

# FMMN 기반 뉴로-퍼지 분류기와 응용

## FMMN-based Neuro-Fuzzy Classifier and Its Application

곽근창\*, 전명근, 유정웅  
충북대학교 전기전자공학부

Keun-Chang Kwak, Myung-Geun Chun, Jeong-Woong Ryu  
School of Electrical and Electronic Engineering  
Chungbuk National University  
(Tel : 043-261-2422, Fax : 043-263-2419)

### ABSTRACT

In this paper, an Adaptive neuro-fuzzy Inference system(ANFIS) using fuzzy min-max network(FMMN) is proposed. Fuzzy min-max network classifier that utilizes fuzzy sets as pattern classes is described. Each fuzzy set is an aggregation of fuzzy set hyperboxes. Here, the proposed method transforms the hyperboxes into gaussian membership functions, where the transformed membership functions are inserted for generating fuzzy rules of ANFIS. Finally, we applied the proposed method to the classification problem of iris data and obtained a better performance than previous works.

### I. 서론

최근에 시스템 모델링의 뉴로-퍼지 기법에 대한 연구가 점점 인기있는 분야로 각광을 받게 되었다. 일반적으로, 이 접근은 구조동정과 파라미터 동정인 두 개의 동정 형태로 구성되며 연속적으로 수행되어진다. 특히 구조동정의 입력공간분할에 대한 문제는 좀 더 효율적으로 분할하여 퍼지 규칙을 생성하기 위해 많은 논문에서 여러 가지 방법들이 제안되어져 왔다.

일반적인 분할의 형태는 그리드 분할, 트리 분할, 스캐터 분할로 나누어질 수 있으며 그리드 분할의 대표적인 논문은 jang의 ANFIS[3], Lin의 NN-FLCS[4], wang의 직교최소자승학습[5]이 있다. 이와 같은 논문들은 우수한 성능을 발휘하나

입력공간이 증가함에 따라 규칙의 수가 지수적으로 증가하는 단점을 가지고 있다. 트리 분할은 각각의 영역에 대응되는 결정 트리에 따라 특징화되어질 수 있도록 입력공간을 분할하는 방법으로 CART[6], ID3[7], 퍼지 k-d트리[8] 등이 있다. 이 방법은 어느 정도 그리드 분할의 단점을 개선하기 위해 적응적인 분할은 하지만 여전히 입력공간 증가에 따른 문제점을 안고 있다. 스캐터 분할은 입력공간이 다차원일 경우 수치적인 데이터로부터 클러스터를 추정하여 퍼지 규칙을 생성하는 방법으로 Bezdek의 FCM[9], Yager의 mountain 클러스터링[10], Chiu의 차감 클러스터링[11] 등이 있다. 이는 클러스터를 구함에 있어 오랜 시간이 요구되며 클러스터의 수를 미리 정하거나 여러 초기 파라미터 설정에 따라 시스템에 많은 변화를 가지고 온다.

따라서, 본 논문에서는 Simpson[12]에 의해 제

안된 퍼지 Min-Max 네트워크를 구축하여 퍼지 하이퍼 박스를 생성한 다음 퍼지 규칙을 자동적으로 생성하도록 한다. 파라미터 동정은 Jang의 방법과 같이 하이브리드 학습을 사용하였다. 이렇게 함으로써 ANFIS의 입력공간이 증가함에 따라 규칙의 수가 지수적으로 증가하는 문제점을 해결할 수 있다. 제안된 뉴로-퍼지 시스템은 분류기로서 사용되어 시뮬레이션의 예로서 iris 데이터 분류에 적용되어진다. 이렇게 제안된 분류기는 이전의 연구들과 비교하여 유용성을 설명한다 [13-15].

## 2. 퍼지 Min-Max 네트워크에 근거한 ANFIS

본 절에서는 퍼지 min-max 네트워크에 근거한 ANFIS의 구축과정에 대해 설명한다. 먼저, 3차원 하이퍼박스에서 min과 max 포인트의 설명은 그림 1과 같이 나타내어진다. 먼저 각 차원에 대해 0과 1사이로 정규화하고 각각의 하이퍼박스 퍼지 집합에 대한 소속함수는 포인트  $x \in R^n$ 가 하이퍼박스 안에 포함되어 있는 정도를 나타낸다.

$R^n$ 에서 각 차원의 값의 범위를 가정하고  $n$ 차원의 공간은 단위 하이퍼박스인  $I^n$ 에서 크기를 조절한다.  $I^n$ 은 전체 집합이 되고  $I^n$ 에서  $j$ 번째 하이퍼박스  $B_j = \{V_j, W_j\}$ 에 대한 소속함수는 다음과 같다.

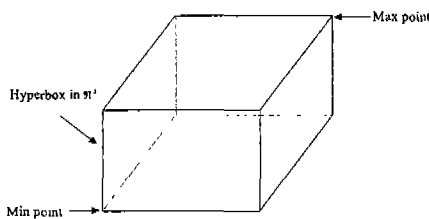


그림 1. min-max 하이퍼박스

$$y_j = m_{B_j}(x) = \left[ 1 - \frac{1}{\alpha_j} \sum_{i=1}^n \max(0, \min(1, \gamma(v_{ji} - x_i))) \right] \times \left[ 1 - \frac{1}{\beta_j} \sum_{i=1}^n \max(0, \min(1, \gamma(x_i - w_{ji})) \right] \quad (1)$$

여기서  $X = (x_1, x_2, \dots, x_n) \in I^n$ 은 입력이고  $V_j = (v_{j1}, v_{j2}, \dots, v_{jn})$ 은  $B_j$ 에 대한 min 포인트이다.  $W_j = (w_{j1}, w_{j2}, \dots, w_{jn})$ 은  $B_j$ 에 대한 max 포인트이다.  $\alpha_j$ 는  $x_i > w_{ji}$ 인 차원의 수,  $\beta_j$ 는  $x_i < v_{ji}$ 인 차원의 수, 그리고  $\gamma$ 는  $x$ 와 하이퍼박스 사이의 거리가 증가함에 따라 얼마나 빨리 소속값이 감소하는지를 조절하는 민감도 파라미터이다.

min-max 포인트는 신경회로망에서 가중치가 되고 활성화함수는 소속함수가 된다. 그리고 각 하이퍼박스는  $n$ 개의 입력노드,  $2n$ 개의 가중치, 하나의 출력 노드를 가진 신경회로망이 된다. 각 입력 노드에서 출력 노드까지 두 연결선을 가지며 하나의 연결선은 각 차원에 대해 min값을 표현하고, 다른 연결선은 max값을 표현한다.  $i$ 번째 입력 노드와  $j$ 번째 출력 노드 사이의 연결은  $v_{ji}$ 와  $w_{ji}$ 이며  $j$ 번째 노드의 출력값  $y_j$ 는 식(1)을 이용하여 계산한다. 퍼지 min-max는 winner-take-all 과정을 이용한다. 가장 큰 값을 가지는 출력 노드는 winner로 정하고 그 하이퍼박스에 대한 클래스 레벨은 클래스로 나타낸다.

퍼지 min-max 학습과정은 다음과 같이 확장 과정과 축소과정으로 이루어지며, 학습집합  $D$ 는  $M$ 개 쌍  $\{X_k, d_k\}$ 의 집합으로 구성된다.

여기서  $X_k = (x_{k1}, x_{k2}, \dots, x_{kn}) \in I^n$ 은 입력 패턴이며  $d_k \in \{1, 2, \dots, m\}$ 은  $m$ 클래스 중 하나의 index이다. 학습 과정은  $D$ 로부터 순서화된 쌍을 선택하고, 입력 데이터를 포함시키는 것을 필요로 한다면 같은 클래스에 대해 확장할 수 있는 하이퍼박스를 찾음으로서 시작한다. 하이퍼박스는 확장 기준에 확인이 되지 않는다면, 새로운 하이퍼박스는 생성되고 신경회로망에 더해진다. 이 성장 과정은 클래스가 비선형적으로 분리할 수 있도록 한다. 하이퍼박스 확장 과정은 중첩된 하이퍼박스를 찾았을 때 같은 클래스를 표현하는 하이퍼박스는 문제가 되지 않는다. 그러나, 다른 클래스를 표현하는 하이퍼박스 사이에 중첩이 발생된다면, 각각의 중첩된 하이퍼박스를 축소함으로써 제거되어진다.

이렇게 얻어진 최종적인 하이퍼박스는 사다리꼴 소속함수이기 때문에 비선형성이 강한 시

시스템일 경우에는 적합하지 않으므로 가우시안 소속함수의 형태로 변형시킨다. 가우시안 소속함수의 중심은 min값과 max값의 중심으로 사용하고 폭은 중심으로부터 min값 혹은 max값까지의 거리를 사용하면 된다. 얻어진 소속함수는 ANFIS의 2층에서 스캐터 분할에 의해 퍼지 규칙이 생성되며 하이브리드 학습을 이용하여 전제부 및 결론부 파라미터를 조정한다.

따라서, ANFIS 구조는 그림2와 같이 나타내어지며 구조동정은 점선부분과 같이 퍼지 min-max 네트워크에 의해 수행되어짐으로서 기존의 ANFIS가 가지는 입력공간이 증가할 때 규칙의 수가 증가하는 단점을 개선시킬 수 있으며, 이 구축된 시스템은 분류기로서 사용하였다.

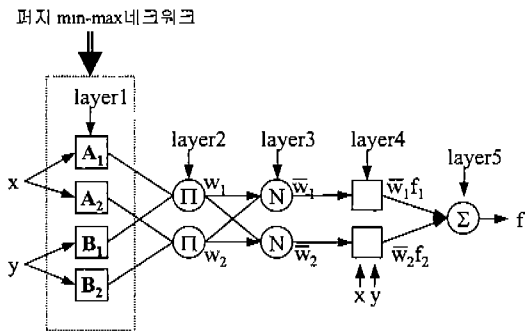


그림 2. 제안된 뉴로-퍼지 시스템의 구조

워크를 본 논문에서 제안된 방법을 이용하여 3개의 하이퍼 박스를 가우시안 소속함수로 변형한 다음 그림 3과 같은 형태로 ANFIS에 3개의 퍼지 규칙을 생성하였다. 100epoch동안 시뮬레이션 한 결과 그림 4와 같이 한 개의 데이터(42번째)만 오류를 나타내고 있음을 알 수 있다. 이전 논문인 퍼지 min-max 네트워크에 의한 분류는 48개의 하이퍼박스를 생성하였고 2개의 데이터에 대해서 잘못된 분류를 보였다.

표 1은 시뮬레이션 결과로 이전 연구 결과와 제안된 방법의 분류 결과를 나타내고 있으며, 제안된 방법인 FMN-ANFIS는 이전의 다른 논문에 비해 좋은 결과를 보이고 있음을 알 수 있다.

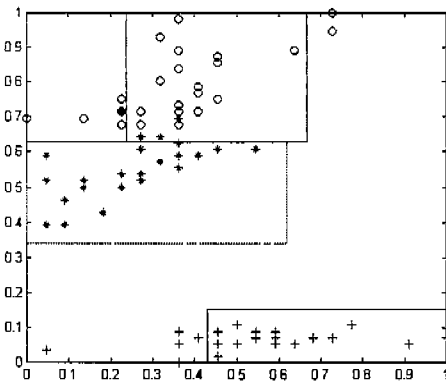


그림 3. 퍼지 min-max 네트워크에 의한 iris 검증 데이터의 분류

### III. 시뮬레이션 및 결과

본 논문에서 제안한 퍼지 min-max 네트워크에 근거한 ANFIS의 성능을 알아보기 위해 Fisher의 iris 데이터 분류에 적용하였다. 이 데이터 집합은 150개의 쌍으로 된 4차원 패턴들로 이루어졌다. 또한 3개의 클래스를 포함하며 각 클래스는 50개의 데이터로 이루어진다. 학습데이터는 임의로 75개를 선택하고 검증데이터는 나머지 75개의 데이터를 이용하였다. 그림 3은 학습데이터를 이용하여 퍼지 min-max 네트워크를 구성한 다음 검증 데이터에 대해 시뮬레이션 한 결과이다. 그림에서 알 수 있듯이 3개의 하이퍼 박스를 생성하며 4개의 데이터(32, 39, 42, 60번째)가 오류된 분류를 나타내었다. 이 구축된 퍼지 min-max 네트

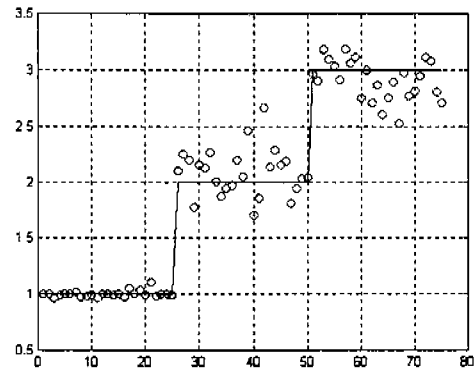


그림 4. 제안된 방법에 의한 분류 결과

#### IV. 결론

본 논문에서는 새로운 적응 뉴로-퍼지 시스템을 구축하기 위해 Simpson에 의해 제안된 퍼지 Min-Max 네트워크를 이용하였다. 먼저, 퍼지 Min-Max 네트워크를 구축하여 하이퍼박스를 생성한 다음 비선형 특성에 맞도록 가우시안 소속함수로 변형하고 이 소속함수는 ANFIS에서 퍼지 규칙을 자동적으로 생성하도록 하였다. 파라미터 동정은 jang에 의해 제안된 역전파 알고리즘과 최소자승법을 이용한 하이브리드 학습을 사용하였다. 시뮬레이션의 예로서 iris 데이터 분류에 적용하여 이전의 논문에 비해 우수한 성능을 보였으며, 이렇게 함으로서 ANFIS의 입력공간이 증가함에 따라 규칙의 수가 지수적으로 증가하는 문제점을 해결할 수 있고, 기존의 퍼지 min-max 네트워크에 비해 적은 하이퍼박스의 수로도 분류 성능을 좋게 할 수 있음을 알 수 있었다.

표 1. 분류성능의 비교

방법	오류 인식수
Bayes classifier[13]	2
k-nearest neighbor[13]	4
Fuzzy k-NN[14]	4
Fisher ratios[13]	3
Ho-Kashyap[13]	2
Perceptron[15]	3
Fuzzy Perceptron[15]	2
Fuzzy min-max network[12]	2
The proposed method	1

#### 참고문헌

- [1] J. S. R. Jang, C. T. Sun, "Neuro-Fuzzy and Soft Computing", Prentice Hall, 1997.
- [2] C. T. Lin, C. S. G. Lee, "Neural Fuzzy Systems", Prentice Hall, 1996.
- [3] J. S. R. Jang, "ANFIS : Adaptive-Network based Fuzzy Inference System", *IEEE Trans. System Man Cybernetics*, Vol. 23, No. 3, pp. 665-685, 1993.
- [4] C. T. Lin, C. S. G. Lee, "Reinforcement structure parameter learning for neural network-based fuzzy logic control system", *IEEE Trans. on Fuzzy System*, Vol. 2, pp. 4-63, 1994.
- [5] L. X. Wang, J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-square learning", *IEEE Trans. Neural Networks* 3, pp. 807-814, 1992.
- [6] J. S. R. Jang, "Structure determination in fuzzy modeling : A Fuzzy CART Approach", *Proceedings of IEEE International Conference on Fuzzy Systems*, 1994.
- [7] C. Z. Janikow, "Fuzzy Decision Trees", *IEEE Transaction on Man Cybernetics*, Vol. 28, No. 1, 1998.
- [8] C. T. Sun, "Rule-based structure identification in an Adaptive Network-Based Fuzzy Inference System", *IEEE Transaction on Fuzzy system*, Vol. 2, No. 1, 1994.
- [9] J. Bezdek, R. Hathaway, M. Sabin and W. Tucker, "Convergence theory for fuzzy c-means: Counterexamples and repairs", CRC Press, Vol. 3, Chap. 8, 1987.
- [10] R. R. Yager, D. P. Filev, "Generation of Fuzzy Rules by Mountain Clustering", *Journal of Intelligent and Fuzzy System* Vol. 2, pp. 209-219, 1994.
- [11] S. Chiu, "Fuzzy Model Identification based on Cluster Estimation", *Journal of Intelligent Fuzzy System* Vol.2, NO. 3, 1994.
- [12] P. K. Simpson, "Fuzzy Min-Max Neural Networks-Part1 : Classification", *IEEE Trans. on Neural Networks*, Vol. 3, No. 5, pp. 776-786, 1992.
- [13] S. Fahlman, C. Lebiere, "The cascade-correlation learning architecture", Carnegie Mellon University, 1990.
- [14] J. Bezdek, S. Chuah and D. Leep, "Generalized k-nearest neighbor rules", *Fuzzy Sets and Systems*, Vol. 18, pp. 237-256, 1986.
- [15] J. Keller, D. Hunt, "Incorporating fuzzy membership functions into the perceptron algorithm", *IEEE Trans. Patt. Anal. Intell.*, Vol. 7, pp. 693-699, 1985.