최대경사법에 의한 퍼지모델의 조정

이건명* 김창범 이광형

한국과학기술원 전산학과

Tuning of the Fuzzy Model based on the Gradient Descent Method

Keon-Myung Lee ChangBum Kim Hyung Lee-Kwang

Dept. of Computer Science

Korea Advanced Institute of Science and Technology

요 약

조건부와 결론부에 언어항(퍼지집합)을 갖는 퍼지규칙은 언어적 지식표현에 유용한 형태이다. 이러한 퍼지규칙으로 구성되는 퍼지모델이 실제 응용되기 위해서는 대상시스템의 입출력 특성을 잘 반영할 수 있도록 조정하는 작업이 필요하다 본 논문에서는 퍼지규칙의 결론부에 언어항으로 삼각 퍼지숫자를 허용하는 퍼지모델을 대상으로 하여, 퍼지추론에서의 비퍼지화 연산을 유리식 형태의 함수로 변환하는 방법을 제안한다. 또한 이 함수를 이용하여 최대경사법에 기초한 퍼지규칙의 조건부와 결론부에 있는 언어항을 조정하는 방법을 제안한다.

1 서론

퍼지모델링은 시스템의 특성을 퍼지추론을 이용하여 기술하는 방법으로, 복잡한 비선형 시스템을 언어적으로 표현할 수 있는 것이 특징이다. 일반적으로 퍼지모델링은 구조동정(structure identification)과 파라미터 조정(parameter tuning)의 두가지 과정을 거치게 된다. 구조동정에서는 입력변수, 입력변수간의 관계, 규칙의 개수, 입출력 공간에서의 분할(partition), 모델 파라미터의 초기값 등의 구조에 관한 요소를 결정한다 파라미터 조정 단계에서는 데이타와 모델의 출력간의 오차를 최소화하도록 모델 파라미터를 조정한다

구조동정을 위해 템플레이트(template) 방법, 군집화(clustering) 방법, 의사결정트리(decision tree) 방법 등이 제안되어 있지만, 구조동정은 본질적으로 자동화하기 어려운 문제로서 시스템 개발자의 개입이 요구된다[6]. 이에 반해 파라미터 조정은 비교적 자동화가 용이하여, 학습에 기반한 방법, 수학적 계획법(mathematical programming), 휴리스틱에 기반한 방법등이 도입되어 성과를 보이고 있다[1][7][8].

본 논문은 파라미터 조정을 위해 학습에 기반한 방법으로서 최대경사법(gradient descent method)을 사용하는 방법에 중점을 두고 있다. 우선, 기존의 최대경사법에 기반한 방법과 이들이 적용될 수 있는 퍼지모델에 대해서 간단히 살펴본 다음, 기존에 다루고 있지 않았던 규칙의 조건부와 결론부에 언어항(linguistic term)을 가지는 퍼지규칙으로 구성된 퍼지모델의 조정방법을 다룬다. 여기에서는 무게중심(center of gravity, centroid)법에 의한 비퍼지화(defuzzification) 연산을 유리식 형태의 함수로 변환하는 방법을 제시하고, 이 함수를 이용하여 규칙의 조건부와 결론부를 조정하는 최대경사법에 기반을 둔 방법을 제안한다.

2 기존 연구

파라미터 조정을 위한 몇가지 최대경사법에 의한 방법이 제안되어 왔다[1][2][3][4][5] 이들 방법은 대상이 되는 퍼지모델의 형태(퍼지규칙베이스를 구성하는 퍼지규칙의 모양)에 따라 다음과 같이 결론부가 상수인 퍼지규칙으로 구성되는 퍼지모델, 결론부가 입력변수의 선형방정식인 퍼지규칙으로 구성되는 퍼지모델, 선형퍼지집합을 결론부에 갖는 규칙에 진리값(truth value)를 가지는 퍼지모델로 구별될 수 있다.

2.1 결론부가 상수인 퍼지규칙

결론부가 상수인 퍼지규칙은 다음과 같은 형태를 가진다. 여기에서 X_j와 Y는 각각 입력변수와 출력변수를 나타낸다 조건부의 A_{ij}는 입력변수 X_j에 대한 언어항으로 소속함수 (membership function)에의해 표현되다 결론부의 w_i는 임의의 상수값이다.

$$R_i : \text{IF } X_1 \text{ is } A_{1i} \text{ and } \ldots \text{ and } X_m \text{ is } A_{mi} \text{ THEN } Y = w_i$$

이와 같은 퍼지규칙으로 구성된 퍼지모델에서의 추론은, 무게중심법을 사용하여 비퍼지화를 할 경우 다음과 같이 수행된다.

$$y = \frac{\sum_i \mu_i w_i}{\sum_i \mu_i}$$

$$\mu_i = f(A_{1i}(x_1), \ldots, A_{mi}(x_m))$$

위식에서 y는 비퍼지화된 추론결과를 나타내고, $A_{ji}(x_j)$는 언어항 A_{ji}에 대한 입력 x_j의 소속정도(적합도)를 나타낸다 μ_i는 규칙 R_i의 입력 $(x_1, x_2, \ldots, x_m)$에 대한 적합도(matching degree)를 나타내고, 조건부의 각 언어항에 대한 적합도를 합성하는 함수 f에는 최소값(minimum) 연산이나 곱셈(multiplication) 연산이 사용된다.

f 최소값 연산

$$\mu_i = f(A_{1i}(x_1), \ldots, A_{mi}(x_m)) = \min_j \{A_{ji}(x_j)\}$$

f : 곱셈 연산

$$\mu_i = f(A_{1i}(x_1), \ldots, A_{mi}(x_m)) = \prod_j A_{ji}(x_j)$$

이러한 퍼지모델에 대한 조정방법은 조건부에 대칭/비대칭인 삼각 퍼지숫자, 시그모이드(sigmoid) 또는 종형(bell-shaped) 퍼지집합을 갖는 퍼지규칙으로 구성된 퍼지모델에 대해 제안되어 있다. 이들 모델에 대한 최대경사법에 의한 조정은 다음과 같이 이루어진다. 식에서 a_{ij}는 조정될 파라미터를 나타내고, y_i^d는 i번째 학습데이타의 입력에 대한 기대

출력을 나타내고, η는 학습률(learning rate)를 나타낸다.

$$E = \frac{1}{2}\sum_{i=1}^{p}(y_i - y_i^d)^2$$
$$a_{ij}(t+1) = a_{ij}(t) - \eta\frac{\partial E}{\partial a_{ij}}$$

2.2 결론부가 선형방정식인 퍼지규칙

Sugeno-Takagi 퍼지규칙이라고 불리는 결론부가 입력변수의 선형방정식 (linear equation)인 퍼지규칙은 다음과 같은 형태를 가진다. 여기에서 k_{ji}는 상수를 나타내고, x_i는 입력변수를 나타낸다

$$R_i : \text{IF } X_1 \text{ is } A_{1i} \text{ and } \dots \text{ and } X_m \text{ is } A_{mi} \text{ THEN } Y = g_i(\,)$$
$$g_i(x_1,\dots,x_n) = k_{0i} + k_{1i}x_1 + \cdots + k_{mi}x_m$$

위와같은 퍼지규칙으로 구성된 퍼지모델에서 추론은, 무게중심법을 사용할 때 다음과 같이 수행된다.

$$y = \frac{\sum_i \mu_i g_i}{\sum_i \mu_i} \tag{1}$$
$$\mu_i = f(A_{1i}(x_1),\dots,A_{mi}(x_m))$$

비퍼지화 연산을 하는 식 (1)이 미분가능하기 때문에, 조건부의 언어항에 대한 파라미터와 결론부의 계수(coefficient)에 대해 아래와 같이 최대경사법을 적용할 수 있다.

$$E = \frac{1}{2}\sum_{i=1}^{p}(y_i - y_i^d)^2$$
$$a_{ij}(t+1) = a_{ij}(t) - \eta\frac{\partial E}{\partial a_{ij}}$$

2.3 결론부가 선형 퍼지집합이고 진리값을 갖는 퍼지규칙

결론부에 두가지의 선형 퍼지집합만을 허용하고, 규칙에 진리값이 부여된 퍼지규칙으로 다음과 같은 형태를 갖는다.

R_i (IF X_1 is A_{1i} and ... and X_m is A_{mi} THEN $Y = B_k$) is τ_i

$\tau_i \in [0,1]$: 진리값(truth value)

$B_k,\ k = 1,2$: 선형소속함수

$$B_1(y) = \frac{-y+\max(Y)}{\max(Y)-\min(Y)}$$
$$B_2(y) = \frac{y-\min(Y)}{\max(Y)-\min(Y)}$$

위 식에서 $\max(Y)$과 $\min(Y)$는 각각 출력공간(output space)에서의 최대값과 최소값을 나타낸다 이러한 퍼지규칙으로 구성된 퍼지모델에서의 추론은 다음과 같이 수행한다.

$$y = \frac{\sum_{k=1}^{2}\mu_k' B_k^{-1}(\mu_k')}{\sum_{k=1}^{2}\mu_k'} \tag{2}$$
$$\mu_i' = \frac{\sum_{i=1}^{m}\mu_i\tau_i}{\sum\mu_i}$$
$$\mu_i = f(A_{1i}(x_1),\dots,A_{mi}(x_m))$$

식 (2)가 미분가능하기 때문에 최대경사법에 의해 규칙의 진리값과 조건부에 있는 언어항의 파라미터를 조정할 수 있다.

앞에서 언급한 세가지 모델에서 보는 바와 같이, 최대경사법이 어떤 퍼지모델의 조정에 적용되기 위해서는 오차에 대한 식(E)이 미분가능해야 한다 즉, 비퍼지화 연산을 하는 식이 미분가능해야 한다 그런데 규칙의 조건부와 결론부에 언어항을 가지는 퍼지규칙이 시스템 개발자가 전문지식을 용이하게 표현할 수 있음에도 불구하고, 비퍼지화 연산식에 대한 미분방법이 개발되지 않아 결론부에 언어항을 갖는 퍼지규칙으로 구성된 퍼지모델의 조정에 최대경사법을 사용할 수 없었다

퍼지모델링에서 삼각 퍼지숫자(triangular fuzzy number)를 가지는 퍼지규칙은 흔히 사용된다. 본 논문에서는 규칙의 결론부에 삼각 퍼지숫자를 갖는 퍼지모델에 최대경사법을 사용하는 방법을 제안한다.

3 규칙의 결론부에 언어항을 갖는 퍼지모델의 조정

3.1 퍼지모델

조건부와 결론부에 언어항을 가지는 퍼지규칙은 다음과 같은 형태를 가진다. 이때 언어항 A_{ji}과 B_{ji}에 대한 소속함수는 중심점, 폭 등 몇개의 파라미터에 의해 결정되는 함수이다

$$R_i : \text{IF } X_1 \text{ is } A_{1i} \text{ and } \dots \text{ and } X_m \text{ is } A_{mi} \text{ THEN}$$
$$Y_1 \text{ is } B_{1i} \text{ and } \dots \text{ and } Y_p \text{ is } B_{pi}$$

결론부의 언어항으로는 다음과 같은 삼각 퍼지숫자 $Tri(\,)$로 제한한다 아래식에서 c는 삼각 퍼지숫자의 중심점에 해당하고, $c-l$과 $r-c$는 각각 왼쪽 폭과 오른쪽 폭에 해당한다.

$$Tri(l,c,r;h) = \begin{cases} \frac{h}{c-l}(x-l) & x \in [l,c] \\ \frac{h}{c-r}(x-r) & x \in (c,r] \\ 0 & \text{otherwise} \end{cases}$$

한편, 출력공간의 가장 왼쪽의 언어항과 가장 오른쪽 언어항에는 다음과 같은 사다리꼴(trapezoidal) 퍼지숫자 $Trap_l(\,)$, $Trap_r(\,)$이 허용된다.

$$Trap_l(l,c,r;h) = \begin{cases} h & x \in [l,c] \\ \frac{h}{c-r}(x-r) & x \in (c,r] \\ 0 & \text{otherwise} \end{cases}$$

$$Trap_r(l,c,r;h) = \begin{cases} \frac{h}{c-l}(x-l) & x \in [l,c] \\ h & x \in (c,r] \\ 0 & \text{otherwise} \end{cases}$$

편의상 출력변수가 하나이고, 출력변수에 대해 n개의 언어항이 정의되어 있다고 하자

$$R_i : \text{IF } X_1 \text{ is } A_{1i} \text{ and } \dots \text{ and } X_m \text{ is } A_{mi} \text{ THEN } Y \text{ is } B_i$$

출력변수에 정의된 언어항을 중심점의 위치에 따라 다음과 같이 N_i로 나타내자.

$$N_i = (l_i, c_i, r_i, h_i),\ i = 1,\dots,n$$
$$c_{i-1} < c_i$$

이때, 이들 언어항들은 다음의 제한조건이 만족하도록 정의되어야 한다. 이들 제한조건이 강한 것은 아니기 때문에 대부분의 퍼지모델이 만족시킬 만한 것이다

$$c_{i-1} < c_i \tag{3}$$
$$l_{i-1} < l_i \tag{4}$$
$$r_{i-1} < r_i \tag{5}$$
$$r_{i-1} \le l_{i+1} \tag{6}$$

반면, 조건부의 언어항에는 제한이 없기 때문에 삼각 퍼지숫자, 사다리꼴 퍼지숫자, 종형 퍼지숫자등 임의의 퍼지집합이 올 수 있다

3.2 비퍼지화 연산

퍼지추론에서 추론결과는 규칙의 조건부와 입력의 적합도와 결론부의 언어항에 의해서 결정된다 즉, 퍼지추론에서 결론부의 합성및 비퍼지화 연산은 적합도를 파라미터로 하는 함수로 볼 수 있다. 따라서 이러한 함수식를 구할 수 있고 이 함수식이 미분가능하다면, 최대경사법을 퍼지모델의 조정에 적용할 수 있다

식 (3)-(6)의 제한조건을 만족하는 퍼지모델에 대해서, 각 퍼지규칙의 부분적인 추론결과를 합성하여 무게중심법으로 비퍼지화하는 함수를 구할 수 있다 단, 이때 퍼지모델에서의 추론방법은 Max-Min 추론이나 Max-Product 추론을 전제로 한다

여기에서는 Max-Min 추론이 사용되는 경우의 비퍼지화 함수 $C(\ )$에 대해서 기술한다. 편의상 다음의 표기법을 사용한다 α_i은 언어항 N_i에 대한 적합도이다 즉, α_i는 결론부에 언어항 N_i를 가지는 규칙들의 적합도 중에서 최대값이다 $C_{\alpha_i}(N_i)$와 $A_{\alpha_i}(N_i)$는 각각 언어항 N_i를 소속정도가 α_i 이상되는 부분을 잘라낸 퍼지집합 $\{(y, \min\{N_i(y), \alpha_i\}) \mid y \in Supp(N_i)\}$의 무게중심점과 면적이다. I_i는 언어항 N_i와 N_{i+1}이 겹쳐서 만들어지는 퍼지집합이다. 그런데 결론부의 언어항은 삼각 퍼지숫자이기 때문에 N_i와 N_{i+1}가 겹쳐서 생기는 퍼지집합 I_i는 다음과 같은 삼각형 모양이다.

$$I_i = N_i \cap N_{i+1} = Tri(l_{i+1}, c_i^I, r_i,\ h_i^I)$$
$$c_i^I = \frac{h_i r_i (c_{i+1} - l_{i+1}) - l_{i+1} h_{i+1} (c_i - r_i)}{h_i (c_{i+1} - l_{i+1}) - h_{i+1} (c_i - r_i)}$$
$$h_i^I = \max\{\frac{h_{i+1}}{c_{i+1} - l_{i+1}} (\frac{h_i r_i (c_{i+1} - l_{i+1}) - l_{i+1} h_{i+1} (c_i - r_i)}{h_i (c_{i+1} - l_{i+1}) - h_{i+1} (c_i - r_i)} - l_{i+1}), 0\}$$

추론결과를 무게중심법에 의해 비퍼지화하는 식 $C(\alpha_1, \alpha_2, \cdot\cdot, \alpha_n)$은 다음과 같다 아래식에서 $N_i^{\alpha_i}$은 언어항 N_i에서 소속정도가 α_i이상인 것을 α_i로 해서 만든 퍼지집합을 나타낸다

$$C(\alpha_1, \alpha_2, \ldots, \alpha_n) = \frac{\int x S(x) dx}{\int S(x) dx} \tag{7}$$
$$S = N_1^{\alpha_1} \cup N_2^{\alpha_2} \cup \cdots \cup N_n^{\alpha_n}$$

식 (7)이 적분을 포함하기 때문에 복잡해 보이지만, 식(3)-(6)의 조건이 만족되면 식 (7)은 다음과 같은 식으로 변환될 수 있다.

$$\begin{aligned} & C(\alpha_1, \alpha_2, \ldots, \alpha_n) \\ &= \frac{\sum_{i=1}^{n} C_{\alpha_i}(N_i) A_{\alpha_i}(N_i) - \sum_{i=1}^{n-1} C_{\alpha_i \wedge \alpha_{i+1}}(N_i \cap N_{i+1}) A_{\alpha_i \wedge \alpha_{i+1}}(N_i \cap N_{i+1})}{\sum_{i=1}^{n} A_{\alpha_i}(N_i) - \sum_{i=1}^{n-1} A_{\alpha_i \wedge \alpha_{i+1}}(N_i \cap N_{i+1})} \\ &= \frac{\sum_{i=1}^{n} C_{\alpha_i}^u(N_i) - \sum_{i=1}^{n-1} C_{\alpha_i \wedge \alpha_{i+1}}^u(N_i \cap N_{i+1})}{\sum_{i=1}^{n} A_{\alpha_i}(N_i) - \sum_{i=1}^{n-1} A_{\alpha_i \wedge \alpha_{i+1}}(N_i \cap N_{i+1})} \\ &= \frac{\sum_{i=1}^{n} C_{\alpha_i}^u(N_i) - \sum_{i=1}^{n-1} C_{\alpha_{\triangle_i}}^u(I_i)}{\sum_{i=1}^{n} A_{\alpha_i}(N_i) - \sum_{i=1}^{n-1} A_{\alpha_{\triangle_i}}(I_i)} \end{aligned} \tag{8}$$
$$\alpha_{\triangle_i} = \alpha_i \wedge \alpha_{i+1} \wedge h_i^I$$

적합도에 따른 언어항의 면적 $A_\alpha(\cdot)$은 다음과 같이 구해진다.

$N_i = Tri(l_i, c_i, r_i; h_i)$일때

$$A_{\alpha_i}(N_i) = \frac{l_i - r_i}{2h_i} \alpha_i^2 + (r_i - l_i)\alpha_i$$

$N_i = trap_l(l_i, c_i, r_i; h_i)$일때 $(i = 1)$

$$A_{\alpha_i}(N_i) \frac{(c_i - r_i)}{2h_i} \alpha_i^2 + (r_i - l_i)\alpha_i$$

$N_i = trap_r(l_i, c_i, r_i; h_i)$일때 $(i = n)$

$$A_{\alpha_i}(N_i) = \frac{-(c_i - l_i)}{2h_i} \alpha_i^2 + (r_i - l_i)\alpha_i$$

$I_i(i = 1, \cdots, n-1)$에 대한 면적

$$A_{\alpha_{\triangle_i}}(I_i) = \frac{l_{i+1} - r_i}{2h_i^I} \alpha_{\triangle_i}^2 + (r_i - l_{i+1})\alpha_{\triangle_i}$$

언어항에 대한 무게중심 $C_\alpha(\ )$은 다음과 같은 식에 의해 계산된다

$$C_{\alpha_i}(N_i) = \frac{C_{\alpha_i}^u(N_i)}{A_{\alpha_i}(N_i)}$$

$N_i = Tri(l_i, c_i, r_i; h_i)$일때

$$C_{\alpha_i}^u(N_i) = \frac{(c_i - r_i)^2 - (c_i - l_i)^2}{6h_i^2} \alpha_i^3 + \frac{(c_i - r_i)r_i - (c_i - l_i)l_i}{2h_i} \alpha_i^2 + \frac{(r_i^2 - l_i^2)}{2} \alpha_i$$

$N_i = trap_l(l_i, c_i, r_i; h_i)$일때 $(i = 1)$

$$C_{\alpha_i}^u(N_i) = \frac{(c_i - r_i)^2}{6h_i^2} \alpha_i^3 + \frac{(c_i - r_i)r_i}{2h_i} \alpha_i^2 + \frac{(r_i^2 - l_i^2)}{2} \alpha_i$$

$N_i = trap_r(l_i, c_i, r_i; h_i)$일때 $(i = n)$

$$C_{\alpha_i}^u(N_i) = -\frac{(c_i - l_i)^2}{6h_i^2} \alpha_i^3 - \frac{(c_i - l_i)l_i}{2h_i} \alpha_i^2 + \frac{(r_i^2 - l_i^2)}{2} \alpha_i$$

$I_i(i = 1, \cdots, n-1)$에 대한 무게중심

$$C_{\alpha_{\triangle_i}}^u(I_i) = -\frac{(c_i^I - l_{i+1})^2 - (c_i^I - r_i)^2}{6(h_i^I)^2} \alpha_{\triangle_i}^3$$
$$-\frac{(c_i^I - l_{i+1})l_{i+1} - (c_i^I - r_i)r_i}{2h_i^I} \alpha_{\triangle_i}^2 - \frac{1}{2}(l_{i+1}^2 - r_i^2)\alpha_{\triangle_i}$$

위에서 $C_{\alpha_i}^u(N_i)$, $C_{\alpha_{\triangle_i}}^u(I_i)$, $A_{\alpha_i}(N_i)$, $A_{\alpha_{\triangle_i}}(I_i)$와 식이 유리식으로 표현되기 때문에 식 (8)는 유리식이다 따라서 비퍼지화하는 식 $C(\ )$는 미분가능한 식이다

삼각 퍼지숫자만 결론부의 언어항에 나타나는 경우, 제한조건 (3)-(6)을 만족하지 않는 경우에 대해서도 이론적으로는 비퍼지화하는 유리 함수식를 구할 수 있지만 수식이 매우 복잡해진다 또한 사다리꼴 퍼지숫자로 결론부가 구성되는 경우에도 마찬가지로 비퍼지화하는 함수를 구할 수 있긴 하지만 수식이 복잡해진다

3.3 규칙의 결론부 조정

비퍼지화 함수식 $C(\alpha_1, \alpha_2, \ldots, \alpha_n)$가 미분가능하기 때문에, 퍼지모델의 조정에 최대경사법을 사용할 수 있다.

모델에 대한 오차를 다음과 같이 정의한다. y_i는 i번째 학습데이타의 입력에 대한 기대 출력이고, y_i^d는 퍼지모델의 실제 출력이다

$$E = \frac{1}{2} \sum (y_i^d - y_i)^2$$

퍼지모델의 출력 y_i가 함수 $C(\cdot)$에 의해 계산될 수 있기 때문에, 퍼지규칙의 결론부에 대해 최대경사법을 다음과 같이 적용할 수 있다. 단, 파라미터를 조정할 때 결론부에 대한 제한조건 (3)-(6)을 만족시키도록 해야 한다

$$c_i^{(t+1)} = c_i^{(t)} - \eta \frac{\partial E}{\partial c_i} = c_i^{(t)} - \eta \frac{\partial E}{\partial C} \frac{\partial C}{\partial c_i}$$

$$l_i^{(t+1)} = l_i^{(t)} - \eta \frac{\partial E}{\partial l_i} = l_i^{(t)} - \eta \frac{\partial E}{\partial C} \frac{\partial C}{\partial l_i}$$

$$r_i^{(t+1)} = r_i^{(t)} - \eta \frac{\partial E}{\partial r_i} = r_i^{(t)} - \eta \frac{\partial E}{\partial C} \frac{\partial C}{\partial r_i}$$

$$\frac{\partial E}{\partial C} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial C} = (y_i - y_i^d)$$

3.4 규칙의 조건부 조정

Max-Min 추론이나 Max-Product 추론을 할 때는 규칙의 조건부에 있는 모든 언어항이 추론결론에 영향을 미치는 것은 아니다 조건부의 언어항을 조정할 때는 해당 언어항의 입력과의 적합도가 출력에 어떻게 영향을 미치는지 알아야 한다.

조건부 언어항의 조정을 위해, 언어항의 적합도의 영향을 명확히 나타낼 수 있는 (그림 1)과 같은 퍼지추론 네트워크를 고려한다. 제한조건 (3)-(6)를 만족하고 결론부에 삼각 퍼지숫자를 갖는 퍼지모델은 퍼지추론 네트워크로 나타낼 수 있다.

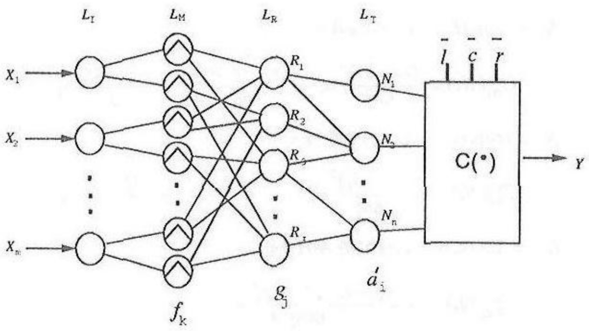

그림 1: 퍼지추론 네트워크

퍼지추론 네트워크에서 L_I층의 노드는 입력변수에 해당하고, L_M층의 노드는 조건부의 언어항에 나타내고 입력과의 적합도를 계산하는 역할을 한다. L_R층의 각 노드는 규칙의 조건부에 해당하고 이전 층에서 계산한 언어항과 입력의 적합도를 최소값을 구하거나 모두 곱하거나 하여 합성한다. L_T층의 노드는 출력변수에 대해 정의된 언어항을 나타내고, L_R층에서 계산된 적합도 중에서 최대값을 구하는 일을 한다. L_R층의 노드는 자신에 해당하는 규칙의 결론부에 있는 언어항을 나타내는 L_T층의 노드와 연결된다. f_k, g_j과 α_i는 각각 L_M층의 k번째 노드의 출력, L_R층의 j번째 노드의 출력, L_T층의 i번째 노드의 출력을 나타낸다.

퍼지모델에서 규칙의 조건부에 있는 언어항은 최대경사법을 사용하여 다음과 같이 조정한다.

$$a_k^{(t+1)} = a_k^{(t)} - \eta \frac{\partial E}{\partial a_k} = a_k^{(t)} - \eta \sum_i \sum_j \frac{\partial E}{\partial y}\frac{\partial y}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial g_j}\frac{\partial g_j}{\partial f_k}\frac{\partial f_k}{\partial a_k} \quad (9)$$

위 식 (9)의 편미분에서 $\frac{\partial E}{\partial C}$과 $\frac{\partial y}{\partial \alpha_i}$는 수학적 미분식으로 계산되고, $\frac{\partial \alpha_i}{\partial g_j}$, $\frac{\partial g_j}{\partial f_k}$은 다음과 같이 계산된다.

$$\frac{\partial \alpha_i}{\partial g_j} = \begin{cases} 1 & g_j = \alpha_i \text{ 이고 해당 노드간에 연결이 있을때} \\ 0 & \text{otherwise} \end{cases}$$

조건부 언어항의 합성에 최소값 연산을 사용할 때 : $g_j = \min_s\{f_s\}$

$$\frac{\partial g_j}{\partial f_k} = \begin{cases} 1 & f_k = g_j \text{ 이고 해당 노드간에 연결이 있을 때} \\ 0 & \text{otherwise} \end{cases}$$

조건부 언어항의 합성에 곱셈을 사용할 때 : $g_j = \prod_s f_s$

$$\frac{\partial g_j}{\partial f_k} = \begin{cases} \prod_{k \neq s} f_s & f_k = g_j \text{이고 해당 노드간에 연결이 있을 때} \\ 0 & \text{otherwise} \end{cases}$$

조건부의 언어항을 조정할 때 오차의 영향은 퍼지추론 네트워크에서 다음과 같이 역전파해 간다. 우선 L_T층의 노드에서 $\frac{\partial E}{\partial y}\frac{\partial y}{\partial \alpha_i}$의 값을 계산하여 이값을 L_R에 전파시킨다. L_R층에서는 $\frac{\partial \alpha_i}{\partial g_j}$을 구하고, $\sum_i(\frac{\partial E}{\partial y}\frac{\partial y}{\partial \alpha_i})\frac{\partial \alpha_i}{\partial g_j}$를 계산한다. L_M 층에서는 $\frac{\partial g_j}{\partial f_k}$을 계산하고 $\sum_j(\sum_i(\frac{\partial E}{\partial y}\frac{\partial y}{\partial \alpha_i})\frac{\partial \alpha_i}{\partial g_j})\frac{\partial g_j}{\partial f_k}$을 구한다. 이 값을 이용하여 조건부의 언어항을 조정한다.

4 결론

본 논문에서는 기존의 최대경사법에 의해 조정이 가능했던 퍼지모델에 살펴보고, 결론부에 언어항을 갖는 퍼지모델로서 결론부가 삼각 퍼지숫자를 허용하는 모델에 대한 최대경사법에 의한 퍼지모델의 조정방법을 제안했다.

결론부에 삼각 퍼지숫자를 갖는 퍼지모델에 대해 퍼지추론의 비퍼지화하는 연산을 유리식 형태의 함수로 변환하는 방법을 제시했다. 이렇게 유리함수로 비퍼지화 연산을 나타냄으로써 퍼지추론 결과를 이산화(discretization)하지 않은채 수행할 수 있다. 또한 이 함수가 미분가능하게 때문에 조건부와 결론부의 언어항을 조정하는데 최대경사법을 적용할 수 있게 해준다.

퍼지규칙의 조건부에 있는 각 언어항의 적합도를 합성하는데 최소값 연산을 사용하면, 최대경사법을 사용하여 조건부의 언어항을 조정할 때 오차가 증가하는 경우가 발생할 수 있다. 현재 이러한 문제의 해결방법에 대한 연구가 진행중이다.

참고 문헌

[1] François Guély, Patrick Siarry, "Gradient Descent Method for Optimizing Various Fuzzy Rule Bases", *2nd Int. Conf. on Fuzzy Systems 1993*, pp. 907-917, 1993.

[2] Hiroyoshi Nomura, Isao Hayashi, Noboru Wakami, "A Self-Tuning Method of Fuzzy Control by Descent Method", *IFSA'91 Brussels, Vol: Engineering*, pp.155-158, 1991.

[3] J.S.Roger Jang, "Fuzzy Controller Design without Domain Experts", *2nd Int. Conf. on Fuzzy Systems 1992*, pp.289-296, 1992.

[4] Hidetomo Ichihashi, "Iterative Fuzzy Modeling and a Hierachical Network", *IFSA'91 Brussels, Vol: Engineering*, pp.49-52, 1991.

[5] Shin-ichi Horikawa, Takeshi Furuhashi, Yoshiki Uchikawa, "On Fuzzy Modeling Using Fuzzy Neural Networks with the Back-Propagation Algorithm", *IEEE Trans. on Neural Networks*, Vol.3, No.5, pp.801-806, 1992.

[6] Ronald Yager, "Structures for Building Flexible Neural Fuzzy Systems", *IFSA'93 Seoul, Plenary Talk*, 1993.

[7] Tomohiro Takagi, Michio Sugeno, "Fuzzy Identification of Systems and Its Applications to Modeling and Control", *IEEE SMC, Vol.SMC-15, No.1*, pp.116-132, 1985.

[8] M. Sugeno, G.T. Kang, "Structure Identification of Fuzzy Model", *Fuzzy Sets and Systems, Vol.28, No.1*, pp.15-33, 1988.