

# 유전자 알고리즘 최적화 신경망을 이용한 학습

윤여창

우석대학교 정보보안학과

e-mail : [yoonyc@core.ws.ac.kr](mailto:yoonyc@core.ws.ac.kr)

## A Learning Using GA Optimized Neural Networks

YeoChang Yoon

Dept. of Information Security, Woosuk University

### 요 약

시스템 분석에 주로 사용하는 자료 중에는 비선형 자료와 시계열 등이 있다. 이들 자료는 그 함축적인 관계가 매우 복잡하여 전통적인 통계분석 도구로 분석하는데 어려움이 많다. 본 연구에서는 현실 세계에서 다양하게 나타나는 복잡성을 다루기 위하여 하이브리드 진화 신경망 모델링 접근 방법으로 자료를 모형화하고 이를 통한 학습의 적합도를 살펴본다. 비선형 자료 등을 모형화하기 위한 학습은 역전파 신경망 기법을 이용한다. 학습의 효율을 높이기 위해서 격자감소 학습 알고리즘과 함께 이용하는 유전자 알고리즘은 네트워크 구조를 최적화시킬 수 있는 초기가중값을 이용한 전역최소값을 찾는 데 이용한다. 학습 결과를 통해 제안된 하이브리드형 접근방법의 학습이 보다 효율적임을 살펴보기 위하여 유전자 알고리즘으로 최적화된 신경망 학습 알고리즘을 비선형 모의자료의 학습에 적용하여 보았다.

### 1. 서론

복잡한 시스템에서 예측을 통한 의사결정은 과거 자료들의 패턴을 분석한 후 미래의 경향은 과거패턴에 따른다는 가정을 통해서 이루어진다. 대부분의 시계열자료는 특정 시점 구간에서 관측한 자료이다. 이러한 자료는 체계적이며 시스템적인 움직임을 보이고 있으며 이에 대한 많은 연구가 시계열 예측모형으로 개발되어 오고 있다. 여기서 시계열모형의 다양한 유형들이 이용된다고 하여도 복잡한 현실 문제에 비선형모형을 적용할 때 많은 어려움이 있다[1]. 이러한 제약들을 해결하기 위하여 비선형자료는 주로 인공신경망, SVM(Support Vector Machine)등의 데이터마이닝 기법들을 이용하여 분석하고, 보다 로보스트한 비모수적 모델링에는 역전파 신경망(Backpropagation neural network, 이하 BPNN)을 이용한다[2]. 그러나 오차평면이 울퉁불퉁한 경우에 격자감소법을 이용하는 처리절차로는 지역최소값을 쉽게 벗어날 수 없기 때문에[3-5], BPNN 은 불일치하거나 예측할 수 없는 결과를 보일 수 있다. 더욱이 시계열분석에서 중요한 역할을 하는 입력자료들의 불확실성들을 다루는데 있어서 여러 가지의 제약들이 있다. 이 연구에서는 비선형 예측모형 뿐만 아니라 네트워크 구조의 입력자료를 최적화하고 전역 최적값을 구하기 위한 초기가중값을 구하기 위하여[6], 문제해결과 함수의 최적화 방법에 잘 적용할 수 있는 도구인 유전자 알고리즘(Genetic Algorithm, 이하 GA)을 이용한다. 비선형자료를 학습시키기 위해서는 격자감소 알고리즘과 BPNN 을 이용한다. 제안된 하이브리드 진화 신경망 모델링 접근방법(HENN)은 비선형모형으로 발생시킨 모의자료에 적

용해 본다. 그리고 제안한 방법으로 예측한 결과는 BPNN 모형을 이용한 결과와 비교해 본다.

### 2. BPNN 을 이용한 시계열 예측모형

시계열예측은 과거 관측값들을 이용한 예측 그리고 자료의 함축된 관계를 설명하기 위한 모형분석에 중요한 분야이다. 모형은 시계열을 이용하여 미래를 추론하는데 이용한다. 시계열  $\{u_i\}$  ( $i = 1, 2, t, \dots, n$ )를 수식으로 표현한 예측모형은 다음 식(1)과 같다.

$$u_{t+1} = f(u_t, u_{t-1}, \dots, u_{t-p+1}). \quad (1)$$

여기서  $p$  는 입력자료의 길이이다.  $f: R^p \rightarrow R$  는 고차 비선형 관계를 표현하고 있고 BPNN 으로 모형화한다. 즉  $f: R^n \rightarrow R^m$ 로 나타낸다. 신경망을 나타내는 BPNN 의 전형적인 표현식은

$$f(x) = BPNN_{[n;nh_1, nh_2, \dots, m]W}(x), \quad (2)$$

이다. 여기서 BPNN 모형의 구조를 정의하는 모수들인  $n$  과  $m$  은 각각 입력과 출력패턴의 차원이다.  $nh_1, nh_2$ 은 각 은닉층의 노드의 개수이고  $W$  는 연결 가중값이다. 식(1)에 의하면  $n=p, m=1$  이다. 이때 BPNN 시계열모형은 다음과 같다.

$$u_{t+1} = BPNN_{[p;nh_1, nh_2, \dots, 1]W}(u_t, u_{t-1}, \dots, u_{t-p+1}). \quad (3)$$

현실 응용문제에 BPNN 을 이용할 때 역전파 알고리즘은 전통적인 방법을 따른다. 주요 연구들에서 분석자들은 네트워크 구조(입력노드의 개수, 은닉층과 은닉노드 개수 등)와 초기가중값등을 사전에 정의한다. 이 모수들을 정의하는 것은 지역최소값 문제를 해결하기 위해서 중요한데, 이 연구에서는 복잡한 해공간에서 전역 최적해를 찾기 위해 유전자 알고리즘을 이용한다. 자세한 내용은 다음 절에서 설명한다.

### 3. BPNN 시계열모형의 GA 최적화 학습

이 장에서는 신경망의 연결 가중값과 네트워크 구조 간의 관계를 살펴본다.

#### 3.1 네트워크 구조의 최적화

이 단계에서 입력층을 포함하는 네트워크 구조에서 모수들은 GA 에 의해 생성한다. 특히 2 개의 은닉층을 갖는 BPNN 을 이용하는 경우에 모수 구조는  $p, nh_1$  그리고  $nh_2$ 이다.

단계가 진행되는 동안 모수 구조들은 가중값이 생성되는 동안 찾아지며 최적화는 다음 단계에 이루어진다. 랜덤하게 생성된 모수집합  $\{p, nh_1, nh_2\}$ 으로 학습을 시작하며, 이어지는 후속 세대들은 최적의 구조를 찾기 위하여 유전형질 전달 재생산, 유전교차 교배, 돌연변이 등과 같은 유전연산자를 이용하여 생성할 수 있다. BPNN 의 구조적 정보는 이 단계에서 사용된다. 각 모수집합의 적합성은 직접적으로 예측분석을 통하여 추정할 수 있다. 적합성 추정은 가중값들이 최적화된 후의 다음 번 단계에서 실행한다.

모든 단계에서 생성된 네트워크 구조의 새로운 세대를 갖게 되는데 각 네트워크구조의 가중값들은 다음 세대에서 선택되고 최적화된다.

#### 3.2 초기가중값의 선택

네트워크구조 최적화 단계에서 생성된 모든 네트워크 구조에 대하여 연결 가중값들은 네트워크 연결구조에 따라 구해진다. 이때 표준 GA 의 진화 처리과정은 거의 유사한 최적값을 구하기 위하여 현재 네트워크의 초기가중값들을 개선시킨다.

이 경우에 각 가중값의 적합성은 현재 네트워크 구조와 실제 자료에 대한 가중값에 의해 정의된 BPNN 모형의 계산결과를 비교함으로써 직접 추정한다.

#### 3.3 최적화 초기가중값과 학습

GA 가 발견적 학습에 의한 해가 찾아진 후에 전역탐색을 수행하지 못하는 경우가 있기 때문에, 현재 네트워크에 대한 최적 가중값을 개선시키기 위해 역전파 알고리즘을 이용한다. 선행 연구에서 주요 이슈인 과다적합 문제를 해결하기 위하여, BPNN 모형은 먼저 검증자료에 대하여 각 학습 반복주기의 끝에서 주어진 예측 정확도 판단을 위한 검정을 한다.

#### 3.4 적합성 검정

진화 처리과정 중에 생성된 BPNN 들은 적합성을 위하여 측정된 자료에 대한 대안적 검정을 해야 한다. 수집된 자료는 각각 학습자료, 검증자료의 두 그룹으로 나눈다. 가장 좋은 BPNN 은 예측값과 측정값 사이에 최소 오차를 보여주는 것이다. 그래서 적합성은

$$u_{NN} = BPNN(p, nh_1, nh_2, \dots, 1|W),$$

$$fitness = \frac{1}{n} \sqrt{\sum_{i=1}^n (u_{NN,i} - u_{target,i})^2}, \quad i = 1, 2, \dots, n, \quad (4)$$

와 같이 계산된다. 여기서  $u_{NN}$ 은 BPNN 시계열모형의 출력값이고  $u_{target}$ 은 목표값,  $n$ 은 학습자료의 전체 개수이다.

모든 네트워크 구조의 가중값들을 최적화하면, 각 네트워크 구조의 적합성이 학습자료와 검증자료의 두 경우에 대하여 추정될 수 있다. 적합성이 만족되면 사전에 정의된 중단조건에 따라 학습 중단된다.

### 4. 알고리즘

제안된 진화 신경망학습 접근의 주요 단계별 설명은 다음과 같다[7,8].

- I. 네트워크구조의 모수를 랜덤하게 생성한다.
- II. 아래 단계에 따라 각 네트워크 구조의 적합성을 검정한다.
  - i. 식(1)의 모형의  $p$  값에 따라 학습자료와 검증 자료를 구분한다.
  - ii. 주어진 모형에 대응하는 가중값 집합을 랜덤하게 생성한다.
    - i). 식(3)과 같은 네트워크 구조에 대응하는 각 가중값 집합의 적합성을 추정한다.
    - ii). 만약 가중값 진화의 학습 중단조건이 만족되면 단계 ii 로 간다. 그렇지 않으면 가중값 집합의 새로운 모수를 생성시키기 위하여 유전자 연산자를 실행시키고, 단계 i로 간다.
    - iii). 전통적인 격자감소 알고리즘을 이용하여 진화시킨 초기가중값을 적용한 BPNN 을 학습한다.
  - iii. 검증자료에 대하여 학습한 BPNN 을 검정한다. 그리고 식(3)에 따르는 네트워크 구조의 적합성을 계산한다.
  - iv. 만약 모든 네트워크 구조의 적합성이 계산되면 단계 III로 간다. 그렇지 않으면 다음 번 네트워크 구조의 적합성을 계산하기 위하여 단계 ii 로 간다.
- III. 만약 구조 진화의 중단조건이 만족되면 알고리즘은 끝나고 최적의 학습과 예측결과와 함께 BPNN 모형을 선택한다. 그렇지 않으면 단계 IV로 간다.
- IV. 네트워크 구조의 새로운 모수를 생성시키도록 유전자 연산자를 실행한다. 그리고 단계 II로 간다.

### 5. 응용 예

다음과 같은 비선형 함수를 추정하기 위한 네트워크를 학습한다고 하자.

$$f(x) = e^{-(x-1)^2} + e^{-(x+1)^2}, \quad x \in [-2.5, 2.5] \quad (5)$$

식(5)에서 11개의 모의 학습자료  $x$ 는 구간  $[-2.5, 2.5]$  의 균등분포에서 발생하며, 발생시킨 목표값  $f(x)$ 는 평균 0, 분산 0.005인 정규분포를 따르는 오차를 포함하고 있다.

네트워크 구조는 한 개의 입력값과 그에 대응하는 한 개의 출력값 그리고 4개의 은닉노드로 이루어진 1-4-1 네트워크를 이용한다. 입력과 출력노드는 선형이며 은닉층에는 시그모이드형 변환함수를

이용한다.

초기가중값은 균등분포  $U(0,1)$  을 따르는 확률난수 (RND)를 이용한 식  $20 \times (RND - 0.5)$  에서 주어진 구간 안에서 발생시킨다.

비교를 위해 HENN과 BPNN을 이용한 응용 예에서 이용하는 구조는 3층 신경망 구조이다. 여기서 학습의 중단조건으로 설정한 최종 학습 허용오차는  $\epsilon = 0.01$  이고, 초기가중값의 범위는  $|W| < c$  이다. 여기서 HENN인 경우는  $c = 1$  이고, BPNN에서는  $c = 0.001 \sim 10$  이다. 각각의 경우에 오차 판단기준은 식(4)의 MSE이다.

본 연구에서 고려하는 비선형함수를 이용한 신경망 학습의 결과는 표 1과 같다. 이 결과에서 볼 수 있듯이 HENN은 인위적으로 7회 학습을 하였고 각 학습에 대하여 주어진 비선형 함수에 거의 접근시킬 수 있었다.

즉 초기가중값의 범위 조정을 통해 초기값을 작은 범위에서 발생시킨다는 전제 하에서 살펴본다고 하여도, BPNN을 적용할 때보다 HENN을 적용 시에, 학습량과 학습의 정도에서 더 효율적이고 좋은 성능을 보이고 있음을 알 수 있다.

<표 1> 신경망 학습결과

#	HENN		c	BPNN[8]	
	학습량	MSE		학습량	MSE
1	182	0.00995	0.001	1227	0.012571
2	210	0.01024	0.01	452	0.010102
3	1000	0.05214	0.1	684	0.010874
4	179	0.00091	1	576	0.011545
5	183	0.01021	2	635	0.012221
6	102	0.01001	5	701	0.012567
7	131	0.01008	10	817	0.012983

## 6. 결론

이 연구에서는 복잡한 비선형자료를 모형화 하기 위하여 BPNN 을 학습하기 위한 진화 신경망 모델링 접근방법에 대하여 살펴보았다. 진화체제는 입력자료와 발견적 학습을 통한 최적화 초기가중값 뿐만 아니라 네트워크 구조의 진화적 진역탐색을 수행시킬 수 있다. 이 방법을 이용한 HENN 학습과 예측측면에서 둘 다 유의적인 개선이 이루어졌다. 실제적인 응용은 비선형 함수로부터 발생시킨 모의 자료를 이용하였다. 제안한 방법은 복잡한 비선형자료를 분석하고 예측하기 위한 유용한 도구로서 사용할 수 있음을 알 수 있었다.

## 참고문헌

[1] G.E.P. Box and G. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco, CA, 1970  
 [2] G.P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, 2003, Vol. 50, pp. 159-175  
 [3] B.K. Wong, T.A.E. Bodnovich, and Y. Selvi, "A Bibliography of Neural Network Applications Research:

1988-1994," Expert Systems, 1995, Vol. 12, pp. 253-261  
 [4] B. Curry and P. Morgan, "Neural Networks: a Need for Caution," OMEGA, International Journal of Management Sciences, 1997, Vol. 25, pp. 123-133  
 [5] P. Cortez, M. Rocha and J. Neves, "A Lamarckian Approach for Neural Network Training," Neural Processing Letters, 2001, Vol. 15, pp. 105-116  
 [6] D.E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA, 1989  
 [7] C.X. Yang and Y.F. Zhu, "Time Series Analysis Using GA Optimized Neural Networks," Third International Conference on Natural Computation, 2007, pp. 270-276  
 [8] Y. Yoon, "A Comparison on the Learning Effect of Simulated Nonlinear Data Using a Modified Genetic and Backpropagation Algorithm," Proceedings of the Korea Computer Congress, 2005, Vol.32, No.1(B), pp. 694-696