Neurocomputing 174 (2016) 617–626

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Regularized online sequential extreme learning machine with adaptive regulation factor for time-varying nonlinear system

XinJiang Lu *, Chuang Zhou, MingHui Huang, WenBing Lv

State Key Laboratory of High Performance Complex Manufacturing, School of Mechanical & Electrical Engineering, Central South University, Hunan 410083, China

ARTICLE INFO

Article history:
Received 15 April 2015
Received in revised form
18 August 2015
Accepted 21 September 2015
Dr. G.-B. Huang
Available online 21 October 2015

Keywords:
Extreme learning machine
Adaptive regulation factor
Leave-one-out cross validation
modeling

ABSTRACT

In order to more accurately model time-varying nonlinear systems, we propose a regularized online sequential extreme learning machine with adaptive regulation factor (ROSELM-ARF). The construction of a new objective function allows for the online updating of both the model coefficient as well as the regulation factor, while negating the influence of the cumulate error. This differs from the traditional regularized online sequential extreme learning machine (ReOS-ELM) which only updates the model coefficient. The development and application of a two-step solving method is used to determine the optimal parameters, where the optimal regulation factor is derived using the proposed fast and online leave-one-out cross validation (FOLOO) method. The computational performance could be drastically improved by using the proposed FOLOO method as compared to using the existing leave-one-out cross validation (LOO) method. The application of the proposed method in the modeling of two practical cases is done in order to demonstrate its effectiveness. The experimental results indicate that the proposed method provides a more accurate model than several conventional modeling methods, while also improving the computational performance.

1. Introduction

Data driven modeling methods have typically been used to model complex systems, which tend to be highly nonlinear with variations in time and are comprised of many unknown dynamic factors. A well-known data-driven modeling method is the neural networks (NN) method, which has been widely applied in industry [1–8]. However, training algorithms of these NN methods are complex and usually far slower than required [9]. As a result, the framework of the NN methods was recently used to develop the extreme learning machine (ELM) in order to improve the computational performance of the single-hidden-layer feedforward neural networks. The ELM is computationally attractive due to the random generation of both its input weights and node biases and its ability to analytically determine its output weights [9–14]. The ELM also has a short training time and uniform approximation ability for any system, making it applicable for a wide range of uses [9–25].

The ELM has initially been developed for offline learning [11], however, recent efforts have focused on developing an online learning machine, including the online sequential learning machine (OSELM) proposed by Liang et al. [12], the fuzzy OSELM method proposed by Rong which combined the positive aspects of the fuzzy interface and the OSELM [13]. The OSELM also incorporated the use of a forgetting mechanism [18–20], and a timeless OSELM which was designed to have a shorter training time [21]. Despite the successful application of these offline and online ELM methods, they all continue to minimize the empirical risk while neglecting the structural risk, which can potentially lead to a poor generalization and over-fitting [22].

Recently, some regularized OSELM (ReOS-ELM) methods have been developed which were designed to consider both the empirical risk and the structural risk. Models built utilizing these methods were shown to have improved generalization capabilities [22–25]. However, while these methods consider the time-varying nature of the empirical risk, they neglect those of the structural risk resulting from a constant regulation factor in all time, greatly limiting their effectiveness when modeling unknown time-varying nonlinear systems. Time-varying systems should have different structural risks in different nonlinear regions. The structural risk is decided by the regulation factor, indicating that the regulation factor should also vary over time with the time-varying nonlinear dynamics.

Here we propose a ReOS-ELM with adaptive regulating factor (ROSELM-ARF) for the modeling of unknown time-varying nonlinear systems. A new objective function is first constructed to online update not only the model coefficient but also the

* Corresponding author.
E-mail address: luxj@csu.edu.cn (X. Lu).

http://dx.doi.org/10.1016/j.neucom.2015.09.068
0925-2312/

regulation factor, avoiding the influence of the cumulate error. This approach differs from the traditional regularized online sequential extreme learning machine (ReOS-ELM) in which only the model coefficient is updated. A two-step solving method is then developed in order to determine the optimal parameters, where a fast and online leave-one-out (FOLOO) method is proposed in order to derive the optimal regulation factor. This FOLOO method greatly improves the computational performance as compared to the existing LOO method.

The organization of this paper is as follows: an overview of the ReOS-ELM is given in Section 2; the details of the proposed ReOS-ELM with adaptive regulating factor are described in Section 3; experimental support and comparisons are presented in Section 4; the overall conclusions are discussed in Section 5.

2. Review of regularized online sequential extreme learning machine (ReOS-ELM)

Consider the following nonlinear time-varying system

$$\begin{cases} \boldsymbol{w}_{k+1} = g[\boldsymbol{w}_k, \boldsymbol{u}_k, \boldsymbol{q}_k] \\ \boldsymbol{y}_k = f^{\sim}[\boldsymbol{w}_k, \boldsymbol{q}_k] \end{cases} \tag{1}$$

where $\boldsymbol{w}_k$, $\boldsymbol{u}_k$ and $\boldsymbol{y}_k$ represent the state, input and output of the systems, g and $f^{\sim}$ are unknown nonlinear functions, and $\boldsymbol{q}_k$ denotes the time-varying parameter vector.

A general I/O representation of the system (1) can be described by the following equation:

$$\boldsymbol{y}_k = f(\boldsymbol{x}_k) \tag{2}$$

where f is an unknown function, $\boldsymbol{x}_k = [\boldsymbol{y}_{k-1}, \boldsymbol{y}_{k-2}, \cdots, \boldsymbol{y}_{k-r}, \boldsymbol{u}_{k-1}, \boldsymbol{u}_{k-2}, \ldots, \boldsymbol{u}_{k-r}]$, and r represents system order. Here, the ReOS-ELM is trained to represent the dynamic behavior of unknown system (1), as shown in Fig. 1.

For N arbitrary distinct samples $\left(\boldsymbol{x}_j, \boldsymbol{y}_j\right)$, $j=1,\ldots,N$, in the ReOS-ELM, the system model could be represented as

$$\boldsymbol{H} \cdot \boldsymbol{\beta} = Y \tag{3}$$

Where $\boldsymbol{H} = \left[\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_N\right]^{\mathrm{T}}$, $\boldsymbol{h}_j = \left[h_{j1}, h_{j2}, \ldots, h_{jN^{\sim}}\right]^T$, $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_{N^{\sim}}]^{\mathrm{T}}$, $\boldsymbol{Y} = \left[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N\right]^{\mathrm{T}}$, and $\tilde{N}$ is the neurons number. Any bounded nonlinear function could be chosen as the active function $h(\cdot)$ [30]. For example, the radial basis function (RBF) is chosen as the active function $h_{js} = \exp(-||\boldsymbol{x}_j - \boldsymbol{c}_s||/b_s)$ $(s=1,2,\ldots,\tilde{N})$, where $\mathbf{c}_s$ and b_s are the center vector and the width factor. The only unknown variable that needs to be determined in this model is coefficient $\boldsymbol{\beta}$.

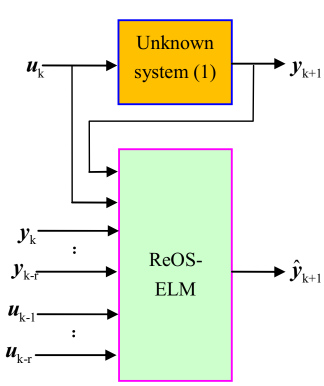

Fig. 1. ReOS-ELM model for I/O representation.

The objective function can then be represented as follows [22]:

$$\min_{\boldsymbol{\beta}} \quad L = ||\boldsymbol{\beta}||^2 + \gamma \sum_{j=1}^{N} \varepsilon_j^2 \tag{4}$$

$$s.t. \quad \boldsymbol{y}_j = h(\boldsymbol{x}_j) \cdot \boldsymbol{\beta} + \varepsilon_j, \ \ j = 1, 2, \ldots, N \tag{5}$$

where γ is a constant regulation factor that is typically given prior to optimization. Solving this optimization problem leads to the derivation of the online form of this solution, given as follows [22]:

$$\begin{cases} \boldsymbol{L}_k = \boldsymbol{L}_{k-1} + \boldsymbol{H}_k^{\mathrm{T}} \boldsymbol{H}_k \\ \boldsymbol{\beta}^k = \boldsymbol{\beta}^{k-1} + \boldsymbol{L}_k^{-1} \boldsymbol{H}_k^{\mathrm{T}} \left(\boldsymbol{Y}_k - \boldsymbol{H}_k \boldsymbol{\beta}^{k-1}\right) \end{cases} \tag{6}$$

the initial values are given as:

$$\begin{cases} \boldsymbol{L}_0 = \boldsymbol{H}_0^{\mathrm{T}} \boldsymbol{H}_0 + \frac{1}{\gamma} \boldsymbol{I} \\ \boldsymbol{\beta}^0 = \boldsymbol{L}_0^{-1} \boldsymbol{H}_0^{\mathrm{T}} \boldsymbol{Y}_0 \end{cases}$$

where $\boldsymbol{Y}_0 = \left[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{N_0}\right]^{\mathrm{T}}$, $\boldsymbol{H}_0 = \left[\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{N_0}\right]^{\mathrm{T}}$, $\boldsymbol{H}_k = \left[\boldsymbol{h}_{N_0 + \sum_{i=1}^{k-1} m_i + 1}, \boldsymbol{h}_{N_0 + \sum_{i=1}^{k-1} m_i + 2}, \ldots, \boldsymbol{h}_{N_0 + \sum_{i=1}^{k-1} m_i + m_k}\right]^{\mathrm{T}}$ is produced by new samples $\left[\boldsymbol{x}_{N_0 + \sum_{i=1}^{k-1} m_i + 1}, \boldsymbol{x}_{N_0 + \sum_{i=1}^{k-1} m_i + 2}, \ldots, \boldsymbol{x}_{N_0 + \sum_{i=1}^{k-1} m_i + m_k}\right]^{\mathrm{T}}$ as they arrive in a series of chunks, $m_i (i = 1, \ldots, k)$ is the sample number of the ith chunk, $N_0 + \sum_{i=1}^{k-1} m_i$ is the number of the old samples, $\boldsymbol{Y}_k = \left[\boldsymbol{y}_{N_0 + \sum_{i=1}^{k-1} m_i + 1}, \boldsymbol{y}_{N_0 + \sum_{i=1}^{k-1} m_i + 2}, \ldots, \boldsymbol{y}_{N_0 + \sum_{i=1}^{k-1} m_i + m_k}\right]^{\mathrm{T}}$ is the new output vector, and $\boldsymbol{\beta}^k$ is the model coefficient at the kth time.

Despite the widespread application of ReOS-ELM, it still has the following disadvantages:

- As the sample number N increases, the value of the second norm term $\gamma \sum_{j=1}^{N} \varepsilon_j^2$ in the optimization (4) becomes larger, however, the first norm term $||\boldsymbol{\beta}||^2$ is unaltered. This disrupts the balance between these two norm terms, which leads to the empirical risk and the structural risk being unbalanced. This means that these methods consider the time-varying nature of the empirical risk and pay no attention to those of the structural risk;
- The built model has a low capability for adapting to the time-varying dynamics of the system due to the regulation factor being held constant at all times.

3. Regularized online sequential extreme learning machine with adaptive regulation factor (ROSELM-ARF)

In order to model an unknown time-varying nonlinear system, a ROSELM-ARF is proposed, as shown in Fig. 2. Using this method allows for both the model coefficient $\boldsymbol{\beta}$ and the regulation factor γ to be updated in real time, according to the modeling error in order to track the time-varying dynamics of the system. The influence of the cumulate error is also taken into consideration. As the proposed method considers the time-varying nature of both the empirical risk and the structural risk, it is capable of accurately modeling the time-varying nonlinear system.