Journal of the Korea Academia-Industrial cooperation Society
Vol. 13, No. 11 pp. 5379-5388, 2012

http://dx.doi.org/10.5762/KAIS.2012.13.11.5379

HCM 클러스터링 알고리즘 기반 비퍼지 추론 시스템의 비선형 특성

박건준[1], 이동윤[2*]
[1]원광대학교 정보통신공학과, [2]중부대학교 전기전자공학과

Nonlinear Characteristics of Non-Fuzzy Inference Systems Based on HCM Clustering Algorithm

Keon-Jun Park[1] and Dong-Yoon Lee[2*]
[1]**Department of Information Communication Engineering, Wonkwang University**
[2]**Department of Electrical Electronic Engineering, Joongbu University**

요 약 비선형 공정에 대한 퍼지 모델링에서, 퍼지 규칙은 일반적으로 입력 변수 선택, 공간 분할 수 및 소속 함수에 의해 형성된다. 비선형 공정에 대한 퍼지 규칙의 생성은 차원이 증가할수록 규칙의 수가 지수적으로 증가하는 문제를 가지고 있다. 이를 해결하기 위해, 입력 공간의 퍼지 분할에 의한 퍼지 규칙을 생성함으로써 복잡한 비선형 공정을 모델링 할 수 있다. 따라서 본 논문에서는 HCM 클러스터링 알고리즘을 이용하여 입력 공간을 분산 형태로 분할함으로써 비퍼지 추론 시스템의 규칙을 생성한다. 규칙의 전반부 파라미터는 HCM 클러스터링 알고리즘에 의한 소속행렬로 결정된다. 규칙의 후반부는 다항식 함수의 형태로 표현되며, 각 규칙의 후반부 파라미터들은 표준 최소자승법에 의해 동정된다. 마지막으로, 비선형 공정으로는 널리 이용되는 데이터를 이용하여 비선형 특성 및 성능을 평가한다. 본 실험을 통해 고차원의 비선형 시스템은 매우 적은 수의 규칙을 가지고 모델링할 수 있었다.

Abstract In fuzzy modeling for nonlinear process, the fuzzy rules are typically formed by selection of the input variables, the number of space division and membership functions. The Generation of fuzzy rules for nonlinear processes have the problem that the number of fuzzy rules exponentially increases. To solve this problem, complex nonlinear process can be modeled by generating the fuzzy rules by means of fuzzy division of input space. Therefore, in this paper, rules of non-fuzzy inference systems are generated by partitioning the input space in the scatter form using HCM clustering algorithm. The premise parameters of the rules are determined by membership matrix by means of HCM clustering algorithm. The consequence part of the rules is represented in the form of polynomial functions and the consequence parameters of each rule are identified by the standard least-squares method. And lastly, we evaluate the performance and the nonlinear characteristics using the data widely used in nonlinear process. Through this experiment, we showed that high-dimensional nonlinear systems can be modeled by a very small number of rules.

Key Words : Non-Fuzzy Inference Systems, Hard C-Means Clustering Algorithm, Scatter Partition of Input Space, Rule Generation, Nonlinear Characteristics

1. 서론

퍼지 모델의 성능은 퍼지 규칙의 구성 방법에 의존하며 보다 좋은 성능을 위해서는 퍼지 규칙의 동정이 필연적이다. 퍼지시스템 이론의 발전으로 퍼지모델 동정 알고리즘의 접근 방식도 향상되었다. 초기 퍼지 모델의 동정연구로는 언어적 접근 방식[1,2]과 퍼지 관계 방정식에 기초한 접근방식[3,4]이 제안되었다. 언어적 접근방식에서, Tong은

*Corresponding Author : Dong-Yoon Lee
Tel: +82-10-8377-6117 email: dylee@joongbu.ac.kr
접수일 12년 08월 28일 수정일 12년 09월 27일 게재확정일 12년 11월 08일

논리적 조사 방법에 의해 가스로 공정을 동정하였고[5], Xu와 Zailu는 이 방법의 수정으로 더 좋은 결과를 얻는 방법과 결정 테이블에 기초한 자기 학습 알고리즘을 제안하였다. 이 알고리즘은 필요한 컴퓨터 용량 및 계산시간 때문에 고계다변수 시스템의 적용에 문제점을 발생시켰다[6,7]. 퍼지 모델링에서 퍼지 규칙의 생성은 입력 공간의 분할 및 소속 함수에 의해 결정된다[8, 9]. 또한 고차원의 비선형 공정을 모델링하는 것은 규칙수의 한계를 갖고 있다[10]. 이와 같이, 비선형 공정에서 퍼지 모델링하는 것은 많은 시행착오를 거쳐 진행된다. 전반부 및 후반부 동정에서 전체 입력 공간을 지역 공간으로 퍼지 분할하고 각 지역 공간을 표현하는 것은 많은 어려움이 있다.

본 논문에서는 비선형 공정에 대해 퍼지 모델을 동정하기 위해 분산 형태의 공간 분할 및 퍼지 추론 방법에 의한 비퍼지 추론 시스템의 입출력 특성을 분석한다. 규칙은 HCM 클러스터링 알고리즘에 의해 입력 공간을 분산 형태로 분할하고 각각의 분할된 공간이 하나의 규칙을 갖도록 형성한다. 전반부 파라미터의 동정에는 HCM 클러스터링 알고리즘[11]에 의한 소속 행렬로 결정된다. 후반부 동정에서 퍼지 추론 방법은 간략추론, 선형추론, 2차식 추론 및 변형된 2차식 추론에 의해 수행되며, 표준 최소자승법을 사용하여 후반부 파라미터를 동정한다. 비선형 공정으로 적용하기 위해 Box와 Jenkins가 사용한 가스로 공정 데이터[12]를 모델링함으로써 입출력 공간 특성 및 성능을 분석한다.

본 연구는 서론에 이어 제2장에서는 비퍼지 추론 시스템의 전반부 및 후반부 동정에 대해 다루며, 제3장에서는 비선형 공정으로 적용하고, 마지막으로 제4장에서는 결론을 다룬다.

2. 비퍼지 추론 시스템

퍼지 모델링은 if-then 형식으로 비선형 공정을 기술하며, 구체적으로 입출력 데이터의 상호관계에 의해 설정된 입출력 변수로부터 확립된다. 퍼지모델의 동정은 전반부와 후반부의 동정으로 나누어진다. 퍼지 규칙의 전반부 동정은 전반부 입력 변수의 선택과 입력 공간의 퍼지 분할 결정, 그리고 입력공간의 파라미터 결정이 필요하다. 후반부 동정은 후반부 변수의 선택과 후반부 변수의 파라미터를 결정하는 것이다.

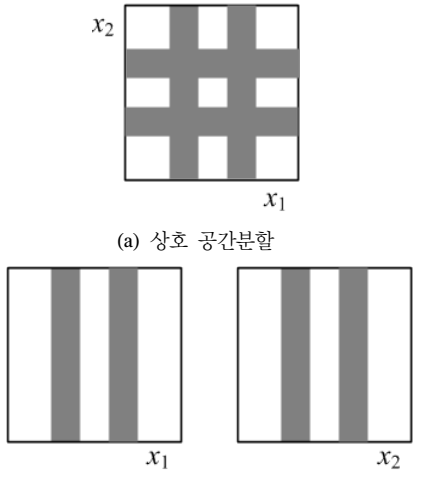
2.1 전반부 동정

퍼지 모델 동정에서 입력 공간의 분할은 일반적으로 그림 1과 같이 격자 형태로 행해지며, 각 분할된 공간은 퍼지 규칙을 의미한다. 특히, 그림 1(a)와 같이 입력 공간은 상호 관계에 의해 분할하는 형태를 많이 사용하고 있으며, 그림 1(a)는 입력 공간이 2차원 공간일 때 각 입력에 대해 3개로 분할된 지역 공간을 보여준다. 분할된 지역 공간은 퍼지 규칙을 형성하고 지역 공간의 수는 퍼지 규칙 수가 된다. 각각의 입력 공간에 대해 공간 분할 수가 같은 경우 상호 공간 분할에 의한 퍼지 규칙의 수는 m^k가 된다. 여기서 k는 차원의 수이고, m은 각 입력 공간의 분할 수이다. 이러한 상호 공간 분할 방식은 차원이 증가할수록 규칙 수가 지수적으로 증가하는 단점이 있다.

개별 공간 분할 방식은 그림 1(b)와 같이 각 입력 차원의 개별적인 공간 분할에 의해 각각의 분할된 공간이 규칙이 된다. 각각의 입력 공간에 대해 공간 분할 수가 같은 경우 개별 공간 분할에 의한 퍼지 규칙의 수는 $m*k$가 된다.

분산 형태의 입력 공간 분할 방식은 그림 2에서와 같이 HCM 클러스터링 알고리즘에 의해 수행되고 클러스터의 수만큼 입력 공간이 분할된다. 이러한 분할 방법은 클러스터의 수만큼 입력 공간이 분할되고, 각 분할된 지역 공간은 규칙 수가 된다. 따라서 분산 형태의 입력 공간 분할 방식에서 규칙의 수는 클러스터의 수가 된다.

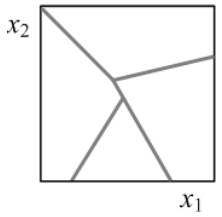
HCM 클러스터링 알고리즘에 의한 입력 공간 분할 방식은 소속 행렬에 의해 공간 분할이 결정되기 때문에 이치논리에 의한 규칙이 생성되고, 그 경계면은 명확히 구분된다.



(a) 상호 공간분할

(b) 개별 공간 분할

[그림 1] 2차원 입력 공간의 격자 분할
[Fig. 1] Grid partition of 2-D input space



[그림 2] 2차원 입력 공간의 분산 분할
[Fig. 2] Scatter partition of 2-D input space

본 논문에서는 데이터들간의 거리를 기준으로 하여 근접한 정도를 측정하고, 이를 바탕으로 데이터를 분류하는 HCM 클러스터링 알고리즘[11]을 사용한다. 클러스터링 알고리즘이란 데이터 내부의 비슷한 패턴, 속성, 형태 등의 기준을 통해 데이터를 분류하여 내부의 구조를 찾아내는 것이다. HCM 방법은 n개의 데이터를 c개의 그룹으로 분류하고 데이터의 거리가 최소인 각 그룹의 중심을 찾는다. 또한 클러스터의 소속을 "0", "1"로 나타내는 이치논리를 사용한다. HCM 클러스터링 알고리즘의 수행과정은 다음과 같다.

[단계 1] 클러스터의 개수 ($2 \le c < n$)를 결정하고, 소속행렬 $\mathbf{U}$를 초기화한다.

$$\mathbf{M}_c = \left\{ \mathbf{U} \mid u_{ik} \in \{0,1\}, \sum_{i=1}^{c} u_{ik} = 1,\ 0 < \sum_{k=1}^{n} u_{ik} < n \right\} \quad (1)$$

여기서, $u_{ik}(i = 1,2,...,c;\ k = 1,2,...,n)$ 는 소속행렬의 파라미터이다.

[단계 2] 각각의 클러스터에 대한 중심벡터 $\mathbf{v}_i$를 구한다.

$$\mathbf{v}_i^{(r)} = \{v_{i1}, v_{i2}, ..., v_{ij}\},\ v_{ij} = \frac{\sum_{k=1}^{n} u_{ik} \cdot\ x_{kj}}{\sum_{k=1}^{n} u_{ik}} \quad (2)$$

[단계 3] 각각의 클러스터 중심과 데이터와의 거리를 계산하여 새로운 소속행렬 $\mathbf{U}^{(r)}$을 생성한다.

$$d_{ik} = d(\mathbf{x}_k - \mathbf{v}_i) = \| \mathbf{x}_k - \mathbf{v}_i \| = \left[\sum_{j=1}^{m} (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (3)$$

$$u_{ik}^{(r+1)} = \begin{cases} 1 & d_{ik}^{(r)} = \min\{d_{jk}^{(r)}\} \text{ for each } j \neq i \\ 0 & otherwise \end{cases} \quad (4)$$

여기서, d_{ik}는 k번째 데이터 표본 x_k와 i번째 클러스터 중심 $\mathbf{v}_i$의 기학학적 거리이다.

[단계 4] 만일 식 (5)를 만족한다면 종료하고, 그렇지 않으면 $r = r+1$로 놓고 [단계 2]로 간다.

$$\| \mathbf{U}^{(r+1)} - \mathbf{U}^{(r)} \| \le \varepsilon(\text{tolerance level}) \quad (5)$$

2.2 후반부 동정

퍼지 모델의 후반부 동정도 전반부와 마찬가지로 구조동정과 파라미터 동정으로 나눠진다. 후반부 구조로는 퍼지추론에 의해 구별되는 네 가지 구조에 대해 다룬다. 또한, 최대 피벗팅 알고리즘을 가지는 가우스 소거법에 의한 표준 최소자승법을 이용하여 후반부 파라미터를 동정한다.

2.2.1 간략 추론(구조 1)

후반부가 단일 상수항만을 가지는 것으로, 이와 같은 추론법을 간략 추론법이라 한다. 제안된 모델은 아래와 같은 형태를 가지는 구현규칙들로 구성된다.

$$R^j : If\ \mathbf{x}_k\ is\ H_j\ then\ y_j = a_{j0} \quad (6)$$

여기서 R^j는 $j(j=1,\cdots,n)$번째 규칙, $\mathbf{x}_k(k=1,\cdots,d)$는 입력변수, H_j는 HCM 클러스터링 알고리즘에 의한 j번째 규칙의 소속 정도, a_{j0}는 상수이다.

각 규칙의 전반부 적합도 w_{ji} 는 HCM 클러스터링 알고리즘에 의한 소속 행렬에 의해 얻어지며 다음과 같다.

$$w_{ji} = u_{ji} \quad (7)$$

따라서, 추론된 값 y^*는 가중 평균에 의해 다음과 같다.

$$y^* = \frac{\sum_{j=1}^{n} w_{ji}\ y_j}{\sum_{j=1}^{n} w_{ji}} = \sum_{j=1}^{n} w_{ji}\ a_{j0} \quad (8)$$

여기서, 각 규칙의 적합도의 합은 1이다.

후반부 파라미터 동정에서 전반부 입력변수 및 파라미터가 주어지면, 성능지수를 최소화하는 최적 후반부 파라미터를 결정할 수 있다.

후반부의 파라미터는 a_{j0}로써 입출력 데이터가 주어졌

을 때 최소자승법에 의해 구해진다. 최소자승법에 의한 후반부 파라미터의 동정은 식(9)에 의해 구해진다.

$$\hat{a} = (X^T X)^{-1} X^T Y \tag{9}$$

$$E = [\epsilon_1, \cdots, \epsilon_m]^T, \quad a^T = [a_{10}, \cdots, a_{n0}],$$
$$X = [x_1, x_2, \cdots, x_m]^T, \quad x_i^T = [w_{1i}, \cdots, w_{ni}],$$
$$Y = [y_1, y_2, \cdots, y_m]^T. \tag{10}$$

2.2.2 선형 추론(구조 2)

후반부가 일차 선형식으로 표현된 것으로 선형 추론법이라 한다. 이 모델은 다음 식의 형태를 가지는 구현규칙들로 구성된다.

$$R^j : If \ \mathbf{x}_k \ is \ H_j \ then \ y_j = a_{j0} + \sum_{k=1}^{d} a_{jk} x_k \tag{11}$$

여기서 a_{jk}는 후반부의 파라미터이다. 추론된 값 y^*는 다음과 같다.

$$y^* = \frac{\sum_{j=1}^{n} w_{ji} \ y_j}{\sum_{j=1}^{n} w_{ji}} = \sum_{j=1}^{n} w_{ji} \ (a_{j0} + a_{j1} x_{1i} + \cdots + a_{jk} x_{ki}) \tag{12}$$

여기서, w_{ji}는 식(7)과 같다. 최소자승법에 의한 후반부 파라미터의 동정은 간략 추론과 같은 방법으로 구해진다.

2.2.3 2차식 추론(구조 3)

2차식 추론법은 후반부 구조가 2차식 함수의 다항식 형태로 표현되고 다음과 같은 구현 규칙으로 구성된다.

$$R^j : If \ \mathbf{x}_k \ is \ H_j \ then \ y_j = a_{j0} + \sum_{k=1}^{d} a_{jk} x_k + \sum_{k=1}^{d} a_{j,d+k} x_k^2 + \sum_{k=1}^{d} \sum_{l=k+1}^{d} A_{jz} x_k x_l \tag{13}$$

여기서, z는 입력 변수들의 조합의 수이다.

모델의 추론된 값 y^*는 마찬가지로 다음과 같이 구하여진다.

$$y^* = \sum_{j=1}^{n} w_{ji} \ (a_{j0} + \sum_{k=1}^{d} a_{jk} x_{ki} + \sum_{k=1}^{d} a_{j,d+k} x_{ki}^2 + \sum_{k=1}^{d} \sum_{l=k+1}^{d} A_{jz} x_{ki} x_{li}) \tag{14}$$

2.2.4 변형된 2차식 추론(구조 4)

변형된 2차식 추론법은 입력 변수의 2차항이 생략된 2차식 추론법의 변형된 형태로써 다음과 같은 규칙으로 구성된다.

$$R^j : If \ \mathbf{x}_k \ is \ H_j \ then \ y_j = a_{j0} + \sum_{k=1}^{d} a_{jk} x_k + \sum_{k=1}^{d} \sum_{l=k+1}^{d} A_{jz} x_k x_l \tag{15}$$

모델의 추론된 값 y^*는 마찬가지로 다음과 같이 구하여진다.

$$y^* = \sum_{j=1}^{n} w_{ji} \ (a_{j0} + \sum_{k=1}^{d} a_{jk} x_{ki} + \sum_{k=1}^{d} \sum_{l=k+1}^{d} A_{jz} x_{ki} x_{li}) \tag{16}$$

3. 비선형 공정으로의 적용

본 논문에서는 제안된 퍼지 모델의 평가를 위해 비선형 공정에 대한 성능 평가의 척도로 사용되고 있는 가스로 공정[12]을 사용한다. 모델의 평가 기준인 성능지수는 수치 데이터인 가스로 공정에 대해서 MSE(Mean Squared Error)를 이용한다.

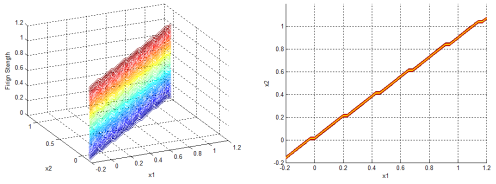
$$PI = \frac{1}{m} \sum_{i=1}^{m} (y_i - y_i^*)^2 \tag{17}$$

Box와 Jenkins가 사용한 가스로 시계열 데이터를 이용하여, 입출력 데이터인 가스 흐름률과 연소된 이산화탄소 농도의 가스로 공정을 퍼지 모델링한다. 입력이 가스 흐름률이고 출력이 이산화탄소 농도인 1입력 1출력의 가스로 시계열 입출력 데이터 296쌍을 모의실험을 위해 입력으로 [$u(t-3)$, $u(t-2)$, $u(t-1)$, $y(t-3)$, $y(t-2)$ $y(t-1)$]을, 출력으로 $y(t)$를 구성하여 사용한다. 또한 데이터 집합은 학습과 테스트 데이터로 나누어 추론에 의한 모델링을 수행한다.

먼저, 대표적으로 사용되는 $u(t-3)$과 $y(t-1)$을 입력으로 사용하여 2입력 1출력 시스템을 구성하여 2차원 시스템

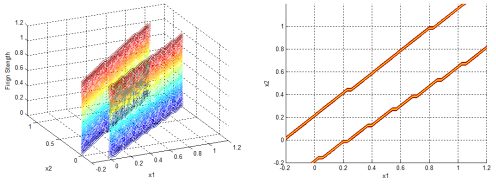
을 모델링한다.

그림 3부터 그림 11까지는 2차원 입력 공간에서 HCM 클러스터링 알고리즘을 이용하여 클러스터의 수를 2개에서 10개까지 하나씩 증가함으로써 생성된 소속 행렬 및 분산 분할된 입력 공간을 보여준다. 소속 행렬에 의한 경계선은 이치논리에 의한 경계를 나타낸다. 비선형 공정 데이터의 분포에 따른 각 클러스터의 중심을 기준으로 대각선 형태로 공간이 분할되는 것을 알 수 있으며 클러스터의 수가 증가할수록 데이터가 밀집된 부분에서 세밀하게 공간이 분할되는 것을 알 수 있다.



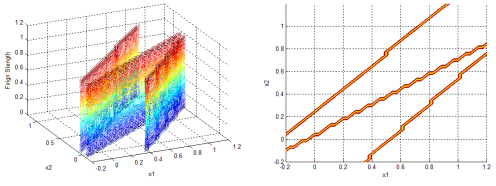
(a) 소속 행렬 (b) 공간 분할

[그림 3] 소속 행렬 및 분산 분할(2 클러스터)
[Fig. 3] Membership matrix and scatter partition (2 clusters)



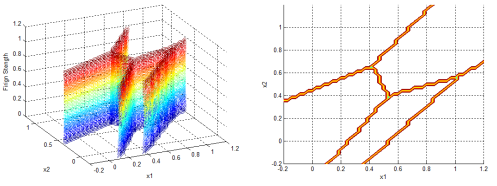
(a) 소속 행렬 (b) 공간 분할

[그림 4] 소속 행렬 및 분산 분할(3 클러스터)
[Fig. 4] Membership matrix and scatter partition (3 clusters)



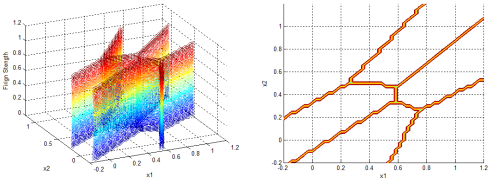
(a) 소속 행렬 (b) 공간 분할

[그림 5] 소속 행렬 및 분산 분할(4 클러스터)
[Fig. 5] Membership matrix and scatter partition (4 clusters)



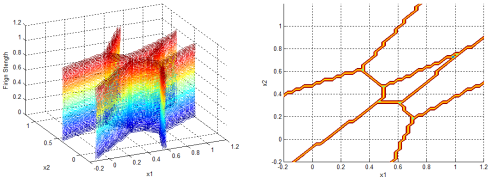
(a) 소속 행렬 (b) 공간 분할

[그림 6] 소속 행렬 및 분산 분할 (5 클러스터)
[Fig. 6] Membership matrix and scatter partition (5 clusters)



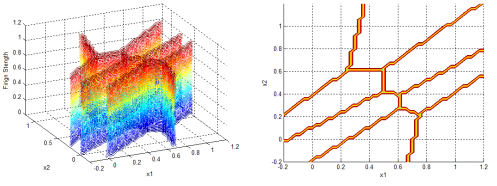
(a) 소속 행렬 (b) 공간 분할

[그림 7] 소속 행렬 및 분산 분할 (6 클러스터)
[Fig. 7] Membership matrix and scatter partition (6 clusters)



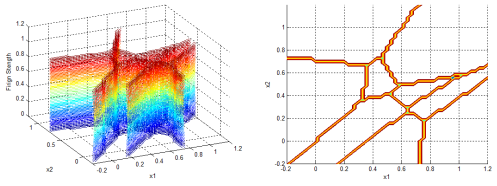
(a) 소속 행렬 (b) 공간 분할

[그림 8] 소속 행렬 및 분산 분할 (7 클러스터)
[Fig. 8] Membership matrix and scatter partition (7 clusters)



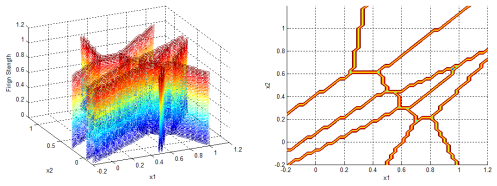
(a) 소속 행렬 (b) 공간 분할

[그림 9] 소속 행렬 및 분산 분할 (8 클러스터)
[Fig. 9] Membership matrix and scatter partition (8 clusters)



(a) 소속 행렬 (b) 공간 분할

[그림 10] 소속 행렬 및 분산 분할 (9 클러스터)
[Fig. 10] Membership matrix and scatter partition (9 clusters)



(a) 소속 행렬 (b) 공간 분할

[그림 11] 소속 행렬 및 분산 분할 (10 클러스터)
[Fig. 11] Membership matrix and scatter partition (10 clusters)

표 1은 HCM 클러스터 알고리즘에 의한 클러스터의 수 및 추론 방법에 의한 학습 데이터와 테스트 데이터에 대한 성능 지수를 보여준다. 여기서, PI는 학습 데이터에 대한 성능 지수를, E_PI는 테스트 데이터에 대한 성능 지수를 보여준다.

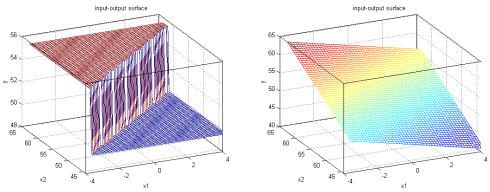
[표 1] 2입력 시스템에 대한 성능지수
[Table 1] PI for 2 input system

클러스터 수	Type	PI	E_PI
2	구조 1	3.459	4.104
	구조 2	0.022	0.335
	구조 3	0.022	0.337
	구조 4	0.022	0.347
3	구조 1	1.671	2.161
	구조 2	0.021	0.346
	구조 3	0.020	0.339
	구조 4	0.021	0.354
4	구조 1	1.054	1.706
	구조 2	0.020	0.347
	구조 3	0.018	0.337
	구조 4	0.019	0.350
5	구조 1	1.026	1.747
	구조 2	0.018	0.347
	구조 3	0.016	0.324
	구조 4	0.017	0.346
6	구조 1	0.898	1.857
	구조 2	0.019	0.341
	구조 3	0.015	0.352
	구조 4	0.016	0.344
7	구조 1	0.825	1.957
	구조 2	0.016	0.327
	구조 3	0.014	0.309
	구조 4	0.015	0.325
8	구조 1	0.713	1.460
	구조 2	0.018	0.302
	구조 3	0.013	0.300
	구조 4	0.015	0.302
9	구조 1	0.803	1.673
	구조 2	0.017	0.330
	구조 3	0.013	0.374
	구조 4	0.016	0.329
10	구조 1	0.623	1.379
	구조 2	0.016	0.300
	구조 3	0.010	0.338
	구조 4	0.014	0.311

표 1로부터 간략 추론 보다는 선형 추론, 2차식 추론, 변형된 2차식 추론이 더 좋은 성능을 보여준다. 또한, 클러스터의 수가 증가할수록, 즉 규칙 수가 많을수록 일반적으로 더 좋은 성능을 보여주는 것을 알 수 있다. HCM 클러스터링 알고리즘에 의한 비퍼지 추론 모델은 입력 공간을 분산 형태로 8개의 규칙을 가지고 2차식 추론을 이용한 경우가 가장 좋은 성능을 보여준다. 이때의 성능 지수는 PI는 0.013이고 E_PI는 0.300이다.

비퍼지 추론 모델의 경우 학습데이터에 의한 근사화 능력과 테스트 데이터에 의한 일반화 능력이 대체로 균형을 잘 잡는 것을 알 수 있으며, 다소 근사화 능력에서 좋은 성능을 갖는 것을 알 수 있다.

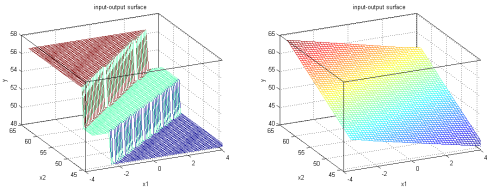
그림 12부터와 그림 20까지는 표 1의 방법을 이용하여 입력 공간의 분할 수 및 추론 방법에 따른 입출력 특성 평면을 보여준다. 간략 추론의 경우 경계면이 확연하게 구분되는 것을 볼 수 있으며, 각 분할된 지역 공간은 간략 추론에 의한 상수항의 평면 특성 보여준다. 선형 추론, 2차식 추론, 변형된 2차식 추론의 경우 각 지역 공간에서 선형, 2차식, 변형된 2차식의 입출력 특성 보여준다. 또한, 각각의 지역 공간은 지역 공간이 서로 중첩되지 않는 독립된 입출력 특성을 보여준다.



(a) Simplified (b) Linear

(c) Quadratic (d) Modified quadratic

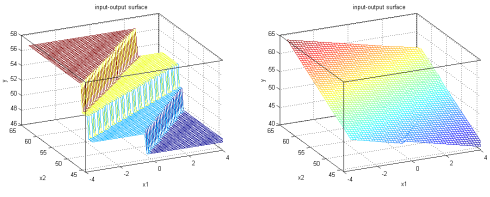
[그림 12] 입출력 공간 평면 (2 클러스터)
[Fig. 12] Input-output space (2 clusters)



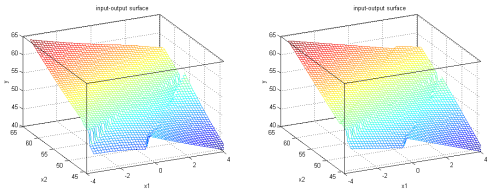
(a) Simplified (b) Linear

(c) Quadratic (d) Modified quadratic

[그림 13] 입출력 공간 평면 (3 클러스터)
[Fig. 13] Input-output space (3 clusters)

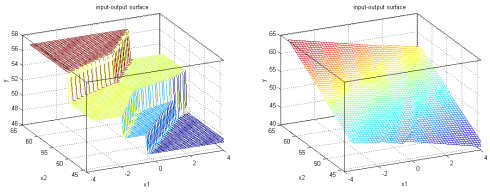


(a) Simplified (b) Linear

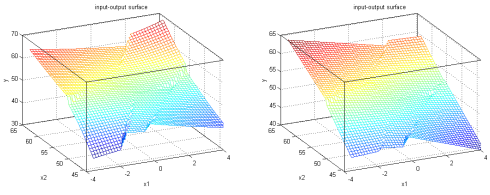


(c) Quadratic (d) Modified quadratic

[그림 14] 입출력 공간 평면 (4 클러스터)
[Fig. 14] Input-output space (4 clusters)

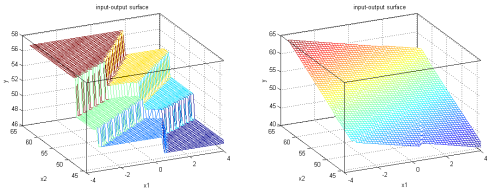


(a) Simplified (b) Linear

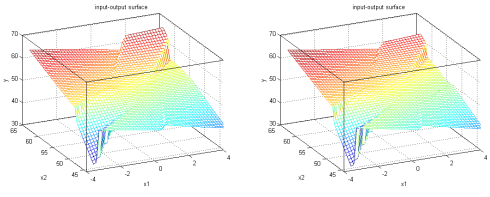


(c) Quadratic (d) Modified quadratic

[그림 15] 입출력 공간 평면 (5 클러스터)
[Fig. 15] Input-output space (5 clusters)

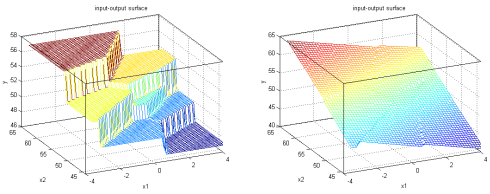


(a) Simplified (b) Linear

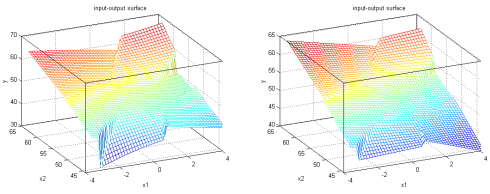


(c) Quadratic (d) Modified quadratic

[그림 16] 입출력 공간 평면 (6 클러스터)
[Fig. 16] Input-output space (6 clusters)

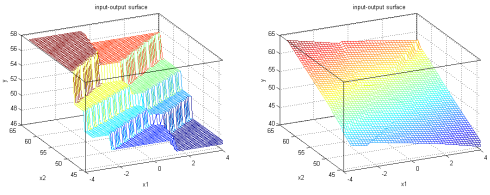


(a) Simplified (b) Linear

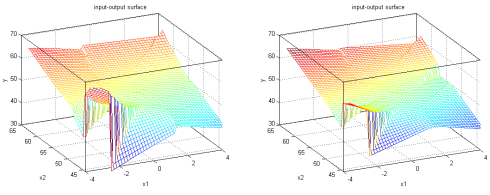


(c) Quadratic (d) Modified quadratic

[그림 17] 입출력 공간 평면 (7 클러스터)
[Fig. 17] Input-output space (7 clusters)

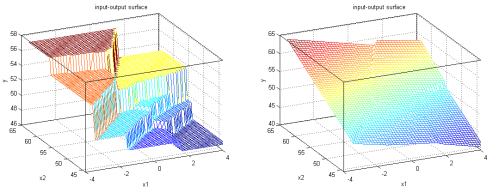


(a) Simplified (b) Linear

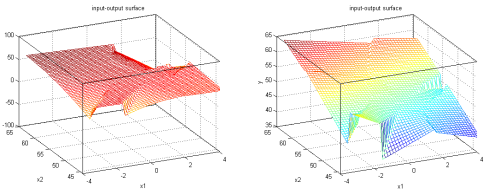


(c) Quadratic (d) Modified quadratic

[그림 18] 입출력 공간 평면 (8 클러스터)
[Fig. 18] Input-output space (8 clusters)

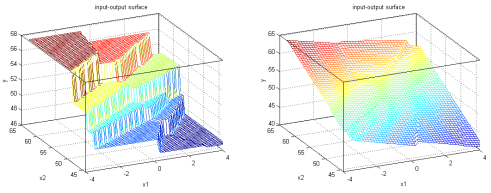


(a) Simplified (b) Linear

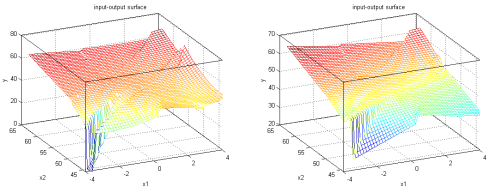


(c) Quadratic (d) Modified quadratic

[그림 19] 입출력 공간 평면 (9 클러스터)
[Fig. 19] Input-output space (9 clusters)

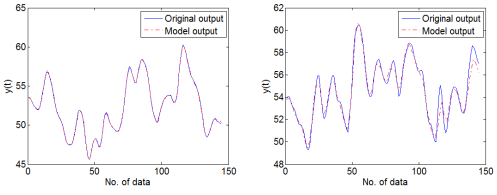


(a) Simplified (b) Linear



(c) Quadratic (d) Modified quadratic

[그림 20] 입출력 공간 평면 (10 클러스터)
[Fig. 20] Input-output space (10 clusters)



(a) training data (b) testing data

[그림 21] 모델 출력 (8 클러스터, 2차식)
[Fig. 21] Model output (8 clusters, Quadratic)

그림 21은 표 1에서 선택된 모델인 8개의 규칙수를 가지고 후반부 구조가 2차식 추론인 모델에 대하여 학습 데이터와 테스트 데이터에 대한 원 출력과 모델출력을 보여준다. 그림에서 보는 바와 같이 원 출력과 유사한 출력을 갖는 것을 알 수 있다.

[표 2] 6입력 시스템에 대한 성능지수
[Table 2] PI for 6 input system

클러스터 수	Type	PI	E_PI
2	구조 1	3.5586	4.815
	구조 2	0.0146	0.185
	구조 3	0.0090	0.178
	구조 4	0.0095	0.242
3	구조 1	1.9112	2.902
	구조 2	0.0137	0.203
	구조 3	0.0047	1.547
	구조 4	0.0064	0.366
4	구조 1	2.1242	3.058
	구조 2	0.0119	0.243
	구조 3	0.0027	1.494
	구조 4	0.0042	5.820
5	구조 1	1.8754	3.055
	구조 2	0.0111	0.192
	구조 3	0.0011	3.069
	구조 4	0.0025	5202.400
6	구조 1	1.3705	2.713
	구조 2	0.0093	0.226
	구조 3	0.0003	32.801
	구조 4	0.0018	142.372
7	구조 1	1.4659	2.721
	구조 2	0.0093	0.260
	구조 3	0.0004	7.327
	구조 4	0.0011	0.923
8	구조 1	1.4655	2.632
	구조 2	0.0082	0.226
	구조 3	2.8251E-17	23.430
	구조 4	0.0007	5.415
9	구조 1	1.2586	2.521
	구조 2	0.0089	0.199
	구조 3	2.83E-17	52.315
	구조 4	0.0003	6.059
10	구조 1	1.2740	2.532
	구조 2	0.0087	0.185
	구조 3	4.5155E-18	24.049
	구조 4	0.000099	48.026

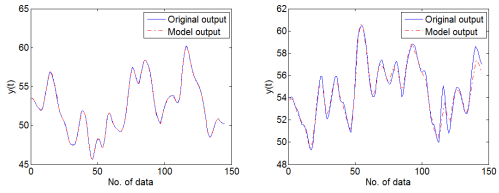
다음으로, 고차원의 비선형 시스템을 모델링하기 위하여 전체 입력을 사용하여 6입력 1출력 시스템을 구성하여 6입력 시스템을 모델링한다.

앞에서와 같은 방법으로 HCM 클러스터링 알고리즘을 이용하여 입력 공간을 분산 형태로 분할하고 클러스터의 수와 추론 방법을 이용하였으며, 학습 데이터와 테스트 데이터에 대한 성능 지수는 표 2에서 보여준다.

표 2로부터 간략 추론인 경우 근사화 능력과 일반화 능력이 좋지 않은 결과를 가져왔다. 선형 추론의 경우는 근사화 능력과 일반화 능력이 균형을 이루면서 성능이 개선된 것을 알 수 있다. 2차식 추론, 변형된 2차식 추론의 경우 클러스터의 수가 증가할수록 매우 불안정한 결과를 보여주며 근사화 능력은 매우 좋으나 일반화 능력이 매우 안 좋은 결과를 보여준다.

HCM 클러스터링 알고리즘에 의한 비퍼지 추론 모델은 입력 공간을 분산 형태로 10개의 규칙을 가지고 선형 추론을 이용한 경우가 가장 좋은 성능을 보여준다. 이 때의 성능지수는 PI는 0.0087이고 E_PI는 0.185이다.

그림 22는 표 2에서 10개의 규칙 수를 가지고 후반부 구조가 선형 추론인 모델에 대하여 학습 데이터와 테스트 데이터에 대한 원 출력과 모델출력을 보여준다. 그림에서 보는 바와 같이 원 출력과 유사한 출력을 갖는 것을 알 수 있다.



(a) training data (b) testing data

[그림 22] 모델 출력 (10 클러스터, 선형)
[Fig. 22] Model output (8 clusters, linear)

4. 결론

본 논문에서는 HCM 클러스터 알고리즘에 의한 규칙을 형성함으로써 비선형 공정에 대해 비퍼지 추론 시스템을 구축하여 특성을 분석하였다. 제안된 비퍼지 추론 시스템은 입력 공간을 분산 형태로 분할하고 분할된 지역 공간은 서로 중첩되지 않는 독립된 공간으로 규칙을 형성하였으며, 각각의 지역 공간은 간략 추론, 선형 추론, 2차식 추론 및 변형된 2차식 추론을 이용하여 표현하였다.

저차원의 비선형 시스템의 동정은 일반적으로 근사화 능력과 일반화 능력이 대체로 균형을 잘 잡는 것을 알 수 있으며, 다소 근사화 능력에서 좋은 성능을 갖는 것을 알 수 있다. 고차원의 비선형 시스템의 동정은 매우 적은 수의 규칙으로 모델링 할 수 있었으며, 선형 추론인 경우

매우 좋은 결과를 얻을 수 있었다. 하지만, 2차식 추론과 변형된 2차식 추론은 매우 불안한 성능을 보여줌으로써 사용함에 있어 주의가 필요하다.

References

[1] R.M. Tong, "Synthesis of fuzzy models for industrial processes", Int. J. Gen. Syst., Vol. 4, pp.143-162, 1978.

[2] W. Pedrycz, "An identification algorithm in fuzzy relational system", Fuzzy Sets Syst., Vol. 13, pp.153-167, 1984.

[3] W. Pedrycz, "Numerical and application aspects of fuzzy relational equations", Fuzzy Sets Syst., Vol. 11, pp.1-18, 1983.

[4] E. Czogola and W. Pedrycz, "On identification in fuzzy systems and its applications in control problems", Fuzzy Sets Syst., Vol. 6, pp.73-83, 1981.

[5] R. M. Tong, "The evaluation of fuzzy models derived from experimental data", Fuzzy Sets Syst., Vol. 13, pp.1-12, 1980.

[6] C. W. Xu, "Fuzzy system identification", IEEE Proceeding Vol. 126, No. 4, pp.146-150, 1989.

[7] C. W. Xu and Y. Zailu, "Fuzzy model identification self-learning for dynamic system", IEEE Trans. on Syst. Man, Cybern., Vol. SMC-17, No. 4, pp.683-689, 1987.

[8] Keon-Jun Park, Dong-Yoon Lee, "Characteristics of Fuzzy Inference Systems by Means of Partition of Input Spaces in Nonlinear Process", The Korea Contents Association, Vol. 11, No. 3, pp. 48-55, 2011. 3.

[9] Keon-Jun Park, Dong-Yoon Lee, "Characteristics of Input-Output Spaces of Fuzzy Inference Systems by Means of Membership Functions and Performance Analyses", The Korea Contents Association, Vol. 11, No. 4, pp. 74-82, 2011. 4.

[10] Keon-Jun Park, Dong-Yoon Lee, “Nonlinear Characteristics of Fuzzy Inference Systems by Means of Individual Input Space”, The Korea Academia-Industrial Cooperation Society, Vol. 12, No. 11, pp. 5164-5171, 2011. 11.

[11] P. R. Krishnaiah and L. N. Kanal, editors. Classification, pattern recognition, and reduction of dimensionality, volume 2 of Handbook of Statistics. North-Holland, Amsterdam, 1982.

[12] Box and Jenkins, "Time Series Analysis, Forcasting and Control", Holden Day, SanFrancisco, CA.

박 건 준(Keon-Jun Park) **[정회원]**

- 2005년 2월 : 원광대학교 제어계측공학과 (공학석사)
- 2010년 8월 : 수원대학교 전기공학과 (공학박사)
- 2010년 10월 ~ 2012년 8월 : 원광대학교 공과대학 2단계BK21 Post-Doc
- 2012년 9월 ~ 현재 : 원광대학교 공업기술개발연구소 리서치펠로우 연구교수

<관심분야>
컴퓨터 및 인공지능, 퍼지추론 시스템, 신경망, 유전자 알고리즘 및 최적화이론, 지능시스템 및 제어

이 동 윤(Dong-Yoon Lee) **[정회원]**

- 1990년 2월 : 연세대학교 전기공학과 (공학석사)
- 2001년 2월 : 연세대학교 전기전자공학과 (공학박사)
- 2002년 3월 ~ 현재 : 중부대학교 전기전자공학과 교수

<관심분야>
시큐리티시스템, 퍼지추론 시스템, 인공지능