

Leaky Least Mean Mixed Norm Algorithm

Mohammed Abdul Nasar

CVR College of Engineering /ECE, Hyderabad, India

Email: nasar.mohammed@gmail.com

Abstract— In this paper, the leakage-based variant of the Least Mean Mixed Norm (LMMN) algorithm, the leaky Least Mean Mixed Norm (LLMMN) algorithm, is derived. The proposed algorithm will help mitigate the weight drift problem experienced in the conventional Least Mean Square (LMS) and Least Mean Fourth (LMF) algorithms. The aim of this paper is to derive the LLMMN adaptive algorithm and perform the transient analysis using the energy conservation relation framework. Finally, simulation results are carried out to support the theoretical findings, and show improved performance obtained through the use of LLMMN over the conventional LMMN algorithm in a weight drift environment.

Index Terms—Adaptive filters, weight drift, leaky least mean mixed norm.

I. INTRODUCTION

In our everyday life the Adaptive filters are used in a variety of areas such as plant modelling or system identification, noise cancelation and adaptive equalization, to name a few. The theory about adaptive filters, advantages and applications are widely described in the literature [1].

The LMF algorithm outperforms the LMS in non-Gaussian noise environments. Another example of adaptive filter algorithms is the leaky Least Mean Square (leaky LMS) algorithm [4]. This leaky LMS algorithm was first introduced to overcome the weight drift problem in LMS adaptive filters. The LMF algorithm also suffers from the weight drift problem under the same conditions as LMS algorithm,. The LMMN algorithm [5] is found to provide a better performance in both Gaussian and Non-Gaussian environments than either LMS or the LMF and hence the LMMN algorithm will behave identically in such a scenario. Therefore, in this paper, a new variant of the LMMN is introduced which overcomes the weight drift instability from occurring using the leakage technique.

In this paper, the leaky LMMN algorithm is derived and its transient analysis is carried out using energy conservation concept [1] and the conditions for the mean and mean square stability of the algorithm are also derived.

II. THE LEAKY LEAST MEAN MIXED NORM ALGORITHM UPDATE RECURSION

The output of an FIR channel of length M in the presence of an additive noise can be written as follows [1]:

$$y_k = u_k w_0 + n_k \quad (1)$$

where u_k is a zero mean stationary input process with variance σ_u^2 , $\{n_k\}$ is a stationary noise process with zero mean and variance σ_n^2 , and w_0 is weight vector for M taps impulse response for unknown input.

In the case of the LMMN algorithm, the cost function to be minimized is given by [5]

$$J_{LMMN}(w) = \delta E[e_n^2] + (1 - \delta) E[e_n^4] \quad (2)$$

with α being the mixing parameter and e_n is the error between the output of the unknown system and the adaptive filter and is defined as

$$e_n = y_n - u_n w_n \quad (3)$$

The LMMN cost function is modified and leakage parameter α is introduced to obtain the proposed method same way as was done for the leaky LMS [1]. Therefore, the cost function that we wish to minimize is given as

$$J(w) = \alpha \|w\|^2 + \{\delta E[e_n^2] + (1 - \delta) E[e_n^4]\} \quad (4)$$

The corresponding update equation of the leaky LMMN is then

$$w_{n+1} = w_n + \mu e_n \{\delta + (1 - \delta) e_n^2\} u_n \quad (5)$$

The LMMN update equation is obtained for $\alpha = 0$ in (5) and μ is the step size.

III. PERFORMANCE ANALYSIS OF THE LEAKY LMMN ALGORITHM

In this section, the proposed leaky LMMN algorithm is analyzed in both the mean and the mean-square sense using the frame work of fundamental energy relation. Consequently, all the assumptions applicable for long filters are employed. In addition, the following assumptions are used:

A1 The noise sequence $\{v_n\}$ is i.i.d. with zero odd order moments and variance $\sigma_v^2 = E[v_n]^2$

A2 The sequence v_n is independent of u_j, w_k for all j, k .

A3 The regressor u_j are i.i.d Gaussian random vectors with covariance matrix $R_u = E[u_n^T u_n] > 0$.

A4 The random variables d_n, u_n and v_n have zero means.

The mean and mean square stability conditions are derived and learning curves are constructed to calculate MSD and EMSE.

A. Mean Behaviour

Defining the weight error vector w_n as follows:

$$w_n = c - w_n \tag{6}$$

Results in following update equation

$$w_{n+1} = (1 - \mu\alpha)w_n + \mu\alpha c - \mu u_n^T e_n [\delta + (1 - \delta) \|e_n\|^2] \tag{7}$$

taking the expectations on both sides and using the assumptions, we get

$$E[w_{n+1}] = (1 - \mu\alpha)E[w_n] + \mu\alpha c - \mu E[u_n^T e_n [\delta + (1 - \delta) \|e_n\|^2]] \tag{8}$$

To solve $E[u_n^T e_n [\delta + (1 - \delta) \|e_n\|^2]]$, we will make use of above assumptions to solve the equation:

$$\begin{aligned} E[u_n^T e_n [\delta + (1 - \delta) \|e_n\|^2]] &= \delta E[u_n^T e_n] + (1 - \delta) E[u_n^T e_n^3] \\ &= \{\delta + 3(1 - \delta)(\sigma_v^2 + \zeta)\} RE[w_n] \end{aligned} \tag{9}$$

Using (9) in (8), we get

$$E[w_{n+1}] = [I - \mu\{\alpha I + (\delta + 3(1 - \delta)(\sigma_v^2 + \zeta)R)\}]E[w_n] + \mu\alpha c \tag{10}$$

Where ζ is defined as

$$\zeta = E[e_a^2(n)] \tag{11}$$

The range of step-size values for which w_n remains bounded is given as

$$0 < \mu < \frac{2}{\alpha + [\delta + 3(1 - \delta)(\sigma_v^2 + \zeta)]\lambda_{\max}(R)} \tag{12}$$

Where $\lambda_{\max}(R)$ is the largest eigen value of R.

B. Behavior of Mean Square Error

In this section the performance of the leaky LMMN algorithm in the mean-square sense is analyzed. The assumption used is that the adaptive filter is long enough to justify the following:

A5 The norm of the input regressor ($\|u_n\|^2$) can be assumed to be uncorrelated with e_n^6

1) Error and Performance Measures:

We are interested in studying the time-evolution and the steady-state values of $E[\|w_n\|^2]$ and $E[e_a^2(n)]$ which give the EMSE and the MSD, respectively.

For some symmetric positive definite weighting matrix A, the weighted *a priori* estimation error is defined as

$$e_a^2(n) = u_n^T A w_n \tag{13}$$

the standard *a priori* estimation error is obtained when A = I from the above equation as,

$$e_a(n) = e_a^\Sigma(n) = u_n w_n \tag{14}$$

It is easy to see that the estimation error, e_n , and the *a priori* estimation error, $e_a(n)$, are related via

$$e_n = e_a(n) + v_n \tag{15}$$

Thus, using A3 and (14), the EMSE can be set up as follows:

$$\zeta_n = E[e_a^2(n)] = E[\|w_n\|_R^2] \tag{16}$$

2) Time evolution of the weighted variance $E[\|w_n\|_A^2]$:

In this section, the time evolution of the weighted variance $E[\|w_n\|_A^2]$ is derived for the leaky LMMN algorithm using the framework of fundamental weighted-energy conservation relation. Thus, by using (8) and (14), one can arrive at:

$$\begin{aligned} E[\|\tilde{w}_{n+1}\|_\Sigma^2] &= (1 - \mu\alpha)^2 E[\|\tilde{w}_n\|_\Sigma^2] + \mu\alpha c \|\Sigma\|^2 \\ &+ \mu^2 \text{tr}(R) H_v - 2\mu(1 - \mu\alpha) E[e_a(n) e_a^\Sigma(n)] H_G \\ &+ 2\mu\alpha c^T \Sigma J E[\tilde{w}_n] \end{aligned} \tag{17}$$

Where

$$J = [I - \mu\{\alpha I + [\delta + 3(1 - \delta)(\sigma_v^2 + \zeta)]R\}] \tag{18}$$

The transient behavior of the weighted variance $E[\|w\|_\Sigma^2]$ is shown above for any constant weight matrix Σ . Proper choice of the weight matrix Σ gives various performance measures. Now depending upon the correlation of the input the analysis is further divided into two parts.

3) Transient Analysis

When the input data is correlated, i.e., R is a non-diagonal matrix, different weighting matrices will appear on both sides of the equation. Therefore, we again resort to the technique given in [1]. At the end, we get

$$\underbrace{\begin{bmatrix} A_{n+1} \\ E[w_{n+1}] \end{bmatrix}}_{w_{n+1}} = \underbrace{\begin{pmatrix} F_1 & F_2 \\ 0 & J \end{pmatrix}}_{F_n} \underbrace{\begin{bmatrix} A_n \\ E[w_n] \end{bmatrix}}_{w_n}$$

Where F_1, F_2, A_n and L_n given as

$$A_n = \begin{bmatrix} E[\|w_n\|^2] \\ E[\|w_n\|_R^2] \\ E[\|w_n\|_{R^2}^2] \\ \vdots \\ E[\|w_n\|_{R^{M-1}}^2] \\ E[\|w_n\|_{R^M}^2] \end{bmatrix} \quad (19)$$

$$L_n = \mu H_U \begin{bmatrix} \text{tr}(R) \\ \text{tr}(R^2) \\ \text{tr}(R^3) \\ \vdots \\ \text{tr}(R^M) \end{bmatrix} + \mu \alpha^2 \begin{bmatrix} \|c\|^2 \\ \|c\|_R^2 \\ \|c\|_{R^2}^2 \\ \vdots \\ \|c\|_{R^{M-1}}^2 \end{bmatrix} \quad (20)$$

$$F_2 = 2\mu \alpha c^T J \begin{bmatrix} I \\ R \\ R^2 \\ \vdots \\ R^{M-1} \end{bmatrix} \quad (21)$$

$$F_1 = \begin{bmatrix} k_1 & -k_2 & 0 & 0 & \dots & 0 \\ 0 & k_1 & -k_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & k_1 & -k_2 \\ k_2 p_0 & k_2 p_1 & \dots & \dots & k_2 p_{M-2} & k_1 + k_2 p_{M-1} \end{bmatrix}$$

Where

$$k_1 = (1 - \mu\alpha)^2 \quad (22)$$

$$k_2 = 2\mu(1 - \mu\alpha)H_G$$

$$H_U = \mu M \delta^2 \sigma_v^2 + \mu M (1 - \delta)^2 \xi_v^6 + 2\mu \delta (1 - \delta) \xi_v^4$$

the first and second entries of the state vector w_{n+1} show the development of $E[\|w_n\|^2]$ and $E[e_a^2(n)]$. The learning curve of the filter then becomes

$$E[e^2(n)] = E[e_a^2(n)] + \sigma_v^2 \quad (23)$$

B. Mean Square Stability

One of the conditions for the mean square stability of the leaky LMMN algorithm to be convergent, is obtained from the block structure of F_k as shown in (20-21). To obtain the mean square stability of the leaky LMMN the

same approach of mean convergence on the step size is used as shown in (12).

$$0 < \mu < \frac{1}{\lambda_{\max}(G_1^{-1}G_2)} \quad (24)$$

Where $\lambda_{\max}(G_1^{-1}G_2)$ is the largest eigenvalue of $G_1^{-1}G_2$ with

$$G_1 = 2(\alpha I + H_G^* B) \quad (25)$$

$$G_2 = \alpha(\alpha I + 2H_G^* B)$$

Where in from above equations

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \\ -p_0 & -p_1 & \dots & \dots & -p_{M-2} & -p_{M-1} \end{bmatrix} \quad (26)$$

and

$$H_G^* = \delta + 3(1 - \delta)\{\sigma_v^2 + \vartheta\}$$

For the proposed method to converge in both the mean and mean square sense (12) and (24) are combined the condition obtained is

$$\mu_{\max} = \left\{ \begin{array}{l} \frac{2}{\alpha + H_G^* \lambda_{\max}(R)}, \alpha > \frac{H_G^* \lambda_{\max}(R)}{4} \\ \frac{1}{\lambda_{\max}^{G_1^{-1}G_2}}, otherwise \end{array} \right\} \quad (27)$$

IV. Experimental Results

The simulations are based on a system identification setup where the regression vector u_n is a Gaussian vector with filter of length 5.

The weight drift process is carried out in the same way as was done in [7]. With a filter vector of $[0.7071 - 0.7071]^T$ and regressor vector $\pm[0.5, -0.5]$ and choosing it to be equal probable making input matrix as singular, the proposed algorithm is compared with the LMMN in a weight drifting environment. The output and the quantization noise are grouped together and modeled as a Gaussian vector with mean of $[0.49, -0.49]^T$ and a variance of 10^3 . The adaptive filter coefficients and the regressors are set to 10 quantization bits with a step size of 0:0156 and the product of the step size and the leakage factor was set at 0:002. By taking the infinite norms of the updated weight vectors the experiment is run over 10^4 samples..

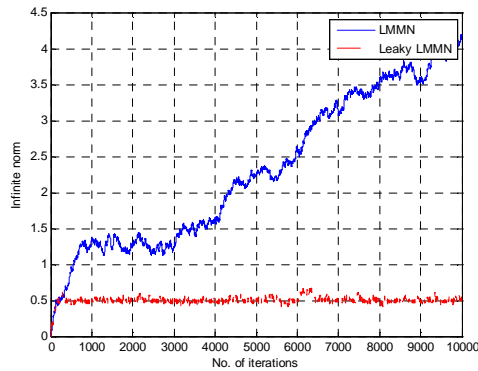


Figure 1. Comparison of LMMN and Leaky LMMN in a weight drift environment.

The above Fig(1) shows that the parameter drift causes the adaptive filter weights to blow up while in the case of the leaky LMMN they remain bounded.

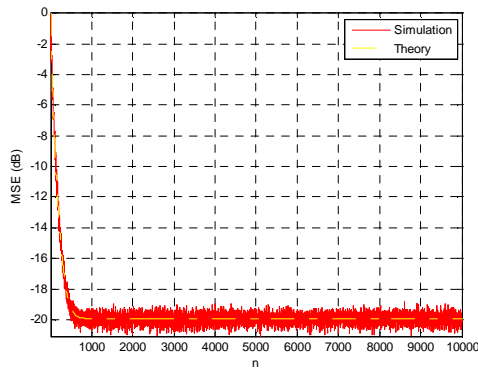


Figure 2. MSE learning curve of the LLMMN in Gaussian noise environment for white data

The white Gaussian data was used with the step size 0.01 and leakage factor is 0.001 to compare the theoretical findings with the simulation result, while the number of trials set at 500 and number of samples were 10^4 . The randomly generated normalized system weight vector with the number of taps set at 5 is taken and the Gaussian with variance 0.1 was chosen. The Fig. 2 shows a good match between the theoretical and simulation results.

CONCLUSIONS

In this paper, a new adaptive algorithm, the leaky LMMN algorithm, is presented. The expressions were derived for the transient analysis of the algorithm and also derived the conditions for the mean and mean square stability. Monte Carlo simulations were performed which match well with the theoretical values. Finally, the advantage of the leaky LMMN over the conventional LMMN in a weight drift environment is also shown.

ACKNOWLEDGMENT

I would like to acknowledge the support provided by King Fahd University of Petroleum & Minerals and CVR College of Engineering.

REFERENCES

- [1] "A.H. Sayed, Fundamentals of Adaptive Filtering. New York, NY:Wiley Interscience,2003".
- [2] E. Wallach and B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and its Family," IEEE Trans. Information Theory, vol. IT-30, no. 2, March 1984.
- [3] "A. Zerguine, "Convergence Behaviour of the Normalized Least Mean Fourth Algorithm," Proceedings of the 34th Annual Asilomar Conf. on Signals, Systems, and Computers, pp. 275-278, Oct.29-Nov.1, 2000".
- [4] R.D. Gitlin, H.C. Meadors and H.B. Weinstein, "The Tap Leakage Algorithm: An Algorithm for the Stable Operation of a Digitally Implemented Fractionally-Spaced Adaptive Equalizer," Bell Sys. Tech. Journal, vol. 61, pp. 1817-1840, Oct. 1982.
- [5] J. A. Chambers, O. Tanrikulu and A. G. Constantinides. "Least mean mixed norm adaptive filtering". Electronic Letters, 30(19):1574-1575, 1994.
- [6] A. Zerguine, M. K. Chan, T. Y. Al-Naffouri, M. Moinuddin, and C. F.N. Cowan, " Convergence and Tracking Analysis of a Variable Normalized LMF (XE-NLMF) Algorithm", Signal Processing, Vol. 89, No. 5, pp. 778-790, May 2009.
- [7] "V.H. Nascimento and A.H. Sayed, "Unbiased and Stable Leakage-Based Adaptive Filters", IEEE Trans. Signal Processing, vol. 47, No.12, pp. 3261-3276, December 1999".