

모멘텀 정규화 경사도 알고리즘과 그 평균 수렴 특성

正會員 金 海 正*

A Momentum Normalized Gradient Algorithm and Its Mean Convergence Characteristics

Hae Jung Kim* Regular Members

要 約

성능 향상을 위하여 스칼라 因數 α_{740} 을 媒介變數化하여 갱신항을 첨가한 非線形 適應 알고리즘인 MNG 알고리즘의 特性을 分析한다. 수렴 特性의 分析에서 平均 필터계수 벡터에 대하여 전이행렬의 固有값이 記述된다. 그 알고리즘이 安定하기 위한 範圍도 證明된다. 본 알고리즘의 時間 常數도 유도되고, 기존의 LMS 방법, LFG 방법, NLMS 방법의 計算量도 比較해 본다. 컴퓨터 模擬實驗(simulation)에서 채널 등화기에 이용하여 MNG 알고리즘의 性能 特性을 보여준다. MNG 알고리즘이 LMS, LFG, NLMS 알고리즘보다 계산량이 증가하는 반면 收斂速度에서 현저한 向上을 보여준다.

ABSTRACT

This paper analyzes the properties of an MNG algorithm that corresponds to the nonlinear adaptive algorithm with additional update terms, parameterized by the scalar factor α_{740} to improve its performance. The analysis of convergence leads to eigenvalues of the transition matrix for the mean filter coefficient vector. Regions in which the algorithm becomes stable are demonstrated. The time constant is derived and the computational complexity of this algorithm is compared with those of the conventional LMS, LFG and NLMS algorithms. In the computer simulation a channel equalizer is utilized to demonstrate the performance feature of the MNG algorithm. The MNG algorithm has more computational complexities but faster convergence speed than LMS and LFG algorithms.

*西京大學校 컴퓨터工學科

Department of Computer Engineering, Seo

Kyeong University

論文番號 : 94352

接受日字 : 1994年 12月 6日

I. 序 論

本 論文에서 제안된 모멘텀 정규화 최소자승오차 알고리즘 (momentum normalized gradient algorithm : MNG)을 연구하게 된 동기는 기존의 최소평균자승 알고리즘 (least mean squares algorithm : LMS)(1), 정규화 최소평균자승 알고리즘(normalized least mean square algorithm : NLMS)(2, 3), 선형 필터화 경사도 알고리즘(linearly filtered gradient algorithm : LFG)(4, 5)보다 빠른 수렴 속도에 대한 필요성에서 비롯되었다. 선형 필터화 경사도 알고리즘은 Proakis (4)에 의하여 제안되었고, Roy 및 Shynk (5)에 의하여 모멘텀 최소평균자승알고리즘(momentum LMS algorithm)이란 이름으로 같은 알고리즘이 더욱 일반적이고 엄밀하게 분석되었다.

LMS 알고리즘은 나쁜 조건의 공분산 행렬을 가진 입력 신호에 대하여 수렴 문제가 간단하지 않다. 그러나 NLMS 알고리즘은 LMS 알고리즘에 비하여 입력신호에 대한 선형적 지식을 필요로 하지 않는다. NLMS 알고리즘의 추적 동특성은 LMS 알고리즘보다 입력신호 분포의 변동에 훨씬 덜 민감하다.

많은 알고리즘 설계자들은 LMS 알고리즘에 비하여 NLMS 알고리즘의 이점을 알고 있지만, NLMS 알고리즘은 실제로 널리 사용되지 못하고 있다. 그 이유로서는, 입력 전력으로 나뉘셈해야 하는 계산상의 복잡성이 하드웨어 곱셈기를 가지는 요즈음의 디지털 신호처리(DSP) 들과 양립되기 어렵다는 것이다. 그러나 급수 전개, 표 검색 등을 이용함으로써 허용되는 정밀성 내에 있도록 이 나뉘셈을 구하는 다양한 방법이 실현될 수 있다.

MNG 알고리즘은 NLMS 알고리즘(2, 3)으로부터 미소량의 모멘텀 매개변수를 부가하여 유도된다. 필터계수의 변화량을 계산할 때 그 이전의 변화량의 일부를 부가하여 필터계수 변화량을 같은 방향으로 유지시킨다. 이 부가항이 수렴시 모멘텀 효과를 나타낸다. 전필터계수 변화량을 현재의 갱신량에 더하기 때문에 필터계수 벡터에 대하여 보다 빠르고 圓滑한(smooth) 수렴특성을 나타낸다. 과거의 필터계수 벡터 $W(n-1)$ 등의 기억공간과 부가적인 곱셈으로 인하여 그 특성이 개선된다. MNG 알고리즘의 계산량은 NLMS 알고리즘에 비하여 증가한다. 그러나 알고리즘이 광역 최소값(global minimum)에 접근할 때 모멘텀 항이 성능을 저하시켜 오조정량(misadjustment)을 증가시킬 수 있다. 이런 문제를 극복하기

위하여 필터계수 벡터의 변화량이 충분히 작을 경우 모멘텀 항을 제거하여 $\alpha_1=0$ 로 둘 수 있다.

本 論文에서는 고유값 해석 방식을 이용하여 MNG 알고리즘의 수렴 특성을 해석하고, 채널 등화기에 응용하여 제안된 알고리즘의 타당성을 검토한다. 2장에서는 MNG 알고리즘을 서술하고 이를 수렴특성의 해석에 적합한 형태로 나타낸다. 3장에서는 평균수렴에서 안정하기 위한 제한조건을 살펴본다. 4장에서는 평균 필터계수 순환식에서 복소수 고유값의 발생에 대하여 새로운 결과를 記述한다. 5장에서는 시정수에 관하여 근사적인 결과식을 보여주고, MNG 알고리즘의 산술적 연산량이 LMS, NLMS, 및 LFG 알고리즘의 연산량과 비교된다. 6장에서는 컴퓨터 모의실험(simulation)을 실시하여 채널 등화기에 적용한 결과 MNG 알고리즘이 LMS 알고리즘보다 계산량이 증가하는 반면 수렴 특성에서 향상된 수렴 속도와 원활한 수렴효과를 나타내는 것을 보여준다. 7장에서는 지금까지의 내용과 앞으로의 과제를 요약한다.

2. 모멘텀 정규화 경사도 알고리즘

경사도 하강 알고리즘(gradient descent algorithm)의 일종으로서 모멘텀 정규화 경사도 알고리즘(MNG algorithm)은 다음과 같은 제2차 필터계수 갱신식으로 표현될 수 있다.

$$W(n+1) = W(n) - \mu \nabla(n) [X^T(n)X(n)]^{-1} + \alpha_1 [W(n) - W(n-1)] \quad (2.1)$$

위 식에서 $W(n) = [w_0(n), \dots, w_{N-1}(n)]^T$ 이다. 자승평균의 앙상블 평균(ensemble average)을 이용하지 않고, 확률적 경사도 알고리즘(stochastic gradient algorithm)(6, 7)에서 위 경사도의 추정치인 순시값을 적용하면 위 경사도는 다음과 같이 근사화 된다.

$$\nabla(n) \approx \begin{bmatrix} \partial e^2(n)/\partial w_0 \\ \partial e^2(n)/\partial w_1 \\ \vdots \\ \partial e^2(n)/\partial w_{N-1} \end{bmatrix} \quad (2.2)$$

$$= -2e(n)X(n)$$

위 식에서 위첨자 T는 벡터의 전치를 의미한다. 스칼라 출력 오차는 다음 식으로 주어진다.

$$c(n) = d(n) - X^T(n)W(n) \quad (2.3)$$

위 식에서 $X(n)=[x(n), \dots, x(n-N+1)]^T$ 은 입력신호 벡터이고, $d(n)$ 은 理想應答이고 $y(n)=W^T(n)X(n)$ 은 $d(n)$ 의 추정값이다. 스텝 크기 μ 는 수렴 속도와 안정성을 제어하는 적용상수이고, 매개변수 α_1 은 경사도에 부가된 모멘텀 항의 축척 계수가 된다.

$\alpha_1=0$ 이면 NLMS 알고리즘이 된다. 알고리즘의 특성상 $\alpha_1 \rightarrow 1$ 일 때 저역통과 필터의 구조를 가지고, $\alpha_1 \rightarrow -1$ 일 때 고역통과 필터의 구조를 가진다. 식 (2.1)에서 경사도를 알아내기가 어렵기 때문에 식 (2.2)에서는 유한한 입력 신호를 이용하여 각 반복순환에서 경사도가 불완전하게 추정되고 있으며 이로 인하여 경사도 잡음이 발생한다[8]. 입력신호 공분산 행렬은 $R=E[X(n)X^T(n)]$ 로 정의하고, 고유벡터 분해 $R=QQT^T$ 가 성립하고, 대각선 행렬 $\Lambda=diag\{1, 2, \dots, N\}$ 에서 위 입력신호 공분산 행렬의 고유벡터는 $0 < \lambda_{\min} \leq \lambda_i \leq \lambda_{\max}$ 로 배열되어 있다. 행렬 Q 의 역은 $QQ^T=I$ 인 특성이 성립하는 직교 고유벡터이고, I 는 단위 행렬이다. 이상응답과 입력신호 벡터사이의 상호 공분산 벡터 $P=P=E[X(n)d(n)]$ 로 정의한다. 입력신호가 넓은 의미에서 정제된(wide-sense stationary) 상태에 있고 서로 상관관계에 있지 않다(uncorrelated)고 가정한다.

필터계수오차 벡터를 $V(n)=W(n)-W_0$ 라고 할 때, $W_0=R^{-1}P$ 은 이산적 Wiener-Hopf 필터의 最適 계수 벡터가 되고, 다음 식이 성립한다.

$$\begin{aligned} V(n+1) - V(n) &= 2\mu e_0(n)[X^T(n)X(n)]^{-1}X(n) \\ &\quad - 2\mu[X^T(n)X(n)]^{-1} \\ &\quad \cdot X(n)X^T(n)V(n) + \alpha_1[V(n) - V(n-1)] \end{aligned} \quad (2.4)$$

위 식에서 최적 출력오차는 $e_0(n)=d(n)-W_0^T X(n)$ 이다. 그러나 위의 Wiener-Hopf 필터에서 선형위상 특성을 보장하지는 않는다.

회전 필터계수 오차벡터를 $\tilde{V}(n)=Q^T V(n)$ 이라고 두고, 식 (2.4)에 Q^T 를 곱하면 다음 식이 성립한다.

$$\begin{aligned} \tilde{V}(n+1) - \tilde{V}(n) &= 2\mu e_0(n)[X^T(n)X(n)]^{-1} \\ &\quad \tilde{X}(n) - 2\mu[X^T(n)X(n)]^{-1} \\ &\quad \tilde{X}(n)\tilde{X}^T(n)\tilde{V}(n) \\ &\quad + \alpha_1[\tilde{V}(n) - \tilde{V}(n-1)] \end{aligned} \quad (2.5)$$

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

적용 매개변수 공간에서 위 알고리즘이 최소값에 도달하면 학습은 끝난다. 그러나 국소(local) 최소값에 도달하면, 출력 오차는 용인될 수 없을 정도로 크다. 위 알고리즘이 허용될 수 있는 최적값에 도달하기 전에 학습을 중지한다면, 위 매개변수를 변화시키거나 필터계수의 초기값을 바꿈으로써 그 문제를 해결할 수 있다.

3. 平均 收斂 特性

단위원 내에 있도록 특성 다항식의 영점의 수를 결정하는 것은 선형 차분 방정식의 안정성 문제와 밀접히 관련된다. 이 문제는 Bezout 행렬 또는 Bezoutian 에 의하여 해결될 수 있다. 복소수 다항식 $p(\zeta)$ 를 가정하고, 단위원에 관한 $p(\zeta)$ 의 inertia 를 $\iota(p)=\{\iota_+(p), \iota_-(p), \iota_0(p)\}$ 라고 정의하자.

$\iota_+(p)$ 은 열린 단위원 내에 있는 다항식의 영점의 수를 의미하고, $\iota_-(p), \iota_0(p)$ 는 각각 단위 원주外部 와 단위 원주에 있는 다항식의 영점의 수를 의미한다. B 를 Bezout 행렬이라고 두자[9]. Q 가 회전 행렬이라고 하면, 회전 Bezout 행렬 $\tilde{B}=QB$ 가 Hermite 행렬이다. 다항식 $p(\zeta)$ 가 안정하기 위한 필요 충분 조건은 회전 Bezout 행렬 \tilde{B} 가 양 확정적(positive definite)이라는 것이다[10, 11]. 또한, C^N 에서 모든 ζ 에 대하여 다항식이 항상 양의 값을 가질 때, 다항식 형태가 양 확정적으로 분류된다.

식 (2.5)에서 평균을 취하면, 다음 식이 성립한다.

$$\begin{aligned} E[\tilde{V}(n+1) - \tilde{V}(n)] &= -2\mu[\text{tr}R]^{-1}AE[\tilde{V}(n)] \\ &\quad + \alpha_1E[\tilde{V}(n) - \tilde{V}(n-1)] \end{aligned} \quad (3.1)$$

위 식 (3.1)에서 $A=E[X(n)X^T(n)]$ 은 고유값의 대각선 행렬이다. 필터의 임펄스 응답이 최적화 되어 있다면, 오차 신호는 입력신호와 상관관계에 있지 않다. 이 때, $E[e_0(n)\tilde{X}(n)]=0$ 이 성립한다[1]. $W(n)$ 과 $X(n)$ 은 독립적이다. $2N$ 차원 상태 벡터 $\tilde{U}(n)$ 를

$$\tilde{U}(n)=[\tilde{V}(n-1)\tilde{V}(n)]^T \quad (3.2)$$

으로 정의하고, 상수 행렬 A 와 B 를 다음과 같이 정의하자.

$$\begin{aligned} A &= -\alpha_1 I \\ B &= (1 + \alpha_1)I - 2\mu [\text{tr}R]^{-1}A \end{aligned} \quad (3.3)$$

위 식 (3.3)을 이용하면, 식 (3.1) 은 다음과 같은 순환식으로 고쳐 쓸 수 있다.

$$E[\tilde{U}(n+1)] = C E[\tilde{U}(n)] \quad (3.4)$$

식 (3.4) 에서 행렬 C 는 블록 전이 행렬(block transition matrix)로서 2N 차수(2Nth order)의 크기를 가진다. 행렬 C 는 다음과 같이 정의 된다.

$$C = \begin{bmatrix} 0 & I \\ A & B \end{bmatrix} \quad (3.5)$$

여기서 0 는 $N \times N$ 영행렬이다. 벡터 $E[\tilde{U}(n+1)]$ 가 수렴하기 위하여, 위 행렬 C 의 모든 고유값이 단위원 내에 있어야 한다. 본 알고리즘이 수렴할 조건을 스텝 크기와 모멘텀 매개변수와 관련하여 구하도록 하자. C 의 임의의 고유값을 ξ 라고 하고 그것에 해당하는 고유 벡터를 $q^T = [q_1^T \ q_2^T]$ 라고 하자. 여기서 q_1^T , q_2^T 는 N차수 열벡터이다. 그러면 $Cq = \xi q$ 의 방정식을 얻는다. 이를 다시 쓰면

$$\begin{bmatrix} 0 & I \\ A & B \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \xi \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \quad (3.6)$$

이므로

$$Aq_1 + \xi B q_1 = \xi^2 q_1 \quad (3.7)$$

을 얻는다. 식 (3.7) 에서 상수 행렬을 치환하면 아래와 같이 다시 쓸 수 있다.

$$\begin{aligned} -\alpha_1 q_1 + \xi [(1 + \alpha_1)I - \mu(1 - \alpha_1) \\ [\text{tr}R]^{-1}A] q_1 &= \xi^2 q_1 \end{aligned} \quad (3.8)$$

식 (3.8) 은 N 개의 스칼라 방정식으로 다시 쓸 수 있다.

$$\begin{aligned} \xi^2 - [(1 + \alpha_1) - 2\mu [\text{tr}R]^{-1} \lambda_i] \xi \\ + \alpha_1 = 0, \quad i=0, 1, 2, \dots, N-1 \end{aligned} \quad (3.9)$$

위 식에서 2N개의 근을 결정하기 위하여, 다음의 2차형(quadratic form)

$$\xi_i^2 - (1 + \alpha_1 - 2\mu [\text{tr}R]^{-1} \lambda_i) \xi_i + \alpha_1 = 0 \quad (3.10)$$

을 조사할 필요가 있다. 수정된 Schur-Cohn 기준 [12]을 이용하여, 위 식의 다항식의 안정성을 판정하기 위하여, 다음의 안정성 조건을 유도할 수 있다.

$$\text{조건 1 : } 0 < \mu < \frac{(1 + \alpha_1)[\text{tr}R]}{\lambda_{\max}}$$

$$\text{조건 2 : } -1 < \alpha_1 < 1$$

α_1 이 0 에 접근할 때, 위 2개의 제한 조건은 다음과 같은 안정성 조건으로 귀결된다.

$$\text{조건 1 : } 0 < \mu < \frac{[\text{tr}R]}{\lambda_{\max}}$$

위 안정도 한계는 정규화 최소평균자승 알고리즘의 안정성 조건이 되는 것을 알 수 있다.

위 조건에서 $\alpha_1 < 0$ 일 때, α_1 과 μ 가 안정하기 위한 범위는 다음과 같이 쓸 수 있다.

$$\begin{aligned} -1 + 2\mu [\text{tr}R]^{-1} \lambda_i < \alpha_1 < 0 \\ 0 < \mu < (1 - |\alpha_1|) [\text{tr}R] / \lambda_i \end{aligned} \quad (3.11)$$

$\alpha_1 \rightarrow 0$ 일 때, 안정도 한계는 $0 < \mu < [\text{tr}R] / \lambda_{\max}$ 로 증가한다. $\alpha_1 \rightarrow -1$ 일 때, 안정도 한계는 0 으로 감소한다. α_1 가 음수일 경우 알고리즘이 안정하더라도 그 성능을 저하시킨다. 즉, 전 필터계수 갱신항이 음의 방향에 있을 때 ($\alpha_1 [W(n) - W(n-1)] < 0$), 그 필터계수 갱신항의 음의 분수를 그 다음의 필터계수 갱신항에 첨가하기 때문에 수렴 속도가 저하될 수 있다.

$\alpha_1 > 0$ 및 $-1 + \mu [\text{tr}R]^{-1} \lambda_i > 0$ 일 때, 다음과 같은 부등식이 성립한다.

$$-1 + \mu[\text{trR}]^{-1} \lambda_i < \alpha_1 < 1$$

$$1 / \lambda_i [\text{trR}]^{-1} < \mu < (1 + \alpha_1) / \lambda_i [\text{trR}]^{-1} \quad (3.12)$$

$\alpha_1 \rightarrow 0$ 일 때 안정성 한계는 $[\text{trR}] / \lambda_{\min} < \mu < (1 + \alpha_1) [\text{trR}] / \lambda_{\max}$ 로 감소하고, $\alpha_1 \rightarrow 1$ 일 때 안정도 한계는 $[\text{trR}] / \lambda_{\min} < \mu < 2[\text{trR}] / \lambda_{\max}$ 이 된다. $\alpha_1 \rightarrow -1$ 일 때 안정도 한계는 $\mu \rightarrow 0$ 가 된다.

$\alpha_1 > 0$ 및 $-1 + \mu[\text{trR}]^{-1} \lambda_i < 0$ 일 때, 다음과 같은 부등식이 성립한다.

$$0 < \alpha_1 < 1$$

$$\mu < [\text{trR}] / \lambda_i \quad (3.13)$$

$\alpha_1 \rightarrow 0$ 일 때 안정도 한계는 $\mu < [\text{trR}] / \lambda_{\max}$ 이고, $\alpha_1 \rightarrow 1$ 일 때 안정성 한계는 $\mu < [\text{trR}] / \lambda_{\max}$ 이 된다.

NLMS 알고리즘의 전이행렬은 대칭적이고, 그 고유값은 실수가 된다. 그러나 모멘텀 정규화 LMS 알고리즘의 전이 행렬은 대칭적이지 않고, 그 고유값은 복소수가 된다.

4. 안정성 영역

특성 방정식의 근이 실수가 되기 위하여, 그 판별식은 다음과 같이 음수가 되어서는 안된다.

$$(1 + \alpha_1 - 2\mu[\text{trR}]^{-1} \lambda_i)^2 - 4\alpha_1 \geq 0 \quad (4.1)$$

위 식을 정리하면 다음 식과 같이 된다.

$$4\lambda_i^2 [\text{trR}]^{-2} \mu^2 - 4(1 + \alpha_1) \lambda_i [\text{trR}]^{-1} \mu + (1 - \alpha_1)^2 \geq 0 \quad (4.2)$$

위 식에서 2차 함수를 $f(\mu)$ 라고 놓으면, 다음과 같은 성질을 가진다. $f(\mu)$ 의 제2 도함수 $f''(\mu) = 8\lambda_i^2 [\text{trR}]^{-2}$ 은 strictly positive 이기 때문에 $f(\mu)$ 는 concave upward 하다.

$f(\mu)$ 은 $\frac{(1 + \alpha_1) \text{trR}}{2\lambda_i}$ 에서 극값을 가진다는 것을 알 수 있다.

$f(\mu)$ 는 다음과 같은 경우에 음수가 되지 않고, 특성 방정식이 실근을 갖는다.

i) $\alpha_1 < 0$

ii) $\alpha_1 \geq 0$ 일 때, $\mu \leq \mu_{i,1}$ 혹은 $\mu \geq \mu_{i,2}$

여기서 $\mu_{i,1}$ 및 $\mu_{i,2}$ 는 다음과 같다.

$$\mu_{i,1} = \frac{(1 - \sqrt{\alpha_1})^2 \text{trR}}{2\lambda_i}$$

$$\mu_{i,2} = \frac{(1 + \sqrt{\alpha_1})^2 \text{trR}}{2\lambda_i}$$

위의 ii)에서 $0 \leq \alpha_1 < 1$ 이기 때문에

$$0 < \mu_{i,1} \leq \mu_{i,2} < \frac{2\text{trR}}{\lambda_i} \text{ 이 성립한다.}$$

앞의 결과로부터 알고리즘의 i 번째 모드 of 안정성에 대한 조건은 다음과 같다.

$$0 < \mu < \mu_{i,3} = (1 + \alpha_1) \text{trR} / \lambda_i \quad (4.3)$$

$\alpha_1 < 0$ 일 때, 특성 방정식은 전 범위에 걸쳐 실근을 갖는다.

$\alpha_1 \geq 0$ 일 때, 특성 방정식은 다음과 같이 제한된 범위에서 실근을 가진다.

$$(0 < \mu < \mu_{i,1}) \cup (\mu_{i,2} < \mu < \mu_{i,3}) \quad (4.4)$$

위 식에서 \cup 는 union 연산자이다. 전체 모드 of 안정도에 대한 조건은 식 (4.4)의 영역의 교집합에 의하여 주어진다.

5. 時間 常數 및 演算量

각 순간값은 기하 수열이 되고, i 번째 mode of 기하비(geometric ratio)를 가지고 있다. 제안한 알고리즘이 수렴하기 위하여 기하비의 크기는 모든 mode 에 대하여 1 보다 작아야 한다. 따라서 위 안정한 조건을 만족할 경우, 초기 조건에 관계없이 충분히 큰 반복 회수에 대하여 모든 mode 는 0 에 접근하기 때문에 가장 낮은 최적해에 접근한다는 것을 알 수 있다.

일반적으로, 시간 단위(time unit)를 한 반복 주기(cycle) 기간으로 가정한다면, 위 기하 수열은 지수 포락선으로 근사화될 수 있다(8). 시간 상수 τ_i 는 i 번째 모드 of 크기가 초기 값의 $1/e$ 로 減衰하는데 요구되는 시간으로 정의된다. 저속의 적응성과 큰 값의 시간 상수를 가정할 때, τ_i 는 다음과 같은 근사화 식으로 표현할 수 있다.

일반적으로 위 기하 수열이 지수 포락선으로 근사화된다. 저속의 적응성과 큰값의 시정수를 가정할 때, τ_i 는 다음과 같은 근사화 식으로 표현할 수 있다(8).

$$\tau_i = \frac{-1}{\ln |\zeta_i|} \quad (5.1)$$

위 식에서 특성방정식의 근 ζ_i 가 실수이고 양수의 제곱근만을 유지한다고 가정하면, i 번째 모드에서 시간 상수 τ_i 는 전이 행렬의 그 대응 고유값과 연관되고, 다음과 같은 관계식을 가진다[5].

$$\tau_i \approx 1/(1-\xi_i) \quad (5.2)$$

스텝 크기 μ 및 모멘텀 매개변수 α_1 이 작다고 가정한다면, LFG 알고리즘의 시간 상수에 대한 근사적인 표현식을 유도하는 것이 가능하다.

$$\xi_i = 1 + \alpha_1 - 2\mu [\text{tr} R]^{-1} \lambda_i - \frac{\alpha_1}{1 + \alpha_1 - 2\mu [\text{tr} R]^{-1} \lambda_i} \quad (5.3)$$

식 (5.2) 및 식 (5.3)으로부터 α_1^2 항을 무시하고 $2\mu \ll 1$ 를 가정하면 시간 상수는 다음과 같이 주어진다.

$$\tau_i \approx \frac{1 + \alpha_1}{1 + 2\alpha_1} \cdot \frac{[\text{tr} R]}{2\mu \lambda_i} \quad (5.4)$$

MNG 알고리즘은 1을 적절히 조정함으로써 LMS 알고리즘 보다 수렴속도를 개선할 수 있다. 그러나 이는 오조정량을 증가시킬 수 있으므로, 서로 상충하는 수렴속도와 오조정량이 실제 필터설계에서 허용되도록 안정한 범위 내에서 최적의 α_1 을 선정하여야 한다.

MNG 알고리즘의 計算量(computational complexity)을 기존의 알고리즘의 계산량과 비교하여 보자. MNG 알고리즘의 산술 연산(arithmetic operation)의 수가 sign 알고리즘 [13], LMS 알고리즘, LFG 알고리즘, NLMS 등의 산술 연산의 수와 비교된다. Sign 알고리즘의 곱셈 연산수는 N 이고, LMS 알고리즘의 곱셈 연산수는 $2N$ 이고, LFG 알고리즘의 곱셈 연산수는 $3N$ 이고, NLMS 알고리즘의 곱셈 연산수는 $3N$ 이고, MNG 알고리즘의 곱셈 연산수는 $4N$ 이다. 여기서 N 은 적응필터 계수의 크기(size)를 나타낸다. MNG 알고리즘의 곱셈 연산 수는 sign 알고리즘에 비하여 4배 증가하고,

LMS 알고리즘에 비하여 2배 증가한다. MNG 알고리즘의 곱셈 연산수는 LFG 및 NLMS 알고리즘에 비하여 4/3배 증가한다. MNG 덧셈 연산 수는 sign 알고리즘에 비하여 5/2배 증가하고, LMS 알고리즘에 비하여 5/2배 증가하고, LFG 알고리즘에 비하여 5/4배 증가하고, NLMS 알고리즘에 비하여 5/3배 증가한다. MNG 알고리즘은 필터 구조에서 곱셈기 및 덧셈기가 많았으므로 계산량이 증가한다는 것을 위 결과로부터 알 수 있다.

6. 컴퓨터 모의 실험

MNG 알고리즘을 이용한 채널 동화기(channel equalizer)의 블록도가 그림 1에 도시되어 있다. MNG 필터의 입력이 되는 채널의 출력은 다음 식에 따라서 생성된다.

$$x(n) = G(z)s(n) + \nu(n) \quad (6.1)$$

위 식에서 $s(n)$ 은 채널 입력이고, $\nu(n)$ 은 첨가잡음(additive noise)이다. 이들은 상호 상관관계가 없고 백색 수열이다. 채널 입력 $s(n)$ 은 이진 신호(binary signal)인데, 각 신호는 동일한 확률을 가진 신호였다. 첨가 잡음은 $\nu(n)$ 인데, 균일한 분포(uniform distribution)를 가지고 있었다. 따라서 위 백색 랜덤 수열(white random sequences)은 평균이 0이고 분산이 각각 $\text{var}(s)=1.0$ 과 $\text{var}(\nu)=0.001$ 이었다. 모의실험에 사용된 채널의 임펄스 응답 $g(i)$ 는 다음과 같이 정의 되는 raised-cosine pulse에 상응하는 제2차 FIR 필터였다 [14][15].

$$g(i) = \frac{k}{2} \left[1 + \cos\left(\frac{2\pi(i-1)}{h}\right) \right], \quad i = 0, 1, 2 \quad (6.2)$$

위 식에서 스케일 인수(scale factor) h 는 왜곡이 h 에 따라 증가하는 채널에 의하여 생성되는 크기 왜곡의 양을 제어한다. 이로부터 적응필터의 입력신호 상관행렬 R 의 고유값 擴張比(eigenvalue spread) 혹은 條件數(condition number)를 결정할 수 있다 [16][17]. 적응 필터의 계수의 수는 $N=21$ 이었다. 이 상응하는 채널 입력에서 11개의 샘플이 지연(retard)되었다. 이로 인하여 21개의 탭 동화기(tap equalizer)에 대하여 중심 탭이 11번째 계수가 되는 인과적인 동화기(causal equalizer)가 된다.

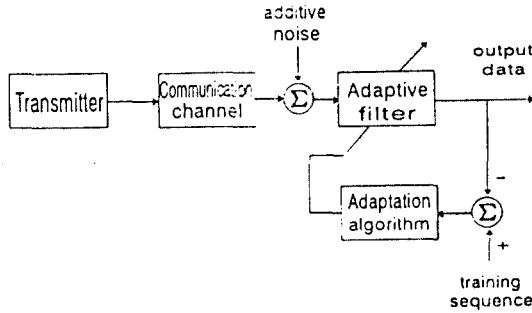


그림 1. MNG 알고리즘을 이용한 채널 등화기
Fig. 1. Channel equalization using the MNG algorithm

그림 2~그림 4는 채널 등화기(channel equalizer)의 모의실험에서 LMS알고리즘, LFG 알고리즘, QFG 알고리즘에 대한 평균자승오차 학습 곡선(MSE learning curve)을 보여준다. 이 곡선은 컴퓨터를 사용하여 50회의 독립 시행을 실시하여 출력 오차 $e(n)$ 의 자승을 평균하여 얻은 앙상블 평균(ensemble average) 수렴특성을 보여준다. 이 경우에 적응

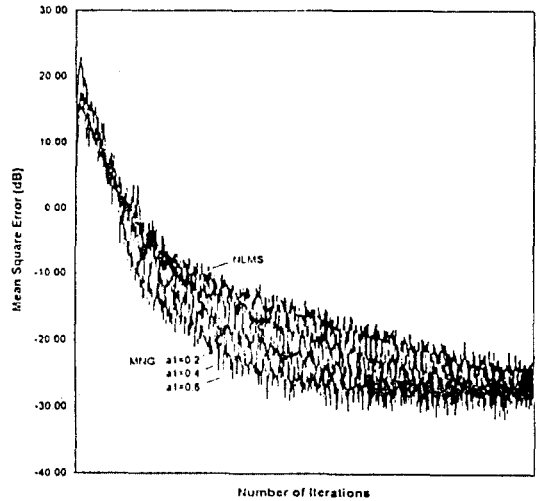


그림 3. NLMS, $\alpha_1=0.2, 0.4, 0.6$ 을 가진MNG 알고리즘에 대한 MSE 학습 곡선
Fig. 3. MSE learning curves for the NLMS, and MNG algorithms with $\alpha_1=0.2, 0.4, 0.6$

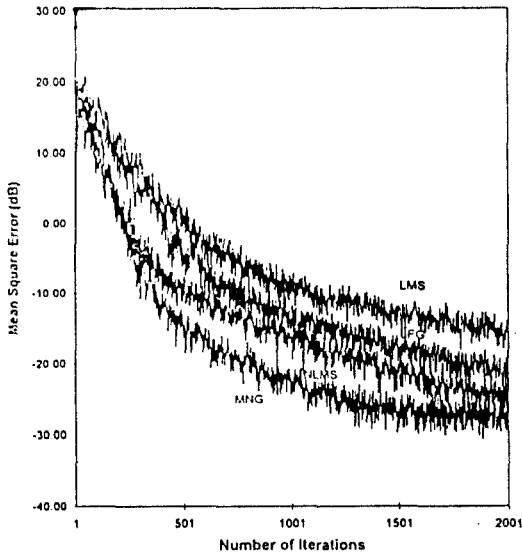


그림 2. LMS, NLMS, $\alpha_1=0.4$ 을 가진 MNG 및 LFG 알고리즘에 대한 MSE 학습 곡선
Fig. 2. MSE learning curves for the LMS, NLMS, MNG and LFG algorithms with $\alpha_1=0.4$

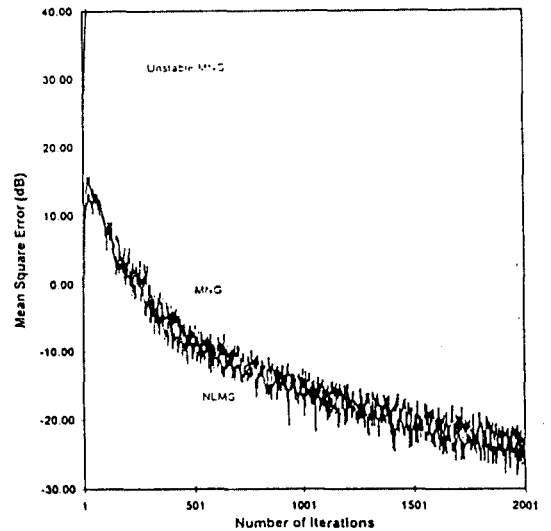


그림 4. NLMS 및 $\alpha_1 = -0.2, 1.0$ 등을 가진MNG 알고리즘에 대한 MSE 학습 곡선
Fig. 4. MSE learning curves for the NLMS, and MNG algorithms with $\alpha_1=-0.2$ and 1.0

필터 계수의 수 N 은 21개이었다. 또한 채널 시스템 $G(z)$ 에서 상수 k 를 1.0 이라고 가정하였다.

그림 2 에서는 조건 $h=3.0, \mu=0.05$ 아래에서 매개변수 $\alpha_1=0.4$ 인 MNG 알고리즘과 NLMS 알고리즘, LFG 알고리즘 및 LMS 알고리즘이 도시되어 있는데, MNG 알고리즘의 수렴속도가 크게 향상된 것을 알 수 있다. $h=3.0$ 일 때 固有擴張比는 8.2 가 되었고 이에 상응하는 채널 계수는 $g(0)=0.25, g(1)=1.0, g(2)=0.25$ 이었다.

그림 3에서는 조건 $h=3.0, \mu=0.05$ 아래에서 여러가지 매개변수를 가진 MNG 알고리즘 및 NLMS 알고리즘을 보여주고 있는데, 모멘텀 매개변수가 없는 NLMS 알고리즘보다 모멘텀 매개변수가 있는 MNG 알고리즘의 수렴 속도가 증가하였고, 또한 모멘텀 매개변수가 클수록 수렴 속도가 향상되는 것을 알 수 있다.

그림 4 에서는 조건 $h=3.0, \mu=0.05$ 아래에서 MNG 알고리즘 및 NLMS 알고리즘을 보여주고 있는데, $\alpha_1=1$ 인 MNG 알고리즘이 불안정하다는 것을 보여주고 있고, 이는 MNG 알고리즘이 $\alpha_1 \geq 1$ 일 때 불안정하다는 것은 평균 순환 해석에서 추론할 수 있기 때문에 그 해석 결과와 모의 실험결과가 서로 일치함을 알 수 있다. 또한 α_1 의 값이 음수 일 때 성능이 저하되는 것을 보여주고 있는데, 이는 앞에서 추론한 해석 결과와 일치되는 결과를 나타내고 있다. 전 필터계수 갱신항이 음의 방향에 있을 때 그 필터계수 갱신량의 음의 분수를 그 다음의 필터계수 갱신항에 첨가하기 때문에 수렴 속도가 저하되는 것이다.

7. 結 論

스칼라 변수 α_1 에 의하여 매개변수화된 정규화 최소 평균자승 방법인 MNG 알고리즘을 제안하였고, 이에 관하여 收斂 結果를 살펴보았다. 평균 수렴특성을 해석할 때 평균자승 순환식을 유도하고 알고리즘이 안정하기 위하여 고려되어야 할 필요조건에 대하여 알아 보았다. 또한 그 해석에서 그 전이 행렬의 고유값이 복소수 혹은 실수가 되는 범위를 규명하였다. 수렴 특성을 해석할 때 본 알고리즘이 안정하기 위한 범위로서 제한조건을 보여 주었다. 이 때 $\mu_1 > (1 + \alpha_1)[\text{tr}R]/\lambda_1$ 에서는 본 알고리즘이 안정함을 알 수 있었고, 그 결과는 컴퓨터 모의 실험을 통하여 증명되었다. 적절한 모멘텀 매개변수 α_1 가 알고리즘의 수렴속도를 향상시킬 수 있지만, 수렴할 때 평균자승오차가 증가하는 경우가 발생할 수 있다. 이

경우에는 알고리즘이 수렴 상태에 도달하기 전에 모멘텀 항을 제거하여 오조정량을 감소시킬 수 있다.

채널 동화기에 적용한 컴퓨터 모의실험에서는 MNG 알고리즘이 기존의 LMS 알고리즘, sign 알고리즘, NLMS 알고리즘보다 수렴속도에서 현저한 향상을 보여 주었고, 비스트 오차 효과를 방지하도록 원활한 (smooth) 수렴특성을 나타내었다. 그러나 단점으로서 MNG 알고리즘의 계산량이 LMS, LFG, NLMS, sign 알고리즘의 계산량에 비하여 증가하고, 부가적인 기억 공간을 필요로 한다는 것을 알 수 있었다.

디지털 通信 시스템에서 평균자승오차를 최소화하기 위한 채널 동화기는 초기의 학습기간(initial training period)을 필요로 하였다. 그러나 이와 같은 초기의 학습기간도 필요 없고 필터의 입력 data 흐름에 대하여 동화기로서 언제든지 직접 수행될 수 있는 보다 효율적인 동화기와 carrier recovery 알고리즘 등은 앞으로 연구되어야 할 것으로 사료된다.

참 고 문 헌

1. B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985.
2. J. I. Nagumo and A. Noda, "A learning method for system identifications," *IEEE Trans. Automat. Control*, vol.12, pp. 282- 287, June 1967.
3. R. R. Bitmead and B. D. O. Anderson, "Performance of adaptive estimation algorithms in dependent random environments," *IEEE Trans. Automat. Control*, vol.25, no.4, pp.788-793, August 1980.
4. J. G. Proakis, "Channel identification for high speed digital communications," *IEEE Trans. Automat. Contr.*, vol. AC-19 pp.916-922, Dec. 1974.
5. S. Roy and J. J. Shynk, "Analysis of the momentum LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, pp.2088-2098, Dec. 1990.
6. W. A. Gardner, "Learning characteristics of

- stochastic gradient descent algorithms : A general study, analysis, and critique." *Signal Processing*, vol. 6, pp.113-133, 1984.
7. B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive Antenna Systems." *Proc. IEEE*, vol.55, no.12, pp.2143-2159, Dec.1967.
 8. B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *Proc. IEEE*, vol.64, no.8, pp.1151- 1162, Aug. 1976.
 9. P. Lancaster and M. Tismenetsky, *The Theory of Matrices with Application*, Orlando, Florida : Academic Press, 1985.
 10. J. B. Foley and F. M. Boland, "A note on convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-36, no. 7, pp. 1087-1089, July 1988.
 11. F. M. Boland and J. B. Foley, "Stochastic convergence of the LMS algorithm in adaptive systems," *Signal Processing*, (Elsevier Science Publishers B.V. (North-Holland)), vol. 13, no. 4, pp. 339-352, December 1987.
 12. E. I. Jury, *Theory and Application of the z-Transform Method*, New-York : Wiley, 1964.
 13. S. H. Cho, "Convergence analysis for efficient adaptive digital filtering algorithms and structures," *Ph.D. dissertation*, University of Utah, August 1989.
 14. B. Mulgrew and C. F. N. Cowan, *Adaptive Filters and Equalizers*, Kluwer Academic Publishers, 1988.
 15. B. Mulgrew and P. M. Grant, *Adaptive Filters*, Englewood Cliffs, NJ : Prentice- Hall, 1985.
 16. S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ : Prentice-Hall, 1991.
 17. S. Roy, "Design of Linear Phase FIR Filters Using the LMS Algorithm." *IEEE Trans. Signal Processing*, vol.41, pp.1685- 1689, Apr. 1993.



金海正 (Hae Jung Kim) 정희원

1984년 : 慶北大學校 電子工學科 졸업 (학사)

1986년 : 漢陽大學校 大學院 電子工學科 졸업 (석사)

1995년 : 漢陽大學校 大學院 電子工學科 졸업 (박사)

1995년~현재 : 西京大學校 컴퓨터工學科 전임강사

※ 관심분야 : Adaptive Filtering, Digital and Statistical Signal Processing, Data Communication, Parameter Estimation and Adaptive Control, Neural Network, Biomedical Engineering, etc.