Expert Systems with Applications 38 (2011) 15224–15233

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Integration of support vector regression and annealing dynamical learning algorithm for MIMO system identification

Chia-Nan Ko

Department of Automation Engineering, Nan Kai University of Technology, Tsaotun, Nantou 542, Taiwan

ARTICLE INFO

Keywords:
System identification
Support vector regression
Radial basis function network
Annealing dynamical learning algorithm
Particle swarm optimization

ABSTRACT

This paper presents a robust approach to identify multi-input multi-output (MIMO) systems. Integrating support vector regression (SVR) and annealing dynamical learning algorithm (ADLA), the proposed method is adopted to optimize a radial basis function network (RBFN) for identification of MIMO systems. In the system identification, first, SVR is adopted to determine the number of hidden layer nodes, the initial structure of the RBFN. After initialization, ADLA with nonlinear time-varying learning rate is then applied to train the RBFN. In the ADLA, the determination of the learning rate would be an important work for the trade-off between stability and speed of convergence. A computationally efficient optimization method, particle swarm optimization (PSO) method, is adopted to simultaneously find optimal learning rates. Due to the advantages of SVR and ADLA (SVR-ADLA), the proposed RBFN (SVR-ADLA-RBFN) has good performance for MIMO system identification. Two examples are illustrated to show the feasibility and superiority of the proposed SVR-ADLA-RBFNs for identification of MIMO systems. Simulation results are provided to demonstrate the effectiveness of the proposed algorithm.

1. Introduction

System identification is a vital work in industry applications such as control design, plant diagnosis, and system monitoring. Recently, identification of nonlinear MIMO systems have been used widely in various fields (Chan, Bao, & Whiten, 2006; Felici, van Wingerden, & Verhaegen, 2007; Goethals, Pelckmans, Suykens, & De Moor, 2005a, Goethals, Pelckmans, Suykens, & De Moor, 2005b; Huang, Benesty, & Chen, 2006; Majhi & Panda, 2010; Rouss & Charon, 2008; Vieira, Santos, Carvalho, Pereira, & Fileti, 2005). However, it should be pointed out that structural identification and parameter estimation of nonlinear MIMO systems are rather difficult issues in system identification. Therefore, experts have put much effort in this research field. Cardinal spline functions to model MIMO Hammerstein systems have been adopted (Goethals et al., 2005a, 2005b). Wang, Ding, and Liu (2007) have introduced a hierarchical least squares algorithm for identifying MIMO ARX-like systems based on the hierarchical identification principle. An artificial neural network model for system identification by expanding the input pattern by Chebyshev polynomials has been proposed (Purwar, Kar, & Jha, 2007). A systematic way that SVR integrating least squares regression has been proposed to identify MIMO systems (Fu, Wu, Jeng, & Ko, 2009). A neural inverse dynamic NARX model has been adopted to perform MIMO system identification (Anh & Phuc, 2010).

In the neural network, RBFNs have received considerable applications, such as function approximation, prediction, recognition, etc. (Chuang, Jeng, & Lin, 2004; Sing, Basu, Nasipuri, & Kundu, 2007; Xu, Xie, Tang, & Ho, 2003; Yu, Gomm, & Williams, 2000). Since RBFNs have only one hidden layer and have fast convergence speed, they are widely used for nonlinear system identification recently (Apostolikas & Tzafestas, 2003; Chen, Hong, Luk, & Harris, 2009; Falcao, Langlois, & Wichert, 2006; Fu et al., 2009; Li & Zhao, 2006). Besides, the RBFNs are often referred to as model-free estimators since they can be used to approximate the desired outputs without requiring a mathematical description of how the outputs functionally depend on the inputs (Kosko, 1992).

When utilizing RBFNs, the number of hidden layer nodes, the initial parameters of the kernel, and the initial weights of the networks must be determined first. However, a systematic way to determine the initial structure of RBFNs has not been established yet. In many cases, these parameters are determined according to the experience of the designer or just chosen randomly. For example, in Falcao et al. (2006), Manrique, Rios, and Rodriguez-Paton (2006), and Sarimveis, Alexandridis, Mazarakis, and Bafas (2004), the number of hidden layer nodes is fixed according to the choice of the designer first. Then different kinds of algorithms such as the least squares method, the gradient descent method, and the genetic algorithm are used to optimize the parameters. However, such kind of improper initialization usually results in slow convergence speed and poor performance of the RBFNs. Meanwhile, a learning rate serves as an important role in the procedure of training RBFNs. Generally, the learning rate is selected as a time-invariant constant

E-mail address: t105@nkut.edu.tw

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.05.078

by trial and error (Chuang et al., 2004; Chuang, Su, & Hsiao, 2002; Fu et al., 2009; Hsieh, Sun, Lin, & Liu, 2008). However, there still exist several problems of unstable or slow convergence. Some researchers have engaged in exploring the learning rate to improve the stability and the speed of convergence (Song, Zhang, & Sun, 2008; Yoo, Park, & Choi, 2007; Yu, 2004).

Recently, support vector machine (SVM) has been successfully used in various fields due to the potential capability of handling classification tasks in the case of high dimensionality and sparsity of sampling data (Camps-Valls, Munoz-Mari, Martinez-Ramon, Requena-Carrion, & Rojo-Alvarez, 2009; Goethals et al., 2005a, 2005b; Manel et al., 2006; Suykens, 2001; Vapnik, 1998). In some research (Espinoza, Suykens, & De Moor, 2005; Fu et al., 2009; Gao, Dai, Zhu, & Tang, 2007; Lima, Coelho, & Von Zuben, 2007), SVR algorithm has been adopted for nonlinear system identification. In this paper, in order to overcome the above problems of training RBFNs, first, an SVR method with Gaussian kernel function (Gao et al., 2007; Hua & Zhang, 2006; Vapnik, 1995) is adopted to determine the initial structure of the RBFNs for identifying nonlinear systems. This means that the proposed method is to use the SVR method to determine the number of hidden layer nodes and the initial parameters of the kernel. After initialization, an annealing robust concept (Chuang et al., 2002; Fu et al., 2009) with dynamical learning algorithm (ADLA) is then applied to train the RBFN (ADLA-RBFN), in which PSO method is adopted to find optimal learning rates during learning procedure. Two simulation examples will be given to illustrate the feasibility and efficiency of the proposed SVR-based ADLA-RBFNs (SVR-ADLA-RBFNs) for identification of MIMO systems.

This paper is organized as follows. Section 2 describes the RBFNs for identification of nonlinear MIMO systems. In Section 3, an ADLA based on SVR is introduced to train RBFNs, in which a nonlinear time-varying evolution concept is induced. In Section 4, a population-based stochastic searching method and a fitness function evaluating populations of PSO are presented. Section 5 provides the proposed algorithm and flowchart for SVR-ADLA-RBFNs using the PSO approach. Simulation results of system identification for two MIMO examples are illustrated to evaluate the SVR-ADLA-RBFNs in Section 6. Section 7 brings conclusions for the main contributions of this paper.

2. RBFNs for identification of nonlinear MIMO systems

In general, the input–output relation of a nonlinear MIMO system can be expressed as

$$\mathbf{y}(t+1) = \mathbf{f}(\mathbf{y}(t), \mathbf{y}(t-1), \ldots, \mathbf{y}(t-n_y), \mathbf{x}(t), \mathbf{x}(t-1), \ldots, \mathbf{x}(t-n_u)), \tag{1}$$

where $\mathbf{x}(t) = [x_1(t) \ \cdots \ x_m(t)]^{\mathrm{T}}$ is the input vector, $\mathbf{y}(t) = [y_1(t) \ \cdots \ y_p(t)]^{\mathrm{T}}$ is the output vector, n_u and n_y are the maximal lags in the input and output, respectively, and $\mathbf{f}(t) = [f_1(t) \ \cdots \ f_p(t)]^{\mathrm{T}}$ denotes the nonlinear relation to be estimated.

One can use a neural network to estimate the input–output relation of a nonlinear MIMO system. In this paper, an RBFN will be adopted since it has a simple structure as shown in Fig. 1. When the Gaussian function is chosen as the radial basis function, an RBFN can be expressed in the form

$$\hat{y}_j(t+1) = \sum_{i=1}^{L} G_i w_{ij} = \sum_{i=1}^{L} w_{ij} \exp\left(-\frac{\|\mathbf{x}-\mathbf{m}_i\|^2}{2\sigma_i^2}\right) \quad \text{for} \quad j = 1, 2, \ldots, p, \tag{2}$$

where $\mathbf{x}(t) = [x_1(t) \ \cdots \ x_m(t)]^{\mathrm{T}}$ is the input vector, $\hat{\mathbf{y}}(t) = [\hat{y}_1(t) \ \cdots \ \hat{y}_p(t)]^{\mathrm{T}}$ is the output vector of the RBFN, w_{ij} is the synaptic weight, G_i is the Gaussian function, $\mathbf{m}_i$ and σ_i are the center and width of G_i, respectively, and L is the number of the Gaussian functions, which is also equal to the number of hidden layer nodes.

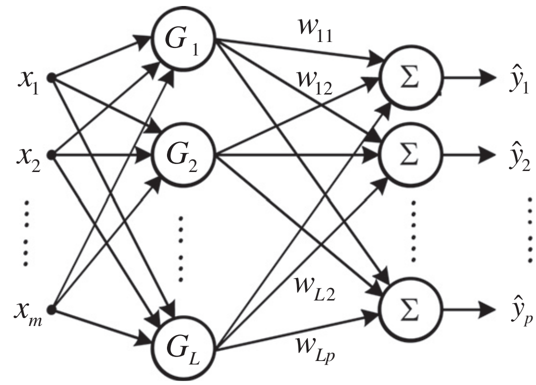

Fig. 1. The structure of a radial basis function network.

Given a set of training input–output pairs $(\mathbf{x}^{(k)}, \mathbf{y}^{(k)})$, $k = 1, 2, \ldots, N$, where $\mathbf{x}^{(k)}(t) = \left[x_1^{(k)}(t) \ \cdots \ x_m^{(k)}(t)\right]^{\mathrm{T}}$ and $\mathbf{y}^{(k)}(t) = \left[y_1^{(k)}(t) \ \cdots \ y_p^{(k)}(t)\right]^{\mathrm{T}}$, the identification problem of the nonlinear MIMO system is to determine the values of L, w_{ij}, $\mathbf{m}_i$, and σ_i to minimize the following performance index

$$J = \sum_{k=1}^{N} \|\mathbf{y}^{(k)} - \hat{\mathbf{y}}^{(k)}\|^2, \tag{3}$$

where $\hat{\mathbf{y}}^{(k)}$ is the corresponding output of the RBFN when the input of the network is $\mathbf{x}^{(k)}$.

It is very difficult, if not impossible, to solve the above problem directly. In usual cases, the initial values of L, w_{ij}, $\mathbf{m}_i$, and σ_i are chosen first. Then a training algorithm is applied to the RBFN to search for the optimal combination of these values in an iterative manner. However, as mentioned above, there is no way to choose the initial values of L, w_{ij}, $\mathbf{m}_i$, and σ_i systematically. Therefore, in the following section, an SVR approach will be proposed to serve for this purpose.

3. SVR-ADLA-RBFNs

3.1. SVR-based initial structure of RBFNs

When approximating an unknown function, SVR method (Vapnik, 1995) can be adopted to build the initial structure of RBFNs. Meanwhile, assume that a set of basis functions, $g_l(\mathbf{x})$, $l = 1, 2, \ldots, M$, is given. Then the problem of function approximation is transformed into finding the parameters of the following basis linear expansion

$$f(\mathbf{x}, \boldsymbol{\theta}) = \sum_{l=1}^{M} \theta_l g_l(\mathbf{x}) + b, \tag{4}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_M)$ is a parameter vector to be identified and b is a constant to be determined.

From SVR method, one can find that the solution is to find $f(\mathbf{x}, \boldsymbol{\theta})$ in Eq. (4) that minimizes

$$R(\boldsymbol{\theta}) = \frac{1}{N} \sum_{k=1}^{N} L_\varepsilon\left(y_1^{(k)} - f(\mathbf{x}^{(k)}, \boldsymbol{\theta})\right), \tag{5}$$

subject to the constraint

$$\|\boldsymbol{\theta}\|^2 < C, \tag{6}$$

where $L_\varepsilon(\cdot)$ is the ε-insensitive loss function defined as