Neurocomputing 149 (2015) 215–223

Contents lists available at ScienceDirect

Neurocomputing

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

Improved incremental Regularized Extreme Learning Machine Algorithm and its application in two-motor decoupling control

Jin-Lin Ding [a,*], Feng Wang [a,b], Hong Sun [a], Li Shang [a]

[a] Suzhou Vocational University, Suzhou, Jiangsu Province 215104, China
[b] School of Electrical and Information Engineering, Jiangsu University, Jiangsu, Zhenjiang 212013, China

ARTICLE INFO

Article history:
Received 15 August 2013
Received in revised form
28 January 2014
Accepted 4 February 2014
Available online 30 September 2014

Keywords:
II-RELM
Decoupling control
Cholesky factorization without square root
Generalized inverse
Vector mode
Changing rate of learning error

ABSTRACT

Regularized Extreme Learning Machine (RELM) is an ideal algorithm for regression and classification due to its fast training speed and good generalization performance. However, how to obtain the suitable number of hidden nodes is still a challenging task. In order to solve the problem, a new incremental algorithm based on Cholesky factorization without square root is proposed in this paper, which is called the improved incremental RELM (II-RELM). The method can automatically determine optimal network structure through gradually adding new hidden nodes one by one. It achieves less computational cost and better accuracy through updating output weights. Finally, neural network generalized inverse (NNGI) based on II-RELM is applied to two-motor synchronous decoupling control. Simulation indicates that the proposed algorithm has excellent performance in prediction control. It realizes the decoupling control between velocity and tension.

1. Introduction

Inverse method is an effective way for feedback linearization of complex nonlinear system in multivariable decoupling control. However, the method is excessively dependent on accurate mathematical model and difficult to solve, which restricts its application. To overcome these defects, Dai et al. [1,2] put forward neural network inverse (NNI) control strategy, which builds a dynamic inverse model of the original system based on neural network (NN) and integrator. In [3] Dai and Liu proposed NN α-th order inverse system, and applied it in the two-motor variable frequency speed-regulating control. Neural network generalized inverse (NNGI) is presented in [4–6], which can transform the multi-input multi-output (MIMO) nonlinear system into a number of single-input single-output (SISO) linear subsystems with open-loop stability. The constructed pseudo-linear composite system (PLCS) realizes the decoupling control between speed and tension. Experimental results show that the method has good static and dynamic decoupling performance. In [7–11] some methods of building NNI are also presented. On-line adjustment control of two motor speed-regulating system based on NNGI is applied in [12]. The proposed method achieves strong stability and robustness.

Despite the many achievements that have been acquired in NNI, the slow training speed of traditional NN restricts its application in control system. Extreme learning machine (ELM), a novel learning algorithm for single hidden-layer feed forward neural networks (SLFNs), has been proposed in [13–22], and has been applied in various fields [23–26]. Within the framework of ELM, the input weights and hidden biases are randomly chosen, and the output weights are analytically determined by using Moore–Penrose generalized inverse. Compared with the traditional gradient-based learning algorithms, ELM not only learns much faster with higher generalization performance, but also avoids many difficulties that are faced by gradient-based learning methods such as slow learning rate, the existence of local minima and over tuning issues. ELM has far less computational burden than evolutionary algorithms. Its performance in terms of accuracy is also satisfactory [17].

Although the ELM model has many advantages, it may lead to over-fitting problem [14,15]. Compared with the original ELM algorithm, regularized ELM (RELM) based on the structural risk minimization principle and weighted least square has significantly improved generalization performance in most cases without increasing training burden. In [15] RELM achieves better robustness than ELM, and overcomes the over-fitting issue of ELM.

However, it is still unknown how to choose the optimal number of hidden nodes in RELM. A common approach to searching such fixed network size is by trial and error. This approach, although straightforward, is computationally expensive and does not guarantee that the selected network size will be close to optimal. In [27] researchers have proved in theory that for SLFNs with additive nodes, input weights and biases of hidden nodes may be frozen once they have been tuned. They need not be further adjusted when new nodes are added.

* Corresponding author.
E-mail address: djinlin@jssvc.edu.cn (J.-L. Ding).

http://dx.doi.org/10.1016/j.neucom.2014.02.071
0925-2312/

In order to solve the determination problem of hidden nodes when RELM is applied to chaotic time series prediction, incremental RELM algorithm based on Cholesky factorization (CF-RELM) is proposed [28], which can determine the optimal hidden nodes automatically. Simulation results show that the method has higher prediction accuracy and less computational cost. In [29], Huang et al. further proved that an incremental ELM (I-ELM) can enable SLFNs work as universal approximators. Such a new network is fully automatic in the learning process, and no manual tuning of control parameters is needed. In [30] Huang and Chen provided a modified incremental algorithm based on a convex optimization method to further improve the convergence rate of I-ELM by allowing properly adjusting the output weights of the existing hidden nodes when a new hidden node is added. To obtain compact network architecture, enhanced I-ELM (EI-ELM) is presented [31]. At each learning step, EI-ELM picks the optimal hidden node among several randomly generated hidden nodes, which leads us to the smallest residual error. Compared with the original I-ELM, EI-ELM achieves faster convergence rate and much more compact network architecture. In [32] an error-minimization-based method (EM-ELM) is proposed, which can grow hidden nodes one by one or group by group. The output weights will be incrementally updated which significantly reduces the computational complexity.

The paper puts forward a new incremental RELM algorithm, denoted by improved incremental RELM (II-RELM). The method can automatically determine optimal network structure through gradually adding new hidden nodes one by one, and has less computational cost and better accuracy through updating output weights than conventional I-RELM. II-RELM is applied to approximate the NNGI of the two-motor synchronous system (TMSS), and to connect identified inverse model with the original system. The formed PLCS realizes the decoupling control of tension and speed. In this paper, there are two major contributions. (1) II-RELM based on Cholesky factorization without square root is proposed. The algorithm has faster learning speed and less computational cost than traditional I-RELM. (2) II-RELM is applied to TMSS.

This paper is organized as follows. Section 2 analyzes the proposed II-RELM. In Section 3 mathematical model of TMSS is described. Section 4 presents simulation of NNGI based on II-RELM. Section 5 demonstrates the experimental results of TMSS based on II-RELM. Section 6 is the conclusion.

2. Proposed II-RELM algorithm

2.1. Brief of RELM

In this section, we briefly describe the essence of RELM [15,16].

Given N distinct training samples $\{(X_i, t_i) | X_i \in R^n, t_i \in R^m\}_{i=1}^N$, standard SLFNs with L hidden nodes and activation function $g(X)$ are mathematically modeled as

$$H\beta = T \tag{1}$$

where

$$H(a_1, \dots, a_L, b_1, \dots, b_L, X_1, \dots, X_N) = \begin{bmatrix} g(a_1 \cdot X_1 + b_1) & \cdots & g(a_L \cdot X_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(a_1 \cdot X_N + b_1) & \dots & g(a_L \cdot X_N + b_L) \end{bmatrix}_{N \times L} \tag{2}$$

$$\beta = \left[\beta_1^T \dots \beta_L^T\right]_{L \times m}^T, \quad T = \left[t_1^T \dots t_L^T\right]_{N \times m}^T \tag{3}$$

$a_j = [a_{j1}, \dots, a_{jn}]^T$ is the weight vector connecting the jth hidden node and the input nodes, $\beta_j = [\beta_{j1}, \dots, \beta_{jm}]^T$ is the weight vector connecting the jth hidden node and the output nodes, $g(x)$ is the activation function, and b_j is the threshold of the jth hidden node. a_j and b_j can be arbitrarily assigned.

In order to improve the generalization ability of the traditional SFLNs based on ELM, Huang et al. proposed the equality constrained optimization-based ELM in [16]. In their approach structural risk regarded as the regularization term is introduced. The so-called RELM can adjust the proportion of structural risk and empirical risk through the parameter γ. The regression problem for the proposed constrained optimization can be formulated as

$$\begin{aligned} &\text{Minimize} \quad E(W) = \tfrac{1}{2}\gamma\varepsilon^T\varepsilon + \tfrac{1}{2}\beta^T\beta \\ &\quad \text{Subject to} \quad H\beta - T = \varepsilon \end{aligned} \tag{4}$$

where γ is the proportion parameter of two kinds of risks, and ε is the regression error. To solve the above optimization problem Lagrange function is constructed as follows:

$$\xi(\beta, \varepsilon, \alpha) = \tfrac{1}{2}\gamma\varepsilon^T\varepsilon + \tfrac{1}{2}\beta^T\beta - \alpha(H\beta - T - \varepsilon) \tag{5}$$

Letting the partial derivatives of $\xi(\beta, \varepsilon, \alpha)$ with respect to β, ε and α be zero, we have

$$\begin{cases} \beta^T = \alpha H \\ \gamma\varepsilon^T + \alpha = 0 \\ H\beta - T - \varepsilon = 0 \end{cases} \tag{6}$$

From Eq. (6) we obtain

$$\beta = \left(\frac{1}{\gamma}I + H^T H\right)^{-1} H^T T \tag{7}$$

where I is the identity matrix, and the corresponding regression model of RELM is

$$Y = h(X)\beta = h(X)\left(\frac{1}{\gamma}I + H^T H\right)^{-1} H^T T \tag{8}$$

2.2. Improved calculation method of output weights

On the basis of Huang's RELM, the improved RELM algorithm is proposed by us. In Huang's RELM, the weight vector β is solved by inverse operation of high-order matrix. There exists large amount of calculation, and it reduces modeling efficiency. In order to simplify computation, a new method based on Cholesky factorization without square root [33] is introduced.

From Eq. (6) we obtain

$$\left(\frac{1}{\gamma}I + H^T H\right)\beta = H^T T \tag{9}$$

Let

$$\begin{aligned} A &= \frac{1}{\gamma}I + H^T H \\ B &= H^T T \end{aligned} \tag{10}$$

Then Eq. (9) can be rewritten as

$$A\beta = B \tag{11}$$

Since A is a symmetric positive definite matrix [28], it can be uniquely factored as

$$A = LDL^T \tag{12}$$

where

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \tag{13}$$

$$L = \begin{bmatrix} l_{11} & & \\ \vdots & \ddots & \\ l_{n1} & \dots & l_{nn} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1/l_{11} & & \\ & \ddots & \\ & & 1/l_{nn} \end{bmatrix} \tag{14}$$