

論文2003-40CI-5-2

확장 칼만필터를 이용한 온라인 퍼지 모델링 알고리즘에 대한 연구

(A Study on On-line modeling of Fuzzy System via Extended Kalman Filter)

金 殷 泰 *

(Euntai Kim)

요 약

본 논문에서는 퍼지 모델의 온라인 동정 알고리즘을 제안한다. 본 논문에서 고려하는 퍼지 모델은 후건부가 싱글톤인 퍼지 시스템으로 퍼지 기저함수의 선형 합으로 표현된다. 온라인 동정을 위해서 제곱 코사인 소속함수를 제안한다. 제곱 코사인 함수는 다른 소속함수에 비해 적은 파라미터를 갖으며 전 구간에서 미분 가능한 특징을 갖는다. 퍼지 모델의 파라미터는 그레디언트 하강법과 확장칼만필터를 이용하여 온라인으로 결정한다. 끝으로 컴퓨터 모의 실험을 통하여 제안한 방법의 타당성을 확인한다.

Abstract

In this paper, an explanation regarding on-line identification of a fuzzy system is presented. The fuzzy system to be identified is assumed to be in the type of singleton consequent parts and be represented by a linear combination of fuzzy basis functions. For on-line identification, squared-cosine membership function is introduced to reduce the number of parameters to be identified and make the system consistent and differentiable. Then the parameters of the fuzzy system are identified on-line by the gradient search method and Extended Kalman Filter. Finally, a computer simulation is performed to illustrate the validity of the suggested algorithms.

Keywords : 확장칼만필터, 온라인 퍼지 모델링, 타입 2

I. 서 론

지난 20년간 Zadeh교수에 의해 제안된 퍼지 논리는 공학 및 과학 분야에서 널리 이용되었으며 최근에는 신경회로망, 유전자 알고리즘 등과 결합하여 소프트 컴퓨팅이라는 이론으로 발전하였다. 이는 기존의 인공지능의 개념을 대체하는 개념으로 모호한 정보를 처리해

야 하는 필요성에 의해서 대두된 것이다.

이러한 퍼지 논리의 응용 중 가장 대표적인 것이 퍼지 모델링이다^[1]. 퍼지 모델링이란 샘플 데이터로부터 전체의 시스템이나 함수를 퍼지 시스템으로 다시 구현하는 것으로 이러한 퍼지 모델링 기법에 대해서는 이미 많은 연구결과가 보고되어 있다. 이는 퍼지 모델이 기존의 선형 모델에 비해 묘사력이 뛰어나며, 신경망 모델과 달리 구조가 명백하고 오퍼레이터에게 직관적으로 설득력이 있기 때문이다.

대표적인 연구를 보면 Pedrycz는^[2,3]의 논문에서 퍼지 관계식을 이용한 설정 방식을 제안하였고 Sugeno등

* 正會員, 延世大學校 電氣電子工學部

(Yonsei University, School of Electrical and Electronic Engr.)

接受日字:2001年7月25日, 수정완료일:2003年8月5日

은 TSK (Takagi-Sugeno-Kang) 퍼지 시스템과 설정 방법을 제안하였다^[4,5]. Sugeno는 또 Yasukawa와 함께 질적 모델링이라는 방법을 제안하였고^[6] 이 밖에도 여러 연구자들이 TSK 모델이나 뉴로 퍼지 모델 등에 대하여 연구결과를 발표하였다^[7-10].

그러나 이러한 방법들은 거의 대부분 오프라인(off-line)방식으로 구현되는 알고리즘으로 실시간 처리를 요하는 비선형 적응 제어나 적응 신호처리에서는 사용할 수 없는 단점을 가지고 있다. 또 온라인 구현을 목표로 '연속퍼지 모델링 방식' 이 제안되었지만^[11] 이 방법 역시 진정한 의미의 온라인 방식 퍼지 모델이라 할 수 없다.

이러한 문제를 극복하기 위해 본 논문에서는 온라인 퍼지 모델링 방식을 제안한다. 우선 퍼지 시스템의 파라미터 설정을 용이하게 하기 위하여 제곱 코사인(squared-cosine) 소속함수를 정의하고 그레디언트 하강법, 확장 칼만필터 (extended Kalman filter), 하이브리드 방식의 세 가지 방법을 이용하여 온라인 퍼지 모델링을 수행한다. 제곱 코사인 소속함수는 삼각형 함수보다 부드럽고 가우시안 소속함수보다 적은 파라미터를 가지고 있어서 빠른 퍼지 모델의 설정에 도움이 된다.

II. 퍼지 시스템 일반

1. 퍼지 시스템의 종류

Mamdani가 퍼지 논리를 실제 시스템에 응용한 이래, 구조나 소속함수 등의 변화에 따라 여러 형태의 퍼지 시스템이 발표되었다. 이러한 형태들은 크게 다음의 세 가지 종류로 요약할 수 있다.

타입 I : (언어형 후건부)

$$R^n : \text{If } x_1 \text{ is } A_1^n \text{ and } x_2 \text{ is } A_2^n, \dots, x_m \text{ is } A_m^n, \\ \text{then } y \text{ is } C^n$$

타입 II : (싱글톤 후건부)

$$R^n : \text{If } x_1 \text{ is } A_1^n \text{ and } x_2 \text{ is } A_2^n, \dots, x_m \text{ is } A_m^n, \\ \text{then } y \text{ is } \theta^n$$

타입 III : (선형 후건부)

$$R^n : \text{If } x_1 \text{ is } A_1^n \text{ and } x_2 \text{ is } A_2^n, \dots, x_m \text{ is } A_m^n, \\ \text{then } y = a_0^n + a_1^n x_1 + a_2^n x_2 + \dots + a_m^n$$

A_i^n 과 C^n 는 퍼지 변수이고 θ^n 는 싱글톤이며 a_i^n 는 후건부 계수이다.

위에서 타입 II는 타입 I과 타입 III의 중간형태로 볼 수 있고 Sugeno 등에 의해 제안된 타입 III는 흔히 TSK (Takagi-Sugeno-Kang) 퍼지 시스템이라 한다. 본 논문에서는 다음과 같이 후건부가 싱글톤 형태인 타입 II 퍼지시스템 γ 을 이용한다.

$$R^{n_1 n_2 \dots n_m} : \text{If } x_1 \text{ is } A_1^{n_1} \text{ and } x_2 \text{ is } A_2^{n_2}, \dots, x_m \text{ is } A_m^{n_m} \\ \text{then } y \text{ is } \theta^{n_1, n_2, \dots, n_m} \quad (1) \\ (n_1 = 1, \dots, M_1, n_2 = 1, \dots, M_2, \dots, n_m = 1, \dots, M_m)$$

여기서 $\mathbf{x} \equiv [x_1, x_2, \dots, x_m]$ 과 y 는 퍼지 시스템의 입력 출력 변수이고 퍼지시스템 (1)은 다음과 같은 함수의 역할을 한다.

$$\gamma : U \subset R^m \rightarrow V \subset R$$

$$\text{여기서 } U = U_1 \times U_2 \times \dots \times U_m \subset R^m$$

여기서 싱글톤 퍼지화기, 곱하기 T-놈, 중심평균의 비퍼지화기를 이용하면 퍼지 시스템 (1)은 다음의 식으로 표현된다.

$$y_m = \gamma(\mathbf{x}) = \gamma(x_1, x_2, \dots, x_m) \\ = \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \left\{ \frac{\prod_{i=1}^m A_i^{n_i}(x_i)}{\sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \prod_{i=1}^m A_i^{n_i}(x_i)} \right\} \theta^{n_1, n_2, \dots, n_m} \quad (2-1)$$

$$= \sum_{n=1}^M \left\{ \frac{\Omega^n(\mathbf{x})}{\sum_{n=1}^M \Omega^n(\mathbf{x})} \right\} \theta^n = \sum_{n=1}^M \{ \xi^n(\mathbf{x}) \} \theta^n \quad (2-2)$$

여기서 편의상 다음의 표현을 사용한다.

$$\mathbf{n} \equiv (n_1, n_2, \dots, n_m) \quad (2-3)$$

$$\mathbf{M} \equiv (M_1, M_2, \dots, M_m) \quad (2-4)$$

$$\sum_{n=1}^M \equiv \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \quad (2-5)$$

$$\Omega^n(\mathbf{x}) \equiv \prod_{i=1}^m A_i^{n_i}(x_i) \quad (2-6)$$

여기서 $\xi^n(\mathbf{x}) \equiv \frac{\Omega^n(\mathbf{x})}{\sum_{n=1}^M \Omega^n(\mathbf{x})}$ 를 퍼지 기저함수라

한다.

2. 제곱 코사인 소속함수

타입 II 퍼지 시스템은 일반적으로 가우시안 소속함수나 삼각형 소속함수, 사다리꼴 소속함수를 이용한다. 본 논문에서는 새로이 다음의 <그림 1>과 같은 식 (3)의 제곱코사인 소속함수를 제안한다.

$$A_i^n(x_i) = \begin{cases} \cos^2 \frac{\pi(x_i - \sigma_i^n)}{2d_i^n} & \text{if } \sigma_i^{n-1} \leq x_i < \sigma_i^n \\ \cos^2 \frac{\pi(x_i - \sigma_i^n)}{2d_i^n} & \text{if } \sigma_i^n \leq x_i \leq \sigma_i^{n+1} \\ 0 & \text{그밖에} \end{cases} \quad (3)$$

여기서 $d_i^n \equiv \sigma_i^{n+1} - \sigma_i^n$ ($n=1, \dots, M_i$)

제안한 제곱 코사인 소속함수는 다음과 같은 특징을 갖는다.

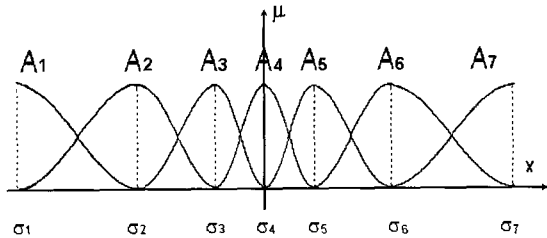


그림 1. 입력영역과 제곱코사인 소속함수
Fig 1. The universe of discourse and squared-cosine membership functions.

(1) 삼각형이나 사다리꼴 소속함수와 달리 제곱코사인 소속함수는 전 지역에서 미분 가능하다. 특히 사다리꼴 접합부인 $\sigma_i^{n-1}, \sigma_i^n, \sigma_i^{n+1}$ 에서도 소속함수는 미분가능하며 이는 다음과 같은 과정으로 확인할 수 있다.

$$\lim_{x_i \rightarrow \sigma_i^n -} \frac{dA_i^n(x_i)}{dx_i} = \lim_{x_i \rightarrow \sigma_i^n +} \frac{dA_i^n(x_i)}{dx_i}$$

$$\lim_{x_i \rightarrow \sigma_i^{n-1} -} \frac{dA_i^n(x_i)}{dx_i} = \lim_{x_i \rightarrow \sigma_i^{n-1} +} \frac{dA_i^n(x_i)}{dx_i}$$

$$\lim_{x_i \rightarrow \sigma_i^{n+1} -} \frac{dA_i^n(x_i)}{dx_i} = \lim_{x_i \rightarrow \sigma_i^{n+1} +} \frac{dA_i^n(x_i)}{dx_i}$$

(2) 제곱코사인 소속함수의 경우, 퍼지 기저함수 $\xi^n(\mathbf{x})$ 는 삼각형 함수의 경우와 같이 단순히 소속함수의 곱으로 표현된다.

$$\sum_{n=1}^M \Omega^n(\mathbf{x}) = \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \prod_{i=1}^m A_i^{n_i}(x_i) = 1,$$

$$\xi^n(\mathbf{x}) = \frac{\Omega^n(\mathbf{x})}{\sum_{n=1}^M \Omega^n(\mathbf{x})} = \Omega^n(\mathbf{x}) = \prod_{i=1}^m A_i^{n_i}(x_i).$$

따라서 퍼지 출력은 식 (4)와 같이 단순화된다.

$$y_m = \mathcal{J}(\mathbf{x}) = \mathcal{J}(x_1, x_2, \dots, x_m) = \sum_{n=1}^M \{ \xi^n(\mathbf{x}) \} \theta^n$$

$$= \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \left\{ \frac{\prod_{i=1}^m A_i^{n_i}(x_i)}{\sum_{n_1=1}^M \sum_{n_2=1}^M \dots \sum_{n_m=1}^M \prod_{i=1}^m A_i^{n_i}(x_i)} \right\} \theta^{n_1, n_2, \dots, n_m}$$

$$= \sum_{n=1}^M \left\{ \frac{\Omega^n(\mathbf{x})}{\sum_{n=1}^M \Omega^n(\mathbf{x})} \right\} \theta^n = \sum_{n=1}^M \{ \Omega^n(\mathbf{x}) \} \theta^n \quad (4)$$

식 (2-2)과 비교하면 식 (4)는 퍼지 기저함수의 분모가 없는 형태로 빠른 퍼지 추론과 온라인 학습에 적합하다는 것을 알 수 있다. 이 특성의 증명은 부록에서 한다.

식 (3) 가우시안 함수는 하나의 소속함수를 나타내기 위하여 중앙값과 폭의 두 개의 값을 필요로 하지만 제곱코사인 소속함수는 경계값 σ_i^n 만으로 표현할 수 있고 소속함수간 불필요한 겹침이 발생하지 않는다.

III. 온라인 동정

본 논문에서는 그레디언트 하강법과 확장 칼만필터 방식, 그리고 이 둘을 결합한 세 가지 방식을 이용하여 퍼지 시스템을 온라인으로 모델링한다. 편의상 그레디언트 하강법(the gradient search algorithm)을 GS, 확장 칼만필터법 (extended Kalman filter)을 EKF, 이 두 가지 방법을 결합한 하이브리드 방식을 HYB로 표시한다. 각각의 방법을 설명하고 장단점을 비교하기 전에 퍼지 시스템의 초기 구현에 대하여 설명한다.

1. 퍼지 모델의 초기 구현

퍼지 모델에서 각 축에 대한 소속함수의 수는 미리 주어진 것으로 하고 i 축에서의 집합의 수를 M_i 로 표시한다. 이제 M_i 개의 제곱코사인형 소속함수가 i 축

입력공간 U_i 을 모두 포함하도록 설계한다. 즉 전진부 퍼지집합의 변수 $\sigma_i^1, \dots, \sigma_i^{M_i}$ 을 U_i 위에 등간격으로 위치시킨다. 이 변수들은 온라인 모델링 과정에서 변하게 되지만 그 순서는 변하지 않도록 한다.

이 같은 전진부 초기 구현이후 후진부는 임의의 값을 이용하도록 하고 전체 퍼지 시스템은 다음의 식(5)와 같이 초기 구현된다.

$$R^{n_1 n_2 \dots n_m} : \text{If } x_1 \text{ is } A_1^{n_1} \text{ and } x_2 \text{ is } A_2^{n_2}, \dots, x_m \text{ is } A_m^{n_m}$$

then y is $\theta^{n_1, n_2, \dots, n_m}$ (5)

($n_1=1, \dots, M_1, n_2=1, \dots, M_2, \dots, n_m=1, \dots, M_m$)

또 이는 식 (6)과 같이 입출력 형태로 나타난다.

$$y_m = \mathcal{J}(\mathbf{x}, \boldsymbol{\theta}) = \mathcal{J}(x_1, x_2, \dots, x_m, \boldsymbol{\theta})$$

$$= \sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \theta^{n_1, n_2, \dots, n_m} \prod_{i=1}^m A_i^{n_i}(x_i, \sigma_i^{n_i-1}, \sigma_i^{n_i}, \sigma_i^{n_i+1})$$
 (6)

여기서 전진부 파라미터 σ_i^k ($i=1, \dots, m, k=1, \dots, M_i$)과 후진부 파라미터 $\theta^{n_1, n_2, \dots, n_m}$ ($n_i=1, \dots, M_i, i=1, \dots, m$)는 온라인 모델링 과정에서 그 값을 설정하게 된다. 편의상 설정할 파라미터를 모두 모아 다음과 같은 하나의 벡터로 표현한다.

$$\boldsymbol{\theta} = (\boldsymbol{\theta}_c^T, \boldsymbol{\theta}_p^T)^T (L \times 1 \text{ 행렬})$$

여기서 $\boldsymbol{\theta}_c^T = (\theta^{1,1, \dots, 1}, \dots, \theta^{M_1, M_1, \dots, M_m})$

→ 후진부 파라미터 ($L_c \equiv \prod_{i=1}^m M_i$ 개의 성분)

$$\boldsymbol{\theta}_p^T = (\sigma_1^1, \dots, \sigma_1^{M_1}, \sigma_2^1, \dots, \sigma_2^{M_2}, \dots, \sigma_m^1, \dots, \sigma_m^{M_m})$$

→ 전진부 파라미터($L_p \equiv \sum_{i=1}^m (M_i - 2)$ 개의 성분)

(σ_i^1 과 $\sigma_i^{M_i}$ 는 입력 영역의 끝점으로 고정시킨다.)

파라미터 $\boldsymbol{\theta}$ 의 총 개수는 $L \equiv L_c + L_p = \prod_{i=1}^m M_i + \sum_{i=1}^m (M_i - 2)$ 이다.

2. 그레디언트 하강법(GS)

그레디언트 하강법은 가장 널리 사용되는 온라인 학습방법으로 다음의 식 (7)과 같이 오차의 그레디언트의 음의 방향으로 파라미터를 설정한다.

$$\hat{\boldsymbol{\theta}}(t+1) = \hat{\boldsymbol{\theta}}(t) - \eta \frac{\partial E}{\partial \boldsymbol{\theta}(t)} + \alpha (\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t-1)) \quad (7)$$

여기서 t 는 학습순간을 나타내며 η 는 학습률, α 는 모멘트를 나타낸다. 전진부 파라미터 $\hat{\boldsymbol{\theta}}_p(t)$ 과 후진부 파라미터 $\hat{\boldsymbol{\theta}}_c(t)$ 에 대하여 각기 다른 학습률 η_p 과 η_c 을 이용한다. 오차함수는 다음과 같이 정의된다.

$$E = \frac{1}{2} (y_d(t) - y_m(t))^2$$

여기서 y_d 는 원하는 출력 값이고 y_m 는 퍼지모델의 출력이다. 식 (7)에서 파라미터에 대한 오차함수의 그레디언트는 다음과 같이 나타낼 수 있다.

$$\frac{\partial E}{\partial \boldsymbol{\theta}(t)} = -(y_d(t) - y_m(t)) \frac{\partial y_m(t)}{\partial \boldsymbol{\theta}(t)}$$

여기서 $\frac{\partial y_m(t)}{\partial \boldsymbol{\theta}(t)} = \frac{\partial \mathcal{J}}{\partial \boldsymbol{\theta}(t)}$ 는 수학적으로 전개가능

하며 수치적으로도 계산할 수 있다. 그레디언트 하강법을 나타내는 식 (7)은 벡터로 표현되어 있지만 실제로는 성분간의 덧셈과 곱셈만을 이용하여 구현한다. 따라서 그레디언트 하강법은 다른 온라인 학습 방법에 비해 연산의 부담이 적고 쉽게 구현할 수 있다는 장점을 갖는다. 반면 단점으로는

- (1) 학습률이 고정되어있고 알고리즘이 국부 최소점에 빠지기 쉽기 때문에 학습이 늦음
- (2) 성능이 경험에 의해 결정되는 학습률과 모멘텀에 크게 의존함
- (3) 알고리즘이 과거의 데이터보다는 현재의 데이터에 크게 의존하기 때문에 측정 노이즈에 민감함 등을 들 수 있다.

3. 확장 칼만필터 (EKF)

비선형 학습방법 중 또 하나의 방식은 비선형 시스템을 특정점을 중심으로 선형화하고 선형 최적 설정(linear optimal estimation)을 적용하는 방식으로 그 대표적인 방식이 확장 칼만필터 방식이다. 확장 칼만필터는 매 샘플링 순간에 비선형시스템의 선형화된 모델에 칼만필터 방식을 적용하는 것으로 선형 칼만필터 방식의 확장으로 볼 수 있다. 본 절에서는 이 같은 확장 칼만필터를 퍼지모델의 온라인 설정에 사용한다.

우선 다음과 같은 비선형 이산 시스템을 고려한다.

$$\mathbf{x}(t+1) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{w}(t) \quad (8)$$

$$\mathbf{y}(t+1) = \mathbf{h}(\mathbf{x}(t)) + \mathbf{v}(t) \quad (9)$$

여기서 $\mathbf{x}(t)$ 과 $\mathbf{y}(t)$ 는 각 상태변수와 관측변수를 나타내고 $\mathbf{f}(\cdot)$ 과 $\mathbf{h}(\cdot)$ 는 시불변 비선형 함수이다.

또 $\mathbf{w}(t)$ 와 $\mathbf{v}(t)$ 는 평균이 0이고 분산이 각각 $\mathbf{Q}(t)$ and $\mathbf{R}(t)$ 인 독립 가우시안 백색잡음이다. 확장 칼만필터는 $t-1$ 순간까지의 관측값에 근거하여 $\hat{\mathbf{x}}(t|t-1)$ 을 추정하며 그 알고리즘은 다음과 같은 반복형 식으로 표현된다^[12, 13].

$$\hat{\mathbf{x}}(t+1|t) = \mathbf{f}(\hat{\mathbf{x}}(t|t)) \quad (10-1)$$

$$\hat{\mathbf{x}}(t|t) = \hat{\mathbf{x}}(t|t-1) + \mathbf{K}(t)[\mathbf{y}(t) - \mathbf{h}(\hat{\mathbf{x}}(t|t-1))] \quad (10-2)$$

$$\mathbf{K}(t) = \mathbf{P}(t|t-1) \mathbf{H}(t)^T [\mathbf{H}(t) \mathbf{P}(t|t-1) \mathbf{H}(t)^T + \mathbf{R}(t)]^{-1} \quad (10-3)$$

$$\mathbf{P}(t+1|t) = \mathbf{F}(t) \mathbf{P}(t|t) \mathbf{F}(t)^T + \mathbf{Q}(t) \quad (10-4)$$

$$\mathbf{P}(t|t) = \mathbf{P}(t|t-1) - \mathbf{K}(t) \mathbf{H}(t) \mathbf{P}(t|t-1) \quad (10-5)$$

여기서 행렬 $\mathbf{F}(t)$ 과 $\mathbf{H}(t)$ 는 다음과 같이 정의되며 식 (10)의 연산이 가능하도록 적당한 크기를 갖는다.

$$\mathbf{F}(t) = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right)_{\mathbf{x} = \hat{\mathbf{x}}(t|t)} \quad (11-1)$$

$$\mathbf{H}(t) = \left(\frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right)_{\mathbf{x} = \hat{\mathbf{x}}(t|t-1)} \quad (11-2)$$

확장 칼만필터 역시 $\mathbf{f}(\mathbf{x}(t))$ 과 $\mathbf{h}(\mathbf{x}(t))$ 의 선형 모델에서 상태를 추정하는 방식이므로 그래디언트 하강법에서와 같이 국부 최소점에 빠질 수도 있다. 그럼에도 불구하고 그래디언트 방식에 비해서 훌륭한 수렴성을 갖고 있는 것이 보고되어 있다^[12, 13].

확장 칼만필터를 퍼지 시스템 (6)의 동정에 적용하기 위해 퍼지 시스템을 다음과 같은 상태공간 방정식과 관측식으로 표현한다.

$$\boldsymbol{\theta}(t+1) = \boldsymbol{\theta}(t) \quad (12-1)$$

$$y_d(t) = \mathcal{J}(\boldsymbol{\theta}(t)) + v(t) \quad (12-2)$$

$$= y_m(t) + v(t)$$

여기서 t 는 시간을 나타낸다. $v(t)$ 는 분산이 λ 인 백색잡음으로 가정한다. 확장 칼만필터를 식 (12)의 퍼

지 시스템 설정에 적용하면 다음과 같은 반복형 식을 얻을 수 있다.

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \mathbf{K}(t) [y_d(t) - \hat{y}_m(t)] \quad (L \times 1) \quad (13-1)$$

$$\mathbf{K}(t) = \frac{\mathbf{P}(t-1) \mathbf{H}(t)^T}{\mathbf{H}(t) \mathbf{P}(t-1) \mathbf{H}(t)^T + \lambda} \quad (L \times 1) \quad (13-2)$$

$$\mathbf{P}(t) = \mathbf{P}(t-1) - \mathbf{K}(t) \mathbf{H}(t) \mathbf{P}(t-1) \quad (L \times L) \quad (13-3)$$

여기서 $\mathbf{P}(t|t) = \mathbf{P}(t+1|t)$, $\hat{\boldsymbol{\theta}}(t|t) = \hat{\boldsymbol{\theta}}(t+1|t)$ 이므로 편의상 $\mathbf{P}(t) = \mathbf{P}(t|t)$, $\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t|t)$ 로 표시하였다. $\hat{y}_m(t)$ 는 시간 $t-1$ 에서의 파라미터에 근거한 퍼지시스템의 출력을 나타낸다. 즉

$$\hat{y}_m(t) = \mathcal{J}(\mathbf{x}(t), \hat{\boldsymbol{\theta}}(t-1)).$$

식 (11-2)로부터 $\mathbf{H}(t)$ 는 다음과 같이 주어진다.

$$\mathbf{H}(t) = \left(\frac{\partial \mathcal{J}}{\partial \boldsymbol{\theta}} \right)_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}(t-1)} = \left(\frac{\partial \mathcal{J}}{\partial \theta_1}, \frac{\partial \mathcal{J}}{\partial \theta_2}, \dots, \frac{\partial \mathcal{J}}{\partial \theta_L} \right)_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}(t-1)} \quad (1 \times L)$$

여기서 $\frac{\partial \mathcal{J}}{\partial \theta_i}$ ($i=1, \dots, L$)는 부록 B와 같다. 또 분산 λ 는 일반적으로 미리 알 수 없으므로 추정하여 사용할 수도 있다. 그러나 본 논문에서는 편의상 λ 를 1로 하였고 이는 모든 샘플데이터를 같은 중요도로 취급하는 것을 의미한다.

확장 칼만필터의 구현할 때, 퍼지 규칙이나 입력의 수가 많은 경우 행렬 $\mathbf{P}(t)$ ($L \times L$)이 너무 커서 식 (13)의 반복식의 구현이 어려운 경우가 발생한다. 이 경우 모든 파라미터 $\boldsymbol{\theta}$ ($L \times 1$)를 모두 같이 설정하기는 어렵고 이를 몇 개의 부벡터로 나누고 이 부벡터를 순서대로 설정하여야 한다. 예컨대 $\boldsymbol{\theta} = (\boldsymbol{\theta}_c^T, \boldsymbol{\theta}_p^T)^T$ 에서 $\hat{\boldsymbol{\theta}}_c$ 값을 이용하여 먼저 $\hat{\boldsymbol{\theta}}_p$ 을 설정한 후 새로 설정한 $\hat{\boldsymbol{\theta}}_p$ 가 실제의 $\boldsymbol{\theta}_p$ 라는 가정하에 $\hat{\boldsymbol{\theta}}_c$ 를 설정하여야 한다.

4. 하이브리드 방식 (HYB)

그래디언트 하강법은 스칼라 합과 곱으로만 구성되어 있어 연산의 부담이 적지만 성능이 학습률과 모델

템에 의존하는 문제를 가지고 있다. 반면 확장 칼만 필터 방식은 그레디언트 하강법보다 안정하지만 연산이 너무 많은 단점을 갖는다. 하이브리드 방식은 두 가지 방식을 결합한 것으로 확장 칼만 필터보다 연산의 부담이 적고 그레디언트 하강법보다 좋은 성능을 갖는다. 식 (6)에서 본 바와 같이 퍼지 모델은 후진부 파라미터에 대해서는 선형이며 전진부 파라미터에 대해서는 비선형이다. 따라서 하이브리드 방식에서는 후진부에 대해서는 확장칼만 필터를 이용하고, 전진부에 대해서는 그레디언트 방식을 이용한다. 후진부의 경우 선형이므로 확장 칼만 필터는 그레디언트 하강법보다 빠른 수렴을 보장할 수 있고, 전진부의 경우 확장 칼만 필터 역시 근사치만을 제공하므로 그레디언트 방식을 이용하여 연산의 부담을 줄이는 방식이다.

즉 하이브리드 방식은 다음과 같이 구현한다.

파라미터가 $\theta = (\theta_c^T, \theta_p^T)^T$ 일 때

$$\hat{\theta}_p(t+1) = \hat{\theta}_p(t) - \eta_p \frac{\partial E}{\partial \theta_p(t)} + \alpha (\hat{\theta}_p(t) - \hat{\theta}_p(t-1)) \quad (L_p \times 1) \quad (14-1)$$

$$\hat{\theta}_c(t) = \hat{\theta}_c(t-1) + K_c(t) [y_d(t) - \hat{y}_m(t)] \quad (L_c \times 1) \quad (14-2)$$

$$K_c(t) = \frac{P_c(t-1) H_c(t)^T}{H_c(t) P_c(t-1) H_c(t)^T + 1} \quad (L_c \times 1) \quad (14-3)$$

$$P_c(t) = P_c(t-1) - K_c(t) H_c(t) P_c(t-1) \quad (L_c \times L_c) \quad (14-4)$$

IV. 컴퓨터 모의실험

본 장에서는 몇 개의 예를 통해 제안한 알고리즘의 타당성을 확인한다. 예제는 기존의 논문 [6-9]에서 이용하여 제안한 온라인 방식과 기존의 방식의 성능을 비교한다. 비교를 위해 비교의 척도는 원 논문에서 사용한 척도를 사용하며 따라서 예마다 다르지만 어느 경우나 척도 값이 작을수록 모델링 성능이 좋은 것이다.

1. 예제 1

모델링할 시스템은 참고문헌 [6-9]에서 사용한 식 (15)의 비선형 함수이다.

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2, \quad 1 \leq x_1, x_2 \leq 5 \quad (15)$$

100개의 샘플을 $1 \leq x_1, x_2 \leq 5$ 영역에서 무작위로 추출하여 사용하며, 제안한 3가지 방법에 의하여 퍼지 모델을 온라인으로 구성한다. 본 예제에서 입력공간은 각축이 5개의 퍼지 규칙을 갖는 것으로 하고 이후의 예제에서도 이를 이용한다. 즉 $M_1 = M_2 = \dots = M_m = 5$ 이다. 그레디언트 하강법의 경우 η_p, η_c, α 를 각 0.2, 0.6, 0.2로 하고 확장 칼만 필터의 경우 $P(0)$ 는 $10I$ 로 한다. 하이브리드 방식의 경우 $\eta_p, \alpha, P_c(0)$ 는 각 0.5, 0.2, $10I$ 로 한다. 본 예에서는, 다음과 같은 평균제곱오차 (MSE, Mean Squared Error)를 성능지수로 사용한다.

$$\text{성능지수} = \text{MSE} = \overline{e^2} = \frac{1}{n} \sum_{i=1}^n (y_d(i) - y_m(i))^2$$

<표 1>에서 본 논문에서 사용한 3가지 성능지수와 다른 방식과를 비교하였다. [6, 7]의 논문은 수천 세대 (epoch)에 걸쳐 만든 오프라인 모델인 반면 본 논문의 모델은 1세대 학습으로 만든 온라인 모델이다.

여기서 온라인 학습방식인 확장 칼만 필터 방식이

표 1. 성능비교 (예제 1)
Table 1. COMPARISON OF PERFORMANCE (EXAMPLE 1).

모델	성능지수
Kim et al ^[6]	0.0197
Sugeno and Yasukawa ^[7]	0.0790
그레디언트 하강법	0.0247
확장 칼만 필터	0.0193
하이브리드	0.0255

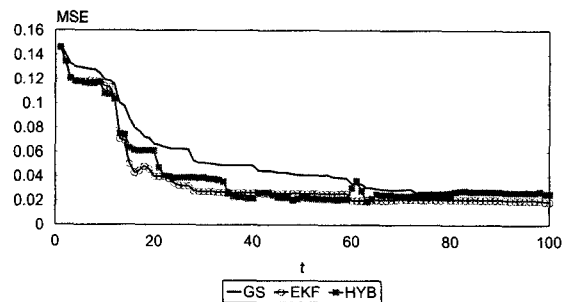


그림 2. 성능지수 대 시간(예제 1)
Fig. 2. MSE vs. time(Example 1).

기존의 오프라인 모델보다 우수한 성능을 보이는 것은 주목할 만 한다. <그림 2>는 시간에 대한 성능 지수의 변화를 나타낸다. 이 그림은 모델링은 온라인으로 진행되었지만 참고를 위해 모든 데이터를 이용하여 평균제곱오차를 계산한 것이다. <그림 3>은 원 함수와 퍼지 모델로 구성된 시스템간을 비교하고 있다.

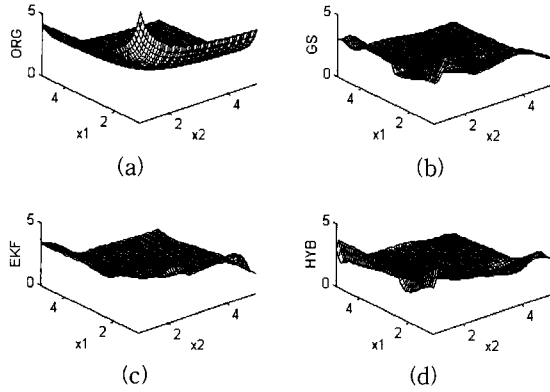


그림 3. 원함수와 모델링 함수 (예제 1) (a) 원함수 (b)그레디언트 하강법 (c) 확장칼만필터 (d) 하이브리드 법

Fig. 3. Original system and on-line reconstructed fuzzy systems. (Example 1) (a) Original (b) GS (c) EKF (d) HYB

2. 예제 2

다음과 같이 두 개의 입력과 하나의 출력을 갖는 함수를 고려한다.

$$y = (2 + x_1^{1.5} - 1.5 \sin(3x_2))^2, \quad 0 \leq x_1, x_2 \leq 3 \quad (16)$$

본 예는 [9]에서 인용한 것으로 예제 1의 경우처럼 $0 \leq x_1, x_2 \leq 3$ 구역에서 100개의 데이터를 얻어 온라인 퍼지 모델링을 수행한다. 그레디언트 하강법의 경우 η_p, η_c, α 를 각 0.005, 0.085, 0.005로 한다. 하이브리드 방법의 경우 η_p 과 α 를 모두 0.001로 한다. 이 예는 [9]에서 인용한 것이므로 [9]에서 사용된 다음의 성능지수를 그대로 사용한다.

$$\text{성능지수} = \frac{\sqrt{\sum_{i=1}^n (y_d(i) - y_m(i))^2}}{\sum_{i=1}^n |y_d(i)|}$$

다음의 표는 [9]의 결과와 제안된 방법을 비교한 것이다.

표 2. 성능비교(예제2)

Table 2. COMPARISON OF PERFORMANCE (EXAMPLE 2).

모델	성능지수
Lin and Cunningham ^[9]	0.00497
그레디언트 하강법	0.08400
확장칼만필터	0.02090
하이브리드	0.02500

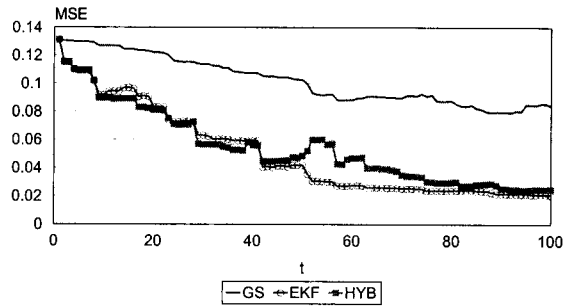


그림 4. 성능지수 대 시간 (예제 2)

Fig. 4. PM vs. time (Example 2).

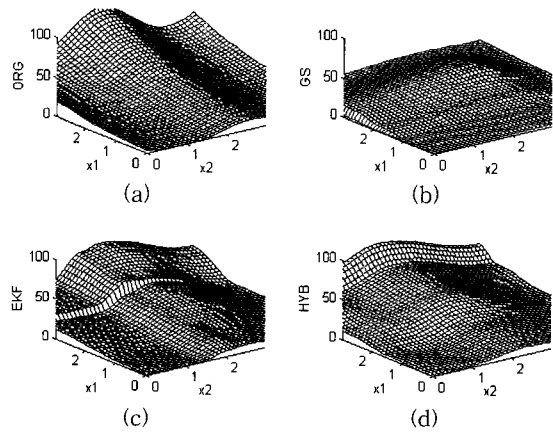


그림 5. 원함수와 모델링 함수 (예제 2) (a) 원함수 (b) 그레디언트 하강법 (c) 확장 칼만필터 (d) 하이브리드 법

Fig. 5. Original system and on-line reconstructed fuzzy systems(Example 2). (a) Original (b) GS (c) EKF (d) HYB

본 예제에서는 제안된 방법이 기존의 방법을 능가하지는 못했다. 그러나 [9]의 결과는 2800세대의 학습을 통해서 얻어진 것이고 본 논문의 방식은 온라인 (1세대)을 통해 얻어진 것이므로 본 논문에서 제안된 방식은 실시간 데이터를 처리하여야 하는 분야에서 적당하

다고 할 수 있다. <그림 4>는 시간에 대한 성능 지수의 변화를 나타낸다. <그림 5>는 원 함수와 퍼지 모델로 구성된 시스템간을 비교하고 있다.

V. 결 론

본 논문에서는 퍼지 모델의 온라인 동정법을 제안하였고 컴퓨터 모의 실험을 통하여 타당성을 확인하였다. 우선 제곱코사인 소속함수를 제안하여 파라미터의 수를 줄이고 온라인 학습을 쉽게 하였다. 이후 세 가지 학습 방법(그라디언트 하강법, 확장 칼만필터법, 하이브리드법)을 제안하고 이를 이용하여 온라인 퍼지 모델링을 수행하였다.

본 논문에서 제안된 방식은 비선형 적응 제어나 신호처리와 같이 실시간의 데이터 처리를 필요로 하는 곳에 적용될 수 있는 방법으로 많은 응용이 가능하다.

부 록

$x = (x_1, x_2, \dots, x_m)$ 에 대하여, $\sigma_1^{k_1} \leq x_1 < \sigma_1^{k_1+1}$, $\sigma_2^{k_2} \leq x_2 < \sigma_2^{k_2+1}$, \dots , $\sigma_m^{k_m} \leq x_m < \sigma_m^{k_m+1}$ 인 경우, i 번째의 x_i 는 한 순간 다음의 두 개의 소속함수만을 활성화화시킨다.

$$A_i^{k_i}(x_i) = \cos^2 \frac{\pi(x_i - \sigma_i^{k_i})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})}$$

$$A_i^{k_i+1}(x_i) = \cos^2 \frac{\pi(x_i - \sigma_i^{k_i+1})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})}$$

$$A_i^p(x_i) = 0 \quad (p=1, \dots, M_i, \quad p \neq k_i, \quad k_i+1)$$

증

$$\sum_{n_i=1}^{M_i} A_i^{n_i}(x_i) = A_i^{k_i}(x_i) + A_i^{k_i+1}(x_i) = \cos^2 \frac{\pi(x_i - \sigma_i^{k_i})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})} + \cos^2 \frac{\pi(x_i - \sigma_i^{k_i+1})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})}$$

$$= \cos^2 \frac{\pi(x_i - \sigma_i^{k_i})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})} + \cos^2 \left\{ \frac{\pi(x_i - \sigma_i^{k_i})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})} - \frac{\pi}{2} \right\}$$

$$= \cos^2 \frac{\pi(x_i - \sigma_i^{k_i})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})} + \left\{ -\sin \frac{\pi(x_i - \sigma_i^{k_i})}{2(\sigma_i^{k_i+1} - \sigma_i^{k_i})} \right\}^2 = 1$$

따라서

$$\sum_{n_1=1}^{M_1} \sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \prod_{i=1}^m A_i^{n_i}(x_i)$$

$$= \sum_{n_1=1}^{M_1} A_1^{n_1}(x_1) \left(\sum_{n_2=1}^{M_2} \dots \sum_{n_m=1}^{M_m} \prod_{i=2}^m A_i^{n_i}(x_i) \right)$$

$$= \sum_{n_1=1}^{M_1} A_1^{n_1}(x_1) \sum_{n_2=1}^{M_2} A_2^{n_2}(x_2) \left(\sum_{n_3=1}^{M_3} \dots \sum_{n_m=1}^{M_m} \prod_{i=3}^m A_i^{n_i}(x_i) \right) =$$

$$\dots$$

$$= \prod_{i=1}^m \left\{ \sum_{n_i=1}^{M_i} A_i^{n_i}(x_i) \right\} = 1$$

참 고 문 헌

[1] J. C. Bezdek, "Editorial: Fuzzy Models - What Are They, and Why ?", IEEE Trans. Fuzzy Systems, vol. 1, No. 1, pp. 1-6, Feb. 1993.

[2] W. Pedrycz, "An identification algorithm in fuzzy relational systems," Fuzzy Sets and Systems, vol. 13, pp.153-167, 1984.

[3] W. Pedrycz, Fuzzy Control and Fuzzy Systems, John Wiley and Sons, New York, 1993.

[4] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," IEEE Trans. Systems Man Cybernet, vol. 15, No. 1, pp.116-132, 1985.

[5] M. Sugeno and G. T. Kang, "Structure identification of fuzzy model", Fuzzy Sets and Systems, vol. 28, pp 15-33, 1988.

[6] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling", IEEE Trans. Fuzzy Systems, Vol. 1, No. 1, pp 7-31, Feb. 1993.

[7] E. Kim, M. Park, S. Ji and M. Park, "A new approach to fuzzy modeling," IEEE Trans. Fuzzy Systems, vol. 5, No. 3, pp. 328-337, Aug, 1997.

[8] E. Kim, M. Park, S. Kim and M. Park, "A transformed input-domain approach to fuzzy modeling," IEEE Trans. Fuzzy Systems, vol. 6, no. 4, pp. 596-604, 1998.

[9] 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. Neural Networks, vol. 9, no. 5, pp 756-765, 1998.
- [10] Y. Lin and G. A. Cunningham III, "A new approach to fuzzy-neural modeling," IEEE Trans. Fuzzy Systems, Vol. 3, No. 2, pp 190-197, May, 1995.
- [11] M. Sugeno and K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system", Fuzzy Sets and Systems, vol. 42, pp 315-334, 1991.
- [12] G. Chen, Approximate Kalman Filtering, World Scientific Publishing Co., 1993.
- [13] B. D. O. Anderson and J. B. Moore, Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall, 1979.

 저 자 소 개



金 殷 泰(正會員)

1992년 2월 : 연세대학교 전자공학과 졸업(공학사). 1994년 2월 : 연세대학교 전자공학과 석사과정 졸업(공학석사). 1999년 2월 : 연세대학교 전자공학과 박사과정 졸업(공학박사). 1999년 3월~2002년 2

월 : 국립한경대학교 제어계측공학과 조교수. 2002년 3월~현재 : 연세대학교 전기전자공학부 조교수. 1998년~현재 : IEEE TFS, IEEE SMC, IEEE CAS, FSS등에서 심의위원으로 활동중. <주관심분야 : Computational intelligence, 지능형 서비스로보틱스>