진화전략을 이용한 뉴로퍼지 시스템의 학습방법

정 성 훈
한성대학교, 정보전산학부
전화 : 02-760-4344 / 핸드폰 : 016-203-5815

Training Algorithms of Neuro-fuzzy Systems Using Evolution Strategy

Sung Hoon Jung
School of Information and Computer Engineering, Hansung University
E-mail : shjung@hansung.ac.kr

Abstract

This paper proposes training algorithms of neuro-fuzzy systems. First, we introduce a structure training algorithm, which produces the necessary number of hidden nodes from training data. From this algorithm, initial fuzzy rules are also obtained. Second, the parameter training algorithm using evolution strategy is introduced. In order to show their usefulness, we apply our neuro-fuzzy system to a nonlinear system identification problem. It was found from experiments that proposed training algorithms works well.

I. 서론

뉴로퍼지 시스템은 신경망 구조에 퍼지추론 기능을 합성한 것으로서 지능제어, 패턴인식 및 시계열예측 등 많은 분야에 적용되어 왔다 [1,2,4]. 뉴로퍼지 시스템의 성능에 있어서 중요한 것은 구조적으로 몇 개의 중간 노드를 갖느냐와 (즉 몇 개의 퍼지규칙으로 표현하느냐) 학습데이터로부터 어떻게 효율적으로 학습하느냐에 있다. 기 발표된 논문에서는 뉴로퍼지 시스템의 구조학습이나 파라메터 학습 시 주로 수정된 오류역전파 알고리즘이나 유전자 알고리즘 등이 많이 사용되어 왔다. 그러나 수정된 오류역전파 알고리즘은 학습시간이 많이 걸리고 학습률 (learning rate)를 실험적으로 설정해야하는 등의 단점이 있으며 유전자 알고리즘은 구조학습에는 적합하나 파라메터 학습시 유전형 (genotype) 개체로 표현함으로 세부적인 탐색이 어려운 문제점이 있었다 [1,4].

본 논문에서는 구조학습을 위하여 학습데이터의 연관관계를 이용한 알고리즘을 소개하며, 최적화 기법중 하나인 진화전략 (evolution strategy)를 사용한 파라메터 학습 방법을 제안한다. 진화전략을 이용한 학습방법은 해를 표현형 (phenotype)의 실수 개체로 표현함으로 유전자 알고리즘을 사용한 방법보다 세부적인 탐색이 가능한 장점이 있으며 수정된 오류역전파 알고리즘보다는 학습시간이 빠른 장점이 있다. 제안한 진화전략 학습방법의 성능을 보이고자 간단한 비선형 시스템 동정에 응용한 결과 학습시간이나 학습결과에 있어서 우수함을 확인하였다.

II. 뉴로퍼지 시스템

제안한 뉴로퍼지 시스템의 구조가 그림 1에 나와있다. 그림 1에서 보듯이 제안한 뉴로퍼지 시스템은 크게 뉴로퍼지 시스템모듈 (NFSM: Neuro-fuzzy System Module) 과 진화전략 학습모듈 (ESTM: Evolution Strategy Training Module)로 구현되어 있다. 뉴로퍼지 시스템 모듈은 3층 신경망의 구조를 갖으며 각각의 입출력 링크는 하나의 퍼지 언어항이 할당된다.

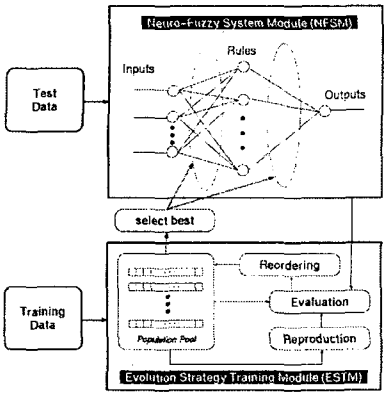

그림 1. 뉴로퍼지 시스템의 구조

중간노드 (다른 말로 규칙 노드)와 출력노드에서는 퍼지 추론의 t-norm 연산과 t-conorm 연산을 수행하며, 출력노드는 또한 최종 출력 전에 추론된 퍼지 출력 값을 비 퍼지화 한다. 본 논문에서는 t-norm 연산으로 min 연산을 t-conorm 연산으로는 max 연산을 사용하였다. 또한 비 퍼지화 동작으로 LGM (level-grading method) 방법을 사용하였다 [3]. LGM은 max 연산과 비 퍼지화 연산을 합성하여 계산함으로 결국, 중간노드는 min 연산 출력노드는 LGM 비 퍼지화 연산을 수행하게 된다. 뉴로퍼지 시스템의 동작은 알고리즘 1과 같다.

알고리즘 1 (뉴로퍼지 시스템 동작)

PP : N개의 부모개체와 N개의 자식개체로 이루어진 개체군

f : 어떤 세대에서 최대 적합도

f_t : 학습 종료 판단용 적합도

1. 각 모듈을 초기화하고 학습데이터를 설정
2. NFSM 을 초기화
3. ESTM 을 초기화
4. 학습 데이터를 설정
5. ESTM을 이용하여 학습
6. **Do**
7. N개의 부모 개체를 평가
8. N개의 자식 개체를 생성하기 위하여 N개의 부모 개체를 돌연변이 시킴
9. N개의 자식 개체를 평가하고 PP 에 삽입
10. 개체군 PP에서 적합도가 좋은 N개를 선택하여 다음세대의 부모개체로 설정
11. **Until** $f < f_t$
12. 최대 적합도를 갖는 개체의 파라메터를 NFSM에 설정
13. 테스트 데이터를 NFSM에 적용하여 출력을 발생
14. 테스트 데이터를 NFSM에 적용
15. NFSM의 출력을 얻음
16. NFSM의 출력을 이용하여 응용에 따라 사용

뉴로퍼지 시스템의 구현에 있어서 중요한 것 중의 하나가 중간노드의 수를 결정하는 것이다. 본 논문에서는 중간노드의 수를 결정하기 위한 구조학습으로 알고리즘 2를 사용하였다.

알고리즘 2 (구조학습 방법)

r_j : j번째 규칙

k : 생성된 규칙 개수

t_i : i번째 학습 데이터

N : 총 학습 데이터

m_{ij} : t_i 와 r_j 사이의 정합 정도

m^d : 정합판단 값

$M(\cdot, \cdot)$: 정합 함수

$R(t_j)$: 학습데이터 t_j 로부터 생성된 규칙

1. $R(t_1)$ (첫 번째 규칙을 생성)
2. $k = 1$
3. **for** $i = 2$ **to** N (각 학습데이터를 테스트)
4. **for** $j = 1$ **to** k (기 생성된 규칙들과 테스트)
5. $m_{ij} = M(t_i, r_j)$
6. **end for**
7. **if** $\max(m_{i1}, \ldots, m_{ik}) < m^d$ **then**
8. k = k + 1
9. $R(t_i)$ (k번째로 생성된 규칙 r_k)
10. **end if**
11. **end for**

즉, 각 학습데이터를 이용하여 퍼지규칙을 생성하고 기 생성된 퍼지규칙과 새로운 학습데이터간에 정합정도를 판단하여 기 생성된 퍼지규칙과의 정합정도의 최대치가 정합판단 값 m^d 보다 적을 경우에 해당 학습데이터를 위한 퍼지규칙을 생성한다. 그렇지 않을 경우에는 퍼지규칙을 생성하지 않는다. 그러므로 알고리즘 2는 필요한 퍼지규칙의 수뿐만 아니라 퍼지규칙도 함께 생성하는데, 이 퍼지규칙을 초기 퍼지규칙으로 사용하면 학습속도가 빨라진다. 정합판단 값 m^d는 사용자가 선택해야할 초기 파라메터 중 하나이다.

ESTM 은 NFSM을 학습시키는 역할을 담당한다. 개

체군에서 하나의 개체는 NFSM 의 링크 퍼지언어항에 대한 정보와 각 퍼지 언어항의 파라메터 (예로, Gaussian 함수를 퍼지 언어항의 멤버함수로 사용했다면, 표준편차와 평균값)를 갖는다. 그러므로 하나의 개체는 NFSM 이 동작하는데 필요한 모든 정보를 갖고 있다. 예를 들어 퍼지언어항의 개수를 3개로 (***NB, ZO, PB)*** 사용했다면 입,출력 링크 각각은 이 3개중 하나의 퍼지언어항이 할당된다. 또한 입,출력 별로 각각의 퍼지언어항에 대한 표준편차와 평균값이 개체 내에 표현된다. 그러므로 학습이 종료되면 모든 입,출력 링크에 학습된 퍼지언어항이 할당되며 각 입,출력 별 퍼지언어항들의 표준편차와 평균이 설정된다. 그림 2 는 두 개의 입력 x_1, x_2 과 두 개의 중간노드 h_1, h_2 그리고 하나의 출력 y_1을 갖는 뉴로퍼지 시스템에서 3개의 퍼지언어항이 사용되었을 경우에 개체가 표현된 모습을 보여준다. 그림에서 (χ, σ)는 퍼지언어항의 소속함수로 가우시안 함수를 사용했을 때 소속함수의 중심과 표준편차를 의미한다.

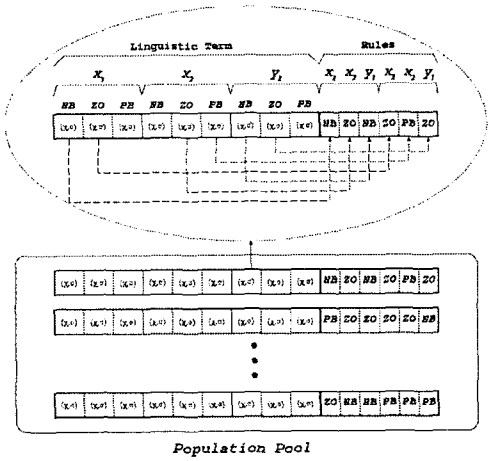

그림 2 . 개체군에서 개체의 표현

ESTM에서 각 개체는 각 개체의 파라메터를 이용하여 NFSM을 설정하고 학습데이터를 NFSM에 가하여 출력을 구해 이를 원하는 출력과 비교함으로서 개체의 적합도가 평가된다. 알고리즘 3은 개체군에서 개체를 평가하는 방법이다.

알고리즘 3 (개체 평가)

N : 총 학습데이터 수

P : 총 개체 수

O : NFSM의 출력 수

$NFSM^i$: i번째 개체를 이용하여 설정된 NFSM

o_j^k : j번째 학습데이터 적용시 $NFSM^i$ 의 k번째 실제 출력

t_j^k : j번째 학습데이터의 k번째 목표 출력

e_i : i번째 개체의 MSE (mean square error)

f_i : i번째 개체의 적합도

for $i=1$ **to** P

$NFSM^i$ 설정

$e_i = 0$

for $j=1$ **to** N

j번째 학습데이터를 $NFSM^i$에 적용

$NFSM^i$의 실제출력 o_j 을 구함

에러계산, $e_j = \sum_{k=1}^{O}(t_j^k - o_j^k)^2$

end for

i번째 개체의 MSE 를 계산, $e_i = \frac{1}{N}\sum_{j=1}^{N} e_j$

i번째 개체의 적합도를 설정, $f_i = \frac{1}{1+e_i}$

end for

부모개체로부터 돌연변이에 의하여 자식개체를 생성하는 방법은 다음과 같다. 먼저 돌연변이 되어야할 파라메터로는 각 부모개체의 입출력 퍼지 언어항에 대한 중심과 표준편차와 각 부모개체의 퍼지규칙 부분이 있다. 퍼지언어항에 대한 파라메터는 실수범위에서 돌연변이 하여야 하나 퍼지 규칙부분은 퍼지언어항의 개수 내에서 정수로 돌연변이 되어야만 한다. 진화전략에서는 돌연변이를 위하여 가우시안 랜덤노이즈를 이용한다. 다음 수식 (1)은 퍼지 언어항의 중심과 표준편차를 돌연변이 시키기 위한 수식이다.

$$\begin{aligned} \chi_o &= \chi_p + N(0, \sigma_t) \\ \sigma_o &= \sigma_p + N(0, \sigma_t) \end{aligned} \qquad (1)$$

수식 (1)에서 χ_o, σ_o 는 자식개체 소속함수의 중심과 표준편차이며 χ_p, σ_p 는 부모개체 소속함수의 중심과 표준편차이다. σ_t 는 자식개체를 생성하기 위한 가우시안 랜덤노이즈의 표준편차이다. σ_t는 효율적인 탐색을 위하여 해당 개체의 적합도 f_i 및 최대적합도 $f^{\max}$ 그리고 해당 입출력의 최대 최소값과 퍼지언어항의 개수 N_{lt}을 고려하여 식 (2)와 같이 결정된다.

$$\sigma_t = \left(1 - \frac{f_i}{f^{\max}} + \sigma_t^{\min}\right) \times \frac{(\max - \min)}{N_{lt}} \qquad (2)$$

식 (2)에서 $\sigma_t^{\min}$은 최대적합도를 갖는 부모개체도 돌연변이에 의하여 진화되게끔 사용자가 선택하는 최소 표준편차이다.
퍼지규칙의 경우에는 식 (3)에 의하여 자식개체의 퍼지규칙이 결정된다.

$$r_o = r_p + N(0, \sigma_r) \quad (3)$$

식 (3)에서 σ_r 은 퍼지규칙 돌연변이를 위한 표준편차로 효율적인 탐색을 위해 다음 식 (4) 처럼 계산된다.

$$\sigma_r = \left(1 - \frac{f_i}{f^{\max}} + \sigma_r^{\min}\right) \quad (4)$$

식 (4)에서 $\sigma_r^{\min}$은 사용자가 선택할 최소 표준편차이다. 퍼지 규칙의 경우 퍼지언어항의 개수 내에서 정수의 값을 가져야 한다. 즉 퍼지 언어항의 개수가 3개일 경우 0, 1, 2 중에서 하나의 값을 갖는다. 그러므로 추가적으로 식 (5)와 같은 조작이 필요하다.

$$r_o^* = \begin{cases} 0 & \text{if } r_o < 0 \\ N_{lt} - 1 & \text{if } r_o > N_{lt} - 1 \\ (integer) r_o & Otherwise \end{cases} \quad (5)$$

III. 실험결과

본 논문에서 제안한 뉴로퍼지 시스템의 성능을 보이기 위하여 식 (6)과 같은 비선형 시스템 동정에 응용하였다. 식 (7)은 실험에 사용한 입력이다.

$$y(k+1) = \frac{y(k)}{1+y^2(k)} + u^3(k) \quad (6)$$

$$u(k) = \sin(2\pi k/25) \quad , 0 \le k < 100 \quad (7)$$

실험시 설정한 파라메터는 다음과 같다. 학습데이터로 식 (6)에 무작위적인 입력을 가해 기록한 200개의 데이터를 사용하였고 퍼지언어항은 5개를 사용했으며 정합판단값 m^d는 0.4로, 학습종료 MSE 는 0.001 로, 부모 및 자식개체 수는 각각 40개로 $\sigma_t^{\min}$는 0.00005로, $\sigma_r^{\min}$은 0.005로 설정했다. 그림 2의 출력 결과는 입력을 수식 (7)과 같이 주었을 때 뉴로퍼지 시스템과 실제 시스템의 출력을 보여준다. 학습시간은 PentiumIII 400Mhz CPU 에 Linux 운영체제 하에서 25분 10초가 걸렸다. 그림 2에서 보듯이 빠른 학습에도 불구하고 학습성능이 좋음을 볼 수 있다.

IV. 결론

본 논문에서는 뉴로퍼지 시스템의 새로운 학습방법으로서 구조학습방법과 진화전략을 사용한 파라메터 학습방법을 제안하였다. 구조학습으로 필요한 퍼지규칙의 수뿐만 아니라 초기규칙을 생성할 수 있었으며 진화전략의 빠른 최적화 기능을 통하여 빠른 시간안에 뉴로퍼지 시스템이 학습됨을 보았다. 이러한 진화전략을 이용한 방법은 기존의 수정된 오류역전파 알고리즘을 이용한 학습보다 빠르며 지역해에 빠질 가능성이 적으며 기존의 유전자 알고리즘을 이용한 학습보다 세부적으로 탐색이 가능한 장점을 갖는다. 앞으로 학습시 사용되는 다수의 파라메터의 적절한 선택을 위한 이론적인 연구가 필요하다 하겠다.

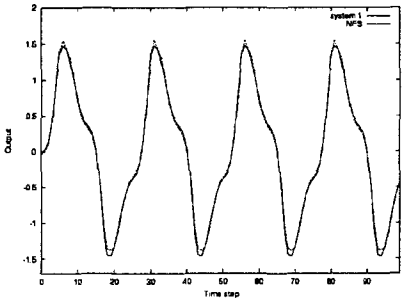

그림 3. 시스템 동정 결과

참고문헌(또는 Reference)

[1] W. A. Farag, V. H. Quintana, and G. Lambert-Torres, "A Genetic-Based Neuro-Fuzzy Approach for Modeling and Control of Dynamical Systems," *IEEE Trans. on Neural Network*}, vol. 9, pp. 756--767, Sept. 1998.

[2] K. S. Narendra and K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE Trans. on Neural Networks*, vol. 1, pp. 4--27, Mar. 1990.

[3] S. H. Jung, K. H. Cho, T. G. Kim, and K. H. Park, "Defuzzification Method for Multishaped Output Fuzzy Sets}," *Electronics Letters*, vol. 30, pp. 740--742, Apr. 1994.

[4] T. L. Seng, M. B. Khalid, and R. Yusof, "Tuning of a Neuro-Fuzzy Controller by Genetic Algorithm}," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 29, pp. 226--236, Apr. 1999.