

論文95-32B-3-14

다층신경회로망을 이용한 선형시스템의 식별

(Linear System Identification Using Multi-layer Neural Network)

趙奎翔*, 金庚基*

(Gyu-Sang Cho, and Kyong-Gi Kim)

요약

본 논문에서는 역전파 알고리즘으로 학습된 다층 신경회로망을 이용한 새로운 방식의 선형 시스템의 파라미터 식별법을 제안한다. 선형시스템의 시스템 파라미터는 $\partial \dot{x}(t)/\partial x(t)$ 로 표시되고, 입력 파라미터는 $\partial \dot{x}(t)/\partial u(t)$ 로 표시되므로, 미지의 선형시스템을 학습한 신경회로망의 입력에 대한 출력의 미분관계를 뉴런의 출력값과 웨이트의 함수로 나타내어 파라미터값을 추정한다. 이 방법에 대한 관계식을 유도하며, 2층 신경회로망과 3층 신경회로망에 적용한 시뮬레이션 결과를 통해 제안한 방법의 타당성을 제시한다.

Abstract

In this paper, a novel approach is proposed which identifies linear system parameters using a multilayer feedforward neural network trained with backpropagation algorithm. The parameters of linear system can be represented by $\partial \dot{x}(t)/\partial x(t)$ and $\partial \dot{x}(t)/\partial u(t)$. Thus, its parameters can be represented in terms of the derivative of output with respect to input of trained neural network which is a function of weights and output of neurons. Mathematical representation of the proposed approach is derived, and its validity is shown by simulation results on 2-layer and 3-layer neural network.

I. 서론

신경회로망은 임의의 함수를 근사화할 수 있는 능력이 있어 비선형성을 갖는 플랜트의 동적특성을 학습할 수 있다^[1,2]. 이러한 이론적인 기반으로 제어시스템에 도입이 되어서 신경회로망을 제어기와 식별기로 사용하는 많은 연구결과들이 발표되었다. 식별방법에 대해서 현재까지 발표된 주요한 연구들은 플랜트의 입출력

샘플을 신경회로망의 학습데이터로 사용하여 시스템을 식별하며, 식별된 신경회로망의 출력을 이용하는 방식으로 사용되고 있다^[3].

FNN(Feedforward Neural Network)을 사용한 예를 보면, Psaltis^[4] 등은 플랜트의 역다이나믹스의 단점을 극복하기 위해, 일반학습과 특별학습 구조를 갖는 시스템을 제안하였는데, 일반학습 구조는 넓은 동작영역에서 플랜트의 역다이나믹스를 학습하기 위한 구조이고, 특별학습은 기준신호 범위내의 동작영역에서 학습할 수 있는 구조이다. Kawato^[5] 등은 역다이나믹스의 모델의 목표출력이 필요치 않고, 학습과 제어가 온라인으로 수행되며, 플랜트를 통한 여러 역전파가 필

* 正會員, 漢陽大學校 電子工學科

(Dept. of Electronic Eng., Hanyang Univ.)

接受日字 : 1994年 11月 10日

요치 않은 구조를 제안하였는데, FNN은 플랜트의 역다이나믹스를 학습하도록 피드백 에러 학습 구조가 사용되며, 플랜트에 순방향 제어 신호를 생성하는 역할을 할 수 있는 구조이다. Nguyen^[6] 등은 두개의 FNN을 사용하는 제어 방식을 제안하였는데, 하나는 비선형 플랜트의 특성을 학습하는 에뮬레이터로 사용되고, 다른 하나는 제어기로써 사용된다. 이 구조에서의 특징은, 플랜트 출력과 목표출력의 에러를 에뮬레이터를 통해 역전파시켜 신경회로망 제어기를 학습시킨다는 것이다. Narendra^[7] 등은 미지의 비선형시스템의 식별과 제어문제에 대한 일반화된 신경회로망 모델을 제안하였고, 시스템 식별모델로써 직렬-병렬, 병렬 학습모델을 소개하였다. Chen^[8] 은 SISO 피드백 선형화 시스템을 제어하기 위한 자기동조 제어시스템에 역전파 방식으로 학습시킨 신경회로망을 도입하여 비선형시스템의 추종문제에 좋은 결과를 보였다.

한편, 신경회로망을 선형 동적 시스템 파라미터 식별에 사용한 연구가 있었는데, Chu등^[9]은 입력, 상태, 상태도함수를 측정하여 선형 동적시스템을 Hopfield 네트워크를 이용하여 식별하였다. Hopfield 네트워크는 경우에 따라서는 대역 최적해에 도달하지 못하는 결점이 지적되고 있다. 이런 경우에는 계산적으로 복잡하고 상당히 부정확한 결과를 낼 수 있는 호프필드 네트워크를 사용하는 것보다는 다층신경망을 이용하는 것이 정확한 결과를 내는데 도움이 된다. S. Bhama^[10] 등은 이러한 단점을 지적하고 단층 신경회로망으로 구성된 네트워크로 미지의 선형 동적 시스템의 파라미터를 경사도기법을 이용하여 추정하는 알고리즘을 제시하였다. 단층 신경회로망의 구조적인 제한으로 임의의 함수를 매핑하는 능력이 결여되어 있다.

본 논문에서는, 선형 시스템의 입력, 상태, 상태도함수를 다층 신경회로망의 학습 데이터로 사용하여 미지의 선형시스템을 식별한 후, 식별된 신경회로망의 입력에 대한 출력의 편미분을 뉴런 출력과 웨이트의 함수로써 나타내어 선형시스템의 파라미터를 추정하는 새로운 접근법을 제안한다. 이 방법에 대한 수학적인 표현식을 유도하며, 2층 신경회로망과 3층 신경회로망에 적용하여 2차선형 시스템의 파라미터를 추정한 결과를 시뮬레이션 과정에서 제시하기로 한다.

II. 다층신경회로망

1. 신경회로망 모델

일반적으로 다층 신경회로망은 입력층과 출력층 사이에 하나 또는 그 이상의 은닉층으로 구성이 된 네트워크이다. 각 층은 뉴런이라 하는 계산 유니트로 구성

이 되어 있다. 첫째층의 출력은 둘째층의 입력으로 작용하고, 둘째층의 출력은 셋째층의 입력으로 작용하며, 이와같은 방식으로 출력층에 이르게 된다. 은닉층에서의 뉴런은 인접한 다른층의 뉴런들과 완전하게 결선되어 있다. 로 사상하는 연산자인 뉴런의 출력은 다음과 같다.

$$o = \Gamma\left(\sum_{j=1}^n w_j x_j + w_o\right) \quad (1)$$

여기서 $x^T = [x_1, x_2, \dots, x_n]^T$ 는 입력 벡터이고, $w^T = [w_1, w_2, \dots, w_n]^T$ 는 뉴런의 웨이트 벡터이다. w_o 는 바이어스를 나타낸다. 함수 $\Gamma(\cdot)$ 는 보통 단조 증가 연속함수 $\Gamma: R \rightarrow (-1, 1)$ 로 사상하는 함수(tanh) 또는 $\Gamma: R \rightarrow (0, 1)$ 로 사상하는 함수(sig)가 많이 사용되며 사상의 범위를 제한하지 않기 위해서는 선형(항등)함수가 사용된다. 이 함수들은 다음과 같이 표현된다.

$$\begin{aligned} \text{sig}(u) &= \frac{1 - e^{-u}}{1 + e^{-u}} \\ \text{tanh}(u) &= \frac{e^u - e^{-u}}{e^u + e^{-u}} \\ \text{idn}(u) &= u \end{aligned} \quad (2)$$

여기서 u 는 이 함수에 대한 입력이다.

뉴런이 $l=0, 1, 2, \dots, L$ 층으로 구성이 되어 있는 경우에, l 층에서 j 번째 뉴런의 출력은 다음과 같이 나타낸다.

$$\begin{aligned} o_j^l &= \Gamma_j^l\left(\sum_{i=1}^{N_{l-1}} w_{ji}^l o_i^{l-1} + w_{j0}^l\right) \\ &= \Gamma_j^l(u_j^l) \end{aligned} \quad (3)$$

여기서, $u_j^l = \sum_{i=1}^{N_{l-1}} w_{ji}^l o_i^{l-1} + w_{j0}^l$ 이고, $[w_j^l]^T = [w_{j0}^l, w_{j1}^l, \dots, w_{jN_{l-1}}^l]^T$ 은 l 번째 층에서 j 번째 뉴런에 관계된 웨이트 벡터이다. N_{l-1} 은 $(l-1)$ 층의 뉴런의 수를 나타낸다. 입력층은 $l=0$ 로 표시하고, 출력층은 $l=L$ 로 표시한다. 모든 다른 층들은 은닉층 또는 중간층이라고 한다. 바이어스 w_{j0}^l 은 뉴런의 출력이 항상 1이 되는 웨이트이다.

2. 역전파 알고리즘

역전파 알고리즘은 신경회로의 출력 NN_f 와 목표값 $f(x)$ 와의 에러를 최소화하는 경사도법을 사용한다. 신경회로망의 입력 x 에 대한 출력과 목표출력과의 에러는 다음과 같이 정의된다.

$$e(x) = f(x) - NN_f \quad (4)$$

에러를 최소화하도록 웨이트값의 갱신이 이루어지는데, 출력층에서의 웨이트의 갱신은 다음 식에 따라 이루어진다.

$$w_{kj}^t(k+1) = w_{kj}^t(k) + \Delta w_{kj}^t(k) \quad (5)$$

여기서, $\Delta w_{kj}^t(k)$ 는

$$\Delta w_{kj}^t(k) = \eta \delta_k^t o_j^{t-1} \quad (6)$$

$$\delta_k^t = (t_k - o_k) \Gamma'(u_k^t)$$

이다.

은닉층에서의 웨이트의 갱신은 다음 식에 따라 이루어진다.

$$w_{ji}^t(k+1) = w_{ji}^t(k) + \Delta w_{ji}^t(k) \quad (7)$$

여기서, $\Delta w_{ji}^t(k)$ 는

$$\Delta w_{ji}^t(k) = \eta \delta_j^t o_i^{t-1} \quad (8)$$

$$\delta_j^t = \sum_{k=1}^{N_{t+1}} [\delta_k^{t+1} w_{kj}^{t+1}(k+1)] \Gamma'_t(u_k^t)$$

이다.

III. 선형동적시스템

1. 시스템의 입출력 표현식

동적 시스템을 나타내는 미분방정식은 다음과 같다.

$$\frac{dx(t)}{dt} = \dot{x}(t) = \phi[x(t), u(t)] \quad (9)$$

$$y(t) = \psi[x(t)] \quad (10)$$

여기서, 상태변수 $x(t)$ 는 $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ 이고, 입력 $u(t)$ 는 $u(t) = [u_1(t), u_2(t), \dots, u_p(t)]^T$, 출력 $y(t)$ 는 $y(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T$ 이다. 식(9),(10)에서의 ϕ 와 ψ 는 $\phi: R^n \times R^p \rightarrow R^n$ 으로 정의 되고 $\psi: R^n \rightarrow R^m$ 으로 정의된다. 상태벡터 $x(t)$ 는 시간 t 에서 시스템의 상태를 나타내고, 구간 $[t_0, t]$ 에서 정의되는 입력 u 와 시간 $t_0 < t$ 에서의 상태에 의해 결정된다.

(9),(10)식으로 표현되는 선형시불변 연속시간 시스템은 다음과 같이 나타낼 수 있다.

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (11)$$

$$y(t) = Cx(t) \quad (12)$$

여기서 A는 $(n \times n)$ 인 시스템 행렬이고, B는 $(n \times p)$

인 입력행렬, C는인 $(m \times n)$ 출력행렬이다.

2. 시스템의 추정

제어문제에 있어서 식별 모델을 설정하는 것이 중요한 과정이다. 측정된 신호로써 수학적인 모델의 범주에서 시스템의 동적특성을 결정하는 과정이므로 설정한 모델이 제어 대상 시스템을 식별할 수 있는 시스템 식별가능성(System Identifiability)이 있는 구조가 되어야 한다^[3]. 즉, 주어진 모델 구조로써 대상시스템을 적절하게 나타낼 수 있어야 한다. Hornik, Funahashi등은 한개의 은닉층을 갖고 있고, 은닉층은 비선형성을 갖는 뉴런으로 구성된 신경회로망은 임의의 연속 함수를 사상할 수 있다고 증명하였다^[1,2]. 이와 같은 이론적인 토대로 신경회로망이 미지의 선형 또는 비선형 시스템을 식별하여 식별기로써 사용할 수 있으며, 시스템의 파라미터를 구할 수 있는 근거가 된다.

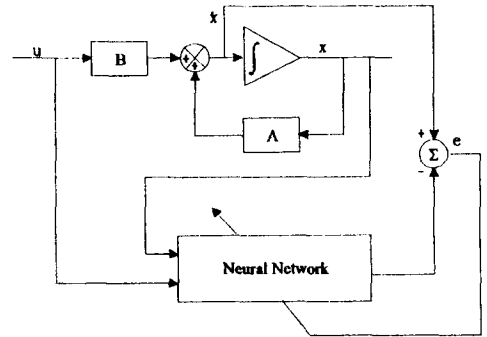


그림 1. 함수 ϕ 를 추정하는 신경회로망 식별 시스템의 구조

Fig. 1. Neural network identification architecture for estimation of function ϕ .

연속시간 시스템 $\dot{x}(t) = \phi[x(t), u(t)]$ 에서 ϕ 가 미지인 경우에 신경회로망으로 시스템을 근사화 하기 위해서는 $((x(t), u(t)) \rightarrow (\dot{x}(t)))$ 쌍을 이용하여 온라인 방식으로 학습하여야 한다. 그러므로 ϕ 에 대한 추정문제는 신경회로망 $NN_f[x(t), u(t), W]$ 의 파라미터 W의 추정문제가 된다. 추정에러 $e(t)$ 는

$$e(t) = \dot{x}(t) - \hat{\dot{x}}(t) \quad (13)$$

이므로 이것은

$$e(t) = f(x(t), u(t)) - NN_f[x(t), u(t), W] \quad (14)$$

$$\cong 0$$

가 되고 이것은 다음과 같이 된다.

$$f(x(t), u(t)) \cong NN_f[x(t), u(t), W] \quad (15)$$

신경회로망의 입력패턴에 대한 출력쌍(즉, $(x(t), u(t)) \rightarrow (\hat{x}(t))$)은 신경회로망의 각 층의 뉴런출력과 웨이트간의 관계로 나타낼 수 있다.

$$\frac{\partial \hat{\mathbf{x}}(t)}{\partial \mathbf{u}(t)} = \begin{bmatrix} \frac{\partial \hat{x}_1(t)}{\partial u(t)} \\ \frac{\partial \hat{x}_2(t)}{\partial u(t)} \\ \vdots \\ \frac{\partial \hat{x}_n(t)}{\partial u(t)} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (18b)$$

IV. 다층신경회로망에서 동적시스템의 파라미터 추정

동적 시스템을 학습한 신경회로망의 입력, 뉴런의 출력, 웨이트들은 시스템의 동적특성을 나타내므로(식 15), 동적 시스템의 출력에 대한 입력의 편미분 관계를 신경회로망의 출력에 대한 입력의 편미분 관계로 나타낼 수 있다. 이 장에서는, 이러한 관계를 이용하여, 선형 시스템의 파라미터를 추정할 수 있는 관계식을 신경회로망의 웨이트-뉴런출력의 함수로써 표현되는 식으로 유도하기로 한다.

차수가 n 인 연속시간 시스템은 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \dot{x}_1(t) &= a_{11}x_1(t) + a_{12}x_2(t) + \dots + a_{1n}x_n(t) + b_1u(t) \\ \dot{x}_2(t) &= a_{21}x_1(t) + a_{22}x_2(t) + \dots + a_{2n}x_n(t) + b_2u(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n1}x_1(t) + a_{n2}x_2(t) + \dots + a_{nn}x_n(t) + b_nu(t) \end{aligned} \quad (16)$$

식(16)에서 $x_1(t), x_2(t), \dots, x_n(t), u(t)$ 에 대한 $\dot{x}_i(t)$ 의 편미분은 각각 다음과 같다.

$$\begin{aligned} \frac{\partial \dot{x}_1(t)}{\partial x_1(t)} &= a_{11} \\ &\vdots \\ \frac{\partial \dot{x}_1(t)}{\partial x_n(t)} &= a_{1n} \\ \frac{\partial \dot{x}_1(t)}{\partial u(t)} &= b_1 \end{aligned} \quad (17)$$

이와 같은 방식으로 $x_1(t), x_2(t), \dots, x_n(t), u(t)$ 에 대해서 $\dot{x}_2(t), \dots, \dot{x}_n(t)$ 의 편미분을 취한 것을 행렬식으로 표현하면 다음의 (18a)와 (18b)와 같은 형태가 된다.

$$\frac{\partial \dot{\mathbf{x}}(t)}{\partial \mathbf{x}(t)} = \begin{bmatrix} \frac{\partial \dot{x}_1(t)}{\partial x_1(t)} & \frac{\partial \dot{x}_1(t)}{\partial x_2(t)} & \dots & \frac{\partial \dot{x}_1(t)}{\partial x_n(t)} \\ \frac{\partial \dot{x}_2(t)}{\partial x_1(t)} & \frac{\partial \dot{x}_2(t)}{\partial x_2(t)} & \dots & \frac{\partial \dot{x}_2(t)}{\partial x_n(t)} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \dot{x}_n(t)}{\partial x_1(t)} & \dots & \dots & \frac{\partial \dot{x}_n(t)}{\partial x_n(t)} \end{bmatrix} \quad (18a)$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

위의 (18a)와 (18b)를 시스템 행렬과 입력 행렬로 간단히 표현하면

$$\frac{\partial \dot{\mathbf{x}}(t)}{\partial \mathbf{x}(t)} = A \quad (19a)$$

$$\frac{\partial \dot{\mathbf{x}}(t)}{\partial \mathbf{u}(t)} = B \quad (19b)$$

가 된다. 위의 식에서 알 수 있듯이, 시스템 행렬 파라미터는 상태 $x(t)$ 에 대한 상태도함수 $\dot{x}(t)$ 의 편미분으로 나타낼 수 있고, 입력 행렬 파라미터는 입력 $u(t)$ 에 대한 상태도함수 $\dot{x}(t)$ 의 편미분으로 나타낼 수 있다.

신경회로망은 임의의 층 $l=0, 1, \dots, (L-1), L$ 로 구성되어 있다. 0층은 입력층으로서 뉴런의 수는 n_0 개이고, 첫번째 은닉층 1층의 뉴런은 n_1 개, 출력층에 인접한 은닉층 $L-1$ 층의 뉴런의 수는 $n_{(L-1)}$ 개, 출력층 L 층의 뉴런의 수는 n_L 로 구성되어 있다. 출력층 $l=L$ 층과 출력층에 인접한 은닉층 $l=(L-1)$ 층 사이에는 다음과 같은 관계가 성립한다. L 층의 첫번째 뉴런과 $L-1$ 층의 첫번째 뉴런의 관계를 연쇄법칙을 적용하면 다음과 같이 나타낼 수 있다.

$$\frac{\partial o_1^L}{\partial o_1^{(L-1)}} = \frac{\partial o_1^L}{\partial u_1^L} \frac{\partial u_1^L}{\partial o_1^{(L-1)}} \quad (20)$$

가 된다. 여기서 $\frac{\partial o_1^L}{\partial u_1^L}$ 와 $\frac{\partial u_1^L}{\partial o_1^{(L-1)}}$ 는 각각 다음과 같다.

$$\frac{\partial o_1^L}{\partial u_1^L} = \Gamma_1^L(u_1^L) \quad (21)$$

$$\begin{aligned} \frac{\partial u_1^L}{\partial o_1^{(L-1)}} &= \frac{\partial}{\partial o_1^{(L-1)}} \left(\sum_{i=1}^{n_{(L-1)}} w_{1i}^L o_i^{(L-1)} \right) \\ &= \frac{\partial}{\partial o_1^{(L-1)}} (w_{11}^L o_1^{(L-1)} + w_{12}^L o_2^{(L-1)} + \dots \\ &\quad + w_{1n_{(L-1)}}^L o_{n_{(L-1)}}^{(L-1)}) \\ &= w_{11}^L \end{aligned} \quad (22)$$

여기서 $\Gamma(\cdot)$ 는 식(2)의 시그모이드 함수 중의 한가지를 사용한다. 식(21), (22)으로 부터 식(20)은 다음과 같이 된다.

$$\frac{\partial o_1^L}{\partial o_1^{(L-1)}} = \Gamma_1^L(u_1^L) w_{11}^L \quad (23)$$

같은 방식으로, L 층의 j 번째 뉴런과 $(L-1)$ 층의 i 번째 뉴런에서는 다음과 같이 된다.

$$\frac{\partial o_j^L}{\partial o_i^{(L-1)}} = \frac{\partial o_j^L}{\partial u_j^L} \frac{\partial u_j^L}{\partial o_i^{(L-1)}} = f_j^L(u_j^L) w_{ji}^L \quad (24)$$

은닉층 h 층과 층과 $h-1$ 층에서 h 층의 j 번째 뉴런과 $h-1$ 층의 i 번째 뉴런의 관계는

$$\frac{\partial o_j^h}{\partial o_i^{(h-1)}} = \frac{\partial o_j^h}{\partial u_j^h} \frac{\partial u_j^h}{\partial o_i^{(h-1)}} = f_j^h(u_j^h) w_{ji}^h \quad (25)$$

가 된다. 그리고 입력층에 인접한 은닉층 1층과 입력층 0층 사이에서 1층의 j 번째 뉴런과 0층의 i 번째 뉴런에서는 다음과 같은 관계가 성립한다.

$$\frac{\partial o_j^1}{\partial o_i^0} = \frac{\partial o_j^1}{\partial u_j^1} \frac{\partial u_j^1}{\partial o_i^0} = f_j^1(u_j^1) w_{ji}^1 \quad (26)$$

(24),(25),(26)식에서 알 수 있듯이 임의의 층 l 와 $l-1$ 층 사이에서는 층만 다르게 표현되고 식은 같은 형태가 성립한다.

다층 신경회로망의 구조적인 특성상 입력층에서 출력층까지는 각층마다 여러개의 뉴런이 구성되어 있고 각각의 뉴런을 연결하는 웨이트가 있으므로, 출력층 ($l=L$)의 j 번째 뉴런에 대한 입력층 ($l=0$)의 i 번째 뉴런의 편미분은 은닉층에서의 뉴런들이 매개가 되는 연쇄법칙을 적용하면 다음과 같이 나타낼 수 있다.

$$\frac{\partial o_j^L}{\partial o_i^0} = \sum_{p=1}^{n_{L-1}} \sum_{q=1}^{n_{L-2}} \dots \sum_{m=1}^{n_1} \sum_{n=1}^{n_0} \left(\frac{\partial o_j^L}{\partial u_p^L} \frac{\partial u_p^L}{\partial o_q^{(L-1)}} \frac{\partial o_q^{(L-1)}}{\partial u_n^{(L-2)}} \times \dots \times \frac{\partial o_p^3}{\partial o_n^2} \frac{\partial o_n^2}{\partial u_m^1} \frac{\partial o_m^1}{\partial o_i^0} \right) \quad (27)$$

윗 식(27)은 식(24),(25),(26)로 부터

$$\frac{\partial o_j^L}{\partial o_i^0} = \sum_{p=1}^{n_{L-1}} \sum_{q=1}^{n_{L-2}} \dots \sum_{m=1}^{n_1} \sum_{n=1}^{n_0} \left[(f_j^L(u_j^L) \cdot w_{jp}^L)(f_p^{(L-1)}(u_p^{(L-1)}) \cdot w_{pq}^{(L-1)}) \dots (f_n^2(u_n^2) \cdot w_{nm}^2)(f_m^1(u_m^1) \cdot w_{mi}^1) \right] \quad (28)$$

가 된다. 신경회로망의 입력은 선형시스템의 상태벡터 $x_1(t), x_2(t), \dots, x_n(t)$ 와 제어입력 $u(t)$ 이고, 신경회로망의 목표출력은 상태도함수 $\dot{x}_1(t), \dot{x}_2(t), \dots, \dot{x}_n(t)$ 이므로 $x_i(t)$ 에 대한 출력 $\dot{x}_j(t)$ 의 편미분은 식(18a)로 부터 다음과 같이 된다.

$$\frac{\partial \dot{x}_j(t)}{\partial x_i(t)} = a_n = \frac{\partial o_j^L}{\partial o_i^0} = \sum_{p=1}^{n_{L-1}} \sum_{q=1}^{n_{L-2}} \dots \sum_{m=1}^{n_1} \sum_{n=1}^{n_0} \left[(f_j^L(u_j^L) \cdot w_{jp}^L) \dots (f_n^2(u_n^2) \cdot w_{nm}^2)(f_m^1(u_m^1) \cdot w_{mi}^1) \right] \quad (29)$$

$u(t)$ 에 대한 $\dot{x}_j(t)$ 의 편미분은 식(18b)로 부터 다음과 같이 된다.

$$\begin{aligned} \frac{\partial \dot{x}_j(t)}{\partial u(t)} &= b_j = \frac{\partial o_j^L}{\partial o_{n_0}^0} \\ &= \sum_{p=1}^{n_{L-1}} \sum_{q=1}^{n_{L-2}} \dots \sum_{m=1}^{n_1} \sum_{n=1}^{n_0} \left[(f_j^L(u_j^L) \cdot w_{jp}^L) \dots (f_p^{(L-1)}(u_p^{(L-1)}) \cdot w_{pq}^{(L-1)}) \dots (f_n^2(u_n^2) \cdot w_{nm}^2) (f_m^1(u_m^1) \cdot w_{mn}^1) \right] \end{aligned} \quad (30)$$

여기서 n_l 는 신경회로망에 인가하는 입력의 수이다. $u(t)$ 는 신경회로망의 n_l 번째 입력이 된다.

식(29)와 (30)은 신경회로망이 임의의 층을 갖는 경우에 대해 일반적인 입력-출력의 편미분관계를 유도한 식이다. 대부분 다층 신경회로망은 입력층, 은닉층, 출력층을 각각 한 층으로 구성하거나, 한개의 입력층, 두개의 은닉층, 한개의 출력층으로 구성된 시스템을 사용하게 된다. 한개의 은닉층을 갖는 2층 신경회로망에서는 (29),(30)식을 다음과 같이 나타낼 수 있다.

$$\begin{aligned} \frac{\partial \dot{x}_j(t)}{\partial x_i(t)} &= a_n = \frac{\partial o_j^O}{\partial o_i^I} = \sum_{h=1}^{n_H} \left[\frac{\partial o_j^O}{\partial o_h^H} \frac{\partial o_h^H}{\partial o_i^I} \right] \\ &= \sum_{h=1}^{n_H} \left[(f_j^O(u_j^O) \cdot w_{jh}^O)(f_h^H(u_h^H) \cdot w_{hi}^H) \right] \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{\partial \dot{x}_j(t)}{\partial u(t)} &= b_j = \frac{\partial o_j^O}{\partial o_{n_I}^I} = \sum_{h=1}^{n_H} \left[\frac{\partial o_j^O}{\partial o_h^H} \frac{\partial o_h^H}{\partial o_{n_I}^I} \right] \\ &= \sum_{h=1}^{n_H} \left[(f_j^O(u_j^O) \cdot w_{jh}^O)(f_h^H(u_h^H) \cdot w_{hn_I}^H) \right] \end{aligned} \quad (32)$$

여기서, O, H, I 는 각각 출력층, 은닉층, 입력층을 나타낸다. 은닉층, 출력층에 사용하는 뉴런의 함수가 모두 tanh인 경우에는 $f_j^O(\bullet) = f_h^H(\bullet) = f(\bullet)$ 가 되고 $f(u) = (1+0)(1-0)$ 가 되므로 (31),(32)식은 다음과 같이 된다.

$$\begin{aligned} a_n &= \frac{\partial o_j^O}{\partial o_i^I} \\ &= \sum_{h=1}^{n_H} \left[((1+o_j^O)(1-o_j^O) \cdot w_{jh}^O)((1+o_h^H)(1-o_h^H) \cdot w_{hi}^H) \right] \end{aligned} \quad (33)$$

$$\begin{aligned} b_j &= \frac{\partial o_j^O}{\partial o_{n_I}^I} \\ &= \sum_{h=1}^{n_H} \left[((1+o_j^O)(1-o_j^O) \cdot w_{jh}^O)((1+o_h^H)(1-o_h^H) \cdot w_{hn_I}^H) \right] \end{aligned} \quad (34)$$

두개의 은닉층을 갖는 3층 신경회로망의 경우에는 (29),(30)를 각각 다음과 같이 표현할 수 있다.

$$\begin{aligned} a_{ij} &= \frac{\partial o_j^O}{\partial o_i^I} \\ &= \sum_{h=1}^{n_H} \sum_{k=1}^{n_H} \left[((f_j^O(u_j^O) \cdot w_{jh}^O)(f_h^H(u_h^H) \cdot w_{hk}^H) \cdot w_{ki}^H) \right. \\ &\quad \left. (f_k^H(u_k^H) \cdot w_{ki}^H) \right] \\ &= \sum_{h=1}^{n_H} \sum_{k=1}^{n_H} \left[((1+o_j^O)(1-o_j^O) \cdot w_{jh}^O)((1+o_h^H)(1-o_h^H) \cdot w_{hk}^H) \cdot w_{ki}^H) \right. \\ &\quad \left. ((1+o_k^H)(1-o_k^H) \cdot w_{ki}^H) \right] \end{aligned} \quad (35)$$

$$\begin{aligned}
 b_i &= \frac{\partial o_i^p}{\partial o_{n_i}^p} \\
 &= \sum_{k=1}^{n_k} \sum_{h=1}^{n_h} [((I_{k_i}^p(u_i^p) \cdot w_{ik}^p) (I_{h_i}^{H_i}(u_{k_i}^{H_i}) \cdot w_{h_i k_i}^H) \\
 &\quad (I_{h_i}^{H_i}(u_{k_i}^{H_i}) \cdot w_{h_i k_i}^H)] \quad (36) \\
 &= \sum_{k=1}^{n_k} \sum_{h=1}^{n_h} [((1+o_i^p)(1-o_i^p) \cdot w_{ik}^p) ((1+o_{h_i}^{H_i})(1-o_{h_i}^{H_i}) \cdot w_{h_i k_i}^H) \\
 &\quad ((1+o_{h_i}^{H_i})(1-o_{h_i}^{H_i}) \cdot w_{h_i k_i}^H)]
 \end{aligned}$$

이상에서 유도한 바와같이 선형시스템의 파라미터는 신경회로망의 뉴런출력과 웨이트의 함수로 구할 수 있고 2층신경회로망인 경우에는 식(33),(34)로 나타낼 수 있으며, 3층신경회로망인 경우는 식(35),(36)으로 나타낼 수 있다.

V. 시뮬레이션

시뮬레이션 과정에서는, 제안한 선형시스템의 파라미터 추정법을 2층 신경회로망(은닉층 1개)인 경우와 3층신경회로망(은닉층 2개)인 경우에 대해서 각각 수행하기로 한다. 식별 대상 시스템은 2차 선형 시스템이다.

1. 2층 신경회로망

2층 신경회로망이 학습할 선형시스템은 다음과 같다.

$$\dot{x}(t) = \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1.5 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = [1 \ 0] x(t)$$

이 시스템에 대한 입력은 $u(t) = \sin(2t) + \cos(t)$ 이다. 샘플링 시간은 0.5초로 설정하였으며 시스템의 응답은 0초에서 20초까지의 상태벡터 값을 측정하여 신경회로망을 학습한다.

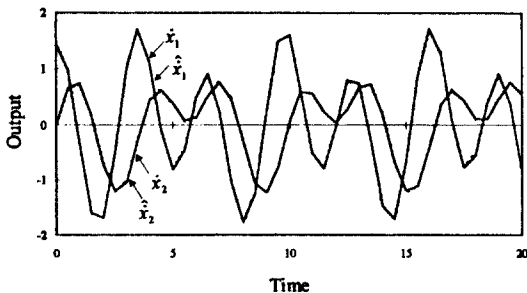


그림 2. 학습된 2층 신경회로망의 출력과 선형시스템의 출력

Fig. 2. Outputs of the trained 2-layer NN and the linear system.

신경회로망은 입력층, 은닉층, 출력층으로 구성을 하며 입력층의 수는 3개, 은닉층의 뉴런의 수는 15개, 출력층의 뉴런은 2개이다. 첫번째 입력에는 상태벡터 $x_1(t)$ 을 인가하고 두번째 입력에는 상태벡터 $x_2(t)$ 를

인가한다. 세번째 입력에는 $u(t)$ 를 인가한다. 신경회로망의 첫번째 출력 뉴런의 목표출력은 상태도함수 $x_1(t)$ 이고, 두번째 뉴런의 목표출력은 상태도함수 $x_2(t)$ 이다. 은닉층에는 tanh함수를 사용하며, 출력층에는 선형함수(idn)을 사용한다. 신경회로망의 역전파 학습비는 0.05이고, 각 패턴(40패턴)을 10,000회 학습하였고, 각 패턴의 평균에러는 0.001이다.

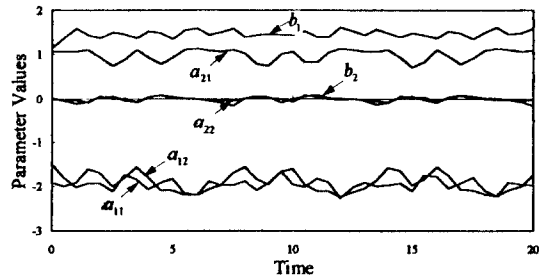


그림 3. 선형 시스템을 학습한 2층 신경회로망으로부터 추정된 파라미터 값

Fig. 3. Estimated parameter values from the system identified 2-layer NN.

그림 2는 선형시스템의 상태 도함수 $x_1(t)$ 과 $x_2(t)$ 의 출력과 신경회로망이 학습한 출력을 나타낸 것이다. 신경회로망의 출력값과 목표 출력은 에러값이 0.001인 상태로 수렴하였으므로 그래프상에서 두 값의 차이는 매우 작게 나타난다. 그림 3은 시스템의 동적특성을 학습한 신경회로망으로부터 추정된 파라미터의 값들이다. 학습을 한 후에, 매 시간마다 주어지는 각 입력 패턴 $x_1(t), x_2(t), u(t)$ 을 인가하여 식(33),(34)로 구한 파라미터값들을 나타낸 것이다. a_{11} 의 값은 $-1.75 \sim 2.25$, a_{12} 는 $-1.6 \sim 2.2$, a_{21} 은 $0.75 \sim 1.1$, a_{22} 는 $-0.15 \sim 0.09$, b_1 은 $1.35 \sim 1.55$, b_2 는 $-0.09 \sim 0.08$ 값의 범위를 나타냈다. 그림 3에서의 각 파라미터 값들이 고정되지 않고 진동을 나타내는데, 이것은 파라미터의 값들을 결정하는 요소(뉴런의 출력, 웨이트)들의 영향을 받기 때문이다. 학습을 마친 상태에서는 웨이트의 변화는 거의 없지만, 뉴런의 출력은 주어지는 입력의 값에 영향을 받게 된다. 입력이 변화는 주기적으로 이루어지므로, 정확한 파라미터값을 구하기 위해서는 입력으로 주어지는 데이터의 최소한 한 주기의 평균치를 구하는 것이 보다 정확한 값을 얻을 수 있는 방법이다. 그림 4는 그림 3에서 구한 파라미터들에 평균을 취해서 나타낸 것이다. 각 파라미터 값은, a_{11} 은 -1.95 , a_{12} 는 -1.92 , a_{21} 은 0.99 , a_{22} 는 0.01 , b_1 은 1.46 , b_2 는 -0.01 로 수렴하였다. 이것은 그림 3에서의

진동이 있는 부분을 완화시킨 결과로 실제 파라미터값에 근접한 결과를 보였다. 그림 5는 이 시스템을 학습할 때, 신경회로망의 학습횟수에 따른 오차의 평균을 나타낸 것이다.

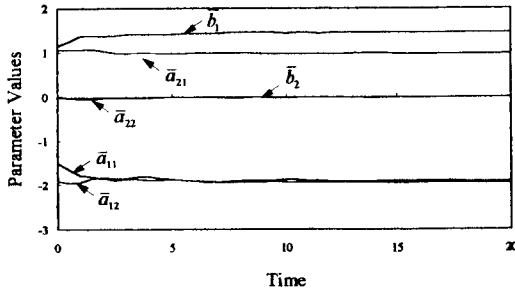


그림 4. 2층 신경회로망으로 부터 추정한 시스템의 파라미터값의 평균값

Fig. 4. Average values of estimated parameter from the system identified 2-layer NN.

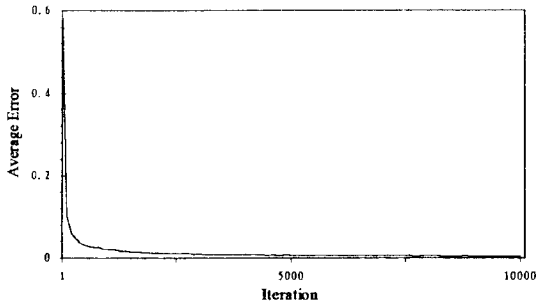


그림 5. 2층 신경회로망으로 시스템 식별시의 각 패턴의 평균오차

Fig. 5. Average sum-squared error during plant identification using 2-layer NN.

2. 3층 신경회로망

3층 신경회로망으로 학습할 선형 시스템은 다음과 같다.

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} x(t) + \begin{bmatrix} 1.5 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = [1 \ 0] x(t)$$

이 시스템에 대한 입력은 $u(t) = \sin(t)$ 이며 샘플링 시간은 0.5초이며 시스템의 응답은 0초에서 20초까지의 상태벡터 값을 측정하여 신경회로망을 학습한다.

신경회로망은 입력층, 두개의 은닉층, 출력층으로 구성을 하며 입력의 수는 3개, 첫번째와 두번째 은닉층의 뉴런의 수는 각기 10개, 출력층의 뉴런은 2개이다. 입력층의 첫번째 노드에는 상태벡터 $x_1(t)$ 을 인가하고

두번째 노드에는 상태벡터 $x_2(t)$ 를 인가한다. 세번째 노드에는 $u(t)$ 를 인가한다. 신경회로망의 첫번째 출력 뉴런의 목표출력은 상태도함수 $\dot{x}_1(t)$ 이고, 두번째 뉴런의 목표출력은 상태도함수 $\dot{x}_2(t)$ 이다. 출력층과 두 은닉층 모두 tanh 함수를 사용하며, 신경회로망의 역전파 학습비는 0.01이고, 각 패턴(40패턴)을 100,000회 학습하였고, 각 패턴의 평균오차는 0.0005이다. 그림 6은 선형시스템의 상태도함수 $\dot{x}_1(t)$ 와 $\dot{x}_2(t)$ 의 출력과 100,000회의 학습을 마친후, 신경회로망에 각 패턴 입력에 대한 출력을 시간 0에서 20초까지의 응답을 나타낸 것이다. 그림 7은 식(35), (36)을 적용하여, 매 시간마다 주어지는 각 입력 패턴에 대한 파라미터값들을 나타낸 것이다. a_{11} 의 값은 $-0.18 \sim 0.3$, a_{12} 는 $0.96 \sim 1.4$, a_{21} 은 $-1.83 \sim -1.96$, a_{22} 는 $-2.82 \sim -3.1$, b_1 은 $-0.4 \sim 0.01$, b_2 는 $1.94 \sim 1.98$ 값의 범위를 나타냈다. 그림 8은 그림 7에서 구한 파라미터들을 평균하여 나타낸 것으로, 각 파라미터의 값은 a_{11} 은 0.05, a_{12} 는 1.12, a_{21} 는 -1.93 , a_{22} 는 -2.93 , b_1 은 -0.15 , b_2 는 1.96값으로 수렴하였다. 그림 9는 이 시스템을 학습할 때, 신경회로망의 학습횟수에 따른 오차의 평균을 나타낸 것이다.

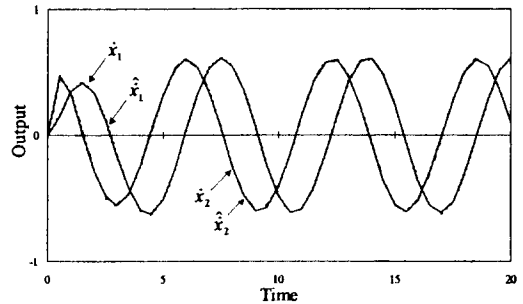


그림 6. 학습된 3층 신경회로망의 출력과 선형시스템의 출력

Fig. 6. Outputs of the trained 3-layer NN and the linear system.

본 논문의 방법과 Bhama^[10]방법을 방법론적인 측면에서 비교하여 보면, Bhama의 방법은 다층신경회로망을 이용하여 1000회정도의 학습으로 원하는 파라미터 값을 얻을 수 있었다. 이 방법이 신경회로망의 학습시간 측면에서 상당한 잇점이 있을 수 있다. 그러나 임의의 함수(선형, 비선형함수)를 학습하는데 있어서 다층 신경회로망의 제한적인 능력으로 인해 제대로 학습하지 못하는 경우가 있기 때문에, 학습능력이 인정되는 다층 신경회로망을 사용하는 방법이 타당하리라 라

고 생각한다^[1,2,3].

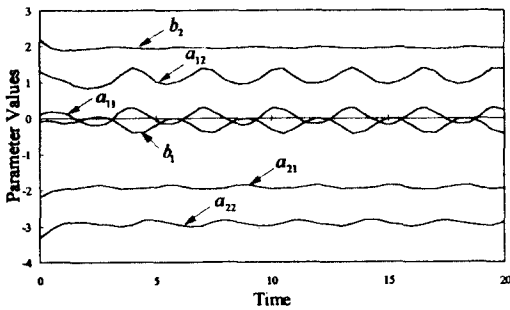


그림 7. 선형 시스템을 학습한 3층 신경회로망으로부터 추정한 파라미터 값

Fig. 7. Estimated parameter values from the system identified 3-layer NN.

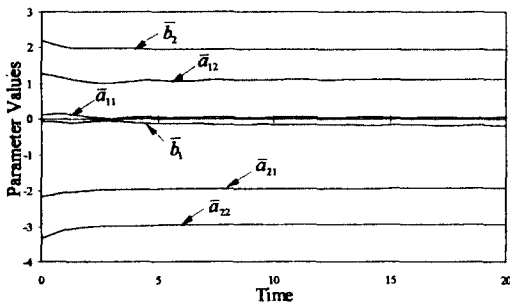


그림 8. 선형 시스템을 학습한 3층 신경회로망으로부터 추정한 파라미터 값

Fig. 8. Average values of estimated parameter from the system identified 3-layer NN.

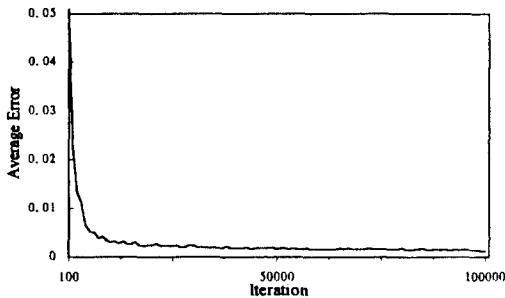


그림 9. 3층 신경회로망으로 시스템 식별시의 각 패턴의 평균에러

Fig. 9. Average sum-squared error during plant identification using 3-layer NN.

VI. 결론

본 논문에서는 상태공간 모델로 표현되는 선형 시스템의 파라미터를 다층 신경회로망의 웨이트와 뉴런의 출력함수와의 관계로써 추정하는 새로운 방법을 제안하였으며, 시뮬레이션 과정을 통하여 제안한 방식으로 선형시스템의 파라미터를 추정할 수 있다는 것을 입증하였다.

2층 신경회로망과 3층 신경회로망에 대해서 2차 선형시스템의 식별 문제를 시뮬레이션 과정을 통해 결과를 제시하였다. 2층 신경회로망으로 추정하여 평균을 취한 파라미터는 최대 4%(절대값 오차 0.08), 최소 1%(절대값 오차 0.01)오차를 보였으며, 3층 신경회로망의 경우에는 최대 15%(절대값 오차 0.15), 최소 3.5%(절대값 오차 0.07)의 오차를 보였다.

기존의 신경회로망을 이용한 시스템 식별방법은 시스템의 입출력 동작특성을 신경회로망이 학습하는 비파라미터적인 방식^[3]을 주로 사용하였으나, 본 논문의 식별방법은 파라미터적인 방법과 비파라미터적인 식별방법을 동시에 사용할 수 있는 장점이 있다.

본 논문에서 제안한 방법을 상태공간 모델에만 적용하였으나, 전달함수 형태의 모델(ARMA)등에서도 확장하여 적용할 수 있다. 또한, 신경회로망이 비선형시스템의 동적특성을 학습할 수 있는 점을 고려하면 비선형시스템을 선형모델로 근사화하는 방법에도 적용가능하리라 생각한다.

참고 문헌

- [1] K.I. Funahashi, "On the approximate realization of continuous mappings by neural networks", *Neural Networks*, No2, pp.183-192, 1989.
- [2] K. Hornik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks*, No2, pp.359-366, 1989.
- [3] K.J. Hunt, D.Sbarbaro, R.Zbikowski, P. J.Gawthrop, "Neural networks for control systems - A survey", *Automatica*, Vol. 28, No. 6, pp1083-1112, 1992.
- [4] D.Psaltis, A.Sideris, A.A.Yamamura, "A multilayered neural network controller", *IEEE Control Systems Magazine*, Vol. 8, No. 2, pp17-21, 1988.
- [5] M. Kawato, Y.Uno, M.Isobe, R. Suzuki, "Hierarchical neural network model for voluntary movement with application to

- robotics", *IEEE Control Systems Magazine*, Vol. 8, No. 2, pp8-17, 1992.
- [6] D.H.Nguyen, B.Widrow, "Neural networks for self-learning control systems", *IEEE Control Systems Magazine*, Vol. 10, No. 2, pp18-23, 1990.
- [7] K.S. Narendra, K. Pathasarathy, "Identification and control of dynamical systems using neural networks", *IEEE Trans. on Neural Networks*, Vol. 1, No. 1, pp4-27, 1992.
- [8] F. Chen, "Back-propagation neural networks for nonlinear self-tuning adaptive control", *IEEE Control Systems Magazine*, Vol. 10, No. 2, pp44-48, 1990.
- [9] S. Reynold Chu, R. Shoureshi, M. Tenorio, "Neural networks for system identification", *IEEE Control Systems Magazine*, Vol. 10, No. 2, pp31-35, 1990.
- [10] S.Bhama, H.Singh, "Single layer neural networks for linear system identification using gradient descent techniques", *IEEE Trans. on Neural Networks*, Vol. 4, No. 5, pp884-888, 1993.

 저 자 소 개



趙奎翔(正會員)

1963년 3월 31일생. 1986년 2월 한양대학교 전자공학과(학사). 1989년 2월 한양대학교 산업대학원(석사). 1989년 3월 ~ 현재 한양대학교 대학원 전자공학과(박사과정 재학). 주관심분

야는 자동제어, 신경회로망, 컴퓨터 비전

金庚基(正會員) KITE journal of electronics engineering Vol. 1, No2, pp 50 - 56, Dec. 1990

현재 한양대학교 전자공학과 교수