

개선된 유전자 알고리즘과 역전파 신경망 알고리즘을 이용한 비선형 모의자료의 학습비교

윤여창

우석대학교 e-정보공학과

e-mail : yoonyc@core.woosuk.ac.kr

A Comparison on the Learning Effect of Simulated Nonlinear Data Using a Modified Generic and Backpropagation Algorithm

YeoChang Yoon

Dept. of e-Information Engineering, Woosuk University

요 약

본 논문에서는 개선된 유전자 알고리즘과 역전파 신경망 알고리즘의 특징을 살펴보고, 비선형 모의자료를 이용하여 개선된 유전자 알고리즘 기반의 신경망 학습 효과와 역전파 신경망 알고리즘을 이용한 신경망 학습 효과를 비교해 본다. 유전자 알고리즘을 이용한 신경망 학습에는 개선된 신경망 제어기를 이용한다. 역전파 알고리즘을 이용한 신경망 학습에는 일반화 성능향상을 위한 인자들의 결합효과를 이용한다. 모의실험을 통하여 두 가지의 학습에서 학습 수렴의 정도와 학습 속도 등을 비교하는 모의실험 결과를 개선된 유전자 알고리즘과 신경망 알고리즘의 학습 결과와 함께 제시한다. 모의실험의 결과로서 유전자 알고리즘을 적용한 개선된 신경망 제어기를 통한 학습 결과가 일반 신경망 학습 결과보다 초기 가중값을 작은 범위에서 발생시킬 때 수렴 정확도 및 학습 속도에서 좋은 결과를 나타내 주고 있다.

1. 서 론

전형적인 유전자 알고리즘은 다윈의 자연선택(natural selection)에 근거한 생물학적 진화론을 이용한 일종의 계산 모형(computational model)으로써 매우 강력하며 최적의 성능을 보이는 알고리즘 중의 하나로 알려져 있다 [1,2]. 유전자 알고리즘의 주요 특성으로는 단순히 전형적인 역전파 알고리즘에서 주로 이용하는 격자정보에만 의존하지 않으며, 모집단과 그 탐색전략에서의 개별적인 정보전달에 그 특성이 있다. 따라서 전통적인 방법을 이용하였을 때에 해결하지 못했던 복잡한 비선형 문제들을 주로 다루는 경우에 유전자 알고리즘은 문제해결에 보다 적합하며, 단순하거나 또는 매우 포괄적이고 강력한 응용문제의 적용에 많은 특성들을 갖고 있다고 알려져 있다 [1,2,3,4].

유전자 알고리즘은 다양한 분야에서 이용되고는 있지만 그 이론적인 배경이 상대적으로 약하고, 매우 많은 개체 수와 세대를 필요로 하는 경우가 대부분이다. 그리고 유전교차를 통하여 항상 좋은 방향의 개체를 생성시키지는 않는다는 것이 극복할 문제점 중의 하나이다.

신경망 알고리즘을 이용한 학습은 일반적으로 역전파 알고리즘으로 수행되고 있다. 이 알고리즘은 복잡한 비선형 함수와 패턴인식 그리고 자가 학습능력 등에서

전반적으로 학습이 수렴될 수 있는 특성을 갖고 있기 때문에 경제 시계열 분석, 전산응용 관련 그리고 자동제어 분야 등과 같이 다양한 분야에서 주요 연구 과제로 대두되어 오고 있다.

역전파 알고리즘도 다양한 분야에서 넓게 이용되고는 있지만 주요 학습 알고리즘으로써의 극복할 문제점들을 많이 갖고 있다 [2]. 역전파 신경망 알고리즘은 수렴 속도가 일반적으로 느릴 뿐만 아니라 지역 최소값에 수렴되는 경향이 있는 격자감소법을 이용하는 교사 학습 알고리즘의 일종이라는 것에 그 이유가 있을 수 있다.

본 논문에서는 개선된 유전자 알고리즘과 역전파 신경망 알고리즘을 이용하여 비선형 함수를 따르는 모의자료를 이용한 신경망 학습의 효과를 살펴 보고, 이를 통한 수렴의 정확도 및 학습 속도 등을 비교해 보고자 한다. 유전자 알고리즘의 개선을 위해서는 유전교차율의 적용방법을 개선시킨 자동적응조정 알고리즘을 이용한다. 개선된 역전파 신경망 알고리즘은 초기 가중값을 적절하게 적용하면서 신경망 학습시의 수렴의 정도를 높여주고, 최소한의 단순 구조를 갖는 일반화 네트워크로 추정할 수 있게 한다. 따라서 두 가지 알고리즘의 학습 결과로서 나타나는 추정 모형은 최소한의 구조를 갖게 되며 모의실험을 통하여 추정의 정도를 각각 비교한다.

2. 개선된 학습 알고리즘

2.1 개선된 유전자 알고리즘

이 논문에서 고려하는 개선된 신경망 제어기의 구조는 그림 1과 같다. 다층 신경망은 신경망 제어기를 구성하며 네트워크의 가중값들은 유전자 알고리즘을 통하여 최적화 된다 [5].

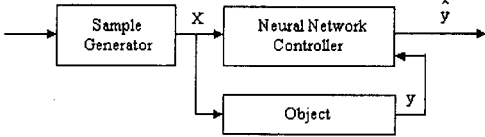


그림 1. 신경망 제어기를 갖는 유전자 학습 구조.

간단한 구조의 전형적인 유전자 알고리즘은 지금까지 다양하게 개발되어 왔지만 전역 최소값으로의 수렴과 최적해로의 탐색에서 정교한 적합능력에 좋은 효율을 보이지 않고 있다 [5]. 이러한 단점을 보완하기 위하여 이 논문에서는 유전자 알고리즘에서 유전교차율과 돌연변이율에 대한 몇 가지의 개선된 방법을 이용한 개선된 알고리즘을 제시한다.

유전자 알고리즘에서 중요한 두 가지 인자들은 유전교차율(crossover rate, P_c)과 돌연변이율(mutation rate, P_m)이다. 이들 인자들은 유전자 알고리즘을 이용한 학습에서 보다 효과적인 탐색을 하는데 직접적인 영향을 준다. 여기서 간단한 유전자 알고리즘과 유전교차율 그리고 돌연변이율은 일반적으로 연구자들의 경험적인 학습 결과에 의하여 주어지게 된다. 이 연구에서 각 인자들에 적용시킨 확률은 $P_c = 0.5 \sim 1$, $P_m = 0.001 \sim 0.5$ 구간의 상수이다 [6]. 여기서 두 인자들을 서로 다른 구간의 값으로 선택한 이유는 알고리즘의 탐색 효율에 대하여 연구자의 경험적인 탐색 영향을 반영하였기 때문이다. 그리고 이 논문에서의 유전교차율은 다음 식 1과 같이 자동 적응조정 알고리즘을 이용한다.

$$P_c = \begin{cases} k_1(f_{\max} - f'_c)/(f_{\max} - \bar{f}), & f'_c \geq \bar{f} \\ k_3, & f'_c < \bar{f} \end{cases} \quad (1)$$

여기서 f'_c 은 유전자 교차가 이루어지기 전에 부모가 갖고 있는 값이다. \bar{f} 는 모집단의 평균값이고 f_{\max} 는 모집단의 최대값이다. 그리고 $k_1=1$, $k_3=0.5$ 이다. 돌연변이율은 1점 교차와 2점 교차 후에 얻어지는 고정된 상수값이다. 이 연구에서는 수렴 속도를 높이기 위하여 학습 준비 단계에서는 2점 교차를 이용한다. 그리고 그룹에서 최대 적합 가능 값이 다른 어떤 값보다 더 클 때에는 1점 교차를 적용한다. 이와 같이 개선하면 유전자 알고리즘을 이용한 탐색에서 더 높은 속도 효율을 볼 수가 있다.

다층 신경망의 네트워크 가중값들을 최적화하기 위한 유전자 알고리즘을 이용한 각각의 신경망 학습 단계는 다음과 같다.

알고리즘 2.1: 개선된 유전자 알고리즘:

단계 1. 입력자료/출력자료 설정:

$$u_p/u_d, \quad p=1,2,\dots,n, \quad d=1,2,\dots,n.$$

단계 2. 유전자 알고리즘의 모수 초기화.

단계 3. 주어진 모집단 개수 n 에 대한 모집단 초기화.

단계 4. 모집단의 유전자 연결 개수 및 길이 그리고 강도를 설정하기 위한 신경망의 초기 가중값 생성.

단계 5. 모집단의 각 유전자 연결을 해석하고, n 개 그룹의 네트워크 입력값들 $(y_p)_j$, $j=1,2,\dots,n$ 의 각 가중값에 일치하는 n 개 그룹의 네트워크 출력값들을 계산한다.

단계 6. 신경망의 목적함수를 정의하고 그 결과를 유전자 알고리즘에 적합하도록 변경시킨다.

단계 7. 유전자 알고리즘 공간 내에서 선택 실행.

단계 8. 선택된 유전교차, 돌연변이, 모수들의 계산.

단계 9. 주어진 효율이 만족될 때까지 단계 5에서 단계 9까지를 반복하여 최적의 가중값 그룹을 구한다.

2.2 개선된 신경망 알고리즘

본 논문에서 제시하는 개선된 역전파 신경망 학습 과정의 전체적인 흐름도는 그림 2와 같다. 그림 2에서 가중값은 설정된 각 구간에서 발생시킨다. 발생된 가중값을 이용하여 네트워크를 학습시키고, 주어진 허용오차를 만족할 때까지 네트워크를 단순 모형에서부터 점차로 확장시키는 학습과정을 통하여 일반화 네트워크를 구축해 나간다 [7].

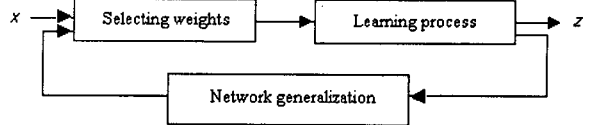


그림 2. 개선된 신경망 학습과정.

학습과 모형 확장이라는 두 가지 단계의 학습과정은 모든 패턴이 함께 학습될 때까지 반복한다. 이와 같이 일반화 네트워크를 위한 은닉노드의 개수가 최소 개수로 확장될 수 있고, 궁극적으로 해에 수렴될 수 있는 신경망을 구할 수 있도록 개선한 학습과정의 개요는 다음과 같다.

알고리즘 2.2: 개선된 신경망 학습과정:

입력 : 출력 패턴은 각각 입력 패턴 N_p 와 연결된다.

출력 : 주어진 허용오차 ϵ 을 만족하고 가능한 최소 개수의 은닉노드로 학습된 네트워크.

단계 1: 먼저 $l=1$ 로 설정한다. 학습 자료로부터 한 개의 패턴을 선택한다. 허용 오차 ϵ 를 만족하기 위해 선택된 패턴을 이용하여 은닉노드 하나로 신경망을 학습한다. 초기 가중값은 주어진 설정범위 내에서 발생된다.

단계 2: 만약 $l < N_p$ 가 만족되면 특정 범주에 맞는 다음 번 패턴을 선택한다. 그리고 $l=l+1$ 로 설정하고 단계 3으로 간다. 그렇지 않으면 끝.

단계 3: 만약 학습 알고리즘이 ϵ 을 만족하는 범위 안에서 $E(l)$ 을 감소시킬 수 있으면 단계 2로 간다. 그렇지 않으면 단계 4로 간다.

단계 4: 최종적으로 변화된 신경망의 가중값을

저장한다. 은닉노드의 개수를 하나 증가시키고, 새롭게 변화된 모형에 대한 초기 가중값을 주어진 설정범위 내에서 다시 설정한다. 단계 3으로 간다.

적용된 학습 알고리즘의 단계 3은 LM(Levenberg-Marquardt) 알고리즘을 이용한다.

3. 실증분석

다음과 같은 비선형 함수를 추정하기 위한 네트워크를 학습한다고 하자.

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

식 2에서 11개의 모의 학습자료 x는 구간 [-2.5, 2.5]의 균등분포에서 발생되며, 발생된 목표값 f(x)는 평균 0, 분산 0.005인 정규분포를 따르는 오차를 포함하고 있다.

먼저 개선된 유전자 알고리즘을 이용한 모의실험을 위한 주요 실행 단계는 다음과 같다.

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

초기 가중값은 균등분포 U(0, 1)를 따르는 확률난수(RND)를 이용한 식 20×(RND-0.5)에서 주어진 구간 안에서 발생시킨다.

단계 2. 위에서 적용한 네트워크에 대한 가중값의 개수는 8이고 각 가중값의 강도는 10비트이다. 따라서 각 개체의 유전자 스트링의 길이는 L = 10×8 = 80 비트이다.

단계 3. 부모 모집단의 개수는 50이다. 그리고 돌연변이율은 P_m = 0.03이다. 채택한 값이 1.3보다 작은 경우에 2점 돌연변이를 이용한다. 채택한 값이 1.3보다 큰 경우에 1점 돌연변이를 이용한다.

단계 4. 신경망 학습에서 오차 최적화를 위한 적합식은 다음과 같이 최소제곱법을 이용하고 이를 통해 목표값과 신경망 출력값과의 최적의 탐색을 수행한다.

$$f = \frac{1}{2J} = \frac{1}{\sum_{k=1}^4 [d(k) - y(k)]^2} \quad (3)$$

여기서 J는 목적함수이고 y는 네트워크의 출력값이다. d는 목표값이다

단계 5. 유전자 알고리즘의 실행은 논문에서 제시하는 개선된 유전자 알고리즘을 이용하여 처리한다.

역전파 신경망 알고리즘을 이용한 실증분석에서 이용하는 구조 또한 3층 신경망 구조이다. 여기서 학습의 중단시점으로 설정한 최종 학습 허용오차는 ε = 0.01이고, 초기 가중값의 범위는 |c| < w이다. 여기서 w = 0.001~10이다. 또한 오차 판단기준은 MSE이다.

본 연구에서 고려하는 비선형함수를 이용한 신경망 학습의 결과는 표 1과 같다. 이 결과에서 볼 수 있듯이 개선된 유전자 알고리즘 기반의 신경망은 인위적인 비선형 함수에 거의 접근시킬 수 있었다.

즉 초기 가중값의 범위 조절을 통해 초기값을 작은

범위에서 발생시킨다는 전제 하에서 살펴보면, 전형적인 신경망 알고리즘을 적용할 때보다 유전자 알고리즘을 적용 시에, 학습량과 학습의 정도에서 더 효율적이고 좋은 성능을 보이고 있음을 알 수 있다.

표 1. 초기값의 발생 범위에 따른 학습결과

w	개선된 유전자 알고리즘		개선된 신경망 알고리즘	
	학습량	MSE	학습량	MSE
0.001	182	0.00755	1227	0.012571
0.01	210	0.00756	452	0.010102
0.1	173	0.00792	684	0.010874
1	179	0.00804	576	0.011545
2	183	0.00798	635	0.012221
5	1028	0.08113	701	0.012567
10	1317	0.07042	817	0.012983

4. 결론

본 논문에서는 개선된 유전자 알고리즘과 역전파 신경망 알고리즘을 이용하여 비선형 모의자료를 이용한 신경망 학습의 효과를 살펴 보고, 이를 통한 수렴의 정확도 및 학습 속도 등을 비교해 보았다. 유전자 알고리즘의 개선을 위해서는 유전자차의 적용방법을 개선시킨 자동적응조정 알고리즘을 이용하였으며, 역전파 신경망 알고리즘의 개선을 위하여 초기 가중값의 발생 범위를 적절하게 변화 적용시키면서 신경망 학습시의 수렴의 정도를 높여주고, 최소한의 단순 구조를 갖는 일반화 네트워크로 추정할 수 있게 하였다.

개선된 유전자 알고리즘 기반의 신경망 제어기의 구조는 보다 단순한 구조를 갖고 있으며, 초기 가중값의 범위 조정의 전제 하에서, 신경망 학습 능력과 더불어 유전자 알고리즘의 보다 빠른 수렴능력과 학습의 정도를 보여주고 있다.

참고문헌

- Holland, J.H., 'Emergence: From Chaos to Order', Oxford University Press, 2000.
- Mahony R.E. and Williamson R.C., "Prior Knowledge and Preferential Structures in Gradient Descent Learning Algorithms," Journal of Machine Learning Research, Vol. 1, pp. 311-355, 2001.
- Gen M. and Cheng R., Genetic Algorithms and Engineering Design, Wiley-Interscience, 1997.
- Langdon W.B. and Poli R., Foundations of Genetic Programming, Springer, 2002.
- Shengsong Mei, Zhuo Huang, "A Neural Network Controller Based on Genetic Algorithms," IEEE International Conference on Intelligent Processing Systems, Vol. 2, pp.1624-1628, 1997.
- Mitchell M., An Introduction to Genetic Algorithms: Complex Adaptive Systems, The MIT Press, 1998.
- Liu, D., Chang, T. S. and Zhang, Y., "A New Learning Algorithm for Feedforward Neural Networks," Proceedings of IEEE, International Symposium on Intelligent Control, Sep. 5-7, Mexico, pp. 39-44, 2001.