Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

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

An adaptive second order fuzzy neural network for nonlinear system modeling☆

Hong-Gui Han [a,b,*], Lu-Ming Ge [a,b], Jun-Fei Qiao [a,b]

[a] College of Electronic and Control Engineering, Beijing University of Technology, Beijing 100124, China
[b] Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

ARTICLE INFO

Article history:
Received 28 February 2016
Received in revised form
20 May 2016
Accepted 16 July 2016
Communicated by Prof. Zidong Wang

Keywords:
Nonlinear system modeling
Fuzzy neural network
Adaptive second-order algorithm
Fast convergence

ABSTRACT

In this paper, an adaptive second order algorithm (ASOA) has been developed to train the fuzzy neural network (FNN) to achieve fast and robust convergence for nonlinear system modeling. Different from recent studies, this ASOA-based FNN (ASOA-FNN) has the quasi Hessian matrix and gradient vector which are accumulated as the sum of related sub matrices and vectors, respectively. Meanwhile, the learning rate of ASOA-FNN is designed to accelerate the learning speed. In addition, the convergence of the proposed ASOA-FNN has been proved both in the fixed learning rate phase and the adaptive learning rate phase. Finally, several comparisons have been realized and they have shown that the proposed ASOA-FNN has faster convergence speed and more accurate results than that of some existing methods.

1. Introduction

Most real processes are characterized by nonlinear and time-varying behavior, and thus there always exist modeling inaccuracies [1,2]. In fact, nonlinear system model imprecision may come from actual uncertainties about the plant or from the plant dynamics. However, for a nonlinear system, it is a challenge to obtain an accurate and faithful mathematical model due to its strongly nonlinear behavior, high degree of uncertainty, or time-varying characteristics [3,4]. Therefore, lots of attention has been paid to effective process modeling techniques for nonlinear system [5]. Recent research results show that the fuzzy neural networks (FNN), combining the capability of fuzzy reasoning to handle uncertain information and that of artificial neural networks to learn from input-output datasets, has been recognized as a powerful tool for nonlinear system modeling [6]. Based on the neural implementation of a fuzzy system, the fuzzy rules derivation and the parameters automatic tuning can be realized by using the learning and optimization methods [7].

In fact, the learning algorithm is a fundamental element in the use of FNN. Therefore, many algorithms have already been developed in recent years. Among these algorithms, the backpropagation (BP) algorithm, a powerful training technique, is probably the most frequently used technique [8]. In the BP algorithm, the descent approach is used to minimize the error function, and thus the local minima is always reached associated with excessively long training time for convergence [9]. To improve the convergence speed, some adaptive learning algorithms have been studied. For example, Meng et al. introduced a decreasing function to update learning rate in [10]. And a fuzzy conjugate gradient algorithm has been developed to speed up the convergence rate of FNN in [11]. However, as widely reported in the literature, these BP algorithms [9–11] still have limited accuracy and often suffer from local minima problem.

To improve the accuracy and avoid tracking into local minima, the evolutionary algorithms, which can find the global solution to optimize the overall parameters of FNN, have been widely used [12–14]. For example, Tzeng proposed an efficient genetic algorithm (GA) to adjust the parameters of FNNs. In this algorithm, the parameters such as the weights and membership functions can be characterized by minimizing a quadratic measure of the error function. The simulation results illustrate the advantages of high approximation accuracy and good generalization performance [12]. Then, a two-phase learning method, based on the typical GA, was

☆This work was supported by the National Science Foundation of China under Grants 61533002 and 61225016, Beijing Nova Program under Grant Z131104000413007, China Postdoctoral Science Foundation under Grant 2014M550017, pH.D. Program Foundation from Ministry of Chinese Education under Grants 20121103120020 and 20131103110016, Collaborative Innovation Program under Grant ZH14000177, Project supported by Beijing Postdoctoral Research Foundation under Grant 2015ZZ-03, and Beijing Municipal Education Commission Foundation under Grants km201410005001 and KZ201410005002.

* Corresponding author at: College of Electronic and Control Engineering, Beijing University of Technology, Beijing 100124, China.
E-mail addresses: Rechardhan@sina.com (H.-G. Han), geluming1992@126.com (L.-M. Ge), isibox@sina.com (J.-F. Qiao).

http://dx.doi.org/10.1016/j.neucom.2016.07.003
0925-2312/

developed to reduce the generated error of FNN in [13]. The first phase roughly estimated the optimal fuzzy weights, and the second phase provided better estimate for the shape of the membership function. Moreover, Kuo et al. designed a particle swarm optimization (PSO), according to fuzzy neural network, to determine the relationship between the radio frequency identification signals and the position of a picking cart [14]. Furthermore, some other evolutionary algorithms (such as artificial bee colony (ABC) algorithm, differential evolution (DE), Hierarchical fair competition parallel genetic algorithm, and so on) were employed to optimize the parameters of FNN in [15–18]. However, even though these evolutionary algorithms can find the global solution to train the parameters of FNN [12–18], the searching processes are extremely time-consuming [19,20]. In order to improve the learning speed and the solution accuracy, a species-based improved electromagnetism-like mechanism (SIEM) is proposed to optimize the parameters of FNN. This SIEM algorithm combines the advantages of electromagnetism-like mechanism and gradient-descent technique to obtain fast convergence and low computational complexity [21]. Zhao et al. proposed a new gradient learning approach for FNN, which considers both nonlinear parameters in the rule premises and linear parameters in the rule consequents. As a result, the learning speed and the accuracy of the proposed method can be improved simultaneously [22]. Moreover, a fast gradient training algorithm was used to train the parameters of FNN in [23], and a gradient learning algorithm with a dynamic learning rate was derived to adjust the parameters of the fully connected FNN in [24]. In addition, some other learning algorithms have been proposed in [25–27]. These gradient learning algorithms [21–27] lead to significantly improved modelling performance of FNN. However, most of the former algorithms have not examined the algorithm convergence that ensures the system modeling performance.

To avoid the aforementioned problems, in this paper, a hybrid learning algorithm, which is a combination of both adaptive second-order algorithm (ASOA) and adaptive learning rate strategy, is applied to train the parameters of FNN. This ASOA-based FNN (ASOA-FNN) has fast convergence in learning process and high efficiency to deal with nonlinear system modeling. The proposed ASOA-FNN owns four major advantages over other networks as follows: Firstly, a new ASOA is developed for training the parameters of FNN. One of the key factors of the ASOA strategy is that the quasi Hessian matrix and gradient vector are accumulated as the sum of related sub matrices and vectors, respectively. Due to the advantages of fast convergence and powerful searching ability of the proposed ASOA, this ASOA-FNN can reduce the computational complexity of the learning process and reach smaller testing error with much faster speed. Secondly, with the purpose of achieving effluent modeling performance, an adaptive learning rate strategy has been carried out for ASOA-FNN. In fact, when the iterative sequence is near the solution set, the learning rate may be smaller than the machine precision, and so it will lose its role [28,29]. Thus, this proposed adaptive learning rate strategy can accelerate the convergence of ASOA-FNN in the learning process. Thirdly, as the convergence is important to the applications, this proposed ASOA-FNN has been specifically designed with these in mind. The convergence analysis and convergence conditions of ASOA-FNN are demonstrated theoretically in details.

The remainder of this paper is organized as follows. Section 2 defines the problem and necessary conditions. The details of ASOA-FNN is introduced in Section 3. Section 4 discusses the convergence of ASOA-FNN and gives the convergence conditions. Section 5 presents the results by utilizing three types of dynamic system modeling, respectively. The models include the Mackey-Glass time-series prediction, the classical MIMO nonlinear system modeling, and the sludge bulking prediction in wastewater treatment process. Finally, Section 6 concludes the paper.

2. Problem definition

For a multi-input and multi-output system, the nonlinear form of the system is defined by

$$\mathbf{y} = f(\mathbf{x}), \tag{1}$$

where $\mathbf{y}=[y_1,y_2,\ldots,y_Q]^T$ and $\mathbf{x}=[x_1,x_2,\ldots,x_k]$ are the output and input of the nonlinear system, respectively.

To model the MIMO nonlinear system, a MIMO FNN is introduced. An illustrative FNN structure is shown in Fig. 1(a MIMO for example). The mathematical description of this MIMO FNN is given below:

$$\hat{\mathbf{g}} = \mathbf{W}\mathbf{v}, \tag{2}$$

where

$$\hat{\mathbf{g}} = [\hat{g}_1, \hat{g}_2, \cdots, \hat{g}_Q]^{\mathbf{T}}, \ \ \mathbf{W} = \left[\mathbf{w}^1, \mathbf{w}^2, \cdots, \mathbf{w}^Q\right]^{\mathbf{T}}, \tag{3}$$

$\hat{g}_q$ is the output of the qth neuron in the output layer, $\mathbf{W}$ is the parameter matrix, $\mathbf{w}^q = [w_1^q, w_2^q, \ldots, w_P^q]$ are the weights between the qth neuron in the output layer and the normalized layer, Q is the number of neurons in the output layer, $\mathbf{v}$ is the output of the normalized layer, and for a fuzzy model

$$\hat{g}_q = \mathbf{w}^q\mathbf{v} = \frac{\sum_{l=1}^{P} w_l^q e^{-\sum_{i=1}^{k}\frac{(x_i-c_{il})^2}{2\sigma_{il}^2}}}{\sum_{j=1}^{P} e^{-\sum_{i=1}^{k}\frac{(x_i-c_{ij})^2}{2\sigma_{ij}^2}}}, \tag{4}$$

where P is the number of neurons in the normalized layer, v_l is the output of the lth normalized neuron, and $\mathbf{v}=[v_1, v_2,\ldots,v_P]^{\mathbf{T}}$, and

$$v_l = \frac{\phi_l}{\sum_{j=1}^{P}\phi_j} = \frac{e^{-\sum_{i=1}^{k}\frac{(x_i-c_{il})^2}{2\sigma_{il}^2}}}{\sum_{j=1}^{P} e^{-\sum_{i=1}^{k}\frac{(x_i-c_{ij})^2}{2\sigma_{ij}^2}}}, \ \ j = 1, 2, \cdots, P; \ \ l = 1, 2, \cdots, P \tag{5}$$

the number of neurons in the radial basis function (RBF) layer is equal to the number of neurons in the normalized layer, and φ_j is the output value of the jth RBF neuron

$$\phi_j = \prod_{i=1}^{k} A_j^i(x_j) = \prod_{i=1}^{k} e^{-\frac{(x_i-c_{ij})^2}{2\sigma_{ij}^2}} = e^{-\sum_{i=1}^{k}\frac{(x_i-c_{ij})^2}{2\sigma_{ij}^2}},$$
$$i = 1, 2, \cdots, k; \ \ j = 1, 2, \cdots, P, \tag{6}$$

$\mathbf{c}_j=[c_{1j},c_{2j},\ldots,c_{kj}]$ and $\boldsymbol{\sigma}_j=[\sigma_{1j}, \sigma_{2j},\ldots,\sigma_{kj}]$ are the vectors of centers and widths of the jth RBF neuron, respectively, and

$$u_i = x_i, \ \ (i = 1, 2, \cdots, k), \tag{7}$$

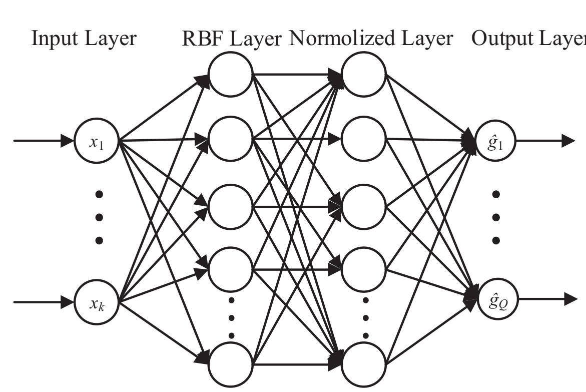

Fig. 1. The structure of FNN.

Please cite this article as: H.-G. Han, et al., An adaptive second order fuzzy neural network for nonlinear system modeling, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.07.003