Comparative Evaluation of Generalized ADALINE using Variable Learning Rate Parameter

Savita[1], Divneet Singh Kapoor[2] and Amit Kumar Kohli[1]

[1] Electronics and Communication Engineering Department, Thapar University, Patiala, Punjab, India

[2] Electronics and Communication Engineering Department, Chandigarh University, Gharuan, Mohali, Punjab, India

Abstract

The ADAptive LINear Element (ADALINE) neural network uses Least Mean Square (LMS) learning rule. This paper presents a comprehensive comparison of three different variable learning rate (VLR) parameter LMS algorithms, for the generalized ADALINE neural network paradigms. These algorithms are used to adjust the weights of the ADALINE neural network, which are tested under three different applications: adaptive prediction, system identification and noise cancellation. The performance analysis of the algorithms for different scenarios is discussed with the help of computer simulations. The simulation results show that in the initial stage, the mean square error between the network output and the desired output is less for the algorithm that has the fastest convergence rate. Therefore, the major advantage is faster convergence of synaptic weights towards the optimum solution, in addition to the better tracking performance in different applications.

Keywords: *ADALINE neural network, variable step-size LMS algorithm, system identification, adaptive prediction, noise cancellation.*

1. Introduction

An artificial neural network is a processor, which is made up of simple processing units called as neurons, which are parallel distributed for storing the knowledge acquired from the experience [1]. Many neural networks use ADAptive LINear Element (ADALINE), sometimes referred to as ADAptive LInear NEuron [2] or simply, adaptive neurons. The ADALINE is a multiple input and single output unit, which receives the input from several neurons. The link between the input and output neurons possesses weighted interconnections, which change during the training of neuron.

The generalized ADALINE neural network has an output feedback, and it also remembers the past input/output data, as the current system output is dependent on the current inputs, past inputs and past outputs. In ADALINE weight adjustment, the generalization of adaptive learning is done by adding a momentum term [3]. The momentum term introduces a smoothening effect in the learning curve, as it affects only during the convergence period. The effect of the momentum term diminishes, when the learning error is significantly reduced while training of the neuron [4]. The simulations for the synaptic weight adjustment of the ADALINE neural network are discussed in this paper. The simulation results show that the VLR algorithms that have faster convergence rate produce less learning error than the other algorithms. The system model for generalized ADALINE network is discussed in Section 2. The VLR algorithms for evaluation under three different scenarios are discussed in Section 3. The computer simulation results are presented in Section 4. In Section 5, some conclusions are drawn and future scope of the work is given.

2. System Model for Generalized ADALINE

The ADALINE is a multiple input and single output unit, which receives the input from several neurons. The ADALINE structure is shown in Fig. 1.

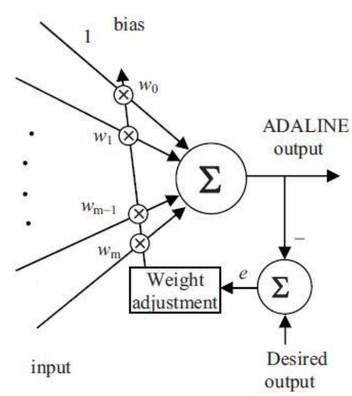

Fig. 1. ADALINE structure

The bias input is not present in the generalized ADALINE structure. The generalized ADALINE output is defined as

$$y = \mathbf{u}^T \mathbf{w} \tag{1}$$

where, **u** is the input vector and **w** is the weight vector, defined as

$$\mathbf{u} = \left[u_1, u_2, \ldots\ldots, u_p\right]^T$$
$$\mathbf{w} = \left[w_1, w_2, \ldots\ldots, w_p\right]^T \tag{2}$$

The error between the desired output and ADALINE output is given as

$$e(n) = d(n) - y(n) \tag{3}$$

where, *d* is the desired output and *n* is the iteration index.

The learning algorithm for the generalized ADALINE neural network is based on the LMS algorithms, which minimizes the error defined above, in the mean square sense. The weight adjustment for the system is defined as

$$\Delta\mathbf{w}(n) = \mathbf{w}(n+1) - \mathbf{w}(n) = \eta(n)e(n)\mathbf{u}(n) \tag{4}$$

where, $\eta(n)$ is the Variable Learning Rate (VLR) parameter, usually in the range of (0,1). The algorithms for deciding the VLR are discussed in the next section. The weight adjustment of generalized ADALINE structure also has a momentum term added to it given as

$$m(n) = \bar{\alpha}\Delta\mathbf{w}(n-1) \tag{5}$$

where, $\bar{\alpha}$ is a small non-negative number. In this paper, the generalized ADALINE neural network is used for three different application scenarios.

a) Adaptive Prediction

In this application, the neural network is used to predict the signal, which is shifting in frequency after some arbitrary time delay.

b) System Identification

In this application, the neural network is used to adapt to the system in the environment to the different frequency signals at the input.

c) Noise Cancellation

The neural network in this application is used to extract the original signal from the signal received with additive noise (both colored and white in terms of statistics).

3. VLR Algorithms for Evaluation

The training of the neural network is governed by the Widrow-Hopf Delta Learning Rule [1], which are LMS algorithms, for adapting the synaptic weights. The key factor for practical implementation of the LMS algorithm is the step-size or learning rate parameter. For rapid convergence of the algorithm, the learning rate has to be large enough, but this also increases the steady state mean-square error [5]. In order to increase the performance of the algorithm, the VLR is used rather than fixed learning rate. In this manuscript, for the aforementioned application scenarios, the comparative evaluation is done for the following three VLR algorithms. The adaptive variable learning rate parameter used for weight adjustments of the neural network are given as

- $$\eta(n+1) = \alpha\eta(n) + \gamma e^2(n) \tag{6}$$

with $0 < \alpha < 1,\ \gamma > 0$, where $\eta(n)$ is the learning rate, n is the iteration index, $e(n)$ is the error defined in Eq.(3), α, γ are the control parameters [6].

- $$\eta(n+1) = \eta(n) + \rho e(n)e(n-1)\mathbf{u}^T(n)\mathbf{u}(n-1) \tag{7}$$

where, ρ is a small positive constant that controls the adaptiveness of the algorithm [7].

- $$\eta(n+1) = \alpha\eta(n) + \gamma p^2(n) \tag{8}$$

where, $p(n) = \beta p(n-1) + (1-\beta)e(n)e(n-1)$, and α, β, γ are those parameters, which control the convergence behavior of the algorithm [8].

The above stated algorithms are incorporated in the neural network synaptic weight adjustments for the aforementioned applications, and the comparative evaluation is discussed in the next section.

4. Simulation Results

The ADALINE neural network is used in the application of adaptive prediction, system identification and noise cancellation, in which the adaptiveness pertaining to the tracking of the input signal(s) is governed by the learning algorithms stated in the previous section. The computer simulations are carried out to evaluate the performance of the VLR algorithms for each of the application scenario. The parameter $\bar{\alpha}$ in Eq. (5) for adding the momentum is taken to be 0.1 in all the application scenarios.

Case – I: Adaptive Prediction

An input signal is to be predicted with the help of ADALINE neural network. The weight adjustments of the neural network according to the application are performed by the algorithms specified in the Eq. (6), (7), and (8). A sinusoidal signal of 2 kHz frequency, with 12.5 μs sampling time, is taken in conjunction with the sinusoidal signal of 4 kHz frequency, with the total duration of 7.5 s. The target signal and the predicted signal are shown in Fig.2. The ADALINE neural network with 4 synapses is considered for the adaptive prediction of the input signal. The algorithms are implemented with the parameters α = 0.97, γ = 0.05 in Eq. (6), ρ = 0.001 in Eq. (7), and α = 0.97, γ = 0.001, β = 0.99 in Eq. (8).

ISSN (Online): 2230-7893
www.IJCEM.org

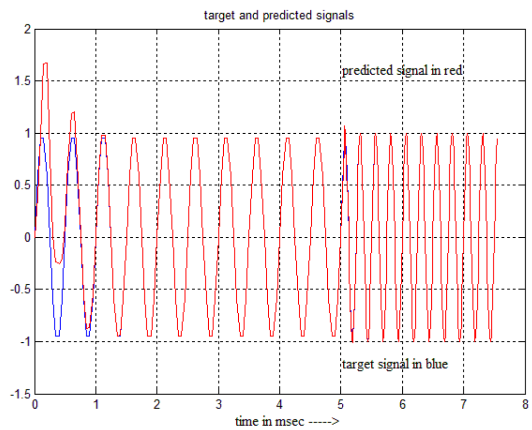

Fig. 2. Target and predicted signals for adaptive prediction.

The comparative results of the squared prediction error in dB, averaged over 100 iterations, for the three algorithms are shown in Fig. 3. It can be demonstrated that the algorithm specified in Eq. (8) results in less prediction error, relatively to the others, as it has faster convergence rate than the other algorithms [8].

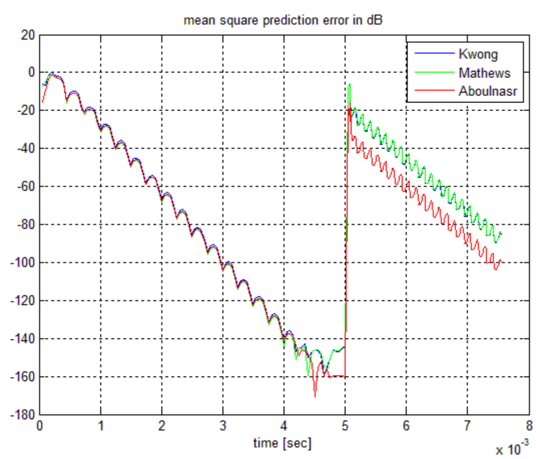

Fig. 3. The mean square prediction error of VLR algorithms for adaptive prediction.

Case – II: System Identification

The ADALINE neural network is used to model a system excited with different frequency signals for each iteration. The system dimensionality is taken to be 3. The system changes the coefficients from [1 -0.6 0.4] to [0.9 -0.5 0.7] after 4 seconds. For the underlying application, the algorithms are implemented with the parameters $\alpha = 0.99$, $\gamma = 0.01$ in Eq. (6), $\rho = 0.01$ in Eq. (7), and $\alpha = 0.97$, $\gamma = 0.001$, $\beta = 0.99$ in Eq. (8). The weights of the neural network are assumed to be random for the initial setting. The input signal and the predicted signal for each algorithm are shown in Fig. 4. The experiment is performed 100 times, and the mean square estimation error results (in dB) on the basis of ensemble average are plotted in Fig. 5. In this application, the algorithm specified in Eq. (8) outperforms the other algorithms in terms of mean square estimation error.

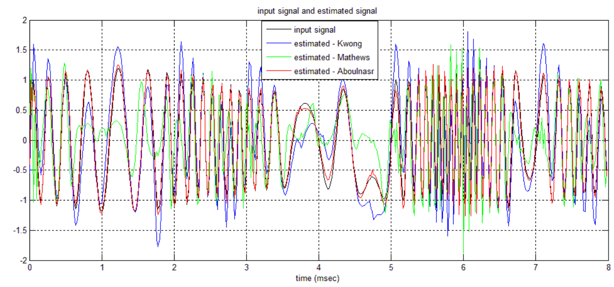

Fig. 4. Input and estimated signals for different algorithms.

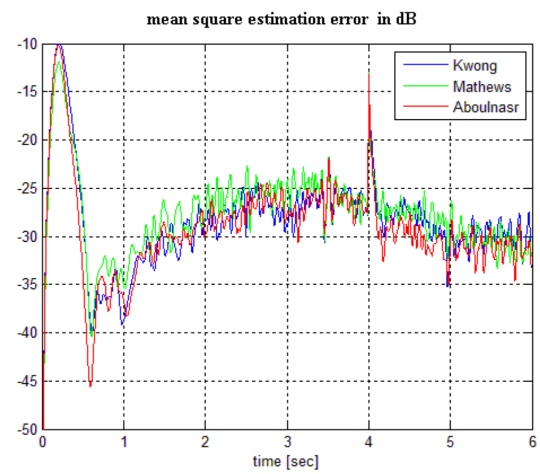

Fig. 5. The mean square estimation error (dB) of VLR algorithms.

Case – III: Noise Cancellation

In this application, ADALINE neural network is used to extract the target signal which is corrupted by the noise (both white and colored). The original input signal, noisy input signal, and the types of noise are depicted in Fig. 6.

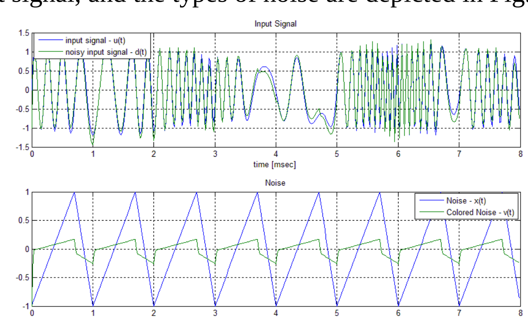

Fig. 6. Input signal, noisy input signal, and types of noise.

The parameters for implementing the algorithms used for the application of noise cancellation are taken as $\alpha = 0.95$, $\gamma = 0.01$ in Eq. (6), $\rho = 0.005$ in Eq. (7), and $\alpha = 0.95$, $\gamma = 0.01$, $\beta = 0.99$ in Eq. (8). The mean square error (in dB) is calculated for this application under different algorithms. The simulations are averaged over 100 iterations, and the comparison is shown in Fig. 7.

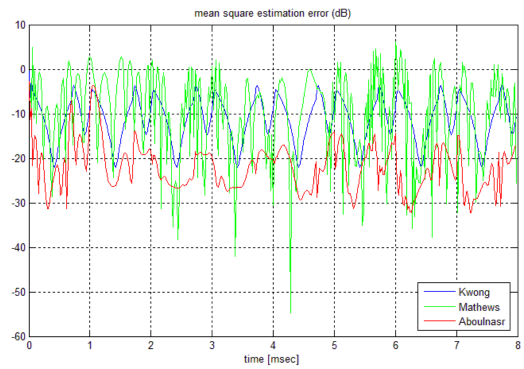

Fig. 7. Mean square error (dB) for noise cancellation scenario.

If we compare Kwong's algorithm [6] with Mathews's algorithm [7], then it may be inferred that under low signal-to noise ratio and correlated environment, the initial convergence rate of Mathews' algorithm is very high, but the tracking performance is relatively poor. On the other hand, under high signal-to-noise ratio and correlated environment, the convergence rate as well as the tracking performance of Mathews' algorithm is approximately similar to the Kwong's algorithm. Under uncorrelated environment, the Mathews's algorithm exhibits high convergence rate only under high signal-to-noise ratio conditions, while giving high mean square error in the tracking mode. However, the algorithm of Aboulnasr [8] supersedes both former algorithms in performance.

5. Concluding Remarks and Future Scope

In this paper, the performance of three different VLR algorithms applied to the weight adjustments of the ADALINE neural network for the different scenarios are compared. It is observed from the analysis, as shown in Fig. 3, Fig. 5 and Fig. 7, that the algorithm given by Aboulnasr *et al.* in [8] results in less estimation and prediction errors for the different applications. In addition, this algorithm has faster convergence rate [8], as compared to other algorithms under consideration. Future work includes the development of gradient based VLR algorithms for ADALINE neural networks.

References

[1] S. Haykin, Neural Networks: A Comprehensive Foundation, Second Edition, Prentice Hall, 1999.

[2] B. Widrow, and M. E. Hoff Jr., "Adaptive Switching Circuits," in 1960 IRE WESCON Convention Record, 1960, Part-4, New York: IRE, pp. 96-104.

[3] Wenle Zhang, "System Identification Based on a Generalized ADALINE Neural Network," in Proc. of the 2007 American Control Conference, July 11-13, 2007, New York, pp. 4792-4797.

[4] Wenle Zhang, "On Training Optimization of the Generalized ADLINE Neural Network for Time Varying System Identification," in Chinese Control and Decision Conference (CCDC), 2009, China, pp. 775-780.

[5] Yonggang Zhang, Ning Li, Jonathon A. Chambers, and Yanling Hao, "New Gradient-Based Variable Step Size LMS Algorithms," EURASIP Journal on Advances in Signal Processing, Hindawi Publishing Corporation, 2008.

[6] Raymond H. Kwong, Edward W. Johnston, "A Variable Step Size LMS Algorithm," IEEE Transactions on Signal Processing, Vol. 40, No. 7, July 1992.

[7] V. John Mathews, Zhenhua Xie, "A Stochastic Gradient Adaptive Filter with Gradient Adaptive Step Size," IEEE Transactions on Signal Processing, Vol. 41, No. 6, June 1993.

[8] Tyseer Aboulnasr, K. Mayyas, "A Robust Variable Step-Size LMS-Type Algorithm: Analysis and Simulations," IEEE Transactions on Signal Processing, Vol. 45, No. 3, March 1997.

[9] Jayaraj. V, Jegathesh Amalraj. J, "Using Clustering and Indexing to Enhance Customer Relationship Management Based on Customer and Product Value Estimation - A Neural Networks Approach" International Journal of Computational Engineering and Management IJCEM, Vol. 15 Issue 4 July 2012.

Savita received her B.E. degree in Electronics and Communication Engineering from Thapar University, Patiala, in 2009. She has worked as a Lecturer in Swami Parmanand College, Mohali from June 2009 - July 2011. She is currently pursuing her M.E. in Electronics and Communication Engineering from Thapar University, Patiala. Her research areas include neural networks, and signal processing.

Divneet Singh Kapoor received the B.E. degree in Electronics and Communication Engineering from Thapar University, Patiala (Punjab), and M.E. degree in Electronics and Communication Engineering from Thapar University, Patiala (Punjab), in 2009 and 2011 respectively. He is currently working as Assistant Professor in Electronics and Communication Engineering Department of Chandigarh University, Mohali. His research interests include wireless communication and signal processing.

Amit Kumar Kohli received the B.Tech. (Honor) degree from Guru Nanak Dev Engineering College, Ludhiana, the M.E. (Highest Honor) degree from Thapar University Patiala (previously Thapar Institute of Engineering and Technology) and the Ph.D. degree from Indian Institute of Technology, Roorkee, India, in 2000, 2002 and 2006 respectively, all in electronics and communication engineering. He is presently Senior Assistant Professor in Electronics and Communication Engineering Department of Thapar University Patiala, India. His research interests include signal processing and its applications, modeling of fading dispersive channels, wireless and high data rate advanced communication systems, neural networks, bio-signal processing and adaptive system design. He is reviewer of distinguished international journals like IEEE, Springer, Elsevier and Taylor & Francis etc. He received National Scholarship, State Scholarship, MHRD Fellowship, Institution Medal, and Gold Medal consecutively for academic performance.