A Time-Varying Normalized Mixed-Norm LMS-LMF Algorithm

Azzedine Zerguine
Electrical Engineering Department
King Fahd University of Petroleum & Minerals
KFUPM, Dhahran 31261, Saudi Arabia.
E-mail:azzedine@kfupm.edu.sa

Abstract

The normalized least mean square (NLMS) algorithm is known to result in a faster convergence than the least mean square (LMS) algorithm but at the expense of a larger steady-state error. A time-varying normalized mixed-norm LMS-least mean fourth (LMF) algorithm is presented in this work to preserve the fast convergence of the NLMS algorithm while resulting in a lower steady-state error. The simulation results show that a substantial improvement, in both convergence time and steady state error, can be obtained with this mixed-norm algorithm.

1 Introduction

Due to its simplicity, the least mean-square (LMS) [1] algorithm is the most widely used algorithm for adaptive filters in many applications. The least mean-fourth (LMF) [2] algorithm was also proposed later as a special case of the more general family of steepest descent algorithms [1] with $2k$ error norms, k being a positive integer.

But for both of these algorithms, the convergence behavior depends on the condition number, ratio of the maximum to the minimum eigenvalues of the input signal autocorrelation matrix $\mathbf{R} = E[\mathbf{x}_n\mathbf{x}_n^T]$, where $\mathbf{x}_n$ is the input signal.

To remove the dependency of the convergence of the LMS algorithm on the condition number, the normalized least-mean square (NLMS) [3] was introduced. Great improvement in convergence is obtained through the use of the NLMS algorithm over the LMS algorithm at the expense of a larger steady-state error.

A mixed-norm algorithm[4], combining both the LMS and the LMF algorithms, will suffer as well from the problem of the eigenvalue spread dependency. To circumvent this problem, a normalized version of the mixed-norm LMS-LMF algorithm must therefore be used.

It is well known that fast convergence and lower steady-state error are two conflicting parameters in general adaptive filtering. The NLMS algorithm results in the fastest convergence but only at the expense of a high steady-state error [5]. A promising solution to this conflict is a time-varying normalized mixed-norm LMS-LMF algorithm. In this mixed-norm algorithm and during the transient state, the NLMS algorithm is used to speed up the algorithm's convergence. However when the steady-state is reached, the algorithm automatically switches from the NLMS to the normalized LMF (NLMF) [6], thanks to a built-in "gear shifting" property, to secure a lower steady-state error.

In this work, the performance of the time-varying normalized mixed-norm LMS-LMF algorithm is evaluated. It is shown that great improvement in both convergence and steady state-state error is obtained through the use of this algorithm.

2 Proposed Algorithm

The mixed-norm LMS-LMF algorithm is based on the minimization of the following cost function [7]:

$$J_n = \alpha E[e_n^2] + (1-\alpha)E[e_n^4], \tag{1}$$

where α is a positive mixing parameter in the interval [0,1]. The error is defined as $e_n = d_n + w_n - \mathbf{x}_n^T\mathbf{c}_n$, where d_n is the desired value, $\mathbf{c}_n$ is the filter coefficient of the adaptive filter, $\mathbf{x}_n$ is the input vector and w_n is the additive noise. A major drawback of this algorithm is, however, the choice of the mixing parameter that is hard to fix a priori for an unknown system.

In [7], a self-adapting LMS-LMF algorithm with a time-varying weighting factor was proposed. This time-variation of the weighting factor was achieved by allowing for a variable mixing factor that is updated every iteration using the modified variable step-size (MVSS) algorithm proposed in [8]. The variable weight mixed-norm LMS-LMF algorithm was defined to minimize the following performance measure:

$$J_n = \alpha_n E[e_n^2] + (1-\alpha_n)E[e_n^4], \tag{2}$$

where α_n, chosen in $[0,1]$ such that the unimodal character of the above cost is preserved, is a time-varying parameter updated according to [8]:

$$\alpha_{n+1} = \delta\alpha_n + \gamma p_n^2, \tag{3}$$

and

$$p_n = \beta p_{n-1} + (1-\beta)e_n e_{n-1}. \quad (4)$$

The parameters δ and β, both confined to the interval [0,1], are exponential weighting parameters that govern the averaging time constant, i.e., the quality of estimation of the algorithm, and $\gamma > 0$. Note that the algorithm defined by (1) is restored when $\delta = 1$ and $\gamma = 0$, which forces α_n to have a fixed value.

Based on this motivation, the weight mixed-norm LMS-LMF algorithm for recursively adjusting the coefficients of the system is expressed in the following form:

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \mu[\alpha_n e_n + 2(1-\alpha_n)e_n^3]\mathbf{x}_n, \quad (5)$$

where μ is the step size.

As mentioned earlier and because of its reliance on the LMS and the LMF, the algorithm defined by (5) will be affected by the eigenvalue spread of the autocorrelation matrix of the input signal. To overcome this problem, a normalized version of this algorithm can be set up and resulting in the following weight update equation:

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \bar{\mu}[\alpha_n e_n + 2(1-\alpha_n)e_n^3]\frac{\mathbf{x}_n}{||\mathbf{x}_n||^2}, \quad (6)$$

where $||\mathbf{x}_n||^2$ is the Euclidean norm of the input vector $\mathbf{x}_n$ and $\bar{\mu}$ is the step size.

3 Convergence Analysis

Throughout our ensuing convergence analysis, the following commonly-used assumptions [1]-[2] are made:

A.1 *The noise sequence $\{w_n\}$ is statistically independent of the input signal sequence $\{x_n\}$ and both sequences have zero mean.*

A.2 *The noise w_n has zero odd moments.*

A.3 *The weight error vector, to be defined later, is independent of the input $\mathbf{x}_n$.*

Examining the mean behavior of Equation (6) under the above assumptions, sufficient conditions for convergence of the proposed algorithm in-the-mean can be derived and are stated as follows.

Proposition 1 *For the algorithm defined by (6) to converge in-the-mean, a sufficient condition is that $\bar{\mu}$ be chosen in the following range:*

$$0 < \bar{\mu} < \frac{2}{E[\alpha_n] + 2(1-E[\alpha_n])[3\sigma_w^2 + 1]}, \quad (7)$$

where σ_w^2 is the noise power and $E[\alpha_n]$ is the mean of the mixing parameter.

It is very clear that if $\alpha_n = 1$, both the NLMS algorithm and its step size range, that is $0 < \bar{\mu} < 2$, are recovered.

4 Steady-State Analysis

In general the adaptation scheme defined in (6) can be written in the following form:

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \bar{\mu}\frac{\mathbf{x}_n}{||\mathbf{x}_n||^2}\mathbf{f}_e(n), \quad (8)$$

where $\mathbf{f}_e(n)$ denotes a general scalar function of the output estimation error e_n and in our case $\mathbf{f}_e(n) = \alpha_n e_n + 2(1-\alpha_n)e_n^3$.

An important measure for an adaptive algorithm is its steady-state mean-square-error (MSE), which is defined as:

$$\begin{aligned} MSE &= \lim_{n\to\infty} E\left[e_n^2\right] \\ &= \lim_{n\to\infty} E\left\{\left[w_n - \mathbf{x}_n^T\mathbf{v}_n\right]^2\right\}. \end{aligned} \quad (9)$$

Under the often realistic assumption **A.1**, we find that the MSE is equivalently given by:

$$MSE = \sigma_w^2 + \lim_{n\to\infty} E\left\{\left[\mathbf{x}_n^T\mathbf{v}_n\right]^2\right\}. \quad (10)$$

4.1 Fundamental Energy Conservation Relation

In this section the steady-state error analysis of the proposed algorithm is carried out using the concept of feedback approach [9] to derive the energy conservation relation. First let us define the following a-priori estimation error, $e_a(n) = \mathbf{x}_n^T\mathbf{v}_n$, and the a-posteriori estimation error, $e_p(n) = \mathbf{x}_n^T\mathbf{v}_{n+1}$. It is easy to show that the estimation error, e_n, and the a-priori error, $e_a(n)$, are related by $e_n = e_a(n) + w_n$. Also, it is easy to show that the a-poteriori error, $e_p(n)$, can be expressed in the following form:

$$e_p(n) = e_a(n) - \frac{\bar{\mu}}{\hat{\mu}_n}\mathbf{f}_e(n), \quad (11)$$

where $\hat{\mu}_n = 1/\left\|\mathbf{x}_n\right\|^2$. Substituting (11) into (8) results into the following update relation:

$$\mathbf{v}_{n+1} = \mathbf{v}_n - \hat{\mu}_n\mathbf{x}_n\left[e_a(n) - e_p(n)\right]. \quad (12)$$

By evaluating the energies of both sides of the above equation, the following new relation is obtained:

$$\left\|\mathbf{v}_{n+1}\right\|^2 + \hat{\mu}_n||e_a(n)||^2 = \left\|\mathbf{v}_n\right\|^2 + \hat{\mu}_n||e_p(n)||^2. \quad (13)$$

4.2 Steady-State Performance Analysis

Recall that we are interested in evaluating the MSE of the adaptive filter once its reaches the steady state. To do so, we resort to (13) and note that in steady-state $E[\left\|\mathbf{v}_{n+1}\right\|^2] = E[\left\|\mathbf{v}_n\right\|^2]$, so that by taking expectations of both sides of (13) we obtain the equality:

$$E\left[\hat{\mu}_n||e_a(n)||^2\right] = E\left[\hat{\mu}_n||e_p(n)||^2\right]. \quad (14)$$

Since $e_p(n)$ is itself a function of $e_a(n)$, Equation (14) collapses to the following fundamental error variance relation in terms of the desired but unknown $e_a(n)$ only

$$E\left[\hat{\mu}_n \|e_a(n)\|^2\right] = E\left[\hat{\mu}_n \|e_a(n) - \frac{\bar{\mu}}{\hat{\mu}_n}\mathbf{f}_e(n)\|^2\right]. \quad (15)$$

The above equation can now be solved for the steady-state excess mean-square-error $(EMSE)$ defined as:

$$\begin{aligned}\zeta &= \lim_{n\to\infty} E\left\{\left[\mathbf{x}_n^T\mathbf{v}_n\right]^2\right\} \\ &= \lim_{n\to\infty} E\left[e_a^2(n)\right].\end{aligned}$$

Observe from (10) that the desired MSE is given by $MSE = \sigma_w^2 + \zeta$, so that finding ζ is equivalent to finding the MSE. Equation (15) can be re-written as:

$$2\bar{\mu}E\left[e_a(n)\mathbf{f}_e(n)\right] = \bar{\mu}^2 E\left[\|\mathbf{x}_n\|^2 \mathbf{f}_e^2(n)\right]. \quad (16)$$

Substituting $\mathbf{f}_e(n)$ in (16) and taking assumptions $\mathbf{A.1}$ and $\mathbf{A.3}$ into account, results in:

$$[a - \bar{\mu}b]\, E\left[\frac{e_a^2(n)}{\|\mathbf{x}_n\|^2}\right] = \bar{\mu}cE\left[\frac{1}{\|\mathbf{x}_n\|^2}\right], \quad (17)$$

where $a = 2\left[E[\alpha_n] + 6(1 - E[\alpha_n])\right]$, $b = \left[E[\alpha_n^2] + 12E[\alpha_n(1-\alpha_n)]\sigma_w^2 + 36E[(1-\alpha_n)^2]\xi_w^4\right]$, $c = \left[E[\alpha_n^2]\sigma_w^2 + 4E[\alpha_n(1-\alpha_n)]\xi_w^4 + 4E[(1-\alpha_n)^2]\chi_w^6\right]$, $\chi_w^6 = E[w_n^6]$, and $\xi_w^4 = E[w_n^4]$.

Two cases can be considered for the evaluation of the expression $E\left[\frac{e_a^2(n)}{\|\mathbf{x}_n\|^2}\right]$. These are detailed in the ensuing analysis. But before doing so, let us state one more assumption needed for the derivations.

$\mathbf{A.4}$ *In steady state,* $\bar{\mu}^2\|\mathbf{x}_n\|^2$ *is statistically independent of* $e_a^2(n)$.

Assumption $\mathbf{A.4}$ in fact becomes very realistic for long filter lengths. Furthermore, the case of larger values of $\bar{\mu}$ is considered in assumption $\mathbf{A.4}$.

Case 1: Under assumption $\mathbf{A.4}$, expression $E\left[\frac{e_a^2(n)}{\|\mathbf{x}_n\|^2}\right]$ becomes:

$$E\left[\frac{e_a^2(n)}{\bar{\mu}^2\|\mathbf{x}_n\|^2}\right] = E\left[e_a^2(n)\right] E\left[\frac{1}{\bar{\mu}^2\|\mathbf{x}_n\|^2}\right], \quad (18)$$

so that (17) leads to the excess MSE for the proposed algorithm:

$$\zeta_{proposed} = \frac{\bar{\mu}c}{a - \bar{\mu}b}. \quad (19)$$

Similarly the excess MSE for the NLMS algorithm will be:

$$\zeta_{NLMS} = \frac{\bar{\mu}\sigma_w^2}{2 - \bar{\mu}}. \quad (20)$$

Case 2: Under the following approximation ([1], pp. 443):

$$E\left[\frac{e_a^2(n)}{\|\mathbf{x}_n\|^2}\right] \approx \frac{E\left[e_a^2(n)\right]}{E\left[\|\mathbf{x}_n\|^2\right]}, \quad (21)$$

expression (17) leads to the excess MSE for the proposed algorithm

$$\zeta_{proposed} = \frac{\bar{\mu}c}{a - \bar{\mu}b}E\left[\frac{1}{\|\mathbf{x}_n\|^2}\right] tr\left\{\mathbf{R}\right\}. \quad (22)$$

And in the case of the NLMS algorithm this can be found to be equal to:

$$\zeta_{NLMS} = \frac{\bar{\mu}\sigma_w^2}{2 - \bar{\mu}}E\left[\frac{1}{\|\mathbf{x}_n\|^2}\right] tr\left\{\mathbf{R}\right\}. \quad (23)$$

In the above two cases, the ratio of the excess MSE of the two algorithms for the same step size $\bar{\mu}$ is given by:

$$\frac{\zeta_{proposed}}{\zeta_{NLMS}} = \frac{[2 - \bar{\mu}]\, c}{[a - \bar{\mu}b]\,\sigma_w^2}, \quad (24)$$

where it is difficult to draw a decisive conclusion about the behavior of the two algorithms, and only special cases are considered and unfortunately due to space limitations these are not reported here.

5 Simulation results

The performance of the proposed algorithm, the time-varying normalized mixed-norm LMS-LMF algorithm, is compared with that of the NLMS algorithm. Experiments are carried out where an unknown system is to be identified under noisy conditions. The unknown system is a non-minimum phase channel. The input signal x_n to both the unknown system and the adaptive filter is obtained by passing a zero-mean white Gaussian sequence through a channel that is used to vary the eigenvalue spread of the autocorrelation matrix of the input signal. The example considered for the sequence $\{x_n\}$ has an eigenvalue spread of 68.9. The additive noise, w_n, is a zero-mean and uniformly distributed. The signal to noise ratio is set to be equal to 20 dB and the performance measure considered is the normalized weight error norm $10log_{10}||\mathbf{c}_n - \mathbf{c}_{opt}||^2/||\mathbf{c}_{opt}||^2$. Results are obtained by averaging over 600 independent runs. The proposed algorithm is implemented with the parameters $\delta = 0.97$, $\beta = 0.98$, $\gamma = 10^{-2}$, $\alpha_0 = 0.8$ and $p_0 = 0$.

Figure 1 compares the fastest convergence characteristics of both the proposed algorithm and the NLMS algorithm. It can be seen from this figure that the proposed algorithm converges as fast as the NLMS algorithm but results in a lower weight mismatch. An improvement of 25 dB is obtained through the use of the proposed algorithm.

Also, as shown in Fig. 2, the proposed algorithm outperforms the NLMS algorithm, for the lowest steady-state error reached by the later, thanks to its built-in gear-shifting mechanism which gives it an extra degree of freedom in this region.

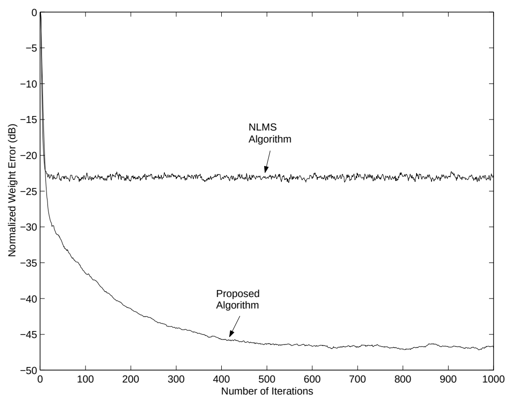

Figure 1: Fastest convergence characteristics of the proposed and NLMS algorithms, $\mu_{proposed} = 0.55$, $\mu_{NLMS} = 1.00$.

The fast convergence obtained by the proposed algorithm can be justified by the fact that when far from the optimum solution, this algorithm exhibits faster convergence than the NLMS algorithm by automatically increasing the step size (gear-shifting property).

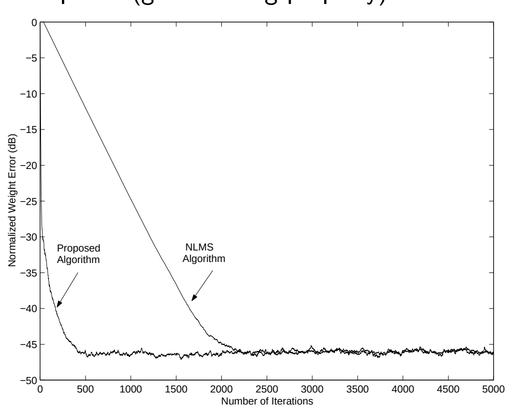

Figure 2: Convergence characteristics of the proposed and NLMS algorithms for the lowest steady-state error reached by the NLMS algorithm, $\mu_{proposed} = 0.65$, $\mu_{NLMS} = .0055$.

Finally, from the viewpoint of computational load the proposed algorithm requires an additional seven multiplications and three additions when compared to the fixed mixed-norm algorithm defined by (1), and only eleven multiplications and six additions when compared to the NLMS algorithm. The small computational over head of the proposed algorithm is therefore well worth the gain in the steady-state error reduction it brings about.

6 Conclusion

In this work, a normalized time-varying mixed-norm algorithm is proposed where a combination of the LMS and the LMF algorithms is incorporated using the concept of variable step-size LMS adaptation. It is found that the proposed algorithm has the fast convergence property of the NLMS algorithm while resulting in a lower steady-state error, therefore eliminating the conflict between these two parameters, i.e., fast convergence and low steady-state error. Finally, the consistency of the performance of the proposed algorithm has been confirmed by other simulation results which are not reported here.

Acknowledgment The author acknowledges the support of KFUPM.

References

[1] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1991.

[2] E. Walach and B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and its Family," *IEEE Trans. Inf. Theory, vol.* IT-30, pp. 275-283, Feb. 1984.

[3] J. I., Nagumo and A. Noda, "A learning method for system identification," *IEEE Transactions, Automatic control*, AC-12, pp. 282-287, 1967.

[4] T. Y. Al-Naffouri, A. Zerguine, and M. Bettayeb, "Convergence properties of mixed-norm algorithms under general error criteria," *Proceedings of the 1999 IEEE ISCAS '99*, pp. 211-214, 1999.

[5] M. Tarrab and A. Feuer, "Convergence and Performance Analysis of the Normalized LMS Algorithm with Uncorrelated Gaussian Data," *IEEE Trans. Inf. Theory, vol.* IT-34, pp. 680-691, July 1988.

[6] A. Zerguine, "Convergence behavior of the normalized least mean fourth algorithm," *Proc. 34th Annual Asilomar Conf. Signals, Syst., Comput.*, pp. 275-278, 2000.

[7] A. Zerguine and T. Aboulnasr, "Convergence analysis of the variable weight mixed-norm LMS-LMF adaptive algorithm," *Proc. 34th Annual Asilomar Conf. Signals, Syst., Comput.*, pp. 279-282, 2000.

[8] T. Aboulnasr and K. Mayyas, "A Robust Variable Step-Size LMS-Type Algorithm: Analysis and Simulations," *IEEE Trans. Signal Processing, vol.* SP-45, No. 3, pp. 631-639, March, 1997.

[9] M. Rupp and A. H. Sayed, "A Time-Domain Feedback Analysis of Filtered-Error Adaptive Gradient Algorithms," *IEEE Trans. Signal Proc., vol.* 44, pp. 1428-1439, June 1996.