STABILITY BOUNDS ON STEP-SIZE FOR THE PARTIAL UPDATE LMS ALGORITHM

Mahesh Godavarti and Alfred O. Hero III
Department of Electrical Engineering & Computer Science
The University of Michigan
Ann Arbor, MI 48105
mgodavar@eecs.umich.edu, hero@eecs.umich.edu

ABSTRACT

Partial updating of LMS filter coefficients is an effective method for reducing the computational load and the power consumption in adaptive filter implementations. Only in the recent past has any work been done on deriving conditions for filter stability, convergence rate, and steady state error for the Partial Update LMS algorithm. In [5] approximate bounds were derived on the step size parameter μ which ensure stability in-the-mean of the alternating even/odd index coefficient updating strategy. Unfortunately, due to the restrictiveness of the assumptions, these bounds are unreliable when fast convergence (large μ) is desired. In this paper, tighter bounds on μ are derived which guarantee convergence in-the-mean of the coefficient sequence for the case of wide sense stationary signals.

1. INTRODUCTION

Partial updating of the LMS adaptive filter has been proposed to reduce computational costs [2, 3, 4]. In this era of mobile computing and communications, such implementations are also attractive for reducing power consumption. However, theoretical performance predictions on convergence rate and steady state tracking error are more difficult to derive than for standard full update LMS. Accurate theoretical predictions are important as it has been observed that the standard LMS conditions on the step size parameter fail to ensure convergence of the partial update algorithm.

The two types of partial update LMS algorithms that are prevalent in the literature have been described in [5]. They are referred to as the "Periodic LMS algorithm" and the "Sequential LMS algorithm". An attempt was made to recalculate the bounds on the step-size parameter for both mean and mean-square convergence. Due to simplifying assumptions, the bounds derived turned out to be the same as those for the standard LMS algorithm. It was shown, however, that these bounds fail to predict situations where the Sequential LMS algorithm is unstable when implemented with these standard step-size constraints.

In this paper we derive bounds on the step-size parameter which ensures convergence in mean for the special case involving alternate even and odd coefficient updates. The bounds are based on extremal properties of the matrix 2-norm and the Bauer-Fike theorem. For simplicity we assume all sequences to be real and stationary and we make the standard independence assumptions used in the analysis of LMS [1]. It is shown that as the input signal becomes more correlated the bounds become much tighter than the bound approximation derived in [5].

The organization of the paper is as follows. First in Section 2, a motivating example is shown which illustrates the need for this work. This is followed by a brief description and analysis of the partial update algorithm in Section 3. In Section 4 verification of the theoretical analysis is carried out via simulations. Finally conclusions and directions for future work are indicated in Section 5.

2. MOTIVATING EXAMPLE

Consider a 2-tap adaptive filter with alternating update of the first and second coefficients $w_{1,k}$ and $w_{2,k}$. For odd k, the updates are given by

$$\begin{bmatrix} w_{2,k+1} \\ w_{1,k+1} \end{bmatrix} = \begin{bmatrix} w_{2,k} \\ w_{1,k} \end{bmatrix} + \begin{bmatrix} \mu e_k x_{k-1} \\ 0 \end{bmatrix} \tag{1}$$

And for even k, the updates are given by

$$\begin{bmatrix} w_{2,k+1} \\ w_{1,k+1} \end{bmatrix} = \begin{bmatrix} w_{2,k} \\ w_{1,k} \end{bmatrix} + \begin{bmatrix} 0 \\ \mu e_k x_k \end{bmatrix} \tag{2}$$

e_k is the error signal given by $e_k = d_k - W_k^T X_k$ where $W_k = [w_{1,k}\ w_{2,k}]$ and $X_k = [x_k\ x_{k-1}]$. d_k is the desired response.

Now make the standard assumption [1] that there exists a coefficient vector W_{opt} such that $d_k = W_{opt}^T X_k + n_k$ with $\{n_k\}$ a zero mean i.i.d sequence independent of the input sequence $\{x_k\}$. Then defining $V_k = [v_{1,k}\ v_{2,k}]^T = W_k - W_{opt}$, for odd k we have the following update equation.

$$\begin{aligned} \begin{bmatrix} v_{2,k+2} \\ v_{1,k+2} \end{bmatrix} &= \\ F \begin{bmatrix} v_{2,k} \\ v_{1,k} \end{bmatrix} &+ \begin{bmatrix} \mu n_k x_{k-1} \\ -\mu^2 n_k x_{k+1} x_k x_{k-1} + \mu n_{k+1} x_{k+1} \end{bmatrix} \end{aligned} \tag{3}$$

where the elements of F are

$$\begin{aligned} f_{11} &= 1 - \mu x_{k-1}^2 \\ f_{12} &= -\mu x_{k-1} x_k \\ f_{21} &= -\mu x_{k+1} x_k + \mu^2 x_{k+1} x_k x_{k-1}^2 \\ f_{22} &= 1 - \mu x_{k+1}^2 + \mu^2 (x_{k+1} x_k)(x_{k-1} x_k) \end{aligned}$$

Assuming $\{v_k\}$ and $\{x_k\}$ to be uncorrelated with each other and X_{k+1} to be independent of X_k with $E[x_k^2] = 1$, $E[x_k x_{k-1}] =$

This work was supported in part by a Graduate Fellowship to the first author from the Department of Electrical Engineering and Computer Science at the University of Michigan and by the Department of Defense Research & Engineering (DDR&E) Multidisciplinary University Research Initiative (MURI) on "Low Energy Electronics Design for Mobile Platforms" managed by the Army Research Office (ARO) under ARO grant DAAH04-96-1-0337. Corresponding author: Mahesh Godavarti, mgodavar@eecs.umich.edu, (313)-763-0564 (voice), (313)-763-8041 (fax)

ρ and $E[x_k x_{k-2}] = \rho^2$ and taking expectations of the update equations we obtain

$$E\begin{bmatrix} v_{2,k+2} \\ v_{1,k+2} \end{bmatrix} = \begin{bmatrix} 1-\mu & -\mu\rho \\ \mu\rho(\mu-1) & 1-\mu+\mu^2\rho^2 \end{bmatrix} E\begin{bmatrix} v_{2,k} \\ v_{1,k} \end{bmatrix} \quad (4)$$

It can be easily verified that for $\rho \approx 1$ and $\mu \approx 0$ the necessary and sufficient condition on μ for stability of the recursion (4) is given by

$$0 < \mu < \frac{2(1-\rho)}{\rho^2} \quad (5)$$

whereas, using the update equations for expected values of coefficient error in [5], the condition in [5] for convergence is

$$0 < \mu < \frac{2}{1+\rho}. \quad (6)$$

As $(1-\rho^2)/\rho^2 < 1$ for $\rho^2 > 1/2$, we have that $\frac{2(1-\rho)}{\rho^2} < \frac{2}{1+\rho}$ so that if the upper bound in condition (6) is used to set μ in partial update LMS, divergence occurs.

3. ALGORITHM DESCRIPTION AND ANALYSIS

It is assumed that the filter is a standard FIR filter of even length, N. For convenience, we start with some definitions. Let $\{x_k\}$ be the input sequence and let $\{w_{i,k}\}$ denote the coefficients of the adaptive filter. Define

$$\begin{aligned} W_{e,k} &= [w_{2,k}\ w_{4,k}\ w_{6,k}\ \dots\ w_{N,k}]^T \\ W_{o,k} &= [w_{1,k}\ w_{3,k}\ w_{5,k}\ \dots\ w_{N-1,k}]^T \\ X_{e,k} &= [x_{k-1}\ x_{k-3}\ \dots\ x_{k-N+1}]^T \\ X_{o,k} &= [x_k\ x_{k-2}\ \dots\ x_{k-N+2}]^T \\ W_k &= [w_{1,k}\ w_{2,k}\ \dots\ w_{N,k}]^T \\ X_k &= [x_k\ x_{k-1}\ x_{k-2}\ \dots\ x_{k-N+1}]^T \end{aligned}$$

where the terms defined above are for the instant k. In addition, Let d_k denote the desired response. In typical applications d_k is a known training signal which is transmitted over a noisy channel with unknown FIR transfer function.

In this paper we assume that d_k itself obeys an FIR model given by $d_k = W_{opt}^T X_k + n_k$ where W_{opt} are the coefficients of an FIR model given by $W_{opt} = [w_{1,opt}\ \dots\ w_{N,opt}]^T$. Here $\{n_k\}$ is assumed to be a zero mean i.i.d sequence that is independent of the input sequence $\{x_k\}$. This is a standard assumption used in the analysis of the standard LMS algorithm [1] which can be shown to be reasonable for jointly stationary x_k and d_k.

The coefficient updates for odd k in the partial update LMS algorithm considered here are given by

$$\begin{bmatrix} W_{e,k+1} \\ W_{o,k+1} \end{bmatrix} = \begin{bmatrix} W_{e,k} \\ W_{o,k} \end{bmatrix} + \begin{bmatrix} \mu e_k X_{e,k} \\ 0 \end{bmatrix} \quad (7)$$

and for even k

$$\begin{bmatrix} W_{e,k+1} \\ W_{o,k+1} \end{bmatrix} = \begin{bmatrix} W_{e,k} \\ W_{o,k} \end{bmatrix} + \begin{bmatrix} 0 \\ \mu e_k X_{o,k} \end{bmatrix} \quad (8)$$

where e_k is the error and is defined to be $e_k = d_k - W_k^T X_k$

We also define coefficient error vectors as

$$\begin{aligned} V_{e,k} &= W_{e,k} - W_{e,opt} \\ V_{o,k} &= W_{o,k} - W_{o,opt} \\ V_k &= W_k - W_{opt} \\ V_k^{eo} &= \begin{bmatrix} V_{e,k} \\ V_{o,k} \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} W_{e,opt} &= [w_{2,opt}\ w_{4,opt}\ w_{6,opt}\ \dots\ w_{N,opt}] \\ W_{o,opt} &= [w_{1,opt}\ w_{3,opt}\ w_{5,opt}\ \dots\ w_{N-1,opt}] \end{aligned}$$

Assuming that $\{x_k\}$ is a WSS random sequence, we analyse the convergence of the mean coefficient error vector $E[V_k]$. For regular LMS algorithm the recursion for $E[V_k]$ is given by

$$E[V_{k+1}] = (I - \mu R)E[V_k] \quad (9)$$

where I is the N-dimensional identity matrix and $R = E\left[X_k X_k^T\right]$ is the input signal correlation matrix. The necessary and sufficient condition for stability of the recursion is given by

$$0 < \mu < 2/\lambda_{max} \quad (10)$$

where λ_{max} is the maximum eigen-value of the input signal correlation matrix R.

For odd k, combining the even and odd update equations and writing them in terms of $V_{\cdot,k}$, we obtain

$$\begin{aligned} V_{k+2}^{eo} &= \quad (11) \\ FV_k^{eo} &+ \begin{bmatrix} \mu n_k X_{e,k} \\ -\mu^2 n_k X_{o,k+1} X_{e,k+1}^T X_{e,k} + \mu n_{k+1} X_{o,k+1} \end{bmatrix} \end{aligned}$$

where the elements of F are

$$\begin{aligned} f_{11} &= I - \mu X_{e,k} X_{e,k}^T \\ f_{12} &= -\mu X_{e,k} X_{o,k}^T \\ f_{21} &= -\mu X_{o,k+1} X_{e,k+1}^T + \mu^2 X_{o,k+1} X_{e,k+1}^T X_{e,k} X_{e,k}^T \\ f_{22} &= I - \mu X_{o,k+1} X_{o,k+1}^T + \mu^2 X_{o,k+1} X_{e,k+1}^T X_{e,k} X_{o,k}^T \end{aligned}$$

We next make the standard assumptions that V_k and X_k are mutually uncorrelated and that X_k is independent of X_{k-1} [1]. These assumptions are somewhat restrictive but greatly simplify the analysis. Taking expectations, using the independence assumption on the sequences X_k, n_k, the mutual independence assumption on X_k and V_k, and simplifying we obtain

$$E[V_{k+2}^{eo}] = \left(I - \mu R'\right) E[V_k^{eo}] \quad (12)$$

where

$$I - \mu R' = \begin{bmatrix} I - \mu R_e & -\mu R_{eo} \\ \mu R_{oe}(\mu R_e - I) & I - \mu R_o + \mu^2 R_{oe} R_{eo} \end{bmatrix} \quad (13)$$

and $R_e = E\left[X_{e,k} X_{e,k}^T\right]$, $R_o = E[X_{o,k} X_{o,k}]$, $R_{eo} = E\left[X_{e,k} X_{o,k}^T\right]$, and $R_{oe} = E\left[X_{o,k} X_{e,k}^T\right] = R_{eo}^T$. Under the assumption of even integer N and real w.s.s. $\{x_k\}$ it can be shown that $R_e = R_o$.

For even k, combining the even and odd update equations and writing them in terms of $V_{\cdot,k}$, we obtain

$$\begin{aligned} V_{k+2}^{eo} &= \quad (14) \\ F'V_k^{eo} &+ \begin{bmatrix} -\mu^2 n_k X_{e,k+1} X_{o,k+1}^T X_{o,k} + \mu n_{k+1} X_{e,k+1} \\ \mu n_k X_{o,k} \end{bmatrix} \end{aligned}$$

where the elements of F' are

$$\begin{array}{rcl} f'_{11} & = & I - \mu X_{e,k+1} X_{e,k+1}^T + \mu^2 X_{e,k+1} X_{o,k+1}^T X_{o,k} X_{e,k}^T \\ f'_{12} & = & -\mu X_{e,k+1} X_{o,k+1}^T + \mu^2 X_{e,k+1} X_{o,k+1}^T X_{o,k} X_{o,k}^T \\ f'_{21} & = & -\mu X_{o,k} X_{e,k}^T \\ f'_{22} & = & I - \mu X_{o,k} X_{o,k}^T \end{array}$$

Taking expectations, and using the same assumptions as above, we obtain

$$E\left[V_{k+2}^{eo}\right] = (I - \mu R'')E\left[V_k^{eo}\right] \tag{15}$$

where

$$I - \mu R'' = \begin{bmatrix} I - \mu R_e + \mu^2 R_{eo} R_{oe} & \mu R_{eo}(\mu R_o - I) \\ -\mu R_{oe} & I - \mu R_o \end{bmatrix} \tag{16}$$

It can be shown that under the above assumptions on X_k, V_k and d_k, the convergence conditions for even and odd update equations are identical. We therefore focus on (12). If we want to write the update equations for the regular LMS algorithm in the same form as (12) we would have

$$E\left[V_{k+1}^{eo}\right] = \begin{bmatrix} I - \mu R_e & -\mu R_{eo} \\ -\mu R_{oe} & I - \mu R_o \end{bmatrix} E\left[V_k^{eo}\right] \tag{17}$$

which is the same as (9) only expressed in a different form. It should be noted here that even though

$$R \neq \begin{bmatrix} R_e & R_{eo} \\ R_{oe} & R_o \end{bmatrix}$$

the matrix on the right is the correlation matrix for a permuted form of the input signal and therefore is also an input signal correlation matrix with the same eigenvalues as R.

Now to ensure stability of (12), the eigenvalues of $I - \mu R'$ should lie inside the unit circle. To estimate the eigenvalues of $I - \mu R'$ we employ the Bauer-Fike theorem [6, p. 321] which states that if α' is an eigenvalue of $A + E \in \mathbf{C}^{n \times n}$ and $M^{-1}AM = diag(\alpha_1, \ldots, \alpha_n)$ then

$$\min_{\alpha \in \alpha(A)} |\alpha - \alpha'| \leq \kappa_p(M)\|E\|_p \tag{18}$$

where $\|\cdot\|_p$ denotes any of the p-norms and $\kappa_p(M) = \|M\|_p\|M^{-1}\|_p$. For convenience, we will choose $p = 2$.

Now writing $I - \mu R'$ as $A + E$ where

$$A = \begin{bmatrix} I - \mu R_e & -\mu R_{eo} \\ -\mu R_{oe} & I - \mu R_o \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & 0 \\ \mu^2 R_{oe} R_e & \mu^2 R_{oe} R_{eo} \end{bmatrix}$$

we have

$$\min_{\alpha \in \alpha(A)} |\alpha - \alpha'| \leq \|E\|_2 \tag{19}$$

that is so because $\kappa_2(M) = 1$ on account of A being an Hermitian matrix which admits a matrix of orthogonal eigenvectors M. Now E can be written as $E = BC$ where

$$B = \begin{bmatrix} 0 & 0 \\ \mu^2 R_{oe} & 0 \end{bmatrix} \tag{20}$$

$$C = \begin{bmatrix} R_e & R_{eo} \\ R_{oe} & R_o \end{bmatrix} \tag{21}$$

Using the properties of the matrix $2-$norm [6, pp. 56-57] we obtain

$$\|E\|_2 \leq \mu^2 \frac{N}{2} \max_{i,j} |b_{ij}| \lambda_{max} \leq \mu^2 \frac{N}{2} R(0) \lambda_{max} \tag{22}$$

where λ_{max} is the largest eigenvalue of the matrix C which is the correlation matrix of the permuted input signal. If we let $\beta = \frac{N}{2} R(0)\lambda_{max}$ we have the simple bound

$$\min_{\alpha \in \alpha(A)} |\alpha - \alpha'| \leq \mu^2 \beta \tag{23}$$

For a given α' if we define $\alpha^0 = \arg\min_{\alpha \in \alpha(A)} |\alpha - \alpha'|$ then we have

$$|\alpha^0 - \alpha'| \leq \mu^2 \beta \tag{24}$$

Using the property $|a| - |b| \leq |a - b|$ we have for $j = 1, \ldots, N$

$$|\alpha_j^0| - \mu^2\beta \leq |\alpha_j'| \leq |\alpha_j^0| + \mu^2 \beta \tag{25}$$

Now, invoking the necessary and sufficient condition for stability of (12)

$$|\alpha_j'| < 1 \quad \forall j = 1, \ldots, N. \tag{26}$$

we obtain the sufficient condition

$$|\alpha_j^0| + \mu^2\beta < 1 \quad \forall j = 1, \ldots, N. \tag{27}$$

Since the set of α^0's is a subset of the set of α's and since $\alpha_j = 1 - \mu\lambda_j$, (27) can be ensured by the simpler condition

$$|1 - \mu\lambda_j| + \mu^2\beta < 1 \quad \forall j = 1, \ldots, N. \tag{28}$$

A sufficient condition on μ which satisfies (28) and hence ensures stability is

$$0 < \mu < \min\left(\frac{\lambda_j}{\beta}, \frac{1}{\lambda_j}\right) \quad \forall j = 1, \ldots, N. \tag{29}$$

which leads to

$$0 < \mu < \min\left(\frac{\lambda_{min}}{\beta}, \frac{1}{\lambda_{max}}\right) \tag{30}$$

Recall that $\beta = \frac{N}{2}R(0)\lambda_{max}$, which for w.s.s. $\{x_k\}$ can be rewritten as $\beta = \frac{1}{2}tr(R)\lambda_{max}$, as the trace of R satisfies $tr(R) = NR(0)$. Hence, as $N \geq 2$, $\frac{\lambda_{min}}{\beta} = \frac{\lambda_{min}}{tr(R)/2} \frac{1}{\lambda_{max}} \leq \frac{1}{\lambda_{max}}$ and we have

$$0 < \mu < \frac{\lambda_{min}}{\beta} = \frac{\lambda_{min}}{\lambda_{max}} \frac{2}{tr(R)} \tag{31}$$

which is the sufficient condition ensuring convergence of (12) and is the main result of this paper.

This condition when applied to the motivating example gives us the bound on μ: $0 < \mu < (1 - \rho)/(1 + \rho)$ which satisfies (5). It should be noticed that as the signal becomes more correlated $\lambda_{min} \to 0$ making the bound tighter.

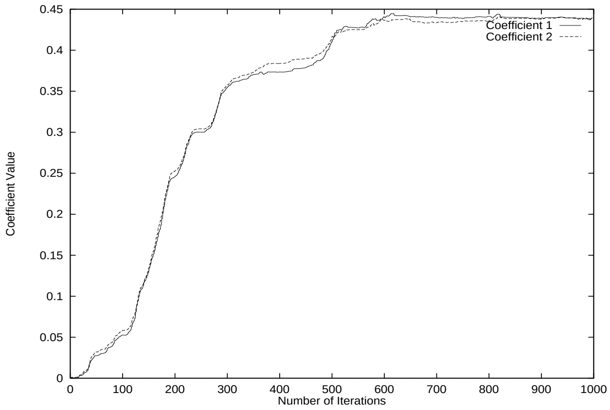

Figure 1: Trajectory of $w_{1,k}$ and $w_{2,k}$ for $\mu = 0.005$

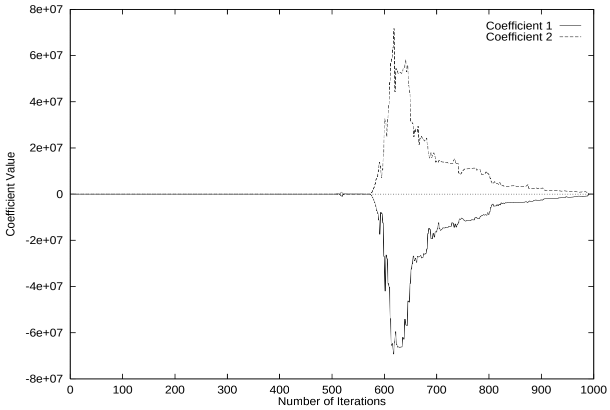

Figure 2: Trajectory of $w_{1,k}$ and $w_{2,k}$ for $\mu = 0.5$

4. SIMULATIONS

We have plotted the evolution trajectory of the 2-tap filter considered in Section 2 for $\rho = 0.99$ and $W_{opt} = [0.4\ 0.5]$ in Figures 1 and 2. For Figure 1 μ was chosen according to condition (31) and for Figure 2 μ was chosen according to (10) which is the condition given in [5] for convergence in-the-mean. For simulation purposes we set $d_k = W_{opt}^T S_k + n_k$ where $S_k = [s_k\ s_{k-1}]^T$ is a vector composed of the w.s.s. AR process $\{s_k\}$ with variance equal to 1 and AR coefficient $\rho = 0.99$, and $\{n_k\}$ is a white sequence, with variance equal to 0.01, independent of $\{s_k\}$. We set $\{x_k\} = \{s_k\} + \{v_k\}$ where $\{v_k\}$ is a white sequence, with variance equal to 0.01, independent of $\{s_k\}$. As can be seen from Figure 2 stricter conditions are needed for convergence in mean than those given by (10).

5. CONCLUSION

We have analyzed the alternating odd/even partial update LMS algorithm and we have derived stability bounds on step-size parameter μ for wide sense stationary signals based on extremal properties of the matrix 2-norm. While these may not be the weakest possible bounds, they do provide the user with a useful sufficient condition on μ which ensures convergence in the mean. The analysis also leads directly to an estimate of mean convergence rate. Mean-square convergence analysis was not undertaken in this paper as the primary motivation was to show that current bounds on step-size are not sufficient to guarantee convergence. Theoretical analysis in the manner considered here for the general case of "Sequential LMS Algorithm" is more complicated but feasible.

6. REFERENCES

[1] S. T. Alexander, *Adaptive signal processing*, Springer-Verlag, New York NY, 1986.

[2] M. J. Gingell, B. G. Hay, and L. D. Humphrey, "A Block Mode Update Echo Canceller Using Custom LSI," GLOBECOM Conference Record, Vol. 3, pp. 1394-1397, Nov. 1983.

[3] M. Godavarti, "Implementation of a G.165 Line Echo Canceller on Texas Instruments' TMS320C3x and TMS320C54x chips," ICSPAT Conference Proceedings, pp. 65-69, Sep. 1997.

[4] D. Messerschmitt, D. Hedberg, C. Cole, A. Haoui and P. Winship, "Digital Voice Echo Canceller with a TMS32020," Application report: SPRA129, Texas Instruments.

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

[6] G. H. Golub and C. F. Van Loan, "Matrix Computations," 3^{rd} edition, The Johns Hopkins University Press, 1996.