

트리구조 퍼지시스템

정창호¹, 강성훈², 박주영³, 박대희⁴

¹고려대 대학원 전산학과, ²고려대 대학원 정보공학과,
³고려대 제어계측공학과, ⁴고려대 전산학과

Tree-Structured Fuzzy System

Changho Jung¹, Seonghoon Kang², Jooyoung Park³, Daihee Park⁴

ABSTRACT

Conventional fuzzy systems have serious problems in dealing with the nonlinear approximations on high-dimensional spaces due to the explosive increase of the number of fuzzy IF-THEN rules. In order to avoid such problems, this paper proposes a tree-structured fuzzy system in which semi-local basis functions form its basic elements, and develops a training algorithm for the proposed system based on the evolution program and LMS rule. The experimental studies demonstrate the effectiveness of the developed methodology.

I. 서론

퍼지 논리는 기존의 논리 체계보다 인간 사고의 기본적인 언어적 정보를 보다 적절하게 표현할 수 있으므로, 실세계의 근사적이고 불확실한 현상을 기술하는데 효과적으로 이용될 수 있다. 이러한 퍼지 논리에 기초한 방법론은 시스템의 특성이 복잡하여 기존의 정량적인 방법으로는 해석할 수 없거나, 획득된 정보가 정성적이고 부정확하며 불확실한 경우에 우수한 성능을 보인다. 퍼지 시스템을 설계하는 과정은 구조식별(structure identification)과 매개변수 조정(parameter tuning)의 두 단계로 나눌 수 있다^[7,10]. 구조식별 단계에서는 시스템의 입출력 변수를 가려내고 입출력 관계를 퍼지 규칙 베이스로 표현한다. 그리고, 매개변수 조정단계에서는 목표 시스템과의 오차를 줄이기 위해 각 매개변수를 조정한다.

퍼지 시스템의 설계에 있어서 퍼지 소속함수의 개수 및 형태를 결정하는 일, 퍼지 규칙의 개수 및 조합을 결정하는 일, 그리고 규칙 결론부의 실수 함을 정밀하게 조정하는 일 등은 많은 시간과 노력이 요구되었다. 이러한 단점은 기본적으로 학습 기능(learning mechanism)의 결여에 의해 파생된다고 할 수 있겠는데, 최근에는 학습 능력을 갖고 있는

신경망 이론과 유전자 알고리즘을 이용하여 서로의 장점을 결합한 융합(fusion) 연구가 활발하게 이루어지고 있다^[1,3,4]. 그러나 함수근사가 고차원 공간에서 이루어질 경우 발생하는 퍼지규칙 개수의 폭발적 증가문제는 기존의 융합 연구로는 대처하기 힘든 측면을 갖는다.

본 논문에서는 함수 공간 $C[-L, L]^r$ 의 근사를 위해 초기 단계에서는 변수 x_1, \dots, x_r 중 하나에만 의존하는 베이스 함수의 일차 결합을 사용하고, 학습이 진행되는 과정에서 그 구조 및 매개변수가 적응적으로 변하는 기존의 반국소 베이스에 필요에 의해 새로운 변수가 추가되어 점차로 다변수에 의존하는 트리 형태의 베이스 함수가 형성되도록 한다. 그리고 트리 형태의 형성에 있어서 반국소 베이스 함수의 모양과 연결 강도는 진화 프로그래밍과 LMS(least mean square) 규칙을 이용하여 결정한다.

본 논문의 순서는 다음과 같다: 2장에서는 본 논문을 통해 제안되는 트리구조 퍼지시스템을 자세히 설명한다. 그리고, 3장에서는 시계열 예측 문제를 위한 시뮬레이션을 통해 그 결과를 관련 연구와 비교 분석한다. 마지막으로 4장에서는 결론을 제시한다.

II. 트리구조 퍼지시스템

본 논문에서는 반국소 유닛(semi-local units)을 기본으로 하는 트리구조 퍼지 시스템을 제안하고 이를 효과적으로 학습하기 위하여 진화 프로그램을 활용하고자 한다.

1) 초기 트리의 구조

본 논문에서 다루는 트리구조 시스템에서는, 고차원 공간 위에서 정의된 베이스 함수가 여러 변수를 동시에 포함할 경우에는 필요한 함수의 수가 매우 커질 수밖에 없음을 착안하여, 전체 시스템이 초기에는 깊이가 1 이고 각 베이스 함수가 하나의 변수에만 의존하는 초기 형태를 취하다가 학습이 진행되어 감에 따라 트리가 확장되어 가는 방식을 취하도록 한다. 따라서, 초기 트리에서는 우리가 근사하고자 하는 함수의 출력 y 를, 다음의 (1)식에서 보여진 바와 같이 하나의 변수에만 의존하는 반국소 베이스 함수의 일차 결합 $y[1]$ 로 근사한다.

$$y[1] = \sum_{i=1}^r \sum_{j=1}^{J_i} w_{ij} \phi_{ij}(x_i) \quad (1)$$

(1) 식에서 소개된 깊이 1의 구조는 베이스 함수 ϕ_{ij} 의 일차 결합 꼴이므로, 일정 조건 아래서 LMS 규칙을 사용하면 최적 값으로 수렴해 가는 연결 강도 w_{ij} 를 얻을 수 있다. 여기에서 제시하는 트리구조의 원형이라고 할 수 있는 Sanger의 LMS 트리^[6]에서는 비슷한 구조로부터 트리 형성이 시작되는데, 이때 베이스 함수 $\phi_{ij}(x_i), j=1, \dots, J_i$ 로는 일정한 간격으로 이동되어진 가우시안, sin 등의 함수가 사용된 바 있다. 그러나, LMS 트리에서와 같이 각 축마다 미리 정해진 베이스 함수(fixed basis functions)를 사용하는 경우에는 각 베이스 함수의 모양이나 개수를 미리 결정하는 작업이 선행 되어져야 한다.

본 논문에서 제안하는 트리구조 시스템은 초기의 트리구조가 (1)식의 응답 특성으로부터 출발하되, 각 베이스 함수의 모양, 개수 및 연결 강도를 진화 프로그램을 통해 최적에 가깝게 결정하는 특징을 갖는다. 즉, 각 변수 x_i 에 관한 베이스 함수 ϕ_{ij} 가 미리 정해지지 않고, 진화 프로그램을 이용한 학습 과정 중 평가 함수를 극대화하는 방향으로 결정되는 것이다.

지금까지 소개한 깊이 1의 초기 트리가 가우시안 베이스 함수를 갖는 경우에는, 출력 노드의 응답은 다음 식에서 보여지는 바와 같이 가우시안의 가중평균(weighted average)으로 정의할 수 있다:

$$y[1] = \frac{\sum_{i=1}^r \sum_{j=1}^{J_i} w_{ij} \exp[-\frac{1}{2}(\frac{x_i - \mu_{ij}}{\sigma_{ij}})^2]}{\sum_{i=1}^r \sum_{j=1}^{J_i} \exp[-\frac{1}{2}(\frac{x_i - \mu_{ij}}{\sigma_{ij}})^2]} \quad (2)$$

즉 (2)식은 결론부가 실수값을 갖는 단순 퍼지 추론 모델(simplified fuzzy reasoning model)의 입출력식이다.^[8]

다음에서는, 가우시안 베이스 함수를 이용한 트리구조 시스템에 대하여, 베이스 함수의 개수 J_i , 가우시안 베

이스의 중심(centroids) μ_{ij} , 유연계수(smoothing factors) σ_{ij} , 및 연결 강도 w_{ij} 가 진화 프로그램에 의하여 최적에 가깝게 결정되는 과정을 자세히 설명한다.

2) 진화 프로그램을 이용한 학습

(2)식으로 표현되는 초기 트리에 있어서 우리가 결정해야 할 부분은 $J_1, \dots, J_r, \mu_{ij}, \sigma_{ij}, w_{ij}, i=1, \dots, r, j=1, \dots, J_i$ 이다. 본 논문에서는 이들 중 연결 강도 값 w_{ij} 를 제외한 전부분을 베이스 함수의 기하학적 특성에 따라 일정 길이 L의 r배인 rL 비트의 이진 스트링(binary string) 개체로 부호화한다. 이들 중 (i+1)번째 자리부터 (i+L)번째 자리까지의 스트링은 매개변수 $J_i, \mu_{ij}, \sigma_{ij}, j=1, \dots, J_i$ 를 결정하는데 사용된다.

다음에서는 깊이 1의 초기 트리에 대하여, 학습 데이터에 대한 오차와 일반화 능력(generalization capability)을 반영한 평가 함수를 최대화시키는 탐색 과정을 설명한다. 이 탐색 과정에는 기존의 유전자 알고리즘보다 우수한 성능을 가진 것으로 알려진 진화 프로그램^[9]을 사용한다. 개체의 분류는 적합도의 크기에 따라 행해지는데, 본 논문에서 사용되는 개체 s_m 의 적합도 $E(s_m)$ 를 다음과 같이 정의하였다:

$$E(s_m) = \frac{1}{\alpha_E(\text{학습오차}) + \alpha_B(\text{베이스 함수의 개수})} \quad (3)$$

이상에서 설명한 진화 프로그램을 초기 트리의 학습에 적용하면 다음과 같은 계산 순서가 얻어진다:

• 진화 프로그램을 이용한 트리의 학습

- [1] 이진 난수 발생에 의해, 길이 rL을 갖는 이진 스트링 개체들로 이루어지는 초기 집단 $S(t) = \{s_1(t), \dots, s_M(t)\}$ 를 구성함($t=0$).
- [2] 집단 내의 각 개체가 갖는 이진 스트링을 앞에서 설명한 방법에 따라 베이스 함수로 변화시킴.
- [3] 학습 데이터를 이용하여 각 개체가 표현하는 트리 구조 시스템을 LMS 규칙으로 학습시킴.
- [4] LMS 규칙에 따른 학습이 미리 정해진 횟수에 도달하거나, 학습 오차의 변화량이 정해진 값 이하가 되면, LMS 규칙에 의한 학습을 종료하고 단계 [5]로 넘어감.
- [5] 집단 내의 각 개체 $s_m(t), m=1, \dots, M$ 에 대하여 적합도 $E(s_m)$ 을 계산.
- [6] 각 개체가 갖는 적합도를 이용하여 개체 집단 $S(t)$ 의 개체들을 확률적 경쟁 선택 방법(stochastic tournament selection)방법으로 우수 개체를 구성하고 부모 개체, 소멸 개체로 분류.
- [7] 부모 개체에 대하여 다중 돌연변이 연산을 수행하여 자식 개체들을 생성하고 여기에 우수개체를 복제하여 추가함으로써, 다음 세대의 개체 집단 $S(t+1)$ 를 구성함.
- [8] 정해진 반복 세대수에 이를 때까지 단계 2에서 단계 7까지를 반복함.

위의 과정 중 단계 [2]에서는 각 개체에 해당하는

$$J_1, \dots, J_r, \mu_{11}, \dots, \mu_{rj}, \sigma_{11}, \dots, \sigma_{rj}$$

를 결정하게 되고, 단계 [3]-[4]에서는 LMS 규칙에 의한 학습을 통해 연결 강도 w_{ij} 를 구한다. 이같은 과정을 거치면 전체 시스템의 구조와 매개변수가 모두 결정된다. 각 개체가 나타내는 트리구조 시스템이 갖는 적합도는 식 (3)에서와 같이 학습 오차와 베이스스 함수의 개수를 반영하여 결정하는데, 이 값은 진화 프로그램의 개체 분류 과정인 단계 [6]에서 사용된다. 이상에서 설명한 학습 과정을 통하여 최고의 적합도를 갖는 개체를 선택할 경우에는, 적합도의 정의가 갖는 특성으로부터 학습 데이터에 대한 오차의 최소화와 overfitting의 방지를 동시에 기할 수 있다.

3) 트리의 확장

앞부분에서 설명한 방법을 통하여 얻은 깊이 1의 트리구조 시스템의 성능이 만족스럽지 못할 경우에는 연결 강도에 대한 변화 압력이 가장 큰 베이스스를 따라 트리를 확장시킨다. 베이스스 함수 ϕ 와 출력을 잇는 연결 강도가 갖는 변화 압력은 LMS 규칙을 참조하여 다음과 같이 정의한다. (4)식에서 위첨자 l 는, l 번째 학습 데이터를 의미하는 index이다.

$$\sum_{j=1}^n |(y^l - \hat{y}^l) \phi(x^l)|^2 \quad (4)$$

여기서 새롭게 확장된 트리구조 시스템은

$$\begin{aligned} &(\text{기존의 트리구조}) + (\text{변화 압력이 가장 큰 베이스스}) * \\ &(\text{새롭게 추가된 부트리}) \end{aligned} \quad (5)$$

형태의 식으로 표현될 수 있다. (5)식을 전개하면, 각 기본항은 1개 이상의 변수를 포함하는 가우시안 형태를 띠게 된다.

이때 새롭게 추가되는 부트리(subtree)는, 초기 트리와 마찬가지로 가우시안의 일차 결합(linear combination) 형태를 띠게 된다. 따라서 rL 비트로 이루어지는 이진 스트림을 이용하여, 새롭게 추가되는 부트리가 갖는 베이스스 함수의 모양과 개수를 나타낼 수 있다. 추가된 부트리 부분을 최적화 시키는 과정은, 초기 트리를 최적화 시키는 데 사용되었던 알고리즘이 사실상 그대로 적용될 수 있다. 다만 단계 3에서 등장하는 “각 개체가 표현하는 트리구조 시스템”은 깊이 1의 초기 트리 혹은 새롭게 추가되는 부트리를 의미하는 대신에, 식 (5)에서 설명된 전체 시스템으로 해석되어야 한다. 그리고, 전체 트리구조 시스템의 출력 노드가 갖는 응답은 (2)식에서와 같이 전체 트리구조가 갖는 기본항 즉 가우시안 베이스스에 대한 가중평균 형태로 얻어진다. 이와 같은 트리 확장은, 주어진 데이터를 이용한 학습 과정에서 주기적으로 적용하되 전체 시스템의 학습 오차가 충분히 만족스러운 값을 갖게 될 때까지 계속한다.

III. 시뮬레이션

식 (6)의 지연 미분 방정식의 해로부터 얻어지는 시계열 데이터의 예측 문제인 Mackey-Glass time-series는, 예측 알고리즘의 성능 평가를 위한 주요한 적용 문제의 하나로

$$\text{인식되어 왔다: } \frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \quad (6)$$

본 장에서는 트리구조 시스템의 성능 분석을 위하여, τ 가 30일 때 Mackey-Glass 시계열의 입력 벡터 $\{x(t), x(t-1), \dots, x(t-8)\}$ 로부터 다음 시간 스텝의 출력값 $x(t+1)$ 을 예측하는 문제를 생각한다. 미분방정식 (6)의 해는 4차 Runge-Kutta 방법을 이용한 수치 해석적 풀이로부터 얻어졌으며, 이 과정에서 사용된 시간 단계(time step)와 초기 데이터 $x(0)$ 값으로는 각각 0.1과 0.8이 사용되었다. t 가 100일 때부터 1099일 때까지 1000개의 입출력 쌍이 본 논문의 트리구조 시스템 및 비교 대상 방법론을 위한 학습 및 시험용의 데이터로 추출되었다. 이들 중 700개의 입출력 쌍들이 학습용 데이터(training data)로 사용되었으며 300개의 입출력 쌍들이 시험용 데이터(checking data)로 사용되었다. 본 논문의 방법론을 적용한 실험에서 사용한 주요 파라미터들은 표 1과 같다.

표 1. 트리구조의 학습에 사용된 주요 파라미터

최대 세대 반복 횟수	15
개체 집단의 크기	15
부모 개체의 개수	10
각 변수에 할당된 유전인자의 비트수	10
돌연변이 연산자의 확률	0.04
교배 연산자의 확률	0.6
LMS 규칙을 위한 학습률	0.1
LMS 규칙을 위한 최대 학습 반복 횟수	100
학습 오차를 위한 가중치	1
베이스스 개수를 위한 가중치	0.001
수목확장 횟수	9

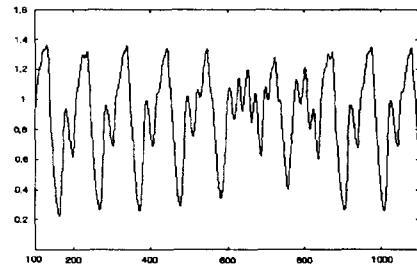


그림 2. 학습 및 시험용 데이터를 추출하는 데 사용한 곡선

그리고, 성능 비교의 대상으로는 같은 문제를 다룬 [8,9]의 퍼지 시스템이 선택되었다.

그림 2에서는 학습 및 시험용 데이터들을 추출하는 데 사용한 곡선이 보여졌다. 그리고, 트리구조 퍼지시스템과 비교 대상으로 선택된 퍼지 시스템이 제공하는 예측 능력이 각각 그림 3과 4를 통해 요약되었다. 예측 결과들과 그림 2의 실제 데이터 사이에는 매우 미세한 차이만이 존재하므로, 이들의 성능을 시각적으로 비교하는 데에는 이들 그림에서와

같이 잔류오차(residual error)를 이용하는 것이 효과적이다. 그리고, 전반적인 예측 성능을 하나의 지수로 나타낼 경우에는 다음에서 정의하는 NMSE(normalized mean square error) 값이 널리 쓰인다.

$$NMSE = \frac{E[(\hat{x}(t) - x(t))^2]^{1/2}}{E[(x(t) - E(x(t)))^2]^{1/2}} \quad (7)$$

이들 그림은, 트리구조 시스템이 학습 및 시험 데이터 구간에서 모두 다 Wang과 Mendel에 의해 제안된 방법 [8,9]보다 작은 크기의 NMSE와 현격하게 저하된 잔류오차를 제공함을 보여준다(그림 3의 tm NMSE = 0.0445, chk NMSE = 0.0475, 그림 4의 tm NMSE = 0.0845, chk NMSE = 0.0792). 따라서, 본 논문에서 제안한 트리구조 퍼지시스템은, 고차원의 입력공간을 갖는 비선형 근사 문제를 다룸에 있어서 주어진 입력력 데이터를 기억하는 학습 능력과 학습에 사용되지 않은 입력이 들어왔을 때 출력을 정확하게 예측할 수 있는 일반화 능력(generalization capability)을 고루 갖추고 있는 방법론임을 알 수 있다.

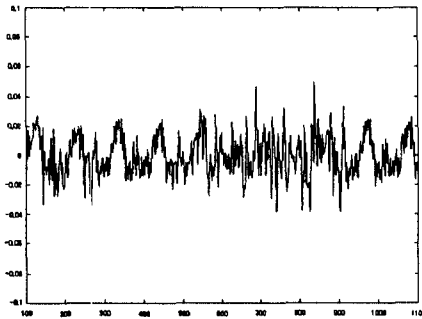


그림 3 수목구조 퍼지 시스템을 이용할 때의 잔류오차와 NMSE 값

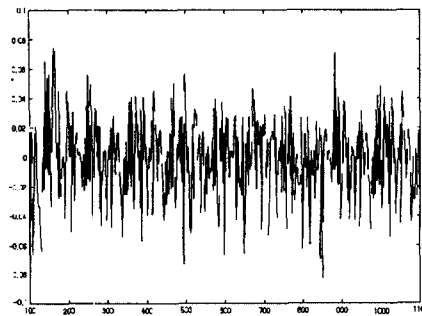


그림 4. [8,9]의 퍼지 시스템을 이용할 때의 잔류오차와 NMSE 값

IV. 결론

본 논문에서는, 고차원 입력 공간 위의 비선형 근사를 효과적으로 수행하기 위하여 1) 학습이 진행됨에 따라 구조가 확장되는 트리구조 지능 시스템을 제안하고, 2) 진화 프

로그래프와 LMS 규칙을 이용하여 트리 확장 및 학습을 조절하는 방법론을 제안하였다. 마지막으로 개발된 방법론이 본 논문에서 목표로 삼는 실제적 문제, 즉 고차원의 입력 변수들이 어느 정도 중복성이 있는 정보를 취하는 경우에 충분한 효과를 발휘하는지를 살펴보기 위하여, Mackey-Glass 시계열 예측 문제를 대상으로 시뮬레이션을 행하였다. 그 결과, 본 논문의 방법론이 주어진 문제를 효과적으로 다룰 수 있음을 알 수 있었다. 앞으로 시도해야 할 향후 연구 과제로, 학습 과정에서 구조를 바꾸는 여타 방법론에 대한 연구 및 비교 분석으로 추가적인 성능 향상을 기하는 것이 필요하다.

참고문헌

- [1] M. A. Lee and H. Takagi, "Integrating design stages of fuzzy systems using genetic algorithms," Second IEEE International Conference on Fuzzy Systems, vol. 1, pp. 612-617, 1993.
- [2] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Second, Extended Edition, Springer-Verlag, New York, 1994.
- [3] K. C. Ng and Y. Li, "Design of sophisticated fuzzy logic controllers using genetic algorithms," Proceedings of Third IEEE International Conference on Fuzzy Systems, pp. 1708-1712, 1994.
- [4] D. Park, A. Kandel, and G. Langholz, "Genetic-based new fuzzy reasoning models with application to fuzzy control," IEEE Transactions on Systems, Man, and Cybernetics, vol. 24, no. 1, pp. 39-47, 1994.
- [5] T. D. Sanger, "A tree-structured algorithm for reducing computation in networks with separable basis functions," Neural Computation, vol. 3, no. 1, pp. 67-78, 1991.
- [6] T. D. Sanger, "A tree-structured adaptive network for function approximation in high-dimensional spaces," IEEE Transactions on Neural Networks, vol. 2, no. 2, pp. 285-293, 1991.
- [7] M. Sugeno and T. Yasukawa, "A fuzzy logic based approach to qualitative modeling," IEEE Transactions on Neural Networks, vol. 1, no. 1, pp. 7-31, 1993.
- [8] L. Wang, *Adaptive Fuzzy Systems And Control: Design And Stability Analysis*, Prentice Hall, Englewood Cliffs, New Jersey, 1994.
- [9] L. Wang and J. M. Mendel, "Generating fuzzy rules from numerical data, with applications," IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, no. 6, pp.1414-1427, 1992.
- [10] R. Yager, "Structures for building flexible neural fuzzy systems," IFSA '93, Plenary Talk, 1993.