2012 International Conference on Information and Network Technology (ICINT 2012)
IPCSIT vol. 37 (2012) © (2012) IACSIT Press, Singapore

Modified NLMS Algorithm Using Adaptive Learning Rate in Nonstationary Environment

Manish Sawale [+] and Ram Yadav

Maulana Azad National Institute of Technology, Bhopal, India

Abstract. In this paper a modified Normalized Least Mean Square (NLMS) algorithm is derived using past weight vectors and adaptive learning rate for nonstationary channels. The proposed minimization criterion minimizes the mean square error (MSE) of currently updated weight vector and past weight vector. The MSE of the modified NLMS algorithm is plotted for various values of adaptive learning rate from zero to one. The result shows that as the adaptive learning rate decreases the convergence rate increases. The result also shows that the modified NLMS algorithm with active weight detection is capable of determining changes in position and strength of the weight vector coefficients in time-varying channel.

Keywords: NLMS algorithm, nonstationary channel, adaptive learning rate, mean square error, past weight vector.

1. Introduction

A wide variety of recursive algorithms have been developed for the operation of linear adaptive filters. The choice of one algorithm over another is determined by one or more of the factors such as, rate of convergence, misalignment, computational complexity [1-2]. For nonstationary environment the tracking factor is important to track the statistical variations. Thus the desired objective can be achieved by altering the minimization criterion of an algorithm. The NLMS algorithm is designed to overcome the difficulty of gradient noise amplification problem suffered by Least Mean Square (LMS) algorithm. The conventional affine projection (AP) algorithm is modified using the set-membership affine projection (SM-AP) algorithm. The minimization criterion of SM-AP is derived using set-membership filtering to reduce the computational complexity of the conventional affine projection [3]-[4]. The modified criterion were developed for improving numerical stability using Leaky recursive least squares (LRLS) and leaky least mean squares (LLMS) algorithms [5].

A new update formula was designed called the momentum LMS (MLMS) by adding a momentum term which contains previous weight vectors [6]. The method reduces the misalignment, but adding a momentum term is not based on a specific criterion. An improved NLMS was designed using a specific criterion for stationary channel [7]. In this paper, we develop a modified minimization criterion of the NLMS algorithm for nonstationary channel using the adaptive learning rate which minimizes the summation of each squared Euclidean norm of difference between the currently updated weight vector and past weight vectors. The new minimization criterion of our algorithm which is based on past weight vectors and adaptive learning rate for nonstationary channel shows improved convergence speed as the adaptive learning rate decreases.

There are various ways in which the nonstationary or time-varying channel can be modelled. The channel can vary with respect to either the strength of each active coefficient, the position of each active coefficient, or both. The least squares (LS) method is used to detect the position of the active coefficient in the unknown channel and the NLMS algorithm works out the strength of the active weights. Channels that

[+] Corresponding author. Tel.: + +919893095600.
E-mail address: manish.sawale@gmail.com.

have a time-varying nature require a window which must adjust according to the recent channel. There are several types of windows that can be used to track the nonstationary coefficients. The exponential window is used in this paper as it has the advantage that more importance is placed on the most recent samples [8].

2. The Constrained Minimization Criterion of Conventional NLMS Algorithm

The NLMS algorithm is built around the transversal filter in which the filter coefficients or weights are updated at p+1 iteration with respect to the squared Euclidean norm of the input vector $\mathbf{x}(p)$ at iteration p. We assume the transversal filter as a finite impulse response (FIR) model wherein the impulse response sequence is defined as $\mathbf{w}^0$ called an R by 1 optimal weight vector. Thus, the measured output of the FIR model is expressed as follows [1]:

$$d(i) = y(i) + \eta(i) \tag{1}$$

Where $y(i) = \mathbf{x}_i \mathbf{w}^0$, is a 1 by N input vector, and $\eta(i)$ is noise measurement at time instant i. The value of N is much larger than the value of R the impulse response sequence of the FIR model.

The conventional NLMS algorithm is a renowned adaptive filtering algorithm. The algorithm is derived from the following constrained minimization criterion:

$$\min_{\mathbf{w}_i} \| \mathbf{w}_i - \mathbf{w}_{i-1} \|^2 \qquad \text{subject to } d(i) = \mathbf{x}_i \mathbf{w}_i \tag{2}$$

Where $\mathbf{w}_i$ is an estimated weight vector at time instant i.

The weight update formula of the conventional NLMS algorithm is derived from the constraint minimization criterion by using the method of Lagrange multiplier [9].

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \frac{\mathbf{x}_i^T}{\mathbf{x}_i \mathbf{x}_i^T} e(i) \tag{3}$$

Where $e(i) = d(i) - \mathbf{x}_i \mathbf{w}_{i-1}$ and $(.)^T$ indicates the transpose operation.

3. The Proposed NLMS Algorithm for Nonstationary Channel

3.1. The proposed minimization criterion and update formula

We propose an improved minimization criterion for the NLMS algorithm which reuses past n weight vectors, as follows:

$$\min_{\mathbf{w}_i} \sum_{k=1}^{n} \rho^{k-1} \| \mathbf{w}_i - \mathbf{w}_{i-k} \|^2 \quad \text{subject to } d(i) = \mathbf{x}_i \mathbf{w}_i \tag{4}$$

where ρ is adaptive learning rate and $0 < \rho < 1$

Using the method of Lagrange multiplier, we rewrite the minimization criterion (4) as follows:

$$J = \sum_{k=1}^{n} \rho^{k-1} \| \mathbf{w}_i - \mathbf{w}_{i-k} \|^2 + \lambda g(i) \tag{5}$$

Where $g(i) = d(i) - \mathbf{x}_i \mathbf{w}_i$ and λ is a Lagrange multiplier. By using the differential relationships of

$$\frac{\partial J}{\partial \mathbf{w}_i} = 2\sum_{k=1}^{n} \rho^{k-1} (\mathbf{w}_i - \mathbf{w}_{i-k}) - \lambda \mathbf{x}_i^T = 0 \tag{6}$$

and

$$\frac{\partial J}{\partial \lambda} = d(i) - \mathbf{x}_i \mathbf{w}_i = 0 \tag{7}$$

From (6), we get

$$\mathbf{w}_i = \alpha \sum_{k=1}^{n} (\rho^{k-1} \mathbf{w}_{i-k}) + \alpha \frac{\lambda}{2} \mathbf{x}_i^T \tag{8}$$

Where

$$\alpha = \left(\sum_{k=1}^{n} \rho^{k-1} \right)^{-1} \tag{9}$$

Substituting (8) into (7) we get,

$$d(i)-\mathbf{x}_i\left(\alpha\sum_{k=1}^{n}\left(\rho^{k-1}\mathbf{w}_{i-k}\right)+\alpha\frac{\lambda}{2}\mathbf{x}_i^T\right)=0 \tag{10}$$

Therefore,

$$\alpha\frac{\lambda}{2}=\frac{1}{\mathbf{x}_i\mathbf{x}_i^T}\left(d(i)-\mathbf{x}_i\alpha\sum_{k=1}^{n}\left(\rho^{k-1}\mathbf{w}_{i-k}\right)\right) \tag{11}$$

By substituting (11) into (8), the proposed NLMS algorithm becomes:

$$\begin{aligned}\mathbf{w}_i &= \alpha\sum_{k=1}^{n}\rho^{k-1}\mathbf{w}_{i-k} + \\ &\cdot\frac{\mathbf{x}_i^T}{\mathbf{x}_i\mathbf{x}_i^T}\left(d(i)-\mathbf{x}_i\alpha\sum_{k=1}^{n}\left(\rho^{k-1}\mathbf{w}_{i-k}\right)\right)\end{aligned} \tag{12}$$

We can write this as:

$$\mathbf{w}_i=\bar{\mathbf{w}}_{i-1,n}+\frac{\mathbf{x}_i^T}{\mathbf{x}_i\mathbf{x}_i^T}\left(d(i)-\mathbf{x}_i\,\bar{\mathbf{w}}_{i-1,n}\right) \tag{13}$$

Where

$$\bar{\mathbf{w}}_{i-1,n}=\alpha\sum_{k=1}^{n}\left(\rho^{k-1}\mathbf{w}_{i-k}\right) \tag{14}$$

When ρ is one, the proposed NLMS algorithm simply becomes as follows:

$$\mathbf{w}_i=\frac{1}{n}\sum_{k=1}^{n}\mathbf{w}_{i-k}+\frac{\mathbf{x}_i^T}{\mathbf{x}_i\mathbf{x}_i^T}\left(d(i)-\mathbf{x}_i\frac{1}{n}\sum_{k=1}^{n}\mathbf{w}_{i-k}\right) \tag{15}$$

Comparing the conventional NLMS algorithm (3) with the weight update formula using adaptive learning rate (15), we find that the proposed algorithm is composed of the averaged past n weight vectors that need additional calculations. As the past weight vectors increases the number of calculations also increases but the additional calculations are not numerically complex and can be performed simply by addition and division operations. It can be seen that, when n=1, the proposed NLMS algorithm including the computational complexity, is identical to the conventional NLMS.

3.2. Modeling of nonstationary channel

There are various ways in which the nonstationary or time-varying channel can be modelled. The channel can vary with respect to either the strength of each active weight, the position of each active weight, or both. The variation of the strength of the known channel taps is one approach for modelling a nonstationary condition. We have used random walk model to incorporate nonstationary channel. The random walk theory is defined as following:

$$\mathbf{w}_{k+1}=\sum\mathbf{w}_k * random\_number * \sigma_\delta \tag{16}$$

Where * denotes multiplication and $\mathbf{w}_k$ is the response of the channel at time index k.

$$\begin{aligned}\mathbf{w}_{k+1}-\mathbf{w}_k=\delta_k=\left[\delta_{0,k},\delta_{1,k},\cdots,\delta_{n-1,k}\right]= \\ \mathbf{w}(k) * random\_number\_2 * \sigma_k\end{aligned} \tag{17}$$

here $\delta_{i,k}$ is zero mean, with variance σ_δ^2 (measures the speed of parameter change) and independent of $\delta_{j,k}$ $\{i\neq j\}$; δ is sometimes known as a Gaussian white noise;

$$\sigma_\delta=\frac{random\_number * \sigma_\delta - 1}{random\_number\_2} \tag{18}$$

Where

$$E\left[\delta_{i,k}\right]=0;$$
$$E\left[\delta_{i,k}\right]=\sigma_\delta^2 \qquad \{\forall_i\};$$
$$E\left[\delta_{i,k},\delta_{j,k}\right]=0.$$

3.3. Active tap detection using exponential sliding window

In adaptive estimation applications, the channel is characterised by a time domain impulse response. In order to detect a weight, a formula known as the LS activity measure is used:

$$X_i=\frac{\sum_{k=i+1}^{N}\left[d_k\ x_{k-i}\right]^2}{\sum_{k=i+1}^{N}x_{k-i}^2} \tag{19}$$

Where d=desired signal, i=tap index and N= number of samples.

Now in order for the tap to be determined as active (as opposed to inactive), the value of X_i must be above a certain minimum value called the active weight threshold condition. This is shown in the following threshold formula:

$$X_i(N)>\sigma_d(N)\log(N) \tag{20}$$

Where σ_d^2 is the variance of d_k .

If the value is found to be below this criterion, then it can be discarded. So in summary, the LS activity criterion (19), (20) works out or detects the position of the active taps in the unknown channel or echo path and the modified NLMS algorithm works out the strength of the active taps.

By theory for a stationary channel, the length of the window which tracks the channel is the length of the number of samples. However, channels that have a time-varying nature require a window which must adjust to the recent channel. There are several types of windows that can be used to track the nonstationary taps. These can be rectangular, triangular, exponential, and such. The problem with the simple rectangular-shaped window is that it allows the least recent samples to have an equivalent rating to the most recent. The problem with triangular window tracking technique is that it assumes that the nature of the channel has a linear relation with respect to the past samples. Hence we have chosen exponential window which has the advantage of more importance being placed on the most recent samples. The equivalent equations to (19), (20) for the exponential sliding window approach are:

$$X_i=\frac{\sum_{k=i+1}^{N}\left[\beta^{N-k}\ d_k\ u_{k-i}\right]^2}{\sum_{k=i+1}^{N}\beta^{N-k}\ u_{k-i}^2} \tag{21}$$

$$\sigma_d^2=\frac{\sum_{k=1}^{N}\beta^{N-k}d_k^2}{\sum_{k=1}^{N}\beta^{N-k}} \tag{22}$$

and

$$X_i>2\sigma_d\log\left[1-\frac{N-k}{R}\right] \tag{23}$$

where β is the exponential factor and obeys the limit of $0<\beta<1$. The smaller the size of β is equivalent to a smaller effective window length. This means that more importance to the more recent samples will be achieved. In contrast, if the exponential factor β is large, the effective window length will be larger and importance to recent samples will be less.

4. Results

In this section, the simulation results of the proposed algorithm (13) are compared with the conventional NLMS algorithm (3) for nonstationary channels. The channel is assumed as FIR model with channel length of 36, i.e., R is 36, and three active weights are chosen, at the 4th weight, the strength is 1; at the 15th weight, the strength is –2; and at the 26th weight, the strength is 3. The input and disturbance signals used are random and of length 2000 with zero mean white Gaussian signals and a variance of 1.0.

In Fig. 1, Fig. 2 and Fig. 3 the MSE of proposed NLMS algorithm for various values of ρ are shown at various samples. The results show that as the adaptive learning rate decreases the convergence rate increases. It can be seen in Fig 4 and Fig 5, that the improved NLMS with active tap detection is capable of finding the position of the nonzero weights of the channel and determining the strengths for each of these coefficients. The value of ρ is kept as one for these calculations.

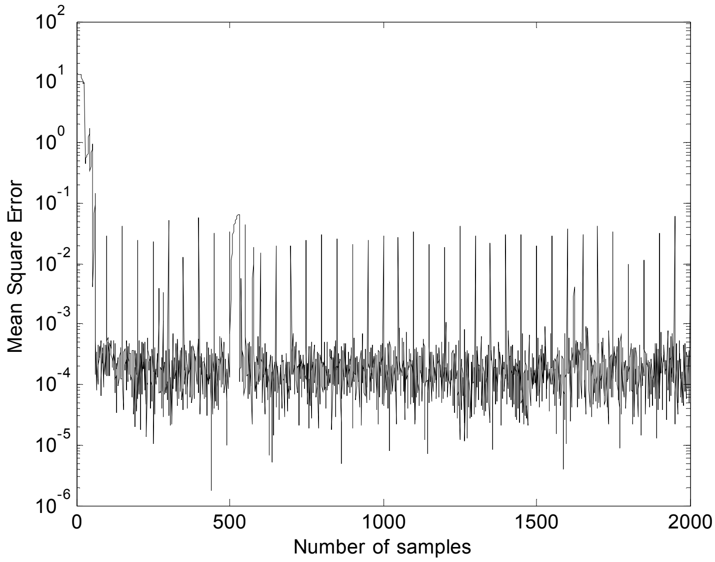

Fig. 1: MSE of proposed NLMS algorithm with ρ =1 for various input samples

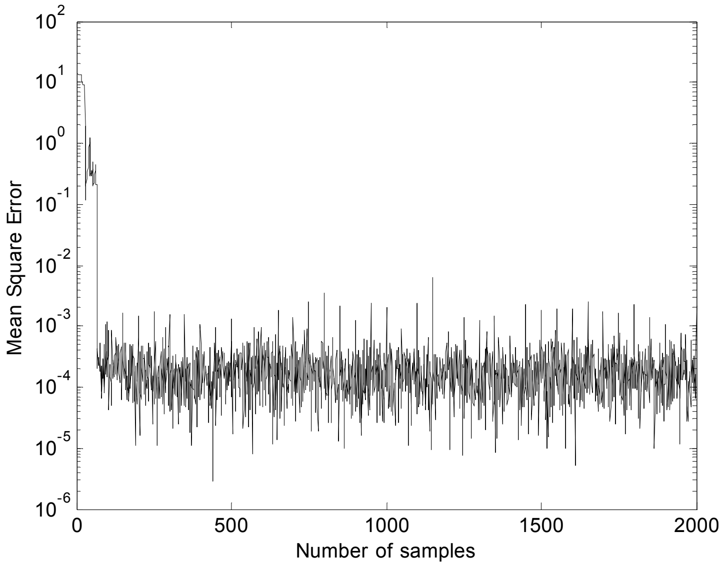

Fig. 2: MSE of proposed NLMS algorithm with ρ =0.1 for various input samples

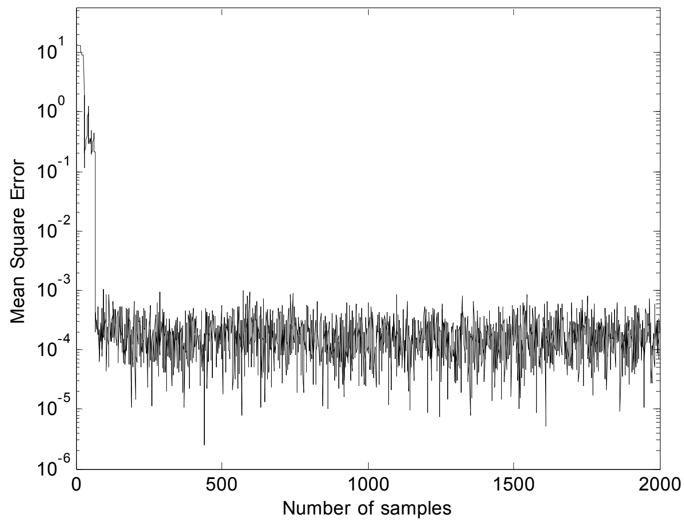

Fig. 3: MSE of proposed NLMS algorithm with ρ =0.01 for various input samples

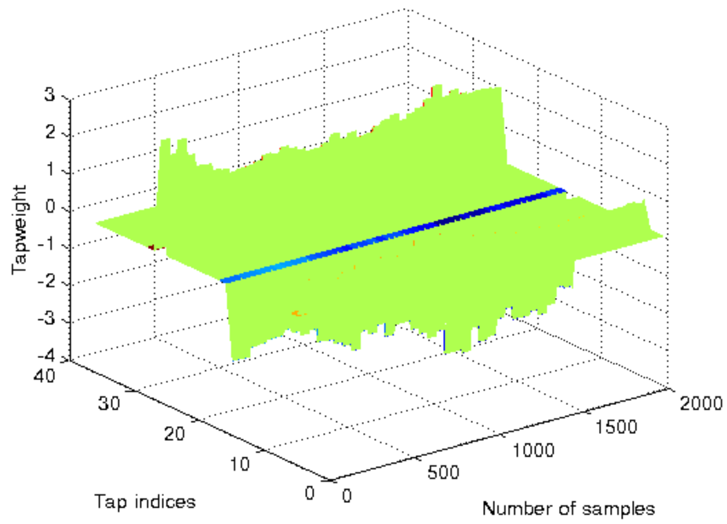

Fig. 4: Detection of actual active taps in nonstationary channel

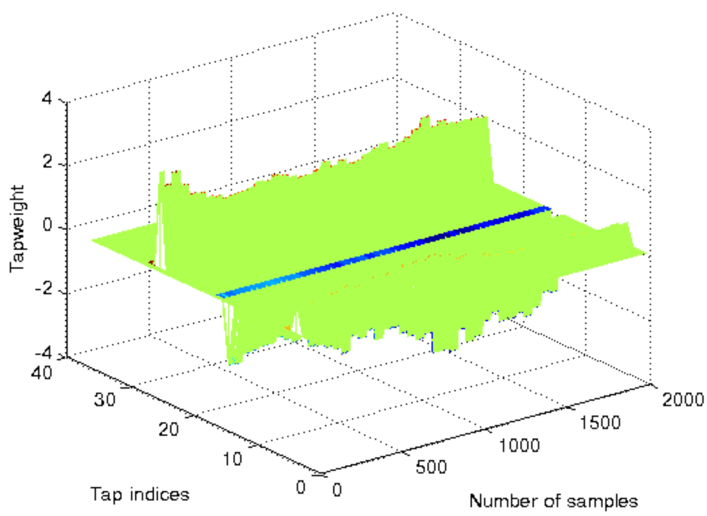

Fig. 5: Detection of active taps using proposed NLMS algorithm in nonstationary channel

5. References

[1] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, 2003.

[2] C.W. Lee, H. Cho, S.W. Kim, Comments on Stereo echo cancellation algorithm using adaptive update on the basis of enhanced input signal vector, *Signal Processing*, 2008, **88** (4): 1079-1081.

[3] S. L. Gay and S. Tavathia, The fast affine projection algorithm, in *Proc. IEEE ICASSP*, Detroit, Michigan, 1995, pp. 3023 – 3026.

[4] S. Werner, S. Diniz, S. Kapoor, Y. F. Huang, Set-membership affine projection algorithm, *IEEE Signal Process. Letter*, 1998, **5** (5): 111-114.

[5] E. Horita, K. Sumiya, H. Urakami, S. Mitsuishi, A leaky RLS algorithm: its optimality and implementation, *IEEE Trans. Signal Processing*, 2004, **52** (10): 2924-2936.

[6] S. Roy, J.J. Shynk, Analysis of the momentum LMS algorithm, *IEEE Trans. Signal Processing*, 1990, **38** (12): 2088-2098.

[7] Cho, H., Lee, C.W., Kim, S.W., Derivation of a new normalized least mean squares algorithm with modified minimization criterion, *Signal Processing*, 2008, **89** (4): 692-695.

[8] J. Homer, I. Mareels, R. R. Bitmead, B. Wahlberg, and A. Gustafsson, LMS estimation via structural detection, *IEEE Trans. Signal Processing*, 1998, **46** (10): 2651-2663

[9] S. Haykin, *Adaptive Filter Theory*, fourth ed., Prentice-Hall, 2002.