

비선형 시스템의 근사화를 위한 직교 신경망의 수정 기법에 관한 연구

A study on Modified Method of Orthogonal Neural Network for Nonlinear system approximation

김성식* · 이영석**

Sung-Sik Kim and Young-Seog Lee

*경북대학교 대학원 전기공학과

**영진전문대학 전기과

요 약

최근 Yang과 Tseng이 제안한 직교 신경망(ONN)은 직교 함수를 이용하여 신경망을 구성한 것으로서, 다층 신경망이 가지는 층의 구조에 대한 어려움 없이 전체 구조를 결정할 수 있다는 장점을 가지고 있다. 또한 요구되는 정확성을 기준으로 직교 함수의 급수를 증가시키므로써 학습하는 동안에 전체 구조를 변형하는 것이 가능하고 가중치의 학습 알고리즘이 오차 역전파법 학습 알고리즘에 비해 간단하며 수렴 속도가 빠르다는 장점이 있다. 그러나 이러한 직교 신경망은 구조의 골격이 되는 직교 함수를 변형할 수 없는 구조를 가진다는 문제점이 있다. 본 논문에서는 입력 변환을 이용하여 직교 함수를 변형할 수 있는 새로운 알고리즘과 입력 변환의 조건을 제시한다. 그리고 입력 변환을 이용하여 직교 함수를 학습할 수 있는 구조를 가지는 수정된 직교 신경망(MONN)을 제안한다. 제안한 수정된 직교 신경망을 이용하여 비선형 시스템을 식별하기 위해 식별기 구조를 설정하고 목적을 달성하기 위한 수정된 직교 신경망의 학습 알고리즘을 유도한다. 사례연구를 통하여, 본 논문에서 제안한 수정된 직교 신경망의 비선형 시스템 모형화 능력, 입력 변환의 유용성을 다층 신경망, 직교 신경망과 비교하여 검증한다.

ABSTRACT

This paper presents a Modified Orthogonal Neural Network(MONN), new modified model of Orthogonal Neural Network(ONN) based on orthogonal functions, and applies it to nonlinear system approximator. ONN proposed by Yang and Tseng, doesn't have the problems of traditional multilayer feedforward neural networks such as the determination of initial weights and the numbers of layers and processing elements. And training of ONN converges rapidly. But ONN cannot adapt its orthogonal functions to a given system. The accuracy of ONN, in terms of the minimal possible deviation between system and approximator, is essentially dependent on the choice of basic orthogonal functions. In order to improve ability and effectiveness of approximate nonlinear systems, MONN has an input transformation layer to adapt its basic orthogonal functions to a given nonlinear system. The results show that MONN has the excellent performance of approximate nonlinear systems and the input transformation makes the ability of MONN better than one of ONN.

1. 서 론

일반적인 함수 근사화 기법은 시스템을 수학적으로 모형화하여 근사화를 이루어 왔다. 그러나 플랜트가 고차의 미분방정식으로 표현되는데 따른 복잡성과 플랜트의 비선형성등으로 인해 만족할 만한 근사화 성능을 나타내기에는 어려움이 있다. 이러한 어려움들을 극복하고자 인공 신경망을 이용한 근사화 기법이 도입되었다. 비선형 시스템의 근사화에 많이 사용

되는 다층 신경망은 초기 가중치를 설정하기 위해 난수를 발생시켜야 하는 어려움과 중간층 수와 각 층의 뉴런 수를 결정하는데 있어서 많은 어려움이 있다.

최근 Yang과 Tseng이 제안한 직교 신경망(Orthogonal Neural Network)[1]은 비선형 함수의 모형화에 많이 사용되어 오던 직교 함수를 이용하여 신경망을 구성한 것으로서, 다층 신경망이 가지는 어려움 없이 전체 구조를 결정할 수 있는 커다란 장점을 가지고 있다. 또한 직교 신경망은 요구되는 정확성을 기준으로 직교

함수의 급수를 증가시키므로 학습하는 동안에 전체 구조를 변형하는 것이 가능하고 가중치의 학습 알고리즘이 오차 역전파와 학습 알고리즘에 비해 간단하며 수렴 속도가 빠르다는 장점이 있다. 그러나 이러한 직교 신경망은 골격이 되는 직교 함수를 변형할 수 없는 구조를 가진다는 문제점이 있다.

시스템을 직교 신경망을 이용하여 모형화할 때, 최소의 오차는 직교 신경망 구조의 골격이 되는 직교 함수와 함수의 차수에 의해서 결정된다. 따라서 시스템이 주어지면 이 시스템에 적당한 직교 함수의 형태를 결정하는 것이 필요하다. 그러나 시스템의 수학적 모형을 모르는 경우 그 시스템에 적당한 직교 함수를 설정한다는 것은 거의 불가능하다고 할 수 있다. 그러므로 직교 신경망에 사용되는 직교 함수는 학습을 통하여 적당한 형태로 변형될 수 있어야 더욱 우수한 특성을 나타낼 수 있을 것이다.

본 논문에서는 입력 변환을 이용하여 직교 함수를 변형할 수 있는 새로운 알고리즘과 입력 변환의 조건을 제시한다. 그리고 입력 변환을 이용하여 직교 함수를 적당한 형태로 학습할 수 있는 구조를 가지는 수정된 직교 신경망(MONN)을 제안한다. 또한, 제시한 입력 변환의 조건을 만족하는 입력 변환 함수를 설정한다. 제안한 직교 신경망을 이용하여 비선형 시스템을 근사화하기 위한 학습 알고리즘을 유도한다. 사례 연구를 통하여, 본 논문에서 제안한 수정된 직교 신경망의 비선형 시스템 근사화 능력과 입력 변환의 유용성을 다층 신경망, 직교 신경망과 비교하여 검증한다.

2. 문제의 정의

식 (1)과 같은 이산 시간 비선형 시스템을 고려한다.

$$y(k+1) = f(y(k), y(k-1), \dots, y(k-n_y+1), u(k), u(k-1), \dots, u(k-n_u+1)) \quad (1)$$

여기서 $y(k)$ 는 시스템의 출력을, $u(k)$ 는 시스템의 입력을 각각 나타낸다. 시스템의 출력은 n_y 개의 과거 시스템 출력과 n_u 개의 과거 시스템 입력에 의해서 결정된다. 이러한 비선형 시스템의 동특성에 대한 수학적 모형은 모르지만 시스템의 과거 운행 자료로부터 입력 출력 정보는 가지고 있다고 가정한다.

비선형 시스템은 직교 함수의 급수를 이용하여 급수의 계수를 결정하므로써 모형화할 수 있다. 그러나 계수를 결정하기 위해서는 비선형 함수의 수학적 모형을 알아야 한다는 어려움이 있다. 최근 Yang과 Tseng은 비선형 시스템의 수학적 모델을 모르는 경우

더라도 신경망의 학습 알고리즘을 이용하여 시스템의 과거 입출력 자료로부터 직교 함수 급수의 계수를 구하여 시스템을 모형화 할 수 있는 직교 함수로 구성된 직교 신경망을 제안하였다.

그러나 이러한 직교 신경망은 비선형 시스템에 가장 적절한 직교 함수의 선정에 어려움이 있다. 따라서 직교 함수를 변형하기 위한 알고리즘이 필요하며, 학습을 통하여 직교 함수를 변형할 수 있는 새로운 구조의 수정된 직교 신경망이 필요하다. 그리고 수정된 직교 신경망을 이용하여 비선형 시스템을 근사화하기 위해서는 적절한 학습 알고리즘을 유도할 필요가 있다.

3. 기존의 직교 신경망

3.1 직교 함수

직교 함수의 성질에 의해 임의의 함수 $f(x), f: [a, b] \rightarrow R$ 에 대해서 식 (2), (3)을 만족하는 $F(x)$ 가 존재한다[2,3].

$$F(x) = w_1 \phi_1(x) + w_2 \phi_2(x) + \dots + w_n \phi_n(x) \quad (2)$$

$$\lim_{n \rightarrow \infty} \int_a^b (f(x) - F(x))^2 dx = 0 \quad (3)$$

여기서, $\{\pi_1(x), \pi_2(x), \pi_3(x), \dots\}$ 는 직교 함수로 된 집합으로 각각의 직교 함수는 식 (4)를 만족하여야 하며 급수의 계수는 식 (5)와 같다.

$$\int_a^b \phi_i(x) \phi_j(x) dx = \begin{cases} 0 & i \neq j \\ A_i & i = j \end{cases} \quad (4)$$

$$w_i = \frac{1}{A_i} \int_a^b f(x) \phi_i(x) dx, \quad i = 1, 2, 3, \dots \quad (5)$$

이러한 성질에 근거하여 직교 함수의 급수는 시스템을 모형화하거나 제어하는데 많이 이용되어 왔다[1, 4-10].

3.2 직교 신경망

그림 1은 Yang과 Tseng[1]이 제안한 m 개의 입력층, N 개의 중간층, 그리고 1개의 출력층으로 구성된 직교 함수를 이용한 직교 신경망을 보인 것이다. 이러한 형태의 신경망을 수식으로 나타내면 식 (6)과 같다.

$$y = \sum_{i=1}^N w_i \Phi_i(X) = W^T \Psi(X) \quad (6)$$

여기서,

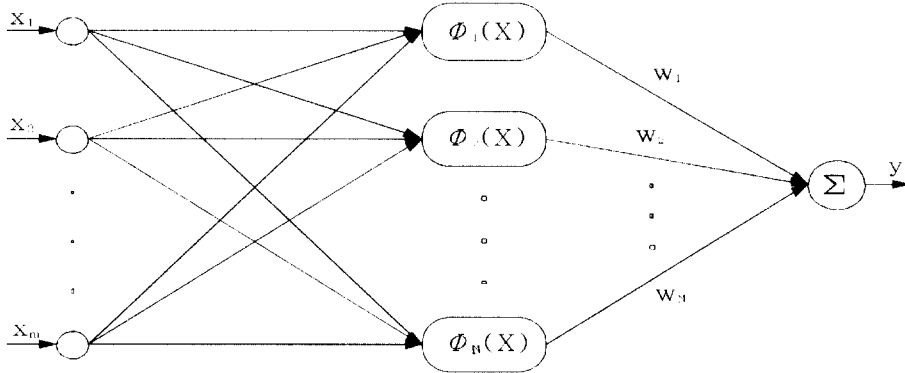


그림 1. Yang과 Tseng이 제안한 직교 신경망.
Fig. 1. Orthogonal neural network proposed by Yang and Tseng.

$$\Psi(X) = [\Phi_1(X) \ \Phi_2(X) \ \dots \ \Phi_N(X)]^T$$

$$X = [x_1 \ x_2 \ \dots \ x_m]^T$$

$$W = [w_1 \ w_2 \ \dots \ w_N]^T$$

이다.
각 입력 변수에 대해 n 차의 직교 함수로 신경망을 구성한 경우 중간층의 수 N 은 식 (7)과 같다.

$$N = n^n \quad (7)$$

직교 신경망은 요구되는 정확성을 기준으로 직교 함수의 급수를 증가시키므로 직교 신경망의 구조를 결정할 수 있다. 또한, 중간층과 출력층 사이에만 가중치가 존재하기 때문에 목적 함수를 식 (9)와 같이 설정하였을 때, 학습 알고리즘이 식 (10)과 같이 간단하며 속도도 빠르다. 그리고 가중치의 초기치는 모두 0으로 설정하므로 신경망의 초기화가 쉽다.

$$e = y - \hat{y} \quad (8)$$

$$J = \frac{1}{2} e^2 \quad (9)$$

$$\Delta W = \eta \frac{\partial J}{\partial W} = \eta e \frac{\partial e}{\partial W} = \eta e \Psi(X) \quad (10)$$

그러나 기존의 직교 신경망은 구조의 골격이 되는 직교 함수가 변형될 수 없는 구조를 가지고 있기 때문에 직교 신경망의 우수한 성능을 나타내게 하기 위해서는 시스템마다 적당한 직교 함수를 설정해 주어야 하지만 시스템의 수학적 모형을 모르는 상황에서는 매우 어렵다.

4. 수정된 직교 신경망

4.1 직교 함수의 변형

입력 변환을 이용함으로써 하여 직교 함수를 변형할 수 있음을 정리 1에서 보이고 이에따른 입력 변환의 조건을 제시한다.

정리 1. $\pi(x)$ 가 구간 $[a, b]$ 에서 서로 직교인 함수일 때, 연속함수 $T(\theta)$, $T: [\alpha, \beta] \rightarrow [a, b]$ 가 구간 $[\alpha, \beta]$ 에서 미분가능하고, $\frac{dT(\theta)}{d\theta} > 0$ 을 만족하면, $\sqrt{\frac{dT(\theta)}{d\theta}} \pi\{T(\theta)\}$ 또한 구간 $[\alpha, \beta]$ 에서 서로 직교인 함수이다.

증명 $x=T(\theta)$ 라 하면, 함수 $T(\theta)$ 가 구간 $[\alpha, \beta]$ 에서 미분가능하고 $\frac{dT(\theta)}{d\theta} > 0$ 이므로 역함수 $\theta=p(x) = T^{-1}(x)$ 가 존재하며

$$\frac{d\theta}{dx} = \frac{dp(x)}{dx} = p'(x) > 0$$

$$\frac{dx}{d\theta} = \frac{dT(\theta)}{d\theta} = T'(\theta) > 0$$

을 만족한다.

$$p'(x) = \frac{1}{T'(x)}$$

$$d\theta = \frac{1}{T'(\theta)} dx = p'(x) dx$$

$$\int_{\alpha}^{\beta} \sqrt{T'(\theta)} \phi_i\{T(\theta)\} \sqrt{T'(\theta)} \phi_j\{T(\theta)\} d\theta$$

$$= \int_{\alpha}^{\beta} \frac{1}{p'(x)} \phi_i(x) \phi_j(x) p'(x) dx$$

$$= \int_{\alpha}^{\beta} \phi_i(x) \phi_j(x) dx$$

$$= \begin{cases} 0 & i \neq j \\ A & i = j \end{cases}$$

따라서, $\sqrt{\frac{dT(\theta)}{d\theta}} \phi_i\{T(\theta)\}$ 는 구간 $[\alpha, \beta]$ 에서 서로

직교인 함수이다.

4.2 수정된 직교 신경망

수정된 직교 신경망의 구조는 그림 2와 같이 입력층, 출력층 그리고 J, K, L의 3개의 중간층, 모두 5개의 층으로 되어있다. 입력층 I는 m개의 뉴런으로 되어있으며, 중간층 J에서는 입력 변환 $T_i(x_i)$ 를 행하고 그 미분 $T'_i(x_i)$ 을 구한다. 그리고 중간층 K에서는 입력 변환된 값과 그 미분값으로부터, 미분값을 모두 곱한 값의 제곱근 ρ 와 직교 함수 $\Phi_k(X_T)$ 의 출력을 구하고, 중간층 L은 ρ 와 $\Phi_k(X_T)$ 의 출력을 곱하여 변형된 직교 함수의 출력을 구한다. 그리고 출력층 O는 변형된 직교 함수에 가중치를 곱한 값을 모두 더하여 전체 출력을 구한다. 이를 수식으로 나타내면 식 (11)과 같다.

$$y = W^T \Psi(X_T) \prod_{i=1}^m \sqrt{T'_i(x_i)}$$

$$= \rho W^T \Psi(X_T) \tag{11}$$

여기서,

$$\rho = \prod_{i=1}^m \sqrt{T'_i(x_i)}$$

$$W = [w_1 \ w_2 \ \dots \ w_N]^T$$

$$\Psi(X_T) = [\Phi_1(X_T) \ \Phi_2(X_T) \ \dots \ \Phi_N(X_T)]^T$$

$$X_T = [x_{T1} \ x_{T2} \ \dots \ x_{TN}]^T$$

$$x_{Ti} = T_i(x_i)$$

그리고 X_T 는 x_i 가 입력 변환된 값으로 이루어진 벡터

이고 N 은 전체 직교 함수의 개수로서 식 (12)와 같다.

$$N = n^m \tag{12}$$

여기서, m 은 입력 변수의 수, n 은 각 변수에 대한 단일변수 직교 함수의 차수를 각각 나타낸다.

본 논문에서는 식 (13)과 같은 단일 변수 레젠드레 다항함수(Legendre polynomial function)[1-3]의 곱으로된 다변수 함수를 기본적인 직교 함수 $\Pi_k(X_T)$ 로 사용하며, 레젠드레 다항함수의 도함수는 식 (14)와 같다.

$$L_1(x) = 1, L_2(x) = x$$

$$L_{i+1}(x) = a_i x L_i(x) - b_i L_{i-1}(x) \tag{13}$$

$$l_1(x) = 0, l_2(x) = 1$$

$$l_{i+1}(x) = a_i (L_i(x) + x l_i(x)) - b_i l_{i-1}(x) \tag{14}$$

$$a_i = 2 - 1/i, \ b_i = 1 - 1/i$$

$$-1 \leq i \leq 1 \tag{15}$$

여기서, L_i 는 레젠드레 다항함수, l_i 는 그 도함수를 나타내고 식 (15)은 레젠드레 다항함수가 직교성을 가지는 입력 범위를 나타낸다.

그림 3은 레젠드레 다항함수의 5차까지 그래프로 나타낸 것이다.

4.3 입력 변환 함수

많이 알려진 함수들 중에 입력 변환의 조건을 만족하는 것으로는 하이퍼볼릭 탄젠트 함수($\tanh(\cdot)$)와 탄젠트 함수($\tan(\cdot)$)등이 있다.

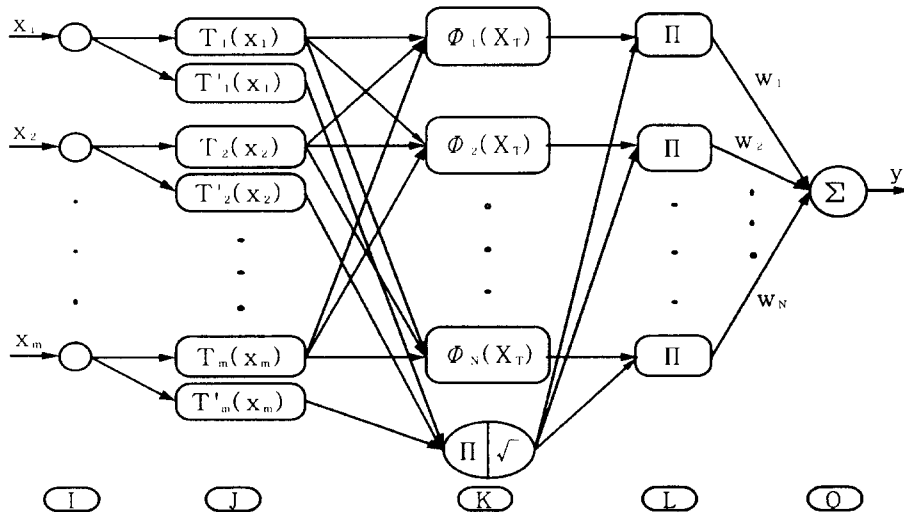


그림 2. 수정된 직교 신경망.
Fig. 2. Modified orthogonal neural network.

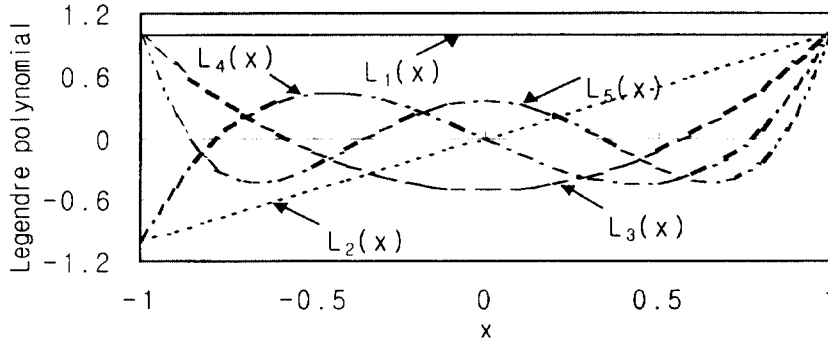


그림 3. 레젠드레 다항함수.
Fig. 3. Legendre polynomials.

$\tanh(x) = \frac{1 - \exp(-ax)}{1 + \exp(-ax)} \in (-1, 1)$ 가 감소하며, 탄젠트 함수는 반대로 기울기가 증가하는 특성을 가지고 있다. 그러나 이러한 함수들은 한 가지의 특성만을 나타내므로 미지의 시스템을 모형화하고 제어하기 위한 수정된 직교 신경망의 입력 변환 함수로서는 적절하지 못하다. 즉 시스템에 둘 중 어떠한 특성의 입력 변환이 더 적당한지 모르는 상태에서 한 가지 특성만을 가진 함수를 입력 변환으로 설정한다는 것은 옳지 못한 설정이라 할 수 있다.

따라서, x_i 에 대한 입력 변환 함수를 식 (16)과 같이 두 가지 특성을 모두 나타낼 수 있는 간단한 3차 다항 함수로 설정한다.

$$T_i(x_i) = \alpha_i x_{is}^3 + \beta_i x_{is} \quad (16)$$

여기서,

$$x_{is} = \frac{x_i - x_{i, \min}}{D_i/2}$$

$$D_i = x_{i, \max} - x_{i, \min}$$

$$x_{i, m} = \frac{x_{i, \max} + x_{i, \min}}{2}$$

본 논문에서는 레젠드레 다항함수를 기본 직교 함수로 사용하기 때문에 식 (16)의 α 와 β 는 $T_i(x_i)$ 의 형태를 결정하는 매개 변수로서 $T_i(x_i)$ 가 정의역 안에서 증가함수가 되도록, 그리고 치역은 $[-1, 1]$ 이 되도록 하는 범위 안에서 학습이 이루어져야 한다. 그러기 위해서 α 와 β 는 식 (17)과 (19)를 만족해야 한다.

$$\alpha_i + \beta_i = 1 \quad (17)$$

$$T_i'(x_i) = \frac{1}{D_i/2} [3\alpha_i x_{is}^2 + \beta_i] > 0,$$

$$\text{for } x_{i, \min} \leq x_i \leq x_{i, \max} \quad (18)$$

$$-0.5 < \alpha_i < 1, \quad 0 < \beta_i < 1.5 \quad (19)$$

그림 4는 본 논문에서 설정한 입력 변환 함수의 형태를 α 값의 변이에 따라 나타낸 것으로서 $0 < \alpha < 1$ 일 때는 $\tan(\cdot)$ 함수의 비슷한 형태를, $-0.5 < \alpha < 0$ 일 때는 시그모이드 함수의 형태를 나타낸다. 이러한 입력 변환을 이용하면 직교 함수를 주어진 비선형 시스템에 적당한 형태로 변형할 수 있다. 그리고, 그림 5는 위에서 설정한 입력 변환을 적용하여 레젠드레 다항함수를 변형한 형태 중 $D=2, x_m=0$ 일 때 서로 다른 α 와 β 에 대하여

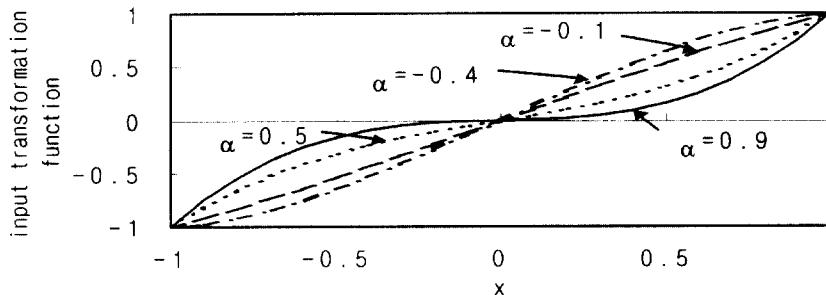


그림 4. 설정한 입력 변환 함수.
Fig. 4. Selected input transformation functions.

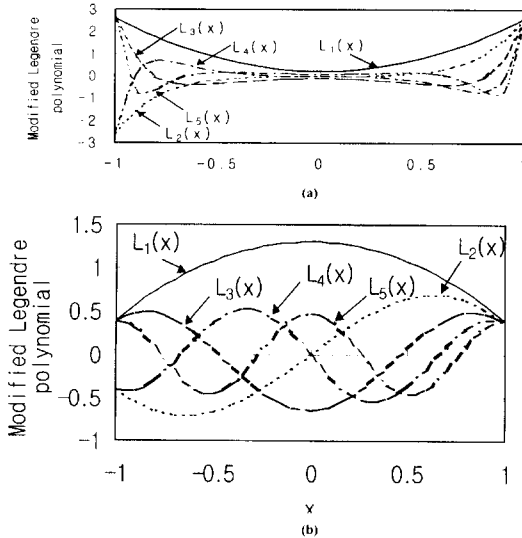


그림 5. 렌첸드레 당향함수에 설정한 입력변환을 적용한 함수.
 (a) $\alpha=0.8, \beta=0.2$
 (b) $\alpha=-0.3, \beta=1.3$

Fig. 5. Transformed functions of Legendre polynomials by selected input transformation functions.
 (a) $\alpha=0.8, \beta=0.2$
 (b) $\alpha=-0.3, \beta=1.3$

$\sqrt{\frac{dT(x)}{dx}} L_i\{T(x)\}$ 을 나타낸 것이다.

5. 시스템의 구성과 학습

플랜트의 동특성을 모형화하기 위한 목적 함수는 식 (21)과 같으며, 식별기의 학습은 목적 함수를 최소화하도록 하는 것이다.

$$e_j^i(k+1) = y^j(k+1) - \hat{y}^j(k+1) \quad (20)$$

$$J_I = \frac{1}{2} \sum_{j=1}^N e_j^i(k+1)^2 \quad (21)$$

여기서 $y^j(k)$ 는 플랜트의 출력, $\hat{y}^j(k)$ 은 식별기의 출력, $e_j^i(k)$ 는 플랜트와 식별기의 출력 오차를 각각 나타낸다.

식별기를 최적화하기 위한 가중치 $W_i = [w_{i1} \ w_{i2} \ \dots \ w_{iN}]^T$ 의 증분량 ΔW_i 는 식 (22), (23)과 같이 구한다.

$$\begin{aligned} \frac{\partial J_I}{\partial W_i} &= \sum_{j=1}^N \frac{\partial J_I}{\partial e_j^i(k+1)} \frac{\partial e_j^i(k+1)}{\partial W_i} \\ &= - \sum_{j=1}^N e_j^i(k+1) \rho_i^j(k) \Psi_i(X_T^j(k)) \end{aligned} \quad (22)$$

$$\begin{aligned} \Delta W_i &= \eta_{w_i} \left(- \frac{\partial J_I}{\partial W_i} \right) \\ &= \eta_{w_i} \sum_{j=1}^N e_j^i(k+1) \rho_i^j(k) \Psi_i(X_T^j(k)) \end{aligned} \quad (23)$$

여기서,

$$\begin{aligned} \Psi_i(X_T^j(k)) &= [\text{Phi}_{i1}(X_T^j(k)) \ \text{Phi}_{i2}(X_T^j(k)) \ \dots \ \text{Phi}_{iN}(X_T^j(k))]^T \\ \rho_i^{20j}(k) &= \prod_{i=1}^m \sqrt{T_{ii}'(x_i^j(k))} \end{aligned}$$

이며, η_{w_i} 는 식별기 가중치의 학습률로서 양의 작은 값이다.

식별기의 입력 변환 계수인 α_{ii} 와 β_{ii} 에 대한 목적 함수의 변화율을 구하면 식 (24), (25)와 같다.

$$\begin{aligned} \frac{\partial J_I}{\partial \alpha_{ii}} &= \sum_{j=1}^N \frac{\partial J_I}{\partial e_j^i(k+1)} \frac{\partial e_j^i(k+1)}{\partial \alpha_{ii}} \\ &= - \sum_{j=1}^N e_j^i(k+1) \left(\rho_i^j(k) \sum_{p_1=1}^{n_i} \sum_{p_2=1}^{n_i} w_{i,p_1 \dots p_2} \right) \end{aligned} \quad (24)$$

$$\begin{aligned} &\prod_{q=1, q \neq i}^m \phi_{iqp_q}(x_{Tq}^j(k)) \times \frac{\partial \phi_{iip_i}(x_{Ti}^j(k))}{\partial x_{Ti}^j(k)} \frac{\partial x_{Ti}^j(k)}{\partial \alpha_{ii}} \\ &+ \frac{\rho_i^j(k)}{T_{ii}'(x_i^j(k))} W_i^T \Psi_i(X_T^{20j}(k)) \frac{\partial}{\partial \alpha_{ii}} T_{ii}'(x_i^j(k)) \end{aligned}$$

$$\begin{aligned} \frac{\partial J_I}{\partial \beta_{ii}} &= \sum_{j=1}^N \frac{\partial J_I}{\partial e_j^i(k+1)} \frac{\partial e_j^i(k+1)}{\partial \beta_{ii}} \\ &= - \sum_{j=1}^N e_j^i(k+1) \left(\rho_i^j(k) \sum_{p_1=1}^{n_i} \sum_{p_2=1}^{n_i} w_{i,p_1 \dots p_2} \right) \end{aligned} \quad (25)$$

$$\begin{aligned} &\prod_{q=1, q \neq i}^m \phi_{iqp_q}(x_{Tq}^j(k)) \times \frac{\partial \phi_{iip_i}(x_{Ti}^j(k))}{\partial x_{Ti}^j(k)} \frac{\partial x_{Ti}^j(k)}{\partial \beta_{ii}} \\ &+ \frac{\rho_i^j(k)}{T_{ii}'(x_i^j(k))} W_i^T \Psi_i(X_T^{20j}(k)) \frac{\partial}{\partial \beta_{ii}} T_{ii}'(x_i^j(k)) \end{aligned}$$

또한, α_{ii} 와 β_{ii} 는 식 (17)을 항상 만족하여야 하므로 α_{ii} 와 β_{ii} 의 증분량 $\Delta \alpha_{ii}$ 와 $\Delta \beta_{ii}$ 는 식 (26)을 만족하여야 하며 식 (24)와 (25)로부터 식 (27), (28)과 같이 구할 수 있다.

$$\Delta \alpha_{ii} + \Delta \beta_{ii} = 0 \quad (26)$$

$$\Delta \beta_{ii} = \frac{\eta_{\beta_{ii}}}{2} \left(\frac{\partial J_I}{\partial \alpha_{ii}} - \frac{\partial J_I}{\partial \beta_{ii}} \right)$$

$$= \frac{\eta_{T_i}}{2} \sum_{j=1}^N e_i^j(k+1) \left(\rho_i^j(k) \sum_{p_1=1}^{n_1} \sum_{p_2=1}^{n_2} w_{T_i, p_1 \dots p_2} \right. \\ \left. \prod_{q=1, q \neq i}^m \phi_{i, q, p_q}(x_{T_i}^j(k)) \right. \\ \left. \times \frac{\partial \phi_{i, p_q}(x_{T_i}^j(k))}{\partial x_{T_i}^j(k)} \frac{\partial x_{T_i}^j(k)}{\partial \beta_{T_i}} [x_{T_i}^j(k)^3 - x_{T_i}^j(k)] \right. \\ \left. + \frac{3x_{T_i}^j(k)^2 - 1}{3\alpha_{T_i} x_{T_i}^j(k)^2 + \beta_{T_i}} \rho_i^j(k) W_{T_i}^T \Psi_{T_i}(X_{T_i}^j(k)) \right) \quad (27)$$

$$\Delta \beta_{T_i} = -\frac{\eta_{T_i}}{2} \left(\frac{\partial J_{T_i}}{\partial \beta_{T_i}} - \frac{\partial J_{T_i}}{\partial \beta_{T_i}} \right) = -\Delta \alpha_{T_i} \quad (28)$$

여기서, η_{T_i} 는 식별기의 입력 변화 계수인 α_{T_i} 와 β_{T_i} 의 학습률로서 양의 작은 값이다.

6. 사례연구와 검토

제안한 수정된 직교 신경망의 비선형 시스템 모형화 능력, 3차 다항함수로 설정한 입력 변화의 유용성을 검증한다.

제안한 알고리즘을 식 (29)와 같이 표현되는 비선형 시스템 모델[11]에 적용하여 근사화 성능을 조사하였다.

$$y(k+1) = \frac{y(k)y(k-1)[y(k)+2.5]}{1+y^2(k)+y^2(k-1)} + u(k) \quad (29)$$

식별기는 $y(k)$, $y(k-1)$, $u(k)$ 를 입력으로하고, 각 입력에 대하여 직교 함수를 5차까지로 구성하여 학습하였다.

표 1. 식별기의 구조와 오차 제곱 평균

Table 1. Structure and mean square error of identifiers

	구조(가중치수)	오차 제곱 평균
다층 신경망	3-50-50-1 (2700)	4.81410×10^{-3}
기존의 직교신경망	3-125-1 (125)	6.65255×10^{-3}
수정된 직교 신경망	3-6-126-125-1 (131)	5.66231×10^{-6}

표 2. 수정된 직교 신경망을 이용한 식별기의 입력 변환 계수

Table 2. Input transformation parameters of MONN

	α	β
$y(k)$	-0.0987895378	1.0987895378
$y(k-1)$	-0.1047325766	1.1047325766
$u(k)$	0.0078421515	0.9921578485

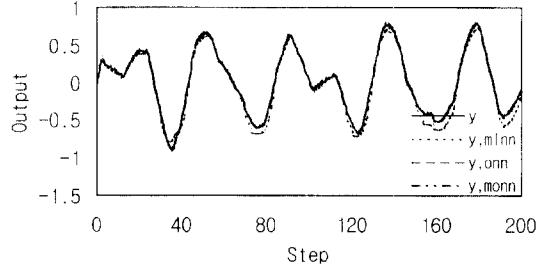


그림 6. 플랜트의 출력과 신경망을 이용한 식별기의 출력.
Fig. 6. Output of the plant and identifier.

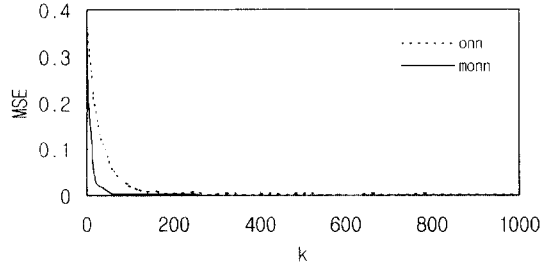


그림 7. 직교신경망을 이용한 식별기와의 오차제곱평균 비교.
Fig. 7. Mean Squared Error comparison with ONN identifier.

표 1은 학습 후, 기존의 방식과 수정된 직교 신경망으로 학습한 식별기의 오차 제곱 평균(MSE)를 비교한 것이다. 그리고 표 2는 식별기의 학습 후, 식별기로 사용한 수정된 직교 신경망의 입력 변화 계수를 나타낸 것이다.

그림 7는 임의의 입력에 대한 플랜트의 출력과 식별기의 출력을 보인 것이다. 그리고 그림 8은 기존의 방식과 수정된 직교 신경망을 이용한 식별기의 오차 제곱 평균의 변화를 비교한 것으로서 수정된 직교 신경망의 학습이 더 빠르다는 것을 알 수 있다.

표 1과 그림 6, 7로부터 제안한 수정된 직교 신경망의 비선형 시스템 모형화 능력이 기존의 다층 신경망과 직교 신경망에 비해 더 우수하다는 것을 알 수 있다.

7. 결 론

직교 함수의 급수를 이용하여 구성한 직교 신경망은 다층 신경망이 가지는 초기 가중치를 설정함에 있어서 랜덤성을 만족시켜야되는 어려움과 중간층의 수와 각 층의 뉴런 수를 결정하는데 있어서의 어려움과 같은 문제를 해결할 수 있다. 그리고 기존의 직교 신경망은 직교 신경망의 정밀도를 결정하는 직교 함수의 형태를 변형할 수 없는 구조를 가지고 있다.

본 논문에서는 입력 변환을 이용하여 직교 함수의 형태를 변형하는 방법을 제시하였고 여기에 수반되는 입력 변환의 조건을 제시하였다. 그리고 이러한 입력 변환을 이용하여 직교 함수를 적당한 형태로 학습할 수 있는 새로운 구조를 가지는 수정된 직교 신경망을 제안 하였다. 제안한 직교 신경망은 기존의 직교 신경망이 가지는 여러 문제점을 해결하였다.

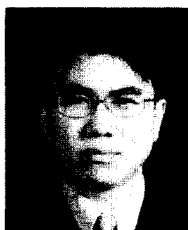
본 논문에서 제안한 입력 변환을 만족하는 3차 다항함수를 수정된 직교 신경망의 입력 변환 함수로 설정하였으며, 비선형 시스템을 식별하기 위해 학습 알고리즘을 유도하였다.

사례연구를 통하여 본 논문에서 제안한 수정된 직교 신경망의 비선형 시스템 모형화 능력, 입력변환 능력 그리고 학습 알고리즘의 유용성을 검증함으로써 제안한 수정된 직교 신경망은 입력 변환을 이용하여 직교 함수를 변형함으로써 기존의 직교 신경망과 다층 신경망에 비해 비선형 시스템을 모형화하는 능력이 더 우수하다는 결론을 얻을 수 있었다.

참고문헌

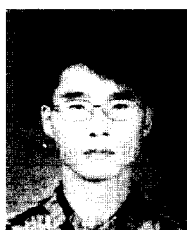
[1] Shiow-Shung Yang and Ching-Shiow Tseng, "An Orthogonal Neural Network for Function Approximation," *IEEE Trans. Systems, Man, and Cybernetics*, **26**(5), pp. 779-785, 1996.
 [2] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, New York : Interscience Publishers, pp.

49-111.
 [3] F. B. Hildebrand, *Advanced Calculus for Applications*, Englewood Cliffs, NJ : Prentice-Hall.
 [4] Y. Fu and G. A. Dumont, "An Optimum Time Scale for Discrete Laguerre Network," *IEEE Trans. Automatic Control*, **38**(6), pp. 934-938, 1993.
 [5] Peter S. C. Heuberger, Paul M. J. Van den Hof and Okko H. Bosgra, "A Generalized Orthonormal Basis for Linear Dynamical Systems," *IEEE Trans. Automatic Control*, **40**(3), pp. 451-465, 1995.
 [6] Pertti M. M kil , "Approximation of Stable Systems by Laguerre Filters," *Automatica*, **26**(2), pp. 333-345, 1990.
 [7] Guy A. Dumont, Christos C. Zervos and Gerry L. Page, "Laguerre-based Adaptive Control of pH in an Industrial Bleach Plant Extraction Stage," *Automatica*, **26**(4), pp. 781-787, 1990.
 [8] C. Zervos, P. R. B langer and G. A. Dumont, "On PID Controller Tuning using Orthonormal Series Identification," *Automatica*, **24**(2), pp. 165-175, 1988.
 [9] P. M. M kil, "Laguerre Series Approximation of Infinite Dimensional Systems," *Automatica*, **26**(6), pp. 985-995, 1990.
 [10] Paul M. J. Van den Hof, Peter S. C. Heuberger and J zsef Bokor, "System Identification with Generalized Orthonormal Basis Functions," *Automatica*, **31**(12), pp. 1821-1834, 1995.
 [11] K. S. Narendra and K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE Trans. Neural Networks*, **1**(1), pp. 4-27, 1990.



이영석(Sung-Sik Kim)

1991년 : 경북대 공대 전기공학과 졸업
 1994년 : 동 대학원 전기공학과 졸업(석사)
 1996년 : 동 대학원 전기공학과 박사과정 수료
 1997년 9월~현재 : 영진전문대 전기과 전임강사



김성식(Young-Seog Lee)

1996년 : 경북대 공대 전기공학과 졸업
 1998년 : 동 대학원 전기공학과 졸업 (석사)