

가속신경회로망에 의한 암반의 물성 추정 연구

김 남 수¹⁾, 양 형 식²⁾

1) 전남대학교 자원공학과 박사과정

2) 전남대학교 자원공학과 교수

1. 서론

지하 구조물의 안정성 확보와 경제적인 시공을 위하여 상세하고 합리적인 암반분류가 필요하다. 설계 초기에는 제한적인 정보와 암반의 불확실성에 따라 암반분류의 신뢰도가 떨어진다. 이러한 불확실한 지질 정보를 근사하게 추론할 수 있는 방법으로서 인공지능(Artificial intelligence) 특히 인공신경망(Artificial neural network)이 있다.

신경망은 세포들간의 연결가중치(connection weight)를 조정하는 학습을 통해 얻어진 내적 지식을 이용하여, 새로운 상황에 일반화시켜 우리가 원하는 자료를 추론하며, 학습되지 않았거나 학습된 자료의 일부 파손, 왜곡된 자료의 입력시에도 연결 가중치에 저장된 자신의 내적 지식에 따라 적절한 결과를 출력한다. 그러나 현재까지 많이 사용되고 있는 신경망 학습방법인 오류 역전파 학습알고리즘(Error back-propagation learning algorithm, EBP)은 학습시간이 지나치게 길고, 지역 극소점에 빠질 가능성이 많은 등 학습알고리즘 자체의 한계점 때문에 학습속도의 개선을 위한 노력이 이루어지고 있다.

본 연구에서는 가속화된 신경망 학습알고리즘을 제시하고, 개선된 학습알고리즘을 이용하여 신경망을 학습시켜서 기존의 신경망 학습알고리즘과 학습시간, 인식률 등을 비교하였다.

2. 인공신경회로망

2.1 오류역전파 회로망

입력층, 출력층, 한 개 이상의 은닉층(hidden layer)을 가진 다층 perceptron이 제안되었다.(Figure 1) 다층 Perceptron에서는 효율적인 학습방법으로 신경망의 연결강도의 값을 구하는 일반화 델타 법칙(generalized delta rule)을 이용한 역전파(back-propagation) 학습 알고리즘을 사용하였다.

역전파 학습알고리즘에서는 두 단계의 학습이 진행된다. 첫 단계는 입력패턴이 신경망의 입력층에 주어졌을 때 각 층에 대해서 입력함수와 활성화함수를 이용하여 출력층에서의 출력값과 오차를 산출하는 전방향 진행(forward pass)이고, 두 번째 단계는 원하는 출력과 실제 출력과의 차이를 계산해서 이를 역전파 시키면서 각 층 사이의 연결강도를 조정하는 역방향 진행(reverse pass)이다. 연결강도 조정후 다시 입력하여 계산하면 처음 시도 때보다 적은 오차 값을 내게 되고 총 오차의 합이 정해진 오차의 기준치에 도달할 때까지 이 두 단계는 계속 반복 실행된다.

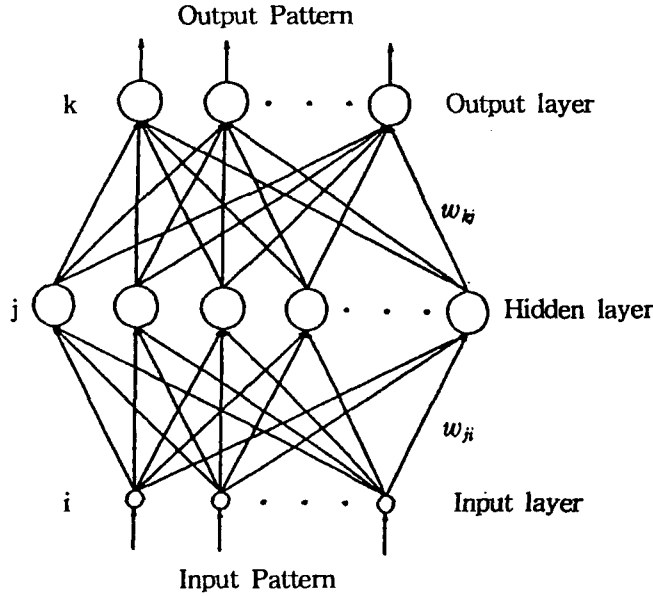


Figure 1. The structure of multi-layer perceptron

2.2 가속 신경회로망

학습 속도 개선 방법 중에 경험적인 방법으로 오차함수를 조정하여 보다 급격한 기울기를 갖도록 조정하는 연구가 있다. 오류 역전파 학습알고리즘에서 연결강도의 값을 변경하는 방법은 상위층에서 역전파된 오차를 해당 처리기 j의 출력 o_{pj} 에 대한 Sigmoid 함수의 미분값으로 곱하는 것이다.

$$\Delta_p w_{kj} = \eta \delta_{pk} o_{pj} = \eta (t_{pk} - o_{pk}) o_{pk} (1 - o_{pk}) o_{pj} \quad (1)$$

이때 미분값 $o_{pk}(1 - o_{pk})$ 는 o_{pk} 값이 0.0이나 1.0에 가까워지면 그 값이 0이 된다. 그런데, 이러한 경우에 오차가 더 클 수 있음에도, 이 오차의 적은 부분만을 역전파하여 전체적으로 학습속도를 느리게 한다. 이런 문제점을 해결하기 위해서 δ_{pj} 를 변경하는 새로운 알고리즘을 사용한다.

$$\begin{aligned} \Delta_p w_{ji} &= \eta \delta_{pj} o_{pi} \\ \delta_{pj} &= \sum_k \delta_{pk} w_{kj} (f'_j(\text{net}_{pj}) + 0.1) \end{aligned} \quad (2)$$

위와 같은 학습규칙의 변경이 타당하기 위해서는 해당 오류함수가 단조증가 또는 단조감소하여야 한다. 따라서 해당 오류 함수를 다음과 같이 유도한다.

$$\begin{aligned} \Delta_p w_{kj} &= -\eta \frac{\partial E_p}{\partial w_{kj}} = -\eta \frac{\partial E_p}{\partial o_{pk}} \cdot \frac{\partial o_{pk}}{\partial \text{net}_{pk}} \cdot \frac{\partial \text{net}_{pk}}{\partial w_{kj}} \\ &= -\eta \frac{\partial E_p}{\partial o_{pk}} \cdot f'_k(\text{net}_{pk}) \cdot o_{pj} \end{aligned} \quad (3)$$

이때 식 (1)과 식 (3)은 같아야 하므로 다음과 같은 식이 성립한다.

$$\frac{\partial E_p}{\partial o_{pk}} = -[(t_{pk} - o_{pk})(f'_k(\text{net}_{pk}) + 0.1)] / f'_k(\text{net}_{pk}) \quad (4)$$

그러므로 (4) 식의 양변을 적분하면 새로운 오류 함수를 얻을 수 있다.

$$E_p = \sum_k \left[\frac{1}{2} \cdot (t_{pk} - o_{pk})^2 + 0.1 \left[\frac{1}{2} \cdot (1 + t_{pk}) \ln \frac{1 + t_{pk}}{1 + o_{pk}} + \frac{1}{2} \cdot (1 - t_{pk}) \ln \frac{1 - t_{pk}}{1 - o_{pk}} \right] \right] \quad (5)$$

변경된 알고리즘의 Sigmoid 함수는 Figure 2와 같다.

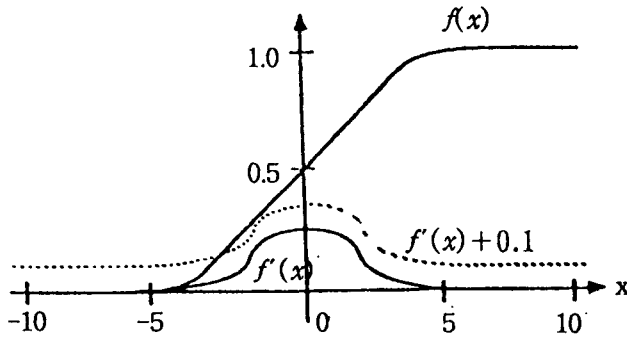


Figure 2. Altered Sigmoid function

3. 시뮬레이션

3.1 회로망의 성능 검증

가속화된 신경망과 일반적인 표준 EBP(general error back-propagation) 신경망의 학습속도와 인식률을 비교하기 위하여 두 가지 잘 알려진 XOR 문제와 문자인식(character recognition) 문제에 적용하였다. XOR 문제의 속도 비교에서 실험에 사용된 신경망의 구조는 3층 역전파 신경망으로 구성되었으며, 각각의 입력층, 은닉층, 출력층 처리기의 수는 $2 \times 2 \times 1$ 이다. 문자인식 문제에 사용된 신경망의 구조는 $64 \times 32 \times 5$ 이고 3층 역전파 신경망으로 구성되었다. 이때 학습율을 0.1, 0.3, 0.5, 0.7, 0.9로 변경하면서 각각의 학습율에 10번씩 실험하였으며 각 학습 실험에 대해 10번씩 인식실험을 하였다. 학습종료 조건으로서 허용오차를 0.1로 하였다.

Table 1은 XOR, Table 2는 문자인식 문제에 대하여 시스템 오차를 0.1로 설정하여 학습한 결과이고, Table 3은 시스템 오차를 0.001로 설정하여 XOR, 문자인식 문제를 가속 학습알고리즘으로 학습한 결과이다. 학습률을 변경시켰을 경우 각 조건의 학습에 사용된 반복횟수와 허용 오차에 도달하는 학습시간, 표준편차를 이용한 인식률을 나타내고 있고, 실험을 통해 나타난 결과는 다음과 같다.

첫째, 학습률이 커질수록, 표준 학습알고리즘에서는 학습의 반복횟수가 줄어들면서 학습시간이 단축되었다. 가속 학습알고리즘에서는 학습률 η 의 변화와 무관하게 학습시간이 모두 1 sec 이하였다.

둘째, 표준 학습알고리즘에서 XOR 문제는 입력층-은닉층-출력층 처리기의 수가 작아서 학습시간에 크게 구애받지 않지만, 문자인식과 같이 각 처리기의 수가 커지면 학습시간은 수배로 증가한다. 그러나 가속 학습알고리즘은 각 처리기의 수가 상당히 커져도 1초 이내에 학습이 이루어졌다.

셋째, 대체적으로 표준 학습알고리즘보다 가속 학습알고리즘의 인식률이 떨어진다. 즉 목적패턴과 출력패턴과의 오차가 더 큰데, 그것은 가속 학습알고리즘의 학습시간이 비교할 수 없을 정도로 작기 때문에 학습 허용 오차를 0.1에서 0.001로 줄여서 학습하더라도, 학습시간은 표준 학습알고리즘보다 짧은 시간이 소요되며 인식률도 향상시킬 수 있다.

Table 1. Comparison of general EBP and accelerated EBP for XOR problem (system error 0.1)

	General EBP			Accelerated EBP		
	epoch	learning time(sec)	recognition	epoch	learning time(sec)	recognition
$\eta = 0.1$	3120	10.7	0.024	98.3	<1	0.138
$\eta = 0.3$	987.3	3.3	0.027	60.1	<1	0.135
$\eta = 0.5$	562.0	1.8	0.024	53.5	<1	0.131
$\eta = 0.7$	421.6	1.5	0.024	88.6	<1	0.134
$\eta = 0.9$	328.1	1.0	0.023	64.4	<1	0.169

Table 2. Comparison of general and accelerated EBP for character recognition problem (system error 0.1)

	General EBP			Accelerated EBP		
	epoch	learning time(sec)	recognition	epoch	learning time(sec)	recognition
$\eta = 0.1$	1148.9	57.2	0.012	13.8	<1	0.069
$\eta = 0.3$	388.5	19.1	0.012	12.2	<1	0.074
$\eta = 0.5$	243.3	11.8	0.011	11.7	<1	0.064
$\eta = 0.7$	211.8	9.8	0.013	11.1	<1	0.065
$\eta = 0.9$	178.3	7.6	0.012	10.4	<1	0.069

Table 3. Comparison of accelerated EBP for XOR and character recognition problems (system error 0.001)

	XOR			Character recognition		
	epoch	learning time(sec)	recognition	epoch	learning time(sec)	recognition
$\eta = 0.1$	105.2	1.1	0.015	16.6	<1	0.01
$\eta = 0.3$	102.8	<1	0.013	15.3	<1	0.01
$\eta = 0.5$	98.3	<1	0.013	14.4	<1	0.01
$\eta = 0.7$	98.0	<1	0.013	14.1	<1	0.01
$\eta = 0.9$	80.0	<1	0.014	13.2	<1	0.01

이상과 같이 표준 학습알고리즘보다 가속 학습알고리즘을 이용하면 학습의 반복횟수가 줄어들어서 학습시간이 단축되고, 짧은 시간에 더 작은 허용 오차를 이용하여 학습시키면 인식률을 향상시킬 수 있어서 가속 학습알고리즘은 신경망의 학습과 산출에 더 좋은 결과를 나타내었다.

3.2 암반물성 추정

다양한 암질과, 예측 불가능한 지질조건이 존재하는 암반 상황에서 한정된 지질 자료를 가지고 암반의 특성을 분류하고 암반의 파괴 특성을 예상하는데는 한계가 있다. 이와 같이 주관적이고 경험에 의한 암반분류 방식을 설계에 효율적으로 이용하고, 부족한 지질 데이터를 유추하며, 암반의 파괴거동을 추정하기 위해서, Lee와 Sterling(1992)은 실제 시공된 기존의 시공 데이터를 통해 학습이 가능한 신경망 모델을 개발하였다.

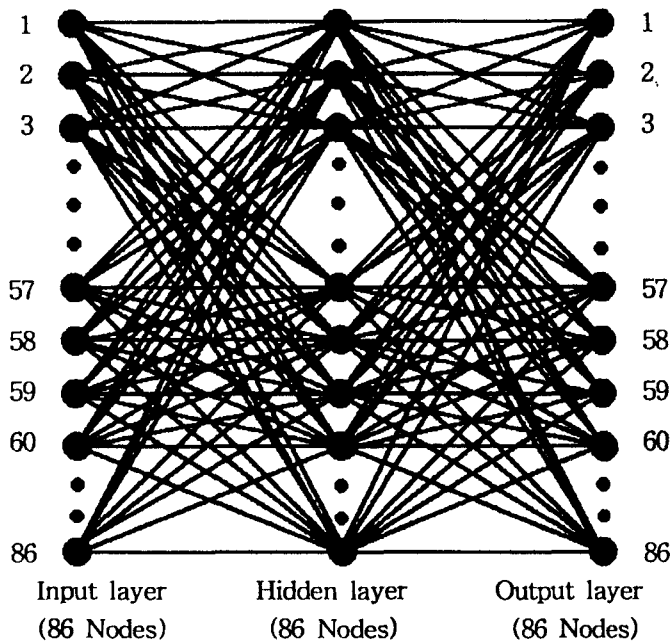


Figure 3. The structure of neural networks for experiment

그러나 이 신경망 모델은 학습시간이 지나치게 길어서 하나의 패턴을 학습시키려면 수십 시간이 소요되어 새로운 데이터를 추가하여 다시 신경망을 학습시키기가 곤란하였다. 본 연구에서는 가속화된 학습알고리즘을 이용하여 Lee와 Sterling의 신경망 모델과 동일한 문제에 적용하여 학습 속도와 인식능력을 비교하였다. 또 문 외(1993)의 연구 결과와도 비교하였다.

수치실험에 사용된 데이터베이스는 스칸디나비아 반도의 실제 시공 사례인 Cecil의 자료와, Barton 등의 자료를 기초로 Lee와 Sterling이 제작성한 데이터베이스이다. 신경망은 지도학습의 3층 역전파 학습알고리즘을 사용하였고, 입력패턴, 목적패턴은 86개의 항목으로 구성되어 은닉층을 포함, Figure 3와 같이 $86 \times 86 \times 86$ 으로 구성하였다.

본 실험에서는 하나의 패턴 오차를 0.0001로 가정하고 시스템 오차가 95개의 패턴오차를 합한 0.0095에 도달할 때 학습을 종료하였다. 그외의 모든 가정과 조건은 Lee와 Sterling 및 문 외의 방법과 같이 하였다.

신경망의 학습 결과, 기존의 연구들은 같은 시스템 기종에서 시스템 오차가 0.0095에 도달할 때까지의 학습시간이 35시간 이상이였으나, 본 실험에서는 단 10분 정도가 소용되어서 비교할 수 없는 학습시간의 단축이 이루어졌다.

Table 4. Comparison of learning time

	Lee & Sterling	Lee's	Author's
System configuration	PC 386/25MHz	IBM PC 486DX2-66MHz 16M RAM	IBM PC 486DX2-66MHz 8M RAM
Operating system		IBM OS/2 Ver 2.1	MS DOS Ver 6.22
System error	0.0001	0.0095	0.0095
Learning time	310 Hr	35 Hr	10 min

인식결과는 문 등의 연구 결과와는 거의 일치하였고, Lee와 Sterling과는 대체적으로 일치하였다.

이상과 같이 현장에서 계속적인 자료의 수집이 있을 때, 단지 입력패턴에 참가하여 신경망을 학습시키더라도 수분 안에 부족한 지질 자료와 파괴거동을 예측할 수 있기 때문에 가속된 신경망 학습알고리즘은 더욱 더 유용할 것으로 판단된다.

4. 결론

본 연구에서는 암반의 설계시 필요한 암반물성의 자료가 부족한 경우 인공지능에 의하여 자료를 예측하기 위하여 가속화된 신경망을 적용하였다. 기존의 신경망 모델은 학습에 수십시간이 필요하여 좋은 결과를 얻기 위해서는 대단한 인내가 필요하였다. 그래서 신경망의 활성화함수인 Sigmoid 함수를 경험적으로 수정하여 만든 가속화된 다층 역전파 학습알고리즘으로 학습한 결과 학습에 소요되는 시간이 수십시간에서 수분으로 단축되었고, 추론된 자료도 거의 일치하였다.

따라서 가속 인공지능망은 현장 자료의 수집으로 데이터베이스가 확장되더라도 빠른 시간에 쉽게 학습이 가능하며, 몇 개의 지질자료만 있어도, 합리적인 지질 자료와 파괴거동 등을 예측할 수 있을 것으로 기대된다.

참고문헌

- Cecil III O. S., 1970, *Correlation of rock bolt-shotcrete support and rock quality parameters in Scandinavian tunnels*, Ph.D. Thesis, University of Illinois, Urbana.
- Fahlman, S. E., 1989, Fast-learning variations on back-propagation : An empirical study, in *Proceedings of the 1988 Connectionist Models Summer School*, Morgan Kaufmann, 186-190.
- Moon, H.K. and Lee, C.W., 1993, Development of artificial neural networks for stability assessment of tunnel excavation in discontinuous rock masses, *J. Korean Soc. Rock Mechanics*, 3: 63-79 (in Korean)
- Lee C. and Sterling R., 1992, Identifying probable failure modes for underground opening using a neural network, *Int. J. Rock Mech. Min. Sci. & Geomech. Abstr.*, 29:49-67.
- Yang, H.S., Kim, N.S., Lee H.K. and Kim H.Y., 1995, Rock mass classification by rock mechanics expert system (ROMES), *J. Korean Soc. Rock Mechanics*, 5:41-47 (in Korean)
- Zhang Q., Song J. and Nie X., 1991, Application of Neural Network Models to Rock Mechanics and Rock Engineering, *Int. J. Rock Mech. Min. Sci. & Geomech. Abstr.*, 28:535-540.

Table 5. Nonzero output of neural networks for experiment

	Input	Output	Lee & Sterling's	Moon & Lee's
▲ OVERBURDEN				
2. medium	1	1.00	0.41 - 1.0	1.00
▲ INTACT ROCK STRENGTH				
5. sound	1	1.00	0.41 - 1.0	1.00
▲ AVERAGE SPACING/SPAN				
15. $0.1 < s/a < 1.0$	0	0.01	0.21 - 0.30	0.00
16. $s/a < 1.0$	0	1.00	0.41 - 1.00	1.00
▲ DISCONTINUITY TIGHTNESS				
20. open	0	0.76	0.41 - 1.00	0.54
▲ DISCONTINUITY PERSISTENCY				
22. continuous	0	0.99	0.41 - 1.00	0.95
23. discontinuous	0	0.06	0.01 - 0.10	0.51
▲ DISCONTINUITY TYPE				
24. joint	0	0.97	0.41 - 1.00	0.98
27. fault or shear zones	0	0.10	0.41 - 1.00	0.00
▲ DISCONTINUITY FILLINGS				
28. none	0	0.35	0.41 - 1.00	0.77
▲ DISCONTINUITY PLANENESS				
33. plane	0	0.89	0.41 - 1.00	0.96
▲ DISCONTINUITY ROUGHNESS				
37. smooth	0	0.33	0.41 - 1.00	0.17
▲ DISCONTINUITY DIP				
40. within 60	1	1.00	0.41 - 1.00	1.00
41. within 90	1	1.00	0.41 - 1.00	1.00
▲ DISCONTINUITY STRIKE				
42. in 30 deg	0	0.83	0.41 - 1.00	0.99
43. in 60 deg	0	0.59	0.41 - 1.00	0.86
▲ SHEAR ZONING OR FAULTING				
49. category E	0	0.81	0.00	1.00
▲ GROUND WATER CONDITION				
55. dry or minor inflow (<5 l/min)	0	0.98	0.01 - 0.10	0.99
56. medium inflow	0	0.00	0.41 - 1.00	0.00
▲ MODE OF INSTABILITY				
△ LOCATION				
59. roof	0	0.07	0.41 - 1.00	0.01
△ MODES				
76. dome-/vault-shaped formation	0	0.02	0.41 - 1.00	0.02