

가변 축척 매개변수를 가진 변형 확률적 경사도 기반 필터의 해석

김해정*

Analysis of a Modified Stochastic Gradient-Based Filter with Variable Scaling Parameter

Haejung Kim* *Regular Member*

요 약

본 논문은 변형 확률적 경사도 기반 (MSGB) 필터를 제안하고 그 필터가 최적화 문제에 대한 해가 될 수 있음을 보여준다. 갱신항으로 첨가된 가변적 축척 매개변수를 가진 비선형 적응 필터인 MSGB 필터의 특성을 분석한다. 가변 매개변수의 MSGB 필터는 가변 축척 매개변수를 통하여 고정 매개변수의 MSGB 필터와 무매개변수의 MSGB 필터를 연결하는 역할을 한다. 그 안정성 영역과 오조정량도 살펴본다. 시스템 식별에 응용하여 컴퓨터 모의실험을 실행하여 MSGB 필터의 개선된 성능 특성을 보여준다.

Key Words : MSGB filter, Variable parameter, Initial factor, Forgetting factor, System identification

ABSTRACT

We propose a modified stochastic gradient-based (MSGB) filter showing that the filter is the solution to an optimization problem. This paper analyzes the properties of the MSGB filter that corresponds to the nonlinear adaptive filter with additional update terms, parameterized by the variable scaling factor. The variably parameterized MSGB filter plays a role in connecting the fixed parameterized MSGB filter and the null parameterized MSGB filter through variably scaling parameter. The stability regions and misadjustments are shown. A system identification is utilized to perform the computer simulation and demonstrate the improved performance feature of the MSGB filter.

1. 서론

기존의 확률적 경사도 필터는 그 구조가 단순하기 때문에 상대적으로 저속의, 데이터 의존적 수렴 거동을 가진 단점이 있으며, 나쁜 조건의 입력 상관 행렬을 가진 입력 신호에 대하여 수렴 문제가 간단하지 않다[1]. 그러나 정규화 확률적 경사도 필터는 기존의 확률적 경사도 필터에 비하여 입력신호에 대한 선형적 지식을 필요로 하지 않는다[2-4]. 정규화 확률적 경사도 필터의 추적 동특성은 기존의 확

률적 경사도 필터보다 입력신호 분포의 변동에 훨씬 덜 민감하다. 정규화 확률적 경사도 필터의 상관관계가 있거나 상관관계가 없는 입력 데이터에 대하여 기존의 필터보다 더 빠른 수렴 속도를 보여준다[2][5].

본 논문에서는 가변 축척 수열에 의하여 매개변수화된 변형 확률적 경사도 기반 필터를 제안한다. 변형 확률적 경사도 필터는 정규화 확률적 경사도 필터로부터 가변 축척 매개변수를 부가하여 구성된다. 변형 확률적 경사도 기반 필터는 초기 상태에서 필터계수의 변화량을 계산할 때 그 이전의 변화량의

* 서경대학교 정보통신공학과 (hjkim@skuniv.ac.kr)

논문번호 : KICS2006-01-043, 접수일자 : 2006년 1월 8일, 최종논문접수일자 : 2006년 8월 21일

일부를 부가하여 필터계수 변화량을 동일한 방향으로 유지시키며 전필터계수 변화량을 현재의 갱신량에 더하기 때문에 빠르고 원활한 수렴특성을 나타낸다. 변형 확률적 경사도 기반 필터는 초기상태에서 고정 매개변수의 변형 확률적 경사도 기반 필터의 장점을 이용하고, 정상상태에서 무매개변수의 변형 확률적 경사도 기반 필터의 장점을 이용한다.

본 논문에서는 변형 확률적 경사도 기반 필터를 제안하고, 그 필터의 타당성을 검토하고, 이를 시스템 식별에 응용한다. 제2절에서는 복소수 공간에서 변형 확률적 경사도 기반 필터를 수식으로 표현하고 복소수 필터계수와 복소수 양태행렬 등을 서술한다. 제3절에서는 고유값 해석방식을 이용하여 평균적 수렴 특성을 해석하고 다양한 안정성 조건을 서술한다. 제4절에서는 오조정량을 서술한다. 제5절에서는 시스템 식별에 적용하여 컴퓨터 모의실험을 실시하고 수렴특성이 개선되었음을 보여준다. 제6절에서는 지금까지의 내용을 간단히 요약한다.

II. 변형 확률적 경사도 기반 필터

복소수 영역에서 변형 확률적 경사도 기반 필터(modified stochastic gradient-based filter: MSGB filter)는 다음과 같은 최적화 문제에 대한 해결 방안이 된다.

$$\begin{aligned} & \text{minimize } \|W(n+1) - W(n)\|_p \\ & \text{subject to } d(n) - W^H(n+1)X(n) = 0 \end{aligned}$$

위 식에서 필터의 계수는 복소수이고, $W(n) = W_R(n) + jW_I(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^T$ 이다. 또한 아래 첨자 R 과 I 는 실수 부분과 허수 부분을 표시한다. 위 첨자 H 는 벡터의 Hermitian 연산을 지칭하고, 위첨자 T 는 벡터의 전치 연산을 의미한다. 여기서 $\|\cdot\|_p$ 는 L_p norm을 의미한다. 확률적 경사도 기반 필터는 필터 계수의 L_p norm에서 최소 변화량(minimum disturbance)을 제공하여, 시간의 진행에 따라 입력 데이터와 이상적 응답 사이의 필터링 관계를 만족시킨다. 이와 같이 제한된 최적화 문제를 해결하기 위하여 복소수 값을 가지는 일반적 경우에 다음과 같은 Lagrange 곱셈기 방법을 사용한다.

$$\begin{aligned} \mathcal{J}(n) = & \\ & \|\delta W(n+1)\|^2 + \text{Re}[\lambda^H(d(n) - W^H(n+1)X(n))] \quad (1) \end{aligned}$$

위 식의 비용 함수 $\mathcal{J}(n)$ 를 최소화하는 필터 계수의 최적값은 정규화 확률적 경사도 기반 필터로 알려진 제1차 필터계수 갱신식으로 표현될 수 있다[2-4].

복소수 영역의 MSGB 필터는 경사도 하강 필터의 일종으로서 다음과 같은 제2차 필터계수 갱신식으로 표현할 수 있다.

$$\begin{aligned} W(n+1) & \\ = & [1 + \gamma(n)]W(n) - \frac{\mu}{\delta + X^H(n)X(n)} \nabla_{W^H} \xi(n) + \gamma(n)W(n-1) \quad (2) \end{aligned}$$

위 식에서 δ 는 양의 값을 가진다. 자승오차의 앙상블 평균(ensemble average)을 이용하지 않고, 복소수 영역의 MSGB 필터에서 위 경사도의 추정값인 순시값을 적용하면 위 경사도는 다음과 같이 근사화된다.

$$\nabla_{W^H} \xi(n) \approx -2e^H(n)X(n) \quad (3)$$

복소수 영역에서 출력 오차는 다음 식으로 주어진다.

$$\begin{aligned} e(n) & = e_R(n) + je_I(n) \\ & = d(n) - W^H(n)X(n) \quad (4) \end{aligned}$$

입력신호 벡터는 $X^T(n) = [x(n), x(n-1), \dots, x(n-N+1)]$ 이고, $d(n)$ 은 이상적 응답이고 $y(n) = W^H(n)X(n)$ 은 이상적 응답의 추정값이다. 또한, 변이인수 μ 는 수렴 속도와 안정성을 제어하는 적응상수이다. 가변적 매개변수(variable parameter)는 경사도에 부가된 항의 축척 계수 $\gamma(n)$ 가 된다. 가변적 매개변수의 일례를 들면 다음과 같은 다항식으로 표현될 수 있다.

$$\gamma(n) = \alpha\beta^n \quad (5)$$

위 식에서 α 를 초기인수(initial factor)라고 부르며, β 를 망각인수(forgetting factor)라고 부르는데 그 범위는 $0 \leq \beta \leq 1$ 로 둔다.

가변적 매개변수 $\gamma(n)$ 이 0일 때 MSGB 필터를 무매개변수의 MSGB (null parameterized MSGB: NPMSGB) 필터라고 부른다. 망각 인수 β 가 0 이 되면 $\gamma(n) = 0$ 가 되어 상시적(常時的) NPMSGB 필터가 된다. 망각인수가 $0 < \beta < 1$ 의 범위에 있을 때 $\lim_{n \rightarrow \infty} \gamma(n) = 0$ 이 되어 점근적(漸近的) NPMSGB 필터가 된다. NPMSGB 필터는 정규화 확률적 경사도 필터가 되므로, 정규화 확률적 경사도 필터는

MSGB 필터의 특수한 경우인 상시적 NPMSGB 필터로 볼 수 있다.

가변적 매개변수 $\gamma(n)$ 이 상수 α 가 될 때, MSGB 필터를 고정 매개변수의 MSGB(fixed parameterized MSGB: FPMSGB) 필터라고 부른다. 망각 인수 β 가 상수 1일 때 $\gamma(n) = \alpha$ 가 되어 상시적 FPMSGB 필터가 된다. 망각인수가 $0 < \beta < 1$ 의 범위에 있을 때 $\lim_{n \rightarrow \infty} \gamma(n) = \alpha$ 가 되어 점근적 FPMSGB 필터가 된다. FPMSGB 필터는 상수의 매개변수를 가졌기 때문에 정규화 momentum LMS 필터라고도 부른다.

망각 인수가 $0 < \beta < 1$ 의 범위에 있을 때 MSGB 필터를 가변 매개변수의 MSGB(variably parameterized MSGB: VPMSGB) 필터라고 부른다. 망각 인수가 $0 \leq \beta \leq 1$ 의 범위에 있을 때 MSGB 필터를 넓은 의미의 VPMSGB(wide-sense VPMSGB: WS VPMSGB) 필터라고 부른다. WS VPMSGB 필터는 VPMSGB, NPMSGB와 FPMSGB의 필터들을 모두 함께 포함하므로 포괄적 개념(comprehensive concept)을 가진 명칭이다.

필터의 특성상 가변적 축척 매개변수 $\gamma(n)$ 이 1에 접근할 때 저역통과 필터의 구조를 가지고, $\gamma(n)$ 이 -1에 접근할 때 고역통과 필터의 구조를 가진다. 식 (3)에서 경사도를 알아내기가 어렵기 때문에 유한한 입력 신호를 이용하여 각 반복순환에서 경사도가 불완전하게 추정되고 있으며 이로 인하여 경사도 잡음이 발생하여 증폭되므로 어려움을 겪게 된다[6]. 그러나 MSGB 필터는 입력신호의 Euclidean norm의 자승으로 정규화되므로 이런 문제가 극복될 수 있다.

병진적 필터계수를 $V(n) = W(n) - W_0$ 라고 할 때, $W_0 = R^{-1}P$ 는 이산적 Wiener-Hopf 필터의 최적 계수벡터가 되고, 다음 식이 성립한다.

$$V(n+1) = \frac{2}{\delta + X^H(n)X(n)} e_o^H(n)X(n) + [1 + \gamma(n)] - 2 \frac{\mu}{\delta + X^H(n)X(n)} X(n)X^H(n) V(n) - \gamma(n) V(n-1) \quad (6)$$

위 식에서 최적 출력오차는 $e_o(n) = d(n) - W_0^H(n)X(n)$ 이다. 그러나 위의 Wiener-Hopf 필터에서 선형위상 특성을 보장하지는 않는다. 최적 출력오차가 입력신호에 직교하므로, 비용함수가 최소값에 도달하기 위한 필요충분조건을 충족시킨다. 이는 직교성 정리를 구성하고 있다고 볼 수 있기 때문에, 본 필터는 최적 조건 하에서 동작한다고 말할 수 있다.

등척적 필터계수를 $\tilde{V}(n) = Q^H V(n)$ 이라고 두고, 위 식에 Q^H 를 곱하면 다음 식이 성립한다.

$$\tilde{V}(n+1) = \frac{2}{\delta + X^H(n)X(n)} e_o^H(n)\tilde{X}(n) + [1 + \gamma(n)] - 2 \frac{\mu}{\delta + X^H(n)X(n)} \tilde{X}(n)\tilde{X}^H(n) \tilde{V}(n) - \gamma(n)\tilde{V}(n-1) \quad (7)$$

위 식에서 $\tilde{X}(n)$ 은 등척적 입력 벡터이고, $\tilde{X}(n) = Q^H X(n)$ 이다.

가변적 축척 매개변수의 공간에서 위 필터가 최소값에 도달하면 학습은 끝난다. 그러나 국소 최소값에 도달하면, 출력 오차는 용인될 수 없을 정도로 크다. 위 필터가 허용될 수 있는 최적값에 도달하기 전에 학습을 중지한다면, 위 매개변수를 변화시키거나 필터계수의 초기값을 바꿈으로써 그 문제를 해결할 수 있다.

III. 평균적 수렴 특성

양태 행렬 Q 를 unitary 행렬이라고 가정한다면, $E[\tilde{X}(n)\tilde{X}^H(n)]$ 은 고유값의 대각선 행렬 Λ 가 된다. 필터의 임펄스 응답이 최적화 되어 있다면, 오차 신호는 입력신호와 상관관계에 있지 않으므로 직교성 정리가 성립한다고 가정한다. 필터계수 $W(n)$ 과 입력 신호 $X(n)$ 은 서로 독립적이라고 가정한다. 또한, 입력 신호 $X(n)$ 은 랜덤 분포를 가진다고 가정한다. MSGB 필터에서 δ 도 0에 접근한다고 가정한다.

3.1 무매개변수의 MSGB 필터

망각 인수 β 가 0일 때 MSGB 필터는 무매개변수의 MSGB (NPMSGB) 필터가 된다. NPMSGB 필터는 $\gamma(n)$ 이 0이 된다. 따라서, NPMSGB 필터는 정규화 확률적 경사도 필터가 되므로, 정규화 확률적 경사도 필터는 MSGB 필터의 특수한 경우인 NPMSGB 필터로 볼 수 있다. 식 (7)로부터 평균을 취하여 NPMSGB 필터의 평균적 순환식을 구하고 극한 값을 구하면 다음과 같은 식을 구할 수 있다.

$$E[\tilde{V}(n+1)] = [I - 2\mu[\text{tr} R]^{-1} \Lambda] E[\tilde{V}(n)] \\ \lim_{n \rightarrow \infty} [I - 2\mu[\text{tr} R]^{-1} \Lambda]^n = 0 \\ \text{diag}[\lim_{n \rightarrow \infty} (1 - 2\mu[\text{tr} R]^{-1} \lambda_i)^n] = 0 \quad (8)$$

위 식으로부터 NPMSGB 필터가 안정하기 위한 조건은 다음과 같다.

조건1: $0 < \mu < \frac{\text{tr}R}{\lambda_{\max}}$

NPMSGB 필터의 안정도 한계는 정규화 확률적 경사도 필터의 안정성 조건이 된다.

3.2 고정 매개변수의 MSGB 필터

β 가 상수 1일 때 MSGB 필터는 고정 매개변수의 MSGB (FPMSGB) 필터가 된다. FPMSGB 필터에서 매개변수 $\gamma(n)$ 는 상수 α 가 된다. FPMSGB 필터는 상수의 매개변수를 가지기 때문에 정규화 momentum LMS 필터로 볼 수 있다. 따라서 정규화 momentum LMS 필터는 MSGB 필터의 특수한 경우로 해석할 수 있다. 비정규화 momentum LMS 필터의 해석은 논문 [7]에서 발표되었으므로 여기서는 FPMSGB 필터에 국한시켜 해석한다. 식 (7)으로부터 평균을 취하면, 다음 식이 성립한다.

$$E[\tilde{V}(n+1) - \tilde{V}(n)] = -2\mu[\text{tr}R]^{-1}AE[\tilde{V}(n)] + \alpha E[\tilde{V}(n) - \tilde{V}(n-1)] \quad (9)$$

2N 차원 상태 신호 $\tilde{U}(n)$ 를 다음과 같이 정의한다.

$$\tilde{U}(n) = [\tilde{V}(n-1) \ \tilde{V}(n)]^T \quad (10)$$

상수 행렬 A 와 B 를 다음과 같이 정의하자.

$$A = -\alpha I$$

$$B = (1+\alpha)I - 2\mu[\text{tr}R]^{-1}A \quad (11)$$

위 식을 이용하면, 식 (9)는 다음과 같은 순환식으로 고쳐 쓸 수 있다.

$$E[\tilde{U}(n+1)] = CE[\tilde{U}(n)] \quad (12)$$

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

$$C = \begin{bmatrix} 0 & I \\ A & B \end{bmatrix} \quad (13)$$

여기서 0 는 NxN 영행렬이다. 평균적 상태 신호 $E[\tilde{U}(n)]$ 이 수렴하기 위하여, 위 행렬 C 의 모든 고유값이 단위원 내에 있어야 한다. FPMSGB 필터가 안정하고 수렴할 조건을 변이 인수와 축척 매개변수와 관련하여 구하도록 하자. C 의 임의의 고유

값을 ζ 라고 하고 그것에 해당하는 고유벡터를 $q^T = [q_1^T \ q_2^T]$ 라고 하자. 여기서 q_1^T, q_2^T 는 N차원 열벡터이다. 그러면 $Cq = \zeta q$ 의 방정식을 얻는다. 이를 다시 쓰면

$$\begin{bmatrix} 0 & I \\ A & B \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \zeta \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \quad (14)$$

이므로

$$Aq_1 + \zeta B q_2 = \zeta^2 q_1 \quad (15)$$

을 얻는다. 위 식에서 상수 행렬을 치환하면 아래와 같이 다시 쓸 수 있다.

$$-\alpha q_1 + \zeta[(1+\alpha)I - \mu(1-\alpha)[\text{tr}R]^{-1}A]q_2 = \zeta^2 q_1 \quad (16)$$

위 식은 N 개의 스칼라 방정식으로 다음과 같이 쓸 수 있다.

$$\zeta^2 - [(1+\alpha) - 2\mu[\text{tr}R]^{-1}\lambda_i]\zeta + \alpha = 0, \quad i = 0, 1, \dots, N-1 \quad (17)$$

위 식에서 2N개의 근을 결정하기 위하여, 다음과 같은 2차 다항식을

$$\zeta^2 - (1+\alpha - 2\mu[\text{tr}R]^{-1}\lambda_i)\zeta + \alpha = 0 \quad (18)$$

을 조사할 필요가 있다. 위 이차 다항식의 안정성을 판정하기 위하여, 수정된 Schur-Cohn 기준(criterion)을 이용하여 검증하면 [8], 다음과 같은 안정성 조건(stability condition)을 유도할 수 있다.

조건 2 : $0 < \mu < \frac{(1+\alpha)\text{tr}R}{\lambda_{\max}}$

조건 3 : $-1 < \alpha < 1$

초기 축척 인수 α 가 0 에 접근할 때, MSGB 필터에서 위 제한 조건 2과 3는 위와 같은 NSGB 필터의 안정성 조건 3으로 귀결한다.

초기 축척 인수 α 가 음수일 때, 위 안정성 조건에서 α 과 μ 가 안정하기 위한 범위는 다음과 같이 쓸 수 있다.

$$\begin{aligned} -1 + 2\mu[\text{tr}R]^{-1}\lambda_i &< \alpha < 0 \\ 0 < \mu < (1-|\alpha|)\frac{\text{tr}R}{\lambda_i} \end{aligned} \quad (19)$$

초기 축척 인수 α 가 0에 접근할 때, 안정도의 한계는 $0 < \mu < [\text{tr} R]/\lambda_{\max}$ 로 증가한다. 그러나 α 가 -1에 접근할 때, 안정도의 한계는 0으로 감소한다. 또한 α 가 음수일 경우 필터가 안정하더라도 그 성능을 저하시킨다. 즉, 전 필터계수 갱신량이 음의 방향에 있을 때, 그 필터계수 갱신량의 음의 분수를 그 다음의 필터계수 갱신량에 첨가하기 때문에 수렴 속도가 저하될 수 있다.

초기 축척 인수 α 가 양수일 때 및 $-1 + \mu[\text{tr} R]^{-1}\lambda_i > 0$ 일 때, 다음과 같은 부등식이 성립한다.

$$-1 + \mu[\text{tr} R]^{-1}\lambda_i < \alpha < 1$$

$$\frac{\text{tr} R}{\lambda_i} < \mu < (1 + \alpha)\frac{\text{tr} R}{\lambda_i} \quad (20)$$

초기 축척 인수 α 가 0에 접근할 때 안정도 한계는 $\text{tr} R/\lambda_{\min} < \mu < \text{tr} R/\lambda_{\max}$ 으로 감소하고, 초기 축척 인수 α 가 1에 접근할 때 안정도 한계는 $\text{tr} R/\lambda_{\min} < \mu < 2\text{tr} R/\lambda_{\max}$ 가 된다.

초기 축척 인수 α 가 양수일 때 및 $-1 + \mu[\text{tr} R]^{-1}\lambda_i < 0$ 일 때, 다음과 같은 부등식이 성립한다.

$$0 < \alpha < 1$$

$$\mu < \frac{[\text{tr} R]}{\lambda_i} \quad (21)$$

초기 축척 인수 α 의 값에 무관하게 안정도 한계는 $\mu < [\text{tr} R]/\lambda_{\max}$ 가 된다.

3.3 가변 매개변수의 MSGB 필터

가변 매개변수의 MSGB (VPMSGB) 필터는 초기상태에서 고정 매개변수의 MSGB (FPMSGB) 필터에 근접하기 때문에 FPMSGB 필터의 특성을 가진다. VPMSGB 필터를 설계할 때, 초기상태의 필터 특성에 영향을 주는 것은 FPMSGB 필터이다. 초기상태에서의 VPMSGB 필터의 안정성은 FPMSGB 필터의 안정성으로 근사화될 수 있다.

VPMSGB 필터는 정상상태에서 NPMSGB 필터에 근접하기 때문에 NPMSGB 필터의 특성을 가진다. VPMSGB 필터를 설계할 때, 정상상태의 필터 특성에 영향을 주는 것은 FPMSGB 필터이다. 정상상태에서의 VPMSGB 필터의 안정성은 NPMSGB 필터의 안정성에 근사적으로 접근한다.

가변 매개변수의 MSGB (VPSGB) 필터는 가변

축척 매개변수(variably scaling parameter)를 통하여 초기상태에서 FPMSGB 필터를 모방하고 정상상태에서 NPMSGB 필터를 모방하므로, FPMSGB 필터와 NPMSGB 필터를 연결하는 교량의 역할을 한다는 것을 알 수 있다.

IV. 오조정량

오조정량(誤調整量)은 상대적인 초과 평균 지수 오차를 말하고, 이는 MSGB 필터에 의하여 계산된 정상 상태의 해가 Wiener-Hopf 필터계수로부터 얼마나 떨어져 있는지를 표시하는 정량적 지수가 된다.

초기에 가변 매개변수의 MSGB (VPMSGB) 필터는 FPMSGB 필터처럼 동작한다. 이때 고정 매개변수를 모방한 가변 매개변수가 그 필터의 수렴속력을 증가시키고 오조정량도 증가시킨다. 초기상태에서의 MSGB 필터는 Wiener-Hopf 필터의 최적 계수값에 도달하기 전이므로 필터의 수렴속도의 증가는 장점으로 작용하여 주요한 의미를 지니지만, 초기상태에서 오조정량의 증가는 부차적인 문제로서 아무런 의미를 가지지 않으며 어떤 흠결이 되는 것은 아니다.

초기에 VPMSGB 필터의 오조정량 Ξ_{VPMSG}^0 은 FPMSGB 필터의 오조정량 Ξ_{FPMSG} 으로 근사화할 수 있으므로 다음과 같이 쓸 수 있다.

$$\Xi_{VPMSG}^0 = \Xi_{FPMSG}$$

$$= \frac{\mu \sum_{i=1}^N \frac{(1+\alpha)[\text{tr} R]^{-1}\lambda_i}{1-\alpha^2-2\mu[\text{tr} R]^{-1}\lambda_i}}{1-\mu \sum_{i=1}^N \frac{(1+\alpha)[\text{tr} R]^{-1}\lambda_i}{1-\alpha^2-2\mu[\text{tr} R]^{-1}\lambda_i}} \quad (22)$$

변이 인수가 작고 축척 매개변수도 작다고 가정하면, $2\mu[\text{tr} R]^{-1}\lambda_i \ll 1$ 이 성립하게 되고, 위 식은 근사적으로 다음과 같이 쓸 수 있다.

$$\Xi_{VPMSG}^0 = \Xi_{FPMSG}$$

$$\approx (1+\alpha)\mu \sum_{i=1}^N \frac{[\text{tr} R]^{-1}\lambda_i}{1-2\mu[\text{tr} R]^{-1}\lambda_i} \quad (23)$$

위 식으로부터 FPMSGB 필터의 오조정량은 축척 매개변수가 양수일 때 그 축척 매개변수에 비례한다. 정상상태에서 가변 매개변수의 MSGB(VPMSGB) 필터는 NPMSGB 필터처럼 동작한다. 이때 무매개변수를 모방한 가변 매개변수가 그 필터의 오조정량을

감소시키고 수렴속력도 감소시킨다. 정상상태에서의 MSGB 필터는 Wiener-Hopf 필터의 최적 계수값에 도달한 후이므로 필터의 오조정량의 감소는 장점으로 작용하여 주요한 의미를 지니지만, 수렴속도의 감소는 부차적인 문제로서 아무런 문제를 일으키지 않는다.

정상상태에서 VPMSGB 필터의 오조정량 Ξ_{VPMSG}^{∞} 은 NPMSGB 필터의 오조정량 Ξ_{NPMSG} 으로 근사화할 수 있으므로 다음과 같이 쓸 수 있다.

$$\Xi_{VPMSG}^{\infty} = \mu \sum_{i=1}^N \frac{[\text{tr} R]^{-1} \lambda_i}{1 - 2\mu[\text{tr} R]^{-1} \lambda_i} \quad (24)$$

변이 인수가 작다고 가정한다면, $2\mu[\text{tr} R]^{-1} \lambda_i \ll 1$ 이 성립하게 되고, 위 식은 다음과 같이 쓸 수 있다.

$$\Xi_{VPMSG}^{\infty} \approx \mu \sum_{i=1}^N [\text{tr} R]^{-1} \lambda_i = \mu \quad (25)$$

위 식으로부터 정상상태에서 VPMSGB 필터의 오조정량 Ξ_{VPMSG}^{∞} 은 변이인수 μ 에 비례한다.

VPMSGB 필터는 탭 필터계수 주위에 작용하는 역방향 케환 환상회로(loop)는 저역통과 필터처럼 작용하고 그 평균적 시간 상수도 변이인수 및 축척 매개변수에 반비례한다. 변이인수에 작은 값을 할당할 때 적응과정은 느리게 진행하고 탭 필터계수에 대한 경사도 잡음의 효과는 크게 여파된다. 이는 결과적으로 오조정량을 감소시키는 효과를 초래한다. 그러므로 MSGB 필터는 외부환경에 가변적으로 적응함으로써 향상된 성능을 실현할 수 있음을 알 수 있다.

VPMSGB 필터는 초기에 FPMSGB 필터처럼 동작하여 수렴속력을 증가시킨다. 이는 VPMSGB 필터는 NPMSGB 필터에 없는 가변적 매개변수로 인하여 계산량이 증가하고 부가적인 기억 공간을 필요로 하기 때문이다. VPMSGB 필터는 정상상태에서 NPMSGB 필터처럼 동작하여 오조정량을 증가시킨다. VPMSGB 필터는 그 가변적 매개변수로 인하여 수렴 속력을 증가시키고 오조정량을 감소시켜 결과적으로 수렴 속력과 오조정량을 개선시키는 탁월한 거동 특성을 가진다.

V. 시스템 식별

MSGB 필터를 시스템 식별(system identification)에 응용할 때, 이상 응답 $D(z)$ 과 입력 신호 $X(z)$ 의 관계는

$$D(z) = G(z)X(z) + V(z) \quad (26)$$

이 된다. 위 식에서 $V(z)$ 는 첨가 잡음이고 입력신호 $X(z)$ 과 상호 상관 관계를 갖지 않고, $G(z)$ 는 미지의 시스템(unknown system)이다. 다음의 2차 시스템이 모의실험에 사용되었다.

$$G(z) = k \frac{1 - 0.3z^{-1}}{1 - 0.3z^{-1} + 0.02z^{-2}} \quad (27)$$

위 식에서 $z = 0.1$, $z = 0.2$ 에서 극점을 가지고 $z = 0.3$ 에서 영점을 가진다. 스칼라량 k 는 $G(z)X(z)$ 가 단위 분산보다 크도록 선택되었다. 입력 $x(n)$ 과 잡음 $v(n)$ 은 uniform 분포를 가진 백색 랜덤 수열이고 그 평균이 0 이고 분산이 각각 $\text{var}(X) = 1$, $\text{var}(V) = 0.001$ 이다. 그 때 각 입력 신호의 평균 전력은 1에 수렴한다.

아래의 그림들은 시스템 식별에 응용한 컴퓨터 모의실험에서 MSGB 필터를 사용하여 자승오차를 구하고 이에 대한 학습 곡선을 보여준다. 이 곡선은 컴퓨터를 사용하여 여러 번의 독립 시행을 실시하여 최적 값에 도달한 필터의 출력 오차 $e(n)$ 의 자승을 얻은 결과이다. 이 경우에 변이 인수 μ 는 0.001, 0.002, 0.005 이었고 필터계수의 차수는 N 은 8개였다. 미지의 시스템 $G(z)$ 에서 단위 분산보다는 약간 크도록 하기 위하여 스칼라량 k 를 1로 선정하였다.

그림 1에서 NPMSGB 필터가 변이 인수 0.002와 축척 매개변수 0을 가질 때, 그 필터의 자승오차의 학습곡선을 보여준다. 그림 2에서 VPMSGB 필터가 변이 인수가 0.002이고 축척 매개변수의 초기인수가 $\alpha = 0.8$ 이고, 그 망각인수가 $\beta = 0.985$ 가지고 있을 때 그 필터의 자승오차의 학습곡선을 보여준다. 도시된 VPMSGB 필터는 이전에 도시된 NPMSGB 필터에 비하여 동일한 조건에서 MSGB 필터의 수렴속도가 현저히 개선된 것을 알 수 있다.

그림 3에서 VPMSGB 필터가 변이 인수 0.001을 가지고 축척 매개변수의 초기인수가 $\alpha = 0.9$, 그 망각인수가 $\beta = 0.985$ 을 가질 때, 그 필터의 자승오차의 학습곡선을 도시한다. 그림 4에서 VPMSGB 필터가 변이 인수가 0.001를 가지고 축척 매개변수의 초기인수가 $\alpha = 0.9$, 그 망각인수가 $\beta = 0.900$ 를 가질 때, 그 필터의 자승오차의 학습곡선이 도시되어 있다. 그림 3의 가변 MSGB 필터는 그림 4의 필터에 비하여 인수 및 축척 매개변수의 초기인수 α 를 고정시키고 그 망각인수 β 를 변경시켜 수렴속력을 개선시킨 것을 보여준다.

그림 5에서 변이 인수 μ 가 0.005 일 때, FPMSGB 필터는 축척 매개변수의 초기인수가 $\alpha = 1.0$ 을 가지고 있으며, VPMSGB 필터는 그 초기인수가 $\alpha = 1.0$ 이고, 그 망각인수가 $\beta = 0.99$ 을 가지고 있을 때, 그 필터의 자승오차의 학습곡선을 각각 보여 주고 있다. FPMSGB 필터는 정규화 momentum LMS 가 되고, 초기 인수 α 가 1이 되면 FPMSGB 필터가 그 안정성 조건을 만족시키지 않게 되므로 불안정하게 되어 발산한다는 것을 보여주고 있다. 그러나 변이인수와 초기인수의 동일한 조건 하에서 망각인수 $\beta = 0.99$ 를 도입하여 축척 매개변수를 가변시켜 주면 불안정한 FPMSGB 필터도 안정한 영역에 들어가게 되고 수렴속력도 개선되며 원활한 수렴 특성을 얻을 수 있다. VPMSGB 필터에서 망각인수가 크면 FPMSGB 필터의 특성에 근접하여 그 필터의 수렴속도가 증가하는 것을 알 수 있다. VPMSGB 필터는 FPMSGB 필터의 안정성 조건을 벗어난 영역에서도 안정적으로 동작하도록 유도할 뿐만 아니라 적절한 매개변수를 설정하여 수렴속력도 개선하고 오조정량도 줄일 수 있다.

앞 절에서 기술된 안정도 조건을 만족하는 MSGB 필터에서 최적의 가변 매개변수를 구하는 방법은 위 안정도 조건을 만족하는 범위 내에서 그 최적의 매개변수를 경험적 방법으로 찾아내어야 한다. 즉, 이전의 절에서 가변 매개변수에 대한 상한($\lim_{n \rightarrow \infty} \gamma(n)$)과 하한

($\lim_{n \rightarrow \infty} \gamma(n)$) 에서 안정성 조건 및 안정한 범위를 해석적인 방법을 이용하여 기술하였고, 본 모의실험에서는 변이인수가 주어졌을 때 적절한 가변 매개변수를 경험적 방법으로 구하여 보여 주었다.

VPMSGB 필터는 FPMSGB 필터에 비하여 가변적 매개변수로 인하여 계산량이 증가하고, 부가적인 기억 공간을 필요로 한다. 그러나 VPMSGB 필터는 초기상태에서 FPMSGB 필터처럼 동작하여 수렴속력을 증가시키고, 정상상태에서 NPMSGB 필터처럼 동작하여 오조정량을 감소시킨다. 그 가변적 매개변수로 인하여 초기에 수렴 속력을 증가시키고 나중에 오조정량을 감소시켜 결과적으로 수렴 속력과 오조정량을 개선시키는 효과가 있다.

VPMSGB 필터가 가변적 축척 매개변수의 공간에서 최소값에 도달하지 못하여 학습이 완료되지 못하였다. 예를 들면, 국소 최소값에 도달하여 출력 오차가 용인될 수 없을 정도로 커지기도 하였다. 또한 열악한 환경에서 여러 가지 이유로 위 필터가 허용될 수 있는 최적값에 도달하지 못할 정도로 학습할 수 없는 경우도 발생하였다. 본 논문에서 제시된 각종 매개변수들을 변화시키거나, 필터계수의 초기값을 바꾸거나, 또는 가변적 매개변수를 또 다른 함수로 표현함으로써 그 문제를 해결할 수 있었다.

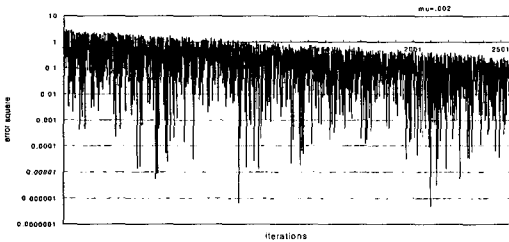


그림 1. NPMSGB 필터에 대한 자승 오차의 학습 곡선 ($\gamma = 0$)
Figure 1. Error squared learning curve for NPMSGB filter with $\gamma = 0$

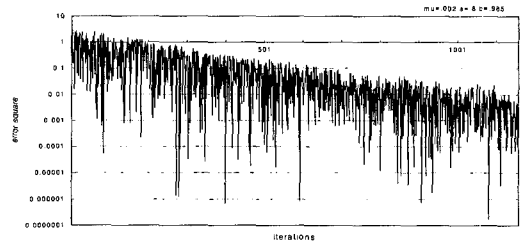


그림 2 VPMSGB 필터에 대한 자승 오차의 학습 곡선 ($\alpha = 0.8, \beta = 0.985$)
Figure 2. Error squared learning curve for VPMSGB filter with $\alpha = 0.8, \beta = 0.985$

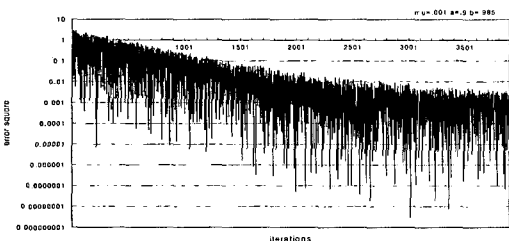


그림 3 VPMSGB 필터에 대한 자승 오차의 학습 곡선 ($\alpha = 0.9, \beta = 0.985$)
Figure 3. Error squared learning curve for VPMSGB filter with $\alpha = 0.9, \beta = 0.985$

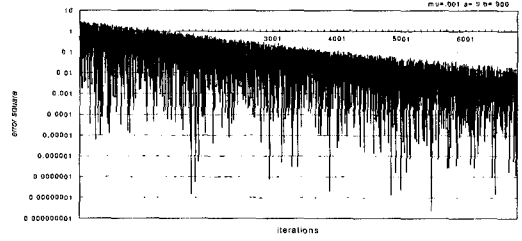


그림 4 VPMSGB 필터에 대한 자승 오차의 학습 곡선 ($\alpha = 0.9, \beta = 0.900$)
Figure 4. Error squared learning curve for VPMSGB filter with $\alpha = 0.9, \beta = 0.900$

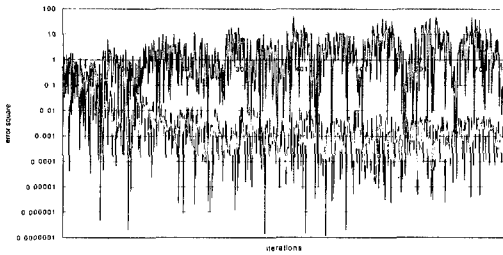


그림 5. FP 및 VP MSGB 필터에 대한 자승 오차의 학습 곡선 ($\alpha = 1.0, \beta = 0.99$)
 Figure 5. Error squared learning curves for FPMSGB and VPMSGB filters with $\alpha = 1.0, \beta = 0.99$

VI. 결론

본 논문에서 복소수 영역에서 MSGB 필터를 제안하였고, 이는 최적화 문제에 대한 해결 방안이 될 수 있음을 살펴보았다. MSGB 필터는 가변적 매개변수의 상태에 따라 NPMSGB, FPMSGB, VPMSGB 필터 등으로 분류하였다. 평균 수렴특성을 해석하면서, NPMSGB, FPMSGB, VPMSGB 등의 세가지 MSGB 필터에 관하여 다양한 안정성 조건을 구하고 서술하였다. VPMSGB 필터는 초기 상태에서 FPMSGB 필터의 특성을 가지고 그 필터의 안정성도 FPMSGB 필터의 안정성으로 근사화될 수 있음을 알 수 있었다. VPMSGB 필터는 정상상태에서 NPMSGB 필터의 특성을 가지고 그 필터의 안정성도 NPMSGB 필터의 안정성으로 근사화될 수 있음도 알 수 있었다. 세가지 MSGB 필터에 관한 오조정량을 필터 내부의 여러 가지 매개변수로 표현하였다. VPMSGB 필터는 초기 상태에서 FPMSGB 필터를 모방하고 그 필터의 오조정량은 FPMSGB 필터의 오조정량으로 근사화될 수 있음을 알 수 있었다. VPMSGB 필터는 정상상태에서 NPMSGB 필터를 모방하고 그 필터의 오조정량도 NPMSGB 필터의 오조정량으로 근사화될 수 있음도 알 수 있었다. 세가지 MSGB 필터를 시스템 식별에 응용하여 컴퓨터 모의실험을 실행하여 자승 오차의 학습곡선을 보여주었다. VPMSGB 필터는 초기 상태에서 FPMSGB 필터처럼 동작하여 그 필터의 수렴속력을 증가시키고, 정상 상태에서 NPMSGB 필터처럼 동작하여 그 필터의 오조정량을 감소시켰는데, 그 결과 수렴속력과 오조정량을 개선한 효과가 있음을 알 수 있었다.

참고 문헌

- [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] A. E. Albert and L. S. Gardner, Jr., Stochastic Approximation and Nonlinear Regression, Cambridge, MA: MIT Press, 1967.
- [4] 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.
- [5] S. C. Douglas, "A family of normalized LMS algorithms," IEEE Trans. Signal Processing Letters, vol. SP-1, pp.49-251, March. 1994.
- [6] 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.
- [7] S. Roy and J. Shynk, "Analysis of the momentum LMS algorithm," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 38, no.12, pp.2088-2098, December. 1990.
- [8] E. I. Jury, Theory and Application of the z-Transform Method, New-York: Wiley, 1964.

김 해 정 (Haejung Kim)

정회원



1984년 2월 경북대학교 전자공학과 졸업

1986년 8월 한양대학교 전자공학과 석사

1995년 2월 한양대학교 전자공학과 박사

1997년 3월~1999년 2월 서경대학교 컴퓨터공학과 조교수

1999년 3월~현재 서경대학교 정보통신공학과 조교수
 <관심분야> 신호처리, 통신공학