

클러스터링 기반 뉴로-퍼지 모델링 학습

Neuro-Fuzzy Modeling Learning method based on Clustering

김승석*, 광근창*, 이대종*, 김성수*, 유정웅*, 김주식**, 김용태***

*충북대학교 전기전자컴퓨터공학부 컴퓨터 정보통신 연구소, **특허청, 한경대학교 정보제어과

S.S. Kim*, K.C. Kwak*, D.J. Lee*, S.S. Kim*, J.W. Ryu*, J.S. Kim**, Y.T. Kim**

*School of Electrical Electronic and Computer Eng, Chungbuk National University, **KIPO, ***Dept. Information Control Engineering, HanKyong National Univeristy***

E-mail : powerkimss@chungbuk.ac.kr

요 약

본 논문에서는 클러스터링과 뉴로-퍼지 모델링을 동시에 실시하는 학습 기법을 제안하였다. 클러스터링을 이용하여 뉴로-퍼지 모델링을 실시하는 일반적인 경우, 클러스터링 학습을 실시한 후 학습된 파라미터를 뉴로-퍼지 모델의 초기 파라미터로 설정하고 모델을 다시 학습하는 방법을 취한다. 즉 클러스터링에서 클러스터의 수를 구하고 파라미터를 최적화함으로써 초기 구조동정과 파라미터 동정을 실시하며 이를 다시 뉴로-퍼지 모델에서 세부적인 파라미터 동정을 실시하는 것이다. 또한 모델에서의 학습은 출력데이터의 오차를 이용한 오차미분기반 학습으로 전제부 소속함수 파라미터를 수정하는 방법을 이용한다. 이 경우 클러스터링의 영향과 모델의 영향이 각각 별개로 고려될 수 있다. 따라서 본 논문에서는 클러스터링을 전제부 소속함수로 부여하고 클러스터링의 학습에 뉴로-퍼지 모델을 이용하면서 또한 모델의 학습에 클러스터링을 직접 적용하는 클러스터링 기반 뉴로-퍼지 모델링을 제안하였으며 이 경우 클러스터링의 학습과 모델의 학습이 동시에 이루어지며 뉴로-퍼지 모델에서 클러스터링의 효과를 직접적으로 확인할 수 있다. 제안된 방법의 유용성을 시뮬레이션을 통하여 보이고자 한다.

1. 서론

주어진 데이터를 이용하여 뉴로-퍼지 모델링을 실시하는 경우 입력데이터 공간의 적절한 분할을 통하여 소속함수를 부여하고 적절한 규칙을 부여하는 것은 시스템의 성능에 직접적으로 영향을 주고 있다[1-2]. 전체 입력공간을 격자분할하여 소속함수를 부여하는 방법은 모든 데이터 공간을 모델의 학습에 이용할 수 있는 장점을 가지고 있지만 입력의 차원이나 소속함수의 수가 증가하는 경우 모델의 규칙이 지수 함수적으로 증가하는 문제점을 가지고 있다[1]. 이러한 문제점을 보완하기 위하여 입력공간 상에서 데이터의 분포를 이용하여 클러스터를 형성하고 이를 뉴로-퍼지 모델의 전제부 소속함수 생성과 규칙을 부여하는 방법이 활발하게 연구되어지고 있다

[3-4][6-8]. 클러스터링에 의한 뉴로-퍼지 모델의 생성 및 학습은 먼저 클러스터링 학습 후에 선택된 클러스터의 수로 모델의 구조동정을 실시하며 학습된 파라미터를 모델의 초기 파라미터로 하여 다시 모델을 학습하여 파라미터 동정을 실시하는 순차적인 과정을 거친다[1][3]. 클러스터링 후의 클러스터링 파라미터에 의한 데이터의 형태 표현은 뉴로-퍼지 모델의 학습과정을 통하여 무의미하게 되는 경우가 발생한다[3]. 또한 클러스터링 학습과 뉴로-퍼지 모델의 학습 방법이 서로 다른 형태를 취하므로 별개의 학습 알고리즘으로 순차적인 모델링만이 가능하였다[3-4]. 하나의 알고리즘의 완료되면 그 구조나 파라미터를 가지고 다음 학습 알고리즘의 초기치를 이용하는 기법의 경우 종종 이전 알고리즘에 추정된 결과가 이후 알고리즘에 의하여 변동함으로써 무의미하게 되

는 경우도 발생한다[1][3]. 또한 입출력 데이터가 존재하는 경우에도 일반적인 클러스터링 기법의 경우에는 단지 입력 데이터에 대하여 알고리즘을 수행하므로써 입출력에 대한 인과관계를 정확하게 표현하지 못하였다[1]. 이러한 문제점을 개선하기 위하여 제안된 방법에서는 클러스터링과 뉴로-퍼지 모델을 하나의 학습 알고리즘으로 통합하여 학습을 실시하였다. 자기 구성 클러스터링 기법을 이용하여 자율적으로 클러스터링의 구조 탐색과 파라미터 최적화를 실시하며 최적화된 파라미터를 이용하여 뉴로-퍼지 모델을 구성한다. 각 규칙에 의하여 추정된 출력을 다시 클러스터링 학습에 영향을 주도록 하였다. 입출력 인과관계가 포함된 출력 오차를 클러스터링의 유사도 측정에 포함하여 클러스터링 기법이 입력공간 뿐만 아니라 출력 공간까지 고려하는 교사학습(supervised learning)으로 확장하였다. 제안된 방법을 유용성을 시뮬레이션을 통하여 보이고자 한다.

2. 자기-구성 클러스터링

클러스터링 기법은 크게 두 가지 학습 형태로 나눌 수 있다[1]. 먼저 클러스터의 구조나 수를 알지 못하는 경우 데이터의 특성을 이용하여 클러스터의 수를 구하는 방법이 있다[5]. 또한 클러스터의 수를 알고 있는 경우 이들 파라미터를 최적화하는 기법이 있다[7]. 제안된 자기-구성 클러스터링은 주어진 데이터의 형태나 구조를 모를 경우 클러스터의 수와 이들 파라미터를 동시에 추정하고 최적화하는 알고리즘이다.

클러스터 중심과 각 데이터 간의 유사도 측정은 제안된 방법에서는 다음과 같은 함수를 이용하였다.

$$r_{ij} = \exp\left(-\frac{1}{2}(x_i - c_j)^T(\zeta \cdot \Sigma_j)^{-1}(x_i - c_j)\right) \quad (1)$$

여기서 c_j 는 j 번째 클러스터 중심이고, Σ_j 는 j 번째 클러스터 공분산이며 ζ 는 사전에 지정해주는 임계치이다.

인접 클러스터의 영향을 제거하고 클러스터의 수를 추정하는 경우 수렴 속도를 개선하기 위하여 유사도를 다음과 같이 제한하였다[6].

$$r_{ij} = \begin{cases} r_{ij}, & r_{ij} > \zeta \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

식(1)에서 ζ 를 포함하지 않은 경우, 식(2)에서 ζ 의 설정이 너무 크면 같은 클러스터의 외곽 데이터가 데이터 추정에 영향을 미치지 않으며 너무 적은 경우 ζ 가 필요가 없어지는 문제점을 개

선할 수 있다. 알고리즘의 초기 모든 데이터가 클러스터의 중심으로 설정된다. 즉 모든 데이터가 모든 클러스터의 중심이 될 수 있는 후보가 되면 알고리즘이 진행되는 동안 점차 특정한 지점들로 클러스터 후보 파라미터들이 수렴한다. 유사도를 이용한 중심의 추정 방법은 식(3)과 같이 추정한다.

$$c_j = \frac{\sum_{i=1}^N r_{ij} x_i}{\sum_{i=1}^N r_{ij}} \quad (3)$$

추정된 중심을 이용하여 Maximum Likelihood Estimation (MLE)를 이용하기 위한 확률은 식(4)와 같이 이용한다[7].

$$p_{ij} = \frac{1}{(2\pi)^{d/2} |\Sigma_j|^{1/2}} \times \exp\left\{-\frac{1}{2}(x_i - c_j)^T \Sigma_j^{-1} (x_i - c_j)\right\} \quad (4)$$

이를 이용한 분할행렬은 사전 확률 $p(x_i)$ 와 곱하여 식(5)와 같다.

$$u_{ij} = \frac{PW_{ij}}{\sum_{i=1}^N PW_{ij}} \quad (5)$$

여기서 사전확률 $p(x_i)$ 는 식(6)과 같으며 알고리즘 초기에 균등하게 설정한다.

$$p(x_i) = \frac{1}{N} \sum_{i=1}^N u_{ij} \quad (6)$$

반복적인 학습동안 알고리즘의 학습 속도 개선과 수렴보장을 위하여 변형된 Subtractive 알고리즘 개념을 도입하였다[1]. 학습이 진행되는 동안 클러스터 중심은 데이터 특정한 지점들로 수렴하는데 이들 수렴되는 클러스터의 중심들의 밀도를 대상으로 밀집한 정도가 심한 클러스터 군을 하나의 통합된 클러스터 중심으로 합병한다. 알고리즘의 진행은 현재 가장 큰 누적밀도 c_{new} 를 식(7)과 같이 선택한다.

$$c_{new} = c_j, \quad \text{if } u_{ij} = \max\left(\sum_{i=1}^N u_{ij}\right) \quad (7)$$

새로운 중심이 선택되면 인근 클러스터의 중심은 소멸함수를 통하여 제거된다. 또한 소멸함수의 범위가 넓어짐에 따라 발생할 수 있는 주변 클러스터의 영향을 줄이기 위하여 범위를 새로운 클러스터의 누적 밀도에서 0.9 이상의 누적 밀도를 가지는 클러스터로 제한하였다. 학습 알고리즘은 초기화 과정에서 모든 데이터를 클러스터 중심으로 설정하고 각종 파라미터를 초기화한다. 반복적인 학습 알고리즘은 식(1)과 식(2)를 이용

하여 누적 유사도를 구한 후 이를 이용하여 식 (3)과 같이 중심을 추정하고 추정된 중심을 이용하여 MLE를 만족하는 공분한 행렬을 구한다. 또한 밀집된 클러스터를 제거함으로써 알고리즘의 수렴을 개선하기 위하여 변형된 Subtractive 알고리즘을 수행한다. 이를 반복적으로 수행하며 이들 파라미터는 뉴로-퍼지 모델의 전제부로 직접 구성되어 전체 모델의 학습에 이용된다.

2. 통합 학습 알고리즘

본 논문에서 이용된 뉴로-퍼지 모델은 Tagaki-Sugeno-Kang (TSK) 퍼지 모델로써, 클러스터링에 의하여 모델의 전제부 소속함수가 구성되면 전방경로 학습은 최소자승법(Least Square Estimation)을 이용하여 결론부 파라미터를 추정하며 그림 1과 같다[1][3].

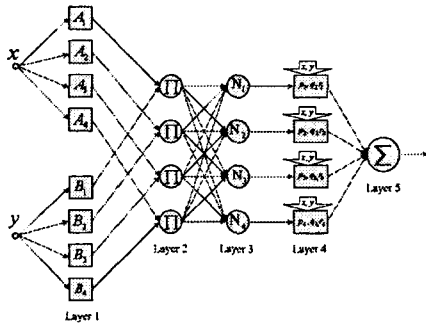


그림 1. TSK 뉴로-퍼지 모델

각 규칙에 의한 입출력 관계는 식(8)과 같이 표현할 수 있다.[1][3-4].

$$\text{RULE } i: \text{ If } x \text{ is } A_i \text{ and } y \text{ is } B_i \text{ (8)}$$

$$\text{Then } f_i = p_i x + q_i y + r_i$$

여기서 입력 x, y 에 대하여 A_i, B_i 는 i 번째 클러스터의 입력차원에 대한 전제부 소속함수이고 p_i, q_i, r_i 는 i 번째 클러스터에 연결된 i 번째 규칙에 의하여 결론을 추론하는 결론부 파라미터이다. 추정된 결론부 파라미터는 규칙에 대하여 상수항(bias) r_i 를 포함하여 각각 다른 출력을 추정한다. 제안된 방법에서는 각 규칙에 의하여 연결된 출력을 기준치와 비교하여 오차 err_i 를 발생시킨다. 각 규칙에 의하여 발생된 오차는 다시 클러스터링의 유사도 측정에 영향을 주도록 하여 입력공간에서 실시하는 클러스터링 기법을 입출력 공간을 동시에 고려하는 비교사 학습으로 확장하였다. 또한 오차에 의한 영향이 유사도 측정에 심한 영향을 미치는 것을 방지하기 위하여 가중치 t 를 식(8)과 같이 부여하였다.

$$r' = err \cdot t + (1 - t) \quad (9)$$

여기서 err 과 t 는 $[0, 1]$ 이다. 이를 이용한 가중

치 r_{ij} 는 식(9)와 같이 적용된다.

$$r_{ij} = r_{ij} \times r' \quad (10)$$

전체 모델의 출력에서, 특정 규칙에 의하여 추정된 규칙이 기준치와 비교하여 적은 오차를 발생하는 경우 해당 규칙에 연결된 클러스터의 유사도는 크게 변하지 않는 반면 오차가 크게 발생한 규칙의 출력은 유사도 측정에 영향을 크게 주어 다음의 반복학습에서 불리한 조건을 받게 된다. 클러스터링에 의한 전제부 소속도는 규칙에 의하여 결론부와 연결이 된다. 각 규칙에 의한 결론부 출력의 영향이 다시 클러스터링 수행에 영향을 주며 이를 그림 2에 나타내었다.

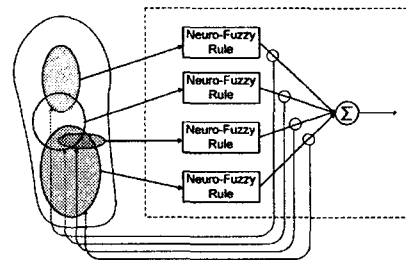


그림 2. 제안된 모델의 학습 개념

학습에는 각 규칙에 의한 개별 출력이 다시 전제부 클러스터링에 이용되며 최종 성능평가는 가중평균법에 의한 최종출력을 이용한다.

3. 시뮬레이션 및 결과

제안된 방법의 유용성을 보이기 위하여 본 논문에서는 대표적인 Bench-mark 데이터인 자동차 연료 예측 문제를 이용하였다. 입력 변수가 이산적인 값과 연속적인 값을 모두 가지며 입력의 구성은 다음과 같다.

실린더의 수(No. of cylinders)	: discrete
배기량(Displacement)	: continuous
마력(Horsepower)	: continuous
무게(Weight)	: continuous
가속(Acceleration)	: continuous
연식(Model year)	: discrete

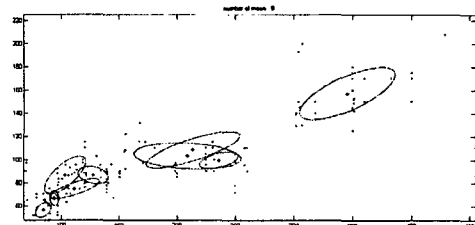


그림 3. 학습 후 클러스터링 결과

클러스터링 기법을 이용하므로 전입력차원을 모두 사용하여도 클러스터의 수에 따른 규칙만 증가하고 입력차원의 증가와는 무관하다. 입력데

이터를 이용하여 모델링을 실시하였을 때 학습된 클러스터의 분포는 그림 3과 같다. 그림 4와 그림 5에서 학습데이터에 의한 출력과 검증데이터에 의한 출력을 나타내었다.

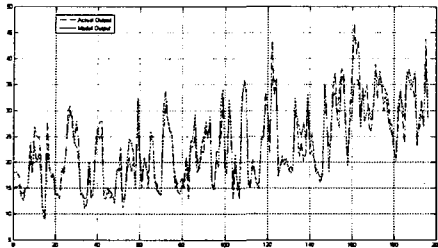


그림 4. 학습데이터에 의한 출력

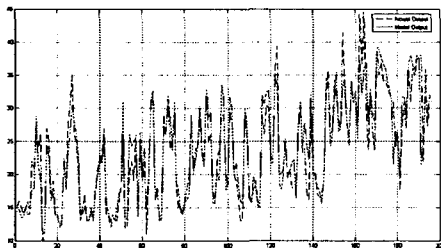


그림 5. 검증데이터에 의한 출력

표에서 이전의 연구 결과와 함께 성능을 비교하였다.

	규칙 수	학습오차	검증오차
Abonyi[9]	2	2.94	3.01
Babuska[10]	4	2.84	2.94
Proposed	9	2.86	2.91

표 1. 결과비교

제안된 방법의 모델링이 학습데이터에 의한 오차 및 검증데이터에 대한 오차가 동시에 감소하는 것을 알 수 있다.

4. 결론 및 향후 연구과제

본 논문에서는 클러스터링 기법을 이용한 뉴로-퍼지 모델의 새로운 학습 방법을 이용하여 모델링을 실시하였다. 기존의 미분기반 학습 방법에서 클러스터링 기반 학습으로 전환하였으며 클러스터링 기법에서는 입출력 인과관계를 고려한 확장된 유사도 측정방식으로 모델의 성능을 개선하고자 하였다.

향후 연구과제로는 클러스터링 기법에서 발생하는 입력의 수가 증가할수록 연산량이 증가하는 문제를 해결하는 것과 뉴로-퍼지 모델의 새로운 학습 규칙을 개발하며 생성된 모델을 이용하여 좀더 실제적인 문제에 적용하는 것등이다.

6. 참고문헌

[1] J-S. R. Jang, C. T. Sun, E. Mizutani, *Neuro-Fuzzy and Soft Computing : A Computational Approach to Learning and Machine Intelligence*, Printice Hall, 1997.

[2] H. B. Verbruggen, H. J. Zimmermann, R. Babuska, *Fuzzy Algorithm for Control*, Kluwer Academic Publisher, 1999.

[3] 김승석, 김성수, 유정웅, "새로운 클러스터링 알고리즘을 적용한 향상된 뉴로-퍼지 모델링", *대한전기학회 논문지*, Vol. 53D, No. 7, pp. 536-543, 2004.

[4] 김승석, 김성수, 유정웅, "개선된 퍼지 클러스터링", *한국퍼지및지능시스템학회논문지*, Vol. 15, No. 1, pp. 6-11, 2005.

[5] R. R. Yager, D. P. Filev, "Generation of Fuzzy Rules by Mountain Clustering", *Jounal of Intelligence and Fuzzy System*, Vo. 2, pp. 209-219, 1994.

[6] Ching-Chang Wong, Chia-Chong Chen, Mu-Chun Su, "A Novel Algorithm for Data Clustering", *Pattern Recognition*, Vol. 34, No. 2, pp. 425-442, 2001.

[7] Roy L. Streit, Tod E. Luginbuhl, "Maximum Likelihood training of Probabilistic Neural Networks", *IEEE Trans on Neural Networks*, Vol. 5, No. 5, pp. 764-783, 1994.

[8] Jeff Bilms, "A Gentle Tutorial of the EM Algorithm and its Application to Paramter Estimation for Gaussian Mixture and Hidden Markov Models, ICSI TR-97-021, April 1998.

[9] Janos Abonyi, J. A. Roubos, M. Oosterom, F. Szeifert, " Compact TS-Fuzzy Models through Clustering and OLS plus FIS Model Reduction", *The 10th IEEE International Conference on Fuzzy Systems*, Vol. 3, pp. 1420-1423, 2001.

[10] Janos Abonyi, R. Babuska, F. Szeifert, "Modified Gath-Geva Fuzzy Clustering for Identification of Tagaki-Sugeno Fuzzy Models", *IEEE Trans on SMC*, Vol. 33, No. 5, pp. 612-621.