파라메타 추정을 위한 지수 함수적 Sliding Window RLS 방식

최봉열* *　　이지태**　　변증남+

*경북대학교(전자공학과) **(화학공학과) +한국과학기술원 (전기및 전자공학과)

Exponentially Weighted Sliding Window RLS Method for Parameter Estimation

Bong Yeol CHOI, Jietae Lee, and Zeungnam Bien

Abstract

In this paper, we propose a new method of parameter estimation which modifies the exponentially weighted RLS (WRLS) method using asliding window. It can be used in adaptive control for a class of linear time varying system. Computer simulations show that our new method tracks rapidly time varying parameters more effectively than WRLS.

1. 서　론

최근 많이 이용되는 적응제어를 공정에 도입하기 위해서는 먼저 그 공정의 dynamics를 특정시키는 파라메타에 대한 정확한 값의 추정이 요구된다. 그러나 대부분의 공정의 파라메타들의 값을 모를뿐아니라 환경이 변함에따라 시간적으로 변화되어진다. 그래서 시변계의 파라메타 추정이 필요로 하게 되었으며 여기에 대한 연구가 계속되어져왔다. 시변계의 추정에 있어서 수렴속도는 매우 중요한 요소인데 만약 빠른 수렴속도가 필요하지 않는다면 단지 하나의 Recursive vector 알고리즘을 사용하여 수렴이 가능한 추정 알고리즘을 사용하면 된다. 그러나 이것은 단지 많은 시간이 흐른후 진짜 파라메타 값에 도달하게 되므로 결코도움이 되지 않는다.

Gradient 방식은 공정의 파라메타에 수렴되는 속도가 상당히 늦어 시간에 대해 천천히 변화하는 공정에 적당하고 수시로 갑자기 변하는 공정의파라메타 추정에는 적당하지 못하다.

최소자승(LS) 계열의 방식은 출력오차의 제곱의 합을 최소화하는 최적 특징으로인해 빠른 수렴속도를 가지고 있다. 그중 RLS는 알고리즘에 사용된 이득행열이 영으로 접근하게 되어 시변계의 추정에는 부적당한 단점을 가지고 있다. 시변계에 적용하기 위해서 과거의 data를 잊어버리기 위해 window를 사용한 방식으로 지수함수적 가중 RLS(WRLS)와 Sliding window RLS(SWRLS)이 있다. WRLS방식은 성능이 가중계수의 선택에 매우민감하고 모든 가중함수 값에대해 빈약한 성능을 나타낸다. SWRLS 방식은 WRLS방식에 비해 수렴속도가 개선되었으나 추정초기치로 인한 오차가 시간이 지남에따라 감소하지 않기때문에 시불변계의 추정에있서 정확한 값으로 수렴하지 못하는 단점을 가지고 있다. 본 논문에서는 WRLS 방식에 Sliding window를 사용하여 수렴속도가 개선된 새로운 알고리즘을 제안하였다(SWWRLS).

이 알고리즘은 최근 N개의 data를 지수함수적으로 가중하여 사용하고 과거의 data를 사용하지 않기때문에 수렴속도를 빠르게 할 수 있으며 추정초기치로 인한 오차가 지수함수적으로 감소하기 때문에 정상상태 오차가 영이 되며 성능이 가중계수의 선택에 둔감하게 된다. 제안된 알고리즘이 수렴함을 수식적으로 입증하였으며 시뮬레이션을 통해 그 타당성을 증명해 보였다.

2. 파라메타 추정 이론

추정하려는 공정을 다음과 같은 차분 방정식으로 표현한다.

$$A(q^{-1})\, y(k) = B(q^{-1})\, u(k) \qquad (1)$$

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n}$$

$$B(q^{-1}) = b_o + b_1 q^{-1} + \ldots + b_m q^{-m}$$

여기서 u(k)와 y(k)는 각각 공정의 입, 출력 수열이며 추정하려고 하는 공정에 대하여 다음과 같은 가정을 한다.

1)다항식 $A(q^{-1})$, $B(q^{-1})$의 차수 n,m을 알고 있다.

2)공정은 최소위상이다. 즉, B(z)의 모든 영점들은 단위원 내에 존재한다.

식(1)을 다시 표현하면 다음과 같다.

$$y(k) = \phi(k-1)^T \theta \qquad (2)$$

$$\phi(k-1) = [y(k-1),\ y(k-2),\ \ldots\ y(k-n),\ u(k),\ldots\ u(k-m)]^T$$

$$\theta = [a_1,\ \ldots\ a_n,\ b_o,\ \ldots\ b_m]^T$$

여기서 $\phi(k)$, θ는 각각 공정의 측정벡터, 파라메타 벡터 이다.

다음과 같이 추정 출력과 출력오차를 정의 한다.

$$\hat{y}(k) = \phi(k-1)^T \hat{\theta}(k) \qquad (3)$$

$$e(k) = y(k) - \hat{y}(k) \qquad (4)$$

여기서 $\hat{\theta}(k)$ 는 파라메타 벡터 θ의 추정치이다.

미지의 파라메타벡터를 추정하기 위해 평가함수 J(k)를 다음과 같이 정의하고 J(k)가 최소가 되도록 $\hat{\theta}(k)$를 결정한다.

$$J(k) = \sum_{i=K-N+1}^{K} \lambda^{k-i} e(i)^2 \qquad (5)$$

여기서 λ 는 $0<\lambda<1$인 가중계수이며 N은 Window의 크기이다.

이것을 $\hat{\theta}(k)$에 대하여 미분하고 그결과를 영으로 두면 다음식이 성립한다.

$$\hat{\theta}(k) = \left[\sum_{K-N+1}^{K} \lambda^{k-i} \phi(i-1)\phi(i-1)^T\right]^{-1} \left[\sum_{K-N+1}^{K} \lambda^{k-i} y(i)\phi(i-1)\right] \qquad (6)$$

알고리즘을 쉽게 적용하기 위해 On-line방식으로 고치면 다음과 같다.

$$\hat{\theta}(k+1) = \hat{\theta}(k) + F(k+1)[e^o(k+1)\phi(k) - e^*(k+1)\lambda^{N+1}\phi(k-N-1)] \qquad (7)$$

$$e^o(k+1) = y(k+1) - \phi^T(k+1)\hat{\theta}(k)$$

$$e^*(k+1) = y(k+1) - \phi^T(k+1)\hat{\theta}(k)$$

여기서 gain matrix F(k+1)은 다음과 같다.

$$F(k+1) = [\lambda F(k)^{-1} + \phi(k)\phi^T(k) - \lambda^{N+1}\phi(k-N-1)\phi^T(k-N-1)]^{-1} \qquad (8)$$

다음과 같이 C(k)을 정의 한다.

$$C(k) = [\lambda F(k)^{-1} + \phi(k)\phi^T(k)]^{-1} \qquad (9)$$

식(8),(9)에 Matrix Inversion Lemma를 적용하면 다음식이 성립한다.

$$C(k) = 1/\lambda\ [F(k) - F(k)\phi(k)\phi^T(k)F(k)/(\lambda + \phi^T(k)F(k)\phi(k))] \qquad (10)$$

$$F(k+1) = C(k) + \lambda^{N+1} C(k)\phi(k-N-1)\phi^T(k-N-1) C(k)/(1-\lambda^{N+1}\phi^T(k-N-1)C(k)\phi(k-N-1)) \qquad (11)$$

식 (7),(10),(11)이 제안된 파라메타 추정방식이다.

3. 수 렴

제안된 알고리즘이 수렴함을 증명하기위해 $\phi(k)$가 Second moment ergodic(s.m.e)하고 wide sense stationary(w.s.s)하다고 가정한다.

다음과 같이 공정의 파라메타 수정오차 $\tilde{\theta}(k)$를 정의한다.

$$\tilde{\theta}(k) = \theta - \hat{\theta}(k) \qquad (12)$$

그러면 식(7),(8),(12)로 부터 다음식이 성립한다.

$$\tilde{\theta}(k+1) = \lambda F(k+1)F(k)^{-1}\tilde{\theta}(k) = \lambda^{k+1} F(k+1)F(0)^{-1}\tilde{\theta}(0) \qquad (13)$$

$\phi(k)$가 s.m.e 하면 다음식이 성립한다

$$\lim_{K\to\infty} 1/k \sum_{i=0}^{K-1} \phi(i)\phi^T(i) = E[\phi(k)\phi^T(k)] = C \quad a.s. \qquad (14)$$

여기서 C는 constant matrix이다.

식(14)로부터

$$F(k+1)^{-1} = \sum_{i=k-N+1}^{K} \lambda^{k-i}\phi(i-1)\phi^T(i-1) = (1-\lambda^N)C/(1-\lambda) \quad a.s. \qquad (15)$$

$$F(k+1) = (1-\lambda)/(1-\lambda^N)\, C^{-1} \quad a.s. \qquad (16)$$

$\phi(k)$가 w.s.s하면 C^{-1}의 norm은 유한한 값을 가지게 되므로 F(k+1)의 norm도 유한한 값을 가지게 된다. K가 증가함에따라 λ^{k+1}이 영으로 수렴하기 때문에 $\theta(k+1)$은 식(13)에 의해 영으로 수렴하게 된다.

4. 시뮬레이션

본 논문에서는 다음과 같은 공정 파라메타에 대하여 WRLS방식과 제안된 SWWRLS 방식을 컴퓨터 시뮬레이션을 통하여 비교하였다.

$$\theta = \begin{cases} [0.7\ -0.5\ 0.3\ -0.4\ -0.2 \\ \quad 0.6\ 0.3\ -0.4\ 0.2\ 0.3] & k<100 \\ [0.8\ -0.3\ 0.4\ -0.2\ -0.4 \\ \quad 0.8\ 0.2\ -0.5\ 0.3\ 0.2] & k>100 \end{cases}$$

그림1은 100번째 iteration후의 추정오차 $\tilde{\theta}(k)$의 norm를 나타내고 있다. 여기서 λ 는 0.95이고 F(o)는 10000 I 이다. 파라메타의 초기치는 $\hat{\theta}(o)$=[o o ... o]이고 window크기는 N=20이다. 그림1에서 알수있듯이 SWWRLS방식이 빨리 수렴하고있다.

추정성능을 측정하기 위해 다음을 정의 하였다.

$$cost=1/200 \sum_{k=100}^{499} \| \tilde{\theta}(k) \|^2$$

그림2에서 보는바와 같이 SWWRLS 방식이 가중계수 λ 가 큰 영역에서 좋은 성능을 보일뿐만 아니라 λ값에 매우 둔감함을 알수있다. 여기서 출력에 Gaussian noise(0,0.02)가 섞여있다.

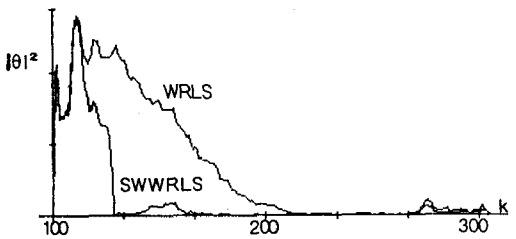

Fig. 1. $|\theta|^2$ trajectories.

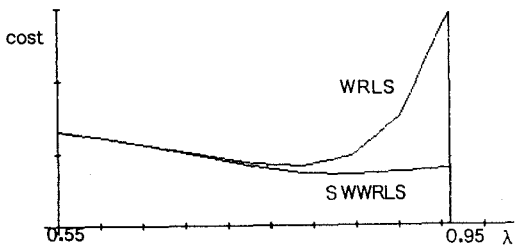

Fig. 2. Cost versus λ.

5. 결 론

본 논문에서는 WRLS 방식에 Sliding window를 사용하여 수렴속도가 개선된 새로운 파라메터 추정방식을 제안하였다. 제안된 추정방식은 시변계에서의 파라메터 추정성능이 WRLS보다 개선되었으며 가중계수의 선택에 둔감한 특성을 보였다. 수식적으로 안정함을 보였으며 시뮬레이션을 통해 WRLS와 비교 분석해 보았다.

References

[1] X.Xianya and R.J.Evans, "Discrete-time adaptive control for deterministic time varying systems," Automatica, Vol.20, no.3, pp309-319, 1984.

[2] S.Haykin, Adapive filter theory, Prentice-Hall, 1985.

[3] G.C.Goodwin, and K.S.Sin. Adaptive filtering prediction and Control Prentice-Hall, 1983.

[4] K.Tsakalis and P. Ioannou, "Adaptive Control of Linear time varying plamts", Automatica, Vol.23, pp459-468, 1987.