KOSPI 200 예측에 있어서 개입시계열모형과 인공신경망모형의 성과비교

양유모*, 하은호**, 오경주***

*대우일렉트로닉스 품질신뢰성연구소 연구원
** 연세대학교 정보통계학과 교수
*** 한성대학교 경영학부 교수

Abstract

많은 경제 시계열 자료 중에서 주가는 국내외 경제상황은 물론 정부정책 등 시장 외적인 영향에 가장 민감하게 반응한다. 하지만, 지금까지의 주가예측에 있어서는 이러한 외부의 영향, 즉 개입(Intervention)이 발생했을 때 주가의 변동에 능동적으로 대처하는 모형이 부재하였다. 실제로 이러한 개입사실을 예측모형에 반영하지 않는다면, 주가예측 있어 그 예측력을 따진다는 것은 무의미하다고 판단된다. 따라서, 개입시점을 발견하고, 이 개입효과를 측정하여 이를 모형에 반영한다면 좋은 예측결과를 얻을 수 있을 것이다.

이 연구에서는 이상점 탐지절차를 이용하여 개입 시점을 발견하고 개입의 효과가 개입시점에만 영향을 주는 모형과 효과가 일정기간 지속되는 모형으로 두개의 개입시계열모형을 구축하고, 이러한 두 모형의 예측성과와 인공신경망모형을 이용한 예측성과를 비교하였다. 초단기예측(개입 직후 예측)에 있어서 개입의 효과가 지속되는 경우에는 개입시계열이 인공신경망보다 좋을 결과를 나타내긴 했지만 그 차이는 크지 않았으며, 개입의 효과가 시점에만 영향을 준 경우에는 인공신경망의 결과가 더 우수한 것으로 나타났다. 단기예측(개입 후 20 일 후의 예측)에 있어서는 개입 효과의 지속여부에 상관없이 인공신경망이 개입 시계열모형보다 우수한 것으로 나타났다.

I. 서론

자본주의가 시작된 이래 가장 많은 예측과 분석이 일어난 분야는 주식시장일 것이다. 지금도 주식시장에서 많은 사람들이 주가 패턴을 이용하여 주가를 예측하거나 이를 이용해 이익을 얻으려 하고 있다. 하지만 불행하게도 계속적인 이익을 얻을 수 있는 투자전략이나 주가 예측기법은 아직 개발되지 못한 것이 현실이다.[장우진, 2002]

주가를 예측한다는 작업은 그렇게 쉬운 일이 아니다. 오늘날의 금융 시장 움직임은 실로 매우 복잡다단하다. 실제로 주가를 제대로 예측한다는 것은 매우 어려운 일로 수많은 예측 방법들이 새로이 나타나고, 사라지고, 겨우 명맥만 유지하고 있다.

일반적으로 경제 시계열이 가지는 특징은 불충분한 자료의 수, 외부 잡음의 상당한 개입, 일정 시간 간격이 아닌 자료 수집, 비정상성을 가지는 경우가 많다는 점이다. 당연히 이들은 경제 시계열 분석과 예측을 근본적으로 어렵게 만드는 요인이 되고 있다.[김상락 외, 1997]

이 연구에서는 경제 시계열이 가지는 특

징 중에 외부 잡음의 개입 측면에 중점을 둔 시계열 분석에 초점을 두고자 한다. 특히 주가를 예측함에 있어서 주가에 변동에 영향을 주는 많은 요인들이 있다. 환율이나 금리의 변동으로 주가가 급락 또는 급상승할 수 있으며, 정부의 정책 발표 또한 주가에 큰 영향을 미치고 있다. 하지만 개입 변수를 결정함에 있어서 개입이 발생한 시점을 알아야만 한다는 단점이 있다. 그러므로 이 연구에서는 개입 발생 시점을 결정할 때 이상점 탐지 방법(Outlier Detection Method)을 사용하였다. 또한 카오스(Chaos) 방법을 이용한 인공신경망(Artificial Neural Network)방법을 이용하여 주가를 예측하였다.

본 연구에서는 실제 자료(KOSPI 200 지수)를 이용하여 두 모형의 예측성과를 비교하고자 하였다.

II. 문헌 연구

1. 통계학적 방법론에 의한 연구

주식 가격에 대한 예측 가능성에 대한 기존 연구는 크게 주가의 랜덤워크(Random walk)가설에 의한 연구, 주가의 자기 상관(Autocorrelation)을 이용한 연구, 시계열 모형을 이용한 연구 등으로 나눌 수 있다.

1.1 랜덤워크 가설에 의한 연구

Kendall은 주가 움직임의 패턴을 분석하였고, 주가에는 어떤 예측 가능한 패턴이 없음을 처음 발견하였다. 즉 주가는 랜덤워크를 따른 다는 것을 처음으로 발견하였다. [Kendall, 1953]

주가가 랜덤워크를 따른 다는 가설은 다우 이론(Dow theory)을 중심으로 하는 소위 기술적 분석(Technical analysis)과는 정면으로 상치된다. 기술적 분석은 주가의 패턴을 찾아 미래의 주가를 예측하려는 분석법으로 기본적으로 랜덤워크를 부인한다. 따라서 초창기의 랜덤워크 가설에 대한 검정은 기술적 분석이 유효한가에 많은 초점이 주어졌다. 그러나 최근에는 주가의 움직임이 비정상적, 비이성적이기보다는 효율적, 이성적 현상에 기인한다는 주장이 점점 설득력을 얻고 있다. [장우진, 2002]

1.2 주가의 자기 상관에 의한 연구

주가의 자기 상관에 대한 연구는 Moore(1964)가 처음으로 시도하였다. Moore의 연구에서는 1951년~1958년까지 29개의 개별 주식 가격에 대한 자기 상관을 계산하였다. Fama(1965)는 1958년~1962년까지의 30개의 Dow Jones 산업 평균 일별 지수의 자기 상관을 구하였다. 이 중에서 11개 산업에서 유의적인 자기 상관이 관찰되었으나 전체적으로는 유의적이지 못한 결과를 얻었다. Fisher(1966)는 과거 주가 수익률 패턴을 이용해 분산된 Portfolio의 월별 수익에서 양의 자기 상관이 관찰되었으나, 전체적으로는 유의적이지 못한 결과를 얻었다. French & Roll(1986)은 미국 주식 시장의 경우 주가 수익률은 일반적으로 상관계수가 0에 가까우나, 소규모 기업의 경우는 시계열 상관계수가 통계적으로 유의한 것으로 관찰되었다.

1.3 시계열 모형을 이용한 연구

시계열 모형을 이용하여 주가를 예측함에 있어서, 현재 발표된 대부분의 연구들은 다변량 시계열 모형을 이용하여 분석을 하였다. 남우식(1992)은 한국의 월별 종합주가지수를 예측할 때 주가와 관련된 주식 거래량을 입력 시계열로 신택하여 전이함수 모형과 개입 시계열 모형을 이용하여 주가를 분석하였다. Wei(1990)은 홍콩의 중국 반환 시점을 중심으로 홍콩의 일별 주가지수를 예측하였으나 유의적인 결과를 얻지는 못했다. 개입 모형과 이상값 모형의 응용방법으로 박완철(1995)은

1993년 8월 12일에 실시한 금융실명제라는 외부개입이 회사채 수익률에 미치는 효과를 분석하였다.

2. 인공신경망을 이용한 연구

현재 국내외적으로 인공신경망 모형이 주가의 변동을 예측하기 위해 학술적으로나 실무적으로 많이 이용되고 있다. 초기 주식시장 예측과 관련된 연구들은 주가지수 추정에 초점을 맞추었다.

Lapedes & Farber(1987)은 10주간의 주식 종가를 가지고 단순 자기회귀 모형을 학습시켜 다음 주의 주가지수를 예측하는 방법으로 하여 인공신경망이 단순이동평균법보다 나은 예측력을 갖고 있음을 보였다. Kimoto 외(1990)의 연구는 인공지능을 이용한 주식시장 예측의 초기의 대표적인 연구의 하나로서 몇 개의 학습 알고리즘을 선택하여 동경주식시장 지수(TOPIX) 예측 시스템이 구축되었다. 이 예측시스템에 사용된 신경망 모형은 3층의 완전히 연결된 구조를 갖고 1개월 후를 예측하는 모형으로 자료를 이동 시뮬레이션 방법으로 예측하였다. 노종래와 이종호(1991)는 주식관련 자료와 경제관련 자료를 사용하여 월별 종합 주가지수를 예측하려 하였다. 월별 예측의 경우는 77%의 높은 예측 결과를 얻었지만, 주별 예측에서는 50%에도 못 미치는 결과를 얻었다. 이재유(1993)는 주간 신용잔고, 고객예탁금과 같은 자료를 입력변수로 하여 종합주가지수와 개별 주식들의 주가를 예측하는 2중 은닉층을 갖는 역전파 인공신경망을 구축하였다. 연구 결과 통계적 모형의 결과가 25%에 머무르는 예측결과를 보인데 반하여, 인공신경망 예측의 정확도가 80%를 넘는 좋은 결과를 가져왔다. 박종엽(1995)은 KOSPI의 4주 후 상승 혹은 하락을 예측하기 위해 4종류의 독립적인 모듈화된 신경망을 구성하였다. 이들 4종류의 신경망은 그 학습기간을 달리하며 학습한 후 한번의 예측치를 산출하였다. 각 신경망의 예측치를 산술평균한 결과, 신경망은 79.46%, 다중회귀모형은 69.23%의 예측력을 나타내어 신경망이 다중회귀분석보다 우수한 결과를 나타냄을 보였다. 김경재(1997)는 한국주가지수 200 선물의 가격의 변동방향을 여러 가지 인공지능 기법을 이용하여 예측하였다. 여러 예측 모델에 사용될 변수를 선정함에 있어서 각각의 데이터의 성격이나 모델의 특성이 반영될 수 있도록 유전자 알고리즘을 사용하였다. 유전자 알고리즘/인공신경망의 결합모델과 유전자 알고리즘/사례기반추론의 결합모델의 데이터 전처리방식을 달리하여 실험한 후 그 예측성과를 비교하였다. 그 결과 유전자 알고리즘/사례기반추론의 결합모델의 예측성과가 더 우월한 결과를 나타냈다.

이처럼 현재까지도 주가 지수를 예측하기 위한 노력은 계속 되어오고 있으며, 계속적인 연구를 통하여 많은 예측기법들이 나타나고 있고 새로운 방법에 의한 예측성과를 측정함에 있어서 통계학적 방법론(ARIMA, Regression 등)과 비교하여 인공신경망의 예측성과가 우수하다는 결론이 많아지고 있다.[지원철, 1995]

III. 실증 분석 및 결론

1. 제안 모형

우리 나라의 주가의 변화에 있어서 어떤 개입이 발생한 경우에는 주가가 급격한 변화를 보이는 경향이 크다. 그러나 대체적으로 그 개입의 효과는 오래 지속되지 못하고 바로 사라지는 경향이 크다. 이 연구에서는 주가의 예측에 있어서 단기 예측의 성과 측정을 위해, 예측시점은 개입이 발생한 직후 10일의 자료를 예

측하는 것과 장기 예측의 성과 비교를 위해 개입이 발생한 후 20일이 지난 후의 10일의 예측을 하는 것으로 한다. 예측오차를 측정하는 방법으로는 예측변수들의 크기나 양이 예측에 중요한 영향을 미치는 경우에 사용하는 MAPE(Mean Absolute Percentage Error)를 이용하여 예측성과를 비교하였다.

2. 실증 분석

1990년 2000년까지의 KOSPI 200지수의 개입 이후 예측성과를 측정하기 위해 일년 단위의 일자료를 이용하여 첫 번째로 이상점 탐지를 통해 각 해에 발생한 이상점을 탐지하고 가장 마지막 이상점이 탐지된 시점을 개입 시점으로 가정하고 그 시점 이전의 60일의 자료를 가지고 분석을 시도하였다. 개입 직후의 10일의 자료를 예측하고, 개입이 발생한 후 20일이 지난 후, 10일의 주가지수의 예측을 시도하였다. 시계열 분석은 개입 시계열분석을 이용하였고, 인공신경망은 Chaos 상관차원을 이용한 역전파 인공신경망과 일반화 회귀 신경망을 이용하여 분석을 하였다.

2-1. 이상점 탐지

Chang & Tiao(1983)가 제안한 반복탐지 절차를 이용하여 KOSPI 200 지수의 이상점을 탐지하였다.

<표 1> 이상점 탐지 결과

년도	탐지시점	이상점의 종류
1990	없음	•
1991	없음	•
1992	없음	•
1993	시점 189	AO
1994	시점 252	IO
1995	시점 151	IO
1996	없음	•
1997	시점 242	AO
1998	시점 281	AO
1999	없음	•
2000	시점 41, 71	AO

이상점이 탐지된 년도 중에서 1993년, 1995년, 1997년을 개입시계열 모형과 인공신경망 모형을 이용하여 분석하였다. 개입발생시점에만 영향을 주는 모형은 [개입 모형A], 개입발생 이후 영향을 주는 모형을 [개입 모형B]라 하겠다.

2-2. 개입 직후 10일의 자료 예측

개입 직후의 자료의 예측은 개입 직후의 자료의 형태를 모르기 때문에 모든 모형의 예측성과를 비교하였다.

<표 2> 예측성과 비교(단기 예측)

년도	모형	MAPE
1993	[개입 모형A]	4.614
	[개입 모형B]	2.953
	인공신경망	**1.957**
1995	[개입 모형A]	1.947
	[개입 모형B]	5.514
	인공신경망	1.972
1997	[개입 모형A]	7.831
	[개입 모형B]	5.442
	인공신경망	**3.774**

모든 년도에서 인공신경망의 예측성과가 우수한 것으로 나타났으나 1995년, 개입의 시점에만 영향을 주는 경우에는 [개입 모형A]인 시계열 모형의 예측성과가 인공신경망과 비슷한 결과를 나타냈다.

2-3. 개입 후 20일이 지난 후 10일의 예측

개입 후 20일이 지난 후의 예측성과는 개입 후의 자료의 형태를 알기 때문에 모형에 맞는 예측성과를 비교하였다.

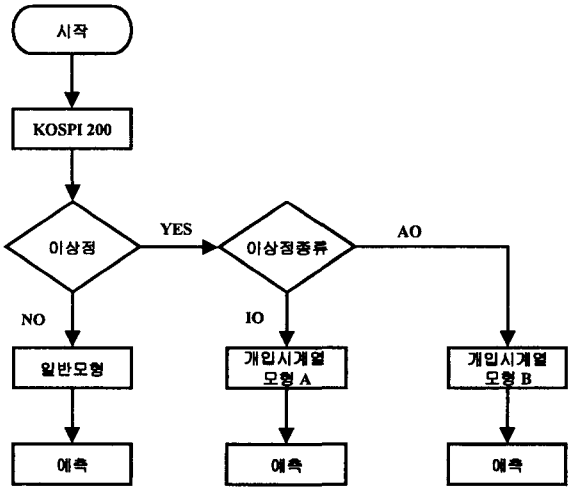
<표 3> 예측성과 비교(장기 예측)

년도	모형	MAPE
1993년	[개입 모형B]	3.011
	회귀신경망	2.698
1995년	[개입 모형B]	4.487
	회귀신경망	3.470
1997년	[개입 모형B]	7.763
	회귀신경망	14.295

개입이 발생하고 20일이 지난 후에 예측성과를 비교하는 것은 과연 개입의 효과가 어느 정도까지 지속되는 지에 대해서 알아보기 위한 것이다. 실증 분석을 한 모든 자료에서 개입이 발생하고 20일이 지난 후의 자료를 보면 어느 정도 일정 시간이 지난 후에는 개입의 효과가 사라지며 장기간의 예측에 있어서는 개입 시계열이 큰 효과가 없다는 것을 알 수 있다.

IV. 연구 과제

실제로 주가 예측을 위한 개입 시계열 모형을 사용하기 위해서는 선행되어야 할 것이 있다. 개입 시계열 모형에서 가장 중요한 것은 개입 시점을 찾는 것이다. 우리는 개입 시점을 알고 있다는 전제하에 개입 시계열 모형을 사용할 수 있다. 그리고 개입 이후의 동태모형의 형태를 알고 있어야 정확한 예측을 할 수 있다. 하지만 복잡하고 변화가 많은 주가의 움직임을 예측할 때 언제 어느 때에 개입이 발생할 지 모르는 것이 문제이고 주가가 어떻게 움직일 지 쉽게 예측할 수 없는 것도 하나의 문제이다. 그렇기 때문에 주가지수의 예측에 있어서 개입에 의한 주가의 예측을 하기 위해서는 다음과 같은 자동화 절차가 필요할 것이다.



<그림 1 > 이상점 탐지 및 예측의 자동화

KOSPI 200지수의 변동을 측정하여 이상점을 탐지하는 절차를 만들어서 이상점이 발견된다면 개입시계열 모형을 이용할 수 있을 것이다. 그리고 이상점의 종류에 따라 모형이 다르므로 지금까지의 KOSPI 200지수의 개입 발생 후의 자료의 분포를 파악한다면 앞으로의 개입이 발생했을 경우, 개입의 원인을 발견하여 예전의 사례를 기반으로 예측모형을 설정할 수 있을 것이다. 그리고 위의 자동화 절차를 만든다면, KOSPI 200지수의 변동에 따른 위험을 감소시킬 수 있는 장점이 있을 것이다.

참고문헌

이중협 외 [SAS/ETS를 이용한 시계열분석과 그 응용], 1994

조신섭 외 [시계열 분석], 1999

최병선, [단변량 시계열 분석], [다변량 시계열 분석], 세경사

Dan w. Patterson, [Artificial Neural Networks-Theory and Applications], Prentice Hall, 1995

김상락 외, 1997, 한국주가지수 수익률의 변동 특성에 관한 연구 - R/S 분석을 중심으로 - 한국재무관리학회, 재무관리연구

김경제, 1997, 범주형 전처리과정을 이용한 이공지능 기법에 의한 주가지수선물가격 예측

장우진, 2002, 「인공지능을 이용한 금융시장 예측연구」

박종엽, 1995, Predicting Korea Composite Stock Price Index Using Modular Neural Network, Master Thesis in KAIST

이재유, 1993, 인공신경망을 이용한 주가예측

지원철, 1995, 시계열 예측 의사 결정 지원시스템 : ARMA 모형 수립을 위한 신경망 접근방법

김상락 외, 1997, 한국주가지수 수익률의 변동특성에 관한 연구 - R/S 분석을 중심으로 - 한국재무관리학회, 재무관리연구

Box, G. E. P. & Tiao, G. C, 1975, Intervention analysis with applications to economic and environmental problems, Journal of America Statistics Association 70, 70-79

Chang, I. and Tiao, G. C., 1983, Estimation of time series in the presence of outliers, Technical Report 8, University of Chicago, Statistics Research Center, Chicago

Fama, E, 1965, The Behavior of Stock Market Prices, Journal of Business 38, 34-105

French, K, and R Roll, 1986, Stock Return Variances The Arrival of information and the Reaction of Traders, Journals of Financial Economics 17, 5-26

Fox, A. J., 1972, Outlier in time series, J. Roy. Statist. Soc., Ser. B, 43, 350-362

Kimoto, A.K, Yoda, M. and Takeoka, M, 1990, Stock market prediction system with modular neural network, Proceedings of the I. Joint Conference on Neural Networks, San Diego, California,

Lapedes, A. and Faver, R., 1987, Nonlinear Signal Processing Using Neural Networks. Prediction and System Modeling, Las Almos National Laboratory, LA-UR-87-2662