Stock Price Prediction using Hidden Markov Model

Nguyet Nguyen
Youngstown State University, Ohio, USA
Email: ntnguyen01@ysu.edu

March 21, 2016

Abstract

The stock market is an important indicator which reflects economic strengths and weaknesses. Stock trading will have great returns if the economy is strongly growing, but it may have negative returns if the economics is depressing. An accurate stock price prediction is one significant key to be successful in stock trading. The hidden Markov model (HMM) is a signal prediction model which has been used to predict economic regimes and stock prices. In this paper we use HMM to predict the daily stock price of three stocks: Apple, Google and Facebook. Our approach is significantly different from the previous studies because we do not only base off historical stock prices, but also on the chances of the likelihoods of the model of the most recent data set with the historical data set. We first use the Akaike information criterion and Bayesian information criterion to compare HMM performances using numbers of states from two to four. We then use the models to predict close prices of these three stocks using both a single observation data and multiple observation data. The results showed that HMM using multiple observation data worked better in stock price predictions.

Keyword: Hidden Markov Model, stock prices, observations, states, regimes, economics, predictions, AIC, BIC, likelihood.

1 Introduction

Stock traders always wish to buy a stock at a low price and sell it will a higher price. However, when the best time is to buy or sell a stock is a challenging question. Stock investments can have a huge return or a big loss due to the high volatilities of stock prices. There are many models that were used to predict stock executions such as the “exponential moving average" (EMA) or the “head and shoulders" methods. Recently, researchers have applied the hidden Markov model for forecasting stock prices. Hassan and Nath [2] used HMM to forecast stock price for interrelated markets. Kritzman, Page, and Turkington [7] applied HMM with two states to predict regimes in market turbulence, inflation, and industrial production index. Guidolin, M. and Timmermann [14] used HMM with four states and multiple observations to study asset allocation decisions based on regime switching in asset returns. Ang and Bekaert [16] applied regime shift model (another name of HMM) for international asset allocation. Nguyen [17] used HMM with both single and multiple observations to forecast economic regimes and stock prices. B. Nobakht, C.-E. Joseph and B. Loni [8] implemented HMM by using multiple observation data (open, close, low, high) prices of a stock to predict its close prices. In our previous work [19], we used HMM for single observation data to predict regimes of some economic indicators and make stock selections based on the performances of these stocks during the predicted regimes. In this study, we explore the new approach of HMM in stock price prediction by adding the difference of likelihoods of the model for recent data set and for the past data set. We use both HMM for multiple independent variables and for a single variable to forecast stock prices. These three popular stocks: Apple Inc., Alphabet Inc., and Facebook, Inc., were chosen to implement the model. We limit numbers of states of the HMM to a maximum of four states and use the Akaike information criterion test to compare the model's performances with two, three, or four states. The prediction process is based on the work of Hassan and Nath [2]. The authors used HMM with the four observations: close, open, high, and low price of some airline stocks to predict their future close price using 4 states. They used HMM to find a day in the past that was similar to the recent day and used the price change in that day and price of the recent day to predict a future close price. Our approach is different from their work in three modifications. The first difference is that we use the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to test the HMM's performances with a number of states from two to four. The second modification is that we use both single observation (close price) and multiple observations (open, close, high, low prices) to predict future close price and compare the results. The third difference is that we do not only consider the stock prices of a “similar day" in the past, but also the sign of difference of the likelihoods of the model of the recent day, and the “similar" day in the past to estimate future stock prices.

This paper is organized as follows: Section 2 gives a brief introduction about the HMM and its three main problems and corresponding algorithms. Section 3 describes the HMM model selections and data collections for stock price prediction. Section 4 presents the experiment results and analyzes their performances, and Section 5 gives conclusions.

2 A brief introduction of the Hidden Markov Model

The Hidden Markov Model, HMM, is a signal detection model which was introduced in 1966. The assumption of the model is that observation sequences were derived by a hidden state sequence which is discrete data and satisfies the first order of a Markov process. HMM was developed from a model for a single observation to a model for multiple observations. The applications of HMM also were expanded widely in many fields such as speech recognition, biomathematics, and financial mathematics. In our previous paper [19], we described HMM for one observation, its algorithms, and applications in stock selection. In this paper, we present HMM for multiple observations and its corresponding algorithms. We assume that the observations data are independent and have the same length. The basic elements of a HMM for multiple observations are:

- Number of observations, L
- Length of observation data, T
- Numbers of states, N
- Number of symbols per state, M
- Observation sequence, $O = \{O_t^{(l)}, t = 1, 2, \ldots, T, l = 1, 2, \ldots, L\}$
- Hidden state sequence, $Q = \{q_t, t = 1, 2, \ldots, T\}$
- Possible values of each state, $\{S_i, i = 1, 2, \ldots, N\}$
- Possible symbols per state, $\{v_k, k = 1, 2, \ldots, M\}$
- Transition matrix, $A = (a_{ij})$, where $a_{ij} = P(q_t = S_j | q_{t-1} = S_i),\ i, j = 1, 2, \ldots, N$
- Vector of initial probability of being in state (regime) S_i at time $t = 1$, $p = (p_i)$, where $p_i = P(q_1 = S_i),\ i = 1, 2, \ldots, N$
- Observation probability matrix, $B = \{b_i(k)\}$, where

$$b_i(k) \equiv b_i(O_t = v_k) \equiv P(O_t = v_k | q_t = S_i),\ i = 1, 2, ..., N,\ k = 1, 2, \ldots, M.$$

In summary, the parameters of a HMM are the matrices A and$B$, and the vector p. For convenience, we use a compact notation for the parameters, given by

$$\lambda \equiv \{A, B, p\}.$$

If the observation probability assumes the Gaussian distribution then $b_i(k) = b_i(O_t = v_k) = \mathcal{N}(v_k, \mu_i, \sigma_i)$, where μ_i and σ_i are the mean and variance of the distribution corresponding to the state S_i, respectively, and $\mathcal{N}$ is Gaussian density function. For convenience, we write $b_i(O_t = v_k)$ as $b_i(O_t)$. Then, the parameters of HMM are

$$\lambda \equiv \{A, \mu, \sigma, p\},$$

where μ and σ are vectors of means and variances of the Gaussian distributions, respectively. With the assumption that the observations are independent, the probability of observation, denoted by $P(O|\lambda)$, is

$$P(O|\lambda) = \prod_{l=1}^{L} P(O^{(l)}|\lambda).$$

There are three main problems of the HMM:

1. Given the observation data and the model parameters, compute the probabilities of the observations.
2. Given the observation data and the model parameters, find the best corresponding state sequence.
3. Given the observation calibrate HMM parameters.

The first problem can be solved by using the forward or backward algorithm [9, 10], the second problem was solved by using the Viterbi algorithm [12, 13], and the Baum Welch algorithm [11] was developed to solve the last problem. In the paper, we only use the algorithms to solve the first and the last problem. We first use Baum Welch algorithm to calibrate parameters for the model and forward algorithm to calculate the probability of observation to predict trending signals for stocks. In this section, we introduce the forward algorithm and the Baum-Welch algorithm for HMM with multiple observations. These algorithms are written based on [9–12, 18]. The backward algorithm [4] for multiple observation is presented in the appendix.

2.1 Forward algorithm

We define the joint probability function as

$$\alpha_t^{(l)}(i) = P(O_1^{(l)}, O_2^{(l)}, ..., O_t^{(l)}, q_t = S_i|\lambda), t = 1, 2, ..., T \textit{ and } l = 1, 2, ..., L,$$

then we calculate $\alpha_t^{(l)}(i)$ recursively. The probability of observation $P(O^{(l)}|\lambda)$ is just the sum of the $\alpha_T^{(l)}(i)'s$.

2.2 Baum-Welch algorithm

Baum-Welch algorithm is an algorithm to calibrate parameters for the HMM given the observation data. The algorithm was introduced in 1970 [5] to estimate the parameters of HMM for a single observation. Then, in 1983, the algorithm was extended to calibrate

Algorithm 1 : The forward algorithm

1. Initialization $P(O|\lambda) = 1$
2. For $l = 1, 2, ..., L$ do
   (a) Initialization: for i=1,2,..., N
   $$\alpha_1^{(l)}(i) = p_i b_i(O_1^{(l)}).$$
   (b) Recursion: for $t = 2, 3, \dots, T$, and for $j = 1, 2, \dots, N$, compute
   $$\alpha_t^{(l)}(j) = \left[\sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij}\right] b_j(O_t^{(l)}).$$
   (c) Calculate:
   $$P(O^{(l)}|\lambda) = \sum_{i=1}^{N} \alpha_T^{(l)}(i).$$
   (d) Update:
   $$P(O|\lambda) = P(O|\lambda) * P(O^{(l)}|\lambda).$$
3. Output: $P(O|\lambda)$.

HMM's parameters for multiple independent observations, [4]. In 2000, the algorithm was developed for multiple observations without the assumption of independence of the observations, [6]. In this paper we use HMM for independent observations, so we will introduce the Baum-Welch algorithm for this case. The Baum-Welch method, or the expectation modification (EM) method, is used to find a local maximizer ,λ^*, of the probability function $P(O|\lambda)$.

In order to describe the procedure, we define the conditional probability

$$\beta_t^{(l)}(i) = P(O_{t+1}^{(l)}, O_{t+2}^{(l)}, .., O_T^{(l)}|q_t = S_i, \lambda),$$

for $i = 1, ..., N, l = 1, 2, ..., L$. Obviously, for $i = 1, 2, ..., N$ $\beta_T^{(l)}(i) = 1$ and we have the following backward recursive:

$$\beta_t^{(l)}(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}^{(l)}) \beta_{t+1}^{(l)}(j), \ t = T-1, \ T-2, ..., 1.$$

We then defined $\gamma_t^{(l)}(i)$, the probability of being in state S_i at time t of the observation $O^{(l)}, l = 1, 2, ..., L$ as:

$$\gamma_t^{(l)}(i) = P(q_t = S_i|O^{(l)}, \lambda) = \frac{\alpha_t^{(l)}(i)\beta_t^{(l)}(i)}{P(O^{(l)}|\lambda)} = \frac{\alpha_t^{(l)}(i)\beta_t^{(l)}(i)}{\sum_{i=1}^{N} \alpha_t^{(l)}(i)\beta_t^{(l)}(i)}.$$

The probability of being in state S_i at time t and state S_j at time $t+1$ of the observation $O^{(l)}, l = 1, 2, ..., L$ as:

$$\xi_t^{(l)}(i, j) = P(q_t = S_i, q_{t+1} = S_j|O^{(l)}, \lambda) = \frac{\alpha_t^{(l)}(i) a_{ij} b_j(O_{t+1}^{(l)}) \beta_{t+1}^{(l)}(j)}{P(O^{(l)}, \lambda)}.$$

Clearly,

$$\gamma_t^{(l)}(i) = \sum_{j=1}^{N} \xi_t^{(l)}(i, j).$$

Note that the parameter λ^* was updated in Step 2 of the Baum-Welch algorithm to maximize the function $P(O|\lambda)$ so we will have $\triangle = P(O, \lambda^*) - P(O, \lambda) > 0$.

If the observation probability $b_i(k)^*$, defined in Section 2, is Gaussian, we will use the following formula to update the model parameter, $\lambda \equiv \{A, \mu, \sigma, p\}$

$$\mu_i^* = \frac{\sum_{l=1}^{L} \sum_{t=1}^{T-1} \gamma_t^{(l)}(i) O_t^{(l)}}{\sum_{l=1}^{L} \sum_{t=1}^{T-1} \gamma_t^{(l)}(i)}$$

$$\sigma_i^* = \frac{\sum_{l=1}^{L} \sum_{t=1}^{T} \gamma_t^{(l)}(i)(O_t^{(l)} - \mu_i)(O_t^{(l)} - \mu_i)'}{\sum_{l=1}^{L} \sum_{t=1}^{T} \gamma_t(i)}.$$

Algorithm 2 : Baum-Welch for L independent observations $O = (O^{(1)}, O^{(2)}, ..., O^{(L)})$ with the same length T

1. Initialization: input parameters λ, the tolerance tol, and a real number $\triangle$

2. Repeat until $\triangle < tol$

    - Calculate $P(O, \lambda) = \Pi_{l=1}^{L} P(O^{(l)}|\lambda)$ using the forward algorithm (**??**)
    - Calculate new parameters $\lambda^* = \{A^*, B^*, p^*\}$, for $1 \leq i \leq N$

$$p_i^* = \frac{1}{L}\sum_{l=1}^{L} \gamma_1^{(l)}(i)$$

$$a_{ij}^* = \frac{\sum_{l=1}^{L}\sum_{t=1}^{T-1} \xi_t^{(l)}(i,j)}{\sum_{l=1}^{L}\sum_{t=1}^{T-1} \gamma_t^{(l)}(i)},\ 1 \leq j \leq N$$

$$b_i(k)^* = \frac{\sum_{l=1}^{L}\sum_{t=1}^{T} |_{O_t^{(l)}=v_k^{(l)}} \gamma_t^{(l)}(i)}{\sum_{l=1}^{L}\sum_{t=1}^{T} \gamma_t^{(l)}(i)},\ 1 \leq k \leq M$$

    - Calculate $\triangle = P(O, \lambda^*) - P(O, \lambda)$
    - Update

$$\lambda = \lambda^*.$$

3. Output: parameters λ.

3 Describe the model

The Hidden Markov Model have been widely used in the financial mathematics area to predict economic regimes ([7], [14], [16], [15], [17]) or predict stock prices ([2], [8], [17]). In this paper, we explore a new approach of HMM in predicting stock prices. In this section, we discuss how to use the Akaike information criterion, AIC, and the Bayesian information criterion, BIC, to test the HMM's performances with different numbers of states. We will then present how to use HMM to predict stock prices. First, we will describe preferred data and the AIC and BIC for HMM with a selected numbers of states.

3.1 Overview of Data Selections

We choose three stocks that are actively trading in the stock market to examine our model: Apple Inc. (AAPL), Alphabet Inc. (GOOGL), and Facebook Inc. (FB). The daily stock prices (open, low, high, close) of these stocks and information of these companies can be found from finance.yahoo.com. We use daily historical prices of these stocks from January 4, 2010 to October 30, 2015 in this paper.

3.2 Model Selection

Choosing a number of hidden states for the HMM is a critical task. We first use two common criteria: the AIC and the BIC to evaluate the performances of HMM with different numbers of states. The two criteria are suitable for HMM because in the model training algorithm, the Baum-Welch algorithm, the EM method was used to maximize the log likelihood of the model. We limit numbers of states from two to four to keep the model simple and feasible to stock prediction. The AIC and BIC are calculated using the following formulas respectively:

$$AIC = -2\ln(L) + 2k,$$

$$BIC = -2\ln(L) + k\ln(M),$$

where L is the likelihood function for the model, M is number of observation points, and k is the number of estimated parameters in the model. In this paper, we assume that the distribution corresponding with each hidden state is a Gaussian distribution, therefore the number of parameters, k, is formulated as $k = N^2 + 2N - 1$, where N is numbers of states used in the HMM.

To train HMM's parameters we use a historical observed data of a fixed length T,

$$O = \{O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)},\ t = 1, 2, ..., T\},$$

where $O^{(i)}$ with $i = 1, 2, 3$, or 4 represents the open, low, high or close price of a stock respectively. For the HMM with single observation, we use only close price data,

$$O = O_t,\ t = 1, 2, ..., T$$

where O_t is stock close price at time t. We run the model calibrations with a different time length, T, and see that the model works well for $T \geq 80$. In the results below, we use a blocks of $T = 100$ trading days of stock price data, O, to calibrate HMM's parameters and calculate the AIC and BIC numbers. Thus, the total number of observation points in each BIC calculation is $M = 400$ for four observation data and $M = 100$ for one observation data. For convenience, we do 100 calibrations for 100 blocks of data by moving the block of data forward. (We take off the price of the oldest day on the block and add the price of the following day to the recent day of the block). The training data set is from January 16, 2015 to October 30, 2015.

The first block of stock prices of 100 trading days from January 16, 2015 to June 6, 2015 was used to calibrate HMM's parameters and calculate corresponding AIC and BIC. Let $\mu^{(O)}$ and $\sigma^{(O)}$ are the mean and standard deviation of observation data, O, respectably. We choose initial parameters for the first calibration as follows:

$$
\begin{aligned}
A =&(a_{ij}),\ a_{ij} = \frac{1}{N}, \\
p =&(1, 0, .., 0), \\
\mu_i =&\mu^{(O)} + Z,\ Z \sim \mathcal{N}(0, 1), \\
\sigma_i =&\sigma^{(O)},
\end{aligned}
\tag{1}
$$

where $i,\ j = 1, .., N$ and $\mathcal{N}(0, 1)$ is the standard normal distribution.

The second block of 100 trading day data from January 17, 2015 to June 7, 2015 was used for the second calibration and so on. The HMM calibrated parameters from the current step is used as initial parameters for the next calibration. We continue the process until we get 100 calibrations. We plot the AICs and BICs of the 100 calibrations of these three stocks (AAPL, FB, and GOOGL) on Figures [1-3]. On Figures [1-3], the graph of AIC is located on the left and BIC is located on the right. The lower AIC or BIC is the better model calibration. However the Baum-Welch algorithm only finds a local maximizer of the likelihood function. Therefore we do not expect to have the same AICs or BICs if we run the calibration twice. The results on Figures [1-3] show that the calibration performances of models with different numbers of states are differed from one simulation to another. Based on the AIC results, the HMM with four states is outperformed the HMM with two or three states in most of the calibrations, but when the number of simulations (or calibrations) increase the AIC of these three models converged to the same number. Based on the BIC, the HMM with four states is the best candidate for Apple stock and HMM with three states is suitable for Google and Facebook stocks. We will use HMM with numbers of states from two to four to predict prices of the three stocks in the next section to see if the model selections in this section are consistent with the model performances in stock predictions.

3.3 Stock Price Prediction

In this section, we will use HMM to predict stock prices and compare the predictions with the real stock prices. We will forecast stock prices of GOOGL, APPL, and FB using

Figure 1: AIC (left) and BIC (right) for 100 parameter calibrations of HMM using Apple, AAPL, stock daily trading prices.

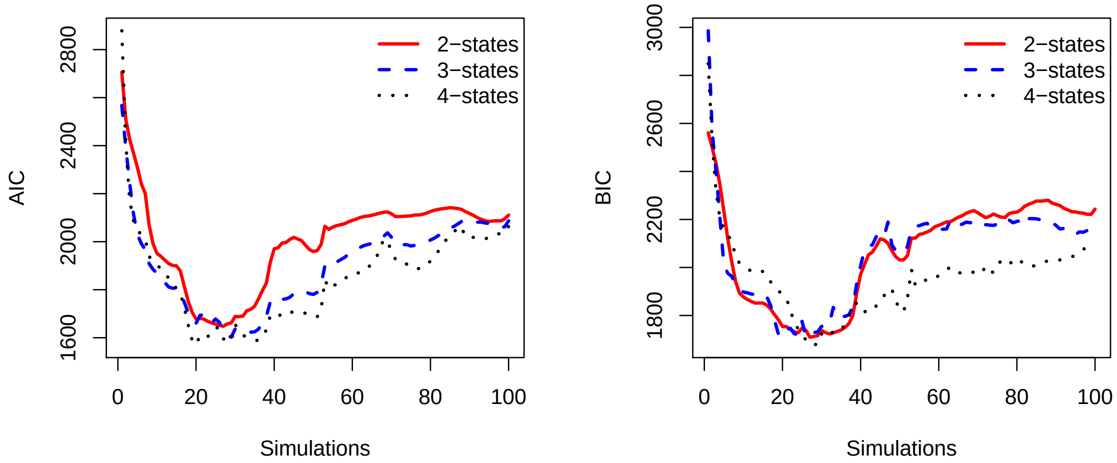


Figure 2: AIC (left) and BIC (right) for 100 parameter calibrations of HMM using Google, GOOGL, stock daily trading prices

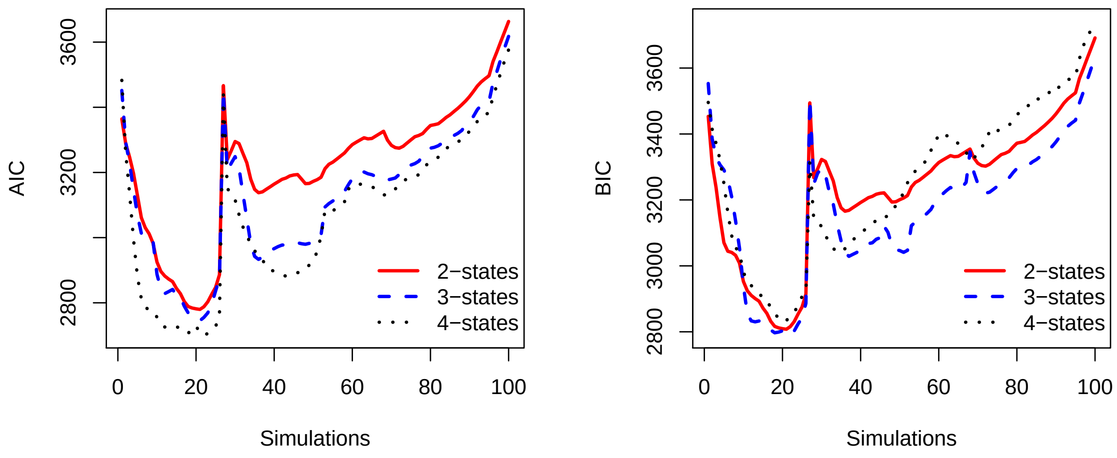
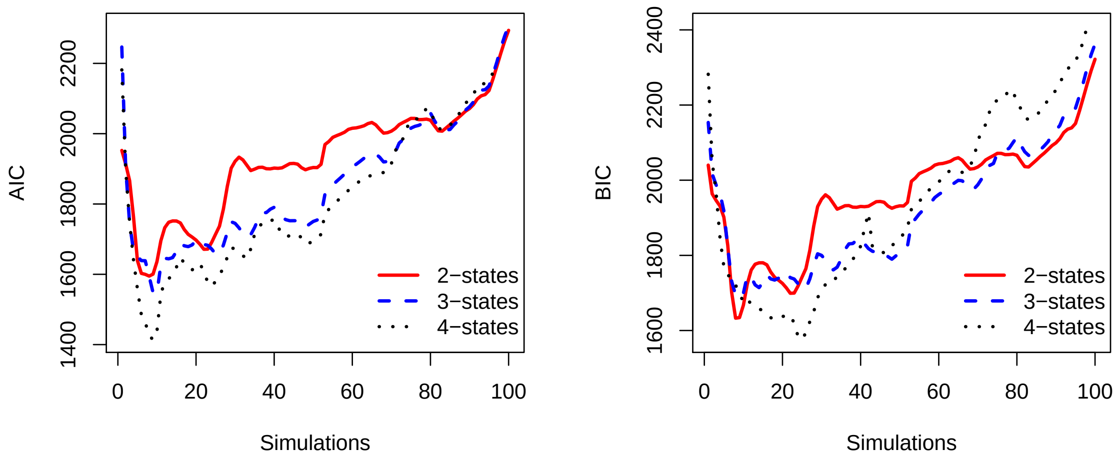


Figure 3: AIC (left) and BIC (right) for 100 parameter calibrations of HMM using Facebook, FB, stock daily trading prices.



HMM with different numbers of states and calculate the mean absolute percentage errors of the estimations.

We first introduce how to predict stock prices using HMM. The prediction process can be divided by three steps. Step one: calibrate HMM's parameters and calculate likelihood of the model. Step 2: find a day in the past that has a similar likelihood to that of the recent day. Finally, step 3: use the difference of stock prices on the "similar" day in the past to predict future stock prices. This prediction approach is based on the work of Hassan and Nath [2]. However, our procedure is different from their method in that we put the sign of the difference of the likelihood of recent day with likelihood of the "similar" day in the past into the calculation of future stock prices. Furthermore, we do not only use HMM with multiple observations (open, low, high, close price), as in [2], but also the HMM with one observation data (close price) to predict future close price. By using both single and multiple observations to predict stocks' close prices, we can examine which method performed better. The HMM with a single observation is a special case of HMM with multiple observations, thus we will explain in detail how to use HMM with multiple observations.

Suppose that we want to predict tomorrow close price of a stock, A. In the first step, we choose a block of T of the four daily prices of stock A: open, low, high, and close, $(O = \{O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}, t = T-99, T-98, ..., T\})$, to calibrate HMM's parameters, λ, and then calculate probability of observation, $P(O|\lambda)$. We assume that the observation probability $b_i(k)$, defined in Section 2, is Gaussian distribution, so the matrix B, in the parameter $\lambda = \{A, B, p\}$, is a 2 by N matrix of means, μ, and variances, σ, of the N normal distributions, where N is the number of states

In the second step, we move the block of data backward by one day to have new observation data $O^{new} = \{O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}, t = T-100, T-99, ..., T-1\}$ and calculate $P(O^{new}|\lambda)$. We keep moving blocks of data backward day by day until we find a data set O^*, $(O^* = \{O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}, t = T^* - 99, T^* - 98, ..., T^*\})$ such that $P(T^*|\lambda) \simeq P(O|\lambda)$. In the final step, we estimate the stock's close price for time $T+1$, $O_{T+1}^{(4)}$, by using the following formula:

$$O_{T+1}^{(4)} = O_T^{(4)} + (O_{T^*+1}^{(4)} - O_{T^*}^{(4)}) * sign(P(O^*|\lambda) - P(O|\lambda)). \tag{2}$$

After the first prediction of the stock's close price for the time $T+1$ we update observation data O by moving it forward one day, $O = \{O_t^{(1)}, O_t^{(2)}, O_t^{(3)}, O_t^{(4)}, t = T-98, T-97, ..., T+1\}$, to predict stock price for the day $T+2$. The calibrated HMM's parameters in the first prediction were used as the initial parameters for the second prediction. We repeat the three step prediction process for the second prediction and so on. For HMM with single observation, we use $O = O_t^{(4)}$ where $O^{(4)}$ is the stock close price set.

We present results of using two state HMM to predict close prices of the three stocks: AAPL, GOOGL, and FB for 100 days in Figures 10, 11, and 12. The graphs of using three and four stats HMM for predicting prices of these stocks cab be found in the appendix. We also compare the performances of HMM with different numbers of states by calculating the mean absolute percentage error, MAPE, of the estimations:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|M_i - P_i|}{M_i},$$

where N is the number of predicted points, M is market price, and P is predicted price of a stock. The results were shown in Table 1. Based on the numerical results in Table

Table 1: Comparison of MAPE of stock price predictions of Apple, Google, and Facebook, using HMM with different numbers of states and numbers of observation data.

Stock	Observations	2 States	3 States	4 States
AAPL	1	0.017	0.019	0.016
	4	0.013	0.009	0.009
GOOGL	1	0.018	0.017	0.021
	4	0.009	0.010	0.010
FB	1	0.021	0.020	0.018
	4	0.015	0.015	0.013

1, we see that the HMM using multiple observation data is better than the HMM using single observation data in stock price predictions. The numbers of states used in the HMM model also affects the accuracy of the prediction. HMM with four states is the best model for forecasting prices of Apple and Facebook stocks while the HMM with two or three states is better in predicting Google stock prices. The results are consistent with the results of the AIC and BIC in Section 3.2. However, AIC and BIC measure the HMMs' parameter calibrations while MAPE evaluates model stock price prediction.

Figure 4: HMM prediction of Apple stock daily close prices from 6/12/2015 to 10/30/2015 using two state HMM with one observation and four observations

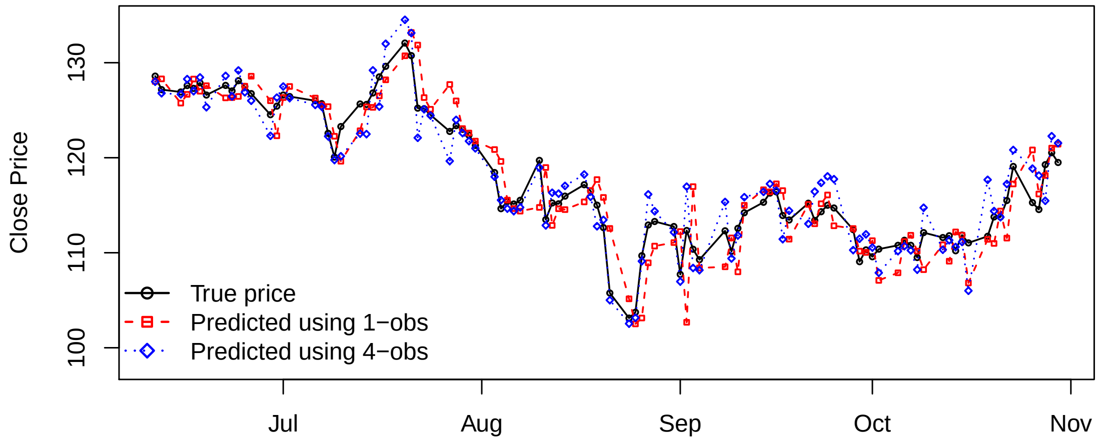


Figure 5: HMM prediction of Google stock daily close prices from 6/12/2015 to 10/30/2015 using two state HMM with one observation and four observations

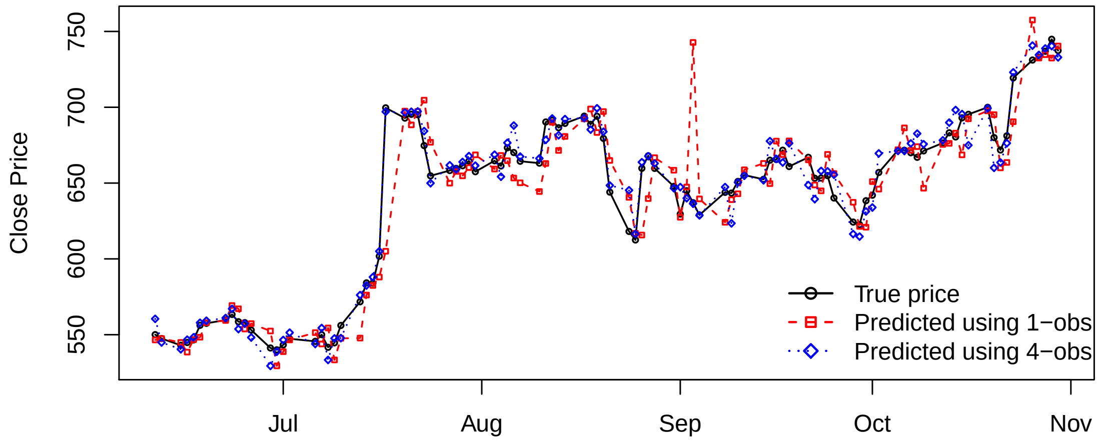
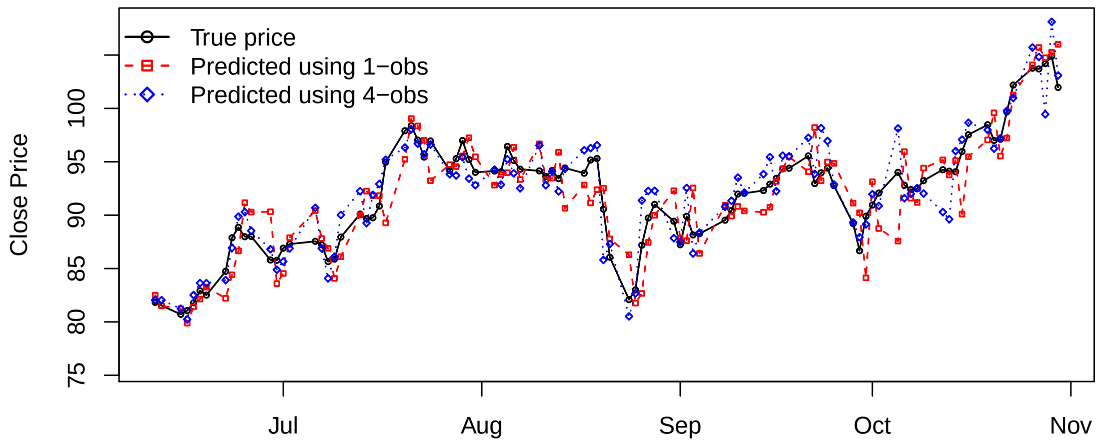


Figure 6: HMM prediction of Facebook stock daily close prices from 6/12/2015 to 10/30/2015 using two state HMM and one observation and four observations



4 Conclusions

Stock performances are an essential indicator of the strengths and weaknesses of the stock's corporation and the economy in general. There are many factors that will drive stock prices up or down. In this paper, we use Hidden Markov Model, HMM, to predict close prices of three stocks: AAPL, GOOGL, and FB. We first use the AIC and BIC criterions to examine the performances of HMM numbers of states from two to four. The results show that when the number of calibrations increase, HMM with different numbers of states gave very similar AICs or BICs. The performances of the models were alternating by the simulations and stocks. We then use the models to predict stock prices and compare their predictions by plotting the forecasted prices versus the market prices and evaluating the mean absolute percentage error, MAPE. The numerical results show that the HMM with four states is the best candidate to predict prices of AAPL and FB stocks. Two or three states HMM is better in forecasting GOOGl stock prices. The results indicate that the HMM is the potential model for stock trading since it captures well the trends of stock prices.

References

[1] Aditya Gupta, Bhuwan Dhingra, Bhuwan Dhingra; Stock Market Prediction Using Hidden Markov Models. *IEEE Engineering and Systems (SCES), 2012 Students*

Algorithm 3 : The backward algorithm

1. Initialization $P(O|\lambda) = 1$
2. For $l = 1, 2, ..., L$ do
   (a) Initialization: for i=1,2,..., N

   $$\beta_T^{(l)}(i) = 1.$$

   (b) Recursion: for $t = T-1, T-2, \ldots, T$, and for $j = 1, 2, \ldots, N$, compute

   $$\beta_t^{(l)}(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}^{(l)}) \beta_{t+1}^{(l)}(j).$$

   (c) Calculate:

   $$P(O^{(l)}|\lambda) = \sum_{i=1}^{N} \beta_1^{(l)}(i).$$

   (d) Update:

   $$P(O|\lambda) = P(O|\lambda) * P(O^{(l)}|\lambda).$$
3. Output: $P(O|\lambda)$.

Figure 7: HMM prediction of Apple stock daily close prices from 6/12/2015 to 10/30/2015 using three state HMM with one observation and four observations

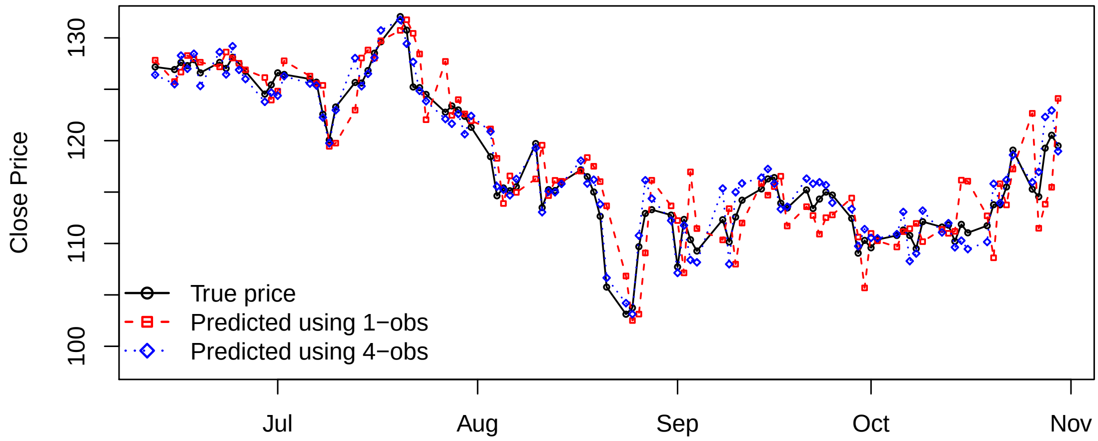


Figure 8: HMM prediction of Google stock daily close prices from 6/12/2015 to 10/30/2015 using three state HMM with one observation and four observations

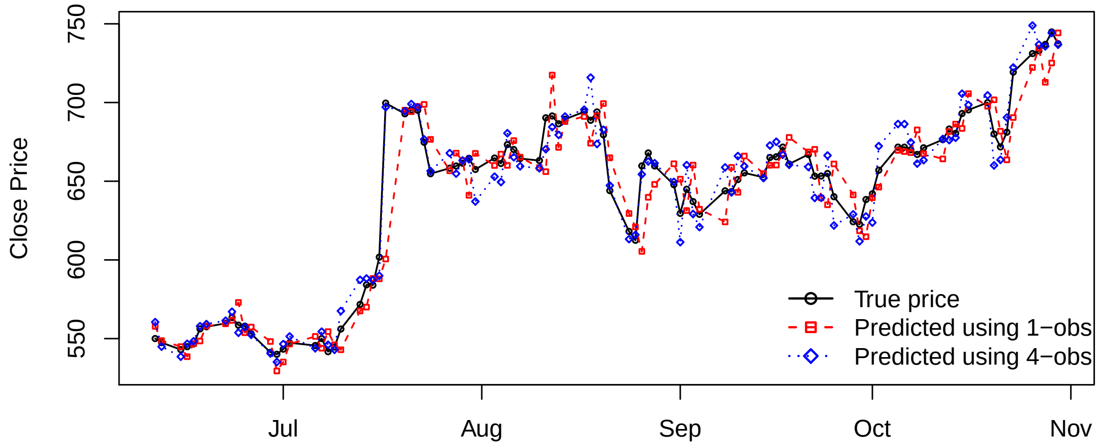


Figure 9: HMM prediction of Facebook stock daily close prices from 6/12/2015 to 10/30/2015 using three state HMM with one observation and four observations

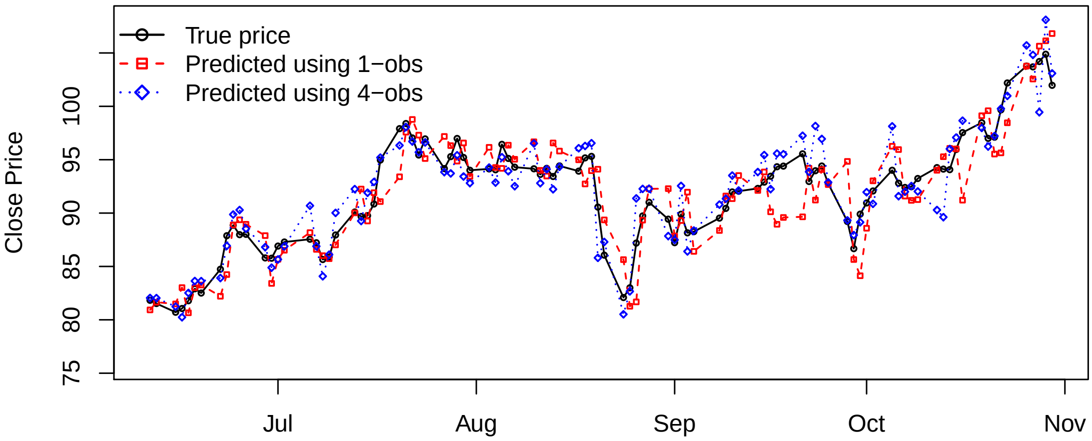


Figure 10: HMM prediction of Apple stock daily close prices from 6/12/2015 to 10/30/2015 using four state HMM with one observation and four observations

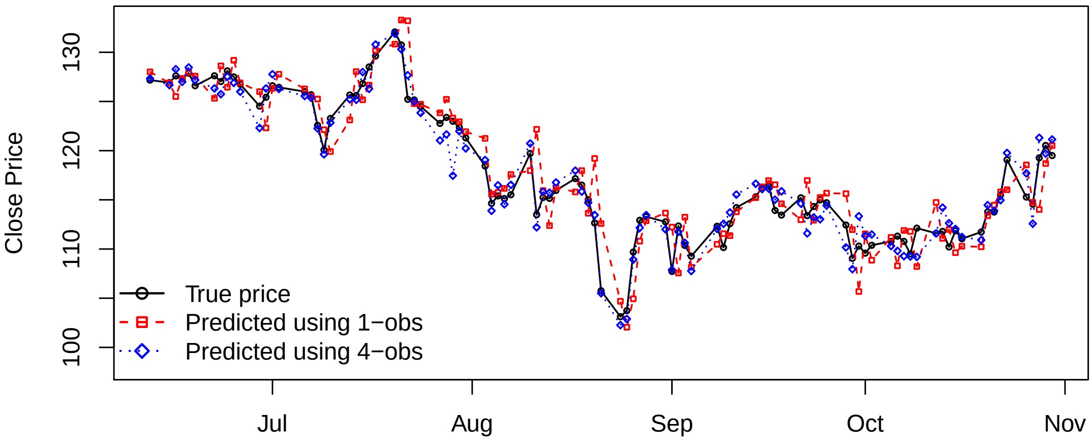


Figure 11: HMM prediction of Google stock daily close prices from 6/12/2015 to 10/30/2015 using four state HMM with one observation and four observations

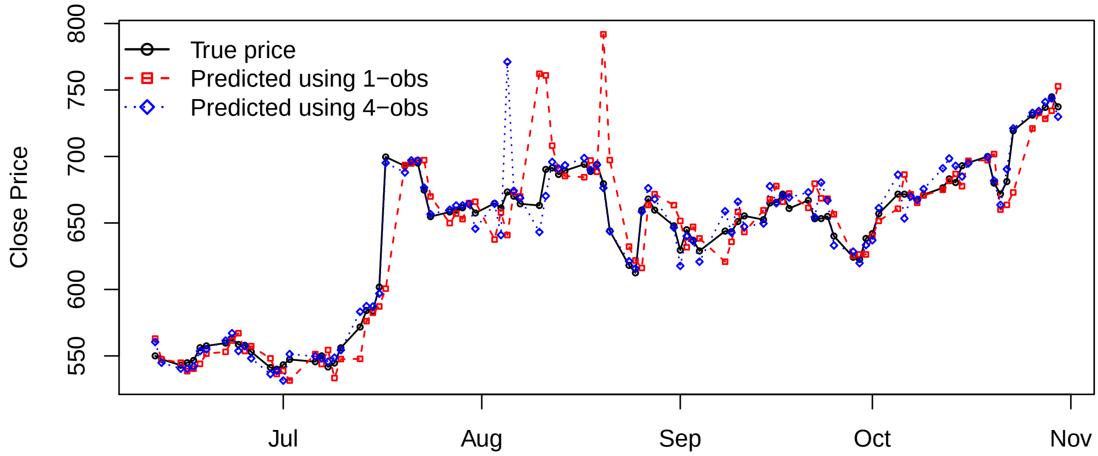
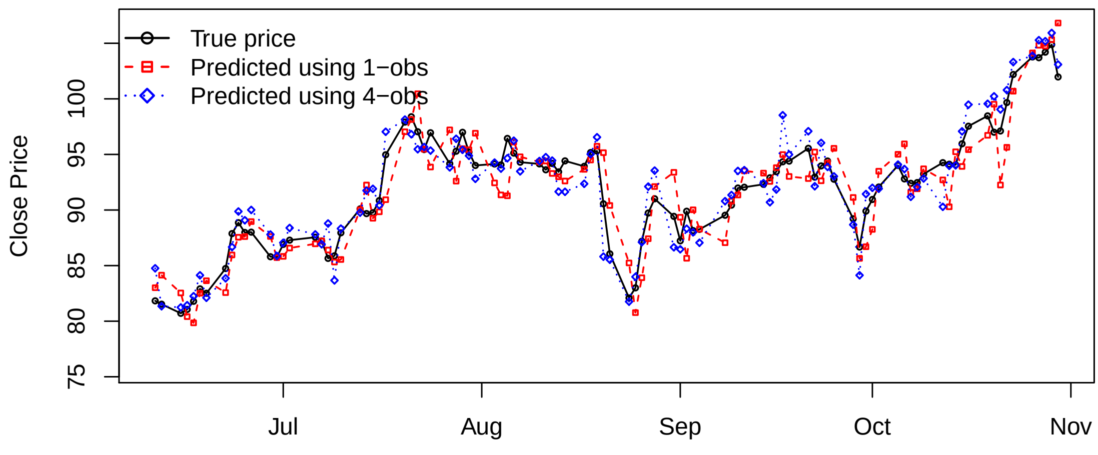


Figure 12: HMM prediction of Facebook stock daily close prices from 6/12/2015 to 10/30/2015 using four state HMM with one observation and four observations



Conference on **2012, 978-1-4673-0456-6**

[2] Hassan, Md. R.; Nath, B.; Stock Market Forecasting Using Hidden Markov Models: A New approach. *Proceeding of IEEE fifth International Conference on Intelligent Systems Design and Applications* **2005**.

[3] Baum, L.E.; Petrie, T.; The Annals of Mathematical Statistics. *Statistical inference for probabilistic functions of finite state Markov chains* **1966**, *37*, 1554-1563.

[4] Levinson, S.E.; Rabiner, L.R.; Sondhi, M.M. An introduction to the application of the theory of probabilistic functions of Markov process to automatic speech recognition. *Bell System Technical Journal* **1983**, *62(4)*, 1035-1074.

[5] Baum, L.E.; Petrie, T.; Soules, G.; Weiss, N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics* **1970**, *41(1)*, 164-171.

[6] Li, X.; Parizeau, M.; Plamondon, R. Training Hidden Markov Models with Multiple Observations - A Combinatorial Method. *IEEE Transactions on PAMI* **2000**, *22(4)*, 371-377.

[7] Kritzman, M.; Page, S.; Turkington, D. Regime Shifts: Implications for Dynamic Strategies. *Financial Analysts Journal* **2012**, *68 (3).*

[8] Nobakht, B.; Joseph, C. E.; Loni, B.; Stock market analysis and prediction using hidden markov models. *Student Conference on Engg and Systems(SCES)* **2012**, 1-4.

[9] Baum, L. E.; Egon, J. A.; An inequality with applications to statistical estiation for probabilistic functions of Markov process and to a model for ecnogy. *Bull. Amer. Meteorol. Soc* **1967**, *73*, 360-363.

[10] Baum L. E.; Sell, G. R.; Growth functions for transformations on manifolds. *Pac. J. Math* **1968**, *27.2*, 211-227.

[11] Rabiner, L. R.; A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *IEEE* **1989**, *77.2*, 257 - 286.

[12] Forney, G. D.; The Viterbi algorithm. *IEEE* **1973**, *61*, 268-278.

[13] Viterbi, A. J.; Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. *IEEE Trans. Informat. Theory* **1967**, *IT-13*, 260-269.

[14] Guidolin, M. and Timmermann, A.; Asset Allocation under Multivariate Regime Switching. *SSRN* **2006**, *FRB of St. Louis Working Paper No. 2005-002C.*

[15] Chen, C.; How well can we predict currency crises? Evidence from a three-regime Markov-Switching model. *JEL, 2005, Classification Code: F31, F37.*

[16] Ang, A. and Bekaert, G.; International Asset Allocaion with Regime Shifts. *The review of financial studies* **2002**, *Vol. 15, No. 4*, 1137-1187.

[17] Nguyen, N.; Probabilistic Methods in Estimation and Prediction of Financial Models. *Electronic Theses, Treatises and Dissertations* **2014**, 9059.

[18] Petrushin, V. A.; Hidden Markov Models: Fundamentals and Applications. *Online Symposium for Electronics Engineer* **2000**.

[19] Nguyen, N and Nguyen, D; *Hidden Markov Model for Stock Selection*, Journal of Risks in special issue: Recent Advances in Mathematical Modeling of the Financial Markets, Risks 2015, 3(4), 455-473; doi:10.3390/risks3040455.

[20] Akaike, H.; A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (6): doi:10.1109/TAC.1974.1100705, MR 0423716 **1994**, 716-723