

새로운 성능지수 함수에 대한 직강하 적응필터

전 병 을, 박 동 조

한국과학기술원 전기 및 전자공학과

Novel Steepest Descent Adaptive Filters Derived from New Performance Function

Jun, Byung-Eul and Park, Dong-Jo

Dept. of Elect. Eng., Korea Advanced Institute of Science and Technology

Abstract

A novel steepest descent adaptive filter algorithm, which uses the instantaneous stochastic gradient for the steepest descent direction, is derived from a newly devised performance index function. The performance function for the new algorithm is improved from that for the LMS in consideration that the stochastic steepest descent method is utilized to minimize the performance index iteratively. Through mathematical analysis and computer simulations, it is verified that there are substantial improvements in convergence and misadjustments even though the computational simplicity and the robustness of the LMS algorithm are hardly sacrificed. On the other hand, the new algorithm can be interpreted as a variable step size adaptive filter, and in this respect a heuristic method is proposed in order to reduce the noise caused by the step size fluctuation.

I 서론

폴렌트 모델링, 잡음상쇄(noise canceling) 및 채널 동화(channel equalization) 등을 위한 적응필터로 가장 널리 사용되고 있는 것은 최소 평균 자승(LMS : least mean square) 알고리즘 이다[1][2]. LMS 적응필터는 원하는 값(desired value)과 필터 출력 사이의 오차의 자승(square error)을 직강하법(steepest descent method)에 의해 반복적으로(iteratively) 줄여가는데, 알고리즘이 간결하고 계산량이 적을 뿐더러 수치적으로 강인하여(numerically robust) 디지털 구현이 용이한 반면 수렴속도가 느린 것이 결함이다. 수렴속도가 빠른 알고리즘으로는 최소 자승법(least square method)에 의한 적응필터 종류가 있으나, 이들은 계산량이 많고, 수치적인 안정화(numerical stabilization)에 대한 세심한 주의가 필요하므로 사용에 어려움이 있을 수 있다. 그러므로 여러 가지의 적응 예를 통하여 안정성 및 강인성이 입증된 LMS 알고리즘의 수렴성을 개선하여 향상된 성능을 갖는 적응필터를 얻으려는 시도가 계속되고 있다. 특히 계산량을 크게 늘리지 않으면서 LMS 필터의 성능을 개선하고자 하는 방식은 대략 두가지로 분류할 수 있는데, 한가지는 LMS 알고리즘의 설계상수인 보폭(step size)을 필터의 상태 및 필터링 환경에 따라 적절히 변화시키는 방식[4][5]을 들 수 있고, 다른 한가지로는 LMS 필터가 필터의 자승오차에 대한 직강하 알고리즘이라는 것으로부터, 직강하법을 적용시키는 대상이 되는 성능지수(performance index) 함수에 대한 변형을 시도하는 것으로, 예를 들면 오차의 자승을 일반화한 오차의 착수승을 성능지수로 선택한 경우[3] 등이 있다. 본 논문에서는 위에서 분류한 두 방법 중 후자의 관점에서 볼 때, LMS 필터에서 성능지수로 사용하는 오차의 자승함에 오차자승의 지수함수가 더해진 새로운 형태의 성능지수 함수를 생각하고, 이 성능지수 함수에 대해 직강하법을 적용하여 새로운 적응필터를 유도한다. 이렇게 얻어진 적응필터는 LMS 알고리즘 성능개선의 첫번째 방식인 가변보폭(variable step size) 개념으로 설명할 수도 있는데, 특히 필터링 초기 단계나 필터의 환경이 갑자기 변하는 등의 요인에 의해 필터링 오차가 큰 상태에서 알고리즘의 안정성

(stability)을 보장할 수 있는 설계상수의 결정에 이러한 해석을 이용할 수 있고, 또 이 가변보폭 개념으로부터 제안된 필터의 성능을 개선하기 위한 몇가지 방안을 제시할 수 있다. 제시한 적응필터 알고리즘의 성능분석 단계에서는, 우선 필터링 초기 상태에 대한 해석으로부터 알고리즘을 안정화하는 설계 상수의 결정에 대해 생각해 보고, 필터가 최적해 가까이 수렴된 상태에 대한 통계적인 오차 특성에 대해 알아본다. 다음에는 폴렌트 모델링 예제 및 적응 동화기(adaptive equalizer) 예제에 대한 컴퓨터 시뮬레이션을 통해 수식적인 해석에 대한 결과와 비교하고, 수식적인 해석에서 다루기 어려웠던 기타의 특성들에 관해 살펴본다.

II LMS 적응필터

개선된 형태의 적응필터를 제안하기에 앞서 LMS 알고리즘 및 그 특성에 대해 간략히 살펴보자. 우선 전형적인 적응필터 응용예의 하나인 그림 1과 같은 모델링 문제를 생각하자. 여기에서 모델링 예제를 고려한 것은 문제를 특수한 경우에 대해 축소하기 위한 것이 아니고 단지 설명의 편의를 위한 것이므로, 앞으로 전개되는 이야기는 폴렌트 모델링 이외의 적응필터링 문제에 대해서도 적절한 접근에 의해 적용될 수 있다. 그림 1에서 폴렌트는 다음 식(1)과 같이 FIR(finite impulse response) 필터에 백색잡음(white noise)이 더해진 형태로 표현할 수 있다고 가정한다.

$$d(t) = W^*T U(t) + \zeta(t) \quad (1)$$

식(1)에서 $W^* \in R^N$, $U \in R^N$ 는 각각 FIR 필터의 계수(혹은 가중치) 및 입력 벡터이고, ζ 는 백색잡음이다. 식(1)의 폴렌트를 다음 식(2)의 FIR 필터로 모델링 하고자 한다.

$$y(t) = W^T(t)U(t), \quad W \in R^N \quad (2)$$

식(2)의 필터가 식(1)로 주어진 폴렌트를 적절히 묘사하도록 하는 필터 계수를 찾기 위한 한 방법으로 적응 필터를 사용할 수 있다. 폴렌트와 필터 출력 사이의 오차를 $e(t) = d(t) - y(t)$ 로 정의하면, LMS 적응필터는 성능지수함수

$$J_{LMS} = \frac{1}{2}e^2(t) \quad (3)$$

를 직강하법에 의해 반복적으로 최소화하는 알고리즘으로 다음과 같이 쓸 수 있다.

$$W(t+1) = W(t) + \mu e(t)U(t) \quad (4)$$

식(4)로 주어지는 LMS 알고리즘은 보폭(μ)을 다음 식(5),(6)의 범위에서 사용할 때에 필터의 평균오차 및 평균자승오차가 수렴하는 것으로 알려져 있다[6][7][8].

$$0 < \mu < \frac{1}{\lambda_i}, \quad i = 1, \dots, N \quad (5)$$

$$\sum_{i=1}^N \frac{2\mu\lambda_i}{1-\mu\lambda_i} < 1, \quad (6)$$

위 식에서 λ_i 는 $R = E[U(t)U^T(t)]$ 로 정의되는 필터입력 $U(t)$ 의 공분

산 행렬(covariance matrix)와 i 번 제 고유값이고, 여기에서 규정한 범위에서 보폭이 사용될 때 LMS 필터의 평균 오차 및 평균자승오차는 다음과 같이 계산된다[6].

$$E[\epsilon(t)] = 0 \quad (7)$$

$$E[\epsilon^2(t)] = E[\zeta^2(t)] + \text{tr}(RK(t)) \quad (8)$$

위 식(8)에서 $K(t)$ 는 FIR 필터의 계수오차($V(t) = W(t) - W^*$)의 공분산행렬($K(t) = E[V(t)V^T(t)]$)로 정의되었다. 식(7),(8)은 필터의 통계적특성(statistic characteristics)을 분석하는데 일반적으로 사용되는 기본 가정(fundamental assumption) [6]에 따라 계산된 것으로, 이 기본 가정에 따라 계산되는 LMS 필터의 성질들을 독립이론(independent theory)[6] 이라 하는데, 앞으로 제시할 적응 필터의 분석에서도 이 가정을 이용할 것이다. 식(8)에서 우변 첫째 항은 최적해(optimal solution or Wiener solution)에 의해서도 추정할 수 없는 최소 평균자승오차(MMSE : minimum mean square error)이고, 두 번째 항은 알고리즘에 따라서 달라질 수 있는 초과 평균자승오차(EMSE : excess mean square error)이다. 초과 평균자승오차를 최소 평균자승오차로 나누어 준 값은 적응필터의 평균자승오차 특성을 나타내는 척도로 사용할 수 있는데, 이를 미조정량(misadjustments) 이라 하고 다음과 같이 정의한다.

$$M = \frac{EMSE}{MMSE} \quad (9)$$

LMS 적응필터가 최적해에 가까이 수렴된 상태에서의 미조정량은 다음과 같이 주어진다[6][7].

$$M_{LMS} = \frac{\sum_{i=1}^N \mu \lambda_i / (1 - \mu \lambda_i)}{2 - \sum_{i=1}^N \mu \lambda_i / (1 - \mu \lambda_i)} \quad (10)$$

식(10)은 $\mu \lambda_i \ll 1$ 일 경우에 다음과 같이 간략하게 쓸 수 있다.

$$M_{LMS} \simeq \frac{\mu}{2} \sum_{i=1}^N \lambda_i \quad (11)$$

III 새로운 성능지수 및 적응필터 제안

LMS 적응필터가 성능지수 함수로 자승오차를 사용하는데 반해, 여기에서는 자승오차에 지수함수에 더해진 형태의 식(12)과 같은 새로운 형태의 성능지수를 생각한다.

$$J_{NSD}(t) = \frac{1}{2} \epsilon^2(t) + \beta e^{-\frac{1}{2} \gamma \epsilon^2(t)} \quad (12)$$

새로이 제시한 성능지수 식(12)에서 우변 첫째 항은 포물선으로 두번째 항이 더해짐으로서 주로 필터오차가 0인 부근에서 포물선의 특성을 변화시킬 수 있게 되는데, 특히 식(12)에서 상수 β, γ 가 다음 조건

$$0 < \beta \gamma < 1 \quad (13)$$

을 만족하도록 정해지면 성능지수 식(12)는 항상 컨벡스 함수(convex function)가 되면서 필터오차가 0 부근의 곡률을 현저히 감소시키게 된다. 직강하법에서 곡률의 감소는 경사도 잡음(gradient noise)을 감소시키게 되어 정상상태에서 필터오차를 줄일 수 있도록 한다. 다시 말하면 새로 도입한 식(12)의 성능지수 함수는 오차가 큰 구간에서는 LMS 알고리즘의 특성을 그대로 유지하고, 오차가 작아짐에 따라 임의의 경사도(stochastic gradient)를 사용하는데 따른 필터오차를 줄일 수 있는 자유도를 준다. 식(12)의 성능지수를 반복적으로 최소화해가는 직강하법에 의한 새로운 적응필터는 다음 식(14),(15)로 쓸 수 있다.

$$W(t+1) = W(t) - \mu_o \nabla W(t) \quad (14)$$

$$\begin{aligned} \nabla W(t) &= \frac{\partial J_{NSD}(t)}{\partial W(t)} \\ &= - \left(1 - \beta \gamma e^{-\frac{1}{2} \gamma \epsilon^2(t)} \right) \epsilon(t) U(t) \end{aligned} \quad (15)$$

위 식으로 나타난 필터를 본 논문에서는 편의상 '새로운 직강하 적응필터(new steepest descent adaptive filter)', 혹은 'NSD 적응필터' 등으로 부르기로 한다.

식(15)에서 상수 $\beta \gamma$ 는 $\epsilon(t) \simeq 0$ 에서 성능지수 함수의 곡률을 감소시키기 위해서 가능한 클수록 좋으나, 성능지수가 컨벡스 특성(convexity)을 유지하여야 하므로 대략 다음과 같은 범위에서 정한다.

$$0.9 < \beta \gamma < 1 \quad (16)$$

새로이 얻은 적응필터 알고리즘을 필터오차의 확률분포를 성능지수로 하여 유도한 결과[3]와 비교해 보면, 여기에서 제시한 필터는 오차가 큰 상태(주로 필터링 초기)의 거동에 있어서는 LMS와 거의 다를 바가 없어 그 특성 예측이 용이하므로 안정도를 보장하는 보폭의 크기를 쉽게 정할 수 있다. 반면에 참고문헌[3]에서 제시한 필터들은 오차의 내재급 이상을 성능지수로 할 경우에 초기거동의 예측 및 해석이 힘들기 때문에 안정도를 보장하는 보폭의 크기를 정하기 어렵다. 한편 제안된 알고리즘은 LMS 필터에서 보폭을 변화시키는 한 방법으로 해석할 수도 있는데, 이러한 관점에서 식(14),(15)로 표시된 필터를 다음과 같이 쓸 수 있다.

$$W(t+1) = W(t) - \mu(t) \epsilon(t) U(t) \quad (17)$$

$$\mu(t) = \mu_o \left(1 - \beta \gamma e^{-\frac{1}{2} \gamma \epsilon^2(t)} \right) \quad (18)$$

식(17),(18)의 알고리즘은 참고문헌 [5]에서 제시한 것의 한 변형으로 생각할 수도 있었는데, 여기에서 제시한 알고리즘이 필터의 오차자승($\epsilon^2(t)$)을 지수함수에 사용하고 있는데 반해, [5]의 알고리즘은 LMS 필터에서 사용하는 성능지수의 경사도 노름(gradient norm)에 해당하는 $\|\epsilon(t)U(t)\|$ 을 사용하고 있다. 수직 원리(orthogonal principle)에 따르면 필터의 상태가 최적해에 접근함에 따라 필터 성능지수 함수의 경사도가 0에 접근하는데, LMS 필터에 대해 [5]에서 제시한 보폭변화 방법을 사용하면 참고문헌 [9]에서 밝힌 바와 같이 성능지수 함수가 바뀌게 되므로 $\epsilon(t)U(t)$ 은 더 이상 [5]에서 제시한 알고리즘의 성능지수에 대한 경사도가 아니다. 특히 [5]에서 제시한 가변보폭의 효과는 필터오차가 작을수록 크게 나타나게 되기 때문에, 필터의 수렴이 진행되어 최적해에 접근함에 따라 위에서 지적한 사실은 더욱 명확해지게 되므로 [5]의 알고리즘이 성능지수의 경사도에 대한 지수함수를 사용하고 있다고 보는 건 해는 잘못된 것이다. 한편 $\|\epsilon(t)U(t)\|$ 을 계산하기 위해서는 필터 탭 수만큼의 곱셈이 필요한데 LMS 알고리즘의 계산량을 50%정도 증가시키는 반면, 자승오차 계산은 단 한번의 곱셈이면 되므로 계산량을 거의 증가시키지 않는다. 또한 필터의 차수가 충분히 커서 실제 시스템을 작은 오차로 근사화 할 수 있다면, 자승오차를 사용하더라도 필터가 수렴된 상태에서 보폭의 크기를 매우 작게 하여 오차특성이 개선된다. 만약에 대상 시스템이 FIR 필터로 잘 묘사될 수 없다면, 본 논문에서 논하고 있는 FIR 필터의 구조를 가진 적응필터를 사용하기에 부적합한 문제이므로 다른 접근법을 찾는 것이 타당하다. 그러므로 위에서 대상시스템을 FIR 필터에 의해 잘 묘사할 수 있다고 한 가정은 현실적으로 유효하다. 다음으로는 제안된 적응필터의 성능을 개선할 수 있는 방안에 대해 생각해 보자. 식(17),(18)로 표현된 필터에서 가변보폭 $\mu(t)$ 가 현재시간의 오차(instantaneous error)만에 의해 계산되므로 요동(fluctuation)이 심하게 되고, 이는 보폭의 변화에 따른 오차를 추가적으로 발생시키게 되는데, 이 오차 성분은 보폭의 변화 정도를 작게 함으로서 줄일 수 있다. 제안된 알고리즘의 기본적인 특성을 유지하며 보폭의 변화정도를 줄일 수 있는 간단하고 계산량이 적은 방법으로 보폭 변화의 요인이 되는 오차자승에 포함된 고주파 성분용 식(19)과 같은 필터를 통하여 줄이는 방법을 생각할 수 있다.

$$\bar{\epsilon}^2(t) = \rho \bar{\epsilon}^2(t-1) + (1-\rho) \epsilon^2(t) \quad (19)$$

앞서 제시한 알고리즘의 변형된 형태인 위 식(19)에서 ρ 는 응용 예에 따라 다르겠으나 대략 다음과 같은 정도에서 정해지면 된다.

$$0.8 < \rho < 1.0 \quad (20)$$

IV 제안된 필터의 특성분석

식(14),(15) 혹은 식(17),(18)로 표시된 적응필터 알고리즘의 특성을 수식적인 검토를 통해 살펴보기로 하자. 우선 필터링 초기 또는 필터링 환경의 급격한 변화 등으로 인하여 오차가 큰 상태에서는 식(15) 혹은 식

(18)에서 지수함수를 포함하는 항의 값이 1 에 비해 상대적으로 작아지게 되어 이를 무시할 수 있으므로 필터식을 다음과 같이 쓸 수 있다.

$$W(t+1) \simeq W(t) + \mu_o \epsilon(t) U(t) \quad (21)$$

위 식(21)은 보폭이 μ_o 인 LMS 적응필터를 나타내고 있는데, LMS 알고리즘에서는 보폭 μ_o 값을 적절히 정해 줌으로서 적응필터의 평균오차 및 평균자승오차에 대한 수렴을 보장할 수 있다. LMS 알고리즘의 평균오차 및 평균자승오차가 수렴하기 위한 보폭에 대한 필요충분 조건으로 알려진 것이 식(5),(6) 인데, 이 식은 실제로 사용하기 어렵게 되어 있으므로 다음과 같은 충분조건 식(22)[7][8], 또는 식(23)[10], 혹은 식(6)에서 $\mu_o \lambda_i \ll 1, \forall i$ 를 가정하여 근사화한 다음 식(24) 등이 보폭 결정에 유용하게 사용될 수 있다.

$$0 < \mu_o < \frac{2}{3tr(R)}, \quad (22)$$

$$0 < \mu_o < \frac{2}{2\lambda_{max} + tr(R)} \quad (23)$$

$$0 < \mu_o < \frac{2}{tr(R)}, \quad (24)$$

위 식(24)에서 λ_{max} 는 필터입력에 대한 공분산행렬의 고유치의 최대값을 나타낸다. 다음으로는 필터의 수렴이 진행되어 최적해에 근접한 상태에 도달했을 경우에 제시한 필터의 통계적인 특성을 살펴보기로 하자. 적응필터 식(14) 혹은 식(17)의 양변에서 최적해(W^*)를 빼주면 필터의 가중치와 최적해의 차이에 대한 다음 식(25)을 얻을 수 있다.

$$V(t+1) = V(t) + \mu_o \left(1 - \beta\gamma e^{-\frac{1}{2}\gamma^2 t}\right) \epsilon(t) U(t) \quad (25)$$

위 식(25)은 필터계수의 오차가 필터링의 진전에 따라 변화하는 동력학을 나타내는 비선형 차분방정식(nonlinear differential equation)으로, 우변에 포함된 지수함수를 전개하면 다음과 같이 쓸 수 있다.

$$\begin{aligned} V(t+1) &= V(t) + \\ &\mu_o \zeta(t) \left(1 - \beta\gamma \sum_{n=0}^{\infty} \frac{(-1)^n \gamma^n}{2^n n!} \zeta^{2n}(t)\right) U(t) - \\ &\mu_o \left(1 - \beta\gamma \sum_{n=0}^{\infty} \frac{(-1)^n \gamma^n}{2^n n!} (2n+1) \zeta^{2n}(t)\right) \cdot \\ &U(t) U^T(t) V(t) + H.O.T.'s \end{aligned} \quad (26)$$

필터링이 진전되어 최적해에 가까이 접근해 있다고 즉 $V(t) \simeq 0$ 로 가정하면, 식(26)은 $V(t)$ 의 고차항을 무시할 수 있으므로 $V(t)$ 에 대한 1차 식이 되고, 다음에 식의 양변에 평균을 취하면 제시한 적응필터 알고리즘의 평균오차에 대한 동적인 거동을 기술하는 다음 식(27)을 얻을 수 있다

$$\begin{aligned} E[V(t+1)] &= (I - \mu_o R) E[V(t)] + \\ &\mu_o \beta\gamma \sum_{n=0}^{\infty} \frac{(-1)^n \gamma^n}{2^n n!} (2n+1) E[\zeta^{2n}(t)] \cdot R \cdot E[V(t)] \end{aligned} \quad (27)$$

위 식의 계산에서는 LMS 알고리즘의 통계적인 특성 분석에서 사용하고 있는 기본가정[6] 및 플렌트 잡음 $\zeta(t)$ 의 홀수차 모멘트(odd moment)가 모두 0이라는 가정을 사용하였다. 다음에는 식(26)로부터 필터 가중치 오차의 공분산 행렬에 관한 동적 방정식을 유도한다. 식(26)을 이용하여 가중치 오차 $V(t)$ 의 공분산 행렬의 동력학 식(28)을 얻을 수 있다.

$$\begin{aligned} K(t+1) &= K(t) - \\ &\mu_o \alpha_1(t) (RK(t) + K(t)R) + \mu_o^2 \alpha_2(t) R + \\ &\mu_o^2 \alpha_3(t) E[U(t)U^T(t)V(t)V^T(t)U(t)U^T(t)] + \\ &H.O.T.'s \end{aligned} \quad (28)$$

위 식에서 사용한 계수 $\alpha_1, \alpha_2, \alpha_3$ 는 다음과 같이 정의한다.

$$\alpha_1(t) = 1 - \beta\gamma \sum_{n=0}^{\infty} \frac{(-1)^n \gamma^n}{2^n n!} (2n+1) E[\zeta^{2n}(t)] \quad (29)$$

$$\begin{aligned} \alpha_2(t) &= E[\zeta^2(t)] - 2\beta\gamma \sum_{n=0}^{\infty} \frac{(-1)^n \gamma^n}{2^n n!} E[\zeta^{2n+2}(t)] + \\ &\beta^2 \gamma^2 \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{n+m} \gamma^{n+m}}{2^{n+m} n! m!} E[\zeta^{2n+2m+2}(t)] \end{aligned} \quad (30)$$

$$\begin{aligned} \alpha_3(t) &= 1 - 2\beta\gamma \sum_{n=0}^{\infty} \frac{(-1)^n \gamma^n}{2^n n!} (2n+1) E[\zeta^{2n}(t)] + \\ &\beta^2 \gamma^2 \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^{n+m} \gamma^{n+m}}{2^{n+m} n! m!} (2n+1) \cdot \\ &(2m+1) E[\zeta^{2n+2m}(t)] \end{aligned} \quad (31)$$

식(28)에서 우변 네제 항의 6차 모멘트는 $U(t), V(t)$ 가 서로 독립적(independent)이고, 필터 입력 $U(t)$ 가 실수 값만을 갖으며 평균이 0인 가우시안 프로세스(Gaussian process) 라고 하면, 가우시안 모멘트의 분할정리(Gaussian moment factoring theorem)에 의해 다음과 같이 쓸 수 있다.

$$E[U U^T V V^T U U^T] = (2RK(t) + tr\{RK(t)\})R \quad (32)$$

위 식의 좌변에서 변수 t 는 편의상 생략하였다. 이제 필터의 평균자승오차를 계산하기로 하자. 식(8)에서와 같이 제시한 적응필터 알고리즘의 평균자승오차는 다음과 같이 쓸 수 있다.

$$MSE(t) (= E[\epsilon^2(t)]) = E[\zeta^2(t)] + tr\{RK(t)\} \quad (33)$$

식(33)에서 우변의 마지막 항이 초과평균자승오차(EMSE)로, 다음과 같이 다시 쓸 수 있다.

$$\begin{aligned} EMSE(t) &= tr\{RK(t)\} = tr\{Q\Lambda Q^T Q X(t) Q^T\} \\ &= \sum_{i=1}^N \lambda_i x_i(t) = \lambda^T x(t) \end{aligned} \quad (34)$$

위 식에서 행렬 Q 는 필터입력 $U(t)$ 의 공분산 행렬 R 을 대각선 행렬로 바꾸어 주는 역할을 하며 $Q^T Q = I$ 인 성질을 갖고, 벡터 $\lambda, x(t)$ 는 각각 필터 입력의 공분산행렬 R 의 고유치 및 필터계수오차의 공분산행렬 $X(t)$ 의 대각선 요소로 구성된다. 위에서 계산한 필터계수 오차의 공분산행렬 $K(t)$ 에 대한 동력학을 나타내는 식(28)에 식(32)을 이용한 뒤, 행렬 Q 를 이용하여 상사변환(similarity transformation) 하면 식(35)을 얻는다.

$$\begin{aligned} X(t+1) &= X(t) - \mu_o \alpha_1(t) (\Lambda X(t) + X(t) \Lambda) + \\ &\mu_o^2 \alpha_2(t) \Lambda + \mu_o^2 \alpha_3(t) (2\Lambda X(t) \Lambda + tr\{\Lambda X(t)\} \Lambda) \end{aligned} \quad (35)$$

식(34)에서 보면 초과 평균자승오차는 $X(t)$ 의 대각선 요소에만 영향을 받는 것을 알 수 있는데, $X(t)$ 의 i 번째 대각선 요소는 다음과 같이 나타낼 수 있다.

$$\begin{aligned} x_i(t+1) &= (1 - 2\mu_o \alpha_1(t) \lambda_i + 2\mu_o^2 \alpha_3(t) \lambda_i^2) x_i(t) + \\ &\mu_o^2 \alpha_3(t) \lambda_i \sum_{j=1}^N \lambda_j x_j(t) + \mu_o^2 \alpha_2(t) \lambda_i \end{aligned} \quad (36)$$

$X(t)$ 의 각 대각선 요소에 대한 식(36)을 모아서 간략하게 다음의 벡터 식으로 바꾸어 쓸 수 있다.

$$x(t+1) = Bx(t) + \mu_o^2 \alpha_2(t) \lambda \quad (37)$$

여기에서 행렬 B 의 각 요소(elements)는 다음과 같이 정의된다.

$$b_{ij} = \begin{cases} 1 - 2\mu_o \alpha_1(t) \lambda_i + 3\mu_o^2 \alpha_3(t) \lambda_i^2, & i = j \\ \mu_o^2 \alpha_3(t) \lambda_i \lambda_j, & i \neq j \end{cases} \quad (38)$$

위에서 계산된 평균자승오차 식(37)로부터 수렴된 상태에서의 초과 평균자승오차를 및 미조정잡음을 계산하면 다음 식(39) 및 식(40)과 같다.

$$EMSE(\infty) = \frac{\frac{\mu_0}{2} \sum_{i=1}^N \alpha_2 \lambda_i / (\alpha_1 - \mu_0 \alpha_3 \lambda_i)}{1 - \frac{\mu_0}{2} \sum_{i=1}^N \alpha_3 \lambda_i / (\alpha_1 - \mu_0 \alpha_3 \lambda_i)} \quad (39)$$

$$M_{NSD} = \frac{\frac{\mu_0}{2} \sum_{i=1}^N \alpha_2' \lambda_i / (\alpha_1 - \mu_0 \alpha_3 \lambda_i)}{1 - \frac{\mu_0}{2} \sum_{i=1}^N \alpha_3 \lambda_i / (\alpha_1 - \mu_0 \alpha_3 \lambda_i)} \quad (40)$$

여기에서 $\alpha_2' = \alpha_2 / E[\zeta^2(t)]$ 로 정의 되었다.

V 시뮬레이션

앞 장에서는 제시한 적응필터 알고리즘의 성능특성을 수식적인 분석을 통해 살펴왔다. 여기에서는 앞에서 계산한 이론적인 값들을 확인하고, 또한 수식적으로 살펴보기 힘들었던 사항들을 수치적인 계산예를 통해 살펴보기 위한 목적으로 시뮬레이션을 수행하였다. 시뮬레이션 예제로는 그림 1에 나타난 폴란트 모델링 문제 및 적응 동화기(adaptive equalizer) 문제를 사용하였다. (1) 예제 1 : 그림 1의 폴란트 모델링 문제에서 폴란트의 동력학을 표현하는 FIR 필터의 계수 즉 적응필터의 최적해는 $W^* = [1 \ 3 \ 5 \ 7 \ 9]^T$ 로 주어지고, 폴란트 잡음은 평균이 0 이고 분산 (σ^2)이 0.01 인 정규분포를 갖는 백색잡음(white noise) 이라고 하자. 폴란트에 가해지는 입력은 이른치 계산의 편의를 위하여 평균이 0 이고, 분산이 1 인 백색 가우시안 프로세스로 가정하였다. 폴란트를 모델링하는 적응필터의 차수는 5차로 하였으며, LMS 는 보폭을 달리하는 두가지 경우에 대해, 제시한 적응필터는 기본형 및 변형된 형태에 대해 시뮬레이션을 수행한다. 시뮬레이션에서 사용한 각 필터의 설계상수로, LMS 필터에서는 보폭 0.026, 0.13 두가지를, 제시한 필터에서는 기본형의 경우 $\mu_0 = 0.13, \gamma = 5, \beta\gamma = 0.95$ 을 사용하였고, 변형된 형태에서는 $\rho = 0.9, \bar{e}^2(0) = 0$ 을 추가적으로 사용하였다. 그림 2는 예제 1의 문제에 대해 위에서 정한 상수들을 이용하여 독립적으로 100회의 시뮬레이션을 수행하여 평균한 결과로, 그림 2.a는 적응 알고리즘에 의해 추정된 필터계수와 최적해와의 차에 대한 노름(norm)을 취하여 정규화 한 것으로 모델링 문제에서 모델 계수의 오차를 나타내는 지표로 볼 수 있는데, LMS는 보폭이 커지면 수렴은 빨라지나 오차가 커지고, 보폭이 작아지면 수렴후의 오차는 작으나 수렴이 늦는 일반적인 특성을 보이고 있는 반면, 제시한 알고리즘 및 그에 대한 변형은 빠른 수렴속도를 가지며 오차도 작은 결과를 보인다. 그림 2.b는 평균자승오차(MSE)를, 그림 2.c는 제안된 필터를 가변보폭 LMS 필터로 보기 위한 것으로 각 알고리즘의 보폭 및 사실상의 보폭(effective step size)을 나타낸다. 표 1은 각 알고리즘이 최적해 가까이 수렴된 정상상태(converged steady state)에서 살펴본 미조정률(M_{ss}) 및 모델오차($\|W^* - W_{ss}\|/\|W^*\|$)의 실험치 및 이론치를 비교한 것으로, 실험치의 경우는 그림 2에 나타난 각 경우에 대한 시뮬레이션 결과에서 정상상태에 도달한 것으로 판단되는 600회(step) 이후의 값에 대한 평균을 취하여 얻었다. 제시한 알고리즘에 대한 이론치와 실험치의 차이가 LMS의 경우보다 크게 나타나고 있는데, 이는 제시한 필터의 분석에서 LMS 필터에 대한 분석의 경우보다 더 많은 가정을 사용했기 때문으로 볼 수 있었다.

(2) 예제 2 : 앞의 예제 1의 경우는 주로 계산치와 실험치의 비교를 위한 목적의 시뮬레이션이었다고 한다면, 두번째의 예제는 앞의 경우 보다 실제 문제에 가깝다고 여겨지는 적응동화기(adaptive equalizer)를 생각하기로 한다. 예제로 사용할 적응동화기의 구조는 그림 3과 같은 구조를 갖는데, 이는 참고문헌[6]에서 LMS 및 RLS 필터의 시뮬레이션 예제로 사용한 것으로, 간략히 기술하면 다음과 같다. 불규칙 잡음 발생기 (1) 및 (2)는 각각 ± 1 값만을 갖는 불규칙 신호 및 분산 0.001 이고 평균 0인 신호를 발생시키고, 시간지연은 7 탭을, 적응필터는 11 탭을 사용하였으며, 채널의 임펄스 응답은 다음과 같이 주어진다.

$$h_n = \begin{cases} \frac{1}{2} \left[1 + \cos\left(\frac{2\pi}{2.5}(n-2)\right) \right], & n = 1, 2, 3 \\ 0, & \text{others.} \end{cases} \quad (41)$$

수행된 시뮬레이션은 3 가지로 나누어 볼 수 있고, 시뮬레이션 결과는 서로 독립적인 200회의 결과를 평균하여 얻은 평균자승오차로 나타내었다. 첫번째 시뮬레이션은 앞에서 제시한 NSD 필터에 대한 수식적인 분석 과정에서 적절한 값의 범위를 제시하지 못했던 설계 파라메터인 γ 의 변화에 따른 특성을 살펴 봄으로서, 구체적인 값을 결정하는데 도움을 주기 위한 목적으로 수행되었다. 그림 4는 NSD 필터의 설계 상수 중에서

$\mu_0, \beta\gamma$ 는 각각 0.1, 0.95로 고정하고 γ 를 5, 10, 20, 30, 40, 50으로 변화시키며 적응동화기의 응답특성을 살펴본 것으로, γ 가 커짐에 따라 수렴속도가 조금씩 빨라지고 정상상태의 오차특성도 γ 가 40 까지는 거의 변하지 않고 유지되었으나, 50이 되면서 부터는 수렴속도의 증가에 비해 오차가 크게 증가하기 시작하였다(표 2 참고). 표 2는 각 시뮬레이션에서 정상상태로 볼 수 있는 900회(step) 이후의 값을 평균하여 얻은 것이다. 다음 시뮬레이션은 LMS, RLS(recursive least square) 및 NSD 필터의 수렴 및 오차 특성을 비교하기 위한 목적으로 수행되었으며, 각 필터의 설계 상수는 LMS에서는 보폭이 0.015, 0.1 인 두 경우를 생각하였고, RLS에서는 참고문헌[6]에 따라 공분산 행렬의 초기 값을 주 대각선 요소에 만 0.04이 되도록 하였다. NSD는 첫번째 시뮬레이션 중에서 $\gamma = 40$ 인 경우를 고려하였다. 그림 5는 두번째의 시뮬레이션 결과로, RLS가 수렴 및 오차 특성에 있어서 가장 우수하고, NSD는 LMS와 RLS의 중간정도에 해당하는 것을 알 수 있다. 마지막으로 NSD의 기본형 및 변형에 대한 시뮬레이션 결과를 비교한 것으로, NSD에서의 설계 상수는 $\mu_0 = 0.1, \beta\gamma = 0.95, \gamma = 40$ 로 하였고, 변형된 NSD 필터에서는 추가적으로 $\rho = 0.95, \bar{e}^2(0) = 0$ 를 사용하였다. 시뮬레이션 결과는 그림 6에 수록하였는데, 변형의 경우가 약간 우수함을 알 수 있다.

VI 결론

여러 응용예에 널리 사용되어 성능이 입증된 최소평균자승(LMS) 적응필터의 특성을 계산량을 크게 늘리지 않으면서, 또한 알려진 LMS 필터의 우수한 안정도 특성을 희생시키지 않으면서, 개선하고자 하는 노력의 일환으로 성능이 향상된 새로운 적응 필터를 제안하였다. 새로운 적응 필터는 LMS 필터에서 사용한 성능지수 함수의 최적해 부근에서의 특성을 스토캐스틱 경사도를 사용하는 직강하법 알고리즘에 보다 적절한 형태로 개선함으로써 얻어졌다. 이 필터는 가변보폭 적응 알고리즘으로도 해석할 수 있는데, 이러한 관점에서 가변보폭의 심한 변동에 의해 추가적으로 생기는 오차를 줄일 수 있는 간략한 개선 방안을 생각하였다. 새롭게 얻은 적응필터에 대한 분석으로 부터, 제시한 알고리즘은 필터오차가 큰 상태에서는 LMS 알고리즘과 같이 볼 수 있으므로 LMS 필터가 가지는 안정도 특성을 유지하며, 필터 오차가 작은 상태에서는 LMS의 경우 보다 더 좋은 오차특성을 갖는 것을 밝혀냈고, 이러한 특성들은 시뮬레이션을 통하여 다시 확인할 수 있었다. 또한 제시한 알고리즘에서 사용자가 정해 주어야 할 설계상수의 선정에 관해서도 어느 정도의 범위로서 개략적인 지침(guide line)을 제시하였고, 수식적으로 정하기 어려운 값인 γ 에 대해서는 시뮬레이션을 통하여 성능에 미치는 현상을 살펴봄으로서 구체적인 값의 선정에 힌트를 얻을 수 있도록 하였다.

References

- [1] B. Widrow and M. E. Hoff, Jr. "Adaptive switching circuits," *IRE WESCON Conv. Rec.*, pt. 4, pp. 96-104, 1960.
- [2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, N. J. : Prentice-Hall, 1985.
- [3] E. Walach and B. Widrow, "The least mean fourth(LMF) adaptive algorithm and its family," *IEEE Trans. Inf. Theory*, vol. IT-19, pp. 59-68, 1984
- [4] R. W. Harris, D. M. Chabries, and F. A. Bishop, "A variable step(VS) adaptive filter algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 309-316, Apr. 1986.
- [5] S. Karni and G. Zeng, "A new convergence factor for adaptive filters," *IEEE Trans. Circuits Syst.*, vol. CAS-36, pp. 1011-1012, Jul. 1989.
- [6] S. Haykin, *Adaptive Filter Theory, second edition*. Englewood Cliffs, N. J. : Prentice-Hall, 1991.

- [7] L. L. Horowitz and K. D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 722-736, Jun., 1981.
- [8] A. Feuer and 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] D. J. Park, "New performance function and variable step size LMS algorithm derived by Karni and Zeng," *Electron. Lett.*, vol. 27, pp. 2182-2183, 7th Nov., 1991.
- [10] V. A. Gholkar, "Mean square convergence analysis of LMS algorithm," *Electron. Lett.*, vol. 27, pp. 1705-1706, 27th Sep., 1990.

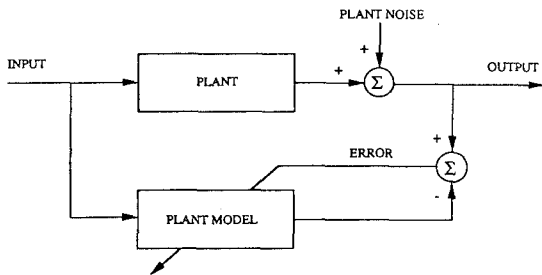
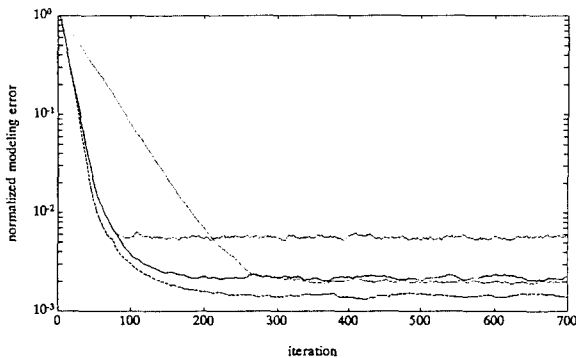


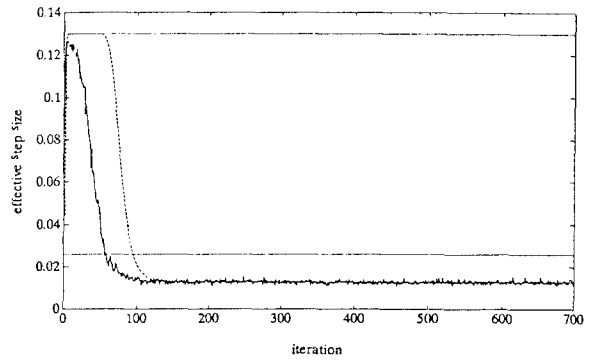
그림 1 적응 필터에 의한 플랜트 모델링

표 1. 예제 1 에서 시뮬레이션 결과와 이론치의 비교

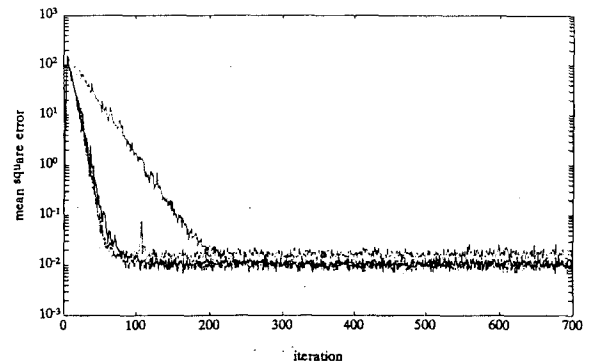
	M_{ss}		$\ W^* - W_{ss}\ / \ W^*\ $	
	Simulation	Theory	Simulation	Theory
LMS # 1	9.9 %	7.2 %	1.99e-3	1.39e-3
LMS # 2	68.9	59.6	5.65e-3	4.02e-3
NSD	9.6	5.0	2.20e-3	1.16e-3
NSDM	4.4	—	1.45e-3	—



(a)



(b)



(c)

그림 2 플랜트 모델링에 대한 시뮬레이션

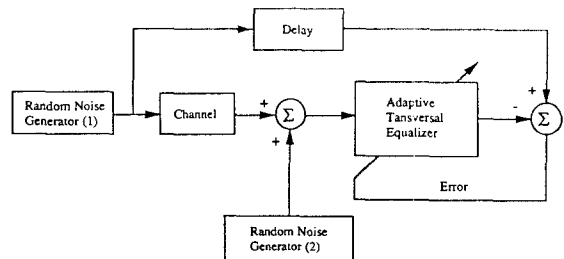


그림 3 적응 동화기의 구조

표 2. 적응 등화기 시뮬레이션의 정상상태 오차 비교

LMS # 1	LMS # 2	RLS	NSD, $\gamma = 40$	NSDM
1.57e-3	5.34e-3	1.39e-3	1.56e-3	1.57e-3
NSD, $\gamma = 5$	NSD, $\gamma = 10$	NSD, $\gamma = 20$	NSD, $\gamma = 30$	NSD, $\gamma = 50$
1.56e-3	1.49e-3	1.48e-3	1.51e-3	1.65e-3

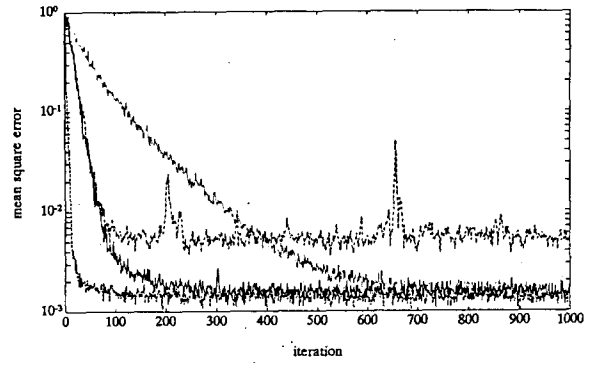


그림 5 적응 등화기에 대한 시뮬레이션(2)

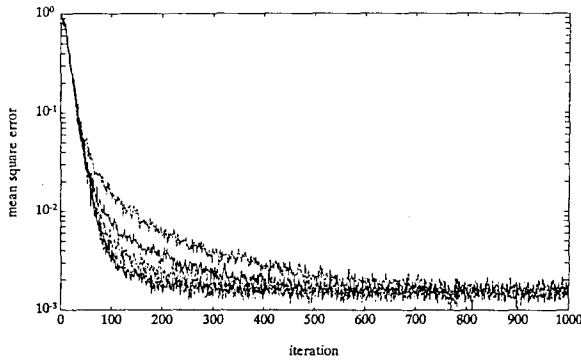


그림 4 적응 등화기에 대한 시뮬레이션(1)

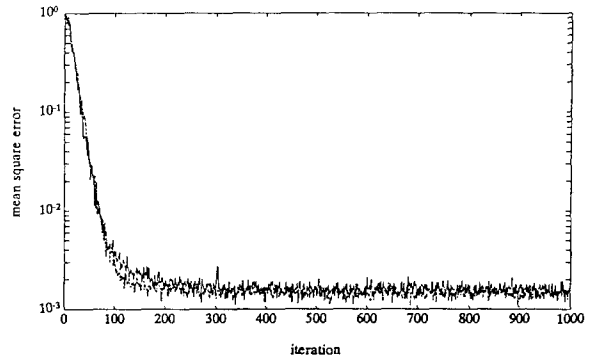


그림 6 적응 등화기에 대한 시뮬레이션(3)