

Random Tabu 탐색법을 이용한 신경회로망의 고속학습알고리즘에 관한 연구

Fast Learning Algorithms for Neural Network Using Tabu Search Method with Random Moves

양 보 석*, 신 광 재**, 최 원 호**
Bo-Suk Yang*, Kwang-Jae Shin**, Won-Ho Choi**

요 약

본 연구에서는 종래에 학습법으로 널리 이용되고 있는 역전파학습법의 문제점으로 지적되어온 학습에 많은 시간이 걸리는 점과 국소적 최적해에 해가 수렴하여 오차가 충분히 작게 되지 않는 등의 문제점을 해결하기 위해, Hu에 의해 고안된 random tabu 탐색법을 이용하여 신경회로망의 연결강도를 최적화하는 학습알고리즘을 새로이 제안하였다. 그리고 이 방법을 배타적 논리합 문제에 적용하여 기존의 역전파학습법과 학습상수 α, β 에 tabu 탐색법을 이용한 결과와 비교 검토하여 본 방법이 국소적 최적해에 수렴하지 않고 수렴정도를 개선할 수 있음을 확인하였다.

ABSTRACT

A neural network with one or more layers of hidden units can be trained using the well-known error back propagation algorithm. According to this algorithm, the synaptic weights of the network are updated during the training by propagating back the error between the expected output and the output provided by the network. However, the error back propagation algorithm is characterized by slow convergence and the time required for training and, in some situation, can be trapped in local minima.

A theoretical formulation of a new fast learning method based on tabu search method is presented in this paper. In contrast to the conventional back propagation algorithm which is based solely on the modification of connecting weights of the network by trial and error, the present method involves the calculation of the optimum weights of neural network. The effectiveness and versatility of the present method are verified by the XOR problem. The present method excels in accuracy compared to that of the conventional method of fixed values.

I. 서 론

계층형 신경회로망은 학습능력이나 비선형사상능력을 가지고 있고, 그 특징을 이용하여 패턴인식이나 동정 및

*부산수산대학교 공과대학 기계공학과

**부산수산대학교 대학원

제어 등에의 적용이 시도되어 성과를 올리고 있다[1]. 현재, 그 학습법으로 널리 이용되고 있는 것이 역전파학습법으로 최급강하법이나 공액경사법 등의 최적화 방법이 적용되고 있지만, 학습에 많은 시간이 걸리는 점, 경우에 따라서는 국소적 최적해(local optimization)에 해의 수렴이 이루어져 오차가 충분히 작게 되지 않는 경우 등이 문제점으로 지적되고 있다[2].

신경회로망의 학습알고리즘은 최적설계문제에서 사용되고 있는 최적화 이론을 적용할 수 있고, 대국적 최적해를 구하는 알고리즘을 적용하면 신경회로망의 학습정도 향상을 도모할 수 있다는 것은 쉽게 유추되므로 많은 연구가 진행되고 있다. 최적화 이론에서 대국적 최적해를 구하는 알고리즘으로서 random 탐색법, simulated annealing법, 유전적 알고리즘(genetic algorithm)법, 그리고 tabu 탐색법 등이 있다. 또, 시그모이드 함수의 기울기를 최적화하는 알고리즘이나 의사역행렬을 이용한 고속학습법이 제안되었다[3].

본 논문에서는 Hu[4]에 의해 고안된 random 탐색법과 조합된 tabu 탐색법(random tabu 탐색법)을 연결강도를 구하는 학습알고리즘으로 사용하여 국소적 최적해에 수렴하는 것을 방지하고, 수렴정도를 개선하는 새로운 방법을 제안한다. 이 방법을 배타적 논리합 문제에 적용하여 역전파법 및 학습상수(α, β)에 tabu 탐색법을 이용한 경우와 비교한다. 그리고, 각 파라미터가 오차함수의 수렴에 미치는 영향을 조사한다.

II. 신경회로망의 개요

신경회로망이란, 인간의 신경세포(neuron)의 모델을 네트워크화 하고, 고도병렬분산처리를 실현한 정보처리 시스템의 총칭이다. 신경회로망은 그 구조에 의해 계층형과 상호결합형으로 대별된다. 그 중 계층형 신경회로망에 관해서는 진단, 문자인식, 로봇 관절(manipulator)의 제어, 패턴인식 등의 응용에도 많은 연구가 활발히 진행되고 있다.

그 배경으로는 (1) 입출력관계를 학습시키는 것만으로 비선형성이 강한 사상을 네트워크 내의 연결강도로서 구축할 수 있다고 하는 간편성, (2) 학습패턴을 순조롭게 내삽하고 미학습의 입력에 대해서도 타당한 값을 출력할 수 있는 일반화(generalization)능력, (3) 학습이 종료하면 간단한 연산을 하는 것만으로 출력이 얻어지는 응용시의 신속성, (4) 입력변수, 즉 입력층 유니트의 증가에 용이하게 임의의 연속함수를 임의의 정밀도로서 근사화할 수 있다고 하는 높은 사상(mapping)능력 등의 여러가지 뛰어난 특징을 가지고 있기 때문이다[2].

신경회로망의 학습이란, 정보처리 시스템의 목적에 맞도록 유니트간의 연결강도를 조절하는 것이다. 따라서 학습은 신경회로망에 있어서 가장 중요한 것이며, 최적 분류 패턴을 하기 위하여 연결강도를 수정하는 것이다. 이러한 연결강도의 수정에 의한 학습을 수행하기 위해서 평가기준이 필요하며, 외부에서 교사신호를 입력신호에 대한 이상적인 출력을 비교평가하는 교사있는 학습방식과 평가기준을 내장해서 외부에서 하나하나 교사신호를 부여하지 않는 교사없는 학습방식으로 구분할 수 있다.

계층형 신경회로망은 유니트가 연결되어 그림 1과 같이 구성된다. 이 네트워크에서 입력유니트 i 의 출력 I_i 와 중간층 유니트 j 의 출력 H_j 그리고 출력층 유니트 k 의 출력 O_k 의 관계는 다음과 같다.

$$\begin{aligned}
 H_j &= f\left(\sum_i W_{ji} \cdot I_i + W_{j0}\right) \\
 O_k &= f\left(\sum_j V_{kj} \cdot H_j + V_{k0}\right)
 \end{aligned}
 \tag{1}$$

여기서

W_{ji} : 입력층 유니트 i 에서 중간층 유니트 j 로의 연결강도

V_{kj} : 중간층 유니트 j 에서 출력층 유니트 k 로의 연결강도

W_{j0} : 중간층 유니트 j 의 오프셋,

V_{k0} : 출력층 유니트 k 의 오프셋

$f(x)$: 시그모이드 함수 $f(x) = 1/(1 + \exp(-2x/u_0))$

u_0 : 시그모이드 함수 온도

교사있는 학습방식에서 널리 쓰이는 평가함수로는 출력(O)과 교사신호(T)와의 차의 제곱, 즉 제곱 오차 함수(E)가 이용된다.

$$E = \sum_p E_p = \sum_p \sum_k (T_{kp} - O_{kp})^2 / 2 \quad (2)$$

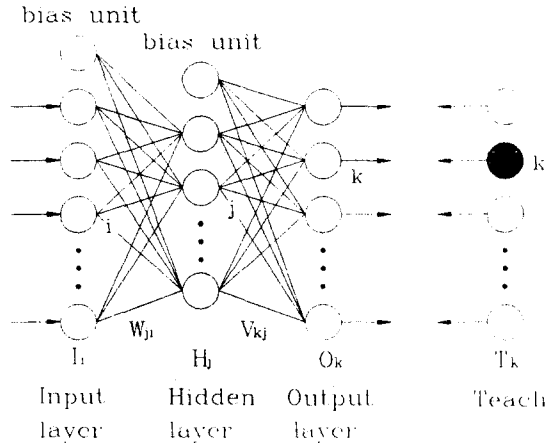


그림 1 신경회로망의 모델
Fig. 1 Neural network model

III. Random Tabu 탐색법

Hu[4]는 Glover가 고안한 tabu 탐색법을 개량하고, 그 방법을 일반적인 제약조건이 있는 최적화 문제에 확장하였다. 어떤 목적함수 $f(x)$ 를 구간 $[a, b]$ 의 제약조건하에서 최소화 하는 문제의 경우, tabu 탐색법에서는 step 수, count 수의 2개의 새로운 정수를 정의한다. Step수는 탐색을 하는 근방영역의 개수이고, count수는 하나의 근방영역을 탐색하는 회수의 상한치를 나타낸다. 제약조건을 만족하는 해의 제 1근사해(초기치)를 x_0 로 하고, x_0 주위에 근방영역 $N(x_0, h_i)$ 을 다음과 같이 설정한다(그림 2).

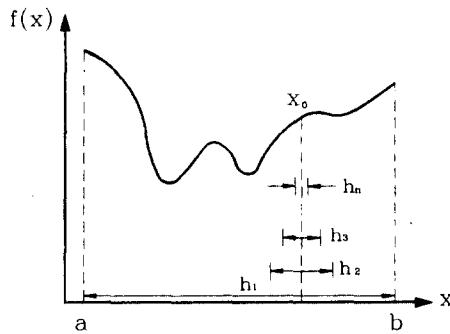


그림 2 근방영역의 설정
Fig. 2 Set of the neighbours

$$N(x_0, h_1), N(x_0, h_2), \dots, N(x_0, h_r)$$

여기서 $h_i(i=1, 2, \dots, r)$ 는 step폭, r 은 step 수, P 는 step비이다. Step h_i 의 집합 H 는 다음과 같이 정의된다.

$$H = \{h_1, h_2, \dots, h_r\} \tag{3}$$

여기서 $h_1 = b - a$

$$h_2 = h_1 P$$

$$h_3 = h_2 P$$

⋮
⋮

$$h_r = h_{r-1} P$$

Tabu 탐색법은 각각의 근방영역 중에서 x 를 임의(random)로 발생시키고, 그것에 의한 $f(x)$ 가 $f(x_0)$ 보다 적으면, 그 점을 그 영역 내의 최소점으로서 기억하고 전 근방영역에서 선택된 최소점들 중에서 가장 작은 x 를 제 2차 근사해 x_1 으로 하고, x_1 주위에 다시 근방영역을 설정하여 탐색을 반복하는 것이다. 기본적인 방법은 Hooke-Jeeves들이 제안한 직접탐색법의 패턴탐색법과 격자탐색을 조합한 방법에 가깝지만[5], 개량 step폭 대신에 근방영역을 설정한 것과 그 크기가 다른 것을 복수개 설정하는 점에서 다르다.

이 방법의 특징으로서는 다음과 같은 것을 들 수 있다.

- (1) 탐색영역을 크기가 다른 여러 영역으로 분할하여 해를 탐색하기 때문에, 맹목적인 탐색을 피할 수가 있고, 탐색에 요하는 회수를 줄일 수가 있다.
- (2) Random 탐색을 사용하기 때문에, 국소적 최적해에 빠지는 것을 막을 수가 있고, 대국적 최적해를 구하는 것이 가능하다.
- (3) 다른 최적화 방법과 조합하여 사용하므로서 보다 정도가 좋고, 고속으로 최적해를 구하는 것이 가능하다.

IV. Random Tabu 탐색법을 이용한 신경회로망의 학습법

신경회로망의 학습은 주어진 입력패턴에 대해서 교사패턴에 가까운 출력패턴을 발생하도록 연결강도의 수정을 하고 이를 통해 결과적으로 평가기준을 만족하는 연결강도를 구하는 것이다. 따라서 평가기준이 되는 목적함수를 식 (2)의 출력과 교사신호의 제곱오차함수로 하고 변수를 입력층과 중간층의 연결강도(W_{ji}), 중간층과 출력층의 연결강도(V_{kj})로 할 때, 제약조건을 만족하고 제곱오차를 최소로 하는 연결강도를 구하는 최적화문제로 다음과 같이 정의할 수 있다.

$$\begin{aligned} E &= \sum_p E_p = \sum_p \sum_k (T_{kp} - O_{kp})^2 / 2 \\ &= f(W_{ji}, V_{kj}) \rightarrow \text{Minimize} \\ a &\leq W_{ji} \leq b, \quad a \leq V_{kj} \leq b \end{aligned} \tag{4}$$

네트워크를 학습시키기 위해 step수와 count수 등의 파라미터를 정한 후 연결강도를 초기화하고 근방영역을 설정한다. 그리고, 연결강도의 행렬 각 요소를 각각의 근방영역 중에서 random하게 발생시키고 이 연결강도에 의한 제곱오차합을 기존의 연결강도에 의한 제곱오차합과 비교하여 적으면 이 연결강도를 그 근방영역에서의 오차최소점으로 기억하고 전 근방영역에서 선택된 최소점들 중에서 가장 작은 오차를 가지는 연결강도를 근사최적 연결강도로 하고, 이 연결강도의 각 요소들 주위에 근방영역을 설정하여 탐색을 반복하는 것이다. 이 과정에서 크

기가 다른 근방영역에서의 탐색이 국소적 최적해에의 수렴을 방지하고 random 탐색에서의 맹목적인 탐색을 피하도록 하고 있다. 본 연구에서 제안한 random tabu 탐색법을 이용한 신경회로망의 구체적인 학습은 다음과 순서로 이루어 진다.

- (1) 네트워크에 입력패턴과 교사패턴을 격납한다.
- (2) 파라미터의 값을 정한다. 즉 중간층 유니트수, 시그모이드 함수의 온도(u_0), step비, step수, count수, 오차 한계치, 최대반복수 그리고 초기 난수발생 경계치를 정한다.
- (3) 연결강도를 초기 난수발생 경계치 내에서 난수로 초기화 한다.
- (4) Step폭 집합 H 를 정의한다.
- (5) Step폭 h_i 에 대한 연결강도의 근방영역을 정의하고 random 이동이 발생되어 진다.
- (6) Random 이동된 연결강도에 의한 오차와 기존의 연결강도에 의한 오차값을 비교한다. 오차가 작아지지 않았다면 그 근방영역에서 다시 random 이동하여 오차값을 비교한다. 단, 하나의 근방영역내에서의 random 이동 횟수는 정해진 count수 까지만 반복이 가능하다. 오차가 작아졌다면 기존의 연결강도를 random 이동된 연결강도로 교체하고 새로운 step폭 h_i 를 정한 후 (5)의 단계로 돌아간다. 단, step수 만큼 반복하고 오차값이 오차한계를 만족하면 종료하고 만족하지 않으면 (5)의 단계로 돌아간다.

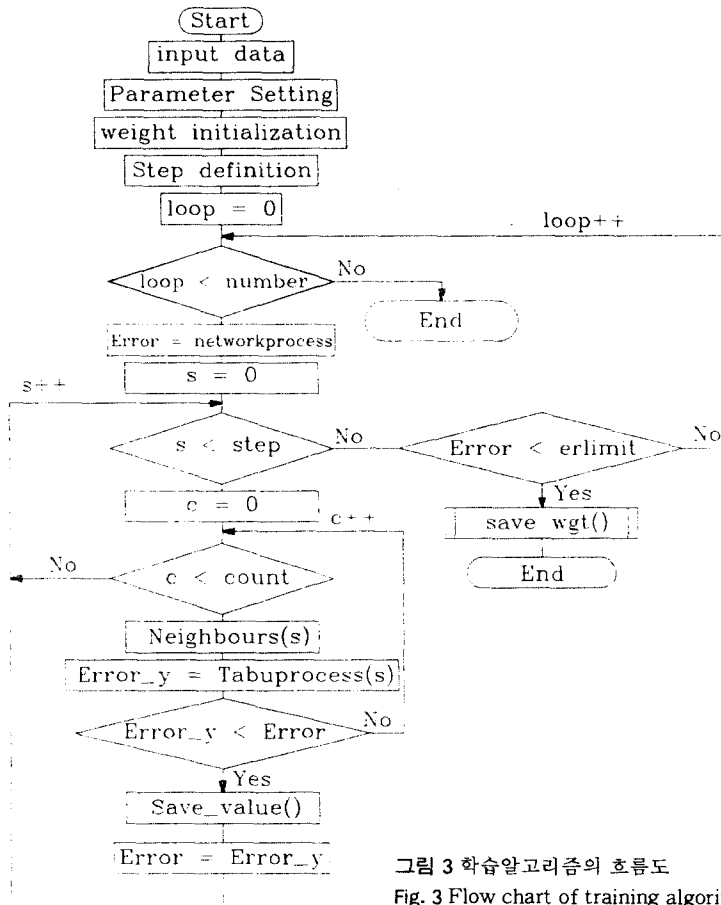


그림 3 학습알고리즘의 흐름도
Fig. 3 Flow chart of training algorithm

그림 3은 이 학습알고리즘의 흐름도(flow chart)를 나타내고 있다.

V. 신경회로망의 학습에의 적용 예

본 연구에서는 적용 예로서 배타적 논리합(XOR)을 이용하였다. 그림 4는 XOR 학습용 신경회로망이다. 표 1은 중간층의 유니트수를 각각 2, 6 및 10개로 변화시켰을 때의 학습결과를 기존의 오차역전파법과 비교한 것이다. 좌측에는 입출력관계가 있는 논리연산자이고, 우측은 random tabu 탐색법에 의한 학습을 30회 반복한 후의 출력값과 오차역전파 알고리즘에 의해 10,000회 학습한 후의 출력값을 비교하여 나타내고 있다. 본 방법에 의한 학습이 학습반복회수가 오차역전파법에 비해 훨씬 적은데도 불구하고 학습결과가 더 좋은 결과를 보임을 알 수 있다.

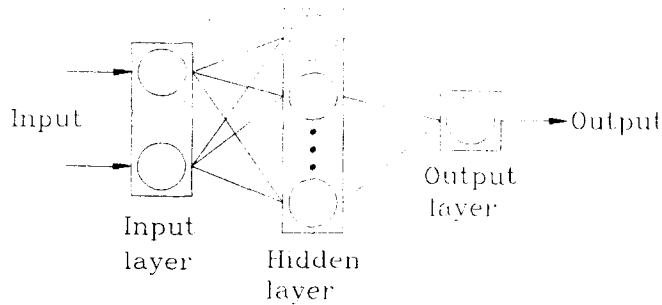


그림 4 XOR 학습용 신경회로망의 구조
Fig. 4 Neural network for training of XOR

표 1 학습결과

Table 1. Learning results

입력신호	교사신호	본 방법			오차역전파법			
		2 개	6 개	10 개	2 개	6 개	10 개	
0	0	0.0	0.00328	0.00332	0.00262	0.00892	0.00385	0.00399
1	0	1.0	0.99592	0.99692	0.99700	0.98974	0.99320	0.99349
0	1	1.0	0.99733	0.99631	0.99653	0.99159	0.99330	0.99405
1	1	0.0	0.00192	0.00120	0.00325	0.00801	0.00814	0.00759

그림 5는 기존의 오차역전파법에서 학습상수 α , β 를 각각 0.1 및 0.2 그리고 시행착오에 의해 가장 양호한 수렴 특성을 보인 0.8에 대한 결과와 tabu 탐색법을 이용하여 학습상수 α , β 를 최적화한 오차역전파법[6]의 결과 및 본 학습법에 의한 제곱오차의 수렴특성을 보이고 있다. 학습상수 α , β 가 각각 0.1과 0.2로 한 기존의 오차역전파법의 경우 국소적 최적해에 수렴되어 반복회수를 10,000회 반복하여도 만족할만한 수렴특성을 보이지 않는다. 시행착오에 의해 가장 수렴특성이 좋은 학습상수 0.8의 경우도 제곱오차가 0.1에 도달하는데에 약 2,000회의 반복을 필요로 하고 있다. 또한 학습상수 α , β 를 설계변수로 하여 본 tabu탐색법을 적용한 경우는 약 750회 반복하여 제곱오차가 0.1에 도달한다. 그러나 본 random tabu 탐색법을 이용한 학습법의 경우, 6회의 반복으로 동일한 정도의 오차에 수렴하여 기존의 방법보다 월등히 수렴속도가 빠르고 국소적 최적해에 수렴하지 않음을 알 수 있다.

그림 6은 연결강도의 초기치 1.0, 중간층 유니트수 10, step비 0.5, step수 5 및 count수 10으로 한 경우의 시그모이드 함수의 기울기에 따른 일정오차에 수렴하는데 필요한 반복수로 오차함수의 수렴정도를 나타내고 있다. 시그모이드 함수의 기울기는 시그모이드 함수의 온도(u_0)의 역수에 상당한다. 시그모이드 함수의 기울기가 클수록

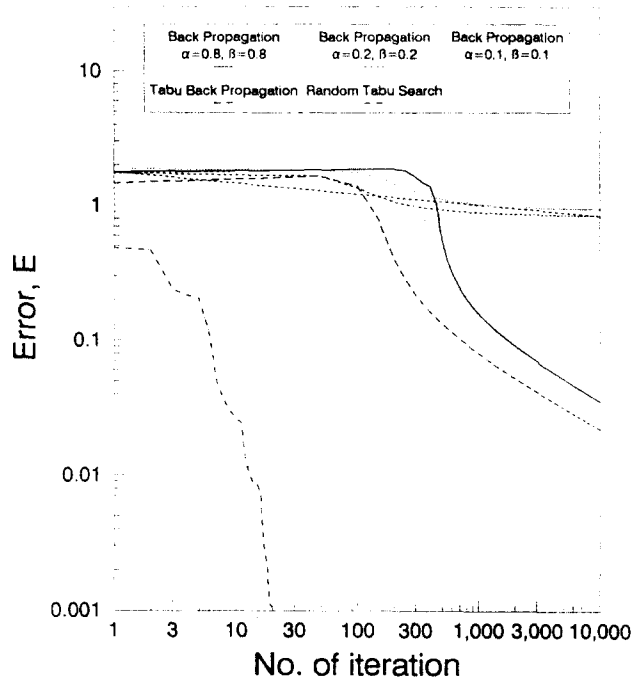


그림 5 오차역전파와 random tabu 탐색법에 의한 학습

Fig. 5 Training a neural network by the error back propagation and random tabu search method

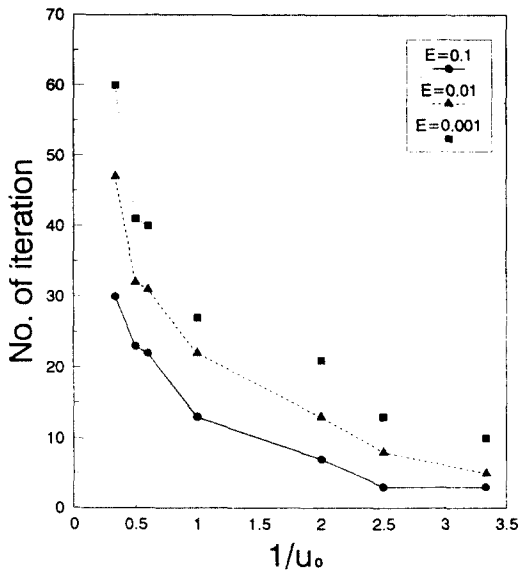


그림 6 시그모이드함수의 기울기(1/u₀)에 따른 수렴특성
Fig. 6 Convergence characteristics due to the slope of sigmoid function

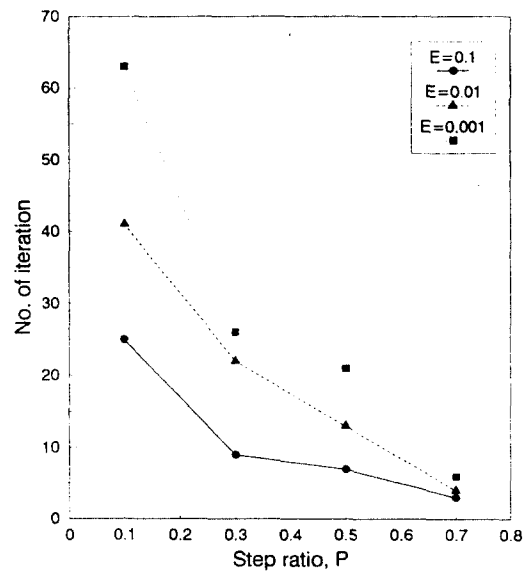


그림 7 Step비에 따른 수렴특성
Fig. 7 Convergence characteristics due to the step ratio

수렴정도는 양호하게 되나 상대적으로 일반화 능력이 떨어질 우려가 있다.

그림 7은 그림 6과 동일한 조건에 $u_0=0.5$ 로 한 경우 근방영역의 폭을 단계별로 줄이는 비율, 즉 step비(P)의 변화에 따른 학습정도를 나타낸 것이다. Step비가 클수록 일정오차에 수렴하는데에 필요한 반복회수가 적어져 수렴속도가 빠르게 된다.

Tabu 탐색법에 있어서 중요하게 고려되는 파라미터는 step수와 count수이다. 이들을 각각 변경했을 때의 수렴 성능을 비교한 결과를 그림 8과 그림 9에 표시한다. 계산은 중간층 유니트수 3, step비 0.5, $u_0=0.5$ 로 하였으며, 그림 8에서 count수는 10, 그림 9에서 step수는 10으로 하였다. 이들 결과로부터 다음을 알 수 있다.

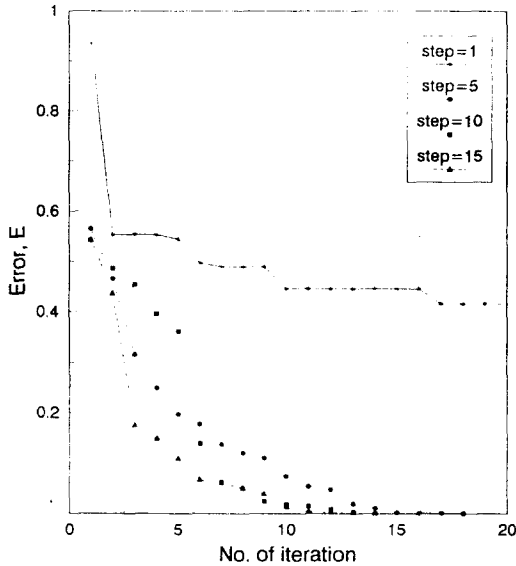


그림 8 Step수에 따른 오차
Fig. 8 Error due to the step number

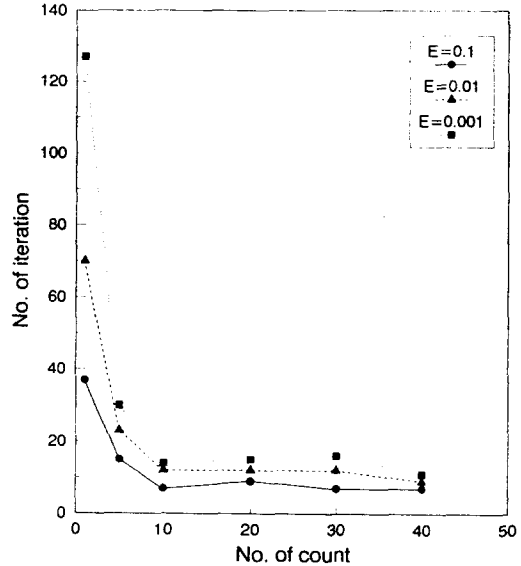


그림 9 Count수에 따른 수렴특성
Fig. 9 Convergence characteristics due to the count number

- (1) Step수가 증가하면, 보다 수렴정도가 좋고 상세한 해의 탐색이 가능하게 된다. 그러나, step수가 너무 많이 증가하여 근방영역의 폭이 아주 작아지게 되면 이 근방영역 중에서의 탐색에 의한 평가함수의 개선이 크게 의미가 없어지는 경우가 있을 수 있다.
- (2) Count수가 증가하면, 일정오차에 도달하는데 필요한 반복수가 급격히 감소하게 되고 count수가 10이상에서는 거의 포화하는 특성을 보이며 오차가 최소, 즉 대국적 최적해에 도달할 확률이 증가하게 된다.

여기서 주의하지 않으면 안되는 것은 이 2개의 정수를 무한히 크게 하면, 불필요한 탐색시간이나 프로그램시 불필요한 메모리 영역을 사용해 버리기 때문에 이들 정수의 결정에는 시행착오를 요한다.

VI. 결 론

Random 탐색법과 조합된 tabu 탐색법을 이용하여 입력층과 중간층 그리고 중간층과 출력층의 연결강도를 변수로 하여 제곱오차를 최소로 하는 새로운 학습알고리즘을 구성하여 국소적 최적해에 수렴하는 것을 방지하고, 수렴정도를 개선하는 새로운 방법을 제안하였고 배타적 논리합 문제에 적용하여 기존의 오차역전파법과 수렴특

성을 비교하고 여러 파라미터에 대하여 검토하였다. 그 결과를 정리하면 다음과 같다.

- (1) 기존의 방법과 달리 국소최적해에 빠지는 것을 방지할 수 있다.
- (2) 대국적 최적해로의 수렴속도가 기존의 방법에 비해 비약적으로 빠르다.
- (3) Step비가 증가할 수록 수렴속도가 빠르다.
- (4) Step수와 count수가 증가할 수록 대국적 최적해로의 수렴에 필요한 학습반복회수가 감소한다.

참 고 문 헌

1. 박민용, 최항식 譯 “뉴로 컴퓨터”, 大英社, 1991.
2. Casimir K., John G., and Garrett P. : “Neural Computing”, Neural Ware, Inc, 1989.
3. Shin Morishita 외 3인, Tabu探索法によるニューラルネットワークの新しい 學習法, 日本機械學會講演論文集, No. 930-42. A pp. 183, 1993.
4. Hu, N., International Journal Numerical Methods in Engineering, Vol. 35, pp. 1055-1070, 1992.
5. Jasbir S. Arora., Introduction to Optimum Design, McGraw Hill, 1990.
6. 김유신, 양보석, 최성필, Tabu탐색법을 이용한 신경회로망의 학습개선에 관한 연구, 대한기계학회 추계학술대회, 1994.