International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V68I6P514

Volume 68 Issue 6, 117-130, June 2022

Original Article

Prediction of Maize Price using Hidden Markov Chain Model: An Application on Grand market of Lome (TOGO)

Marc Marin Agbodjan [1,⋆], Joseph Mung'atu[2], Anthony Wanjoya[2]

[1] *Department of Mathematics, Pan African University, Institute for Basic Science, Technology and Innovation, Nairobi, Kenya*

[2] *Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya*

Received: 26 April 2022 Revised: 04 June 2022 Accepted: 09 June 2022 Published: 28 June 2022

Abstract

Around the world, the Hidden Markov Models (HMM) are the most popular methods in the machine learning and statistics for modeling sequences, especially in speech recognition domain. According to the number of patent applications for speech recognition technology form 1988 to 1998, the trend shows that this method has become very mature. The Hidden Markov Model (HMM) is a stochastic model where the modeled system is considered as a Markov process with unknown parameters. The challenge is to determine the hidden parameters of the observed parameters. It provides a probabilistic framework for time series forecast modeling. If the causal factors are not directly observed and have the properties of the Markov chain, therefore, a pair of observations and its causative factors are hidden Markov model. The proposed model was applied to a variation in the price of cereal products on the Togolese market between January 2015 and December 2021. The price of these cereal products was influenced by several factors, such as the international recession, government policy, climate, etc. The Forward-Backward, Viterbi and Baum-Welch algorithm was used to estimate the estdimated parameters were used to calculate the expected state of price for those grain product to find the most likely sequence and our proposed model would be coded using Matlab.

Keywords - *Hidden Markov Chain(HMM), Transition Probability Matrix, Emission Probability, Maize, Estimated parameters*

In most sub-Saharan countries, economic development depends heavily on the agricultural sector as Togo, a country located to the west of Ghana, to the east by Benin, to the north by state, and also the south by the Atlantic ocean; with a component of 56,600 km2 and a population of around 8.5 million; the agricultural sector occupies 65% of the working population and 40% of the gross domestic product. Most of the working population is therefore directly involved in agricultural activities and also agriculture, therefore, constitutes a backbone of Togo's economy due to the value of agricultural products. The physiographic conditions, soil quality, and climate of the region are remarkably favorable for the assembly of assorted food crops, vegetables, and fruits, essential for human and animal livelihoods. Thus, for crops grown worldwide for human and animal consumption, cereal crops constitute the majority. they're grown in large quantities and function as a source of energy food over the opposite kind of crop. Cereals are stapled crops of great socio-economic importance, which give food and play a very important role in the agricultural gross domestic product. The demand for agricultural products, especially cereals, has increased due to the increase in the world's population. Cereal crops like wheat, rice, and corn provide about 1 two-thirds of all energy in human diets; but despite everything, There are many factors, which act in numerous directions, and affect the agricultural production sector. Most farmers depend totally on nature to produce their agricultural products; then certain products are produced in a very very particular season. Thus, factors like availability of machine power, accessibility of fine quality seeds and planting material, post-harvest technology, marketing facilities, and infrastructure influence the price of the raw materials produced. one of the foremost problems of agriculture in Togo is that of storage and conservation of products which is at the origin of strong price fluctuations between seasons. This problem is aggravated by the dearth of infrastructure which makes it difficult to access certain regions, especially the foremost production sites; agriculture also suffers from the difficulties of farmers' access to credit and land problems maintained, among other things, by a high population density and low recourse to irrigation. Marketing of commodities and an unorganized marketing system are other major problems that make a good gap between producer and consumer. The facilities for transporting goods from their regions of production to other parts of the state or country are insufficient. Heavy rainfall and sometimes annual flooding also affect production moreover thanks to the price of products. The demand for a specific product and therefore the insufficient supply of it can cause price fluctuation. the continual increase in the price of products creates an unwanted effect on the quality of living of ordinary people. This study attempts to model the price movements of assorted agricultural products specifically cereal products and particularly to make forecasts. There are many well-known forecasting methods for solving real-world problems; samples of these forecasting methods include Autoregressive Moving Average, Autoregressive Integrated Moving Average, Fuzzy statistic, Artificial Neural Networks, Empirical-SVR-Hybrid Decomposition, empirical mode Decomposition: intrinsic mode functions - hybrid and ANN Machine support-vector. Each forecasting model has its own advantages and its specialty for solving 2 complex real-world problems. In this paper, to clarify the interrelationships of our model, the Hidden Markov chain is one of the strong tools that are developed to unravel complex real-world problems just like the value prediction of our cereals products price. Over the past some years, machine learning has become an important subject in the trending technological field. It is a subfield of computer science and artificial intelligence that focuses on the design of systems that can learn and make decisions and predictions based on data, instead of explicitly programmed instructions. Overall it enables computers to act and make data-driven decisions rather than being explicitly programmed to carry out a certain task. One of the important uses of machine learning has been the growth of its potential in predicting the stock market; hence our use of the Hidden Markov Model is the one in which we observe a sequence of emissions, but do not know the sequence of states the model went through to generate the emissions. Analysis of hidden Markov models seek to recover the sequence of states from the observed data. The hidden Markov model was proved appropriate for modeling the dynamic system and can be beneficial for finding unknown parameters of HMM.

2 LITERATURE REVIEW

Hidden Markov models (HMM) were early used in the application of pattern recognition such as speech, gesture and handwriting with many successful research results. In contrast, the application of hidden Markov models in stock market forecasting is still quite new but researchers could already show positive results.For example, the modelling of daily return series with HMMs has been investigated by several authors. One of the first and most influential works was done by Rydén et al. in 1998 where the authors showed that the temporal and distributional properties of daily returns series are well reproduced by two- and three-state HMMs with normal components.

Hassan and Nath (2005) presented Hidden Markov Models (HMM) approach for forecasting slock price at the interrelated markets. Any fluctuation in market influences personal and corporate financial lives, and the economic health of a country. Hassan et al, (2005) consider 4 input features for a stock, which is the opening price, closing price, highest price, and the lowest price. The next day's closing price is taken as the target price associated with the four input features. The idea behind new approach in using HMM is that the using of training dataset for estimating the parameter set (A, B, π). Using the trained HMM, likelihood value for current day's dataset is calculated. For instance the likelihood value for the day is then from the past dataset using the HMM locate those instances that would produce the same or nearest to the likelihood value. Assuming that the next day's stock price should follow about the same past data pattern, from the located past day(s) simply calculate the difference of that day's closing price and next to that day's closing price. Thus the next day's stock closing price forecast is established by adding the above difference to the current day's closing price. The results show potential of using HMM for time series prediction.

Maryam Farshchian and Majid Vafaei Jahan's proposed Hidden Markov Model was proved appropriate for modelling the dynamic system and it has been shown how Baum-Welch and Progressive algorithm can be beneficial for finding unknown parameters of HMM. In implementation phase Baum-Welch was used to train model to learn progressively and incrementally and calculates error possibility after each transition. In their research they have compared their predicted values with actual data for different industries and got maximum 81% accuracy Poonam Somani et al have shown how Hidden Markov Model is used to locate patterns from past data sets that matched the current days stock price behavior. Baum Welch algorithm was used to train the data set for developing model. Input features were Open, Close, High, Low and stock data was stored in 3-d vector form in training phase to calculate the fractional change in testing phase for analysis purpose. Once the model was trained and testing was done, result analysis and accuracy was done by using Mean Absolute Percentage Error (MAPE).

Research in the field of prediction using Hidden markov model is a new breakthrough, because most of them are used for bio technology or recognition or stock market trading startegy. Previous literature suggests that there has been no or few research on the agriculture commodities(Maize, Soyabeans, Rice) market price scientifically. This research was intended to study time-series trends using Hidden Markov Model to give a scientific description by finding the most probable next price state and the most likely sequence of our commodities price

3 METHODOLOGY

3.1 *Basics of Markov Model*

A Markov chain is a stochastic process, but it differs from a general stochastic process in that a Markov chain must be "memory-less." That is, (the probability of) future actions are not dependent upon the steps that led up to the present state. This is called the Markov property. While the theory of Markov chains is important precisely because so many "everyday" processes satisfy the Markov property, there are many common examples of stochastic properties that do not satisfy the Markov property.

Definition

The stochastic process $\{X(t), t \in T\}$ is said to exhibit Markov dependence if for a finite (or countable infinite) set of points $(t_0, t_1, \cdots, t_n, t)$, $t_0 < t_1 < t_2 < \cdots < t_n < t$ where $t, t_r \in T$ $(r = 0, 1, 2, \ldots, n)$.

$$P(X(t) \leq x \mid X(t_n) = x_n , X(t_{n-1}) = x_{n-1}, \ldots, X(t_0) = x_0)$$
$$= P[X(t) \leq x \mid X(t_n) = x_n] = F[X_n, x; t_n, t]$$

Transition Matrice

A transition matrix P_t for Markov chain $\{X\}$ at time t is a matrix containing information on the probability of transitioning between states. In particular, given an ordering of a matrix's rows and columns by the state space S, the $(i,j)^{\text{th}}$ element of the matrix P_t is given by

$$(P_t)_{i,j} = \mathbb{P}(X_{t+1} = j \mid X_t = i)$$

This means each row of the matrix is a probability vector, and the sum of its entries is 1 . Transition matrices have the property that the product of subsequent ones describes a transition along the time interval spanned by the transition matrices. That is to seition $P_0 \cdot P_1$ has in its $(i,j)^{\text{th}}$ position the probability that $X_2 = j$ given that $X_0 = i$. And, in general, the $(i,j)^{\text{th}}$ position of $P_t \cdot P_{t+1} \cdots P_{t+k}$ is the probability $\mathbb{P}(X_{t+k+1} = j \mid X_t = i)$. The k-step transition matrix is $P_t^{(k)} = P_t \cdot P_{t+1} \cdots P_{t+k-1}$ and, by the above, satisfies

$$P_t^{(k)} = \begin{pmatrix} \mathbb{P}(X_{t+k} = 1 \mid X_t = 1) & \mathbb{P}(X_{t+k} = 2 \mid X_t = 1) & \ldots & \mathbb{P}(X_{t+k} = n \mid X_t = 1) \\ \mathbb{P}(X_{t+k} = 1 \mid X_t = 2) & \mathbb{P}(X_{t+k} = 2 \mid X_t = 2) & \ldots & \mathbb{P}(X_{t+k} = n \mid X_t = 2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{P}(X_{t+k} = 1 \mid X_t = n) & \mathbb{P}(x_{t+k} = 2 \mid X_t = n) & \ldots & \mathbb{P}(X_{t+k} = n \mid X_t = n) \end{pmatrix}.$$

Properties

A variety of descriptions of either a specific state in a Markov chain or the entire Markov chain allow for better understanding of the Markov chain's behavior. Let P be the transition matrix of Markov chain $\{X_0, X_1, \ldots, X_N\}$.

- A state i has period $k \geq 1$ if any chain starting at and returning to state i with positive probability must take a number of steps divisible by k. If $k = 1$, then the state is known as aperiodic, and if $k > 1$, the state is known as periodic. If all states are aperiodic, then the Markov chain is known as aperiodic.
- A Markov chain is known as irreducible if there exists a chain of steps between any two states that has positive probability.
- An absorbing state i is a state for which $P_{i,i} = 1$. Absorbing states are crucial for the discussion of absorbing Markov chains.
- A state is known as recurrent or transient depending upon whether or not the Markov chain will eventually return to it. A recurrent state is known as positive recurrent if it is expected to return within a finite number of steps, and null recurrent otherwise.
- A state is known as ergodic if it is positive recurrent and aperiodic. A Markov chain is ergodic if all its states are. Irreducibility and periodicity both concern the locations a Markov chain could be at some later point in time, given where it started. Stationary distributions deal with the likelihood of a process being in a certain state at an unknown point of time. For Markov chains with a finite number of states, each of which is positive recurrent, an aperiodic Markov chain is the same as an irreducible Markov chain.

3.2 *Basics of HMM*

HMM is a stochastic model where the system is assumed to be a Markov Process with hidden states. HMM gives better accuracy than other models. Using the given input values, the parameters of the HMM (λ) denoted by A, B and π are found out.

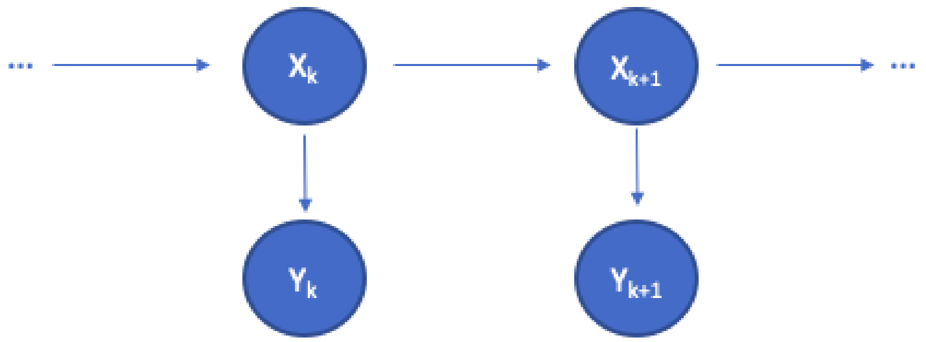


Figure 1: Graphical representation of the dependence structure in a hidden Markov model

HMM consists of :

- ➢ A set of hidden or latent states (S)
- ➢ A set of possible output symbols (O)
- ➢ A state transition probability matrix (A)
- ➢ Probability of making transition from one state to each of the other states
- ➢ Observation emission probability matrix (B)
- ➢ Probability of emitting/observing a symbol at a particular state
- ➢ Prior probability matrix (π)

➢ probability of starting at a particular state

An HMM is defined as λ =(S, O, A, B, π) where :
$S = S_1, S_2, \ldots, S_N$ is a set of N possible states ; $O = O_1, O_2, \ldots, O_M$ is a set of M possible observation symbols ; A is an (NxN) state Transition Probability Matrix (TPM) ; B is an (NxM) observation or Emission Probability Matrix (EPM) and π is an N dimensional initial state probability distribution vector and A, B and π should satisfy the following conditions:

$$\sum_{j=1}^{N} a_{ij} = 1 \quad \text{where} \quad 1 \leq \text{i} \leq \text{N}$$

$$\sum_{j=1}^{M} b_{ij} = 1 \quad \text{where} \quad 1 \leq \text{i} \leq \text{N}$$

$$\sum_{i=1}^{N} \pi_i = 1 \quad \text{where} \quad \pi_i \geq 0$$

The main problems of HMM are: Evaluation, Decoding, and Learning.

Evaluation problem

Given the HMM λ = {A, B,π} and the observation sequence $O = O_1, O_2, ..., O_M$, the probability that model λ has generated sequence O is calculated. Often this problem is solved by the Forward Backward Algorithm (Rabiner, 1989) (Rabiner, 1993).

Forward Algorithm :

We define the forward variable $\alpha_t(i)$ as $\alpha_t(i) = P(O_1, O_2, ..., O_t; q_t = S_i \mid \lambda)$
His interpretation is the probability of observing the sequence $O_1, O_2, ..., O_t$ up to time t and state S_i at time t, under the model λ. So the algorithm is :

1. Initialisation:

$$\alpha_1(i) = \pi_i b_i(O_1), 1 \leq i \leq N$$

2. Induction:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \quad \begin{array}{l} 1 \leq t \leq T-1 \\ 1 \leq j \leq N \end{array}$$

3. Termination:

$$\mathbf{P}(O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

Backward Algorithm :

Similarly to forward, we define a backward variable $\beta_t(i)$ as $\beta_t(i) = P(O_{t+1}, O_{t+2}, ..., O_T \mid q_t = S_i, \lambda)$ The interpretation of the backward variable is instead the probability of observing the sequence $O_{t+1}, O_{t+2}, ..., O_T$ from time $t+1$ to the end, given state S_i at time t under the model λ.So the algorithm is :

1. Initialisation: $\beta_T(i) = 1, 1 \leq i \leq N$ 2. Induction:

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

$$t = T-1, T-2, \ldots, 1, 1 \leq i \leq N$$

Decoding problem

Given the HMM λ = {A, B,π} and the observation sequence $O = O_1, O_2, ..., O_M$, calculate the most likely sequence of hidden states that produced this observation sequence O. Usually this problem is handled by Viterbi Algorithm (Rabiner,1989) (Rabiner,1993).

Viterbi Algorithm

To find the best state sequence $Q = \{q_1, q_2, \ldots, q_T\}$ for a given observation sequence $O = \{o_1, o_2, \ldots, o_T\}$, one need to define the quantity

$$\delta_t(i) = \max_{q_1, q_2, \ldots, q_{t-1}} P[q_1, q_2, \ldots, q_t = s_i, o_1, o_2, \ldots, o_t \mid \lambda]$$

which means the highest probability along a single path, at time t, which accounts for the first t observations and ends in state s_i. By induction one have:

$$\delta_{t+1}(j) = \left[\max_i \delta_t(i) a_{ij}\right] b_j(o_{t+1}).$$

To be able to retrieve the state sequence, one need to keep track of the argument which maximized $\delta_{t+1}(j)$, for each t and j. This is done via the array $\psi_t(j)$. The complete procedure for finding the best state sequence can now be stated as follows:

1. Initialization:

$$\delta_1(i) = \pi_i b_i(o_1), 1 \leq i \leq N$$
$$\psi_1(i) = 0$$

2. Recursion:

$$\begin{aligned} \delta_t(j) &= \max_{1 \leq i \leq N}\left[\delta_{t-1}(i) a_{ij}\right] b_j(o_t), \quad & 2 \leq t \leq T \\ & & 1 \leq j \leq N \\ \psi_t(j) &= \operatorname{argmax}_{1 \leq i \leq N}\left[\delta_{t-1}(i) a_{ij}\right], & 2 \leq t \leq T \\ & & 1 \leq j \leq N \end{aligned}$$

3. Termination:

$$P^* = \max_{1 \leq i \leq N}[\delta_T(i)]$$
$$q_T^* = \operatorname{argmax}_{1 \leq i \leq N}[\delta_T(i)]$$

4. Path (state sequence) backtracking:

$$q_t^* = \psi_{t+1}\left(q_{t+1}^*\right), \quad t = T-1, T-2, \ldots, 1$$

Learning problem

Given some training observation sequences $O = O_1, O_2, \ldots, O_M$ and general structure of HMM (numbers of hidden and visible states), determine HMM parameters $\lambda = \{A, B, \pi\}$ that best fit data. The most common solution for this problem is Baum-Welch algorithm (Rabiner,1989) (Rabiner,1993) which is considered as the traditional method for training HMM.

Baum-Welch algorithm

The third, last, and at the same time most challenging problem with HMMs is to determine a method to adjust the model parameters $\lambda = \{A, B, \pi\}$ to maximize the probability of the observation sequence given the model. There is no known analytical solution to this problem, but one can however choose λ such that $P(O \mid \lambda)$ is locally maximized using an iterative process such as the Baum-Welch method, a type of Expectation Maximization (EM) algorithm. Baum-Welch is however the most commonly used procedure, and will therefore be the one utilized in this the first investigation of HMM.

To be able to re-estimate the model parameters, using the Baum-Welch method, one should start with defining $\xi_t(i,j)$, the probability of being in state s_i at time t, and state s_j at time $t+1$, given the model and the observations sequence. In other words the variable can be defined as:

$$\xi_t(i,j) = P\left(q_t = s_i, q_{t+1} = s_j \mid O, \lambda\right)$$

the forward and backward variables as follows:

$$\begin{aligned}\xi_t(i,j) &= \frac{\alpha_t(i) a_{ij} b_j\left(o_{t+1}\right) \beta_{t+1}(j)}{P(O \mid \lambda)} = \\ &= \frac{\alpha_t(i) a_{ij} b_j\left(o_{t+1}\right) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j\left(o_{t+1}\right) \beta_{t+1}(j)}\end{aligned}$$

which is a probability measure since the numerator is simply $P\left(q_t = s_i, q_{t+1} = s_j, O \mid \lambda\right)$ and denominator is $P(O \mid \lambda)$.

As described bellow, $\gamma_t(i)$ is the probability of being in state S_i at time t, given the observation sequence and the model. Therefore there is a close relationship between $\gamma_t(i)$ and $\xi_t(i,j)$. One can express $\gamma_t(i)$ as the sum of all $\xi_t(i,j)$ over all existing states as follows:

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

By summing $\gamma_t(i)$ over time one get a number which can be interpreted as the number of times that state s_i is visited. This can also be interpreted as the number of transitions from state s_i. Similarly, summation of $\xi_t(i,j)$ over time can be interpreted as the number of transitions from state S_i to state S_j. By using these interpretations, a method for the re-estimation of the model parameters π, A, B for the HMM is as follows:

$$\begin{aligned}\bar{\pi}_i &= \gamma_1(i) \\ \bar{a}_{ij} &= \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \\ \bar{b}_j\left(v_k\right) &= \frac{\sum_{\substack{t=1 \\ s \cdot o_t = v_k}}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}.\end{aligned}$$

3.3 *Model Conception*

In this paper, Maize monthly price value data during 6 years period is considered. Three observing symbols " L " , " M " and " H " have been used: " L indicates Low " , " M indicates Moderate ", " H indicates High ". To define those 3 Observations, we define Δ_1 and Δ_2 as :

Δ_1= mean price $(t) - (0.5)*$ standart deviation(t)

Δ_2= mean price $(t) + (0.5)*$ standart deviation(t)

So :

If This month price > Δ_2 the observation is High.

If This month price > Δ_1 the observation is Low.

If This Δ_1 < month price < Δ_2 the observation is Moderate.

We also have three hidden states assumed and are denoted by the symbol : S_1, S_2, S_3 .

If new monthly price value – previous monthly price value > 0, then observing state is I (Increasing) and the symbol is S_1.
If new monthly price value – previous monthly price value < 0, then observing state is D (Decreasing) and the symbol is S_2.
If new monthly price value – previous monthly price value = 0, then observing state is U (Unchanged) and the symbol is S_3.

The States are not directly observable. The situations of the market are considered hidden. Given a sequence of observation we can find the hidden state sequence that produced those observations.

3.4 *Database*

The dataset is a collection of existing information coded in a form suitable for use and processing. Our data set consists of historical value data in .csv format extracted from the INSEED stock. The csv file contains monthly data on the prices of the various commodities in my study, namely the price of maize, over the period from January 2015 to December 2021. These monthly prices are based on the averages of the prices collected daily.

3.5 *HMM Prediction*

Once the observable and hidden states were defined, the model was implemented in MATLAB. The algorithm is presented in pseudo code below :

- Load data
- Get observed sequence and hidden sequence given the data
- Get model :
  - Estimate the transition and emission matrix with the MATLAB function hmmestimate()
  - Train the model with the MATLAB function hmmtrain()
- For each upcoming Month :
  - Make a prediction using the model and the function hmmviterbi()

4 RESULTS AND DISCUSSION

We chose the true values of maize prices in " Grand Market of Lome " fron January 2015 to December 2021 yielding 84 months of observations. In this model, we have three states, on the assumption that the state space is S = (S1, S2, S3) where S1 = Increasing, S2 = Decreasing and S3 = Unchanged.

Transition Matrix

To find the price movement, we need to find the state transition probability by calculating the number of months that both month day and second month are increasing, we could find the probability from Increasing to Increasing. Then we get the number of months that first month is Decreasing and second month is Decreasing, and Unchanged; we could find the probability from Decreasing to Decreasing and Unchanged to Unchanged. To finish we get the number of months that first month is Decreasing and second month is Increasing, we get the number of months that first month is Increasing and second month is Decreasing, we get the number of months that first month is Decreasing and second month is Unchanged and we get the number of months that first month is Unchanged and second month is Decreasing.

$$S_1 \Rightarrow S_1 = 0.64 \; ; S_1 \Rightarrow S_2 = 0.34 \; ; S_1 \Rightarrow S_3 = 0.02$$
$$S_2 \Rightarrow S_2 = 0.5 \; ; S_2 \Rightarrow S_1 = 0.47 \; ; S_2 \Rightarrow S_3 = 0.03$$
$$S_3 \Rightarrow S_3 = 0.25 \; ; S_3 \Rightarrow S_1 = 0.75 \; ; S_1 \Rightarrow S_2 = 0$$

we get the transition matrix as follows :

$$A = \begin{array}{c} \\ I \\ D \\ U \end{array} \begin{array}{c} \begin{array}{ccc} I & D & U \end{array} \\ \begin{pmatrix} 0.64 & 0.34 & 0.02 \\ 0.47 & 0.50 & 0.03 \\ 0.75 & 0 & 0.25 \end{pmatrix} \end{array}$$

The explanation ce can have of the transition matrix is that , we have 64% chance that the price of Maize will increase next month if this month it has Increase; 34% chance that the price of Mais Decrease next month if this month it has Increase; 2% chance that the Corn price will remain Unchanged next month if this month it has Increase.

47% chance that the price of Maize will Increase next month if this month it has Decrease; 50% chance that the price of But Decrease next month if this month it has Decrease; 3% chance that the price of Mais will remain Unchanged next month if this month it has Decrease

75% chance that the price of Maize will Increase next month if this month remains Unchanged from the month it preceded, 0% chance that the price of Maize Decrease next month if this month is remains Unchanged from previous month, 25% chance that Maize price will remain Unchanged next month if this month remained Unchanged from previous month.

From this Transition matrix we have we can easily make his flow diagram who is just below, Fig 2

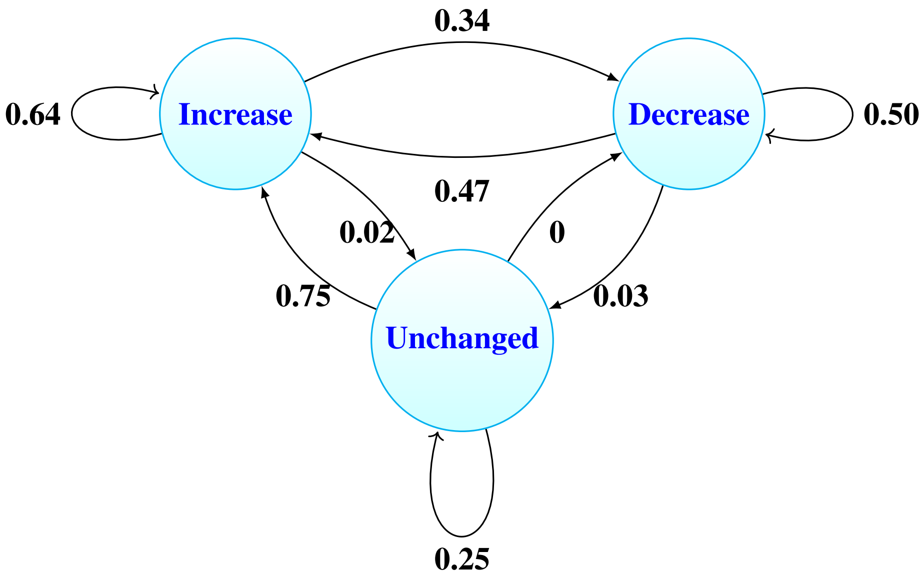
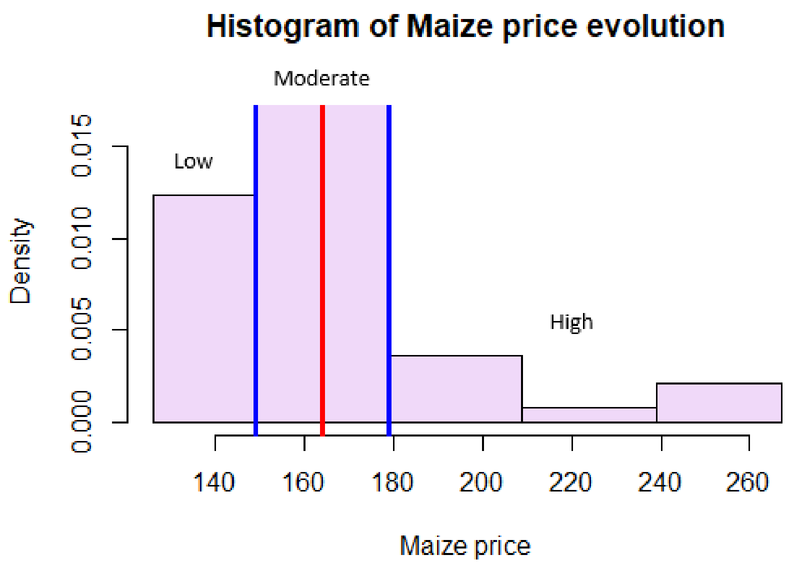


Figure 2: Diagram of Transition Matrix

Emission probabilities

To find the emission probabilities or observable state matrix we need to find the Mean deplacement. We will define our Mean displacement like :



From $[126-149]$ the observation is Low (L), between $[150-180]$ the observation is Moderate (M) and finally, from $[181-267]$ the observation is High (H). From this we have our emission probabilities as :

$$b_{S_1}(L)=0.23,\ b_{S_1}(M)=0.52,\ b_{S_1}(H)=0.25$$
$$b_{S_2}(L)=0.34,\ b_{S_2}(M)=0.53,\ b_{S_2}(H)=0.13$$
$$b_{S_3}(L)=0.5,\ b_{S_3}(M)=0.5,\ b_{S_3}(H)=0$$

we get our Emission matrix as :

$$B=\begin{matrix} & L & M & H \\ I & 0.23 & 0.52 & 0.25 \\ D & 0.34 & 0.53 & 0.13 \\ U & 0.50 & 0.50 & 0 \end{matrix}$$

The initial state distribution (π) is also define as :

$$\pi=(0.57,0.38,0.05).$$

The model of our HMM is define by : $\lambda=(A,B,\pi)$

From this Observable state matrix we have we can easily make his flow diagram who is just below,

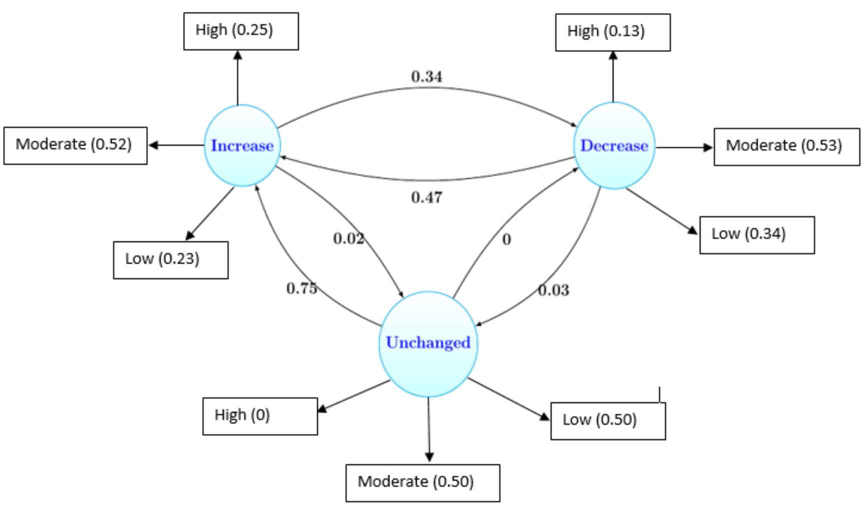


Figure 3: Diagram of HMM

Prediction

This section shows how to analyze hidden Markov models; and to analyse we need :

(1) To Generate a Test Sequence

(2) To Estimate the State Sequence

(3) To Estimate Posterior State Probabilities

(4) To Change the Initial State Distribution

(5) To Estimate Transition and Emission Matrices

To generate a random sequence of states and emissions from the model, we use in matlab **hmmgenerate**: [seq,states] = hmmgenerate(T,TRANS,EMIS). The input T is the length, TRANS the transition matrix and EMIS the Emission probability. The output seq is the sequence of emissions and the output states is the sequence of states; hmmgenerate begins in state 1 at step 0, makes the transition to state i1 at step 1, and returns i1 as the first entry in states. In our work i choose T = 20, so the Output of few randomly generated sequences and states is given below:
Sequences : $L->H->M->M->M->H->H->H->L->L->H->M->L->M->M->L->L->L->L->H$
States : $1->2->2->1->2->1->2->1->1->2->1->2->3->1->1->2->2->2->2->1->1$
The Forward-Backward algorithm gives us the most probable state at each position in the sequence of symbols For example, we can say that state 1 is the most likely to generate the initial symbol L of the sequence O(20) because it has the probability at time = 1 ("$\gamma_t(1) = 0.56694$).

Applying the Baum-Welch algorithm is a bit more complex. Our goal is to reestimate the parameters of this model from the EM algorithm. To do this, we will first use the functions **hmmestimate** or **hmmtrain** to estimate the transition and emission matrices A and B given a sequence seq of emissions.

The function hmmestimate requires that you know the sequence of states that the model went through to generate seq. The following takes the emission and state sequences and returns estimates of the transition and emission matrices:
$[\hat{A},\hat{B}]$ = hmmestimate(seq, states) . From that we have then :

$$\hat{A} = \begin{bmatrix} 0.33 & 0.67 & 0 \\ 0.56 & 0.33 & 0.11 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} 0.30 & 0.30 & 0.40 \\ 0.45 & 0.33 & 0.22 \\ 1 & 0 & 0 \end{bmatrix}$$

By comparing the outputs with the original transition and emission matrices, TRANS and EMIS, we can easily see that there is a big difference and that there is no real convergence... Let's try in a second step to generate 100 observation sequences and 1000 with the **hmmgenerate** function and then re-estimate the transition and emission matrices.
With [seq,states] = hmmgenerate(100,A,B) we have then :

$$\hat{A} = \begin{bmatrix} 0.58 & 0.40 & 0.02 \\ 0.37 & 0.59 & 0.04 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} 0.16 & 0.54 & 0.30 \\ 0.32 & 0.53 & 0.15 \\ 0.67 & 0.33 & 0 \end{bmatrix}$$

With [seq,states] = hmmgenerate(1000,A,B) we have then :

$$\hat{A} = \begin{bmatrix} 0.65 & 0.32 & 0.03 \\ 0.46 & 0.52 & 0.02 \\ 0.64 & 0 & 0.36 \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} 0.24 & 0.51 & 0.25 \\ 0.36 & 0.51 & 0.13 \\ 0.5 & 0.5 & 0 \end{bmatrix}$$

We can first notice that the results obtained are better for 1000 simulations, so the re-estimates converge better when there are more observations... The parameters that we tried to estimate with the sequence O(20) did not converge to the true values. Therefore, in order to illustrate the proper functioning of this algorithm, we must simulate at least 1000 observable symbols.

Viterbi algorithm determine the most likely state sequence to generate the observable symbol sequence. To generate the symbol sequence O(20) = L- H- M- M- M- H- H- H- L- L- H- M- L- M- M- L- L- L- L-H, it is more likely to initially start from state y_1 = 1 then go to state y_2 = 2 , so immediately, and ending at time T = 20 in state $y_2 0$ = 1 (1-2-2-1-2-1-2-1-1-2-1-2-3-1-1-2-2-2-1-1) and in this case, the most likely sequence of states agrees with the sequence 56.30% of the time.

5 CONCLUSION

During this thesis, we have presented the notion of MMC as a possible extension of Markov chains, a notion capable of expressing systems more complex dependencies. It has been explained that this dependence comes from the fact that each state of the chain is linked to a law of probability, which allows to send the symbols. Understanding MMCs is simplified by the way the parameters of the model are built. Algorithms for making inference on MMCs have been coded on MATLAB from the recursive formulas obtained in the development of Chapter 3. This increases the flexibility of the parameterization of MMCs because the mathematical indexing is slight and the majority of the formulas developed are often reduced to mere sums. This way coding in MATLAB is much simplified. We used the Forward-Backward algorithm and the Viterbi algorithm to to evaluate and provide information on the difference between the states of the model to from a sequence of observed symbols. Thanks to the results obtained in the chapter 4, we could indeed see that over time, certain states are more likely than others to generate symbols. Finally, we were also able to understand how the Baum-Welch algorithm works. The latter, which was used for the re-estimation and the adjustment of the parameters of the model from an observed sequence of symbols, allowed us to observe that it strongly depends on the number of data (symbols of states). It was indeed possible to illustrate from a long sequence of symbols (1000 symbols), that there is a convergence of the parameters re-estimated by this algorithm. Our Research did not predict the price of our commodities exactly but predicted with probability that it could be the next state (Increase,Decrease,Unchanged) knowing the current state with other auxiliary observations, like what the possibility that the state either (Low, Moderate or High) .

ACKNOWLEDGMENT

This work was supported by the Pan African University

REFERENCES

[1] Nurul Nnadiah Zakaria,Mahmod Othman,Rajalingam Sokkalingam,Hanita Daud,Lazim Abdullah and Evizal Abdul Kadir(2019), *"Markov Chain Model Development for Forecasting Air Pollution Index of Miri, Sarawak"* MDPI.

[2] Aparna Bairagi, *"Markov Chain Modelling for Prediction on Future Market Price of Potatoes with Special Reference to Nagaon District." IOSR Journal of Business and Management (IOSRJBM) 19.12 (2017): 25-31*

[3] Gongsin Isaac Esbond, Fumilayo W. O. Saporu, *Introducing the Use of Markov Chain Model for Informed Decision Making for Investors in a Stock Market Environment, International Journal of Science and Research (IJSR)*

[4] F. O. Mettle, Enoch Quay, Laryea Ravenhill, (2014), *A methodology for stochastic analysis of share prices as Markov chains with finite states, SpringerPlus 3(1):657*

[5] Otieno S, Otumba EO, Nyabwanga RN, *Application of Markov Chain to Model and Forecast Stock Market Trend: A Study of Safaricom Shares in Nairobi Securities Exchange, Kenya, International Journal of Current Research. Vol. 7, Issue, 04, PP. 14712-14721, April; 2015*

[6] Zhu. X. And Xu. X. (2012).*"Markov chain analysis and prediction on the fluctuation cycle of vegetable price", Journal of System and Management Sciences ISSN 1816-6075 (Print), 1818-0523 (Online), Vol. 2 (2012) No. 3, pp. 40-49*

[7] Jasinthan, P., Laheetharan. A. And Satkunanathan. N. *"A Markov Chain Model for Vegetable Price Movement in Jaffna." Sri Lankan Journal of Applied Statistics, Vol. (16-2) , IASSL, ISSN 2424-6271 93*

[8] Gwan Seon Kim , Mehdi Nemati , Steven Buck , Nicholas Pates and Tyler Mark (2020),*Recovering Forecast Distributions of Crop Composition: Method and Application to Kentucky Agriculture*, MDPI journals

[9] AGBAM, Azubuike Samuel (F,cfia; F,cilrm) (2020), *Application of Markov Chain (MC) Model to the Stochastic Forecasting of Stocks Prices in Nigeria: The Case Study of Dangote Cement*, International Journal of Applied Science and Mathematical Theory,Vol 6. No.1 2020

[10] Zhang D, Zhang X., *Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method*, International Journal of Business and Management. Vol. 4, No. 6. June; 2009

[11] Felix Okoe Mettle, Enoch Nii Boi Quaye and Ravenhill Adjetey Laryea, *A methodology for stochastic analysis of share prices as Markov chains with finite states*, Mettle et al. a SpringerOpen Journal 2014, 3:657

[12] Lyndhurst Collins, *An Introduction to Markov Chains Analysis*, https://ibis.geog.ubc.ca/courses/geob479/spatial_analysis/1-intro-markov-chain-analysis2.pdf

[13] Analisa Teknik Pengenalan Sinyal Wicara Dengan , (2008).*Hidden Markov Model Neural Network. Jurnal litek , vol 5, no 2, 64-67*

[14] Marwala, (2000).*Probabilistic System Analyses: Part I. University of Witwatersrand, Johannesburg*

[15] Andersson Cuellar, Josephine and Fransson, Linus. 2016.*Algorithmic Trading Based on Hidden Markov Models — Hidden Markov Models as a Forecasting Tool When Trying to Beat the Market. Bachelor's thesis. University of Gothenburg. Available: https://gupea. ub.gu.se/bitstream/2077/44767/1/gupea*$_2$077$_4$4767$_1$