

A Smoothing Method for Stock Price Prediction with Hidden Markov Models

Soonho Lee¹⁾ · Chang Hyuck Oh²⁾

Abstract

In this paper, we propose a smoothing and thus noise-reducing method of data sequences for stock price prediction with hidden Markov models, HMMs. The suggested method just uses simple moving average. A proper average size is obtained from forecasting experiments with stock prices of bank sector of Korean Exchange. Forecasting method with HMM and moving average smoothing is compared with a conventional method.

Keywords : Hidden Markov Model, Moving Average, Smoothing Data, Stock Price, Training and Prediction

1. 머리말

일반적으로 주식가격은 높은 상호 의존성을 가지는 경제, 사회, 정치, 심리 등의 복합적 요소의 영향을 받아 결정되므로 주식가격 추세는 복잡하고 변동성이 커 주식가격을 분석하고 예측하는 것은 매우 어려운 일로 알려져 있다. 이러한 어려움을 극복하기 위한 방안으로 주식가격의 분석과 예측을 위하여 시계열모형, 회귀모형 등의 다양한 확률모형이 적용되어 왔으며 최근에는 인공신경망, 퍼지시스템, 서포트벡터머신 등의 기법뿐만 아니라 음성인식과 영상인식에 널리 사용되고 있는 숨은마코프모형(HMM)도 사용되고 있다.

음성인식 분야에서 활용이 시작된 숨은마코프모형은 유전공학을 포함한 다양한 분야에서 시계열의 패턴을 인식하는 도구로 널리 이용되고 있으며(Rabiner & Juang, 1993; Boys 외, 2000), 금융 관련 데이터 분석에서도 사용이 활발해지고 있다. Knab 외(2002)는 HMM을 이용하여 은행고객을 군집화하였으며, Crowder 외(2006)는 HMM을 이용하여 채권 포트폴리오에서 디폴트 상태의 발생을 모형화하였다. Banachewicz 외(2006)은 공변량 HMM을 이용하여 채무이행 불능 상태를 분석하는 방법을 연구하

1) 경북 경산시 대동 영남대학교 통계학과 석사과정

E-mail : yangpaggang@ynu.ac.kr

2) 교신저자 : 경북 경산시 대동 영남대학교 통계학과 교수

E-mail : choh@yu.ac.kr

였다. Hassan & Nath(2005)와 Zhang(2004)는 HMM을 이용하는 주식가격 예측방법을 제시하였고, Hassan 외(2007)은 유전알고리즘과 혼성 인공신경망/HMM을 적용하고, HMM 부분에서 가중평균을 사용하여 주식가격을 예측하는 방법을 연구하였다.

Hassan & Nath(2005)에서는 과거 주식가격의 일별 시가, 고가, 저가, 종가의 네 가지 가격의 패턴을 HMM으로 분석하여 종가를 예측하고, Hassan 외(2007)에서는 이들 네 개의 가격에 대하여, 유전알고리즘은 HMM의 모형모수의 초기화를 위하여, 인공신경망은 4개 가격의 관측열을 변환하여 추세성이 제거된 관측열을 얻기 위하여, HMM은 가중평균을 이용한 예측방법을 위하여 사용되었다. 본 논문에서는 HMM의 모수의 추정에 사용되는 훈련용 데이터에 대하여 시가, 고가, 저가, 종가의 각각을 이동평균으로 평활하여 Hassan & Nath(2005)의 방법으로 종가를 예측할 때의 예측 정확도의 향상 효과를 조사한다. 평활은 데이터에 포함된 소음을 제거해 주는 효과를 가지며 이로 인해 예측의 정확성을 높일 것으로 기대된다. 실제 주식 데이터 열에 관한 분석을 통해 적정한 이동평균의 크기를 알아본다. 평활한 데이터 열을 Hassan 외(2007)에서 가중평균을 이용한 HMM 모형에 대하여 적용한다. 또한 제시된 방법과 ARIMA모형에 의한 예측방법의 비교도 이루어진다. 2절에서는 숨은마코프모형의 소개와 예측방법을 다루며, 3절에서는 한국 증권선물거래소의 유가증권 시장에 상장된 7개의 은행과, 은행과 관련된 3개의 금융지주회사의 주식가격에 대한 실험과 결과를, 마지막 절에서는 토의와 결론을 다룬다.

2. 숨은마코프모형과 주식가격 예측방법

2.1 숨은마코프모형

숨은마코프모형은 이중 확률과정 $\{(X_t, Y_t), t = 1, 2, \dots\}$ 으로 정의된다. 관측되지 않는 마코프연쇄 $\{X_t\}$ 는 상태공간 $Q = \{1, \dots, N\}$ 에 대하여 전이확률 행렬 $A = (a_{ij})_{N \times N}$ 과 초기확률분포 $\Pi = (\pi_1, \dots, \pi_N)$ 를 가진다고 하자. 여기서 $a_{ij} = P(X_{t+1} = j | X_t = i)$ 이며 $\pi_i = P(X_1 = i)$ 이다. 관측가능 확률과정 $\{Y_t\}$ 은 마코프연쇄 $\{X_t\}$ 가 주어진 조건에서 독립이며, 각 t 에 대하여 Y_t 의 조건분포는 오직 X_t 에만 의존한다고 가정한다. 즉, 각 시점에서의 관측값의 확률분포는 그 시점에서의 마코프연쇄의 상태에만 의존하며 관측값 사이는 독립으로 가정한다. 관측값이 연속형 분포를 따를 때 연속형 숨은마코프모형이라고 부른다. 마코프연쇄의 상태가 $X_t = j$ 로 주어진 경우 관측치 Y_t 의 조건부분포는 성분의 개수가 M 인 혼합정규분포를 따르는 경우를 고려하면 조건부 확률밀도 함수 $b_j(y_t)$ 는 다음으로 주어진다.

$$b_j(y_t) = \sum_{m=1}^M w_{jm} \mathbf{n}(y_t; \mu_{jm}, \Sigma_{jm}). \quad (1)$$

여기서 $w_{jm} > 0$ 이며 각 j 에 대하여 $w_{j1} + \dots + w_{jM} = 1$ 이다. 또한 $\mathbf{n}(\cdot; \mu, \Sigma)$ 은 평균 벡터가 μ , 분산공분산 행렬이 Σ 인 다변량 정규확률밀도함수를 나타낸다. 관측치 분

포에 관한 모수를 $B = (w_{11}, \dots, w_{NM}, \mu_{11}, \dots, \mu_{NM}, \Sigma_{11}, \dots, \Sigma_{NM})$ 로 나타내고 이와 같은 숨은마코프모형을 $\lambda = (A, B, \Pi)$ 로 나타내기로 한다.

숨은마코프모형에는 ‘평가문제’, ‘해독문제’, 그리고 ‘추정문제’의 세 개의 규범적 문제가 있다. 크기가 T 인 관측벡터열 $y = \{y_1, \dots, y_T\}$ 가 주어졌다고 하자. 이 때 평가 문제는 모형 λ 가 주어진 경우 관측값 y 의 확률 $P(Y=y|\lambda)$ 를 효율적으로 계산하는 것에 관한 것이며 전진절차 혹은 후진절차의 해법이 있다. 추정문제는 주어진 모형과 관측열에 대하여 모형을 개정하는 문제이며 이를 위한 반복 해법인 Baum-Welch 알고리즘은 EM 알고리즘의 하나이며 국소 최우주정치로 수렴함이 알려져 있다. 숨은마코프모형에서의 전진절차 또는 후진절차와 Baum-Welch 알고리즘에 관한 자세한 내용은 Rabiner(1989)를 참조하기 바란다.

2.2 주식가격 예측방법

확률과정 $Y_1, Y_2, \dots, Y_T, Y_{T+1}, \dots$ 은 $N=4, M=3$ 인 숨은마코프모형 $\lambda = (A, B, \Pi)$ 의 관측가능 확률과정이라고 하자. 각 t 에 대응되는 확률변수 Y_t 는 $d=4$ 개의 성분을 가진 벡터 $Y_t = (Y_{t1}, Y_{t2}, Y_{t3}, Y_{t4})$ 로 주어진다고 하자. 여기서는 기호를 간단하게 하기 위하여 Y_{t4} 를 C_t 로 나타내기로 한다.

주어진 관측열 y_1, \dots, y_T 를 훈련데이터집합이라고 하자. 훈련데이터집합과 초기모형을 이용하여 모수를 추정하여 모형을 개정한다. HMM의 모수를 초기화하기 위하여 K-평균 알고리즘을 사용하며 모수의 추정에는 Baum-Welch 알고리즘을 사용한다. 날 $h (\geq T)$ 에 대하여 관측치 y_1, \dots, y_h 가 주어진 경우에 $h+1$ 번째 날에 대한 종가 C_{h+1} 는 Γ_{h+1} 로 예측된다. Hassan & Nath(2005)는 다음의 예측 방법을 제시하였다.

예측방법 I:

단계 1. $t = 1, \dots, h-1$ 에 대하여 날 h 의 로그우도와 차이가 최저인 날 τ 를 구한다.

$$\tau = \arg \min_{1 \leq t \leq h-1} |\log P(y_h|\lambda) - \log P(y_t|\lambda)|. \quad (2)$$

단계 2. 날 $h+1$ 의 종가 C_{h+1} 에 대한 예측 Γ_{h+1} 은 다음 식으로 주어진다.

$$\Gamma_{h+1} = C_h + (C_{\tau+1} - C_\tau). \quad (3)$$

한편, Hassan 외(2007)은 예측방법 I을 개선하는 하나의 방법으로 HMM의 초기값 설정에는 유전알고리즘을, 입력변수 변환에는 인공신경망을, 그리고 HMM 예측에서의 개선으로는 가중평균을 사용하였다. HMM 부분의 가중평균을 사용하는 방법으로, 현재의 날과 비슷한 행태를 가지는 과거의 날들을 찾고 이들에 대하여 더 가까운 과거의 날에 더 큰 가중치를 부여하는 방법을 제시하였다. 다음은 HMM 예측에서 가중평균을 사용하는 절차이다.

예측방법 II:

단계 1. $t = 1, \dots, h-1$ 에 대하여 h 번째 날의 로그우도와 차이가 최저인 날 τ 를 식 (2)

를 이용하여 구한다.

단계 2. 날 h 와 τ 에 대한 로그우도의 차이를 $L_{h\tau} = |\log P(y_h|\lambda) - \log P(y_\tau|\lambda)|$ 라고 틀 때, $t=1, \dots, h-1$ 에 대하여 로그우도와 차이가 적은 날의 집합을 구한다. 즉,

$$\Lambda_{h\tau} = \{t : |\log P(y_h|\lambda) - \log P(y_t|\lambda)| \leq L_{h\tau} + \epsilon, t = 1, \dots, h-1\}. \quad (4)$$

여기서 $\epsilon \geq 0$ 이며 그 크기는 데이터에 따라 적당하게 정해진다.

단계 3. 날 $h+1$ 의 종가 C_{h+1} 에 대한 예측 Γ_{h+1} 은 다음 식으로 주어진다.

$$\Gamma_{h+1} = C_h + \sum_{i=1}^g w_{\tau(i)} (C_{\tau(i+1)} - C_{\tau(i)}). \quad (5)$$

단, $\Lambda_{h\tau}$ 의 모든 원소는 $\tau(1) < \tau(2) < \dots < \tau(g)$ 이며, $w_{\tau(1)}, \dots, w_{\tau(g)}$ 는 합이 1인 양 수 가중치이며 다음으로 주어진다.

$$w_{\tau(i)} = \frac{v_{\tau(i)}}{v_{\tau(1)} + \dots + v_{\tau(g)}}, \quad i = 1, \dots, g. \quad (6)$$

단, $v_{\tau(i)} = \exp\{(h+1) - \tau(i) + 1\}^{-1}$ 로 주어진다.

예측방법 II는 Hassan 외(2007)에서 HMM의 부분만을 사용한 것이며 식 (4)에서 $\epsilon = 0$ 으로 하는 경우는 예측방법 I로 축소된다. 한편, 예측방법 I과 II의 식 (2), (4)에서 로그우도 계산은 HMM에 대한 전진 혹은 후진절차에 의해 구해진다. 식 (6)의 가중치는 가까운 과거에서부터 먼 과거로 갈 때 지수적 비율로 감소하도록 정해진다.

2.3 이동평균을 이용한 시계열의 소음 제거

주어진 시계열에 대하여 이동평균하여 얻은 시계열은 원래 시계열을 평활하는 효과를 가져온다. 이동평균은 단기간의 변동을 평활하여 추세성이나 계절성의 효과를 돋보이게 한다. 즉, 이동평균은 소음효과를 제거하여 주된 효과를 살펴보는 방법이다. 본 논문에서는 중심이동평균으로 관측열을 평활하는 방법을 사용한다. 관측열 y_1, \dots, y_T 이 주어졌을 때의 이동평균기간이 $2l+1$ 인 중심이동평균 열은 다음 식으로 주어진다.

$$\hat{y}_t = \frac{y_{t-l} + \dots + y_t + \dots + y_{t+l}}{2l+1}, \quad t = 1, \dots, T. \quad (7)$$

단, $l_t = \max \{a : t-a > 0, t+a < T, \text{and } a \leq l\}$ 이며 관측열의 처음 부분과 마지막 부분에서 평균크기를 조절하기 위한 것이다. 또한 l 과 a 는 음이 아닌 정수이다.

3. 실험 및 결과

3.1 실험 데이터

HMM을 이용한 주식가격 예측 방법을 한국 증권선물거래소의 유가증권 시장에 적용하기 위하여 이 시장에 상장된 은행 및 시중은행과 관련된 금융지주회사의 주식가격 데이터를 사용하였다. 분석에 사용된 은행은 대구은행, 제주은행, 한국외환은행, 국민은행, 부산은행, 중소기업은행, 전북은행이고 금융지주회사는 신한금융지주회사, 우리금융지주회사, 하나금융지주회사이다. 표 1은 훈련 및 검증 자료집합에 관한 정보이다. 분석에 사용되는 변수는 Hassan & Nath(2005)에서와 같이 주식시장에서의 매일의 시가, 고가, 저가, 종가의 4가지 가격이다. 예측되는 값은 다음 날의 종가이다.

표 1. 훈련 및 검증 데이터에 관한 정보

훈련 데이터		검증 데이터	
시작	끝	시작	끝
2005. 1. 3	2006. 8. 2	2006. 8. 3	2006. 11. 29

3.2 HMM모형의 설정

HMM의 상태의 개수는 관측변수벡터의 차원과 같은 4로 정한다. 상태이동에 관하여는 좌에서 우로만 이동하는 Bakis 모형 즉, 전이확률에서 $a_{ij} = 0$, $i > j$ 인 모형을 가정한다. 관측변수에 대하여는 계산의 편의성을 위하여 각 성분변수인 시가, 고가, 저가, 종가는 서로 독립이라고 가정하며 각각은 일변량 혼합정규분포를 따른다고 가정한다. 식 (1)에서의 혼합정규분포의 성분의 개수는 상태의 개수와 시가, 고가, 저가, 종가 각각의 훈련데이터 열에 대한 히스토그램과 이를 가격의 두 개씩을 조합한 경우에 대한 히스토그램을 바탕으로 인지적으로 판단하여 $M=3$ 으로 정하였다. Baum Welch 알고리즘을 이용하여 모수를 개정하며, 모수의 초기화를 위하여는 K-평균 군집화 알고리즘을 사용한다. 우도의 계산은 HMM의 평가문제를 위한 전진 또는 후진 절차를 사용한다. 한편 식 (4)의 우도차이의 값은 $\epsilon = 0.01$ 로 하였다.

3.3 효율성 측도

예측방법을 비교하기 위한 효율성 측도로서 평균절대백분율오차 MAPE를 사용한다. 검증을 위한 관측된 주식 가격열 C_{T+1}, \dots, C_{T+r} , 과 이에 대응되는 예측된 주식가격 열 $\Gamma_{T+1}, \dots, \Gamma_{T+r}$ 에 대하여 MAPE는 다음으로 주어진다.

$$\text{MAPE} = \frac{1}{r} \sum_{i=1}^r \left| \frac{C_{T+i} - \Gamma_{T+i}}{C_{T+i}} \right| \times 100.$$

3.4 결과

표 2. 예측방법 I의 회사별 이동평균크기에 따른 평균절대백분률오차

회사	이동평균크기($2l+1$)								
	1	3	5	7	9	11	13	15	31
대구은행	2.18	1.72	1.86	1.63	1.66	1.55	1.59	1.55	1.97
제주은행	1.66	1.57	1.17	1.34	1.28	1.15	1.05	1.17	1.66
한국외환은행	2.26	1.64	1.60	1.62	1.55	1.55	1.52	1.49	2.36
국민은행	1.79	1.46	1.57	1.25	1.29	1.30	1.46	1.45	1.49
신한금융지주	1.63	0.95	0.98	1.0	0.98	0.99	0.93	0.94	1.21
우리금융지주	1.59	1.63	1.58	1.36	1.38	1.33	1.24	1.27	1.90
부산은행	1.92	1.32	1.34	1.27	1.52	1.27	1.31	1.22	1.98
중소기업은행	1.86	1.56	1.34	1.10	1.20	1.17	1.22	1.17	2.07
전북은행	2.49	1.83	1.51	1.27	1.71	1.48	1.49	1.43	2.16
하나금융지주	1.73	1.25	1.36	1.32	1.25	1.31	1.30	1.20	1.28

표 3. 예측방법 II의 회사별 이동평균크기에 따른 평균절대백분률오차

회사	이동평균크기($2l+1$)								
	1	3	5	7	9	11	13	15	31
대구은행	1.47	1.48	1.36	1.29	1.29	1.28	1.23	1.22	1.46
제주은행	1.01	0.99	0.97	0.97	0.98	0.97	0.97	0.97	1.0
한국외환은행	1.51	1.43	1.44	1.44	1.44	1.45	1.46	1.46	1.51
국민은행	1.24	1.27	1.16	1.11	1.12	1.17	1.17	1.20	1.22
신한금융지주	0.87	0.85	0.83	0.89	0.85	0.85	0.85	0.84	0.87
우리금융지주	1.16	1.16	1.32	0.91	0.84	0.90	0.88	0.89	1.15
부산은행	1.21	1.17	1.16	1.13	1.13	1.13	1.14	1.15	1.21
중소기업은행	1.08	1.02	1.0	0.87	0.80	0.78	0.77	0.80	1.08
전북은행	1.20	1.20	1.19	1.18	1.18	1.19	1.18	1.18	1.22
하나금융지주	1.15	1.13	1.14	1.13	1.12	1.13	1.12	1.13	1.13

표 4. 예측방법 II와 ARIMA모형에 대한 평균절대백분률오차

회사	2l + 1 = 9		ARIMA
	2l + 1 = 9	ARIMA	
대구은행	1.29		1.43
제주은행	0.98		0.95
한국외환은행	1.44		1.47
국민은행	1.12		1.17
신한금융지주	0.85		0.85
우리금융지주	0.84		1.21
부산은행	1.13		1.15
중소기업은행	0.80		1.04
전북은행	1.18		4.60
하나금융지주	1.12		1.13

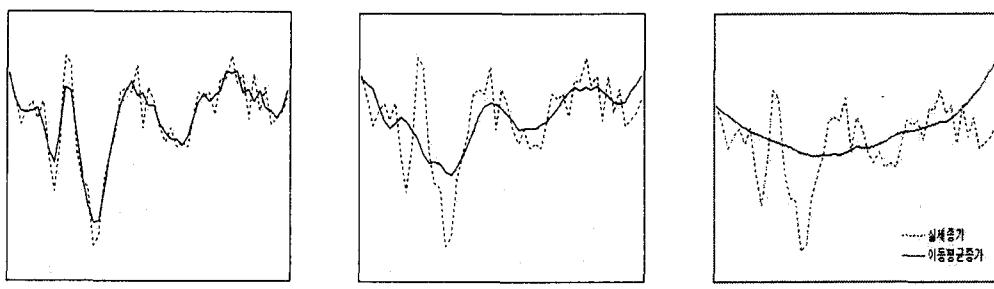
표 2는 주식거래소에 상장된 7개의 은행과, 은행과 관련된 3개의 금융지주회사의 주식에 대한 예측방법 I에 의한 실험결과이다. 이들 두 개의 표에서는 식 (7)의 여러 가지 이동평균의 크기 $2l+1$ 에 대하여 각 회사 주식의 예측에 대한 MAPE를 구하였다. 이동평균의 크기가 1인 경우는 원자료를 이용한 예측에 대한 것이다. 전체적으로 이동평균의 크기가 커짐에 따라 MAPE는 줄어들다가 다시 커지는 경향을 나타내고 있으며 대체적으로 이동평균의 크기가 7에서 15 사이인 경우에 MAPE가 가장 작은 값의 경향을 나타내고 있으며 31 이상은 커지는 경향을 나타낸다. 이는 이동평균의 크기가 지나치게 크면 그림 1에서 보는 것과 같이 주어진 시계열을 지나치게 평활화하는데서 기인한다고 판단된다.

한편, 표 3은 주식거래소에 상장된 7개의 은행과, 은행과 관련된 3개의 금융지주회사의 주식가격에 대한 예측방법 II에 대한 실험 결과이며 예측방법 I과 유사한 결과를 보이고 있고 $2l+1 = 9$ 일 때 MAPE가 가장 작아 보이는 경향을 나타낸다. 다만, MAPE의 감소 정도가 예측방법 I에 비해 크지 않음을 볼 수 있다. 이는 예측방법 II에서는 다음 날의 주식 가격을 예측할 때 과거 데이터 중 시가, 고가, 저가, 종가의 형태가 유사한 날들에 대하여 종가의 차이를 가중평균하여 노이즈를 제거하는 효과를 가지고 있기 때문으로 보인다. 한편 전통적인 ARIMA(p, q, r)모형과 예측방법 II로 예측한 것에 대한 MAPE의 결과를 표 4에 실었다. ARIMA모형은 먼저 p, q, r (p =자동화 모수, q =이동평균 모수, r =누적개수)의 값을 구한 후 각 주식 가격열에 대하여 ARIMA(p, q, r)을 적용하여 예측한다. 표 4의 결과를 보면 MAPE로 비교할 때 이동평균크기 $2l+1 = 9$ 인 방법의 예측력이 ARIMA모형에 의한 예측방법의 예측력에 비해 전체적으로 우수함이 나타났다. 특히 전북은행의 경우에는 ARIMA모형에 의한 예측에서의 MAPE가 예측방법 II에 의한 것보다 월등히 크게 나타남을 알 수 있다.

4. 토의 및 결론

이 논문에서는 시계열 예측을 위한 숨은마코프모형의 적용에서 입력 데이터 열의 노이즈를 제거하는 방법으로 데이터 열을 이동평균하여 평활 데이터 열을 생성하였다. Hassan & Nath(2005)의 모형인 예측모형 I과 Hassan 외(2007)의 모형에서 HMM 부분만을 적용하는 예측모형 II에 대하여 평활된 데이터를 적용하는 경우와 그렇지 않은 경우에 대하여 MAPE를 이용하여 예측력 정확도를 비교하였다. 비교 결과, 이동평균한 평활 데이터 열의 예측 정확도가 원 데이터 열의 그것보다 높은 것으로 나타났다. 주식거래소에 상장된 7개의 은행과, 은행과 관련된 3개의 금융지주회사의 주식 가격에 대하여 여러 가지 크기의 이동평균에 대한 예측정확도를 구한 결과 이동평균 크기는 대체로 9가 적절한 것으로 나타났다.

Hassan 외(2007)은 Hassan & Nath(2005)의 HMM을 이용한 주식가격 예측방법의 예측력을 높이는 방법으로 유전알고리즘과 인공신경망을 혼성 사용하는 방법을 제시



$2l+1 = 3$ $2l+1 = 9$ $2l+1 = 31$
 그림 1. 대구은행 종가에 대하여 이동평균크기($2l+1$)와 평활된 시계열

하고, 그들의 방법이 전통적인 시계열 방법인 ARIMA모형과 비교하여 예측력이 비슷하며, ARIMA모형을 사용할 때 모형요소를 설정하는 등의 어려움이 없으므로 자신들이 제시한 방법의 우월성을 주장하였다. 그러나 본 논문에서의 예측방법 II는 Hassan 외(2007)에서의 유전알고리즘 및 인공신경망을 적용하는 대신에 데이터 열에 대하여 단순히 이동평균을 통한 평활하는 방법을 제시하였다. 제시한 방법이 유전알고리즘이나 인공신경망을 이용하는 것에 비하여 단순하며 직관적인 이동평균을 사용하였음에도 불구하고 예측력은 ARIMA모형에 의한 방법에 필적하거나 오히려 나은 것으로 나타남을 보였으며 이는 Hassan 외(2007)에서 얻은 결과보다 더 나은 결과이다.

감사의 글

본 논문의 내용의 개선과 향상을 위해 여러 가지 제안과 지적을 해 주신 심사위원께 감사드립니다.

참고문헌

1. Achewicz K., Vaart, A. V. D., and Lucas, A. (2006). Modeling portfolio defaults using hidden Markov models with covariates. *Tinbergen Institute Discussion Papers with number 06-094/2*.
2. Boys, R. J., Henderson, D. A., and Wilkinson, D. J. (2000). Detecting homogeneous segments in DNA sequences by using hidden Markov models. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 49, 269-285.
3. Crowder, M., Davis, M., and Giampieri, G. (2005). A hidden Markov model of default interaction. *Quantitative Finance*, 5, 27 - 34.
4. Hassan, M. R. and Nath, G. B. (2005). Stock market forecasting using hidden Markov model. *Proceedings of the 5th International Conference on Intelligent Systems Design and Applications*, 192-196.
5. Hassan, M. R., Nath, G. B., and Kirley, M. (2007). A fusion model of

- HMM, ANN and GA for stock market forecasting. *Expert Systems with Applications: An International Journal*, 33, 171-180.
- 6. Knab, B., Schliep, A., Steckemetz, B., and Wichern, B. (2002). Model-based clustering with hidden Markov models and its application to financial time-series data. *Technical Report*.
 - 7. Landen, C. (2000). Bond pricing in a hidden Markov model of the short rate. *Finance and Stochastics*, 4, 371-389.
 - 8. Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE*, 77, 257-286.
 - 9. Rabiner, L. R. and Juang, L. Y. (1993). Fundamentals of speech recognition. *Prentice-Hall Inc.*
 - 10. Zhang, Y. (2004). Prediction of financial time series with hidden Markov models. *Master's thesis*, Simon Fraser University.

[2007년 11월 접수, 2007년 11월 채택]