax: +91 11 26581005.

E-mail addresses: jayadeva@ee.iitd.ac.in (Jayadeva), reshmaiidt@gmail.com (R. Khemchandani), chandras@maths.iitd.ac.in (S. Chandra).

derivatives or partial derivatives helps determine the region of operation of a nonlinear device, and using samples of such information can make the representation of the device characteristics more accurate and compact. Additionally, a technique that uses such information can benefit by requiring fewer samples to obtain a representation with a desired accuracy. Further, Lamers and Kok [10] used variances of the first order derivatives as a measure of nonlinearity in the training data.

It is often easy to obtain the derivative information as compare to information regarding the functional value. For instance, it is easier to determine the velocity of a moving vehicle as compared to its absolute position.

Recently, Lázaro et al. [11,12] proposed a SVM based approach for the simultaneous learning of a function and its derivatives. This approach employs an iterative weighted least squares (IRWLS) procedure, which has earlier been applied in the regular SVM literature for classification and regression [17].

In this paper, the problem of simultaneous learning of a function and its derivative is formulated in a regularized least squares support vector machines framework. The main advantage of the proposed approach is that it obtains the solution by inverting a single positive definite matrix instead of solving a quadratic programming problem (QPP). The algorithm is simple and fast as the solution is obtained by solving a structured system of linear equations, instead of a single QPP of large size.

The paper is organized as follows: Section 2 briefly dwells on Lázaro et al. approach. Section 3 describes Regularized Least Squares Support Vector Regression for the Simultaneous Learning of a Function and its Derivatives. Section 4 deals with the experimental results, while Section 5 is devoted to the concluding remarks.

2. The SVR approach of Lázaro et al.

In this section, we first describe the model of Lázaro et al. [11,12] for the simultaneous learning of a real valued function and its derivative. Consider a set of M data samples $\{(x^{(i)}, y_i, y_i') \in \mathbf{R} \times \mathbf{R} \times \mathbf{R}, i = 1, 2, \dots, M\}$, where $x^{(i)}$ denotes the i th sample, and y_i (respectively, y_i') denotes the value of the function (respectively, its derivative) at the i th sample. The proposed SVR determines a regressor of the form $f(x) = w^T \phi(x) + b$, where $\phi : \mathbf{R} \rightarrow \mathbf{R}^d$, and $d > 1$ is the dimension of a suitable feature space. Lázaro et al. [11] proposed the following optimization problem for finding the desired regressor

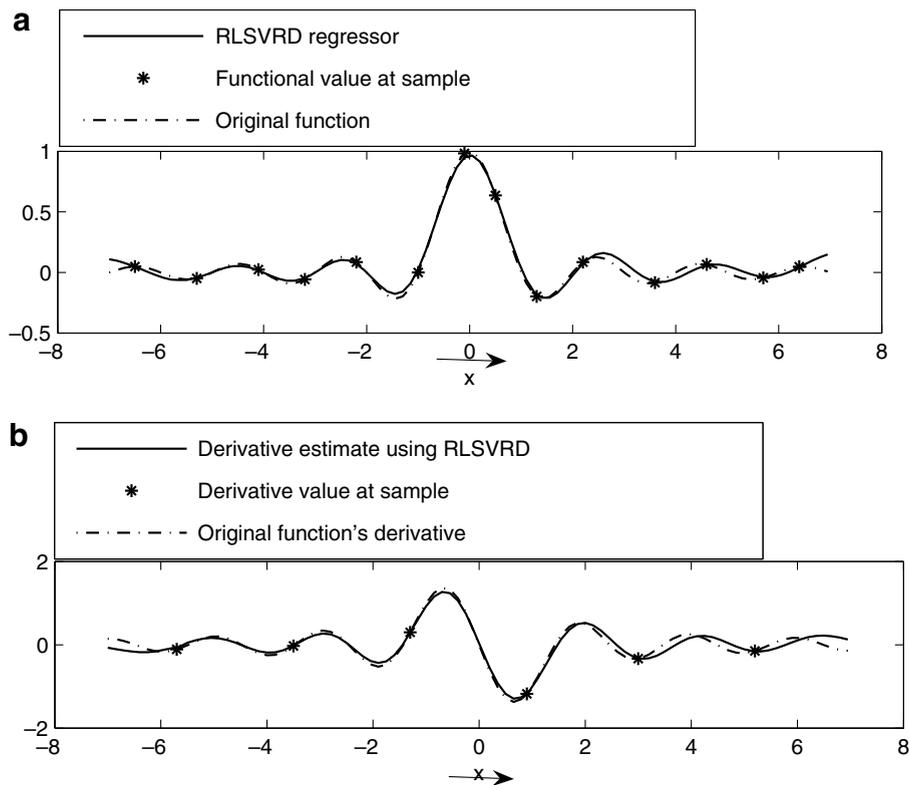


Fig. 1. Plot of (a) the RLSVRD regressor obtained by using 14 points with only function information, and six points with only derivative information, and (b) the RLSVRD derivative computed from the learnt regressor by using (32). Training samples were chosen randomly from the interval $[-7, 7]$, and are depicted as $*$. The Gaussian kernel with $\sigma = 0.9$ was used.

متن کامل مقاله

دریافت فوری ←

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