

제어상자를 이용한 단순 신경망의 개선된 학습과정

(An Improved Learning Process of Simple Neural Networks using the Controller Box)

윤 여 창[†]

(Yeochang Yoon)

요약 본 연구에서는 시계열자료를 예측하기 위해 적용한 $n \times n \times 1$ 신경망 구조에서 초기값의 시각적인 선택을 통한 개선된 학습과정을 제안한다. 적용된 Easton[1]의 제어상자는 시각적인 면과 실용적인 적용측면에서 다차원 구조를 논의하기에는 제한적이지만, 적은 개수의 은닉노드를 갖는 단순한 신경망구조에서는 초기 가중값들의 동적인 선택을 통하여 가능한 빨리 효과적인 학습이 이루어질 수 있게 할 수 있다. 신경망 학습의 오차 판단기준은 기존의 평균제곱오차(MSE)를 고려한다. 실증연구에는 모의생성된 ARMA(1,0) 자료와 담배생산량 자료를 이용한다.

Abstract In this work, we propose an improved learning process of $n \times n \times 1$ neural network system for time series forecasting. We apply Easton[1]'s controller box to assign initial weights for simple neural networks. The system provides the dynamic selection of initial weights on the controller box to get the learning off to as fast a start as possible. For the criterion of learning and validation, we consider MSE. The empirical studies carried out in artificial ARMA(1,0) and tobacco series.

1. 서론

시계열자료 분석이나 신경망에서 중요시되는 두가지 공통 사항으로, Lewbel[2]은 함수형태의 결정(신경망인 경우에 각 층별 노드의 개수, 층의 개수, 변환함수의 형태등을 결정하는 문제)과 모수의 추정문제(신경망의 가중값)를 들고 있다.

시계열자료에 대한 신경망 학습에서, 가중값을 구하기 위해 사용되는 초기값은 중요한 역할을 한다. 일반적으로 신경망의 학습은 특정 범위의 초기값에서 시작한다. Ripley[3]는 랜덤하게 선택된 몇 개의 초기값들을 각각 이용하면서, 변환함수를 통과한 값과 목표값 사이에 가장 잘 적합될 수 있는 초기값을 찾아서 학습을 실행하는 방법과, Kushner[4], White[5] 등의 전역최소값을 구하는 처리절차들을 이용하여 초기값의 문제를 해결할

수 있다고 하였다. Smith[6]에서는 전역최소값으로 수렴시키기 위해서 수많은 실행을 반복하는 시행착오를 겪고 있다. 따라서 초기값을 적절히 찾을 수 있는 방법이 제시되면, 효과적인 신경망학습이 될 수 있다.

본 논문에서는 초기값을 선택하기 위한 대안적인 방법으로써, 회귀진단을 위해 Easton[1]이 제시한 제어상자(controller box)를 이용한다. 제어상자는 다차원 공간을 2차원 평면에 함축적으로 표시하고 있다. 이를 신경망에 도입하면 차원을 무한히 늘릴수 없다는 제한이 있지만 동적으로 다차원의 적절한 초기값들을 적용시킬 수 있는 방안이 될 수 있다. 학습의 전 과정에 적용될 수 있는 이 방법은 목표값과 실제의 출력값과의 MSE를 작게 해주는 임의의 초기값들을 제어상자 위에서 동적으로 변화시면서 시행착오를 최소한으로 줄이고 적절한 초기값을 찾을 수 있게 된다. 여기서 선택된 초기값을 이용하여 신경망 학습을 하면, 전체 학습시간이 줄어들 뿐만 아니라 선형 시계열자료인 경우에 효과적으로 적합시킬 수 있다. 만약에 초기값이 잘못 선택되어 지역 최소값에 수렴되는 경우에는 학습과정 중이라도 제어상

· 이 논문은 2001년도 우석대학교 학술연구조성비에 의하여 연구되었음.

† 정 회 원 : 우석대학교 전산통계학과 교수
yoonyc@core.woosuk.ac.kr

논문접수 : 1998년 4월 2일

실사완료 : 2000년 11월 9일

자를 통하여 새로운 초기값을 동적으로 설정하여 학습 시간을 줄일 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 신경망이론의 기본 개념을 설명한다. 또한 제어상자에 의한 초기값 설정 방법과 은닉노드에 대한 논의를 한다. 3장에서는 선형 시계열모형으로 적합될 수 있는 담배생산량 자료를 이용하여 신경망 분석을 한다. 4장에서는 본 논문의 결론과 향후 연구방향을 기술하고 있다.

2. 신경망의 구조와 학습과정

2.1 신경망의 구조

은닉층과 출력층의 각 노드가 n 개의 입력노드를 갖는 신경망 구조는 $n \times n \times 1$ 신경망 구조는 (그림 1)과 같다.

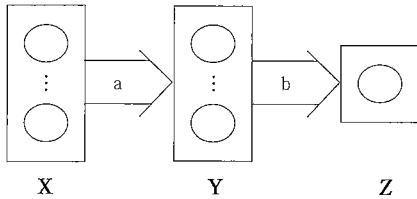


그림 1 $n \times n \times 1$ 신경망 구조

여기서 은닉노드는 $y = g(a_0 + \sum a_i x_i)$ 이고, 출력노드는 $z = g(b_0 + \sum b_j y_j)$ 이며, 다음 식(1)과 같은 변환함수를 이용한다.

$$g(u) = [1 + \exp(-u)]^{-1}, 0 < g(u) < 1 \tag{1}$$

2.2 학습과정

Smith[6]에서 학습과정은 변환함수를 통과한 값과 목표값 사이에 가장 잘 적합되는 가중값을 찾는 과정이며, 올바른 가중값을 어떻게 찾아야 하는가가 학습의 중요한 문제다. 이때 n 번째 예제에 대하여 출력값 (Z_n)과 목표값 (t_n)간의 학습오차는 다음과 같은 오차판단기준으로 설명된다.

$$MSE = \frac{1}{2N} \sum_{n=1}^N (z_n - t_n)^2 \tag{2}$$

여기서 N 은 예제들의 전체 개수이다. MSE는 특이값에 따라 변이가 심하므로 안정적이지 못하지만 통계분석이나 신경망에서 다루고 있는 오차의 판단기준으로 주로 사용되고 있다.

본 연구에서 적용한 역전파 알고리즘의 개요는 <표 1>과 같다. <표 1>에서 처리절차 forward는 변환함수를 현재 가중값에 적용하여 계산한다. 처리절차 back에서는 각 예제에 대한 가중값의 변화량을 결정한다. 처리

표 1 역전파 알고리즘의 개요

```

Repeat
{
    for(n in 1:examples)
    {
        forward()
        back()
    }
    changeweight()
}
    
```

절차 changeweight에서는 각 예제에 대한 가중값들을 누적시킨 결과를 이용하여 현재 가중값들을 각각 바꿔준다. 이와같은 과정을 반복하면서 변환함수를 통과한 Z_n 과 t_n 간의 오차가 최소화 되는 가중값들을 구하기 위하여, 각 노드에 대한 가중값의 변화가 한계값 이하가 될 때까지 반복하면서 적응적 학습률(adaptive learning rate)을 이용하여 학습을 진행시킨다. 오차의 판단기준인 MSE를 이용한 학습과정은 Smith[6]와 같다.

2.3 제어상자와 초기값

2.3.1 제어상자의 구성

Easton[1]은 동적 회귀진단기법을 이용하여 적절한 회귀모형을 탐색하고 있다. 이 연구에서는 최소제곱법에 의해 구한 p 차 다중회귀계수의 추정값 $b = (XX)^{-1}Xy$ 의 b_i 를, 오차의 한계범위 $b_i \pm \delta_i$ 안에서 변화시키면서 적절한 회귀모형을 탐색하고 있다. 예를들어 2^p 의 요인실험인 경우를 고려하자. 적절한 회귀모형의 탐색은 추정값의 허용된 오차한계 내에서 오차를 동적으로 변화시키면서 수행된다. 이때 p 개의 회귀계수에 대한 변화($b_1 \pm \delta_1, b_2 \pm \delta_2, \dots, b_i \pm \delta_i, \dots, b_p \pm \delta_p$)를 살펴보기 위해서는 p 차원 도형을 고려해야 한다.

Easton[1]의 제어상자는 (그림 2)와 같이 $p=3$ 인 3차원 도형안의 수많은 점들을 효과적으로 2차원에 표현시

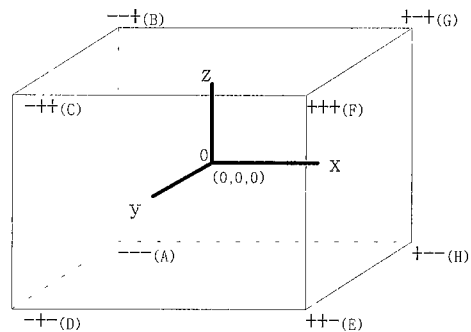


그림 2 3차원 도형

킬 수 있으며, 다차원공간을 2차원 표면에 표현시키는 방법은 다음과 같다.

먼저 (그림 2)의 시작 꼭지점을 A라고 할 때 6면체의 각 모서리 (A,B,C,D,E,F,G,H)를 따라가면서, 중점 O와 연결시킨 2³개의 이등변삼각형들은 △OAB, △OBC, △OCD, △ODE, △OEF, △OFG, △OGH, △OHA이다. 중점 O로 부터 각 모서리의 위치가 양의 방향에 위치하면 (+)기호, 음의 방향에 위치하면 (-)기호를 부여한다. 결국 p=n차원 도형은 (그림 3)과 같이 2ⁿ개의 이등변삼각형들로 구성된 제어상자로 2차원 평면화 된다.

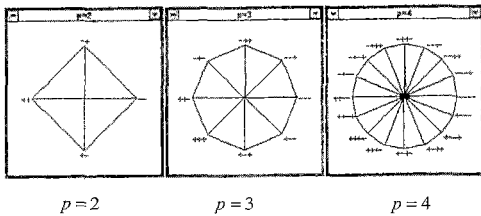


그림 3 p차원 도형에 대한 제어상자

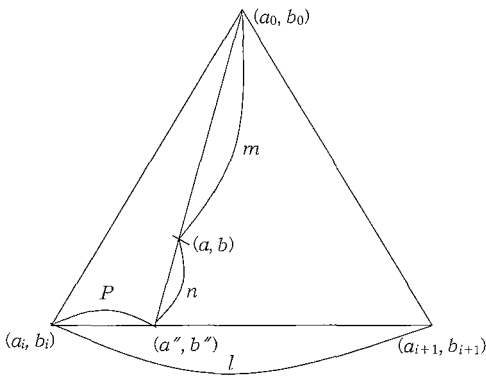


그림 4 제어상자의 첫번째 이등변삼각형

제어상자의 한 점을 다시 p차원 공간의 점으로 표현하는 방법은 다음과 같다. 먼저 p=2인 제어상자에서 선택된 점의 위치를 가중값으로 인식하는 방법은 다음과 같다. 2² 요인실험인 경우이므로 제어상자에 나타나는 이등변삼각형의 개수는 4개이고 중심점은 상대좌표 (0, 0)이다. 중심점에서 각 꼭지점까지의 상대거리를 k라고 할때 시작 꼭지점의 (-k, -k)좌표 에서부터 차례로 (-k, k), (k, k), (k, -k)에서 다시 시작 꼭지점으로 되돌아 오는 순환경로를 이등변삼각형들의 각 밑변으로 한다. 그러면 양변의 길이가 k이고 밑변의 길이는 2k인 이등

변삼각형이 2²개 생기며, 그 첫번째 삼각형은 (그림 4)와 같다.

(그림 4)에서 이등변삼각형의 각 꼭지점 (a₀, b₀), (a_i, b_i), (a_{i+1}, b_{i+1}),은 컴퓨터 화면상의 절대좌표로 주어지고, 선택된 점의 절대위치 (a, b)도 알고 있으므로, 절대좌표 (a*, b*)와 m, n, l의 길이는 다음과 같이 구할 수 있다.

$$\|m\| = \sqrt{(a-a_0)^2 + (b-b_0)^2}$$

$$\|n\| = \sqrt{(a-a^*)^2 + (b-b^*)^2}$$

$$\|l\| = \sqrt{(a_i-a^*)^2 + (b_{i+1}-b^*)^2}$$

$$\|l\| = \sqrt{(a_{i+1}-a^*)^2 + (b_{i+1}-b^*)^2}$$

따라서 시작 꼭지점 (a_i, b_i)와 절대좌표 (a*, b*)의 상대 길이를 c라고 하면 절대좌표 (a*, b*)의 상대좌표를 구할 수 있다.

$$l:l' = 2:c, c = \frac{2 \cdot l'}{l}, t = -k+c$$

그러므로 절대좌표 (a*, b*)의 상대좌표는 (-k, t)이다. 따라서 선택된 점의 절대위치 (a, b)의 상대좌표 (x, y)는 이등변삼각형의 꼭지점과 밑변을 이은 사선을 내분한 점이므로 다음 식(3)과 같다.

$$(x, y) = \left(\frac{-m \cdot k}{m+n}, \frac{m \cdot k \cdot t}{m+n} \right) \tag{3}$$

이와같은 과정을 p=3,4,...로 확장시키면, 제어상자에서 첫번째 이등변삼각형에서 선택된 절대좌표의 위치는 다음과 같이 상대좌표로 표현된 신경망 가중값으로 인식할 수 있게 된다.

p=3인 경우: $(x, y, z) = \left(\frac{-m \cdot k}{m+n}, \frac{-m \cdot k}{m+n}, \frac{m \cdot k \cdot t}{m+n} \right) \tag{4}$

p=4인 경우: $(x, y, z, w) = \left(\frac{-m \cdot k}{m+n}, \frac{-m \cdot k}{m+n}, \frac{-m \cdot k}{m+n}, \frac{m \cdot k \cdot t}{m+n} \right) \tag{5}$

p=3인 경우에 제어상자의 3차원 상대좌표로 표현된 가중값의 예는 (그림 5)와 같다. 이 그림에서 선택된 제어상자의 삼각형 △ODE의 상대좌표는 (-5.58, 8.81, -8.81)이다.

경험적으로 6차원이 넘는 경우에는 제한된 화면 크기에서 최소 64개의 개별 삼각형에서 가중값의 움직임을 고려해야 하기 때문에 시각적인 효율성이 떨어진다. 따라서 본 연구에서도 신경망의 가중값들은 시각적 효과를 고려하여 최대 6개까지 고려하게 된다. 이는 한 층의 노드 개수로 표현하면 최대 5개까지 허용함을 의미한다.

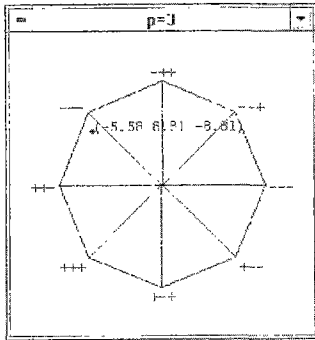
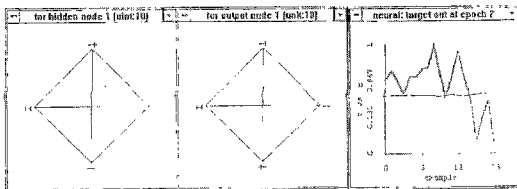


그림 5 $p=3$ 제어상자의 가중값

2.3.2 제어상자에 의한 초기값의 선택

$n \times n \times 1$ 구조의 신경망 모형에서 Easton[1]의 제어상자를 적용하여 보자. 여기서는 신경망을 통한 학습의 가장 단순한 경우로써, 은닉노드가 1개이고 입력노드가 1개인 경우인 $1 \times 1 \times 1$ 구조를 고려한다. 이때 가중값 a 는 (a_0, a_1) 의 2차원이고, 가중값 b 도 (b_0, b_1) 의 2차원이다. 이들 가중값들은 (그림 6)의 (a), (b)와 같이 각각 2차원 제어상자로 표시될 수 있다.

본 논문에서 학습의 시작단계로서 주어지는 초기값들은 $0.2 \times (X - 0.5)$, $X \sim U(0,1)$ 이다. 그리고 제어상자를 통해 선택할 수 있는 초기값의 허용 범위는 경험적으로 $|x| \leq 10$ 을 설정하였다.



(a) 가중값 a (b) 가중값 b (c) 적합되지 않은 상태

그림 6 $1 \times 1 \times 1$ 구조의 제어상자와 학습의 시작단계

제어상자에서 선택할 수 있는 초기값들은 다차원의 p 차 초기값 공간의 2^p 개 평면에서 표시된다. (그림 6c)의 가는 선은 시스템에서 현재 설정된 초기값이고 굵은 선은 목표값이다. (그림 7)은 직관적으로 선택한 초기값과 그 반영결과를 보여주고 있다. 이때의 초기값으로 학습을 하게 되면 초기값의 설정에 따른 수많은 시행착오를 줄일 수 있게 되고, 따라서 수렴하기까지의 전체 학습시간도 줄어들 수 있다.

따라서 단순 신경망모형을 이용한 학습에서 초기 가

중값의 설정을 단순히 smith[6]의 $0.2 \times (X - 0.5)$, $X \sim U(0,1)$ 와 Russo[7]의 $\pm(2.4/\text{은닉노드 개수})$ 를 이용하는 것보다 (그림 7)의 제어상자를 통한 초기값의 실시간 적용이 보다 효과적일 수 있다.

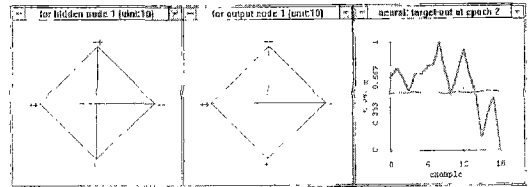


그림 7 $1 \times 1 \times 1$ 구조의 초기값 선택화면

2.4 은닉노드와 초기값

Haykin[8]은 신경망에서 은닉노드의 개수를 결정할 때, 입력노드 개수의 지수승으로 은닉노드 개수를 증가시키는 것이 효과적이라고 하였다. Smith[6]에서는 자료를 학습자료와 검증자료로 나눈 후에 교차 타당성의 검토를 하면서 은닉노드의 개수를 정하기도 한다. 이 방법은 수렴시키기 위한 모든 가능한 은닉노드 개수에 따라 신경망을 학습시켜야 하기 때문에 엄청난 학습시간이 필요하다. 그러나 많은 경우의 신경망 학습에서 은닉노드의 개수는 경험적으로 입력노드 개수의 2배수로 하고 있다. 여기서 최적 은닉노드의 개수는 현실문제에서 매우 어려운 과제로 남아있다[6].

시계열자료를 다루는 본 논문에서도 은닉노드를 충분한 개수만큼 증가시키는 것이 이론적으로는 가능하나, 앞 절에서 설명한 것처럼 현실적인 측면에서는 거의 불가능하다. 따라서 초기값에 더 큰 비중을 두고 학습을 하면 은닉노드의 충분한 개수에는 크게 영향을 받지 않고 효과적인 학습결과를 얻을 수 있다.

은닉노드의 개수에 따른 학습량의 변화를 살펴보기 위하여 <표 2>와 같은 가상의 입력자료를 이용하여 수렴되기까지의 변화되는 학습시간은 각각 (그림 8~9)와 같다.

(그림 8~9)의 왼쪽 화면은 학습 결과이고 오른쪽 화면은 MSE가 수렴되는 과정이다. 이 실험에서 수렴되기까지의 한계오차는 10^{-5} 이다. 학습시간은 은닉노드의 개수가 1일때 114이며 은닉노드의 개수가 2일때는 121이다.

표 2 가상의 입력자료

예제	1	2	3	4	5
입력값	1.1	2.1	3.1	5.1	6.1
목표값	2.1	3.1	5.1	6.1	6.1

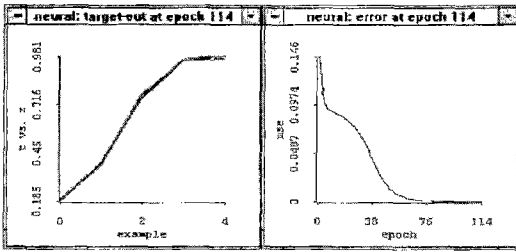


그림 8 1×1×1 구조일 때의 학습시간

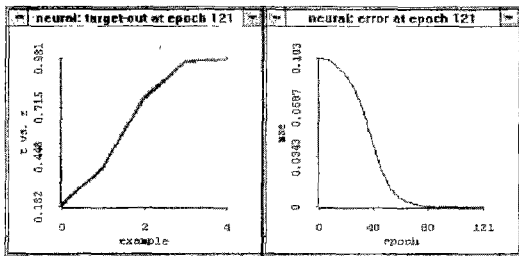


그림 9 1×2×1 구조일 때의 학습시간

(그림 8)에서 수렴속도가 빠른 이유는 주어진 자료가 1개의 은닉노드라도 설정될 수 있는 단순한 경우이기 때문이다. (표 2)와 같은 자료는 단조증가의 형태를 갖고 있으며 이는 신경망의 변환함수와 유사하다. 은닉노드의 개수가 3 이상인 경우의 학습에서는 은닉노드의 개수가 2일때와 유사한 결과를 보이고 있기 때문에 여기서는 제시하지 않았다.

(표 2)와 같은 가상의 입력자료에서 제어상자를 이용하여 초기값을 설정하기 위한 초기 화면은 (그림 10)과 같다. 제어상자 위에서 초기값을 변화시키면서 적합되는 과정을 살펴보면 (그림 11~15)와 같다. (그림 11)은 가중값 a 를 4사분면에, 가중값 b 를 1사분면에 위치시켰을 경우이다. 그리고 (그림 12~14)는 가중값 a 를 고정시켰을 때, 가중값 b 를 2사분면, 3사분면, 4사분면에 각각 위치시켰을 경우이다.

(그림 14)의 가중값에 대한 제어상자에서 제4분면에 위치한 초기값을 점차적으로 움직여 보면 (그림 15)와 유사하게 된다. 이 그림에서 선택된 $a=(0.86, -6.51)$, $b=(2.39, -6.15)$ 를 초기값으로 하여 27번 학습한 결과는 (그림 16)과 같고 $MSE=9.83E-4$ 이다. 여기서 변화된 가중값들은 다음과 같다.

지금까지 설명한 경우와는 다르게 <표 3>과 같이 좀 더 복잡한 $ARMA(1,0) : Z_t=0.8Z_{t-1} + \epsilon_t$ 모형으로 부터 생성된 시계열자료를 이용하여 학습시간과 초기값의 적용

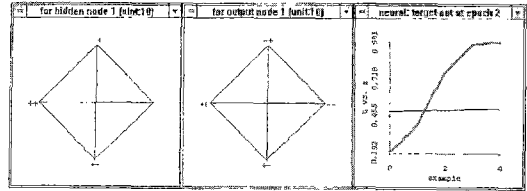


그림 10 초기값을 선택하기 위한 시작화면

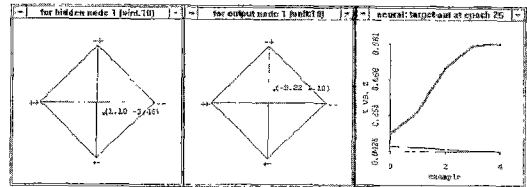


그림 11 초기값 선택에 따른 변화 과정(1)

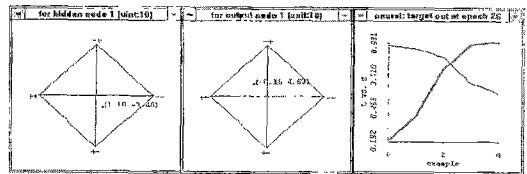


그림 12 초기값 선택에 따른 변화 과정 (2)

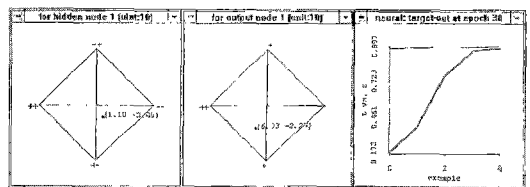


그림 13 초기값 선택에 따른 변화 과정 (3)

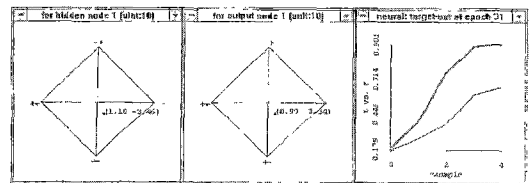


그림 14 초기값 선택에 따른 변화 과정 (4)

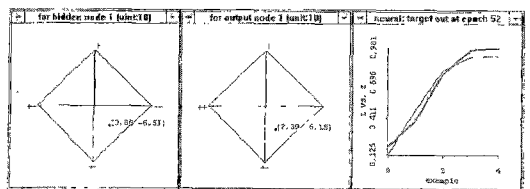


그림 15 초기값 선택에 따른 변화 과정 (5)

$$\hat{a} = \begin{bmatrix} 1.8674 \\ -3.9890 \end{bmatrix}, \hat{b} = \begin{bmatrix} 1.5871 \\ -3.5421 \end{bmatrix}$$

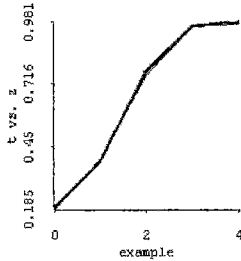
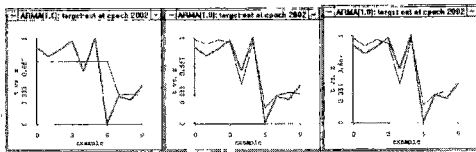


그림 16 은닉노드가 1개인 경우에 초기값을 적용한 학습결과

효과를 살펴보자. 모형은 t 시점의 자료가 $t-1$ 시점의 자료에만 영향을 받을 수 있도록 $1 \times n \times 1$ 구조로 신경망을 설정하였다. 은닉노드의 개수 n 을 1에서 3까지 변화시키면서 신경망 학습을 하고 학습시간이 2000번에서의 결과는 각각 (그림 17)과 같다.

표 3 시계열자료의 입력값

예제	1	2	3	4	5	6	7	8	9	10
입력값	.69	.89	.79	.87	.97	.62	1.0	0.0	3.2	2.8
목표값	.89	.79	.87	.97	.62	1.0	0.0	3.2	.28	.48



(a) n=1 (b) n=2 (c) n=3

그림 17 $1 \times n \times 1$ 구조의 학습결과

(그림 18)은 은닉노드가 1개인 경우에 (그림 17)이후의 학습과정을 계속 진행시켜서 학습시간이 100,000번인 순간을 보여주고 있는데, 학습시간을 충분히 증가시켜도 더 이상의 변화가 없다. 지금까지의 결과에 의하면 (그림 17)과 같이 은닉노드의 개수를 증가시키면 점차로 좋은 적합상태를 볼 수가 있지만 만족할 만한 결과를 주지는 못하고 있다. 이와같은 경우에는 초기값을 새롭게 변화시켜서 학습을 시작한다.

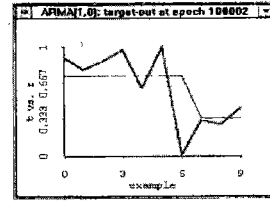


그림 18 은닉노드가 1개일 때의 최종 학습결과

(그림 19)는 은닉노드가 1개인 경우의 초기값을 설정하는 화면이다. 이렇게 하면 (그림 17)과는 전혀 다른 학습시간을 보이게 되며 은닉노드의 개수에도 영향을 덜 받게 된다.

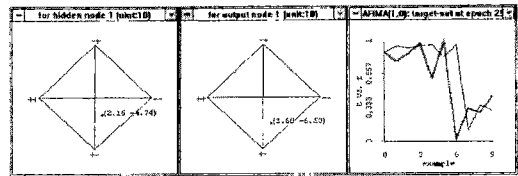


그림 19 (그림 18)에서 초기값이 선택된 경우

이렇게 선택한 값을 학습의 초기값으로 택한 경우에는 단 2100번의 학습시간에도 (그림 20)과 같은 결과를 얻을 수 있다.

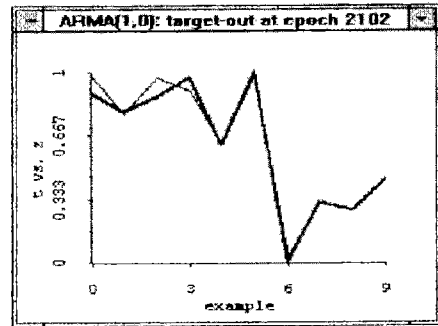


그림 20 선택된 초기값으로 학습한 결과

그러므로 지금까지 논의한 바와 같이 은닉노드의 개수가 복잡한 문제를 해결할 수 있을 정도로 충분하다면 신경망 학습의 초기 가중값 결정은 중요한 문제가 될 수 있다. 그리고 은닉노드의 개수가 부족한 상태에서는 초기값만으로는 적절한 결과를 유도할 수 없게 된다. 그러나 적합한 신경망구조 하에서 초기값을 시각적인 제어상자를 이용하여 선택하게 되면 학습효과를 높일 수 있는 하나의 방법으로 제시될 수 있다.

3. 예 제

신경망에 적용하기 위한 시계열자료는 Wei[9]에 나타나 있는 미국의 연간 담배생산량 자료이다. 1871년부터 1984년 까지의 담배생산량 자료를 이용하여 신경망 학습을 하고 그 예측결과를 전통적인 시계열 예측방법인 Box-Jenkins[10] 예측방법과 비교한다. 학습자료는 1871년부터 105년간 자료이고 검정자료는 1975년부터 1984년까지의 10년간 자료이다. 2×2×1 신경망 구조에서 2개씩의 순차적인 입력자료를 이용한다. 담배생산량 자료의 유형은 시간이 지남에 따라 추세를 보인다. 즉 평균과 분산이 일정하지 않은 비정상 시계열자료이다. 여기서 시차(lag)를 1, 5, 10으로 구분하였을 때에, 일반적인 t와 t-1시점 그리고 변동주기를 고려한 t와 t-5, t-10시점의 세가지 예제로 구분하여 SAS결과와 비교한다[11]. 통계적 모형식별과 모수추정의 결과는 2차 자기회귀모형(AR(2))을 따르며 다음 식(6)과 같다.

$$AR(2): Z_t = .1214 + 1.3284Z_{t-1} - .7112Z_{t-2} + a_t \quad (6)$$

(그림 21)은 담배생산량 자료를 이용하여 신경망학습을 100회 수행한 후의 10년 예측한 결과이다. 학습결과 변한 최종 가중값인 은닉노드의 가중값 와 출력노드의 a_{ij} 가중값 b_j 는 다음과 같다.

$$A = \begin{bmatrix} .819 & 1.352 \\ -2.266 & -1.469 \\ -.998 & -1.889 \end{bmatrix} \quad b = \begin{bmatrix} 2.432 \\ -2.989 \\ -3.193 \end{bmatrix}$$

이 가중값을 이용하여 10년 까지의 신경망 예측결과를 Box-Jenkins[10] 예측과 비교한 결과는 <표 4>와 같다.

<표 4>에서 신경망의 예측결과만 살펴볼 때, 시차가 1인 경우는 5년 예측기간을 제외한 기간에서, 시차가 5

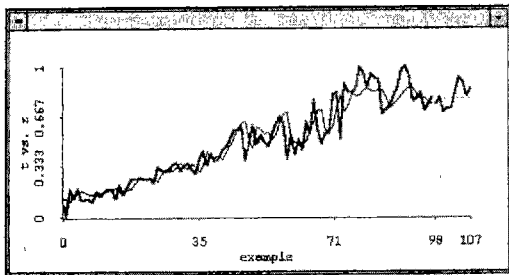


그림 21 100회 학습후의 10년 예측결과 : 원 자료는 굵은 실선으로 표시되고 검정자료는 98 이후이다. 가 는 점선은 신경망학습의 결과이다. 여기서 예측을 위한 변동주기는 10인 경우이다.

표 4 담배생산량 자료의 MSE : 세가지 변동주기에서 신경망 결과가 통계적 분석결과보다 작을 경우에는 *로 표시한다.

예측 기간	통계적 방법	2×2×1 신경망		
		lag=1	lag=5	lag=10
1년	0.0002	0.0001*	1.402e-6*	1.002e-6*
3년	0.0013	0.0008*	1.0011*	0.0014
5년	0.0015	0.0032	0.0026	0.0072
7년	0.0074	0.0049†	0.0054†	0.0052*
10년	0.0110	0.0056*	0.0061*	0.0095*

인 결과보다 MSE가 작고, 시차가 10인 경우 보다는 모든 예측기간에서 MSE가 작다. 신경망의 예측결과를 통계적 예측결과와 비교할 때, 신경망은 시차가 1,5인 경우에 5년 예측을 제외한 다른 예측기간에서 통계적 예측결과보다 MSE가 작다. 시차가 10인 경우에는 3,5년 예측을 제외한 다른 예측기간에서 통계적 예측결과보다 MSE가 작다. 시차가 1인 경우에는 단기예측과 장기예측에서 MSE가 더 작지만 중기예측에는 통계적 방법이 보다 효율적이다.

4. 결 론

본 연구에서는 단순 신경망모형을 이용한 신경망학습의 효율을 높이기 위한 방안으로, 제어상자를 통한 초기값의 효과적인 선택문제와 은닉노드의 개수문제를 논의 하였다.

초기값의 선택을 단순히 난수에 의존하지 않고 허용된 범위에서 시각적인 가중값의 적용효과를 살펴 보면서 실시간으로 선택하여, 학습시간의 단축을 유도하고 지역최소값에서 벗어날 수 있는 대안적인 방법을 제시 하였다. 은닉노드의 설정문제는 단순히 개수만을 증가시키지 않고, 적절한 개수에서도 효율적인 학습이 이루어질 수 있도록 하였다. 제안된 제어상자의 이용은 단순한 신경망구조 하에서 6차원 이하의 공간 점을 표시하는 대안적인 방법이라 할 수 있고 다른 분야에도 쉽게 적용 가능하다.

그러나 본 연구에서 이용한 제어상자는 다차원 공간을 제한적인 평면으로 나타내는 방법이기 때문에 모형 설정과 가중값 선택의 제약이 있을 수 있으며, 예제에서 살펴본 바와 같이 선형 시계열모형이라는 단순구조에서만 논의될 수 있는 방법이라는 한계가 있다. 이에대하여 시각적으로 더욱 개선된 선택 방법에 대한 향후 연구가 필요하며, 이를 이용하여 신경망 구조를 비선형 시계열

자료와 같은 확장된 문제에 적용해 볼 필요가 있다.

참 고 문 헌

[1] G.S.Easton, "A Simple Dynamic Graphical Diagnostic Method for Almost Any Model," Journal of the American Statistical Association, 89, pp. 201-207, 1994.

[2] A.Lewbel, "Comments on Artificial Neural Networks : An Econometric Perspective," Econometric Reviews, 13(1), pp.99~103,1994.

[3] B.D.Riply, "Neural Networks and Related Methods for Classification," Journal of Royal Statistical Society, B, 56(3),pp.409-456,1994.

[4] H.Kushner, "Asymptotic Global Behaviour for Stochastic Approximation and Diffusions with Slowly Decreasing Noise Effects: Glogal Minimization via Monte Carlo," SIAM Journal of Control Optimization, 47, pp.169-185, 1987

[5] H.White, "Learning in Artificial Neural Networks : A Statistical Perspective," Neural Computation, 1, pp.425-464,1988.

[6] M.Smith, "Neural Networks for Statistical Modeling," New York, Van Nostrand Reinhold,1993.

[7] A.P.Russo, "Neural Networks for Sonar Signal Processing," IEEE Conference on Neural Networks for Ocean Engineering, Tutorial Nn. 8, 1991.

[8] S.Haykin, "Neural Networks: A Comprehensive Foundation," Macmillan/IEEE Press, 1994.

[9] W.W.S.Wei, "Time Series Analysis : Univariate and Multivariate Methods," Addison-Wesley Publishing Co. 1990.

[10] G.E.P.Box, G.M.Jenkins, "Time Series Analysis: Forecasting and Control," Holden-Day,1976.

[11] 윤여창, "시계열자료에서 신경망이론의 적용", 성균관대학교 박사학위논문,1995.

[12] 윤여창, 허문열, "Neural Network을 응용한 Time Series", 통계와 신경망,한국과학기술원 인공지능연구센터/한국정보과학회 뉴로컴퓨팅연구회/한국통계학회 통계계산연구회추최 공동워크샵,연세대학교 소프트웨어 응용연구소 주관, pp49-69,1996.

[13] H.Tong, "Non-linear Time Series: A Dynamical System Approach." Oxford, Clarendon press, 1990.



윤 여 창
 1986년 성균관대학교 통계학과 졸업(학사). 1989년 성균관대학교 대학원 통계학과 졸업(석사). 1996년 성균관대학교 대학원 통계학과 졸업(박사). 1995년 ~ 현재 우석대학교 전산통계학과 조교수. 관심분야는 신경망, 인공지능, 시계열자료 분석.

분석.