

퍼지 균등화와 유전알고리즘을 이용한 퍼지 모델링

김승석, 고현주, 전병석, 유정웅
 충북대학교 전기공학과

Fuzzy Modeling Using Fuzzy Equalization and GA

S. S. Kim, H. J. Go, B. S. Jun, J. W. Ryu
 School of Electrical and Electronic Engineering, ChungBuk National University

Abstract - In this paper, we proposed a method of modeling a system using Fuzzy Equalization(FE) and Genetic Algorithm(GA). The initial model is constructed using FE. The antecedent parameters and the rules in fuzzy logic are tuned by GA. The proposed system minimizes the modeling error and the size of structure. The process of building membership functions using PDF(Probability Density Function) and GA tunes the antecedent parameter and rules for minimizing the error and structure. The usefulness of proposed method is demonstrated by applying to Box-Jenkins furnace data.

공간상에서 취득된 수치 데이터가 있을 경우, 이들로부터 확률분포함수를 구한 후 전체 공간 X 에서 이들을 적절히 표현할 수 있는 소속함수들을 원하는 개수만큼 체계적으로 만들어 낼 수 있는 기법이다.

기존의 확률개념에서 사용되는 사건(crisp event)의 개념에서는 A 라는 사건과 이의 확률분포함수 $p(x)$ 가 주어졌을 경우 사건 A 의 확률은 다음과 같이 구해진다.

$$P(A) = \int_A p(x) dx \tag{1}$$

그러나 만약 사건 A 에 대하여 확률분포함수 $p(x)$ 와 더불어 소속함수 $A(x)$ 값이 주어진 퍼지 사건일 경우의 확률은 아래와 같다.

$$P(A) = \int_A A(x)p(x) dx \tag{2}$$

위와 같은 개념에 기초하여 전체 공간 X 에 대하여 확률밀도함수 $p(x)$ 가 주어져 있을 때, 이 공간을 퍼지 집합 $\{A_1, A_2, \dots, A_G\}$ 로 나누는 퍼지사건을 정의한다고 하자. 그러면 이 때, 가장 바람직한 분할 방법은 다음과 같이 각각의 퍼지 사건의 확률이 같도록 아래와 같이 균등화하는 방법이다.

$$P(A_1) = P(A_2) = \dots = P(A_G) = \frac{1}{G} \tag{3}$$

위의 개념을 적용하면 그림 1과 같이 데이터가 조밀한 부분과 그 반대의 경우와 소속함수의 배치가 확률밀도에 근거하여 생성되어 우리의 직관과도 잘 일치함을 알 수 있다.

이를 퍼지 균등화 알고리즘으로 표현하면 다음과 같다.

[단계1] A 에서 소속함수의 수 c 를 정한다.

[단계2] X 의 최저 영역을 x_{min} 을 한다.

[단계3] 다음 식(4)이 만족하는 a 값을 구한다.

$$\int_{x_{min}}^a A_1(x)P(x)dx = \frac{1}{2c} \tag{4}$$

[단계4] 소속함수의 감소부분에 퍼지 사건확률이 다음 식(5)를 만족하는 A_1 지지 상한 영역 b 를 구한다.

$$\int_a^b A_1(x)P(x)dx = \frac{1}{2c} \tag{5}$$

[단계5] A_2 를 시작으로 A_3, A_4 로 가는 삼각 퍼지 집합의 증가 부분에 퍼지 사건의 확률을 계산한다.

$$\epsilon = \int_a^b A_2(x)P(x)dx \tag{6}$$

[단계6] 조건을 만족하는 퍼지 집합의 지지 상한 영역을 결정하여 소속함수의 감소부분을 최적화하여 c 를 구한다.

$$\int_b^c A_2(x)P(x)dx = \frac{1}{c} - \epsilon \tag{7}$$

[단계7] 소속함수를 모두 구할 때 까지 [단계5]와 [단계

1. 서 론

일반적인 퍼지 모델링은 정확한 시스템의 파라미터를 필요로 하지 않으며 전문가 지식이나 또는 경험적 지식을 이용하여 퍼지 규칙과 소속함수를 결정할 수 있다. 그러나 경험적 지식 등은 불안정하거나 체계적이지 못하며 또한 여러 가지 문제점들을 발생시킬 수 있다.

이런 문제점을 해결하기 위하여 수치적인 입력력 데이터를 근거한 퍼지 규칙 생성에 대하여 다양한 연구가 시도되었다. 대표적인 방법으로 뉴로-퍼지 시스템 등이 있다.

이러한 추론 시스템은 대상 시스템의 모델링을 구조동정과 파라미터동정으로 나누며 구조동정은 적당한 규칙과 소속함수의 수, 입력공간 분할을 수행하며, 파라미터 동정은 전제부 및 결론부 파라미터의 조정을 수행한다. 그러나 구조동정에 대한 체계적 방법이 없으므로 실제 시스템 적용에는 많은 문제점을 발생시키고 있다. 일반적인 입력공간 분할 방법[1]으로는 그리드 분할, 트리 분할, 스캐터 분할 등이 있다. 그리드 분할은 가장 직접적인 방법이지만 입력 공간이 증가할 때 규칙의 수가 지수적으로 증가하는 문제점을 가지고 있으며, 트리 분할은 각각의 영역에 대응하는 결정 트리에 따라 특징화 되도록 입력공간을 분할하지만 그리드 분할과 유사한 문제점을 가지고 있다. 또 스캐터 분할은 클러스터의 수와 중심값의 초기 선택에 따라 결과가 민감하게 변동하는 문제점이 있다.

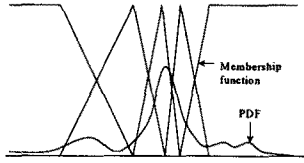
따라서 본 논문에서는 위의 문제점들을 해결하기 위하여 Pedrycz에 의해 제안된 확률밀도함수(PDF)에 의해 소속함수를 생성하는 퍼지 균등화[2]와 적응적 최적화에 많이 응용되는 유전알고리즘[3]을 이용하여 퍼지 모델링을 하였다. 자동적인 퍼지 규칙을 추출하기 위해서 제안된 방법을 Box-Jenkins의 가스르 데이터 모델[4]에 적용하여 유용성을 설명한다.

2. 퍼지 균등화 (Fuzzy Equalization)

2.1 퍼지 균등화에 의한 소속함수 생성

퍼지 균등화는 임의의 전체 공간 X 가 주어져 있고 이

6)을 반복한다.



2.2 전제부 파라미터 변동 범위

범위 설정은 미리 정해진 소속함수의 수를 기준으로 ± 1 의 소속함수 수를 구한 뒤 각 파라미터의 가장 근접한 파라미터를 변동 범위로 정한다. 이는 확률 밀도 함수에 의해 정해진 파라미터는 소속함수의 수가 ± 1 일 경우, 각 소속함수에서 a,b,c의 값이 서로의 영역을 넘는 것을 방지하기 위해서이다. 즉 $a > b$ 이거나 $b > c$ 의 영역을 초과하는 것을 방지한다.

3. 유전알고리즘

적자 생존의 자연계의 진화과정을 모사한 유전알고리즘은 1975년에 John Holland에 의해 개발된 전역적인 최적화 알고리즘이다. 최적화 문제를 정해진 형태의 자료 구조로 표현하고 유전자들을 점차적으로 변형을 하여 세대를 더해 가면서 더 나은 최적화를 이루고자 하는 알고리즘이다.

3.1 유전알고리즘의 기본 연산자

재생산(Representation)은 적자 생존 또는 자연 도태 현상을 모방하여 적합도에 의해 다음 세대로의 약한 개체들을 몰아내고 강한 개체들을 선택하게 하는 알고리즘으로 룰렛 휠 선택 (roulette wheel selection), 순위에 기초한 선택 (ranking-based selection), 토너먼트 선택 (tournament selection) 등이 있다. 즉 적합도가 높은 연산자에 선택을 유리하게 함으로써 다음 세대에 더 나은 결과를 기대한다.

교배 (Crossover)는 탐색공간 상의 가능한 새로운 점을 찾기 위하여 부모 세대로부터 각 염색체를 서로 교환하여 더 좋은 결과를 기대하는 연산자로 일점 교배, 다점 교배, 순환 교배, 균등 교배 등이 있다.

돌연 변이 (mutation)은 알고리즘 중에 국소 최적점 등에 조기에 수렴하는 것을 방지하기 위한 연산자이다.

3.1 엘리티즘(Elitism)

한 세대에서 최적 개체를 다음 세대에서도 생존하도록 보장해주는 방법으로 비교적 강한 선택 메커니즘이다. 이 전략은 이전 세대의 최적 개체를 저장하였다가 재생산, 교배, 돌연변이 등을 거친 후 적합도가 가장 낮은 개체와 교환하는 방법으로 유전알고리즘의 성능을 개선한다.

4. FE와 GA를 이용한 퍼지 시스템

먼저 적용된 퍼지 시스템은 다음과 같은 TSK 퍼지 시스템이다.

$$R^i: \text{If } x \text{ is } A_i \text{ and } y \text{ is } B_i, \text{ Then } f_i = p_i x + q_i y + r_i \quad (8)$$

여기서 x 와 y 는 입력이고, A_i, B_i 는 언어 변수이며 (p_i, q_i, r_i) 는 결론부 파라미터이다.

초기 소속함수의 수에 의한 전제부 파라미터는 퍼지 균등화에 의해 정해지며 다시 퍼지 균등화에 의해 전체부 파라미터 조정 범위가 정해진다. 결론부 파라미터는 유전 알고리즘에 의해 정해진 전제부 소속함수가 정해지면 최소자승법에 의해 추정된다.

적합함수는 모델 형성에 사용되는 학습데이터와 모델

검증에 사용되는 검증 데이터를 모두 고려하여 다음 식과 같이 표현한다.

$$f = \frac{1}{\omega_1 \times MSE_{training} + \omega_2 \times MSE_{checking}} \quad (9)$$

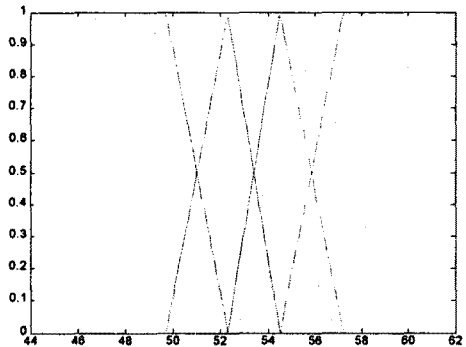
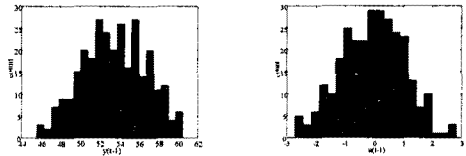
여기서 ω_1, ω_2 는 각 가중치이며, $MSE_{training}$ 은 학습 데이터에 대한 오차이며, $MSE_{checking}$ 은 검증 데이터에 대한 오차이다.

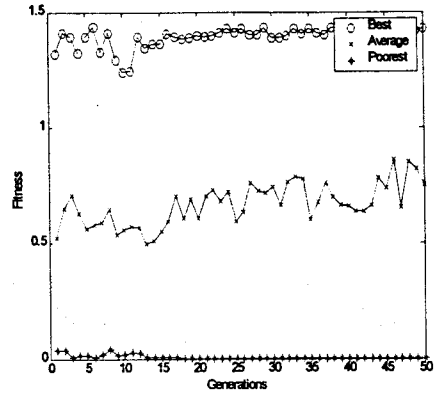
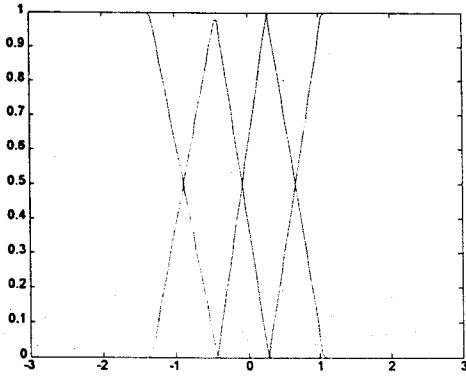
5. 시물레이션

본 논문에서는 제안된 방법을 이용하여 Box-Jenkins의 가스도 데이터 모델에 적용하였다.

Box-Jenkins 가스도 데이터는 가스 흐름률 $u(t)$ 와 이산화 탄소 농도 $y(t)$ 시계열 데이터인 296쌍이 사용된다. 입력으로는 $y(t-1)$ 과 $u(t-4)$ 으로 선택하였으며 출력으로는 $y(t)$ 를 선택하였다. 데이터 쌍을 각각 훈련 및 검증 데이터로 145개의 쌍으로 사용하였다.

적합함수로는 $f = \frac{1}{\omega_1 * MSE_{trn} + \omega_2 * MSE_{chk}}$ 를 사용하여 훈련데이터 오차와 검증 데이터 오차를 동시에 고려하여 학습을 하였다. 학습만을 강조한 경우 $MSE_{training}$ 은 0.076이나 $MSE_{checking}$ 은 1.56이 되어 새로운 데이터에 대해 예상치 못하는 결과를 얻을 수 있다. 또한 유전알고리즘에 의해 규칙의 수가 14개로 선택되어 실험한 결과 $MSE_{training}$ 은 0.095이고, $MSE_{checking}$ 은 0.922이었으며 유전알고리즘에 의해 규칙의 수가 15개로 선택을 경우 $MSE_{training}$ 은 0.096이고, $MSE_{checking}$ 은 0.915으로 검증의 비율을 증가시킴으로서 $MSE_{training}$ 은 약간의 감소를 보였





6. 결 론

본 논문은 퍼지 균등화에 의한 초기 전제부 소속함수 생성 및 조정 범위를 정하고 유전알고리즘을 통하여 전제부 파라미터를 생성하고 결론부 파라미터에 영향을 주지 않는 규칙을 제거하는 TSK 퍼지 추론 시스템을 생성하였다. 이렇게 제안된 방법은 초기에 생성된 전제부 파라미터를 미세 조정하여 좀더 최적의 값을 유도할 수 있었으며 불필요한 규칙을 제거하여 그리드 분할형태의 규칙의 수가 지속적으로 증가하는 문제점을 해결하였다.

시뮬레이션의 예로 Box-Jenkins 가스로 데이터의 모델링에 적용하여 실험하였으며 검증 데이터의 비율이 높아질수록 학습 데이터에 대한 오차가 상대적으로 증가함을 볼 수 있다.

[참 고 문 헌]

- [1] J. S. Jang, "Neuro-Fuzzy and Soft Computing", Prentice Hall, 435-497, 1997
- [2] W. Pedrycz, "Fuzzy Equalization in the construction of fuzzy sets", Fuzzy Sets and Systems, 119, 325-335, 출판년도
- [3] Z. Michalewicz, "Genetic ALgorithms+Data Structures=Evolution Programs", Springer, Third, Revised and Extended Edition, 페이지, 1995
- [4] M. G. Chun, K C. Kwak, J. W. Ryu, W. Pedrycz, "A fuzzy rule extraction method for ANFIS using CFCM and fuzzy equalization", Journal of Advanced Computational Intelligence, Vol 4, No. 5, 2000

