

유전 알고리즘을 이용한 모듈화된 신경망의 비선형 함수 근사화

Nonlinear Function Approximation of Moduled Neural Netowork Using Genetic Algorithm

박현철*, 김성주*, 김종수*, 서재용**, 전홍태*

*중앙대학교 전자전기공학부

**한국기술교육대학교 정보기술공학부

*Hyun Chul Park, *Seong Joo Kim, *Jong Soo Kim, **Jae Yong Seo, *Hong Tae Jeon

*School of Electrical and Electronics Eng., Chung-Ang Univ.

**School of Information and Technology, Korea university of Technology and Education
(anation@ms.cau.ac.kr)

ABSTRACT

Neural Network consists of neuron and synapse. Synapse memorize last pattern and study new pattern. When Neural Network learn new pattern, it tend to forget previously learned pattern. This phenomenon is called to catastrophic inference or catastrophic forgetting. To overcome this phenomenon, Neural Network must be modularized. In this papaer, we propose Moduled Neural Network. Modular Neural Network consists of two Neural Network. Each Network individually study different pattern and their outputs is finally summed by net function. Sometimes Neural Network don't find global minimum, but find local minimum. To find global minimum we use Genetic Algorithm.

Keywords : Neural network, Genetic Algorithm, Nonlinear, Cross over, Mutation

I. 서 론

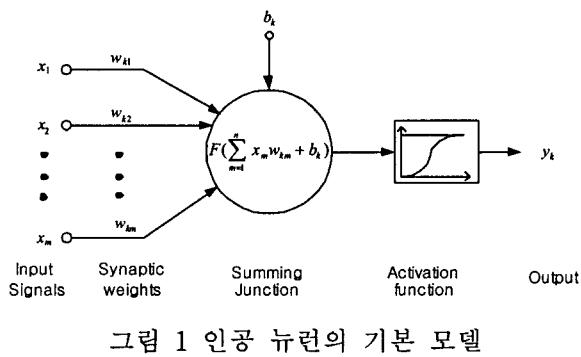
하나의 메모리에 특정 정보가 저장되어있는 기존의 방식과 다리 분산 저장방식을 갖는 신경망은 뛰어난 학습 능력, 일반성, 오류 허용과 같은 특성을 갖는다. 하지만 신경망은 새로운 패턴을 학습하면서 기존의 패턴을 잊어버리는 파멸간섭 현상이 발생하며, 전역적 최소점이 아닌 지역점 최소점에 빠질 가능성을 가지고 있다.

본 논문에서는 이러한 신경망의 파멸간섭 현상을 극복하기위해 모듈화식 설계방식을 도입하였으며, 신경망이 지역적 최소점에 빠지지 않고 전역적 최소점에 도달하도록 유전알고리즘을 이용하였다.

II. 본 론

2.1 신경망과 파멸간섭

인간의 뇌구조를 모델화한 신경망(neural network)은 뉴런(neuron)이라고 부르는 다수의 소연산자들과 그 소연산자들을 결합하는 시냅스(Synapse)라고 부르는 아날로그(analog)의 분산 메모리 소자들로 구성된다. 아날로그의 결합 강도를 시냅스 가중치(weight) 또는 시냅스 결합 강도라고 한다. 신경망 모델의 기본적인 연산 소자인 뉴런은 다입력, 다출력 소자이며 입력신호의 곱과 합의 연산과 비선형 맵핑 연산을 수행하며, 시냅스와 소마(soma)로 구성된다.



시냅스는 과거의 지식을 저장하고 새로운 경험을 학습할 수 있다. 시냅스 연산자는 과거의 경험(w_i)과 뉴런입력(x_i)에 대해서 결합 연산(\circ)을 제공한다. 결국 시냅스 출력

$$z_i = w_i \circ x_i, \quad i = 1, 2, \dots, n \quad (1)$$

여기서, \circ 는 시냅스 연산자로서 내적(inner product)에 의해 구현될 수 있다. 소마 연산자는 뉴런 세포의 내부에서 연산이 수행되며, 결합 연산자 \circ 로 정의하고 합(sum)의 연산으로 구현한다. 소마는 소마집합(somatic aggregation)과 비선형 활성화 함수(nonlinear activation function)로 구성된다. 소마집합(u)은 다음 식 와 같이 나타낼 수 있다.

$$u_i = \sum_{i=0}^L Z_i \quad (2)$$

모든 자연계의 인식 시스템은 과거에 학습한 정보를 점진적으로 잊어버리게 된다. 특히 우리 자신과 같은 인간에게 이러한 현상이 발생한다. 인간의 인식을 모사한 모델은 새로운 정보가 요구될 경우 과거의 정보를 점차 잊어버리는 유사한 패턴을 보이고 있다. 즉, 대부분 자연계의 인식 시스템은 일반적으로 파열적 망각의 속성을 가지고 있다. 불행하게도 분산 연결 망(distributed connectionist networks)의 경우, 이러한 현상은 어떤 환경에서는 부분적으로 발생할 수 있다.

인간의 뇌구조를 모델화한 신경망 역시 학습된 패턴과 다른 새로운 집합을 학습하는 과정에서 이미 학습이 완료된 망의 지식이 없자기 완전

히 제거되는 현상을 발생한다. 이러한 파열간섭은 내부의 분산 표현의 중복에 의해 발생하기 때문에 이러한 중복을 줄임으로써 파열간섭 현상을 감소시키기 위해 분할된 가중치의 단일 집합을 사용하여 망의 일반화된 특성과 파열간섭 현상이 감쇠하도록 하였다.

2.2 유전 알고리즘

유전 알고리즘은 생명체의 자연 진화 법칙을 일반적인 최적화 문제에 적용한 방법으로 노드(node) 영역에 다수의 염색체를 분포시켜 목적 함수 값과 위반 정도에 따라 각각의 염색체 적합도(fitness)를 부여한다. 적합도가 높을수록 교배와 돌연변이 과정에 참여할 확률을 크게 하여 적합도가 높은 염색체에 비슷한 형질을 가진 염색체들이 다음 세대에 보다 많이 나타날 수 있도록 하므로 전체 염색체들이 좋은 방향으로 탐색이 진행된다. 적합도가 높은 최적점을 계속 활용하는 것은 기존의 최적화 방법에서 미분값(gradients)을 이용하는 것과 유사하며 교배와 돌연변이를 시켜가면서 최적점을 찾아내는 것은 기존의 임의 탐색법과 유사하기 때문에 범용성과 유용성이 높은 알고리즘으로 표현된다.

유전 알고리즘에서 개체의 수를 개체군의 크기(population size)라고 한다. 각각의 개체는 염색체(chromosome)를 가지고 있으며 염색체는 복수개의 유전자(gene)의 합으로 구성된다. 유전 알고리즘이 다른 탐색 알고리즘이나 최적화 방법과 구별되는 특징은 다음과 같다.

- ① 유전 알고리듬은 해공간(solution space)에서 단일 해를 사용하는 것이 아니라 해집단(population of solution space)을 이용한다.
- ② 최적화 함수의 미분 가능성이나 연속성 등과 같은 수학적 정보를 필요치 않는다.
- ③ 유전 알고리즘은 전역적(global)인 최적해의 발견을 가능케 한다.

유전 알고리즘은 위와 같은 특징으로 인해 다른 탐색 또는 최적화 방법에 비하여 전역적 해를 구할 가능성이 높다

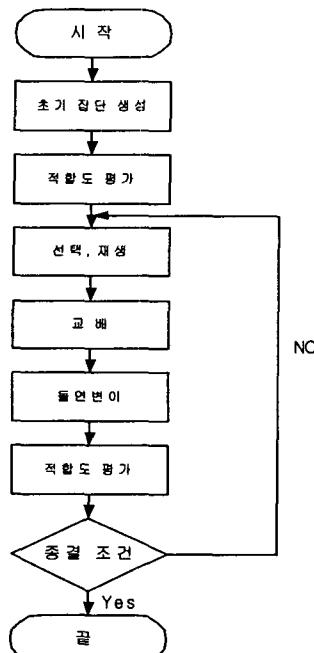


그림 2 유전 알고리즘 흐름도

2.3 모듈화된 신경망

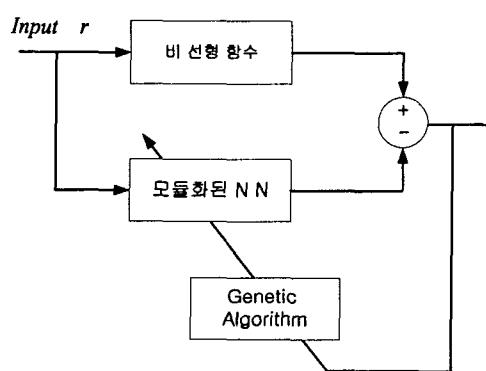


그림 3 모듈화된 신경망-유전알고리즘
제어 시스템

그림3은 모듈화된 유전알고리즘을 이용한 모듈화된 신경망을 통해서 비선형함수를 근사화하는 구조를 나타내었다.

모듈화된 신경망은 일반적으로 모듈화 신경망 구조를 이용해서 창안한 것으로 모듈화 신경망 구조의 연결강도가 게이팅 네트워크를 통해서 이루어지는데 반해서 유전알고리즘을 이용해서 신경망의 연결강도를 갱신하도록 하였다.

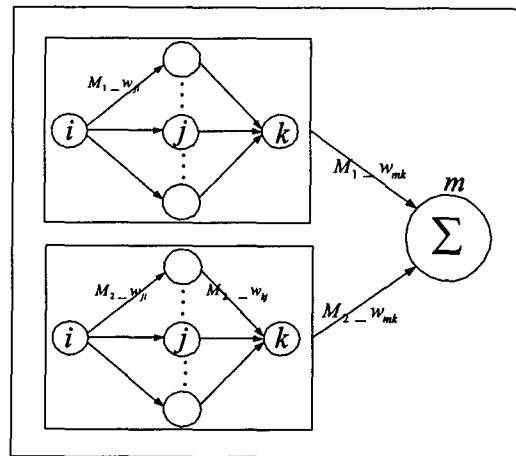


그림4는 모듈화된 신경망 구조를 나타낸 것으로 유전 알고리즘의 초기 개체군을 이루는 각 개체의 초기화는 신경망의 연결강도를 갖도록 코딩하였으며 각 개체군의 적합도는 식(3)으로 구하였다.

$$Fitness = \frac{K_1}{ERROR} + \frac{K_2}{Change\ of\ ERROR} + \frac{K_3}{Energy} \quad (3)$$

2.4 근사화 실험

근사 대상 함수 $f(x)$ 는 구간 $[0, 1]$ 에서 다음과 같이 구간별로 정의된 함수이다.

$$f(x) = \begin{cases} \sin(2\pi \times 2t) + \sin(2\pi \times 2t) & t \leq 0.51 \\ \sin(2\pi \times 5t) + \sin(2\pi \times 10t) & 0.51 < t \leq 1.0 \end{cases} \quad (4)$$

유전 알고리즘에 사용된 제어 파라미터들은 표 1과 같다.

개체군의 크기	교배율	돌연변이율
30	70%	30%

표 1 유전알고리즘 제어 파라미터

근사 대상 함수를 신경망을 이용한것과 모듈화된 신경망과 유전알고리즘을 이용한 본 논문에서 제안된 구조로 실험하였다.

근사 대상함수를 신경망으로 근사화한경우는 역전파 학습 알고리즘을 이용하였으며, 역전파 학습에 의해 연결강도를 갱신해가는 과정은 다음과 같다.

각 입력층에 대한 은닉층의 출력은

각 입력층에 대한 은닉층의 출력은

$$net_{pj}^h = \sum_{i=1}^N w_{ji}^h x_i, \quad i_{pj} = f_j^h(net_{pj}^h) \quad (5)$$

이 되며,

module 1의 출력은

$$net_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj}, \quad o_{pk} = f_k^o(net_{pk}^o) \quad (6)$$

이된다.

module 1과 module 2의 출력값 s_{pm} 은

$$net_{pm}^n = \sum_{m=1}^S w_{mj} o_{pk}, \quad s_{pm} = f_m^n(net_{pm}^n) \quad (7)$$

이된다.

각 노드의 에러를 계산하면 다음과 같다.

모듈의 에러 δ_{pm}^s 은 식(8)이 되며,

$$\delta_{pm}^s = (Y_m - s_{pm}) f_m'(net_{pm}^n) \quad (8)$$

module 1의 출력단과 은닉층의 에러는 식(9)와 식(10)이 된다.

$$\delta_{pk}^o = f_f'(net_{pk}^o) \sum_m \delta_{pm}^s w_{mj} \quad (9)$$

$$\delta_{pj}^h = f_j'(net_{pj}^h) \sum_k \delta_{pk}^o w_{kj}^o \quad (10)$$

위에서 구한 에러를 이용해서 각각의 연결 강도를 갱신해주면 다음과 같이 된다.

모듈의 연결강도, 신경망의 출력단, 신경망의 인식층의 연결강도는 식(11), 식(12), 식(13)이 된다.

$$w_{mj}^s(t+1) = w_{mj}^s(t) + \eta \delta_{pm}^s o_{pk} \quad (11)$$

$$w_{kj}^o(t+1) = w_{kj}^o(t+1) + \eta \delta_{pk}^o i_{pj} \quad (12)$$

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \eta \delta_{pj}^h x_i \quad (13)$$

III. 결 론

본 논문에서는 신경망의 과열현상을 극복하기 위해 모듈화된 신경망을 제안하였으면, 신경망이 가지는 지역적 최소점에 빠지는 현상을 극복하여, 전역적 최소점을 찾기 위한 유전 알고리즘을 이용하였다.

그림 5의 비선형 함수 근사 결과에서처럼 모듈화된 신경망을 이용해서 서로 다른 주기를 갖는 함수를 두 개의 신경망이 분할해서 학습함으로인해 학습함으로서 새로운 입력이 들어왔을 때 기존의 패턴이 파괴되는 현상이 일어나

지 않았으며, 신경망이 지역적 최소점에 빠질 수 있는 현상을 유전 알고리즘의 교배와 돌연변이를 이용해서 전구간을 탐색함으로인해 지역적 최소점을 극복하고 전역적 해를 찾을 수 있었다. 또한 신경망을 역전파 학습 알고리즘을 이용해서 학습한 결과보다 더좋은 결과를 보였다.

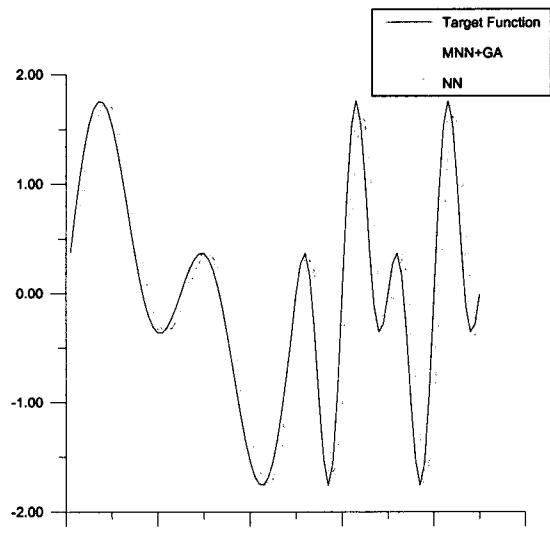


그림 5 비선형 함수 근사 결과

감사의 글 : 본 연구는 과학기술부의 뇌과학연구 프로그램에 의해 지원 받았습니다.

IV. 참고문헌

- [1] Simon Haykin, *Neural Networks - A Comprehensive Foundation*, Macmillian College Publishing Company Inc., 1994.
- [2] R. K. Elsley, "A learning architecture for control based on Back-Propagation neural network", *Proc. of the IEEE Conf. on Neural Networks*, vol. 2 .pp 587-594. 1988
- [3] S. R. Chi, R. Shouresshi, and M. Tenorio, "Neural Networks for system identification ", *IEEE Contr. syst. Mag.*, vol. 10, pp. 31-34, 1990.
- [4] Gasser Auda and Mohamed Kamel, "Modular Neural Network Classifiers: A Comparative Study," *Journal of Intelligent and Robotic Systems*, Vol. 21, pp, 117-129, 1998.