A fast converging improved NLMS algorithm for spares system's identification

Shihab Jimaa*, Luis Weruaga*, Ammar Sadiq*, Hamad Alhammadi*, and Tetsuya Shimamura#

* ECE Department, Khalifa University of Science, Technology and Research, UAE
# Graduate School of Science & Eng., Saitama University, Japan
saj, luis.weruaga@kustar.ac.ae, and shima@sie.ics.saitama-u.ac.jp

Abstract- The Least Mean Square (LMS) and its variant the Normalized LMS (NLMS) adaptive algorithms are widely used in system identification, however they are better in identifying non sparse systems. Hence, in order to improve the performance of the LMS based system identification of sparse systems, an improved adaptive NLMS algorithm is proposed which utilizes the sparsity property of such systems. A sparse system is one whose impulse response consists of many near-zero coefficients. In this paper, a comparative analysis of the LMS, NLMS and the proposed improved NLMS algorithms, to identify an unknown sparse system, is presented. Simulation results demonstrates that the proposed improved NLMS algorithm provides significant performance gains in comparison to the conventional LMS and NLMS algorithms in both convergence rate and steady state behavior.

Keywords*- Adaptive filters, NLMS, System Identification, Sparse systems*

1. Introduction

Adaptive filtering algorithms have been widely applied to solve many problems in digital communication systems such as acoustic echo cancellation, noise cancellation, channel estimation, and system identification [1, 2]. So far, the LMS and the NLMS adaptive algorithms have been the most commonly adopted approaches owing to the clarity of the mean-square-error cost function in terms of statistical concept and the simplicity for computation.

A sparse system is one whose impulse response consists of many near-zero coefficients and very few large coefficients There are many adaptive algorithms for system identification, such as Least Mean Squares (LMS) and Recursive Least Squares (RLS). However, these algorithms have no particular advantage in sparse system identification due to no use of sparse characteristic. In recent decades, some algorithms have exploited the sparse nature of a system to improve the identification performance. To our knowledge, the first of them is Adaptive Delay Filters (ADF), which locates and adapts each selected tap-weight according to its importance. Then, the concept of proportionate updating was originally introduced for echo cancellation application by Duttweiler. The underlying principle of Proportionate Normalized LMS (PNLMS) is to adapt each coefficient with an adaptation gain proportional to its own magnitude. Based on PNLMS, there exist many improved PNLMS algorithms, such as IPNLMS and IIPNLMS. Besides the above mentioned algorithms, there are various improved LMS algorithms on clustering sparse system. These algorithms locate and track non-zero coefficients by dynamically adjusting the length of the filter. The convergence behaviors of these algorithms depend on the span of clusters (the length from the first non-zero coefficient to the last one in an impulse response). When the span is long and close to the maximum length of the filter or the system has multiple clusters, these algorithms have no advantage compared to the traditional algorithms.

Motivated by Least Absolutely Shrinkage and Selection Operator (LASSO) and the recent research on Compressive Sensing (CS), a new LMS algorithm with l_0 norm constraint is proposed in order to accelerate the sparse system identification. Specifically, by exerting the constraint to the standard LMS cost function, the solution will be sparse and the gradient descent recursion will accelerate the convergence of near-zero coefficients in the sparse system. Furthermore, using partial updating method, the additional computational complexity caused by l_0 norm constraint is far reduced.

The paper is organised as follows. Section 2 introduces the discrete time model of the digital communications system. Section 3 gives a brief summary of the NLMS adaptive algorithm. Section 4 gives a brief summary of the proposed algorithm. In section 5 a computer simulation study is given to investigate the performance of the proposed algorithm and to compare it with the performance of the NLMS. Conclusions are given in Section 6. Finally, a list of references is given at the end of the paper.

2. System Model

The discrete time model for the digital communication system considered here is shown in Fig.1. The desired response $d(n)$, providing a frame of reference for the adaptive filter, is defined by:

$$d(n) = w_0{}^H u(n) + v(n) \qquad (1)$$

where $u(n)$ is the input vector, which is common to both the unknown system and the adaptive filter and $v(n)$ is the AWGN with zero mean and variance σ^2. The error signal e(n), involved in the adaptive process, is defined by

$$e(n) = d(n) - y(n) \quad (2)$$

$$= w_0^H(n)\, u(n) + v(n) - \hat{w}^H(n)\, u(n) \quad (3)$$

where w^(n) is the tap-weight vector of the adaptive filter assumed to have a transversal structure. It is also assumed that the adaptive filter has the same number of taps as the unknown system represented by w_0.

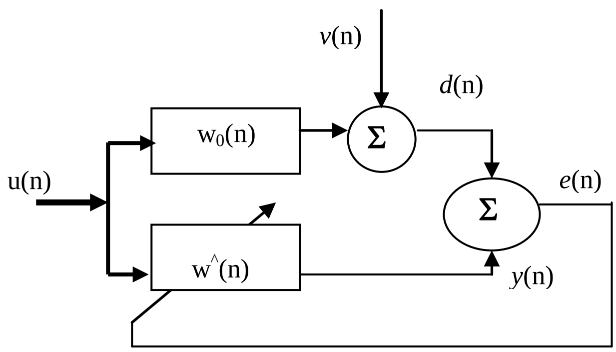

Figure 1 System identification model using Linear transversal adaptive filter

3. NLMS Algorithm

In a transversal adaptive filter, the input vector *U(n)* and the weight vector *W*(n) at the time of n[th] iteration are defined as follows:

$$U_n = [u_n, u_{n-1}, ..., u_{n-M+1}]^T \quad (4)$$

$$W_n = [w_0^{\wedge}, w_1^{\wedge}, ..., w_{M-1}^{\wedge}]^T \quad (5)$$

where u_n is the filter input and $w^{\wedge}_i$, (i=0, 1, …, M-1) is the weight vector which corresponds to the filter length. The filter output is obtained as follows:

$$y_n = W_n^{\wedge T} U_n \quad (6)$$

In the Least Mean Square (LMS) algorithm, the weight vector W_n is updated using the following equation:

$$W_{n+1} = W_n + \mu e_n U_n \quad (7)$$

In (7) μ is a positive constant, which represents the step size parameter in the LMS algorithm. By repeating the updating, the MSE: E[e_n^2] is minimized, where E[·] denotes expectation. To ensure the stability of the adaptive process and the convergence of the MSE, the step size,μ, should satisfy the following condition:

$$0 < \mu < \frac{2}{E[U_n^T U_n]} \quad (8)$$

On the other hand the NLMS algorithm adapts the tap weight vector [3] using a gradient descent algorithm that reduces the squared estimation error at each time instant. The algorithm is given below:

$$W[n+1] = W[n] + \mu\, e_n \frac{U[n]}{\lambda + \| U[n] \|^2} \quad (9)$$

where μ is the algorithm's step-size parameter λ is a very small number added to avoid instability, and e_n is the error signal which is defined as:

$$e_n = d_n - U_n^T W_n^{\wedge} \quad (10)$$

4. The Proposed Algorithm

The results of Improved NLMS Algorithm are presented for the BPSK signal. The unknown system is represented by a 100 tap FIR filter, whose taps are randomly generated. The proposed improved NLMS algorithm is obtained by minimizing the following cost function:

$$L(n) = \frac{1}{2} \cdot e^2(n) + \hat{\gamma} \|w(n)\|_0 \quad (11)$$

Where $\|w(n)\|_0$ denotes the $\mathcal{L}_0$ norm that signifies the number of non-zero entries, and γ is a parameter adopted to adjust the influence of the $\mathcal{L}_0$ norm cost.

$$\gamma = \frac{\hat{\gamma}}{u^T(n) \cdot u(n)} \quad (12)$$

The recursion equation for the proposed algorithm is based on the principle of Newton method or Newton algorithm where you can find successively proper approximations to the function. The recursion equation is shown as follows:

$$W(n+1) = w(n) - \mu \cdot \nabla E \cdot (\nabla^2 E)^{-1} \quad (13)$$

Where ∇E represents the gradient of the error with respect to the tap weights (coefficients) and $\nabla^2 E$ is the second gradient or the "Hessian" of the error function.

$$\nabla E = \frac{\partial\, E(n)}{\partial w} \quad (14)$$

$$= -2 \cdot e(n) \cdot u(n) + \gamma \cdot \beta \cdot \mathrm{sign}(w(n)) e^{-\beta |s(w(n))|}$$

$$E = \frac{\partial^2 E(n)}{\partial w^2} \quad (15)$$

More accurately, the second gradient can be shown as follows, which corresponds to a matrix:

$$\nabla^2 E = \frac{\partial^2 E(n)}{\partial w_i^2\, \partial w_j^2} \quad (16)$$

And the parameter β is a positive value that is used to determine the region of zero attraction In addition, the filter coefficients are updated along the negative gradient direction.

$$w_{new}\ w_{prev} + \text{gradient correcti} + \text{zero attraction}, \tag{17}$$

Zero attraction imposes an attraction to zero on small coefficients. The function of zero attractor leads to the performance improvement of l0-LMS in sparse system identification. To be specific, in the process of adaptation, a tap coefficient closer to zero indicates a higher possibility of being zero itself in the impulse response, when a coefficient is within a neighborhood of zero $(-1/\beta, 1/\beta)$, the closer it is to zero, the greater the attraction intensity is. When a coefficient is out of the range, no additional attraction is exerted. Thus, the convergence rate of those near-zero coefficients will be raised. In conclusion, the acceleration of convergence of near-zero coefficients will improve the performance of sparse system identification since those coefficients are in the majority. A large β means strong intensity but narrow attraction range. Therefore, it is difficult to evaluate the impact of β on the convergence rate. For practical purposes, Bradley and Mangasarian suggested to set the value of β to some finite value like 5 or increased slowly throughout the iteration process for better approximation.

5. Simulation Results

This section presents the computer simulations results of the performance of the proposed improved NLMS algorithm in spares systems in comparison with the standard LMS and NLMS algorithms. The behavior of the error estimation will be shown in the figures below with the comparison of four algorithms, LMS, NLMS, Modified NLMS, and Improved NLMS. Moreover, to obtain the best minimum square error, different values of γ were analyzed. The next figures show how using large and small values of γ can affect the behavior of the MSE along the curve. For a large value of γ, the behavior of the MSE is shown in the figure below:

.

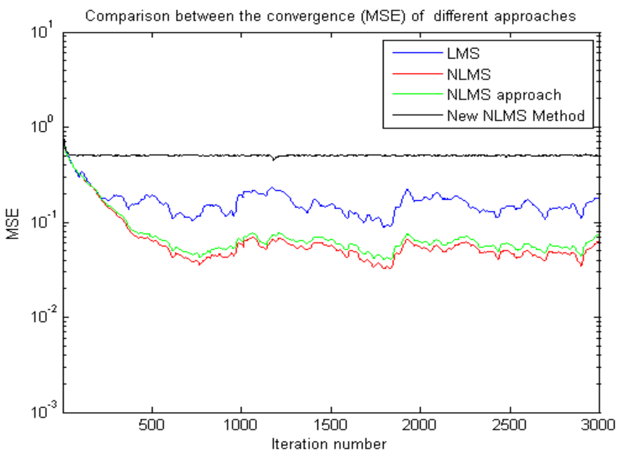

Figure 2: MSE performances comparison between four algorithms where $\boldsymbol{\gamma}$ = 0.005(large value)

According to the figure 2 above, it can be clearly seen that the value of the proposed improved NLMS algorithm takes another turn in providing inaccurate error estimation and this value convey the least performance compared with other methods.

In figure 3, it can be clearly seen that the value of the proposed algorithm takes another turn in providing inaccurate error estimation and this value convey undesired performance, similar to the performance of the standard LMS and NLMS algorithms.

Now, in order to get to the best value of γ, we tried different values and then average these values to get the one that provides minimum square error.

Figure 4 shows how we have obtained the average value of γ between the various trial values. Finally in Figure 5, the performance of the proposed improved NLMS algorithm shows a 10 dB outstanding performance improvement in comparison with the NLMS.

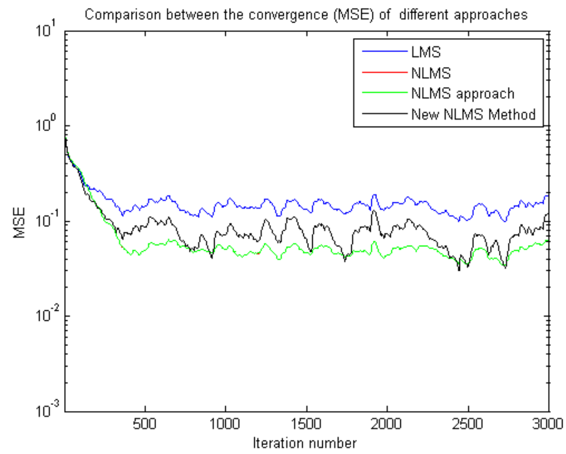

Figure 3 Comparison between the four algorithms where $\boldsymbol{\gamma} =$ 0.000003

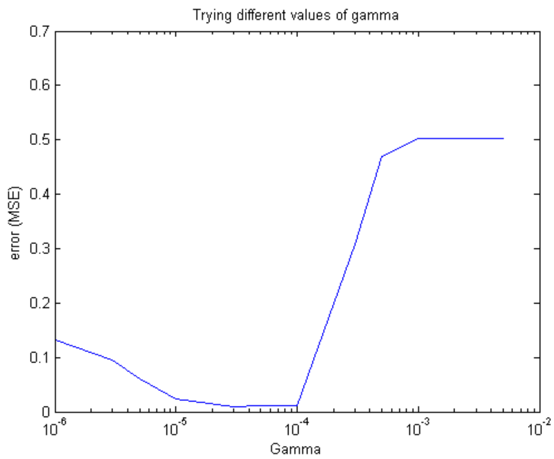

Figure 4: MSE performances of different values of $\boldsymbol{\gamma}$

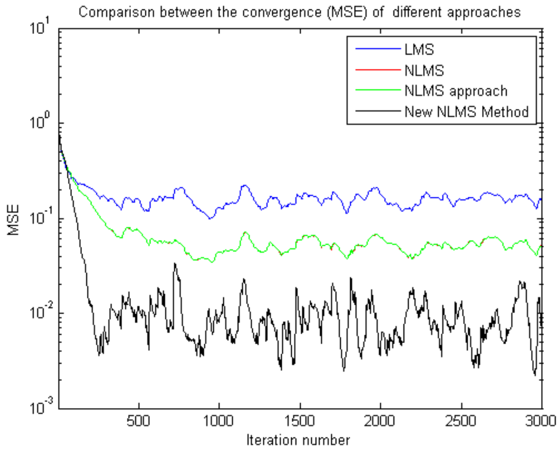

Figure 5: Best value obtained to provide the best error estimation; $\gamma =$ 0.00005

6. Conclusions

In this paper, in order to reach to a better approximation, the parameter of γ was analyzed and many trial and error steps were done in order to reach to a better value to obtain the best error estimation. The error estimation for the proposed improved algorithm was then compared with three various LMS algorithms in order to notice its behavior and successfully it gave a better approximation of the error estimation.

The performance of the proposed improved NLMS algorithm shows a 10 dB outstanding performance improvement in comparison with the LMS and NLMS algorithms.

References

[1] C.F.N.Cowan and P.M.Grant, *Adaptive Filters*, (Prentice Hall, Englewood Cliffs, 1985).

[2] B.Widrow and S.D.Stearns, *Adaptive Signal Processing*, (Prentice Hall, Englewood Cliffs, NJ, 1988).

[3] A. I. Sulyman and A. Zerguine, Convergence and steady-state analysis of a variable step-size NLMS algorithm, ELSEVIER Signal Processing, Vol. 83, , 2003, pp. 1255-1273.

[4] E.Walach and B.Widrow, The Least Mean Fourth (LMF) Adaptive Algorithm and its Family" *IEEE Transactions on Information Theory*, 30(2), 1984, 275-283.

[5] S.A. Jimaa, C.F.N. Cowan and M.J.J. Holt," Adaptive Channel Equalisation Using Least Mean Switched Error Norm Algorithm", IEE 16th Saraga Colloquium on Digital and Analogue Filters and Filtering Systems, London, 9 December 1996.

[6] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms", IEEE Trans. Signal Processing 41 (9), September 1993.

[7] Y. Chen, Y. Hu, and A. O. Hero, "Sparse LMS for system identification,"*Proc IEEE ICASSP*, 2009, pp. 3125–3128.

[8] Y. Gu, J. Jin, and S. Mei, "ℓ0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Proces. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.

[9] K. Shi and P. Shi , "Convergence analysis of sparse LMS algorithms with ℓ1-norm penalty based on white input signal," *Signal Processing*, vol. 90, no. 12, pp. 3289–3293, Dec. 2010.

[10] G. Su, J. Jin, Y. Gu, and J. Wang, "Performance analysis of ℓ0 norm constraint least mean square algorithm," *IEEE Trans. Signal Proces.*, vol. 60, no. 5, pp. 2223–2235, May 2012.