A Novel Adaptive Learning Algorithm for Stock Market Prediction

Lean Yu[1,2], Shouyang Wang[1,2,3], and Kin Keung Lai[3,4]

[1] Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080, China
[2] School of Management, Graduate University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100039, China
{yulean, sywang}@amss.ac.cn
[3] College of Business Administration, Hunan University, Changsha 410082, China
[4] Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
mskklai@cityu.edu.hk

Abstract. In this study, a novel adaptive learning algorithm for feed-forward network based on optimized instantaneous learning rates is proposed to predict stock market movements. In this new algorithm, the optimized adaptive learning rates are used to adjust the weight changes dynamically. For illustration and testing purposes the proposed algorithm is applied to two main stock price indices: S&P 500 and Nikkei 225. The experimental results reveal that the proposed algorithm provides a promising alternative to stock market prediction.

1 Introduction

Stock market prediction is regarded as a challenging task because of high fluctuation and irregularity. There have many studies using artificial neural networks (ANNs) in this area. Back-propagation neural network (BPNN) is the most popular class of ANNs which have been widely applied to time series prediction. The basic learning rule of BPNN is based on the gradient descent optimization method and the chain rule, as initially proposed by Werbos [1] in the 1970s. Since the basic learning rule is based on the gradient descent method, which is known for its slowness and its frequent confinement to local minima [2], many improved BP algorithms are developed such as variable step size, adaptive learning [3-4] and others [5-6]. Generally, these algorithms have an improved convergence property, but most of these methods do not use the optimized instantaneous learning rates. In their studies, the learning rate is set to a fixed value when learning. However, it is critical to determine a proper fixed learning rate for the applications of the BPNN. If the learning rate is large, learning may occur quickly, but it may also become unstable and even will not learn at all. To ensure stable learning, the learning rate must be sufficiently small, but with a small learning rate the BPNN may be lead to a long learning time. Also, just how small the learning rate should be is unclear. In addition, for different structures of BPNN and for different applications, the best fixed learning rates are different.

X. Deng and D. Du (Eds.): ISAAC 2005, LNCS 3827, pp. 443 – 452, 2005.

There are other ways to accelerate the network learning using second-order gradient based nonlinear optimization methods, such as the conjugate gradient algorithm [6] and Levenberg-Marquardt algorithm [7]. The crucial drawbacks of these methods, however, are that in many applications computational demands are so large that their effective use in many practical problems is not viable.

A common problem with the all above mentioned methods is a non-optimal choice of the learning rate even with the adaptive change of the learning rate. A solution is to derive optimal learning rate formulae for BPNN and then allow an adaptive change at each iteration step during the learning process. The resulting algorithm will eliminate the need for a search for the proper fixed learning rate and provide fast convergence.

Due to the highly nonlinearity of neural networks, it is difficult to obtain the optimum learning rate. In this paper, a new method based on matrix and optimization techniques is proposed to derive the optimal learning rate and construct an adaptive learning algorithm. To test the efficiency of the proposed algorithm, two important stock indices, S&P500 and Nikkei225, are used. The rest of this work is organized as follows. In Section 2, the proposed adaptive learning algorithm with optimal learning rate is presented. In order to testing the proposed algorithm, Section 3 gives an experiment and reports the results. Finally, the conclusions are made in Section 4.

2 The Proposed Adaptive Learning Algorithm

Consider a three-layer BPNN, which has p nodes in the input layer, q nodes in the hidden layer and k nodes in the output layer. Mathematically, the basic structure of the BPNN model is described by

$$
Y(t+1) = \begin{bmatrix} y_1(t+1) \\ y_2(t+1) \\ \cdots \\ y_k(t+1) \end{bmatrix} = \begin{bmatrix} f_2[\sum_{i=1}^{q} f_1(\sum_{j=1}^{p} w_{ij}(t)x_j(t) + w_{i0}(t))v_{1i}(t) + v_{10}(t)] \\ f_2[\sum_{i=1}^{q} f_1(\sum_{j=1}^{p} w_{ij}(t)x_j(t) + w_{i0}(t))v_{2i}(t) + v_{20}(t)] \\ \cdots \\ f_2[\sum_{i=1}^{q} f_1(\sum_{j=1}^{p} w_{ij}(t)x_j(t) + w_{i0}(t))v_{ki}(t) + v_{k0}(t)] \end{bmatrix}
$$

$$
= \begin{bmatrix} f_2[\sum_{i=0}^{q} f_1(\sum_{j=0}^{p} w_{ij}(t)x_j(t))v_{1i}(t)] \\ f_2[\sum_{i=0}^{q} f_1(\sum_{j=0}^{p} w_{ij}(t)x_j(t))v_{2i}(t)] \\ \cdots \\ f_2[\sum_{i=0}^{q} f_1(\sum_{j=0}^{p} w_{ij}(t)x_j(t))v_{ki}(t)] \end{bmatrix} = \begin{bmatrix} f_2[V_1^T F_1(W(t)X(t))] \\ f_2[V_2^T F_1(W(t)X(t))] \\ \cdots \\ f_2[V_k^T F_1(W(t)X(t))] \end{bmatrix}
$$

$$
= F_2[V^T(t)F_1(W(t)X(t))] \tag{1}
$$

where $x_j(t)$, $j = 1, 2, \ldots, p$, are the inputs of the BPNN; y is the output of the BPNN; $w_{ij}(t)$, $i = 1, \ldots, q$, $j = 1, \ldots, p$, are the weights from the input layer to the hidden layer; $w_{i0}(t)$, $i = 1, \ldots, q$, are the biases of the hidden nodes; $v_{ij}(t)$, $i = 1, \ldots, q$, $j = 1, \ldots, k$, are the weights from the hidden layer to the output layer; $v_{i0}(t)$ is the bias of the output node; t is a time factor; f_1 is the activation function of the nodes for the hidden layer and f_2 is the activation function of the nodes for the output layer. Generally, the activation function for nonlinear nodes is assumed to be a symmetric hyperbolic tangent