모델기반 방법론을 이용한 환율예측 모형 연구

전진호* · 문석환** · 이채린*

*관동대학교 **영동대학교

A Study of Exchange rate Prediction Model using Model-based

Jin-Ho Jeon* · Seok-Hwan Moon** Chae-Rin Lee*

*Kwan Dong University **Young Dong University

E-mail : jhjy@kd.ac.kr, shmoon@yd.ac.kr, 20081012@kd.ac.kr

요 약

경제적인 국제화가 심화되어 세계경제가 통합화되는 환경에서 기업 및 개인 금융기관 등의 외환거래 참가자들에게 회환거래로 인한 환위험의 회피방안이 무엇보다 절실하다. 이 방안을 마련하기 위해서 본 연구에서는 환율, 주가와 같은 시계열데이터의 모형추정에 적합한 모델을 통해 단기 환율의 예측모형을 추정하고 이를 통해 향 후 예측에 적용한다. 실제의 환율 데이터를 통하여 최적의 모형이 추정된다면 이를 통해 향후의 일정기간의 운동양태의 예측이 가능할 것이다. 은닉마아코프모형의 추정을 위하여 베이지안정보기준을 통해 모형의 상태 수를 정확하게 추정하는지를 확인하였으며 추정되는 모형으로 예측한 결과 실제 운동양태와 예측에 있어 두 곡선의 운동양태가 유사함을 확인하였다.

ABSTRACT

Forex trading participants, due to the intensified economic internationalization exchange risk avoidance measures are needed. In this research, Model suitable for estimation of time-series data, such as stock prices and exchange rates, through the concealment of HMM and estimate the short-term exchange rate forecasting model is applied to the prediction of the future. Estimated by applying the optimal model if the real exchange rate data for a certain period of the future will be able to predict the movement aspect of it. Alleged concealment of HMM. For the estimation of the model to accurately estimate the number of states of the model via Bayesian Information Criterion was confirmed as a model predictive aspect of physical exercise aspect and predict the movement of the two curves were similar.

키워드

환율. 은닉마아코프모델, 베이지안정보기준, 단기예측

Ⅰ. 서 론

새로운 경제 질서 즉, 경제적인 국제화가 심화되어 세계경제가 통합화되고, 내국시장에서도 수많은 외국기업들과 경쟁을 하지 않을 수 없게 되었다. 이러한 개방경제하에서 일국의 비즈니스 투자, 거시경제 정책의 결정 및 운용에 영향을 미치는 가장 중요한 변수 중의 하나가 환율이다. 그러므로 본 연구에서는 환율에 대한 예측 문제에 시계열데이터의 모델링에 적합한 은닉마아코프모델을 통해 환율예측 모델의 추정과 이를 통한 예측문제에 적용하고자 한다.

Ⅱ. 관련 연구

환율, 주식 등과 같은 재무 분석은 기본적 분석과 기술적 분석으로 구분하여 볼 수 있다[1].

기술적 분석방법은 자산가격의 과거 움직임이 미래의 자산가격 움직임을 예측하는 데 의미 있는 정보를 제공한다는 가정에 바탕을 두고 있다. 본 연구에서는 모델기반의 예측을 통해 기술적 분석방법의 유효성을 평가한다[2].

은닉마아코프모델은 환율, 주가 등의 시간적 특징으로 묘사되는 시계열 데이터의 표현 모델링에 적합하다. 이유는 각 상태에서 특징들에 대한

적합한 확률함수를 사용하여 연속적인 값을 갖는 시간적 특징을 갖는 데이터를 쉽게 처리하며 데이터의 묘사가 쉽기 때문이다[3].

III. 은닉마아코프모델(HMM)

1. 은닉마아코프모델

은닉마아코프모델은 은닉(Hidden)상태를 예측하기 위하여 실제적인 관측을 통해서 변화되는 통계적인 특징들을 확률적으로 모델링하는 방법이다. M개의 상태들, $S=(S_1,S_2,...,S_M)$를 갖으며, K 시간의 특징들로 묘사되는 은닉마아코프모델은 $\pi=\{\pi_i\}$, $\pi=P(q_1=S_i)$로 표시되는 상태초기확률, $A=\{a_{ij}\}$, $a_{ij}=P(q_{t+1}=S_j \mid q_t=S_i)$로 표시되는 상태천이 확률, $B=\{b_j(x)\}$, $b_j(x)=P(O_t=x \mid q_t=S_j)$로 표시되는 방출확률 3개의 확률 집합으로 표현된다[4].

2. 모델 상태 수 결정기준

본 연구에서 은닉마아코프모델에서 상태의 수를 결정하기 위해서 휴리스틱 탐색 기준인 베이지안정보기준을 적용한다. 베이지안정보기준(BIC)는 라플라스 근사로부터 유도되며 다음 식(1),(2)으로 표현된다.

$$\log P(X|M) \approx \log P(X|\hat{\theta},M) + \log P(\hat{\theta}|M) + \frac{d}{2}\log(2\pi) - \frac{1}{2}log|A| \quad (1)$$

$$\log P(M \mid X) \approx \log P(X \mid M,\hat{\theta}) - \frac{d}{2}logN \quad (2)$$

위 식(2)에서 d는 모델에서 파라미터의 수이다. N은 데이터 객체들의 수이고 $\hat{\theta}$는 모델 M의 한계우도(ML)의 파라메터 구성이다. 식(2)에서 첫 번째 항은 데이터를 가장 잘 설명할 수 있는 상세한 데이터의 모델을 찾도록 유도하는 성분이다. 두 번째 항은 은 모델 내의 파라메터 개수에 대한penalty 항으로 볼 수 있다[5]. 베이지안정보기준은 이러한 두 항에 상호배타적인 특성이 서로 조화되는 타협점에서 정확하지는 않지만 효율적인 상태 수를 결정한다.

IV. 실 험

환율 데이터를 통하여 예측모형의 유효성을 확인하기 위하여 2009년 1월2일부터 3월30일까지, 2010년 1월4일부터 3월30일까지, 2011년 1월3일부터 3월30일까지의 3개년의 1분기의 데이터의 길이가 60인 실제 환율데이터를 적용하였다. 실험은 예측모형을 생성해보고 이를 통해 실제와 예측되는 곡선의 운동양태를 비교하여 확인해 봄으로서 모형이 최적으로 생성된 것인지를 확인해 보고자 한다.

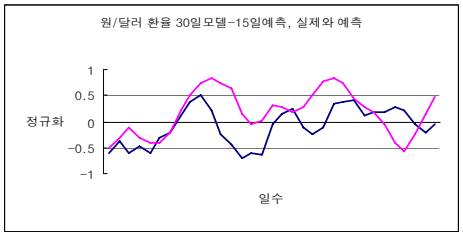

그림1. 30일 데이터 모델-15일 예측 운동양태

그림1은 길이가 30일인 데이터를 통해 모형을 생성한 후 실제와 예측 곡선의 운동양태를 보여주고 있다. 그림1에서 실제와 예측 곡선의 운동양태 모두 초반에 적은 등락을 반복하다 중반에 급격한 등락을 보이고 후반에 하락과 상승을 보이는 유사한 운동양태를 보이고 있다.

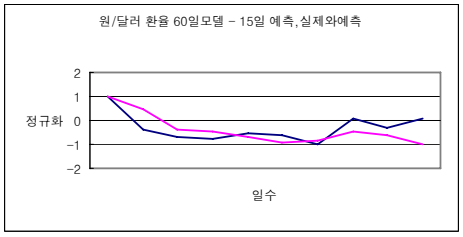

그림2. 60일 데이터 모델-15일 예측 운동양태

그림2는 길이가 60일인 데이터를 통해 모형을 생성한 후 실제와 예측 곡선의 운동양태를 보여주고 있다. 그림2에서 실제와 예측 곡선의 운동양태 모두 초반에 급격한 하락을 보이고 중반에는 변화가 없는 운동양태를 후반에는 상승과 하락하는 유사한 운동양태를 보여주고 있다.

위의 실험결과들에 대하여 즉, 예측결과에 대한 평가방법으로 평균제곱오차(mean square error)를 적용하였다. 즉, 실제의 곡선 운동양태와 예측 곡선의 두 운동양태에 대하여 정규화 후, 두 운동양태의 일별변화의 차이를 측정한다. 두 운동양태의 일별변화의 차이가 적을수록 에러율이 적은 것으로서 유사한 운동양태임을 확인하는 방법을 적용하였다. 정규화 후 실제와 예측 운동양태 곡선 사이의 에러율의 평균치는 0.2에서 0.9 사이에 존재함을 실험을 통해 확인하였다. 중앙값에 해당하는 평균치 임계값을 0.5로 정하였을 때 두 운동양태 사이의 에러율값이 임계값보다 낮을수록 두 운동양태에 유사함을 보여주며, 높을수록 비유사성을 보여준다. 위의 임계값을 기준으로 30일 모델을 통한 15일 예측에서는 불일치의 날짜가 평균 1일로서 90%의 예측정확도를 확인할 수 있었다. 또한 60일 모델을 통한 15일 예측에서는 불일치의 날짜가 평균 3일로서 80%의 예측정확도를 확인할 수 있었다.

위의 결과를 통해 모형 추정과 예측에 적용된 데이터 길이가 길어질 때보다 짧은 기간의 훈련 데이터에서 예측되어진 운동양태들이 실제의 운동양태와 좀 더 유사한 운동양태를 보여주는 것

은 기간이 길어질수록 외환시장에서 발생되는 수많은 불규칙적인 외적변수가 외환시장에서 반영되어 운동양태에 내재되는 것으로 고려된다.

Ⅴ. 결 론

본 연구는 미래의 환율예측을 위하여 환율데이터를 통해 은닉마아코프모델에 적용하여 예측모형을 추정하였으며 예측문제에 적용하였다.

실험에서는 모형의 상태 수를 추정하는 부분에서는 시간과 비용의 절약과 모형의 확장성을 고려하여 베이지안정보기준을 적용하였으며 이를 통한 모형의 생성을 통해 미래의 운동양태를 예측한 결과 실제의 환율 운동양태와 매우 유사한 운동양태를 예측하는 것으로 확인하였다.

앞으로 더 다양한 부분의 일반화된 재무, 경제 관련 데이터에 적용하기 위한 연구를 통하여 일반적이고 유연한 분석, 예측 모형을 세울 수 있는 부분으로 확대가 필요할 것이다.

참고문헌

[1] 신성환, “인공신경망모형과 이동평균법의 환율 예측력 평가”,금융연구, pp103-135, 1995.

[2] L. Rabiner, " A tutorial on Hidden Markov Models and selected applications in speech recognition," Proc. of IEEE77, pp.257-286, 1989.

[3] 전진호, “시계열데이터의 모델기반 클러스터링을 통한 예측모델 결정에 관한 연구”, 단국대학교 박사학위 논문, 2007.

[4] M. Siddiqi, J. Gordon and W. Moore.,"Fast State Discovery for HMM Moel Selection and Learning," In Proc. Int'l Conference on Artificial Intelligence and Statistics, 2007.

[5] 조영희, “시계열데이터의 의미기반 패턴매칭과 예측에 관한 연구”, 단국대학교 박사학위논문, 2009.