

## 퍼지 신경망에 의한 퍼지 회귀분석

### Fuzzy Regression Analysis Using Fuzzy Neural Networks

권기택\*

Kitaek Kwon\*

#### Abstract

This paper propose a fuzzy regression method using fuzzy neural networks when a membership value is attached to each input-output pair. First, a method of linear fuzzy regression analysis is described by interpreting the reliability of each input-output pair as its membership values. Next, an architecture of fuzzy neural networks with fuzzy weights and fuzzy biases is shown. The fuzzy neural network maps a crisp input vector to a fuzzy output. A cost function is defined using the fuzzy output from the fuzzy neural network and the corresponding target output with a membership value. A learning algorithm is derived from the cost function. The derived learning algorithm trains the fuzzy neural network so that the level set of the fuzzy output includes the target output. Last, the proposed method is illustrated by computer simulations on numerical examples.

#### 1. 서론

회귀분석의 목적 중의 하나는 주어진 독립 변수로부터 종속변수의 값을 추정하는 것이다. 추정에 있어서, 인간의 주관적인 평가, 판단 등과 같이 해석 대상의 시스템에 모호성이 포함되어 있는 상황에서는, 회귀분석에 의

한 종속변수의 값을 구간(Interval)이나 퍼지 수(Fuzzy Numbers)로 추정하는 편이 현실적인 경우도 많다. 그래서 Tanaka 등[1]은 모호한 상황을 해석하기 위하여 선형 퍼지 회귀 분석 기법을 제안하고 있다. 이 기법은 입출력 데이터를 이용하여 선형 퍼지 시스템을 모델화한다고 하는 것이다. 즉, 입출력 관계

\* 동양대학교 산업공학과

를 모호하게 동정한 것으로서 모호한 상황 분석에 적합한 기법이다. 또한 전문가에 따른 주관적인 판단을 데이터의 해석 분야에 도입하기 위하여 유사도를 도입한 퍼지 회귀 분석 기법[2]을 정식화하였다. 유사도는 선형 퍼지 회귀식으로부터 얻어진 퍼지수의 레벨 집합(Level Set)에 반영되고, 유사도가 큰 입출력 데이터가 회귀분석에 있어서 중요시된다.

비선형 퍼지 회귀식을 구하는 기법으로서 는 GMDH(Group Method Data Handling)를 이용한 퍼지 GMDH[3], 신경망을 이용한 방법 [4] 등이 제안되어 있다. Ishibuchi 등[4]은 기존의 다층 신경망을 이용하여, 일반적인 데이터(Crisp Data)에 대하여 퍼지 회귀분석을 행하는 기법을 제안하고 있다. 이 기법은 신경망이 가진 비선형 함수의 근사 능력을 이용함으로써, 모호성을 포함하고 있는 복잡한 비선형 시스템을 간단한 학습에 의해 동정할 수 있다. 또한 기존의 다층 신경망을 확장한 퍼지 신경망을 이용하여, 퍼지수 데이터에 대하여 퍼지 회귀분석을 행하는 기법[5]도 제안되어 있다.

현실 세계에서 우리들이 사용하고 있는 데이터에는, 신뢰도가 높은 데이터와 높지 않은 데이터로 섞여 있는 경우도 적지 않다. 종래의 비선형 구간 회귀식 또는 비선형 퍼지 회귀식을 구하는 기법[3-5]은, 주어진 데이터에 대한 신뢰도가 모두 같다고 하는 전제 하에 제안되어 있다. 신뢰도가 높은 데이터와 낮은 데이터가 같은 레벨로 사용됨으로써, 해석 대상이 되는 데이터에 대한 신뢰도가 무시되는 경향이 있다. 예를 들면, 전문가로부터 얻어진 데이터와 비전문가로부터 얻어진

데이터가 같은 신뢰도를 가진다고는 말하기 어려울 것이다. 이와 같은 경우에는 데이터에 대한 신뢰도를 도입하는 것이 자연스러울 것이다.

본 연구에서는, 퍼지 신경망을 이용하여 신뢰도가 도입된 입출력 데이터에 대하여 비선형 퍼지 회귀분석을 행하는 기법을 제안한다. 제안 기법은 입출력 데이터에 대한 신뢰도에 차이가 보이는 경우에 유효하고, 입력벡터가 다차원인 이유로 모델의 함수형을 정할 수 없는 문제도 다룰 수 있다. 또한 퍼지 신경망이 가진 비선형 퍼지 함수의 근사 능력을 이용함으로써 모호성이 포함되어 있는 비선형 시스템을 간단한 학습으로 동정할 수 있다.

본 연구에서는, 먼저 데이터의 신뢰도를 도입한 선형 퍼지 회귀분석 기법을 나타낸다. 여기서 데이터에 대한 신뢰도는 전문가로부터 얻어진다고 하고, 구간 0과 1사이의 소속 함수(Membership Function) 값으로 해석할 수 있다고 한다. 다음으로 각 데이터에 대한 신뢰도를 나타내는 소속함수 값이 붙여진 입출력 데이터에 대하여 퍼지 회귀분석을 행하기 위한 퍼지 연결강도 및 퍼지 임계치를 가진 퍼지 신경망의 구조를 나타낸다. 이 퍼지 신경망은 일반적인 입력벡터에서 퍼지 출력치로 사상하는 비선형 시스템이라고 볼 수 있다. 또한 퍼지 신경망으로부터의 퍼지 출력치와 소속함수 값을 가진 목표 출력치를 이용하여 코스트 함수를 정의하고, 학습 알고리즘을 도출한다. 이 퍼지 신경망의 학습은 소속함수 값을 가진 목표 출력치가 추정된 퍼지 출력치의 대응하는 레벨의 레벨 집합에 포함되도록 행해진다. 마지막으로 수치예를

이용하여 제안 기법에 의한 계산 실험 결과를 나타낸다.

## 2. 신뢰도를 도입한 퍼지 회귀분석

$n$ 입력 1출력 시스템의 입출력 관계에 관한  $m$ 개의 데이터( $x_p, y_p, h_p$ ),  $p=1, 2, \dots, m$ ,가 주어져 있다고 한다. 여기서  $x_p = (x_{p1}, x_{p2}, \dots, x_{pn})$ 는  $p$ 번째의 데이터에 의한  $n$ 차원 입력벡터,  $y_p$ 는 출력치(신경망에서는 목표 출력치),  $h_p$ 는 입출력 쌍( $x_p, y_p$ )의 신뢰도이다. 여기서 데이터에 대한 신뢰도( $h_p$ )는 구간  $[0,1]$ 사이의 값으로 주어져 있다고 하고, 소속함수 값으로 해석할 수 있다고 한다. 이와 같은 데이터에 대한 선형 퍼지 회귀식은 다음과 같다.

$$Y_p = A_0 + A_1 x_{p1} + \dots + A_n x_{pn} \quad (1)$$

여기서  $A_j$ 는 퍼지 계수이다.

퍼지 계수로서는 대칭 삼각형 퍼지수가 사용되는 경우가 많다. 대칭 삼각형 퍼지수  $A$ 는 그 중심( $a^C$ )과 폭( $a^W$ )으로부터  $A = (a^C, a^W)$ 로 표현되고, 소속함수는 다음과 같이 정의된다.

$$\mu_A(x) = \max \{ 1 - |x - a^C| / a^W, 0 \}, \quad a^W > 0 \quad (2)$$

폭( $a^W$ )이 0인 경우의 퍼지수  $A$ 는 일반적인 데이터( $a^C$ )와 일치한다. 여기서 식(1)은 확장 원리에 의해 다음과 같이 계산할 수 있다 [1,2].

$$Y_p = (a_0^C + \sum_{j=1}^n a_j^C x_{pj}, a_0^W + \sum_{j=1}^n a_j^W |x_{pj}|) \quad (3)$$

소속함수 값( $h_p$ )을 가진 입출력 데이터( $x_p, y_p$ )에 대한 퍼지 회귀분석은, Tanaka 등[2]에 의해 제안된 유사도를 도입한 퍼지 회귀 분석과 같은 방법으로 출력치( $y_p$ )가 추정된 퍼지 출력치( $Y_p$ )의  $h_p$ -레벨 집합에 포함되도록 정식화된다. 단, 퍼지수  $Y_p$ 의  $h_p$ -레벨 집합( $[Y_p]_{h_p}$ )은 다음과 같이 정의된다.

$$[Y_p]_{h_p} = \{ y \mid \mu_{Y_p}(y) \geq h_p \} \quad (4)$$

이와 같은 제약 조건의 아래서 추정 퍼지 출력치( $Y_p$ ) 폭의 총합이 최소가 되도록 퍼지 회귀식의 퍼지 계수가 구해진다. 즉 다음과 같은 선형계획 문제가 정식화된다.

$$\min \sum_{p=1}^m h_p (a_0^W + \sum_{j=1}^n a_j^W |x_{pj}|) \quad (5)$$

$$\text{subject to } y_p \in [Y_p]_{h_p}, \quad p=1, 2, \dots, m \quad (6)$$

$$a_j^W \geq 0, \quad j=0, 1, 2, \dots, n \quad (7)$$

이상과 같은 정식화로 부터 신뢰도( $h_p$ )가 큰 데이터가 중요시되고 있다는 것을 알 수 있다.

본 연구에서는 식(1)의 선형 퍼지 회귀식을 비선형 퍼지 회귀식으로 확장하기 위해서, 퍼지 신경망을 이용하여 데이터의 신뢰도를 도입한 퍼지 회귀분석을 행하는 방법을 제안한다.

## 3. 퍼지 신경망의 구조

### 3.1 퍼지 연결강도와 퍼지 임계치

본 연구에서는 기존의 다층 신경망[6]의 연

결강도와 임계치를 삼각형 퍼지수로 확장한다. 삼각형 퍼지수  $A$ 를 그 하한( $a^L$ ), 중심( $a^C$ )과 상한( $a^U$ )으로부터  $A=(a^L, a^C, a^U)$ 로 표현하면, 식(2)는 다음과 같이 나타낼 수 있다.

$$\mu_A(x) = \begin{cases} 0; & x < a^L \\ (x-a^L)/(a^C-a^L); & a^L \leq x \leq a^C \\ (a^U-x)/(a^U-a^C); & a^C < x \leq a^U \\ 0; & a^U < x \end{cases} \quad (8)$$

즉, 삼각형 퍼지수  $A=(a^L, a^C, a^U)$ 의 소속함수 값은 개구간 ( $a^L, a^U$ )에서 양이고, 중심( $a^C$ )에서 1이 된다. 특히 퍼지수가 대칭 삼각형인 경우, 중심( $a^C$ )에 관해서는 다음 식이 성립한다.

$$a^C=(a^L+a^U)/2 \quad (9)$$

### 3.2 입출력 관계의 정의

퍼지 신경망의 구조는 입력층에  $n$  개, 중간층에  $n_2$  개, 출력층에 1 개의 유닛을 가진 3층형으로 한다. 또한 모든 연결강도 및 임계치는 대칭 삼각형 퍼지수로 주어지고 있다. 이와 같은 신경망에 일반적인  $n$ 차원 입력벡터  $x_p=(x_{p1}, x_{p2}, \dots, x_{pn})$ 가 입력되었을 경우, 각 유닛의 입출력 관계는 다음과 같이 정의된다[5].

$$\text{입력층: } o_{pi}=x_{pi}, \quad i=1, 2, \dots, n \quad (10)$$

$$\text{중간층: } O_{pj}=f(Net_{pj}), \quad j=1, 2, \dots, n_2 \quad (11)$$

$$Net_{pj}=\sum_{i=1}^n o_{pi} W_{ji} + \theta_j, \quad j=1, 2, \dots, n_2 \quad (12)$$

$$\text{출력층: } O_p=f(Net_p) \quad (13)$$

$$Net_p=\sum_{j=1}^{n_2} O_{pj} W_j + \theta \quad (14)$$

여기서 연결강도  $W_{ji}, W_j$ 와 임계치  $\theta_j, \theta$ 는 대칭 삼각형 퍼지수이고,  $O_{pj}, O_p, Net_{pj}$ 와  $Net_p$ 는 퍼지수이다. 또한 간단하게 하기 위하여  $x_{pi} \geq 0$ 로 한다. 중간층과 출력층에 있는 함수  $f(\cdot)$ 는 다음과 같은 시그모이드(Sigmoid) 함수이다.

$$f(x)=1/[1+\exp(-x)] \quad (15)$$

식(10)-(14)의 입출력 관계에 의한 퍼지수 연산은 확장원리[7]로부터 정의된다. 구체적인 퍼지 입출력을 계산할 경우에는 레벨 집합마다의 구간 연산이 이용된다. 식(10)-(14)에 의한 퍼지 신경망의 입출력 관계를  $h$ -레벨 집합에 관하여 쓰면 다음과 같다.

$$\text{입력층: } o_{pi}=x_{pi}, \quad i=1, 2, \dots, n \quad (16)$$

중간층:

$$[O_{pj}]_h=[([O_{pj}]_h^L, [O_{pj}]_h^U)] = f([Net_{pj}]_h) = [f([Net_{pj}]_h^L), f([Net_{pj}]_h^U)], \quad j=1, 2, \dots, n_2 \quad (17)$$

$$[Net_{pj}]_h^L = \sum_{i=1}^n [W_{ji}]_h^L \cdot o_{pi} + [\theta_j]_h^L \quad (18)$$

$$[Net_{pj}]_h^U = \sum_{i=1}^n [W_{ji}]_h^U \cdot o_{pi} + [\theta_j]_h^U \quad (19)$$

출력층:

$$[O_p]_h=[([O_p]_h^L, [O_p]_h^U)] = f([Net_p]_h) = [f([Net_p]_h^L), f([Net_p]_h^U)] \quad (20)$$

$$[Net_p]_h^L = \sum_{j=1}^{n_2} [W_j]_h^L \cdot [O_{pj}]_h^L \quad (W_j]_h^L \geq 0) \\ + \sum_{j=1}^{n_2} [W_j]_h^L \cdot [O_{pj}]_h^U + [\theta]_h^L \quad (W_j]_h^L < 0) \quad (21)$$

$$[Net_p]_h^U = \sum_{j=1}^{n_2} [W_j]_h^U \cdot [O_{pj}]_h^U \quad (W_j]_h^U \geq 0) \\ + \sum_{j=1}^{n_2} [W_j]_h^U \cdot [O_{pj}]_h^L + [\theta]_h^U \quad (W_j]_h^U < 0) \quad (22)$$

식(16)-(22)에 있어서, 폐구간이 되는 퍼지 수의  $h$ -레벨 집합은 다음과 같이 표현된다.

$$[A]_h = [[A]_h^L, [A]_h^U] \quad (23)$$

여기서  $[A]_h^L, [A]_h^U$ 는 퍼지수  $A$ 의  $h$ -레벨 집합  $([A]_h)$ 의 하한과 상한이다. 특히 퍼지수  $A$ 가 대칭 삼각형일 경우,  $h$ -레벨 집합의 하한과 상한은 다음과 같다.

$$[A]_h^L = (1-h/2) \cdot a^L + h \cdot a^U / 2 \quad (24)$$

$$[A]_h^U = h \cdot a^L / 2 + (1-h/2) \cdot a^U \quad (25)$$

#### 4. 퍼지 신경망의 학습

##### 4.1 코스트 함수(Cost Function)의 정의

퍼지 신경망으로부터의 퍼지 출력치( $O_p$ )와 소속함수 값을 가진 입출력 쌍( $x_p, y_p, h_p$ )을 이용하여, 비선형 퍼지 회귀분석을 행하기 위한 퍼지 신경망의 코스트 함수를 정의한다. 본 연구에서는 추정된 퍼지 출력치( $O_p$ )의  $h_p$ -

레벨 집합( $[O_p]_{h_p} = [[O_p]_{h_p}^L, [O_p]_{h_p}^U]$ )에 목표 출력( $y_p$ )이 포함되도록 하는 것을 목적으로 한다. 즉,

$$[O_p]_{h_p}^L \leq y_p \leq [O_p]_{h_p}^U \quad (26)$$

가 성립되도록 학습이 행해진다. 코스트 함수는 다음과 같은 세 가지의 경우에 따라 각각 정의된다.

(1)  $[O_p]_{h_p}^L \leq [O_p]_{h_p}^U < y_p$ 가 성립될 경우

$$e_p = \omega \cdot e_p^L + e_p^U \quad (27)$$

여기서  $\omega$ 는 ( $0 < \omega < 1$ )을 만족하는 양의 실수,  $e_p^L$ 은 목표 출력치와 추정된 퍼지 출력치의  $h_p$ -레벨 집합의 하한에 관한 오차 제곱의 총합이고,  $e_p^U$ 는 상한에 관한 오차 제곱의 총합이다. 학습에 있어서 최소화되는 코스트 함수  $e_p^L, e_p^U$ 는 각각 다음과 같이 정의된다.

$$e_p^L = h_p (y_p - [O_p]_{h_p}^L)^2 / 2 \quad (28)$$

$$e_p^U = h_p (y_p - [O_p]_{h_p}^U)^2 / 2 \quad (29)$$

다시 말하여 식(26)의 부등식 관계를 만족시키지 못하고 있는 끝점( $[O_p]_{h_p}^U$ )에 대해서는 오차 제곱합  $e_p^U$ 를 사용하여 코스트 함수로 정의하고, 부등식 관계를 만족시키고 있는 끝점( $[O_p]_{h_p}^L$ )에 대해서는 오차 제곱합  $e_p^L$ 에  $\omega$ 를 곱한 것을 코스트 함수로 하고 있다.

(2)  $y_p < [O_p]_{h_p}^L \leq [O_p]_{h_p}^U$ 가 성립될 경우

$$e_p = e_p^L + \omega \cdot e_p^U \quad (30)$$

이 경우에는 상한( $[O_p]_{hp}^U$ )이 부등식 관계를 만족시키고 있으므로, 오차 제곱합  $e_p^U$ 에  $\omega$ 를 곱하고 있다.

(3)  $[O_p]_{hp}^L \leq y_p \leq [O_p]_{hp}^U$ 가 성립될 경우

$$e_p = \omega \cdot e_p^L + \omega \cdot e_p^U \quad (31)$$

이 경우에는 양끝점( $[O_p]_{hp}^L$ 과  $[O_p]_{hp}^U$ ) 모두 부등식 관계를 만족시키고 있기 때문에 오차 제곱합  $e_p^L$ 과  $e_p^U$ 에  $\omega$ 를 각각 곱하고 있다.

식 (27),(30),(31)에 곱해져 있는  $\omega$ 의 값은 수치 실험에 의하여, 0.01-0.001 사이의 값이 적합하다.  $\omega$ 의 값이 너무 적은 경우(0.001 이하)는 학습의 속도가 늦어지고, 너무 큰 경우(0.01 이상)는 퍼지 신경망으로부터의 퍼지 출력치가 목표 출력치를 포함하는 정도가 낮아진다.

### 4.2 학습 알고리즘의 도출

앞에서 정의된 코스트 함수를 이용하여 퍼지 신경망의 학습 알고리즘을 도출한다. 먼저, 식(27),(30),(31)에서 정의된 코스트 함수( $e_p$ )를 이용하여 중간층과 출력층 사이의 퍼지 연결강도( $W_j$ )에 대한 수정량을 구한다. 연결강도  $W_j$ 가 대칭 삼각형 퍼지수인 것을 고려해서 그 하한( $w_j^L$ )과 상한( $w_j^U$ )의 수정을 행하기로 한다. 이때 중심( $w_j^C$ )은 식(9)로부터 결정된다. 퍼지 연결강도  $W_j$ 의 하한( $w_j^L$ )과 상한( $w_j^U$ )에 대하여 Rumelhart 등[6]과 같은 방법으로 수정량을 각각 다음과 같이 정의한다.

$$\Delta w_j^L(t+1) = \eta(-\partial e_p / \partial w_j^L) + \alpha \Delta w_j^L(t) \quad (32)$$

$$\Delta w_j^U(t+1) = \eta(-\partial e_p / \partial w_j^U) + \alpha \Delta w_j^U(t) \quad (33)$$

여기서  $\eta$ 는 학습 윋수,  $\eta$ 는 학습률(Learning Rate),  $\alpha$ 는 모멘텀항 계수(Momentum Constant)이다.  $\eta$ 는 학습의 진동이 일어나지 않는 범위에서 가능한 한 큰 계수, 즉 1 정도가 적합하다.  $\alpha$ 는 0.9 정도로 취함으로써 같은 해를 빨리 얻을 수 있다. 학습 윋수  $\eta$ 는 학습용 데이터의 수에 따라 다르지만, 본 연구에서의 수치에 정도이면 10,000회 정도로 충분하다.  $\partial e_p / \partial w_j^L$  및  $\partial e_p / \partial w_j^U$ 는 식(10)-(14) 및 식(16)-(22)를 고려함으로써, 식(27), (30),(31)의 코스트 함수로부터 도출할 수 있다. 구체적인 수정량의 도출 방법은 부록에 나타낸다. 또한 식(32), (33)을 이용하여 퍼지 연결강도  $W_j(t+1)$ 의 하한  $w_j^L(t+1)$ , 중심  $w_j^C(t+1)$ 과 상한  $w_j^U(t+1)$ 을 다음과 같이 수정한다.

$$w_j^L(t+1) = w_j^L(t) + \Delta w_j^L(t+1) \quad (34)$$

$$w_j^C(t+1) = w_j^C(t) + \Delta w_j^C(t+1) \quad (35)$$

$$w_j^U(t+1) = w_j^U(t) + \Delta w_j^U(t+1) \quad (36)$$

이와 같은 수정으로부터 퍼지 연결강도  $W_j(t+1)$ 의 하한  $w_j^L(t+1)$ 과 상한  $w_j^U(t+1)$ 과의 역전이 일어날 경우를 고려하여 수정 후의 퍼지 연결강도  $W_j(t+1)$ 를 다음과 같이 설정한다.

$$W_j(t+1) = (w_j^L, w_j^C, w_j^U) \quad (37)$$

여기서

$$w_j^L = \text{Min} \{ w_j^L(t+1), w_j^U(t+1) \} \quad (38)$$

$$w_j^C = (w_j^L + w_j^U) / 2 \quad (39)$$

$$w_j^U = \text{Max} \{ w_j^L(t+1), w_j^U(t+1) \} \quad (40)$$

이다.

입력층 유닛과 중간층 유닛 사이의 퍼지 연결강도( $W_{ji}$ )와 퍼지 임계치( $\theta, \theta_j$ )의 수정도 퍼지 연결강도  $W_j$ 와 같은 방법으로 행할 수 있다(부록 참조).

### 5. 컴퓨터 시뮬레이션

#### 5.1 수치예 1

1입력 1출력 시스템에 관한 학습용 데이터로서 신뢰도(소속함수 값)를 가진 입출력 쌍  $(x_p, y_p, h_p)$ 이 다음과 같이 주어져 있다고 한다.

$$\{(x_p, y_p, h_p)\} = \{(0.1, 0.2, 0.25), (0.1, 0.35, 1.0), (0.1, 0.5, 0.25), (0.3, 0.4, 0.5), (0.3, 0.5, 1.0), (0.3, 0.6, 0.5), (0.5, 0.15, 0.25), (0.5, 0.3, 1.0), (0.5, 0.45, 0.25), (0.7, 0.4, 0.75), (0.7, 0.45, 1.0), (0.7, 0.5, 0.75), (0.9, 0.6, 0.5), (0.9, 0.7, 1.0), (0.9, 0.8, 0.5)\}$$

이와 같은 학습용 데이터를 그림 1에 나타낸다. 그림 1의 학습용 데이터를 사용하여, 앞에서 나타낸 학습 알고리즘에 의한 퍼지 신경망의 학습을 행하였다. 이때 중간층의 유닛 수는 8, 학습률( $\eta$ )은 1.0, 모멘텀항 계수( $\alpha$ )는 0.9,  $\omega$ 는 0.004로 설정하였다. 퍼지 신경망의 학습은 식(27),(30),(31)의 코스트 함수를 이용하여 각각의 패턴에 대하여 10,000회 행하였다. 학습 후의 퍼지 신경망으로부터의 퍼지 출력치를 나타내면 그림 2, 3과 같이 된다. 그림 2 및 그림 3으로부터 추정된 퍼지 출력치( $O_p$ )의  $h_p$ -레벨 집합( $[O_p]_{h_p}$ )이 소속함수 값을 가진 목표 출력치( $y_p$ )를 근사적

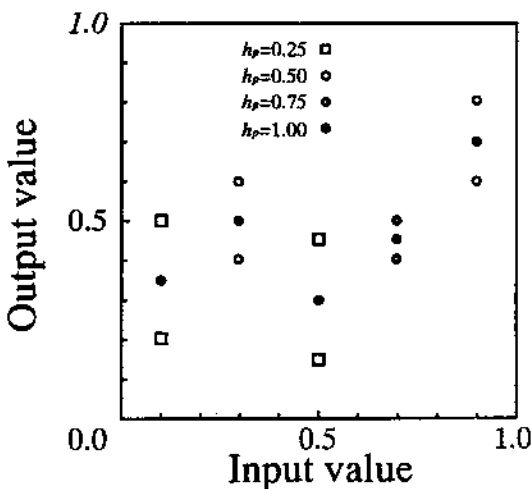


그림 1. 학습용 데이터

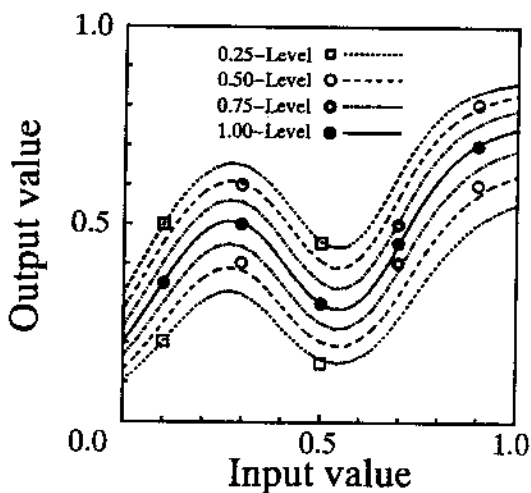


그림 2. 퍼지 신경망으로부터의 퍼지 출력치의 레벨 집합

으로 포함하고 있다는 것을 알 수 있다. 다시 말하여, 각 데이터의 신뢰도를 반영한 비선형 퍼지 회귀분석이 행해져 있다는 것을 알 수 있다.

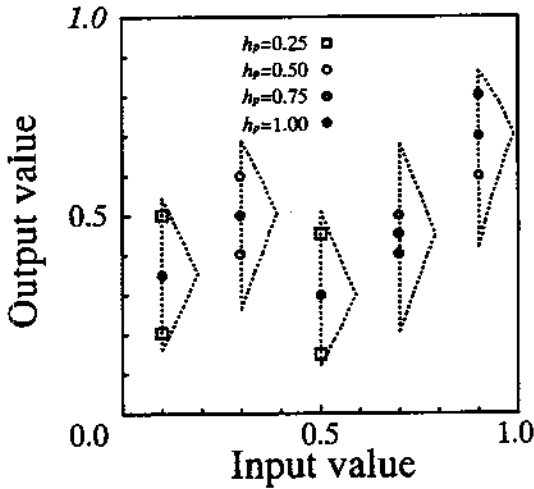


그림 3. 추정된 퍼지수(수치에 1)

5.2 수치에 2

2입력 1출력 시스템에 관한 학습용 데이터가 표 1과 같이 주어져 있다고 한다. 주어진 학습용 입력력 데이터 쌍( $x_{p1}, x_{p2}, y_p, h_p$ )을 이용하여 입력층에 2 개, 출력층에 1 개의 유닛을 가진 퍼지 신경망의 학습을 행하였다. 단, 학습 조건은 수치에 1과 동일하다. 학습 후의 퍼지 신경망으로부터의 퍼지 출력치를 나타내면 그림 4와 같다. 이 결과로부터 입력치가 2차원인 9 개의 샘플에 대한 추정된 퍼지 출력치의  $h_p$ -레벨 집합이 신뢰도를 가진 목표 출력치를 근사적으로 포함하고 있다는 것을 알 수 있다. 또한 주어진  $m$  개의 입력력 데이터 쌍에 대해, 퍼지 신경망의 학습으로부터 최소화되는 코스트 함수( $e$ )의 값을

표 1. 신뢰도를 가진 2입력 1출력 시스템에 관한 학습용 데이터

Sample No.	$p$	입 력		출 력	신뢰도	Sample No.	$p$	입 력		출 력	신뢰도
		$x_{p1}$	$x_{p2}$					$x_{p1}$	$x_{p2}$		
1	1	0.9	0.5	0.2	0.5	6	16	0.5	0.9	0.2	0.75
	2	0.9	0.5	0.3	1.0		17	0.5	0.9	0.3	1.0
	3	0.9	0.5	0.5	0.5		18	0.5	0.9	0.4	0.75
2	4	0.7	0.5	0.4	0.75	7	19	0.5	0.7	0.3	0.5
	5	0.7	0.5	0.5	1.0		20	0.5	0.7	0.5	1.0
	6	0.7	0.5	0.6	0.75		21	0.5	0.7	0.6	0.5
3	7	0.1	0.5	0.1	0.25	8	22	0.5	0.1	0.2	0.5
	8	0.1	0.5	0.3	1.0		23	0.5	0.1	0.3	1.0
	9	0.1	0.5	0.5	0.25		24	0.5	0.1	0.6	0.5
4	10	0.3	0.5	0.4	0.5	9	25	0.5	0.3	0.2	0.25
	11	0.3	0.5	0.5	1.0		26	0.5	0.3	0.5	1.0
	12	0.3	0.5	0.7	0.5		27	0.5	0.3	0.7	0.25
5	13	0.5	0.5	0.4	0.25						
	14	0.5	0.5	0.7	1.0						
	15	0.5	0.5	0.9	0.25						



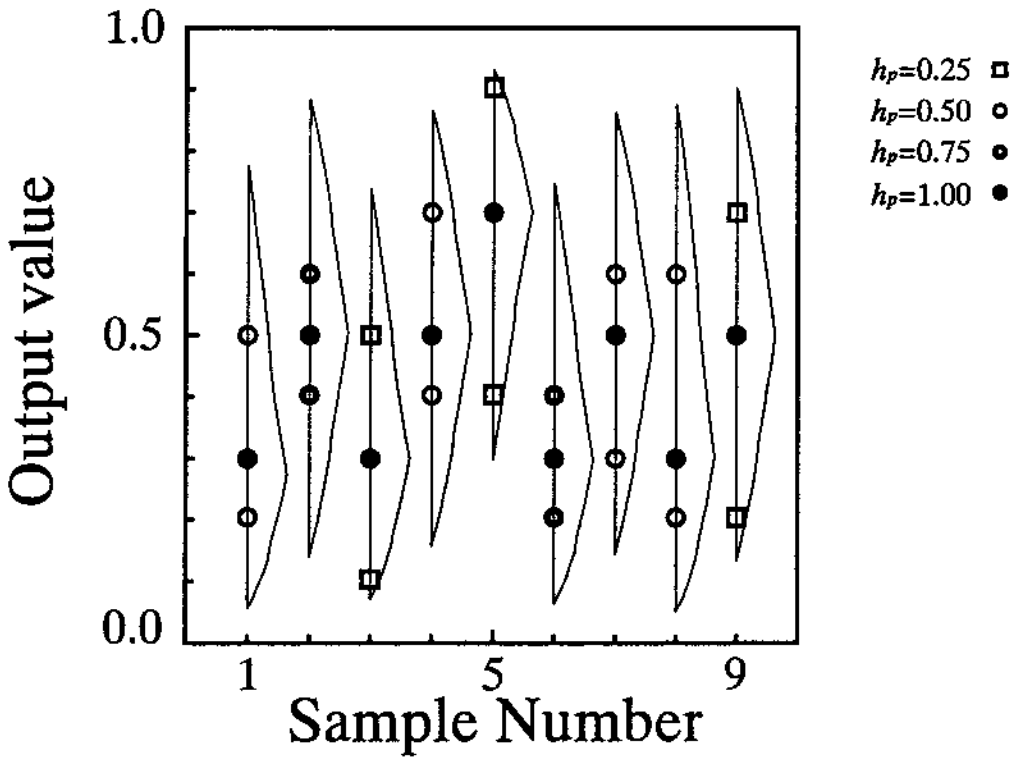


그림 4. 추정된 퍼지수(수치에 2)

표 2에 나타낸다. 코스트 함수  $e$ 는  $m$  개의 입출력 데이터에 대해, 식(27),(30),(31)에 정의된 코스트 함수  $e_p$ 의 총합을 취함으로써 다음과 같이 정의된다.

$$e = \sum_{p=1}^m e_p \quad (41)$$

표 2는 몇 가지 중간층 유닛 수에 대한 학습 결과이다. 표 2로부터, 반복 학습 횟수가 증가함에 따라 학습용 데이터에 대한 적합 능력이 높아지고 있다는 것을 알 수 있고, 중간층 유닛 수가 8, 10인 경우가 학습면에서 가장 적합하다는 것을 알 수 있다.

표 2. 코스트 함수  $e$ 

Iterations	Hidden Units			
	6	8	10	12
1	0.4182	0.9272	0.3285	1.2684
500	0.0808	0.0821	0.0466	0.0392
1000	0.0082	0.0108	0.0084	0.0083
5000	0.0051	0.0036	0.0024	0.0028
10000	0.0036	0.0023	0.0022	0.0024

## 6. 결 론

본 연구에서는 주어진 입출력 데이터에 신

뢰도를 나타내는 소속함수 값이 붙여진 경우에 대하여 퍼지 신경망에 의한 퍼지 회귀분석을 행하는 기법을 제안하였다. 신뢰도를 도입한 비선형 퍼지 회귀분석을 행하기 위해, 퍼지 신경망으로부터의 퍼지 출력치의 레벨 집합과 소속함수 값을 가진 목표 출력치를 이용하여 코스트 함수를 정의하고 학습 알고리즘을 도출하였다. 이 퍼지 신경망의 학습은 주어진 목표 출력치가 퍼지 신경망으로부터의 퍼지 출력치의 대응하는 레벨의 레벨 집합에 포함되도록 행해진다. 제안된 기법은 입출력 데이터에 대한 신뢰도에 차이가 보일 경우에 유효하고, 모델의 함수형을 정할 수 없는 문제도 다룰 수 있다. 또한 인간의 주관적인 판단, 평가 등과 같이 모호성을 포함하는 시스템의 모델화에 적합한 방법이다.

본 연구에서 제안한 기법의 유효성에 관한 실증적인 고찰은 다음의 연구 과제로 하고 싶다.

#### 참 고 문 헌

- [1] Tanaka H., Hayashi I. and Watada J., "Possibilistic Linear Regression Analysis for Fuzzy Data", European Journal of Operational Research, Vol.40, pp.389-396, 1989.
- [2] Tanaka H., Ishibuchi H. and Hwang S. G., "Fuzzy Model of the Number of Staff in Local Government by Fuzzy Regression Analysis with Similarity Relations", Journal of Japan Industrial Management Association, Vol. 41, No.2, pp. 99-104, 1990.
- [3] Hayashi I. and Tanaka H., "The Fuzzy GMDH Algorithm by Possibility Models and Its Application", Fuzzy Sets and Systems, Vol.36, pp.245-258, 1990.
- [4] Ishibuchi H. and Tanaka H., "Fuzzy Regression Analysis Using Neural Networks", Fuzzy Sets and Systems, Vol.50, pp.257-265, 1992.
- [5] Ishibuchi H., Tanaka H. and Okada H., "Fuzzy Neural Networks with Fuzzy Weights and Fuzzy Biases", Proc. of 1993 IEEE International Conference on Neural Network, San Francisco, pp. 1650-1655, 1993.
- [6] Rumelhart D.E., McClelland J.L. and PDP Research Group, Parallel Distributed Processing, Vol.1, MIT Press, Cambridge, 1986.
- [7] Zadeh L.A., "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning: Part 1,2 and 3", Information Sciences, Vol.8, pp.301-357 and Vol. 9, pp.43-80, 1975.

부 록

식(27), (30), (31)에 의해 정의된 코스트 함수( $e_p$ )를 이용하여 퍼지 신경망의 학습 알고리즘을 도출한다. 여기서는 코스트 함수( $e_p$ )를 다음과 같이 일반적으로 표현하기로 한다.

$$e_p = \nu \cdot e_p^L + \mu \cdot e_p^U \tag{A.1}$$

단,  $[O_p]_{hp}^L \leq [O_p]_{hp}^U < y_p$  일 때  $\nu = \omega, \mu = 1, y_p < [O_p]_{hp}^L \leq [O_p]_{hp}^U$  일 때  $\nu = 1, \mu = \omega$  이고,  $[O_p]_{hp}^L \leq y_p \leq [O_p]_{hp}^U$  일 때  $\nu = \omega, \mu = \omega$  이다.

식(32), (33)에서 나타낸 것 같이 퍼지 연결강도의 수정량은 다음과 같이 정의된다.

$$\Delta w_j^L(t+1) = \eta (-\partial e_p / \partial w_j^L) + \alpha \Delta w_j^L(t) \tag{A.2}$$

$$\Delta w_j^U(t+1) = \eta (-\partial e_p / \partial w_j^U) + \alpha \Delta w_j^U(t) \tag{A.3}$$

$$\Delta w_{ji}^L(t+1) = \eta (-\partial e_p / \partial w_{ji}^L) + \alpha \Delta w_{ji}^L(t) \tag{A.4}$$

$$\Delta w_{ji}^U(t+1) = \eta (-\partial e_p / \partial w_{ji}^U) + \alpha \Delta w_{ji}^U(t) \tag{A.5}$$

식(A.1)의 코스트 함수에 의해,  $\partial e_p / \partial w_j^L$  등의 미계수는 각각 다음과 같이 도출된다.

(1)  $\partial e_p / \partial w_j^L$ 의 도출

(a)  $[W_j]_{hp}^L \geq 0$ 일 때

$$\begin{aligned} \frac{\partial e_p}{\partial w_j^L} &= \nu \cdot \frac{\partial e_p^L}{\partial w_j^L} + \mu \cdot \frac{\partial e_p^U}{\partial w_j^L} \\ &= \nu \cdot \frac{\partial e_p^L}{\partial [O_p]_{hp}^L} \cdot \frac{\partial [O_p]_{hp}^L}{\partial [Net_p]_{hp}^L} \cdot \frac{\partial [Net_p]_{hp}^L}{\partial [W_j]_{hp}^L} \cdot \frac{\partial [W_j]_{hp}^L}{\partial w_j^L} \\ &\quad + \mu \cdot \frac{\partial e_p^U}{\partial [O_p]_{hp}^U} \cdot \frac{\partial [O_p]_{hp}^U}{\partial [Net_p]_{hp}^U} \cdot \frac{\partial [Net_p]_{hp}^U}{\partial [W_j]_{hp}^U} \cdot \frac{\partial [W_j]_{hp}^U}{\partial w_j^U} \\ &= -\delta_p^L \cdot [O_{pj}]_{hp}^L \cdot (1-h/2) - \delta_p^U \cdot [O_{pj}]_{hp}^U \cdot h/2 \end{aligned} \tag{A.6}$$

여기서

$$\delta_p^L = \nu \cdot (y_p - [O_p]_{hp}^L) \cdot [O_p]_{hp}^L \cdot (1 - [O_p]_{hp}^L) \quad (\text{A.7})$$

$$\delta_p^U = \mu \cdot (y_p - [O_p]_{hp}^U) \cdot [O_p]_{hp}^U \cdot (1 - [O_p]_{hp}^U) \quad (\text{A.8})$$

이다.

(b)  $[W_j]_{hp}^U < 0$ 일 때

$$\frac{\partial e_p}{\partial w_j^L} = -\delta_p^L \cdot [O_{pj}]_{hp}^U \cdot (1 - h/2) - \delta_p^U \cdot [O_{pj}]_{hp}^L \cdot h/2 \quad (\text{A.9})$$

(c)  $[W_j]_{hp}^L < 0 \leq [W_j]_{hp}^U$ 일 때

$$\frac{\partial e_p}{\partial w_j^L} = -\delta_p^L \cdot [O_{pj}]_{hp}^U \cdot (1 - h/2) - \delta_p^U \cdot [O_{pj}]_{hp}^U \cdot h/2 \quad (\text{A.10})$$

(2)  $\partial e_p / \partial w_j^U$ 의 도출

(a)  $[W_j]_{hp}^L \geq 0$ 일 때

$$\frac{\partial e_p}{\partial w_j^U} = \delta_p^L \cdot [O_{pj}]_{hp}^L \cdot h/2 - \delta_p^U \cdot [O_{pj}]_{hp}^U \cdot (1 - h/2) \quad (\text{A.11})$$

(b)  $[W_j]_{hp}^U < 0$ 일 때

$$\frac{\partial e_p}{\partial w_j^U} = \delta_p^L \cdot [O_{pj}]_{hp}^U \cdot h/2 - \delta_p^U \cdot [O_{pj}]_{hp}^L \cdot (1 - h/2) \quad (\text{A.12})$$

(c)  $[W_j]_{hp}^L < 0 \leq [W_j]_{hp}^U$ 일 때

$$\frac{\partial e_p}{\partial w_j^U} = \delta_p^L \cdot [O_{pj}]_{hp}^U \cdot h/2 - \delta_p^U \cdot [O_{pj}]_{hp}^U \cdot (1 - h/2) \quad (\text{A.13})$$

(3)  $\partial e_p / \partial w_{ji}^L$ 의 도출

(a)  $[W_j]_{hp}^L \geq 0$ 일 때

$$\frac{\partial e_p}{\partial w_{ji}^L} = \delta_{pj}^{LU} \cdot o_{pi} \cdot (1 - h/2) - \delta_{pj}^{UU} \cdot o_{pi} \cdot h/2 \quad (\text{A.14})$$

(b)  $[W_j]_{hp}^U < 0$ 일 때

$$\frac{\partial e_p}{\partial w_{ji}^L} = \delta_{pj}^{LU} \cdot o_{pi} \cdot h/2 - \delta_{pj}^{UL} \cdot o_{pi} \cdot (1 - h/2) \quad (\text{A.15})$$

(c)  $[W_j]_{hp}^L < 0 \leq [W_j]_{hp}^U$  일 때

$$\frac{\partial e_p}{\partial w_{ji}^L} = \delta_{pj}^{LU} \cdot o_{pi} \cdot h/2 - \delta_{pj}^{UU} \cdot o_{pi} \cdot h/2 \tag{A.16}$$

여기서

$$\delta_{pj}^{LL} = \delta_p^L \cdot [W_j]_{hp}^L \cdot [O_{pj}]_{hp}^L \cdot (1 - [O_{pj}]_{hp}^L) \tag{A.17}$$

$$\delta_{pj}^{UL} = \delta_p^U \cdot [W_j]_{hp}^U \cdot [O_{pj}]_{hp}^L \cdot (1 - [O_{pj}]_{hp}^L) \tag{A.18}$$

$$\delta_{pj}^{LU} = \delta_p^L \cdot [W_j]_{hp}^L \cdot [O_{pj}]_{hp}^U \cdot (1 - [O_{pj}]_{hp}^U) \tag{A.19}$$

$$\delta_{pj}^{UU} = \delta_p^U \cdot [W_j]_{hp}^U \cdot [O_{pj}]_{hp}^U \cdot (1 - [O_{pj}]_{hp}^U) \tag{A.20}$$

이다.

(4)  $\partial e_p / \partial w_{ji}^U$ 의 도출

(a)  $[W_j]_{hp}^L \geq 0$ 일 때

$$\frac{\partial e_p}{\partial w_{ji}^U} = \delta_{pj}^{LU} \cdot o_{pi} \cdot h/2 - \delta_{pj}^{UU} \cdot o_{pi} \cdot (1 - h/2) \tag{A.21}$$

(b)  $[W_j]_{hp}^U < 0$ 일 때

$$\frac{\partial e_p}{\partial w_{ji}^U} = \delta_{pj}^{LU} \cdot o_{pi} \cdot (1 - h/2) - \delta_{pj}^{UL} \cdot o_{pi} \cdot h/2 \tag{A.22}$$

(c)  $[W_j]_{hp}^L < 0 \leq [W_j]_{hp}^U$ 일 때

$$\frac{\partial e_p}{\partial w_{ji}^U} = \delta_{pj}^{LU} \cdot o_{pi} \cdot (1 - h/2) - \delta_{pj}^{UU} \cdot o_{pi} \cdot (1 - h/2) \tag{A.23}$$

퍼지 임계치  $\theta, \theta_j$ 의 수정은 퍼지 연결장도  $W_j, W_{ji}$ 와 같은 방법으로 각각 행할 수 있다.