직교함수를 은닉층에 지닌 신경회로망에 대한 연구

권성훈*, 최용준*, 이정훈*, 유석용*, 엄기환*, 손동설**
*동국대학교 전자공학과, **유한대학 전자공학과
전화 : (02) 2260-3332 / 팩스 : (02) 2279-1798

The Study of Neural Networks Using Orthogonal Function System in Hidden-Layer

Sung Hoon Kwon*, Yong Jun Choi*, Jung Hoon Lee*, Seok Yong Yu*, Ki Hwan Eom*, Dong Seol Son**
*Dept. of Electronic Engineering Dongguk University,
**Dept. of Electronic Engineering Yuhan College
E-mail : kwon2@cakra.dongguk.ac.kr

Abstract

In this paper we proposed a heterogeneous hidden layer consisting of both sigmoid functions and RBFs(Radial Basis Function) in multi-layered neural networks. Focusing on the orthogonal relationship between the sigmoid function and its derivative, a derived RBF that is a derivative of the sigmoid function is used as the RBF in the neural network. so the proposed neural network is called ONN(Orthogonal Neural Network). Identification results using a nonlinear function confirm both the ONN's feasibility and characteristics by comparing with those obtained using a conventional neural network which has sigmoid function or RBF in hidden layer.

I. 서론

신경회로망은 인간의 뇌신경세포를 모방한 정보처리 모델로서, 1943년의 W. S. McCulloch와 W. Pitts의 신경회로망 모델과 1958년의 F. Rosenblatt의 퍼셉트론의 제안 등의 연구를 시작으로 1986년에 D. E. Rumelhart 등이 제안한 오차 역전파 알고리즘을 전기로 하여 많은 분야에서 지금까지도 활발히 연구가 되어오고 있다.

신경회로망은 비선형함수 구성기능, 범화능력에 의해 패턴분류, 인식제어 등 많은 분야에 응용되고 있다. 신경회로망의 제어에의 응용도 직접 역모델을 구성하여 조작량을 구하는 방법, 플랜트 동정을 하여 조작량을 구하는 방법 등이 있다. 이러한 문제에 사용되는 신경회로망은 입력층과 출력층 사이에 하나 이상의 은닉층을 갖고 있는 다층 신경회로망으로 비선형 사상능력이나 학습능력이 우수하다고 볼 수 있다.[1]

다층 신경회로망의 은닉층 활성화함수(active function)로는 선형이나 비선형 함수로 주어진 문제에 따라 해결될 수 있도록 선택되어 진다. 일반적으로 역전파 알고리즘을 이용하여 훈련되는 다층 신경회로망의 활성화함수로는 미분이 가능한 시그모이드 함수가 널리 쓰이고 있지만, 가우스 함수 등의 Radial Basis Function(RBF)를 사용하는 경우도 많이 있다. 하지만, 순조롭고 연속적인 응답을 나타내는 포화함수인 시그모이드 함수와 입력에 대해 부분적으로 응답하는 RBF는 각각 적용대상이 다르다. 따라서, 시그모이드 함수와 RBF는 사상대상에 따라 알맞게 선택되어지고 있다. 그러나, 사상대상의 사전정보를 전혀 모르는 경우에는 시그모이드 함수나 RBF의 하나의 함수를 사용하는 것보다는 시그모이드 함수와 RBF를 혼성시켜 사용하는 것이 타당한 근사결과를 얻을 수 있다고 기대된다.

이러한 혼성방식은 일본의 이시다(石田)에 의해 제

안되었으며, 여기서는 은닉층에 공존시킨 두 종류의 함수에 대해서 입력차원이 d일 때에 하나의 가우스 함수가 만드는 초타원 영역을 2^d개의 시그모이드 함수로 만드는 초평면으로 근사화할 수 있음을 나타내었다.[2]

제어시스템의 함수근사화 문제로서의 혼성방식은 시그모이드 함수와 가우스 함수를 활성화함수로 한 범함수 급수전개로 생각할 수 있지만, 단순히 함수를 공존시킨 경우에는 각 함수 사이의 관계가 반드시 분명한 것은 아니다.

본 논문에서는 시그모이드 함수와 시그모이드 함수의 도함수로 유도한 RBF의 직교관계에 착안하여 은닉층에 직교함수를 활성화함수로 갖는 신경회로망을 제안한다. 제안한 방식의 유용성을 확인하기 위하여 비선형 함수의 근사 시뮬레이션에 의해 사상능력을 검토한다.

II. 제안하는 신경회로망

은닉층 활성화함수로서 식(1)과 같은 시그모이드 함수를 갖고, 은닉층을 1개 갖는 구조의 다층 신경회로망(Sigmoid NN, SNN)을 생각한다.

$$f(x,a)=\frac{1.0-\exp\left(-\frac{x}{a}\right)}{1.0+\exp\left(-\frac{x}{a}\right)} \tag{1}$$

단, a는 함수형상을 변화시키는 파라미터로 네트워크를 구성할 때에 이 파라미터 a를 개개의 은닉층 유니트에서 독립적으로 조정하는 것으로 사상특성을 개선할 수가 있다. 이 SNN으로의 입력벡터를 $\mathbf{I}(k)$, 입력층과 은닉층 사이의 가중치를 $\mathbf{W}_s^{hi}$, 은닉층과 출력층의 가중치를 $\mathbf{W}_s^{oh}$로 하면 신경회로망의 출력 $u_s(k)$는 식(2)와 같다.

$$u_s(k)=\sum_j\left\{W^{oh}_{s_j}\, f_j\, [\sum_i W^{hi}_{s_{ji}}\, I_i(k),\ a_j]\right\} \tag{2}$$

여기서, k는 샘플링 상수이고, 입력벡터 $\mathbf{I}(k)$에 문턱값 1을 포함시키고 있다. 이 SNN에 유도연산자 $\frac{\partial}{\partial I(k)}$를 작용시켜 유도한 네트워크는 은닉층 활성화함수에 다음 식으로 정의되는 시그모이드 함수 f의 1차도함수가 사용된다.

$$\Phi(x,a)=\frac{2.0\exp\left(-\frac{x}{a}\right)}{a\left[1.0+\exp\left(-\frac{x}{a}\right)\right]^2} \tag{3}$$

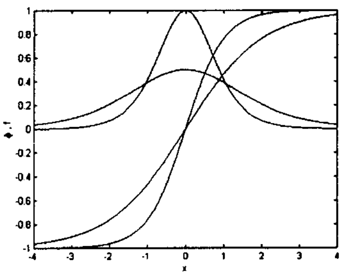

그림 1 함수의 형상
Fig 1. Function shape

그림 1과 같이 함수 Φ는 가우스함수 ($g(x,b)=\exp\left[-\left(\frac{x}{b}\right)^2\right]$)와 형상이 유사하므로 이것을 은닉층 활성화함수로서 이용한 네트워크를 RBFN(Derived RBFN, DRBFN)이라고 한다. 이 RBFN출력 $u_r(k)$를 다음 식으로 정의한다.

$$u_r(k)=\sum_j\left\{W^{oh}_{r_j}\, \Phi_j\, [\sum_i W^{hi}_{r_{ji}}\, I_i(k),\ a_j]\right\} \tag{4}$$

여기에서 $\mathbf{W}_r^{hi}$는 입력층과 은닉층 사이의 가중치, $\mathbf{W}_r^{oh}$는 은닉층과 출력층 사이의 가중치이다.[3]
네트워크에 의한 보다 일반적인 사상이라는 관점에서 한 종류의 함수에 의한 표현보다도 종류가 서로 다른 함수를 혼성시킨 네트워크를 사용하는 것에 의해 타당한 사상특성이 얻어진다고 생각된다. 따라서, 여기에서는 시그모이드함수 f와 유도RBF Φ를 공존시킨 직교함수 신경회로망(Orthogonal NN, ONN)을 제안한다.

제안하는 ONN에 의한 함수의 사상은 직교하는 기저함수를 이용한 범함수급수전개의 일종이라고 생각할 수 있다. 그림 2에 ONN의 구성법을 나타낸다. 본 논문에서는 그림 2와 같이 비선형처리를 하는 은닉층 앞에 버퍼층을 사용하여 시그모이드함수 f_i와 유도RBF Φ_i에 대한 입력을 동일한 값으로 하고, i에 대해서 함수형상 파라미터 a_i도 동일한 값으로 하는 것으로 직교관계를 유지하도록 한다. 이 경우, ONN의 출력 $u_o(k)$는 다음 식과 같이 주어진다.

$$u_o(k) = \sum_j \{ W^{ohf}_{o_j} f_j [\sum_i W^{hi}_{o_n} I_i(k),\ a_j] + W^{oh\Phi}_{o_j} \Phi_j [\sum_i W^{hi}_{o_n} I_i(k),\ a_j] \} \quad (5)$$

여기에서 $\mathbf{W}^{hi}_o$는 입력층과 은닉층 사이의 가중치, $\mathbf{W}^{ohf}_o$는 시그모이드함수 출력을 지닌 은닉층과 출력층 사이의 가중치, $\mathbf{W}_o^{oh\Phi}$는 유도RBF 출력을 지닌 은닉층과 출력층 사이의 가중치이다.

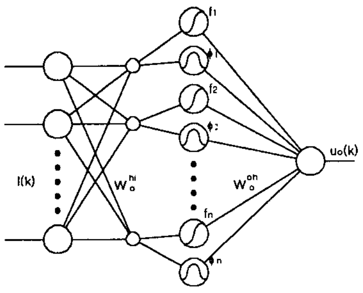

그림 2 혼성 신경회로망의 구성

Fig 2. Network architecture

III. 순방향 동정기에 의한 함수근사

ONN의 사상특성을 검토하기 위해서 그림 3의 순방향 동정기를 구성하여 1입력 1출력의 정적인 함수의 근사 시뮬레이션을 하여 SNN, RBFN과 비교 검토한다. 각 네트워크의 학습은 식(6)으로 정의되는 평가함수 $J(p)$를 최소화하기 위하여 일반화 델타규칙에 의해 식(8)에 따라 각 시행마다 가중치 $\mathbf{W}^{hi}_q$, $\mathbf{W}^{oh}_q$ $(q=o,\ s,\ r)$ 의 조정을 하는 학습형으로 한다.

$$J(p) = \frac{1}{2} \sum_{k=0}^{\rho} e^2(k) \quad (6)$$

$$e(k) = y_d(k) - u_q(k) \quad (7)$$

$$\boldsymbol{W}_q(p+1) = \boldsymbol{W}_q(p) - \eta_w \frac{\partial J(p)}{\partial \boldsymbol{W}_q(p)} \quad (8)$$

여기에서 p는 시행횟수, ρ는 1회시행당 총샘플링수, η_w는 학습계수, $y_d(k)$ 는 목표출력, $u_q(k)$ 는 네트워크 출력 $(q=o,\ s,\ r)$ 이다. 또한, 함수형상 파라미터 a도 최급강하법에 따라 식(6)을 최소화하도록 개개의 i에 대해서 조정하여 함수의 최적형상을 탐색한다.[4][5]

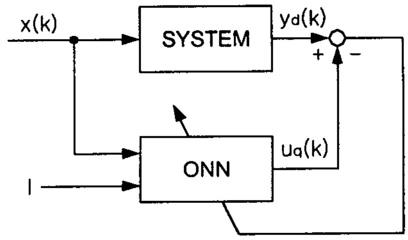

그림 3 순방향 함수동정기의 블록도

Fig 3. Block diagram of direct function identifier

$$a_i(p+1) = a_i(p) - \eta_a \frac{\partial J(p)}{\partial a_i(p)} \quad (9)$$

여기에서 η_a는 학습계수이다. 파라미터 a의 초기값은 유도RBF가 입력 0일 때에 출력 1인 $a=0.5$ 로 했다. 시뮬레이션에서는 1회시행당 총샘플링수를 100, 네트워크 구조는 입력층 2 $(I^T(k)=[x(k)\ \ 1])$, 출력층 1로 했다. ONN에서 은닉층 유니트수를 m으로 했을 때, 네트워크의 가조절 파라미터수가 5m이 된다. 한편, SNN 및 RBFN에서 은닉층 유니트수를 n으로 했을 때, 가조절 파라미터수는 4n이 된다. 따라서, 각 네트워크에서 가조절 파라미터수가 같아지도록 은닉층을 결정하며 여기서는 m=4, n=5로 하고 목표값은 식(10)으로 하였다.

비선형 네트워크의 학습과정에는 가중치의 초기값이 영향을 미치기 때문에, 서로 다른 네트워크의 학습을 비교하기 위해서는 이 초기값문제에 조건을 달 필요가 있다. 이를 위해 각 네트워크의 초기값을 다음과 같이 설정하고, 네트워크 출력의 초기상태를 맞추도록 했다.

(1) ONN을 [-1.0 1.0]의 범위에서 랜덤하게 초기화하고, 입력 $x(k)$를 가한 경우의 SNN 출력 $u_{oinit}(k)$를 준비한다.

(2) SNN(RBFN)을 [-1.0 1.0]의 범위에서 랜덤하게 초기화하고, 입력 $x(k)$를 가해 $u_{oinit}(k)$를 교사신호로 하여 학습한다.

(3) 학습수렴 후의 가중치를 SNN(RBFN)의 초기값으로 하여 목표함수의 동정에 사용한다. ONN의 초기값은 (1)에서 사용한 가중치를 사용한다.

VI. 시뮬레이션

다음의 함수를 목표값으로 하여 시뮬레이션을 했다.

$$y_d(k) = x(k)\cos x(k), \quad [-\pi/2 \le x(k) \le \pi/2] \quad (10)$$

그림 4는 ONN의 학습과정을 식(6)의 평가함수 J에 의해 나타낸 것으로 실선은 가중치의 학습과 함수형상의 탐색을 동시에 한 경우, 점선은 함수형상을 고정하고 가중치만을 학습한 경우이다. 어느 경우에서도 시행이 진행됨에 따라 평가함수 J는 감소하고 있다. 그림 4의 비교로부터 가중치의 학습과 함수형상의 탐색을 동시에 하는 것으로 평가함수를 작게할 수 있다는 것을 알 수 있다.

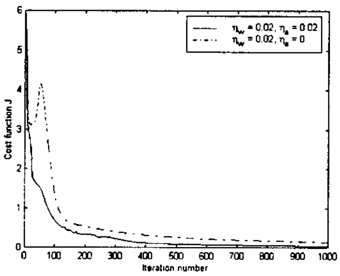

그림 4 ONN에 의한 동정기의 학습과정
Fig 4. Learning process of direct function identifier using ONN

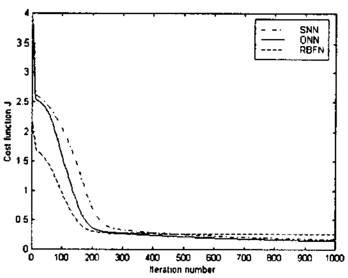

그림 5 각 네트워크에 의한 동정기의 학습과정
Fig 5. Learning process of direct function identifier

그림 5는 ONN, SNN, RBFN의 학습과정을 평가함수 J로 비교한 것으로, 모두 가중치와 함수형상의 탐색을 동시에 하고 있다. 각 네트워크 모두 시행이 진행됨에 따라 학습은 수렴하고 있다. 그림 5에서는 RBFN의 고속학습과 시그모이드함수 네트워크의 학습지연 특징이 잘 나타나 있고, ONN에 의한 학습속도는 그 중간정도인 것을 알 수 있다. 즉, 학습종료후(ρ=1000)에서 평가함수값의 크기를 비교하면, $J_{ONN} \fallingdotseq J_{SNN} < J_{RBFN}$으로 되어있다. 따라서, 학습속도에서는 SNN에 대해서, 사상특성에서는 RBFN에 대해서 ONN이 유리하다는 것을 알 수 있다.

V. 결론

제어대상의 사전정보가 미지인 경우의 사상 및 제어를 위하여 은닉층에 직교함수를 사용하는 직교함수 신경회로망을 제안하였다. 제안하는 직교함수 신경회로망은 은닉층 앞에 버퍼층을 사용하고 은닉층에는 시그모이드 함수와 유도 RBF를 이용한 직교함수를 사용하였다. 제안하는 직교함수 신경회로망의 사상특성을 확인하기 위하여 시뮬레이션 한 결과, 미지의 환경에서 활성화함수를 선택할 필요없이 ONN을 사용함으로서 좋은 성능을 얻을 수 있었다.

참고문헌

[1] 임영도, 이상부, "퍼지 · 신경망 · 유전진화", 도서출판 영과일, 1997.

[2] 石田和子, 亀山啓輔, 小杉幸夫, RBFを含むニューラルネット中間層の適應的構成法, 信學技報, NC95-112, pp. 83-90, 1996.

[3] N. K. Bose, P. Liang, "Neural Network Fundamentals with Graphs, Algorithms, and Applications", McGraw-Hill, 1996.

[4] K. S. Narendra, K. Parthasarthy, "Identification and Control of Dynamical Systems Using Neural Networks", IEEE Trans, Neural Network, vol. 1, No. 1, March 1990, pp. 4-27.

[5] J. G. Kuschewski, S. Hui, S. H. Zak, "Application of Feedforward Neural Networks to Dynamical System Identification and Control", IEEE Trans, Control System Technology, vol. 1, pp. 37-49, 1993. 5.