

퍼지비선형회귀모형

Fuzzy Nonlinear Regression Model

황승국* · 박영만* · 서유진* · 박광박**

Seung Gook Hwang*, Young Man Park*, Yoo Jin Seo* and Kwang Pak Park**

*경남대학교 산업공학과, **경남대학교 수학과

요 약

이 논문은 퍼지비선형회귀모형에 대한 것으로서 유전적 알고리즘을 이용한 퍼지회귀분석모형을 제안한다. 유전적 알고리즘이란 좀 더 나은 퍼지회귀분석을 위하여 입력데이터를 분류하는데 사용되어진다. 이 분할에서 각 데이터는 분류된 데이터그룹에 속하는 멤버쉽함수의 값을 가지게 된다. 데이터 그룹은 각 변수의 영역을 최적으로 분할함에 따라 몇 개의 퍼지선형회귀모형에서 서로 다른 퍼지파라메타를 가지게 된다. 데이터에 대한 최종 퍼지수를 얻기 위하여 각 데이터그룹의 퍼지출력을 구성한다. 이 방법의 유효성은 사례연구에 의하여 보이고자 한다.

ABSTRACT

This paper is to propose the fuzzy regression model using genetic algorithm which is fuzzy nonlinear regression model. Genetic algorithm is used to classify the input data for better fuzzy regression analysis. From this partition, each data can be have the grade of membership function which is belonged to a divided data group. The data group, from optimal partition of the region of each variable, have different fuzzy parameters of fuzzy linear regression model one another. We compound the fuzzy output of each data group so as to obtain the final fuzzy number for a data. We show the efficiency of this method by means of demonstration of a case study.

1. 서 론

가능성 선형시스템을 이용한 선형회귀분석은 Tanaka[1]에 의하여 제안되었으며, 그동안 많은 연구가 이루어졌다[2-7]. 가능성을 이용한 선형회귀분석을 퍼지회귀분석이라 부르고 있다. 이 퍼지회귀분석과 종래의 회귀분석의 차이점을 살펴보면, 종래의 회귀분석은 확률모형을 기초로 하는 반면에 퍼지회귀분석은 가능성모형을 근거로 하고 있다. 또한, 종래의 회귀분석에서는 측정데이터와 모델에 의한 추정치와의 차는 관측오차로 간주하지만, 퍼지회귀모형에서는 모형과 데이터와의 차는 퍼지회귀모형에서의 시스템이 가지는 구조의 퍼지니스에 의존하고 있다고 가정하고 있다[8]. 따라서, 시스템이 가지는 구조의 퍼지니스를 파라메타의 퍼지니스로 변환시킨다는 의미로 퍼지회귀모형의 파라메타는 퍼지수로 나타낸다.

여기서, 퍼지회귀모형을 정식화하기 위하여 Ta-

naka는 대칭삼각형 퍼지파라메타를 가지는 선형계획법을 사용하고 있다. 이것은 선형계획법으로서 시스템의 파라메타를 구한다는 것을 의미한다. 이와 같이 선형계획법을 사용하는 것은 회귀문제를 선형계획법으로 변환하여 문제를 푸는 것에 의하여 퍼지회귀모형의 파라메타를 쉽게 구할 수 있다는 것이 장점이라 할 수 있다.

또한, 뉴럴네트워크를 이용한 퍼지회귀분석이 Ishibuchi[9]에 의하여 제안되어 있다. 이 방법의 특징은 뉴럴네트워크의 방법에 의하여 데이터에 대한 상한값과 하한값의 추정치를 얻을 수 있다는 것이다. 이것은 비선형모형에 적합하다고 할 수 있다. 이 방법외에 유전적 알고리즘[10-12]이라고 불리는 것이 있는데, 이것은 생물학적인 진화이론을 수학적인 알고리즘을 만드는데 응용한 것으로서 비선형모형에 대한 최적해를 얻을 수 있는 방법이다. 유전적 알고리즘의 장점은 교차나 돌연변이와 같은 확률적인 조작에 의하여 국소

*이 논문은 경남대학교 학술연구조성비에 의하여 작성된 것입니다.

해를 피할 수 있는 가능성을 가지고 있다는 것이다.

이 논문에서는 퍼지비선형회귀모형에 대한 것으로서 유전적 알고리즘을 이용한 퍼지회귀모형을 제안하고자 한다. 논문의 구성은 다음과 같다. 첫째로, 퍼지회귀모형과 유전적 알고리즘을 이용한 퍼지회귀모형의 차이를 파악하기 위하여 퍼지회귀분석모형에 대하여 기술한다. 둘째로, 일반적으로 알려져 있는 유전적 알고리즘에 대하여 기술한다. 셋째로, 이 논문에서 유전적 알고리즘을 직접 사용한 곳으로서 주어진 데이터를 분할하는 방법을 제시한다. 넷째로, 유전적 알고리즘을 이용한 퍼지회귀분석방법을 제안한다. 마지막으로 제안한 이 방법의 유효성을 나타내기 위하여 사례연구를 들고자 한다.

2. 퍼지회귀분석

다음으로 퍼지선형시스템을 모형로 하는 퍼지회귀분석에 대해서 기술한다.

대칭삼각형 퍼지수 A 를 $A=(\alpha, c)_L$ 라고 표현하고 그 멤버쉽함수는 다음과 같이 정의된다.

$$\mu_A(x) = L((x - \alpha)/c); c > 0 \quad (1)$$

단, $L(x)$ 는 형함수(type function)이며 $L(x)=1-|x|$ 로 표현한다. 그리고 α 는 중심을 c 는 폭을 나타내고 있다.

다음으로 퍼지수의 연산을 정의하기 위해서 사용되는 확장원리에 대해 기술한다. 함수 $y=f(x_1, x_2, \dots, x_n)$ 가 주어졌을 때 퍼지입력 A_1, A_2, \dots, A_n 에 대한 퍼지출력 $Y=f(A_1, A_2, \dots, A_n)$ 의 멤버쉽함수는 다음과 같이 정의된다.

$$\mu_A(y) = \sup \mu_{A_1}(x_1) \wedge \dots \wedge \mu_{A_n}(x_n) \quad (2)$$

$$\{x_1, x_2, \dots, x_n \mid y = f(x_1, x_2, \dots, x_n)\}$$

단, $\{x_1, x_2, \dots, x_n \mid y=f(x_1, x_2, \dots, x_n)\}$ 는 y 를 고정했을 때 y 의 값을 얻기 위한 x 의 집합을 나타내고, 이 집합 상에서 sup을 취하는 것을 의미하고 있다. 이것이 확장원리라고 불리는 것은 x 의 함수가 퍼지수의 함수로 확장되어 있기 때문이다. 여기서, \wedge 는 최소를 나타낸다.

대칭삼각형 퍼지수의 연산은 다음과 같이 정의된다.

(1) 가법 : 두개의 대칭삼각형 퍼지수 $A_1=(\alpha_1, c_1)_L$ 과 $A_2=(\alpha_2, c_2)_L$ 에 대해서

가법은 다음과 같이 정의된다.

$$(\alpha_1, c_1)_L + (\alpha_2, c_2)_L = (\alpha_1 + \alpha_2, c_1 + c_2)_L \quad (3)$$

(2) 감법 : 두개의 대칭삼각형 퍼지수 $A_1=(\alpha_1, c_1)_L$ 과 $A_2=(\alpha_2, c_2)_L$ 에 대해서 감법은 다음과 같이 정의된다.

$$(\alpha_1, c_1)_L - (\alpha_2, c_2)_L = (\alpha_1 - \alpha_2, c_1 + c_2)_L \quad (4)$$

(3) Scalar배 : 대칭삼각형 퍼지수 $A=(\alpha, c)_L$ 의 Scalar배는 다음과 같이 정의된다.

$$\lambda(\alpha, c)_L = (\lambda\alpha, |\lambda|c)_L \quad (5)$$

다음은 h-level에 의한 대칭삼각형퍼지수의 포함관계를 h-level set을 이용하여 다음과 같이 정의한다.

$$A \supset_h B \Leftrightarrow [A]_h \supset [B]_h \quad (6)$$

여기서, h-level set 는 퍼지집합 A, B에 대하여 각각

$$[A]_h \equiv \{x \mid \mu_A(x) \geq h\} \quad (7)$$

$$[B]_h \equiv \{x \mid \mu_B(x) \geq h\} \quad (8)$$

를 만족한다.

지금, 관측데이터(y_i, x_i), $i=1, \dots, N$ 이 있다고 하자. 여기서, y_i 는 출력변수이고, $x_i=(x_{i1}, \dots, x_{in})'$ 는 입력벡터이다. 퍼지선형회귀분석은 가능성모형에 따르기 때문에 이와 같은 데이터가 주어졌을 때, 퍼지회귀모형을 얻는 정식화를 하기 위하여 다음의 것을 가정한다.

(1) 다음과 같이 가능성선형시스템을 모형로 한다. 즉,

$$Y_i^* = A_0^* + A_1^* x_{i1} + \dots + A_n^* x_{in}, i=1, 2, \dots, N(9)$$

단, $[Y_i^*]_h$ 는 추정퍼지수이고, A_j^* 는 퍼지계수이다.

(2) 데이터에 있어서 출력치 y_i 는 추정퍼지수 Y_i^* 에 포함된다고 한다. 즉,

$$y_i \in_h Y_i^* \Leftrightarrow y_i \in [Y_i^*]_h, i=1, 2, \dots, N. \quad (10)$$

단, $[Y_i^*]_h$ 는 추정퍼지수 Y_i^* 의 h-level set 이다.

식(10)을 퍼지수의 포함관계식을 이용하여 바꾸어 쓰면 다음과 같이 된다.

$$y_i \leq \alpha x_i + (1-h) c |x_i| \quad (11)$$

$$y_i \geq \alpha x_i - (1-h) c |x_i| \quad (12)$$

$$i=1, 2, \dots, N$$

(3) 퍼지선형시스템에서 애매함의 지표인 $J(c)$ 를 다음과 같이 정의하고, $J(c)$ 가 작을수록 좋은 추정모형이라고 한다.

$$J(c) = \sum_i c |x_i| \quad (13)$$

단, $c|x_i|$ 는 추정퍼지수 Y_i^* 의 폭이다.

이상의 가정에 의해 y_i 가 추정퍼지수 Y_i^* 에 포함된다고 하는 조건하에서 추정퍼지폭의 합계인 $J(c)$ 를

최소화하는 계수 $A_j, j=1, 2, \dots, n$ 을 구하는 문제가 다음과 같이 정식화 된다. 여기서, 결정되는 것은 α 와 c 이다. 퍼지선형회귀분석은 식(11),(12)의 제약조건을 기본으로 식(13)의 $J(c)$ 를 최소로 하는 계수를 구하는 다음의 LP문제로 귀착된다.

$$\begin{aligned} \min J(c) &= \sum_i c |x_i| & (14) \\ \text{s.t.} & \text{식(11), } c \geq 0 \end{aligned}$$

3. 유전적 알고리즘

유전적 알고리즘은 생물학적 진화의 과정을 모방한 방법으로서 후보해의 그룹내에서 최적해를 찾는 데 적합하다. 개체는 0 또는 1 이라는 스트링을 가지는 유전자로 구성되어 있다. 이것은 개체가 스트링에 의하여 표현된다는 것을 의미한다. 개체의 집합을 세대라고 한다. 또한, 개체의 수는 세대의 크기라고 한다. 각 스트링은 적합도에 의하여 평가되어진다. 적합도의 양에 따라서 각각의 개체는 생존 또는 선택이 결정된다. 교차 또는 돌연변이와 같은 확률적 조작은 그룹의 진화에 대하여 반복적으로 행하여 진다.

스트링의 비트는 랜덤수에 의하여 초기화 되어 진다. 그리고, 각 개체의 적합도는 평가되어 지고, 각 개체의 선택은 적합도에 의하여 결정된다. 교차에 대한 조작은 스트링의 포인트를 선택하는 것이고, 포인트 주위를 중심으로 하여 스트링의 블록을 변화시키는 것이다. 돌연변이에 대한 조작은 스트링의 값인 0과 1에 대하여 0을 1로 또는 1을 0으로 하는 것이다.

4. 입력데이터의 분할

4.1 변수의 분할

입력데이터의 수를 n 이라 하고 독립변수의 수를 m 이라 하자. 만약 l_j 부분에 대하여 변수 x_j 의 영역을 나누다면, 입력데이터는 $w(=l_1 \times l_2 \times \dots \times l_n)$ 데이터그룹으로 나누어진다. C_k 로서 k 번째의 데이터그룹을 구성하는 x_j 의 퍼지집합에 대하여 기술한다. k 번째의 데이터그룹 G_k 는 식(15)와 같이 표현된다.

$$G_k = (C_{k1} \times C_{k2} \times \dots \times C_{km}) \quad (15)$$

i 번째 입력데이터 $D_i=(x_{i1}, x_{i2}, \dots, x_{in})$ 이 데이터그룹 G_k 에 속하는 멤버쉽함수 $g_{ik}=\mu_{Gk}(x_{i1}, x_{i2}, \dots, x_{in})$ 는 식(16)과 같이 표현된다.

$$\mu_{Gk}(x_{i1}, x_{i2}, \dots, x_{in}) = [\mu_{Ck1}(x_{i1}) \wedge \dots \wedge \mu_{Ckm}(x_{in})] \quad (16)$$

4.2 개체

유전적 알고리즘에 있어서 세대의 개체 I_r 은 각 변수의 가능한 영역 내에서의 분할점들로 나타내어지는 스트링에 의하여 식(17)과 같이 표현되어진다.

$$I_r = P_{11} \dots P_{1s_1} \dots P_{n1} \dots P_{nsn} \quad (17)$$

여기서, P_{jq} 는 변수 x_j 의 q 번째 분할점이고, s_j 는 변수 x_j 의 최대 분할수 이다. t_j 는 변수 x_j 의 한 분할점으로 표현할 수 있는 최소 스트링의 길이 이다. 개체는 비트($\sum_j t_j s_j$)로 표현한다. 예를 들면, 만약 하나의 변수 x_j 가 $[0,1]$ 의 값을 취하고, 가능분할구간이 0.01이라고 하면, 그 때의 t_j 는 7이다. 만약 s_j 가 4이라면 x_j 에 대하여 28비트가 필요한 것이다. 만약, 분할점이 동일하다면 분할의 수는 감소한다

4.3 퍼지파라메타

k 번째 데이터그룹 G_k 의 퍼지파라메타는 선형계획법문제를 푸는 것에 의하여 구하여 진다. 각 개체에 대하여 만약 데이터그룹의 수가 w 개 이라면 w 개의 선형계획법문제를 풀어야 한다. 데이터그룹 G_k 의 선형계획법문제는 식(18)과 같다.

$$\begin{aligned} \text{Min} & \sum_{ij} x_{ij} c_i & (18) \\ \text{s.t.} & \sum_{ij} x_{ij} a_{ij} + (1-h)/g_{ik} \sum_{ij} x_{ij} c_{ij} \geq y_i, i \in S_k \\ & \sum_{ij} x_{ij} a_{ij} - (1-h)/g_{ik} \sum_{ij} x_{ij} c_{ij} \geq y_i, i \in S_k \\ & c_{ij} \geq 0, i \in S_k \end{aligned}$$

여기서, h 는 h 레벨집합을 의미한다. S_k 는 그레이드 g_{ik} 가 ϵ 보다 큰 입력데이터의 집합이고, ϵ 는 그레이드 g 의 하한허용영역이다.

4.4 분할된 데이터그룹에 대한 출력의 합성

데이터그룹 G_k 에 대한 추정출력 Y_{ik} 는 식(19)에 의한 퍼지파라메타 A_{ik} 를 이용하여 구한다.

$$Y_{ik} = A_{0k} + A_{1k}x_{i1} + A_{2k}x_{i2} + \dots + A_{nk}x_{in} \quad (19)$$

g_{ik} 를 이용하여 Y_{ik} 를 합성한다. 최종추정출력 Y_i 는 식(20)으로 구한다.

$$Y_i = \sum_k g_{ik} Y_{ik} \quad (20)$$

5. 유전적알고리즘을 이용한 퍼지비선형회귀 분석의 순서

본 논문에서 제안하고 있는 유전적알고리즘을 이용한 퍼지비선형회귀분석의 순서는 다음과 같다.

- 순서1) 초기화
 개체의 초기화는 일양난수에 의하여 결정한다. 세대의 크기는 d 이다.
- 순서2) 적합도평가
 각 개체에 대하여 다음 사항을 반복한다.
- 1) 각 변수에 대하여 분할점을 분류한다. 만약 분할점이 동일한 값을 가지게 되면 두 값 중에서 한 값을 제거한다.
 - 2) 그레이드 g_{ij} 는 모든 입력데이터 $D_i, i=1, \dots, m$ 와 데이터그룹 $G_j, j=1, \dots, w$ 에 대하여 결정된다.
 - 3) 식(18)의 선형계획법 모형을 이용하여 각 데이터 그룹에 대한 퍼지파라메타를 결정한다.
 - 4) 각 입력데이터에 대하여 퍼지파라메타를 합성한다. 만약 추정된 최종출력에 실패데이터가 포함되어 있지 않다면 h 레벨을 높이고 $h < 1$ 까지 3)을 반복한다.

개체가 입력데이터에 대한 역중합폭은 적합도로 나타낸다.

- 순서3) 선택확률
 각 개체에 대한 선택확률은 자체의 적합도에 따라서 계산된다.
- 순서4) 교차
 선택확률에 따라서 서로 다른 쌍 $d/2$ 개를 선택한다. 각 쌍에 대하여 개체내의 포인트를 선택하고, 그 포인트 주위에 있는 스트링의 블록을 변화시킨다.
- 순서5) 돌연변이
 모든 새로운 개체의 각 비트에 대하여 미리 지정된 확률에 따라 돌연변이의 조작이 진행된다.
- 순서2)에서 순서5)까지의 1사이클을 세대라 한다.

6. 사례연구

Table 1. Input-output data

Work											
No	Sub	H.E	St.	Ar.	CFF	No	Sub	H.E	St.	Ar.	CFF
1	S1	20.33	11	5	0.93	5	S1	5.50	12	7	3.65
	S2	17.83	7	6	5.22		S2	4.50	8	6	0.00
	S3	15.33	4	5	0.00		S3	3.50	5	3	1.78
	S4	11.83	7	6	6.66		S4	1.75	9	4	0.00
	S5	11.50	7	7	0.00		S5	2.75	7	4	5.06
	S6	12.17	12	4	2.56		S6	5.50	7	6	0.00
	S7	17.67	7	5	0.00		S7	8.50	7	3	0.00
	S8	23.17	11	5	0.00		S8	3.75	7	6	6.45
2	S1	16.33	17	7	0.00	6	S1	2.00	4	2	0.93
	S2	12.83	11	6	0.00		S2	4.25	6	5	2.09
	S3	15.17	10	6	7.96		S3	3.50	4	4	10.62
	S4	8.67	12	5	12.37		S4	2.50	6	4	10.49
	S5	15.83	11	5	2.03		S5	4.75	5	2	3.03
	S6	17.67	7	4	7.69		S6	4.50	7	2	15.38
	S7	24.33	9	4	7.69		S7	5.50	9	3	12.82
	S8	12.00	11	6	9.68		S8	2.75	6	6	0.00
3	S1	13.83	14	8	0.93	7	S1	3.50	7	4	1.83
	S2	12.50	8	7	3.13		S2	2.00	4	6	4.16
	S3	13.83	9	6	7.96		S3	1.75	3	2	5.31
	S4	9.17	11	4	19.06		S4	1.25	7	3	6.66
	S5	9.83	11	2	1.00		S5	4.50	3	3	5.06
	S6	9.83	7	4	12.82		S6	5.50	7	2	10.26
	S7	8.67	10	8	10.26		S7	5.25	9	5	10.26
	S8	11.33	9	6	3.23		S8	3.00	5	8	0.00
4	S1	14.63	4	8	1.83	8	S1	3.25	9	1	0.93
	S2	12.33	8	7	3.13		S2	1.25	4	5	2.09
	S3	15.00	8	6	8.87		S3	3.75	3	5	12.40
	S4	9.33	11	3	7.63		S4	1.00	3	3	1.91
	S5	11.33	7	5	4.03		S5	2.25	5	4	4.03
	S6	10.83	9	6	0.00		S6	3.50	10	2	17.95
	S7	18.83	8	6	15.38		S7	9.75	7	4	12.82
	S8	19.66	10	6	3.23		S8	4.50	4	6	6.45

H.E: Human Error, CFF: Critical Fusion Frequency, St.: Stress, Ar.: Arousal

본 논문에서 제안한 방법의 사례연구로서 VDU (Video Display Unit) 작업자의 작업수행도에 대한 데이터[13]를 사용하여 그 유효성을 보이고자 한다.

VDU작업은 일반적으로 자료입력, 자료탐색, 대화작업의 3가지 형태로 구분할 수 있다. 그 중에서 첫 번째인 자료입력은 단순하고 반복적인 작업으로서 건강에 나쁜 영향을 미치며, 많은 인간에러를 유발시킨다고 볼 수 있다. 단조로운 VDU작업의 수행시 작업의 단조로움에 의한 스트레스를 받게 되어 작업수행도를 낮게 하는데 영향을 미치게 된다. 그리고, VDU작업은 작업자의 눈의 피로에 많은 영향을 주고 있다.

따라서, 본 사례연구에서는 VDU작업에 많은 영향을 미친다고 판단되는 인간의 심리·생리적인 요소 중 스트레스, 각성, CFF(Critical Fusion Frequency)의 3가지를 이용하여 VDU작업자에 대한 작업수행도를 인간에러로 평가하고자 하는 것으로서, 8명의 여자 대학생을 피실험자로 하였다.

실험절차로서는 본 실험을 하기 전에 예비실험을 피실험자가 실험내용을 충분히 이해할 수 있도록 실시하였다.

실험내용은 피실험자가 컴퓨터상에서 난수를 이용하여 생성된 1자리 또는 2자리의 숫자를 보고 그것을 연속적으로 덧셈, 뺄셈, 곱셈의 작업을 수행하는 것이다. 피실험자는 그 계산결과를 키보드를 이용하여 입력하고, 수정도 가능하지만 정해진 시간내에 처리해

야 하며, 잘못된 입력은 작업중 자동적으로 기록된다. 실험시간은 40분이며, 이 때 400 또는 600문제를 처리하도록 하여 이 실험의 수행도라 할 수 있는 인간에러를 구하고, 나머지 데이터는 매 실험 종료 후 측정하였으며, 10분의 휴식시간을 가진 후 다시 실험을 반복하여 하루 4회의 실험을 행하여 2일간의 실험결과를 정리한 것이 Table 1 이다.

Table 1에서의 No.는 실험회수를 의미한다. 8명의 피실험자가 8회의 실험을 하였으므로 데이터의 총수는 64개가 된다. 이 데이터를 제안한 방법을 이용하여 회귀분석을 실시하기 위한 모형식은 다음의 식(21)과 같다.

$$Y = A_0 + A_1x_1 + A_2x_2 + A_3x_3 \quad (21)$$

여기서, x_1 은 스트레스, x_2 는 각성, x_3 는 CFF를 의미하며, A_0, A_1, A_2, A_3 는 퍼지파라메타를 나타내며, A_0 는 수정항 이다.

제안한 방법을 이용하여 구한 분할데이터 그룹별 퍼지파라메타는 다음의 Table 2와 같다.

50세대에 대한 계산 후 19번째 세대에서 개체 수 10개 중 9번째 개체(01110100110111101)에서 총 분할 영역수가 $8(x_1 : 2, x_2 : 2, x_3 : 2)$ 일 때 제안한 방법의 폭의 합이 357.21로 최소가 되었다. 이 값은 퍼지회귀 분석에 의한 폭의 합이 493.05인 것에 비하면 비선형적인 방법이 27.5% 성능향상 되었다고 볼 수 있다.

Table 2의 퍼지파라메타를 이용하여 인간에러의 추

Table 2. Fuzzy parameters of partitin data group

No.	A_0	A_1	A_2	A_3
1	(17.72, 4.00)	(-0.30, 0.00)	(-0.36, 1.81)	(0.00, 0.00)
2	(3.63, 0.00)	(-0.02, 0.00)	(0.16, 1.27)	(0.00, 0.00)
3	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
4	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
5	(-0.98, 0.00)	(0.33, 0.00)	(3.50, 1.21)	(0.00, 1.07)
6	(0.22, 0.00)	(0.35, 0.29)	(0.66, 0.00)	(0.00, 0.00)
7	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)
8	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)	(0.00, 0.00)

Table 3. Real values and estimated values for human error

No	HE	하한값	중앙값	상한값	No	HE	하한값	중앙값	상한값
1	20.33	6.93	16.37	25.82	33	5.5	1.35	3.89	6.44
2	17.83	7.83	15.46	23.09	34	4.5	1.4	3.95	6.5
3	15.33	7.2	16.65	26.09	35	3.5	0.26	4.08	7.9
4	11.83	5.23	14.68	24.12	36	1.75	0.29	4.11	7.93
5	11.5	2.86	15.94	29.01	37	2.75	-0.9	4.19	9.29
6	12.17	6.45	15.89	25.34	38	5.5	0.29	4.11	7.93
7	17.67	5.03	16.29	27.55	39	8.5	-0.82	4.27	9.37
8	23.17	7.2	16.65	26.09	40	3.75	2.42	3.69	4.96
9	16.33	7.2	16.65	26.09	41	2	1.39	3.94	6.48
10	12.83	7.2	16.65	26.09	42	4.25	0.26	4.08	7.9

Table 3. Continued

No	HE	하한값	중앙값	상한값	No	HE	하한값	중앙값	상한값
11	15.17	6.06	14.76	23.45	43	3.5	2.87	5.96	9.05
12	8.67	6.61	10.1	13.60	44	2.5	2.19	5.25	8.3
13	15.83	2.26	15.33	28.41	45	4.72	1.36	3.9	6.45
14	17.67	6.4	12.6	18.80	46	4.5	2.5	6.97	11.45
15	24.33	5.76	13.57	21.38	47	5.5	3	6.73	10.47
16	12	5.72	9.22	12.71	48	2.75	1.4	3.95	6.5
17	13.83	11.27	17.09	22.90	49	3.5	1.38	3.92	6.47
18	12.5	6.28	15.72	25.17	50	2	0.23	4.05	7.87
19	13.83	6.06	14.76	23.45	51	1.75	0.21	4.03	7.85
20	9.17	8.82	12.31	15.81	52	1.25	0.19	4.01	7.83
21	9.83	6.91	16.35	25.80	53	4.5	2.44	3.71	4.98
22	9.83	6.76	10.25	13.75	54	5.5	2.18	5.17	8.15
23	8.67	8.2	12.9	17.61	55	5.25	2.18	5.17	8.15
24	11.33	8.42	16.05	23.68	56	3	1.4	3.95	6.5
25	14.67	11	16.82	22.63	57	3.25	1.39	3.94	6.48
26	12.33	6.28	15.72	25.17	58	1.25	-0.85	4.24	9.33
27	15	10.03	15.94	21.86	59	3.75	2.98	6.59	10.2
18	9.33	6.46	12.79	19.12	60	1	0.26	4.08	7.9
29	11.33	3.84	15.1	23.36	61	2.25	1.34	3.89	6.43
30	10.83	9.37	17	24.63	62	3.5	2.66	7.88	13.1
31	18.83	9.89	14.6	19.30	63	9.75	2.34	6.07	9.8
32	19.66	10.59	16.4	22.22	64	4.5	1.3	3.85	6.4

정치값을 구하면 Table 3과 같다. 인간에러에 대한 실제값에 대한 추정치값은 인간의 수행도에 대한 가능성을 파악할 수 있으며, 실제값과 추정치값의 차를 최소로 하는 추정방법인 모형을 개발하므로써 인간수행도에 대한 보다 정확한 정보를 얻을 수 있게 된다. Table 3에서 실제값과 추정치의 중앙값을 비교해 보았을 때 가까운 수치가 많은 것을 보았을 때 추정의 정도가 높다고 할 수 있다.

7. 결 론

본 논문에서 유전적알고리즘을 이용한 퍼지비선형회귀분석법을 제안하였다. 이 제안 방법은 비선형성의 입력데이터의 분할을 사용하므로 기존의 퍼지선형회귀분석에 비하여 유연성과 현실에의 응용성을 가지고 있다고 볼 수 있으며, 사례연구에서와 같이 인간의 정성적인 요인을 이용한 실험에서는 퍼지이론의 필요성이 더욱 강조되므로, 퍼지이론과 퍼지비선형회귀분석을 이용하여 인간의 정성적인 요인을 표현하고, 여기서 인간수행도에 대한 정보를 가능하면 정확하게 얻기 위하여 인간수행도에 대한 모형을 개발하는 데에 사용하는 것이 유효하다는 것을 알 수 있었다.

참고문헌

[1] H. Tanaka, S. Uejima and K. Asai, "Linear re-

gression analysis with fuzzy model", *IEEE Trans. SMC-12*(6), 903-907 (1982).
 [2] H. Tanaka, J. Watada and I. Hayashi, "On three formulations of fuzzy linear regression analysis", *J. of the Society of Instrument and Control Engineering* **22**(10), 1051-1057 (1986).
 [3] H. Tanaka, "Fuzzy data analysis by possibilistic linear models", *Fuzzy Sets and Systems*, **24**(3) 363-375 (1987).
 [4] H. Tanaka, I. Hayashi and J. Watada, "Possibilistic Linear Regression Analysis Based on Possibility Measure", *Proceedings of 2nd IFSA Congress*, 317-320 (1987).
 [5] I. Hayashi and H. Tanaka, "The Fuzzy GMDH Algorithm by Possibility Models", *Proceedings of 2nd IFSA Congress*, 321-324 (1987).
 [6] S.G. Hwang, I. Hayashi and H. Tanaka, "Analysis of Human Judgement by Interval", *Proceedings of IEA 10th International Congress*, 484-486 (1988).
 [7] S.G. Hwang and Y.J. Seo, "Estimation of the Actual Working Time by Interval Linear Regression Models with Constraint Conditions", *J. of the Korean OR/MS Society*. **14**(2), 105-114 (1989).
 [8] H. Ishibuchi and H. Tanaka, "Fuzzy regression analysis using neural networks", *Fuzzy Sets and Systems* **50**(3), 257-266 (1992).
 [9] H. Nomura, I. Hayashi and N. Wakami, "A Self-Tuning Method of Fuzzy Reasoning by Genetic Algorithm", *IFSICC 92*, 236-245 (1992).
 [10] D.E. Goldberg, *Genetic Algorithms*, Addison Wesley (1989).

- [11] Y.J. Seo, Y.M. Park, S.G. Hwang and K.P. Park, "Fuzzy Regression Analysis Using Genetic Algorithm", *Proceedings of 5th IFSA World Congress*, 513-516 (1993).
- [12] Y.C. Hong and S.S. Ohm, "Modeling of Nonlinear Dynamic Systems Using a Modified GMDH Algorithm", *J. of Fuzzy logic and Intelligent Systems* 8(3), 50-55 (1998).
- [13] Y.J. Seo, Y.M. Park, S.G. Hwang and M. Kumashiro, "A Fuzzy Model of Human Performance for VDU Workers", *J. of the Ergonomics Society of Korea* 14(1), 97-104 (1995).

황승국(Seung Gook Hwang)

1981년 : 동아대학교 공업경영학과(IE) 공학사
 1983년 : 동아대학교 공업경영학과(IE) 공학석사
 1991년 : 일본 오사카부립대학 경영공학과(IE) 공학박사
 1991년-현재 : 경남대학교 공과대학 산업공학과 부교수

서유진(Yoo Jin Seo)

1975년 : 동아대학교 공업경영학과(IE) 공학사
 1978년 : 동아대학교 공업경영학과(IE) 공학석사
 1997년 : 일본 동아대학교 학술박사 (정보시스템전공)
 1981년-현재 : 경남대학교 공과대학 산업공학과 교수

박영만(Young Man Park)

1980년 : 서울대학교 산업공학과 공학사
 1982년 : 서울대학교 산업공학과 공학석사
 1987년 : 서울대학교 산업공학과 박사과정수료
 1983년-현재 : 경남대학교 공과대학 산업공학과 부교수

박광박(Kwang-Pak Park)

1969년 : 부산대학교 수학과 학사
 1977년 : 부산대학교 수학과 석사
 1986년 : 경상대학교 수학과 이학박사
 1990년-1991년 : 일본 국립 오사카대학 교환교수
 1980년-현재 : 경남대학교 자연과학대학 수학과 교수