한국산업정보학회논문지 제6권 제2호 (2001. 6)

구간 신경망에 의한 구간 벡터의 식별
(Classification of Interval Vectors by Interval Neural Networks)

권 기택*
(Ki-taek Kwon)

요 약 본 논문에서는 구간 데이터 식별을 위한 구간 신경망의 학습 알고리즘을 제안한다. 제안된 기법은 각 데이터의 속성치가 구간으로 주어져 있는 경우의 패턴 식별 문제에 적용된다. 먼저, 구간 입력 벡터를 다루기 위한 구간 신경망의 구조를 제안하고, 평가 함수를 이용하여 학습 알고리즘을 도출한다. 평가 함수는 구간 신경망으로부터의 구간 출력과 대응하는 목표 출력을 이용하여 정의된다. 마지막으로 컴퓨터 시뮬레이션에 의해 제안 기법의 구간 데이터 식별 능력을 나타내고, 통상의 역전파 신경망을 이용 기법과 비교한다.

Abstract This paper proposes a pattern classification method of interval vectors by interval neural networks. The proposed method can be applied to pattern classification where attribute values of each sample are given as interval numbers. First, an architecture of interval neural networks is proposed for dealing with interval input vectors. Next, a learning algorithm is derived from the cost function. a cost function is defined using the interval output from the interval neural network and the corresponding target output. Last, using numerical examples, the proposed approach is illustrated and compared with other approach based on the standard back-propagation neural networks.

1. 서 론

미지의 패턴이 주어졌을 때, 그 패턴이 어느 범주에 속하는가는 정하는 식별 문제에 있어서, 일반적으로 각 패턴의 속성치는 다차원 속성 공간내의 실수 값으로 나타내어지는 경우가 대부분이다.

그러나 현실 세계에서 우리 인간들이 이용할 수 있는 데이터에는 결락치(missing value)가 포함된 불완전한 데이터, 속성치가 일정하지 않고 변동하고 있는 데이터, 속성치의 정확한 측정이 불가능한 데이터, 속성치를 실수 값으로 표현하기 어려운 데이터 즉, 불확실한 데이터를 표현하기 위한 현실적이고 가장 간단한 방법으로서는, 그 가능한 값의 범위(상한치와 하한치)로 나타내는 구간(interval) 표현 방법도 생각해 볼 수 있다. 따라서 각 샘플의 속성치가 다차원 속성 공간 내의 구간 벡터로서 표현되는 구간 데이터에 대한 식별 기법도 연구되고 있다. 이러한 구간 데이터에 대한 식별 방법으로서는 선형 식별 함수를 이용한 선형계획문제의 정식화가 있다[1]. 그러나 이 기법에서는 구간 데이터를 식별할 수 있는 능력을 갖고 있지만, 식별 가능한 구간 데이터는 선형분리 가능한 데이터에만 유효하다는 문제점을 갖고 있다.

*동양대학교 산업경영공학과 교수

본 논문에서는, 다차원 입력공간에서 출력공간으로 사상을 행하는 비선형 시스템의 근사 능력을 가진 구간 신경망을 이용하여 구간 데이터를 2 개의 군으로 분류하는 방법에 대해서 고찰한다. 먼저, 구간 데이터를 처리할 수 있는 구간 신경망의 구조를 제안하고, 구간 데이터 식별을 위한 학습 알고리즘을 제안한다. 평가 함수는 구간 신경망으로부터의 구간 출력과 목표 출력치를 이용하여 정의되고, 이 평가 함수가 최소화 될 수 있도록 학습이 이루어진다. 학습 알고리즘은 이 평가 함수로부터 도출되고, 학습된 구간 신경망에 의해 미지의 구간 벡터들이 2 개의 군으로 분류된다. 마지막으로 컴퓨터 시뮬레이션에 의해, 제안 기

법의 구간 데이터 2군 식별 능력을 평가하고, 통상의 BP(back-propagation) 신경망[2]을 이용한 결과와 비교 해 본다.

2. 구간 데이터

본 논문에서는 실수를 영 소문자 $a, b, c, \cdots, z$, 구간을 영 대문자 $A, B, C, \cdots, Z$로 표현하기로 하고 구간치는 그 하한치와 상한치를 이용하여 $X=[x^L, x^U]$로 표현하기로 한다. 여기서 첨자 L은 하한치, U는 상한치를 나타낸다.

구간 신경망에 이용될 구간 연산의 합과 곱은 다음과 같다[3].

$$\begin{aligned} A+B &= [a^L,\ a^U]+[b^L,\ b^U] \\ &= [a^L+a^U,\ b^L+b^U] \end{aligned} \quad (1)$$

$$\begin{aligned} A\cdot B &= [a^L, a^U]\cdot[b^L, b^U] \\ &= [\min\{a^Lb^L, a^Lb^U, a^Ub^L, a^Ub^U\}, \\ &\quad \max\{a^Lb^L, a^Lb^U, a^Ub^L, a^Ub^U\}] \end{aligned} \quad (2)$$

식(2)에 있어서 구간 B가 양($0 \le b^L \le b^U$)인 경우, 구간의 곱은 다음과 같이 간단히 계산할 수 있다.

$$\begin{aligned} A\cdot B &= [a^L,\ a^U]\cdot[b^L,\ b^U] \\ &= [\min\{a^Lb^L, a^Lb^U\}, \\ &\quad \max\{a^Ub^L, a^Ub^U\}] \end{aligned} \quad (3)$$

2 개의 군으로 식별하는 문제로서, 제 1군(G1)과 제 2군(G2)으로 분류되는 m 개의 패턴이 주어져 있다고 하고, 각 패턴들은 n 개의 속성을 갖는다고 하자. 이 때, 각 속성치는 구간 값으로 주어져 있다고 하고, 제 p 패턴의 i 번째 구간 속성치를 X_{pi} 로 나타내기로 한다. 아울러, p 번째의 패턴은 n 차원 구간 벡터 $\boldsymbol{X}_p = (X_{p1}, X_{p2}, \cdots, X_{pn})$로 표현하기로 한다.

여기서의 문제는 2 개의 군으로 분류되는 m 개의 구간 벡터 $\boldsymbol{X}_1$, $\boldsymbol{X}_2, \cdots$, $\boldsymbol{X}_m$ 를 이용하여 구간 신경망의 학습을 행하고 미지의 구간 벡터를 식별하는 문제가 된다.

3. 구간 신경망

3.1 구간 신경망의 구조

구간 데이터의 2 군 식별을 행하기 위한 구간 신경망의 구조는 <그림 1>과 같이 입력층에 n 개, 중간층에 n_2 개, 출력층에 1 개의 유니트를 갖는 3층형으로 한다. 여기서, 입력층의 유니트 수는 입력 벡터의 차원과 동일하지만, 중간층의 유니트 수는 임의로 정할 수 있다. 또한, 결합강도 및 임계치는 구간으로 주어져 있다고 한다. 이러한 구간 신경망에 n 차원 구간 입력 벡터 $\boldsymbol{X}_p = (X_{p1}, X_{p2}, \cdots, X_{pn})$가 입력된 경우의 각 유니트의 입출력 관계는 다음과 같다[4].

입력층: $O_{pi} = X_{pi},\quad i = 1, 2, \cdots, n$ (4)

중간층: $O_{pj} = f(Net_{pj}),\ j = 1, 2, \cdots, n_2$ (5)

$$Net_{pj} = \sum_{i=1}^{n} O_{pi} W_{ji} + \Theta_j,\ \ j = 1, 2, \cdots, n_2 \quad (6)$$

출력층: $O_p = f(Net_p)$ (7)

$$Net_p = \sum_{j=1}^{n_2} O_{pj} W_j + \Theta \quad (8)$$

여기서, 결합강도 W_{ji}, W_j와 임계치 Θ_j, Θ는 구간 값이고, $O_{pi}, O_{pj}, O_p, Net_{pj}$와 Net_p도 모두 구간 값이다. 또한 간단하게 하기 위하여 $X_{pi} \ge 0$ 으로 한다. 중간층과 출력층에 있는 함수 $f(\cdot)$는 다음과 같은 시그모이드(sigmoid) 함수이고, <그림 2>와 같은 구간 입출력 함수가 된다.

$$f(x) = 1/\{1+\exp(-x)\} \quad (9)$$

식(4)-(8)에서, 각 유니트의 입출력 관계에 의한 구체적인 계산은 위에서 기술한 구간연산[3]으로부터 다음과 같이 행해진다.

입력층:

$$O_{pi} = [o_{pi}^L, o_{pi}^U] = [x_{pi}^L, x_{pi}^U],\quad i = 1, 2, \cdots, n \quad (10)$$

중간층:

$$O_{pj} = [o_{pj}^L, o_{pj}^U] = [f(net_{pj}^L), f(net_{pj}^U)],\quad j = 1, 2, \cdots, n_2 \quad (11)$$

$$net_{pj}^L = \sum_{\substack{i=1 \\ w_{ji}^L \ge 0}}^{n} o_{pi}^L w_{ji}^L + \sum_{\substack{i=1 \\ w_{ji}^L < 0}}^{n} o_{pi}^U w_{ji}^L + \theta_j^L,\quad j = 1, 2, \cdots, n_2 \quad (12)$$

$$net_{pj}^U = \sum_{\substack{i=1 \\ w_{ji}^U \geq 0}}^{n} o_{pi}^U w_{ji}^U + \sum_{\substack{i=1 \\ w_{ji}^U < 0}}^{n} o_{pi}^L w_{ji}^U + \theta_j^U, \quad (13)$$

$$j = 1, 2, \cdots, n_2$$

출력층:

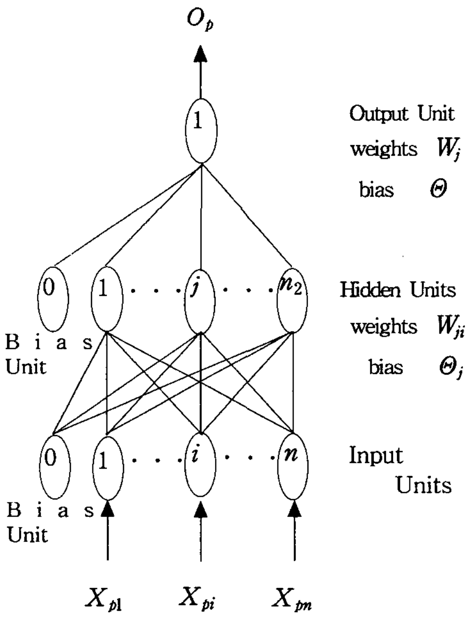
$$O_p = [o_p^L, o_p^U] = [f(net_p^L), f(net_p^U)] \quad (14)$$

$$net_p^L = \sum_{\substack{j=1 \\ w_j^L \geq 0}}^{n_2} o_{pj}^L w_j^L + \sum_{\substack{j=1 \\ w_j^L < 0}}^{n_2} o_{pj}^U w_j^L + \theta^L$$

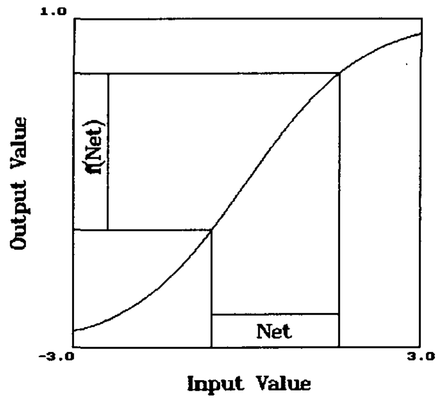
(15)

$$net_p^U = \sum_{\substack{j=1 \\ w_j^U \geq 0}}^{n_2} o_{pj}^U w_j^U + \sum_{\substack{j=1 \\ w_j^U < 0}}^{n_2} o_{pj}^L w_j^U + \theta^U$$

(16)



<그림 1> 구간 신경망의 구조



<그림 2> 각 유니트의 구간 입출력 함수

3.2 학습 알고리즘

구간 입력 패턴 $\mathbf{X}_p = (X_{p1}, X_{p2}, \cdots, X_{pn})$에 대응하는 목표 출력을 다음과 같이 정한다.

$$t_p = \begin{cases} 1, & \mathbf{X}_p \in G1 \\ 0, & \mathbf{X}_p \in G2 \end{cases}$$

(17)

목표 출력치 t_p와 구간 신경망으로부터의 구간 출력치 $O_p = [o_p^L, o_p^U]$를 이용하여 패턴 p에 관한 평가 함수 e_p를 다음과 같이 정한다.

$$e_p = e_p^L + e_p^U$$

(18)

여기서,

$$e_p^L = (t_p - o_p^L)^2/2$$

(19)

$$e_p^U = (t_p - o_p^U)^2/2$$

(20)

이다. 구간 신경망의 학습은 식(18)의 평가 함수가 최소화 되도록 행해지고, 구간 결합강도의 수정량은 다음과 같이 설정한다.

$$\Delta w_{ji}^L(t+1) = \eta(-\partial e_p / \partial w_{ji}^L) + \alpha \Delta w_{ji}^L(t)$$

(21)

$$\Delta w_{ji}^U(t+1) = \eta(-\partial e_p / \partial w_{ji}^U) + \alpha \Delta w_{ji}^U(t)$$

(22)

$$\Delta w_j^L(t+1) = \eta(-\partial e_p/\partial w_j^L) + \alpha\Delta w_j^L(t) \quad (23)$$

$$\Delta w_j^U(t+1) = \eta(-\partial e_p/\partial w_j^U) + \alpha\Delta w_j^U(t) \quad (24)$$

여기서, η는 학습률, α는 모멘텀항 계수, α는 구간 결합강도의 수정 횟수이다. 식(21)-(24)에 의해, 수정 후의 구간 결합강도 W_{ji}, W_j의 하한과 상한은 다음과 같이 된다.

$$w_{ji}^L(t+1) = w_{ji}^L(t) + \Delta w_{ji}^L(t+1) \quad (25)$$

$$w_{ji}^U(t+1) = w_{ji}^U(t) + \Delta w_{ji}^U(t+1) \quad (26)$$

$$w_j^L(t+1) = w_j^L(t) + \Delta w_j^L(t+1) \quad (27)$$

$$w_j^U(t+1) = w_j^U(t) + \Delta w_j^U(t+1) \quad (28)$$

이와 같은 수정에 의해, 구간 결합강도의 하한과 상한의 역전이 일어날 수 있다. 그래서 역전이 일어날 것도 고려해서 구간 결합강도를 다음과 같이 설정한다.

$$W_{ji}(t+1) = [\min\{w_{ji}^L(t+1), w_{ji}^U(t+1)\}, \max\{w_{ji}^L(t+1), w_{ji}^U(t+1)\}] \quad (29)$$

$$W_j(t+1) = [\min\{w_j^L(t+1), w_j^U(t+1)\}, \max\{w_j^L(t+1), w_j^U(t+1)\}] \quad (30)$$

구간 임계치 Θ_j, Θ도 구간 결합강도와 같은 방법으로 수정할 수 있다.

4. 컴퓨터 시뮬레이션

4.1 수치 예 1

2 개의 군으로 분류되는 6 개 샘플의 속성치가 2 차원 구간 벡터로서 <표 1>과 같이 주어져 있다고 한다. 위에서 기술한 구간 신경망의 학습 알고리즘으로 주어진 구간 데이터의 2 군 식별을 시험해 본다. 학습을 위해 중간층의 유니트 수는 6 개로 하였고, 학습률 η는 0.5, 모멘텀항 계수 α는 0.9로 설정하였다. 이와 같은 설정 아래 10,000 회 학습 후의 구간 신경망으로부터의 구간 출력치가 0.45 ~ 0.55 사이인 점을 연결하면 <그림 3>과 같이 된다. 또한, 각 구간 입력 벡터에 대응하는 구간 신경망으로부터의 구간 출력 값을 <표 2>에 나타낸다. <그림 3>과 <표 2>로부터, 제안 기법을 이용한 구간 신경망은 2 개의 군으로 분류되는 구간 데이터를 바르게 식별하고 있다는 것을 알 수 있다.

제안 기법과 통상의 BP 알고리즘을 이용한 기법과 비교를 위해, <표 1>의 구간 데이터를 이용하여 통상의 BP 알고리즘에 의한 신경망의 학습을 행하였다. 학습용 데이터로는 <표 1>에 나타난 6 개의 2차원 구간 벡터 각각의 4 단점이 이용되었고, 학습을 위한 중간층 수, η, α 등의 값은 제안 기법에서 설정된 값과 같게 하였다. 10,000회 학습 후의 출력치가 0.45 ~ 0.55사이인 점을 연결하면 <그림 4>와 같이 나타난다. 이 결과로부터, 통상의 BP 알고리즘을 이용한 기법으로는 구간 벡터의 식별을 바르게 하지 못하였다는 것을 알 수 있다.

4.2 수치 예 2

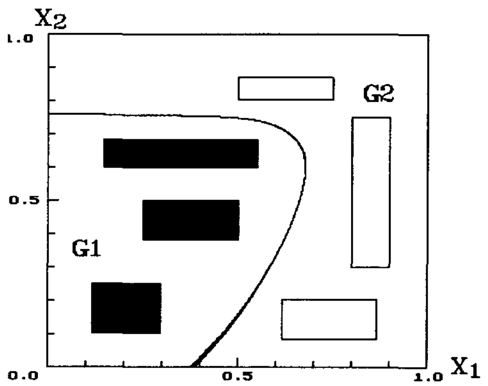
여기서는 제안 기법에 의한 식별 능력이 구간 데이터 뿐만 아니라 실수치인 경우에도 우수한지 시험해 본다. 실수치 데이터가 <표 3>과 같이 주어져 있다고 하고, 제안 기법에 의한 구간 신경망의 학습 알고리즘으로 학습을 행하였다. 학습에서 실수치인 데이터는 하한 값과 상한 값이 같은 구간 데이터로 간주되었다. 예를 들면, 실수치 데이터 $x=0.1$은 구간 데이터 $X=[x^L, x^U]=[0.1, 0.1]$로 간주되어 구간 신경망의 학습에 이용된다. 학습을 위한 중간층 수, η, α 등의 값은 수치 예 1에서 설정된 값과 같게 하였다. 10,000회

<표 1> 구간 데이터

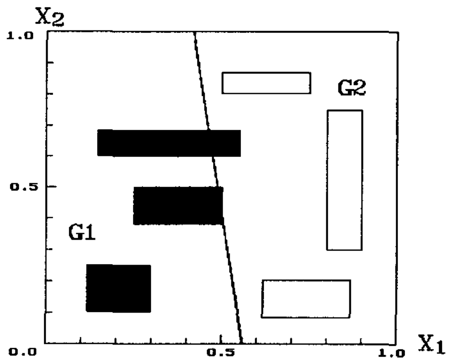
제 1군 (G1)			제 2군 (G2)		
No.	X_1	X_2	No.	X_1	X_2
1	[0.12, 0.30]	[0.10, 0.30]	4	[0.62, 0.87]	[0.08, 0.20]
2	[0.25, 0.50]	[0.38, 0.50]	5	[0.80, 0.90]	[0.30, 0.75]
3	[0.15, 0.55]	[0.60, 0.68]	6	[0.50, 0.75]	[0.80, 0.87]

<표 2> 학습 후의 구간 신경망으로부터의 구간 출력치

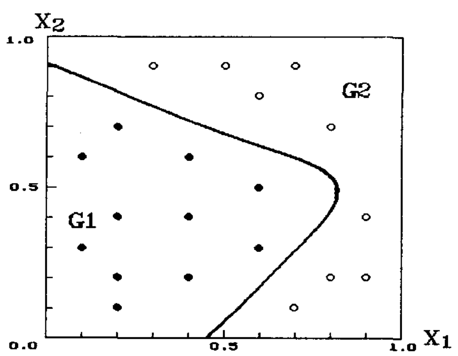
제 1군 (G1) 목표출력 1		제 2군 (G2) 목표출력 0	
No.	구간 출력치	No.	구간 출력치
1	[0.998, 0.999]	4	[0.000, 0.001]
2	[0.999, 0.999]	5	[0.000, 0.002]
3	[0.995, 0.999]	6	[0.000, 0.003]



<그림 3> 제안 기법에 의한 시뮬레이션 결과 (수치 예 1)



<그림 4> BP 알고리즘에 의한 시뮬레이션 결과 (수치 예 1)



<그림 5> 제안 기법에 의한 시뮬레이션 결과 (수치 예 2)

학습 후의 구간 출력치가 0.45 ~ 0.55사이인 점을 연결하면 <그림 5>와 같이 나타난다. 이 결과로부터, 제안 기법을 이용한 구간 신경망은 구간

<표 3> 실수치 데이터

제 1군 (G1)			제 2군 (G2)		
No.	x_1	x_2	No.	x_1	x_2
1	0.1	0.3	12	0.3	0.9
2	0.1	0.6	13	0.5	0.9
3	0.2	0.1	14	0.7	0.9
4	0.2	0.2	15	0.8	0.7
5	0.2	0.4	16	0.9	0.4
6	0.2	0.7	17	0.7	0.1
7	0.4	0.2	18	0.8	0.2
8	0.4	0.4	19	0.9	0.2
9	0.4	0.6	20	0.6	0.8
10	0.6	0.3			
11	0.6	0.5			

데이터뿐만 아니라 실수치인 데이터도 바르게 식별하고 있다는 것을 알 수 있다.

5. 결론

본 논문에서는 구간 신경망을 이용하여 주어진 각 샘플의 속성치가 구간치인 경우에 대하여 식별할 수 있는 기법을 제안하였다. 구간 입력 벡터를 다루기 위한 구간 신경망의 구조를 나타내었고, 구간 데이터의 2 군 식별을 위한 평가 함수를 이용하여 학습 알고리즘을 나타내었다. 학습 알고리즘은 구간 연산에 의해 계산되었고 제안된 평가 함수가 최소화되도록 하였다. 컴퓨터 시뮬레이션에 의해, 제안된 학습 알고리즘은 구간 데이터의 2 군 식별 능력이 우수함을 나타내었고, 통상의 BP 알고리즘에 의한 결과보다 우수하다는 것을 나타내었다. 또한, 제안 기법에 의한 2 군 식별 능력은 구간 데이터뿐만 아니라 실수치 데이터에도 유효하다는 것을 나타내었다.

참 고 문 헌

[1] Ishibuchi H., Tanaka H. and Fukuoka N., " Discriminant Analysis of Multi-Dimensional Interval Data and Its Application to Chemical Sensing", International Journal of General Systems, Vol.16, No.4, pp.311-329, 1990.

[2] Rumelhart D.E., McClelland J.L. and PDP Research Group, Parallel Distributed Processing, Vol.1, Cambridge, MIT Press, 1986.

[3] Alefeld G. and Herzberger J., Introduction to Interval Computations, pp.1-38, Academic Press, 1983.

[4] Kwon K., Ishibuchi H. and Tanaka H., "Nonlinear Mapping of Interval Vectors by Neural Networks", Proc. of IJCNN'93-Nagoya, Vol.1, pp.758-761, 1993.

권 기 택 (Ki-taek Kwon)
대구대학교 산업공학과 졸업
1995년 일본 大阪府立대학
석사 및 박사학위 취득
현재 동양대학교 산업경영
공학과 조교수로 재직중

관심분야 : 퍼지이론, 지능시스템 등