

# 퍼지서포트벡터기계의 시계열자료 패턴분류를 위한 퍼지소속함수에 관한 연구

## On the Fuzzy Membership Function of Fuzzy Support Vector Machines for Pattern Classification of Time Series Data

이수용\*

Soo-Yong Lee

\* 연세대학교 과학기술대학 컴퓨터정보통신공학부

### 요 약

본 논문에서는 FSVM(Fuzzy Support Vector Machine)의 퍼지소속함수를 새롭게 제안한다. SVM의 완화변수(slack-variable)에 퍼지소속함수를 결합하는 FSVM은 주어진 데이터베이스의 특성이 반영되어 안정적으로 분류성능을 향상시킬 수 있는 퍼지소속 함수를 필요로 한다. 시계열 자료의 패턴분류 성능을 비교하기 위하여 SVM, FSVM(1), 그리고 제안하는 FSVM(2) 등의 분류모델들을 비교 실험하였다. 사용한 데이터베이스는 한국금융시장의 시계열 경제지표 지수들이다.

### Abstract

In this paper, we propose a new fuzzy membership function for FSVM(Fuzzy Support Vector Machines). We apply a fuzzy membership to each input point of SVM and reformulate SVM into fuzzy SVM (FSVM) such that different input points can make different contributions to the learning of decision surface. The proposed method enhances the SVM in reducing the effect of outliers and noises in data points. This paper compares classification and estimated performance of SVM, FSVM(1), and FSVM(2) model that are getting into the spotlight in time series prediction.

Key Words : SVM, FSVM, Fuzzy Support Vector Machines, Pattern Classification, Time Series Prediction

### 1. 서론

SVM(Support Vector Machine)은 1995년 Vapnik[12]에 의해 처음 소개되었다. 학습패턴들을 서로 다른 두 개의 Class로 이원 분류할 때 판별기준이 되는 결정함수를 SVM 학습 알고리즘을 이용하여 최적의 판별함수를 찾는 것이다. 즉, 학습패턴들을 고차원의 특징공간으로 사상시킨 후 두 분류 사이의 여백을 최대화시키는 최적의 결정함수(hyperplane)를 찾는 것이다. 이 때 사상에 대한 정보가 없어도 SVM은 특징공간에서 커널이라는 내적함수를 활용하여 원하는 최적의 의사결정함수를 찾는다. 최적의 결정함수는 support vector라는 어떤 학습 데이터들의 결합으로 표현된다.

최근 수 년 동안 SVM은 다양한 분야에서 우수한 패턴분류 모델임을 성공적으로 입증해 왔지만, 각 학습패턴들의 속성들을 확일적으로 학습하는 경향이 있다. 예를 들어, 가장 최근 패턴이 향후 패턴 추세에 영향을 주는 시계열(time series) 데이터들을 학습할 때, 학습패턴들을 순차적으로 학습과정에 영향을 주도록 수정하는 것이 필요하다. 즉 각 학

습패턴들에 따라 최근의 추세형성에 영향을 준 기여정도를 고려하는 것이 바람직하다. 이에 순차적 기여정도에 대한 퍼지소속함수를 정의하고 각 학습패턴에 적용시키면 모든 학습패턴들이 확일적으로 취급되지 않도록 학습시킬 수 있다. 따라서 시계열성 자료들의 학습에서 SVM은 각 데이터들의 순차적 속성이 반영되는 퍼지소속함수와 결합하는 Fuzzy SVM 패턴분류기 모델로의 확장이 필요하다.

C.F.Lin[5]은 퍼지소속함수를 SVM의 slack variable에 적용하여 오분류의 측도인 완화변수(slack variable)들이 퍼지소속함수의 영향을 받아 결정함수(hyperplane)의 기울기를 조정할 때 유연성있는 FSVM을 제안했다. 이에 S.Y.Lee[4,5]은 Lin이 제안한 FSVM의 2차형식의 퍼지소속함수를 고차형식 ( $n \geq 3$ )의 다항함수 차원으로 확장하여 주어진 데이터베이스에 알맞은 퍼지소속함수의 차원을 제시하였으며, 기존의 시계열 분석을 위한 데이터마이닝 모형들과 성능을 비교하여 FSVM이 시계열 패턴분류성능에서 우수함을 입증하였다.

본 논문에서는 시계열자료에 적합한 퍼지소속함수를 새롭게 제안하여 기존의 SVM, FSVM(1)[Lee,4,5]과 본 논문에서 새롭게 제안하는 퍼지소속함수와 결합한 FSVM(2)들의 패턴분류 성능을 비교하였다. 실험에 사용한 시계열 데이터베이스는 한국금융시장의 경제지표 지수들이다.

논문의 구성은 2장에서 최근 수 년 동안 다양한 분야에 연구되고 있는 FSVM의 연구동향을 조사했으며, 3장에서는 FSVM의 개념을 정리하고, 시계열자료들의 패턴 학습에 적

접수일자 : 2007년 10월 16일

완료일자 : 2007년 10월 31일

감사의 글 : 본 연구는 C.G.R. Inc.로 부터 부분적인 지원을 받아 수행되었음.

합한 SVM의 새로운 퍼지소속함수를 제안하였다. 4장에서는 기존의 SVM, FSVM(1), 그리고 FSVM(2)들의 금융지수 패턴분류 성능을 비교한 실험내용들을 정리했으며, 5장에서 본 논문의 결론을 정리하고, 6장에서 향후 연구과제에 대한 설명을 하였다.

### 2. FSVM의 기존 연구

SVM은 최근 몇 년 동안 이론적인 발전뿐만 아니라, 실제 구현되어 다양한 패턴인식 응용분야에 널리 실용되고 있다. 특히 SVM과 Fuzzy 이론을 결합한 FSVM은 데이터마이닝을 비롯한 영상인식, 의료진단 시스템 등의 다양한 패턴분류 모형 연구분야에서 기존의 SVM보다 우수한 성능을 입증하고 있다. 그러나 FSVM을 이용한 금융지수에 대한 지수(index)의 향후 추세를 분류/예측하는 데이터마이닝 분야의 연구는 아직 활발하지 않다.

2001년 T.Inoue[3]는 다원분류를 위해 Fuzzy SVM을 사용했으며, 2002년 C. F. Lin[6]은 입력 패턴들이 서로 각기 차별화된 특성을 반영하도록 각 slack variable에 서로 다른 정도의 소속함수를 부여한 Fuzzy SVM을 제안했다. 2003년 S.Y.Lee[4,5]는 Lin이 제안한 퍼지소속함수의 차원을 일반화하여 한국금융시장의 경제지표인 KOSPI200의 향후 주별 상승/하락 패턴을 예측하였다. 2003년 J.Mill[8]은 support vector를 퍼지화하여 흥채인식하는 FSVM을 발표했다. 2004년 S.Abel[1]는 다원분류 문제를 위해 Fuzzy LP-SVM을 제안했다. 2005년 Y.Wang[13]는 신용불량 고객과 우량고객을 구별하기 위해 이상치(outlier)를 처리하는 Fuzzy SVM를 제안했다. 2005년 E.Spyrou[10]은 퍼지 기저함수로 구성된 커널함수를 이용하여 Fuzzy SVM를 영상처리에 적용하였다. 2006년 H.Xia[14]은 부적절한 특징추출을 위해 Lin[8]이 제안한 Fuzzy SVM을 적용했다. 2006년 X.Shi[9]는 폐암을 발견하기 위해 Lin[8]의 Fuzzy SVM을 CAD에 적용하였다. 2006년 Y.Mao[7]는 폐암 진단을 위해 Fuzzy SVM을 사용했다.

### 3. SVM과 FSVM(Fuzzy SVM)

SVM은 V.N. Vapnik과 AT&T Bell연구소 팀이 개발한 패턴분류 방법으로 최근 수년동안 이론과 응용분야에서 각광을 받아 왔다. 특히 데이터마이닝 분야에도 응용되고 있다. 이 절에서는 V.N. Vapnik [12]이 제안한 SVM과 Lin[6]이 제안한 FSVM에 대해 간단히 소개한다.

#### 3.1. SVM

SVM은 다항식(polynomial), 방사기저함수(Radial Basis Function), 그리고 다층퍼셉트론 분류기(Multi-Layer Perceptron classifiers)의 대안적인 학습 방법으로 학습패턴들을 고차원의 특징공간으로 사상(mapping) 시킨 후 두 분류 사이의 여백(margin)을 최대화시키는 결정함수(hyperplane)를 찾는 것이다. SVM은 신경망등의 인공지능형 모델들과는 다르게 QP(Quadratic Programming) 문제를 해결함으로써 의사결정 함수를 찾는다. 최적의 의사결정함수는 지지벡터(support vector)라는 몇 개의 입력 벡터들의 결합으로 나타낸다. 특히, SVM은 사상에 대한 정보가 없어도 특징공간에서 커널

(kernel)이라는 내적함수를 활용하여 원하는 최적의 의사결정함수를 찾는다. 다음은 SVM에서 사용하는 커널함수이다.

$$K(x_i, x_j) = (x_i^T x_j + 1) : \text{Polynomial Kernel}$$

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|_2^2}{\sigma^2}\right) : \text{RBF Kernel}$$

학습패턴  $S = \{(y_i, x_i) | i = 1, \dots, l\}$ 가 주어졌을 때, 각 훈련데이터  $x_i \in R^N$ 는 두개로 부호화된 부분 중 반드시 한 곳에 속하고, 부호는  $i = 1, \dots, l$ 에 대하여  $y_i \in \{-1, +1\}$ 이다. 입력공간에서 적절한 결정함수의 탐색이 쉽지 않으므로 입력공간의 차원보다 더 높은 차원의 특징공간으로 입력공간을 사상(mapping)시키면 최적의 의사결정함수를 탐색할 수 있게 된다.

$z = \phi(X)$ 이  $R^N$ 에서 특징공간  $Z$ 로의 사상일 때,  $W \cdot Z + b = 0$ 를 만족하는  $(W, b)$ 를 결정함수라 한다. 이때  $W \in Z, b \in R$ 이고,  $X_i$ 는 다음 함수에 의해 분리된다.

$$f(x_i) = \text{sign}(W \cdot Z_i + b) = \begin{cases} +1 & \text{if } y_i = 1 \\ -1 & \text{if } y_i = -1 \end{cases}$$

선형분리 되지 않는 데이터들을 처리하기 위하여, 완화변수(slack variable)  $\xi_i \geq 0$ 를 오분류 척도로 정의하면, 결정함수는  $y_i(W \cdot Z_i + b) \geq 1 - \xi_i, i = 1, \dots, l$ 이다. 여기서  $\xi_i$ 는 의사결정함수를 만족하지 않는  $X_i$ 에 대한 오분류(misclassification) 척도이므로  $\sum_{i=1}^l \xi_i$ 는 훈련집합  $S$ 에 대한 오분류의 척도가 된다. 따라서 최적의 의사결정함수(hyperplane)는 다음과 같이 표현된다.

$$\text{minimize } \frac{1}{2} W \cdot W + C \sum_{i=1}^l \xi_i$$

$$\text{subject to } y_i(W \cdot Z_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l$$

여기서 상수  $C$ 는 결정함수의 마진폭과 분류율의 정도를 조정하는 매개변수이다.

라그랑지안 승수  $a = (a_1, \dots, a_l)$ 에 대하여, QP문제의 처리가 곧 최적의 결정함수를 구하는 해법이 된다.

$$\max W(a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(x_i, x_j)$$

$$\text{subject to } \sum_{i=1}^l y_i a_i = 0, 0 \leq a_i \leq C, i = 1, \dots, l$$

$\phi: R^N \rightarrow Z, \phi(X) = Z$ 일 때, 커널함수  $K(x_i, x_j)$ 가  $Z_i \cdot Z_j = \phi(x_i) \cdot \phi(x_j) = K(x_i, x_j)$ 를 만족하면 최적의 결정함수(hyperplane)은 Mercer의 정리에 의해 다음 두 식을 만족한다.

$$\max W(a) = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j K(x_i, x_j)$$

$$\text{subject to } \sum_{i=1}^l y_i a_i = 0, 0 \leq a_i \leq C, i = 1, \dots, l$$

따라서 결정함수는 다음과 같이 표현된다.

$$f(x_i) = \text{sign}\left(\sum_{j=1}^l a_j h_j K(x_i, x_j) + b\right)$$

### 3.2. FSVM

시계열 데이터들이 추세를 형성할 때 가장 최근 데이터들의 패턴에 영향을 많이 받는다면 순차적 성질을 갖는 퍼지소속함수를 정의하여 각 훈련패턴에 적용하면 패턴학습시 모든 훈련 데이터들이 확률적으로 취급되지 않고 순차 데이터들의 패턴에 따라 영향을 받도록 학습시킬 수 있다. 따라서 SVM 기법은 순차적 속성을 반영하는 퍼지소속함수와 결합하여 확장된 SVM 패턴분류기 모델의 설계가 필요하다. C.F.Lin[6]은 퍼지소속함수를 SVM의 완화변수(slack variable)에 적용하는 FSVM을 제안했다. FSVM의 특징은 SVM이 비선형 분류문제를 해결할 때 퍼지소속함수와 결합한 훈련데이터를 사용함으로 오분류(misclassification)의 단위인 완화(slack) 변수들이 퍼지소속함수의 영향을 받아 결정곡면의 기울기를 조정할 때 유연성을 지니도록 조정하는 것이다. 즉,

$S = \{(y_i, x_i, s_i) | y_i \in \{+1, -1\}, i = 1, \dots, l\}$ 라 하자. 여기서  $\sigma$ 는 임의의 작은 수이며,  $X_i \in R^N$ 는 학습패턴이고,  $y_i \in \{-1, +1\}$ 는 부호화된 목표 데이터이며,  $S_i = \{s_i \in R | 0 \leq s_i \leq 1\}$ 는 퍼지소속 함수값이다.

퍼지소속 함수값  $s_i$ 는 벡터  $X_i$ 가 한  $\alpha_i$ 에 속하는 정도를 표시한 속성이다.  $\xi_i$ 는 SVM에서 오분류에 대한 오차(error)의 척도(measure)이므로  $s_i \xi_i$ 는 서로 다른 가중치를 갖는 새로운 오차의 척도로 변환된다. 따라서 FSVM에서 최적의 결정함수(hyperplane)는 다음 식을 만족하는 해이다.

$$\min \frac{1}{2} WW + C \sum_{i=1}^l s_i \xi_i$$

$$\text{subject to } y_i(W \cdot Z_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l$$

라그랑지안 승수  $\alpha = (\alpha_1, \dots, \alpha_l)$ 에 대하여, 커널함수  $K(x_i, x_j)$ 가  $Z_i \cdot Z_j = \phi(x_i) \cdot \phi(x_j) = K(x_i, x_j)$ 를 만족하면 최적의 결정함수(hyperplane)는 Mercer의 정리에 의해 SVM의 유도과정을 따라서 다음 두 식을 만족한다.

$$\max W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$\text{subject to } \sum_{i=1}^l y_i \alpha_i = 0, 0 \leq \alpha_i \leq s_i C, i = 1, \dots, l$$

### 3.3. Fuzzy Membership Function

C.F.Lin[6]은 시계열 패턴에 대한 퍼지소속 함수  $s_i = f(t_i)$ 를 1차원과 2차원일 때에 한하여 Remark-1을 제안했다.

[Remark-1]

$$s_i = f(t_i) = at_i + b$$

$$s_i = f(t_i) = (1 - \sigma) \left[ \frac{t_i - t_1}{t_i - t_1} \right]^2 + \sigma$$

S.Y.Lee[4,5]는 Lin의 퍼지소속 함수의 차원을 확장하여 차원  $n$ 은 주어진 데이터베이스에 따르는 최적의 차원을 실험을 통해 선택했다. 즉, 소속함수 차원을  $n \geq 3$ 차원으로 다양하게 확장하여 시계열 자료들의 패턴분류 문제에 적용하였는데, 데이터베이스의 특성에 따라  $n \geq 3$ 의 고차 지수일 때 (예를 들어,  $n = 33, n = 187, n = 725, \dots$ ) 오분류율을 낮출 수 있었으며 학습시간을 크게 단축시키는 유용한 결과를 얻었다.

[Remark-2]

$$s_i = f(t_i) = (1 - \sigma) \left[ \frac{t_i - t_1}{t_i - t_1} \right]^n + \sigma \quad (1)$$

본 논문에서는 퍼지소속함수를 (2)와 같이 새롭게 정의하여 SVM과 결합한 새로운 FSVM 패턴분류기 모델을 제안한다.

[Remark-3]

$$s_i = f(t_i) = \frac{1}{1 + \exp\left(a - 2a \left(\frac{t_i - t_1}{t_i - t_1}\right)\right)} \quad (2)$$

$t_1 \leq t_i \leq t_l, a \text{ ascending rate } (0 < a)$

(1)식을 FSVM의 퍼지소속함수로 사용한 패턴분류기 모델을 FSVM(1)이라 하고, (2)식을 FSVM의 퍼지소속함수로 사용한 패턴분류기 모델을 FSVM(2)라 한다.

## 4. 실험

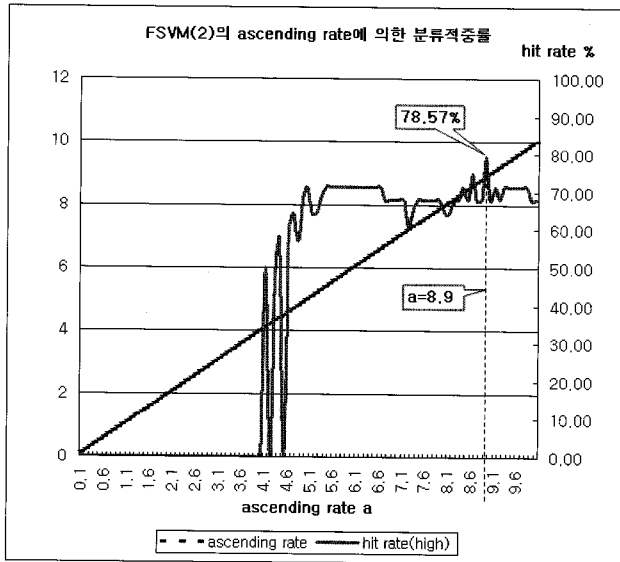
### 4.1. 데이터베이스

FSVM(1)과 FSVM(2) 그리고 기존의 SVM들의 패턴분류 성능을 실험하기 위하여 목표변수는 한국금융지수 KOSPI200 지수의 향후 주별 상승/하락을 예측하도록 설정하고, 학습을 위한 변수들은 국내 및 국외의 89개 금융지수들을 사용하였다. KOSPI, KOSPI200, DOW-JONES, NASDAQ 등의 시가, 고가, 저가, 종가, 거래량, 거래대금, 그리고 회사채수익률, 원달러 환율 등의 기본변수 24개와 기본변수들의 일별 증감 변수 24개, 그리고 기술적 지표 41개들로서 총 89개의 변수들이다. 1999년 1월부터 2003년 7월까지 주별 데이터 237개로 데이터베이스를 구성하였다.

### 4-2. 실험 내용

기존의 SVM과 2가지 퍼지소속함수에 따른 각각의 FSVM(1)과 FSVM(2)의 시계열 자료에 대한 패턴분류 성능을 비교하였다. 주어진 데이터베이스에 따라서 FSVM(1)의 경우 퍼지소속함수 (1)의 차원을  $n = 187$ 로 선택하였고, FSVM(2)에 대한 파라미터인 ascending rate  $a = 8.9$ 로 선택하였다. 그림-1은 FSVM(2)의 ascending rate  $a$ 의 최적화 지수를 찾은 실험결과이다. 실험에서  $0 < a < 5$ 에서는 분류적중률이 거의 0이었지만,  $a = 8.9$ 일 때 FSVM(2)가 최고 78.57%의 분류적중률을 나타내었다.

표 1. FSVM(2)의 ascending rate에 의한 분류



다음 실험 표-2는 SVM, FSVM(1), 그리고 FSVM(2)의 분류적중률 성능 비교 결과이다. 모델들 모두 RBF 커널을 사용했으며, 퍼지소속함수의 파라미터는 FSVM(1)은  $n = 187$ , FSVM(2)는  $a = 8.9$ 를 선택하였고, 학습패턴은 209개, 테스트 패턴은 총 24개로 선택하였는데 + class 12개와 - class 12개로서 각 class별로 50%씩의 비율로 구성하였다. 결정함수의 마진폭과 분류적중률을 조정하는 C는 100~2,000범위 각 500단계별로 실험하여 FSVM(1)과 FSVM(2)가 78.57%의 분류적중률을 나타내었다.

표 2. RBF커널

Classifier	SV	SVC	STD	C	hit rate
SVM	145	19	3.1	1500	60.71
FSVM(1)	11	5	9	2000	78.57
FSVM(2)	8	0	4.9	500	78.57

표-3의 실험 내용은 표-2의 동일한 실험환경에서 RBF 커널함수를 Polynomial 커널함수로 교체한 실험결과이다. 실험내용을 정리하면 FSVM(1)이 75.00%로서 FSVM(2)보다 높은 적중률을 보였다. 그러나 학습된 모델들이 실제 +class를 예측해야 하는 경우에 FSVM(1)은 구해진 최적의 파라미터 집합 2가지들이 서로 상반된 +class와 -class로 예측하였다. Polynomial 커널의 경우 FSVM(1)이 높은 적중률을 제시했지만, C=500일 때 향후 상승(+class)를 예측하였고, C=1500일 때는 향후 하락(-class)를 예측하는 불안정적인 결과를 보였다.

표 3. Polynomial커널

Classifier	SV	SVC	STD	C	hit rate	Prediction
SVM	115	17	3	1000	53.57	+ class
FSVM(1)	11	5	1	500	75.00	+ class
	11	5	1	1500		- class
FSVM(2)	4	0	1	100	68.86	+ class

다음 실험은 RBF커널을 사용하고, 테스트 데이터수를 15개로 구성했을 때, 표-5에서 최고 분류적중률의 파라미터 집합수가 SVM은 5가지, FSVM(1)은 4가지를 보였지만, FSVM(2)의 경우는 최고 분류적중률을 제시하는 파라미터 집합들을 총 80가지 제시함으로써 FSVM(2)가 SVM과 FSVM(1) 보다 안정적인 패턴분류 모델로 활용될 것으로 기대되었다.

표 4. RBF커널

Classifier	SV	SVC	STD	C	hit rate	prdeiction
SVM	186	0	1.2	100	66.67	+ class
	186	0	1.2	500		+ class
	186	0	1.2	1000		+ class
	186	0	1.2	1500		+ class
	186	0	1.2	2000		+ class
FSVM(1)	11	5	8.8	500	80.00	+ class
	11	5	7.9	1500		+ class
	11	5	10	1500		+ class
	11	5	9.2	2000		+ class
FSVM(2)	7	0	5	500	80.00	+ class
	7	0	5.2	500		+ class
	7	0	5.3	500		+ class
	...	...	...	...		+ class
	...	...	...	...		+ class
	...	...	...	...		+ class
	4	0	8.9	2000		+ class
	4	0	8.9	2000		+ class
	4	0	9.1	2000		+ class
	4	0	9.2	2000		+ class

## 5. 결 론

본 논문에서는 SVM의 오분류 척도인 완화변수(slack variable)  $\xi_i$ 에 주어진 시계열 데이터베이스의 속성이 반영 되도록 작용하는 퍼지소속함수  $s_i$ 를 새롭게 정의하여 결합한  $s_i, \xi_i$ 가 오분류 척도로 적용되는 FSVM(2) 모델을 제안하였다.

한국금융지수인 KOSPI200의 향후 주별 상승/하락에 대한 SVM, FSVM(1), 그리고 FSVM(2)의 패턴분류 성능을 실험하였는데 결과는 FSVM(1)과 FSVM(2)가 기존의 SVM보다 우수한 패턴분류 적중률을 보였으며, FSVM(2)가 FSVM(1)보다 안정적으로 많은 최적의 파라미터 집합들을 제시하였다. 또한 FSVM은 주어진 데이터베이스에 대하여 polynomial커널 함수 보다 RBF커널함수를 사용했을 때 안정적인 분류적중률을 보였다. 실험에서 FSVM(1)의 퍼지소속함수에 대한 다항함수의 차원은  $n = 187$ 로 선택하였고, FSVM(2)의 ascending rate은  $a = 8.9$ 로 최적화하였다.

## 6. 향후 연구과제

데이터베이스의 모든 패턴들이 확일적으로 취급되는 기존의 SVM 패턴분류기 모델이 퍼지소속함수와 결합함으로써 각 학습패턴들의 속성들이 서로 다르게 적용되는 Fuzzy Support Vector Machine으로 발전되도록 다양한 퍼지소속함수에 대한 연구가 필요하다.

## 참 고 문 헌

[1] S. Abe, Fuzzy LP-SVMs for Multiclass Problems, *ESANN'2004 proceedings*, pp. 429-434, Bruges (Belgium), 28-30 April 2004.

[2] N. Cristianini, An Introduction to Support Vector Machines, *Cambridge University Press*, 2000.

[3] T. Inoue, S. Abe, Fuzzy support vector machines for pattern classification, *Neural Networks Proceedings*. Vol. 2, pp. 1449-1454, 2001.

[4] S.Y.Lee, S.Y.Sohn, C.E.Kim, Y.B.Lee, Fuzzy Support Vector Machine for Pattern Classification of Time Series Data of KOSPI200 Index, *Journal of Fuzzy Logic and Intelligent Systems*, Vol. 14, No. 1, pp.52-56, 2004.

[5] S.Y.Lee, A Comparative Study of Model Design and Performance Test in Financial Index Pattern Classification System, *Doctoral Thesis*, Yonsei University, 2004.

[6] C. F. Lin, "Fuzzy Support Vector Machines", *IEEE Transactions on Neural Networks*, Vol.13, No.2, March 2002.

[7] Y. Mao, Multiclass Cancer Classification by Using Fuzzy Support Vector Machine and Binary Decision Tree With Gene Selection, *Journal of Biomedicine and Biotechnology*, Vol. 2, pp.160 - 171, 2005.

[8] J. Mill, A. Inoue, An application of fuzzy support vectors, *Fuzzy Information Processing Society, NAFIPS. 22nd International Conference of the North American*, pp.302- 306, 2003.

[9] X. Shi, H.D. Cheng, and L. Hu, Mass Detection and Classification in Breast Ultrasound Images Using Fuzzy SVM, *Proceeding of JCIS*, 2006.

[10] E. Spyrou, G. Stamou, Y. vrithis, S. Kollias, Fuzzy support vector machines for image classification fusing MPEG-7 visual descriptors, *2nd European Workshop, Integration of Knowledge, Semantics and Digital Media Technology*, 2005.

[11] F. E. H. Tay, "Application of support vector machines in financial time series forecasting", *Omega* 29, pp. 309-317, 2002.

[12] V. N. Vapnik, *Statistical Learning Theory*, Wiley-Interscience Pub., 1998.

[13] Y. Wang, S. Wang, and K. K. La, A New Fuzzy Support Vector Machine to Evaluate Credit Ris, *IEEE Transactions on Fuzzy Systems*, Vol. 13, No. 6, pp. 820-831, 2005.

[14] H. Xia, B. Q. Hu, Feature selection using fuzzy support vector machines, *Fuzzy Optimal Decision Making*, vol. 5, pp.187 - 192, 2006.

## 저 자 소 개



이 수 용 (Soo-Yong Lee)

2004년~현재, 연세대학교 과학기술대학  
컴퓨터정보통신공학부 조교수.

2004년 8월 : 연세대학교 대학원 컴퓨터과  
학.산업시스템공학과 졸업,  
공학박사.

1992년 2월 : 경희대학교 대학원 수학과 졸  
업, 이학박사.

관심분야 : Fuzzy SVM, Pattern Classification,  
Application Mathematics, Data Mining  
Time Series Prediction,

Phone : 02-3293-3293

E-mail : 0691@yonsei.ac.kr