New LMS Algorithms Based on the Error Normalization Procedure

ZAYED RAMADAN and ALEXANDER POULARIKAS
Electrical and Computer Engineering Department
The University of Alabama in Huntsville
Huntsville, AL 35899
USA

Abstract:– This paper introduces several new least mean-square (LMS) algorithms based on error normalization procedure. Different minimization approaches and techniques were used in developing the proposed algorithms. Some of these algorithms are selected and applied to an adaptive noise canceller setup with different stationary noise power levels. Simulation results, carried out using a real speech, clearly demonstrate the superiority of the proposed algorithms over other standard LMS algorithms. Smaller values of steady-state excess mean-square error and better tracking capabilities are obtained using these new algorithms.

Key-Words: Adaptive filters, Adaptive noise cancellers, Error nonlinearity, Error normalization LMS, LMS algorithms.

1 Introduction

The LMS algorithm is a stochastic gradient algorithm in that it iterates each tap weight of the transversal filter in the direction of the instantaneous gradient of the squared error signal with respect to the tap weight in question. The simplicity of the LMS algorithm coupled with its desired properties, has made it and its variants an important part of the adaptive techniques.

Because of the successful use of the LMS algorithm in modeling unknown systems [1,2], in linear prediction [3,4], in adaptive noise canceling [5,6], in adaptive antenna systems [7], in channel equalization [8], and in many other areas [9–11], improvements of the algorithm are constantly being sought. The present work introduces several new algorithms based on the introduction of the error norm.

2 Least-Perturbation Approach

The relation

$$e(i) = d(i) - \mathbf{x}(i)\,\mathbf{w}(i-1) \tag{1}$$

is defined as the a priori output error, and

$$e_{ps} = d(i) - \mathbf{x}(i)\,\mathbf{w}(i) \tag{2}$$

is defined as the a posteriori output error ($i \geq 0$). Furthermore, we define the following quantities: $d(i)$ is the desired signal, the data $\mathbf{x}(i) = [x(i)\ x(i-1)\ \ldots\ x(i-N+1)]$ is a row vector, and the filter coefficients $\mathbf{w}(i)=[w_0(i)\ w_1(i)\ \ldots\ w_{N-1}(i)]^T$ is a column vector.

If we apply the following minimization problem [11]:

$$\min_{\mathbf{w}} \ \left\|\mathbf{w}(i) - \mathbf{w}(i-1)\right\|^2 \quad \text{subject to}$$

$$e_{ps}(i) = \left(1 - \mu \frac{\|\mathbf{x}(i)\|^2}{\|\mathbf{e}(i)\|^2}\right) e(i), \tag{3}$$

we obtain the *error normalized* LMS algorithm:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \mu(i)\,\mathbf{x}^H(i)\,e(i)\ , \tag{4}$$

where

$$\mu(i) = \frac{\mu}{\left\|\mathbf{e}(i)\right\|^2}\ , \tag{5}$$

and $\mathbf{e}(i) = [e(i)\ e(i-1)\ \ldots\ ]^T$ is a continuous increasing in length vector or a vector with varying energy but constant length N. If we had introduced the constraint:

$$e_{ps}(i) = \left(1 - \mu \frac{\|\mathbf{x}(i)\|^2}{\|\mathbf{e}(i)\|^2\ \|\mathbf{x}(i)\|^2}\right) e(i), \tag{6}$$

we would have obtained the following *error-data normalized* LMS algorithm:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \frac{\mu}{\|\mathbf{e}(i)\|^2 \|\mathbf{x}(i)\|^2} \mathbf{x}^H(i)\ e(i) \tag{7}$$

3 Regularization Method

To develop the regularized Newton method we add a positive constant to the Hessian matrix and, hence the recursive equation becomes [11]:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \mu[\varepsilon I + \nabla_W^2 J(\mathbf{w}(i-1)]^{-1} \times [\nabla_w J(w(i-1)]^H \tag{8}$$

where w(−1) = initial guess and $i \geq 0$.

Applying the simplification for the LMS algorithm, (8) becomes

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \mu[\varepsilon I + R_X]^{-1} \times [P_{dx}(i) - R_X \mathbf{w}(i-1)] \tag{9}$$

where

$$R_X = \mathbf{x}^H(i)\,\mathbf{x}(i) \text{ and } P_{dx}(i) = d(i)\,\mathbf{x}^H(i)$$

Expanding (9) and applying the matrix inversion theorem we obtain the ε-normalized LMS algorithm. If, however, instead of the constant ε we introduced the constant $\varepsilon\|\mathbf{e}(i)\|^2$, the *ε-error normalized* LMS algorithm takes the form:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \frac{\mu}{\varepsilon\ \|\mathbf{e}(i)\|^2 + \|\mathbf{x}(i)\|^2} \mathbf{x}^H(i)\quad e(i) \tag{10}$$

If we introduce two positive constants α and γ in (9), we obtain the following equation:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \mu\,[\,\alpha\,\|\mathbf{e}(i)\|^2\ I + \mathbf{x}^H(i)\,\gamma\,\mathbf{x}(i))]^{-1} \times \mathbf{x}^H(i)\,[d(i) - \mathbf{x}(i)\ \mathbf{w}(i-1)]^H \tag{11}$$

Applying the matrix inverse theorem to $[\alpha\|\mathbf{e}(i)\|^2\ I + \mathbf{x}^H(i)\,\gamma\,\mathbf{x}(i))]^{-1}$, we obtain the *ratio-error-data* LMS algorithm

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \frac{\mu}{\alpha\|\mathbf{e}(i)\|^2} \times \left[1 - \frac{\gamma\|\mathbf{x}(i)\|^2}{\alpha\|\mathbf{e}(i)\|^2 + \gamma\,\|\mathbf{x}(i)\|^2}\right] \times \mathbf{x}^H(i)\,[d(i) - \mathbf{x}(i)\,\mathbf{w}(i-1)] \tag{12}$$

this can be rewritten in a more simplified form as:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \left[\frac{\mu}{\alpha\,\|\mathbf{e}(i)\|^2 + \gamma\|\mathbf{x}(i)\|^2}\right] \mathbf{x}^H(i)\ \ e(i) \tag{13}$$

If we let γ=0 and α=1 in (13), we obtain (4).

4 Other LMS-Type Algorithms

If we apply the following minimization:

$$\min_{\mathbf{w}}\ [\,\alpha\|\mathbf{e}(i)\|^2\ \|\mathbf{w}\| + E\{|\,d - \mathbf{x}\mathbf{w}\,|^2\}\,],$$

we obtain the *error-leaky* LMS algorithm

$$\mathbf{w}(i) = (1 - \mu\alpha\|\mathbf{e}(i)\|^2)\,\mathbf{w}(i-1) + \mu\mathbf{x}^H(i)\ e(i) \tag{14}$$

provided that $(1 - \mu\alpha\|\mathbf{e}(i)\|^2)$ is less than 1. This algorithm constraints the step size constant μ in the range

$$0 < \mu < \frac{2}{\alpha \|\mathbf{e}\|^2_{max} + \lambda_{max}} \qquad (15)$$

where λ_{max} is the maximum eigenvalue of the data correlation matrix $\mathbf{R}_x$.
If we, next, apply the minimization procedure

$$\min_{\mathbf{w}} \quad E\,[\,(\gamma \|\mathbf{e}(i)\|^2 - |\,\mathbf{xw}\,|^2\,)^2\,]$$

where γ is a positive constant, we obtain the *error-constant-modulus* algorithm:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + \mu \mathbf{x}^H(i)\,[\,\mathbf{x}(i)\mathbf{w}(i-1)] \quad \times \quad [\,\gamma\,\|\mathbf{e}(i)\|^2 - |\mathbf{x}(i)\,\mathbf{w}(i-1)|^2\,] \qquad (16)$$

If, on the other hand, we apply the minimization procedure

$$\min_{\mathbf{w}} \quad E\,[\,(\gamma - \delta\,|\,\mathbf{xw}\,|^2\,)^2\,]$$

we obtain the ratio-constant-modulus algorithm:

$$\mathbf{w}(i) = \mathbf{w}(i-1) + [\,\mu'\,\mathbf{x}^H(i)\,\mathbf{x}(i)\,\mathbf{w}(i-1)\,] \quad \times \quad [\,\gamma - \delta\,|\mathbf{x}(i)\,\mathbf{w}(i-1)|^2\,] \qquad (17)$$

where $\mu' = 2\mu\delta$.

In our previous studies [12,13], we have shown that an LMS-based algorithm with an *ε–error normalized* step-size parameter

$$\mu(k) = \frac{\mu}{\varepsilon + \|\,\mathbf{e}(k)\,\|^2}\,, \qquad (18)$$

and that with a *modified ε–error normalized* step-size parameter

$$\mu(k) = \frac{\mu}{1 + \mu\,\|\,\mathbf{e}(k)\,\|^2}\,, \qquad (19)$$

perform much better than other LMS algorithms under diverse signal to noise conditions.

5 Simulations Results

The *error normalized* LMS algorithm (4), the *ratio-error-data* LMS algorithm (13), and the NLMS algorithms were applied to an adaptive noise canceller shown in Fig.1.

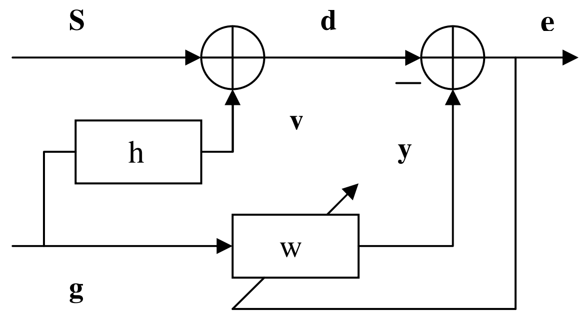

Fig. 1: Typical adaptive noise canceller (ANC).

A typical ANC is composed of two inputs: primary input and reference input. The primary input signal **d** consists of the original speech, **S**, corrupted by an additive noise **v**. The noise source is represented by **g**, and the transmission path from the noise source to the primary input is represented by the low pass filter, h. The input to the adaptive filter is the reference noise signal **g** that is correlated with **v**, but uncorrelated with **S**. The filter weights **w** are adapted by means of an LMS-based algorithm to minimize the power in the output signal. This minimization is achieved by processing **g** via the adaptive filter to provide an estimate of **v**, ($\mathbf{y} = \hat{\mathbf{v}}$), and then subtracting it from **d** to get **e** which represents an estimate of the original speech **S**.

Simulation results of the three algorithms are shown in Table 1. The same value of the step-size (μ=0.1) was used in all the algorithms to achieve a compromise between small excess mean-square error (EMSE) and high initial rate of convergence for a wide range of noise variances. In the *ratio-error-data* LMS algorithm (13), we used α=0.7 and γ=0.3 ($\alpha+\gamma$=1), and an error vector with increasing length was used. The order of the adaptive filter is assumed to be N=12.

Simulations were carried out using a male native

Table 1: Comparison of the steady-state EMSE and Misadjustment (M) of the three examined algorithms.

Stationary White noise **g** mean = 0	The NLMS algorithm		The *error normalized* LMS algorithm		The *ratio-error-data* LMS algorithm	
	$EMSE_{ss}$ (dB)	M %	$EMSE_{ss}$ (dB)	M %	$EMSE_{ss}$ (dB)	M %
$\sigma_g^2 = 0.001$	–30.70	8.46	–47.26	0.19	–50.96	0.08
$\sigma_g^2 = 0.01$	–30.70	8.46	–42.49	0.56	–43.76	0.42
$\sigma_g^2 = 0.1$	–30.70	8.46	–43.20	0.48	–44.34	0.37

speech sampled at a frequency of 11.025 kHz. The number of bits per sample is 8 and the total number of samples is 33000 or 3 sec of real time. The simulation results are presented for stationary noise environments in which the noise ***g*** was assumed to be zero mean white Gaussian with three different variances. The steady-state excess mean-square error (EMSE) and the misadjustment (M), [9], of the three algorithms are shown in Table1.

Figure 2 compares the performance of the proposed algorithms with that of the NLMS for the case when $\sigma_g^2 = 0.01$. The figure shows plots of the EMSE in dB for that noise level of the three algorithms. Performance improvement of the proposed algorithms over the NLMS is clear. Figure 3 demonstrates the superiority of the proposed algorithms by plotting the excess (residual) error (S–e) of the three algorithms. It was also confirmed by listening tests which demonstrated higher quality of the recovered speech and less signal distortion and reverberation than that if the NLMS is used. However, the *ratio-error-data* LMS algorithm, (13), shows a slight performance enhancement over the *error normalized* LMS algorithm, (4).

6 Conclusion

Several new least mean-square (LMS) algorithms based on error normalization procedure were introduced. Computer simulations of some of these proposed algorithms, carried out using a real speech in an adaptive noise canceller setup, were also shown. The performance of the proposed algorithms outperforms that of the standard NLMS algorithm by achieving less signal distortion, reverberation and excess mean-square error.

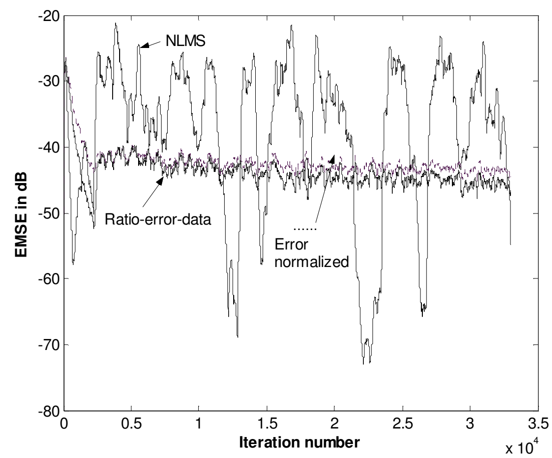

Fig. 2: EMSE of the three examined algorithms ($\sigma_g^2 = 0.01$, Table 1).

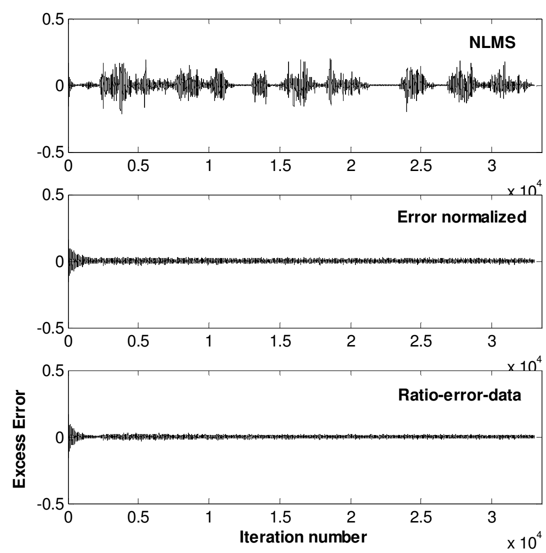

Fig. 3: Excess error (S–e) of the three examined algorithms ($\sigma_g^2 = 0.01$, Table 1).

References:

[1] B. Farhang-Boroujeny, "On statistical efficiency of the LMS algorithm in system modeling", *IEEE Trans. Signal Process*. vol. 41, pp. 1947-1951, 1993.

[2] M. Mboup, M. Bonnet, and N. Bershad, "LMS coupled adaptive prediction and system identification: A statistical model and transient mean analysis", *IEEE Trans. Signal Process.* vol.42, pp 2607-2615, 1994.

[3] J. Makhoul, "Linear prediction: A tutorial review", *Proc. IEEE*, vol. 63, pp. 561-580, 1975.

[4] P. Prandoni and M. Vetterli, " An FIR cascade structure for adaptive linear prediction", *IEEE Trans. Signal Process*., vol 46, pp 2566-2571, 1998.

[5] S. Ikeda and A. Sugiyama, "An adaptive noise canceller with low signal distortion for speech codecs," *IEEE Trans. on Signal Processing,* vol. 47, no. 3, pp 665- 674, March 1999.

[6] J. E. Greenberg " Modified LMS algorithms for speech processing with an adaptive noise canceller," *IEEE Trans. on Speech and Audio Processing*, vol. 6, no. 4, July 1998.

[7] L. C. Crodara, "Improved LMS algorithm for adaptive beam forming", *IEEE Trans. Antenna Propag*., vol. 38, pp 1631-1635, 1990.

[8] S. Qureshi, "Adaptive equalization data transmissions", *IEEE Commun. Mag*., vol. 20, pp. 2-16, 1982.

[9] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1985.

[10] B. Farhang-Boroujeny, *Adaptive Filters Theory and Applications*, John Wiley & Sons, New York, N.Y. 1998. .

[11] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley Interscience, Hoboken, N.J., 2003.

[12] Z. Ramadan and A. Poularikas, " Performance analysis of a new variable step-size LMS algorithm with error nonlinearities", *Proc. of the 36th IEEE Southeastern Symposium on System Theory (SSST)*, March 2004, pp 384–388.

[13] Z. Ramadan and A. Poularikas, "An Adaptive Noise Canceller Using Error Nonlinearities in the LMS Adaptation" *Proc. of the IEEE SoutheastCon*, March 2004, pp 359–364.