MSE ANALYSIS OF THE M-MAX NLMS ADAPTIVE ALGORITHM

T. Aboulnasr
School of Information Technology and Engg.
University of Ottawa
Ottawa, Ontario, Canada K1N 6N5
aboulnas@eng.uottawa.ca

K. Mayyas
Department of Electrical Engg.
Jordan University of Science and Technology
Irbid, Jordan 221 10
mayyas@just.edu.jo

Abstract

In this paper, we provide a mean square analysis of the M-Max NLMS (MMNLMS) adaptive algorithm introduced in [1]. The algorithm selects, at each iteration, a specified number of coefficients that provide the largest reduction in the error. It is shown that while the MMNLMS algorithm reduces the complexity of the adaptive filter, it maintains the closest performance to the full update NLMS filter for a given number of updates. The stability of the algorithm is shown to be guaranteed for the extreme case of only one update/iteration. Analysis of the MSE convergence and steady state performance for i.i.d. signals is also provided for that extreme case.

1 Introduction

Several algorithms were proposed to reduce the complexity cost of the NLMS algorithm [1,2,3]. In this paper, we focus on two recent algorithms, namely the MMNLMS algorithm [1], and the Max-NLMS algorithm [2]. The M-Max algorithm updates M coefficients out of N at each iteration. Those M coefficients are the ones associated with the M largest $|x(n-i+1)|$, $i = 1,..,N$, at that iteration. The algorithm update equation can be written as [1]

$$w_i(n+1) = \begin{cases} w_i(n) + \frac{\mu}{\mathbf{X}^{\mathbf{T}}(n)\mathbf{X}(n)} e(n)x(n-i+1)\ , \\ \text{if } i \text{ corresponds to one of the first} \\ M \text{ maxima of } |x(n-i+1)|,\ i = 1,..,N \\ w_i(n)\ , \quad \mathit{otherwise} \end{cases} \tag{1}$$

The complexity of the algorithm, excluding the overhead of calculating $\mathbf{X}^{\mathbf{T}}(n)\mathbf{X}(n)$ [1], is $N+M+1$ multiplications, $N+M$ additions, a single division, and $2\log_2(N)+2$ comparisons at most.

The Max-NLMS algorithm is described by [2]

$$w_i(n+1) = \begin{cases} w_i(n) + \frac{\mu}{x(n-i+1)} e(n) \\ \text{if } |x(n-i+1| = \max|x(n-j+1)|, \\ \quad j = 1,..N \\ w_i(n)\ , \quad \mathit{otherwise} \end{cases} \tag{2}$$

For $\mu = 1$, the algorithm provides the minimum absolute change in the filter coefficients at each sample time subject to zero a posteriori error. At each iteration, the algorithm updates only one coefficient which is associated with maximum absolute value of the input data window. The algorithm requires $N+1$ multiplications, N additions, a single division, and 3 comparisons.

We will show that the M-Max algorithm reduces the complexity of the NLMS while preserving performance as close as possible to the regular NLMS. A study of the MMNLMS convergence properties is also presented. Simulation results are provided to illustrate the advantages in performance of the M-Max algorithm compared to the Max-NLMS for $M = 1$.

2 The M-Max NLMS versus the full update NLMS

It is shown in [4] that the full update NLMS algorithm results in the minimum possible value of the squared error $e^2(n+1)$ with $\mu = 1$. We show here that the MMNLMS algorithm, with its step size constrained to be $\frac{\mu}{\mathbf{X}^{\mathbf{T}}(n)\mathbf{X}(n)}$, leads to closest possible performance to the full NLMS algorithm when both are used with same step size value μ.

The Talyor series expansion of $e(n+1)$, the error at time instant $n+1$, in terms of $e(n)$, the error at time instant n, is given by [4]:

$$\begin{aligned} e(n+1) &= e(n) + \sum_{j=1}^{N} \frac{\partial e(n)}{\partial w_j(n)} \Delta w_j \\ &+ \frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 e(n)}{\partial w_j(n)\partial w_k(n)} \Delta w_j \Delta w_k + \ldots \end{aligned} \tag{3}$$

where $\Delta w_j = w_j(n+1) - w_j(n)$. The error $e(n)$, given by $e(n) = d(n) - \mathbf{X}^{\mathbf{T}}(n)\mathbf{W}(n)$, is a linear function of $w_j(n)$, $j = 1,2,...,N$, therefore Eq.(3) reduces to

$$e(n+1) = e(n) + \sum_{j=1}^{N} \frac{\partial e(n)}{\partial w_j(n)} \Delta w_j \tag{4}$$

For the full update NLMS algorithm, Δw_j is given by

$$\Delta w_j = \frac{\mu}{\mathbf{X^T}(n)\mathbf{X}(n)} e(n)x(n-j+1)\ ,\ \ j=1,2,...,N \tag{5}$$

Substituting Eq.(5) in Eq.(4) and squaring the error $e(n+1)$ results in

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

The best instantaneous performance of the full update NLMS algorithm is achieved when $\mu = 1$, which leads to $e(n+1) = 0$ [4]. Consider a partial update of a set of M out of the N coefficients; $w_{i_1},\ w_{i_2}, \ldots,\ w_{i_M}$. At any iteration, the quantity Δw_j is given by

$$\Delta w_j = \begin{cases} 0 & ,\ \text{if}\ j \neq i_k \\ \frac{\mu}{\mathbf{X^T}(n)\mathbf{X}(n)} e(n)x(n-j+1) & ,\ \text{if}\ j = i_k, \end{cases} \tag{7}$$

where $k = 1,2,...,M$. Substituting Eq.(7) in Eq.(4) and squaring the error $e(n+1)$ results in

$$e^2(n+1) = e^2(n)(\ 1 - \frac{\mu}{\mathbf{X^T}(n)\mathbf{X}(n)} \sum_{j=1}^{M} x^2(n-i_j+1)\ )^2 \tag{8}$$

It is clear that for $M = N$, Eq.(8) reduces to Eq.(6) of the full update NLMS. In other words, as $\sum_{j=1}^{M} x^2(n-i_j+1)$ approaches $\mathbf{X^T}(n)\mathbf{X}(n)$, the convergence speed of the partial update algorithm will approach that of the full update NLMS. In the MMNLMS algorithm, the M coefficients to be updated are chosen to correspond to the M largest $x^2(n-j+1)$, $j = 1,\ ,2, \ldots, N$, thus resulting in the largest $\sum_{j=1}^{M} x^2(n-i_j+1)$ at the nth iteration. Clearly, this results in $e^2(n+1)$ being the smallest possible for a given M, i.e., the closest to $e^2(n+1)$ for a full update NLMS.

3 Mean square analysis of the M-Max NLMS

In this section, we will study the convergence properties of the MMNLMS algorithm. To make the analysis tractable, we will only consider the case for $M = 1$. Our objective is to show that the algorithm is guaranteed to converge for $M = 1$ (provided μ is chosen in the stability region) and that it will converge to the same steady state error as the full update NLMS. It thus follows that the algorithm will converge for $M > 1$ (we already know it converges for $M = N$).

Assuming that $x(n)$ is drawn independently from a known probability density function, and defining the error vector $\mathbf{V}(n) = \mathbf{W}(n) - \mathbf{W}^*$, then for $M = 1$, Eq.(1) becomes

$$V_i(n+1) = \begin{cases} (1-\mu(n)x^2(n-i+1))V_i(n) \\ -\mu(n)\sum_{j=1,j\neq i}^{N} x(n-i+1)x(n-j+1). \\ V_j(n) + \mu(n)x(n-i+1)e^*(n) \\ \quad \text{if } i \text{ corresponds to the maximum} \\ \text{of} |x(n-i+1)|,\ i = 1,..,N \\ \\ V_i(n) \qquad \textit{otherwise} \end{cases} \tag{9}$$

where $\mu(n) = \frac{\mu}{\mathbf{X^T}(n)\mathbf{X}(n)}$. In [2], it is shown that the autocorrelation matrix of the Max-NLMS algorithm in Eq.(2) can have negative eigenvalues for certain classes of input signals identified in [2]. This causes the divergence of the Max-NLMS algorithm irrespective of the step size value used. Following the same approach in [2], and assuming that for high order adaptive filters $\mathbf{X^T}(n)\mathbf{X}(n) \approx N\sigma_x^2$, it can be easily seen that the autocorrelation matrix governing the evolution of the mean error weight vector in Eq.(9) is $\mathbf{R} = \frac{\sigma_x^2}{N}\mathbf{I}$, where $\mathbf{I}$ is the $N \times N$ identity matrix. Note that $\mathbf{R}$ is symmetric and positive definite. This guarantees the mean convergence of the MMNLMS algorithm with a proper choice of the step size, irrespective of the type of the probability density function of the input signal (unlike the situation in [2]). To find a limit on the step size of the MMNLMS, we consider the mean square analysis of the algorithm. We assume that $N \geq 2$. Let max be the index of the coefficient to be updated at time instant n, i.e., $w_{max}(n)$ is the coefficient to be updated. Then from Eq.(9), it can be shown that the difference equation of the mean of the $max - th$ coefficient for zero mean i.i.d input signal is

$$E\{V_{max}^2\}(n+1) = (1 - 2\acute{\mu}\sigma_x^2 + \acute{\mu}^2\eta\ )E\{V_{max}^2\}(n) + \acute{\mu}^2\sigma_x^4 \sum_{j=1,j\neq max}^{N} E\{V_j^2(n)\} + \acute{\mu}^2\sigma_x^2\epsilon_{min} \tag{10}$$

where $\eta = E\{x^4(n)\}$, $\epsilon_{min} = E\{e^{*2}(n)\}$, and $\acute{\mu} = \frac{\mu}{N\sigma_x^2}$. For a zero mean independent Gaussian input signal $\eta = 3\sigma_x^4$. In [2], it is shown that with the assumptions used here, the sequence of indices of updated coefficients is a Markov process with a uniform probability of selecting any coefficient for updating. Accordingly, we have $E\{V_{max}^2(n)\} = E\{V_j^2(n)\} = C(n),\ \ \forall j = 1,2,...,N$. Thus, Eq.(10) becomes

$$\begin{aligned} E\{V_{max}^2\}(n+1) &= (1 - 2\acute{\mu}\sigma_x^2 + \acute{\mu}^2[\eta + (N-1). \\ &\quad \sigma_x^4\ \ ])C(n) + \acute{\mu}^2\sigma_x^2\epsilon_{min} \end{aligned} \tag{11}$$

The probability of updating any coefficient at each sample time is $\frac{1}{N}$, therefore for $j \neq max$

$$C(n+1) = \frac{1}{N}(\,(N-1)E\{V_j^2(n+1)\}+E\{V_{max}^2(n+1)\}) \tag{12}$$

Note that for $\forall j \neq max$, $E\{V_j^2(n+1)\} = E\{V_j^2(n)\} = C(n)$, the substituting Eq.(11) in Eq.(12) results in

$$\begin{aligned} C(n+1) &= (1 - 2\frac{\acute{\mu}}{N}\sigma_x^2 + \frac{\acute{\mu}^2}{N}\left[\eta + (N-1)\sigma_x^4\right])C(n) \\ &\quad + \frac{\acute{\mu}^2}{N}\sigma_x^2\epsilon_{min} \end{aligned} \tag{13}$$

To ensure the convergence of Eq.(13), and noting that $\acute{\mu} = \frac{\mu}{N\sigma_x^2}$, the step size μ of the MMNLMS algorithm should be bounded by

$$0 < \mu < \frac{2N\sigma_x^4}{\eta + (N-1)\sigma_x^4} \tag{14}$$

For a zero mean independent Gaussian input signal we have $0 < \mu < 2\frac{N}{N+2}$ The excess mean square error (MSE) $\epsilon_{ex}(n)$ of the adaptive filter when the input is an i.i.d. signal is given by

$$\epsilon_{ex}(n) = \sigma_x^2 \sum_{j=1}^{N} E\{V_j^2(n)\} \tag{15}$$

Assuming that the step size of the MMNLMS algorithm is chosen such that MSE convergence is guaranteed, and using Eq.(13), then the steady state excess MSE $\epsilon_{ex}(\infty)$ of the MMNLMS algorithm has the form

$$\begin{aligned} \epsilon_{ex,M=1}(\infty) &= N\sigma_x^2 C(\infty) \\ &= \frac{\mu\epsilon_{min}}{2 - \frac{\mu}{N\sigma_x^4}(\eta + (N-1)\sigma_x^4)} \end{aligned} \tag{16}$$

For a zero mean independent Gaussian input signal, the excess MSE of the MMNLMS algorithm (M=1) upon convergence is

$$\epsilon_{ex,M=1}(\infty) = \frac{\mu\epsilon_{min}}{2 - \mu\frac{N+2}{N}} \tag{17}$$

In the case when all coefficients are updated at each iteration (M=N), i.e., full update NLMS algorithm, it can be shown that the steady state excess MSE is given by

$$\epsilon_{ex,M=N}(\infty) = \frac{\mu\epsilon_{min}}{2 - \mu\frac{N+2}{N}} \tag{18}$$

Eq.(17) and Eq.(18) show that the case of $M = 1$ and the full update NLMS provide similar misadjustment when applied under same conditions and used with the same step size value. Given that the algorithms are using FIR filters of the same order, it automatically implies they have both reached the same solution. However, and as expected, the full update NLMS algorithm converges faster than when only one coefficient is updated per sample time.

4 Simulations

In this section, we compare the performance of the MMNLMS algorithm for $M = 1$ with the Max-NLMS algorithm in [2]. Both algorithms will pick identical coefficients to update (coefficient with maximum $x(n-i+1)$). However, the update term is slightly different (compare Eq.(2) for the Max NLMS and Eq.(1) for the MMNLMS).

Example 1:

It is shown in [2] that the Max NLMS algorithm diverges for input signals with non-symmetric distribution for any step size. The input signal used in [2] is a continuous approximation to a skewed binary distribution with a p.d.f given by

$$p(x)x = \begin{cases} \frac{1}{\delta}(0.5-\alpha) & ,if\ 1-\beta < x < 1-\beta+\delta \\ \frac{1}{\delta}(0.5+\alpha) & ,if\ -1-\beta < x < -1-\beta+\delta \\ 0 & ,\ otherwise \end{cases} \tag{19}$$

where $0 < \delta \ll \beta$. The unknown system is an FIR system with 10 coefficients and the FIR adaptive filter has $N = 10$. A zero mean white Gaussian noise of 0.01 variance is added to the desired signal. The MMNLMS algorithm (M=1) and the Max-NLMS are used with $\mu = 0.4$ and $\mu_{Max-NLMS} = 0.24$, respectively. The input signal parameters are $\alpha = 0.02$, $\beta = 0.02$, and $\delta = 0.001$ [2]. Results are obtained by averaging over 100 independent runs. As expected from the analysis, Fig.1 shows that while the Max-NLMS algorithm diverges with such input signal, the MMNLMS algorithm maintains stability and achieves convergence even when using a larger step size $\mu = 0.4$ compared to $\mu_{Max-NLMS} = 0.24$. When $\alpha = 0$ and $\beta = 0$, both the Max-NLMS and the M-Max algorithm converge without stability problems.

Example 2:

Here, we consider a white Gaussian input signal with zero-mean and unity variance. Both the adaptive filter and unknown system have length 50. The largest possible step size for the Max-NLMS before divergence is $\mu_{Max-NLMS} = 0.2$. The MMNLMS algorithm, with $M = 1$, is used with $\mu = 0.8$ to obtain the same level of steady state MSE of the Max-NLMS algorithm . The variance of the added noise is 0.0001. Fig.2 shows that the MMNLMS algorithm is twice faster than the Max-NLMs algorithm for this example. The same experiment is repeated for $N = 5$, and results are shown in Fig.3. The step sizes used are $\mu_{Max-NLMS} = 0.2$, and $\mu = 0.56$. The improvement attained by the MMNLMS algorithm tends to be marginal for low filter orders. This is expected since the update equation of the MMNLMS algorithm for $M = 1$ approximates generally that of the Max-NLMS algorithm for sufficiently low N. This can be explained by the fact that for sufficiently small N, the difference between $\mathbf{X}^{\mathbf{T}}(n)\mathbf{X}(n)$ and $x^2(n - max + 1)$ is not significant.

5 Conclusion

In this paper, we presented MSE analysis of the MMNLMS algorithm which belongs to the family of algorithms that update only a subset of the adaptive filter coefficients at each iteration. It was shown that MMNLMS results in the closest performance to full update NLMS for a given number of coefficient updates. Compared to the Max-NLMS, the MMNLMS remains stable for skewed probability density functions and provides better convergence in general. The MMNLMS also allows the flexibility of choosing M larger than one.

References

[1] T. Aboulnasr and K. Mayyas, "Selective Coefficient Update of Gradient-Based Adaptive Algorithms", *Proc. ICASSP'96*, pp. 1929-1932, 1996.

[2] S. C. Douglas, " Analysis and Implementation of the Max-NLMS Adaptive Filter," *in Proc. 29th Asilomr Conf. Signals and Systs. Comput.*, Vol. 1, Pacific Grove, CA, pp. 659-663, Oct. 1995.

[3] S. C. Douglas, "Adaptive Filters Employing Partial Updates," *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, pp. 209-216, March 1997.

[4] W. S. Mikhael, et. al, "Adaptive Filters with Individual Adaptation of Parameters," *IEEE Trans. Circuits Systems*, vol. CAS-33, pp. 677-685, July 1986.

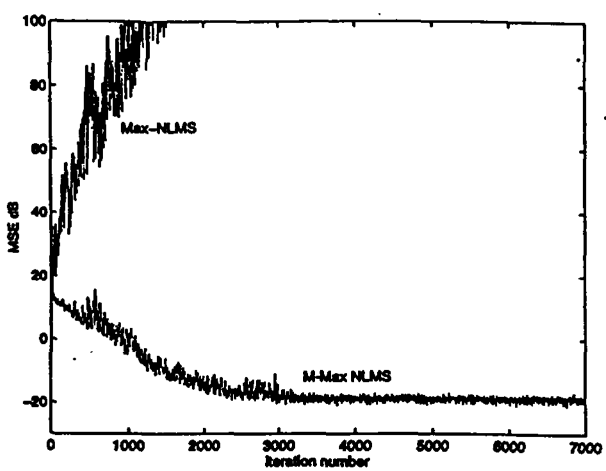

Fig.1 Comparison of MSE between the Max-NLMS, and MMNLMS algorithm (M=1) with $N = 10$ and a skewed binary distribution input .

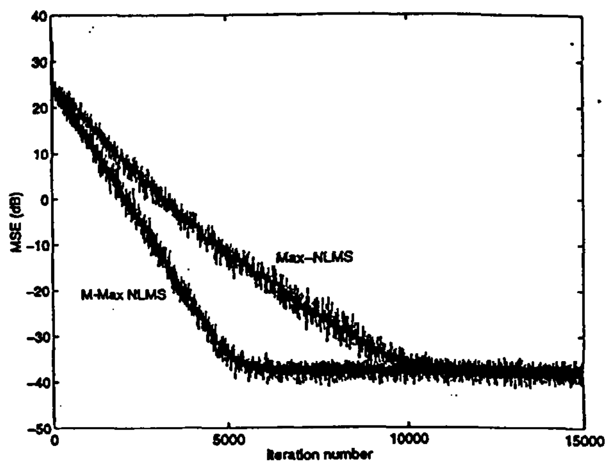

Fig.2 Comparison of MSE between the Max-NLMS, and MMNLMS algorithm (M=1) with $N = 50$ and a white input.

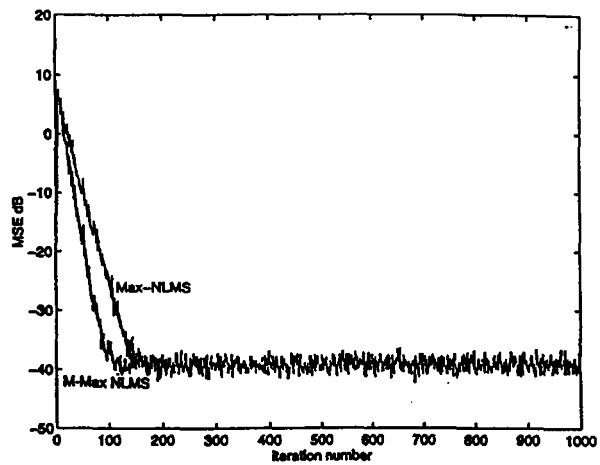

Fig.3 Comparison of MSE between the Max-NLMS, and MMNLMS algorithm (M=1) with $N = 5$ and a white input.