한국컴퓨터정보학회 동계학술대회 논문집 제20권 제1호 (2012. 1)

# 신경망을 이용한 시계열 패널자료의 예측

김인규°
°우송정보대학 컴퓨터정보계열
e-mail:ikkim0056@wsi.ac.kr

# Prediction for Time Series Panel Data using Neural Network

Inkyu Kim°
°Division of Computer Information, Woosong Information College

● 요 약 ●

본 논문은 여러개의 독립적인 시계열로 구성된 시계열 패널 자료를 이용하여 비선형 모형인 GRCA모형과 신경망을 이용하여 예측값을 구하여 서로 비교 분석하고자 한다. 먼저 GRCA모형에 대하여 연구하고 신경망의 구조와 예측값을 구하기 위한 여러 가지 변환함수를 유도한다. 단기 예측에서는 신경망 방법의 예측값이 더 좋았고, 장기예측에서는 비선형모형을 이용한 예측값이 더 좋은 것으로 나타났다.

**키워드:** 신경망(neural network), GRCA모형, 패널자료(panel data), 예측(prediction)

## I. 서론

여러개의 독립적인 시계열로 구성된 시계열 패널 자료는 실제 데이터 분석시 자주 나타나고 있다. 예를 들어 여러명의 혈압을 여러 시점에 걸쳐 기록하는 경우에 개인별 자료는 서로 독립이라 할 수 있고, 각 개인의 혈압기록은 종속 구조를 가진 시계열 자료로 모형화 할 수 있다. 신경망 모형은 수많은 뉴런(neuron)으로 연결되어 있는 뇌의 구조와 같이 서로 다른 노드와 연결된 구조를 가지고 다른 뉴런으로부터 입력을 받으면 뉴런은 활성화되고 다른 뉴런들에게 신호를 통해서 전달된다. 신경망 모형은 이러한 신호의 전달을 수학적으로 모형화한 것이며 이는 여러 응용 분야에 폭넓게 적용되고 있다. 본 논문에서는 이분산성을 따르는 비선형 시계열 자료를 가지고 비선형 모형중 GRCA모형과 자료의 형태에 상관없이 적용할 수 있는 신경망(neural networks)모형을 이용하여 예측을 해서 어느 모형이 시계열 자료에 적합한지를 비교 분석을 해보는 것이다.

## II. 관련 연구

### 2.1 GRCA모형

비선형 시계열 모형인 p차 확률계수 모형(RCA(p))은 아래와 같다.

$$X_t = (\Theta_1 + Z_{t1})X_{t-1} + \cdots + (\Theta_p + Z_{tp})X_{t-p} + \varepsilon_t$$

여기서{ $Z_{tj}$, j=1,⋯,p}는 평균이 0이고 분산이 유한한, 독립이고 동일한 분포를 갖는 확률변수이며, { $\varepsilon_t$ }와 독립이다. 또한 { $\varepsilon_t$ }는 평균이 0이고 분산이 $\sigma^2$인 독립이며 동일한 분포를 갖는 확률변수의 수열이고, 그리고 $\Theta_1, \cdots, \Theta_p$ 는 모수들이다.

Hwang과 Basawa(1998)는 일반화된 p차 확률계수 모형(GRCA(p))을 소개하였는데 그 모형은 다음과 같다.

$$X_t = \Phi'_t X(t-1) + \varepsilon_t$$

여기서 $\Phi_t = (\Phi_{t1}, \cdots, \Phi_{tp})'$는 (p×1)인 확률계수의 벡터, $X(t-1) = (X_{t-1}, X_{t-2}, \cdots, X_{t-p})'$는 시점 t에서의 (p×1)인 과거 관측치의 벡터이고, $\{\varepsilon_t\}$는 확률 오차들의 수열이다.

### 2.2 신경망(Neural Network) 모형

신경망의 구조는 입력층, 은닉층, 출력층으로 구성되어 있으며 일반적으로 신경망의 구조는 입력노드의 개수 × 은닉노드의 개수 × 출력노드의 개수로 나타낸다. 본 논문에서는 아래의 변환 함수를 사용하였다.

$$g(u) = \frac{1}{1+e^{-u}} \quad , \quad 0 < g(u) < 1$$

학습절차에서 C언어를 이용한 역전파 프로그램의 개요는 다음과 같다.

```
for(i=0; i<examples; i++)
{
   forward( );
   back( );
}
changeWeight( );
```

학습규칙은 학습과정에서 가중값의 변화량을 적절하게 결정하는 방법으로서 $w_m$을 가중값 $w$의 m번째 학습 후의 변화된 가중값이라 하면 $w_m$은 다음과 같다.

$$w_m = w_{m-1} + c_m$$

여기서 $c_m$은 m번째 학습량이 끝났을 때 가중값 $w$의 변화량이다. 즉,

$$c_m = \mu c_{m-1} - (1-\mu) e_m d_m \qquad 0 \le \mu < 1.$$

여기서 $e_m$은 가중값이 변화하므로써 최소값으로 수렴되어가는 가중값에 대한 학습 비율로서 다음과 같다.

$$e_m = \begin{cases} e_{m-1} + \chi, & d_m f_m > 0 \\ e_{m-1} \times \phi, & d_m f_m \le 0,\ 0 < \phi < 1. \end{cases}$$

여기서 $\chi$와 $\phi$는 모수이다. 그리고 $d_m$은 오차를 가중값에 대하여 미분한 모든 예제들의 누적합으로 m번째 학습후의 누적합이며, $f_m$은 현재와 과거의 미분값들의 가중 평균이다.

## III. 결론

선형 Box-Jenkins모형과 신경망 방법을 비교분석하는데 있어서 단기예측에서는 선형 Box-Jenkins모형의 예측값이 좀 더 좋은 예측값을 가졌고, 장기예측에서는 신경망 방법의 예측값이 더 좋은 예측값을 갖는 것을 알았다(Tang(1991)). 본 논문에서는 단기예측에서는 신경망 방법의 예측값이 더 좋았고, 장기예측에서는 비선형모형을 이용한 예측값이 더 좋은 것으로 나타났다. 이러한 결과를 통해서 보면 비선형모형은 관찰된 자료의 전체적인 추세와 경향을 그대로 반영하므로 초기의 예측값보다는 중장기의 예측값에 좀더 적절한 예측값을 가짐을 알수 있었다.

## 참고문헌

[1] Hwang, S. Y. and Basawa, I. V. (1998). Parameter estimation for generalized random coefficient autoregressive processes. Journal of Statistical Planning and Inference 68, 323-337.

[2] Nicholls, D. F. and Quinn, B. G. (1982). Random Coefficient Autoregressive Models: An Introduction. Lecture Notes in Statistics, Vol. 11, Springer-Verlag, New York.

[3] Tang, Z., Almeida, C., Fishwick, P.A. (1991). Time Series Forecasting Using Neural Networks vs. Box-Jenkins Methodology, Simulation, 57(5), 303-310.

[4] White, H. (1988). Economic Prediction Using Neural Networks : The Case of IBM Stock Prices. IJCNN, 2, 451-458.