

다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크 설계 및 최적화

박건준, 오성권, 김현기
수원대학교 전기공학과

Design of Fuzzy Relation-based Fuzzy Neural Networks with Multi-Output and Its optimization

Keon-Jun Park, Sung-Kwun Oh, and Hyun-Ki Kim
Dep. of Electrical Engineering, The University of Suwon

Abstract - 본 논문에서는 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크를 설계한다. 퍼지 관계 기반 퍼지뉴럴네트워크는 전체 입력 변수에 따른 입력 공간을 분할함으로써 네트워크를 구성한다. 규칙의 전반부는 앞서 언급한 전체 입력 공간을 분할하여 표현하고, 규칙의 후반부는 다항식으로 표현되며 오류역전파 알고리즘을 이용하여 연결가중치인 후반부 다항식을 학습한다. 또한, 각 입력에 대한 전반부 멤버십함수의 정점과 학습률 및 모멘텀 계수를 유전자 알고리즘을 이용하여 최적 동조한다. 따라서 유전자 알고리즘을 이용하여 퍼지뉴럴네트워크를 최적 설계한다. 마지막으로 제안된 모델은 표준 모델로서 널리 사용되는 수치적인 예를 통하여 평가한다.

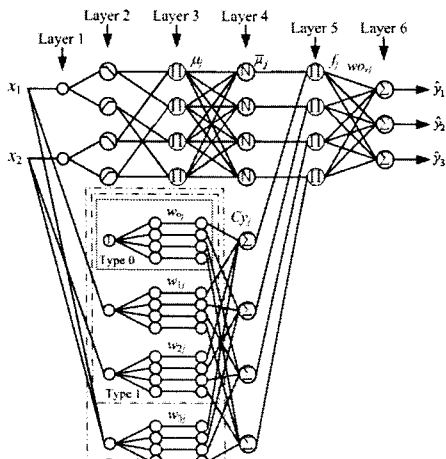
1. 서 론

퍼지 이론과 신경망을 상호 융합한 지능형 모델, 즉 퍼지뉴럴네트워크[1]는 비선형 공정에 대한 복잡성과 불확실성을 제어하기 위해 많은 연구가 이루어지고 있다. 퍼지이론은 1965년 Zadeh에 의해 창안된 퍼지 집합 이론을 이용하여 비선형적이고 복잡한 실 시스템의 특성을 해석하는데 적용함으로써 시스템의 성능 및 기능면에 있어서 기존의 수학적 모델보다 좋은 결과를 가져왔다. 인간의 두뇌가 이루고 있는 신경의 구조를 모방한 신경망은 학습을 통하여 모델이 지능적으로 행동할 수 있도록 하였고 계산처리에 있어서 병렬 분산 처리를 함으로써 계산처리 능력을 향상시키는 장점을 가졌다. 인간의 언어 및 사고에 관련된 애매함을 다루는 퍼지이론과 학습능력을 가진 신경회로망을 상호 결합한 지능형 모델인 퍼지뉴럴네트워크는 기존의 모델에 비해서 학습속도가 빠르고, 수렴특성이 우수한 장점을 가진다. 따라서 본 논문에서는 입력 공간을 상호 관계적으로 분할하고 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크를 설계한다. 설계된 네트워크는 실수코딩 유전자 알고리즘[2]을 이용하여 각 입력 변수에 대한 전반부 멤버십함수의 정점과 학습률 및 모멘텀 계수를 최적적으로 동조한다. 또한, 제안된 퍼지뉴럴네트워크를 적용하기 위하여 다중 패턴 실험 데이터로 Fisher의 Iris 데이터를 이용하여 평가한다.

2. 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크 설계

2.1 퍼지뉴럴네트워크의 구조

전체 입력 변수에 따른 공간 분할에 의한 퍼지뉴럴네트워크 구조는 그림 1과 같다.



<그림 1> 퍼지뉴럴네트워크 구조

그림 1에서 각 입력신호는 각 멤버십함수를 거쳐 적합도가 구해지고 적합도는 연결가중치 w 와 곱해지며 이들이 대수적으로 합해져 j 번째 규칙의 출력 f_j 를 생성한다. 최종 출력값은 w_{0j} 와 곱해진 다음 대수적으로 합해진다.

$$\hat{y}_r = \sum_{j=1}^n f_j \cdot w_{0j} \quad (1)$$

퍼지규칙은 기존의 내부구조가 규칙을 이루지만, 제안된 네트워크는 전체 구조가 퍼지규칙을 이루며, 그 규칙은 다음과 같다.

$$R^j : \text{If } x_1 \text{ is } A_{1j} \text{ and } \dots \text{ and } x_k \text{ is } A_{kj} \quad (2)$$

$$\text{then } Cy_j = g(x_1, \dots, x_k)$$

간략 추론 : $g = w_{0j}$

선형 추론 : $g = w_{0j} + w_{1j} \cdot x_1 + \dots + w_{kj} \cdot x_k$

변형된 2차식 추론 : $g = w_{0j} + w_{1j} \cdot x_1 + \dots + w_{1j} \cdot x_k$

$$+ w_{(k+1)j} \cdot x_1 \cdot x_2 + \dots + w_{(k(k+1)/2)j} \cdot x_{k-1} \cdot x_k$$

여기서, R^j 는 j 번째 퍼지규칙을 나타내며, A_{kj} 는 전반부 퍼지변수로 멤버십함수로 표현된다. w_{kj} 는 후반부 상수이며, 퍼지 규칙내의 연결가중치를 나타낸다.

전반부 멤버십함수는 이웃하는 멤버십함수와 보수적인 관계를 가진 삼각형 멤버십함수를 사용한다. 따라서 j 번째 규칙에 대한 결과는 다음과 같다.

$$f_j = \mu_j \cdot Cy_j \quad (3)$$

제안된 퍼지뉴럴네트워크 구조에서 각 층의 연산과정은 다음과 같다.

[1층] 입력층

[2층] 입력에 대한 적합도 계산 : μ_j

[3층] 각 규칙의 적합도 계산 : $\mu_j = \mu_{A_1}(x_1) \times \dots \times \mu_{A_k}(x_k)$

[4층] 정규화된 적합도 계산 : $\bar{\mu}_j = \mu_j / \sum_{j=1}^n \mu_j = \mu_j$

[5층] 연결가중치와의 곱 : $f_j = \bar{\mu}_j \cdot Cy_j$

[6층] 최종 출력 : $\hat{y}_r = \sum_{j=1}^n f_j \cdot w_{0j}$

2.2 퍼지뉴럴네트워크의 학습

제안된 네트워크의 학습은 퍼지규칙내의 연결가중치 w_{kj} 와 출력층의 연결가중치 w_{0j} 를 가장 적절한 값으로 조정해 나가면서 이루어진다. 학습방법은 오류역전파 알고리즘을 이용하며, 이는 원 출력과 모델 출력간의 차, 즉 오차를 바탕으로 이루어진다.

$$E_p = \frac{1}{2} \sum_{r=1}^s (y_{pr} - \hat{y}_{pr})^2 \quad (4)$$

2.2.1 출력층의 연결가중치 조절

연결가중치의 조절은 다음과 같이 기존의 값에 오차로부터 구해진 조절량 만큼의 변화를 더해 준다.

$$w_{0rj}(t+1) = w_{0rj}(t) + \Delta w_{0rj} \quad (5)$$

$$\Delta w_{0rj} = \eta \left(- \frac{\partial E_p}{\partial w_{0rj}} \right) \quad (6)$$

$$- \frac{\partial E_p}{\partial w_{0rj}} = - \frac{\partial E_p}{\partial \hat{y}_{pr}} \cdot \frac{\partial \hat{y}_{pr}}{\partial w_{0rj}} \quad (7)$$

여기서,

$$- \frac{\partial E_p}{\partial \hat{y}_{pr}} = - \frac{\partial}{\partial \hat{y}_{pr}} \left(\frac{1}{2} \sum_{r=1}^s (y_{pr} - \hat{y}_{pr})^2 \right) = y_{pr} - \hat{y}_{pr} \quad (8)$$

$$\frac{\partial \hat{y}_{pr}}{\partial w_{0rj}} = \frac{\partial}{\partial w_{0rj}} \left(\sum_{j=1}^n f_j \cdot w_{0rj} \right) = f_j \quad (9)$$

따라서,

$$\Delta w_{0rj} = \eta (y_{pr} - \hat{y}_{pr}) f_j \quad (10)$$

또한, 모멘텀을 포함하는 연결가중치 변화량은 다음과 같다.

$$\Delta w_{0rj} = \eta (y_{pr} - \hat{y}_{pr}) f_j + \alpha (w_{0rj}(t) - w_{0rj}(t-1)) \quad (11)$$

2.2.2 퍼지규칙내의 연결가중치 조절

같은 방법으로 연결가중치의 조절은 이루어지며, 간략 추론일 경우에는 다음과 같다.

$$w_{0j}(t+1) = w_{0j}(t) + \Delta w_{0j} \quad (12)$$

$$\Delta w_{0j} = \eta \left(-\frac{\partial E_p}{\partial w_{0j}} \right) \quad (13)$$

$$-\frac{\partial E_p}{\partial w_{0j}} = \sum_{r=1}^s \left(-\frac{\partial E_p}{\partial \hat{y}_{pr}} \cdot \frac{\partial \hat{y}_{pr}}{\partial f_j} \right) \cdot \frac{\partial f_j}{\partial w_{0j}} \quad (14)$$

여기서,

$$-\frac{\partial E_p}{\partial \hat{y}_{pr}} = -\frac{\partial}{\partial \hat{y}_{pr}} \left(\frac{1}{2} \sum_{r=1}^s (y_{pr} - \hat{y}_{pr})^2 \right) = y_{pr} - \hat{y}_{pr} \quad (15)$$

$$\frac{\partial \hat{y}_{pr}}{\partial f_j} = \frac{\partial}{\partial f_j} \left(\sum_{i=1}^n f_i \cdot w_{0i} \right) = w_{0j} \quad (16)$$

$$\frac{\partial f_j}{\partial w_{0j}} = \frac{\partial}{\partial w_{0j}} (\mu_j \cdot w_{0j}) = \mu_j \quad (17)$$

따라서,

$$\Delta w_{0j} = \eta \sum_{r=1}^s \{ (y_{pr} - \hat{y}_{pr}) w_{0j} \} \mu_j \quad (18)$$

또한, 모멘텀을 포함하는 연결가중치 변화량은 다음과 같다.

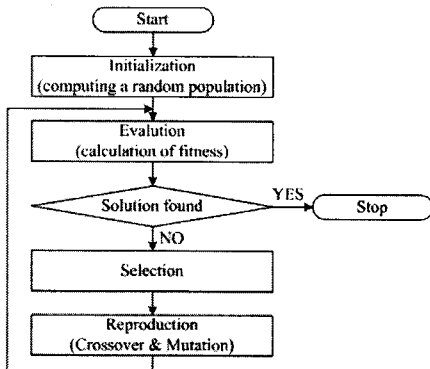
$$\Delta w_{0j} = \eta \sum_{r=1}^s \{ (y_{pr} - \hat{y}_{pr}) w_{0j} \} \mu_j + \alpha \{ w_{0j}(t) - w_{0j}(t-1) \} \quad (19)$$

선형 추론인 경우와 변형된 2차식 추론인 경우에도 같은 방법으로 유도된다.

3. 최적화 알고리즘

유전자 알고리즘[2]은 자연 선택과 유전학에 기반으로 하는 확률적인 탐색방법으로써 탐색과 해의 가능영역들을 균형 있게 이용하기 위하여 생산, 교배, 돌연변이의 과정을 수행하는 일반성 있는 탐색방법으로 비선형 최적화 이론에 탁월한 성능을 발휘하고 있다. 기존의 다른 탐색방법들과 달리 유전자 알고리즘은 해가 될 가능성이 있는 개체집단을 유지하면서 그들 모두가 동시에 최적값을 찾아나가기 때문에 지역 극소에 빠질 위험을 어느 정도 해결할 수 있다는 점과 더불어 모델의 성능지수가 최소가 되는 전역 극소 영역을 찾는 능력을 갖고 있으며, 또한 기존의 방법들과는 달리 선형, 연속, 미분가능 등의 제한이 없기 때문에 다양한 분야에 별다른 제한 없이 적용할 수 있다는 장점을 가진다는 것이 중요한 특징이다.

본 논문에서는 제안된 네트워크에 대한 전반부 멤버쉽함수의 정점과 학습 및 모멘텀을 최적으로 설정해 주기 위하여 실수 코딩 유전자 알고리즘을 이용하여 네트워크의 최적화 과정을 수행하게 된다.



〈그림 1〉 유전자 알고리즘의 수행과정

4. 실험 데이터를 통한 결과 고찰

본 논문에서 제안된 퍼지뉴럴네트워크를 적용하기 위하여 다중 패턴 실험 데이터로 Fisher의 Iris 데이터를 이용하였다. 실험에 사용된 Iris 데이터는 3개의 클래스(setosa, versicolor, virginica)로 구성되며, 각 클래스마다 50개의 패턴으로 구성되어 있고, 각 패턴은 4개의 속성으로 정의된다.

퍼지뉴럴네트워크에 의한 패턴분류는 각 입력 데이터의 최소값과 최대값을 균등분할하여 전체 입력 공간 분할을 시행하였고, 공간 분할에 따른 전반부 멤버쉽함수의 정점을 각각 할당하였으며, 후반부 구조로는 간략추론을 사용하였다. 또한, 유전자 알고리즘을 적용하여 전반부 멤버쉽함수의 정점과 학습률, 모멘텀을 동조하였다. 표 1은 퍼지뉴럴네트워크 및 유전자 알고리즘의 파라미터를 보여준다.

제안된 네트워크를 이용하여 패턴분류를 수행하였으며, 표 2는 각 입력에 대해 멤버쉽함수의 수가 2개인 경우와 3개인 경우에 있어서 간략, 선형, 2차식 추론방법 및 학습에 의한 성능을 보여준다.

〈표 1〉 퍼지뉴럴네트워크 및 유전자 알고리즘의 파라미터

파라미터		값
Fuzzy Neural Network (FNN)	학습횟수	150
	멤버쉽함수 (MF)의 수	각 입력당 2개 / 3개
Genetic Algorithm (GA)	세대수	150
	개체수	60
	교배율	0.65
	돌연변이율	0.1

〈표 2〉 성능평가

(a) FNN의 성능평가			(b) GA에 의한 FNN의 성능평가		
FNN			GA+FNN		
MF 수	추론	오분류율	MF 수	추론	오분류율
2	간략	3.3	2	간략	0.6
	선형	3.3		선형	0.6
	2차식	2.6		2차식	0.6
3	간략	3.3	3	간략	0.6
	선형	2.6		선형	0.6
	2차식	3.3		2차식	0.6

표 3은 기존 모형과의 비교를 보여준다. 대표적인 분류 모형으로 의사 결정 나무(Decision tree) 모형과 로지스틱 판별(Logistic discriminant) 모형을 비교해 보면 제안된 네트워크가 더 좋은 성능을 보여준다.

〈표 3〉 기존 모형과의 비교

모형	오분류율
Decision tree	2.4
Logistic discriminant	3.3
Fuzzy Neural Network	0.6

5. 결 론

본 논문에서는 다중 출력을 가지는 퍼지 관계 기반 퍼지뉴럴네트워크를 설계하였다. 퍼지 규칙의 전반부는 전체 입력공간을 상호 관계에 의해 분할하여 표현하였으며, 규칙의 후반부는 다항식을 이용하여 오류역전파 알고리즘을 이용하여 각 다항식의 계수를 학습하였다. 실수코딩 유전자 알고리즘을 이용하여 각 파라미터를 자동 동조하였으며, 제안된 네트워크는 Fisher의 Iris 데이터를 이용하여 패턴분류에 적용하였다.

실험에서 알 수 있듯이 제안된 네트워크는 기존 모형에 비해 훨씬 좋은 성능을 보여주지만, 제안된 네트워크에서 퍼지 규칙의 특성상 입력 공간이 고차원으로 갈수록 차원의 저주로부터 자유로워지지 못하기 때문에 이에 대한 연구를 수행할 예정이다.

감사의 글

이 논문은 2007년도 정부(과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(F01-2007-000-10058-0).

참 고 문 헌

- [1] Takeshi Yamakawa, "A Neo Fuzzy Neuron and Its Application to System Identification and Prediction of the System Behavior", Proceeding of the 2nd International Conference on Fuzzy logic & Neural Networks, pp. 447-483, 1992.
- [2] D. E. Goldberg, Genetic Algorithm in search, Optimization & Machine Learning, Addison wesley, 1989.
- [3] 진강규, 유전알고리즘과 그 응용, 교우사, 2004.
- [4] 한학용, 패턴인식 개론, 한빛미디어, 2006.
- [5] 오성권, 프로그래밍에 의한 컴퓨터지능, 내하출판사, 2003.
- [6] J.-S. R. Jang, C. -T. Sun, E. Mizutani, Neuro-Fuzzy And Soft Computing, Prentice-Hall, 1997.