A FAST ROBUST LMS ALGORITHM UTILIZING THE DYNAMICS OF A DAMPED PENDULUM

T. F. Haddad and M. A. Khasawneh
Department of Electrical Engineering
P.O.Box 3030, Jordan University of Science & Technology
Irbid 221 10 Jordan, Fax: +962 (2) 295 123

ABSTRACT

In this paper, a new fast LMS-based adaptive algorithm is proposed. It is derived by incorporating a damping force in the LMS update recursion in analogy with the force acting upon a damped planar pendulum. An expression for the evolution and the steady-state behaviour for the mean weight vector is developed. This expression provides a mathematical bound which constrains the parameter that controls the maximum contribution of the introduced damping force. Simulation results show an improved robust performance for the new algorithm as compared with the conventional LMS algorithm in smoothly tracking the optimal solution in correlated and nonstationary power environments, especially, in the presence of plant noise.

1. INTRODUCTION

Although the LMS algorithm is relatively simple to implement in practical applications, it is not always satisfactorily capable of tracking the optimal solution on the MSE surface. Therefore, a lot of work in the literature addresses several modified LMS-based techniques with improved convergence characteristics [1 − 4]. The LMS weight update recursion is given by

$$\mathbf{W}_N(n+1) = \mathbf{W}_N(n) + 2\mu e(n)\mathbf{X}_N(n) \quad (1)$$

where $\mathbf{W}_N(n)$ is the $N \times 1$ weight vector at time n, $\mathbf{X}_N(n)$ is the input signal vector, μ is the updating step size, and $e(n) = d(n) - \mathbf{X}_N^T(n)\mathbf{W}_N(n)$ is the prediction error, where the scalar $d(n)$ is the desired signal.

The new gradient adaptive algorithm proposed in this paper, is developed by introducing a damping force to the LMS update recursion analogous to the force governing the motion of a damped planar pendulum. The introduced damping force is controlled by a scaling function which largely reduces the force effect at steady-state. An analytical expression for the evolution of the mean weight vector is presented to demonstrate its convergence to the optimal steepest descent weight vector. Finally, simulation examples are presented to demonstrate the advantage of implementing the new algorithm in achieving a faster rate of convergence with less misadjustment as compared to the LMS algorithm in noisy correlated and nonstationary power environments.

2. ALGORITHM DERIVATION

The analogy between the motion of a planar pendulum and the convergence of the weights towards the global minimum of the MSE surface presented in [1], will be extended to the case which emulates the dynamics of a damped pendulum, in order to derive a robust and faster adaptive algorithm.

Note that a planar pendulum is affected by a damping force of magnitude [5]

$$F = c\left|\frac{d\theta}{dt}\right| \quad (2)$$

where the damping constant c is positive, and θ is the angle between the pendulum and its steady-state position. Let the filter coefficient $w_i(n)$ be associated with the pendulum mass m; an analogy similar to the one adopted in [1]; then the rate of change of the angle θ can be associated with the rate of change of the MSE surface gradient, consequently

$$\frac{d\theta}{dt} \sim \frac{d\hat{\nabla}\epsilon(n)}{dt} \quad (3)$$

where $\sim$ denotes the analogy between the two quantities, and $\hat{\nabla}\epsilon(n) = -2e(n)\mathbf{X}_N(n)$ is the LMS noisy gradient estimate. The gradient rate of change can be in the limit approximated as follows

$$\frac{d\hat{\nabla}\epsilon(n)}{dt} = \frac{\hat{\nabla}\epsilon(n) - \hat{\nabla}\epsilon(n-1)}{T} \quad (4)$$

where T is the sampling period. The proposed modified LMS recursion is of the form

$$\mathbf{W}_N(n+1) = \mathbf{W}_N(n) + 2\mu e(n)\mathbf{X}_N(n) + k\,\frac{d\hat{\nabla}\epsilon(n)}{dt} \quad (5)$$

where k is a positive scaling parameter. As stated in Eq. (2), the contribution of the damping force has to be positive for all time, irrespective of the sign of $\hat{\nabla}\epsilon(n)$. Thus in the derivation for the new algorithm, the negative sign in the definition of the noisy gradient estimate will be disregarded. After substituting for the estimate of the gradient rate of change in Eq. (4), and replacing the gradient estimate $\hat{\nabla}\epsilon(n)$ by its value, Eq. (5) reduces to

$$\mathbf{W}_N(n+1) = \mathbf{W}_N(n)+(2\mu+\eta)e(n)\mathbf{X}_N(n)-\eta e(n-1)\mathbf{X}_N(n-1) \quad (6)$$

Eq. (6) above represents the final form of the recursion for the new algorithm. Note that to make use of the sign fluctuation in $\hat{\nabla}\epsilon(n)$, the actual, not the absolute, value of the new term is introduced in the LMS update recursion. The scaling factor η is assumed to be a time varying exponential function of the form $\eta = r\ exp\left(-\frac{1}{e^2(n)}\right)$, where r is a positive parameter. It is clear that this proposed form for $\eta(n)$ will effectively reduce the contribution of the introduced damping force for small values of $e(n)$, while accentuating that contribution for large $e(n)$.

3. STABILITY ANALYSIS

Several approaches can be followed to study the stability of any adaptive algorithm. The motivation, however, is to analyze our new algorithm by examining the behaviour of the mean weight vector. As analysis unfolds, it will be shown that the parameter r must satisfy certain conditions in order to achieve a stable convergence for the mean weight vector towards its optimal value.

In order to simplify the analysis, we will assume that both $\mathbf{X}_N(n)$ and $d(n)$ are stationary and uncorrelated with both $\mathbf{W}_N(n)$ and $\eta(n)$. This assumption holds whenever the mean values of $\mathbf{X}_N(n)$ and $d(n)$ are changing more rapidly than the mean values of $\mathbf{W}_N(n)$ and $\eta(n)$, which is usually the case at steady-state. In the sequel, we are going to make use of the following definitions

$$\left.\begin{array}{ll} \mathbf{R}_{NN} = E\{\mathbf{X}_N(n)\mathbf{X}_N^T(n)\}, & \bar{\eta}(n) = E\{\eta(n)\} \\ \mathbf{P}_N = E\{d(n)\mathbf{X}_N(n)\}, & \bar{\mathbf{W}}_N(n) = E\{\mathbf{W}_N(n)\} \end{array}\right\} \quad (7)$$

Using Eqs. (6) and (7) as well as the assumption stated above, the expected value of both sides of Eq. (6) follows immediately as

$$\begin{aligned} \bar{\mathbf{W}}_N(n+1) &= \bar{\mathbf{W}}_N(n) + (2\mu+\bar{\eta}(n))\mathbf{P}_N \\ &- (2\mu+\bar{\eta}(n))\mathbf{R}_{NN}\bar{\mathbf{W}}_N(n) \\ &- \bar{\eta}(n)\mathbf{P}_N + \bar{\eta}(n)\mathbf{R}_{NN}\bar{\mathbf{W}}_N(n-1) \quad (8) \end{aligned}$$

which can be written as

$$\begin{aligned} \bar{\mathbf{W}}_N(n+1) &= [\mathbf{I}_{NN} - (2\mu+\bar{\eta}(n))\mathbf{R}_{NN}]\bar{\mathbf{W}}_N(n) + 2\mu\mathbf{P}_N \\ &+ \bar{\eta}(n)\mathbf{R}_{NN}\bar{\mathbf{W}}_N(n-1) \quad (9) \end{aligned}$$

The input autocorrelation matrix $\mathbf{R}_{NN}$ can be decomposed into $\mathbf{R}_{NN} = \mathbf{Q}_{NN}\mathbf{\Lambda}_{NN}\mathbf{Q}_{NN}^T$, [6], where $\mathbf{\Lambda}_{NN}$ is the matrix of eigenvalues of $\mathbf{R}_{NN}$, and $\mathbf{Q}_{NN}$ is the modal matrix of eigenvectors of $\mathbf{R}_{NN}$. Then, substituting the decomposed form of $\mathbf{R}_{NN}$ in Eq. (9), and pre-multiplying Eq. (9) by $\mathbf{Q}_{NN}^T$, yield

$$\begin{aligned} \acute{\mathbf{W}}_N(n+1) &= [\mathbf{I}_{NN} - (2\mu+\bar{\eta}(n))\mathbf{\Lambda}_{NN}]\acute{\mathbf{W}}_N(n) \\ &+ 2\mu\acute{\mathbf{P}}_N + \bar{\eta}(n)\mathbf{\Lambda}_{NN}\acute{\mathbf{W}}_N(n-1) \quad (10) \end{aligned}$$

where

$$\acute{\mathbf{W}}_N(n) = \mathbf{Q}_{NN}^T\bar{\mathbf{W}}_N(n), \quad \acute{\mathbf{P}}_N = \mathbf{Q}_{NN}^T\mathbf{P}_N \quad (11)$$

Furthermore, since $\eta(n)$ is an exponentially damped function, we can assume that it has a time-invariant mean value. Then the N decoupled scalar equations are

$$\begin{aligned} \acute{\bar{w}}_i(n+1) &= [1-(2\mu+\bar{\eta})\lambda_i]\acute{\bar{w}}_i(n) + 2\mu\acute{p}_i \\ &+ \bar{\eta}\lambda_i\acute{\bar{w}}_i(n-1), \quad 0 \le i \le N-1 \quad (12) \end{aligned}$$

At this point, the z-transform is used to solve for $\acute{\bar{w}}_i(n)$. Let $\gamma_i = 1-(2\mu+\bar{\eta})\lambda_i$ and $\alpha = 2\mu$, then the corresponding z-transform of Eq. (12) is

$$\begin{aligned} \acute{\bar{W}}_i(z) &= \frac{(z^2-\gamma_i z)\acute{\bar{w}}_i(0)}{z^2-\gamma_i z-\bar{\eta}\lambda_i} + \frac{z\acute{\bar{w}}_i(1)}{z^2-\gamma_i z-\bar{\eta}\lambda_i} \\ &+ \frac{\alpha\acute{p}_i z}{(z^2-\gamma_i z-\bar{\eta}\lambda_i)(z-1)} \quad (13) \end{aligned}$$

Then, by using partial fraction expansion, we obtain the following solution for $\acute{\bar{w}}_i(n)$

$$\begin{aligned} \acute{\bar{w}}_i(n) &= \acute{\bar{w}}_i(0)[A_1(r_1)^n + B_1(r_2)^n] + \acute{\bar{w}}_i(1)[A_2(r_1)^n \\ &+ B_2(r_2)^n] + \alpha\acute{p}_i[A_3 + B_3(r_1)^n + C_3(r_2)^n], \\ &0 \le i \le N-1 \quad (14) \end{aligned}$$

where

$$r_{1,2} = \frac{\gamma_i}{2} \mp \frac{1}{2}(\gamma_i^2 + 4\bar{\eta}\lambda_i)^{\frac{1}{2}} \quad (15)$$

and the constants appearing in Eq. (14) are defined as follows; $A_1 = \frac{r_1-\gamma_i}{r_1-r_2}$, $B_1 = \frac{r_2-\gamma_i}{r_2-r_1}$, $A_2 = \frac{1}{r_1-r2}$, $B_2 = \frac{1}{r_2-r1}$, $A_3 = \frac{1}{(1-r_1)(1-r_2)}$, $B_3 = \frac{1}{(r_1-1)(r_1-r_2)}$, $C_3 = \frac{1}{(r_2-1)(r_2-r_1)}$.

Note that, Eq. (14) guarantees a monotonic convergence for the mean weight vector provided that $r_{1,2}$ are real and less than unity in magnitude. To find the steady-state value of $\acute{\bar{w}}_i(n)$, note that due to its exponentially damped behaviour, the steady-state mean value of $\eta(n)$ is zero. Therefore, it can easily be shown that

$$\lim_{n\to\infty} \acute{\bar{w}}_i(n) = \frac{\bar{p}_i}{\lambda_i} \quad (16)$$

Thus, regardless of the filter weights initial conditions, Eq. (14) leads to convergence to the optimal steepest descent weight vector as per Eq. (16), [6].

Under the assumption of stationarity in the input signal, the input autocorrelation matrix is symmetric and positive semi-definite; therefore, having N real non-negative eigenvalues as stated in [7]. Accordingly, it is guaranteed that both $r_{1,2}$ are real, thereby, yielding non-oscillatory weight trajectories. The unity bound condition for stability that restricts $|r_{1,2}|$ can be easily simplified to get

$$0 < (\gamma_i^2 + 4\bar{\eta}\lambda_i)^{\frac{1}{2}} < 2 - \gamma_i \quad (17)$$

Note that the maximum value of γ_i is equal to 1; thus the condition of Eq. (17) can be written as

$$0 < (\gamma_i^2 + 4\bar{\eta}\lambda_i)^{\frac{1}{2}} < 1 \quad (18)$$

Since the lower bound is always satisfied, we will elaborate on solving for the upper bound of Eq. (18). Therefore, as a function of γ_i, Eq. (18) reduces to

$$-(1 - 4\bar{\eta}\lambda_i)^{\frac{1}{2}} < 1 - (2\mu + \bar{\eta})\lambda_i < (1 - 4\bar{\eta}\lambda_i)^{\frac{1}{2}} \quad (19)$$

It is obvious that if $\bar{\eta}$ is set to its steady-state zero value, Eq. (19) reduces to the well-known bound derived in [6] for the LMS step size; $0 < \mu < \frac{1}{\lambda_{max}}$. In order to always have a real bound of convergence, we substitute for the maximum value assumed by the constant $\bar{\eta}$ in Eq. (19), which is the positive constant r, thus yielding

$$-(1 - 4r\lambda_i)^{\frac{1}{2}} < 1 - (2\mu + r)\lambda_i < (1 - 4r\lambda_i)^{\frac{1}{2}} \quad (20)$$

It follows that for $1 - 4r\lambda_i > 0$, r must satisfy

$$0 < r < \frac{1}{4\lambda_i} \quad (21)$$

Now, as far as Eq. (21) and the well-known bound for the LMS convergence are satisfied, we can always guarantee that $1 - (2\mu + r)\lambda_i < (1 - 4r\lambda_i)^{\frac{1}{2}}$. However, the left hand inequality of Eq. (20) will not be satisfied unless the positive parameter r is constrained with an upper value computed according to

$$-(1 - 4r\lambda_i)^{\frac{1}{2}} = 1 - (2\mu + r)\lambda_i \quad (22)$$

Moreover, the upper bound of r is determined by solving Eq. (22), whenever the zero crossing of $1-(2\mu+r)\lambda_i$ is less than the maximum value specified in Eq. (21); that is

$$\frac{1}{\lambda_i} - 2\mu < \frac{1}{4\lambda_i} \quad \text{or} \quad \mu > \frac{3}{8\lambda_i} \quad (23)$$

Note that, the value of r obtained by solving Eq. (22) will always satisfy Eq. (21) for values up to $\mu = \frac{1}{\lambda_{max}}$. Thus a theoretical upper bound on r is provided by Eq. (22) for $\frac{3}{8\lambda_{max}} < \mu < \frac{1}{\lambda_{max}}$; otherwise, the bound of Eq. (21) applies.

4. SIMULATION RESULTS

In this section simulation examples are presented to demonstrate the improved convergence rate with the smoothing property provided by the new proposed algorithm. A comparison with the conventional LMS algorithm is performed in correlated and nonstationary power noisy environments.

Case 1: In this case study, the improved performance of the new algorithm will be demonstrated by identifying a 5-tap FIR system in a noisy environment with a varying-power excitation input. The system noise is a zero-mean white Gaussian sequence of 0.01 variance, where the excitation signal is zero-mean Gaussian with variance varying from 0.25 to 1 and finally to 4 at iterations 1000 and 2000, respectively. Such environment is commonly encountered in case of installing adaptive systems in teleconferencing rooms [8]. Fig. 1(a) represents the results of implementing both the new and the LMS algorithm for the same value of $\mu = 0.001$ and $r = 1$, where in, the new algorithm exhibits a largely increased speed of convergence. Since the new algorithm

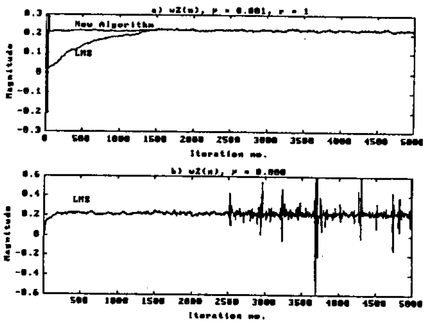

Figure 1: Trajectories for $w_2(n)$.

reduces to the LMS at steady-state, for the same of μ, both will result in an identical weight misadjustment, as shown in Fig. 1(a). On the other hand, increasing the speed of convergence of the LMS algorithm, by setting $\mu = 0.008$, causes large jitters in the trajectories of the system parameters, as shown in Fig. 1(b). Therefore, the proposed algorithm exhibits robustness against variations in the input excitation power with relatively faster rate of convergence as compared with the LMS algorithm.

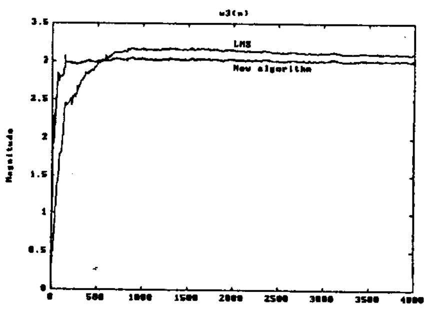

Figure 2: $w_3(n)$ for the same μ with correlated input.

Case 2: In several practical applications, adaptive filters are implemented to perform in correlated and noisy environments. Such environments usually impose limitations on the performance of adaptive algorithms. These limitations are attributed to the sensitivity of adaptive algorithms to the disparity in the eigenvalues of a given input autocorrelation matrix, which is usually the case in correlated environments.

In this simulation example, the behaviour of the new algorithm is examined in identifying the following 5-tap FIR system $h = [-0.5\ 1\ 3\ 2.5\ -2.5]$ with a correlated excitation signal. The correlated signal was obtained by passing a zero-mean white Gaussian noise of unity variance in the all-pole system, [2], $y(n) = 0.44x(n) + 1.5y(n-1) - y(n-2) + 0.25y(n-3)$. The system identifier is corrupted with a 0.0225 variance zero-mean white Gaussian measurement noise, which was generated independent, in the statistical sense, of the system input signal. Fig. 2 depicts the third weight trajectory for the adaptive filter over one realization for $\mu_{LMS} = \mu_{New} = 0.005$ with $r = 0.22$. In this figure we note the superiority of the new algorithm over the LMS in its ability to more rapidly track its pre-assumed optimal value. Increasing the speed of convergence for the LMS algorithm was experimentally achieved by setting $\mu_{LMS} = 0.04$. As shown in Fig. 3, increasing μ_{LMS} causes an increased level of misadjustment in the weight trajectory. The stability of the new algorithm in this critical environment is verified by evaluating the mean square error over 50 realizations, which is shown in Fig. 4 compared to the MSE of the LMS for the same value of $\mu = 0.005$. The LMS mean square error for $\mu = 0.04$ is shown in Fig. 5, which suffers from significant jitters in the entire time span.

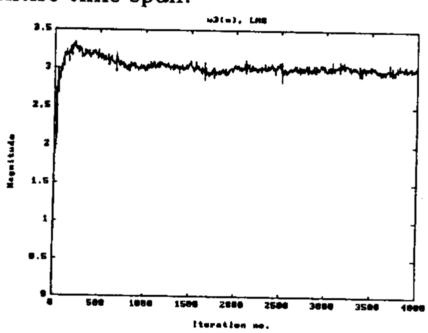

Figure 3: $w_3(n)$, for $\mu_{LMS} = 0.04$ with correlated input.

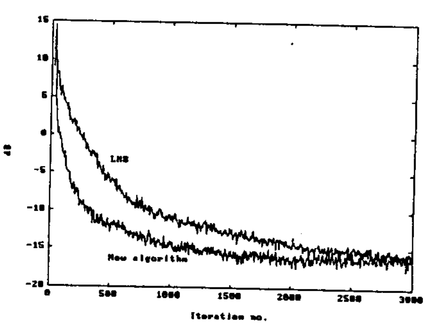

Figure 4: Mean Square Error for the same μ.

5. CONCLUSIONS

In this paper, a new gradient-based adaptive algorithm was derived by embedding a force, analogous to that controling the dynamics of a damped planar pendulum, within the conventional LMS recursion. The proposed algorithm was shown to exhibit a faster rate of convergence with a smaller adaptation step size as compared with the conventional LMS algorithm. This new feature enhances the robustness and the smoothing property in tracking the filter optimal state. A bound on the constant that controls the contribution of the damping force was derived by means of stability analysis assuming typical environmental restrictions, as well as a stable mean weight vector. However, the performance of the new algorithm can be optimized by the proper choice of this constant according to the operating environment. The improved behaviour for the new algorithm was demonstrated by means of simulations under correlated and nonstationary power environments, corrupted with measurement noise.

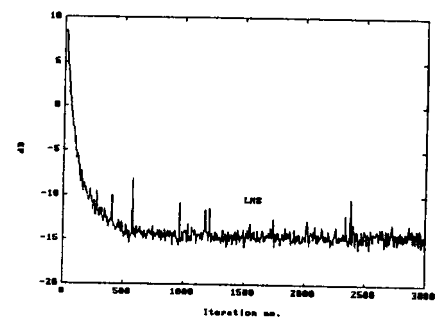

Figure 5: Mean Square Error for $\mu_{LMS} = 0.04$.

References

[1] M. A. Khasawneh, W. E. Alexander, "System Identification Using the Fast LMS-Sin Algorithm," *Proc. IEEE ISCAS'89,* pp. 1736 - 1739, Portland, Oregon.

[2] V. J. Matheus, Zhenhua Xie, "A Stochastic Gradient Adaptive Filter with Gradient Adaptive Step Size," *IEEE Trans. on Signal Processing,* vol. 41, No. 6, June 1993.

[3] C. P. Kwong, "Robust Design of the LMS Algorithm," *IEEE Trans. on Signal Processing,* vol. 40, no.10, oct. 1992.

[4] W. Y. Chen, R. A. Haddad, "A variable Step Size LMS Algorithm ," *Midwest Symposium on Circuits and Systems,* vol.1, Publ. by IEEE, pp. 423 - 426, 1991.

[5] W. E. Boyce, R. C. Diprima, *Elementary Differential Equations and Boundary value problems,* New York: Wiley, 1986.

[6] S. T. Alexander, *Adaptive Signal Processing: theory and applications,* New York: Springer Verlag, 1986.

[7] Harry L. Van Trees, *Detection, Estimation, and Modulation Theory,* vol. 1, New York: Wiley, 1968.

[8] Jia-Sien Soo and K. K. Pang, "A Multistep Size (MSS) Frequency Domain Adaptive Filter, " *IEEE Trans. Signal Processing,* vol. 39, No. 1, pp. 115-121, Jan. 1991.