

직교 신경망을 이용한 비선형 시스템의 제어

김성식*, 이영석, 안대찬, 서보혁
경북대학교 전기공학과

Nonlinear System Control Using Othogonal Neural Network

Sung-Sik Kim, Young-Seog Lee, Dae-Chan Ahn, Bo-Hyeok Seo
Department of Electrical Engineering, Kyungpook National University

Abstract - This paper presents an Orthogonal Neural Network based on orthogonal functions and applies the network to nonlinear system control. The Orthogonal Neural Network doesn't have the problems of traditional feedforward neural networks such as the determination of initial weights and the numbers of layers and processing elements. In this paper, Orthogonal Neural Network is modified already introduced one by input transformation. The results show that the modified neural network has the better performance than existing one and performance of controller using this network is good.

1. 서 론

일반적인 제어 기법은 대상 시스템의 수학적 모델을 특정 동작점에서 선형화한 선형 모델로부터 제어기를 설계한다. 그러나 이러한 기법은 제어대상 플랜트가 고차의 미분방정식으로 표현되는데 따른 복잡성과 플랜트의 비선형성, 그리고 플랜트 모델링의 부정확성으로 인해 만족할만한 제어 성능을 나타내기에는 어려움이 있다. 최근 이러한 어려움들을 극복하기 위하여 인간의 뇌 신경을 모델링한 인공 신경망을 응용한 제어기가 도입되어 많은 주목을 받고 있다. 그러나 기존의 신경망은 초기 가중치를 설정하는데 어려움이 있을 뿐만 아니라 숨은 층의 수와 각 층의 뉴런 수를 결정하는데 있어서도 어려움을 가지고 있다. 직교 함수의 급수를 이용한 직교 신경망은 이러한 어려움 없이 전체 구조를 결정할 수 있다는 커다란 장점을 가지고 있다[1]. 또한 요구되는 정확성을 기준으로 직교 함수의 급수를 늘림으로써 학습하는 동안에 전체 구조를 변형하는 것이 가능하다. 그러나 이러한 직교 신경망에 있어서도 구조의 골격이 되는 직교 함수를 설정하는 데에 있어서 한계가 있다.

본 논문에서는 입력 변환을 갖는 수정된 구조의 직교 신경망을 제안하여 이러한 한계를 해결하였으며 이를 이용하여 식별기와 제어기로 이루어진 비선형 시스템 제어기를 구현하고자 한다. 그리고 사례 연구를 통하여 그 특성의 우수함을 보이고자 한다.

2. 직교 신경망

2.1 기본 이론

다변수 함수 $g(x_1, x_2, x_3, \dots, x_m)$, $g: S^m \rightarrow R$, $S^m = S_1 \times S_2 \times \dots \times S_m$ 에 대하여 식(1),(2)를 만족하는 $G_N(X)$ 가 존재한다[2,3].

$$G_N(X) = \sum_{i=1}^N w_i \phi_i(X) \quad (1)$$

$$\lim_{N \rightarrow \infty} \int_{S_1} \dots \int_{S_m} (g(X) - G_N(X))^2 dx_m \dots dx_1 = 0 \quad (2)$$

여기서,

$$\phi_i(X) = \prod_{k=1}^m \phi_{k,p_k}(x_k) \quad (3)$$

$\phi_{k,p_k}(x_k)$: k 번째 변수 x_k 에 대한 p_k 번째 직교 함수

$$X = [x_1, x_2, x_3, \dots, x_m]^T \quad (4)$$

이다. 그리고 식(3)의 $\phi_i(X)$ 는 공간 S^m 에 대하여 서로 직교인 함수이다.

정리 1.

$\phi_i(x)$ 가 구간 $[a, b]$ 에서 서로 직교인 함수일때, 함수 $h(t)$, $h: [a, \beta] \rightarrow [a, b]$ 가 구간 $[a, \beta]$ 에서 미분가능하고, $\frac{dh(t)}{dt} > 0$ 을 만족하면, $\sqrt{\frac{dh(t)}{dt}} \phi_i\{h(t)\}$ 또한 구간 $[a, \beta]$ 에서 서로 직교인 함수이다.

증명)

$$x = h(t) \quad (5)$$

라 하면, 함수 $h(t)$ 가 구간 $[a, \beta]$ 에서 미분가능하고

$$\frac{dh(t)}{dt} > 0 \text{ 이므로 역함수}$$

$$t = p(x) = h^{-1}(x) \quad (6)$$

가 존재하며

$$\frac{dt}{dx} = \frac{dp(x)}{dx} = p'(x) > 0 \quad (7)$$

을 만족한다.

$$\frac{dx}{dt} = \frac{dh(t)}{dt} = h'(t) > 0 \quad (8)$$

$$p'(x) = \frac{1}{h'(x)} \quad (9)$$

$$dt = \frac{1}{h'(t)} dx = p'(x) dx \quad (10)$$

이므로

$$\begin{aligned} & \int_a^\beta \sqrt{h'(t)} \phi_i\{h(t)\} \sqrt{h'(t)} \phi_j\{h(t)\} dt \\ &= \int_a^\beta \phi_i(x) \phi_j(x) dx \\ &= \begin{cases} 0 & i \neq j \\ A_i & i = j \end{cases} \end{aligned} \quad (11)$$

따라서, $\sqrt{\frac{dh(t)}{dt}} \phi_i\{h(t)\}$ 는 구간 $[a, \beta]$ 에서 서로 직교인 함수이다. ■

2.2 직교 신경망의 구조

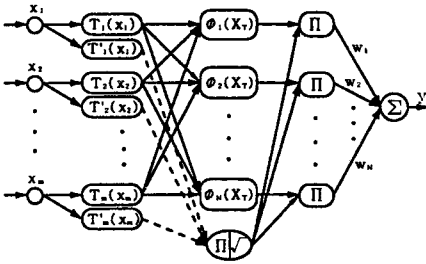


그림 1. 직교 신경망의 구조

본 논문에서 제안하는 직교 신경망의 구조는 그림 1과 같다. 입력단의 입력이 $T(x)$ 에 의해 변환된 값으로 직교 함수 $\phi(X_T)$ 에 입력된다. 그리고 $T(x)$ 의 미분값의 곱을 제공된 취한 값과 직교함수의 출력을 곱하고 여기에 가중치를 곱한 것이 전체 출력이 된다. 이를 수식으로 나타내면 식 (12)와 같다.

$$y = \prod_{i=1}^n T_i(x_i) \sum_{k=1}^m w_k \phi_k(X_T) \quad (12)$$

여기서 X_T 는 x_i 의 입력 변환된 값으로 이루어진 벡터이다. 이러한 구조는 기존의 직교 신경망[1]에 정리 1을 적용할 수 있도록 한 것이다. 그림 1에서 알 수 있듯이 직교 신경망에서는 가중치가 출력단과 연결된 곳에만 존재하기 때문에 학습 규칙이 간단하다. 본 논문에서는 식 (13)의 레젠드레 다항함수(Legendre polynomial function)[1-3]를 기본적인 직교 함수 급수로 사용하였고, 그 도함수는 식 (14)와 같다. 또한 입력 변환은 식 (15)을 사용하였다.

$$\begin{aligned} L_1(x) &= 1, \quad L_2(x) = x \\ L_{i+1}(x) &= a_i x L_i(x) - b_i L_{i-1}(x) \end{aligned} \quad (13)$$

$$\begin{aligned} l_1(x) &= 0, \quad l_2(x) = 1 \\ l_{i+1}(x) &= a_i (L_i(x) + x l_i(x)) - b_i l_{i-1}(x) \end{aligned} \quad (14)$$

여기서, $a_i = 2 - 1/i$, $b_i = 1 - 1/i$

$$T_i(x_i) = \alpha x_i^3 + \beta x_i \quad (15)$$

L_i 는 레젠드레 다항함수, l_i 는 그 도함수를 나타내며, 식 (15)의 α 와 β 는 x_i 의 정의역 안에서 증가함수가 되도록, 그리고 치역은 $[-1, 1]$ 이 되도록 설정하여야 한다. 이러한 입력 변환을 이용해 직교함수 급수를 변형할 수 있다는 장점이 있다.

3. 제어 시스템의 구성과 학습

직교 신경망을 이용하여 비선형 시스템을 제어하기 위

한 제어 시스템은 그림 2와 같이 직교 신경망으로 된 식별기와 제어기로 구성된다. 그리고 식별기와 제어기는 모두 오프-라인으로 학습한다.

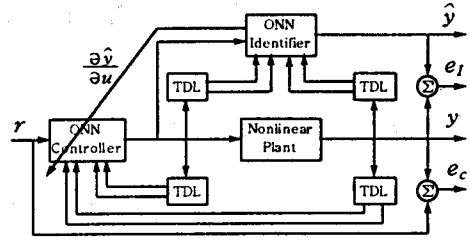


그림 2. 제어 시스템의 구성

3.1 식별기의 학습

식별기의 학습 목적은 식 (16)으로 정의한 목적 함수를 최소화 하는 것이다.

$$J_I = \frac{1}{2} (y - \hat{y})^2 = \frac{1}{2} e_I^2 \quad (16)$$

여기서 y 는 플랜트의 출력, \hat{y} 은 식별기의 출력, e_I 는 플랜트와 식별기의 출력 오차를 각각 나타낸다. 그리고, 가중치의 증분량은 식 (18)로 구한다.

$$\frac{\partial J_I}{\partial w_I} = -e_I \phi_I \quad (17)$$

$$\begin{aligned} \Delta w_I &= \eta \left(-\frac{\partial J_I}{\partial w_I} \right) \\ &= \eta e_I \phi_I \end{aligned} \quad (18)$$

여기서 η 는 학습률로서 아주 작은 양의 값이다.

3.2 제어기의 학습

제어기의 학습은 식별기로부터 시스템의 정보를 받아서 식 (19)로 정의한 목적 함수를 최소화 하는 것이다.

$$J_C = \frac{1}{2} (r - y)^2 = \frac{1}{2} e_c^2 \quad (19)$$

여기서 r 은 제어기의 기준입력이고 e_c 는 기준입력과 플랜트의 출력 오차이다. 제어기 가중치의 증분량을 구하기 위해서는 플랜트의 입력에 대한 출력의 변화량을 알아야 하지만 미지의 플랜트라 가정하였기 때문에 식별기로부터 그 추정치를 받아서 식 (22)과 같이 구한다.

$$\frac{\partial J_C}{\partial w_c} = -e_c \frac{\partial y}{\partial u} \phi_c \quad (20)$$

$$\Delta w_c = \eta \left(-\frac{\partial J_C}{\partial w_c} \right) = \eta e_c \frac{\partial y}{\partial u} \phi_c \quad (21)$$

$$\Delta w_c = \eta e_c \frac{\partial \hat{y}}{\partial u} \phi_c \quad (22)$$

여기서 플랜트의 입력에 대한 출력의 변화량의 추정치 $\frac{\partial \hat{y}}{\partial u}$ 는 $x_1 = u$ 라 하면 식 (23)으로 된다.

$$\frac{\partial \hat{y}}{\partial x_1} = \sum_{p_1=1}^m \dots \sum_{p_n=1}^m w_{p_1 \dots p_n} \phi_{2p_1} \dots \phi_{mp_n} \frac{\partial \phi_{1p_1}}{\partial x_1} \quad (23)$$

3. 사례연구

본 논문의 사례연구에서 설정한 플랜트는 비선형 시스템 제어에서 대표적인 것 중 하나인 식 (24)으로 표현되는 역추(Inverted Pendulum)이다.[4]

$$\begin{cases} (M+m)\ddot{x} + mL\cos\theta\ddot{\theta} - mL\sin\theta\dot{\theta}^2 - u = 0 \\ m\cos\theta\ddot{x} + mL\ddot{\theta} - mg\sin\theta = 0 \end{cases} \quad (24)$$

여기서 M 은 수레의 질량(50g), m 은 추의 질량(30g), l 은 추의 길이(2m)이고 x 는 수레의 위치, θ 는 추의 기울어진 각도, u 는 수레를 미는 힘으로서 플랜트의 입력을 각각 나타낸다. 4차 룬게쿠타(Runge-Kutta) 방법을 사용하여 플랜트의 출력을 얻었으며 이산시간 구간은 0.1초이다. 그리고 식별기는 $y_k, y_{k-1}, y_{k-2}, u_k, u_{k-1}$ 을 입력으로하여 각각 4차 급수까지로 구성하였으며, 제어기는 y_k, y_{k-1}, y_{k-2} 를 입력으로 각 8차 급수까지로 구성하였다.

식별기를 학습하는 데에 있어서 플랜트의 동작 범위를 $-1rad \leq \theta \leq 1rad$, 입력 u 의 범위를 ± 1.5 로 제한하였으며 기준 입력은 θ 를 영이 되도록 하고자 한다.

제어기를 학습하기 위해 식 (25)와 같은 목적 함수를 하나 더 설정하여 입력이 과도한 값을 갖지 않도록 전체 목적 함수를 식 (26)으로 두었다.

$$J_{c,2} = \frac{1}{2} u^2 \quad (25)$$

$$J = k_o J_c + k_u J_{c,2} \quad k_o = 1, \quad k_u = 0.1 \quad (26)$$

그림 3과 4는 식별기 학습 완료 후 정현파 입력에 대한 플랜트의 출력과 식별기의 출력을 비교한 것이다.

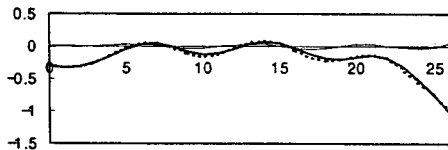


그림 3. 기존의 직교 신경망 학습 결과
점선: 플랜트 출력, 실선: 식별기 출력, 가선: 오차

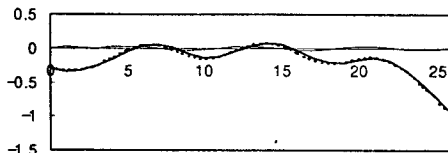


그림 4. 제안한 직교 신경망 학습 결과

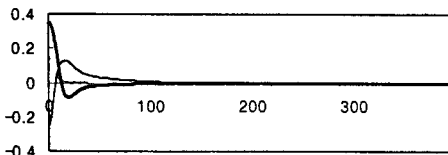


그림 5. 플랜트의 출력

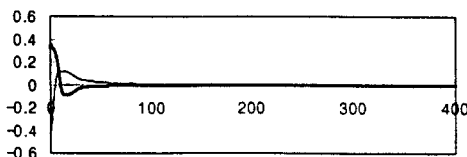


그림 6. 플랜트의 입력

그림 3은 입력 변환이 없는 기존의 직교 신경망으로

학습한 것으로서 오차 제곱 평균이 0.0006522인데 반하여 그림 4는 y_k 에 대하여 $\alpha = -0.1, \beta = 1.1$ 의 입력 변환을 하여 학습한 것으로서 오차 제곱 평균이 0.0004026으로 플랜트와 더 가깝게 학습되었음을 나타낸다. 출력은 모두 rad값으로 하였다.

그림 5와 6은 플랜트의 출력[rad]과 입력[N]을 나타낸다. 플랜트의 출력이 작을 때 좀 더 섬세한 제어가 가능하도록 y_k 에 대하여 $\alpha = -0.4, \beta = 1.4$ 의 입력 변환을 하였으며 결과가 만족할만하다는 것을 알 수 있다.

5. 결 론

본 논문에서는 정리 1이 성립함을 보였고 이를 적용해서 기존의 직교 신경망에 입력 변환을 가지는 수정된 형태의 직교 신경망을 제안하였으며, 이러한 수정된 직교 신경망을 이용한 식별기와 제어기로 구성된 제어 시스템을 구현하였다. 식별기는 플랜트 출력과의 오차를 최소화 하도록 학습하였으며 제어기는 플랜트의 미분 정보를 식별기로부터 받아서 플랜트의 출력과 기준 입력의 오차를 최소화 하도록 학습하였다. 그리고 사례연구로서 비선형 시스템의 대표적인 예 중 하나인 역추 균형문제를 통해 제안한 직교 신경망을 기존의 것과 비교하여 그 성능의 우수함을 보였으며 제어 결과도 만족할만하다는 것을 알 수 있었다.

[참 고 문 헌]

- [1] Shioh-Shung Yang and Ching-Shioh Tseng, "An Orthogonal Neural Network for Function Approximation", *IEEE Trans. Systems, Man, and Cybernetics*, vol. 26, no. 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] 최면승, 박영문, "인공 신경회로망을 이용한 비선형 시스템의 최적추적제어기" *대한전기학회 논문지*, vol 43, no. 1, pp. 161-171, 1994
- [5] K. S. Narendra and K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks", *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 4-27, 1990.
- [6] M. S. Chen and M. T. Manry, "Conventional modeling of the multilayer perceptron using polynomial basis functions", *IEEE Trans. Neural Networks*, vol. 4, no. 1 pp. 164-166, 1993.
- [7] P. Baldi, "Gradient Descent Learning Algorithm Overview : A General Dynamical Systems Perspective", *IEEE Trans. Neural Networks*, vol. 6, no. 1, pp. 182-195, 1995
- [8] 이영석, 김성식, 서보혁, "유전 알고리즘을 이용한 최적의 물 맵핑을 가지는 퍼지 제어기 설계", *대한전기학회 추계학술대회 논문집*, pp. 68-70, 1996.