

Cascade-Correlation Network를 이용한 종합주가지수 예측[†]

지 원철, 박 시우, 신 현정, 신 흥섭
홍익대학교 산업공학과

Abstract

Korea Composite Stock Price Index (KOSPI) was predicted using Cascade Correlation Network (CCN) model. CCN was suggested, by Fahlman and Lebiere [1990], to overcome the limitations of backpropagation algorithm such as step size problem and moving target problem. To test the applicability of CCN as a function approximator to the stock price movements, CCN was used as a tool for univariate time series analysis. The fitting and forecasting performance of CCN on the KOSPI was compared with those of Multi-Layer Perceptron (MLP).

제 1 장 서 론

신경망을 이용한 주가예측에 관한 연구는 최근 국내외에서 활발히 진행되고 있다 [Barr and Mani, 94; Chen, 94; Hiemstra and Haefke, 95; 한재민, 95; 박종엽, 95]. 관련된 연구들을 신경망의 응용형태에 따라 구분하면 크게 두가지로 나눌 수 있다.

첫번째는 신경망을 패턴분류기로 사용하는 것이다. 즉 주가의 움직임을 상승 포함 하락 등 몇개의 계급으로 나누어 미래 주가가 속할 계급을 예측하거나, 기술적 분석에서 많이 사용되는 주가 패턴을 사전에 정의하고 같은 패턴을 미래의 주가 움직임 속에서 찾으려 하는 것이다 [Kamijo and Tanigawa, 90]. 두번째는 신경망을 함수근사기(Function Approximator)의 형태로 이용하는 것으로 과거의 주가나 주가에 영향을 미치는 요인들로부터 미래의 주가를 직접적으로 예측하려는 것이다 [Kimoto et al, 90].

본 연구에서는 신경망을 함수근사기의 형태로 이용하여 한국의 종합주가지수(KOSPI)를 예측하고자 한다. 기존의 연구들에서는 주가에 영향을 미치는 요인으로 거래량, 고객예약금, 사채이자율, 환율, Dow-Jones지수 및 거시경제지표등을 이용하여 다중회귀분석의 형태로 다계층 퍼셉트론(Multilayer Perceptron, MLP)을 이용한 경우가 많았다. 비록 MLP가 가장 많이 사용되고 있는 신경망 모형이지만 학습속도가 느리고 특히 좋은 예측성과를 얻기위해 적절한 크기의 은닉층 수를 결정하는데 많은 어려움이 있다는 단점이 있다. 이러한 MLP의 단점을 극복하기위해 학습방법이나 신경망구조에 변화를 주는 연구가 최근 많이 시도되고 있다 [Wasserman, 95].

본 연구에서는 이러한 연구 결과들 중의 하나인 Fahlman과 Lebiere [90]에 의해 제시된 Cascade Correlation Network (CCN)을 단일변수 시계열분석(univariate time series)의 형태로 이용하여 KOSPI의 예측에 적용하고자 한다. 단일변수 시계열 분석의 형태를 취한 이유는 CCN의 예측성과에 대한 연구결과가 아직 없기 때문에 설명변수의 선택으로 인해 예측성과가 영향받을 것을 방지하고 CCN의 주가예측에의 사용가능성, 즉 함수근사기로서의 성능을 보다 정확히 평가하기 위함이다.

제 2 장 Cascade Correlation Network

2.1 백프로파게이션 학습의 비효율성

Fahlman과 Lebiere [90]는 백프로파게이션 학습의 속도가 느린 원인으로 Step-size결정 문제와 Moving Target 문제의 두가지를 지적하고 있다. Step-size 결정문제는 표준 백프로파게이션 학습에 있어 전역오차함수에 대해 1차 부분도함수만을 사용하여 gradient, 즉 가중치의 갱신량을 계산하려하기 때문에 발생한다. 이 문제를 완화시키기 위해 관성항을 사용하거나 2차 도함수를 사용하는 방법들이 많이 제시되었다.

두번째로 Moving Target 문제는 MLP의 경우

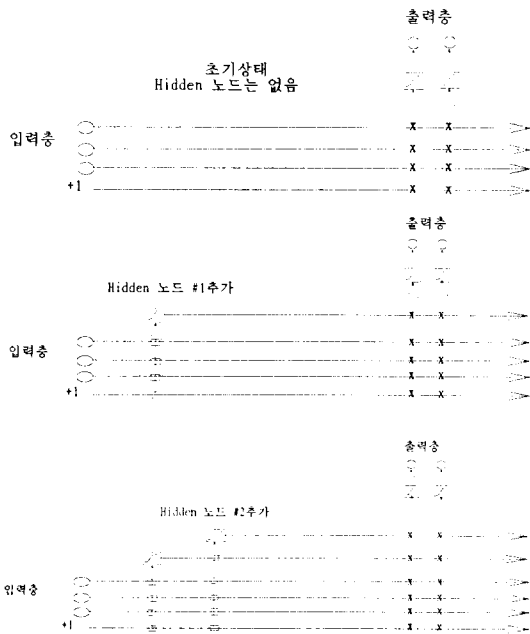
† 본 연구는 한국과학재단의 연구비 지원에 의해 이루어 졌음(과제번호: 951-0905-115-2)

신경망의 구조가 고정되어 있고 변화하는 입출력 패턴에 대해 내부가중치들이 모두 동시에 변화하게 되므로 생기는 문제이다. 즉 신경망 내부의 모든 노드들이 동시에 유용한 'feature detector'가 되려하기 때문에 오히려 학습속도를 늦추고 경우에 따라서는 학습결과와 편차가 심해지는 현상을 보이게 된다는 것이다.

Fahlman과 Lebiere는 Step-size결정 문제는 Quickprop 학습방법에 의해 해결하고 Moving Target 문제에 대해서는 새로운 신경망 구조를 제시함으로써 백프로파게이션학습에 의한 다계층 퍼셉트론의 한계를 벗어나고자 했다.

2.2. Cascade-Correlation Network의 구조

CCN의 구조는 <그림 1>과 같다. CCN의 학습은 은닉층의 노드들이 없는 상태에서 단순히 입출력 노드만을 가지고 시작한다.



<그림 1> Cascade 네트워크의 구조
(□ 표시 가중치는 고정되고 × 표시 가중치는 반복적으로 학습됨)

CCN의 학습을 진행시켜나가면서 오차의 감소가 더 이상 진전되지 않는 시점에서 일단 학습을 멈추고 해당 CCN의 성능을 평가하여 학습결과에 만족하면 학습을 종료하고, 만약 학습결과에 만족하지 못할 경우에는 다음과 같은 학습과정을 반복하면서 은닉층의 노드들을 하나씩 순차적으로 증가시켜 나간다.

새로운 은닉층 노드를 추가하기 위해서는 먼저 4~8개로 구성되는 후보노드군을 생성하여 입력층에 연결하고, 각기 다른 초기 가중치를 갖도

록 한다. 각 후보노드들은 학습자료에 대해 병렬적으로 학습하며, 이때의 학습목적은 후보노드의 출력값(V)과 CCN의 출력노드에서의 오차값(E)들 사이의 상관관계(S)가 최대로 되는 후보노드를 찾는 것이다. S를 최대로 하는 입력가중치(w_i)들의 값을 구하기 위해 $\partial S / \partial w_i$ 를 계산하여야 하는데 이를 수식으로 표현하면 (식 2-1, 2)와 같다. 여기서 p와 o는 입력패턴 및 CCN의 출력노드에 대한 첨자이다.

$$S = \sum_o \left| \sum_p (V_p - \bar{V})(E_{p,o} - \bar{E}_o) \right| \quad (\text{식 2-1})$$

$$\frac{\partial S}{\partial w_i} = \sum_o \sigma_o (E_{p,o} - \bar{E}_o) f'_p I_{i,p} \quad (\text{식 2-2})$$

where

V : 후보노드의 출력값

E : 출력노드에서의 오차값

\bar{V}, \bar{E} : 모든 패턴에 대한 평균

σ_o : 후보 및 출력노드 o의 출력값간의 상관관계의 부호

f'_p : 후보노드 활성화함수의 미분값

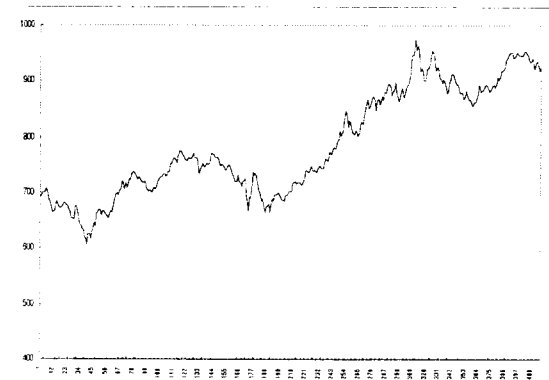
$I_{i,p}$: 후보노드가 입력노드 i로부터의 입력값

후보노드와 입력층 사이의 가중치에 대한 조정을 (식 2-2)에 의해 마친후 S가 최대로되는 후보노드를 새로운 은닉층의 노드로 선택한 후 이를 출력층에 연결한다. 출력 가중치의 학습은 입력가중치를 고정시킨후 Quickprop에 의해 진행된다.

제 3 장. CCN에 의한 종합주가지수 예측

3.1 데이터 및 실험방법

본 연구에 사용된 주가지수는 한국 증권시장의 종합주가지수(KOSPI)로서 1993년 1월 15일부터 1994년 6월 13일 까지의 일일 종가를 신경망의 학습을 위해 사용하였는데 <그림 2>와 같다.



<그림 2> 실험에 사용된 KOSPI 자료

학습된 신경망의 예측성능을 평가하기 위해 학습자료에 연이은 20일간 즉 1994년 6월 14일부터 7월 6일 사이의 KOSPI를 예측하도록 하였다. 학습 및 검증자료의 구성에 있어 공휴일과 일요일은 무시하였으며 신경망의 학습 및 예측은 주가수익률을 이용하지 않고 KOSPI 자체에 대해 실시하였다.

KOSPI의 예측을 위한 입출력값들은 과거 10일간의 KOSPI를 입력값으로 하고 바로 다음날의 KOSPI를 출력값으로 하는 단일변수 시계열분석의 형식을 취하였다. 따라서 예측은 (식 3-1)과 같은 형식으로 이루어 진다.

$$x_t = f(x_{t-1}, x_{t-2}, x_{t-3}, \dots, x_{t-n})$$

where

x_t : t 시점의 종합주가지수 (식3-1)

n : 입력값의 갯수

f : 학습 알고리즘에 해당하는 블랙박스

예측성과 평가를 위해서 Mean Squared Error (MSE)와 Mean Absolute Percentage Error (MAPE)의 2가지 척도를 사용하였다.

$$1. \text{MSE} = \frac{\sum_{i=1}^n (F_i - T_i)^2}{n} \quad (\text{식3-2})$$

$$2. \text{MAPE} = \frac{\sum_{i=1}^n \left| \frac{F_i - T_i}{X_i} \right| \times 100}{n} \quad (\text{식3-3})$$

T_i = 목표출력값
 F_i = 예측값

3.1.2. 학습멈춤조건

CCN에 의한 학습은 사용자가 사전에 학습 회수나 Error Index (EI) 값을 지정하여 학습을 제어할 수 있다. 본 연구에서는 EI 값을 사용하였는데 EI의 정의는 (식 3-4)와 같다. 본 연구의 목적이 CCN의 주가예측에의 적용가능성을 검증하는 것이므로 EI 값을 0.10 부터 0.03까지 0.01씩 줄여가며 8단계로 학습을 시행하였다.

$$\text{ErrorIndex} = \frac{\text{TrueError}}{N\text{OutputValue} \times \text{StdDev}} \quad (\text{식3-4})$$

$$\text{TrueError} = \sum (\text{예측값} - \text{목표출력값})^2$$

$$N\text{OutputValue} = N\text{output} \times N\text{Pattern}$$

$$\text{StdDev} = \text{출력값들의 표준편차}$$

3.2. 실험 결과

3.2.1. CCN에 의한 예측 결과

CCN 학습은 독립적으로 5회 수행하였다. 각 실행시 마다 EI 값을 0.10 부터 0.03까지 0.01씩 줄여가며 학습시켰는데, 사용된 KOSPI자료에 대해 EI의 값을 0.03이하로 떨어뜨리기는 매우 어려웠다. CCN의 학습결과 생성된 은닉층노드의 수를 정리하면 <표 1>과 같다.

<표 1> CCN내에 생성된 은닉층 노드의 수

EI RUN#	0.10	0.09	0.08	0.07	0.06	0.05	0.04	0.03
	1	2	6	14	21	27	34	42
2	2	5	12	22	30	37	45	56
3	3	6	13	19	26	34	41	51
4	3	5	12	19	27	33	41	51
5	2	6	13	22	30	38	47	58

EI = 0.05와 EI = 0.03일 때에 대하여 학습 및 예측결과를 MSE와 MAPE에 의해 정리하면 <표 2>와 같다.

<표 2> CCN의 학습 및 예측 결과

	RUN#	EI = 0.05		EI = 0.03	
		MSE	MAPE	MSE	MAPE
Fitting	1	26.558	0.482	8.460	0.253
	2	23.360	0.471	8.400	0.268
	3	23.338	0.485	8.422	0.260
	4	23.340	0.481	8.394	0.273
	5	23.318	0.464	8.386	0.273
Fore-casting	1	142.062	1.061	305.726	1.093
	2	122.220	0.988	180.944	1.070
	3	139.066	1.014	196.621	1.069
	4	91.547	0.845	160.967	1.907
	5	119.109	1.074	196.261	1.080

3.2.2. MLP에 의한 예측결과

CCN에서 사용된 것과 동일한 방법으로 KOSPI 자료를 사용하여 다계층퍼셉트론을 백프로파게이션 알고리즘에 의해 학습을 시켰다. 다계층퍼셉트론의 경우 은닉층노드의 수를 사전에 지정해 주어야 하므로 5가지 경우에 대해 학습을 시킨 후 결과를 <표 3>에 정리하였다.

3.2.3. 예측결과의 결과

<표 2와 3>에서 볼 수 있듯이 주어진 학습 자료에 대한 Fitting능력은 CCN이 MLP에 비해 월등함을 알 수 있다. 즉 MLP의 은닉층 노드수가 50일 경우에도 CCN의 Fitting능력에 미치지 못했다. 검증자료에 대한 예측성파에 있어서는 MLP가 30~40개의 은닉층 노드를 가질 때 가장 좋은 예측성파를 보였으며 CCN의 경우 EI=0.05일 때 네 번째 시행에서 MLP와 유사한 예측성

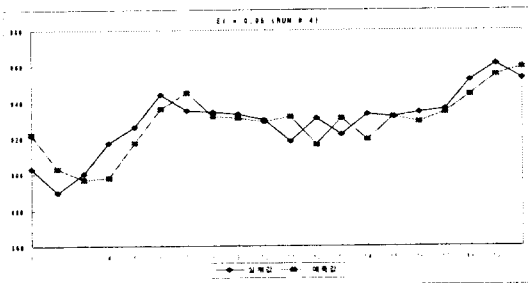
제 4 장 결 론

	은닉층노드 수	MSE	MAPE
Fitting	10	119.113	1.114
	20	103.362	1.071
	30	100.851	1.005
	40	85.093	0.925
	50	79.440	0.900
Fore-casting	10	115.062	0.940
	20	122.120	0.952
	30	119.181	0.898
	40	116.227	0.886
	50	123.454	0.945

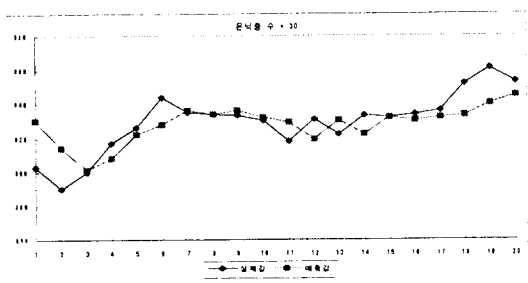
<표 3> 다계층퍼셉트론의 학습결과

과를 보였는데 이 때의 은닉층 노드의 수는 33개가 생성되었다. 이 결과로 부터 유추해 볼 때 CCN의 학습결과로 생성된 은닉층의 노드 수를 MLP의 은닉층의 노드수 결정에 이용할 수 있을 것이다.

CCN의 학습결과를 분석해 보면 Fitting결과는 시행에 관계없이 MSE나 MAPE가 매우 안정적임에 비해 예측결과는 편차가 심한 편이다. 이는 EI=0.05의 값이 CCN으로 하여금 학습자료에 과잉적합(Overfitting)이 되도록 한 것으로 생각해 볼 여지가 있어 추가적인 분석의 필요성이 있다. <그림 3과 4>는 예측구간에 대한 실적자료와 예측결과를 비교하여 나타낸 것인데, 특히 CCN의 경우 KOSPI의 변동패턴을 정확히 하루 늦게 뒤 쫓아가는 것을 볼 수 있다. 이와같은 현상은 Overfitting에 대한 검토의 필요성 및 단일변수 시계열분석이 갖는 한계를 반영하는 것으로 볼 수 있다.



<그림 3> CCN의 예측값과 실제값의 비교



<그림 4> MLP의 예측값과 실제값의 비교

CCN은 은닉층의 노드들이 'High-order feature detector'로서 작용할 수 있는 신경망 구조를 가지고 있고 학습방법도 Quickprop방법을 사용함으로써 주어진 학습자료에의 적응력이 매우 뛰어나는 뿐만 아니라 학습속도에 있어서도 MLP에 대해 크게 개선되었다. 본 연구에서는 CCN을 한국종합주가지수(KOSPI)의 예측에 적용해 본 결과 주가 예측수단으로서 매우 유용함을 입증하였다. 특히, KOSPI자료에 대해 직접 적용했음에도 불구하고 뛰어난 Fitting능력을 보인 것은 CCN이 비안정적 시계열에 대한 유력한 분석 수단이 될 수 있음을 보여 주는 것이다. 하지만 주가예측에 있어 본 연구의 한계를 벗어나기 위해서는 주가에 영향을 주는 요인들을 취사선택하여 예측모형에 포함시키고 주가자료를 수익률자료로 변환시켜 시계열을 보다 안정적으로 변화시키는 것이 필요하다.

[참 고 문 헌]

- 한재민, 성시중, (1995). 백프로과제이션을 이용한 종합주가지수예측, 한국전문가시스템 '95 추계학술대회논문집, pp345-357.
- 박종엽, 한인구, (1995). 인공신경망을 이용한 종합주가지수 예측, 한국전문가시스템 '95 추계학술대회논문집, pp359-371.
- Barr, D. & G. Mani, (1994). Using Neural Nets to Manage Investments, *AI Expert*, pp16 -21.
- Chen, C. H., (1994). Neural Networks for Financial Market Prediction, *Proc. of the 1994 IEEE Int'l Conf. on Neural Networks*, pp1199 -1202.
- Hiemstra, Y. & C. Haefke, (1995). Predicting Quarterly Excess Returns: Two Multilayer Perceptron Training Strategies, *Proc. of the 3rd Int'l Conf. on AI Applications on Wall Street*, pp212-217.
- Jhee, W.C. & M.J. Shaw, (1994). Time Series Prediction Using Minimally-Structured Neural Networks: An Empirical Test. *Proc. of World Congress on Neural Networks*, San Diego, CA., USA, June 5-9.
- K. Kamijo & T. Tanigawa, (1990). Stock Price Pattern Recognition: A Recurrent Neural Network Approach, *IJCNN* pp1215-222.
- T. Kimoto, K. Asakawa, M. Yoda & M. Takeoka, (1990). Stock Market Prediction with Modular Neural Networks, *IJCNN*, pp11-6.
- Wasserman, P.D. (1993). *Advanced Methods in Neural Computing*. VNR, New York