Generalized recurrent neural network for $\epsilon$-insensitive support vector regression

Yan Zhao $^a$, Qingshan Liu $^b$,*

$^a$ Department of Basic Courses, Wannan Medical College, Wuhu 241000, China
$^b$ School of Automation, Southeast University, Nanjing 210096, China

Received 24 July 2010; received in revised form 17 March 2011; accepted 9 March 2012
Available online 23 August 2012

Abstract

In this paper, a generalized recurrent neural network is proposed for solving $\epsilon$-insensitive support vector regression ($\epsilon$-ISVR). The $\epsilon$-ISVR is first formulated as a convex non-smooth programming problem, and then a generalize recurrent neural network with lower model complexity is designed for training the support vector machine. Furthermore, simulation results are given to demonstrate the effectiveness and performance of the proposed neural network.

© 2012 IMACS. Published by Elsevier B.V. All rights reserved.

Keywords: Non-smooth optimization; $\epsilon$-Insensitive support vector regression; Generalized recurrent neural network; Global convergence

1. Introduction

Support vector machines (SVMs) are powerful tools for data classification and regression. In the recent years, many fast algorithms for SVMs have been developed [2]. Mangasarian [14] proposed the finite Newton algorithm for SVMs learning. Keerthi and DeCoste [5] introduced the modified finite Newton algorithm to speed up the finite Newton algorithm for fast solution of large scale linear SVMs.

More recently, as a software and hardware implementable approach, recurrent neural networks for solving linear and nonlinear optimization problems with their engineering applications have been widely developed [6,9,13,15,17]. Compared with traditional numerical optimization algorithms, the neural networks have fast convergence rate in real-time solutions. In 1986, Tank and Hopfield [15] proposed a recurrent neural network for solving the linear programming problems for the first time. In 1988, the dynamical canonical nonlinear programming circuit (NPC) was introduced by Kennedy and Chua [6] for optimization by utilizing a finite penalty parameter, which can generate the approximate optimal solutions. Wang and Xia [17] proposed a primal-dual neural network for solving the linear assignment problems. To get the optimal solutions of non-smooth optimization problems, Forti et al. [4] proposed and investigated the generalized NPC (G-NPC), which can be considered as a natural extension of NPC. In order to reduce the model complexity, some one-layer recurrent neural networks with lower model complexity have been constructed for solving linear and nonlinear programming problems [10,12]. This paper is concerned with a generalized recurrent neural network for the $\epsilon$-insensitive support vector regression. The global convergence of the proposed recurrent neural network is demonstrated through simulation results.

* Corresponding author.
E-mail address: qsliu@seu.edu.cn (Q. Liu).

0378-4754/$36.00$ © 2012 IMACS. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.matcom.2012.03.013
network is guaranteed using the Lyapunov-like method. Compared with the existing neural networks for support vector regression (SVR) learning, the proposed neural network herein has lower model complexity, but is efficient for SVR learning.

2. \(\epsilon\)-Insensitive support vector regression (\(\epsilon\)-ISVR)

Consider a given training data set \(\mathcal{T} = \{(x_i, y_i) : x_i \in \mathbb{R}^n, y_i \in \mathbb{R}, i = 1, 2, \ldots, m\}\), where \(x_i\) is the input data and \(y_i\) is called the observation. We would like to find a linear or nonlinear regression function, \(f(x)\), tolerating a small error in fitting the given data set. This can be achieved by utilizing the \(\epsilon\)-insensitive loss function that sets an \(\epsilon\)-insensitive “tube” around the data, within which errors are discarded. Disregarding the tiny errors that fall within some tolerance, say \(\epsilon\), that may lead to a better generalization ability by utilizing an \(\epsilon\)-insensitive loss function. Also, applying the idea of SVMs [2,8,16], the function \(f(x)\) is made as flat as possible in fitting the training data set by utilizing the \(\epsilon\)-insensitive loss function.

We start with the case of linear function \(f(x) = x^T w + b\), where \(w \in \mathbb{R}^n\) and \(b \in \mathbb{R}\). To obtain a linear predictor, SVR solves the following convex optimization problem:

\[
\min \frac{1}{2} w^T w + c \sum_{i=1}^{m} (\xi_i + \bar{\xi}_i),
\]

\[
\text{s.t. } w^T x_i + b - y_i \leq \epsilon + \xi_i, \quad y_i - w^T x_i - b \leq \epsilon + \bar{\xi}_i,
\]

\[
\xi_i, \bar{\xi}_i \geq 0, \quad i = 1, 2, \ldots, m,
\]

where \(c > 0\) determines the trade-off between the flatness of \(f\) and the amount up to which deviations larger than \(\epsilon\) are tolerated.

Eliminating the slack variables \(\xi_i\) and \(\bar{\xi}_i\), we get the following unconstrained non-smooth optimization problem:

\[
\min L_\epsilon(w, b) = \frac{1}{2} (w^T w + b^2) + c \sum_{i=1}^{m} |w^T x_i + b - y_i|_\epsilon,
\]

where \(|w^T x_i + b - y_i|_\epsilon = \max(0, |w^T x_i + b - y_i| - \epsilon)\) that represents the fitting errors.

Nonlinear SVR can be obtained by using a kernel function and an associated reproducing kernel Hilbert space \(H\). According to the representer theory [7], the optimal function can be expressed as a linear combination of the kernel functions centered in the training samples,

\[
f(x) = \sum_{i=1}^{m} \alpha_i K(x, x_i) + \beta.
\]

The nonlinear support vector regression formulation can be written as follows:

\[
\min L_\epsilon(\alpha, \beta) = \frac{1}{2} \left( \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j K(x_i, x_j) + \beta^2 \right) + c \sum_{i=1}^{m} \max \left(0, \left| \sum_{j=1}^{m} \alpha_j K(x_i, x_j) + \beta - y_i \right| - \epsilon \right).
\]

Assume the kernel matrix \(K = \{K(x_i, x_j)\}_{m \times m}\) and \(K_i\) being the \(i\)th row of \(K\), then (4) is rewritten as

\[
\min L_\epsilon(\alpha, \beta) = \frac{1}{2} \left( \alpha^T K \alpha + \beta^2 \right) + c \sum_{i=1}^{m} \max \left(0, |K_i \alpha + \beta - y_i| - \epsilon \right).
\]

Without loss of generality, we consider only the nonlinear SVR in this paper and the linear SVR as a special case.
دریافت فوری
متن کامل مقاله

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

ISIArticles
مرجع مقالات تخصصی ایران