creative commons

COMMONS DEED

저작자표시-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이 저작물을 영리 목적으로 이용할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공학석사 학위논문

횡단면적 군집분석을 이용한 주식투자 전략: KOSPI 200 주식을 중심으로

KOSPI 200 stock investment strategy using the cross sectional clustering analysis

2014년 8월

서울대학교 대학원

산업공학과

신 승 민

횡단면적 군집분석을 이용한 주식투자 전략: KOSPI 200 주식을 중심으로

지도 교수 장 우 진

이 논문을 공학석사 학위논문으로 제출함

2014년 6월

서울대학교 대학원
산업공학과
신 승 민

신승민의 공학석사 학위논문을 인준함
2014년 8월

위 원 장	조 성 준	(인)
부위원장	장 우 진	(인)
위 원	이 재 욱	(인)

국문초록

횡단면적 군집분석을 이용한 주식투자 전략: KOSPI 200 주식을 중심으로

서울대학교 대학원
산업공학과
신 승 민

본 논문은 2003년 4분기부터 2011년 1분기까지 kospi200에 속했던 기업들을 분석하고, 그 결과를 이용해 분기별 최적의 투자전략을 제시한다. 기업의 분기별 재무제표를 이용해서 횡단면적 군집분석을 실시하고 이를 분석함으로써 기업의 재무구조가 크게 바뀌는 시점과 바뀐다면 어떻게 바뀌는지 찾아냈다. 횡단면적 군집분석의 결과로 기업은 매 분기마다 특정 군집에 속하게 되는데, 이때 속한 군집이 변하게 되면, 예를 들어 LG전자가 2005년 1분기에는 군집2에 속해있고, 2005년 2분기에는 군집3에 속할 경우, 이를 군집간 이동이 발생했다고 한다. 따라서 이 군집간 이동은 곧 기업의 재무구조가 변했음을 의미한다. 재무구조의 변화는 필연적으로 주가에 영향을 주기 때문에, 재무제표 발표일에 해당 주식을 매입하고, 일정기간 후에 매도함으로써 수익을 창출할 수 있다. 이때 각 군집간 이동 유형 마다 초과수익률의 분포를 분석함으로써 '어느 분기에 어떤 식으로 재무구조가 바뀌었을 때' 투자하는 것이 가장 좋은 결과를 기대할 수 있는지 살펴보았다.

Keywords: 횡단면적 군집분석, 투자전략, 재무제표
Student Number: 2012-21061

목 차

제 1 장. 서론.....1

제 2 장. 데이터.....4

2.1. 분기별 재무제표.....4

2.2. 수정주가.....6

2.3. KOSPI200 지수.....7

제 3 장. 연구기법.....8

제 4 장. 군집화(clustering) 결과 및 분석.....10

4.1. 결과.....10

4.2. 분석.....13

4.2.1. 군집별 특징.....13

4.2.2. 군집간 이동.....15

제 5 장. 투자전략.....24

5.1. 투자방법.....24

5.2. 분기별 기간별 초과수익률의 분포.....25

5.3. 분기별 투자전략 제시.....35

제 6 장. 결론.....39

참고문헌.....40

표 목차

표 1. 지표에 따른 재무인자(비율) 목록 4
표 2. 군집별 재무인자, 재무비율의 평균값 14
표 3. 군집별 표준화된 재무인자 재무비율의 평균값 14
표 4. 군집2와 군집3 비교 19
표 5. 군집2와 군집5 비교 19
표 6. 군집4와 군집5 비교 19
표 7. 1분기 군집간 이동 및 머무름 횟수 20
표 8. 2분기 군집간 이동 및 머무름 횟수 20
표 9. 3분기 군집간 이동 및 머무름 횟수 21
표 10. 4분기 군집간 이동 및 머무름 횟수 21
표 11. 1분기 기간별 평균 초과수익률 26
표 12. 2분기 기간별 평균 초과수익률 27
표 13. 4분기 기간별 평균 초과수익률 28
표 14. 3분기 기간별 평균 초과수익률 29
표 15. 1분기 투자대안별 초과수익률의 평균, 표준편차 31
표 16. 2분기 투자대안별 초과수익률의 평균, 표준편차 32
표 17. 3분기 투자대안별 초과수익률의 평균, 표준편차 33
표 18. 4분기 투자대안별 초과수익률의 평균, 표준편차 34

그림 목차

그림 1. 군집의 개수에 따른 silhouette coefficient 값.....9

그림 2. 표준화된 시가총액과 부채비율을 축으로 한 산포도.....10

그림 3. 표준화된 자산과 매출액을 축으로 한 산포도.....11

그림 4. 군집별 개체 수의 비율.....11

그림 5. 연구기간 동안 분기마다 군집에 속한 개체 수의 변화.....12

그림 6. 연구기간 동안의 군집간 이동과 군집 내 머무름 개요.....17

그림 7.연도별 특정 군집간 이동횟수.....22

그림 8. 1분기 투자대안별 초과수익률 분포.....31

그림 9. 2분기 투자대안별 초과수익률 분포.....32

그림 10. 3분기 투자대안별 초과수익률 분포.....33

그림 11. 4분기 투자대안별 초과수익률 분포.....34

그림 12. 1분기 투자대안별 초과수익률의 평균-표준편차.....36

그림 13. 2분기 투자대안별 초과수익률의 평균-표준편차.....36

그림 14. 3분기 투자대안별 초과수익률의 평균-표준편차.....37

그림 15. 4분기 투자대안별 초과수익률의 평균-표준편차.....37

제 1 장. 서론

전통적인 증권분석 방법은 크게 기술적 분석(technical analysis)과 기본적 분석(fundamental analysis)으로 나뉜다. 여기서 기술적 분석은 과거 주식의 가격이나 거래량과 같은 자료를 이용해서 발견한 주가 변화의 추세를 가지고 미래의 주가를 추정하는 방법이고, 기본적 분석은 기업의 재무요인이나 경제요인 등을 고려해 추정한 내재가치와 현재 기업의 주가를 비교하여 투자를 결정하는 방법이다. 위의 두 방법은 모두 뚜렷한 한계를 갖고 있는데, 먼저 기술적 분석은 전제 조건인 과거 주가추세나 패턴의 반복을 장담할 수 없다는 점과 동일한 과거 자료로도 사람마다 서로 다른 해석이 가능하다는 점, 그리고 시장변화의 근본적인 원인을 분석할 수 없다는 점이다. 반대로 기본적 분석의 경우에는 재무제표에 근거한 주식가격 평가는 불충분하다는 점, 기업마다 회계처리 방법이 다르기 때문에 기업간 비교평가나 기업을 장기간에 걸쳐 비교할 수 없다는 점, 그리고 주가는 기업의 내재가치뿐 아니라 계량화 할 수 없는 심리적 요인들에 의해서도 결정된다는 점 등이 한계로 여겨진다. 때문에 현명한 투자자들은 이 두 방법을 동시에 고려함으로써 각각의 한계를 최소화 시킨다. 예를 들면 기본적 분석을 통해 투자할 종목을 고르고 기본적 분석을 통해 매매시점을 파악하는 식이다.

본 논문 역시 기술적 분석과 기본적 분석을 적절히 활용하였다. 먼저 기업들의 분기별 재무제표를 이용한 군집분석(clustering)을 통해 '언제' 기업의 재무현황이 바뀌는지, 바뀐다면 '어떻게'(단기부채가 많아지는지, 수익이나 규모가 커지는지 등) 바뀌는지를 분석했다. 이 결과를 바탕으로 특정 기업이 재무현황이 바뀔 때 주식을 사고 일정기간 후에 파는 투자전략을 취함으로써, KOSPI200 지수 대비 어느 정도의 초과수익률을 얻을 수 있는지 살펴보았다. 최종적으로 과거 초과수익률 분포로부터 '어느 분기에 어떤 식으로 기업의 재무현황이 바뀌었을 때' 최고의 결과를 예상할 수 있는지를 분석해 분기별 최적의 투자전략을 제시한다.

본 논문은 2003년 4분기부터 2011년 1분기까지 KOSPI200에

속했던 기업들을 대상으로 분기를 단위로 연구를 진행하였다. 또한 횡단면적(Cross-sectional) 군집분석을 하였기 때문에 같은 기업일지라도 시점이 다르면 다른 개체로 취급하였다. 즉 '2003년 2분기 삼성전자'와 '2003년 3분기 삼성전자'는 다른 개체이다. 군집분석의 변수로는 기업의 재무현황을 파악함에 있어서 중요하다고 생각되는 11개의 주요 재무인자와 재무비율을 사용하였다. 그 결과 총 4659개의 개체를 군집분석을 통해 5개의 군집(cluster)으로 나눌 수 있었다.

먼저 이렇게 나누어진 군집별 특징과 대표기업을 알아보았다. 또한 횡단면적 군집분석을 하였기 때문에 특정기업의 군집간 이동을 볼 수 있었다. 예를 들면, '2003년 3분기 삼성전자'가 군집1에 속해있고 '2003년 4분기 삼성전자'가 군집3에 속해있다면 '2003년 4분기에 삼성전자가 군집1에서 군집3으로 이동하였다'고 정의할 수 있다. 전체 연구기간 동안 어떤 군집에서 어떤 군집으로의 이동이 발생했고, 또 그 이동은 몇 번이나 발생했는지를 살펴보았다. 기업의 입장에서 군집간 이동이 어떠한 의미를 갖는지도 알아보았고, 이런 군집간 이동이 특정기간에만 발생할 수 있을 것으로 판단하여 연도별, 분기별 이동횟수를 살펴보고 분석하였다.

특정기업이 군집간 이동을 했을 경우에는 기업 내 재무현황에 큰 변화가 발생했다고 볼 수 있으므로 주가가 유의미한 변화를 보일 것으로 예상했다. 따라서 재무제표 발표일에 기업의 주식을 매수하고 일정기간 후에 매도함으로써, 주가 변화에 따른 매매차익을 얻을 수 있다. 이때 단순히 수익률을 구할 경우에는 각 사례(case)마다 그 당시 시장의 호황과 불황에 영향을 받을 수 있으므로 KOSPI200지수 대비 초과수익률을 구함으로써 시장의 호황, 불황에 대한 영향을 최소화 하였다. 이렇게 분기별, 기간별로 군집간 이동 시 얻을 수 있는 초과수익률을 구했다. 일정수준 이상의 초과수익률을 보이는 투자안의 경우 초과수익률의 분포를 구해보았다. 이때 분포에서 특정 초과수익률이 이상치(outlier)로 판단될 경우 제거하고 다시 평균과 표준편차를 구했다. 이렇게 구한 평균과 표준편차를 이용해서 분기별로 초과수익률-위험 산포도를

산출하여 투자자 개인의 성향에 따른 투자전략을 제안하였다.

최근 재무제표에 대한 관심이 늘어나고 있고, 이를 데이터마이닝을 통해 분석하려는 시도도 늘어나고 있다. 하지만 이런 시도들이 대부분 기업채권평가나 분식회계를 잡아내는데 집중되어 있고[1,2,3]투자전략을 제시하려는 시도는 많지 않았다. 또한 투자전략을 제시하는 연구라 할지라도 대부분 데이터마이닝 기법으로 분류분석(classification analysis)을 사용하였고[4], 군집분석을 사용한 연구는 거의 없었다. 따라서 본 논문은 재무제표를 이용한 군집분석을 통해 분기별 투자전략을 제시하는 논문으로서 의의를 갖는다.

이상에 대한 연구를 진행한 본 논문의 구성은 다음과 같다. 2장에서는 연구를 위해서 사용한 데이터에 대해서 자세히 살펴볼 것이고, 3장에서는 위 데이터를 분석하기 위해 이용한 연구기법(fuzzy c-means)에 대해서 알아볼 것이다. 4장에서는 실행한 군집분석의 결과를 보여주고 이에 대해서 자세히 분석해 볼 것이다. 5장에서는 수익률 분석에 따른 분기별 최적의 투자전략을 제시하고, 마지막으로 6장에선 지금까지의 결과를 요약하고, 개선할 점과 함께 앞으로의 발전 방향을 제시한다.

제 2 장 데이터

2.1. 분기별 재무제표

본 논문에서는 기본적 분석(fundamental analysis)의 관점에서 주식종목을 분류, 분석하기 위해 기업의 분기별 재무제표를 이용하였다. 때문에 2003년 4분기부터 2011년 1분기까지 KOSPI200에 속해있던 기업들의 분기별 재무제표 데이터를 구했다. 규모, 수익성, 효율성, 레버리지, 그리고 유동성이라는 5가지의 지표가 기업의 재무현황을 파악하기에 중요하다고 판단하여, 분기별 재무제표로부터 이 지표들을 반영할 수 있는 11개의 재무인자와 재무비율을 구했다[5].

지표	재무인자(비율)
규모	시가총액 총자산 매출액
수익성	총자산이익률(=순이익/총자산) 자기자본이익률(=순이익/자기자본)
효율성	총자산회전율(=매출액/총자산) 재고자산회전율(=매출원가/재고자산) 매출액순이익률(=순이익/매출액)
레버리지	장기부채-자기자본비율(장기부채/자기자본) 총부채비율(총부채/총자산)
유동성	유동비율(=유동자산/유동부채)

[표1] 지표에 따른 재무인자(비율) 목록

규모를 반영하는 재무인자로는 전 상장 주식을 시가로 평가한 금액을 나타내는 시가총액과 자본과 부채를 합친 금액으로 기업의 전체 재산을 나타내는 총 자산, 그리고 기업의 매출액을 선택했다.

기업의 수익성을 파악하기 위해서 2개의 재무비율을 선택했는데 총 자산이익률과 자기자본이익률이다. 총 자산이익률은 ROA(Return On

Assets)라고도 불리며 수익성을 살펴보는 대표적인 지표로 '총자산에서 당기순이익이 차지하는 비중'이다. ROA가 높다는 것은 자산에 비해 이익이 많다는 의미로서, 자산을 기준으로 볼 때 수익성이 높다고 할 수 있다. 자기자본이익률은 경영자가 기업에 투자된 자본을 이용하여 이익을 어느 정도 올리고 있는가를 나타내는 기업의 이익창출능력이다. 따라서 자기자본이익률이 높은 기업은 자본의 효율적 사용으로 이익을 많이 내는 기업으로 주가도 높게 형성되는 경향이 있어 투자지표로 사용된다.

기업이 얼마나 자산을 효율적으로 사용하는지를 살펴보기 위해서 총자산회전율, 재고자산회전율, 매출액순이익률을 산출했다. 총자산회전율은 기업의 총자산이 1년에 몇 번이나 회전하였는가를 나타내므로 기업이 소유하고 있는 자산들을 얼마나 효과적으로 이용하고 있는가를 보여준다. 때문에 총자산 회전율이 높으면 유동자산, 고정자산 등이 효율적으로 이용되고 있다는 것을 뜻하고, 반대로 낮으면 과잉투자와 같은 비효율적인 투자를 하고 있다는 것을 의미한다. 재고자산회전율은 재고자산 보유수준의 과부족을 판단하는데 가장 적합한 지표로서 기업이 얼마나 효율적으로 재고관리를 하는지 보여준다. 일정한 표준 비율은 없으나 일반적으로 이 비율이 높으면 자본수익률이 높아지고, 매입채무가 감소하여 재고손실을 막아 기업측에 유리하다. 그러나 지나치게 높을 경우 재고 부족으로 계속적인 생산 및 판매활동에 지장을 초래할 수도 있다. 마지막으로 매출액순이익률은 매출액과 순이익과의 관계를 표시해주는 비율로서 기업활동의 총체적인 능률을 판단하는 지표이다. 보통 비율이 높을수록 좋다.

레버리지는 '지렛대'라는 의미로 금융계에선 차입을 뜻한다. 본 연구에서는 이를 반영하기 위해 장기부채-자기자본 비율과 총부채비율을 살펴보았다. 장기부채-자기자본비율은 기업의 장기지급능력을 파악하기 좋다. 이 비율은 단순히 높거나 낮다고 해서 좋거나 나쁘다고 단정지을 수 없다. 오히려 기업의 재무전략에 대한 태도를 보여준다고 할 수 있다. 예를 들어 철광업이나 통신사처럼 안정적인 사업모델을 가진 기업은 부채를 통한 자본조달을 선호할 것이고, IT제품과 같이 1년 이내 수명주기를 제품을 보유한 기업의 경우, 낮은 부채비율을 유지하려 할 것이다. 총부채비율은 총부채를 총자산으로 나눈 비율로 장기부채-자기자본 비율과 의미가 비슷하지만, 단기부채까지 고려했다는 점에서 다르다. 이를 이용해 기업이 지닌 부채의 성격을 파악하기에 용이하다.

마지막으로 기업의 유동성 판단을 위해 사용된 유동비율은 기업이 보유하는

지급능력, 또는 그 신용능력의 판단을 위해 쓰이는 것으로 신용분석적 관점에서 가장 중요하다. 이 비율이 클수록 기업의 재무유동성이 크다고 할 수 있다. 보통 200% 이상으로 유지되는 것이 이상적이다.

위에서 설명한 11개의 값을 갖는 특정 기업의 특정 분기에 발표된 재무제표가 하나의 개체가 된다. 즉 '2005년 1분기 현대차'와 '2005년 2분기 현대차'는 서로 다른 개체이고 각각은 시가총액, 총부채비율, 유동비율 등 11개의 값을 갖는다. 데이터 전처리 후에 총 259기업과 4659개의 개체를 얻을 수 있었다.

이때 군집분석의 변수로 사용되는 재무인자와 재무비율들의 단위가 서로 다르기 때문에 그대로 비교하는 것은 무의미하다. 따라서 각각의 11개 변수의 평균과 표준편차를 구한 후, 개체가 갖는 값을 표준화($Z = \frac{(X_i - \mu)}{\sigma}$) 시켰다.

해당 재무제표 자료는 금융자료 제공 사이트인 에프앤가이드(www.fnguide.com)로부터 얻었다.

2.2. 수정주가

본 논문의 투자전략은 재무제표 발표일에 특정기업이 군집간 이동이 발생한 경우, 그 기업의 주식을 매수하고 일정기간 후에 매도하여 매매차익을 얻는 방식이다. 때문에 2003년 4분기부터 2011년 1분기까지 KOSPI200에 속해있던 기업들의 일일 수정주가데이터를 구했다. 이때 투자전략 방식의 특성상 매수시점과 매도시점의 주가만 필요하므로 전처리를 통해 그 외 시점의 수정주가는 제거하였다.

수정주가란 주가에 연속성을 부여하기 위해 일정한 수정을 가하는 것이다. 주가는 유무상증자, 배당, 액면분할 등이 생길 때 연속성을 잃고 단층을 보이게 되는데 이럴 경우 이전주가와 현재주가를 그대로 비교하기에는 어려움이 있다. 때문에 본 논문에서도 여러시점의 주가를 비교하기 위해서 수정주가를 사용했다.

재무제표 발표일은 금융감독원 전자공시시스템인 dart.fss.or.kr에서 가져왔다.

2.3. KOSPI200 지수

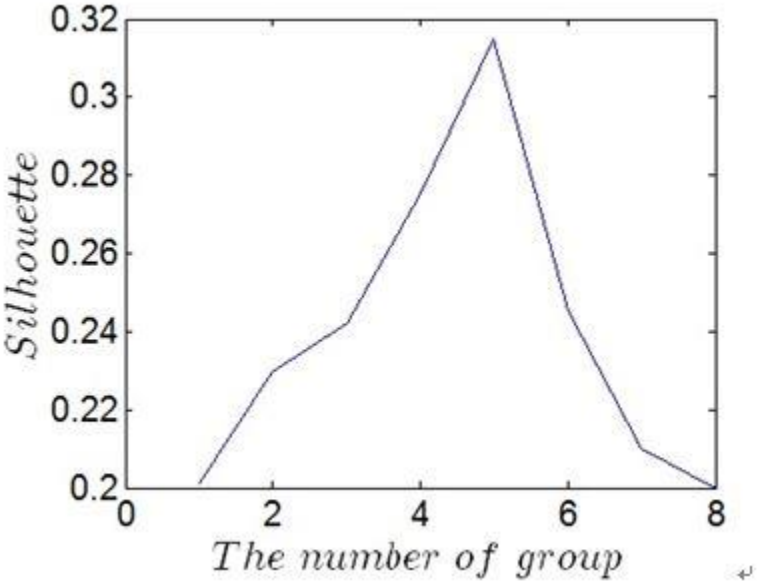
본 논문에서 제시하는 투자전략의 특성상, 그 당시 시장이 호황과 불황에 영향을 받을 수밖에 없다. 때문에 이에 대한 영향을 최소화하기 위해서 단순히 매매차익으로 인한 수익률을 고려하는 대신에 KOSPI200 지수 대비 초과수익률을 구해 투자전략을 비교하는 방식을 택했다. 이를 위해서 2003년 4분기부터 2011년도 1분기까지 KOSPI200 지수를 구했고, 수정주가와 마찬가지로 매수시점과 매도시점 외의 자료는 전처리를 통해 제거하였다.

제 3 장. 연구기법

본 연구에서는 기본적 분석의 관점에서 주식종목을 분류, 분석하기 위해 횡단면적 군집분석을 사용하였다. 군집분석이란 유사성의 개념을 바탕으로 비슷한 성질을 지닌 개체들을 모아 몇 개의 군집으로 분류하고 이를 분석하는 방법이다. 군집분석 방법으로는 FCM 군집분석(Fuzzy C-Means clustering)[6]을 사용하였다. FCM은 각각의 개체를 각 군집에 속하는 정도를 측정하여 분류하는 알고리즘이다. 개체의 유사성은 변수들의 유클리디언 거리 값을 이용하고, 특이점은 대부분의 군집분석은 개체가 하나의 군집에 속해 있다고 가정하는데 FCM은 각 개체가 여러 군집에 비율로 속해있다고 가정한다. 본 연구에서는 가장 큰 비율로 속해있는 군집에 개체가 속해있다고 정의하였다.

FCM을 수행하기 위해서는 입력자료로 데이터, 군집의 개수, 반복 횟수가 필요하다. 이때 데이터는 구했고 반복 횟수는 500으로 임의로 정하였다. 마지막으로 군집분석 시에는 군집의 개수가 결과에 가장 큰 영향을 미치기 때문에 임의로 개수를 정하지 않고, silhouette coefficient를 이용하였다. Silhouette coefficient는 군집분석 성과평가에 널리 이용되는 방법으로, -1 에서 1사이의 값을 갖는데 1에 가까울수록 군집분석이 잘 행해졌다는 의미이다. 군집의 개수를 1~8까지 변화시키면서 값을 측정해 보았다.

그 결과 군집의 개수가 5개 일 때 silhouette coefficient 값이 가장 높은 것을 볼 수 있었고, 따라서 군집의 개수는 5개로 정하였다. 본 연구에서 FCM은 R package의 e1071 library를 사용하여 구현했다. 해당 package는 http:// cran.r-project.org/ web/ packages/ e1071/ e1071.pdf 에서 찾아볼 수 있다.

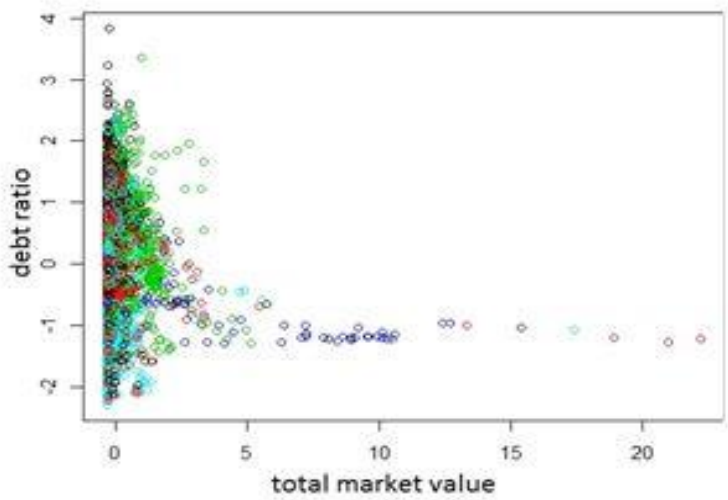


[그림1] 군집의 개수에 따른 silhouette coefficient 값

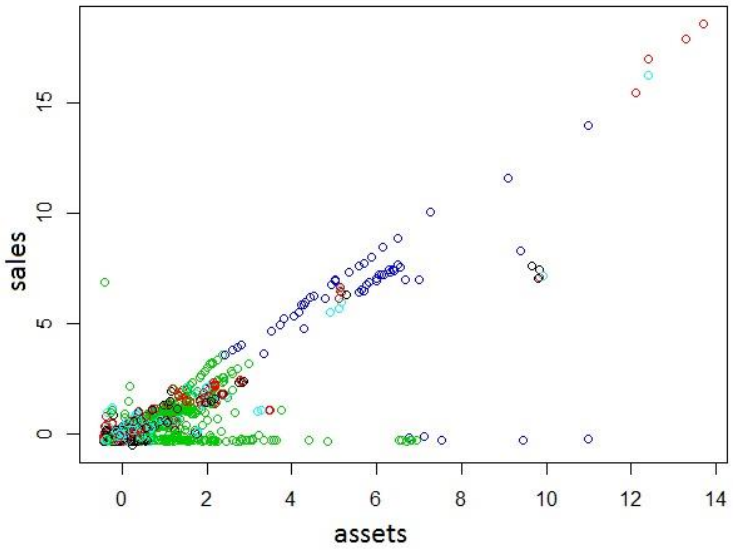
제 4 장. 군집화(clustering) 결과 및 분석

4.1. 결과

총 4659개의 개체를 표준화시킨 11개의 주요 재무인자, 재무비율을 변수로 FCM을 통해 군집분석을 실시하였다. 그 결과 각각의 개체에 1~5까지의 군집번호가 부여되었다. 간단히 말하자면, 11차원에서 4659개의 개체를 5개의 군집으로 나눈 것이다. 이때 11차원을 시각화 하기 어렵기 때문에 군집화에 큰 영향을 미친 것으로 판단되는 주요 변수를 축으로 2차원 산포도를 그려보았다.

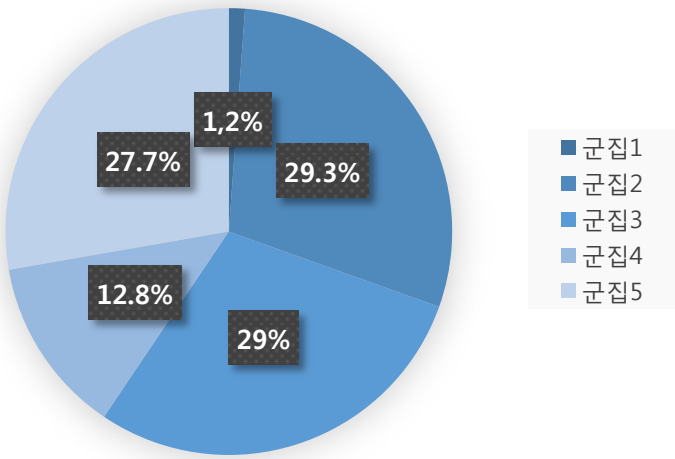


[그림2] 표준화된 시가총액과 부채비율을 축으로 한 산포도



[그림3] 표준화된 자산과 매출액을 축으로 한 산포도

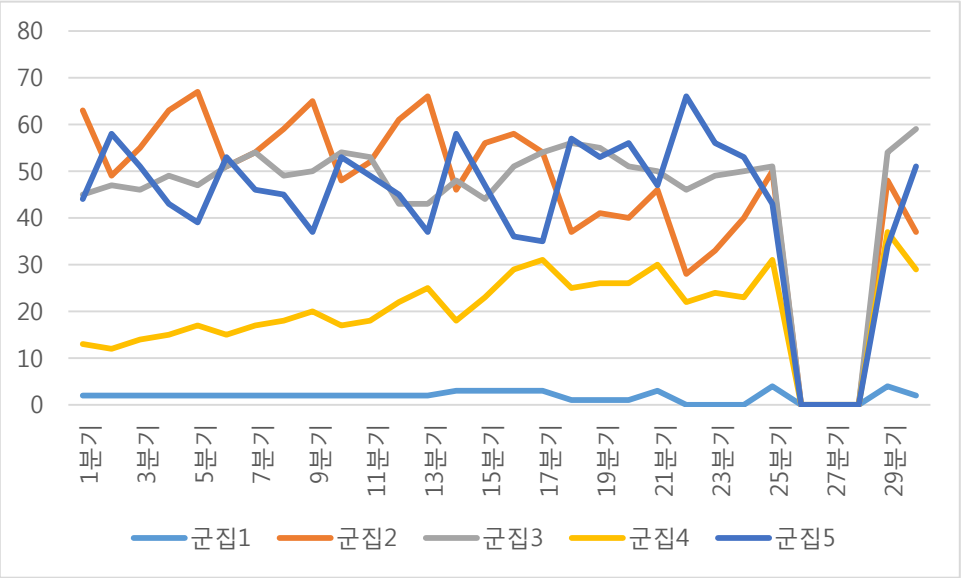
좀 더 자세히 살펴보면, 2003년 4분기부터 2011년도 1분기까지 전체 연구기간 동안 군집1에는 54개, 군집2에는 1365개, 군집3에는 1349개, 군집4에는 597개, 군집5에는 1292개의 개체가 속해있었다.



[그림4] 군집별 개체 수의 비율

군집화한 결과는 위 그림과 같다. 전체 연구기간 동안 군집1에

속한 개체는 1.2%로 매우 적었고, 군집2,3,5는 29.3%, 29% 27.7%로 대다수의 개체들이 이 세 개의 군집에 속해있었음을 알 수 있었다. 군집4의 경우 12.8%가 속해 있었다. 위 결과를 이용해 우리는 매 분기마다 평균적으로 군집1에는 2개, 군집2,3,5에는 58개, 군집4에는 27개의 기업이 속해있다고 예상할 수 있다. 하지만 특정 군집에 특정 시기의 개체들만 몰려 있어서 위와 같은 결과가 나올 경우, 예를 들어 2008년 9월에 발생한 금융위기 때문에 그 이후의 개체들이 좋지 않은 재무상황을 나타내는 특정 군집에 몰려있는 거라면, 단순히 [그림4]에 나온 비율대로 분기마다 군집별 개체수가 일정하다고 보장할 수 없다. 그렇게 되면 과거의 추세가 현재에도 계속된다고 가정하에 투자전략을 제시하는 본 논문의 의미가 퇴색된다. 때문에 군집별 개체 수를 전체 연구기간 동안 크게 보는 것이 아니라 매 분기마다 각 군집에 속한 개체 수가 일정하거나 혹은 추세를 갖는지 봐야 한다. 그리고 그 결과는 아래와 같다.



[그림5] 연구기간 동안 분기마다 군집에 속한 개체 수의 변화

위 그림을 보면 먼저 26분기~28분기(2010년 1분기~2010년 3분기)동안 데이터가 없는 것을 볼 수 있는데, 이는 데이터를 수집하고 전처리하는 과정에서 이 기간 동안의 데이터가 빠진 부분이 많아 전체를 삭제하였기 때문이다. 실제 데이터를 이용하지 못한 점이 아쉽지만 25분기의 결과와 29분기의 결과가 큰 차이를 보이지 않는 점을 보아 본 논문의 결과에 큰 영향을 미치지 않을 것으로 판단하여 연구를 계속 진행하였다. [그림5]를 살펴보면 KOSPI200내에서 평균적으로 2개의 기업이 군집1에, 40~60개의 기업이 군집2에, 45~52개의 기업이 군집3에, 15~30개의 기업이 군집4에, 그리고 37~55개의 기업이 군집5에 속해있다. 여기서 특이점은 군집4에 속하는 기업의 수가 점차 늘어난다는 점, 그리고 군집2에 속하는 기업의 수가 점차 줄어든다는 점, 마지막으로 금융위기에 영향을 받은 2009년 1분기~2009년 3분기(22분기~24분기)동안 군집1에 속하는 기업이 없다는 점이다. 뒤에서 자세히 알아볼 것이지만 군집1이 보통 삼성전자, 현대차, posco 등 다른 기업들에 비해 훨씬 규모가 큰 기업들이 속해있는 것을 봤을 때, 금융위기에 따른 영향으로 위와 같은 결과가 나타났다고 볼 수 있다.

4.2. 분석

4.2.1. 군집별 특징

위에서 얻은 결과를 이용해 군집별 특징을 알아보았다. 먼저 군집별로 군집분석에 사용된 변수인 재무인자, 재무비율의 평균값은 아래와 같다[표2]. 이때 각 재무인자, 재무비율마다 단위가 다르기 때문에 적절한 비교를 위해 표준화된 평균 재무인자, 재무비율의 값을 구했다[표3]

군집	규모			수익성		효율성 지표			레버리지 지표		유동성지표
	시가총액 (천억)	자산(억)	매출액 (억)	총자산이익률	자기자본 이익률	총자산회전율	재고자산 회전율	매출액 순이익률	장기부채 자기자본비율	총부채비율	유동비율
1	531	580	362	0.06	0.08	0.51	67.9	0.12	17	0.27	1.34
2	8.3	10.9	5.66	0.05	0.08	0.78	7.58	0.07	0.55	0.43	1.45
3	10.9	10.8	6.33	0.04	0.06	0.46	16	0.12	0.53	0.21	4.17
4	90.1	137.4	45.4	0.04	0.09	0.72	22.08	-0.07	21.84	0.56	1.14
5	11.1	22.9	6.16	0.00	0.00	0.47	6.82	-0.19	3.84	0.64	0.93

[표2] 군집별 재무인자, 재무비율의 평균값

군집	규모			수익성		효율성 지표			레버리지 지표		유동성지표
	시가총액	자산	매출액	총자산이익률	자기자본 이익률	총자산회전율	재고자산 회전율	매출액 순이익률	장기부채 자기자본비율	총부채비율	유동비율
1	**6.6317**	**6.6168**	**7.5145**	0.3085	0.1250	-0.0824	0.6173	0.0425	0.2769	-0.8610	-0.0933
2	-0.2376	-0.3177	-0.2071	0.1756	0.1089	0.1828	-0.0533	0.0256	-0.0839	-0.0460	-0.0799
3	-0.2034	-0.3194	-0.1926	0.1173	0.0140	-0.1242	0.0407	0.0415	-0.0843	**-1.2066**	**0.2794**
4	**0.8373**	**1.2227**	**0.6545**	0.0375	0.1481	0.1243	0.1080	-0.021	0.3833	0.6114	-0.1196
5	-0.2004	-0.1718	-0.1962	**-0.3385**	**-0.2035**	-0.1177	-0.0618	-0.063	-0.0118	**1.0620**	-0.1480

[표3]군집별 표준화된 재무인자 재무비율의 평균값

(중요하게 살펴볼만한 값에 밑줄을 긋고 굵은 글씨체로 바꿈)

군집분석의 변수로 사용된 재무인자와 재무비율들의 단위가 서로 다르기 때문에 단순히 평균값이 아닌, 표준화된 평균값을 살펴봐야 한다. 그러므로 지금부터 언급하는 변수의 값은 모두 표준화된 변수값을 의미한다.

표를 봤을 때, 가장 먼저 눈에 띄는 것은 군집1에서 규모를 반영하는 시가총액, 자산, 매출액의 값이 나머지 군집에 비해서 월등하게 큰 것을 볼 수 있다. 즉 군집1에는 KOSPI200내 기업 중에서도 규모가 상당히 큰 기업들이 속할 것으로 생각할 수 있고, 실제 삼성전자, 현대차, POSCO, 한국전력이 여기에 속한다.

다음으로 특이한 점은 군집4에서 규모를 반영하는 재무인자 값들이다. 비록 군집1 보다는 작지만 나머지 군집2, 군집3, 그리고 군집5 보다는 유의할만하게 큰 것을 볼 수 있다. 즉 군집1에 들어갈 만큼 큰 규모를 갖진 않지만 그래도 KOSPI200내에서 큰 규모를 지니는 기업들이 군집4에 속해있을 것으로 보이고, 대표기업으로는 기아차, SK하이닉스, 현대건설, LG전자 등이 있다.

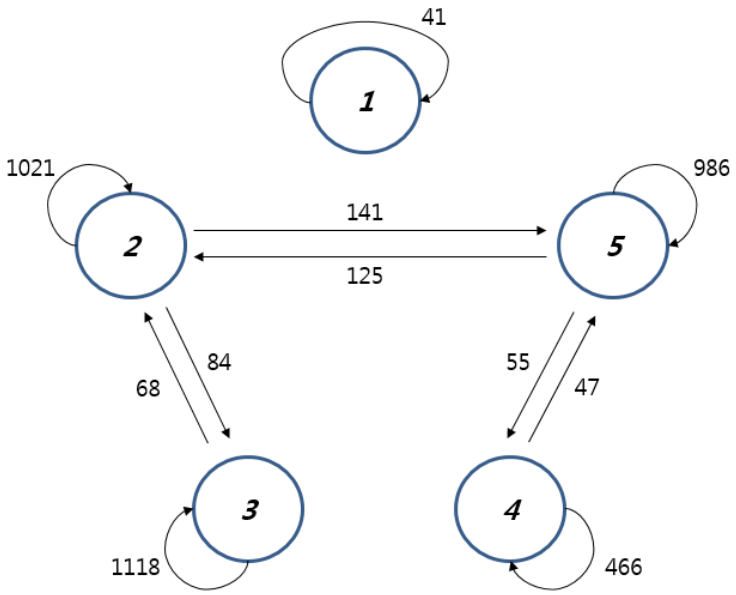
군집2, 군집3, 군집5는 사실상 규모면 에서는 거의 차이가 없다. 눈에 띄는 부분은 군집3의 총부채비율이 나머지 군집들에 비해서 훨씬 낮고, 유동비율이 다른 군집들보다 훨씬 높다는 점이다. 이를 봤을 때, 군집3에 속해있는 기업들의 재무현황은 부채가 거의 없고 유동성이 좋으므로 매우 안정적인 기업군이라고 볼 수 있다. 대표기업은 한국유리, 고려제강, KCC, 남양유업 등이다.

마지막으로 군집5의 경우 수익성을 나타내는 재무비율들이 다른 군집들과 다르게 유일하게 음수를 갖고, 총부채비율 역시 훨씬 크다. 이를 보았을 때, KOSPI200내에서 상대적으로 기업의 재무상황뿐 아니라 기업운영, 그리고 수익성이 좋지 않다는 걸 알 수 있다. 한화, 쌍용차, 금호산업, 오리온 등이 여기 속해있다.

4.2.2. 군집간 이동

본 논문에서는 횡단면적 군집분석을 사용했기 때문에 '군집간

이동’과 ‘군집 내 머무름’이라는 특이한 개념이 존재한다. 예를 들자면 ‘2005년 2분기 삼성전자’라는 개체가 군집3에 속해있는데 ‘2005년 3분기 삼성전자’라는 개체가 군집4에 속해있는 경우, ‘삼성전자가 2005년 3분기에 군집3에서 군집4로 이동했다’고 정의한다. 반대로 예를 들어 위의 상황에서 2005년 2분기는 동일하고 ‘2005년 3분기 삼성전자’가 군집3에 속해있다면 ‘삼성전자는 2005년 3분기에 군집 내에서 머물렀다’라고 정의한다. 기업이 KOSPI200에서 퇴출당하거나 아니면 KOSPI200으로 처음 들어왔을 때에는 비교할 수 있는 군집이 없으므로 당연히 ‘군집간 이동’과 ‘군집 내 머무름’이 일어나지 않는다. 이때 2003년 4분기부터 2011년 1분기까지의 발생한 ‘군집간 이동’과 ‘군집 내 머무름’을 나타낸 도표는 아래와 같다. 연구기간이 총 27분기인데 발생횟수가 27번도 되지 않을 경우, 분기마다 1번도 발생할 것으로 기대하지 못한다는 것이므로 그런 이동과 머무름은 의미가 없다고 판단하여 제외시켰다.



[그림6] 연구기간 동안의 군집간 이동과 군집 내 머무름 개요

위의 [그림6]을 보았을 때, 특이한 점은 군집1에서는 다른 군집으로의 이동이 없고, 각각 군집2와 군집3 사이, 군집2와 군집5 사이, 그리고 군집4와 군집5 사이에서만 군집간 이동이 발생했다는 점이다. 군집2와 군집4 사이, 군집3과 군집5 사이, 그리고 군집3과 군집4 사이에서는 군집간 이동이 없는 것이다.

본 논문에서 투자전략을 제시하는 시점은 재무제표 발표일에 특정기업의 군집간 이동이 발생했을 때이다. 군집간 이동이 발생했다는 것은 기업 내 재무상황이 크게 변했다고 볼 수 있으므로 주가가 변할 것으로 예상되기 때문이다. 따라서 군집 내 머무름보다 군집간 이동에 초점을 맞추어 연구를 진행하였다.

그렇다면 과연 특정 군집간 이동의 발생이 기업의 재무상황이 구체적으로 어떻게 변하는 것을 의미하는지 살펴보았다. 이때 위의 [그림6]을 보면 군집간 이동은 대부분 군집2와 군집3 사이, 군집2와 군집5 사이, 그리고 군집4와 군집5 사이에서만 발생한다. 따라서 지금부터는 이 세 가지 경우에 대해서 알아보자.

먼저 특정 기업이 군집2에서 군집3으로 이동했다고 가정하자. 뒤에 나오는 [표4]를 보면, 자산은 거의 비슷한 데 비해서 총부채비율이 크게 낮아진 것으로 보아 기업의 부채가 상당히 낮아진 것을 알 수 있다. 이때 유동비율이 상승한 사실에서 기업이 장기부채가 아닌 단기부채를 많이 상환하였음을 알 수 있다. 즉, 군집2에서 군집3으로의 이동은 기업의 단기부채가 줄어 재무상태가 안정적으로 변하였음을 의미한다.

다음으로 기업이 군집2에서 군집5로 이동하는 경우를 살펴보자. [표5]를 보면 수익성을 의미하는 재무비율들이 하락하였고, 장기부채자기자본 비율은 크게 차이가 나지 않는데 비해서 총부채비율은 현격히 커진 것을 봤을 때, 기업이 가치하락과 수익성의 악화로 인해 단기부채에 의존하려 하는 상황임을 추측할 수 있다.

마지막으로 군집4에서 군집5로 기업이 이동한다고 가정해보자. [표6]을 보면 이 경우 거의 모든 값이 달라진 것을 볼 수 있다. 기업의 규모는 훨씬 작아졌고, 수익성 역시 크게 하락했다. 또

효율성도 떨어졌고, 기업의 부채상환도 나빠졌다. 즉 모든 지표들을 고려해 봤을 때, 이러한 군집간 이동은 기업의 상황이 매우 좋지 않음을 의미한다.

군집	규모			수익성		효율성 지표			레버리지 지표		유동성지표
	시가총액	자산	매출액	총자산이익률	자기자본 이익률	총자산회전율	재고자산 회전율	매출액 순이익률	장기부채 자기자본비율	총부채비율	유동비율
2	-0.2376	-0.3177	-0.2071	0.1756	0.1089	0.1828	-0.0533	0.0256	-0.0839	**-0.0460**	**-0.0799**
3	-0.2034	-0.3194	-0.1926	0.1173	0.0140	-0.1242	0.0407	0.0415	-0.0843	**-1.2066**	**0.2794**

[표4] 군집2와 군집3 비교

군집	규모			수익성		효율성 지표			레버리지 지표		유동성지표
	시가총액	자산	매출액	총자산이익률	자기자본 이익률	총자산회전율	재고자산 회전율	매출액 순이익률	장기부채 자기자본비율	총부채비율	유동비율
2	-0.2376	-0.3177	-0.2071	**0.1756**	**0.1089**	0.1828	-0.0533	0.0256	-0.0839	**-0.0460**	-0.0799
5	-0.2004	-0.1718	-0.1962	**-0.3385**	**-0.2035**	-0.1177	-0.0618	-0.063	-0.0118	**1.0620**	-0.1480

[표5] 군집2와 군집5 비교

군집	규모			수익성		효율성 지표			레버리지 지표		유동성지표
	시가총액	자산	매출액	총자산이익률	자기자본 이익률	총자산회전율	재고자산 회전율	매출액 순이익률	장기부채 자기자본비율	총부채비율	유동비율
4	**0.8373**	**1.2227**	**0.6545**	**0.0375**	**0.1481**	**0.1243**	**0.1080**	-0.021	**0.3833**	**0.6114**	-0.1196
5	**-0.2004**	**-0.1718**	**-0.1962**	**-0.3385**	**-0.2035**	**-0.1177**	**-0.0618**	-0.063	**-0.0118**	**1.0620**	-0.1480

[표6] 군집4와 군집5 비교

기업에는 계절성도 존재하고, 또 연초에는 새로운 프로젝트를 시작하기도 하므로 이에 따른 재무현황에 변화도 있을 것이다. 즉 위에서 살펴본 군집간 이동이 1분기, 2분기, 3분기, 그리고 4분기에 동일한 횟수로 발생하진 않는다. 따라서 분기별로 군집간 이동횟수를 구했고, 이는 아래와 같다. 열은 현재 재무제표 발표일의 군집번호이고, 행이 이전 재무제표 발표일의 군집번호이다. 예를 들어 밑의 표의 (3,2)의 값이 14인 것의 의미는 '1분기에 군집3에서 군집2로 이동한 개체가 14개' 라는 뜻이다. 또한 (i,i)의 값은 기업의 군집이 변하지 않았음을 의미하므로 군집간 머무름을 뜻한다.

1분기	1	2	3	4	5
1	11	0	0	6	0
2	0	267	31	1	104
3	0	14	323	1	0
4	1	4	5	124	36
5	0	10	0	5	253

[표7] 1분기 군집간 이동 및 머무름 횟수

2분기	1	2	3	4	5
1	10	0	0	0	0
2	0	237	15	2	5
3	0	16	284	1	0
4	0	2	1	105	1
5	0	35	0	14	294

[표8] 2분기 군집간 이동 및 머무름 횟수

3분기	1	2	3	4	5
1	10	0	0	6	0
2	0	235	17	2	14
3	0	21	241	2	0
4	0	2	1	113	6
5	0	36	2	14	228

[표9] 3분기 군집간 이동횟수

4분기	1	2	3	4	5
1	10	0	0	6	0
2	0	282	21	1	17
3	0	17	270	3	1
4	5	0	0	124	4
5	0	44	0	22	211

[표10] 4분기 군집간 이동횟수

위 결과에서 특이한 점은 1분기에서 군집2에서 군집5로의 이동과, 군집4에서 군집5로의 이동이 다른 분기에 비해 훨씬 많이 일어났다는 점이다. 1분기에 군집2에서 군집5로의 이동횟수는 104번인데 비해 나머지 분기에서는 5, 15, 17번 일어났다. 또한 군집4에서 군집5로의 이동은 1분기에 36번 발생했는데 나머지 분기에서는 1번, 6번, 4번 발생했다.

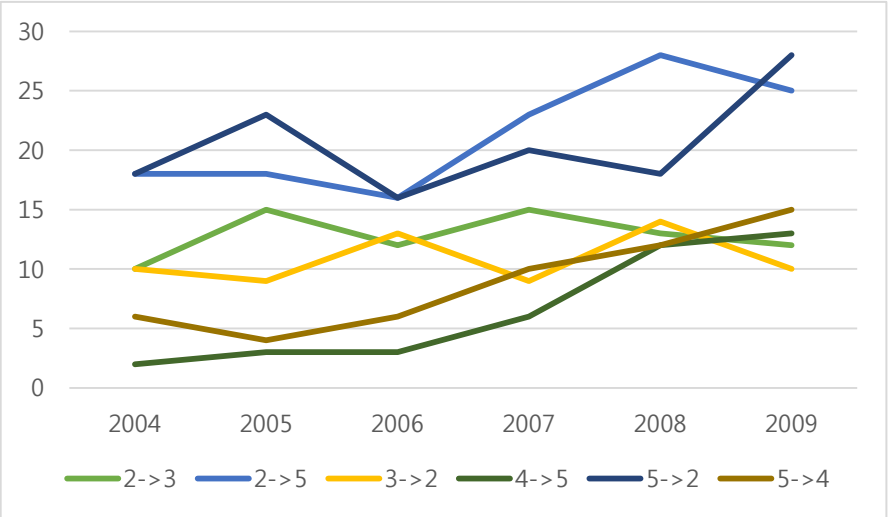
앞에서 봤듯이 군집2에서 군집5로의 이동은 기업의 가치하락과 단기부채의 상승을 의미한다. 보통 기업의 새로운 프로젝트는 연초에 시작된다. 이 경우 프로젝트 초기에는 아직 수익이 창출되지 않을 것이고 투자를 위해서 단기부채에 의존할 것이므로 위와 같이 1분기에 군집2에서 군집5로의 이동이 집중된 현상은 충분히 설명 가능하다.

군집4에서 군집5로의 이동은 기업의 규모와 수익성에 큰 하락이

있고, 효율성과 부채상황 역시 나빠짐을 의미한다. 한마디로 기업이 큰 위기에 빠졌다고 할 수 있다. 보통 장기부채의 만기는 연말이나 연초로 되어있는데, 이때 기업의 수익이 하락하여 자금부족이 생기는 경우, 장기부채 상환 실패로 기업은 단기부채에 의존하게 된다. 이런 이유로 군집4에서 군집5로의 이동이 다른 분기보다 1분기에 많이 발생한다고 설명할 수 있다.

본 논문의 투자전략은 특정 분기에 재무제표가 발표되었을 때 기업에게 특정 군집간 이동(예를 들면 군집3에서 군집2로의 이동)이 일어나면, 해당 기업의 주식을 매수하고 일정기간 후에 매도하여 그 매매차익을 얻는 방법이다.

위에서는 분기별 군집간 이동횟수를 알아보았는데, 만약 특정 시점에만 이런 군집간 이동이 많이 발생했던 거라면 앞으로도 발생할거라고 장담하기 어렵다. 그럴 경우 본 논문에서 제시하는 투자전략은 아무런 의미를 갖지 못한다. 때문에 특정 시점이 아닌 전체 연구기간 동안 매년 특정 군집간 이동들이 일정수준 이상 발생했어야 투자전략으로써 의미가 있다. 따라서 이를 살펴보았고 결과는 다음과 같다.



[그림7] 연도별 특정 군집간 이동횟수

참고: 2->3은 군집2에서 군집3으로의 이동을 의미

위에서 말했다시피, 2010년의 데이터는 수집, 전처리 과정에서 제거 되었고, 2011년 데이터의 경우 1분기까지만 있어서 연도별로 비교하기 적당하지 않다고 판단하여 2004년부터 2009년까지의 데이터만 분석했다.

2->3, 3->2, 5->2는 특정한 패턴을 보이지 않고, 매년 거의 일정하게 발생하는 것을 볼 수 있다. 하지만 2->5와 4->5는 2008년에 가까워 질수록 이전보다 더 많이 발생한 것을 볼 수 있는데, 이는 금융위기 때문에 재무상황이 나빠지거나 위기에 처한 기업들이 많았음을 보여준다. 또한 5->4 역시 4->5의 추세에 따라서 많이는 아니지만 횟수가 증가했다.

위 그림을 보았을 때, 4->5와 5->4의 경우에는 2008년에 발생한 금융위기 때문에 발생횟수가 늘어난 것으로 보인다. 때문에 투자전략을 수립하기에 적절하지 않아 보이기도 하지만, 이는 2010년 이후의 데이터로 분석을 보충해야 확실해 질것이다. 본 논문에서는 일단 4->5와 5->4 역시 투자전략에 포함시켰다.

제 5 장. 투자전략

5.1. 투자방법

발표되었을 때 군집간 이동이 발생 한다면, 해당 기업의 주식을 매수하고 일정기간 후에 매도하여 매매차익을 얻는 것이다. 그렇다면 지금부터는 과연 어떠한 군집간 이동이 일어났을 때 더 나은 결과를 얻을 수 있는지를 알아보겠다. 즉 이번에 2014년 1분기 재무제표가 발표 되었을 때, 2->3인 기업도 있을 것이고 4->5인 기업도 있을 것이다. 이때 어떤 군집간 이동이 일어난 기업의 주식을 매수하고 매도해야 더 적은 위험을 가지고 더 많은 수익을 얻을 수 있는지 살펴본다.

이를 위해 특정 군집간 이동이 발생했을 경우의 초과수익률을 구해 분석해보았다. 단순한 수익률이 아닌 KOSPI200지수 대비 초과수익률을 구하는 이유는 군집간 이동이 발생한 시기에 따라 그 당시 시장의 호황, 불황에 오는 영향력을 받아 수익률이 더 높아지거나, 낮아질 수 있기 때문이다. 따라서 본 논문에서는 초과수익률을 사용함으로써 시장으로부터의 영향력을 최소화 시켰다.

이전에 비슷한 주제로 진행한 연구에서는 투자전략을 특정기업에 군집간 이동이 발생했을 때, 이 기업의 주식을 재무제표 발표일에 매수하고 다음 분기 재무제표 발표일에 매도하는 방법이었다[7]. 이 연구에서는 단순히 특정 군집간 이동이 발생했을 경우, 위의 투자전략을 사용해서 발생한 초과수익률을 구해서 비교하였다.

이와 다르게 본 논문에서는 분기별 기간별로 초과수익률을 구해서 분기별로 특정군집간 이동 시 투자전략을 비교하였다. 이렇게 분석방법을 바꾼 이유는 크게 세 가지이다.

첫 번째는 <군집화 분석>에서 보았듯이 분기별로 특정 군집간 이동횟수가 같지 않기 때문이다. 2->5, 4->5의 경우 나머지 분기보다 1분기에 훨씬 많이 발생한 것을 볼 수 있었다.

두 번째는 분기별 재무제표 발표일 사이의 시간이 동일하지 않기

때문이다. 각 분기는 3개월이다. 따라서 단순히 생각하면 분기별 재무제표 발표일 사이 기간은 3개월로 생각되지만 실제로는 그렇지 않다. 금융감독원 전자공시시스템 사이트인 dart.fss.or.kr에서 조사한 결과 보통 1분기 재무제표 발표일은 5월10~20일 사이였고, 2분기 재무제표 발표일은 8월10~20일, 3분기 재무제표 발표일은 11월10~20일, 마지막으로 4분기 재무제표 발표일은 3월25~4월5일 이었다. 따라서 1분기와 2분기 재무제표 발표일 사이의 기간은 평균 92일, 2분기와 3분기 재무제표 발표일 사이의 기간은 평균 92일, 3분기와 4분기 재무제표 발표일 사이는 평균 135일, 그리고 4분기과 1분기 재무제표 발표일 사이의 기간은 평균 46일이다. 따라서 분기별로 구분하지 않고 특정 군집간 이동들의 초과 수익률을 비교하는 것은 투자기간이 다른 투자대안들을 비교하는 것과 같다. 때문에 분기별로 나누어 초과수익률을 비교하는 것이 타당하다.

세 번째 이유는 재무제표 발표일에 변한 재무제표 수익성 지표를 보고 유입되는 투기성 투자 때문이다. 투기성 투자의 경우 장기간 기업의 주식을 보유하고 있는 것이 아니라, 일정기간 보유한 후에 일정수준 이득을 보면 바로 주식을 매도하여 주가가 하락하게 된다. 때문에 단순히 다음 재무제표 발표일에 매도하는 것이 아니라 기간별로 예를 들면 10일, 20일 후에 주식을 매도했을 때 발생하는 초과수익률을 구해 가장 높은 수익률이 기대되는 기간 동안 주식을 보유하는 것이 현명하다.

위의 세 가지 이유로 다음 절에서는 분기별 기간별 초과수익률에 대해서 살펴보도록 하겠다.

5.2. 분기별 기간별 초과수익률의 분포

먼저 분기마다 기간별로 특정 군집간 이동이 발생했을 때 얻을 수 있는 KOSPI200대비 초과수익률의 분포(평균과 표준편차)를 분석해보았다. 기간은 주말을 제외하고 주식 매매가 가능한 10일 간격으로 나누어서 초과수익률을 구했다.

From	To	10	20	30	빈도
2	3	−3.19%	−4.26%	−4.70%	**23**
2	4	1.77%	3.96%	−1.22%	1
2	5	**0.73%**	−0.53%	−2.98%	**17**
3	2	−0.09%	2.11%	**3.54%**	**17**
3	4	−3.15%	−10.73%	−8.22%	3
3	5	4.99%	5.28%	10.13%	1
4	1	−1.99%	−2.18%	**1.71%**	**5**
4	5	1.96%	2.56%	0.73%	**4**
5	2	**0.63%**	0.44%	−0.38%	**44**
5	4	**1.94%**	0.47%	−1.10%	**22**

[표11] 4분기 기간별 평균 초과수익률

표의 네 번째 행에서 굵게 표시된 0.73%의 의미는 다음과 같다.
4분기 재무제표가 발표 되었을 때, 지난 3분기에는 군집2에 속했었고, 이번 4분기에는 군집5에 속한 기업들을 찾는다. 이 기업들의 주식을 재무제표 발표일에 매수하고 투자가능일 10일이 지난 시점에 매도하면, 매매차익으로 KOSPI200대비 약 0.73% 높은 수익률을 기대할 수 있다.
연구기간 동안 4분기가 총 8번 있었는데, 그 반(4)보다 빈도가 낮다면 사실상 투자전략으로써 무의미하다고 판단하여 그 이상이 되는 빈도를 굵게 칠하였다. 그 이후에 굵게 칠하여진 사례(case) 중에서 초과수익률이 양수이고, 총 기간 중에 가장 높은 초과수익률을 굵게 표시하였다.
아래에서 1, 2, 3 분기에 관찰한 초과수익률을 보여주는 [표12, 13, 14] 역시 위와 같은 방식으로 해석할 수 있다.

From	To	10	20	30	40	50	60	빈도
1	4	0.18%	1.12%	2.26%	4.36%	**5.27%**	1.89%	**4**
2	3	0.59%	1.70%	2.60%	**3.12%**	1.38%	−1.12%	**24**
2	4	−10.21%	0.75%	−9.09%	3.08%	2.72%	0.91%	1
2	5	−4.63%	−3.67%	−4.93%	−4.78%	−5.02%	−5.76%	**93**
3	2	−4.13%	−3.47%	−5.43%	−2.42%	−4.01%	−3.72%	**11**
4	1	−1.98%	3.01%	−5.13%	1.61%	4.85%	6.27%	1
4	2	−5.58%	−4.89%	−1.14%	−2.42%	2.85%	8.82%	1
4	3	4.68%	−1.65%	−1.34%	−3.66%	−4.51%	−5.37%	3
4	5	1.39%	**1.44%**	0.20%	−1.39%	−1.47%	−1.65%	**27**
5	2	−0.52%	**2.05%**	−1.02%	1.80%	−1.04%	−5.19%	**8**
5	4	−7.81%	−5.35%	−4.07%	−0.32%	−0.25%	−2.85%	3

[표12] 1분기 기간별 평균 초과수익률

From	To	10	20	30	40	50	60	빈도
2	3	3.60%	3.25%	3.10%	3.91%	**4.18%**	3.13%	**15**
2	4	20.57%	22.40%	28.85%	11.08%	19.57%	19.35%	2
2	5	−1.85%	−2.91%	−11.12%	−20.16%	−19.02%	−16.76%	**5**
3	2	**0.46%**	−1.34%	−4.77%	−9.99%	−5.82%	−5.47%	**16**
3	4	−1.98%	0.29%	1.12%	6.44%	5.67%	8.11%	1
4	2	−6.94%	−1.02%	−17.64%	−43.53%	−24.52%	−17.01%	2
4	3	9.75%	0.85%	−13.82%	−41.19%	−14.53%	−11.34%	1
4	5	−6.25%	−6.19%	−8.05%	−16.68%	−26.42%	−26.17%	1
5	2	0.59%	2.80%	2.24%	4.86%	**5.94%**	3.92%	**35**
5	4	−0.71%	−1.02%	−1.43%	−0.30%	**1.67%**	−0.18%	**14**

[표13] 2분기 기간별 평균 초과수익률

From	To	10	20	30	40	50	60	70	80	90	빈도
2	3	5.69%	5.62%	3.10%	8.23%	8.19%	9.92%	9.01%	**10.06%**	9.57%	**17**
2	4	−2.45%	−1.48%	5.57%	10.30%	5.11%	10.23%	13.68%	1.22%	1.70%	2
2	5	−4.41%	−3.53%	2.29%	2.92%	−0.45%	1.03%	5.52%	5.46%	**8.28%**	**14**
3	2	**1.36%**	−2.08%	0.12%	−3.28%	−7.02%	−3.92%	−3.82%	−2.23%	−1.09%	**21**
3	4	0.16%	1.63%	5.28%	−6.99%	−3.70%	−7.80%	−4.03%	−7.89%	−5.84%	2
4	2	1.61%	6.27%	3.37%	10.45%	12.93%	10.44%	9.44%	2.59%	2.55%	2
4	3	2.85%	−0.89%	−1.09%	−4.15%	−2.16%	−0.10%	1.83%	−2.44%	−3.25%	1
4	5	−5.21%	2.54%	6.22%	10.47%	9.33%	**17.37%**	11.44%	14.69%	14.18%	**6**
5	2	−0.39%	−3.04%	−1.34%	−3.31%	−4.94%	−7.21%	−2.83%	−0.40%	−1.13%	**36**
5	3	3.28%	14.04%	18.82%	19.43%	13.48%	14.26%	25.71%	29.34%	27.86%	2
5	4	0.01%	−1.30%	−0.63%	−0.20%	−5.50%	−0.45%	**0.34%**	−1.09%	−0.19%	**13**

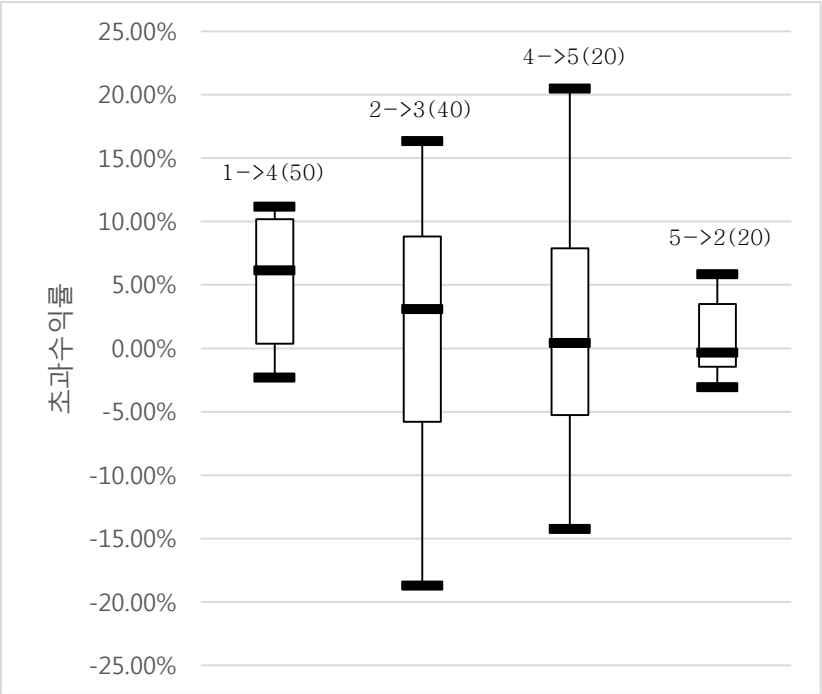
[표14] 3분기 기간별 평균 초과수익

[표11,12,13,14]로부터 분기별로 투자전략으로써 의미가 있는 군집간 이동과, 이때 가장 높은 초과수익률을 얻으려면 주식을 며칠 동안 보유한 후에 매도해야 할지 알 수 있었다. 이것들이 본 논문에서 제시하는 투자대안들이다. 하지만 이것만으로는 투자전략을 세우기에는 부족하다. 투자를 할 때 고려할 것은 단순히 높은 수익률만이 아니기 때문이다. 통칭 위험이라고 불리는 초과수익률의 변동성 역시 고려해야만 한다. 이를 위해서 발생 가능한 초과수익률의 분포를 고려하는 것이 아니라, 위에서 얻어낸 투자대안들로부터 얻어지는 초과수익률의 분포만 고려하였다.

이때 초과수익률의 분포는 상자-수염 그래프(Box-Whisker's plot)을 이용해 분석해보았다. 각 투자대안들의 분포에서 25 percentile(Q1), 50 percentile(median), 75 percentile(Q3), IQR(InterQuartile Range) 등의 값을 구한 후, Inner-fence(Q1-IQR ~ Q3+IQR)의 밖에 있는 극단치(outlier)를 제거하였다. 그런 후에 다시 초과수익률의 평균과 표준편차를 구함으로써 투자전략의 신뢰성을 높였다.

이 과정에서 극단치 제거 후에 초과수익률의 평균이 음수가 된 투자대안이 2개 있었다. 3분기에 발생한 5->4(70일)과 4분기에 발생한 5->2(10일)은 각각 kospi200 지수 대비 -4% 그리고 -0.9%의 초과수익률을 보였다. 이와 같은 경우에는 투자전략으로써 의미가 없다고 판단하여 투자대안리스트에서 제거하였다.

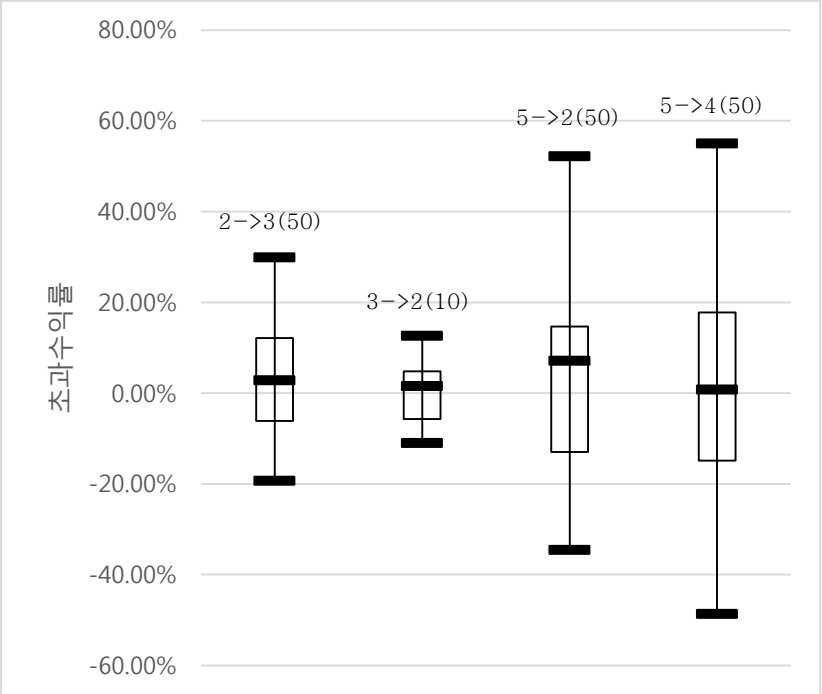
분기별 투자대안들의 초과수익률 분포를 상자-수염 그래프를 이용해 나타내 보았고, 그 결과는 다음과 같다.



[그림8] 1분기 투자대안별 초과수익률 분포

1 분기	1-4(50)	2-3(40)	4->5(20)	5->2(20)
평균	5.27%	1.03%	1.44%	0.15%
표준편차	6.13%	10.25%	9.75%	2.89%

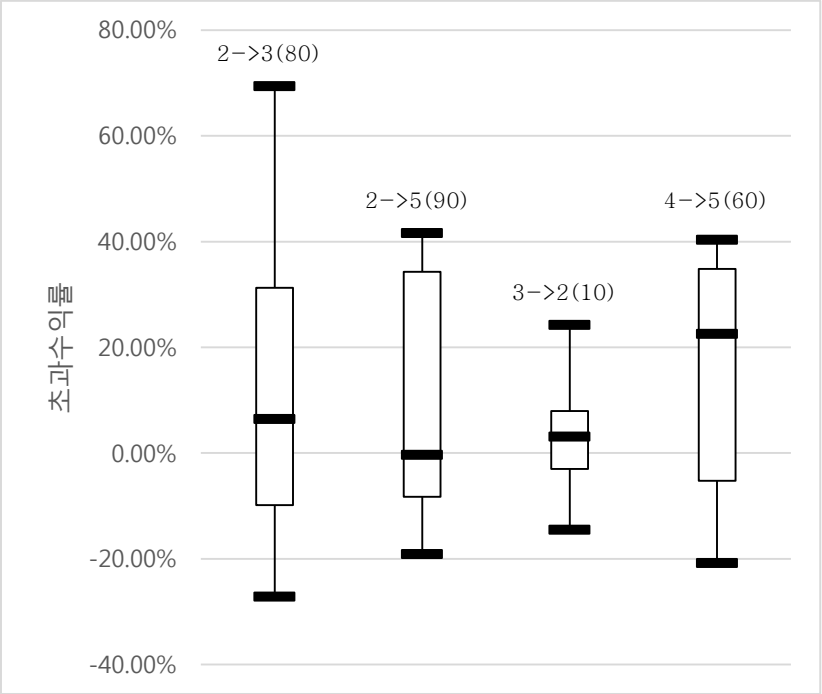
[표15] 1분기 투자대안별 초과수익률의 평균, 표준편차



[그림9] 2분기 투자대안별 초과수익률 분포

2 분기	2->3(50)	3->2(10)	5->2(50)	5->4(50)
평균	4.18%	0.46%	3.73%	1.67%
표준편차	12.43%	6.65%	19.98%	25.27%

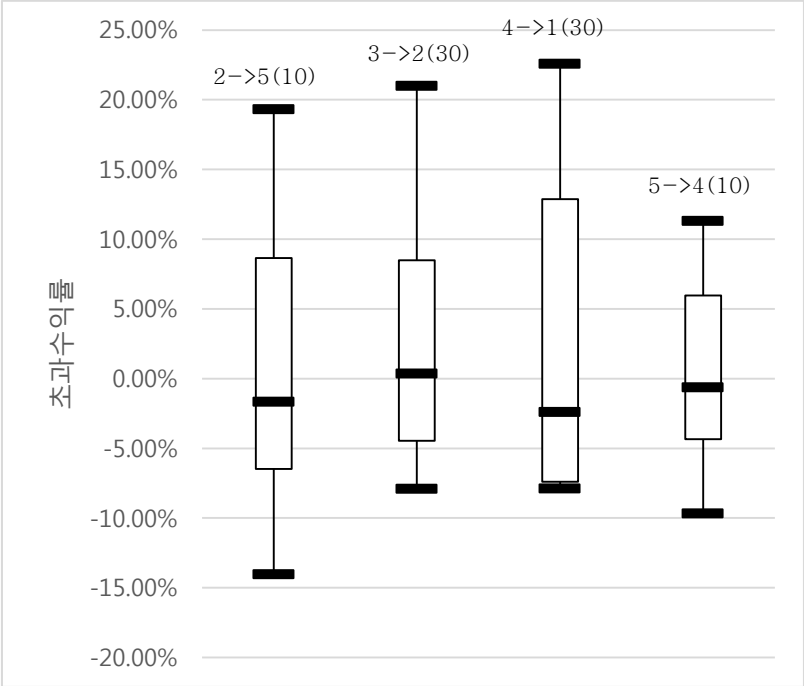
[표16] 2분기 투자대안별 초과수익률의 평균, 표준편차



[그림10] 3분기 투자대안별 초과수익률 분포

3 분기	2->3(80)	2->5(90)	3->2(10)	4->5(60)
평균	10.05%	8.28%	3.11%	17.37%
표준편차	27.86%	20.76%	9.43%	21.55%

[표17] 3분기 투자대안별 초과수익률의 평균, 표준편차



[그림11] 4분기 투자대안별 초과수익률 분포

초과수익률	2->5(10)	3->2(30)	4->1(30)	5->4(10)
평균	0.73%	2.55%	1.71%	0.14%
표준편차	9.03%	7.77%	12.45%	5.68%

[표18] 4분기 투자대안별 초과수익률의 평균, 표준편차

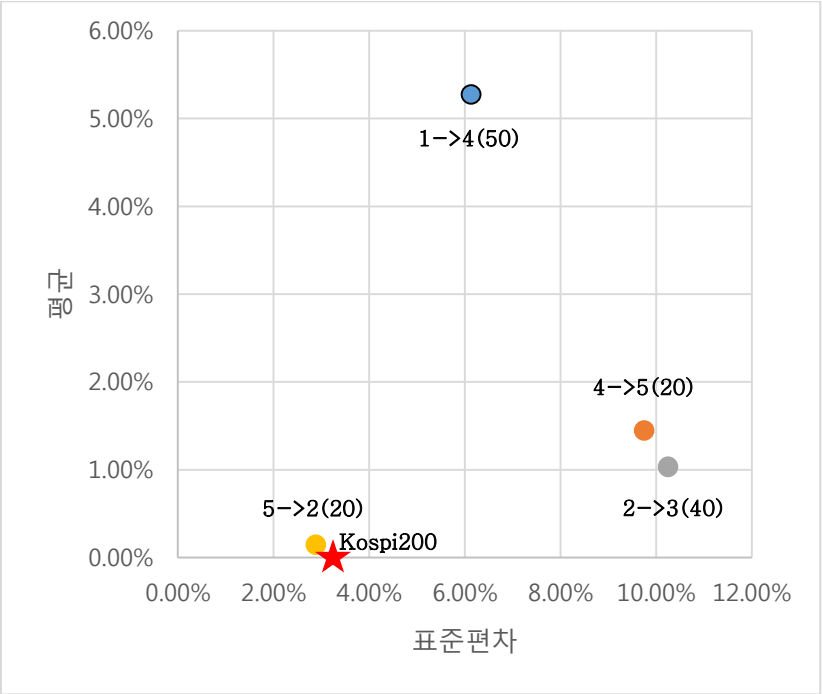
5.3. 분기별 투자전략 제시

위에서 구한 투자대안별 초과수익률의 분포를 이용해서 투자자에게 최적인 투자전략을 찾아보았다. 이때 각 투자대안의 비교에 있어서 지배원리(Dominance Principle)를 이용했다. 이성적인 투자자들은 얻을 수 있는 기대수익률뿐 아니라 감수해야 할 위험을 모두 고려해서 투자대상을 선택한다. 얻을 수 있는 기대수익률은 초과수익률의 평균으로 볼 수 있고, 감수할 위험은 초과수익률의 표준편차로 고려할 수 있다. 이렇게 평균과 표준편차를 모두 고려할 때, 최적의 투자대상을 찾기 위한 방법으로 지배원리가 있다.

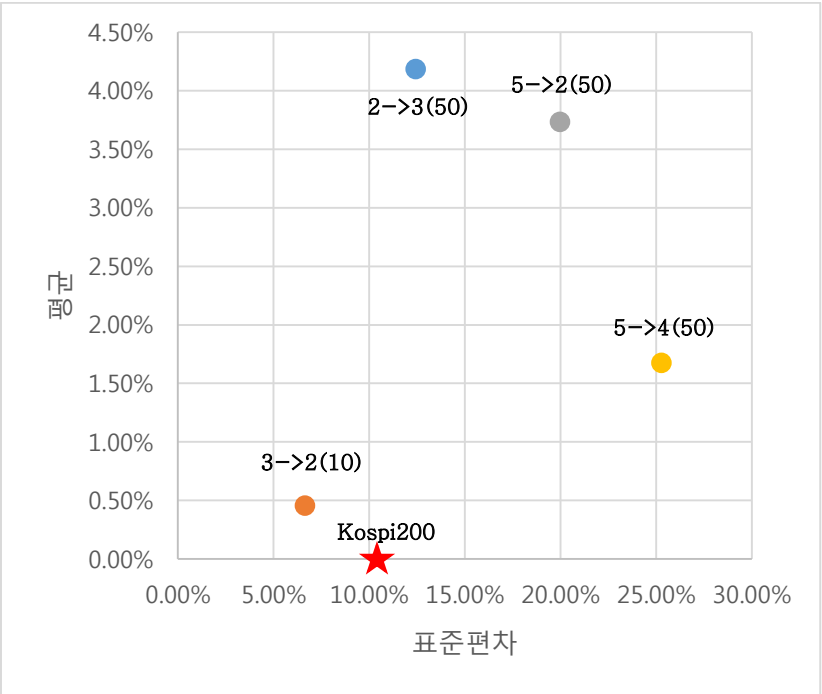
지배원리는 매우 직관적인 의사결정 방법이다. 투자대상 A와 B가 동일한 기대수익률을 갖고, A가 B보다 더 작은 표준편차 값을 갖는다고 가정하자. 이때 이성적인 투자자는 동일한 수익률을 더 낮은 위험으로 기대할 수 있는 A를 선택할 것이다. 이때 'A가 B를 지배한다'고 한다. 반대로 A와 B의 표준편차 값은 같고, B의 기대수익률이 A보다 더 높을 경우, 투자자는 같은 위험으로 더 높은 수익률을 기대할 수 있는 B를 선택할 것이다. 그리고 이때 'B가 A를 지배한다'고 한다.

즉, 지배원리란 같은 위험 하에서는 기대수익률이 높은 것이, 같은 기대수익률 하에서는 위험이 작은 것이 그렇지 않은 투자대상을 지배한다는 것이다. 보통 이 지배원리에 입각하여 선택되는 투자안을 효율적 투자대상(Efficient investment)이라고 한다.

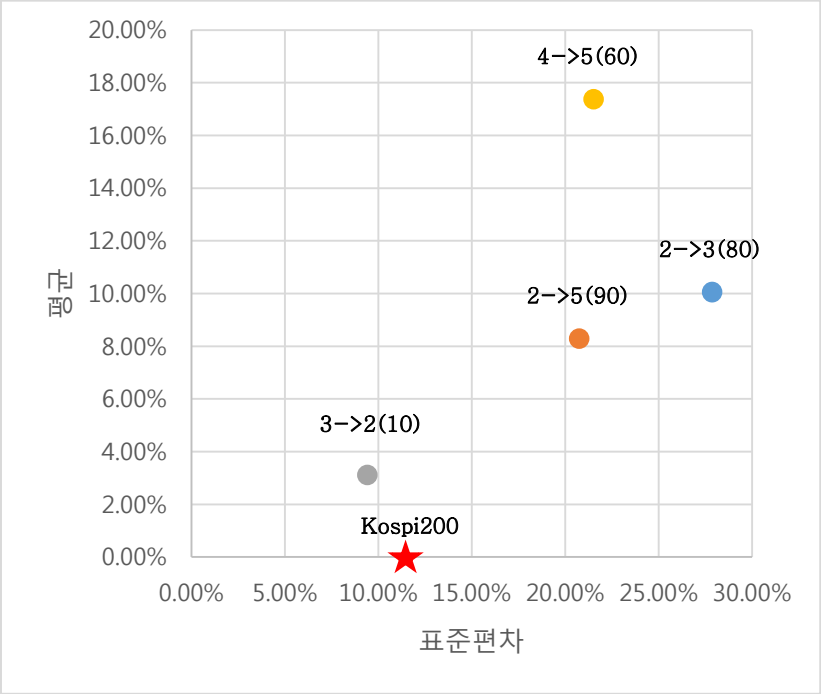
복수의 투자대상이 있는 경우, 지배원리를 이용해 하나의 효율적 투자대상을 찾을 수도 있지만 2개 이상의 효율적 투자대상이 남는 경우도 있다. 예를 들면 투자대상 A가 B보다 더 높은 기대수익률을 갖지만, 또 A가 B보다 표준편차가 높은 경우가 존재한다. 이럴 경우 A와 B의 우열을 가리기 어렵다. 이렇듯 효율적인 투자대상이 2개 이상 존재하는 상황에서 어떤 것을 최적의 투자안으로 선택하는지는 순전히 투자자가 선호하는 위험 수준에 달려있다.



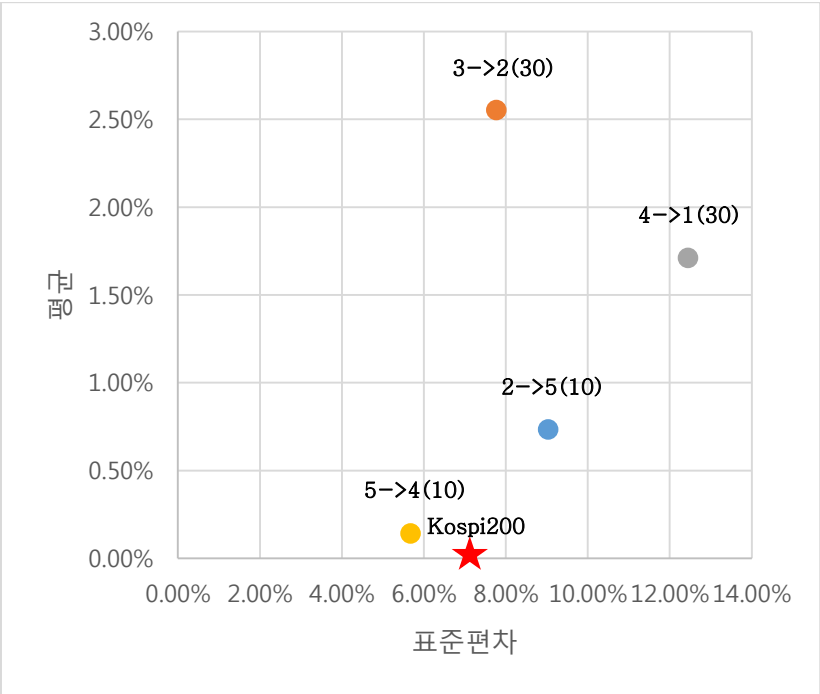
[그림12] 1분기 투자대안별 초과수익률의 평균-표준편차



[그림13] 2분기 투자대안별 초과수익률의 평균-표준편차



[그림14] 3분기 투자대안별 초과수익률의 평균-표준편차



[그림15] 4분기 투자대안별 초과수익률의 평균-표준편차

1분기 재무제표가 발표되었을 때, 1->4인 기업이 존재한다면 지배원리에 의해서 4->5나 2->3인 기업의 주식은 매수하지 않는다. 또한 4->5인 기업이 존재하는 경우 지배원리에 의해 2->3인 기업의 주식은 매수하지 않는다.
1->4, 4->5, 그리고 2->3인 기업 중 하나와 5->2인 기업이 동시에 존재할 경우에는 각각이 모두 효율적 투자대상이므로, 단순히 투자자의 성향에 따라서 최적의 투자전력을 선택한다. 공격적으로 1->4, 4->5, 그리고 2->3인 기업 중 하나의 주식을 매수하여 높은 수익률을 기대할 수도 있고, 반대로 안정적으로 5->2인 기업의 주식을 매수하여 낮지만 일정수준의 수익률을 기대할 수 있다. 또는 적당한 분산투자를 통해 일정수준의 위험을 감수하면서 또 일정수준의 수익률을 기대할 수 있다.

2분기 재무제표가 발표되었을 때, 2->3인 기업이 존재한다면 지배원리에 의해서 5->2나 5->4인 기업의 주식은 매수하지 않는다. 또 5->2인 기업이 존재하는 경우 지배원리에 의해 5->4인 기업의 주식은 매수하지 않는다. 2->3, 5->2, 그리고 5->4인 기업 중 하나와 3->2인 기업이 동시에 존재할 경우, 1분기에서와 마찬가지로 투자자가 선호하는 위험수준에 따라서 그에 맞는 투자전략을 수립할 수 있다.

3분기 재무제표가 발표되었을 때, 지배원리에 의해 4->5인 기업이 존재하는 경우에는 2->3인 기업의 주식을 매수하지 않는다. 3->2와 2->5, 그리고 4->5(혹은 2->3)인 기업이 동시에 존재하는 경우, 투자전략간에 우열을 가릴 수 없으므로 앞에서와 같이 투자자의 성향에 맞는 투자를 한다.

4분기 재무제표가 발표되었을 때에는 3->2인 기업의 주식을 매수하는 전략이 2->5, 혹은 4->1인 기업의 주식을 매수하는 전략을 지배한다. 더 적은 위험을 감수하면서 더 높은 초과수익률을 기대할 수 있기 때문이다.
반면에 2->5, 4->1, 3->2인 기업 중 하나와 5->4인 기업이 동시에 존재할 경우에는 하나의 투자전략이 나머지 하나를 지배하지 못하므로 투자자가 원하는 초과수익률, 감당할 수 있는 위험의 크기 등을 고려하여 투자한다.

제 6 장. 결론

본 논문에서는 기술적 분석과 기본적 분석을 적절히 활용하여 분기별 투자전략을 제시하였다. 먼저 기업들의 재무현황을 파악하기 위해서 11개의 재무인자와 재무비율을 변수로 횡단면적 군집분석을 실시하였고, 이를 통해 4659개의 개체를 5개의 군집으로 나누었다. 이때 각 군집에 속해있는 개체들은 비슷한 재무구조를 갖는다. 대부분의 기업들이 별다른 이유 없이 재무구조에 큰 변화를 주지 않기 때문에 시간이 지나도 동일한 군집에서 계속 머무른다. 예를 들면 A라는 기업이 2005년 1분기에 군집1에 있을 경우, 다음 분기에도 군집1에 속해 있을 것으로 생각하는 것이 일반적이다. 이는 [표7,8,9,10]에서 군집간 머무름 횟수가 군집간 이동 횟수보다 월등하게 많다는 것에서 확인할 수 있다. 때문에 기업의 군집간 이동이 발생한다면 이는 그 기업의 재무상태가 크게 변했다고 판단할 수 있고 주가에 영향을 줄 것을 예상할 수 있다. 다음으로 생각해 볼 점은 과연 이 재무현황의 변화가 주가에 좋은 영향을 미칠 것인지, 아니면 나쁜 영향을 미칠 것인지 이다. 실제적으로 군집간 이동을 한 기업의 주식을 샀을 때 주가가 떨어진다면 이는 좋은 투자전략이 아니다. 때문에 이 부분에서 과거의 데이터를 이용한 기술적 분석이 필요하다. 본 논문에서는 분기별로 군집간 이동이 발생했을 때의 초과수익률 분포를 분석, 비교하였고 이를 통해서 각 분기마다 투자자 성향에 맞는 최적의 투자전략을 제시했다.

[그림11,12,13,14]를 살펴보면 본 논문에서 제시하는 투자전략은 꽤나 매력적이다. 참고치(Benchmark)인 kospi200과 비교했을 때, 더 높은 수익률을 기대할 수 있기 때문이다. 하지만 한계점이 존재한다. 실질적으로 분석할 수 있는 데이터가 많지 않았다는 점이다. 본 논문에서는 기업들의 분기별 재무제표를 사용하였는데 실질적으로 군집간 이동은 자주 발생하는 이벤트가 아니기 때문에, 특정 군집간 이동이 발생한 횟수 자체가 그리 많지 않다. 물론 상자-수염 그래프를 이용해 극단치를 제거하거나 연도별 분기별 군집간 이동 횟수를 살펴봄으로써 투자전략의 신뢰성을 고려했지만,

분석할 샘플이 많지 않았던 점은 아쉽다. 때문에 추후 연구에서는 관찰기간을 더 늘린다면 더 의미 있는 결과를 얻을 수 있을 것이다. 또 단순히 관찰기관을 늘리는 것이 아니라 관찰기업 자체를 늘릴 수도 있다. 즉, 미국 S&P500에 속하는 기업을 대상을 연구를 진행한다면 충분히 좋은 결과를 얻을 수 있을 것이다.

참고문헌

1. Hyunchul Ahn, k-j.k, Corporate Bond Rating Using Various Multiclass Support Vector Machines, Asia Pacific Journal of Information Systems, 2009. 19(2): p.157-178
2. 안현철, 김, 한인구, 단분류 Support Vector Machine을 이용한 한국 기업의 지능형 기업채권평가모형, 경영학연구, 2006. 35(5): p.1479-1496
3. Efstathios Kirkos, Charalambos Spathis, Yannis Manolopoulos, Data Mining Techniques for the detection of fraudulent financial statements, Expert Systems with Apllication 32 (2007) 995-1003
4. Oyo, Ado Ekiti, Forecasting portfolio investment using data mining, African journal of computing & ICT (2012)
5. Richard A. Brealey, S.C.M., Franklin Allen, 핵심재무관리. 2ed. 2011
6. James C. Bezedek, Robert Ehrlich, William Full, FCM: The Fuzzy C-Means Clustering Algorithm, Computuers & Geosciences, 1984
7. 조풍진, 이창주, 신승민, 재무인자들의 clustering 연구 및 투자전략 제시, working paper

Abstract

In this paper, we analysis the enterprises in KOSPI 200 from the fourth quarter of 2003 to the first quarter of 2011 and suggest the quarterly optimal investment strategy. We can find the time when the financial structure of the enterprise changes and how it changes by the cross-sectional clustering analysis using quarterly financial statements. After the clustering analysis enterprise belongs to particular cluster in each quarter. At this time, if the cluster that the enterprise belongs to changes, for instance, LG Electronics belonged to cluster2 in the first quarter of 2005 and belonged to cluster3 in the second quarter of 2005, then we say 'there was the movement between clusters'. The movement between clusters means that the financial structure of the enterprise changes. Because the change of the financial structure understandably affects stock price, we can make a profit by buying the stocks at the publication date of the financial statements and selling the stocks after some time. Finally we analysis 'when and how the financial structure changes' have the best result by computing the excess return of the each movement type.

Keywords: cross-sectional clustering analysis, investment strategy, financial statement

Student Number: 2012-21061