LEARNING ALGORITHM OF STAGE CONTROL NBP NETWORK

Yan Lixiang Qin Zheng
(Xi'an JiaoTong University, Xi'an 710049)

Abstract This letter analyzes the reasons why the known Neural Back Promulgation (NBP) network learning algorithm has slower speed and greater sample error. Based on the analysis and experiment, the training group descending Enhanced Combination Algorithm (ECA) is proposed. The analysis of the generalized property and sample error shows that the ECA can heighten the study speed and reduce individual error.

Key words Neural Back Promulgation (NBP) network; Training group descending; Enhanced Combination Algorithm (ECA)

I. Introduction

Neural Back Promulgation (NBP) net is a crucial realization technology in some engineering system such as artificial intelligence, decision support system, and the NBP learning algorithm is a key problem of NBP net[1]. The learning algorithm decides the efficiency and effect of an NBP net in resolving real problems. So the NBP net learning algorithm has been a focus problem in artificial intelligence, NBP net and decision support system for a long time[1]. Most of the known learning algorithms make use of the learning method group by group in learning process[2-4]. With high convergence speed[5], these algorithms can meet the total error requirements, but for slow learning samples, bigger sample error will lead to bigger variance, and finally affect the generalized properties of whole NBP net[3,6]. If we reduce the error to improve the generalized properties, the learning speed will be slower than that of Refs.[7,8]. Based on the research on these known algorithms, we proposed an improved learning algorithm which combines group learning and batch learning, decreases the training set gradually in the training process and retrains the samples which have slow learning speed. Experiment shows that comparing with known algorithms, this algorithm can obtain better generalization properties, higher learning speed while reduce the error greatly.

II. The Betterment Algorithm

During the learning process of NBP net, most of the samples can converge with high speed, only a few samples have slower learning speed. With group learning algorithms, even if the total sample error can meet the error requirement, the slower learning samples still have bigger errors, so they affect the generalized properties of the total net. If we improve the generalized properties by means of reducing error, the learning speed will be slow. In the experiment, we find that combining group learning and batch learning method, decreasing

1 Manuskript received date: April 4, 2003; revised date: July 7, 2003.
Supported by the National Defense Research item “Data fusion” of Tenth Five-Year Plan 102010203
the train set gradually and retraining the samples with slower learning speed can obtain a better generalization properties and higher learning speed.

The process is as follows (see Fig.1): Divide the process into two stages: First train all samples in sample set, if some samples can meet the fixed error requirement $\epsilon_1$, exit the learning process. When all the samples in this sample set can meet the fixed error requirement $\epsilon_1$, the NBP net trains all the samples in the sample set again, until the total net error is less than the fixed error $\epsilon_1$.

There are two methods to form a training set: One is to decrease the training set immediately when some sample reached the fixed error $\epsilon_1$, another is to reorganize a new training set when fixed number samples reached the fixed error $\epsilon_1$. In order to get higher parallel efficiency in a parallel program, the second method and the node number in a parallel computer are often used.

### III. Enhanced Combination Algorithm

After analyzing the advantages and disadvantages of all kinds of algorithms and based on the training set descending method, we propose a new combination algorithm which uses the advantages of the known algorithms. It is called Enhanced Combination Algorithm (ECA). The process is as follows:

1. Assign the initial value for matrix $W_1$ and $W_2$, where $W_1$ is the weight matrix between the hidden layer and input layer, and $W_2$ between the output layer and hidden layer, clear the sample learning marks after completing sample learning. $\gamma$ is assigned an initial value 1 and set $k$ at 1, $t$ at 0.

2. Read the training scheme, they are denoted by

   \[ [I(1), D(1)], [I(2), D(2)], [I(3), D(3)], \cdots, [I(p), D(p)] \]

   For sample $[I(k), D(k)]$,