A Practical Radial Basis Function Network and Its Applications

S. Q. Yang[1, 2] C. Y. Jia[1]

(1 Dalian Maritime University ; 2 Dalian Naval Academy)

Abstract : Artificial neural networks have become important tools in many fields. This paper describes a new algorithm for training an RBF network. This algorithm has two main advantages: higher accuracy and a too stable learning process. In addition, it can be used as a good classifier in pattern recognition.

Key Words RBF Network Chaos Classification

1. INTRODUCTION

Recently, artificial neural networks (ANNs) have become important tools in many fields, such as statistical data analysis, pattern recognition, signal processing, automatic control, forecasting, artificial intelligence and so on. The basic problem in the above fields is function approximation. At present, there are two main methods for function estimation, one is the traditional methods named " parametric approaches " or " model-based approaches " and the other is ANNs methods named "nonparametric approaches" for its many general properties. With regard to the former, the latter allows one strongly simplify the definition of the model based on a priori assumptions about the signal source and is very suitable for the nonlinear signals [1].

In function approximation, the most popular ANNs are BP networks, RBF networks and wavelet networks. However, the existing networks suffer from many drawbacks. The training processes of BP algorithm often settle in undesirable local minima of the error surface and converge too slowly. The wavelet networks also have one primary problem: When the training data is presented in the form of clustering, this is especially true for the data coming from functions of high dimensions, the wavelet networks will include a number of unnecessary hidden nodes. To solve this problem, the algorithms of clustering and sub-networks must be used [2~5].

However, RBF networks also have their defects. In our study, we find that the results of RBF networks greatly rely on the parameters of the basis functions. When using the gradient descent algorithm, the squared error of the network often disperses if the parameters are not properly selected. The oscillation of the error in a large scale is the main problem of RBF networks.

In this paper, we propose a new algorithm for training an RBF network. In function estimation, this algorithm overcomes the above disadvantage. It has higher accuracy and a too stable learning process. Furthermore, we study its ability of pattern classification, the result shows that the RBF network using our algorithm can be used as a classifier.

2. THE IMPROVED ALGORITHM

The structure of an RBF network is shown in figure 1.

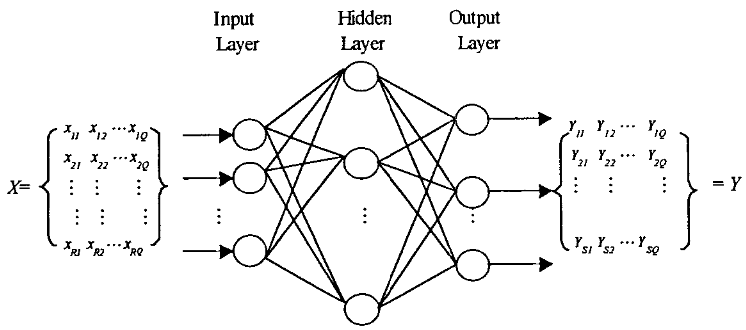

Figure 1 The structure of an RBF network

In Figure 1, X is an R-by-Q matrix a column of which is an input vector. Correspondingly, Y is an S-by-Q matrix a column of which is a target vector. Assume that $X=\{x_1, x_2,...,x_Q\}$, $Y=\{y_1, y_2,...,y_Q\}$, $G(\cdot)$ is the Gaussian-type RBF of the hidden nodes. Then, we have the nonlinear function $f_i(\cdot)$ realized by an RBF network as follows:

$$\left.\begin{array}{l} y_{ik} = f_i(x_k) = \sum_{j=1}^{M} w_{ij} G(x_k, x_{cj}, \alpha \bullet \sigma_j) + b_i \\ \\ k = 1,2,\dots,Q \qquad i = 1,2,\dots,S \end{array}\right\} \quad \cdots\cdots\cdots\cdots (1)$$

Where y_{ik} is the i-th element of the target vector y_k , M is the number of the hidden layer neurons, w_{ij} is the weight which connects the i-th node in output layer to the j-th node in hidden layer, b_i is the bias of the i-th node in output layer. x_{cj} and $\alpha \cdot \sigma_j$ are the center and spread of the j-th RBF, respectively. And α is a factor of the spread.

Now, the problem can be described as: Given X, Y and M, find $f_i(w_{ij}, x_{cj}, \alpha, \sigma_j)$ in *(1)*. To solve this problem we select E in (2) as the error function:

$$E(x_{cj}, \alpha \bullet \sigma_j, w_{ij}, b_j) = \frac{1}{Q} \sum_{k=1}^{Q} \sum_{i=1}^{S} (y_{ik} - f_i(x_k))^2 \quad \cdots\cdots\cdots\cdots\cdots (2)$$

Thus, the new algorithm may be summarized as follows:

1) Use one of the popular clustering algorithms (such as j-means clustering algorithm) to determine the center x_{cj} , σ_j in (1) in terms of X, Y and M. Let $\alpha=1$.
2) Use the least-square method to gain w_{ij}, b_j which minimize the error E in (2).
3) Use the gradient-descent approach to adjust α according to (3).

$$\alpha = \alpha - \eta \bullet \frac{\partial E}{\partial \alpha} \quad \cdots\cdots\cdots\cdots (3)$$

Where η is the learning rate. Then turn to step 2) to loop for learning until the error goal is gained.

Finally, if the error goal can not be gained, you can choose a bigger M, and then train the network again.

3. EXPERIMENTAL RESULTS OF COMPUTER SIMIULATION

We test our algorithm on two aspects. One is the function approximation. The other is classification.

A. THE RESULT ON FUNCTION APPOXIMATION

To test the approximation ability of our algorithm, we apply it to fitting the well-known chaotic Lorenz system (4).

$$\begin{cases} \dot{x} = -(8/3)x + yz \\ \dot{y} = -10y + 10z \\ \dot{z} = -yx + 28y - z \end{cases} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (4)$$

The differential equations are solved numerically using a forth-order Runge-Kutta integration with a fixed step size $\Delta\, t$=0.01s. The initial point is chosen near the attractor and the transient points are discarded. We use the former 500 points of x-coordinate to learn and the latter 1000 points to test. The form of the function that will be fitted is $x_k = f(x_{k-1}, x_{k-2}, x_{k-3})$. The results are presented in Figure 2, " —" is

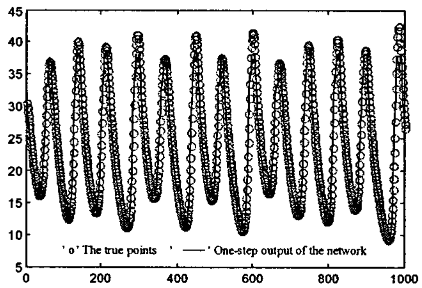

Figure 2 The result of fitting a chaotic system

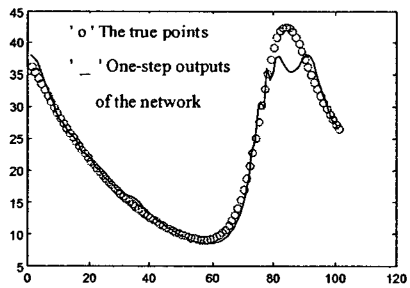

Figure 3 The enlarge of one part of the Figure 2

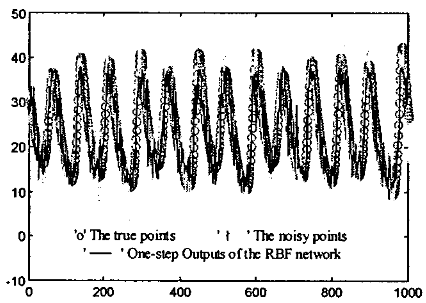

Fig 4 The result of fitting a noisy chaotic system

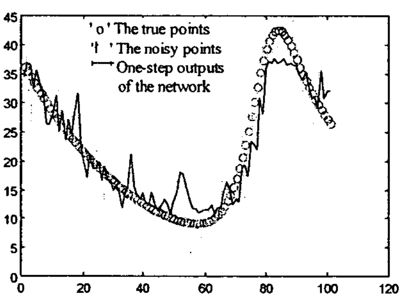

Fig 5 The enlarge of one part of the Fig 4

composed of one-step outputs of the RBF network, " o " are the true points. To show the situation clearly, 100 points from the 900-th to 1000-th are display in Figure 3.

Similarly, Fig 4 and 5 show the results when the signal is contaminated with white noise. Here SNR=5. Apparently, the results of the approximation are satisfactory.

B. THE RESULT ON CLASSIFICATION

In order to test the network the property of classification, we design a training set $\{ P,T \}$ and testing set $\{ P_t, T_t \}$. There are three kinds of input vectors: One is the vector whose elements gradually increase. Another is the vector the second element of which is the least one. The other is the vector whose elements gradually decrease.

$$\text{Here, } P=\begin{pmatrix} 1\,2\,5\,3\,3\,6\,7\,5\,6 \\ 1\,2\,6\,1\,1\,3\,1\,2\,5 \\ 2\,3\,7\,8\,4\,9\,1\,1\,2 \end{pmatrix}, T=\begin{pmatrix} 1\,1\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1\,1\,1 \end{pmatrix}, \text{ and } P_t=\begin{pmatrix} 1\,1\,5\,4\,5\,4\,5\,6\,6\,3\,6\,7\,5\,8 \\ 1\,2\,5\,1\,1\,1\,1\,5\,4\,4\,4\,4\,4\,3 \\ 3\,3\,7\,6\,7\,9\,0\,1\,2\,5\,7\,1\,1\,1 \end{pmatrix},$$

$$T_t=\begin{pmatrix} 1\,1\,1\,0\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,1\,1\,1 \end{pmatrix}.$$ After training the network with $\{ P , T \}$, we test the network with $\{ P_t , T_t \}$. As a result, the rate of recognition is 100%. Of course, the rate of recognition to $\{ P , T \}$ is also 100%. Therefore, the network has the ability in pattern classification.

4 CONCLUSION

In this paper, we describe a novel algorithm for training an RBF network. The network using this new algorithm is very suitable for learning functions from experimental data, and it has the ability in classification. This algorithm has two main advantages: higher accuracy and a more stable training process. There is, however, a shortcoming of the RBF network, which is shared by the wavelet network. That is, much training time are needed. To solve this problem will be our future work.

REFERENCE

[1] L. Bruzzone et al., Structure neural networks for signal classification , Signal Processing , Vol. 64 , pp. 271-290 , 1998.

[2] Jun Zhang et al., Wavelet neural networks for function leaning , IEEE Trans. Signal Processing , Vol. 43 , No. 6 , pp. 1485-1497 , June 1995 .

[3] Simon Haykin et al., Detection of signals in chaos , Proc. IEEE , Vol. 83 , No. 1 , pp 95-122 , Jan. 1995.

[4] E. Uchino et al. , Nonlinear modeling and filtering by RBF network with application to noisy speech signal , Information Sciences , Vol. 101 , pp 177-185 , 1997.

[5] Li Biao et al., A new method of object recognition using a neural network, Systems Engineering and Electronics, Vol 21 , No. 2 , pp 39-42 , 1999. (in Chinese)

Shao-qing Yang , received the Ph. D. degree in automatic control engineering in 2000 from Harbin Institute of Technology. Now he is a postdoctoral researcher at Dalian Maritime University. His research interests are primarily in the area of chaotic signal analysis, computer vision and image understanding, data fusion and target recognition.

Chuan-ying Jia, Professor of Dalian Maritime University.

Correspond Address: 222 Xiao Long Street 1, Dalian, Liaoling, China. PC 116018. *TEL: 0411_2685405. Email: sqyang@online.ln.cn*