

# 시스템 식별을 위한 Erlang 함수의 망각 인자를 가진 개선된 RLS 알고리즘

## An Improved New RLS Algorithm with Forgetting Factor of Erlang Function for System Identification

석진욱, 최경삼, 이종수, 조성원  
(Jin-wuk Seok, Kyung-Sam Choi, Jong-soo Lee, and Seong-won Cho)

**Abstract** : In this paper, we present an effective RLS algorithm with forgetting factor of Erlang function for the system identification. In the proposed algorithm, the forgetting factor decreases monotonically in the first stage, and then it increases monotonically in the second stage in contrary to the conventional forgetting factor RLS Algorithms. In addition, annealing effect and an asymptotically stability of the proposed algorithm is discussed based on the analysis of convergence property on. Simulation results for the system identification problem indicate the superiority of the proposed algorithm in comparison to the RLS algorithms such as NLMS and Kalman filter based algorithm.

**Keywords** : RLS, system identification, forgetting factor, Erlang function, annealing effect

### I. 서론

시스템 식별(system identification)등에서 광범위하게 사용되는 파라미터 추정(parameter estimation) 알고리즘은 최소자승법을 기본개념으로 망각 인자(forgetting factor)알고리즘이나 칼만 필터(Kalman filter) 알고리즘 등으로 발전되어왔다. 파라미터 추정 알고리즘에서 중요한 접근 방법인 steepest descent 방식의 가중 RLS 알고리즘은 일반적으로 다음과 같다[1][2].

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \mu L(t)(y(t) - \varphi^T(t)\hat{\theta}(t)) \quad \mu \in (0,1) \quad (1)$$

(1)이  $\mathfrak{R}^n$  위에서 정의된 식이라고 하면,  $\hat{\theta}(t) \in \mathfrak{R}^n$ 는 시간  $t$ 일 때 추정될 파라미터 벡터이며,  $y(t) \in \mathfrak{R}$ 는 시간  $t$ 에서 미지 시스템의 출력이며,  $\varphi(t) \in \mathfrak{R}^n$ 는  $y(t)$ 를 제외한 입출력 특성벡터(input-output quantity vector)로서,  $(y(t), \varphi(t))^T$ 는 전체 입출력의 모든 정보이다.  $\mu \in \mathfrak{R}$ 는 적응속도(speed of adaptation) 혹은 step size이며  $L(t) \in \mathfrak{R}^n$ 는 랜덤과정(random process)으로서 적응이득 벡터이다.

일반적으로 적응이득 벡터  $L(t)$ 를 어떻게 선택하는가에 따라 (1)의 기본 알고리즘들은 각각 NLMS(Normalized LMS), 망각 인자 알고리즘, 칼만 필터 알고리즘등으로 구분되며 그 중에서도 망각 인자 알고리즘의 접근 방식은 일반적인 RLS알고리즘과 비교하여 평가함수 자체를 다음과 같이 놓음으로써 구하게 된다.

$$V_t(\theta) = \sum_{s=1}^t \lambda^{t-s} (y(s) - \varphi^T(s)\hat{\theta}(s))^2 \quad (2)$$

(2)에서  $\lambda$ 는 망각 인자로서 만일  $\lambda = 1$ 이면 일반적인 RLS알고리즘을 유도할 수 있다. 그러나,  $\lambda < 1$ 의 경우,  $t \rightarrow \infty$

의 크기가 크면 클수록 손실항  $y(s) - \varphi^T(s)\hat{\theta}(s)$ 의 크기는 평가함수  $V_t(\theta)$ 에 영향을 미치지 못하게 된다. 즉, 초기 손실항의 크기는 무시되며 알고리즘이 어느정도 진행된 다음부터의 손실항의 크기에 따라 파라미터 추정이 효과적으로 이루어지게 된다. 따라서, 망각 인자의 도입은 추정된 파라미터가 불안정한 초기상태의 영향을 가급적 적게 받도록하여 보다 올바른 추정이 가능하도록 하는데 목적이 있으나 망각 인자의 적절한 크기 선정이 어려우며, 또한 망각 인자의 도입 자체가 알고리즘의 수렴성을 불안하게 만드는 문제점을 가지고 있다. 따라서 망각 인자는 보통 1 보다 약간 작은 0.99나 0.95 등의 값을 취하거나 지수함수적으로 감소하는 함수를 도입하여 사용하기도 한다.

본 논문은 망각 인자의 효과를 극대화 하며 동시에 망각 인자를 도입함으로써 발생하는 알고리즘의 불안정성을 해소하는 것이 목적이며 동시에 안정적이고 정확한 추정이 가능한 RLS알고리즘을 제안하고자 한다.

### II. 망각 인자를 사용한 RLS알고리즘

망각 인자를 사용하는 RLS알고리즘을 서술하기 전에 먼저 다음의 가정을 둔다.

가정 1 :

1) 미지 시스템의 입출력 특성벡터(input-output quantity vector)  $\varphi(t) \in \mathfrak{R}^n$ 는 평균 입출력 특성벡터  $E_t \varphi(t) \in \mathfrak{R}^n$ 의  $\epsilon$ -neighborhood 주위에 충분히 조밀하게 분포하고 있다. i.e.

$$\forall t \in Z^+, \varphi(t) \in \mathfrak{R}^n, \exists \gamma > 0 \text{ s.t.} \\ \varphi(t) \in B^0(\varphi_0, \epsilon) = \{ \varphi(t) \in \mathfrak{R}^n \mid \|\varphi(t) - \varphi_0\| < \gamma \} \quad (3)$$

(3)에서  $Z^+$ 는 0과 양의 정수의 집합이며  $\|\cdot\|$ 는 Euclidean norm이다.

접수일자 : 1997. 11. 19., 수정완료 : 1999. 4. 9.  
석진욱 : 홍익대학교 과학기술 연구소  
최경삼, 이종수, 조성원 : 홍익대학교 공과대학 전자전기 공학부

2) 미지 시스템의 모델  $M_1$ 과 시스템  $\zeta_1$ 은 다음의 구조를 가지고 있다.

$$M_1 : A(q)y(t) = B(q)u(t)$$

$$\zeta_1 : y(t) + a_1y(t-1) + a_2y(t-2) \cdots + a_ny(t-n)$$

$$= b_1u(t-k) + b_2u(t-k-1) \cdots + b_mu(t-k-m) \quad (4)$$

3) 미지 시스템에 가해지는 Noise  $e(t)$ 는 IID(Incremental Independent Distribution) 확률과정(stochastic process)이다.

가정 1의 1)은 논의의 대상을 BIBO(Bounded Input, Bounded Output) 시스템으로 한정하는 것이다. 두 번째 가정은 미지 시스템의 모델과 시스템구조를 확정하는 것이다. 따라서 가정 1의 1)과 두 번째 가정을 종합하여 미지시스템의 추정 시스템을 다음과 같이 놓을 수 있다.

$$y(t) = \varphi^T(t)\hat{\theta}(t) + e(t) \quad (5)$$

(5)에서  $\hat{\theta}(t)$ 는 시간  $t$ 에서의 시스템 파라미터 벡터(system parameter vector)로서 가정 1의 2)로부터 다음과 같이 주어진다.

$$\hat{\theta}(t) = (-\hat{a}_1(t), -\hat{a}_2(t), \dots, -\hat{a}_n(t), \hat{b}_1(t), \hat{b}_2(t), \dots, \hat{b}_m(t))^T \quad (6)$$

위에서,  $\hat{a}_k(t)$ ,  $\hat{b}_l(t)$ 는 미지의 추정 시스템의 계수  $a_k$ ,  $b_l$ 의 시간  $t$ 에서의 추정치이며,  $e(t)$ 는 가정 1의 3)으로 부터 임의의 양수  $k, l$ 에 대하여  $Ee(t-k)e(t-l) = \delta_{kl}$ 의 특징을 가진다. 여기서 함수는 dirac-delta함수로서  $k = l$ 이면  $\delta_{kl} = 1$ 이고  $k \neq l$ 이면  $\delta_{kl} = 0$ 이다. 또한 입출력 특성벡터  $\varphi(t)$ 는 다음과 같다.

$$\varphi(t) = (y(t-1), y(t-2), \dots, y(t-n),$$

$$u(t-k), u(t-k-1), \dots, u(t-k-m))^T \quad (7)$$

따라서 시스템 파라미터 벡터  $\theta(t)$ 와 입출력 특성벡터  $\varphi(t)$ 는 다음의 domain을 가진다.

$$\hat{\theta}(t) \in D_M \subset \mathfrak{R}^{n+m}, \varphi(t) \in \mathfrak{R}^{n+m} \quad (8)$$

가정 1의 조건하에서, 미지 시스템의 추정 시스템을 구하기 위한 망각 인자를 가진 RLS알고리즘은 다음과 같다[3][4].

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \varphi^T(t)\hat{\theta}(t-1)]$$

$$L(t) = \frac{P(t-1)\varphi(t)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)}$$

$$P(t) = \frac{1}{\lambda(t)} \left[ P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)} \right] \quad (9)$$

(2)에서 망각 인자는 평가함수  $V_i(\theta)$ 가  $\lambda$ 의 지수함수 꼴로 나타나므로  $\forall t \in Z'$ 에 대하여  $\lambda(t) = \lambda$ 로 놓아진다. 이 때 망각 인자는 보통 0.98에서 0.995 사이의 값을 선택한다[3]-[5]. 그러나 망각 인자가 시간에 불변인 일정한 값으로 놓아질 경우에는 추정 알고리즘에서의 파라미터 벡터는 매우 약한 위상에서 확률측도로 약 수렴(weak convergence in probability measure)한다. 따라서 어느 정도 알고리즘이 진행된다 하더라도 일반적으로

사용되는 위상인  $\mathfrak{R}^n$  공간상에서 보면 약간의 외란을 가지고 추정 파라미터가 움직이는 것을 관찰 할 수 있다 [6].

망각 알고리즘의 안정성을 강화하기 위해 망각 인자  $\lambda(t)$ 를 시간에 대한 증가함수로 놓을 수 있다[4].

$$\lambda(t) = 1 - \lambda_0[1 - \lambda(0)] \quad (10)$$

(10)의 경우 다음의 과정을 거쳐 회귀방정식(recursive equation)으로 치환할 수 있다.

$$\lambda(t-1) = 1 - \lambda_0[1 - \lambda(0)]$$

$$\lambda_0^{-1} = \frac{\lambda(t-1) - 1}{1 - \lambda(0)}$$

위 식을 (10)에 대입하면

$$\lambda(t) = \lambda_0\lambda(t-1) + (1 - \lambda_0) \quad (11)$$

이와 같이 지수 함수적으로 감소하는 망각 인자 알고리즘은 시불변 망각 인자를 갖는 RLS 알고리즘보다 안정성이 뛰어난 장점을 가지고 있으나 망각 인자를 도입함으로써 갖는 특성은 시불변 망각 인자 알고리즘보다 떨어지게 된다.

### III. Erlang 함수망각인자를 사용한 RLS알고리즘

본 논문에서 제안하는 RLS알고리즘의 망각 인자는 Erlang 함수를 사용하여, 어느 일정시간이 될 때까지는 단조감소하다가 일정 시간 이후부터 단조증가하여 1에 접근하는 방식이다. 제안한 알고리즘은 다음과 같다.

step 0 :

$$P^{-1}(t_0) = \sum_{k=1}^{t_0} \varphi(k)\varphi^T(k) \quad \text{set } t = t_0 = 0$$

$$\hat{\theta}(t_0) = P(t_0) \sum_{k=1}^{t_0} \varphi(k)y(k)$$

마지막 iteration time으로 임의의 큰 수  $M_i \in Z'$ 를 놓는다.

step 1 :

$$\lambda(t) = 1 - \alpha \cdot t \cdot \exp(-\beta t) \quad \text{where } \alpha, \beta \ll 1$$

$$t \leftarrow t+1$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \varphi^T(t)\hat{\theta}(t-1)]$$

$$L(t) = \frac{P(t-1)\varphi(t)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)}$$

$$P(t) = \frac{1}{\lambda(t)} \left[ P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)} \right] \quad (12)$$

step 2 :

If  $t > M_i$ , then stop else goto step 1

(12)를 통해 나타나는 망각 인자  $\lambda(t)$ 는  $\alpha$ 가 0.02이고,  $\beta$ 는 0.05일 때 그림 1과 같다.

제안한 알고리즘은 망각 인자의 형태에 의해 annealing 효과와 수렴성(convergence)을 동시에 보장할 수 있다.

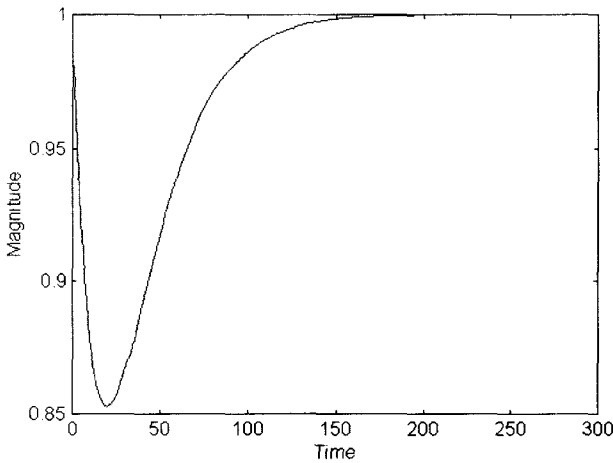


그림 1. 제안한 알고리즘의 망각인자.  
Fig. 1. The forgetting factor in proposed algorithm.

즉, 알고리즘이 적용되는 초기 시간대에서 망각 인자가 감소함으로써 추정하고자 하는 파라미터 벡터가 데이터에 더욱 민감하게 작용하게 한다. 이때 각 데이터에는 가정에서 IID 백색잡음이 존재한다고 가정하였으므로 파라미터 벡터는 백색잡음의 영향을 많이 받게 된다. 이것은 망각 인자의 값을 1에 근사한 어떤 값으로 고정하였을 때와 비교하면, 백색잡음의 영향을 더욱 많이 받는 것으로 볼 수 있다. 그러므로 기존의 알고리즘에 일종의 noise를 첨가시켜 annealing 효과를 얻는 것과 동치가 된다. 또한 망각 인자의 감소를 통해 파라미터 벡터가 불안정하게 움직일 가능성이 커지게 됨에 따라 최소화하고자 하는 평가함수의 값이 이전의 파라미터 벡터를 대입했을 때 보다 증가할 수 있게 된다. 이를 통해 국소최소점을 벗어날 가능성이 발생한다. annealing 효과는 이와 같이 평가함수가 multi modal한 경우, 국소최소점을 벗어나기 위해 평가함수의 에너지 경사면을 잡음이나 그 밖의 방법으로 올라가는(hill-climbing) 경우 발생한다[8].

본 논문에서 제안한 알고리즘이 annealing 효과를 가지고 있음을 증명하기 위하여 다음의 가정을 둔다.

가정 2 : RLS 알고리즘에 의해 추정된 파라미터  $\hat{\theta}(t)$ 는 참값  $\theta(t)$ 에 대하여 유한한 범위내에 있다. i.e.

$$\|\hat{\theta}(t) - \theta(t)\| < \delta \Leftrightarrow \hat{\theta}(t) \in D_M \quad (13)$$

가정 2는 RLS 알고리즘이 기본적으로 다음과 같은 선형 최적 추정 임을 감안할 때 타당하며, 가정 1의 ii)와 iii)을 통해 유도된 (8)을 살펴볼 때 자명하다[3]-[5].

$$\hat{\theta}(t) = \left( \sum_{k=1}^{t_0} \varphi(k)\varphi^T(k) \right)^{-1} \cdot \left( \sum_{k=1}^{t_0} \varphi(k)y(k) \right) \quad (14)$$

보조정리 1 : 가정 1, 2의 조건이 만족되고  $\varphi_0 = 0$ 이 만족될 경우,

행렬  $\varphi(t)\varphi^T(t)$ 의 행렬 norm [7]  $\|\varphi(tk)\varphi^T(t)\|$ 은 다음의

부등식을 만족한다.

$$\|\varphi(t)\varphi^T(t)\| < \frac{\gamma^2}{n+m} \quad (15)$$

(15)에서 행렬 norm은 벡터 norm  $\|\cdot\|$ 에서 유도된 것으로서 다음과 같이 정의된다.

$$\|\varphi(tk)\varphi^T(t)\| \equiv \sup_{x \in \mathbb{R}^{n+m}} \|\varphi(tk)\varphi^T(t)x\|_{\|\cdot\|_2=1}$$

보조정리 1의 증명 : 가정 1의 첫 번째 가정과  $\varphi_0 = 0$ 에 의하여  $\varphi(t)$ 는 다음과 같이 놓을 수 있다.

$$\|\varphi(t)\| < \left\| \frac{\gamma}{\sqrt{n+m}} (1, 1, \dots, 1)^T \right\|$$

따라서 행렬 norm의 정의에 의하여

$$\begin{aligned} & \sup_{x \in \mathbb{R}^{n+m}} \|\varphi(t)\varphi^T(t)x\|_{\|\cdot\|_2=1} \\ & < \frac{\gamma^2}{(n+m)^2} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 \end{pmatrix} \frac{1}{\sqrt{n+m}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{\gamma^2}{(n+m)^2} \cdot \sqrt{n+m} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{\gamma^2}{n+m} \end{aligned}$$

보조정리 2 : 망각 인자가 최소가 되는 시간을  $t_{\min}$ 이라고 하고  $t \in (0, t_{\min})$ 에 대하여  $P(t)$ 의 1 역행렬을  $\bar{R}(t)$ 라 놓으면,  $\bar{R}(t)$ 는 다음 부등식을 만족한다.

$$\|\bar{R}(t)\| < \frac{1 - \lambda_{\max}^{t+1}}{1 - \lambda_{\max}} \cdot \frac{\gamma^2}{n+m}$$

보조정리 2의 증명 : 역행렬에 대한 보조정리에 의하여

$$P^{-1}(t) = \bar{R}(t) = \lambda(t)\bar{R}(t-1) + \varphi(t)\varphi^T(t) \quad (16)$$

(16)을 다시 정리하면

$$\begin{aligned} \bar{R}(t) &= \lambda(t)\bar{R}(t-1) + \varphi(t)\varphi^T(t) \\ &= \lambda(t)\lambda(t-1)\bar{R}(t-2) + \lambda(t)\varphi(t-1)\varphi^T(t-1) \\ &\quad + \varphi(t)\varphi^T(t) \\ &= \sum_{k=0}^t \left[ \prod_{s=0}^{k-1} \lambda(t-s) \right] \left\{ u(k-1) + \delta(k) \right\} \varphi(t-k)\varphi^T(t-k) \quad (17) \end{aligned}$$

(17)에 norm을 취하면

$$\begin{aligned} & \|\bar{R}(t)\| \\ &= \left\| \sum_{k=0}^t \left[ \prod_{s=0}^{k-1} \lambda(t-s) \right] \left\{ u(k-1) + \delta(k) \right\} \varphi(t-k)\varphi^T(t-k) \right\| \\ &\leq \sum_{k=0}^t \left\| \left[ \prod_{s=0}^{k-1} \lambda(t-s) \right] \left\{ u(k-1) + \delta(k) \right\} \varphi(t-k)\varphi^T(t-k) \right\| \quad (18) \end{aligned}$$

(18)에서  $\left\{ \prod_{s=0}^{k-1} \lambda(t-s) \right\} u(k-1) + \delta(k)$  는, 제안한 알고리즘에 의해 나타나는 망각 vector중  $t \in (0, t_{\min})$ 에 대하여 최대값을  $\lambda_{\max}$ 라 하면 다음과 같다.

$$\left\{ \prod_{s=0}^{k-1} \lambda(t-s) \right\} u(k-1) + \delta(k) < \lambda_{\max}^{-k}$$

따라서, (18)은 Minkovsky 부등식을 사용하여 다음과 같이 정리된다

$$\begin{aligned} & \|\bar{R}(t)\| \\ & \leq \sum_{k=0}^t \left\| \left\{ \prod_{s=0}^{k-1} \lambda(t-s) \right\} u(k-1) + \delta(k) \right\| \varphi(t-k) \varphi^T(t-k) \\ & < \sum_{k=0}^t \lambda_{\max}^{-k} \varphi(t-k) \varphi^T(t-k) \\ & \leq \sum_{k=0}^t \lambda_{\max}^{-k} \cdot \|\varphi(t-k) \varphi^T(t-k)\| \end{aligned} \quad (19)$$

보조정리 1에 의하여 (19)는 다음과 같다.

$$\begin{aligned} \|\bar{R}(t)\| & < \sum_{k=0}^t \lambda_{\max}^{-k} \cdot \|\varphi(t-k) \varphi^T(t-k)\| \\ & < \sum_{k=0}^t \lambda_{\max}^{-k} \cdot \frac{\gamma^2}{n+m} = \frac{1-\lambda_{\max}^{-t-1}}{1-\lambda_{\max}^{-1}} \cdot \frac{\gamma^2}{n+m} \end{aligned} \quad (20)$$

정리 1 : 제안한 알고리즘은 가정 1, 2의 조건이 만족될 경우 annealing 효과를 갖는다.

증명 : 정리 1은 평가함수를  $V_t(\theta) = \sum_{s=1}^t (y(s) - \varphi^T(s)\hat{\theta}(s))^2$ 로 놓을 경우 시간 t에서의 추정 시스템의 추정 파라미터  $\hat{\theta}(t)$ 와 t로부터 k시간 이전의 추정 파라미터  $\hat{\theta}(t-k)$ 간의 거리가 유한시간동안 수렴하지 않고 발산한다는 것과 동치이다. 따라서 추정 파라미터 수열  $\{\hat{\theta}(t)\}_{t=0}^{t=t_{\max}}$ 는 망각 인자가 감소하는 동안 Cauchy 수열이 아님을 증명하면 된다. 제안한 알고리즘에서  $L(t)$ 는 다음과 같이 정리된다.

$$\begin{aligned} L(t) & = \frac{P(t-1)\varphi(t)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)} \\ & = \frac{1}{\lambda(t)} \cdot P(t-1)\varphi(t) - \frac{1}{\lambda(t)} \cdot \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)\varphi(t)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)} \\ & = \frac{1}{\lambda(t)} \cdot \left[ P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\lambda(t) + \varphi^T(t)P(t-1)\varphi(t)} \right] \varphi(t) \\ & = P(t)\varphi(t) = \bar{R}^{-1}(t)\varphi(t) \end{aligned} \quad (21)$$

따라서, 제안한 알고리즘에서 파라미터 계산식은 다음과 같다.

$$\hat{\theta}(t-1) = \hat{\theta}(t-1) + \bar{R}^{-1}(t)\varphi(t) \left[ y(t) - \varphi^T(t)\hat{\theta}(t-1) \right] \quad (22)$$

제안한 알고리즘에서 파라미터 계산식에 의해 생성되는 수열  $\{\hat{\theta}(t)\}_{t=0}^{t=t_{\max}}$ 이 Cauchy 수열인가를 알아보기 위해  $t, t-k \in (0, t_{\min})$ 인  $t, t-k$ 를 택한다. 이 때 두 파라미터 벡터의 차이는 다음과 같다.

양변에 norm을 취하고 삼각 부등식을 적용하면,

$$\begin{aligned} & \|\hat{\theta}(t) - \hat{\theta}(t-k)\| \\ & = \|\hat{\theta}(t) - \hat{\theta}(t-1) + \hat{\theta}(t-1) - \hat{\theta}(t-2) \cdots + \hat{\theta}(t-k+1) - \hat{\theta}(t-k)\| \\ & \leq \|\hat{\theta}(t) - \hat{\theta}(t-1)\| + \|\hat{\theta}(t-1) - \hat{\theta}(t-2)\| + \cdots + \|\hat{\theta}(t-k+1) - \hat{\theta}(t-k)\| \\ & = \sum_{s=0}^{k-1} \|\hat{\theta}(t-s) - \hat{\theta}(t-s-1)\| \\ & = \sum_{s=0}^{k-1} \bar{R}^{-1}(t-s)\varphi(t-s) \left[ y(t-s) - \varphi^T(t-s)\hat{\theta}(t-s-1) \right] \end{aligned} \quad (23)$$

(23)에 Minkovsky 부등식을 적용하면,

$$\begin{aligned} & \|\hat{\theta}(t) - \hat{\theta}(t-k)\| \\ & \leq \sum_{s=0}^{k-1} \|\bar{R}^{-1}(t-s)\varphi(t-s) \left[ y(t-s) - \varphi^T(t-s)\hat{\theta}(t-s-1) \right]\| \\ & \leq \sum_{s=0}^{k-1} \|\bar{R}^{-1}(t-s)\| \cdot \|\varphi(t-s)y(t-s) - \varphi(t-s)\varphi^T(t-s)\hat{\theta}(t-s-1)\| \end{aligned} \quad (24)$$

가정 2에 의해 모든 t에 대하여  $t \in (0, t_{\min})$

$\|\varphi(t-s)y(t-s) - \varphi(t-s)\varphi^T(t-s)\hat{\theta}(t-s-1)\| < \delta$ 를 만족하는 양수  $\delta > 0$ 가 존재한다. 또한 보조정리 2에 의해

$\|\bar{R}(t)\| < \frac{1-\lambda_{\max}^{-t-1}}{1-\lambda_{\max}^{-1}} \cdot \frac{\gamma^2}{n+m}$ 이므로 (24)는 다음과 같다.

$$\begin{aligned} & \|\hat{\theta}(t) - \hat{\theta}(t-k)\| \\ & \leq \sum_{s=0}^{k-1} \|\bar{R}^{-1}(t-s)\| \cdot \|\varphi(t-s)y(t-s) - \varphi(t-s)\varphi^T(t-s)\hat{\theta}(t-s-1)\| \\ & \leq \sum_{s=0}^{k-1} \frac{1-\lambda_{\max}^{-t-s+1}}{1-\lambda_{\max}^{-1}} \cdot \frac{n+m}{\gamma^2} \cdot \delta = (1-\lambda_{\max}^{-1}) \cdot \frac{n+m}{\gamma^2} \cdot \delta \cdot \frac{k}{1-\lambda_{\max}^{-t-s+1}} \end{aligned} \quad (25)$$

위 식에서 t를 제외한 다른 계수들은 모두 주어진 상수로서 임의로 선택할 수 없다. 그런데 t는 최소한 k보다 커야 하므로 0으로 감소할 수 없으며 오히려 증가하여야 하는데 이 경우에는  $\lambda_{\max} < 1$ 이므로  $t \rightarrow t_{\max}$ 에 따라  $\frac{1}{1-\lambda_{\max}^{-t-s+1}} \uparrow$ 이 되므로  $\|\hat{\theta}(t) - \hat{\theta}(t-k)\|$ 를 임의의 어떤 양수  $\varepsilon$  이내로 만들 수 없다. 따라서  $\{\hat{\theta}(t)\}_{t=0}^{t=t_{\max}}$ 는 Cauchy 수열이 아니다. ■

정리 1은 제안된 알고리즘 뿐 아니라 일반적인 시불변 망각 인자를 사용하였을 때 알고리즘이  $\mathcal{R}^n$  위상에서 불안정함을 증명한다. 즉, 망각 인자는 전체시간  $t \in (0, t_{\min})$ 에 대하여 일정한 값이므로,  $\lambda_{\max}$ 를 시불변 망각 인자  $\lambda$ 로 바꾸면 마찬가지로의 결론을 얻게된다.

정리 1에서 알 수 있는 것은 또한 망각 인자  $\lambda$ 가 1로 단조증가할 때  $\lambda_{\max}$ 는 시간의 함수로서  $t > t-k$ 에 대하여 시간  $t$ 의 함수로서 주어진다. 즉,  $\lambda_{\max} = \lambda(t)$ 가 된다. 그러므로 정리 1의 결과는 망각 인자  $\lambda$ 가 1로 단조증가할 경우 알고리즘이  $\mathcal{R}^n$ 위상에서 파라미터의 수렴성이 보장되는 결론을 유도할 수 있다.

정리 2 : 제안한 알고리즘은  $t > t-k \in (t_{\min}, \infty)$ 인  $\theta$ 에 대하여 가정 1, 2의 조건이 만족되는 경우 추정 파라미터 수열은 수렴성이 보장된다.

증명 : 정리 2는 추정 파라미터 수열  $\{\hat{\theta}(t)\}_{t=t_{\min}}^{\infty}$ 는 Cauchy 수열이라는 것과 동치이다. 따라서 임의의 양수  $\epsilon$ 에 대하여  $t \uparrow \infty$ 이면  $\|\hat{\theta}(t) - \hat{\theta}(t-k)\| < \epsilon$ 가 될 수 있음을 증명하면 된다.

먼저 보조정리 1의 결과는 정리 2의 가정을 주었을 경우  $\lambda_{\max} = \lambda(t)$ 이므로 다음과 같이 나타낸다.

$$\|\bar{R}(t)\| < \frac{1 - \lambda^{n+m}(t)}{1 - \lambda(t)} \cdot \frac{\gamma^2}{n+m} \quad (26)$$

따라서,  $\|\hat{\theta}(t) - \hat{\theta}(t-k)\|$ 는 다음 식과 같이 나타낸다.

$$\begin{aligned} & \|\hat{\theta}(t) - \hat{\theta}(t-k)\| \\ & \leq \sum_{s=0}^{k-1} \|\bar{R}^{-1}(t-s)\| \cdot \|\varphi(t-s)y(t-s) - \varphi(t-s)\varphi^T(t-s)\hat{\theta}(t-s-1)\| \\ & \leq \sum_{s=0}^{k-1} \frac{1 - \lambda(t)}{1 - \lambda^{s+1}(t)} \cdot \frac{n+m}{\gamma^2} \cdot \delta = \frac{1 - \lambda(t)}{1 - \lambda^{k+1}(t)} \cdot \frac{n\delta^{n+m}}{\gamma^2} \end{aligned} \quad (27)$$

(27)에서  $\frac{n\delta^{n+m}}{\gamma^2}$ 는 상수로서 주어진 값이다. 따라서  $\|\hat{\theta}(t) - \hat{\theta}(t-k)\|$ 를 임의의 작은 양수  $\epsilon < 1$ 이내로 보낼 수 있는가 하는 것은  $\frac{1 - \lambda(t)}{1 - \lambda^{k+1}(t)}$  항이  $t \uparrow \infty$ 에 따라 0으로 수렴할 수 있는가를 살펴보면 알 수 있다. 먼저 임의의 양수  $h$ 를  $\lambda^{s+1}(t) < h < 1$ 가 되도록 선택한다. 그러면

$$\frac{1 - \lambda(t)}{1 - \lambda^{s+1}(t)} < \frac{1 - \lambda(t)}{1 - h} \quad (28-a)$$

그런데 가정에서  $\lambda > h$ 에 따라  $\lambda(t) > h$ 이므로 임의의 양수  $\epsilon$ 에 대하여

$$\frac{1 - \lambda(t)}{1 - h} \cdot \frac{n\delta^{n+m}}{\gamma^2} > \epsilon \Leftrightarrow \lambda(t) > 1 - \epsilon \frac{\gamma^2(1-h)}{n\delta^{n+m}} \quad (28-b)$$

여기서  $h \geq 1 - \frac{n\delta^{n+m}}{\gamma^2}$ 를 만족하는  $\lambda(t)$ 를 취할 수 있다. 따라서 제안된 알고리즘에 의해 만들어지는 수열  $\{\hat{\theta}(t)\}_{t=t_{\min}}^{\infty}$ 는 Cauchy 수열이다. ■

정리 2의 결론은 망각 인자  $\lambda(t)$ 가 1로 단조증가하는 알고리즘이면 안정성이 보장된다는 것을 의미한다.

IV. 사례연구

실험은 제안한 알고리즘과 기존의 여러 알고리즘을 시스템 식별에 적용하여 수행하였다. 그림 2의 시스템 식별구조에서, 바라는 신호(desired signal)는 FIR필터의 형태를 갖는 미지의 시스템에서 출력되며 평균이 0인 백색 잡음이 noise로서 미지 시스템에 작용한다고 가정하였다. 미지 시스템의 입출력상태는 그림 2와 같다.

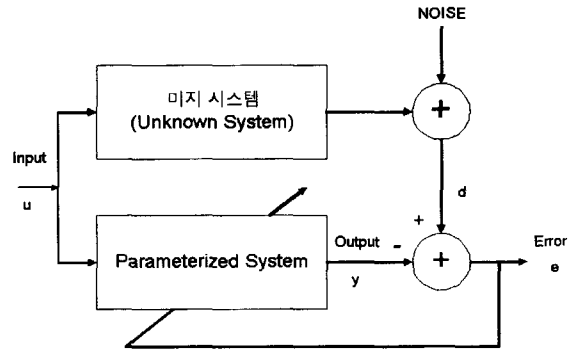


그림 2. 시스템 식별구조. Fig. 2. The structure of system identification.

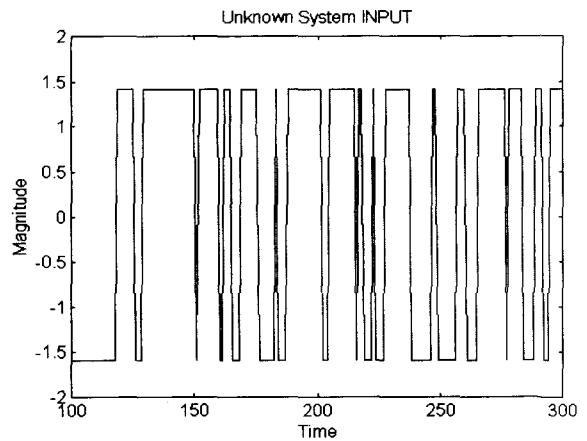


그림 3. 미지 시스템의 입력파형. Fig. 3. Input for unknown system.

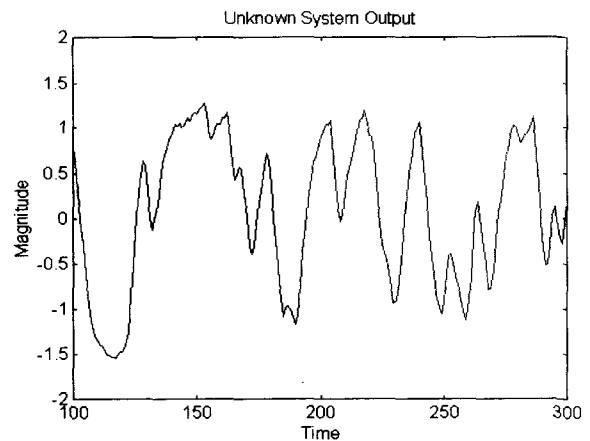


그림 4. 미지 시스템의 출력파형. Fig. 4. Output of unknown system.

실험은 미지 시스템을 다음과 같은 FIR 시스템으로 간주하여 각 파라미터  $a_1, a_2, b_3, b_4$ 를 추정하여 추정된 FIR시스템의 출력과 실제출력과의 오차를 비교한다.

$$y(t) + a_1y(t-T) + a_2y(t-2T) = b_3u(t-3T) + b_4u(t-4T) \quad (29)$$

(29)로 주어지 추정 시스템(estimated system)은 akaike의 FPE(Final Prediction Error Criterion) test에서 각각 표 1과 같은 결과를 나타내어 잘 추정된 시스템임을 알 수 있다. 참고로 (29)이외의 몇몇 시스템의 경우, 각 4개 추정모델과 추정방법에 대한 FPE test의 합산평균 결과는 표 2와 같다.

표 1. 각 estimation방법에 의한 FPE test 결과.  
Table 1. The result of FPE test for each estimation method.

Estimation Method	Akaike's FPE Test Result
IV4 Method	0.001532
ARX Model	0.001731
PEM Method	0.001708
ARMAX MODEL	0.001731

표 2. 각 Estimation System에 대한 FPE test 결과.  
Table 2. The result of FPE test for each estimation system.

$y(t) + a_1y(t-T) + a_2y(t-2T) = b_3u(t-3T) + b_4u(t-4T) + b_5u(t-5T)$			
IV4	ARX Model	PEM Method	ARMAX
0.001686	0.001868	0.001843	0.001869
$y(t) + a_1y(t-T) + a_2y(t-2T) = b_2u(t-T) + b_3u(t-2T)$			
IV4 Model	ARX Model	PEM Method	ARMAX
0.007344	0.008595	0.008784	0.008872
$y(t) + a_1y(t-T) + a_2y(t-2T) + a_3y(t-3T) = b_2u(t-T) + b_3u(t-4T)$			
IV4 Method	ARX Model	PEM Method	ARMAX
0.001583	0.002594	0.002569	0.002595
$y(t) + a_1y(t-T) + a_2y(t-2T) + a_3y(t-3T) = b_2u(t-T) + b_3u(t-2T)$			
IV4 Method	ARX Model	PEM Method	ARMAX
0.005400	0.007764	0.007687	0.007808

(29)에서  $T$ 는 표본시간(sampling time)으로서 0.08초로 놓았다. 실험환경은 Pentium100 IBM 호환기종에서 행하였다. 실험에 사용된 각 알고리즘의 계수들은 다음과 같다. 먼저 일반적인 망각 인자법에서 망각 인자  $\lambda$ 는 다음과 같다.

$$\lambda = 0.98 \quad (30)$$

한편 지수함수적으로 감소하는 망각 인자법에서 망각 인자  $\lambda$ 와 계수와 망각 인자  $\lambda_0$ 는 다음과 같이 구성된다[4].

$$\lambda(t) = \lambda_0 \lambda(t-1) + (1 - \lambda_0) \quad (31)$$

(31)에서  $\lambda_0 = 0.98, \lambda(0) = 0.95$ 이다. NLMS법에서의  $\gamma$ 와 칼만 필터 알고리즘에서 추정된 파라미터 벡터  $\hat{\theta}(t)$

표 3. 시스템 식별문제에 대한 각 알고리즘의 실험결과.

Table 3. The experimental results of each algorithm for system identification.

RLS Method	Squared Error
Proposed Algorithm	2.066341
Kalman Filtering Algorithm	2.507428
Conventional Forgetting Factor Method	2.088428
Exponential Forgetting Factor Method	2.073088
Normalized LMS (NLMS)	4.120870

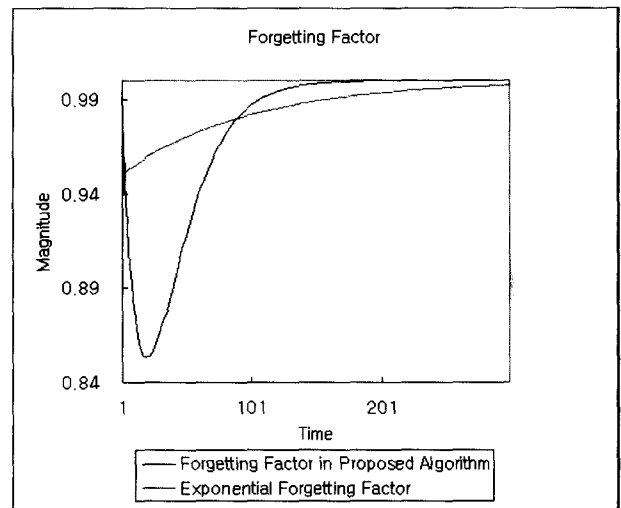


그림 5. 각 알고리즘에서의 forgetting factor.  
Fig. 5. Forgetting factor for each algorithm.

의 초기 공분산 행렬은 다음과 같이 놓았다.

$$P(0) = \gamma I_{\gamma=0.2} \quad (32)$$

제안한 알고리즘에서 Erlang함수의 망각 인자는 다음과 같다.

$$\lambda(t) = 1 - \alpha \cdot t \cdot \exp(-\beta t) \quad (33)$$

(33)에서  $\alpha = 0.02, \beta = 0.05$ 이다. 제안한 알고리즘과 지수함수적으로 증가하는 망각 인자의 시간에 따른 변화는 그림 5와 같다.

실험에서의 평가항목은 다음과 같은 오차의 자승합이다.

$$\sum_{k=0}^M (d(k) - y(k))^2 \quad (34)$$

위 식에서  $d(k)$ 는 주어진 미지 시스템의 출력이며  $y(t)$ 는 추정된 시스템의 미지 시스템과 동일 입력을 인가하였을 때의 출력이다.

### V. 결론

본 연구에서는 Erlang 함수를 사용한 새로운 망각 인자를 가진 RLS알고리즘을 제안하였다. 제안한 알고리즘은 초기 망각 인자가 감소할 때는 annealing효과를 보

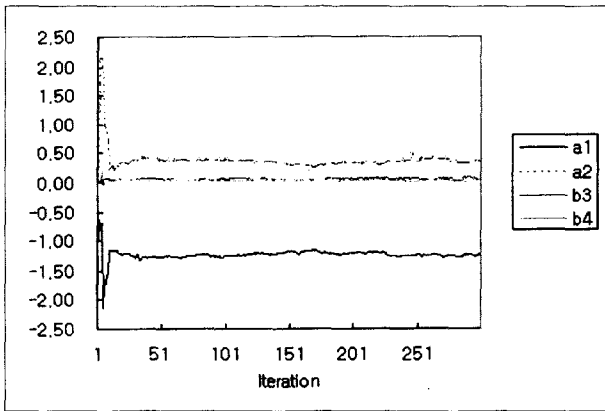


그림 6. 칼만 필터 알고리즘에서의 파라미터 수렴 상태.  
 Fig. 6. The convergence property of system parameter in Kalman filter algorithm.

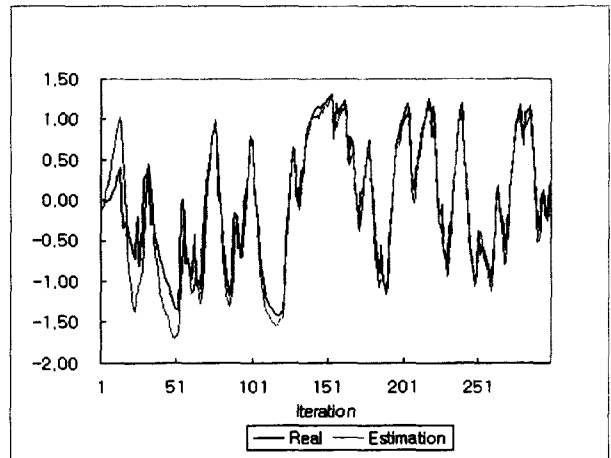


그림 9. NLMS 알고리즘에서의 추정된 시스템의 출력.  
 Fig. 9. The output property of estimated system in NLMS algorithm.

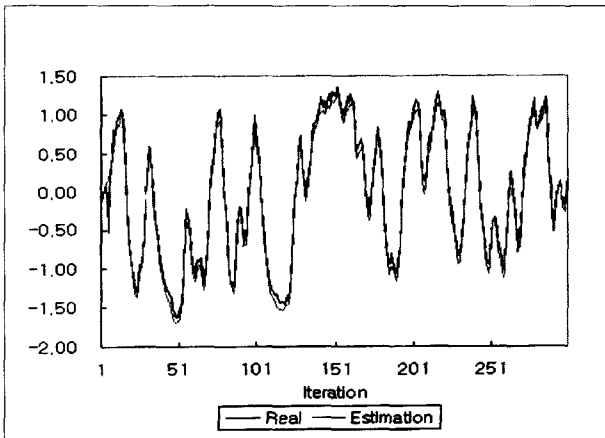


그림 7. 칼만 필터 알고리즘에서의 추정된 시스템의 출력.  
 Fig. 7. The output property of estimated system in Kalman filter algorithm.

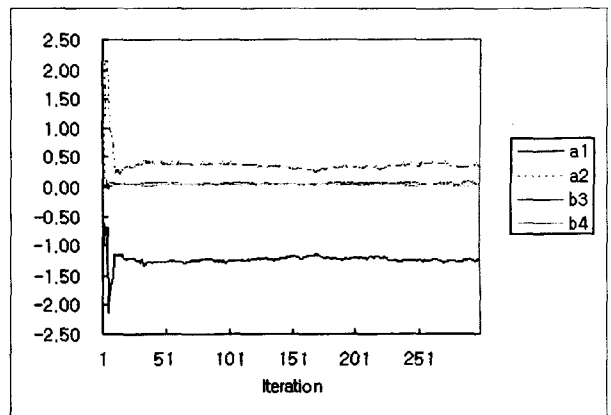


그림 10. Exponential forgetting factor 알고리즘에서의 파라미터 수렴 상태.  
 Fig. 10. The convergence property of system parameter in exponential forgetting factor algorithm.

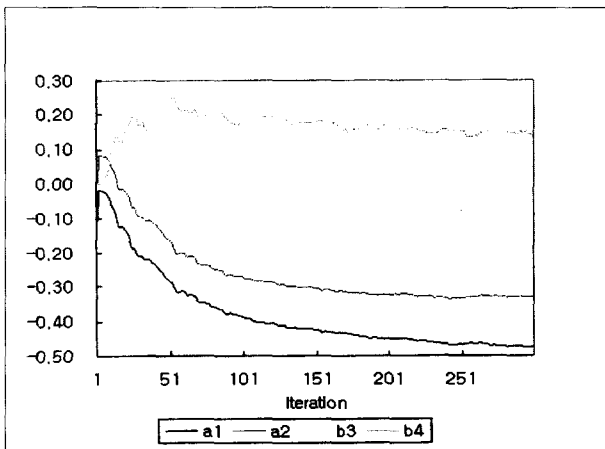


그림 8. NLMS 알고리즘에서의 파라미터 수렴 상태.  
 Fig. 8. The convergence property of system parameter in NLMS algorithm.

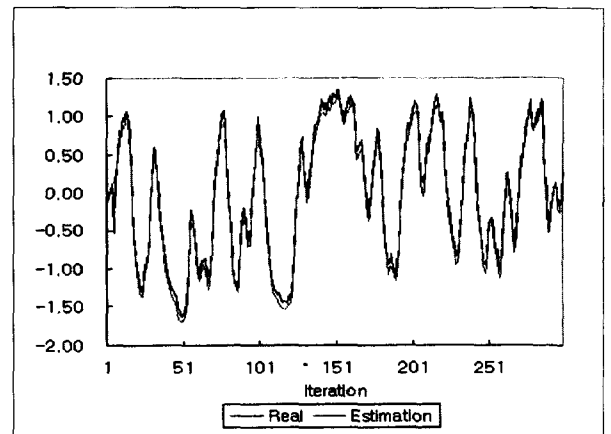


그림 11. Exponential forgetting factor 알고리즘에서의 추정된 시스템의 출력.  
 Fig. 11. The output property of estimated system in exponential forgetting factor algorithm.

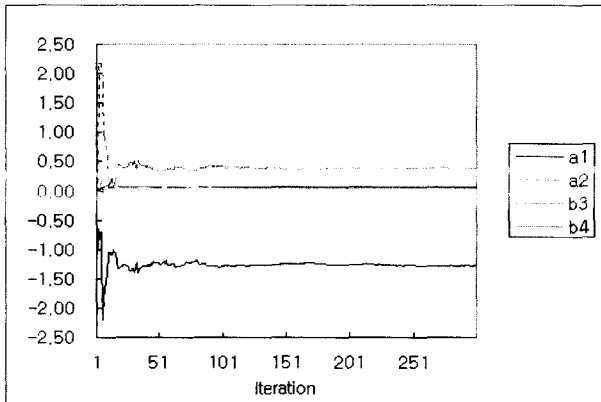


그림 12. 제안된 알고리즘에서의 파라미터 수렴상태.  
Fig. 12. The convergence property of system parameter in the proposed algorithm.

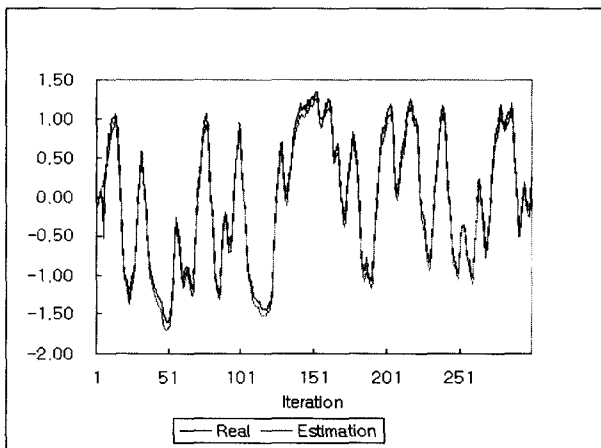


그림 13. 제안한 알고리즘에서의 추정된 시스템의 출력.  
Fig. 13. The output property of estimated system in the proposed algorithm

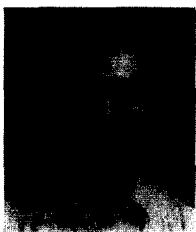
이지만 이후 망각 인자가 1로 단조증가함에 따라 추정된 파라미터가  $\mathfrak{R}^n$  위상에서 수렴해 들어감을 이론적으로 증명하였으며 컴퓨터 모의실험을 통한 결과, 추정된 시

스템의 출력응답은 대체로 비슷한 모습을 보여주고 있으나 표 3에서 나타나듯이 포인트 100에서 300까지의 오차의 자승합은 본 논문에서 제안한 방법이 여러 RLS 알고리즘 계열들의 결과보다 시스템 식별에서 보다 나은 결과를 가져다 줄 수 있음을 알 수 있다. 또한 각 파라미터의 수렴결과를 보았을 때 제안된 알고리즘의 경우 칼만 필터 알고리즘이나 NLMS 알고리즘보다 더 좋은 결과를 나타내고 있음을 알 수 있었다.

본 사례연구의 비교대상인 칼만 필터 알고리즘이나 NLMS 알고리즘 등은 시스템식별 분야 이외에도 여러 분야에 알고리즘의 수정없이 거의 그대로 이용되는 알고리즘이다. 따라서 제안한 알고리즘 역시 칼만 필터 알고리즘이나 NLMS 알고리즘 등이 응용되는 분야에서 마찬가지로 널리 응용될 수 있으리라 전망된다.

**참고문헌**

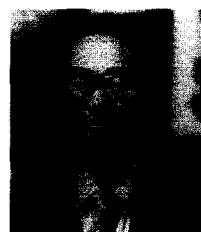
- [1] L. Guo, L. Ljung, "Performance analysis of general tracking algorithms," *IEEE Trans. Automatic Control*, vol. 40, no. 8, August, pp. 1388-1402, 1995.
- [2] L. Guo, L. Ljung, "Exponential stability of general tracking algorithms," *IEEE Trans. Automatic Control*, vol. 40, no. 8, August, pp. 1376-1387, 1995.
- [3] L. Ljung, *System Identification : Theory for the User*. Prentice Hall : New York, 1987.
- [4] T. Soderstrom, P. Stoica, *System Identification*, Prentice Hall : New York, 1989.
- [5] N. Kalouptsidis, S. Theodoridis, *Adaptive System Identification and Signal Processing Algorithms*. Prentice Hall : New York, 1993.
- [6] H. Kushner, A. Schwartz, "Weak convergence and asymptotic properties of adaptive filters with constant gains," *IEEE Trans. Information Theory*, vol. IT-30, no. 2, March, pp. 172-184, 1984.
- [7] M. Vidyasagar, *Nonlinear System Analysis*, Prentice Hall : New Jersey, 2nd Ed., 1993.
- [8] R. Otten, L. van Ginneken, *The Annealing Algorithm*, Kluwer Academic Publishers : Boston, 1989.



**석진욱**

1969년 6월 26일생. 1993년 2월 홍익대학교 전기제어 공학과 졸업(공학석사). 1998년 8월 동 대학원 전기공학과 졸업(공학박사). 1998년-현재 홍익대학교 과학기술연구소 재직. 관심분야는 신경회로망, 확률제어, 비선형제어 등.

선형제어 등.

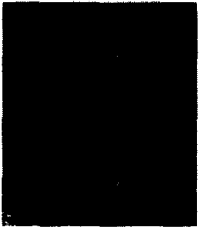


**최경삼**

1940년 11월 3일생. 1963년 연세대 전자공학과 졸업. 1968년 동 대학원 전기공학과 졸업(공학석사). 1973년 동 대학원 전기공학과 졸업(공학박사). 1970년-현재 홍익대 공대 전자전기공학부 조교수, 부교수, 교수. 주 관심분야는 이산시간 제어 시스템, 시스템 식별, 로봇틱스, 인공지능 제어 등.

스, 인공지능 제어 등.



**이 종 수**

1957년 5월 3일생, 1980년 서울대 공대 전기 공학과 졸업. 1982년 서울대 대학원 계측제어공학과 졸업(공학석사), 1988년 서울대 대학원 제어계측 공학과 졸업(공학박사). 1985년-현재 홍익대학교 조교수, 부교수, 주 관심분야는 로봇제어, 비선형제어

**조 성 원**

제어·자동화·시스템공학 논문지 제2권, 제2호, 참조.