신경회로망을 이용한 최적제어 설계방식

권성훈*, 이인재*. 정지원*, 김한웅*, 엄기환*
*동국대학교 전자공학과
Email : kwon2@cakra.dongguk.ac.kr

Optimal Control Design Method Using the Neural Network

Sung-Hoon Kwon*, In-Jae Lee*, Ji-Won Jung*, Han-Woong Kim*, Ki-Hwan Eom*
*Dept. of Electronic Eng., DongGuk Univ.
Email : kwon2@cakra.dongguk.ac.kr

요 약

최근 여러 분야에서 필요성이 증대되고 있는 최적제어, 최적설계 및 최적추정에 대하여 신경회로망을 이용하는 방식을 제안하였다. 제안한 방식은 3점 탐색 알고리즘을 사용하여 성능지표의 최소값을 찾아내어 이 때의 가중치를 신경회로망의 가중치 초기값으로 설정하여 다시 학습시킴으로서 보다 효율적인 학습결과를 얻을 수 있었다. 제안한 방식을 최적제어에 대하여 시뮬레이션하여 유용성을 확인하였다.

I. 서론

최근의 최적제어, 최적설계 및 최적추정은 이공학 분야 뿐만 아니라 사회과학 분야 등 여러 분야에서 필요성이 증대되고 있으며, 여기에 필요한 목적함수의 수치계산법의 필요성은 점점 커지고 있다. 일반적으로 목적함수는 비선형 다변수로 구성된 한 개의 범함수로서 이 함수의 극치(최소치 또는 최대치)를 산출하는 것은 쉽지 않으며, 이에 대한 알고리즘이 여러 가지 개발되어 있다. [2]

그러나 이들은 보조 방정식의 산출 및 프로그램 등의 복잡성과 정확성 및 계산속도 등 여러 가지 장단점이 있다. 특히 반복계산법이 필요하므로 컴퓨터에 입력하는 프로그램의 작성에 많은 노력이 필요하다. [1][3]

본 논문에서는 신경회로망을 이용한 최적설계 방식을 제안한다. 제안한 방식은 신경회로망의 학습을 위하여 직접탐색 알고리즘인 3점 탐색 알고리즘을 이용하여 가중치의 초기값을 탐색하여 최적화 한다. 제안한 방식의 유용성을 확인하기 위하여 레귤레이터 문제에 적용하여 시뮬레이션 한다.

II. 신경회로망

신경회로망(Neural Network)은 생물학적 뉴런의 구조와 신호의 전달방법을 공학적인 모델링의 구조로 응용한 공학적인 하나의 도구이다. 신경회로망의 구조는 간단한 처리요소들(processing elements)이 생물학적인 신경에서와 같이 다양한 방법으로 연결되어 있다. 처리요소를 절점(node)이라 부르는데 이 절점에 여러 다른 절점들이 연결 고리로 연결되어 회로망을 형성한다. 절점에는 입력과 연결된 고리도 있고 출력과 연결된 고리도 있다. 입력을 통해서 들어온 수치를 다양한 수

학적인 방법으로 처리하고 처리된 값을 응용하고자 하는 시스템의 특성에 적합한 함수에 통과시킨다. 여기에서 발생되는 수치가 각 절점의 출력이 되어 다른 절점들과 연결되거나 신경회로망의 출력값이 되기도 한다. [6]

역전파 신경회로망(Back-Propagation Neural Network)은 선형 뉴런의 입력층과 출력층 사이에 하나 이상의 중간층, 즉 비선형 뉴런층으로 구성된 은닉(hidden)층을 갖는 다층(multi-layer) 구조의 회로망이다. 입력층의 입력신호들은 은닉층의 뉴런을 거쳐 출력층으로 전달되는 전방향 회로망이다. 출력층의 출력신호는 신경회로망이 추종하고자 하는 목표출력 신호와 비교되어 오차신호를 발생하고, 이 오차의 제곱을 최소화하도록 출력층으로부터 은닉층을 거쳐 입력층으로 역전파 방식으로 연결강도를 조정한다. [4]

하나의 은닉층을 가진 일반적인 다층 신경회로망의 구조는 다음과 같다.

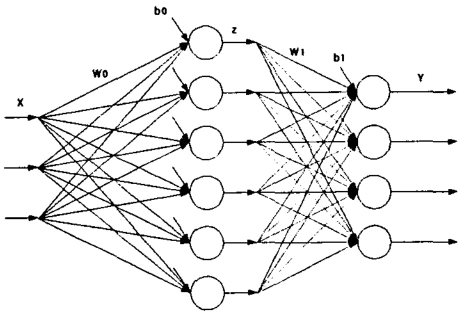

그림 1. 다층 신경회로망의 구조

III. 3점 탐색 알고리즘[5]

통신계나 전기계 및 기계계나 장치 등의 최적설계에 적용되는 성능지표(performance index 또는 cost function) J는 일반적으로 식(12)와 같은 2차형식함수(quadratic function)로 구성되는 경우가 많다.

$$J = X \cdot R \cdot x, \quad x = [x_1,\ x_2,\ \cdots,\ x_n] \quad (12)$$

여기서 x는 설계에 필요한 독립변수 $x_1, x_2, \cdots, x_n$로 구성된 벡터(vector)이며, 변수가 n개일 때는 n차의 벡터이고, R은 $n \times n$개의 변수를 갖는 정정행렬(positive definition matrix)이다. 그리고, 취급한 함수는 설정한 탐색범위 내에서 극치가 1개만 존재하는 unimodality형인 경우의 단일변수와 다변수(4개 변수)로 된 성능지표이며 적용하는 3점 탐색법은 다음과 같다.

(1) 1차 벡터인 경우

함수(성능지표) J의 변수 x가 1차 $x = [x_1]$인 경우이며, 그림 1과 같이 탐색의 반복횟수를 w라하고 각 회수에 따라 설명하면 다음과 같다.

(i) 초기탐색(w=0)

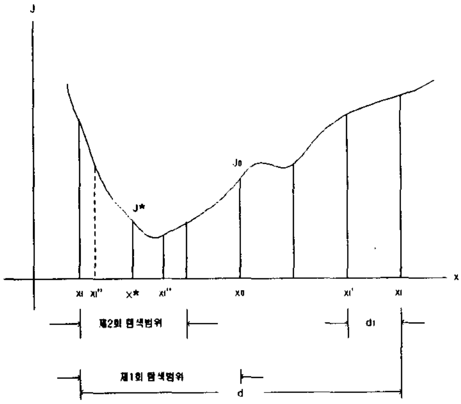

그림 2. 1차원 함수의 탐색법

그림 2와 같이 탐색범위를 x_i, x_f로 하고, 그 크기를 $d = x_f - x_i$로 표시하면 그 중앙점 x_0에서 성능지표 J를

$$J_0 = f(x_0) \quad (13)$$

로 계산하여 J_0와 x_0를 기억시킨다.

(ii) 제1회 계산(w=1)

제1회의 분할구간

$$d_1 = \frac{d}{2} \times \frac{1}{3} \quad (14)$$

를 계산하고, x_0부터 $\pm 2d_1$인 점을 각각 x_i', x_f'로 표시하여, $J_i = f(x_i)$과 $J_f = f(x_f)$를 계산하여 J_0, J_i 및 J_f의 세 값에서 최소인 것을 $J^* = f(x^*)$로 기술하고, J^* 및 x^*를 기억시킨다.

(iii) 제2회 계산(w=2)
다시 분할구간 d_2는

$$d_2 = \frac{d_1}{3} \tag{15}$$

를 계산하고 x^*점을 중신으로 $\pm 2d_2$의 곳에 위치한 x_i'', x_f''의 점에서 J의 값 $J_i = f(x_i'')$와 $J_f = f(x_f'')$를 계산하여 J^*와 J_i, J_f를 비교하여 이들 중에서 최소치 J를 $J^* = f(x^*)$로 기술하고 J^* 및 x^*를 기억시킨다. 그리하여 제1회 탐색범위의 $\frac{1}{3}$로 축소된 탐색범위 내에서 탐색하게 된다.

(iv) 3회 이상의 계산(w>3)
제i번째의 탐색은 분할구간

$$d_i = \frac{d_{i-1}}{3} = \frac{d}{2 \cdot 3^i} \tag{16}$$

이므로, 전항과 같은 방법으로 J의 계산을 반복해서 설정한 계산횟수로 종료한다.
(iii)항의 방법을 반복하여 J^*를 계산하였을 때 J의 총 계산횟수 y는

$$y = 3 \times w \tag{17}$$

이며, 탐색정도(즉, 최종 계산시 탐색범위)는

$$d_w = \frac{d}{2} \times \left(\frac{1}{3}\right)^w = \frac{d}{2 \times 3^w} \tag{18}$$

으로 w=10인 경우에는 $8.467 \times 10^{-6} \times d$가 되어 정밀한 탐색정도가 얻어진다.

특히, 최소치 x^*가 경계점 x_i나 x_f 상의 극한점에 위치하는 경우에 x가 다음 식(19)와 같은 반복계산에 의하여 x_0로부터 출발해서 x_f에 수렴함을 알 수 있다. 즉, 최초의 중심점 x_0로부터 w회의 반복계산으로 도달할 수 있는 거리 x_d는

$$x_d = \frac{d}{2} \times \left(\frac{2}{3} + \frac{2}{3} + \cdots + \frac{2}{3^w}\right)$$

$$= \frac{d}{2} \times \frac{2}{3} \times \left(\frac{1 - \left(\frac{1}{3}\right)^w}{1 - \left(\frac{1}{3}\right)}\right) \tag{19}$$

이고, $w = \infty$일 때 $x_d = \left(\frac{d}{2}\right)$로 되어 x_i점 또는 x_f점에 도달함을 알 수 있다.

이상의 과정을 검토하면 탐색범위로 설정한 구간 내의 모든 점을 접근하여 최소값의 탐색이 가능하고, 탐색구간을 벗어나는 경우는 존재하지 않으므로 정확한 값을 구할 수 있어 신뢰도가 인정되고, 안정한 상태의 유지가 가능하다.

(2) 2차 벡터인 경우
독립변수가 2개인 경우에 x점은 그림 2에서 평면상에 표시 되므로

$$x = [x_1, x_2] \tag{20}$$

이 평면상의 점에 대응한 J의 값에서 최소치를 구하는 문제가 된다.

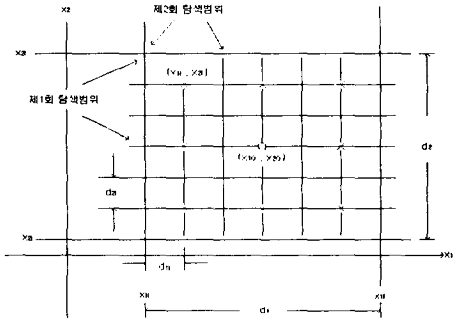

그림 3. 2변수 벡터의 탐색과정

(1)에서 설명한 것을 2차원으로 확장하여 탐색범위 x_{1i}, x_{2i}와 x_{1f}, x_{2f}의 평면 사각형을 중앙점(그림 3에서 ○점으로 표시) x_{10}, x_{20}와 그림 3에서 ×로 표시된 8개의 점을 포함한 9점에서 J를 계산하고, 그 중에서 J의 최소치

$$J^* = f(x_1^*, x_2^*) \tag{21}$$

를 탐색하여 J^* 및 x_1^*, x_2^*를 기억시키고 제1회(w=1)의 탐색을 마친다.

제2회는 J^*를 중심으로하여 길이가 $\frac{1}{3}$로 축

소된 탐색범위에서 같은 방법으로 최소치를 계산하여 설정된 w회의 반복계산을 한다.

(3) n차 벡터의 경우

탐색변수를 식(22)의 벡터로 표시하고,

$$x = [x_1,\ x_2,\ x_n] \tag{22}$$

이에 대응하는 탐색방법은 (2)절의 방법을 유추 적용하여 축차적으로 범위를 축소하여 설정한 반복횟수 w에서 종료한다.

이 때, 성능지표의 계산횟수 y는

$$y = 3^N \times w \tag{23}$$

이고, 정도 Z(최후의 탐색범위)는

$$Z_k = \frac{d_k}{2} \times \left(\frac{1}{3}\right)^{w-1} \tag{24}$$

이다. 여기서 d_k는 x의 k번째 성분에 대한 초기 탐색범위의 길이이다.

Ⅲ. 시뮬레이션

시뮬레이션에 적용한 plant는 다음과 같다.

$$\begin{aligned} x_1(t+1) &= x_1(t) + 0.01x_2(t) \\ x_2(t+1) &= x_2(t) + 0.01\{0.2(1.0 - x_1^2(t))x_2(t) \\ &\qquad - x_1(t) + u(t)\} \end{aligned} \tag{25}$$

그리고, 성능지표는 다음과 같다.

$$J = \sum_t [x_1^2(t) + x_2^2(t) + u^2(t-1)] \tag{26}$$

제안한 신경회로망의 구조는 표 1과 같다.

표 1

은닉층의 뉴런수	20 개
학습율	0.0005
초기값	[-0.1, 0.1]
step size	5 msec

시뮬레이션에 적용한 기준입력은 sin(t)이고, step입력은 1로 하였다.

우선 random하게 설정된 가중치를 가지고 역전과 신경회로망으로 학습시킨 결과를 그림 4에 나타낸다.

이 때 그림 5에서 보여진 성능지표의 최소값을 구하는데 일반적으로 행하는 미분법을 이용하지 않고 3점 탐색법으로 최소값을 탐색하여 이 때의 가중치를 신경회로망의 초기 가중치로 하여 다시 학습시킨 결과를 그림 6에 나타낸다.

그리고, random하게 설정했을 때와 최적화시킨 가중치를 설정했을 때의 setp입력에 대한 응답곡선을 그림 7과 그림 8에 각각 나타내었다.

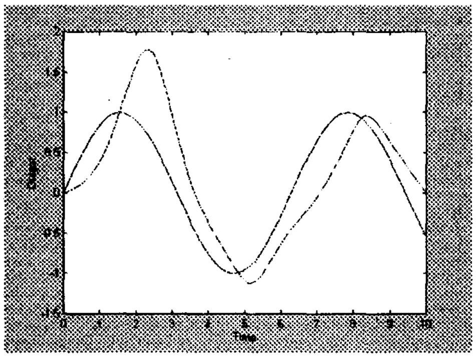

그림 4. random한 가중치 설정

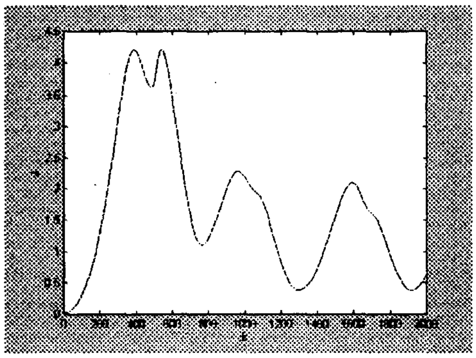

그림 5. random한 가중치에서의 성능지표

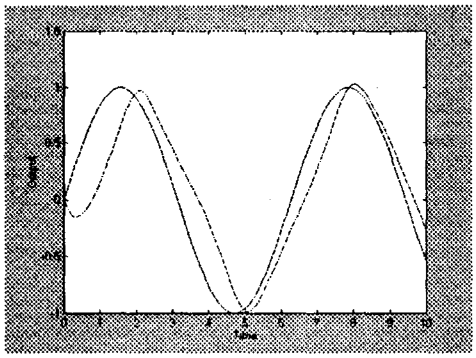

그림 6. 최적화된 가중치 설정

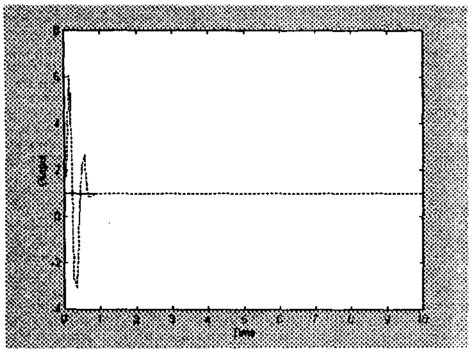

그림 7. random한 가중치에서의 step입력에 대한 응답곡선

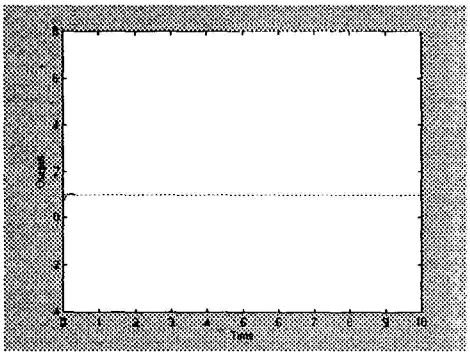

그림 8. 최적화된 가중치에서의 step입력에 대한 응답곡선

IV. 결론

신경회로망을 이용한 최적제어 설계방식을 제안하였다. 제안한 방식은 신경회로망의 학습을 위하여 gradient법을 사용하지 않고 직접탐색 알고리즘인 3점탐색 알고리즘을 사용하였다. 제안한 방식을 이용하여 2차원 형식의 단일변수와 다변수로 구성되는 성능지표의 최소치를 탐색하여 신경회로망의 가중치 초기값을 random한 설정과 최적화된 설정과의 결과를 비교 검토한 결과는 다음과 같다.

입력값 추적에 있어서 random하게 가중치를 선정한 것보다 빠르고 작은 오차로 추적하는 것을 확인할 수 있었다.

V. 참고문헌

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

[2] E .S Plunmer. Optimal Control of Terminal Processes Using Neural Network. IEEE Trans. Neural Networks, 7:408-418, 1996..

[3] D. Rumelhart, G. Hilton, and R. Williams. Learning Representations by Error Propagation. In Parallel Distributed Processing, volume 1, pp 318-362. MIT Press, 1986

[4] Johan A.K. Suykens, Joos P. L. Vandewalle, and Vart L. R. Cc Moor, Artificial Neural Networks for Modelling and Control of Non-Linear Systems. Kluwer Academic Publishers

[5] 김주홍. 최적설계를 위한 3점 탐색 알고리즘의 제안. 한국통신학회 논문지. volume 16, No 7. 1991.

[6] 김대수. 신경망 이론과 응용. 하이테크 정보. 1992.