The Journal of Natural Science,
Pai Chai University, Korea
Vol.8 No2: 63~67, 1996

GENEARL CONVERGENCE ANALYSIS OF THE LVCMS ALGORITHM

Seung-Hyon Nam and Yong-Hoh Kim
* Department of Electronic Engineering, Pai Chai University

LVCMS 알고리즘에 대한 일반적인 수렴 특성 분석

남승현, 김용호
배재대학교 전자공학과

Adaptive algorithms based on the higher order error criterion such as the LVCMS and the LMF show performance degradation if input signal containes additive noise with a heavier-tailed density. Conventional analysis often neglects higher order terms in the recusion and may not suit for predicting exact behavior of these higher order algorithms. This paper presents a new convergence analysis which containes all the higher order terms in the recursion. The analysis shows that the higher order terms, which are often neglected, does not affect the upper bound on the step size but the misadjustment. However, the effect decreases sharply proportional to the square of the step size.

LVCMS나 LMF와 같이 고차의 통계에 근거한 적응 신호 알고리즘들은 입력신호에 꼬리가 두툼한 확률밀도를 갖는 잡음이 섞여있는 경우 현저한 성능 저하를 보여준다. 일반적인 알고리즘 분석 방법에서는 순환식에서 2차 보다 큰 power를 갖는 항들을 모두 무시하기 때문에 정확한 알고리즘의 성능을 예측하지 못할 수도 있다. 이 논문에서는 새로운 분석방법을 통하여 무시된 항들의 영향이 무엇인가를 알아본다. 분석 결과 무시된 항들은 스텝 크기의 상한선에는 영향을 미치지 못하나 misadjustment를 증가시킬 수 있는 것으로 나타났다. 그러나 그 영향은 스텝 크기의 제곱에 비례하여 감소한다.

Key words : adaptive algorithm, LVCMS, higher order criterion, convergence, misadjustment

I. INTRODUCTION

The least variance subject to a constraint on the mean square error (LVCMS) adaptive filtering algorithm was introduced by Gibson and Gray in [1]. The LVCMS algorithm is motivated by the steepest descent method like the LMS and LMF algorithms [2,3]. In [1], the convergence in the mean coefficient error of the LVCMS algorithm was analyzed without assumptions on the density function of the input data as in [2] for the LMS algorithm and in [3] for the LMF algorithm. The resulting upper bound on the step size parameter is quite loose and the actual step size should be chosen much smaller than the upper bound to ensure the convergence of the MSE in practice. This is more evident if additive noise has a heavier-tailed density. Recently, the LMS, LMF, and LVCMS algorithms are considered as special cases of a more general adaptive algorithm [2], and the convergence analysis of the general algorithm is given [4]. The resulting upper bound is much tigher than the previous one by a factor of more than three. However, the analysis neglected the terms higher than the power of two in the recusion. This paper presents a new convergence analysis which containes all the higher order terms in the recusion, and investigate its effect on the upper bound on the step size and the misadjustment values.

II. THE LVCMS ALGORITHM AND THE GENERALIZED ERROR CRITERION

The adaptive signal processing configuration of interest is depicted in Fig. 1. The input data vector at time k is given by $X(k)=[x_1(k), x_2(k), \cdots, x_N(k)]^T$, and $X(k), k=0,1,2,\cdots$, are assumed to be uncorrelated. The error signal at time k is given by

$$\begin{aligned} \varepsilon(k) &= d(k) - W^T(k)X(k) \\ &= n(k) - (W(k) - W^*)^T X(k) \\ &= n(k) - V^T(k)X(k) \end{aligned} \quad (1)$$

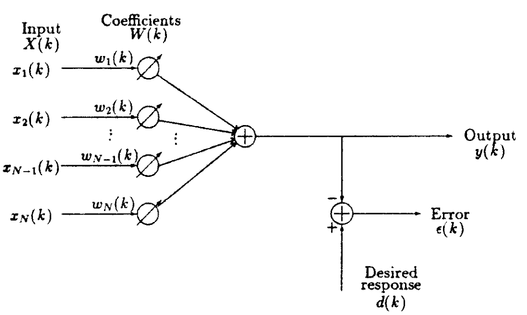

Fig 1. Adaptive Signal Processing Configuration.

where W^* is the vector of optimal coefficients, $W(k)$ is the coefficient vector at time k so that

$$V(k) = W(k) - W^* \quad (2)$$

is the coefficient error vector at time k and $n(k)$ is the noise such that

$$n(k) = d(k) - X^T(k)W^* \quad (3)$$

The noise $n(k)$ is assumed to be white and to have a symmetrical probability density function with zero mean and finite higher order moments. Further, the noise $n(k)$ is assumed to be independent of the input data vector $X(k)$.

The LVCMS algorithm minimizes $E\{[\varepsilon^2(k) - E[\varepsilon^2(k)]]^2\}$ subject to a constraint on the MSE $E[\varepsilon^2(k)] = \sigma_o^2$ and has the adaptation rule [1]

$$\begin{aligned} W(k+1) = \; & W(k) + 4\mu\varepsilon(k)X(k) \\ & -2\mu(2\sigma_o^2 + \lambda)\varepsilon(k)X(k) \end{aligned} \quad (4)$$

where λ is a Lagrange multiplier which is nonpositive.

Now we consider the general error criterion considered in [5] is

$$\begin{aligned} H_g(W, a, b, c, d) = \; & aE[\varepsilon^2(k) - E(\varepsilon^2(k))]^2 \\ & + bE[\varepsilon^2(k)]^2 \\ & + cE[\varepsilon^2(k)] \\ & + d. \end{aligned} \quad (5)$$

Taking partial derivative of $H_g(W, a, b, c, d)$ with respect to W yields the instantaneous gradient-based coefficient adaptation rule is given by

$$\begin{aligned} W(k+1) = \; & W(k) + 4\mu b\varepsilon^3(k)X(k) \\ & + 2\mu c\varepsilon(k)X(k). \end{aligned} \quad (6)$$

We should note that only b and c are included in the generalized algorithm (5) since the instantaneous estimate of the gradient is used. The generalized algorithm in (5) corresponds to the LMS algorithm if $(b, c) = (0, 1)$, the LMF algorithm if $(b, c) = (1, 0)$, and the LVCMS algorithm if $(b, c) = (1, -(2\sigma_o^2+\lambda))$.. In the following analyses, the general adaptation rule (6) is used although the result does not apply to the LMF algorithm.

III. GENERAL CONVERGENCE ANALYSIS

We now assume that the input data $X(k)$ is Gaussian and uncorrelated in time, so that it is also independent. Therefore, we can apply the independence assumptions widely used in stochastic analyses of the LMS type algorithm [6].

Subtract W^* from both sides of (6) to obtain the recursion for $V(k)$,

$$\begin{aligned} V(k+1) = {} & V(k) + 4\mu b \varepsilon^3(k) X(k) \\ & + 2\mu c \varepsilon(k) X(k), \end{aligned} \tag{7}$$

and then use (1) for $\varepsilon(k)$ to obtain

$$\begin{aligned} V(k+1) = {} & V(k) + 4\mu b X(k) \sum_{i=0}^{3} \binom{3}{i} n^i(k) \\ & (-X^T(k) V(k))^{3-i} \\ & + 2\mu c (n(k) - X^T(k) V(k)) X(k). \end{aligned} \tag{8}$$

Since $R = E[X(k) X^T(k)]$ is symmetric, one can define the unitary matrix U as

$$E\{UX(k)X^T(k)U^T\} = \Gamma = diag(\gamma_1, \gamma_2, \cdots, \gamma_N). \tag{9}$$

Then

$$\begin{aligned} & UV(k+1)V^T(k+1)U^T = UV(k)V^T(k)U^T \\ & + 16\mu^2 b^2 \left\{ \sum_{i=0}^{3} \binom{3}{i} n^i(k) (-X^T(k)V(k))^{3-i} \right\}^2 \\ & \quad UX(k)X^T(k)U^T \\ & + 4\mu^2 c^2 [n(k) - X^T(k)V(k)]^2 \\ & \quad UX(k)X^T(k)U^T \\ & + 4\mu b \left\{ \sum_{i=0}^{3} \binom{3}{i} n^i(k) (-X^T(k)V(k))^{3-i} \right\} \\ & \quad \{UX(k)V^T(k)U^T + UV(k)X^T(k)U^T\} \\ & + 2\mu c [n(k) - X^T(k)V(k)] \\ & \quad \{UX(k)V^T(k)U^T + UV(k)X^T(k)U^T\} \\ & + 16\mu^2 bc \left\{ \sum_{i=0}^{3} \binom{3}{i} n^i(k) (-X^T(k)V(k))^{3-i} \right\} \\ & \quad [n(k) - X^T(k)V(k)] UX(k)X^T(k)U^T. \end{aligned} \tag{10}$$

Assuming that $V(k)$ lies inside a certain bounded domain Ω around the optimal coefficient vector W^*, that is, $V(k) \in \Omega$. Then taking the conditional expectation of $UV(k)V^T(k)U^T$ in (10) with respect to $V(k)$ gives

$$\begin{aligned} & E_{V_k}\{UV(k+1)V^T(k+1)U^T\} = UV(k)V^T(k)U^T \\ & - 2\mu \bar{a} E\{X^T(k)V(k)[UX(k)V^T(k)U^T \\ & + UV(k)X^T(k)U^T]\} \\ & + 4\mu^2 \bar{b} E\{(X^T(k)V(k))^2 UX(k)X^T(k)U^T\} \\ & + 4\mu^2 \bar{c} E\{UX(k)X^T(k)U^T\} \\ & + 16\mu^2 \bar{d} E\{(X^T(k)V(k))^4 UX(k)X^T(k)U^T\} \\ & + 16\mu^2 b^2 E\{(X^T(k)V(k))^6 UX(k)X^T(k)U^T\} \\ & - 4\mu b E\{(X^T(k)V(k))^3 [UX(k)V^T(k)U^T \\ & + UV(k)X^T(k)U^T]\}. \end{aligned} \tag{11}$$

where $E_{V_k}[\ ]$ denotes the conditional expectation and

$$\begin{aligned} \bar{a} &= 6bE(n^2(k)) + c \\ \bar{b} &= 60b^2E(n^4(k)) + 24bcE(n^2(k)) + c^2 \\ \bar{c} &= 4b^2E(n^6(k)) + 4bcE(n^4(k)) \\ & \quad + c^2E(n^2(k)) \\ \bar{d} &= 15b^2E(n^2(k)) + bc. \end{aligned} \tag{12}$$

In obtaining (11), assumptions that the $n(k)$ is independent of the input vector $X(k)$ and has zero mean are used. Note that the traditional convergence analysis neglects terms higher than the power of 2 in (11).

Let

$$\tilde{C}(k) = UV(k)V^T(k)U^T. \tag{13}$$

It can be shown that

$$E_{V_k}\{X^T(k)V(k)[UX(k)V^T(k)U^T + UV(k)X^T(k)U^T]\} = \Gamma\tilde{C}(k) + \tilde{C}(k)\Gamma. \quad (14)$$

and using the fourth moment expansion for Gaussian random variables, that

$$E_{V_k}\{(X^T(k)V(k))^2 UV(k)V^T(k)U^T\} = 2\Gamma\tilde{C}(k)\Gamma + tr[\Gamma\tilde{C}(k)]\Gamma. \quad (15)$$

After substituting (14) and (15) into (11), we may obtain the recursion

$$\begin{aligned} E_{V_k}[\tilde{C}(k+1)] = & \tilde{C}(k) - 2\mu\bar{a}[\Gamma\tilde{C} + \tilde{C}\Gamma] \\ & +4\mu^2\bar{b}[2\Gamma\tilde{C}\Gamma + tr(\Gamma\tilde{C})\Gamma] \\ & +4\mu^2\bar{c}\Gamma + 16\mu^2\bar{d}B_1(k) \\ & +16\mu^2 B_2(k) - 4\mu B_3(k) \end{aligned} \quad (16)$$

where

$$\begin{aligned} B_1 &= E_{V_k}\{(X^T(k)V(k))^4 UX(k)X^T(k)U^T\} \\ B_2 &= E_{V_k}\{(X^T(k)V(k))^6 UX(k)X^T(k)U^T\} \\ B_3 &= E_{V_k}\{(X^T(k)V(k))^3[UX(k)V^T(k)U^T \\ & \quad + UV(k)X^T(k)U^T]\}. \end{aligned} \quad (17)$$

Note that matrices $B_1(k), B_2(k),$ and $B_3(k)$ are positive definite if R is positive definite. Hence, those are bounded above by some positive definite matrices for $V(k) \in \Omega$. Hence, (16) becomes

$$\begin{aligned} E_{V_k}[\tilde{C}(k+1)] \le & \tilde{C}(k) - 2\mu\bar{a}[\Gamma\tilde{C} + \tilde{C}\Gamma] \\ & +4\mu^2\bar{b}[2\Gamma\tilde{C}\Gamma + tr(\Gamma\tilde{C})\Gamma] \\ & +4\mu^2\bar{c}\Gamma + 16\mu^2 B \end{aligned} \quad (18)$$

where

$$dB_1(k) + B_2(k) \le B \qquad \text{for } V(k) \in \Omega \quad (19)$$

for some constant B. Since the recursion (18) holds for all $V(k) \in \Omega$, one can average it over all possible values of $V(k) \in \Omega$. Hence, one has

$$\begin{aligned} \bar{C}(k+1) \le & \bar{C}(k) - 2\mu\bar{a}[\Gamma\bar{C} + \bar{C}\Gamma] \\ & +4\mu^2\bar{b}[2\Gamma\bar{C}\Gamma + tr(\Gamma\bar{C})\Gamma] \\ & +4\mu^2\bar{c}\Gamma + 16\mu^2 B \end{aligned} \quad (20)$$

since $\bar{C}(k) = E[\tilde{C}(k)]$.

From [4], we have the recusion obtained by neglecting the higher order terms in recursion (11), which is rewritten as,

$$\begin{aligned} C(k+1) = & C(k) - 2\mu\bar{a}\{\Gamma C(k) + C(k)\Gamma\} \\ & +4\mu^2\bar{b}\{2\Gamma C(k)\Gamma + tr[\Gamma C(k)]\Gamma\} \\ & +4\mu^2\bar{c}\Gamma \end{aligned} \quad (21)$$

where

$$C(k) = E[UV(k)V^T(k)U^T]. \quad (22)$$

Note that recursion (20) is identical to recursion (21) used in the convergence analysis except for the last constant term. The constant term does not affect the upper bound on the step size μ. To derive the expression for misadjustment, one should use recusion (21) instead of (20). Therefore, the resulting expression for misadjustment may not predict the actual misadjustment values effectively due to additional constant terms in (20). Effect of additional constant term on the misadjustment decreases sharply with μ since it is of second order in μ.

V. CONCLUSIONS

A general convergence analysis of the LVCMS algorithm has been presented under the assumption of uncorrelated Gaussian input data. The general error criterion, which admits the LMS and LMF criteria as well as the LVCMS criterion, was considered. The results so obtained is applied to the LMF, and LVCMS algorithms. The general convergence algorithm presented here include all the higher order terms in the recusion, which are often neglected for convience of analysis. The analysis shows that the upper bound on the step size does not affect by the neglected higher order terms. If the expression for the misadjustment is derived, the predicted value is smaller than the actual value due to the neglected higher order terms. The difference, however, will be proportional to the square of the step size and may be ignored for small step sizes. Therefore, it is reasonable to ignore the higher order terms in performance analysis of the adaptive algorithms which use higher order error criterion.

Acknowledgement

This study was financially supported by a central research fund in 1995 from Pai Chai University.

REFERENCES

1. J. D. Gibson and S. D. Gray, "MVSE adaptive filtering subject to a constraint on MSE," IEEE Trans. Circuits and Systems, vol. CAS-35, no. 5, pp. 603-608, May 1988.
2. B. Widrow, "Adaptive filters," in Aspects of Network and System Theory, R. Kalman and N. DeClaris, Eds., New York: Holt, Rinehart, and Winston, 1971, pp. 563-587.
3. E. Walach and B. Widrow, "The least mean fourth algorithm and its family," IEEE Trans. Inform. Theory, vol. IT-30, pp. 275-283, Mar. 1984.
4. Seung H. Nam and Insung Lee, "Convergence analysis of the LMF algorithm family for the uncorrelated Gaussian Data", The 2nd APCC, Osaka, Japan, pp. 545-548, June 1994.
5. J. D. Gibson and S. D. Gray, "MVSE adaptive filtering subject to a constraint on MSE," TCSL Res. Rep. no. 86-05-01, Dept. Elec. Eng. Texas A&M Univ., Aug. 5, 1986.
6. J. E. Mazo, "On the independence theory of equalizer convergence," Bell Sys. Tech. J., vol. 58, no. 5, pp. 963-993, May-June 1979.