Neural Networks 63 (2015) 79–86

Contents lists available at ScienceDirect

Neural Networks

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

Fast Clustered Radial Basis Function Network as an adaptive predictive controller

Dino Kosic

Department of Automatic Control, Faculty of Electrical Engineering, University of Banja Luka Patre 5, 78000 Banja Luka, Bosnia and Herzegovina

ARTICLE INFO

Article history:
Received 29 May 2014
Received in revised form 18 November 2014
Accepted 21 November 2014
Available online 28 November 2014

Keywords:
Artificial neural network
Radial basis function network
Least squares
Adaptive predictive control

ABSTRACT

This paper presents a novel artificial neural network with the Radial Basis Function (RBF) as an activation function of neurons and clustered neurons in the hidden layer which has a high learning speed, thus it is called Fast Clustered Radial Basis Function Network (FCRBFN). The weights of the network are determined by solving a number of linear equation systems. In addition, new training data can be given to the network on-line and the re-training is done at high speed using the Least Squares method. In order to test the validity of the FCRBFN, it is applied to 4 classical regression applications, and also used to build the functional adaptive predictive controller. Experimental results show that, compared with other methods, the FCRBFN with a small amount of hidden neurons could achieve good or better regression precision and generalization, as well as adaptive ability at a much faster learning speed.

1. Introduction

Artificial Neural Networks (ANN) have numerous applications, including identification and modeling of nonlinear systems and various types of controllers, including adaptive predictive controllers, in control systems (Akpan & Hassapis, 2011; Broomhead & Lowe, 1988; Chen & Chen, 1995; Ge, Yang, & Lee, 2008; Hagan, Demuth, & De Jesus, 2002; Kim et al., 2004; Na, Ren, Shang, & Guo, 2012). There are several types of ANNs and learning algorithms, from the standard backpropagation, which is a specific type of a gradient descent method, to the fast learning algorithms developed specifically to overcome problems with high iteration count and convergence issues (Azizi-Sadjadi & Ren-Jean, 1992; Castillo, Guijarro-Berdinas, Fontenla-Romero, & Alonso-Betanzos, 2006; Chow, Yam, & Cho, 1999; Jiang, Gielen, Zhang, & Luo, 2003; Li, Niu, Wang, & Liu, 2014; Mashor, 2000).

With further development of neural network application in mind, it is of significant interest to experiment and research on various network topologies and learning algorithms. This paper proposes a novel Fast Clustered Radial Basis Function Network (FCRBFN). Since it is based on a standard Radial Basis Function Network (RBFN) its weights are computed by solving a number of linear equations systems. This is performed by using the Least Squares method and the Recursive Least Squares method (Astrom & Wittenmark, 1995; Plackett, 1950), meaning that the FCRBFN has a very high learning speed. It is also easy to verify the topology by determining systems' compatibility through the simple calculation of a matrix rank. This also reduces the time spent on finding the right network topology. In this paper it is shown that the FCRBFN can be effectively used for regression and generalization, that it is easily adapted and could be utilized as an adaptive predictive controller for nonlinear systems. For regression and generalization performances it is compared to 3 other methods: Standard feedforward neural network with a single hidden layer trained with backpropagation algorithm, Extreme Learning Machine (Huang, Zhu, & Siew, 2006) and Support Vector Machine (Chang & Lin, 2011; Cortes & Vapnik, 1995).

Since it shows good accuracy and speed in both regression and generalization, as well as the ability to adapt, the FCRBFN is also used as a predictive controller for standard control problem—the Three Tank Water system. In order to assess its quality in that field, it is compared to the standard Proportional–Integral–Derivative (PID) controller, as well as another adaptive method—Data-driven Model-free adaptive method (Hou & Jin, 2011). The idea was to emphasize controller's adaptive ability by not having it pre-trained, which is a valuable characteristic of both adaptive methods used in testing, unlike more standard approach with recurrent adaptive neurocontrollers (Prokhorov, 2006, 2007) that are based on the extensive pre-training of the network connection weights, while their adaption involves the change of other network parameters.

The rest of the paper consists of Section 2, which contains detailed description of the proposed Fast Clustered Radial Basis Function Network (FCRBFN) with 4 subsections dedicated to

E-mail address: dino.kosic@etfbl.net.

http://dx.doi.org/10.1016/j.neunet.2014.11.008
0893-6080/

the learning process of the network (the initial training and the after-training knowledge addition), the parameter choice recommendation and the comparison with the standard RBF network; Section 3, where the experimental study and discussion are given with side-by-side testing with the other methods and the testing of the FCRBFN's sensitivity to the choice of the parameters; Section 4, that shows the application of the FCRBFN as an adaptive predictive controller; and finally Section 5 with the conclusion of this paper.

2. Fast Clustered Radial Basis Function Network

In this section, a novel artificial neural network is described. The main characteristic of the new network is that it has connections between each neuron in the input layer with only certain amount of neurons in the hidden layer and all neurons in the hidden layer are connected to exactly one neuron from the input layer, unlike in fully connected networks where all neurons in the hidden layer are connected to all neurons in the input layer. All neurons from the hidden layer that are connected to the same neuron in the input layer form so-called Neuron Cluster. Neurons in the input and output layer all have the identity function $a(x) = x$ as the activation function. Activation function of neurons in the hidden layer is of the Radial Basis Function (RBF) type, and it is proposed to use Gaussian function $a(x) = \mathrm{e}^{-\left(\frac{x-c}{\sigma}\right)^2}$. Number of neurons in each neuron cluster may vary, as well as the centers c and the spreads σ of the activation functions of each neuron. Weights of the connections between the input and the hidden layer are all equal to 1. Weights of the connections between the hidden and the output layer are determined using the Least Squares method once to solve a linear equation system for each output, thus making the speed of the learning process very high. Given all of the above, the novel ANN is called Fast Clustered Radial Basis Function Network (FCRBFN). The FCRBFN's structure is given in Fig. 1. The learning process of the FCRBFN is described in detail as follows.

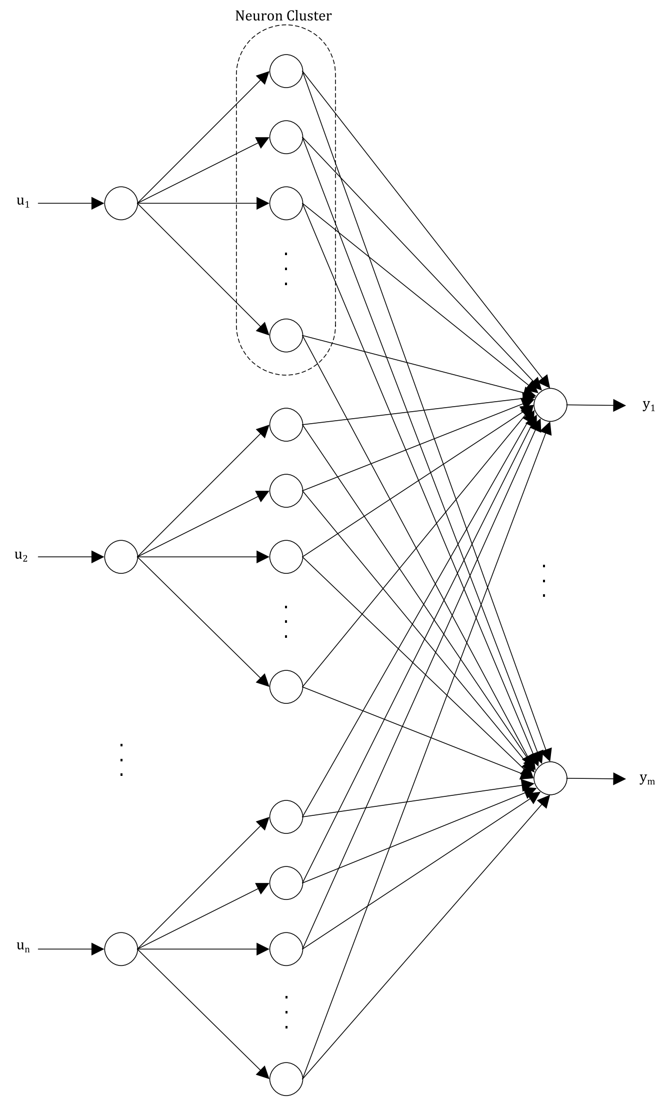

Fig. 1. Structure of the Fast Clustered Radial Basis Function Network.

2.1. Determining the weights

Let the network have n inputs and m outputs. This means there are n neuron clusters, as described earlier, one for each of the inputs. Furthermore, let there be n_i neurons in the ith neuron cluster ($i = 1, 2, \ldots, n$). This means that the total amount of neurons in the hidden layer is given by $n_{tn} = \sum_{i=1}^{n} n_i$.

Let the activation function of the jth neuron in the ith neuron cluster be defined by its center $c_{i,j}$ and its spread $\sigma_{i,j}$ and let the weight of the connection between the jth neuron in the ith neuron cluster and the k-th neuron in the output layer be $w_{i,j,k}$.

Suppose that the input of the network is given by n-dimensional vector $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$ and that the output of the network is given by m-dimensional vector $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix}$, where y_k ($k = 1, 2, \ldots, m$) is the output of the k-th neuron in the output layer and is given by

$$y_k = \sum_{i=1}^{n} \sum_{j=1}^{n_i} w_{i,j,k} \cdot \mathrm{e}^{-\left(\frac{x_i - c_{i,j}}{\sigma_{i,j}}\right)^2}. \tag{1}$$

Next, suppose that there are n_p known pairs of input and desired output vectors $\left({}^{(l)}\mathbf{x} \quad {}^{(l)}\mathbf{y}\right)$ ($l = 1, 2, \ldots, n_p$) for the learning process. If that is applied to Eq. (1), a linear system of n_p equations for each of the n_m outputs is formed. The variables in these systems are weights of the connections between one of the neurons in the output layer and all of the neurons in the hidden layer. This system of systems could be written as

$$\mathbf{\Phi} \cdot \mathbf{W} = \mathbf{Y}. \tag{2}$$

$\mathbf{\Phi}$ is $n_p \times n_{tn}$ matrix $\mathbf{\Phi} = \left[\varphi_{a,b}\right]_{n_p \times n_{tn}}$ where

$$\varphi_{a,b} = \mathrm{e}^{-\left(\frac{{}^{(a)}x_i - c_{i,j}}{\sigma_{i,j}}\right)^2}, \tag{3}$$

with ($i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n_i;\ b = j + \sum_{t=1}^{i-1} n_t$).

$\mathbf{W}$ is $n_{tn} \times m$ matrix $\mathbf{W} = \left[w_{a,b}\right]_{n_{tn} \times m}$ where

$$w_{a,b} = w_{i,j,b}, \tag{4}$$

with ($i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n_i;\ a = j + \sum_{t=1}^{i-1} n_t$).

$\mathbf{Y}$ is $n_p \times m$ matrix $\mathbf{Y} = \left[y_{a,b}\right]_{n_p \times m}$ where

$$y_{a,b} = {}^{(a)}y_b. \tag{5}$$

This system of systems, if compatible, has a solution given by

$$\mathbf{W} = \mathbf{\Phi}_{\mathbf{inv}} \cdot \mathbf{Y}, \tag{6}$$

where $\mathbf{\Phi}_{\mathbf{inv}}$ is pseudo-inverse of $\mathbf{\Phi}$ calculated by

$$\mathbf{\Phi}_{\mathbf{inv}} = (\mathbf{\Phi}^{\mathsf{T}}\mathbf{\Phi})^{-1}\mathbf{\Phi}^{\mathsf{T}}. \tag{7}$$

If $n_p = n_{tn}$, then the given system is quadratic and Eq. (7) could be shortened to standard inverse

$$\mathbf{\Phi}_{\mathbf{inv}} = \mathbf{\Phi}^{-1}. \tag{8}$$