

# 기계학습알고리즘을 이용한 위험회복지수의 개발과 활용

김 선 웅\*

## Development and Application of Risk Recovery Index using Machine Learning Algorithms

Sun Woong Kim\*

### Abstract

Asset prices decline sharply and stock markets collapse when financial crisis happens. Recently we have encountered more frequent financial crises than ever. 1998 currency crisis and 2008 global financial crisis triggered academic researches on early warning systems that aim to detect the symptom of financial crisis in advance.

This study proposes a risk recovery index for detection of good opportunities from financial market instability. We use SVM classifier algorithms to separate recovery period from unstable financial market data. Input variables are KOSPI index and V-KOSPI200 index. Our SVM algorithms show highly accurate forecasting results on testing data as well as training data. Risk recovery index is derived from our SVM-trained outputs. We develop a trading system that utilizes the suggested risk recovery index. The trading result records very high profit, that is, its annual return runs to 121%.

Keywords : Financial Crisis, Early Warning System, SVM, Risk Recovery Index, Trading System

Received : 2016. 08. 01.    Revised : 2016. 09. 18.    Final Acceptance : 2016. 09. 19.

※ This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government(NRF-2014 S1A5A8019528).

\* Associate Professor, Graduate School of Business Information Technology, Kookmin University, 77 Jeongneung-ro, Seongbuk-gu, Seoul, 02707, Korea, Tel : +82-2-910-5471, e-mail : swkim@kookmin.ac.kr

## 1. 서 론

자산가치가 폭락하고 주식시장이 패닉상태에 빠지는 등 금융시장에 충격을 던지는 금융위기(financial crisis)는 최근 그 발생빈도가 높아지고 있다. 2008년 서브프라임 모기지론(sub-prime mortgage loan)으로 촉발된 미국의 금융위기는 빠른 속도로 전 세계 금융시장을 강타하였다. 당시 우리나라도 주가지수가 급락하고 환율이 급등하는 등 큰 혼란을 겪었다. 금융위기는 곧바로 실물경제에도 악영향을 주기 때문에 금융기관 종사자나 투자자뿐만 아니라 모든 국민들에게 영향을 미치게 된다. 우리나라는 이미 1990년대 말 외환위기를 겪으면서 여러 분야에서 심각한 피해를 경험한 바 있다. 이에 따라 학계를 중심으로 금융위기에 대한 연구가 본격화되었으며, 대부분의 연구는 “어떻게 금융위기를 조기에 탐지할 수 있을까”에 대한 모형을 연구하는 조기경보시스템(Early Warning System : EWS)에 초점이 맞추어졌다[El-Shazly, 2011].

하지만 위기는 위험상황이면서 동시에 여러 경제 주체들에게 새로운 기회를 찾을 수 있는 중요한 상황이라는 측면을 간과해서는 안 된다. 최근에 와서 전 세계 경제는 글로벌화가 가속화되면서 위기발생은 일상화되고 있다. 금융위기가 발생하면 투자자뿐만 아니라 경제 정책 당국자들도 패닉 상태에 빠져 합리적인 대응책을 제시하지 못하고 우왕좌왕하게 된다. 그러는 사이 위기는 대부분 빠른 속도로 회복되는 특성을 보이기 때문에, 금융위기 발생의 사전 탐지 못지않게 위기 회복 타이밍을 찾아내는 것도 투자자나 금융정책 입안자 모두에게 대단히 중요하다. 따라서 본 연구는 중요한 의사결정 문제이면서도 그동안 학계에서 소홀히 다루어져왔던 위기회복시스템을 새로이 제안하고, 위기상황에서 기회를 찾아내는 기준을 제시하여 경제 주체들의 의사

결정에 의미 있는 정보를 제공하는 것이 목적이다. 구체적으로 금융위기를 전후한 금융시장의 상황 분석을 통해 위험회복의 정도를 판단할 수 있는 위험회복지수(Risk Recovery Index; RRI)를 개발하고자 한다. 위험회복지수는 금융위기가 발생한 경우 금융시장이 악화된 이후 금융위기가 회복하는 정도를 지수화한 값이다.

위험회복지수를 개발하고 경제 주체들이 실제 시장상황에서 활용하기 위해서는 모형이 조금 더 현실의 상황을 빨리 설명할 수 있어야 한다. 그 동안 조기경보시스템에서는 대부분 분기별이나 월별 거시경제변수를 분석하여 금융위기를 조기에 탐지하려고 노력해왔다. 1980년대까지만 하더라도 경제의 기초여건(fundamentals)이 악화된 상황에서 정부의 잘못된 정책 결정으로 금융위기가 주로 발생했지만, 1990년대 이후로는 경제의 기초여건이나 정책 결정의 오류 없이도 주식시장이나 외환시장에서의 투기적 가수요만으로도 자기실현적 위기(self-fulfilling crisis)가 자주 발생하고 있다. 이러한 금융위기는 빠른 속도로 진행되기 때문에 월간단위나 분기단위로 측정 가능한 거시경제변수로는 적기에 대응할 수 없다.

최근의 금융위기 관련 연구들은 금융위기에 빨리 대응하기 위해서 분기별이나 월별 자료가 아닌 일별 자료와 같은 고빈도 자료(high-frequency data)를 활용하기 시작하였다[Kim et al., 2004a; Kim et al., 2004b; Oh et al., 2006; Yoon and Park, 2014]. 그러나 대부분의 연구가 주식의 평균가격의 변화를 측정하는 KOSPI(Korea Stock Price Index)와 같은 주가지수 자료를 분석하고 있다. 본 연구에서는 주식시장의 추세를 나타내는 주가지수뿐만 아니라 파생상품시장의 흐름을 나타내는 변동성지수(volatility index)의 일별 자료를 이용하여 위험회복지수를 개발하고자 한다. 금융위기가 발생하면 주가가 폭락하고 환율이 큰

폭으로 상승하면서 투자자들의 심리상태가 패닉 상태에 빠지게 된다. 변동성지수는 시장의 패닉 상태를 비교적 잘 설명하는 것으로 알려져 있기 때문에 주식시장과 파생상품시장을 동시에 분석하면 우수한 예측성과를 보이는 금융위기 분석 시스템을 만들 수 있을 것이다.

본 연구는 금융위기가 발생하면서 나타나는 주가지수와 변동성지수의 입력변수들로부터 급변동패턴과 안정을 찾아가는 안정화패턴을 구분하기 위해 패턴분류기법을 이용한다. 구체적으로 패턴분류기로서 Support Vector Machines(SVM)을 이용하며, 과거의 금융위기국면의 금융시장 입력 자료로부터 학습과정을 거쳐 위험회복패턴을 찾는 알고리즘을 찾아낸다. 알고리즘의 결과를 이용해 위험회복지수를 구하고 이를 이용해 새로운 투자전략인 위험회복지수 트레이딩시스템을 제시하고자 한다. 주가나 변동성 관련 자료들은 일별자료나 심지어 분별자료까지의 실시간 고빈도자료의 확보가 가능하기 때문에, 시스템의 활용 측면에서는 상당히 빠른 의사결정에 도움을 줄 것이다. 본 연구에서는 일별자료를 입력 변수로 분석하고 있지만 필요에 따라서는 분단위의 분별 자료로도 확장가능하다.

본 연구는 제 2장에서 선행연구를 조사하며, 제 3장에서는 모형을 제안한다. 제 4장에서는 실험 설계 및 위험회복지수를 도출하며, 이를 주식시장에 적용한 트레이딩시스템의 실증분석을 통해 투자 성과를 분석한다. 마지막 장에서는 결론 및 본 연구의 한계점을 제시한다.

## 2. 선행 연구

금융위기에 대한 국내·외 연구현황을 보면 금융위기를 사전에 탐지하기 위한 조기경보시스템의 개발 및 활용에 집중되고 있으며, 본 연구의 주제인 금융위기 발생 시 기회를 포착하기 위한

위험회복지수에 대한 연구는 찾아보기 힘들다. 초기의 연구들은 경제 환경을 둘러싼 거시경제 변수의 악화로 인해 금융위기가 발생한다는 가정을 바탕으로 장기적 관점에서 월간단위 또는 분기단위의 경제변수를 중심으로 판별분석(discriminant analysis)이나 로짓모형(logit model) 등 통계적 모형을 적용한 연구들이 주류를 이루었다. 1990년대 아시아 외환위기를 겪으면서 EWS 구축 필요성이 증대되었고 학계에서도 EWS에 대한 연구가 본격화되었다. Kaminsky et al.[1998]은 수출입량, 환율변동, 국가간 금리차, 통화량, 주가 등 15개의 거시경제지표에 대하여 noise-to-signal ratio를 최소화하는 임계치를 찾아내어 24개월 전에 외환위기 신호를 탐지할 수 있는 EWS를 제안하였다. Berg and Pattillo[1999]는 패널 로짓모형(panel-logit model)을 이용하여 외환위기의 조기경보시스템을 제안하고 기존의 신호점근법보다 예측능력이 우수함을 보였다. 이 연구를 따라 추가적인 연구가 다수 이루어졌다[Kumar et al., 2003; Fuertes and Kalotychou, 2007; Berg et al., 2008]. 그러나 이 연구들은 금융위기와 안정기의 구조적 변화를 이용하지 아니하였다. Bussiere and Fratzscher[2006]는 다항 로짓모형을 제안하였으나, EWS를 평가하는 데 있어서 ad hoc 임계치를 사용하여 객관성을 결여하고 있다. Candelon et al.[2012]은 Credit-Scoring 접근법을 이용한 최적 임계치 결정모형을 제안하고 12개의 개발도상국을 상대로 분석한 결과, 이자율스프레드가 외환위기를 예측하는 데 가장 중요한 지표가 됨을 밝혔다.

EWS 문제를 학습의 문제로 인식하면서 다양한 인공지능기법들을 이용한 학습모형들이 최근에 와서 개발되기 시작하였다. Kim et al.[2004a]는 1997년의 한국 외환위기 기간의 한국 주식시장의 주가지수 자료에서 4개의 특성변수를 추가로 추출하여 시장국면을 안정국면, 불안정국면,

그리고 위기국면으로 구분한 후 인공신경망기법으로 학습을 시킨 결과 검증데이터에서 좋은 결과를 보여주었다. Kim et al.[2004b]는 EWS를 학습하는데 있어서 logistic discrimination 모형, 의사결정수 모형, SVM, 뉴로퍼지 모형, 그리고 ANN(Artificial Neural Network)모형을 비교한 결과, 학습데이터에서 의사결정수 모형, ANN, SVM이 우수한 성과를 보여주었다. 그러나 이러한 연구들은 모두 금융위기가 언제 나타나는지를 예측하는 데 초점이 맞추어져 있다는 점에서 본 연구의 위기 이후의 대처 시스템과는 차이가 난다.

EWS에 대한 지수접근법도 시도되었다. Boduroglu and Erenay[2007]는 패턴인식기법을 이용하여 터키의 거시경제지표들을 하나로 묶는 경제안정지수를 개발하고 터키의 금융위기 국면에 적용하여 금융위기가 발생하기 전에 거의 100% 가까운 예측력을 보임을 입증하였다. Son et al.[2009]는 한국의 금융위기 국면을 조기에 경보하기 위하여 인공지능모형을 이용한 주식시장 안정지수를 개발하고 실증분석을 통해 금융시장의 불안정국면을 잘 찾아내고 있음을 보여주었다. Oh et al.[2006]는 주식시장의 고빈도자료를 이용하여 금융시장 상황지표(Daily Financial Condition Indicator; DFCI)를 제안하고 인공신경망을 이용하여 1단계로 개별 시장변수들의 sub-DFCI를 만들고 2단계로 유전자알고리즘을 이용하여 이들을 통합한 최종 DFCI를 만들어 한국시장에 적용하여 우수성을 입증하였다. Kim et al.[2009]는 기존의 연구들과는 달리 다량의 데이터가 활용가능한 일반적인 시장의 안정국면에서 학습을 통해 현재의 국면이 안정국면과 다른지를 판단하는 지수를 제안하고 ANN으로 학습을 실시한 후 한국 시장상황에 적용하여 제안한 조기경보시스템의 우수성을 보여주었다. Yoon and Park[2014]는 Kim et al.[2004b]의 방법에 기초하여 ANN 학습을 통한 시장불안정지수를 제안하고, 시장의 불안정 단계에

따라 대응단계를 정하는 시스템을 개발하였다. 이 시스템을 한국과 그리스의 시장에 적용하여 조기경보시스템으로서의 우수성을 보여주었다. 위험회복을 주제로 Duasa and Ahmad[2010]는 말레이시아의 거시경제변수를 분석하였으나, 위기에서 언제 벗어날 것인가를 예측한 것이 아니라 위기를 벗어나는데 평균 어느 정도의 시간이 걸리는지를 분석하였다. 연구자가 기 수행한 연구 중에서는 Kim and Ahn[2010]의 연구논문이 SVM 알고리즘의 활용 측면에서 연관성이 있다. 이 연구는 기계학습 알고리즘을 이용하여 주식투자자들이 주식시장에서 매수-매도를 적절히 실행하는 최적의 트레이딩시스템을 개발하고, 실증분석을 통해 더 나은 투자성과를 실현하였음을 보여주었다. 이 때 시장의 여러 입력변수들을 이용하여 주가가 오를 것인가 내릴 것인가를 예측하는 확률을 계산하고, 이 확률에 유전자 알고리즘을 적용하여 최적 투자 의사결정을 위한 이중 임계치를 결정하여 기존의 투자결정 방법보다 좋은 성과를 시현하였다.

이상에서 검토한 것처럼 본 연구는 기존의 연구들과는 많은 차별성을 보이고 있다. 무엇보다도 본 연구는 위기국면에서 기회를 찾는 데 도움을 줄 수 있는 위험회복지수를 제안한다는 점에서 그동안의 EWS 연구들과는 차별성을 가진다. 본 연구에서 이용하는 입력변수 중에는 그동안의 연구에서 다루어지지 않았던 파생상품시장의 정보를 추가하고 있는데, 파생상품시장을 대표하는 변동성지수는 금융위기 국면에서는 투자자들의 심리상태까지 잘 반영하고 있어서 위기의 발생과 회복의 예측에서 중요한 정보를 제공해 줄 것으로 기대된다. 많은 연구 결과 주가지수와 변동성지수 사이에는 비대칭적 역의 관계(negative asymmetry)가 존재한다는 사실이 밝혀지고 있다[Hibbert et al., 2008; Ederington and Guan, 2010; Kim, 2010].

### 3. 제안 모형

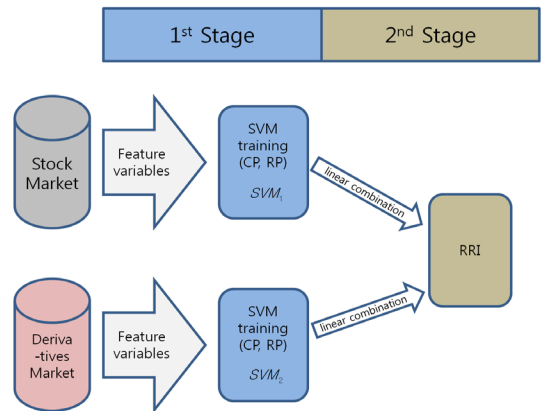
본 연구에서는 한국의 증권시장 관련 자료를 이용하여 금융위기에서와 같이 복잡한 비선형관계를 보이는 분류문제를 해결하는 데 적합한 SVM 기법을 활용한 위험회복지수를 개발하고자 한다. SVM은 러시아의 통계학자인 Vapnik[1995]이 1995년 제안한 분류기법으로, 복잡한 비선형 관계를 갖는 시스템에서 인공신경망과 마찬가지로 효과적인 분류 성능을 보이고 있다. 복잡한 분류 문제에서 우수한 예측 정확도를 보이는 SVM은 다음과 같은 여러 측면에서 강점을 보유하고 있다. SVM의 학습 수행 방법은 ‘구조적 위험(structural risk)’을 최소화하는 방향으로 진행되기 때문에 상대적으로 과최적화(overfitting) 문제에서 자유로울 수 있다. 여기에 더하여, SVM은 지지 벡터(support vector)라 불리는 소수의 데이터만을 학습 과정에서 사용하기 때문에, 적은 양의 학습 데이터로도 우수한 예측성적을 나타낸다는 점에서 금융위기와 같은 특수한 경제 환경에 적용하기에 유용하다[Kim, 2003; Ahn et al., 2005].

위험회복지수 개발을 위한 모형의 구성은 크게 2단계로 구분된다. 모형의 1단계는 SVM을 이용해서 주식시장 데이터와 파생상품시장 데이터 각각에 대해 위기지속구간과 위험회복구간의 분류 예측모형을 구축한다. 모형의 2단계는 주식시장과 파생상품시장 각각에 대해서 모형 1단계에서 구해진 결과를 선형 결합하여 위험회복지수를 산출하는 과정이다.

#### 3.1 SVM 모형의 학습

모형의 1단계는 금융위기구간을 위기지속구간과 위험회복구간으로 구분하고, 이 구간에서 관찰된 주식시장 일별자료와 파생상품시장 일별 자료를 학습자료로 하여 SVM 분류 예측모형을 구축한다. 예측 결과는 특정일이 위기지속일(Crisis

Period : CP)과 위험회복일(Recovery Period : RP)의 두 범주로 분류하고, 그 예측 결과를 0~1사이의 분류 추정확률로 산출할 수 있는 SVM 모형을 적용하였다. 기본적으로 추정확률은 분류기로부터 데이터가 얼마나 멀리 떨어져 있는지에 대한 정보를 얻을 수 있기 때문에, 본 연구의 목적인 위험회복지수를 도출하는 데 유용할 것으로 판단된다. 주가지수의 특성변수에 SVM 분류학습을 한 결과를  $SVM_1$ , 변동성지수의 특성변수에 SVM 분류학습을 한 결과를  $SVM_2$ 로 표시한다. <Figure 1>은 제안된 시스템의 구성도이다.



<Figure 1> Structure of Suggested System

#### 3.2 위험회복지수의 도출

주식시장과 파생상품시장 각각에 대한 1단계의 추정확률이 주어지면, 모형의 2단계는 그 결과를 선형 결합하여 위험회복지수를 산출하는 과정이다. 위험회복지수 RRI는 다음의 선형 결합식으로 도출한다.

$$RRI_t = \omega_1 SVM_{1,t} + \omega_2 SVM_{2,t} \quad (1)$$

여기서  $SVM_{1,t}$ 는 주식시장의 특성변수로부터 예측된 t일의 추정확률이며,  $SVM_{2,t}$ 는 파생상품

시장의 특성변수로부터 예측된  $t$ 일의 추정확률, 그리고  $\omega_1, \omega_2$ 는 주식시장과 파생상품시장 각각의 분류 결과에 대한 가중치로서,  $\omega_1 \geq 0, \omega_2 \geq 0, \omega_1 + \omega_2 = 1$ 을 만족시킨다. 가중치  $\omega_1, \omega_2$ 는 다음과 같은 목적함수에서 예측 오차를 최소화하는 값으로 결정한다. 여기서  $k$ 는 위기지속구간은 0, 위험회복구간은 1의 값을 갖는다.

$$\min \sum_{i=1}^2 \sum_{t=1}^T [\omega_i SVM_{i,t} - k]^2, \quad (2)$$

where  $T$  is the number of training data

## 4. 실험 설계 및 실증 분석

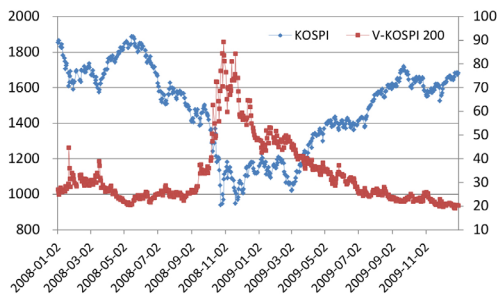
### 4.1 변동성지수 V-KOSPI200

본 연구에서는 기존의 금융위기에 대한 연구들과 달리, 파생상품시장의 대표 지수인 변동성지수를 추가하여 실험을 진행한다. 한국거래소(Korea Exchange)는 2009년 4월부터 한국의 변동성지수인 V-KOSPI200을 실시간으로 발표하고 있으며, 2008년 3월 자료까지 하루의 시가-고가-저가-종가 정보를 소급하여 제공하고 있다. V-KOSPI200은 한국거래소가 1997년 도입하면서 세계적 파생상품시장으로 발전한 KOSPI200 주가지수옵션가격을 이용하여 계산하고 있으며, 이는 미국의 CBOE(Chicago Board Options Exchange)에서

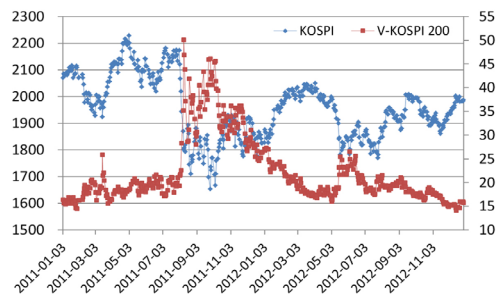
발표하고 있는 변동성지수인 VIX와 계산방법이 동일하다. VIX는 옵션가격에 내재된 기초자산의 미래변동성에 대한 시장의 기대치를 나타낸 지수로서, Whaley[2000]는 이를 투자자공포지수(investor fear gauge)로 해석하였다. <Figure 2>에서 (a)는 2008년 금융위기국면에서 KOSPI가 폭락하면서 변동성지수 V-KOSPI200이 급등하는 특성을 보여주고 있는 학습구간 데이터이며, (b)는 제안된 모형의 검증을 위한 구간에서의 KOSPI와 V-KOSPI200 지수의 움직임 특성을 보여주고 있다.

### 4.2 실험 데이터

본 실험에서는 주식시장 입력변수로 주가지수인 KOSPI, 파생상품시장 입력변수로 변동성지수인 V-KOSPI200의 일별 종가를 사용한다. 학습용 데이터의 금융위기 구간은 2008년부터 2009년까지, 검증용 데이터의 금융위기 구간은 2011년부터 2012년까지의 구간이다. 금융위기 구간의 구분은 증권 전문가의 자문을 받아 주가지수가 폭락하면서 역사적 변동성(historical volatility)이 두 배 이상 증가한 구간을 CP로, 주가지수 폭락국면에서 다시 주가지수의 저점이 직전 저점보다 높아지는 상승국면으로의 전환 시점을 RP로 구분하였다. 학습구간과 검증구간의 구체적인 구분 결과는 다음 <Table 1>과 같다.



(a) 2008~2009

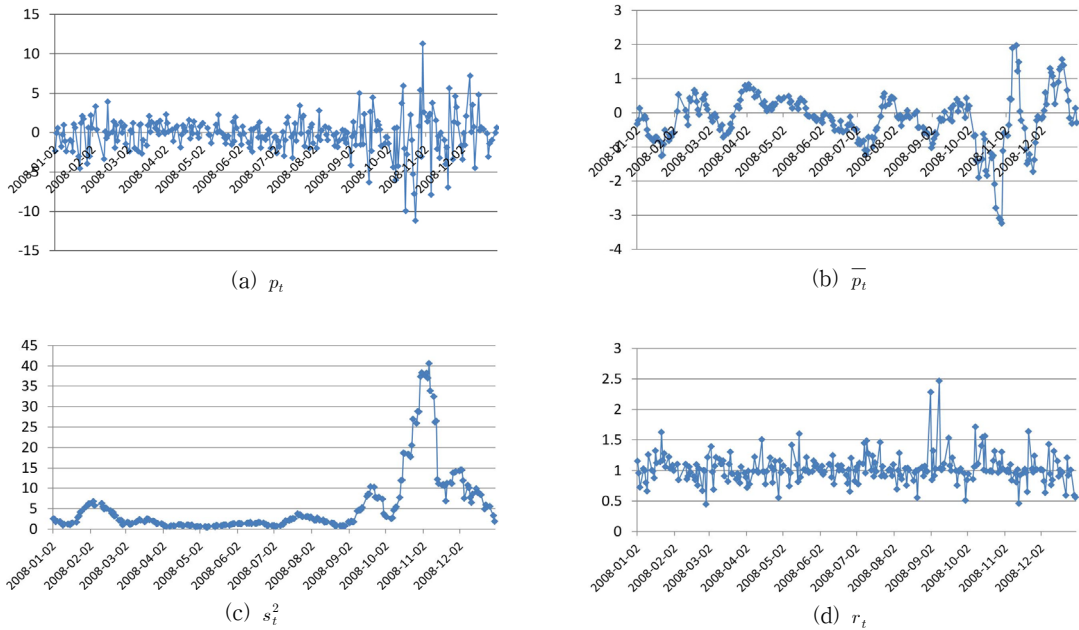


(b) 2011~2012

<Figure 2> Trend of KOSPI and V-KOSPI200

<Table 1> Training and Testing Periods

	CP	RP
Training	2008/05/20 ~ 2008/10/27	2008/10/28 ~ 2009/04/07
Testing	2011/08/02 ~ 2011/09/26	2011/09/27 ~ 2012/02/20



<Figure 3> KOSPI Feature Variables

주식시장 정보인 KOSPI 주가지수에 대한 특성변수는 Kim et al.[2004b]에서 제시한 다음의 4변수( $p_t, \bar{p}_t, s_t^2, r_t$ )를 동일하게 활용한다.

$$p_t = \ln(x_t/x_{t-1}) \times 100, \quad \bar{p}_t = \sum_{i=t-9}^t p_i/10, \quad (3)$$

$$s_t^2 = 1/10 \sum_{i=t-9}^t (p_i - \bar{p}_t)^2, \quad r_t = s_t^2/s_{t-1}^2$$

여기서  $x_t$ 는 t일의 KOSPI 주가지수이다. 식 (3)에서  $p_t$ 는 t일의 KOSPI 주가지수의 일별수익률 (daily return),  $\bar{p}_t$ 는 최근 10일 동안의 수익률의 평균,  $s_t^2$ 은 t일의 최근 10일 동안의 수익률의 분산,  $r_t$ 는 (t-1)일 대비 t일의 수익률 분산의 변동비를 각각 측정하는 변수로서, 2008년 금융위기 구간

에서의 주가지수 특성변수 움직임은 <Figure 3>과 같다.

파생상품시장 정보인 V-KOSPI200 변동성지수에 대한 특성변수도 주가지수와 같이 Kim et al. [2004b]에서 제시한 다음의 4변수( $p_t, \bar{p}_t, s_t^2, r_t$ )를 동일하게 활용한다.

$$p_t = \ln(y_t/y_{t-1}) * 100, \quad \bar{p}_t = \sum_{i=t-9}^t p_i/10, \quad (4)$$

$$s_t^2 = 1/10 \sum_{i=t-9}^t (p_i - \bar{p}_t)^2, \quad r_t = s_t^2/s_{t-1}^2$$

여기서  $y_t$ 는 t일의 V-KOSPI200 주가지수이며, 2008년 금융위기 구간에서의 변동성지수 특성변수 움직임은 <Figure 4>와 같다.

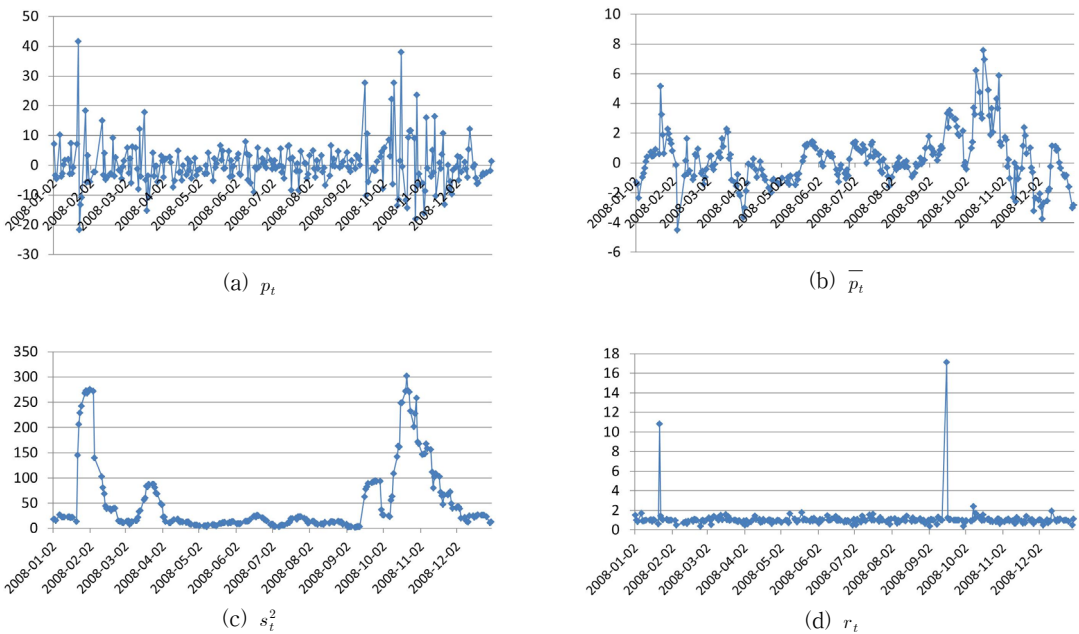
### 4.3 시스템 설계

제안 시스템에서는 SVM의 커널함수(kernel function)로 일반적으로 가장 많이 사용되는 선형 함수(linear function), 가우시안 RBF 함수(Gaussian radial basis function), 그리고 시그모이드 함수(sigmoid function) 등을 적용하였다. 커널 함수  $K(u,v)$ 는 선형함수  $u'v$ , RBF 함수  $e(-\gamma\|u-v\|^2)$ , 시그모이드함수  $\tanh(\gamma u'v + coef)$  형태로 주어지며, 파라미터는 비용 상수인 조정 파라미터(regularization parameter)  $C$ 와 감마(gamma) 파라미터 등이 학습 결과에 영향을 미칠 수 있다[Ahn et al., 2005; Kim and Ahn, 2010]. 이에 따라 본 연구에서도 Kim and Ahn[2010]의 방법을 따라 파라미터 값을 다양하게 변화시키면서 실험을 행하고 가장 우수한 성과를 보이는 파라미터를 선택한다. 구체적으로  $C$ 는 1부터 100까지,  $\gamma$ 는 0.1에서 1까지 다양하게 변화시키면서 최적의 성과를 보이는 파라미터 조합을 찾는다.

먼저, KOSPI 주가지수 데이터에 대한 학습 결과가 가장 좋은 SVM 분류기를  $SVM_1$ , V-KOSPI200 변동성지수 데이터에 대한 학습 결과가 가장 좋은 SVM 분류기를  $SVM_2$ 라고 한다. 우리의 위험회복 지수를 도출하기 위하여 두 분류기를 선형 결합한 통합모형을 구축하여 한국 주식시장의 위험회복의 정도를 판단할 수 있는 식 (1)의  $RRI_t$ 를 제안하고, 이 모형을 활용한 트레이딩시스템의 실증분석을 통해 제안된 위험회복지수의 유용성을 검증한다.

### 4.4 실증 분석 결과

본 연구 제안 시스템의 1단계 과정은 KOSPI와 V-KOSPI200 각각의 학습용 데이터에 위기 지속구간과 위험회복구간으로 구분하는 SVM 분류 학습을 시행하고 검증용 데이터에서 검증하였다. SVM의 실험은 통계패키지 R의 'e1071'을 이용하였다. 다음의 <Table 2>와 <Table 3>은 커널함수와 파라미터의 여러 조건 하에서, KOSPI 주가지수와 V-KOSPI200 변동성지수 각각의 특



<Figure 4> V-KOSPI200 Feature Variables



성변수에 대해서 실험된 SVM 모형의 학습 결과들을 보여주고 있다.

<Table 2>의 KOSPI 주가지수에 대한 학습 결과를 보면, 학습용 데이터에 대한 성과는 RBF 커널함수가 전반적으로 높은 분류 예측력을 보이고 있다. 그러나 검증용 데이터에서 낮은 예측성 과를 보이고 있어 과최적화의 문제점이 지적될 수 있다. sigmoid 커널함수는 학습용 데이터와 검증용 데이터에서 안정적인 예측성 과를 보여주고 있으며, 파라미터 2(gamma)가 중간 값인 0.5에서

전반적으로 높은 예측력을 보여주고 있다. <Table 3>의 V-KOSPI200 변동성지수에 대한 학습 결과는 <Table 2>의 KOSPI 주가지수에 대한 분류 성과보다는 안정적인 성과를 보여주고 있다. V-KOSPI200 변동성지수에 대한 커널함수별 예측성 과는 KOSPI 주가지수에서와 비슷한 패턴을 보여주고 있다. <Table 2>와 <Table 3>으로부터 커널함수는 sigmoid, cost 파라미터 C는 1, gamma 파라미터는 0.5를 선택하고, 이를 이용하여 모형의 2단계인 위험회복지수를 도출한다.

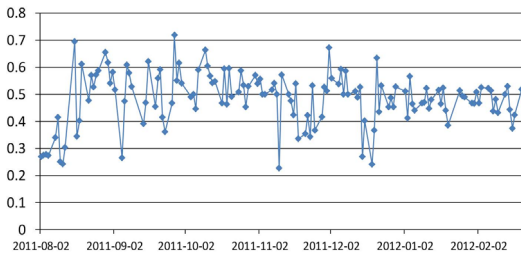
<Table 2> SVM Training and Testing Results on KOSPI

Kernel Function	Parameter 1 (C)	Parameter 2 ( $\gamma$ )	Training Data	Testing Data
linear	1		72.97%	47.10%
	10		72.08%	45.65%
	50		72.08%	45.65%
	100		72.08%	45.65%
radial basis	1	0.1	74.32%	42.75%
		0.5	85.59%	39.13%
		1.0	85.59%	41.30%
	10	0.1	81.98%	34.06%
		0.5	89.64%	51.45%
		1.0	94.59%	53.62%
	50	0.1	86.04%	35.51%
		0.5	92.79%	55.07%
		1.0	95.95%	50.00%
	100	0.1	85.14%	37.68%
		0.5	93.69%	55.07%
		1.0	97.30%	47.83%
sigmoid	1	0.1	71.17%	53.62%
		0.5	60.36%	59.42%
		1.0	55.86%	54.35%
	10	0.1	68.02%	45.65%
		0.5	59.46%	56.52%
		1.0	55.86%	52.17%
	50	0.1	68.02%	46.38%
		0.5	58.56%	57.25%
		1.0	55.86%	52.17%
	100	0.1	66.67%	44.93%
		0.5	58.56%	57.25%
		1.0	55.86%	52.17%

<Table 3> SVM Training and Testing Results on V-KOSPI200

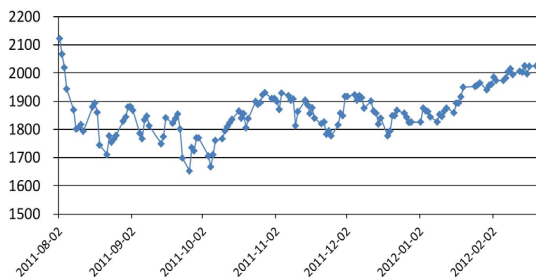
Kernel Function	Parameter 1 (C)	Parameter 2 ( $\gamma$ )	Training Data	Testing Data
linear	1		75.23%	69.57%
	10		75.23%	69.57%
	50		75.23%	69.57%
	100		75.23%	69.57%
radial basis	1	0.1	76.89%	73.19%
		0.5	81.98%	75.36%
		1.0	85.14%	72.46%
	10	0.1	81.53%	70.29%
		0.5	86.49%	71.01%
		1.0	87.39%	65.94%
	50	0.1	81.08%	70.29%
		0.5	87.84%	64.49%
		1.0	90.09%	62.32%
	100	0.1	81.08%	68.84%
		0.5	90.09%	60.87%
		1.0	90.95%	63.77%
sigmoid	1	0.1	68.47%	65.94%
		0.5	60.81%	58.70%
		1.0	53.60%	56.52%
	10	0.1	64.86%	57.25%
		0.5	56.76%	54.35%
		1.0	55.41%	55.07%
	50	0.1	65.32%	55.80%
		0.5	59.91%	50.00%
		1.0	54.95%	55.07%
	100	0.1	63.96%	55.07%
		0.5	57.66%	51.45%
		1.0	54.95%	55.07%

1단계에서 구한 SVM 분류모형으로부터 얻은 추정확률을 이용하여 위험회복지수를 계산하기 위해 식 (2)를 학습용 데이터구간에서 적용한 결과, 분류 오류를 최소화하는 비율은  $\omega_1 = 1, \omega_2 = 0$ 의 경우로, 학습용 데이터 구간에서의 KOSPI 주가지수의 위험회복구간 분류위험이 V-KOSPI200 변동성지수보다 항상 우위를 점하고 있다. <Figure 5>는 검증용 데이터에서의 위험회복지수의 추이를 보여주고 있다.



<Figure 5> RRI Trend in Testing Data

위험회복지수는 금융시장이 금융위기 발생 후 위험의 회복 정도를 지수화한 값으로서, 실무적으로도 다양하게 활용될 수 있다. 2008년도 글로벌 금융위기 상황에서 대부분 국가의 금융정책당국이 행한 위기 대응전략을 보면 사후약방문식의 임기응변으로 일관한 측면이 강했다. 이에 비해 투자의 달인이라는 워렌버핏 같은 투자자들은 금융위기가 발생한 상황에서 정책당국이나 투자자들이 우왕좌왕 하는 사이, 주식시장에서 새로운 투자 기회를 발견하고 과감한 투자를 실행하여



(a) Price Trend of KOSPI

엄청난 수익을 올린 바 있다. 위험회복지수는 일반투자자들에게도 전문적 기관투자가처럼 감정에 흔들리지 않고 금융위기 국면에서 행동할 수 있는 좋은 방향타를 제시할 수 있는데, 다음 절에서는 위험회복지수를 활용한 투자전략을 제시하고 실증분석을 통해 그 유용성을 검토하고자 한다.

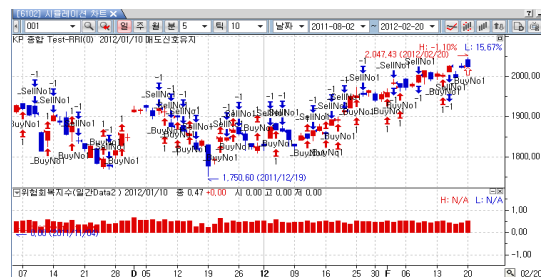
#### 4.5 위험회복지수를 활용한 트레이딩시스템의 성과 분석

본 절에서는 위험회복지수를 투자에 활용하는 트레이딩시스템을 제안하고 실증분석을 통해 운용 성과를 분석하고자 한다. 위험회복지수 트레이딩시스템(RRI-TS)은 위험회복지수가 0.5를 초과하면 위험회복의 신호로, 0.5 이하로 하락하면 위험의 계속 구간으로 판단하여 다음과 같은 매수(Buy)-매도(Sell)의 진입 규칙(Entry Rule)을 포함하고 있다.

$$RRI-TS \text{ Entry Rule} \tag{5}$$

If  $RRI > 0.5 + level$  then Buy;  
 If  $RRI \leq 0.5 - level$  then Sell;  
 where level is a buffer.

<Figure 6>에서 (a)는 검증구간에서의 KOSPI 주가지수 추이를, (b)는 하이투자증권이 제공하고 있는 “에스트레이더” 트레이딩 툴로서, RRI-TS 전략을 프로그래밍 하여 주가 차트에 Buy-Sell 신호를 보여주고 있다.



(b) Buy-Sell Signal of RRI-TS

<Figure 6> KOSPI and Buy-Sell Signal of RRI-TS

<Table 4>는 RRI-TS를 이용하여 검증용 데이터 구간에서 행한 트레이딩의 실증분석 결과이다.

<Table 4> Performance Report on RRI-TS(2011/08/02~2012/02/20)

	Total	Buy	Sell
Total Return	205.43pt	54.74pt	150.69pt
Number of Trading	62	31	31
Number of Profitable Trading	33	16	17
Percent Profitable	53%	52%	55%
Buy-and-Hold	-96.37pt		

전체 기간 138일 동안 총 62회의 거래가 발생하였고, 그중 33회의 거래에서 이익을 얻어 거래 승률은 53%로 나타났다. 제안된 RRI-TS 투자 전략의 전체 거래에서 발생한 총 손익은 +205.43포인트로서, 검증 시작일에 매수포지션을 구축하여 종료일까지 단순 매입-보유하는 벤치마크 전략(Buy-and-Hold Strategy)의 성과(-96.37포인트) 대비 우수한 투자 성과를 보여주고 있다. 본 시뮬레이션에서는 KOSPI 주가지수의 추상적 매매를 가정하여 분석하였기 때문에 거래비용 등은 고려하지 못하였다. 거래에 따르는 수수료나 실제 거래 시 불리하게 체결될 수 있는 가능성인 슬리피지비용(slippage cost) 등을 고려한 실제 매매시스템의 시뮬레이션을 위해 거래 대상을 KOSPI 주가지수가 아닌 실제 투자가 가능한 KOSPI200 주가지수선물로 바꾸어 실험한 결과는 <Table 5>와 같다.

<Table 5> Performance Report on RRI-TS : KOSPI200 Futures

	Total	Buy	Sell
Total Return	21.40pt	5.95pt	15.45pt
Number of Trading	62	31	31
Number of Profitable Trading	35	17	18
Percent Profitable	56%	55%	58%
Buy-and Hold	-8.65pt		

현재 세계적인 시장으로 발전한 한국의 KOSPI 200 주가지수선물시장에서 제안된 트레이딩시스템 RRI-TS의 투자 성과를 분석한 결과, 전체 기간 동안 총 62회의 거래를 통해 21.40포인트의 수익을 시현하였다. KOSPI200 주가지수선물 1계약을 거래하기 위해서는 증거금(margin)으로 약 1,500만 원이 필요하고, 본 시스템의 수익 21.40포인트는 금액으로 환산하면 1,070만 원이 되어 투자 수익률이 +71%에 달하는 높은 성과를 시현하였다. 연율로 환산하면 +121%에 해당하는 수익률로, 벤치마크전략인 Buy-and-Hold 전략의 -49% 대비 우수한 투자성과를 시현하였다.

## 5. 결론 및 연구의 한계점

본 연구에서는 그 동안의 금융위기 연구에서 소외되었던 금융위기 탈출 신호에 대한 모형을 도출하였다. 금융시장은 본질적으로 복잡계이며, 특히 갑작스러운 외부 충격으로 금융시장의 불안정성이 증가하기 시작하여 금융위기로 진행되면 투자자들의 비이성적 행동이 시장을 지배하기 때문에, 주가나 환율 등의 급등락과 같은 불안정성이 극에 달한다. 따라서 본 연구에서는 금융위기와 같은 불안정성이 높아진 금융시장의 데이터에서 우수한 분류 학습 능력을 보이는 SVM 모형을 활용하였다. 불안정한 금융위기 자료에서는 투자자들의 감정이 많이 영향을 미치기 때문에, 행동재무학(behavioral finance)적 특성을 보이는 변동성지수를 입력변수로 추가하여 분석하였으며, 모형의 1단계에서 얻어진 추정확률을 이용하여 위험회복지수를 도출하였다. 마지막으로 위험회복지수를 활용하는 투자전략을 제안하고 실제 주식시장에서 그 시뮬레이션의 성과를 분석하였다.

본 연구의 의의 및 결론은 크게 다음과 같이 요약할 수 있다. 첫째, 본 연구는 그 동안의 금융

위기 연구가 대부분 위기 발생의 조기탐지에 초점이 맞추어진 데 반해서, 본 연구는 금융위기 발생 이후 금융시장이 안정을 찾아가는 과정을 추적하는 위험회복시스템을 도출하였다는 점이다. 이러한 관점에서 본 연구는 향후 금융위기에 대한 새로운 연구방향을 제시하였다고 볼 수 있다. 둘째, 본 연구에서 도출된 금융위기 회복시스템을 실제 투자전략에 응용하여 높은 투자 성과를 얻음으로써, 투자자들에게 새로운 투자 전략을 제시하였다는 점이다. 특히, 금융위기 국면에서는 대부분의 투자자들이 비이성적으로 행동하면서 감정에 휩쓸리기 때문에 큰 손실을 입게 되는데, 본 연구의 결과로 도출된 위험회복지수를 활용하게 되면 금융위기 국면에서 오히려 더 좋은 투자 성과를 얻을 수 있음을 보여주었다. 셋째, 본 연구의 실험 결과를 요약하면, 금융위기의 위기 지속구간과 위험회복구간에 대한 SVM의 커널함수로는 시스모이드 커널함수가 우수한 성과를 시현하였으며, 위험회복지수를 활용한 트레이딩시스템은 연율로 환산한 투자 성과가 121%에 달하여 뛰어난 투자 성과를 기록하였다.

하지만 본 연구는 여러 한계점도 내포하고 있다. 첫째로, 본 연구의 주제와 같은 복잡한 데이터를 분류하는 기계학습으로 다양한 기계학습모형이 알려져 있으나, 본 연구에서는 여러 학습모형과의 비교 검토가 이루어지지 않았다는 점을 들 수 있다. 향후 연구에서는 인공신경망모형 등 SVM보다 더 좋은 성과를 보일 수 있는 여러 모형을 이용한 금융위기의 위험회복시스템에 대한 연구가 이루어져야 한다. 둘째로, 검증용 데이터의 문제점을 지적할 수 있다. 금융위기의 학습용 데이터로는 2008년 글로벌 금융위기 데이터가 선택되어 적절하다고 할 수 있으나, 검증용 데이터로 선택한 2011~2012년의 금융 불안정 구간은 2008년의 금융위기 구간 같은 본격적인 불안정 국간은 아니다. 그러나 그 사이에 본격적인

금융위기 구간이 없었기 때문에, 이 문제점은 향후 나타나게 될 가능성이 있는 본격적인 금융위기에서 데이터를 확보하면 추가적인 분석이 가능할 것이다. 셋째로, 검증 구간에서 제안된 모형의 예측력이 낮게 나타나고 있는데, 이러한 과최적화 문제를 극복하기 위해 교차검증(cross-validation)을 이용해 모형의 파라미터를 선정하고 예측성고가 향상되는지를 분석할 필요가 있다. 마지막으로, 위험회복지수를 도출하는 과정에서 본 연구의 제안과는 달리 변동성지수의 영향력이 제한적인 것으로 나타나고 있는 점은 아쉬운 결과이다. 이러한 결과는 주가지수의 특성변수와 동일한 방법으로 추출된 변동성지수의 특성변수에 기인할 수도 있을 것으로 판단되어, 변동성지수의 새로운 특성변수 추출방법에 대한 깊이 있는 연구가 향후 이루어져야 할 것으로 보인다.

## References

- [1] Ahn, H., Kim, K., and Han, I., "Purchase prediction model using support vector machine", *Journal of Intelligence and Information Systems*, Vol. 11, No. 3, 2005, pp. 69-82.
- [2] Berg, A. and Pattillo, C., "Predicting currency crises : The indicators approach and an alternative", *Journal of International Money and Finance*, Vol. 18, No. 4, 1999, pp. 561-586.
- [3] Berg, J., Candelon, B., and Urbain, J., "A Cautious note on the use of panel models to predict financial crises", *Economics Letters*, Vol. 101, No. 1, 2008, pp. 80-83.
- [4] Boduroglu, I. and Erenay, Z., "A pattern recognition model for predicting a financial

- crisis in Turkey : Turkish economic stability index”, *International Journal of High Performance Computing Applications*, Vol. 21, No. 1, 2007, pp. 5-20.
- [5] Bussiere, M. and Fratzscher, M., “Towards a new early warning system of financial crises”, *Journal of International Money and Finance*, Vol. 25, No. 6, 2006, pp. 953-973.
- [6] Candelon, B., Dumitrescu, E., and Hurlin, C., “How to evaluate an early warning system : Toward a unified statistical framework for assessing financial crises forecasting methods”, *IMF Economic Review*, Vol. 60, No. 1, 2012, pp. 75-113.
- [7] Duasa, J. and Ahmad, N., “Predicting the global crisis recovery period : Lessons from the 1997 crisis”, *Global Economic Review*, Vol. 39, No. 1, 2010, pp. 99-113.
- [8] Ederington, L. and Guan, W., “How asymmetric is U.S. stock market volatility?”, *Journal of Financial Markets*, Vol. 13, No. 2, 2010, pp. 225-248.
- [9] El-Shazly, A., “Designing an early warning system for currency crises : An empirical treatment”, *Applied Economics*, Vol. 43, No. 14, 2011, pp. 1817-1828.
- [10] Fuertes, A. and Kalotychou, E., “Optimal design of early warning systems for sovereign debt crises”, *International Journal of Forecasting*, Vol. 23, No. 1, 2007, pp. 85-100.
- [11] Hibbert, A., Daigler, R., and Dupoyet, B., “A behavioral explanation for the negative asymmetric return-volatility relation”, *Journal of Banking and Finance*, Vol. 32, No. 10, 2008, pp. 2254-2266.
- [12] Kaminsky, G., Lizondo, S., and Reinhart, C., “Leading indicators of currency crises”, *IMF Staff Papers*, Vol. 45, No. 1, 1998, pp. 1-48.
- [13] Kim, D. H., Lee, S. J., Oh, K. J., and Kim, T. Y., “An early warning system for financial crisis using a stock market instability index”, *Expert Systems*, Vol. 26, No. 3, 2009, pp. 260-273.
- [14] Kim, K. J., “Financial time series forecasting using support vector machines”, *Neurocomputing*, Vol. 55, No. 1-2, 2003, pp. 307-319.
- [15] Kim, S. W. and Ahn, H., “Development of an intelligent trading system using support vector machines and genetic algorithms”, *Journal of Intelligence and Information Systems*, Vol. 16, No. 1, 2010, pp. 71-92.
- [16] Kim, S. W., “Negative asymmetric relationship between VKOSPI and KOSPI200”, *Journal of the Korean Data Analysis Society*, Vol. 12, No. 4, 2010, pp. 1761-1773.
- [17] Kim, T. Y., Hwang, C., and Lee, J., “Korean economic condition indicator using a neural network trained on the 1997 crisis”, *Journal of Data Science*, Vol. 2, 2004a, pp. 371-381.
- [18] Kim, T. Y., Oh, K. J., Sohn, I., and Hwang, C., “Usefulness of artificial neural networks for early warning system of economic crisis”, *Expert Systems with Applications*, Vol. 26, 2004b, pp. 583-590.
- [19] Kumar, M., Moorthy, U., and Perraudin, W., “Predicting emerging market currency crashes”, *Journal of Empirical Finance*, Vol. 10, No. 4, 2003, pp. 427-54.
- [20] Oh, K. J., Kim, T. Y., and Kim, C., “An early warning system for detection of financial crisis using financial market volatility”, *Expert*

- Systems*, Vol. 23, No. 2, 2006, pp. 83-98.
- [21] Son, I. S., Oh, K. J., Kim, T. Y., Kim, C., and Do, J. D., "Stock market stability index : An intelligent approach", *Intelligent Data Analysis*, Vol. 13, No. 6, 2009, pp. 983-993.
- [22] Vapnik, V., *The nature of statistical learning theory*, New York, NY : Springer-Verlag, 1995.
- [23] Whaley, R., "The investor fear gauge", *The Journal of Portfolio Management*, Vol. 26, No. 3, 2000, pp. 12-17.
- [24] Yoon, W. J. and Park, K. S., "A study on the market instability index and risk warning levels in early warning system for economic crisis", *Digital Signal Processing*, Vol. 29, 2014, pp. 35-44.

## ■ 저자소개

**김 선 응**

서울대학교에서 경영학사, 한국 과학기술원에서 경영과학 석사 및 경영과학박사를 취득하였고 현재 국민대학교 비즈니스IT 전문대학원 교수로 재직 중이

며, 주요 관심분야는 트레이딩시스템, 자산운용 및 관리, 로보 어드바이저 등이다. 실전주가지수선 물거래 등 투자관련 저서를 출판하였고, Journal of Portfolio Management, Journal of Information Technology Applications & Management, 지능 정보연구, 경영과학, 한국IT서비스학회지, 응용통계연구 등 국내외 학술지에 다수의 논문을 발표하였다.