CONVERGENCE PROPERTIES OF MIXED-NORM ALGORITHMS UNDER GENERAL ERROR CRITERIA

Tareq Y. Al-Naffouri
Dept. of Electrical Eng.
Stanford University
Stanford, CA 94305
U.S.A

Azzedine Zerguine
Dept. of Physics
KFUPM
Dhahran 31261
Saudi Arabia

Maamar Bettayeb
Dept. of Electrical Eng.
KFUPM
Dhahran 31261
Saudi Arabia

Abstract

The convergence properties of mixed-norm algorithms as it applies to echo cancelers under general error criteria is derived for correlated and identically distributed inputs. The convergence analysis of this class of algorithms is carried out using the linearization process of the error nonlinearities. Necessary and sufficient conditions for convergence are derived for the independent input case.

1 Introduction

This work presents the convergence properties of the class of least-mean mixed-norms (LMMN) algorithms. The LMMN algorithm was first introduced in [1] as an application to data echo cancelers with two sections. The motivation behind using two norms for the two sections is that echoes in telephone circuits consist of two distinct components, a near echo and a far one, with completely different characteristics. It thus seems reasonable to use different norm criteria to cancel each echo. The LMMN algorithm is depicted in Fig. 1.

This work will summarize the convergence characteristics of this class of algorithms for both correlated and iid inputs. It will give the conditions under which these algorithms will be well-behaved (i.e. nondivergent). These conditions also serve as a measure of the performance of this class of algorithms.

2 The LMMN algorithm

The LMMN algorithm is described by two sets of the weight-error vectors $\boldsymbol{V}_N(k) = \boldsymbol{W}_N(k) - \boldsymbol{W}_{No}$ and $\boldsymbol{V}_F(k) = \boldsymbol{W}_F(k) - \boldsymbol{W}_{Fo}$ defined by

$$\begin{aligned} \boldsymbol{V}_N(k+1) &= \boldsymbol{V}_N(k) + \mu_N f(e(k))\boldsymbol{X}_N(k) \quad (1) \\ \boldsymbol{V}_F(k+1) &= \boldsymbol{V}_F(k) + \mu_F g(e(k))\boldsymbol{X}_F(k), \quad (2) \end{aligned}$$

where $f(e(k))$ and $g(e(k))$ are the error nonlinearities, $\boldsymbol{W}_{No}$ and $\boldsymbol{W}_{Fo}$ are the true impulse responses of the near-end and far-end sections, respectively, and where μ_N and μ_F are the step sizes of the near-end and far-end sections, respectively. The error is defined by

$$e(k) = n(k) - \boldsymbol{V}_N^t(k)\boldsymbol{X}_N(k) - \boldsymbol{V}_F^t(k)\boldsymbol{X}_F(k), \quad (3)$$

where $n(k)$ is the additive noise.

3 Convergence analysis of the LMMN algorithm

The convergence properties of the LMMN algorithm is carried out using the linearization concept. This linearization idea was first used in [2] in studying the (single-norm) class of algorithms, although a similar idea appeared in [3]. With this linearization, the adaptation equations (1) and (2) become:

$$\begin{aligned} \boldsymbol{V}_N(k+1) &= \boldsymbol{V}_N(k) + \mu_N[f(n(k)) \\ &\quad - f'(n(k))\left(\boldsymbol{V}^t(k)\boldsymbol{X}(k)\right) \\ &\quad + \frac{1}{2}f''(n(k))\left(\boldsymbol{V}^t(k)\boldsymbol{X}(k)\right)^2]\boldsymbol{X}_N(k) \quad (4) \end{aligned}$$

$$\begin{aligned} \boldsymbol{V}_F(k+1) &= \boldsymbol{V}_F(k) + \mu_F[g(n(k)) \\ &\quad - g'(n(k))\left(\boldsymbol{V}^t(k)\boldsymbol{X}(k)\right) \\ &\quad + \frac{1}{2}g''(n(k))\left(\boldsymbol{V}^t(k)\boldsymbol{X}(k)\right)^2]\boldsymbol{X}_F(k) \quad (5) \end{aligned}$$

The convergence analysis now boils down to studying the convergence of (4) and (5) in the mean and in the mean-square sense. Table 1 summarizes the main parameters in the convergence analysis of the LMMN algorithm for a correlated input [4]. Similarly, Table 2 summarizes the main parameters in the convergence analysis of the

0-7803-5471-0/99/$10.00©1999 IEEE

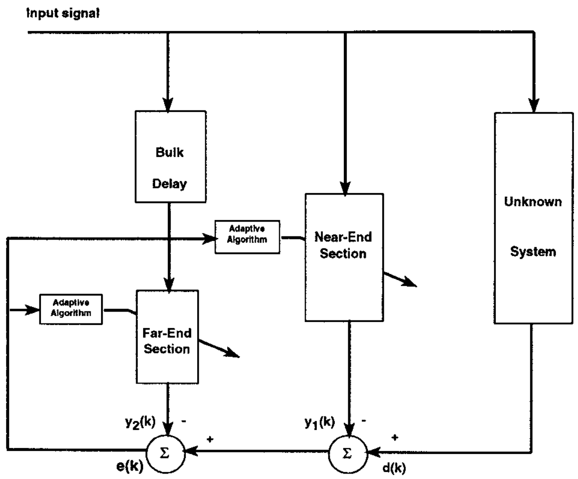

Figure 1: Echo canceler with two sections.

LMMN algorithm for an i.i.d input [4]. Due to space limitations we shall focus on the mean-square convergence of the LMMN algorithm for an iid input. Detailed convergence analysis for the correlated input case can be found in [4].

4 Analysis of the LMMN algorithm in the mean-square sense (iid input)

The mean-square behavior is completely determined by the behavior of the weight-error variances $\eta_N(k) = E[\boldsymbol{V}_N^t(k)\boldsymbol{V}_N(k)]$ and $\eta_F(k) = E[\boldsymbol{V}_F^t(k)\boldsymbol{V}_F(k)]$. They in turn satisfy the following recursion: [4]

$$\begin{bmatrix} \eta_N(k+1) \\ \eta_F(k+1) \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \eta_N(k) \\ \eta_F(k) \end{bmatrix} + \boldsymbol{B}, \quad (6)$$

where

$$\boldsymbol{A} = \begin{bmatrix} 1 - a_1\mu_N + (b_1+c_1)\mu_N^2 & c_1\mu_N^2 \\ c_2\mu_F^2 & 1 - a_2\mu_F + (b_2+c_2)\mu_F^2 \end{bmatrix},$$

and

$$\boldsymbol{B} = \begin{bmatrix} \mu_N^2 d1 \\ \mu_F^2 d2 \end{bmatrix}.$$

The constants a_1,b_1, and c_1 are defined by

$$\begin{aligned} a_1 &= 2\sigma_x^2 E[f'], & (7) \\ b_1 &= (m_{x,4} - \sigma_x^4)E[f'^2 + ff''], & (8) \\ c_1 &= \sigma_x^4 L_N E[f'^2 + ff''], & (9) \end{aligned}$$

where the expectation is done with respect to the additive noise $n(k)$. a_2,b_2, and c_2 are also defined by (7), (8), and(9), respectively, with f replaced by g and N replaced by F. Finally, $m_{x,4}$ is the fourth moment of x and L_N is the length of the near-end section. The mean-square behavior of the LMMN algorithm is completely determined by (6) as the illustrated in the following remarks:

1. The necessary and sufficient condition for the convergence of the matrix recursion (6) and hence for the mean-convergence convergence of the LMMN algorithm is that the eigenvalues of $\boldsymbol{A}$ be absolutely less than one. This will be the case if and only if

$$\frac{\left(\mu_N - \frac{a_1}{2(b_1+c_1)}\right)^2}{\frac{a_1^2(b_2+c_2)+a_2^2(b_1+c_1)}{4(b_1+c_1)^2(b_2+c_2)}} + \frac{\left(\mu_F - \frac{a_2}{2(b_2+c_2)}\right)^2}{\frac{a_1^2(b_2+c_2)+a_2^2(b_1+c_1)}{4(b_1+c_1)(b_2+c_2)^2}} \leq 1 \quad (10)$$

and

$$\begin{aligned} &\left(a_1\mu_N - (b_1+c_1)\mu_N^2\right)\left(a_2\mu_F - (b_2+c_2)\mu_F^2\right) \\ &-c_1c_2\mu_N^2\mu_F^2 \geq 0. \quad (11) \end{aligned}$$

2. Conditions (10) and (11) bear an interesting geometrical interpretation. Clearly, the first inequality means that μ_N and μ_F should be inside the ellipse with center

$$(\mu_{N_e}, \mu_{F_e}) = \left(\frac{a_1}{2(b_1+c_1)}, \frac{a_2}{2(b_2+c_2)}\right),$$

and radii $\left(\frac{a_1^2(b_2+c_2)+a_2^2(b_1+c_1)}{4(b_1+c_1)^2(b_2+c_2)}\right)^{1/2}$ and $\left(\frac{a_1^2(b_2+c_2)+a_2^2(b_1+c_1)}{4(b_1+c_1)(b_2+c_2)^2}\right)^{1/2}$. By a rotation of axes it is shown that the second inequality (11) restricts μ_N and μ_F to be within the hyperbola with axes along the lines $\mu_N = \pm\mu_F$, center

$$(\mu_{N_p}, \mu_{F_p}) = \left(\frac{a_1(b_2+c_2)}{b_1b_2+b_1c_2+b_2c_1}, \frac{a_2(b_1+c_1)}{b_1b_2+b_1c_2+b_2c_1}\right),$$

and radius $\frac{(b_1b_2+b_1c_2+b_2c_1)^2}{\sqrt{2a_1a_2c_1c_2}}$.

3. The convergence conditions (10) and (11) are difficult to check since μ_N and μ_F are coupled in these conditions. The inequalities

$$0 < \mu_N \leq \frac{a_1}{2(b_1+c_1)} \quad (12)$$

$$0 < \mu_F \leq \frac{a_2}{2(b_2+c_2)} \quad (13)$$

represent uncoupled sufficient conditions for convergence.

4. Assuming the convergence conditions (10) and (11) to be satisfied, we see from (6) that the vector $[\eta_N(k)\ \eta_F(k)]^t$ will decay to its steady state

$$\begin{bmatrix} \overline{\eta}_N \\ \overline{\eta}_F \end{bmatrix} = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} = \frac{1}{\Delta}\mathbf{D}\mathbf{B} \qquad (14)$$

where

$$\mathbf{D} = \begin{bmatrix} a_1\mu_N - (b_1+c_1)\mu_N^2 & c_1\mu_N^2 \\ c_2\mu_F^2 & a_2\mu_F - (b_2+c_2)\mu_F^2 \end{bmatrix},$$

where Δ is the determinant of $\mathbf{A}$. This in turn yields the misadjustment of the algorithm which is given by

$$M = \frac{\sigma_x^2}{\sigma_n^2}\left(\overline{\eta}_N + \overline{\eta}_F\right). \qquad (15)$$

5 Concluding remarks

1. Table 1 summarizes the convergence analysis results of the LMMN algorithm for the correlated input case. These results encompass the results of the LMMN algorithm considered in [1] for which $f(e(k)) = e(k)$ and $g(e(k)) = e^3(k)$, that is the LMS and the LMF algorithms, respectively [3].
2. When the two nonlinearities f and g coincide, the LMMN algorithm reduces to single-norm adaptation. In this special case, the results of Table 1 also encompass the results of [2] in which adaptation with a single and general error nonlinearity was considered. It is interesting to note that the single-norm results of [2] in turn describe to the first order many results under a specific choice of the (single) error nonlinearity.
3. The difficulty brought about by the correlated input assumption hindered our ability to see the natural coupling that exists between the near- and far-end sections. This coupling was preserved, however, under the simpler independent input assumption. This can be seen from Table 2 which summarizes the convergence analysis results for an independent input.
4. Notice that the correlated input analysis applies for an independent input as well. We can thus compare the correlated input analysis, as it applies to an independent input, with the independent input analysis. From Tables 1 and 2, it turns out that the correlated input results describe, to the first order and for sufficiently small step sizes, the results of the independent input case.
5. As remarked earlier, when the nonlinearities f and g coincide, the LMMN algorithm reduces to single-norm adaptation. If, in addition, the input is taken to be independent and Gaussian, then our independent input analysis is in agreement with the conditional analysis of [2]. In particular, the mean and mean-square recursions and represent the steady state recursions of those derived in [2]. Thus, the independent input analysis can be thought of as a generalization of the conditional analysis of [2] which is limited to Gaussian inputs.
6. The convergence analysis that is carried out here can be used to arrive at the optimum choice of the nonlinearities f and g; a choice that guarantees fastest convergence for a given steady-state error. Details can be found in [4].

6 Conclusion

The least-mean mixed norm (LMMN) algorithm has recently been proposed for long data echo canceler. This work summarizes the convergence behavior of the LMMN algorithm with a pair of general error nonlinearities and for correlated as well as iid inputs. This convergence study sets up the stage for obtaining the optimum error nonlinearity, a step that is carried out in [4].

Acknowledgment: The authors would like to acknowledge King Fahd University of Petroleum and Minerals for its support.

References

[1] A. Zerguine, C. F. N. Cowan, and M. Bettayeb, "Adaptive Echo Cancellation Using Least Mean Mixed-Norm Algorithm," *IEEE Trans. Signal Processing, vol.* SP-45, No. 5, pp. 1340-1343, May, 1997.

[2] S. C. Douglas and T. H. Y Meng, "Stochastic Gradient Adaptation under General Error Criteria", *IEEE Trans. Signal Processing*, vol. SP-42, no. 6, pp. 1352–1365, June 1994.

[3] E. Walach and B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and its Family", *IEEE Trans. Inf. Theory*, vol. IT-30, no. 2, pp. 275–283, Feb. 1984.

[4] T. Y. Al-Naffouri, *Adaptive Filtering Using the Least-Mean Mixed-Norms Algorithm and its Application to Echo Cancellation*, Master thesis, King Fahd University of Petroleum and Minerals, June 1997.

[5] T. Y. Al-Naffouri, A. Zerguine, and M. Bettayeb, "A Unifying View of Error Nonlinearities In LMS Adaptation," in Proc. *ICASSP*, Seattle, USA, May 12-15, 1998.

	Near-end section	Far-end section
Conditions for convergence of the mean	$0 < \mu_N < \frac{2}{\lambda_{N,max} E[f'(n(k))]}$	$0 < \mu_F < \frac{2}{\lambda_{F,max} E[g'(n(k))]}$
Time constants	$\tau_{N,i} = \frac{1}{\mu_N \lambda_{N,i} E[f'(n(k))]}$	$\tau_{F,j} = \frac{1}{\mu_F \lambda_{F,j} E[g'(n(k))]}$
Conditions for convergence in the mean-square	$0 < \mu_N < \frac{1}{\lambda_{N,max} E[f'(n(k))]}$	$0 < \mu_F < \frac{1}{\lambda_{F,max} E[g'(n(k))]}$
Misadjustment	$M_N = \mu_N \frac{\sigma_x^2}{\sigma_n^2} \frac{E[f^2(n(k))]}{2E[f'(n(k))]} L_N$	$M_F = \mu_F \frac{\sigma_x^2}{\sigma_n^2} \frac{E[g^2(n(k))]}{2E[g'(n(k))]} L_F$

Table 1: Summary of the convergence analysis results for the LMMN algorithm with a correlated input.

	Near-end section	Far-end section
Conditions for convergence of the mean	$0 < \mu_N < \frac{2}{\sigma_x^2 E[f'(n(k))]}$	$0 < \mu_F < \frac{2}{\sigma_x^2 E[g'(n(k))]}$
Time constants	$\tau_N = \frac{1}{\mu_N \sigma_x^2 E[f'(n(k))]}$	$\tau_F = \frac{1}{\mu_F \sigma_x^2 E[g'(n(k))]}$
Conditions for convergence in the mean-square (necessary & sufficient)	$a_1\mu_N - (b_1 + c_1)\mu_N^2 + a_2\mu_F - (b_2 + c_2)\mu_F^2 \geq 0$ $\left(a_1\mu_N - (b_1 + c_1)\mu_N^2\right)\left(a_2\mu_F - (b_2 + c_2)\mu_F^2\right) - c_1 c_2 \mu_N^2 \mu_F^2 \geq 0$	
Conditions for convergence in the mean-square (only sufficient)	$0 < \mu_N \leq \frac{a_1}{2(b_1 + c_1)}$	$0 < \mu_F \leq \frac{a_2}{2(b_2 + c_2)}$
Misadjustment	$\begin{bmatrix} M_N \\ M_F \end{bmatrix} = \frac{1}{\Delta} \frac{\sigma_x^2}{\sigma_n^2} \begin{bmatrix} a_1\mu_N - (b_1 + c_1)\mu_N^2 & c_1\mu_N^2 \\ c_2\mu_F^2 & a_2\mu_F - (b_2 + c_2)\mu_F^2 \end{bmatrix} \begin{bmatrix} \mu_N^2 d_1 \\ \mu_F^2 d_2 \end{bmatrix}$	

Table 2: Summary of the convergence analysis results for the LMMN algorithm with an independent input.