A Robust Variable Step Size LMS-Type Algorithm: Analysis and Simulations

K. Mayyas and T. Aboulnasr
Department of Electrical Engineering
University of Ottawa, Ottawa, Ontario K1N 6N5, Canada

Abstract

This paper presents a robust variable step size LMS-type algorithm with the attractive property of achieving a small final misadjustment while providing fast convergence at early stages of adaptation. The performance of the algorithm is not affected by presence of noise. Approximate analysis of convergence and steady state performance for a zero-mean stationary Gaussian inputs and a nonstationary optimal weight vector is provided. Simulation results clearly indicate its superior performance for stationary cases. For nonstationary environment, our algorithm provides performance equivalent to that of the regular LMS algorithm.

1 Introduction

Since its introduction, the LMS algorithm has been the focus of much study. This is largely due to its simplicity and robustness which have made it widely adopted in many applications. However, the inherent limitation of the LMS of not being able to satisfy the opposing fundamental requirements of fast convergence rate and small misadjustment demanded in most adaptive filtering applications, has always directed researchers at means and alternatives to improve and optimize its performance. One common approach is to employ a time-varying step size in the standard LMS weight update recursion, [1], [2], [3]. Our simulation results show that the performance of these existing variable step size (VSS) algorithms, [1-3], is highly sensitive to the noise disturbance [4], [5]. Since measurement noise is a fact in any practical system, the usefulness of any adaptive algorithm is judged by its performance *in the presence* of this noise.

In this paper we will start by discussing the performance of the algorithm in [1] in noisy conditions as an example. We will show that its performance deteriorates in the presence of measurement noise. We then propose a new VSS LMS algorithm where the step size of the algorithm is adjusted according to an error autocorrelation function. As a result, the algorithm can effectively adjust the step size as in [1] while maintaining the immunity against independent noise disturbance. Another significant feature of the new algorithm is that the addition of a new parameter pertaining to the time-averaging operation allows controlling misadjustment and convergence time more independently without the inherent need to compromize between them as in other VSS algorithms.

2 The Algorithm

In [1], a variable step size LMS is proposed where the step size is proportional to the error energy. The weight update recursion of the algorithm is of the form

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu(n)e(n)\mathbf{X}(n) \qquad (1)$$

and the step size update expression is

$$\mu(n+1) = \alpha\mu(n) + \gamma e^2(n) \qquad (2)$$

where $0 < \alpha < 1, \gamma > 0$, and $\mu(n+1)$ is set to μ_{min} or μ_{max} when it falls below or above one of them, respectively. The algorithm has preferable performance over the fixed step size LMS: at early stages of adaptation, the error is large causing the step size to increase to provide faster convergence speed. When the error decreases, the step decreases thus yielding smaller misadjustment. Unfortunately, the usage of the instantaneous squared error as a measure of the closeness to the optimum results in significant degradation in the presence of measurement noise. This can be deduced by examining Eq.(2). The output error of the identification system is

$$e(n) = d(n) - \mathbf{X}^{\mathbf{T}}(n)\mathbf{W}(n) \tag{3}$$

where the desired signal , $d(n)$, is given by

$$d(n) = \mathbf{X}^{\mathbf{T}}(n)\mathbf{W}^*(n) + \xi(n) \tag{4}$$

where $\xi(n)$ is a zero-mean independent disturbance and $\mathbf{W}^*(n)$ is the time-varying optimal weight vector. The input signal autocorrelation matrix, defined as $\mathbf{R} = E\{\mathbf{X}(n)\mathbf{X}^{\mathbf{T}}(n)\}$, can be represented as $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathbf{T}}$ where $\mathbf{\Lambda}$ is the eigenvalues matrix, and $\mathbf{Q}$ is the modal matrix of $\mathbf{R}$. Using $\acute{\mathbf{V}}(n) = \mathbf{Q}^{\mathbf{T}}\mathbf{V}(n)$, and $\acute{\mathbf{X}}(n) = \mathbf{Q}^{\mathbf{T}}\mathbf{X}(n)$ where $\mathbf{V}(n) = \mathbf{W}(n) - \mathbf{W}^*(n)$ is the translated error vector, then the statistical behavior of $\mu(n+1)$ is determined by substituting Eqs.(3) and (4) in (2) and taking the expected average

$$\begin{aligned} E\{\mu(n+1)\} &= \alpha E\{\mu(n)\} \\ &+\gamma(E\{\xi^2(n)\} + E\{\acute{\mathbf{V}}^{\mathbf{T}}(n)\mathbf{\Lambda}\acute{\mathbf{V}}(n)\}\,) \end{aligned} \tag{5}$$

where we have made use of the common independence assumption of $\acute{\mathbf{V}}(n)$ and $\acute{\mathbf{X}}(n)$ [6]. Clearly, the term $E\{\acute{\mathbf{V}}^{\mathbf{T}}(n)\mathbf{\Lambda}\acute{\mathbf{V}}(n)\}$ influences the closeness of adaptive system to the optimal solution, accordingly $\mu(n+1)$ is adjusted. However, due to $E\{\xi^2(n)\}$, the step size update deteriorates at all stages of adaptation but particularly near the optimum. To avoid this sensitivity to noise, a new measure is proposed. The idea is based on the fact that near the optimum, not only is the energy in the error is small but so is the correlation between successive samples. Thus, we use an estimate of the autocorrelation between $e(n)$ and $e(n-1)$ to control step size updating. The estimate is a time-averaged of $e(n)e(n-1)$ described as

$$p(n) = \beta p(n-1) + (1-\beta)e(n)e(n-1) \tag{6}$$

and the step size update equation is

$$\mu(n+1) = \alpha\mu(n) + \gamma p(n)^2 \tag{7}$$

where limits on $\mu(n+1)$, α, and γ are the same as those of the VSS LMS algorithm. The positive constant β $(0 < \beta < 1)$ is an exponentially weighting parameter that governs the averaging time constant. The autocorrelation of $e(n)$ and $e(n-1)$ serves two objectives; firstly, rejecting the independent noise sequence effect on step size, secondly, the autocorrelation is an efficient measure of the closeness to the optimum. At steady-state $p(n) \approx 0$ leading to $\mu \approx 0$, thus achieving smaller misadjustment values. When the adaptation process is active, $p(n)$ is large leading to larger μ and faster convergence speed. This can be seen if we rewrite the step size in Eq.(7) as

$$\mu(n+1) = \alpha\mu(n) + \gamma E^2\{e(n)e(n-1)\} \tag{8}$$

where we have assumed perfect estimation of the autocorrelation of $e(n)$ and $e(n-1)$. Under the assumption of zero-mean independent noise sequence,

$$\mu(n+1) = \alpha\mu(n) + \gamma\left[\sum_{i=2}^{N}\lambda_i E\{\acute{V}_i(n)\acute{V}_{i-1}(n-1)\}\right]^2 \tag{9}$$

Owing to the averaging operation, the instantaneous behavior of the step size will be smoother and insensitive to independent disturbance noise. Moreover, the step size is basically proportional to a term that can efficiently sense the adaptation state *while being independent of the noise term.*

3 Performance Analysis

Performance analysis of the algorithm will be considered when operating in stationary and nonstationary noisy environments. The weight update equation is the same as in Eq.(1) where $\mu(n)$ is attained by Eqs.(6) and (7). The optimal weight vector is time varying generated by a random walk model as

$$\mathbf{W}^*(n) = \mathbf{W}^*(n-1) + \eta(n-1) \tag{10}$$

where $\eta(n)$ is a stationary noise process of zero-mean and correlation matrix $\sigma_n^2\mathbf{I}$ that accounts for the nonstationarity of the physical system. The measurement noise is a result of nonzero $\xi(n)$ in Eq.(4). Substituting Eqs.(3), (4) and (10) in (1) results in

$$\acute{\mathbf{V}}(n+1) = [\mathbf{I} - \mu(n)\acute{\mathbf{X}}(n)\acute{\mathbf{X}}^{\mathbf{T}}(n)]\acute{\mathbf{V}}(n) + \mu(n)\xi(n)\acute{\mathbf{X}}(n) - \eta(n) \quad (11)$$

where $\epsilon_{min} = E\{\xi^2(n)\}$ is the minimum value of the MSE. From Eq.(11), it can been shown that to ensure convergence of the weight vector mean, $0 < E\{\mu(n)\} < \frac{2}{\lambda_{max}}$, where λ_{max} is the maximum eigenvalue of $\mathbf{R}$. To ensure convergence of the mean square error, it can be shown that a sufficient condition is [1]

$$\frac{E\{\mu^2(\infty)\}}{E\{\mu(\infty)\}} \leq \frac{2}{3tr(\mathbf{R})} \quad (12)$$

from which we can show that γ, α, and β have to satisfy

$$0 < \frac{\gamma\epsilon_{min}^2(1-\beta)}{1-\alpha^2} \leq \frac{1}{3tr(\mathbf{R})} \quad (13)$$

The misadjustment is defined as $M = \frac{\epsilon_{ex}}{\epsilon_{min}}$ where $\epsilon_{ex} = E\{\acute{\mathbf{V}}^{\mathbf{T}}(\infty)\mathbf{\Lambda}\acute{\mathbf{V}}(\infty)\}$ is the steady state excess MSE. In the event of small values of misadjustment, the following expression for algorithm misadjustment can be obtained

$$M \approx \frac{y}{2}tr(\mathbf{R}) + \frac{N\sigma_n^2}{2E\{\mu(\infty)\}\epsilon_{min}} \quad (14)$$

where $E\{\mu(\infty)\} \approx \frac{\gamma(1-\beta)}{(1-\alpha)(1+\beta)}\epsilon_{min}^2$ and $y = \frac{2\gamma\alpha\epsilon_{min}^2(1-\beta)}{(1-\alpha^2)(1+\beta)}$. In a stationary environment $\sigma_n^2 = 0$ and the misadjustment is reduced to $M \approx \frac{y}{2}tr(\mathbf{R})$. Practically, α, γ, and β are selected to produce the same MSE attained by the fixed step size LMS (FSS). In stationary applications, the exponential weighting parameter β is chosen close to 1, bringing down the misadjustment to smaller values. This allows the usage of a larger γ to obtain the same level of misadjustment while maintaining the stability of the algorithm, Eq.(13), where a larger γ will improve the convergence characteristics of the algorithm. Thus, the utilization of β and γ enable more direct control of both convergence speed and final excess MSE without sacrificing one for the other. From Eq.(14), the choice of β in a nonstationary environment should achieve a compromise between acceptable tracking properties and a low level of excess MSE. Since the first term in Eq.(14) is directly proportional to γ and the second term is inversely proportional to γ, the optimum γ for given α and β is the one making both terms in Eq.(14) equal.

4 Simulations

Here, the proposed correlation-based VSS LMS (MVSS) algorithm is compared with: variable step size LMS (VSS) algorithm [1], the stochastic gradient algorithm with gradient adaptive step size (SGA-GAS) [2], and the fixed step size LMS (FSS) algorithm [6]. Parameters of these algorithms are selected to produce a comparable level of misadjustment. Moreover, our choice of these parameters is also guided by the recommended values in their corresponding publication. In all simulations presented here, the desired signal $d(n)$ is disturbed by zero-mean, uncorrelated Gaussian noise of unity variance. Results are obtained by averaging over 200 independent runs.

In this example, both the system and the adaptive filter are excited by a correlated signal $x(n)$ generated by [1]

$$x(n) = 0.9x(n-1) + a(n) \quad (15)$$

where $a(n)$ is a zero-mean, uncorrelated Gaussian noise of unity variance. This type of signals provides flattened elliptical contours which usually cause problems in the tracking capabilities of gradient algorithms. The system to be modeled is a 4-coefficient FIR filter, and the FIR adaptive filter has a dimension $N = 4$. The MVSS algorithm is used with $\alpha = 0.97$, and $\beta = 0.99$. To obtain a final steady state excess MSE of about -28 dB, we used Eq.(14) to determine $\gamma = 8 \times 10^{-4}$. Note that for this example $tr(\mathbf{R}) = 21.0526$. The VSS algorithm is used with $\alpha_{vss} = 0.97$, and to obtain the same level of misadjustment, γ_{vss} is set to 8×10^{-6}. For SGA-GAS , we found experimentally that $\rho = 5 \times 10^{-11}$ provides the desired misadjustment. For all algorithms, $\mu_{max} = 0.008$, and $\mu_{min} = 1 \times 10^{-4}$. The FSS algorithm is used with $\mu_{FSS} = 3 \times 10^{-4}$. Fig.1 shows that for correlated input signals, the MVSS has better convergence than the VSS, SGA-GAS and the FSS algorithms, while providing the same steady state MSE. The reason can be seen from Fig.2: μ remains at μ_{max}

for a longer period providing the fastest convergence speed.
We have also determined that the MVSS is very fast in responding to sudden changes in the system where simulations showed that the convergence rate of MVSS after the change in the system remains the same as at the initial stage.
The previous example was also repeated using a white input signal and a nonstationary optimal weight vector generated according to a random walk model. Simulations illustrated that the proposed MVSS algorithm along with the other algorithms performed as well as the FSS in that case.

5 Conclusion

A new VSS LMS algorithm was introduced. The step size of the algorithm is adjusted according to the square of a time-averaging estimate of the autocorrelation of $e(n)$ and $e(n-1)$. As a result, the algorithm can effectively sense the adaptation process while maintaining the immunity against independent noise. The autocorrelation is estimated recursively requiring only one extra multiplication per iteration. Results show that the proposed algorithm has a significant convergence rate improvement over other VSS algorithms in stationary environment for the same excess MSE.

References

[1] R. H. Kwong, and E. W. Johnston, "A Variable Step Size LMS Algorithm", *IEEE Trans. Signal Processing*, vol. 40, no. 7, pp. 1633-1642 , July 1992.
[2] V. J. Mathews, and Z. Xie, "A Stochastic Gradient Adaptive Filter with Gradient Adaptive Step Size", *IEEE Trans. Signal Processing*, vol. 41, no. 6, pp. 2075-2087 , June 1993.
[3] R. Harris etc. al.,"A Variable Step Size (VSS) Algorithm", *IEEE Trans., Acoust., Speech, Signal Proc.*, vol. ASSP-34, no. 3, pp. 499-510, June 1986.
[4] J. Soo, and K. K. Pang," A Multi Step Size (MSS) Frequency Domain Adaptive Filter", *IEEE Trans. on Signal Processing*, vol. 39, no. 1, pp. 115-121, Jan. 1991.
[5] J. B. Evans, P. Xue, and B. Liu," Analysis and Implementation of Variable Step Size Adaptive Algorithms", *IEEE Trans. on Signal Processing*, vol. 41, no. 8, pp. 2517-2535, Aug. 1993.
[6] L. L. Horowitz and K. D. Senne, " Performance advantage of complex LMS for controlling narrowband adaptive arrays,' *IEEE Trans. Acoust., Speech, and Signal Processing,* vol. ASSP-29, pp. 722-736, June 1981.

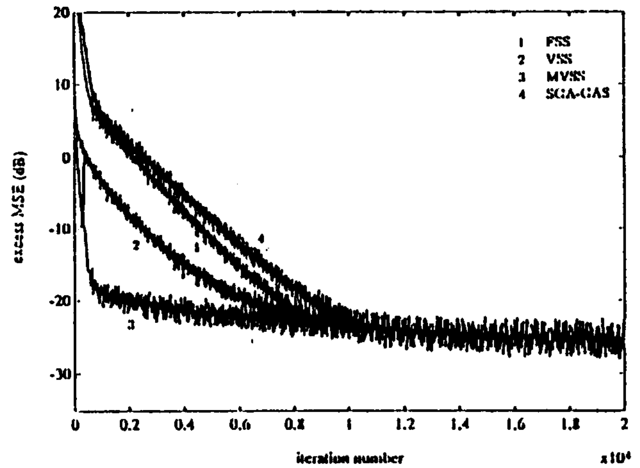

Fig.1 Comparison of excess MSE of various adaptive algorithms.

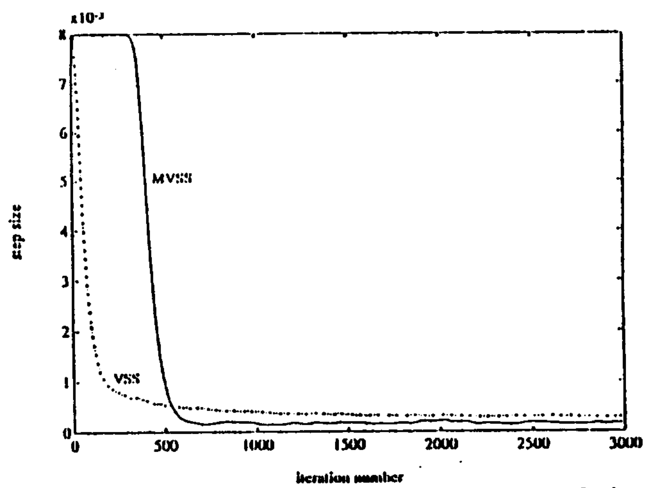

Fig.2 Comparison of step size evolution of the VSS algorithm and the MVSS algorithm.