Journal of the Korean Data & Information Science Society
2011, **22**(1), 37–47

한국데이터정보과학회지

러프 집합을 이용한 코스피 200 주가지수옵션 시장에서의 박스스프레드 전략 실증분석 및 거래 전략

김민식[1] · 오경주[2]

[12]연세대학교 정보산업공학과

접수 2010년 10월 26일, 수정 2010년 12월 26일, 게재확정 2011년 01월 07일

요 약

주가지수 옵션시장에는 많은 투자전략이 개발되어 있다. 그중 차익거래 전략은 시장이 효율성 유지측면에서 매우 중요한 역할을 하고 있다. 본 연구는 이러한 차익거래 전략 중 박스스프레드 전략을 적용하여 과거 옵션 데이터를 통해 사후 검증하고 러프 집합을 이용해 수익성을 향상시키고자 한다. 옵션 데이터는 2002년 1월부터 2006년 12월까지 실제 증권거래소에서 거래되었던 틱 데이터를 기반으로 하고 있으며 비주얼 베이직을 이용해 9시부터 오후 3시까지의 1분 마다의 종가인 1분봉으로 변형하여 분석을 하였다. 박스스프레드 전략은 낮은 위험, 낮은 이익 구조를 가지고 있다. 기존의 전략을 과거 데이터를 기반으로 백 테스팅 해보고 러프 집합을 이용하여 거래 진입 시점을 제한함으로써, 동일 위험 대비 좀 더 높은 수익구조를 만들어 낼 수 있는 전략을 구사한다면 낮은 위험으로 안정적 수익을 취할 수 있다.

주요용어: 러프집합, 박스스프레드, 코스피200 주가지수옵션.

1. 서론

1983년 미국의 S&P100 기반의 주가지수옵션이 생긴 이래 1997년 7월에 우리나라에도 KOSPI200 주가지수를 기반으로 하는 옵션시장이 개설되었다. 현재 대한민국의 옵션 시장은 거래량이나 거래대금 측면에서도 세계 파생상품시장에서 확고한 1위를 차지하고 있다. 옵션이라는 상품은 기초자산이 되는 주식시장과 밀접한 관련을 가지므로 기존의 금융시장과 달리 용도나 역할이 매우 다양하다. 특히 투자자의 입장에서는 위험관리 목적으로 사용할 수도 있는 반면 때론 투기의 용도로 사용할 수도 있는 상품으로 어떠한 목적이냐에 따라 위험성이나 수익성 측면에서 크게 바뀐다고 할 수 있겠다.

KOSPI200 옵션시장은 주가지수인 KOSPI200에 근거하여 가격이 결정되므로 만일 두 가지 상품의 가격 괴리가 생긴다면 효율적 시장의 특성상 시장참여자들은 다양한 방법을 이용하여 그 차이만큼의 이득을 취하려 할 것이다. 이를 통해 짧은 시간 안에 시장은 다시 효율적인 시장으로 회복되어 가격괴리가 사라짐으로써 시장은 다시 균형 상태를 회복한다. 실제로 옵션시장에서는 수많은 차익거래 전략이 존재하고 있으며 누가 먼저 시장의 가격괴리를 찾아내느냐에 따라 수익창출의 기회가 나눠지고 있는 실정이다 (변현우 등, 2009; Kim 등, 2009).

본 연구에서는 이러한 다양한 차익거래 전략 중 박스스프레드 전략을 이용하여 시장의 효율성을 다시 한 번 확인해보고 거래 전략을 변형 및 적용해 봄으로써 실증분석과 함께 수익성 증대를 목적으로 하고

[1] (120-749) 서울특별시 서대분구 신촌동 134번지, 연세대학교 정보산업공학과, 석사과정
[2] 교신저자 : (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 부교수.
E-mail: johanoh@yonsei.ac.k

자 한다. 기존의 박스스프레드 전략은 수학적 분석을 통해 계산된 이론적인 옵션들의 시장가격과 현재 거래되고 있는 가격사이의 차이가 발생하는 경우 포지션을 취하고, 상품의 만기일인 매월 둘째 주 목요일에 포지션을 청산하여 그 차익만큼의 수익을 취하는 구조이다. 하지만 이렇게 계산된 이론가격은 투자자의 판단에 따라 모두 다를 수 있으며 현재 시장가격이 올바른 상황인데도 가격괴리가 발생했다고 판단 할 수 있는 등의 문제점을 가지고 있을 뿐만 아니라, 한번 투자한 자산은 만기일까지 놔둠으로서 자금의 유동성 측면에서도 아쉬운 부분이 많은 실정이다.

이러한 문제를 해결하고 수익성을 향상시키기 위해서는 옵션 상품들의 실제 시장 움직임을 살펴 볼 필요성이 있다. 그리하여 본 논문에서는 과거의 옵션 상품들의 거래 내역을 살펴보고 연구된 전략을 사용하였을 때 수익이 나는지 확인하기 위하여 과거의 가격수준과 다양한 변수요소들을 현재 시점으로 가정하고 실험을 하는 백 테스팅 프로그램을 프로그램 언어인 비주얼 베이직을 이용하여 개발하였다. 또한 러프 집합을 이용해 패턴분석을 통한 포지션 진입 시점을 찾아냄으로서 제안된 거래 모델의 수익성을 향상 시킬 수 있는 방향을 제시하고자 한다.

2장에서는 선행연구를 고찰했으며, 3장에서는 본 연구가 제시하는 모델을 자세히 기술했으며, 4장에서는 연구모델의 실증분석을 행하고, 본 연구에서 제시하는 모델의 유용성과 우수성을 입증하였다. 마지막으로 본 연구의 기대효과 및 향후 연구사항에 대해 서술하였다.

2. 선행연구

2.1. 박스스프레드 차익거래전략

차익거래란 동일한 상품이 서로 다른 가격으로 거래되는 경우 고평가되어 있는 자산은 매도하고 저평가되어 있는 자산은 매수하여 그 차이만큼의 이익을 얻는 구조이다. KOSPI200 옵션시장에서 사용되는 박스스프레드 거래전략은 크레딧박스 (credit box)스프레드 차익거래와 데빗박스 (debit box)스프레드 차익거래도 나뉘며, 두 가지 거래전략 모두 옵션가격만을 이용하여 옵션가격에서 발생하는 가격불균형을 이용하여 이익을 얻는 거래전략이다 (이재하와 한덕희, 2006; 오세헌, 2002).

본 연구에서는 두 가지 박스스프레드 차익거래 전략 중 크레딧박스 스프레드 전략을 사용하였다. 행사가격이 높은 (X_2) 콜옵션 (C_2)을 매수하고 풋옵션(P_2)을 매도하며, 이와 함께 행사가격이 낮은 (X_1) 콜옵션 (C_1)을 매도하고 풋옵션 (P_1)을 매수하는 포지션을 취함으로서 현재는 $-C_2 + P_2 + C_1 - P_1 - Sleep_{now}$의 현금흐름이 발생하고, 옵션의 만기가 되는 T 시점에서는 항상 일정한 $X_1 - X_2 - Sleep_T$의 수익이 발생하는 전략으로 볼 수 있다 (태석준, 2000). 이 때 현재시점의 현금흐름과 만기시점의 수익의 차이가 발생 할 경우 포지션을 취함으로서 이익을 내는 구조이다 (단, 여기서 $Sleep_{now}$와 $Sleep_T$는 실제 거래시 존재하는 슬리피즈 가격으로, 거래를 성사시키기 위해 시장가격보다 한 단계 높은 가격으로 주문을 함으로써 발생되는 손실액이다. 여기서 한 단계 높은 가격이란 거래소에서 제공하는 주문가격단위를 기준으로 하여 현재 거래되고 있는 호가보다 바로 위의 가격단위로 상품을 사겠다고 주문하는 것으로써, 현재 대한민국 옵션시장의 규칙에 따라 유동적으로 변한다. 예를 들어 물건이 만원에 시장에서 거래되고 있는데 살 수있는 단위는 천원단위라고 한다면 만 원짜리를 만 천원을 주고 구입하는 이치와 같다고 할 수 있다.)

2.2. 러프집합

러프집합 이론은 1982년 Pawlak (Pawlak, 1994; Pawlak, 1997)에 의해 처음으로 제시되었다 (노태협 등, 2005). 이는 불확실성으로 가득한 데이터를 속성들 간의 중요성과 상관성을 찾아냄으로서 모든 개체들을 각각의 속성이 제공하는 정보에 따라 집합 짓는 것이다. 즉, 이러한 동질성 관계를 찾는 것이

러프집합 이론의 기초가 된다. 즉 새로운 데이터를 분류할 때 해당 데이터가 가지는 속성들을 과거 데이터들의 속성과 비교하고 분류하여, 과거 데이터들이 구분되어 있는 특정 집합 단위중 하나로 포함시키는 과정을 거치게 된다.

본 연구에서는 이러한 방법론을 이용하여 박스스프레드 전략의 포지션 진입 시점을 각각의 옵션들의 Greek를 속성 변수로 이용하여 집합지음으로써 수익이 났던 시점과 손실이 났던 시점을 구분해 내고자 한다.

2.3. 옵션 Greeks

옵션 Greeks는 옵션 가격의 민감도를 나타내는 것으로써 크게 5가지 Delta, Gamma, Theta, Roh, Vega로 표시된다. Delta는 기초자산 가격에 대한 옵션가격의 민감도를 나타내며, Gamma 는 기초자산 가격에 대한 Delta의 변화를 나타낸다. Theta는 옵션 잔존 만기에 대한 옵션가격의 변화를, Roh는 이자율 변화에 대한 옵션가격의 변화, Vega는 변동성에 대한 옵션가격의 변화를 나타낸다.

3. 제안모델

3.1. 모델의 구성

본 연구에서는 크레딧박스 스프레드 전략의 기초가 되는 4개 옵션의 시장가격들의 움직임을 살펴보고 가격들의 비율움직임에 따른 포지션 진입 · 청산 시점을 찾아본다. 또한 러프집합을 이용하여 과거 데이터들의 움직임을 살펴봄으로써 수익이 나는 진입시점을 찾아내고 손실이 나는 진입시점에는 거래를 제한하는 모델을 제안했다. 따라서 본 연구모델은 크게 3단계로 구성된다.

제 1단계는 증권거래소의 틱 데이터를 1분봉으로 가공한 뒤 (송치우와 오경주, 2009), 크레딧박스 스프레드를 구성하는 4개 옵션의 가격 데이터간의 관계를 살펴보고 특정 움직임이 있는지를 살펴본다. 각각의 가격을 2개씩 한 그룹으로 구성한 뒤 한 그룹을 나머지 그룹으로 나누어 비율을 계산하였다. 본 연구에서는 콜옵션 가격들의 합을 분자 그룹으로, 풋옵션 가격들의 합을 분모 그룹으로 구성하였으며, 이렇게 생성된 비율의 움직임을 살펴보고 중심이 되는 평균값을 찾아보았다.

제 2단계는 1단계에서 구성된 비율의 움직임의 구성을 살펴보고 임의의 밴드를 설정하여 진입 시점과 청산 시점을 찾아보았다. 옵션 가격비율의 평균을 중심으로 하여 위·아래로 특정간격 값들보다 높거나 낮은 경우를 진입밴드로 설정하였으며 (특정 밴드범위를 벗어나는 경우로서 일반적인 범주내에 있지 않기 때문), 또 다른 간격 값들 사이로 가격 비율이 형성되는 경우 (특정 밴드 범위내에 가격비율이 들어오는 경우로서 일반적인 범위내에서 움직이는 경우) 청산하는 전략을 취하였다. 이러한 밴드의 설정은 진입 밴드 2가지, 청산 밴드 2가지로 구성하여 총 4가지의 조합을 구성한 거래 전략을 사용하였다.

제 3단계는 2단계에서 실시한 4가지 조합의 거래 전략 중 수익성이 높은 전략을 선택하여 러프 집합을 이용하여 진입 시점을 찾아낸다. 크레딧박스 스프레드를 구성하는 4가지 옵션의 개별 Greeks 변수 5가지와 각 옵션의 내재변동성을 이용하여, 총 24가지 변수를 사전입력변수로 선택하였다. 그 후 개별변수와 수익·손실 결과와의 관계를 살펴보고 그중 가장 영향을 많이 미치는 변수 5개를 선택한다. 마지막으로 선택된 5개 변수들 중 3개씩을 골라 의사결정을 하는 모델을 구축하였다. 즉 1단계와 2단계를 거쳐 진입·청산시점을 선택하는 모델의 진입시점을 한번 더 판단해 냄으로서 수익의 증대를 이룰 수 있는지를 확인해 보았다.

3.2. 모델구축단계

3.2.1. 크레딧박스 스프레드를 구성하는 옵션가격 비율의 움직임

밴드를 설정하기에 앞서 크레딧박스 스프레드를 구성하는 4가지 옵션에 대하여 각각의 가격정보를 이용하여 같은 시간의 가격끼리의 비율을 찾아보았다. 수식 (3.1)은 본 연구에 사용된 비율을 구하는 식이며 앞으로 p_ratio라 정의한다.

$$p_ratio = (Call^{+n} + Call^{-n})/(Put^{+n} + Put^{-n}) \tag{3.1}$$

C^{+n} : ATM으로부터 n단계 위의 행사가를 가지는 콜옵션
C^{-n} : ATM으로부터 n단계 아래의 행사가를 가지는 콜옵션
P^{+n} : ATM으로부터 n단계 위의 행사가를 가지는 풋옵션
P^{-n} : ATM으로부터 n단계 아래의 행사가를 가지는 풋옵션
ATM : 거래되고 있는 옵션의 행사가격과 행사가격이 같은 상태의 행사가격

이렇게 계산된 p_ratio의 5년간의 결과를 살펴보면 그림 3.1과 같다. 전체적으로 일정한 값을 유지하며 그 값을 중심으로 움직이고 있음을 알 수 있다. 이 중심 값을 기준으로 하여 일정범위만큼을 벗어날 때는 가격 불균형의 상황이 되며 다시 원래 값으로 돌아와 안정을 찾는다고 판단할 수 있다.

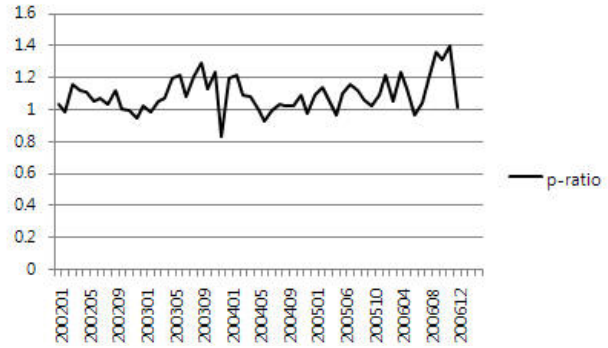

그림 3.1 2002년1월 2006년12월 p_ratio

3.2.2. 밴드설정에 따른 거래 전략 구축

p_ratio의 결과를 바탕으로 하여 포지션 진입·청산의 기준이 되는 밴드를 설정하였다. 진입시점은 p_ratio가 특정 상한 값을 넘어가거나 특정 하한 값보다 아래의 수치를 나타나는 순간을 기준으로 하였으며, 청산시점은 평균을 기준으로 특정 범위 내에 있는 경우를 기준으로 하였다. 수식 (3.2)는 포지션 진입 기준밴드를 나타내고 있으며 수식 (3.3)은 포지션 청산 기준밴드를 나타내고 있다. 본 연구에서는 p_ratio가 많이 몰려있는 가운데 부분을 청산밴드로 하고, 드물게 나타나는 바깥쪽 부분을 진입밴드로 설정하기 위하여 아래 수식 (3.2)와 (3.3)에서의 a, b 를 $a > b$ 하도록 설정하였다.

$$\begin{aligned} Buy_{up} &: mean_{p_ratio} + a \bullet \sigma_{p_ratio} \\ Buy_{dn} &: mean_{p_ratio} - a \bullet \sigma_{p_ratio} \end{aligned} \tag{3.2}$$

$$Sell_{up} : mean_{p_ratio} + b \bullet \sigma_{p_ratio}$$
$$Sell_{dn} : mean_{p_ratio} - b \bullet \sigma_{p_ratio} \quad (3.3)$$

a, b : 임의의 상수 (단, $a > b$)

진입과 청산의 기준을 살펴보면 수식 (3.4)와 같다.

$$\text{진입} : (p_ratio < Buy_{dn}) or (p_ratio > Buy_{up})$$
$$\text{청산} : Sell_{dn} < p_{-ratio} < Sell_{up} \quad (3.4)$$

본 연구에서는 1단계 실험에서 나온 p_ratio의 분포를 살펴보고 따로 동떨어져있는 2 부분과 가장 몰려 있는 -0.5와 0.5 값을 기준으로 임의로 a, b 의 값을 2개씩 설정하여 총 4가지 경우로 거래 시뮬레이션을 하였다. 표3.1은 a, b 에 다른 각각의 케이스를 나타내고 있으며, 하나의 Case마다 수식 (3.1)에서의 n 값을 1부터 4까지 조정하여 4개씩의 거래 시뮬레이션을 하였다. 즉 2단계에서 총 16가지의 거래시뮬레이션이 이루어 졌다. n 값의 변화에 따른 크레딧박스 스프레드의 명칭은 $ATM \pm 1$ 부터 $ATM \pm 4$ 까지로 표시한다. 그림 3.2는 수식 (3.2)와 수식 (3.3)의 밴드를 그림으로 도식화 한 것이며, 진입·청산 밴드와 $ATM \pm n$ 거래 전략의 개념도(n=2 인 경우)이다.

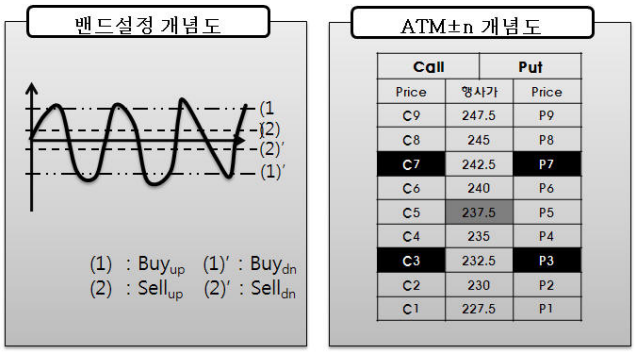

그림 3.2 밴드설정과 ATM ± n 의 개념도(n=2 인 경우)

표 3.1 a, b 값에 따른 거래시뮬레이션

Case	a	b
Case1	2.0	0.25
Case2	2.5	0.25
Case3	2.0	0.5
Case4	2.5	0.5

3.2.3. 러프집합을 통한 진입시점 수정

2단계의 16가지 거래 시뮬레이션을 통해 나온 수익의 결과 값을 이용하여 수익·손실의 결과와 각 옵션들과의 관계를 이용해 진입시점을 제한하고자 러프집합을 적용하였다. 실험은 러프집합 모델을 구축하고 적용할 수 있는 로제타라는 프로그램을 이용하였으며 전체적인 실험과정은 그림 3.3과 같다.

초기 입력변수는 크레딧박스 스프레드 전략을 구성하는 각각의 옵션의 개별 내재변동성 (Implied volatility)과 옵션 Greek인 Delta, Gamma, Vega, Theta, Roh 의 6가지씩의 변수 값을 계산하여 총 24가지 변수를 선정하였다. 선정된 변수들은 옵션 Greek와 내재변동성으로서 이들은 연속형 변수의 속성을 가지고 있다. 그러므로 러프 집합에 적용시키기 위하여 이산화 과정을 필요로 한다 (서완석과 김재련, 2005). 연속형 변수의 이산화 과정은 해당 변수를 위한 경계 값을 생성함으로써 임의로 간격을 구분지어 주는 역할을 한다. 일반적으로 통용되는 방법론은 4가지 정도가 있으며 (노태협 등, 2005; Øhrn, 2001), 본 연구에서는 각 변수들을 독립적이라 가정한 뒤, 결과 값이나 변수값 (조건 변수)에 대한 고려를 하지 않고 범주화 시키는 Equal Frequency Scaling 방법을 적용하였다 (이철희와 서선학, 2005). 또한 실제 실험에 사용하기에 24개의 변수는 너무 많으므로 러프집합을 이용하여 각 시뮬레이션당 3개씩 변수를 선정하여 시뮬레이션에 적용하였다. 즉, Phase1 과 Phase2 과정을 24개의 변수에 대해 한번 적용한 뒤, 변수를 선정하는 과정에서는 하나의 변수와 결과 값과의 관계를 직접 찾아 가장 우수한 예측 결과를 가지는 상위 3개씩의 변수를 선정하는 방법을 사용하였다. 이렇게 나온 변수의 조합을 이용하여 Phase1과 Phase2 과정을 다시한번 적용한 뒤 트레이딩 시뮬레이션을 적용하였다.러프집합을 이용한 시뮬레이션에서는 2단계의 16개 거래 전략의 결과를 종속변수로 하여 2002년부터 2005년까지의 데이터를 트레이딩 구간으로 두고, 2006년 데이터를 테스팅 하였다. 또한 여기서 나온 결과 값을 기준으로 하여 가장 우수한 상위 2가지의 거래 전략을 다시 한번 트레이딩과 테스팅 구간을 조절하며 실험 검증과정을 거쳤다.

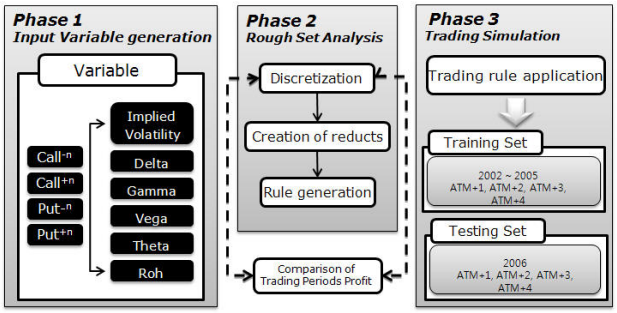

그림 3.3 러프집합 모형 구축 및 실험과정

4. 실증분석

4.1. 크레딧박스 스프레드 비율 계산

2002년 1월부터 2006년 12월까지 크레딧박스 스프레드를 구성하는 4개 옵션의 1분봉 가격 데이터를 이용하여 p_ratio를 계산하였다. 옵션시장이 열리는 아침9시 1분봉부터 장이 마감되는 3시 15분까지의 매분의 종가를 바탕으로 계산되었으며 5년간의 평균인 $mean_{p_ratio}$는 약 1.088의 값을 나타내는 것으로 측정되었다. 또한 분산인 σ_{p_ratio}는 0.108의 값을 가지며 그래프를 그려보았을 때, 대략 정규분포와 비슷한 모습을 나타내고 있음을 알 수 있었다. 그림 4.1은 이를 잘 나타내고 있다.

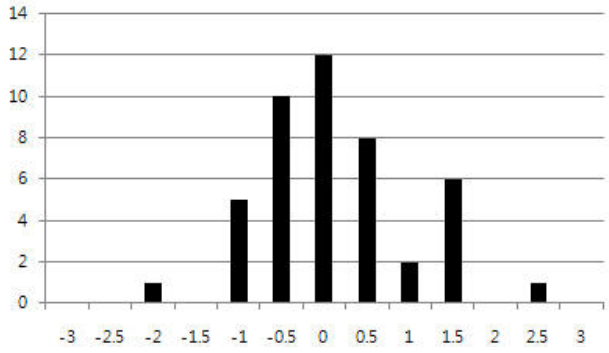

그림 4.1 p_ratio 의 움직임

4.2. 밴드 설정에 따른 크레딧박스 스프레드 전략의 수익 및 손실 결과

크레딧박스 스프레드 전략의 진입·청산 시점은 표 3.1과 같이 총 4가지 Case의 경우를 실험하였다. a 값을 조정함으로써 포지션의 진입 빈도수를 조절하였으며 (a값이 클수록 진입 기회는 줄어듦) b값을 조정함으로써 포지션의 청산 빈도수를 조정 하였다 (b값이 클수록 청산 기회는 늘어남). 2002년 1월부터 2006년 12월까지의 옵션 1분봉 가격 데이터를 이용하여 백 테스팅을 하였다 (표의 profit 부분의 단위는 point로써 1point 는 원화 10만원을 의미함).

표 4.1에서 각각의 Case 내에서의 n 의 값이 증가함에 따라 수익의 변화를 살펴보자. Case1부터 Case4까지 모두, 개별 Case 내에서 n 의 값이 1부터 3까지는 손실을 기록하고 있음을 알 수 있다. 이는 슬리피즈 때문에 이러한 결과가 나오는 데, 그만큼 현재 우리나라 옵션 시장은 차익거래를 이용하여 수익을 내기 쉽지 않다는 것이다. 하지만 n 의 값이 4로 변하는 순간부터 손실이 수익으로 바뀜을 볼 수 있는데 (n의 값이 커짐에 따라 손실의 크기도 줄어듦을 알 수 있다), 이는 옵션의 거래가 되는 KOSPI200 주가지수로부터 행사가가 멀어지면 멀어질수록 가격괴리가 많이 발생한다고 해석할 수 있으며 현재 ATM으로부터 간격이 먼 행사가의 옵션들의 거래 과정에서 시장 불균형의 상황이 많이 나온다고 해석할 수 있다. 본 연구에서는 이러한 변화에 초점을 두고 손실 구조에서 수익구조로 변하는 $ATM \pm 3ATM \pm 4$ 크레딧박스 스프레드 전략을 러프집합에 적용시켜 실험을 하였다.

4.3. 러프 집합을 적용한 크레딧박스 스프레드

2단계 실험 결과를 바탕으로 Case1부터 Case4까지의 $ATM \pm 3ATM \pm 4$ 의 크레딧박스 스프레드 전략의 진입시점을 러프 집합을 이용하여 재설정하였다. 표 4.2와 같이 2002년부터 2005년까지의 데이터를 트레이닝 데이터로 설정하고 2006년 데이터를 테스팅 함으로써 수익률 향상을 살펴보았다.

표 4.3의 Testing 데이터 비교 부분의 기존 수익은 진입·청산 밴드 설정에 따른 각 Case별 $ATM \pm 3ATM \pm 4$ 전략들의 2006년 1년 동안의 실증 거래 분석 결과이다. 이러한 기존의 2단계 실험 결과를 종속변수로 두고 러프 집합을 이용해 전략 포지션의 진입 시점을 예측하고 제한함으로서 8가지의 거래 시뮬레이션 중 Case1의 $ATM \pm 4$ 부분을 제외하고는 7가지 중 5가지 거래 전략은 손실의 폭이 줄어들었음을 알 수 있으며, 2가지 전략은 손실에서 수익구조로 전환됨을 볼 수 있다. 이는 옵션의 Greek와 내재 변동성의 속성이 개별 옵션의 현재 상황을 잘 나타내주고 있으며 이들의 과거 움직임과 패턴을 분석함으로써 현재 주어진 4개 옵션의 가격 비율 상황에서 일정 조건을 만족하면 수익을 낼 수 있음을 증

명한다.

본 연구에서는 러프 집합을 적용한 8가지 거래 전략 중에서 2006년 테스팅 구간 중 가장 성과가 좋은, 손실에서 수익으로 바뀌는 구조를 나타내는 Case2의 $ATM \pm 4$ 전략과, Case4의 $ATM \pm 4$ 을 다시한 번 검증하는 과정을 거쳤다. 실험은 표 4.4와 같이 1년을 단위로 하여 테스팅 구간과 트레이닝 구간을 바꾸어 가면서 러프 집합을 적용하였다.

표 4.1 각 Case별 4가지 크레딧박스 스프레드의 결과

Case	구분	ATMpm1	ATMpm2	ATMpm3	ATMpm4
Case1	Winning Trade (Number)	10	34	58	51
	Total Trade (number)	348	299	220	127
	Win rate (%)	2.87	11.37	26.36	40.16
	Profit (point)	-80.7	-66.65	-26.45	16.75
Case2	Winning Trade (Number)	9	32	55	51
	Total Trade (number)	320	289	198	119
	Win rate (%)	2.81	11.07	27.78	42.86
	Profit (point)	-73.42	-60.33	-13.54	11.11
Case3	Winning Trade (Number)	10	34	51	55
	Total Trade (number)	421	338	246	149
	Win rate (%)	2.38	10.06	20.73	36.91
	Profit (point)	-97.58	-77.87	-38.88	10.32
Case4	Winning Trade (Number)	8	25	50	46
	Total Trade (number)	382	316	225	132
	Win rate (%)	2.09	7.91	22.22	34.85
	Profit (point)	-87.01	-68.45	-28.39	4.34

표 4.2 러프 집합에 사용된 Training 및 Testing 구간

Case	구분	Training	Testing
Case1	$ATM \pm 3$	2002 ~ 2005 (1분봉)	2006 (1분봉)
	$ATM \pm 4$		
Case2	$ATM \pm 3$		
	$ATM \pm 4$		
Case3	$ATM \pm 3$		
	$ATM \pm 4$		
Case4	$ATM \pm 3$		
	$ATM \pm 4$		

표 4.3 러프 집합을 적용한 결과와 적용 전 Testing 구간의 비교

Case	구분	Training : Hit ratio (%)	Testing : Hit ratio (%)	Testing 데이터 비교(point)	
				기존 수익	러프집합적용후수익
Case1	$ATM \pm 3$	97.1	53.5	-11.09	-2.07
	$ATM \pm 4$	96.5	47.0	2.82	-0.57
Case2	$ATM \pm 3$	92.2	79.6	-6.56	-0.64
	$ATM \pm 4$	96.5	49.6	-0.24	0.42
Case3	$ATM \pm 3$	93.8	31.0	-15.33	-6.4
	$ATM \pm 4$	97.4	52.0	-1.04	-0.45
Case4	$ATM \pm 3$	96.6	45.3	-8.71	-2.02
	$ATM \pm 4$	98.2	68.3	-0.48	0.78

표 4.4 러프 집합 검증에 사용된 전략과 Training 및 Testing 구간 (1분봉)

전략	구분	Training	Testing
Case2($ATM \pm 4$)	Case2_01	2003 ~ 2006	2002
	Case2_02	2002, 2004 ~ 2006	2003
	Case2_03	2002 ~ 2003, 2005 ~ 2006	2004
	Case2_04	2002 ~ 2004, 2006	2005
	Case2_05	2002 ~ 2005	2006
Case4($ATM \pm 4$)	Case4_01	2003 ~ 2006	2002
	Case4_02	2002, 2004 ~ 2006	2003
	Case4_03	2002 ~ 2003, 2005 ~ 2006	2004
	Case4_04	2002 ~ 2004, 2006	2005
	Case4_05	2002 ~ 2005	2006

차익거래 모델은 시간의 흐름에 따른 영향을 받지 않으므로 테스팅 구간에 해당하는 년차를 바꿈으로써 자유롭게 실험이 가능하였다. 표4.5에서 나오는 것과 같이 두 가지 전략의 5가지 실험결과를 살펴보면 총 10가지 중에서 6가지나 수익의 상승이 있음을 알 수 있다. 이는 2단계 실험에서 부족했던 부분을 좀 더 심화하여 차익거래의 수익률을 향상 시킬 수 있음을 말해주고 있다. 이중에서도 Case4의 $ATM \pm 4$ 전략의 러프 집합 적용은 매우 우수한 결과를 나타내고 있음을 알 수 있다. 2단계의 밴드설정을 이용한 백 테스팅에서 5년간 총 4.34 point (Case4_01부터 Case4_05까지의 기존수익의 총합계)의 수익을 얻을 수 있었지만 러프 집합을 적용시킴으로써 8.75 point (Case4_01부터 Case4_05까지의 러프집합 적용후 수익의 총합계)의 수익을 냄으로써 대략 2배 정도의 수익증가가 가능함을 볼 수 있다.

표 4.5 러프 집합을 적용한 결과와 적용 전 Testing 구간의 비교

전략	구분	Training : Hit ratio (%)	Testing : Hit ratio (%)	Testing 데이터 비교(point)	
				기존 수익	러프집합적용후수익
Case2($ATM \pm 4$)	Case2_01	96	64	10.71	7.36
	Case2_02	97	70	-0.85	1.86
	Case2_03	97	67	0.02	0.33
	Case2_04	99	68	1.47	0.46
	Case2_05	96	49	-0.24	0.42
Case4($ATM \pm 4$)	Case4_01	100	49	7.36	6.28
	Case4_02	97	66	-0.99	1.81
	Case4_03	98	68	-1.6	0.02
	Case4_04	99	70	0.05	-0.14
	Case4_05	98	68	-0.48	0.78

5. 결론

1997년 대한민국 증권시장에 주가지수옵션 시장이 개설된 이래 근 10년 동안 거래량 및 거래금액 면에서 세계 1위 시장의 자리를 지키며 괄목할 만한 성장을 지속하고 있다. 그 성장세를 다른 각도에서 보면 한국의 금융시장의 변동성이 다른 어느 나라보다도 큼을 암시한다 (Oh, 2004). 옵션시장은 그 목적이 주식시장에서 채워주지 못하는 투자위험의 헷지 성격과 함께 투자자들에게 투기의 목적으로도 매우 큰 매력이 있는 시장이다. 특히 한국 시장은 시장 참여자의 80%가 개인투자자라는데 있어서도 시장의 효율성 유지는 옵션시장이 제 기능을 유지하기 위해 매우 중요한 역할을 한다. 효율적인 시장이란 같은 상품이 같은 시각에 같은 가격으로 팔려야 함을 의미하며 어디선가 조금이라도 이 균형이 깨어지는 순간 투자자들은 그 차이만큼 이득을 보려 할 것이기 때문에 차익거래 전략의 다양성과 수익성은 시장이 가격

괴리와 같은 비효율적 상황에 직면 했을 때 원래 균형을 찾을 수 있다. 이러한 점에서 많은 차익거래 중 하나인 크레딧박스 스프레드 전략을 이용하여 낮은 위험대비 낮은 수익을 좀 더 수익을 창출해내기 위한 연구를 했다는 점에서 의의를 두고 싶다. 또한 개별옵션의 속성을 미리 파악함으로써 시장 불균형인 상태를 남들보다 빨리 찾아낼 수 있으며 수익원을 찾아낸다는 것은 매우 우수한 시장진입 기회를 가질 수 있음을 나타낸다.

본 연구에서는 기존의 크레딧박스 스프레드 전략의 기본 개념인 포지션 진입 후 만기 청산이라는 법칙을 변형하고 밴드를 설정함으로써, 옵션 포지션을 취한 뒤 만기까지 기다리지 않고 언제나 청산함으로서 자금의 유동성을 확보할 수 있다. 뿐만 아니라 러프 집합을 이용하여 옵션의 속성과 수익기회의 관계를 찾아봄으로써 기존의 전략의 낮음 위험의 장점은 가져가되 낮은 수익을 향상시킬 수 있다. 마지막으로 위의 실험결과는 한단위의 포지션만을 사고파는 거래를 반복하여 실험하였다. 이는 같은 포지션의 매수·매도 형태를 유지하며 옵션의 거래 개수를 늘림으로써 수익의 증가도 옵션의 개수에 비례하여 증가한다는 점에 매우 큰 의의를 둔다. 시장의 거래 유동성을 깨지 않는 선에서 많은 개수의 포지션을 잡음으로써 큰 수익을 올릴 수 있을 것이다. 이번 실험에서는 크레딧박스 스프레드 전략의 포지션 최적개수를 찾아보며 더욱 더 안정된 거래 밴드의 경우까지는 고려하지 못하였다. 이 부분은 추후 연구 과제로 남기겠다.

참고문헌

노태협, 유명환, 한인구 (2005). 러프집합이론과 사례기반추론을 결합한 기업신용평가 모형. <정보시스템연구>, **14**, 41-65

변현우, 송치우, 한성권, 이태규, 오경주 (2009). 변동성 지수기반 유전자 알고리즘을 활용한 계층구조 포트폴리오 최적화에 관한 연구. <한국데이터정보과학회지>, **20**, 467-478

서완석, 김재련 (2005). 러프집합이론과 SOM을 이용한 연속형 속성의 이산화. <산업경영시스템학회지>, **28**, 1-7

송치우, 오경주 (2009). 코스피 200 주가지수옵션 데이터의 효율적 가공을 통한 다양한 옵션 전략들의 사후검증에 관한 연구. <한국데이터정보과학회>, **20**, 1061-1073

오세헌 (2002). <KOSPI200주가지수 선물 및 옵션시장의 차익거래전략에 관한 실증 연구>, 석사학위논문, 한국과학기술원, 서울.

이재하, 한덕희 (2006). KOSPI200 옵션시장에서의 박스스프레드 차익거래 수익성. <선물연구>, **14**, 109-137

이철희, 서선학 (2002). 러프집합과 계층적 분류구조를 이용한 데이터마이닝에서 분류지식발견. <한국지능시스템학회논문지>, **12**, 202-209

태석준 (2000). KOSPI 200 옵션 市場에서의 박스 스프레드 戰略에 관한 硏究. <경영연구>, **9**, 243-260

Kim. K. K., Cho. M. H. and Park. E. S. (2008). Forecasting the volatility of KOSPI 200 using data mining. *Journal of Korean Data and Information Science Society*, **19**, 1305-1325

Oh, K. J. (2004). Comparing change-point detection methods to detect the korea economic crisis of 1997. *Journal of Korea Data and Information Science Society*, **15**, 585-592

Øhrn, A., (2001). ROSETTA Technical Reference Manual.

Pawlak, Z. and Sowinski, R. (1994). Rough set approach to multi-attribute decision analysis. *European Journal of Operational Research*, **72**, 443-459

Pawlak, Z. (1997). Rough set approach to knowledge-based decision support. *European Journal of Operational Research*, **99**, 48-57

Journal of the Korean Data & Information Science Society
2011, **22**(1), 37–47

한국데이터정보과학회지

Using rough set to support arbitrage box spread strategies in KOSPI 200 option markets

Min Sik Kim[1] · Kyong Joo Oh[2]

[12]Department of Information and Industrial Engineering, Yonsei University

Received 26 October 2010, revised 26 December 2010, accepted 07 January 2011

Abstract

Stock price index option market has various investment strategies that have been developed. Specially, arbitrage strategies are very important to be efficient in option market. The purpose of this study is to improve profit using rough set and Box spread by using past option trading data. Option trading data was based on an actual stock exchange market tick data ranging from 2001 to 2006. Validation process was carried out by transferring the tick data into one-minute intervals. Box spread arbitrage strategies is low risk but low profit. It can be accomplished by back-testing of the existing strategy of the past data and by using rough set, which limit the time line of dealing. This study can make more stable profits with lower risk if control the strategy that can produces a higher profit module compared to that of the same level of risk.

Keywords: Box spread, KOSPI 200, rough set.

[1] Graduate Student, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea.
[2] Corresponding Author : Associate Professor, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea. E-mail: johanoh@yonsei.ac.kr