

# 함수근사를 위한 로버스트 역전파 알고리즘

김 상 민<sup>†</sup> · 황 창 하<sup>††</sup> · 오 광 식<sup>††</sup>

## 요 약

함수근사는 과학과 공학분야에서 광범위하게 응용된다. 다층 전방향 신경망은 비선형 함수근사를 위한 도구로서 제안되어져 왔으며, 다층 신경망을 학습시키기 위한 학습 알고리즘으로 역전파 알고리즘이 널리 이용되어져 왔다. 그러나 이상치(outlier)를 포함한 학습자료가 존재할 때에는 학습되는 함수는 존재하는 모든 자료 사이를 보간 하므로 이상치가 있는 자료의 위치까지도 보간 하여, 원하지 않는 구조를 파악하게 된다. 따라서 이상치의 영향을 최소화 시키기 위해, 본 논문에서는 로버스트 에너지함수를 유도하여 개량된 로버스트 역전파 알고리즘을 제안한다.

## A Robust Back Propagation Algorithm for Function Approximation

Sangmin Kim<sup>†</sup> · Changha Hwang<sup>††</sup> · Kwangsik Oh<sup>††</sup>

### ABSTRACT

Function approximation from a set of input-output pairs has numerous applications in scientific and engineering areas. Multilayer feedforward neural networks have been proposed as a good approximator of nonlinear function. The back propagation (BP) algorithm allows multilayer feedforward neural networks to learn input-output mappings from training samples. However, the mapping acquired through the BP algorithm may be corrupt when erroneous training data are employed. In this paper we propose a robust BP learning algorithm that is resistant to the erroneous data and is capable of rejecting gross errors during the approximation process.

### 1. 서 론

함수근사는 과학과 공학분야에서 광범위하게 응용된다. 다층 전방향 신경망은 비선형 함수근사를 위한 도구로서 제안되어져 왔다[6, 7, 10]. 이들 신경망에 의해 표현된 모수적 모형은 비선형 모형을 표현한다. 특히 3개의 층으로 구성된 신경망을 이용한 함수근사는 투사지향방법을 이용하는 근사방법과 매우 유사하다.

Friedman & Stuetzle[3]은 함수근사의 일종인 회귀분석에 투사지향 방법을 처음으로 시도하였다. Diaconis & Shahshahani[2]는 어떤 연속함수라도 투사지향 근사 방법으로 근사시킬 수 있음을 보였다. 또한 Hornik 등 [6]은 하나의 은닉층으로 된 신경망을 이용하여 은닉층 내에서 노드수가 충분히 클 때 어떤 가측 Borel함수도 실제로 근사할 수 있음을 보였다. 이와 같이 다층 신경망이 전반적인 근사도구로 이용되어져 왔다.

다층 신경망을 학습시키기 위한 학습 알고리즘으로 역전파 알고리즘이 널리 이용되어져 왔다[9]. 역전파 학습 알고리즘은 최급강하법을 이용하여 오차제곱의 합이 최소화 되도록 신경망의 모수 즉, 가중치

<sup>†</sup> 정 회 원: 대구효성가톨릭대학교 대학원 전산통계학과

<sup>††</sup> 정 회 원: 대구효성가톨릭대학교 정보통계학과 교수

논문접수: 1996년 8월 12일, 심사완료: 1997년 1월 11일

를 반복적으로 변화시켜 나가는 알고리즘이다. 역전파 알고리즘으로 학습되는 신경망은 고도의 비선형 구조를 파악할 수 있다. 즉, 학습되는 함수는 모든 학습점 사이를 보간하여 그 구조를 파악한다. 이때 실제상황에서는 많은 경우에 학습자료에 이상치가 포함되게 된다. 그러므로 학습되는 함수는 모든 자료 사이를 보간하므로 이상치가 있는 자료의 위치까지도 보간 하여, 원하지 않는 구조를 파악하게 된다. 따라서 이상치에 민감하지 않아서, 그 영향을 최소화시키는 로버스트 역전파 알고리즘의 필요성이 대두되었다. 이상치의 영향을 최소화시키기 위해 Chen[1]은 Hampel의 hyperbolic tangent 추정량을 이용한 로버스트 역전파 알고리즘을 제안한 바 있다. 본 논문에서는 Chen의 알고리즘과는 다른 로버스트 에너지 함수를 이용하여 새로운 로버스트 역전파 알고리즘을 제안한다. 제안된 로버스트 역전파 알고리즘을 이상치가 포함된 자료에 적용하여 이상치에 어느 정도로 로버스트한지를 분석, 연구하고자 한다.

**2. 역전파 알고리즘을 이용한 함수 근사와 문제점**

신호처리, 컴퓨터비전, 패턴인식등과 같은 대다수의 공학문제는 입력과 출력의 쌍  $T = \{(X_p, O_p); p = 1, \dots, P\}$ 으로 된 학습자료들로부터 미지의 함수를 근사시키는 문제들을 다루어 왔다. 이때 입력벡터  $X_p = (x_{p1}, \dots, x_{pm})^T$ 와 출력벡터  $O_p = (o_{p1}, \dots, o_{pm})^T$ 는 미지의 함수  $f$ 에 대해  $O_p = f(X_p) + e_p$ 이라고 가정한다. 여기서  $e_p$ 는 오차를 나타내는 확률벡터이다. 함수근사의 목표는 근사오차를 최소화시키는  $f$ 의 추정치  $\hat{f}$ 을 구하는 일이다. 일반적으로 함수근사를 위해서는 한개 이상의 입력노드와 한개의 출력노드를 갖는 신경망이 이용된다. 한편,  $n$ 개의 입력노드와 한개의 출력노드 그리고  $N$ 개의 노드를 가지는 한개의 은닉층으로 구성된 신경망에 대한 보수적 표현은

$$\hat{f}(x_1, x_2, \dots, x_n; W) = \sum_{j=1}^N \beta_j g_j(\sum_{i=1}^n \alpha_{ji} x_i) \quad (1)$$

과 같다. 여기에서  $\beta_j, 1 \leq j \leq N$ 는  $N$ 개의 은닉노드에서 출력노드로 연결된 가중치를 나타내고,  $\alpha_{ji}, 1 \leq i \leq n$ 는  $i$ 번째 입력노드에서  $j$ 번째 은닉층 노드로 연결

된 가중치를 나타낸다. 그리고  $g_j(\cdot)$ 는 은닉층 활성화 함수이다. 출력노드에 대한 활성화함수로는 항등함수를 사용한다. 식(1)에서 사용한 표현은 투사지향회귀와, 분류에서 사용되어진 투사지향 근사와 매우 유사하다[3, 4]. 만일  $g_j(\cdot)$ 함수가 유계함수로 국한되지 않는다면 Diaconis & Shahshahani[2]는 어떠한 다항함수라도 식(1)로 표현될 수 있음을 보였다.

비록 일반적인 역전파 학습 알고리즘이 실제 문제를 해결하는데 많이 이용되기는 하지만, 이 알고리즘의 성공 여부는 입력되는 학습자료에 좌우된다. 즉, 일반적인 역전파 학습 알고리즘은 학습자료에 내재되어 있는 전체적인 오류에 의해 불안정하게 되는 단점을 지니고 있다. 따라서 전체적인 오류에 민감하지 않는, 특히 이상치의 영향을 최소화시키는 알고리즘이 필요하다.

**3. 새로운 로버스트 역전파 알고리즘의 제안**

역전파 알고리즘은 다음과 같은 오차제곱함수를 최소화 시키도록 신경망의 가중치를 반복적으로 조정해 나가는 것이다.

$$E_{LS}(W, T) = \frac{1}{2} \sum_{p=1}^P \sum_{j=1}^m (t_{pj} - o_{pj})^2. \quad (2)$$

이때  $W$ 는 가중치 집합을,  $T$ 는 시간이 지남에 따른 반복회수를,  $t_{pj}$ 는  $p$ -번째 학습패턴에 대해  $j$ -번째 목표치를 나타내고  $o_{pj}$ 는  $p$ -번째 학습패턴에 대한  $j$ -번째 추정된 출력치를 나타낸다. 이때 학습목표는 신경망이 모든 학습패턴을 보간함으로써 추정된  $\hat{f}$ 이 계산되었을 때  $E_{LS}(W, T)$ 가 최소값을 가지도록 하는 것이다[5, 8, 11]. 따라서 이러한 식(2)를 다시 통계물리학에서 자주 이용되는 일반화된 에너지함수 형태로 표현하면

$$E(V, W) = \sum_{p=1}^P \sum_{j=1}^m V_p z(t_{pj}, o_{pj}) + E_{prior}(V) \quad (3)$$

과 같다[10]. 여기서  $z(t_{pj}, o_{pj}) = \frac{1}{2} (t_{pj} - o_{pj})^2$ 이고  $V_p$

는 0 또는 1의 값을 가지는 확률변수이며  $V = \{V_p, p = 1, \dots, P\}$ 이다. 이 값은 입력자료가 이상치인지 표본 인지를 결정하는데 사용된다. 즉  $V_p = 1$ 이면 들어오는

입력자료를 표본으로 간주하고, 그렇지 않으면 이상치로 간주한다. 한편  $E_{prior}(\mathbf{V})$ 는  $\{V_p\}$ 의 사전확률분포에 의해 공헌되어진 에너지의 양을 나타내며, 이것의 자연스러운 선택과  $E_{prior}(\mathbf{V})$ 의 해석을 다음과 같이 생각할 수 있다. 즉,  $E_{prior}(\mathbf{V}) = \eta \sum_{p=1}^P (1 - V_p)$ 일때,  $z(t_{pj}, o_{pj}) < \eta$ 이면,  $V_p = 1$ 이 되어 주어진 학습패턴이 표본으로 간주되고, 그렇지 않으면  $V_p = 0$ 이 되어 이상치로 간주된다.

한편  $\eta$ 는 지정된 반복횟수에 도달할 때 마다 계산되는 우측경계값으로 정의한다. 즉, 지정된 반복횟수에 도달할 때 마다 학습패턴들에 대응되는  $\sum_{j=1}^m z(t_{pj}, o_{pj})$  값들을 구한 후에 정렬하여 삼사분위수( $Q_3$ :75%)와 일사분위수( $Q_1$ :25%)를 계산한다. 그리고 삼사분위수에서 일사분위수를 공제한 값인 사분위범위수( $IQR$ : $Q_3 - Q_1$ )를 계산하여 우측경계값,  $Q_3 + 1.5 * IQR$ 을  $\eta$  값으로 취한다.

우리들의 목표는  $V_p$ 가 이진 값을 가진다는 제약조건 하에서  $\{V_p\}$ 와  $\mathbf{W}$ 에 관해  $E(\mathbf{V}, \mathbf{W})$ 를 최소화시키는 것이다. 그런데 이 문제는 연속형변수와 이산형변수가 혼합된 경우의 최적화 문제이기 때문에 해석적인 해를 구할 수 없을 뿐 아니라, 최급강하법을 사용하여 해를 구하는 것도 쉽지는 않다. 따라서 이런 문제점을 해결하기 위해 Gibbs분포를 사용하며, 그 분포는  $P[\mathbf{V}, \mathbf{W}] = 1/Z e^{-\beta E(\mathbf{V}, \mathbf{W})}$ 로 정의된다.

이때,  $Z$ 는 관계식  $\sum_{\mathbf{V}} \int_{\mathbf{W}} P[\mathbf{V}, \mathbf{W}] = 1$ 을 만족한다.

따라서  $E(\mathbf{V}, \mathbf{W})$ 를 최소화시키는 문제는  $P[\mathbf{V}, \mathbf{W}]$ 를 최대화 시키는 문제로 귀착된다. 그러나 이것 또한 연속형변수와 이산형 변수가 혼합된 경우의 최적화 문제이기 때문에 어려움이 따른다. 따라서 이런 문제에 대한 하나의 해결책으로는  $\mathbf{W}$ 의 주변분포를 구하여  $P[\mathbf{V}, \mathbf{W}]$ 를 최대화 시키는 것이다. 이때  $\mathbf{W}$ 의 주변분포는

$$\begin{aligned}
 P_{margin}(\mathbf{W}) &= \frac{1}{Z} \sum_{\mathbf{V}} \exp(-\beta \sum_p \{ \sum_j^m V_p \frac{1}{2} (t_{pj} - o_{pj})^2 \\
 &\quad + \eta(1 - V_p) \}) \\
 &= \frac{1}{Z} \prod_p \sum_{V_p \in \{0,1\}} \exp(-\beta \{ \sum_j^m V_p \frac{1}{2} (t_{pj} - o_{pj})^2 \\
 &\quad + \eta(1 - V_p) \})
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{Z} \prod_p \{ \exp(-\beta \eta) + \exp(-\beta \{ \sum_j^m \frac{1}{2} (t_{pj} - o_{pj})^2 \}) \} \\
 &= \exp(-P\beta\eta) / Z \prod_p \{ 1 + \exp(-\beta \{ \sum_j^m \frac{1}{2} \\
 &\quad (t_{pj} - o_{pj})^2 - \eta \}) \}.
 \end{aligned}$$

이다. 이제,  $Z_m = Z e^{P\beta\eta}$ 와  $E_p(\mathbf{W}) = -1/\beta \sum_p \log \{ 1 + \exp(-\beta \{ \sum_i^m \frac{1}{2} (t_{pi} - o_{pi})^2 - \eta \}) \}$ 로 두면, 다음과 같은 주변 분포  $P_{margin}(\mathbf{W}) = 1/Z_m \exp(-\beta E_p(\mathbf{W}))$ 를 구하게 되고, 이때  $P_{margin}(\mathbf{W})$ 을 입력패턴에 대해 최대화 시킨다는 것은  $E_p(\mathbf{W})$ 를 최소화 시키는 문제로 귀착된다. 또한, 에너지함수  $z(\mathbf{t}, \mathbf{o})$ 의 값이 작다면  $E_p$ 의 값은 단지  $z(\mathbf{t}, \mathbf{o})$ 가 지닌 값에 불과하고, 반면  $z(\mathbf{t}, \mathbf{o}) \rightarrow \infty$ 이면  $E_p$ 의 값은 일정한 상수로 되며 이때  $z(\mathbf{t}, \mathbf{o})$  값을 크게 하는 이상치를 표본과 다르게 취급할 수 있다.

이제 로버스트 에너지함수  $E_p(\mathbf{W})$ 를 사용하는 로버스트 역전파 알고리즘의 유도과정을 은닉층이 하나인 신경망에 대해서 설명한다. 비선형 활성화함수를 사용하여 다음 층으로 변화량을 전달하는 방법은 다음과 같다. 은닉층의  $j$  번째 노드에 입력되는 값은  $a_{pj} = \sum_i w_{ji} x_{pi}$ 와 같이 주어지고, 이때  $x_{pi}$ 는 입력층에서 은닉층으로의 가중치이고  $x_{pi}$ 는 입력층의  $i$  번째 노드에서  $p$  번째 관측치의 출력이다. 그러면 은닉층의  $j$  번째 노드의 출력값은  $h_{pj} = g_j(a_{pj})$ 로 나타내며, 이때  $g_j$ 는 은닉층의 시그모이드 비선형 활성화 함수이다. 또한 출력층의  $k$  번째 노드에 입력되는 값은  $b_{pk} = \sum_j v_{kj} h_{pj}$ 와 같이 주어지고, 이때  $v_{kj}$ 는 은닉층에서 출력층으로의 가중치이다. 그러면 출력층의  $k$  번째 노드의 출력값은  $o_{pk} = g_k(b_{pk})$ 로 나타내며, 이때  $g_k$ 는 항동함수로서 출력층의 활성화 함수를 나타낸다.

가중치의 변화량은 오차 제곱이 가장 많이 감소하는 방향으로 변화한다. 즉, 수식으로  $\Delta_p w_{ji} \propto \frac{-\partial E_p}{\partial w_{ji}}$  이다.

합성함수 미분공식과 역전파 알고리즘 유도하는 일반적인 방법을 사용하여 로버스트 역전파 알고리즘을 다음과 같이 두 가지 경우에 대하여 구한다.

(1) 출력층 노드의 경우에는

$$\Delta v_{kj} = \alpha \sum_p \frac{1}{1 + \exp(\beta \{ \sum_i^n 1/2(t_{pi} - o_{pi})^2 - \eta \})} \times \{t_{pk} - o_{pk}\} \times h_{pj}$$

과 같다. 그리고 오프셋의 수정은  $\Delta o_k = \beta \delta_{pk}$ 과 같다. 이때  $\alpha$ 와  $\beta$ 는 학습률이다.

(2) 중간층 노드의 경우

$$\Delta w_{ji} = \alpha \sum_k \delta_{pk} v_{kj} x_{pi}$$

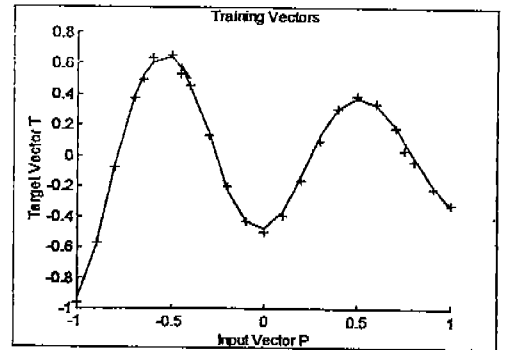
과 같다. 그리고 오프셋의 수정은  $\Delta \theta_j = \beta \delta_{pj}$ 과 같다. 이때  $\alpha$ 와  $\beta$ 는 학습률이다.

4. 실험결과

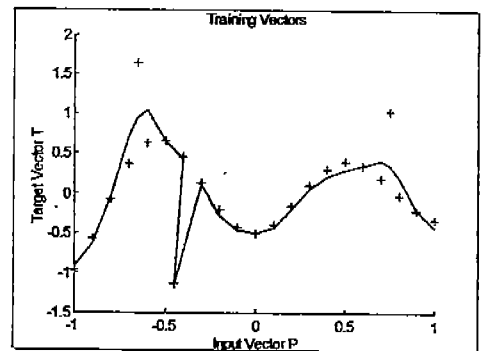
제안된 로버스트 역전파 알고리즘의 성능을 분석하기 위해 비선형 일변량 함수의 근사문제를 생각한다. 근사를 위한 함수로는 2개의 예를 들어 설명하였다. 이상적으로는 각각의 예제에 대해 각기 다른 최적의 신경망구조를 채택해야 하지만, 신경망 모형의 선택문제는 본 논문에서는 배제한다. 신경망의 층의 수와 노드가 많으면 과다한 근사를 하게 되며, 반대로 적으면 단순화되어 제대로 근사를 할 수 없게 된다. 따라서 본 논문에서는 신경망 구조에 대한 여러 실험 중, Chen[1]이 사용한 3개의 은닉층으로 된 신경망 구조를 이용, 차례로 첫번째 은닉층 노드수 21개, 두번째 은닉층 노드수 14개, 세번째 은닉층 노드수 7개로 실험을 하였다. 활성화함수로는 시그모이드함수를 사용하였으며, 출력층에서는 활성화함수를 사용하지 않았다. 한편 제안된 로버스트 역전파 알고리즘을 사용하여 실험한 결과, 적당한 층의 변화와, 노드 수의 변화는 결과에 큰 영향을 미치지 않는다는 사실을 경험적으로 알게 되었다.

첫번째 예제에서 사용한 함수는  $s(x) = 0.65 \sin(2.5\pi x)$ 를 조절하여 사용하였다. 입력으로는  $[-1, 1]$  사이에서 40개의 점을 임의로 추출하여 사용하였으며, 출력으로는  $[-1, 1]$  사이의 값이 나타나도록 하였다. (그림 1)에서는 일반적인 역전파 알고리즘과 로버스트 역전파 알고리즘의 결과가 같다는 것을 설명하고 있다. 그러나 (그림 2)에서 보는 바와 같이 이상치가 존

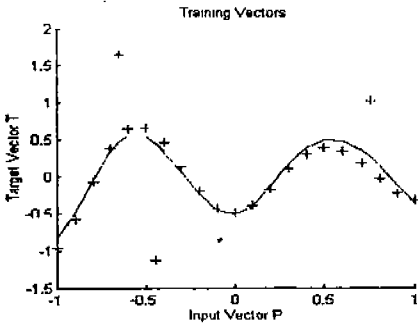
재할 때는 일반적인 역전파 알고리즘으로는 해를 구하기가 어렵지만, 이 논문에서 제안한 로버스트 역전파 알고리즘은 성공적으로 해를 구하고 있음을 (그림 3)에서 볼 수 있다. 이때 사용한 학습률은 다음과 같이 임의로 결정하여 사용하였다. 반복 회수가 600번, 종료하기 위한 오차함의 허용값을 0.001로, 학습률  $\alpha$ 의 값은 0.01로,  $\beta$ 의 값을 0.5로 사용하였으며, 반면  $\eta$ 의 값은 50번 반복할 때 마다 자동으로 결정된다. (그림 4)에서는  $\eta$ 의 값의 변화에 따른 결과를 나타내었다. 또한 (그림 3)에서의 동일한 신경망구조와 자료로 Chen[1]의 알고리즘으로 구한 결과를 (그림 5)에 나타내었다. 제안한 알고리즘과 차이가 없이 좋은 결과를 보이고 있다.



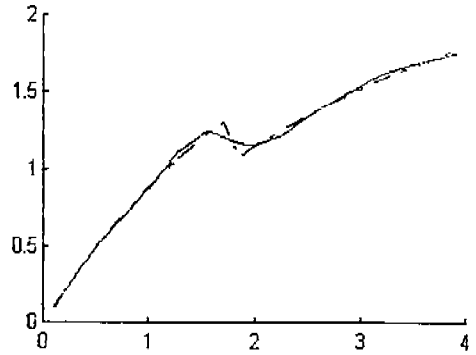
(그림 1) 이상치가 없는 BP와 RBP결과 (Fig. 1) BP and RBP of noiseless



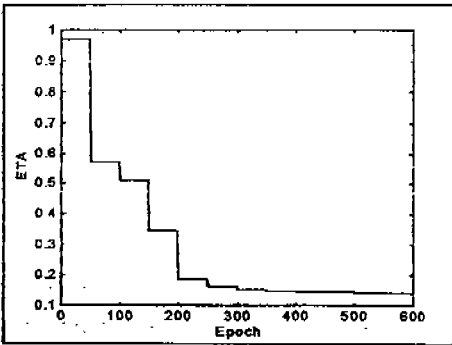
(그림 2) 이상치가 있을 때 BP결과 (Fig. 2) BP result of noisv



(그림 3) 이상치가 있을 때 RBP결과  
(Fig. 3) RBP result of noisy

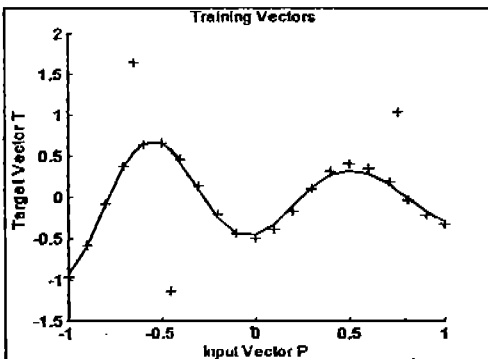


(그림 6) 이상치가 없을 때 BP와 RBP결과  
(Fig. 6) BP and RBP of noiseless

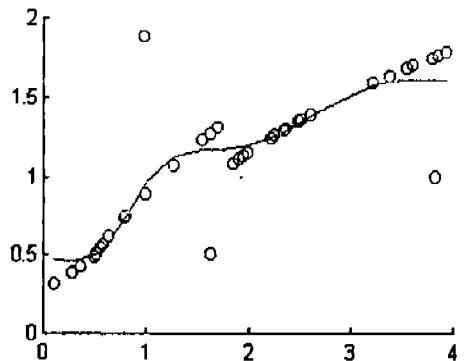


(그림 4)  $\eta$ 값의 변화  
(Fig. 4) Change of  $\eta$  value

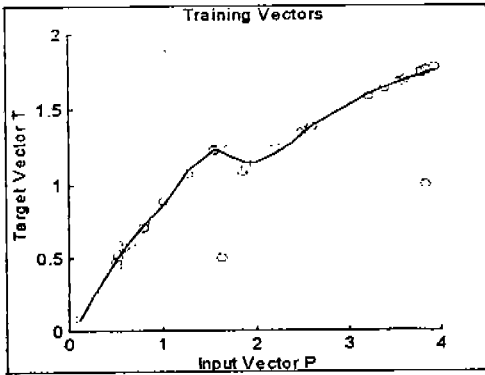
두 번째 예제에서는  $\tanh$ 함수 두개를 접합시켜 사용하였다. 우선  $[0, 4]$  사이에서 40개의 점을 임의로 추출하여, 첫번째 함수로는  $0.5\tanh(2x)$ 를 구간  $[0, 2]$ 에서 잘라내었고 나머지 함수로는  $2.5\tanh(2x)$ 를 구간  $(2, 4]$ 에서 잘라내어 사용하였다. (그림 6)에서는 이상치가 없을 때의 목표함수는 점선으로 표시하였고, 이상치가 없을 때 일반적인 역전파 알고리즘의 결과와 로버스트 역전파 알고리즘의 결과는 동일하다는 것을 실선으로 보였다. (그림 7)에서는 이상치가 3군데 포함되어 있는 자료들을 원으로 표시하였으며, 로버스트 역전파 알고리즘으로 실행한 결과를 실선으로 보이고 있다. 신경망은 하나의 은닉층 만을 사용하였고, 8개의 노드를 사용하였다. 이때 사용한 인수들은



(그림 5) Chen의 RBP결과  
(Fig. 5) RBP of Chen of noisy



(그림 7) 제한한 RBP결과  
(Fig. 7) RBP result of noisy



(그림 8) Chen의 RBP결과  
(Fig. 8) Result of Chen RBP

임의로 결정하여 사용하였으며, 반복회수가 6000번, 종료하기 위한 허용 오차함을 0.001로, 학습률  $\alpha$ 의 값은 0.01로,  $\beta$ 의 값을 3으로 하였으며, 반면  $\eta$ 의 값은 50번 반복할 때 마다 자동으로 결정된다. 또한 (그림 7)에서의 동일한 신경망 구조와 자료로 Chen[1]의 알고리즘으로 구한 결과를 (그림 8)에 나타내었다. 자료에 이상치가 포함되어 있을때 두 그림 모두 이상치의 영향을 최소화 하고 있다. 한편 제안된 로버스트 알고리즘은 586PC에서 MATLAB으로 구현하였다.

### 5. 토 의

다층 전방향 신경망은 비선형 함수근사를 위한 도구로서 제안되어져 왔으며, 다층 망을 학습시키기 위한 학습 알고리즘으로 역전파 알고리즘이 널리 이용되어져 왔다. 일반적인 역전파 알고리즘은 최급강화법을 이용하여 오차합이 최소화 되도록 가중치를 반복적으로 변화시켜 나가는 알고리즘이다. 그러나 학습자료에 이상치가 존재하는 경우에는 신경망에 의해 측정된 함수는 이상치 주변에서 진동하게 된다. 본 논문에서는 이상치에 다소 로버스트한 에너지 함수를 사용하여 로버스트 역전파 알고리즘을 유도하고 2개의 예제에 적용한 결과에 의하면 제안된 알고리즘이 일반적인 역전파 알고리즘 보다 이상치에 의한 영향을 적게 받게 된다. 앞으로 다양한 응용분야에 적용하여 그 성능을 확인할 필요가 있다.

### 참 고 문 헌

- [1] D. S. Chen, and R. C. Jain, "A Robust Back Propagation Learning Algorithm for Function Approximation," IEEE transactions on Neural Networks, 5, 467-479, 1994.
- [2] P. Diaconis, and M. Shahshahani, "On Linear Functions of Linear Combinations," SIAM J.SCI. STAT.COMPUT., 5, 175-191, 1984.
- [3] J. H. Friedman, and W. Stuetzle, "Projection Pursuit Regression," J. Amer. Statis. Assoc., 76, 817-823, 1981.
- [4] J. H. Friedman, "Classification and Multiple Regression through Projection Pursuit," Dep. Statist., Stanford Univ. and Stanford Linear Accelerator Center, Tech. Rep, LCS 12, 1985.
- [5] J. B. Hampshire II, and B. Pearlmutter, "Equivalence Proofs for Multi-Layer Perceptron Classifiers and the Bayesian Discriminant Function," Proc. the 1990 Connectionist Model Summer School, Morgan Kaufmann, 1990.
- [6] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks are Universal Approximators," Neural Networks, 2, 359-366, 1989.
- [7] T. Poggio, and F. Girosi, "Networks for Approximation and Learning," Proceedings of the IEEE, 78, 1481-1497, 1990.
- [8] M. D. Richard, and R. P. Lippmann, "Neural Network Classifiers Estimate Bayesian a Posteriori Probabilities," Neural Computation, 3, 461-483, 1991.
- [9] D. E., Rummelhart, G. E. Hinton, and R. J. Williams, "Parallel Distributed Processing: Explorations in the Microstructure of Cognition.," 1, Chapter 8, MTP Press, 1986.
- [10] L. Xu, and A. L. Yuille, "Robust Principal Component Analysis by Self-Organizing Rules Based on Statistical Physics Approach," IEEE transactions on Neural Networks, 6, 131-143, 1995.
- [11] H. White, "Learning in Artificial Neural Networks

: *A statistical perspective.*" Neural Computation, 1, 425-464, 1989.



**김 상 민**

- 1987년 동국대학교 전자계산학과(공학사)
- 1991년 동국대학교 대학원 전자계산학과(공학석사)
- 1994년~현재 대구효성가톨릭대학교 대학원 전산통계학과(박사과정 수료)

1992년~현재 김천전문대학 전산정보처리과 조교수  
관심분야: 신경망, 패턴인식



**오 광 식**

- 1978년 경북대학교 사범대학 수학교육과(이학사)
- 1981년 고려대학교 경제학석사
- 1989년 고려대학교 이학박사
- 1982년~현재 대구효성가톨릭대학교 정보통계학과 교수

관심분야: 회귀분석 및 진단, 통계그래픽스



**황 창 하**

- 1982년 경북대학교 사범대학 수학교육과(이학사)
- 1984년 서울대학교 대학원 계산통계학과(이학석사)
- 1985년~1987년 한국통신 전임연구원
- 1991년 미국 미시간대학교 통계학과(이학박사)

1992년~1994년 경성대학교 전산통계학과 조교수  
1995년~현재 대구효성가톨릭대학교 정보통계학과 조교수

관심분야: 다변량통계분석, 통계적패턴인식, 신경망 이론