

# 로버스트 역전파 알고리즘을 위한 오차함수

김현철<sup>1)</sup> 이철원<sup>2)</sup>

## 요약

인공신경망 모형을 적합시키는데 사용하는 역전파 알고리즘을 로버스트하게 만드는 새로운 오차함수를 제안했으며, 새 방법의 성능을 확인하기 위해 Liano가 제안한 방법에 따라 모의실험을 수행했다. 실험결과 새 방법은 LMS방법만큼 안정적이었으며, Liano의 LMLS방법보다 더 로버스트했다. 또 실제 사례를 분석함으로써 이 방법이 의미있는 방법임을 보였다. 새 방법은 특히 오차가 없거나 작은 오차를 갖는 표본에 대해서도 좋은 성질을 가짐으로서 대형오차의 유무에 관계없이 항상 사용할 수 있는 방법으로 판명되었다.

## 1. 서론

역전파 알고리즘은 인공신경망 모형을 적합시키는데 사용하는 대표적인 알고리즘이다. 그런데 전통적인 역전파 알고리즘은 로버스트하지 않다. 즉 훈련표본에 이상점이 있을 때 전통적인 역전파 알고리즘은 이 이상점들의 영향으로 왜곡된 결과를 만들어 낸다. 이는 역전파 알고리즘이 함수를 근사시킬 때 평활(smoothing)에 의해 근사시키지 않고 내삽(interpolation)에 의해 함수를 근사시키기 때문이다(Chen과 Jain 1994).

역전파 알고리즘에서 출력층의 가중값은 다음과 같이 계산한 값으로부터 적절히 수정된다.

$$d_p = r_p o_p (1 - o_p) \tag{1.1}$$

여기서  $o_p$ 는  $p$ 번째 훈련표본에서 계산된 출력값이고,  $r_p$ 는 오차이다. 즉 오차  $r$ 이 수정값에 선형적으로 영향을 주게 되어 있어, 오차가 큰 표본점(이상점)이  $d_p$ 에 더욱 큰 영향을 미치게 된다. 이 사실이 전통적인 역전파 알고리즘을 이상점에 대해 로버스트하지 않게 만들고 있다. 이 문제는 그 동안 여러 문헌에서 과적합(over-fitting)의 문제로 다루어져 왔는데 (Freeman 1994, p.100 ; Smith 1993, pp.113-114 등), 이러한 논의는 크게 세 가지로 나누어 볼 수 있다.

첫번째 부류로는 Freeman(1994)을 들 수 있다. Freeman은 한 패턴에 대해 허용할 수 있는 최대오차의 크기를 임의로(예를 들면 0.1) 정한 다음 한 입력 패턴에 대한 오차의 크기가 이 값보다 작을 때는 가중치를 수정하지 않는 방법으로 다른 입력 패턴에 대한 적합도를 떨어뜨리면서까지 특정한 입력 패턴에 과적합되는 것을 막는 방법을 제안했다. 이 방법은 과적합의 문제를 다루는 가장 손쉬운 방법이지만 일정 수준 이상 적합되지 않아 최종 모형

1) (573-701) 전북 군산시 미룡동 산 68번지, 군산대학교 정보통계학과, 조교수  
2) (480-702) 경기도 의정부시 가능 3동 562-1, 경민대학 전자계산과, 조교수

의 적합도가 떨어진다는 문제점을 안고 있다. 특히 Chen과 Jain(1994)은 이상점이 있을 때 역전파 알고리즘에 의해 함수를 근사시키는 과정을 정밀한 수준까지 수렴될 수 있도록 오래 동안 반복시키지 않고 일찍 중단시켜도 여전히 이상점의 영향을 받는다는 사실을 보여주고 있다. 이는 함수를 적합시키는 과정에서 이상점 때문에 생기는 문제가 단순히 과적합이나 과학습(over-learning)의 문제가 아니라는 점을 보여준다.

두번째 부류로 Li 등(1993)은 앞의 식(1.1)에서  $o_p(1 - o_p)$ 에 초점을 두고 이 문제를 해결하려 시도하였다. 이들은  $o_p(1 - o_p)$ 가 전이함수(transfer function)로부터 만들어진다는 점에 착안하여 전이함수를 전통적인 시그모이드함수(sigmoid function)와 다른 새로운 전이함수를 제안하고 있다. 이 방법은 문제의 직접 원인이 되는  $r_p$ 가 선형함수라는 점을 해결하지 않고 있어 적절한 방법으로 보이지 않는다.

세번째 부류는 M-추정량의 개념을 도입하여 해결을 시도한 연구들이다. Chen과 Jain은 식(1.1)에서  $r_p$ 가 선형함수이기 때문에 문제가 된다는 점에 착안하여  $r_p$ 의 영향이 적절하게 정정될 수 있도록  $r_p$ 를 수정하는 오차함수(error function)를 제안하고 있다. 역전파 알고리즘을 로버스트하게 만들고자 할 때 사용할 수 있는 오차함수는 로버스트 통계에서 사용하는 M-추정량의 일반적인 오차함수들이 사용될 수 있다. Chen과 Jain이 채택한 것은 Hampel의 tanh추정량이다. 그들은 Hampel의 tanh추정량을 개량하여 학습이 진행됨에 따라 조율상수(tuning constant)의 값이 달라지게 만들었다. 이 때 조율상수를 정하기 위해서는 붓스트랩 방법으로 잔차에 대한 95% 신뢰구간을 계산해야 하므로 계산량이 폭발적으로 늘어나게 된다는 단점이 있다. 더욱이 이상점의 비율을 미리 알고 있어야 하기 때문에 Beckman과 Cook(1983)이 지적했던, 일부 이상점 검색방법들이 안고 있는 방해효과(masking effect)나 과장효과(swamping effect)를 피할 수 없다. 따라서 이 방법은 로버스트 통계학의 M-추정량을 개선하면서도 또 다른 문제점을 안게 되었다.

Wang 등(1996)은 Chen과 Jain의 방법이 계산량이 지나치게 많아져서 학습과정의 수렴속도를 비교할 때 반복회수는 크게 줄어들지만 계산시간은 오히려 증가한다는 문제점을 해결하기 위해 조율상수의 값을 수정하는 방법을 개선시켰다. 따라서 여전히 Chen과 Jain의 방법이 새로 안게 된 문제점들은 개선되지 않았다.

Liano(1996) 역시 로버스트추정법을 이용하여 이 문제를 해결하려 시도했다. Liano는 전통적인 로버스트추정에서 사용하던 영향함수 대신 새로운 영향함수를 찾았다. 그리고 여기에 대응하는 오차함수를 최소화하는 역전파 알고리즘을 제안했다. 이 방법은 역전파 알고리즘에 그대로 적용시킬 수 있다는 점과 로버스트추정의 개념을 그대로 사용했기 때문에 새로운 문제점이 개입되지 않는다는 점에서 장점이 있다.

따라서 본 연구는 Liano가 시도한 것처럼 새로운 오차함수와 영향함수를 통해 역전파 알고리즘을 로버스트하게 만드는 데 목적이 있다. 본 연구에서 제안하고 있는 오차함수는 Liano가 사용한 것에 비해 훨씬 일찍 상한값에 수렴하는 형태를 갖고 있다. 영향함수 역시 일찍 0에 수렴하는 성질을 갖고 있다. 따라서 대형오차(gross error)에 대해 Liano의 연구결과보다 더욱 로버스트해질 수 있다. 또 이 오차함수들이 얼마나 효율적인 지를 확인하기 위해 Liano가 제시하는 방법에 따라 똑 같은 모형으로 모의실험을 수행하였다.

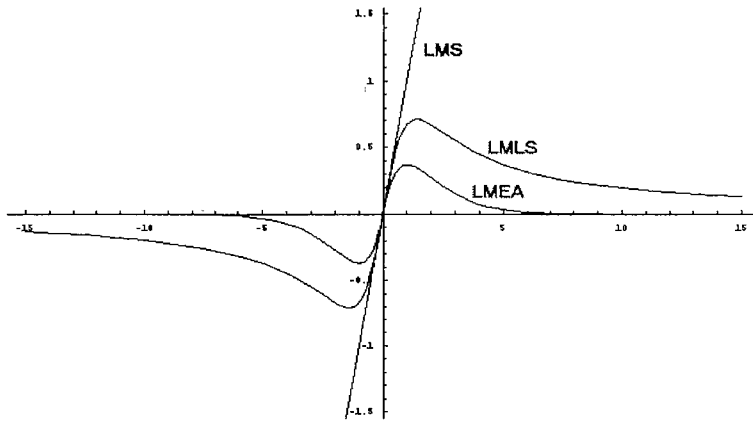


그림 2.1: 몇가지 영향함수의 예

## 2. 영향함수

학습 알고리즘은 어떤 비용함수를 생각하고 이 비용함수를 최소화하는 형태로 만들어진다. 비용함수는 보통 다음과 같이 오차의 함수로 표현된다.

$$E = \frac{1}{N} \sum_{i=1}^N \rho(r_i) \tag{2.1}$$

여기서  $\rho(\cdot)$ 는 오차함수로 대칭이고 연속인 함수이며,  $r_i$ 는  $i$ 번째 입력패턴에 대한 목표값에서 실제 출력값을 뺀 잔차로  $t_i$ 를 목표값이라고 하면,  $r_i = t_i - o_i$ 이다. 또  $N$ 은 훈련패턴의 수이다.

한편  $\mathbf{w}$ 를 가중값 벡터라고 하면  $E$ 를 최소화하는 가중값  $\mathbf{w}$ 는  $E$ 를  $\mathbf{w}$ 에 대해 미분한 값을 0으로 두고 풀면된다.

$$\frac{\partial E}{\partial \mathbf{w}} = \frac{1}{N} \sum_{i=1}^N \psi(r_i) \frac{\partial r_i}{\partial \mathbf{w}} \tag{2.2}$$

여기서  $\psi(r_i) = \frac{\partial \rho(r_i)}{\partial r_i}$ 이며, 이 값은 식(2.2)에서 점  $i$ 가 비용  $E$ 에 미치는 가중값 혹은 영향(influence)의 역할을 하고 있으므로 영향함수라고 한다. 영향함수  $\psi$ 가 가져야할 바람직한 성질로는 오차  $r$ 이 일정한 값 이하일 때는 선형에 가깝고  $r$ 이 커짐에 따라 점점 완만하게 증가하다가  $r$ 이 더욱 커지면 오히려 감소하는 형태이어야 한다는 성질이 있다. 이런 M-추정량을 재감소(redescending) M-추정량이라고 한다.

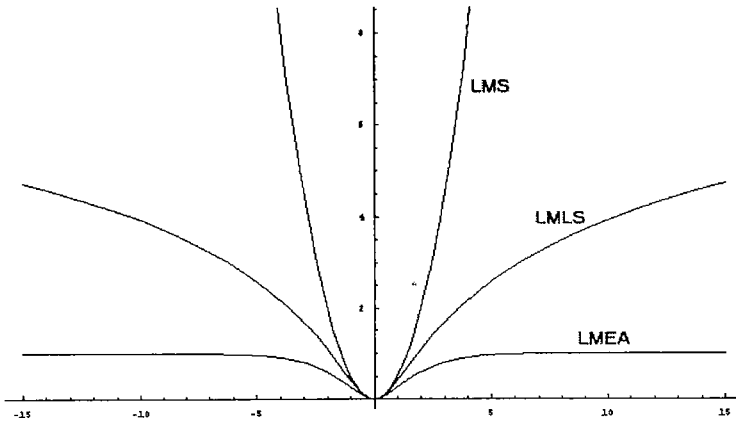


그림 2.2: (그림 2.1)에 대응하는 오차함수들

### 2.1. LMS 방법

LMS(least mean squares)방법은 전통적인 역전파 알고리즘에서 채용하고 있는 것으로 오차함수와 영향함수는 각각 다음과 같이 정의된다.

$$\rho(r) = \frac{1}{2}r^2 \quad (2.3)$$

$$\psi(r) = r \quad (2.4)$$

이 영향함수를 보면 모형으로부터 멀리 떨어진 점일수록 가중값에 더욱 큰 영향을 미치는 것을 알 수 있다. 이것은 최소제곱법의 기본 출발점이기도 하다. 그러나 이상점이 있을 때 이 방법은 심각한 문제를 일으킬 수 있다.

LMS의 장점을 살리면서 동시에 이상점의 문제를 해결하려면 이상점일 가능성이 적은 점들에 대해서는 모형으로부터 멀어질수록 영향이 커지도록하고 이상점일 가능성이 많은 점들에 대해서는 반대로 영향이 작아지도록 하면 된다. 따라서 영향함수가 그림 2.1에 표시된 두 곡선(LMLS와 LMEA)과 유사한 형태가 되는 것이 바람직하다. 이런 영향함수들에 대응하는 오차함수를 그림으로 나타내면 그림 2.2와 같다.

### 2.2. LMLS 방법

Liano가 제안한 LMLS (least mean log squares)방법은 다음과 같은 오차함수와 영향함수를 갖는다.

$$\rho(r) = \log \left( 1 + \frac{1}{2}r^2 \right) \quad (2.5)$$

$$\psi(r) = \frac{r}{1 + \frac{1}{2}r^2} \quad (2.6)$$

이 오차함수는 오차의 분포가 코시분포일 때 이 함수를 사용하는 M-추정량이 최대우도추정량이 된다는 특징을 갖고 있으며, 영향함수의 모양이 앞의 그림에서 보듯이 M-추정량의 바람직한 모양을 갖고 있다.

### 2.3. LMEA 방법

본 연구에서 새로 제안하는 방법은 편의상 LMEA(least mean exponential absolutes)로 이름을 붙였는데, 다음과 같은 오차함수와 영향함수를 사용한다.

$$\rho(r) = 1 - e^{-|r|}(|r| + 1) \tag{2.7}$$

$$\psi(r) = r e^{-|r|} \tag{2.8}$$

이 영향함수의 그림은 Liano의 것과 유사하지만 훨씬 일찍 0으로 수렴하는 성질을 갖고 있어 이상점에 큰 효과를 나타낼 수 있다. 그러나 만약 영향관측값이 있다면, 일반적인 로버스트 통계방법과 마찬가지로, 이 값에 대해서도 오차함수의 반응이 둔하게 되어 적합도가 낮아질 수 있다.

한편 LMEA의 영향함수는 오차가 0에 가까울 때는 LMLS의 영향함수의 1/2정도이고 오차가 0에서 멀어짐에 따라 그 차이가 더욱 커지기 때문에 척도(scale)를 조정할 경우 LMLS에 비해 더욱 로버스트해질 수 있다.

## 3. 모의 실험

앞에서 소개한 LMS, LMLS, LMEA 등은 모두 서로 다른 오차측도들을 최소화한다는 점에서만 다르다. 우리는 Liano의 방법에 따라 모의 실험을 함으로써 새로 제안하는 방법이 얼마나 로버스트한 지를 확인하려고 한다.

### 3.1. 실험방법

LMS, LMLS, LMEA 등의 효율성을 비교하기 위해 네트워크의 모든 요소들은 동일하게 유지한 채(네트워크 구성, 초기가중값, 학습파라미터, 학습알고리즘 등) 오직 오차측도만을 다르게 하여 실험하였다. 이 방법은 Liano가 자신의 방법이 로버스트함을 보이기 위해 사용한 실험방법과 완전히 동일하다.

먼저 네트워크는 하나의 입력노드, 10개의 은닉노드, 그리고 하나의 출력노드를 갖고 있으며, 각 조합에 대해서 수렴할 때까지의 반복회수(epoch)는 각각 400번으로 하였다. 또 은닉노드의 전이함수로 시그모이드함수를 사용했고, 출력노드의 전이함수는 단위함수(identity function)이다. 실험을 위해 준비된 데이터는 간단한 함수를 적합시키는 문제로 사용한 함수는 다음과 같다.

$$y = x^{\frac{2}{3}} \tag{3.1}$$

또 모의실험은 3가지 형태의 서로 다른 표본에 대해 수행되었다. 1) 정확하게 (3.1)식의 관계를 만족하는(오차가 없는) 표본, 2) 정규분포를 따르는 약간의 오차를 수반하는 표본,

그리고 3) 데이터의 5%는 대형오차를 갖고 나머지는 정규분포를 따르는 약간의 오차를 수반하는 표본이 그것이다.

독립변수  $x$ 에 대해 구간  $[-2, 2]$ 를 0.008간격으로 표본을 추출하여 (3.1)식의 계산을 통해 종속변수  $y$ 를 결정하는 방법으로 총 500개의 훈련 패턴이 만들어졌다. 이 패턴들이 1번 표본으로 사용되었다. 2번 표본은 여기에  $N(0.0, 0.1)$ 로부터 얻은 난수를 오차항으로 더해 줌으로써 만들어졌다. 마지막으로 이상점이 있는 3번 표본은 다음과 같이 생성시킨 오차항을 더해서 만들어지는데,

$$F = (1 - \epsilon)G + \epsilon H \quad (3.2)$$

여기서  $\epsilon = 0.05$ ,  $G \sim N(0.0, 0.1)$ ,  $H \sim N(10.0, 0.1)$ 이다.

다만 우리는 학습파라미터들(모멘텀, 학습률)을 특정한 하나의 값에서 실험하면 그 경우에 국한된 결과를 일반화하는 오류를 범할 수 있으므로 Liano와 달리 모멘텀 값은 0.9에서 0.6까지 0.1간격으로, 학습률은 0.00005에서 0.006까지 0.00005간격으로 변화시켜가며 실험하였다. 따라서 각 모형별로 실험조합의 수는 4(모멘텀) x 120(학습률) x 3(데이터형태) = 1440 가지이다.

이렇게 Liano의 방법을 그대로 답습하므로써 우리가 새로 제안하고 있는 방법을 Liano의 방법과 비교하여 더 바람직한 성질을 갖는다는 것을 보이려고 하는 것이다.

### 3.2. 실험 결과

각 표본에 대한 네트워크의 성능을 비교하기 위해 MSE(mean squared error) 값을 측정하여 그림으로 나타내었다. MSE는 다음과 같이 정의된다.

$$MSE = \frac{\sum_{i=1}^N (t_i - o_i)^2}{N} \quad (3.3)$$

여기서 목표값  $t_i$ 는  $x_i$ 에서 함수(3.1)의 실제값이고,  $o_i$ 는  $x_i$ 가 입력값으로 주어졌을 때의 출력값이다.

그림 3.1은 정확한 데이터를 사용한 표본1의 경우이다. 모멘텀과 학습률에 따라 LMS에 의한 오차는 부분적으로 매우 정확한 근사가 이루어지기도 하지만 전반적으로 0.172정도의 MSE값을 보여주고 있으며, MSE값이 주기변동에 가까운 변동을 한다는 점 외에는 전반적으로 안정적인 모습을 보여주고 있다. 그러나 LMLS에 의한 결과는 폭넓은 구간에서 LMS보다 더 좋은 적합도를 보여주면서도 최대 MSE가 5.0을 초과하는 등(그림에서는 상한을 0.5로 절단하였음) 매우 불안정한 모습을 보여주고 있다. 또 LMEA 방법은 학습률이 작을 때 부분적으로 LMS보다 나쁜 결과를 보여주지만 대부분의 경우에 LMS보다 좋은 적합도를 나타내고 있다. 특히 LMLS와 비교했을 때 일부 구간에서 다소 적합도가 떨어지고 있지만 더욱 안정적인 적합도를 보여주고 있다.

그림 3.2에 나타난 표본2의 경우도 표본1의 경우와 크게 다르지 않으며 자료에 오차가 개입됨에 따라 LMEA방법의 안정성이 더욱 두드러졌다.

대형오차를 갖는 표본3의 경우에는 이런 대조가 더욱 강하게 드러난다. LMS방법으로는 근사시키는데 한계가 있음을 보여주고 있으며, 대형오차에 로버스트하게 만들어진 두

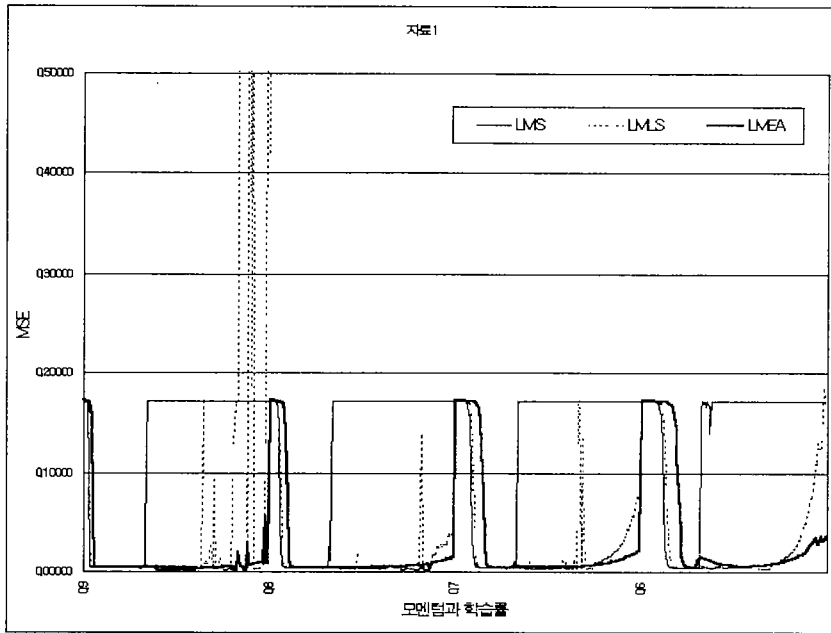


그림 3.1: 표본1에 대한 학습결과 비교

방법의 결과는 LMS방법에 비해 적합도가 크게 개선되었다. 그러나 여전히 LMLS방법은 학습 파라미터의 값에 따라 매우 불안정해지고 있다. 이 경우에도 LMEA방법은 매우 안정적이며, 세가지 형태의 표본 모두에 대해서 MSE값의 상한선이 일정하게 나타났다. 특히 LMEA방법은 학습파라미터의 변화에 영향을 적게 받아 실제 분석을 시도할 때 부딪치는 학습파라미터 선정의 어려움을 크게 완화시킬 수 있는 방법으로 나타났다.

따라서 이 방법은 LMLS의 로버스트한 특징과 LMS의 안정적인 성질을 골고루 갖추고 있는 방법으로 판단된다. 이런 결과는 2.3절에서 언급한 척도를 개선시키지 않고 수행한 실험결과로 척도를 개선할 경우에 더욱 나은 결과를 얻을 수 있을 것이다.

#### 4. 사례분석

지금까지는 인위적으로 만든 실험자료를 가지고 LMS, LMLS, LMEA 방법들을 비교하였다. 그 결과 LMEA가 다른 방법보다 로버스트한 것으로 나타났다. 이제 실제자료를 통하여 위의 세가지 방법을 비교하고자 한다. 여기서 사용한 자료는 Li 등이 사용했던 자료이다. 그들은 종속변수  $y$ 로 용광로의 라이닝 내구력(furnace lining durability)과 독립변수로  $x_1$ (첨가되는 라이닝 시료의 양; quantity of adding lining material),  $x_2$ (블로잉 시간; blowing time),  $x_3$ (용광로의 1회 가열처리 기간; duration of heating),  $x_4$ (망간의 포함량; content of Mn),  $x_5$ (용재의 철 포함양; content of iron in slag),  $x_6$ (처리 속도; operation rate)의 관계를 분석하였다. 이와 같은 관계를 모형화하면 용광로의 라이닝 내구력을 연장할 수 있는 기술적 조건을 조정하는데 도움이 된다고 한다.

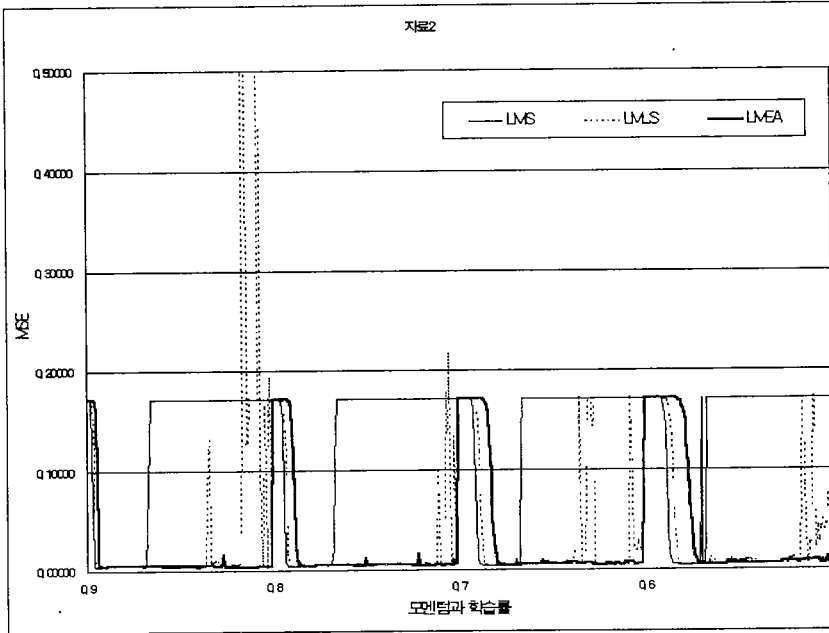


그림 3.2: 표본2에 대한 학습결과 비교

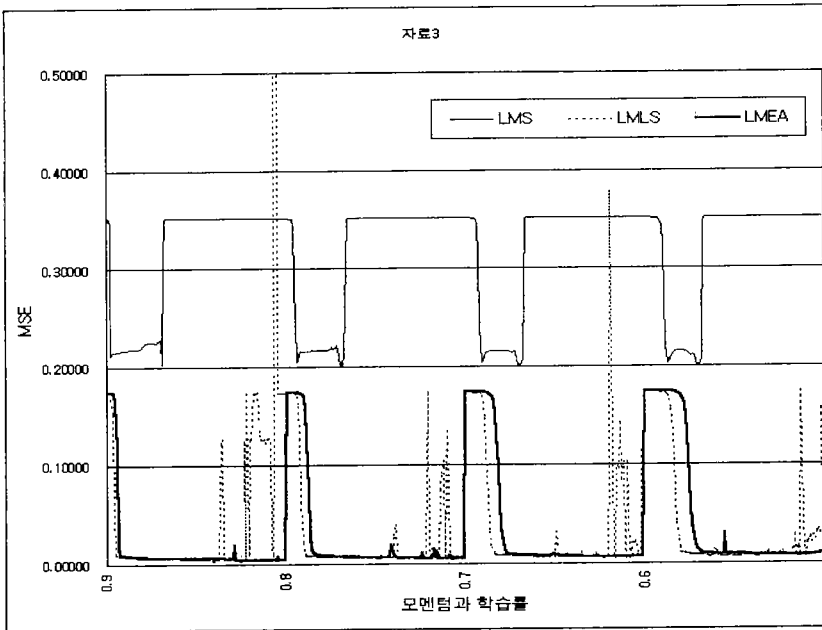


그림 3.3: 표본3에 대한 학습결과 비교



표 4.1: 각 방법에 대한 MSE 비교

방법	LMS	LMLS	LMEA
MSE 평균	1.07267	0.64739	0.52249
MSE 표준편차	0.68654	0.59097	0.26784

#### 4.1. 분석방법

분석방법은 Li 등의 분석방법과 동일하다. 그들은 자료의 단위 차이를 제거하기 위하여 정규화 변환을 하였다. 네트워크는 입력노드가 6개로 각 입력노드에는 독립변수  $x_1 - x_6$ 가 각각 입력되고, 은닉노드는 5개이며 은닉노드의 전이함수는 시그모이드 함수를 이용하였다. 출력노드는 1개로 종속변수  $y$ 가 대응되며 전이함수는 단위함수를 이용하여 네트워크가 수렴하기까지 1000번 반복하였다. 그들은 전체자료 수 33개의 관찰값중에서 랜덤으로 4개(일련번호중 1, 7, 17, 33번)를 선택하여 29개의 훈련표본과 4개의 예측표본으로 분리하여 분석하였다. 우리는 3장의 모의실험과 같이 특정한 학습파라미터의 영향을 제거하기 위하여 모멘텀값은 0.6에서 0.9까지 0.1 간격으로, 학습률은 0.001에서 0.05까지 0.001간격으로 변화시켜가며 200 가지(=4(모멘텀)x50(학습률))로 분석했다.

#### 4.2. 분석결과

각 방법에 평가는 식 (3.3)의 MSE로 한다. 분석결과 각 방법에 대한 MSE는 표 4.1과 같다. 표 4.1에서 평균은 200가지 모형에서 얻은 분석결과의 MSE에 대한 평균값으로 LMEA 방법의 MSE는 0.522로 나타나 LMLS 방법보다 24% 정도 우수하게 나타났다. 이것으로 LMEA 방법이 세 가지 방법 중에 가장 로버스트한 것을 알 수 있다. 또 LMEA방법에 의한 MSE의 표준편차가 다른 방법에 비해 매우 작아 안정성이 높은 것으로 확인되었다. LMS 방법은 MSE가 거의 모든 모형에서 가장 크게 나타나 이상점에 대한 처리능력이 가장 떨어지는 것을 볼 수 있으며 MSE의 변동이 주기변동과도 비슷하게 나타나는 것으로 보아 모멘텀과 학습률에도 큰 영향을 받고 있는 것을 알 수 있다. 이상점의 영향을 제거하고자 하는 LMLS방법과 LMEA의 방법을 비교하면 부분적으로 LMLS가 LMEA보다 MSE가 작은 경우도 있지만 전체적으로 LMEA가 LMLS보다 우수한 것으로 나타나고 있으며 LMEA는 모멘텀과 학습률의 영향을 LMLS, LMS와 비교하여 상대적으로 적게 받는 것으로 나타났다. 따라서 모의실험의 결과와 동일하게 LMEA는 로버스트한 특징과 안정적인 성질을 모두 갖추고 있음을 보여주고 있다.

### 5. 결론

인공신경망으로 함수를 적합시키고자 할 때, LMS방법을 사용하는 전통적인 역전파 알고리즘은 로버스트하지 않다. 이 사실은 훈련표본에 이상점이 있을 때 큰 문제를 일으킬 수

있다. 따라서 많은 저자들이 이런 문제를 해결하기 위해 시도하여 왔다. 우리는 로버스트한 오차함수를 사용함으로써 이 문제를 해결하려 한 일련의 연구를 쫓아서 새로운 오차함수를 제안했다. 그리고 모의실험을 통해 새로 제안한 LMEA방법이 LMS방법만큼 안정적일 뿐 아니라 LMLS방법보다 더욱 로버스트하다는 사실을 확인했다. 특히 정확한 값을 갖는 표본이나 정규분포를 따르는 작은 오차를 수반하는 표본에 대해서도 좋은 결과를 보여줌에 따라 대형오차의 유무에 관계없이 사용할 수 있다는 장점이 발견되었다. 다만 여기서는 영향관측값은 없는 것으로 간주하고 있으므로 만약 영향관측값들이 있을 때의 적합도를 확인하기 위해서는 다른 실험이 필요하다.

### 참고문헌

- [1] 송문섭 (1996). <로버스트통계>, 자유아카데미.
- [2] Beckman, R. J. and Cook, R. D. (1983). Outlier.....s, *Technometrics*, Vol. 25, No. 2, 119-149.
- [3] Chen, S. David and Jain, C. Ramesh (1994). A Robust Back Propagation Learning Algorithm for Function Approximation, *IEEE Transactions on Neural Networks*, Vol. 5, No. 3, 467-479.
- [4] Freeman, J. A. (1994). *Simulating Neural Networks with Mathematica*, Addison Wesley.
- [5] Li, Z., Cheng, Z., Xu, L. and Li, T. (1993). Nonlinear Fitting by Using a Neural Net Algorithm, *Analytical Chemistry*, Vol. 65, 393-396.
- [6] Liano, K. (1996). Robust Error Measure for Supervised Neural Network Learning with Outliers, *IEEE Transactions on Neural Networks*, Vol. 7, No. 1, 246-250.
- [7] Smith, M. (1993). *Neural Networks for Statistical Modeling*, Van Nostrand Reinhold.
- [8] Wang, J. H., Jiang, J. H. and Yu, R. Q. (1996). Robust back propogation algorithm as a chemometric tool to prevent the overfitting to outliers, *Chemometrics and Intelligent Laboratory Systems*, Vol. 34, 109-115.

[ 1998년 10월 접수, 1999년 4월 최종수정 ]

# Robust Error Measure for Back Propagation Algorithm

Hyun Chul Kim<sup>1)</sup> Chul Won Lee<sup>2)</sup>

## ABSTRACT

The conventional back propagation algorithm is non-robust. Many authors suggest some robust back propagation algorithms include M-estimator based methods. Here we suggest a new error function as a robust error measure for back propagation algorithm. It is shown that the new method has good properties by simulation. It can be used as a general error measure for back propagation algorithm whenever there are outliers or not in the sample data.

---

1) Assistant Professor, Department of Informatics and Statistics, Kunsan Nat'l Univ., Kunsan (573-701), Korea  
2) Assistant Professor, Department of Computer Science, Kyungmin College, Uijungbu (480-702), Korea