

# EM 알고리즘에 의한 뉴로-퍼지 시스템의 퍼지 규칙 생성

## Fuzzy rule Extraction of Neuro-Fuzzy System using EM algorithm

김승석 · 광근창 · 유정웅 · 전명근  
충북대학교 전기, 전자 및 컴퓨터공학부

Seoung-Suk Kim and Keun-Chang Kwak, Jeong-Woong Ryu, Myung-Geun Chun

\*School of Electrical and Computer Engineering, Chungbuk National University

E-mail : powerkimss@hotmail.com

### ABSTRACT

본 논문에서는 여러 분야에서 널리 응용되고 있는 적응 뉴로-퍼지 시스템(ANFIS)에서의 효과적인 퍼지 규칙 생성방법을 제안한다. ANFIS의 성능 개선을 위해 구조동정을 수행함에 있어서 전제부 파라미터는 EM(Expectation-Maximization) 알고리즘을 적용하였으며, 파라미터 학습은 Jang에 의한 하이브리드 방법을 적용한다. 여기서 초기의 중심과 분산을 구하기 위해 FCM(Fuzzy c-means) 클러스터링 기법을 사용하였다. 이렇게 함으로서 적은 규칙 수를 가지면서도 효율적인 퍼지 규칙을 얻을 수 있도록 하였다. 이들 방법의 유용함을 보이고자 Box-Jenkins의 가스로 데이터에 적용하여 제안된 방법이 이전의 연구보다 좋은 결과를 보임을 보이고자 한다.

**Keyword** : Expectation Maximization, Fuzzy Rule Extraction, ANFIS, Fuzzy c-Means clustering

### 1. 서론

패턴인식, 신경망 학습, 제어, 신호처리 및 통신 등에서 중요하게 다루어지는 문제 중에 하나가 미지의 데이터로부터 원하는 파라미터의 추정이다. 이전의 연구들에서 다양하게 이루어졌던 Hard Fuzzy C-means (HCM), Fuzzy C-Means (FCM)[1], 퍼지 군등화 방법[2]들을 통하여 주어진 데이터들을 적절히 분할하거나 평균 등을 이용하여 필요한 파라미터를 추정하였다. 이러한 반복적인 파라미터 최적화 기법으로 EM 알고리즘[3][4]이 있다. 본 알고리즘은 Dempster[4]가 다른 연구자들의 생각을 결합시켜 수렴성을 증명하고 EM 알고리즘이라는 용어를 만들어 내기 전까지 여러 다른 연구자들에 의해 독립적으로 연구되어 적용되어 왔다. 파라미터들을 추정하는데 필요한 데이터가 직접 처리가 불가능하거나, 데이터의 일부가 손실되었을 경우 파라미터의 추

정은 더욱 복잡하고 어렵다. EM 알고리즘은 데이터의 기본 분포에서 관측을 통제하는 분포로 다대일 사상이 있을 때 파라미터의 Maximum Likelihood Estimation (MLE)[5]을 통하여 추정을 한다. EM 알고리즘은 예측 단계와 최대화 단계의 두 개의 주요한 단계로 구성된다. 예측에서는 현재의 파라미터를 추정하고 관측의 필요조건이 되므로 미지의 기본 변수들에 관계한다. 최대화 단계에서는 새로운 파라미터를 추정한다. 이 두 단계는 수렴할 때까지 반복된다.

제안된 방법은 초기치에 민감한 여러 알고리즘에 EM에서의 결과를 이용함으로써 원하는 모델의 성능을 개선하고자 한다. 모델의 학습 전의 초기치를 좀더 최적에 설정하므로써 원하는 성능 개선이나 처리시간을 줄일 수 있다. 이러한 과정을 통하여 추정된 파라미터의 값의 유용성을 설명하기 위하여 Radial Basis Function Network

(RBFN)과 Adaptive Neural based Fuzzy Inference System (ANFIS)[6]등의 뉴러-퍼지 모델의 초기 값으로 적용하여 Box-Jenkins의 가스로 데이터[7]에 적용하여 제안된 방법의 유용성을 보이고자 한다.

## 2. EM 알고리즘에 의한 뉴러-퍼지 시스템 구축

본 절에서는 EM 알고리즘을 이용한 ANFIS에서의 퍼지 규칙과 소속함수를 결정하는 구조동정에 대해서 다루고자 한다.

먼저, EM 알고리즘을 적용하기 위해서는 두 단계에 의해 초기 중심과 분산을 구해야 한다. 첫 단계에서는 기존의 K-means 클러스터링에 의해 초기 중심을 구하는 방법대신 주어진 데이터로부터 0과 1사이의 소속도를 가진 유사한 클러스터로 나누는 개선된 형태인 FCM 클러스터링을 적용한다. 또한 두 번째 단계는 가장 가까운 중심에 대한 거리로서 초기 분산 값을 설정한다.

위 두 단계에 의해 얻어진 초기 중심과 분산은 최대우도 추정(Maximum likelihood estimate)에 가장 가깝도록 EM 알고리즘을 이용하여 반복적으로 수행한다. EM 알고리즘은 두 단계과정인 Expectation 단계와 Maximization 단계로 나누어진다. Expectation 단계에서 사후 확률은 식(1)과 같이 나타내어진다.

$$P(\omega_j|x_k, \Phi) = \frac{p(x_k|\omega_j, \Phi)P(\omega_j)}{\sum_k p(x_k|\omega_i, \Phi)P(\omega_i)} \equiv h_{kj} \quad (1)$$

가우시안 밀도함수가  $d$ 차원이라면 식(2)와 같다.

$$p(x|\omega_j, \Phi) = \frac{1}{(2\pi)^{d/2}|\Sigma_j|^{1/2}} \times \exp\left[-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right] \quad (2)$$

$g_j$ 가  $P(\omega_j)$ 라고 한다면 식(3)과 같다.

$$h_{kj} = \frac{g_j |\Sigma_j|^{1/2} \exp[-(1/2)(x_k - \mu_j)^T \Sigma_j^{-1}(x_k - \mu_j)]}{\sum_l g_l |\Sigma_l|^{1/2} \exp[-(1/2)(x_k - \mu_l)^T \Sigma_l^{-1}(x_k - \mu_l)]} \quad (3)$$

Maximization 단계에서, 구성요소 파라미터  $\Phi$ 인 평균, 분산, 가중치를 갱신한다.

$$\mu_j = \frac{\sum_k h_{kj} x_k}{\sum_k h_{kj}} \quad (4)$$

$$\Sigma_j = \frac{\sum_k h_{kj} (x_k - \mu_j)(x_k - \mu_j)^T}{\sum_k h_{kj}} \quad (5)$$

$$g_j = \frac{1}{n} \sum_k h_{kj} \quad (6)$$

그림 1은 EM 알고리즘의 순서도를 나타내고 있다.

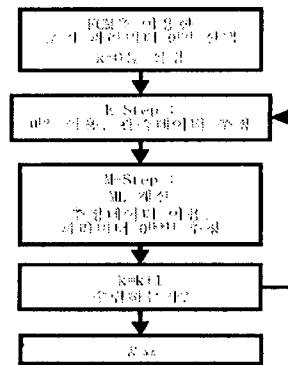


그림 1. EM 알고리즘의 순서도

구조동정이 EM 알고리즘에 의해 결정된 후에 파라미터 동정은 Jang의 방법과 같이 최소사승법과 역전과 알고리즘에 의해 수행되어진다.

여기서 ANFIS 모델링과 학습 메카니즘을 간략히 살펴보자. 먼저 아래와 같이 결론부가 일차 선형방정식인 두 개의 TSK형태의 퍼지 규칙을 갖는 추론시스템에 대해 생각해보자.

$$R^1: \text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ then } f_1 = p_1x + q_1y + r_1$$

$$R^2: \text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ then } f_2 = p_2x + q_2y + r_2$$

(7)

여기서  $x$ 와  $y$ 는 입력이고  $A_i, B_i$ 는 언어적 변수이며  $\{p_i, q_i, r_i\}$ 는 결론부 파라미터이다. 위와 같은 규칙에 대해 적응네트워크 기반 퍼지 시스템의 구조는 그림 2와 같이 나타낼 수 있으며 동작특성과 학습절차는 다음과 같다.

1층의 전제부 소속함수는 식(8)과 같은 가우시안 소속함수를 선택하였다.

$$\mu_{A_i}(x) = \exp\left\{-\left(\frac{x-c_i}{a_i}\right)^2\right\} \quad (8)$$

여기서  $\{a_i, c_i\}$ 는 전제부 파라미터이며,  $a_i$ 는 소속함수의 폭이고  $c_i$ 는 소속함수의 중심이다. 이들 파라미터는 EM 알고리즘으로 초기화된다.

최종적인 추론 값은 식(9)와 같이 가중평균법에 의해 구해진다.

$$y_i^* = \frac{\sum_{i=1}^2 \bar{w}_i f_i}{\sum w_i} = \frac{\sum w_i f_i}{\sum w_i} \quad (9)$$

여기서 각 파라미터는 다음과 같다.

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \quad i=1,2 \quad (10)$$

$$\bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (11)$$

$$\bar{w}_i = \frac{w_i}{w_1 + w_2} \quad (12)$$

문에서는 학습오차와 검증오차가 서로 유사한 값이 나온다는 것을 알 수 있었다.

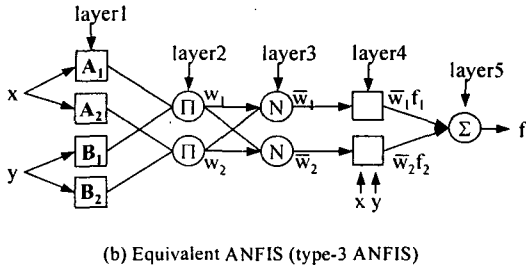


그림 2. ANFIS의 구조

위와 같은 적응네트워크 기반 퍼지 시스템의 결론부 파라미터  $\{p_i, q_i, r_i\}$ 는 식(13)과 같이 최소자승법으로 추정된다.

$$X^* = (A^T A)^{-1} A^T B \quad (13)$$

### 3. 시뮬레이션 및 결과

Box-Jenkins의 가스로 데이터를 이용한 비선형 시스템 모델링을 하기 위해 본 논문에서 제안한 뉴로-퍼지 시스템을 적용하였다. 여기서, 입력은 가스 흐름을  $u(k)$ 와 출력은 이산화탄소 농도  $y(k)$ 인 시계열 데이터로 이루어졌다. 본 논문에서는 이전의 다른 논문과 비교하기 위해 입력을  $y(k-1)$ 과  $u(k-4)$ 로 이용하였으며 출력은  $y(k)$ 를 사용하였다.

그림 3은 입력의 전체 데이터에 대한 분포도와 EM 알고리즘에 의해 구해진 6개의 중심과 분산을 나타내고 있다. 이를 초기치로 하여 ANFIS에 적용한 결과 그림 4와 같이 모델의 추론 값과 원하는 데이터 출력 값을 비교하고 있다. 그림에서 알 수 있듯이 작은 오차를 보이면서 가스로 데이터가 예측되고 있음을 알 수 있었다. 표 1은 시뮬레이션 결과로 모델에 대한 규칙 수와 오차를 비교한 결과이며 제안된 방법이 이전의 다른 논문과 RBFN에 비해 적은 규칙의 수와 입력 수를 가지고도 작은 오차를 보임을 알 수 있다. 또한, 그림 5와 6은 시스템의 출력이 모델 형성을 위한 학습데이터에 편중되지 않고 새로운 데이터에 대해서도 타당한 결과를 보일 수 있는가를 평가하기 위해 데이터 집합을 학습데이터와 검증데이터로 나누어 예측한 결과이다. 그림은 100epoch동안 학습한 결과이며 두 경우 모두 적은 오차를 보이면서 예측하고 있음을 알 수 있었다. 표 2는 제안된 모델에 대한 규칙 수와 오차를 비교한 결과이며 Jang과 Oh는 학습오차는 작은 반면 검증오차가 상대적으로 커지기 때문에 overfitting문제를 발생할 수 있지만 본 논

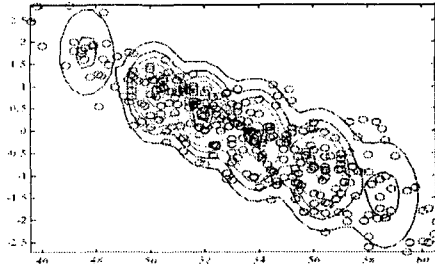


그림 3. 입력 데이터의 분포도

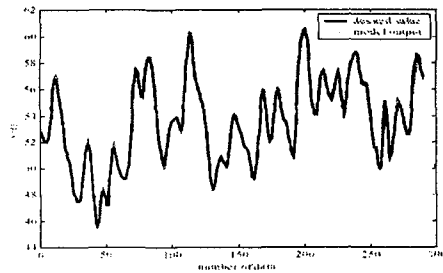


그림 4. 제안된 모델의 추론값과 실제 데이터 출력과의 비교

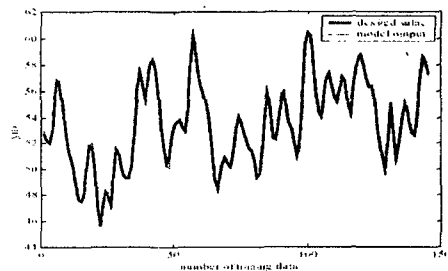


그림 5. 학습 데이터에 대한 제안된 모델의 추론 값과 실제 데이터 출력과의 비교

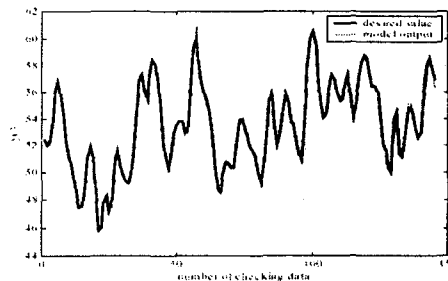


그림 6. 검증 데이터에 대한 제안된 모델의 추론 값과 실제 데이터 출력과의 비교

표 1. 규칙의 수와 오차의 비교

	입력 수	규칙 수	MSE
Tong[8]	2	19	0.469
Pedrycz[9]	2	20	0.776
Xu[10]	2	25	0.328
Sugeno[11]	3	6	0.190
Joo[12]	2	6	0.166
RBFN(EM)	2	6	0.171
ANFIS(EM)	2	6	0.118

표 2. 규칙의 수와 오차의 비교

	입력 수	규칙 수	학습 오차	검증 오차
Jang[13]	2	4	0.135	0.530
Oh[14]	2	4	0.026	0.272
	2	6	0.020	0.264
RBFN(EM)	2	6	0.179	0.188
ANFIS(EM)	2	6	0.110	0.158

## 5. 결론 및 향후과제

본 논문에서는 EM 알고리즘을 이용한 Neuro-Fuzzy 모델링을 실시하였다. 각각 RBFN과 ANFIS의 은닉 층의 소속함수와 전제부 소속함수의 초기치 추정에 EM 알고리즘을 효과적으로 사용하여 각 모델에 대하여 성능을 개선하였다. 향후 연구과제로는 초기치 선택있어, EM 알고리즘에서의 Gaussina pdf를 Neuro-Fuzzy 모델에서의 Gaussian 소속함수로의 타당한 변환과 국부적 최적해에 대한 해결 등이 있다.

## 6. 참고문헌

[1] J. S. R. Jang, C. T. Sun, E. Mizutani, "Neuro-Fuzzy and Softing : A Computational Approach to Learning and Machine Intelligence", *Prentice Hall*, 1997.  
 [2] 김승석, 박근창, 유정웅, 전명근, "퍼지군등화와 언어적 Hedge를 이용한 GA 기반 순차적 모델링", *한국 퍼지 및 지능시스템학회 논문지*, Vol. 11, No. 9, pp. 827-832, 2001.  
 [3] Ethem Alpayd, "Soft vector quantization and the EM algorithm", *Neural Networks*, Volume 11, Issue 3, April 1998, Pages 467-477  
 [4] Todd K. Moon, "*IEEE Signal Processing Magazine*", IEEE, 1996.

[5] Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern Classification", *John Wiley & Sons Inc*, 2001.  
 [6] J. S. Jang, "ANFIS : Adaptive-Networks-based Fuzzy Inference System", *IEEE Trans. on System, Man and Cyernetics*, Vol. 23, No. 3, pp. 665-685, 1993.  
 [7] Reza Langari, Liang Wang, John Yen, "Radial Basis Function Networks, Regression Weights, and the Expectation-Maximization Algorithm", *IEEE Trans on System, Man, And Cyernetic-Part A : Systems and Humans*, Vol. 27, No. 5, pp. 613-623, 1997.  
 [8] R. M. Tong, "The evaluation of fuzzy models derived from experimental data", *Fuzzy Sets and Systems*, Vol. 4, pp. 1-12, 1989.  
 [9] W. Pedrycz, "An identification algorithm in fuzzy relational systems", *Fuzzy Sets and Systems*, Vol. 13, pp. 153-167, 1984.  
 [10] C. W. Xu, Y. Z. Lu, "Fuzzy model identification and self-learning for dynamics systems", *IEEE Trans. on System, Man, and Cybernetic*, Vol. 17, pp. 683-589, 1987.  
 [11] M. Sugeno, K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system", *Fuzzy Sets and Systems*, Vol. 42, pp. 315-334, 1991.  
 [12] Y. H. Joo, H. S. Hwang, K. B. Kim, K. B. Woo, "Fuzzy system modeling by fuzzy partition and GA hybrid schemes", *Fuzzy Sets and Systems*, Vol. 86, pp.279-288, 1997.  
 [13] Jang, "Input Selection for ANFIS learning", in *Proceedings of the IEEE International Conference on Fuzzy Systems*, New Orleans, 1996.  
 [14] S. K. Oh, W. Pedrycz, "Identification of fuzzy systems by means of an auto-tuning algorithm and its application to nonlinear systems", *Fuzzy Sets and Systems*, Vol. 115, pp. 205-230, 2000.