

제2차 MLMS 알고리즘을 이용한 시스템 Identification

(System Identification Using the Second Order MLMS Algorithm)

金海正*, 李斗秀*

(Hae Jung Kim and Doo Soo Lee)

要約

본 논문에서는 스칼라 인수 α 및 β 를 매개변수화하여 갱신항을 첨가한 최소 평균 자승(LMS) 알고리즘의 특성을 해석하고 그 구조를 나타낸다. 수렴 특성의 해석에서 평균 중량벡터에 대하여 전이행렬의 복소수 고유값이 기술된다. 그 알고리즘의 안정한 범위도 증명된다. MADF, Sign, 및 기존의 LMS 알고리즘과 MLMS 알고리즘의 계산량도 비교해 본다. 시스템 identification에 적용한 결과 제2차 MLMS 알고리즘이 LMS 및 제1차 MLMS 알고리즘보다 계산량이 증가하는 반면 수렴속도에서 현저한 향상을 보여준다.

Abstract

This paper analyzes the properties of such algorithm that corresponds to the LMS algorithm with additional update terms, parameterized by the scalar factors α and β , and presents its structure. The analysis of convergence leads to complex eigenvalues of the transition matrix for the mean weight vector. Regions in which the algorithm becomes stable are demonstrated. The computational complexities of MLMS algorithms are compared with those of MADF, sign and the conventional LMS algorithms. In application of the system identification the second order momentum MLMS algorithm has faster convergence speed than LMS and the first order MLMS algorithms.

I. 서론

본 논문에서 제안된 제2차 MLMS 알고리즘(second order momentum least mean square algorithm) 혹은 제2차 필터화 경사도 알고리즘(quadratically filtered gradient algorithm)을 연구하게 된 동기는 기존의 최소평균자승(LMS) 알고리즘^[1] 보다 빠른 수렴 속도에 대한 요구에서 비롯되었다. 선형 필터화 경사도 알고리즘(linearly filtered gradient algorithm)은 디지털 통신

선로에 대한 고속 적응 등화기에 처음 도입되었다.^{[2][3][4]} 제2차 MLMS 알고리즘은 갱신된 중량 벡터를 얻기 위하여 매번 반복시 3개의 전 중량벡터가 결합되는 제3차 적응 알고리즘이다. 최근 제1차 MLMS 알고리즘의 비선형 형태를 사용하여 다층 순방향회환 신경망(multi-layer feed-forward neural networks)의 역전파 학습(back-propagation training) 알고리즘에서 수렴 특성을 향상시켰다.^[5]

제2차 MLMS 알고리즘은 제1차 MLMS 알고리즘^[6] 으로부터 미소량 $\beta[W(n-1) - W(n-2)]$ 을 부가하여 유도된다. 이 부가항이 수렴시 모멘텀 효과를 나타낸다. 그 이전의 중량 변화량을 현재의 갱신량에 더하기 때문에 중량 벡터에 대하여 보다 빠르고 스무딩한 수렴특성을 나

*正會員, 漢陽大學校 電子工學科
(Dept. of Elec. Eng., Hanyang Univ.)
接受日字: 1992年 5月 6日

타낸다. 과거의 중량벡터 $W(n-1)$ 및 $W(n-2)$ 의 기억 공간과 부가적인 곱셈으로 인하여 그 특성이 개선된다. 그러나 알고리즘이 최소값(global minimum)에 접근할 때 모멘텀 항이 성능을 저하시켜 誤調停量(misadjustment)을 증가시킬 수 있다. 이런 문제를 극복하기 위하여 중량 벡터의 변화량이 충분히 작을 경우 모멘텀 항을 제거하여 $\alpha=\beta=0$ 로 둘 수 있다.

본 논문에서는 [7][8][9]에서 기술된 방법을 이용하여 제2차 MLMS 알고리즘의 수렴 특성을 해석하고, 시스템 identification에 응용하여 [2][3][4][6]에서 제기된 성능상의 문제점을 해결한다. II장에서는 제2차 MLMS 알고리즘을 기술하고 이를 수렴특성의 해석에 적합한 형태로 나타낸다. III장에서는 평균 수렴 안정도 조건을 살펴보고, 평균 중량 순환식에서 복소수 고유값의 발생에 대하여 새로운 결과를 記述한다. 또한 MLMS의 산술적 연산량을 MADF(multiplication-free adaptive digital filter),^{[10][11]} sign 알고리즘,^[12] LMS 알고리즘의 연산량과 비교된다. IV장에서는 시스템 identification에 적용한 컴퓨터 모의실험(computer simulation)을 보여준다. 결론은 V장에서 요약된다.

II. OMLMS 알고리즘

제2차 모멘텀 최소평균자승(MLMS)은 다음과 같이 제3차 중량 갱신식으로 표현된다.

$$W(n+1) = W(n) + 2\mu e(n)X(n) + \alpha[W(n) - W(n-1)] + \beta[W(n-1) - W(n-2)], \quad |\alpha| < |\beta| < 1 \quad (1)$$

위식에서 $W(n)=[w_0(n), \dots, w_{N-1}(n)]^T$ 이다. 출력 오차는 다음 식으로 주어진다.

$$e(n) = d(n) - X^T(n)W(n) \quad (2)$$

위식에서 $X(n)=[x(n), \dots, x(n-N+1)]^T$ 은 입력 신호 벡터이고, $d(n)$ 은 理想應答이고 $y(n)=W^T(n)X(n)$ 은 $d(n)$ 의 추정값이다. 스텝 크기는 수렴 속도를 제어하고, 매개변수 α 와 β 는 경사도에 부가된 모멘텀 항의 계수이다. $\alpha=0$ 이고 $\beta=0$ 이면 LMS 알고리즘이 되고, $\beta=0$ 이면 제1차 MLMS이 된다. 알고리즘 특성상 $\alpha \rightarrow 1$, $\beta \rightarrow -1$ 일 때 저역통과 필터의 구조를 가지고, $\alpha \rightarrow -1$, $\beta \rightarrow -1$ 일 때 고역통과 필터의 구조를 가진다.

중량오차 벡터를 $V(n)=W(n) - W^*$ 라고 할 때, $W^* = R^{-1}P$ 은 최적 중량 벡터가 되고, 다음 식이 성립한다.

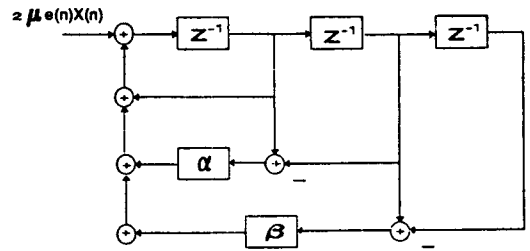


그림 1. 제2차 MLMS 알고리즘의 구조
Fig. 1. The structure of the second order MLMS algorithm.

$$V(n+1) - V(n) = \alpha[V(n) - V(n-1)] + \beta[V(n-1) - V(n-2)] + 2\mu e^*(n)X(n) - 2\mu X(n)X^T(n)V(n) \quad (3)$$

위식에서 최적 출력오차는 $e^*(n)=d(n) - W^{*T}X(n)$ 이다. 입력신호 상관행렬은 $R=E[X(n)X^T(n)]$ 이고, 고유벡터 분해 $R=Q\Lambda Q^T$ 가 성립하고, 대각선 행렬 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, N)$ 은 $R=Q\Lambda Q^T$ 의 고유벡터를 증가하는 순서 $0 < 1 = \lambda_{\min} \leq \lambda_2 \dots, N = \lambda_{\max}$ 로 배열한 것이다. 행렬 Q의 열은 특성 $QQ^T = I$ 을 만족하는 직교 고유 벡터이고, I는 단위행렬이다. 교차상관 벡터 P는 $P=E[X(n)d(n)]$ 이다.

회전 중량오차 벡터를 $\tilde{V}(n)=Q^T V(n)$ 이라고 두고, 식 (3)에 Q^T 를 곱하면 다음 식이 성립한다.

$$\tilde{V}(n+1) - \tilde{V}(n) = \alpha[\tilde{V}(n) - \tilde{V}(n-1)] + \beta[\tilde{V}(n-1) - \tilde{V}(n-2)] + 2\mu e^*(n)\tilde{X}(n) - 2\mu \tilde{X}(n)\tilde{X}^T(n)\tilde{V}(n) \quad (4)$$

위식에서 $\tilde{X}(n)$ 은 회전입력 벡터 $\tilde{X}(n)=Q^T X(n)$ 이다.

III. 수렴특성

1. 안정도 조건

식 (4)의 평균을 취하면, 다음 식이 성립한다.

$$E[\tilde{V}(n+1) - \tilde{V}(n)] = \alpha E[\tilde{V}(n) - \tilde{V}(n-1)] + \beta E[\tilde{V}(n-1) - \tilde{V}(n-2)] - 2\mu \Lambda E[\tilde{V}(n)] \quad (5)$$

위식에서 $E[\tilde{X}(n)\tilde{X}^T(n)]$ 은 고유값의 대각선 행

렬이다. 직교성 원리^[11]에 의해 $E[e^*(n)\tilde{X}(n)]=0$ 이 되고, $W(n)$ 과 $X(n)$ 은 독립적이다. 다음의 3N 차원 상태 벡터

$$\tilde{U}(n+1) = \begin{bmatrix} \tilde{V}(n+1) \\ \tilde{V}(n) \\ \tilde{V}(n-1) \end{bmatrix} \quad (6)$$

를 정의할 때, 다음의 순환식이 성립한다.

$$E[\tilde{U}(n+1)] = \begin{bmatrix} (1+\alpha)I - 2\mu\Lambda & (-\alpha+\beta)I & -\beta I \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} E[\tilde{U}(n)] \quad (7)$$

위식에서 블록 전이 행렬(bolck transition matrix)은 3N 차수를 가진다. 본 알고리즘의 안정도는 다음 행렬식

$$\det \begin{bmatrix} (1+\alpha-\zeta)I - 2\mu\Lambda & (-\alpha+\beta)I & -\beta I \\ I & -\zeta I & 0 \\ 0 & I & -\zeta I \end{bmatrix} = 0 \quad (8)$$

의 근 $\{\zeta\}$ 에 따라 영향을 받는다.

블록 행렬 A, B, C, D에 대하여 다음 식

$$\begin{aligned} \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \det[D] \det[A - BD^{-1}C] \\ &= \det[A] \det[D - CA^{-1}B] \end{aligned} \quad (9)$$

을 이용할 때, D^{-1} 이 존재하기 때문에 다음 특성 방정식을 유도할 수 있다.

$$\zeta^{2N} \det[-\zeta + (1+\alpha) - (\alpha-\beta)\zeta^{-1} - \beta\zeta^{-2}]I - 2\mu\Lambda = 0 \quad (10)$$

식 (3-4)의 3N 개의 근을 결정하기 위하여, 다음의 3차형

$$\zeta^3 - (1+\alpha - 2\mu\lambda_i)\zeta^2 + (\alpha-\beta)\zeta + \beta = 0 \quad (11)$$

을 조사할 필요가 있다. 수정된 Schur-Cohn 검정^[13]을 이용하여, 식 (3-7)의 안정도를 판정할 수 있으며, 다음의 안정도 조건이 유도된다.

$$\text{조건 1: } 0 < \mu < (1+\alpha-\beta)/\lambda_{\max}$$

$$\text{조건 2: } 0 < \mu < 1-\alpha^2 \leq 1$$

$$\text{조건 3: } |-\alpha(1+\beta) + 2\mu\lambda_i| < 1-\alpha^2$$

위 조건 3에서 $\alpha < 0$ 일 때, α 의 범위와 μ 의 범위는 다음과 같다.

$$\begin{aligned} -(1-\beta) + 2\mu\lambda_i/(1+\beta) &< \alpha < 0 \\ \mu &< (1-\alpha^2)/2\lambda_{\max} \end{aligned} \quad (12)$$

$\alpha \rightarrow 0$ 일 때, 안정도 한계는 $0 < \mu < 1/2\lambda_{\max}$ 로 증가한다. $\alpha \rightarrow -1$ 일 때, 안정도 한계는 0으로 감소한다. α 가 음수일 경우 알고리즘이 안정하더라도 근 성능을 저하시킨다. 즉, 전 중량 갱신항이 음의 방향에 있을 때($\alpha(W_i(n) - W_i(n-1)) < 0$) 그 중량 갱신량의 음의 분수를 그 다음의 중량 갱신항에 첨가하기 때문에 수렴 속도가 저하될 수 있다.

$\alpha > 0$ 및 $-\alpha(1+\beta) + 2\mu\lambda_i > 0$ 일 때, 다음의 부등식이 성립한다.

$$\begin{aligned} 0 < -(1-\beta) + 2\mu\lambda_i/(1+\beta) &< \alpha \\ (1-\beta)/2\lambda_{\min} &< \mu \end{aligned} \quad (13)$$

$\beta \rightarrow 0$ 일 때 안정도 한계는 $\min(1/2\lambda_{\min}, (1+\alpha)/\lambda_{\max}) < \mu < \max(1/2\lambda_{\min}, (1+\alpha)/\lambda_{\max})$ 로 감소하고, $\beta \rightarrow 1$ 일 때 안정도 한계는 $0 < \mu < \alpha/\lambda_{\max}$ 이 된다. $\beta \rightarrow 1$ 일 때 안정도 한계는 $1/\lambda_{\min} < \mu < (2+\alpha)/\lambda_{\max}$ 로 된다.

$\alpha > 0$ 및 $-\alpha(1+\beta) + 2\mu\lambda_i < 0$ 일 때, 다음의 부등식이 성립한다.

$$\begin{aligned} 0 < \alpha < 1-\beta + 2\mu\lambda_i/(1+\beta) \\ -(1-\beta^2)/2\lambda_{\min} &< \mu \end{aligned} \quad (14)$$

$\beta \rightarrow 0$ 일 때 안정도 한계는 $\min(1/2\lambda_{\min}, (1+\alpha)/\lambda_{\max}) < \mu < \max(1/2\lambda_{\min}, (1+\alpha)/\lambda_{\max})$ 이고, $\beta \rightarrow 1$ 일 때 안정도 한계는 $0 < \mu < \alpha/\lambda_{\max}$ 이 된다. $\beta \rightarrow 1$ 일 때 안정도 한계는 $0 < \mu < (2+\alpha)/\lambda_{\max}$ 로 된다. β 가 음수일 경우 알고리즘이 안정하더라도 그 성능을 저하시킨다.

LMS 알고리즘의 전이행렬은 대칭적이고, 그 고유값은 실수가 된다. 그러나 MLMS 알고리즘의 전이 행렬은 대칭적이 아니고, 그 고유값은 복소수가 된다.

2. 평균 순환식의 실근 및 허근

특성 방정식의 근이 실수가 되기 위하여 그 판별식이 음수가 되어서는 안된다.

$$\begin{aligned}
 &32\beta\lambda_i^3\mu^3 - 4(\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)\lambda_i^2\mu^2 \\
 &+ 4(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2)\lambda_i\mu \\
 &- \alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 \\
 &- 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta \leq 0 \tag{15}
 \end{aligned}$$

$\beta=0$ 일 때, 식 (15)에서 $-\alpha^2$ 으로 나누고 정리하면, 다음식과 같이 된다.

$$4\lambda_i^2\mu^2 - 4(1+\alpha)\lambda_i\mu + (1-\alpha)^2 \geq 0 \tag{16}$$

(1) $f(\mu)$ 의 2계 도함수

$$f''(\mu) = 192\beta\lambda_i^2\mu - 8(\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)\lambda_i^2 \tag{17}$$

의 근 $\mu^2 = (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2) / 24\lambda_i\beta$ 은 변곡점이 된다.

(2) $f(\mu)$ 은 μ_1, μ_3 에서 극값을 가진다.

i) $\beta > 0$ 일 때, $\mu_1 = (-c_1^2 - 4c_0c_2)^{1/2}$ 에서 극소점이 되고, $\mu^2 = (-c_1^2 + (c_1^2 - 4c_0c_2)^{1/2}) / 2c_0$ 에서 극대점이 된다.

ii) $\beta < 0$ 일 때, $\mu_1 = -(c_1 - (c_1^2 - 4c_0c_2)^{1/2}) / 2c_0$ 에서 극대점이 되고, $\mu^2 = -(c_1 + (c_1^2 - 4c_0c_2)^{1/2}) / 2c_0$ 에서 극소점이 된다.

이 때 c_0, c_1, c_2 의 값은 다음과 같다.

$$\begin{aligned}
 c_0 &= 96\beta\lambda_i^3 \\
 c_1 &= -8(\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)\lambda_i^2 \\
 c_2 &= 4(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2)\lambda_i \tag{18}
 \end{aligned}$$

(3) $f(\mu)$ 는 다음과 같은 경우에 양수가 되지 않는다. 즉, 특성 방정식이 실근을 갖는다.

i) $\beta > 0$ 일 때, $\mu < \mu_{4,1}$ 혹은 $\mu_{4,2} < \mu < \mu_{4,3}$

ii) $\beta < 0$ 일 때, $\mu_{4,1} < \mu < \mu_{4,2}$ 혹은 $\mu_{4,3} < \mu$

특성 방정식의 근 $\mu_{4,1}, \mu_{4,2}, \mu_{4,3}$ 는 부록의 식 (A-9)에 유도되어 있다. 앞의 결과에서 알고리즘의 i 번째 모드의 안정도에 대한 조건은 다음과 같다.

$$0 < \mu < \mu_{4,4} = (1 + \alpha - \beta) / \lambda_i \tag{19}$$

$\beta > 0$ 일 때, 특성 방정식은 다음 범위에서 실근을 갖는다.

$$(0 < \mu < \mu_{4,1}) \cup (\mu_{4,2} < \mu < \mu_{4,3}) \tag{20}$$

$\beta < 0$ 일 때, 특성 방정식은 다음 범위에서 실근을 갖는다.

$$(\mu_{4,1} < \mu < \mu_{4,2}) \cup (\mu_{4,3} < \mu < \mu_{4,4}) \tag{21}$$

위 식에서 \cup 는 union 연산자이다. 전체 모드의 안정도에 대한 조건은 식 (3-15)의 영역의 교집합에 의하여 주어진다.

3. 연산량

MADF에 관한 연구^{[10][11]}에서는 DM(delta modulation)과 연계하여 LMS와 sign 알고리즘을 이용하여 필터 계수가 갱신되었고, DM 출력의 벡터는 계수 갱신 방정식에서 직접 사용되었다. MADF 구조의 해석에서 입력신호는 DPCM(differential pulse-code modulation) 시스템을 사용하여 양자화 되었고 DPCM 출력은 sign 알고리즘을 이용한 적응필터에 의하여 처리되었다.^[12] MADF는 hardware구현시 계산량을 감소시키기 위하여 μ 가 2의 정수배가 되도록 선택되었고 DPCM시스템의 양자화 스텝 크기 Δ 도 2의 정수배가 되도록 선택되었다. 적응 알고리즘에서 스텝 크기 μ 의 곱셈은 비트 이동 연산으로 치환되었다. 양자화 레벨도 많을수록 좋지만, -6Δ 에서 6Δ 까지 범위의 13개의 양자화 레벨이 두개의 비트 이동 연산 및 한개의 덧셈 연산을 사용하여 구현되었다.

표 1에서 MLMS의 산술 연산의 수가 13개의 DPCM 양자화 레벨을 가진 MADF, sign 알고리즘, LMS 알고리즘의 연산 수와 비교되어 있다. 여기서 N은 필터 계수의 크기(size)를 나타낸다. 따라서 MLMS 알고리즘은 필터구조에서 곱셈기(multiplier)가 많아지므로 계산량이 증가함을 알 수 있다.

표 1. 여러가지 알고리즘에 대한 계산량의 비교
Table 1. Comparison of the computational complexity of various algorithm.

알고리즘	곱셈의 수	덧셈의 수
MADF	0	3N+5
Sign	N	2N
LMS	2N	2N
1st order MLMS	3N	4N
2nd order MLMS	4N	6N

MLMS 알고리즘을 이용한 시스템 identification의 블록도가 그림 2에 도시되어 있다. MLMS 필터를 시스템 identification에 응용할 때, 이상응답 $d(n)$ 과 입력신호 $x(n)$ 의 관계는

$$d(n) = G(z)x(n) + v(n) \tag{22}$$

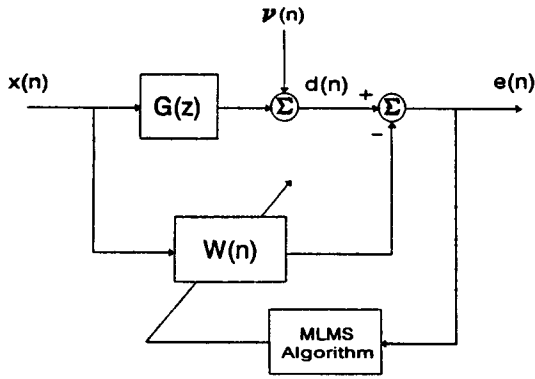


그림 2. 제 2차 MLMS 알고리즘을 이용한 시스템 identification

Fig. 2. System identification using the second order MLMS algorithm.

이 된다. 식 (4-1)에서 $v(n)$ 은 첨가 잡음(additive noise)이고 $x(n)$ 과 상호 상관 관계를 갖지 않고, $G(z)$ 는 미지의 시스템(unknown system)이다. 다음의 2차 시스템이 모의 실험에 사용되었다.

$$G(z) = k(1 - cz^{-1}) / (1 - az^{-1})(1 - bz^{-1}) \quad (23)$$

식 (4-2)에서 $z=a$, $z=b$ 에서 극점을 가지고 $z=c$ 에서 영점을 갖는다. 스칼라 k 는 $G(z)x(n)$ 이 단위 분산을 가지도록 선택되었다. 입력 $x(n)$ 과 잡음 $v(n)$ 은 각각의 평균이 0이고 분산이 $\text{var}(x)=1$, $\text{var}(v)=1$ 인 백색 가우시안 랜덤 수열(white gaussian random sequence)이었다. 그 때 입력신호 상관 행렬 R 은 고유값이 1인 단위 행렬이 된다.

그림 3은 그림 2의 시스템 identification에 적용한 LMS 알고리즘, 제 1차 MLMS 알고리즘($\alpha=0.5$), 및 제 2차 MLMS 알고리즘($\alpha=\beta=0.5$)에 대한 MSE 학습 곡선(learning curve)을 보여준다. 이 곡선은 컴퓨터를 사용하여 200회의 독립 시행을 실시하여 자승 출력 오차 $e^2(n)$ 를 평균하여 얻은 결과이다. 이 경우에 스텝 크기 μ 는 0.0015이었고 계수의 수 N 은 10개이었다. 미지의 시스템 $G(z)$ 에서 $a=0.1$, $b=0.2$, $c=0.5$ 라고 두었고, 상수 k 를 1이라고 가정하였다. 매개변수 α 및 β 는 각각 0.5라고 두었다. 제 2차 MLMS 알고리즘의 수렴 속도는 $\beta > 0$ 일 경우 LMS와 제 1차 MLMS 알고리즘보다 훨씬 빠르다는 것을 알 수 있다. 물론 이 경우에 학습 곡선이 비슷한 수렴특성을 가질 수 있도록 α 의 값을 증가시킬 수도 있다. 이와 같이 모의 실험을 통하여 구해진 최적 중량 계수는 다음과 같다.

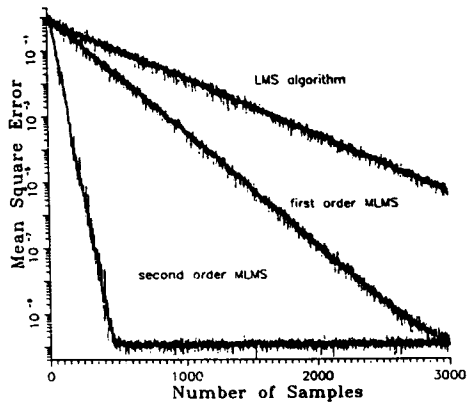


그림 3. 시스템 identification에 대한 MSE 학습 곡선(I)

Fig. 3. MSE learning curves for the system identification(I).

$$W^* = \begin{bmatrix} 1 & -2 \times 10^{-1} & -8 \times 10^{-2} & -2 \times 10^{-2} \\ -4.4 \times 10^{-3} & -9.2 \times 10^{-4} & -1.88 \times 10^{-4} \\ -3.8 \times 10^{-5} & -7.64 \times 10^{-6} & -1.530 \times 10^{-6} \end{bmatrix}^T \quad (24)$$

제 1차 MLMS 알고리즘이 평균 순환 해석에서 알 수 있듯이 $\alpha=1$ 일 때 불안정하다. 또한 α 의 값이 음수 일 때 성능이 저하됨을 알 수 있다. 즉, 전 중량 갱신항이 음의 방향에 있을 때 그 중량 갱신량의 음의 분수를 그 다음의 중량 갱신항에 첨가하기 때문에 수렴 속도가 저하될 수 있다. 불안정한 제 1차 MLMS 알고리즘이 수렴할 수 있도록 적절한 β 의 값을 더하면 그 알고리즘을 안정하게 할 뿐만 아니라 수렴 속도도 현저히 향상시킬 수 있다. $\alpha=1$ 인 불안정한 제 1차 MLMS 알고리즘에서 $-0.998 < \beta < -0.004$ 인 제 2차 MLMS 알고리즘을 적용하였을 때 학습곡선이 안정하였으며, $\beta=-0.085$ 일 때 수렴속도가 가장 빨랐다. 그림 4에서는 스텝 크기 μ 가 0.0015이고 계수의 수 N 이 10인 시스템 identification에 대한 MSE 학습 곡선이 나타나 있다. 이경우에 $\alpha=1$ 인 제 1차 MLMS, $\alpha=1, \beta_1=-0.7$ 및 $\alpha=1, \beta_2=-0.1$ 인 제 2차 MLMS 알고리즘이 각각 도시되어 있다.

V. 결 론

스칼라 변수 α, β 에 의하여 매개변수화된 제 2차 MLMS 알고리즘에 관하여 수렴 결과를 살펴보았다. 평균 수렴 특성을 해석할 때 그 전이 행렬의 고유값이 복소

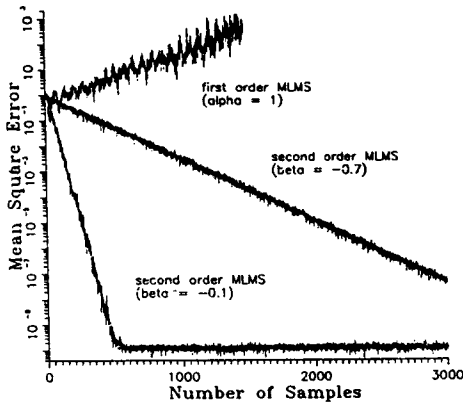


그림 4. 시스템 identification 에 대한 MSE 학습 곡선(II)

Fig. 4. MSE learning curves for the system identification(II).

수 혹은 실수가 될 수 있고, 그 발생 가능 범위를 $\beta > 0$, $\beta < 0$ 경우로 나누어서 구하여 보았다. 평균 수렴 특성을 해석할 때 $\mu < (1 + \alpha - \beta) / \lambda_1$ 에서는 본 알고리즘이 불안정함을 알 수 있었고, 그 결과는 시스템 identification 모의 실험에서 증명되었다. 적절한 β 가 알고리즘의 수렴 속도를 향상시킬 수 있는 반면, 誤調停量(misadjustment)이 증가하는 경우가 발생할 수 있다. 이 경우에는 알고리즘이 수렴 상태에 도달하기 전에 모멘텀 항을 제거하여 오조정량을 감소시킬 수가 있다. 시스템 identification에 적용하였을 때 제 2차 MLMS 알고리즘이 MADF, sign, 및 기존의 LMS 알고리즘보다 수렴속도에서 현저한 향상을 보여었다. 그러나 단점으로서 계산량이 증가함을 알 수 있었다. 또한 제 1차 MLMS 알고리즘이 불안정하거나 발산할 때에도, 제 2차 MLMS 알고리즘을 적용할 경우 학습곡선이 안정하였으며 그 수렴속도가 가장 빨랐다.

부 록

$$a_3 = 32^{-1}(-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta)\beta^{-1}\lambda_1^{-3} \tag{A-1}$$

위 방정식에서 $\mu = y - a_1 / 3$ 으로 치환하면, 다음 식으로 변환된다.

$$y^3 + 3py + q = 0 \tag{A-2}$$

먼저 $y = u + v$ 라 두고, 이를 위식에 대입하면 다음 식을 얻는다.

$$u^3 + v^3 + 3(uv + p)(u + v) + q = 0 \tag{A-3}$$

따라서 $u^3 + v^3 + q = 0$, $uv + p = 0$ 가 되도록, u, v 를 정하면 된다.

$$\begin{aligned} u^3 + v^3 &= -q \\ u^3 v^3 &= -p^3 \end{aligned} \tag{A-4}$$

이므로, u^3, v^3 은 2차 방정식 $t^2 + qt + p^3 = 0$ 의 근이다. 그 근을 A, B로 표현하면, 다음과 같이 나타낸다.

$$\begin{aligned} A &= (-q + (q^2 + 4p^3)^{1/2})/2 \\ B &= (-q - (q^2 + 4p^3)^{1/2})/2 \end{aligned} \tag{A-5}$$

위 식에서 $d = q^2 + 4p^3$ 라 두고 $\omega = (-1 + i 3^{1/2}) / 2$ 라고 나타내면, u 및 v 의 가능한 값은 다음과 같다.

$$\begin{aligned} u &= ((-q + d^{1/2})/2)^{1/3}, \omega((-q + d^{1/2})/2)^{1/3}, \omega^2((-q + d^{1/2})/2)^{1/3} \\ v &= ((-q - d^{1/2})/2)^{1/3}, \omega((-q - d^{1/2})/2)^{1/3}, \omega^2((-q - d^{1/2})/2)^{1/3} \end{aligned} \tag{A-6}$$

u, v 는 관계식 $uv = -p$ 를 만족하기 때문에, 식 (A-2)의 근은 다음과 같이 구해진다.

$$\begin{aligned} y_1 &= A^{1/3} + B^{1/3} \\ y_2 &= \omega A^{1/3} + \omega^2 B^{1/3} \\ y_3 &= \omega^2 A^{1/3} + \omega B^{1/3} \end{aligned} \tag{A-7}$$

이 때 식 (A-1)의 근은 다음과 같이 쓸 수 있다.

$$\begin{aligned} \mu_{1,1} &= A^{1/3} + B^{1/3} - a_1/3 \\ \mu_{1,2} &= \omega A^{1/3} + \omega^2 B^{1/3} - a_1/3 \\ \mu_{1,3} &= \omega^2 A^{1/3} + \omega B^{1/3} - a_1/3 \end{aligned} \tag{A-8}$$

따라서 최종적인 근은 다음과 같다.

식 (15)에서 판별식을 $f(\mu)$ 라고 두면 다음과 같이 쓸 수 있다.

$$\begin{aligned} f(\mu) &= a_0\mu^3 + a_1\mu^2 + a_2\mu + a_3 \\ a_0 &= 1 \\ a_1 &= -8^{-1}(\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)\beta^{-1}\lambda_1^{-1} \\ a_2 &= 8^{-1}(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2)\beta^{-1}\lambda_1^{-2} \end{aligned}$$

$$\begin{aligned} \mu_{s,1} = & 1/(2\lambda_i) + 5\alpha/(12\lambda_i) + \alpha^2/(24\beta\lambda_i) + \beta/(24\lambda_i) \\ & + \{ - [- (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)^3 / (6912\beta^3\lambda_i^3) + \langle (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2) \rangle / (192\beta^2\lambda_i^3) + (-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta^3) / (32\beta\lambda_i^3) + (-i/96)(\alpha^3 - 3\alpha^2\beta - 27\beta^2 + 3\alpha\beta^2 - \beta^3)^{1/3}\beta^2\lambda_i^3 \} \\ & ^{1/3} / 2^{1/3} + \{ - [- (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)^3 / (6912\beta^3\lambda_i^3) + \langle (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2) \rangle / (192\beta^2\lambda_i^3) + (-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta^3) / (32\beta\lambda_i^3) + (i/96)(\alpha^3 - 3\alpha^2\beta - 27\beta^2 + 3\alpha\beta^2 - \beta^3)^{3/2}(3^{1/2}\beta^2\lambda_i^3) \} \\ & ^{1/3} / 2^{1/3} \end{aligned}$$

$$\begin{aligned} \mu_{s,2} = & 1/(2\lambda_i) + 5\alpha/(12\lambda_i) + \alpha^2/(24\beta\lambda_i) + \beta/(24\lambda_i) \\ & + (-1 + i3^{1/2}) \{ - [- (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)^3 / (6912\beta^3\lambda_i^3) + \langle (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2) \rangle / (192\beta^2\lambda_i^3) + (-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta^3) / (32\beta\lambda_i^3) + (-i/96)(\alpha^3 - 3\alpha^2\beta - 27\beta^2 + 3\alpha\beta^2 - \beta^3)^{3/2}(3^{1/2}\beta^2\lambda_i^3) \}^{1/3} \\ & + (-1 + i3^{1/2})^2 \{ - [- (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)^3 / (6912\beta^3\lambda_i^3) + \langle (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2) \rangle / (192\beta^2\lambda_i^3) + (-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta^3) / (32\beta\lambda_i^3) + (i/96)(\alpha^3 - 3\alpha^2\beta - 27\beta^2 + 3\alpha\beta^2 - \beta^3)^{3/2}(3^{1/2}\beta^2\lambda_i^3) \}^{1/3} / (4 \cdot 2^{1/3}) \end{aligned}$$

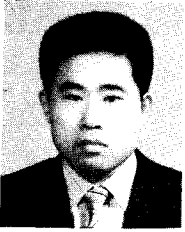
$$\begin{aligned} \mu_{s,3} = & 1/(2\lambda_i) + 5\alpha/(12\lambda_i) + \alpha^2/(24\beta\lambda_i) + \beta/(24\lambda_i) \\ & + (-1 + i3^{1/2}) \{ - [- (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)^3 / (6912\beta^3\lambda_i^3) + \langle (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2) \rangle / (192\beta^2\lambda_i^3) + (-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta^3) / (32\beta\lambda_i^3) + (-i/96)(\alpha^3 - 3\alpha^2\beta - 27\beta^2 + 3\alpha\beta^2 - \beta^3)^{3/2}(3^{1/2}\beta^2\lambda_i^3) \}^{1/3} \\ & + (4 \cdot 2^{1/3}) + (-1 + i3^{1/2}) \{ - [- (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)^3 / (6912\beta^3\lambda_i^3) + \langle (\alpha^2 + 12\beta + 10\alpha\beta + \beta^2)(\alpha^2 + \alpha^3 + 6\beta + \alpha\beta + 4\alpha^2\beta + 10\beta^2 + \alpha\beta^2) \rangle / (192\beta^2\lambda_i^3) + (-\alpha^2 + 2\alpha^3 - \alpha^4 - 4\beta + 8\alpha\beta - 2\alpha^2\beta - 2\alpha^3\beta + 8\beta^2 - 8\alpha\beta^2 - \alpha^2\beta^2 - 4\beta^3) / (32\beta\lambda_i^3) + (i/96)(\alpha^3 - 3\alpha^2\beta - 27\beta^2 + 3\alpha\beta^2 - \beta^3)^{3/2}(3^{1/2}\beta^2\lambda_i^3) \}^{1/3} / (2 \cdot 2^{1/3}) \end{aligned}$$

参 考 文 献

[1] B. Widrow, S. D. Stearns, Adaptive Signal Processing, Englewood Cliffs, NJ:Prentice-Hall, 1985.
[2] J. G. Proakis, "Channel identification for

- high speed digital communications," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 916-922, Dec. 1974.
[3] J. R. Glover, Jr., "Comments on 'Channel identification for high speed digital communications,'" *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 823, Dec. 1975.
[4] J. G. Proakis, "Author's Reply," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 823-824, Dec. 1975.
[5] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, vol. 4, pp. 4-22, Apr. 1987.
[6] S. Roy, J. J. Shynk, "Analysis of the momentum LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, pp. 2088-2098, Dec. 1990.
[7] L. L. Horowitz, K. D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 562-576, June 1981.
[8] A. Feuer, E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 222-230, Feb. 1985.
[9] W. A. Gardner, "Learning characteristics of stochastic gradient descent algorithms: A general study, analysis, and critique," *Signal Processing*, vol. 6, pp. 113-133, 1984.
[10] T. H. Park, D. H. Youn, I. W. Cha, "Multiplication free adaptive digital filters," *Proc. IEEE*, vol. 76, pp. 632-634, May 1988.
[11] S. H. Cho, V. J. Mathews, "Multiplication free adaptive filters," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, New York, pp. 1576-1579, Apr. 1988.
[12] S. H. Cho, "Convergence analysis for efficient adaptive digital filtering algorithms and structures," Ph. D. dissertation, University of Utah, August 1989.
[13] E. I. Jury, Theory and Application of the Z-Transform Method, New-York:Wiley, 1964.

 著 者 紹 介



金 海 正 (正會員)

1960年 10月 26日生. 1984年 경북대학교 전자공학과 졸업(학사). 1986年 한양대학교 대학원 전자공학과 졸업(석사). 1990年~현재 한양대학교 대학원 전자공학과 박사과정 재학중. 주관심분야는 적응 필터, 신경망 이론, 디지털 및 통계 신호처리, 의용공학 등임.

경망 이론, 디지털 및 통계 신호처리, 의용공학 등임.

李 斗 秀 (正會員)

1946年 7月 3日生. 1968年 전북대학교 전자공학과 졸업(학사). 1970年 전북대학교 대학원 전자공학과 졸업(석사). 1973年 전북대학교 대학원 전자공학과 졸업(박사). 현재 한양대학교 전자공학과 부교수. 주관심분야는 디지털 신호처리, 인공지능, 신경망 및 fuzzy 제어 이론, 원격 제어 시스템 등임.