BENIN JOURNAL OF STATISTICS
ISSN 2682-5767
Vol. 6, pp. 70– 83 (2023)

Regime Characteristics Parameter Estimation of Hidden Markov Model on Time Series Data

R. U. Okafor[1,*] **and J. N. Onyeka-Ubaka**[2]

[1,2] *Department of Statistics, Faculty of Science, University of Lagos, Nigeria.*

(*Received: 26 January 2023; Accepted: 02 April 2023*)

Abstract. The hidden Markov model (HMM) is a versatile and well-adapted method for sequential data analysis. It is often adopted to study occurrences where one part of the phenomenon is visible or observable while the other is hidden or unobservable. Markov-switching models are stochastic processes that switch between regimes or states. The switching between the states works on the assumption that the unobserved states depend on a Markov chain. Regime-switching models are suitable for modeling economic time series where the dynamics are characterized differently within periods. These periods could, for example, be periods of financial stability versus periods of financial volatility or periods of economic expansion versus economic recession. Many classical time series models depend on a single set of model parameters that are assumed to be adequate for modeling the dynamics of time series data over these different periods. But another opinion maintains that this classical model assumption is not always logical for real-life data, where time-series data may have varying characteristics, such as the means and variances across these different periods, hence is better modeled using regime-switching models. This work explored the parameter estimation of a hidden Markov model with application to stock price time series using the maximum likelihood estimation method. The results showed different mean and variance values in the four regimes used for the analysis. The findings indicate that regime-switching models may outperform single models in capturing the empirical characteristics of financial time series data.

Keywords: Regime-switching, Parameter estimation, Hidden Markov model, Maximum likelihood estimation, Time-series.

Published by: Department of Statistics, University of Benin, Nigeria

1. Introduction

Hidden Markov Models (HMM) are usually used to represent systems whose states or regimes are hidden or unobservable but determined from observations of the outputs or emissions associated with the states. The system is assumed to be a Markov process with hidden states, where the probability of moving

*Corresponding author. Email: ukagbans@gmail.com

from one state to another depends only on the current state and not on any of the previous states. HMM consists of a set of states and observable outputs or emissions. The transition probabilities connect these states, with each state emitting one of the visible outputs with a probability that depends on that specific state. The probability of observing a sequence of emissions can be determined using the forward-backward approach. This approach is a dynamic programming algorithm that computes the likelihood of being in each state at each time step, given the sequence of emissions. HMMs have a wide range of applications in speech recognition, bioinformatics, natural language processing, and many other fields where hidden structures are estimates from observable data.

Hidden Markov models (HMMs) are regime-switching models known to be double stochastic processes. Regime-switching models seem suitable for modeling economic time series where the dynamics are characterized differently in periods; see Hamilton (1990). These periods could be either financial stability versus periods of crisis or periods of economic expansion versus economic recession. A hidden Markov model is a peculiar example of a state-space model where the hidden variables are discrete and multinomial. Hidden Markov models consist of a stochastic Markov process of unobserved variables and another stochastic process that produces a sequence of the observations given the first process. See Kim (1994) and Ghysels (1998) for more on hidden Markov models. When estimating the parameters of HMMs, it is necessary to establish that the state space is discrete and that the probability distribution of the observable variables belongs to a parametric family (Vidyasagar (2011), Daniel and James (2002)). Lindgren (1978) proved the consistency property of maximum likelihood estimators obtained for the model, which assumes that the observable variable is an independent sequence from a finite mixture distribution. Leroux and Puterman (1992) decided to fit a finite mixture model as an effective strategy for obtaining parameter estimates suitable for initiating the iterative calculations for evaluating the exact maximum likelihood estimates of a hidden Markov model. Baum and Eagon (1967) studied the expectation-maximization algorithm of Dempster *et al.* (1977) for locating a local maximum of the likelihood function for a probability function of a Markov chain. Also, Baum *et al.* (1970) developed the expectation-maximization algorithm, which they applied to general hidden Markov models. Baum and Petrie (1966) studied the large-sample behavior of a sequence of maximum-likelihood estimators for a probabilistic function of a Markov chain. Many classical time series models depend on a single set of model parameters and are assumed to be adequate to model the dynamics of the data over different periods. But some researchers maintain that this classical model assumption is not always logical for real-life data, where time-series data may have varying characteristics, such as the means and variances across different periods, and hence is better modeled using regime-switching models.

This work reviewed the forward-backward algorithm, also known as the Baum-Welch algorithm, which utilizes an iterative algorithm for the parameter estimation of sequential data on the assumption of a discrete probability density within each state. The marginal likelihood of the observations is needed to obtain maximum likelihood estimates of the model parameters. The forward-backward approach is usually adopted to calculate the marginal log-likelihood for hidden Markov models. The aim is to apply the Maximum likelihood parameter estimation method to stock price data for Access Bank Nigeria plc using a regime-switching Markov model to obtain optimal distinct values for the model's parameters. The results of the switching regime HMM show different means and variances in the four states used for the modeling, indicating that real-life time-series data have varying characteristics across different periods. Checking the results generated using the single and regime-switching hidden Markov models, the log-likelihood estimate of the Markov regime-switching model is higher than the single regime model. Also, using another statistical measure, the AIC, the regime-switching model got a lower value, confirming a better model. Comparing the results of the single and regime-switching hidden Markov models indicates that regime-switching models performed better than single models in capturing the empirical characteristics present in financial time series data.

2. Materials and Method

HMM is a type of generative model used to model sequences of observations. The model assumes that there is an underlying, hidden process that generates the observed variable. The HMM consists of an unobservable Markov chain and an observable Markov chain. The unobserved Markov chain represents the underlying process that creates the observations, while the observed Markov chain represents the observation data. The model assumes that the hidden Markov chain is a Markov process, meaning that the probability of transitioning from one hidden state to another depends only on the current state. Hidden Markov models are dependent mixture models comprising the number of states, an initial state probability function, a state transition probability function, and response models. Therefore, they capture dynamic patterns by allowing switching between the states with different equations that describe the time series under various regimes. A significant feature of Markov switching models is the switching controlled by an unobservable state variable that follows a first-order Markov chain.

The hidden Markov model is a statistical model that assumes a Markov process with unobserved states, where the parameter values that make the observed data most likely are determined. See Ankit (2020) and Rubik (2021). These parameters depend on the hidden states, and the joint log-likelihood of the observed variable also depends on the unobserved states. An influential tutorial by Rabiner (1989), based on tutorials by Jack Ferguson in the 1960s, introduced the idea that hidden Markov models are to be characterized by three fundamental problems in real-life application: the evaluation problem, the decoding problem, and the learning problem. But for this work, the concentration is on the learning problem of the HMM, which is the parameter estimation. Problem of learning in hidden Markov models implies estimating the state transition probability and the output emission probability that are most likely for an observed sequence. One of the main challenges in working with HMMs is estimating the model parameters. Because the hidden states of the Markov chain are not directly observable, the likelihood function is difficult to compute in closed form. Algorithms such as the forward-backward (Baum-Welch) algorithm and the Viterbi algorithm are employed to estimate the parameters of HMM. The algorithm works by training the HMM by iteratively estimating the counts. For the learning problem of the hidden Markov model, which is the objective of this work, it entails determining the model's parameters A and B, given an observation sequence O and the set of states of the hidden Markov model. When the model parameters are estimated, the HMM is useful for tasks such as sequence prediction, anomaly detection, and speech recognition.

In hidden Markov model, we assume: (i) As with a first-order Markov chain, the probability of a particular state depends only on the previous state(Markov Assumption): $P(q_i|q_{i-1})$. (ii) The probability of an output observation $\{o_i\}$ depends only on the state that produced the observation $\{q_i\}$ and not on any other states or any other observations(Output Independence): $P(o_i|q_i)$.

The HMM is defined by a set of mathematical equations that describe the probability distributions of the underlying process and the observed events.

Let's denote the components of HMM with the following notations; Q = q_1,q_2,..., q_N as a set of N states, O = o_1,o_2, ..., o_T as a sequence of T observations, each one drawn from a vocabulary V = v_1,v_2,...,v_V, A = a_{11},..., a_{ij},..., a_{NN} as a transition probability matrix, each a_{ij} representing the probability of moving from state i to state j, such that $\sum_{i=1}^{N} a_{ij} = 1, \forall\ i$, $B = b_i(o_t)$ as a sequence of observation likelihoods also called emission probabilities, each expressing the probability of an observation $\{o_t\}$ being generated from a state i and π = π_1, π_2,...,π_N as an initial probability distribution over states. $\{ \pi_i \}$ is the probability that the Markov chain will start in state i, and some states j may have π_j = 0, meaning that they cannot be initial states. Also, $\sum_{i=1}^{n} \pi_i = 1$. The likelihood function of an HMM is used to compute the probability of a sequence of observations given the model parameters.

2.1 *The Transition Probabilities (Forward Probability) and the Observation Probabilities (Backward Probability)*

The Forward-Backward algorithm is a dynamic programming algorithm used for computing the probability of a sequence of observations in a Hidden Markov Model (HMM). The algorithm is on the concept of forward and backward probabilities, which are computed recursively and used to estimate the probability of a particular observation sequence. The algorithm consists of two main steps: the forward pass and the backward pass. Given the current model parameters, the forward pass calculates the probability of the observed data up to a given time step then the backward pass computes the probability of the observed data from a given time step to the end of the sequence.

To compute the likelihood of moving from one hidden state to the other, we use the forward probability denoted as α. Each cell of the forward algorithm $\alpha_t(j)$ represents the probability of being in state j after seeing the first t observations, given the hidden Markov model denoted as λ. The value of each cell $\alpha_t(j)$ is computed by summing over the probabilities of every path that could lead to this cell. The probability for each cell is:

$$\alpha_t(j) = P(o_1, o_2, ..., o_t, q_t = j|\lambda) \tag{1}$$

Since the probability $\alpha_t(j)$ is computed by summing over the extensions of all the paths that lead to the current cell, thus for a given state q_j at time *t*, the value $\alpha_t(j)$ is computed as:

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij} b_j(O_t). \quad (2)$$

where $\alpha_{t-1}(i)$ is the previous forward path probability from the previous time step, a_{ij} is the transition probability and $b_j(O_t)$ is the observation probability for state *j* at time *t*.
The forward algorithm can be computed iteratively as follows:

$$1. \; Initialization: \; \alpha_1(j) = \pi_j b_j(O_1), 1 \leq j \leq N \quad (3)$$

$$2. \; Recursion: \alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij} b_j(O_t); \; 1 \leq j \leq N, \; 1 < t \leq T \quad (4)$$

$$3. \; Termination: \; P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \quad (5)$$

The backward probability denoted as β is the probability of seeing the observations from time $(t+1)$ to the end, given that we are in state *i* at time *t* and given the model λ. This can be expressed as:

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, ..., o_T | q_t = i, \lambda) \quad (6)$$

The backward probability can be computed iteratively in a similar manner to the forward algorithm as:

$$1. \; Initialization: \; \beta_T(i) = 1, \; 1 \leq j \leq N \quad (7)$$

$$2. \; Recursion: \beta_t(i) = \sum_{i=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), \; 1 \leq i, j \leq N, \; 1 < t \leq T \quad (8)$$

$$3.\, Termination:\; P(O|\lambda) = \sum_{i=1}^{N} \pi_j(j) b_j(o_1)\beta_1(j) \tag{9}$$

The probability of being in a particular hidden state at time t, given the observed sequence up to time T, is given by the posterior probability. The algorithm iteratively updates the model parameters, including the transition probabilities, emission probabilities, and initial state probabilities, based on the posterior probabilities computed from the forward-backward algorithm.

2.2 *Parameter Estimation of the Hidden Markov Model*

Parameter estimation involves determining the values of the model parameters based on the observed data. The parameters of an HMM include the transition probabilities, emission probabilities, and initial state probabilities. Several algorithms are known for parameter estimation of an HMM, including the Baum-Welch algorithm (the forward-backward algorithm), the Maximum Likelihood Estimation (MLE) method, and the Expectation-Maximization (EM) algorithm. This work utilized the Baum-Welch algorithm method, one of the well-known methods for estimating the parameters of hidden Markov models.
The Baum-Welch algorithm is an iterative algorithm that updates the model parameters based on the forward and backward probabilities of the observed data. This algorithm estimates the initial state probability, the transition probability, and the emission probability of the HMM by maximizing the likelihood function of the observed data. The Baum-Welch algorithm is an EM algorithm that alternates between the E-step (estimation of the posterior probabilities) and the M-step (updating the model parameters).
The transition probability $\{a_{ij}\}$ and observation probability $b_i(o_t)$ from the observation sequence are derived using the forward and backward probabilities, although the path taken through the model is invisible. The Baum-Welch algorithm solves this by iteratively estimating the counts. It starts with an estimate for the transition and observation probabilities and then uses these estimated probabilities to derive better and better probabilities.
To estimate the transition probability $\{\hat{a}_{ij}\}$, lets use a variant of maximum likelihood estimation method:

$$\hat{a}_{ij} = \frac{expected\, number\ of\ transitions\ from\ state\ \boldsymbol{i}\ to\, \boldsymbol{j}}{expected\, number\ of\ transitions\ from\ state\ \boldsymbol{i}} \tag{10}$$

To compute the expected number of transitions of being in state *i* at time *t* and state *j* at time *t*+1 which is the the joint probability of being in state *i* at time *t* and state *j* at time *j*+1. Let's denote ξ_t as the probability of being in state *i* at time *t* and state *j* at time *j*+1, given the observation sequence and the model as:

$$\xi_t(i,j) = P(q_t = i, q_{t+1} = j|O, \lambda) \tag{11}$$

where $P(q_t = i, q_{t+1} = j|O, \lambda)$ is the probabilities for the model estimation. To generate the maximum likelihood estimates of the model parameters, we calculate the marginal probability of the observations. The forward-backward approach computes the marginal log-likelihood for hidden Markov models. Therefore, the joint probability of being in state *i* at time *t* and state *j* at time *t*+1, denoted as φ_t is the marginal probability of the observations. This is stated as:

$$\varphi_t(i, j) = P(q_t = i, q_{t+1} = j|O, \lambda) \tag{12}$$

Hence,

$$\varphi_t(i, j) = \alpha_t(i)a_{ij}b_j(Ot + 1)\beta_{t+1}(j) \tag{13}$$

where α is the forward probability, β is the backward probabilities, $\{a_{ij}\}$ is the transition probability and $b_j(O_{t+1})$ is the observation probability. But the probability of the observation given the model $P(O|\lambda)$ is the forward or backward probability. This is given as:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_t(j)\beta_t(j) \tag{14}$$

To derive ξ_t which is the probability of being in state *i* at time *t* and state *j* at time (*t*+1), using the laws of conditional independence, divide φ_t by $P(O|\lambda)$. The law of conditionally independent states that;

$$P\{X = a_i|Y = b_j\} = \frac{P\{X = a_i \;\&\; Y = b_j\}}{P\{Y = b_j\}} \tag{15}$$

Therefore, $\xi_t(i, j)$ is given as:

$$\xi_t(i, j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N} \alpha_t(j)\beta_t(j)} \tag{16}$$

The expected number of transitions from state *i* to state *j* is the sum over all *t* of ξ. For the estimate of $\hat{a}_{ij}$, the total expected number of transitions from state *i* is computed. Therefore, to find the total expected number of transitions from state *i*, we sum over all transitions out of state *i*. Thus, the estimate of transition probability $\{a_{ij}\}$ is given as:

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1}\sum_{k=1}^{N} \xi_t(i, k)} \tag{17}$$

2.3 *Re-computing the observation probability*

If $\hat{b}_j(v_k)$ is the probability of a given symbol v_k from the observation vocabulary V , given a state *j,* then;

$$\hat{b}_j(v_k) = \frac{\text{expected number of times in state } j \text{ and observing symbol } v_k}{\text{expeected number of times in state } j} \tag{18}$$

Let's denote $\gamma_t(j)$ as the probability of being in state j at time *t*, given the observations and the model, then $\gamma_t(j)$ is expressed as;

$$\gamma_t(j) = P(q_t = j|O, \lambda) \tag{19}$$

Again, If $\gamma_t(j)$ is conditional on the observation sequence, then, the probability is given as;

$$\gamma_t(j) = \frac{P(\mathrm{q_t}= j, O|\lambda)}{P(O|\lambda\ )} \tag{20}$$

Since $P(q_t = j, O|\lambda)$ is the product of the forward probability and backward probability, then

$$\gamma_t(j) = \frac{\alpha_t(j)\ \beta_t(\mathrm{j})}{P(O|\ \lambda\ )} \tag{21}$$

Therefore, to calculate $\hat{b}_j(v_k)$ using $\gamma_t(j)$, sum $\gamma_t(j)$ for all time steps *t* in which the observation O_t is the symbol v_k for the numerator and sum $\gamma_t(j)$ over all time steps *t* for the denominator. The result is the percentage of the times that the system is in state *j* and saw symbol v_k. Thus;

$$\hat{b}_j(v_k) = \frac{\sum_{t=1\ s.t\ O_t\ =v_k}^{T} \gamma_t(\mathrm{j})}{\sum_{t=1}^{T} \gamma_t(j)} \tag{22}$$

In the Expectation step, the expected state occupancy count γ is computed from transition probability and the expected state transition count ξ is computed from the emission probability. In the maximization step, γ and ξ are used to re-compute new transition and emission probabilities. The equations $\hat{a}_{ij}$ and $\hat{b}_j(v_k)$ are used to re-estimate the transition probability and emission probability from an observation sequence, assuming that we already have a previous estimate of transition probability and emission probability. Also, the initial probability distributions are generated. The forward-backward algorithm starts with some initial estimate of the HMM parameters λ = (*A*, *B*), and then iteratively run the expectation step, and the maximization step. Although in principle the forward-backward algorithm can do completely unsupervised learning of the A and B parameters, in practice the initial conditions are very important. For this reason the algorithm is often given extra information. For example, for HMM-based speech recognition, the HMM structure is often set by hand, and only the emission (B) and (non-zero) A transition probabilities are trained from a set of observation sequences O. These re-estimations are the pivot of the iterative forward-backward algorithm used in the maximum likelihood method of parameter estimation.

3. Results and Discussion

The data used for analysis in this work is the Nigerian Access Bank Plc daily stock prices, extracted from the Nigerian Stock Exchange website. The data spans from March 2012 to March 2022, a period of ten years with 2491 observations. The objective is to train the hidden Markov model to compute the parameters using the closing stock price for a particular day, given the opening stock price for that day and previous days' data. The parameter estimation of the HMM was modeled in R Studio using depmixS4, developed by Ingmar and Maarten (2021).
Assuming the Gaussian distribution for the data and a 4-state switching regime linear model. Let the linear regression model for the observed variable Y_t be expressed as;

$$\mathrm{Y_t} = \hat{\mu}_t + \mathrm{\varepsilon_t} \tag{23}$$

where $\hat{\mu}_t$ is the regression model's prediction of the observed at time *t* and ε_t is the random error in observing Y_t, $\varepsilon_t \sim N(0, \sigma^2)$. Given also that the regression assumptions hold, that is, the residual error ε_t is homoskedastic, and is normally distributed with zero mean and constant variance σ^2. The predicted value of the regression model, $\hat{\mu}_t$ is actually the output of a regression link function, and since we assumed a linear model for the data, the link function $(X_t, \hat{\beta})$ is linear, with

$$\hat{\mu}_t = x_t \,.\, \hat{\beta} \tag{24}$$

Therefore,

$$\mathrm{Y_t} = x_t \,.\, \hat{\beta} + \mathrm{\varepsilon_t} \tag{25}$$

Incorporating the regime-switching variable, to the model, then;

$$Y_t = \hat{\mu}_t S_t + \varepsilon_t \tag{26}$$

where $\mathrm{Y_t}$ (observed variable) is the daily closing price of Access bank stock price, $\mathrm{x_t}$(regression variables) is a single observation of ROC(Access_closets), ROC(Access_opents), ROC(Access_hights), ROC(Access_lowts)), $\hat{\beta}$ (vector) is the fitted coefficients and $\mathrm{S_t}$ is the hidden random variable that is assumed to switch amongst 4 regimes.

Let $\theta = (\pi, A, Y_t)$ be the general parameter vector consisting of three subvectors with parameters for the initial state probabilities, transition probabilities, and response models, respectively. $\mathrm{Y_t}$ is assumed to be normally distributed with mean $\hat{\mu}_t$ and constant variance σ^2 and the HMM is trained to estimate the parameters from the data set. The Access Bank stock closing price data, representing the observed variable, $\mathrm{Y_t}$ was first converted to a time series and checked for stationarity. The results of the time plot and the ACF plot of the stock prices showed that the data is not stationary. The price series were transformed into differenced data, making them "return series" using ROC() function in R. Also, the return series of the opening, closing, low, and high stock prices formed together as a single vector observation representing the regression variable x_t.

Table 1 below shows the descriptive statistics of the daily closing price, opening price, high price, and low price. Figure 1 is the time plot of Access stock closing prices, Figure 2 is the ACF of Access stock closing prices, and Figure 3 is the PACF of Access stock closing prices.

Table 1: Descriptive statistics of Access Bank stock prices(Naira)

	Closing	Opening	High	Low
Observations	2491	2491	2491	2491
Minimum price	3.59	3.48	3.63	3.48
Maximum price	13.45	14.25	13.60	13.15
Mean	7.952	7.950	8.047	7.8498
Std. dev	2.0679	2.0647	2.0795	2.0465
Skewness	0.05458	0.0582	0.06227	0.0494

The summary of the daily Access stock prices, which comprises the open, close, low, and high stock prices spanning ten years, is recorded in Table 1. From the statistics of the stock prices, the mean is not zero for all the attributes, indicating a need to perform a time series analysis on the daily return (log return) of the stock prices.

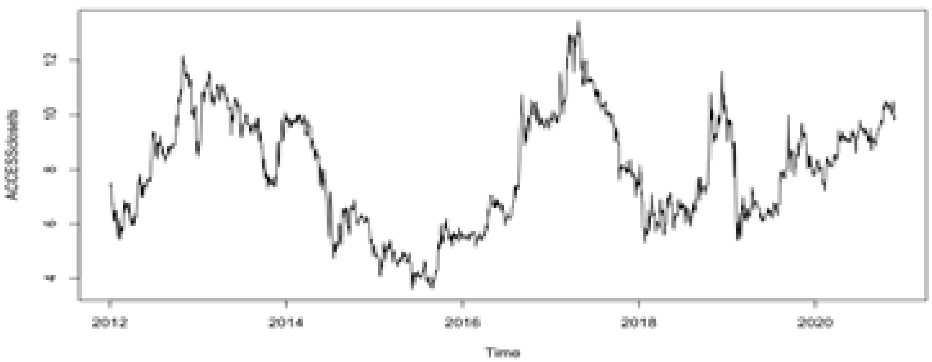

Figure 1: Time plot of Access stock closing prices

The time plot in Figure 1 shows the non-stationarity of the data, as it appears high in some periods and low in others. From the autocorrelation function (ACF) plot of Access stock closing prices in Figure 2, we observe that the time plot decays to zero slowly, which also confirms the non-stationarity of the series. The partial autocorrelation function (PACF) plot of Access stock closing prices in Figure 3 also shows the non-stationarity of the time series. To attain stationarity, transform the series to get differenced data, making it a return series used for the parameter estimation of the HMM.

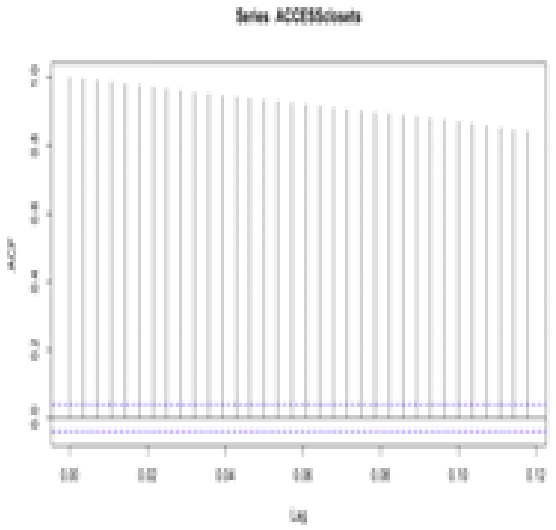

Figure 2: ACF of Access stock closing prices

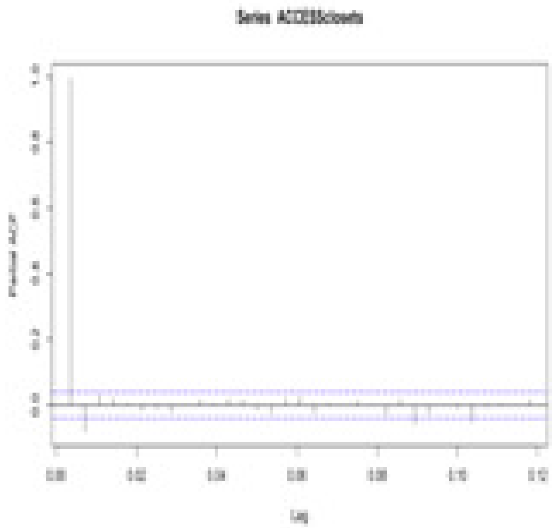

Figure 3: PACF of Access stock closing prices

HMMs are used to model hidden state transitions from the sequential observed data, of which stock price prediction is one. Stock prices depend on factors not visible to the investor, such as business policies and actions, financial situations, managerial choices, and market conditions. The transitions between these underlying factors are unobservable because the changes in the regime are unknown at each time step. These changes affect stock prices (observed data), and this work aims to estimate the parameters of an HMM using this data.
Assuming an HMM for four hidden states, the initial state probabilities model is in Table 2, the transition matrix probabilities are in Table 3, and the response parameters are in Table 4.

Table 2: Initial state probabilities

pr1	pr2	pr3	pr4
1	0	0	0

Table 2 shows the initial probability distributions over the four states. The state probability vector $\{\pi_t\}$ contains the unconditional probability of being in a particular state at time t. For the HMM switching regime $\mathrm{S_t}$ used for the Access stock price data, the initial state probability π_t, is as follows: $\pi_t = P(S_1 = 1),\ P(S_2 = 0),\ P(S_3 = 0),\ P(S_4 = 0)$. The result shows that the initial probability of the Markov chain starts in only state 1, while the chances that the

Markov chain will initiate in states 2, 3, and 4 are zero, meaning that these states cannot be initial states for this model.

Table 3: Transition matrix probabilities

	to S_1	to S_2	to S_3	to S_4
From S_1	0.973	0.018	0.009	0.000
From S_2	0.011	0.975	0.000	0.013
From S_3	0.004	0.000	0.996	0.000
From S_4	0.000	0.022	0.000	0.978

Table 3 is the transition probability matrix, each a_{ij} represents the probability of the process moving from state i to state j, and notice that $\sum_{i=1}^{N} a_{ij} = 1$. The state transition matrix of the 4-state hidden Markov model results for Access stock price data is as follows:

$$A = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix} = \begin{bmatrix} 0.973 & 0.018 & 0.009 & 0.000 \\ 0.011 & 0.975 & 0.000 & 0.013 \\ 0.004 & 0.000 & 0.996 & 0.000 \\ 0.000 & 0.022 & 0.000 & 0.978 \end{bmatrix}$$

A contains the probabilities of transitioning to the next state, conditional upon the current state. The results show that for the assumed four hidden factors that can affect changes in Access stock price, the probability of transition from state 1 to 4, state 2 to 3, state 3 to 4, state 4 to 1, and state 4 to 3 are all zero. But, there are greater chances of the system remaining in a particular state once it moves to that state, as can be seen in the cases of state 1 returning to state 1, state 2 returning to state 2, state 3 returning to state 3, and state 4 returning to state 4.

Table 4: Response Paprameter (Y_t)

		$\sigma(Std.\ Dev.)$
S_1	7.900	4.66
S_2	9.442	0.403
S_3	5.815	0.827
S_4	10.957	0.789

The Gaussian distribution parameters (μ and σ) of the 4-state switching regime linear regression model in Table 4 show distinct values for the Access stock price data. From the regression's equation $Y_t = \hat{\mu}_t S_t + \varepsilon_t$ and since S_t is assumed to switch between regimes 1, 2, 3, and 4, then, from the fitted coefficients of the results, the regime-specific model equations for the Access stock prices at these 4 hidden states are as follows:

$Y_1 = 7.900 + \epsilon_t$ where $\epsilon_t \sim N(0, 0.217)$
$Y_2 = 9.442 + \epsilon_t$ where $\epsilon_t \sim N(0, 0.162)$
$Y_3 = 5.815 + \epsilon_t$ where $\epsilon_t \sim N(0, 0.684)$

$Y_4 = 10.957 + \epsilon_t$ where $\epsilon_t \sim N(0, 0.623)$

The results of the regime-specific model for the Access stock price data show that the model has a state-specific variance for the four states. It indicates a normal distribution for the residual errors of the model around zero mean and also variance that switches between four values depending on which state the underlying Markov process is in.
From the model results, the initial state probabilities indicate that state 1 is the starting state for the process, with negligible chances of starting in states 2, 3, and 4. The transition probabilities show stable states, meaning the likelihood of persisting in any particular state is around 0.9. The resulting model also shows four distinct states with varying responses. The loglikelihood estimate of the Markov Switching Model is -2472.298, AIC is 4990.595, and BIC is 5124.465.

4. Conclusion

Hidden Markov models are known for modeling hidden state transitions from sequentially observed data. Stock predictions exhibit a similar pattern, but stock prices usually depend on various factors generally unobserved by the ordinary investor. The transition between the underlying states changes based on these factors affecting the stock prices. Markov-switching models capture the unobserved movement of data among discrete regimes or states. Regime-switching models are suitable for modeling economic time series with characteristic dynamics in different periods. These periods include, among others, company policies and decisions, financial conditions, management decisions, and market conditions. In cases of stock price predictions, model parameter estimation from a sequence of observations is the first port of call. The objective is to derive the HMM parameters using the forward-backward algorithm applied to Access Bank Nigeria Plc. stock prices. The parameters of HMM are used to predict the stock's closing price for a particular day, given that its opening price is known. Predicting stock market movements is one application of HMM to analyzing time series data. The results of the response model show different means and variances across the four regimes used. This finding indicates that real-life time-series data has varying characteristics across different periods. It contradicts the claim by some researchers that many classical time-series models with a single set of model parameters are adequate to model the dynamics of the data over different periods. The results confirmed that time series data performs better when modeled using regime-switching models, as adopted by other researchers.

Declaration

The authors declared to have no conflict of interest in this paper.

Acknowledgement

The authors acknowledge the anonymous reviewers for their suggestions and constructive criticisms which improved the quality of the article.

References

Ankit, S. (2020). Hidden Markov Models for time series in R studio. https://www.medium.com

Baum, L. E. and Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. Bulletin of the American Mathematical Society, 73(3), 360-363.

Baum, L. E., Petrie, T., Soules, G. and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chain. Ann. Math. Stat., 41, 164 -171.

Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite-state Markov chains. Annals of Mathematical Statistics, 37(6), 1554–1563.

Daniel, J. and James, H. M. (2021). Speech and Language Processing. https://www.stanford.edu

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, 39(1), 1–21.

Ghysels, E. (1998). On the periodic structure of the business cycle. Journal of Business and Economic Statistics, 12, 289–298.

Hamilton, J.D. (1990). Analysis of time series subject to changes in regime. Journal of Econometrics, 45, 39–70.

Ingmar, Visser and Maarten, Speekenbrink (2021). depmixS4: An R package for Hidden Markov models. https://www. cran.r-project.org

Kim, C.J. (1994). Dynamic linear models with Markov-switching. Journal of Econometrics, 60, 1–22.

Leroux, B.G. and Puterman, M. (1992). Maximum-penalized-likelihood estimation for independent and Markov dependent mixture models. Biometrics, 48, 545-558.

Lindgren, G. (1978). Markov regime models for mixed distributions and switching regressions. Scand. J. Statist., 5, 81-91.

Rabiner, L R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2), 257-286.

Rubik (2021). Stock price prediction using Hidden Markov Model. https://www.rubikscode.net

Vidyasagar, M. (2011). Hidden Markov Processes: Theory and Applications to Biology. Princeton university press. https://www.personal.utdallas.edu