

# 완전최소자승법을 이용한 잡음환경하에서 시스템의 적응 역 모델링

황재섭, 김진영, 성경모  
서울대학교 공과대학 전자공학과

## Adaptive Inverse Modelling of Noisy System by Total Least Squares

Jaesub Hwang, Jinyoung Kim, Koengmo Sung  
Dept. of Electronics Eng., Seoul National Univ.

### 요 약

RLS(Recursive Least Squares)나 LMS(Least mean square) 등은 알고리즘 고유의 성질상 잡음이 섞인 시스템에 있어서는 올바른 역 모델링을 할 수 없다. 따라서, 잡음의 영향을 받지 않는 견실한(robust) 모델 추정 알고리즘이 필요하다. 본 논문에서는 잡음환경하에 있는 시스템을 역 모델링하는데 있어서, 잡음의 영향을 줄이기 위해 완전최소자승법을 도입하고 기존의 최소자승법과 비교 실험하였다. 그리고, 이 방법의 적응 알고리즘을 제안하였으며, RLS(Recursive least squares)와 그 성능을 비교하여 타당성을 검토하였다.

### I. 서 론

일반적으로 최소자승 문제에 있어서는 모든 오차가 단지 관측 벡터에만 한정되어 있다는 가정을 그 기본 정의에 내포하고 있다 [1,2]. 그러나 불행하게도 이러한 가정은 많은 실제 경우에 있어서 만족되지 않는다. 예를 들어, 시스템 역 모델링에 있어 표본화 오차, 모델링오차 그리고 장비오차(equipment error) 등은 입력 데이터에 대한 정확한 정보지식의 가능성을 배제한다. 관측벡터에서 발생하는 오차뿐만 아니라 이러한 입력 데이터에서 발생하는 오차를 보상해 주기 위해 고안된 것이 여러 curve fitting 기법 중 하나인 완전최소자승법이다[3]. 이들 두 방법은 모두 최적의 시스템 계수를 구할 때 오차에너지를 최소화시키려는 방향으로 식들을 풀어나간다. 최소자승법을 사용하였을 경우의 최소오차에너지는 항상 완전최소자승법의 그것에 비해 크게 된다. 즉 완전최소자승법을 시스템 역 모델링에 사용하면, 기존의 최소자승법을 사용한 경우보다 더 나은 성능을

기대할 수 있다.

일반적으로 플랜트의 특성은 알 수가 없을 뿐더러 긴 시간에 걸쳐서 천천히 변화하는 성질을 가지기 때문에 역 모델링을 하는데 적응성(adaptivity)은 불가피하다. 그러기 위해서는 완전최소자승법의 적응 알고리즘이 필요하게 되는데, 본 논문에서는 순환적인 완전최소자승법을 제안하여 잡음환경에 놓인 시스템의 적응 역 문제에 기존의 RLS와 비교하여 그 타당성을 검토하였다.

본 논문의 순서는 제 II장에서는 완전최소자승법에 대해서 기술하였다. 제 III장에서는 이를 이용하여 적응 알고리즘을 제안하고 제 IV장에서는 이들 방법의 컴퓨터 모의 실험의 결과를 보였으며, 제 V장에서는 결론을 맺으며 앞으로의 연구방향에 대하여 고찰하고자 한다.

### II. 완전최소자승법에 대한 고찰

완전최소자승법은  $m \times 1$  관측벡터인  $b$ 와  $m \times n$  데이터 행렬  $A$  모두에 있어서 오차가 있을 때 적합한 fitting 방법이다. 이 기법은 행렬  $A$ 의  $i$ 번째 ( $i = 1, 2, \dots, m$ ) 열벡터  $a_i^T$ 에 있어서 공간상의 한 점  $(a_i^T, b_i)$ 에 최적인 부분공간 또는 평면벡터를 fitting하기 위한 것이다. 관측 벡터  $b$ 의 교란량  $r$ 과 데이터 행렬  $A$ 의 교란량  $E$ , 모두를 고려하고 있는 완전최소자승법의 착상에 대해 고찰해 보겠다. 주어진 nonsingular 가중행렬이 아래와 같다고 하자.

$$D = \text{diag}(d_1, d_2, \dots, d_m), \quad d_i > 0, \quad i = 1, 2, \dots, m \quad (1)$$

$$T = \text{diag}(t_1, t_2, \dots, t_{n+1}), \quad t_i > 0, \quad i = 1, 2, \dots, n+1$$

이 때 식 (2)를 최소화하는  $X$ 를 찾는 것이다.

$$\underset{X}{\text{minimize}} [D|E|T]_r^T \quad (2)$$

여기서  $\|\cdot\|_F$ 는  $\|B\|_F^2 = \sum_i |b_i|^2$ 으로 정의되는 Frobenius 노름(norm)이다. 일단 최소로 하는  $[\hat{E}, \hat{r}]$ 를 찾을 수 있다고 가정하자.

$$(A + \hat{E})X = b + \hat{r} \quad (3)$$

식(3)을 만족하는  $X$ 는 완전최소자승 문제를 푼다고 할 수 있다. 즉 각각의 오차정도를 나타내는 행렬  $E$ 와 벡터  $r$ 의 크기를 최소화시킬 수 있는 평면벡터  $X$ 를 구하는 것이 완전최소자승 문제를 푸는 것이 된다. 벡터 공간상의 한 점  $P(a^T, b)$ ,  $P \in \mathbb{R}^{m+1}$ ,  $P(a^T, b) = P(a_1, a_2, \dots, a_m, b)$ 를 지나고 식(4)로 주어진 평면에

$$X_1 a_1 + X_2 a_2 + \dots + X_m a_m - b = 0 \quad (4)$$

수직인 최소 거리를 구하면 식(5)와 같다.

$$PH = \frac{|X_1 a_1 + X_2 a_2 + \dots + X_m a_m - b|}{\sqrt{X_1^2 + X_2^2 + \dots + X_m^2 + 1}} \quad (5)$$

$m$ 번 연속의 부정된 합으로 나타내면 식(6)와 같다.

$$\sum_{i=1}^m \frac{|X_1 a_{i1} + X_2 a_{i2} + \dots + X_m a_{im} - b_i|^2}{X_1^2 + X_2^2 + \dots + X_m^2 + 1} = \sum_{i=1}^m \frac{|a_i^T X - b_i|^2}{X^T X + 1} = \sigma_{min}^2 \quad (6)$$

여기서  $\sigma_{min}^2$ 은 최소의 오차에너지를 나타내는 것이며 완벽한 fitting의 경우에는 영이 된다. 완전최소자승으로 추정된 후의 전체 오차에너지를 최소자승으로 추정된 후의 전체 오차에너지를 비교하기 위하여 간단한 2차원의 경우를 고려해보자.

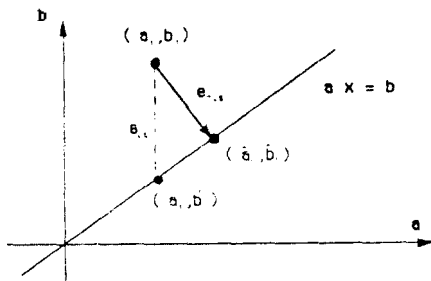


그림 2 : TLS와 LS의 오차거리 비교

그림 2에 나타난 것과 같이 완전최소자승을 사용한 경우의 추정치  $(\hat{a}_1, \hat{b}_1)$ 와 실제 한 점  $(a_1, b_1)$ 와의 오차에너지는 식(7)과 같이 쓸 수 있다.

$$e_{TLS}^2 = (a_1 - \hat{a}_1)^2 + (b_1 - \hat{b}_1)^2 \quad (7)$$

즉,  $a$ 에 대한 오차와  $b$ 에 대한 오차에너지를 합으로 나타낸다. 하지만 최소자승법을 사용한 경우에 있어 오차에너지는 식(8)로 주어지는데

$$e_{LS}^2 = (a_1 - \hat{a}_1)^2 + (b_1 - \hat{b}_1)^2 = (b_1 - \hat{b}_1)^2 \quad (8)$$

$b$ 에 대한 오차에너지만 나타낸다. 그림 2에서 보듯이 피타고라스의 정리를 사용하면 항상 식(9)의 관계를 유지한다.

$$e_{TLS}^2 \leq e_{LS}^2 \quad (9)$$

등호는 fitting된 직선이 벡터와 평행할 경우 뿐이다. 이 경우는 벡터  $b$ 와  $a$ 가 선형종속인 경우에 해당한다. 시스템의 역도행렬에 있어서는 전체 오차에너지를 기존의 최소자승법으로 얻어낸 최소자승계수보다 더욱 작게 가지는 완전최소자승계수를 얻을 수 있다.

### III. 적용 완전최소자승법

제 II장에서 유도한 식(6)를 행렬 형식으로 나타내면 다음과 같이 된다.

$$\sum_{i=1}^m \frac{|a_i^T X - b_i|^2}{X^T X + 1} = \frac{[1|X] \sum_{i=1}^m \begin{bmatrix} -b_i \\ a_i \end{bmatrix} - b_i |a_i^T| \begin{bmatrix} 1 \\ X \end{bmatrix}}{[1|X] \begin{bmatrix} 1 \\ X \end{bmatrix}} = \sigma_{min}^2 \quad (10)$$

식(10)의 분모항에서  $\sum_{i=1}^m \begin{bmatrix} -b_i \\ a_i \end{bmatrix} - b_i |a_i^T|$ 는 자기상관행렬이 된다.  $m \times m$  자기상관행렬  $A$ 은 그 통계적 성질상 대칭행렬이며 nonnegative한 고유치를 가진다. 또한 이 행렬의 distinct 한 고유치  $\lambda_1, \lambda_2, \dots, \lambda_m$ 에 상응하는 고유벡터들인  $P_1, P_2, \dots, P_m$ 은 각각 서로가 선형독립적이며 직교하게 된다[2]. 식(10)에서 행렬  $A$ 의 고유치가  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ 일 때 최소에너지의 범위가 최대고유치  $\lambda_m$ 와 최소고유치  $\lambda_1$  사이에 존재하게 됨은  $m$ 차원 직교변환(Orthogonal transform)을 이용하면 수식적으로 증명할 수 있다[2]. 즉 식(11)과 같이 최소에너지의 범위가 주어진다.

$$\lambda_m \leq \sigma_{min}^2 \leq \lambda_1 \quad (11)$$

식(10)에서 편의상 아래 두 변수를 정의하자.

$$\hat{X}^T = [1|X^T], \quad \hat{X} = \begin{bmatrix} 1 \\ X \end{bmatrix} \quad (12)$$

$$A = \sum_{i=1}^m \begin{bmatrix} -b_i \\ a_i \end{bmatrix} - b_i |a_i^T| \quad (13)$$

식(12)와 (13)을 이용하면 식(10)을 식(14)와 같이 이차형 형태(quadratic forms)로 기술할 수 있다.

$$\frac{\hat{X}^T A \hat{X}}{X^T X} \quad (14)$$

식(11)와 식(14) 그리고  $AP_i = \lambda_i P_i$  관계를 사용하면 식(15)로서 완전최소자승법을 푸는 식(10)의 해석이 가능해진다.

$$\frac{\hat{X}^T A \hat{X}}{\hat{X}^T \hat{X}} = \frac{\hat{P}_n A \hat{P}_n^T}{\hat{P}_n \hat{P}_n^T} = \lambda_n \quad (15)$$

행렬 A의 최소고유치  $\lambda_n$ 를 구하고 상응하는 고유벡터  $\hat{P}_n$ 를 구할 수 있다면 이 고유벡터  $\hat{P}_n$ 는 우리가 원하는 완전 최소자승의 최적치 X의 스케일된 양임을 알 수 있다. 즉, 벡터  $\hat{P}_n$ 를 정규화시키면 역모델링에서 원하는 시스템계수벡터인 X를 얻을 수 있다. 자기상관행렬과 같은 대칭행렬의 대각 고유치 행렬을  $\Lambda$ 라 하고 고유벡터를 행벡터로 가지는 행렬을 P라 하면 다음과 같은 unitary similarity transformation 관계를 이용하여 자기상관행렬의 역행렬에 해당하는 고유치 벡터를 얻을 수 있다.

$$(P^T A P)^{-1} = \Lambda^{-1} \quad (16)$$

식(16)에서 보면, 자기상관행렬의 최소 고유치 및 상응하는 고유벡터를 구하는 문제는 역행렬  $A^{-1}$ 의 최대 고유치 및 그에 상응하는 고유벡터를 구하는 문제로 바꾸어 생각할 수 있다. 자기상관행렬의 역행렬을 구하는 문제는 행렬대수에서 널리 쓰이는 행렬 역변환 보조정리(Matrix Inversion Lemma)를 사용하면 순환적인 계산으로 실 시간구현에 가능하도록 할 수 있으며, 입력의 통계적 성질이 변화하는 것도 고려할 수 있다. 이에 대한 수식 전개는 다음과 같다.

식 (10)에서  $\begin{bmatrix} -b \\ a \end{bmatrix}$ 를 벡터  $b(n)$ 로 두고, 어떤 시간 순간 n일 때의 자기상관행렬을  $A(n)$ 라고 하면 아래와 같이 기술할 수 있다.

$$A(n) = \lambda A(n-1) + B(n)B^T(n) \quad (17)$$

여기서  $\lambda$ 는 지수적 감쇠지수(exponential weighting factor)로서 nonstationary 환경에서 벡터  $B(n)$ 의 통계적 성질 변화의 가능성을 고려한 0과 1사이의 상수이다.

식 (17)에 행렬 역변환 공리를 적용하면 자기상관행렬의 역행렬은 다음과 같이 표현되어 진다.

$$A^{-1}(n) = \lambda^{-1} A^{-1}(n-1) - \frac{A^{-1}(n-1)B(n)B^T(n)A^{-1}(n-1)}{1 + \lambda^{-1}B^T(n)A^{-1}(n-1)B(n)} \quad (18)$$

계산의 편의상,  $P(n)$ ,  $k(n)$ 을 각각 식 (19)와 (20)으로 정의하자.

$$P(n) = A^{-1}(n) \quad (19)$$

$$k(n) = \frac{\lambda^{-1} P(n-1)B(n)}{1 + \lambda^{-1}B^T(n)P(n-1)B(n)} \quad (20)$$

식 (18)은 다음과 같은 방정식으로 된다.

$$P(n) = \lambda^{-1} P(n-1) - \lambda^{-1} k(n)B^T(n)P(n-1) \quad (21)$$

이상에서 식 (29)와 (20)을 사용하면 벡터  $B(n)$ 의 통계적 변화의 가능성을 염두에 두면서도 자기상관행렬  $A(n)$ 의 역행렬을 순환적으로 구할 수 있다. 최종적으로 얻어진 행렬  $P(n)$ , 즉 자기상관행렬의 역행렬에서 최대의 고유치와 상응하는 고유벡터를 구하는 문제는 문헌에 나와 있는 Power 방법을 사용하면 간단하게 해결되어 진다[4].

이렇게 얻어진 고유벡터가 완전최소자승 문제를 푼, 평면을 fitting하는데 있어서 점들간의 거리를 최소로 할 수 있는 평면벡터  $[t|X]$ 를 구한 것이 된다.

#### IV. 모의실험 및 평가

이 장에서는 앞장에서 고찰한 완전최소자승법의 적용 알고리즘을 이용하여, 임의로 만든 특정한 플랜트에 대한 역 모델링의 성능을 컴퓨터 모의실험을 통하여 기존의 RLS와 수렴성 및 오차장도를 평가하였다. 잡음 환경에 시스템이 노출되어 있을 때 최소자승법을 이론적 기반으로 하는 RLS를 사용한 경우와 제안한 적응 완전최소자승법을 사용한 경우에서 있어 기대했던 바의 값이 후자가 더 나은 시스템 역 모델링 결과를 보여주고 있다. 실험에 사용된 시스템에 대한 블록도는 그림 3에 나타나 있다.

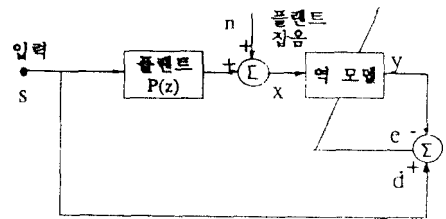


그림 3: 실험에 사용된 시스템 블록도

및 그림에서 플랜트의 전달함수는 아래와 같이 2개의 극점(pole)과 한 개의 영점(zero)를 가진다.

$$P(z) = \frac{1 - 0.4z^{-1}}{1 - 0.6z^{-1} + 0.9z^{-2}} \quad (22)$$

입력은 1과 -1이 무작위하게 연속되는 2000개의 데이터로써 이루어져 있다. 잡음 n은 플랜트의 출력단에 들어가는 백색잡음이라 하겠다. 그리고 적응 역 필터의 가중치 치수는 여러 번의 실험을 거쳐 적당한 12차로 고정시켰다. 먼저 잡음이 없는 이상적인 경우에 있어서는 그림 4에서 보듯이 두 방법은 동일한 결과를 보여준다.

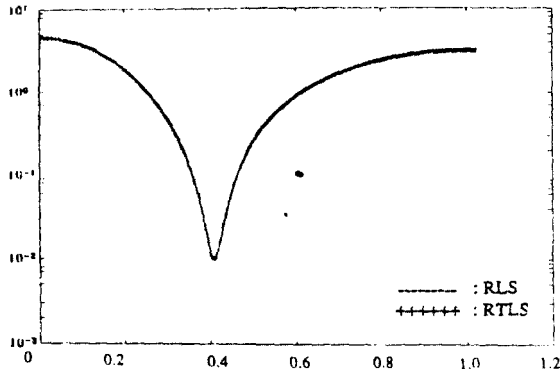
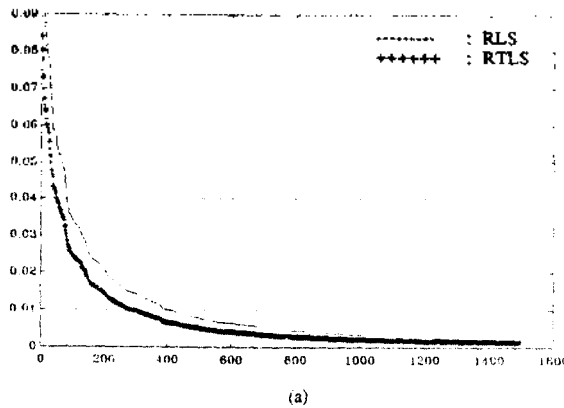
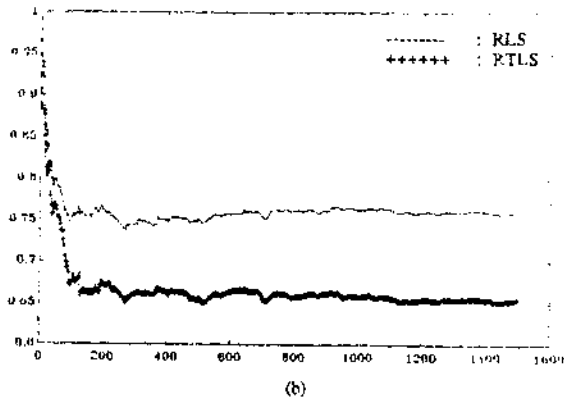


그림 4: 잡음이 없는 경우의 역모델 주파수 특성

그림 4에서 수렴속은 디지털 주파수 특성을 나타내며 수직축은 플랜트의 역 모델링된 주파수 특성을 대수좌표로 나타내었다. 잡음에 따른 두 방법의 수렴 및 오차정도를 반복횟수에 대해 비교할 수 있는 결과가 그림 5에 있다.



(a)



(b)

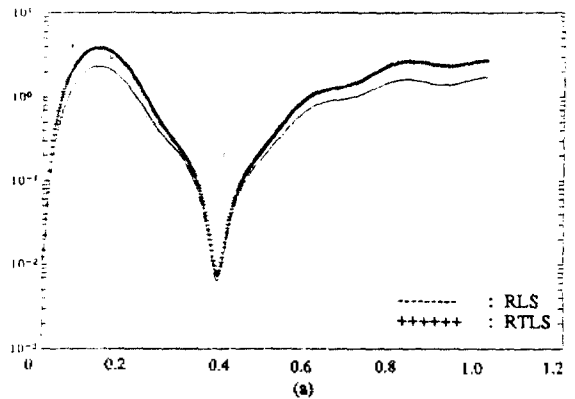
그림 5: 두 방법의 잡음에 대한 수렴성 비교

(a) 잡음이 없는 경우, (b) SNR = 0 dB

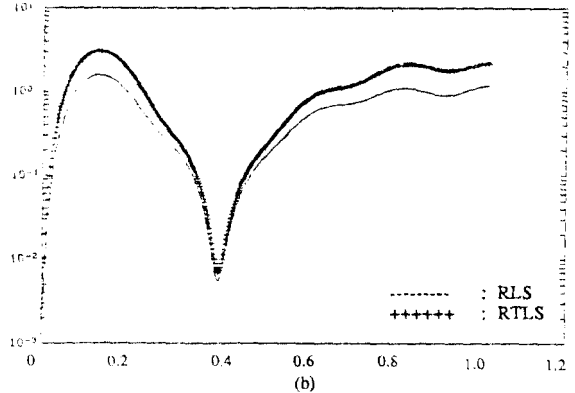
여기서 그림 5.(a)는 잡음이 없는 이상적인 경우, 두 방법에 대한 역 시스템의 파라미터 계수들의 수렴정도와 수렴오차를 보여준다. 위 경우와 같이 잡음이 없을 때 두 방법의 역 시스템에 대한 오차는

반복횟수가 커질수록 영에 수렴하나 그림에서 보아듯이 적용 완전 최소자승법을 사용한 경우가 기존의 RLS를 사용한 경우에 비해 뛰어난 수렴정도를 보여주고 있다. 그림 5.(b)는 신호대 잡음비가 0dB 일때의 수렴정도와 오차를 나타낸 것인데 이와같이 신호대 잡음비가 낮은 경우에는 수렴정도 뿐만 아니라 수렴오차에서도 큰 차이가 보인다.

그림 6은 플랜트 출력단의 잡음을 고려하여 그정도가 신호에 대해서 3.0dB와 같이 바뀌면서 얻은 플랜트의 역 모델링 주파수 특성을 보인 것이다. 그림 6.(b)에서와같이 잡음이 심한 경우 3 dB 이상의 성능차를 보이고 있다.



(a)



(b)

그림 6: 잡음에 대한 역 모델 주파수 특성

(a) SNR = 3 dB, (b) SNR = 0 dB

## V. 결과

본 논문에서는 완전최소자승법을 도입하여 시스템의 역 모델링에 적용하였으며 기존의 최소자승법과 비교, 그 성능을 검토하였다. 잡음환경하에 있는 시스템에 제안한 적용 완전최소자승법을 사용하여 기존의 RLS보다 뛰어난 성능을 얻었다. 두 방법의 성능차는 계산량의 단점을 충분히 보상하리라 본다. 두 기법의 수학적 성능비교

에 대한 연구가 계속되어야 할 것이며, 채널 동화 뿐만 아니라 음성 등의 시스템 계수 추정하는 문제 등 많은 연구의 여지가 있다 하겠다.

### 참고문헌

1. S.Haykin, *Adaptive Filter Theory*, 2 ed. Prentice-Hall, Inc., 1991.
2. A. Basilevsky, *Applied Matrix Algebra in the Statistical Science*, Elsevier Science Pub. Co., 1983.
3. G.H.Golub, C.F.Van Loan, "An Analysis of the Total Least Squares Problem," *SIAM, J.Number Anal.*, Vol.17, pp.883-893, 1990.
4. A. Jennings, *Matrix Computation for Engineers and Scientists*, John Wiley & Sons, 1977.