

플랜트구조와 신경망에뮬레이터의 구조 및 학습시간과의 관계

배 창한, 이광원
아주대학교

A study on interrelation between the structure of a plant and the str neural network emulator and the learning rate

Paе. Chang-Han Lee, Kwang-Won
Aiou University

Abstract - Error-backpropagation has been used in the bulk of practical applications for neural networks. While an emulator, a multilayered neural network, learns to identify the system's dynamic characteristics. There is, however, no concrete theoretical results about the structure of a plant and the structure of a multilayered neural network and the learning rate.

The paper investigates the relation between structure of a plant and a multilayered network and learning rate.

Simulation study shows that the plant signal with a short period and a fast sam time is preferable for learning of the n network emulator.

1. 서 론

미지의 파라미터를 갖는 플랜트에 대한 제어 시스템의 정확한 수학적 모델링은 어려울 뿐만 아니라 플랜트의 구조 변화에 따라 제어도 쉽지 않다. 따라서 플랜트의 모델링이 힘들고 파라미터가 변하는 경우 스스로 학습하면서 적응제어할 수 있는 지능 제어가 요구된다. 이러한 방향으로 신경망 기법을 이용한 신경망 제어가 많이 연구되고 있다. 신경망은 많은수의 작고 간단한 연산 단위인 뉴런으로 구성된다. 각각의 뉴런은 많은 수의 다른 뉴런과 연결 되고 플랜트의 특성은 학습을 통하여 신경망의 연결강도가 조절됨으로써 이루어진다. 신경망은 일반적인 함수를 원하는 만큼 정확히 근사시킬 수 있고 병렬분산처리의 특성상 고속제어시도 사용이 가능하며 불확실한 플랜트 제어에 많이 이용된다. 특히 오차역전파학습을 이용한 다층의 신경망구조는 미지의 비선형동적시스템의 설계시 좋은 결과를 얻을 수 있다.[1][2] 그러나 설계시 필요로 하는 최 의 뉴런의 갯수와 뉴런의 층수를 명확히 결정할수

없고 부적절한 가중치의 설정으로 인한 조기포화 등의 여러가지 단점들도 있다.

본 논문에서는 플랜트의 구조를 여러 가지 형태로 변화 시키고, 또 신경망을 학습시키기 위한 플랜트 입력신호의 주기와 샘플링 시간을 변화해가며 플랜트에 따라 이를 학습하는 데 적당한 신경망에뮬레이터의 구조와 학습시간을 관찰하였다. 그리고 학습된 신경망에뮬레이터를 제어기로 사용하여 플랜트의 제어시스템을 구성하여 시뮬레이션하여 보았다.

2. 학습방법과 시간(학습특성)

2.1 신경망에뮬레이터의 학습방법

신경망에뮬레이터의 학습은 모멘텀항이 첨가된 개선된 오차 역전파 알고리즘에 의하였다. 원하는 출력과 실제 출력과의 오차를 최소로 하기 위해 최급 하강법을 사용하고 출력층의 오차를 역전파하여 중간층의 오차를 구한다음 다층신경망의 시냅스의 가중치를 변화시키는 형태의 학습방법을 사용하였다. 전이함수는

$$f(x) = k \frac{1 - e^{-\frac{\alpha}{k}x}}{1 + e^{-\frac{\alpha}{k}x}} \quad (1)$$

로 출력값의 구간을 $[-k, k]$ 로 하여 신호값의 포화치를 조절할 수 있고 α 로 기울기도 변화시킬 수 있다. 또한 전이함수의 도함수는

$$\frac{df(x)}{dx} = \frac{\alpha}{2} (1 - (\frac{f(x)}{k})^2) \quad (2)$$

의 형태로 전이 함수로부터 쉽게 구할 수 있다.

그림1과 같은 다층의 신경망에뮬레이터의 출력층 오차는

$$E(t) = \frac{1}{2} \sum_k [d_k(t) - o_k(t)]^2 \quad (3)$$

이며 출력층의 k번째 뉴런의 오차는

$$e_k^{[ol]}(t) = - \frac{\delta E(t)}{\delta I_k^{(ol)}(t)}$$

$$\begin{aligned}
 &= -\frac{\delta E(t)}{\delta o_k^{[0]}(t)} \frac{\delta o_k^{[0]}(t)}{\delta I_k^{[0]}(t)} \\
 &= -\frac{\delta E(t)}{\delta o_k(t)} f'[I_k^{[0]}(t)] \quad (4)
 \end{aligned}$$

여기서, $I_k^{[0]}(t)$ 는 출력층의 k번째 뉴런의 액티베이션(activation)이고 $o_k(t)$ 는 출력층의 k번째 뉴런 출력값이다. 출력층이 아닌 다른 s 층에서의 오차는

$$e_j^{[s]}(t) = [\sum_k e_k^{[s+1]}(t) w_{kj}^{[s+1]}(t)] f'[I_j^{[s]}(t)] \quad (5)$$

의 형태로 되어 출력층의 오차가 중간층으로 역전파되게 된다.

최급하강법을 이용한 신경망에물레이터의 가중치 변화는

$$\Delta w_{ji}^{[s]}(t+1) = -\eta \frac{\delta E(t)}{\delta w_{ji}^{[s]}(t)} \quad (6)$$

이며 체인룰(chain rule)을 적용하면

$$\begin{aligned}
 \frac{\delta E(t)}{\delta w_{ji}^{[s]}(t)} &= \frac{\delta E(t)}{\delta I_j^{[s]}(t)} \frac{\delta I_j^{[s]}(t)}{\delta w_{ji}^{[s]}(t)} \\
 &= -e_j^{[s]}(t) x_i^{[s-1]}(t) \quad (7)
 \end{aligned}$$

의 형태가 된다. 그러나 오차의 변화량이 항상 선형의 오차 공간상에 존재하지 않는 경우 가중치의 변화는 수렴하지 못하므로 학습률을 작게하고 전단계의 가중치의 변화량을 현재의 가중치의 변화에 사용해 저역통과필터의 역할을 첨가함으로써 개선된 오차역전파 알고리즘을 구현할 수 있다. 결국 가중치의 변화는

$$\Delta w_{ji}^{[s]}(t) = \eta e_j^{[s]}(t) x_i^{[s-1]}(t) + \beta \Delta w_{ji}^{[s]}(t-1) \quad (8)$$

$$e_j^{[s]}(t) = \frac{\alpha}{2} [\sum_k e_k^{[s+1]}(t) w_{kj}^{[s+1]}(t)] [(1 - \frac{f(I_j^{[s]}(t))}{k})^2] \quad (9)$$

여기서, η 는 학습률이며 β 는 모멘텀 상수이다.[3]

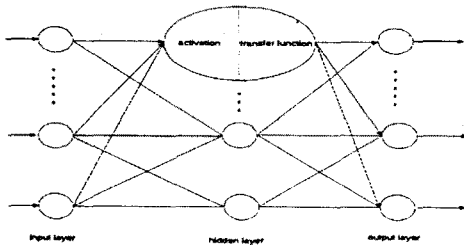


그림1. 다층신경망의 구조

2.2 플랜트 입력신호의 주파수 및

샘플링(sampling)시간과 학습시간과의 관계 신경망에물레이터를 학습시키기 위한 블록도는 그림2와 같다. 플랜트 입력신호는 그림3과 같은 구형파로하고 주기를 0.2초, 0.4초, 1초, 2초, 4초, 10다르게 하였다. 각각의 경우의 학습시간을 확인한 결과 그림 3에서 보듯이 0.2초 주기를 갖는 경우 가장 빠른 학습을 하였다. 또한 0.05초와 0.025초

샘플링시간(sampling time)을 갖는 경우 샘플링간(sampling time)이 짧은 0.025초의 경우가 0.보다 빠른 학습을 하였다. 결과적으로 빠른 주기의 입력신호와 짧은 샘플링시간(sampling time)을 갖는 신호로 학습을 할 경우 좋은 학습효과를 갖는다.

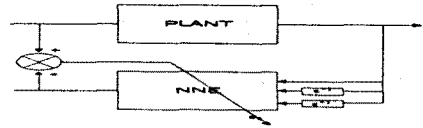


그림2. 신경망학습블럭도

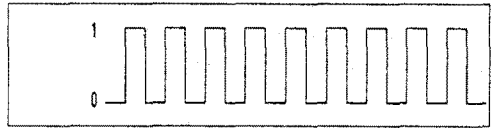


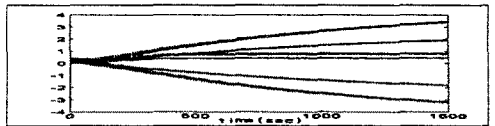
그림3. 입력신호



(a) 0.2초 주기



(b) 2초 주기



(c) 10초 주기

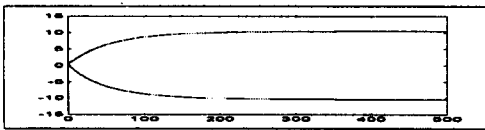
그림4. 입력신호에 따른 가중치의 변화

3. 플랜트의 구조와 신경망에물레이터의 구조

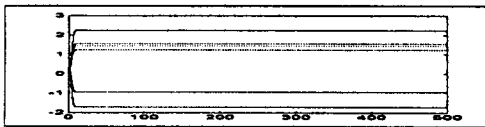
$1/s$, $1/(1+s)$, $1/(1+1.5s+s^2)$ 의 구조를 갖는 플랜트를 $2*1$ 구조, $2*2*1$ 구조, $3*5*1$ 구조로 이루어진 신경망에물레이터로 학습을 시켜 보았다. 플랜트의 출력값과 이를 지연시킨 출력값들을 신경망에물레이터의 입력값으로 하고 그에 따른 신경망에물레이터의 출력값과 플랜트의 입력값을 비교한 오차를 역전파시키게 된다. 그래서 신경망에물레이터의 출력은 플랜트를 학습시키기 위해 사용된 플랜트 입력파형을 추정하게 된다.

그림4에서는 적분기 학습시의 신경망에물레이터의 가중치의 변화를 나타낸다. $2*1$ 구조와 $2*2*1$ 구조의

신경망에물레이터의 경우 큰 학습률을 사용하여 학습의 속도를 높일 수 있었으며 $2*2*1$ 의 구조를 갖는 신경망에물레이터로 충분히 빠른 시간내에 학습을 하였다. 그림5에서는 $1/(1+s)$ 구조의 플랜트학의 경우로 $3*5*1$ 구조의 신경망에물레이터로의 학습시 낮은 학습률이 필요로 하고 따라서 학습시간은 늦어지게 된다. 또한 그림6에서 $1/(1+1.5s+s^2)$ 의 플랜트 학습시 $3*5*1$ 구조의 신경망에물레이터의 학습률은 아주 낮은 값을 사용하여야 한다. 결국 플랜트의 구조가 복잡해짐에 따라서 그리고 신경망에물레이터의 구조가 복잡해짐에 따라서 가중치를 수렴하도록 하는 학습률의 적정값은 작아지게 되고 그에 따라 학습시간이 오래 걸리게 된다. 그러므로 플랜트 구조에 비해 복잡한 구조의 신경망에물레이터를 이용한 학습은 오랜 학습시간을 요하게 된다. 결국 플랜트의 구조에 적합한 신경망에물레이터 구조의 설정이 가장 효과적인 플랜트의 학습을 이루게 한다.

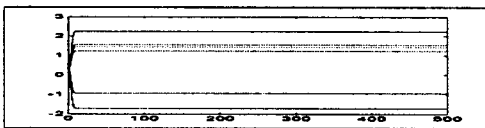


(a) $2*1$ 구조(학습률=0.2)

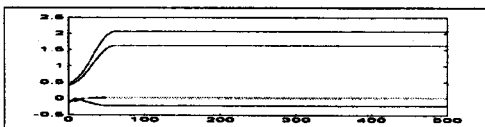


(b) $2*2*1$ 구조(학습률=0.2)

그림5. 적분기(1/s)의 학습



(a) $2*2*1$ 구조(학습률=0.1)



(b) $3*5*1$ 구조(학습률=0.01, 중간층과 출력층사이의 가중치)

그림6. $1/(1+s)$ 의 학습



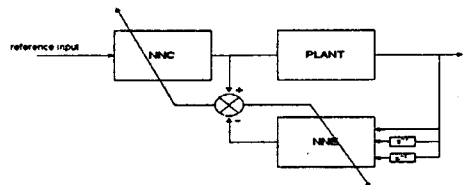
그림7. $1/(1+1.5s+s^2)$ 의 학습 ($3*5*1$ 구조, 학습률=0.001, 중간층과 출력층사이의 가중치)

4. 신경망에물레이터를 이용한 제어특성

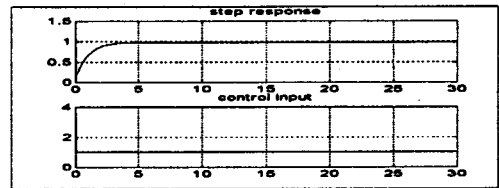
신경망에물레이터는 역플랜트 모델로 학습되며 학습된 신경망에물레이터를 신경망제어기로 사용함으로써 플랜트의 전달함수는 상쇄된다.

4.1 개루프제어

기준입력과 실제출력과의 비교없이 신경망제어기로의 입력신호 즉 기준신호를 계단신호로 사용하면 그림6에서처럼 신경망제어기의 내부의 각각의 뉴런들이 포화현상을 일으키게 되어 원하는 응답특성을 얻기가 어렵고 오프셋에러(offset error)도 생긴다.



(a) 개루프제어블럭선도

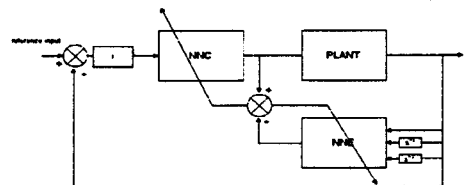


(b) 계단응답

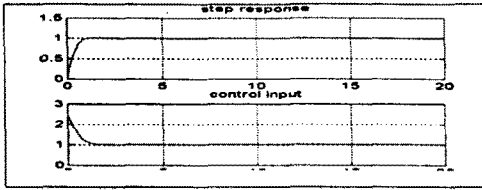
그림8. 개루프제어시

4.2 페루프제어

개루프제어시 발생하는 포화와 오프셋에러(offset error)를 막기위해 페루프제어를 사용하며 오차의 적분값을 신경망에물레이터의 입력신호로 사용한다. 신경망에물레이터의 구조가 복잡한 경우 좋은 제어특성을 얻음을 그림 9에서 확인할수 있다. 그러나 신경망에물레이터의 구조가 복잡하면 학습시간이 오래걸리게 되어 실질적으로 제어시스템의 성능은 저하된다. 신경망의 구조에 따른 계단응답은 그림10에 나타낸다. 학습횟수가 늘어남에 따라서 좋은 응답특성을 갖는다는 것이 그림 11에 나타난다.



(a) 페루프제어 블럭선도



(b) 계단응답

그림9. 페루프제어

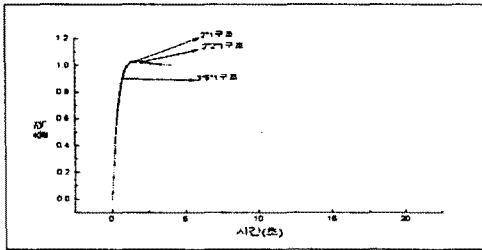


그림10. 신경망에플레이터의 구조에 따른 계단응답(600초 학습)

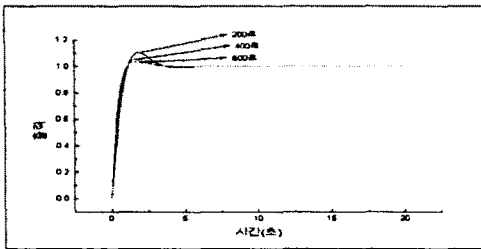


그림 11. 학습횟수에 따른 계단응답(2*1구조)

5.결론

플랜트의 구조와 신경망에플레이터의 구조 및 학습시간과의 관계를 플랜트 입력신호의 주기와 샘플링 시간을 변화시키면서 그리고 각각의 구조를 변화시키면서 살펴보았다.

짧은 주기와 빠른 샘플링시간을 갖는 입력신호의 사용이 플랜트의 학습에 유리하며 플랜트를 학습시키기 위한 신경망에플레이터의 구조가 플랜트에 비해 복잡하면 낮은 학습률을 필요로하게 되고 그에 따라 학습시간이 지연된다. 신경망에플레이터를 이용한 제어특성에서 복잡한 구조의 신경망에플레이터가 좋은 계단응답을 얻을 수 있으나 학습하는 데 걸리는 시간이 오래걸리기 때문에 임의의 플랜트를 학습하는데 높은 학습률을 갖으면서 간단한 구조의 신경망에플레이터를 필요로 한다.

(참 고 문 헌)

- [1] Demetri Psaltis, Athanasios Sider, and A.Yamamura, "A Multilayered Neural Net

Controller", IEEE Control System Maga vol 8, n2, pp.17-21, 1988

- [2] M. Saerens, and A. Soquet, "Neural con based on backpropagation", IEE proceedin vol.138, pp.55-62, 1991.

- [3] Madan M. Gupta Naresh K. Sinh Intelligent Control Systems", IEEE PR 1995