University of Nairobi

School of Mathematics

A Parsimonious Multivariate Markov Chain Model for NSE Stocks

This research project is submitted to the University of Nairobi in partial fulfilment of the requirements for the degree of Master of Science in Social Statistics by:

Anthony K. Mulinge

I56/70554/2011

July 2013

Declaration

This is our original work and has never been presented for any academic award in any other learning institution.

Sign: .... Date: ....

Anthony K. Mulinge

APPROVAL

This project has been submitted for examination for the degree of Master of Science (Social Statistics) of the University of Nairobi with my approval as the candidate's supervisor.

Sign: .... Date: ....

Dr. Ivivi Mwaniki

Dedication

I would like to dedicate this research project to my mother Jacinta Wayua and father Joel Mulinge for teaching me the benefits of education and inspiring me to climb the academic ladder.

Acknowledgements

It would not have been possible to write this masters project without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

First, I would like to thank my wife Gladys and daughter Abigail for their understanding and patience all the time.

My parents, Jacinta Wayua and Joel Mulinge for teaching me the benefits of education and taking me to school despite the hard times.

This research project would not have been possible without the help, support and patience of my supervisor Dr. J. I. Mwaniki, School of Mathematics University of Nairobi. I am extremely grateful for his good advice, support and friendship.

I would like to acknowledge the financial support from the Higher Education Loans Board for paying all my tuition through the postgraduate scholarship programme.

The teaching and non-teaching support from the the School of Mathematics in the University of Nairobi for always being there whenever I needed support.

Above all, I thank God for giving me the will and ability to learn.

Lastly, I take full responsibility for any errors or inadequacies that may remain in this work.

Abstract

This study uses the Parsimonious Multivariate Markov chain model to describe the dependency of transitions of the daily Volume Weighted Average Prices VWAP of Nairobi Securities Exchange prices. The model, unlike the multivariate Markov chain model, can be used for both positively and negatively associated sequence and has relatively fewer parameters. We considered 125 daily volume weighted average price (VWAP) values of three stocks (portfolios) $S_{1,}S_2$ and S_3 in the NSE for a period of 6 months starting 3rd January 2011 to 31st June 2011. From this data we obtained 124 value rates by dividing the VWAP of the day to be calculated with the value of the immediate previous trading day to obtain a three-state (1, 2 and 3 respectively) multivariate Markov chain indicating decrease, no change or increase in price.The transition probability matrices, $P^{(j,i)}$ are estimated through normalization of the transition frequency matrices of the S categorical data sequences. The model parameters $\lambda = \{\lambda_{j,i}\}$ are estimated by minimizing $||BX - X||$ under the vector norm $||\cdot||_\infty$.

Contents

Chapter 1

Introduction

1.1 Background

According to wikipaedia the phrase "stock market prediction" is defined as the act of attempting to determine the future value of a company stock or other financial instrument traded on a stock or securities market exchange. The prediction of stock prices is not easy because of the unpredictable behaviour of the prices. In the Kenyan market, daily stock prices are greatly influenced by a variety of factors including but not limited to day-to-day politics,fuel prices, exchange rate of major international currencies, inflation rates, dividend announcements, introduction of new products and product recall.

The major motivation of attempting to predict stock prices is for optimization of financial gain. In an attempt to forecast future prices traders have used various approaches based upon the following techniques:

1. Fundamental analysis,
2. Technical analysis,
3. Psychological analysis
4. Machine learning methods

The technical analysts are not concerned with any of the company's fundamentals. They seek to determine the future price of a stock based solely on the (potential) trends of the past price -which is a form of time series analysis. Technical analysis paradigm states that all price relevant information is contained in market price itself. Thus, the instant processing of market messages plays specific role, thus leading to permanent interactions among traders. Technical analysis concerns with identifications of both trends and trend reverses using more or less sophisticated procedures to predict future price movements from those of the recent past.

Fundamental analysis entails the examination of the underlying forces that affect the well being of the economy, industry groups, and companies. As with most analysis, the main aim of fundamental analysis is to derive a forecast and profit from future price movements of assets or investment portfolios.

Machine learning methods have been enhanced by technological advancement. This method uses Artificial Neural Networks, which can be thought of as mathematical functions, which simulates (mimics) the functioning of human brains when computers are fed with massive data.

Psychological analysis is based on the belief that the exchange of goods is more or less driven by human psychology and future expectations. Psychologists believe that emotions upon which expectations are built are fear and greed, which in effect compels humans to repeat the mistakes and triumphs of the past. This believe that psychology drives market prices, to a great extend, verifies the use of the technical approach in predicting market prices.

In this study we shall use a Markov chain model technique to predict future prices of selected stocks. Interest to invest in equities is fast gaining momentum at the Nairobi Securities Exchange (formerly Nairobi Stock Exchange) in Kenya. This has been shown by the recent initial public offers (IPOs) including the Safaricom Company, Kenya Electricity Generating Company, Mumias Sugar Company among others which attracted huge investments. These IPOs signified a considerable shift to the more risky and yet more profitable investment options to local and international investors. The Nairobi Securities Exchange (NSE) is the leading securities market in East and Central African region with a market capitalization of about Ksh.1.62 trillion (https://www.nse.co.ke/). With the favourable investment climate in Kenya since the end of post-election violence in 2008 and the peaceful general election in March 2013, investor confidence has greatly improved both in direct foreign investment and active participation in the securities exchange market.

1.2 Problem Statement

Investors need market information so as to know which stocks to buy, when to buy, when to sell and when to wait. However, due to uncertainties in the stock market trends, investors can only maximize returns by seriously studying the history of the listed companies, performance and development prospects of such fundamentals and be familiar with a variety of technical analysis. In our study we shall investigate how the prices of various portfolios affect the future prospects of other portfolios (and themselves) using the parsimonious multivariate Markov chain model.

1.3 Objective of study

1.3.1 The broad objective

The main objective is the application of the parsimonious multivariate Markov chain model to forecast trends of stock prices in the NSE market.

1.3.2 Specific objectives

In this study we shall endeavour to achieve the following specific objectives;

1. Fit stock market trends in a parsimonious multivariate markov chain model
2. To analyze and long-term behavior of individual stocks prices
3. To find out to what degree the selected stock affect each other

1.4 Significance of the study

In the recent years investors have started to show interest in trading in the stock several market indices in order to hedge their market risk. This has been occasioned by the ample investment climate compared to other countries in the region. With the current trend of regional or international integration as envisioned in the Millennium Development Goals (MDGs) and the Kenyan Vision 2030 it is expected that the rush for investment opportunities is going to rise. This study shall apply the idea of studying behavior of multiple sequences to determine which portfolios are likely to be more profitable in future when viewed collectively as opposed to when looked at individually. The study will, therefore, provide a tool to enable investors reap the benefits of diversification of risk by giving an insight on the level "affection" among various investment ventures in the NSE.

Chapter 2

Literature review

Previous studies have verified that stock markets have a Markov property and hence they can be modelled as random walk processes. Markov chains have been widely used in the modelling of many practical systems such as telecommunications, inventory, queuing and manufacturing systems Ching and Ng (2006). The Markov prediction model has become an indispensable tool in modern statistics with many advantages because of its “no after-effect” properties or what is popularly known as the Markovian property which has less demand for historical data.

Svoboda and Lukáś (30th ICMME) used the Markov chain analysis (MCA) to predict the trend of Prague stock exchange PX using the time series of day closing prices for a period of 60 months from January 2004. Their study compared models with different state sets. They applied filtering algorithm which involved omitting subsequently repeated states within the time series data sequence. They finally used the filtered states to get the transition matrix and a matrix of conditional probabilities of growth.

The use of the multivariate Markov chain models in social science was fully embraced early this (third) millennium and thus the numbers of studies on its application are few according to Ersoy & Semra (2011). In their study, they attempted to measure the level of "affection" of the daily closing selling prices for three foreign currencies in Turkey. They concluded that the use of multivariate Markov chain was efficient in measuring the level of affection and the prediction of foreign exchange prices for positively correlated time-discrete and state-discrete sequences.

Ching et al (2007) proposed a parsimonious multivariate Markov chain model for credit risk. Their model provided great deal of flexibility in modelling both the positive and negative associations between time series of credit ratings. Another important feature of their model was its ability to handle of shorter data sequences without compromising the efficiency of their estimates –this curbs the need for long data sequences as is the case with the multivariate Markov models. Their model had a relatively small number of parameters and was more flexible compared to the model on previous work on the same topic by Siu et al (2005). In the former model, the authors proposed a model with two parts- the positively and negatively correlated parts. They also introduced the concept of a normalizing constant in order to obtain the negatively correlated part in their theory. They also made their model more parsimonious by reducing the computational estate experience in the previous models without compromising on the prediction accuracy. Wang and Huang (2013) improved on the idea of Ching et al (2007) by introducing a new convergence condition with a new variability to improve the prediction accuracy and minimize the scale of the convergence condition which according to numerical experiments performs better than the parsimonious multivariate Markov chain model in prediction.

According to Zhang and Zhang (2009), the main difference between the Markov model and other statistical methods like time series and regression analysis is that the former does not need to find mutual laws among the factors from the complex predictor, only to consider the characteristics of the evolution on the history situation of the event itself and to predict changes of the internal state by calculating the state transition probability- which clearly shows that the Markov model has a broader applicability in stock market predictions. Additionally, Markov chain models are effective and easy to construct given the time series data and a finite number of states.

In a study on the prediction closing price trends of the Shanghai Composite index, Zhang and Zhang (2009) observed that an increase in trading days under stable conditions resulted to the convergence of the state probability to a value that is independent of the initial state and more or less stabilized. The study further used past 24 trading-day's closing prices to calculate a forecast of the subsequent day's closing price using a vector formula. After the calculation, they were able to find out that the closing price state interval after each day predicted was consistent with the actual situation.

In determining the relationship between a diverse portfolio of stocks and the entire market Doubleday and Esunge (2011) used Markov chains to show that the portfolio behaved the same way with the entire market. They further observed that when the entire market is viewed as having the Markovian property, the whole market is useful in measuring the behavior of a portfolio of stocks.

Agwuegbo et al (2011) used Markov chains method to analyze the behavior of daily return of the stock market prices of all securities listed in the

Nigeria Stock Exchange. Their study showed that the stock market follows a random walk model and that the stock prices are but martingale and that all what investors can do is to narrow differences between the fairness and otherwise in a way that high chances of small gains may be exchanged with low chances of large gains.

Chapter 3

Methodology

3.1 Assumptions

3.1.1 Assumptions for the parsimonious multivariate Markov chain model

1. The number of sequences $S \geq 2$
2. Sequences are from the same or similar source
3. A finite number of discrete states is considered
4. None of the states is absorbing

5. The model must be convergent

6. Stock markets operate in a discrete time space

3.1.2 Assumptions about the NSE market

The NSE meets the criteria of an efficient market where no trader is presented with an opportunity for making an "abnormal" return, except by chance. This assumption implies that prices are driven by multiple market forces, the fundamentals state of the stock itself, macroeconomic policy, trade and economic degrees and psychological factors of investors.

3.2 Reviews

3.2.1 The Markov chain (MC) model

A MC is a discrete stochastic process with Markov property. For a categorical data sequence X_n we let its state set to be defined by:

$$M = \{1, 2, 3, ...m\}$$

The discrete-time Markov chain with finite discrete states satisfies the following relationship:

$$\begin{aligned}\Pr(X_{n+1} &= \theta_{n+1}|X_0 = \theta_{0,}X_1 = \theta_{1,}X_2 = \theta_2, ..., X_n = \theta_n) \\ &= \Pr(X_{n+1} = \theta_{n+1}|X_n = \theta_n) \qquad (3.1)\end{aligned}$$

which is a one-step transition probability of the Markov chain. They are conditional probabilities of moving from state i at time n to state j at time $n+1$ .These probabilities are given as:

$$P^{(j,i)} = (X_{n+1} = \theta_j|X_n = \theta_i), \qquad \forall \quad i,\ j\ \in\ M \qquad (3.2)$$

$$\text{for } 0 \leq P^{(j,i)} \leq 1;\ \sum_{i=1}^{m} P^{(j,i)} = 1;\ \forall i, j \in M \qquad (3.3)$$

If we assume that $P^{(j,i)}$ are not all zero $\forall j$, we shall have the following propositions:

Proposition 1 *The matrix P has an eigen value equal to 1 and all eigen values of the matrix have a modulus less that or equal 1*

Generally ,for a non-negative matrix we have the following proposition.

Proposition 2 *Let A a non-negative irreducible square matrix of order m. Then:*

1. A has an eigen value $\lambda > 0$ such that $\lambda = \max_k |\lambda_k(A)|$, where $\lambda_k(A)$ denotes the kth eigen value of A. In other words, A has a positive eigen value λ which is equal to its spectral radius.

2. To λ there corresponds an eigen vector $\underline{z}$ of its entries being real and positive such that $A\underline{z}=\lambda\underline{z}$

3. λ is a simple eigen value of A.

From the above propositions we conclude that there exists a $\underline{z} > 0$ such that $P\underline{z}=\underline{z}$. The vector $\underline{z}$ is called a stationary vector probability vector of A and $\underline{z}_i$ is the steady-state probability at state $_i$.

3.2.2 Multivariate Markov chain (MMC) model

In many occasions,data sequences may be correlated and therefore the information of other chains can contribute to explain the captured chain data sequence. This calls for a holistic study of the bahaviour of the categorical sequences using the multivariate Markov chain (MMC) model. A multivariate

Markov chain model is used to represent behaviour of multiple categorical sequences generated by a similar source.

In the MMC model, we shall consider a set $S \geq 2$ categorical sequences each having m possible states in set $M = \{1, 2, 3, ..., m\}$.The state probability distribution of the jth sequence at time $n + 1$ depends on the state probabilities of all the sequences (including itself) at time n. Let $X_{n+1}^{(j)}$be the state probability vector of the jth sequence at time $n + 1$, $X_n^{(i)}$ be the state probability vector of the ith sequence at time n and $X_0^{(i)}$ the initial probability distribution of the ith sequence. The Then, if $P^{(j,i)}$ is the transition probability from the state in sequence i at time n to the the state in sequence j at time $n + 1$, Ching et al (2002) defined the Multivariate Markov Chain model as follow:

$$X_{n+1}^{(j)} = \sum_{i=1}^{s} \lambda_{j,i} P^{(j,i)} X_n^{(i)} \tag{3.4}$$

which in matrix form can be expressed as

$$
\begin{aligned}
X_{n+1}^{(j)} &= \begin{bmatrix} X_{n+1}^{(1)} \\ X_{n+1}^{(2)} \\ \cdot \\ \cdot \\ \cdot \\ X_{n+1}^{(s)} \end{bmatrix} = \begin{bmatrix} \lambda_{1,1}P^{(1,1)} & \lambda_{1,2}P^{(1,2)} & \cdot\ \cdot\ \cdot & \lambda_{1,s}P^{(1,s)} \\ \lambda_{2,1}P^{(2,1)} & \lambda_{2,2}P^{(2,2)} & \cdot\ \cdot\ \cdot & \lambda_{2,s}P^{(2,s)} \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \lambda_{s,1}P^{(s,1)} & \lambda_{s,2}P^{(s,2)} & \cdot\ \cdot\ \cdot & \lambda_{s,s}P^{(s,s)} \end{bmatrix} \times \begin{bmatrix} X_{n}^{(1)} \\ X_{n}^{(2)} \\ \cdot \\ \cdot \\ \cdot \\ X_{n}^{(s)} \end{bmatrix} \\
X_{n+1}^{(j)} &\equiv \Lambda \mathbf{X}_n
\end{aligned}
$$

where $\lambda_{j,i} \geq 0, \ \sum_{i=1}^{s} \lambda_{j,i} = 1 \ ;\forall\ j,\ i = 1, 2, 3, ..., s.$

We notice that although the row sum for Λ do not equal to one (the row sum of $P^{(j,i)}$ is one) we still have the following proposition.

Proposition 3 *Suppose that $P^{(j,i)}$ are irreducible $\forall\ 1 \leq j, i \leq s$, and $\lambda_{j,i} > 0$ for $1 \leq j, i \leq s$. Then there exists a vector*

$$\mathbf{X}_n = \left[\mathbf{X}_n^{(1)},\ \mathbf{X}_n^{(2)},\ \mathbf{X}_n^{(3)}, ...,\ \mathbf{X}_n^{(s)}\right]^T$$

such that $X_n = \Lambda X_n$ and

$$\sum_{i=1}^{m} \left[\mathbf{X}_n^{(j)}\right]^i = 1,\ 1 \leq j \leq s \tag{3.5}$$

where $[\cdot]^i$denote the ith entry corresponding of the corresponding vector.

The vector in the above proposition contains the stationary probability distributions for the VWAP in the three stocks. In other words, for each sequence j, there is a vector $X_*^{(j)}$ which represents the probability distribution for the jth portfolio (stock) in the long-run.

The transition probability matrices, $P^{(j,i)}$ are estimated through normalization of the transition frequency matrices of the S categorical data sequences. The model parameters $\lambda = \{\lambda_{j,i}\}$ are estimated by minimizing $||BX - X||$ under the vector norm $||\cdot||_\infty$.

With the assumption that $\lambda_{j,i} \geq 0$, the MMC model only allows positive correlation among data sequences, that is, cases where an increase in the state probability in any of the sequences at time n can only increase (but never decrease) the state probabilities at time $n+1$. Additionally, the number of parameters for the MMC increases exponentially with increase in the number of categorical sequences. For that reason, Ching et al.(2006) proposed a first-order multivariate Markov chain model for modeling the sales demand of multiple products in a soft drink company. Their model contained $O(S^2m^2 + s^2)$ number of parameters where S is the number of sequences and m is the number of possible states. Their model captured both the intra- and inter-transition probabilities among the sequences.

3.2.3 The Parsimonious Multivariate Markov Chain (PMMC) model

In order to take care of the shortcomings mentioned above and to extend the MMC,Ching et al (2007) considered the following equation to model the case when the state probability vector X_n is negatively associated with a state probability vector Z_{n+1}.

$$Z_{n+1} = \frac{(\mathbf{v} - X_n)}{m - 1} \tag{3.6}$$

where v is a vector of all ones, $(m-1)^{-1}$ is a normalizing constant for $m \geq 2$ states in each sequence.

With notations similar as the ones introduced in the previous section, the improved parsimonious multivariate Markov chain model is defined as

$$X_{n+1}^{(j)} = \Lambda^{+} X_n^{(i)} + (m-1)^{-1} \Lambda^{-} (\mathbf{v} - X_n^{(i)}) \tag{3.7}$$

which in matrix form shall be

$$\begin{bmatrix} X_{n+1}^{(j)} \\ X_{n+1}^{(j)} \\ . \\ . \\ . \\ X_{n+1}^{(j)} \end{bmatrix} = \Lambda^{+} \begin{bmatrix} X_n^{(j)} \\ X_n^{(j)} \\ . \\ . \\ . \\ X_n^{(j)} \end{bmatrix} + (m-1)^{-1} \cdot \Lambda^{-} \begin{bmatrix} \mathbf{v} - X_n^{(j)} \\ \mathbf{v} - X_n^{(j)} \\ . \\ . \\ . \\ \mathbf{v} - X_n^{(j)} \end{bmatrix}$$

where $\Lambda^{+} X_n$ and $(m-1)^{-1} \cdot \Lambda^{-} X_n$ are respectively the positively and negatively correlated parts of the transition probabilities.

$$\Lambda^{+} = \begin{bmatrix} \lambda_{1,1}P^{(1,1)} & \lambda_{1,2}P^{(1,2)} & \cdot & \cdot & \cdot & \lambda_{1,s}P^{(1,s)} \\ \lambda_{2,1}P^{(2,1)} & \lambda_{2,2}P^{(2,2)} & \cdot & \cdot & \cdot & \lambda_{2,s}P^{(2,s)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda_{s,1}P^{(s,1)} & \lambda_{s,2}P^{(s,2)} & \cdot & \cdot & \cdot & \lambda_{s,s}P^{(s,s)} \end{bmatrix} \quad (3.8)$$

$$\Lambda^{-} = \begin{bmatrix} \lambda_{1,-1}P^{(1,1)} & \lambda_{1,-2}P^{(1,2)} & \cdot & \cdot & \cdot & \lambda_{1,-s}P^{(1,s)} \\ \lambda_{2,-1}P^{(2,1)} & \lambda_{2,-2}P^{(2,2)} & \cdot & \cdot & \cdot & \lambda_{2,-s}P^{(2,s)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda_{s,-1}P^{(s,1)} & \lambda_{s,-2}P^{(s,2)} & \cdot & \cdot & \cdot & \lambda_{s,-s}P^{(s,s)} \end{bmatrix} \quad (3.9)$$

$$\sum_{i=-s}^{s} \lambda_{j,i} = 1,\ \lambda_{j,i} \geq 0,\ \forall\ 1 \leq j \leq s,\ 1 \leq |i| \leq s.$$

As observed in the previous section, the rows of $\Lambda^{+}$and Λ^{-} do not necessarily sum one. However, Proposition 3 still holds and there exists two vectors

$$\mathbf{X}_n = \left[\mathbf{X}_n^{(1)},\ \mathbf{X}_n^{(2)},\ \mathbf{X}_n^{(3)}, ...,\ \mathbf{X}_n^{(s)}\right]^T$$

and

$$\mathbf{Z}_n = \left[\mathbf{Z}_n^{(1)},\ \mathbf{Z}_n^{(2)},\ \mathbf{Z}_n^{(3)}, ...,\ \mathbf{Z}_n^{(s)}\right]^T$$

each having a stationary probability vector (X_* and Z_* respectively) as the time (n) increases.

in order to reduce the number of parameters in the model (and hence to make the model more parsimonious) we set $P^{(j,i)} = I$ for $i \neq j$. This idea was adopted, justified and shown to be effective in the forecasting of sales demand by Ching W, Zhang S. and Ng M. (2006).

Lemma 4 *Let $A \in \mathbb{R}^{m\times m}$ be non-negative and irreducible matrix, $B \in \mathbb{C}^{m\times m}$ a complex matrix and λ an eigen value of B. If $|A| > B$, then $\rho(A) > |\lambda|$.*

Application of the PMMC in the NSE market

In this study, we shall consider 125 daily volume weighted average price (VWAP) values of three stocks (portfolios) $S_{1,}S_2$ and S_3 in the NSE for a period of 6 months starting 3rd January 2011 to 31st June 2011. The VWAP is obtained by dividing the turnover per counter by total number of shares traded. To determine the change values, we form data sequences by taking data of the previous day into consideration owing to the structure of the Markov chains.

For this data we obtained 124 value rates by dividing the VWAP of the day to be calculated with the value of the immediate previous trading day. Then we note if the quotient is less than 1, equal to 1 or greater than 1. These three values shall form a three-state (1, 2 and 3 respectively) multivariate Markov chain.

Parameter Estimation for the PMMC model

Preliminary Analysis

Testing for association As stated in section 3.1, one of the assumptions of the MMC model is the correlation of the categorical data sequences. In order to ascertain this assumption for our data we shall we perform simple linear regression and correlation analysis to the VWAP to determine the direction and degree of affection between the six possible combinations of three sequences. In simple regression, we have only two variables which are assumed to have a real linear relationship between them. The two variables are such that one is independent (which for convenience here we shall call Y) while the other one is dependent (X). A simple linear regression equation connecting X and Y is of the form:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon \tag{3.10}$$

where β_0 is the constant coefficient that shows the intersection point, $\beta_1$is the slope coefficient and ε is the random noise or the error term.

The coefficients and can be obtained by solving the following least squares normal equations:

$$\sum Y = n\beta_0 + \beta_1 \sum X \tag{3.11}$$

$$\sum XY = \beta_0 \sum X + \beta_1 \sum X^2 \tag{3.12}$$

The degree of linear relationship between X and Y is measured by the correlation coefficient (r) whose value ranges between $\pm 1$where r values of $+1$ and -1indicate perfect positive and perfect negative correlations respectively while a value of 0 indicates absence of correlation. The square of the coefficient of correlation (called the coefficient of determination or just " R-Squared") gives a measure used in statistical model analysis to assess how well a model explains and predicts future outcomes. The correlation coefficient is defined as:

$$r = \frac{\sum X_i Y_i}{\sqrt{\sum X_i^2 \cdot \sum Y_i^2}} \tag{3.13}$$

The Durbin–Watson statistic The Durbin–Watson statistic (d) is a test statistic used to detect the correlation of error terms after estimating the regression model. This statistic ranges from 0 to 4 and if it is 2, this signifies the presence of autocorrelation (a relationship between values separated from each other by a given time lag) in the residual terms from a regression analysis. An acceptable range is 1.50 - 2.50. Where successive error differences are small, Durbin-Watson is low (less than 1.50); this indicates the presence of positive autocorrelation. Positive autocorrelation is very common. Where successive error differences are large, Durbin-Watson is high (more than 2.50); this indicates the presence of negative autocorrelation. Negative autocorrelation is not particularly common.

If the Durbin-Watson statistic is greater than R- square, it is likely that autocorrelation exists. Autocorrelation indicates that the forecast model could be improved on. In time series with lagged variables, the Durbin-Watson statistic is unreliable as it tends toward a value of 2.0.

This statistic is calculated using the formula:

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=2}^{T} e_t^2} \tag{3.14}$$

Model Parameter Estimation First, we estimate the transition probability $P^{(j,i)}$. If data sequences are given and the state set is $M = \{1, 2, 3, ..., m\}$, $f_{k_j,k_i}^{(j,i)}$ is the frequency from the k_i state in the ith sequence at time n to the k_j state in the jth sequence at time $n+1$ with $k_j, k_i \in M$, the transition frequency matrix can be represented as:

$$F^{(j,i)} = \begin{bmatrix} f_{1,1}^{(j,i)} & f_{1,2}^{(j,i)} & . & . & . & f_{1,s}^{(j,i)} \\ f_{2,1}^{(j,i)} & f_{2,2}^{(j,i)} & . & . & . & f_{2,s}^{(j,i)} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ f_{s,1}^{(j,i)} & f_{s,2}^{(j,i)} & . & . & . & f_{s,s}^{(j,i)} \end{bmatrix}_{\mathrm{m \times m}} \tag{3.15}$$

$P^{(j,i)}$ shall be obtained by normalizing the frequency transition matrix as:

$$P^{(j,i)} = \begin{bmatrix} p_{1,1}^{(j,i)} & p_{1,2}^{(j,i)} & . & . & . & p_{1,s}^{(j,i)} \\ p_{2,1}^{(j,i)} & p_{2,2}^{(j,i)} & . & . & . & p_{2,s}^{(j,i)} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ p_{s,1}^{(j,i)} & p_{s,2}^{(j,i)} & . & . & . & p_{s,s}^{(j,i)} \end{bmatrix}_{\mathrm{m \times m}} \tag{3.16}$$

where

$$p_{k_j,k_i}^{(j,i)} = \begin{cases} \frac{f_{k_j,k_i}^{(j,i)}}{\sum_{k_i=1}^{m} f_{k_j,k_i}^{(j,i)}} & if \quad \sum_{k_i=1}^{m} f_{k_j,k_i}^{(j,i)} \neq 0 \\ \\ \frac{1}{m} & Otherwise \end{cases} \tag{3.17}$$

is the maximum likelihood estimator.

We shall then investigate the existence of a stationary distribution under the Parsimonious multivariate Markov chain model, the rate of convergence and how to speed up the rate of convergence. First, we represent the new multivariate Markov chain model as the following vector-valued difference equation:

$$X_{n+1} = \begin{bmatrix} X_{n+1}^{(1)} \\ X_{n+1}^{(2)} \\ . \\ . \\ . \\ X_{n+1}^{(s)} \end{bmatrix}$$

$$
= \begin{bmatrix} M_{1,1} & M_{1,2} & . & . & . & M_{1,s} \\ M_{2,1} & M_{2,2} & . & . & . & M_{2,s} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ M_{s,1} & M_{s,2} & . & . & . & M_{s,s} \end{bmatrix} \begin{bmatrix} X_n^{(1)} \\ X_n^{(2)} \\ . \\ . \\ . \\ X_n^{(s)} \end{bmatrix} + \tfrac{1}{m-1} \begin{bmatrix} J_{1,-1} & J_{1,-2} & . & . & . & J_{1,-s} \\ J_{2,-1} & J_{2,-2} & . & . & . & J_{2,-s} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ J_{s,-1} & J_{s,-2} & . & . & . & J_{s,-s} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{bmatrix}
$$

$$
\equiv M_s X_n + b
$$

where

$$
M_{j,i} = \begin{cases} (\lambda_{j,i} - \frac{\lambda_{j,-i}}{m-1}) P^{(j,i)} & \text{if} \quad i = j \\ \\ (\lambda_{j,i} - \frac{\lambda_{j,-i}}{m-1}) \mathbf{I} & \textbf{if} \quad i \neq j \end{cases}
$$

$$
J_{j,i} = \begin{cases} \lambda_{j,-i} P^{(j,i)} & \text{if} \quad i = j \\ \\ \lambda_{j,-i} P^{(j,i)} & \text{if} \quad i \neq j \end{cases}
$$

We notice that

$$
X_{n+1} = M_s^2 X_{n-1} + (I + M_s)b = M_s^3 X_{n-2} + (I + M_s + M_s^2)
$$

$$
= M_s^{n+1} X_0 + \textstyle\sum_{k=0}^{n} M_s^k b \qquad \text{, where } I = M_s^0
$$

The parsimonious multivariate Markov chain model has a stationary distribution if for $||M_s|| < 1$ we have

$\lim_{n\to\infty} X_n = \lim_{n\to\infty} \sum_n^k b = \frac{1}{\mathbf{I}-\mathbf{M}_s} b$

We also note that

$||M_s||_\infty \leq \max_{1\leq k\leq s} \left\{ m|\lambda_{j,j} - \frac{\lambda_{j,j}}{m-1}| + \sum_{j\neq i} |\lambda_{j,i} - \frac{\lambda_{j,i}}{m-1}| \right\}.$

The convergence rate of the process to stationary distribution can be speeded up by controlling the value of $||M_s||_\infty$.To achieve this we shall impose an upper bound $\alpha < 1$ and introduce the following additional constraints

$m|\lambda_{j,j} - \frac{\lambda_{j,j}}{m-1}| + \sum_{j\neq i} |\lambda_{j,i} - \frac{\lambda_{j,i}}{m-1}| \leq \alpha$

We observe that if a smaller value of α is chosen, the rate of convergence to the stationary distribution becomes faster. Therefore,we can obtain reasonably accurate estimates for the unknown parameters even when the dataset is short.

One possible way of estimating $\lambda_{(j,i)}$ is given as follows as proposed by Ching W. et al (2005) is to formulate S linear programming problems (Chvatal V. Linear programming ,Freeman,83). In order to avoid gross discrepancies with the data we choose the vector norm $|| \bullet ||_\infty$ to minimize the discrepancies. We have the following optimization problem:

$\min_{\lambda_{j,i}} \left\{ \max_i | \left[\sum_{i=1}^m \lambda_{j,i} P^{(j,i)} \mathbf{X}^{(i)} - \mathbf{X}^{(j)} \right] | \right\}$

Subject to:

$$
\begin{aligned}
\sum_{i=-s}^{s} \lambda_{j,i} &= 1,\ \lambda_{j,i} \geq 0,\ \forall j = 1,2,3,...,s \\
\lambda_{j,i} &\geq 0,\ \forall i = \pm 1, \pm 2, ..., \pm s,\ \forall j = 1,2,3,...,s \\
m|\lambda_{j,j} - \frac{\lambda_{j,-j}}{m-1}| + \sum_{j \neq i} |\lambda_{j,i} - \frac{\lambda_{j,-i}}{m-1}| &\leq \alpha,\ j = 1,2,3,...,s.
\end{aligned}
$$

The PMMC model has a state vector, X_n (at time n) which can be estimated from the sequences by computing the proportion of occurrence of each state within each sequence. This vector shall be defined as :

$$X_n = \left((X_n^{(1)})^T,\ (X_n^{(2)})^T, ..., (X_n^{(s)})^T\right)^T \ , \ \in \mathbb{R}^{sm \times 1}$$

which satisfies: $X_{n+1} = \Lambda^{+} X_n + (m\beta - 1)^{-1} \cdot \Lambda^{-}(\beta v - X_n)$

where Λ^{+} and Λ^{-} are as defined in equation 3.8 and 3.9 respectively.

In the next chapter we shall perform preliminary data analysis to ascertain that the sequences are correlated and then perfom the main analysis to estimate the transition probabilities. Lastly, using linear programming techniques we shall estimate λ.

Chapter 4

Data Analysis and Results

4.1 Preliminary Analysis

4.1.1 Testing relationship of sequences

1. The condition that the VWAP of S_1 is dependent while the VWAP of S_2 is independent:

Table 1: Testing of the Coefficients

S_1	Coeff	Std Err	t	P>|t |	95% CI
S_2	1.187	0.072	16.57	0.000	1.046 1.329
Constant	-21.47	1.906	-11.27	0.000	-25.242 -17.698

According to the findings derived in Table 1, both the constant coefficient in the regression equation and the coefficient of S_2 which represents the independent variable are highly significant. The regression equation shall be estimated by:

$$S_1 = -21.47 + 1.187S_2$$

Table 2: Analysis of Variance

Source	SS	df	MS
Model	513.396	1	513.396
Residual	229.871	123	1.869
Total	743.267	124	5.994

According to Table 2, the equation $S_1 = -21.47 + 1.187S_2$ formed on the basis of data in Table 1 is highly significant (Sig. $0.0000 < \alpha = 0.05$). The results also show that the change in the VWAP of $S_1$is 69.1% depended on the VWAP of S_2 ,implying that there is a 83.1% positive relationship between the VWAPs of S_1 and S_2.The Durbin Watson test statistic is 0.1605 which implies presence of positive first-order serial correlation between error terms.

2. The condition that the VWAP of S_1 is independent while the VWAP of S_2 is dependent:

Table 3: Testing of Coefficients

S_2	Coeff	Std Err	t	P>|t |	95% CI	
S_1	0.582	0.035	16.57	0.000	0.512	0.651
Constant	20.700	0.363	57.03	0.000	19.982	21.419

According to the results derived from Table 3, the constant term and the coefficient of the VWAP of S_1 are highly significant at 95% confidence level.

The regression of the VWAP of S_2 on the VWAP of S_2 can be estimated by the equation:

$$S_2 = 20.700 + 0.582S_1$$

Table 4: Analysis of Variance

Source	SS	df	MS
Model	251.545	1	251.545
Residual	112.628	123	0.916
Total	364.173	124	2.937

The results in Table 4 show that the equation $S_2 = 20.700 + 0.582S_1$ obtained from data in Table 3 is highly significant (Sig $0.0000 < \alpha = 0.05$). Further, the results show that the change in the VWAP of S_2 is 69.1% depended on the VWAP of S_1 ,implying that there is a 83.1% positive relationship between the VWAPs of S_2 and S_1.The Durbin Watson test statistic is 0.1984, which implies presence of positive first-order serial correlation between error terms.

3. The condition that VWAP of S_1 is dependent while the VWAP of S_3 is independent

Table 5: Testing of Coefficients

S_1	Coeff	Std Err	t	P>|t |	95% CI
S_3	8.411	0.302	27.88	0.000	7.812 9.008
Constant	-24.534	1.243	-19.74	0.000	-26.995

The results in Table 5 show that both the constant term and the coefficient of the VWAP of $S_3$are highly significant at 95% level of confidence. The regression of the VWAP of S_1 on the VWAP of S_3 will be estimated by the equation:

$$S_1 = -24.534 + 8.411S_3$$

Table 6: Analysis of Variance

Source	SS	df	MS
Model	641.717	1	641.717
Residual	101.550	123	0.826
Total	743.267	124	5.994

According to the results in Table 6, the regression equation $S_1 = -24.534 + 8.411S_3$ of the VWAP of S_1 on the VWAP of S_3 is highly significant at 95% level of confidence. Further, the results show that the change in the VWAP of S_1 is 86.3% depended on the VWAP of S_3 ,implying that there is a 92.9% positive relationship between the VWAPs of S_1 and S_3.The Durbin Watson test statistic is 0.4673, which implies the presence of positive first-order serial correlation between error terms.

4. Condition when the VWAP of S_1 is independent while the VWAP of S_3 is dependent:

Table 7: Testing of Coefficients

S_3	Coeff	Std Err	t	P>|t |	95% CI
S_1	0.103	0.004	27.88	0.000	0.095 0.110
Constant	3.08	0.038	80.90	0.000	3.005 3.155

The results in Table 7 show that both the constant term and the coefficient of the VWAP of $S_1$are highly significant at 95% level of confidence. The regression of the VWAP of S_3 on the VWAP of S_1shall be estimated by the equation:

$$S_3 = 3.08 + 0.103S_1$$

Table 8: Analysis of variance

Source	SS	df	MS
Model	7.831	1	7.831
Residual	1.239	123	0.010
Total	9.071	124	0.073

The results in Table 8 show that the regression of the VWAP of S_3 on the VWAP of S_1 given by the equation $S_3 = 3.08 + 0.103S_1$ is highly significant at 95% confidence level. The results also show that the change in the VWAP of S_3 is 86.3% depended on the VWAP of S_1 ,implying that there is a 92.9% positive relationship between the VWAPs of S_3 and S_1.The Durbin Watson statistic is 0.5179, which signifies the presence of positive first-order serial correlation between error terms.

5. Condition when VWAP of S_2 is dependent while the VWAP of S_3 is independent:

Table 9: Testing of Coefficients

S_2	Coeff	Std Err	t	P>|t |	95% CI
S_3	4.899	0.362	13.52	0.000	4.181 5.616
Constant	6.405	1.493	4.29	0.000	3.449 9.361

For the results in Table 9, both the constant term and the coefficient of S_3 are highly significant at 95% confidence level. The resulting estimate of the regression of the VWAP of S_2 on the VWAP of S_3 shall be given by:

$$S_2 = 6.405 + 4.899S_3$$

Table 10: Analysis of Variance

Source	SS	df	MS
Model	217.659	1	217.659
Residual	146.513	123	1.191
Total	364.173	124	2.937

The results in Table 10 show that the coefficients of the regression equation $S_2 = 6.405 + 4.899S_3$ are significant at 95% level of confidence. The results also show that the change in the VWAP of S_2 is 59.8% depended on the VWAP of S_3 ,implying that there is a 77.39% positive relationship between the VWAPs of S_2 and S_3.The Durbin Watson statistic of 0.2399, which implies the presence of positive first-order serial correlation between error terms.

6. The condition when the VWAP of S_3 is dependent while the VWAP of S_2 is independent

Table 11: Testing of Coefficients

S_3	Coeff	Std Err	t	P>|t |	95% CI
S_2	0.122	0.009	13.52	0.000	0.104 0.140
Constant	0.873	0.240	3.63	0.000	0.397 1.348

According to the results in Table 11, both the constant term and the coefficient of the VWAP of S_2 are highly significant at 95% level of confidence. From the results, regression of VWAP of S_3on the VWAP of S_2may be estimated using the equation:

$$S_3 = 0.873 + 0.122S_2$$

Table 12: Analysis of Variance

Source	SS	df	MS
Model	5.421	1	5.421
Residual	3.649	123	0.030
Total	9.071	124	0.073

Table 12 shows that the regression equation $S_3 = 0.873 + 0.122S_2$ is highly significant at 95% level of confidence. The results also show that the change in the VWAP of S_3 is 59.8% depended on the VWAP of S_2 ,implying that there is a 77.39% positive relationship between the VWAPs of S_3 and S_2.The Durbin Watson statistic is 0.2272, which implies the presence of positive first-order serial correlation between error terms.

4.1.2 The Durbin-Watson test

1. The condition that the VWAP of S_1 is dependent while the VWAP of S_2 is independent:

Durbin-Watson test

data: fit1

DW = 0.1605, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

2. The condition that the VWAP of S_1 is independent while the VWAP of S_2 is dependent:

Durbin-Watson test

data: fit2

DW = 0.1984, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

3. The condition that VWAP of S_1 is dependent while the VWAP of S_3 is independent

Durbin-Watson test

data: fit3

DW = 0.4673, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

4. Condition when the VWAP of S_1 is independent while the VWAP of S_3 is dependent:

Durbin-Watson test

data: fit4

DW = 0.5179, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

5. Condition when VWAP of S_2 is dependent while the VWAP of S_3 is independent:

Durbin-Watson test

data: fit5

DW = 0.2399, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

6. The condition when the VWAP of S_3 is dependent while the VWAP of S_2 is independent

Durbin-Watson test

data: fit6

DW = 0.2272, p-value < 2.2e-16

alternative hypothesis: true autocorrelation is greater than 0

4.2 Model Parameter Estimation

4.2.1 Transition Frequencies

The transition frequency matrixes $F^{(j,i)}$ for all $j = i$ indicate $intra-sequence$ first-lag state transition frequency counts and $F^{(j,i)}$ for all $j \neq i$ indicate

$inter - sequence$ first-lag state transition frequency counts from the three sequences (S_1, S_2 and S_3) . The transition frequencies are as shown below:

$$F^{(1,1)} = \begin{bmatrix} 43 & 8 & 19 \\ 9 & 1 & 3 \\ 19 & 4 & 17 \end{bmatrix} \qquad F^{(2,1)} = \begin{bmatrix} 27 & 23 & 20 \\ 5 & 5 & 3 \\ 11 & 14 & 15 \end{bmatrix}$$

$$F^{(3,1)} = \begin{bmatrix} 26 & 28 & 16 \\ 6 & 3 & 4 \\ 14 & 8 & 18 \end{bmatrix} \qquad F^{(2,2)} = \begin{bmatrix} 19 & 11 & 13 \\ 15 & 18 & 9 \\ 9 & 13 & 16 \end{bmatrix}$$

$$F^{(1,2)} = \begin{bmatrix} 24 & 4 & 15 \\ 25 & 3 & 14 \\ 22 & 6 & 10 \end{bmatrix} \qquad F^{(3,2)} = \begin{bmatrix} 17 & 15 & 11 \\ 15 & 15 & 12 \\ 14 & 9 & 15 \end{bmatrix}$$

$$F^{(3,3)} = \begin{bmatrix} 20 & 8 & 18 \\ 12 & 18 & 9 \\ 14 & 13 & 11 \end{bmatrix} \qquad F^{(1,3)} = \begin{bmatrix} 26 & 6 & 14 \\ 25 & 4 & 10 \\ 20 & 3 & 15 \end{bmatrix}$$

$$F^{(2,3)} = \begin{bmatrix} 16 & 18 & 12 \\ 14 & 10 & 15 \\ 13 & 14 & 11 \end{bmatrix}$$

4.2.2 Transition Probabilities

Transition probabilities $P^{(j,i)}$ are matrixes are estimated by normalizing the transition frequency matrices in the previous section. The transition probabilities are as shown below:

$$P^{(1,1)} = \begin{bmatrix} 0.614 & 0.114 & 0.271 \\ 0.692 & 0.077 & 0.231 \\ 0.475 & 0.100 & 0.425 \end{bmatrix} \quad P^{(2,1)} = \begin{bmatrix} 0.386 & 0.329 & 0.286 \\ 0.385 & 0.385 & 0.231 \\ 0.275 & 0.350 & 0.375 \end{bmatrix}$$

$$P^{(3,1)} = \begin{bmatrix} 0.371 & 0.400 & 0.229 \\ 0.462 & 0.231 & 0.308 \\ 0.350 & 0.200 & 0.450 \end{bmatrix} \quad P^{(2,2)} = \begin{bmatrix} 0.442 & 0.256 & 0.302 \\ 0.357 & 0.429 & 0.214 \\ 0.237 & 0.342 & 0.421 \end{bmatrix}$$

$$P^{((1,2)} = \begin{bmatrix} 0.558 & 0.093 & 0.349 \\ 0.595 & 0.071 & 0.333 \\ 0.579 & 0.158 & 0.263 \end{bmatrix} \quad P^{(3,2)} = \begin{bmatrix} 0.395 & 0.349 & 0.256 \\ 0.357 & 0.357 & 0.286 \\ 0.368 & 0.237 & 0.395 \end{bmatrix}$$

$$P^{(3,3)} = \begin{bmatrix} 0.435 & 0.174 & 0.391 \\ 0.308 & 0.462 & 0.231 \\ 0.368 & 0.342 & 0.289 \end{bmatrix} \quad P^{(1,3)} = \begin{bmatrix} 0.565 & 0.130 & 0.304 \\ 0.641 & 0.103 & 0.256 \\ 0.526 & 0.079 & 0.395 \end{bmatrix}$$

$$P^{(2,3)} = \begin{bmatrix} 0.348 & 0.391 & 0.261 \\ 0.359 & 0.256 & 0.385 \\ 0.283 & 0.304 & 0.239 \end{bmatrix}$$

4.2.3 State Probability Vectors

The state probability vectors $X_n^{(1)}$, $X_n^{(2)}$ and $X_n^{(3)}$ for the sequences S_1, S_2 and S_3 respectively shall be estimated by the proportion of occurrence of each state within individual sequences as shown below:

$$X_n^{(1)} = \begin{pmatrix} \frac{71}{124} & \frac{13}{124} & \frac{40}{124} \end{pmatrix}^T = \begin{pmatrix} 0.573 & 0.105 & 0.323 \end{pmatrix}^T$$
$$X_n^{(2)} = \begin{pmatrix} \frac{43}{124} & \frac{42}{124} & \frac{39}{124} \end{pmatrix}^T = \begin{pmatrix} 0.347 & 0.339 & 0.315 \end{pmatrix}^T$$

$$X_n^{(3)} = \begin{pmatrix} \frac{46}{124} & \frac{40}{124} & \frac{38}{124} \end{pmatrix}^T = \begin{pmatrix} 0.371 & 0.323 & 0.306 \end{pmatrix}^T$$

Thus, from the above three state probability vectors we can define

$$X_n = \left(X_n^{(1)}, X_n^{(2)}, X_n^{(3)}\right)^T$$
$$= \begin{bmatrix} 0.573 & 0.105 & 0.323 \\ 0.347 & 0.339 & 0.315 \\ 0.371 & 0.323 & 0.306 \end{bmatrix}$$

We note that the vector $X_n \rightarrow$ a stationary probability vector X_* as time (n) increases a shown below:

$$X_* = \begin{bmatrix} 0.46 & 0.228 & 0.317 \\ 0.46 & 0.228 & 0.317 \\ 0.46 & 0.228 & 0.317 \end{bmatrix}$$

Define

$$B_{j,i} = \left[P^{(j,1)} \cdot X_n^{(1)} | P^{(j,2)} \cdot X_n^{(2)} | ... | P^{(j,s)} \cdot X_n^{(s)} | \right],$$

$\forall j = 1, 2, 3, ..., s.$

In our study the matrices will be as follows:

$$B_1 = \left[P^{(1,1)} \cdot X_n^{(1)} | P^{(1,2)} \cdot X_n^{(2)} | P^{(1,3)} \cdot X_n^{(3)} \right]$$
$$= \begin{bmatrix} 0.451 & 0.347 & 0.371 \\ 0.479 & 0.339 & 0.323 \\ 0.420 & 0.315 & 0.306 \end{bmatrix}$$

$$B_2 = \left[P^{(2,1)} \cdot X_n^{(1)} | P^{(2,2)} \cdot X_n^{(2)} | P^{(2,3)} \cdot X_n^{(3)} \right]$$
$$= \begin{bmatrix} 0.573 & 0.335 & 0.371 \\ 0.105 & 0.337 & 0.323 \\ 0.323 & 0.331 & 0.306 \end{bmatrix}$$

$$B_3 = \left[P^{(3,1)} \cdot X_n^{(1)} | P^{(3,2)} \cdot X_n^{(2)} | P^{(3,3)} \cdot X_n^{(3)}\right]$$
$$= \begin{bmatrix} 0.573 & 0.347 & 0.337 \\ 0.105 & 0.339 & 0.334 \\ 0.323 & 0.315 & 0.335 \end{bmatrix}$$

In order for our model to apply for both positively and negatively correlated data sequences we generate a state probability vector $Z_{n+1} = \left[Z_n^{(1)}, Z_n^{(2)}, Z_n^{(3)}\right] = \frac{1}{m-1}(v - X_n)$ which is negatively correlated to X_n, where $(m-1)^{-1}$is a normalizing constant for $m = 3$ for number of states in our data, v is 3×3 matrix with all entries equal to 1 and X_n is the state probability distribution vector for the three sequences. Thus,

$$\mathbf{Z}_n = \begin{bmatrix} 0.214 & 0.448 & 0.338 \\ 0.326 & 0.330 & 0.342 \\ 0.314 & 0.338 & 0.347 \end{bmatrix}^T$$

which is a state distribution probability matrix with each column representing a state probability vector $Z_n^{(1)}$, $Z_n^{(2)}$ and $Z_n^{(3)}$ which is negatively correlated to $X_n^{(1)}$, $X_n^{(2)}$ and $X_n^{(3)}$respectively. We notice that

$$Z_n \to Z_* = \begin{bmatrix} 0.19 & 0.241 & 0.225 \\ 0.19 & 0.241 & 0.225 \\ 0.19 & 0.241 & 0.225 \end{bmatrix}$$

which is a stationary probability matrix as n increases.

Using the definition of $B_{j,i}$ given earlier, we shall get :

$$B_{1*} = \left[P^{(1,1)} \cdot Z_n^{(1)} | P^{(1,2)} \cdot Z_n^{(2)} | P^{(1,3)} \cdot Z_n^{(3)}\right]$$
$$= \begin{bmatrix} 0.274 & 0.326 & 0.314 \\ 0.261 & 0.330 & 0.338 \\ 0.290 & 0.342 & 0.347 \end{bmatrix}$$

$$B_{2*} = \left[P^{(2,1)} \cdot Z_n^{(1)} | P^{(2,2)} \cdot Z_n^{(2)} | P^{(2,3)} \cdot Z_n^{(3)}\right]$$
$$= \begin{bmatrix} 0.214 & 0.332 & 0.314 \\ 0.448 & 0.331 & 0.338 \\ 0.338 & 0.334 & 0.347 \end{bmatrix}$$

$$B_{3*} = \left[P^{(3,1)} \cdot Z_n^{(1)} | P^{(3,2)} \cdot Z_n^{(2)} | P^{(3,3)} \cdot Z_n^{(3)}\right]$$
$$= \begin{bmatrix} 0.214 & 0.326 & 0.331 \\ 0.448 & 0.330 & 0.333 \\ 0.338 & 0.342 & 0.331 \end{bmatrix}$$

4.2.4 Solving optimization problems

In to get $\lambda_{j,i}$ for $j = 1, 2, 3, ..., s$ and $i = \pm 1, \pm 2, \pm 3, ..., \pm s$, we solve the following optimization problem using the Lingo Lindo 14.0 software.

$$\min w_j(\lambda)$$

Subject to:

$$\begin{aligned}
w_j &\geq X_n^{(j)} - B_j\lambda_{j,i} \\
w_j &\geq -X_n^{(j)} + B_j\lambda_{j,i} \\
w_j &\geq 0,\ \sum_{i=-s}^{s}\lambda_{j,i} = 1,\ \lambda_{j,i} > 0, \\
\forall i &= \pm 1, \pm 2, ..., \pm s,\ \forall j = 1, 2, 3, ..., s \\
m|\lambda_{j,j} - \frac{\lambda_{j,-j}}{m-1}| + \sum_{j \neq i} |\lambda_{j,i} - \frac{\lambda_{j,-i}}{m-1}| &\leq \alpha,\ j = 1, 2, 3, ..., s. \\
&\text{for all } B_j
\end{aligned}$$

Problem 5 *We shall solve the following linear programming problem:*

$$min\ w:$$

$$w \geq 0.573 - 0.451\lambda_{1,1} - 0.335\lambda_{1,2} - 0.345\lambda_{1,3}$$

$w \geq -0.573 + 0.451\lambda_{1,1} + 0.335\lambda_{1,2} + 0.345\lambda_{1,3}$

$w \geq 0.105 - 0.479\lambda_{1,1} - 0.335\lambda_{1,2} - 0.349\lambda_{1,3}$

$w \geq -0.105 + 0.479\lambda_{1,1} + 0.335\lambda_{1,2} + 0.349\lambda_{1,3}$

$w \geq 0.323 - 0.420\lambda_{1,1} - 0.337\lambda_{1,2} - 0.342\lambda_{1,3}$

$w \geq -0.323 + 0.420\lambda_{1,1} + 0.337\lambda_{1,2} + 0.342\lambda_{1,3}$

$w \geq 0.214 - 0.274\lambda_{1,-1} - 0.332\lambda_{1,-2} - 0.327\lambda_{1,-3}$

$w \geq -0.214 + 0.274\lambda_{1,-1} + 0.332\lambda_{1,-2} + 0.327\lambda_{1,-3}$

$w \geq 0.448 - 0.261\lambda_{1,-1} - 0.331\lambda_{1,-2} - 0.325\lambda_{1,-3}$

$w \geq -0.448 + 0.261\lambda_{1,-1} + 0.331\lambda_{1,-2} + 0.325\lambda_{1,-3}$

$w \geq 0.338 - 0.290\lambda_{1,-1} - 0.331\lambda_{1,-2} - 0.329\lambda_{1,-3}$

$w \geq -0.338 + 0.290\lambda_{1,-1} + 0.331\lambda_{1,-2} + 0.329\lambda_{1,-3}$

subject to:

$w \geq 0$

$\lambda_{1,1} + \lambda_{1,2} + \lambda_{1,3} + \lambda_{1,-1} + \lambda_{1,-2} + \lambda_{1,-3} = 1$

$\lambda_{1,i} > 0, \forall i = \pm 1, \pm 2, \pm 3$

Solution 6 *Global optimal solution found.*

Objective value: 0.3198829

Objective bound: 0.3198829

Infeasibilities: 0.000000

Extended solver steps: 0

Total solver iterations: 14

Elapsed runtime seconds: 0.12

$Variable$	$Value$
w	0.3198829
$\lambda_{1,1}$	0.2842835
$\lambda_{1,2}$	0.000000
$\lambda_{1,3}$	0.3366718
$\lambda_{1,-1}$	0.000000
$\lambda_{1,-2}$	0.000000
$\lambda_{1,-3}$	0.3790447

Problem 7 *We need to solve the following linear programming problem:*

$min\ w1:$

$w1 \geq 0.347 - 0.348\lambda_{2,1} - 0.335\lambda_{2,2} - 0.335\lambda_{2,3}$

$w1 \geq -0.347 + 0.348\lambda_{2,1} + 0.335\lambda_{2,2} + 0.335\lambda_{2,3}$

$w1 \geq= 0.339 - 0.336\lambda_{2,1} - 0.337\lambda_{2,2} - 0.334\lambda_{2,3}$

$w1 \geq -0.339 + 0.336\lambda_{2,1} + 0.337\lambda_{2,2} + 0.334\lambda_{2,3}$

$w1 \geq 0.315 - 0.3151\lambda_{2,1} - 0.331\lambda_{2,2} - 0.276\lambda_{2,3}$

$w1 \geq -0.315 + 0.3151\lambda_{2,1} + 0.331\lambda_{2,2} + 0.276\lambda_{2,3}$

$w1 \geq 0.326 - 0.327\lambda_{2,-1} - 0.332\lambda_{2,-2} - 0.332\lambda_{2,-3}$

$w1 \geq -0.326 + 0.327\lambda_{2,-1} + 0.332\lambda_{2,-2} + 0.332\lambda_{2,-3}$

$w1 \geq 0.330 - 0.333\lambda_{2,-1} - 0.332\lambda_{2,-2} - 0.333\lambda_{2,-3}$

$w1 \geq -0.330 + 0.333\lambda_{2,-1} + 0.332\lambda_{2,-2} + 0.333\lambda_{2,-3}$

$w1 \geq 0.342 - 0.342\lambda_{2,-1} - 0.334\lambda_{2,-2} - 0.275\lambda_{2,-3}$

$w1 \geq -0.342 + 0.342\lambda_{2,-1} + 0.334\lambda_{2,-2} + 0.275\lambda_{2,-3}$

subject to

$w1 \geq 0$

$\lambda_{2,1} + \lambda_{2,2} + \lambda_{2,3} + \lambda_{2,-1} + \lambda_{2,-2} + \lambda_{2,-3} = 1$

$\lambda_{2,i} > 0, \forall i = \pm 1, \pm 2, \pm 3$

Solution 8 *Global optimal solution found.*

Objective value: 0.1708543

Objective bound: 0.1708543

Infeasibilities: 0.000000

Extended solver steps: 4

Total solver iterations: 36

Elapsed runtime seconds: 0.10

$Variable$	$Value$
$w1$	0.1708543
$\lambda_{2,1}$	0.000000
$\lambda_{2,2}$	0.3261797
$\lambda_{2,3}$	0.1802575
$\lambda_{2,-1}$	0.000000
$\lambda_{2,-2}$	0.00927658
$\lambda_{2,-3}$	0.4842863

Problem 9 *We need to solve the following linear programming problems*

$min\ w2:$

$w2 \geq 0.371 - 0.329\lambda_{3,1} - 0.336\lambda_{3,2} - 0.337\lambda_{3,3}$

$w2 \geq -0.371 + 0.329\lambda_{3,1} + 0.336\lambda_{3,2} + 0.337\lambda_{3,3}$

$w2 \geq 0.323 - 0.388\lambda_{3,1} - 0.335\lambda_{3,2} - 0.334\lambda_{3,3}$

$w2 \geq -0.323 + 0.388\lambda_{3,1} + 0.335\lambda_{3,2} + 0.334\lambda_{3,3}$

$w2 \geq 0.306 - 0.367\lambda_{3,1} - 0.332\lambda_{3,2} - 0.335\lambda_{3,3}$

$w2 \geq -0.306 + 0.367\lambda_{3,1} + 0.332\lambda_{3,2} + 0.335\lambda_{3,3}$

$w2 \geq 0.314 - 0.336\lambda_{3,-1} - 0.331\lambda_{3,-2} - 0.331\lambda_{3,-3}$

$w2 \geq -0.314 + 0.336\lambda_{3,-1} + 0.331\lambda_{3,-2} + 0.331\lambda_{3,-3}$

$$w2 \geq 0.338 - 0.306\lambda_{3,-1} - 0.332\lambda_{3,-2} - 0.333\lambda_{3,-3}$$

$$w2 \geq -0.338 + 0.306\lambda_{3,-1} + 0.332\lambda_{3,-2} + 0.333\lambda_{3,-3}$$

$$w2 \geq 0.347 - 0.317\lambda_{3,-1} - 0.333\lambda_{3,-2} - 0.331\lambda_{3,-3}$$

$$w2 \geq -0.347 + 0.317\lambda_{3,-1} + 0.333\lambda_{3,-2} + 0.331\lambda_{3,-3}$$

$$w2 \geq 0;$$

Solution 10 *Global optimal solution found.*

Objective value: 0.1914343

Objective bound: 0.1914343

Infeasibilities: 0.000000

Extended solver steps: 5

Total solver iterations: 31

Elapsed runtime seconds: 0.08

$Variable$	$Value$
$w2$	0.1914343
$\lambda_{3,1}$	0.000000
$\lambda_{3,1}$	0.000000
$\lambda_{3,3}$	0.5328358
$\lambda_{3,-1}$	0.000000
$\lambda_{3,-2}$	0.4671642
$\lambda_{3,-3}$	0.000000

Thus we have

$$\lambda = \left\{ \begin{array}{cccccc} \lambda_{1,1} & \lambda_{1,2} & \lambda_{1,3} & \lambda_{1,-1} & \lambda_{1,-2} & \lambda_{1,-3} \\ \lambda_{2,1} & \lambda_{2,2} & \lambda_{2,3} & \lambda_{2,-1} & \lambda_{2,-2} & \lambda_{2,-3} \\ \lambda_{3,1} & \lambda_{3,2} & \lambda_{3,3} & \lambda_{3,-1} & \lambda_{3,-2} & \lambda_{3,-3} \end{array} \right\}$$

$$= \left\{ \begin{array}{cccccc} 0.284 & 0.000 & 0.337 & 0.000 & 0.000 & 0.379 \\ 0.000 & 0.326 & 0.180 & 0.0000 & 0.009 & 0.484 \\ 0.000 & 0.000 & 0.533 & 0.000 & 0.467 & 0.000 \end{array} \right\}$$

The parsimonious Multivariate markov chain is of the form

$$X_{n+1}^{(j)} = \Lambda^{+} X_{n}^{(i)} + (m-1)^{-1} \Lambda^{-} (\mathbf{v} - X_{n}^{(i)})$$

hence we have

$$\begin{bmatrix} X_{n+1}^{(1)} \\ X_{n+1}^{(2)} \\ X_{n+1}^{(3)} \end{bmatrix} = \begin{bmatrix} 0.284P^{(1,1)} & 0.000\mathbf{I} & 0.337\mathbf{I} \\ 0.000\mathbf{I} & 0.326P^{(2,2)} & 0.180\mathbf{I} \\ 0.000\mathbf{I} & 0.000\mathbf{I} & 0.533P^{(3,3)} \end{bmatrix} \times$$

$$\begin{bmatrix} X_n^{(1)} \\ X_n^{(2)} \\ X_n^{(3)} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0.000P^{(1,1)} & 0.00\mathbf{I} & 0.379\mathbf{I} \\ 0.000\mathbf{I} & 0.009P^{(2,2)} & 0.484\mathbf{I} \\ 0.000\mathbf{I} & 0.0.467\mathbf{I} & 0.000P^{(3,3)} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{1} - X_n^{(1)} \\ \mathbf{1} - X_n^{(2)} \\ \mathbf{1} - X_n^{(3)} \end{bmatrix}$$

$X_{n+1}^{(1)} = 0.284P^{(1,1)}X_n^{(1)} + 0.337IX_n^{(2)} + \frac{1}{2} \times 0.379(1 - X_n^{(3)})$

$X_{n+1}^{(2)} = 0.326P^{(2,2)}X_n^{(2)} + 0.180IX_n^{(3)} + \frac{1}{2}0.009P^{(2,2)}(1 - X_n^{(2)}) + \frac{1}{2}0.484I(1 - X_n^{(3)})$

$X_{n+1}^{(3)} = 0.533P^{(3,3)}X_n^{(3)} + \frac{1}{2}0.467I(1 - X_n^{(2)})$

Chapter 5

Conclusions and Recommendations

5.1 Discussion of Results

The preliminary tests show that the sequences are all positively related in the six conditions stated above. Thus all our assumptions for the PMMC model have been ascertained.

The Durbin-watson test statistic is less than the "acceptable" level (less than 1.5) which implies that the ordinary least squares method is not the best method to test for level of affection of the three sequences. This qualifies our approach to use a different method to test for level of affection across sequences using the multivariate markov chain.

Form the stationary probability vector vector X_* of X_n we observe that shares prices for the three stocks are eventually up about the probability of 31.7%, non-changing about 22.3% and decreasing in about 46%.Therefore, investors should be less optimistic for the near future because they can only avoid loss with a total probability of about 54.5%.

In the Parsimonious Multivariate Markov Chain model we considered both positively and negatively correlated data sequences.Thus, different perspectives may be formed and different comments can be made compared to the previous Multivariate Markov Chain model. In this study, daily VWAP changes in three stocks in the NSE are taken as the categorical data sequences. The results of the extend to which the three sequences affect each other is discussed below:

$X_{n+1}^{(1)}$ variable in the daily changes of the VWAP of S_1, at time $n+1$ is bound to itself $(X_n^{(1)})$ by 28.4% and is bound to the variable $X_{n+1}^{(2)}$ in the daily changes of the VWAP of S_2 by 33.7% at time n.We also notice that the variable depends on other variables which are negatively correlated to S_3 by

37.9% at time n.

$X_{n+1}^{(2)}$ variable in the daily changes of the VWAP of S_2, at time $n+1$ is bound to itself by 32.6% and is bound to $X_n^{(3)}$ by 18.0% at time n .The varaible is also bound to a varaible negatively associated with the growth of VWAP of S_3at time n by 48.4%.

$X_{n+1}^{(3)}$ variable in the daily changes of the VWAP of S_3, at time $n+1$ is bound to itself ($X_n^{(3)}$) at time n by 53.3% and is bound to a sequence(s) negatively correlated to $X_n^{(2)}$ by 41.7% at time n.

We observe that the prices of S_3 are highly influenced by what happens within the sequence by more than 50% compared to influence from other sequences.We also observe that S_3 is either directly or indirectly affecting the growth prospects of the other two sequences either directly or indirectly.

5.2 Conclusion

We have seen that the parsimonious multivariate markov chain model can be used to predict future prospects of stock prices and how the level of "effection" among various stock. However, the prediction method used here is only a probability forecasting method, the predicted results is expressed as probability of a certain state of stock prices in the future, rather than be in an absolute state. Therefore, it is recommended that a combination methods be used in addition to the prediction using the parsimonious multivariate markov chain model so as to measure the effect of the other factors (forces) that drive stock prices in the NSE.

5.3 Recommendation for Further Studies

The use of Higher-order Parsimonious Multivariate Markov Chain model in stock markets

REFERENCES

[1] Agwuegbo S., Adewole A. and Maduegbuna A. (2010); A Random Walk Model for Stock Market Prices (2010); *Journal of Mathematics and Statistics*, pg 342-346.

[2] Ching W., Fung E., Ng M. (2002): A multivariate Markov chain model for Categorical Data Sequences and its applications in demand predictions; *IMA J Management Maths*. 13 (2002) pg 87-199.

[3] Ching W., Ng m., Fung E. (2007) Higher -order multivariate markov chains and their Application; *Elsevier Journal* (2008)pg 492-507.

[4] Ching, W., Siu, T., Li, L., Li, H. J. T., & Li, W. (2009). An improved parsimonious multivariate Markov chain model for credit risk. *Journal of Credit Risk*, 5, 1-25.

[5] Chao Wang and Ting-Zhu Huang, A New Improved Parsimonious Multivariate Markov Chain Model, *Journal of Applied Mathematics*, vol. 2013

[6] Chva′tal V., (1983) ,Linear Programming;Freeman, New York.

[7] Doubleday K., and Esunge J. (2011) ; Application of Markov Chains to Stock Trends; *Journal of Mathematics and Statistics*, pg 103-106.

[8] Ersoy Ö. and Semra E. (2011); An Application of Multivariate Markov Chain Model on the Changes in Exchange Rates: Turkey Case ;*European Journal of Social Sciences* – Volume 18, pg 542-552.

[9] Siu T, Ching W, Fung E and Ng M (2005); On a multivariate Markov chain model for credit risk measurement, *Quantitative Finance*, vol. 5, no. 6, pp. 543–556

[10] Svoboda and Lukáš ;*30th Proceedings of 30th International Conference Mathematical Methods in Economics*: Application of Markov chain analysis to trend prediction of stock indices

[11] Zhang D., Zhang X., (2009); Study on Forecasting the Stock Market Trend Based on Stochastic Analysis Method;*International Journal of Business Management*

[12] https://www.nse.co.ke/

Appendix

1. \QTR{bf}{R codes for estimating Markov chain model}

```
b=read.csv("NSE_data_states.csv")

attach(b)

c=subset(b,x1==1)

d=subset(c,x4==1)

e=subset(c,x4==2)

f=subset(c,x4==3)

length(d$x1)###entry of the frequencey matrix###

length(e$x1)###entry of the frequencey matrix###

length(f$x1)###entry of the frequencey matrix###

s1=sum(length(d$x1)+length(e$x1)+length(f$x1))

#######

c1=subset(b,x1==2)

d1=subset(c1,x4==1)

e1=subset(c1,x4==2)

f1=subset(c1,x4==3)

length(d1$x1)###entry of the frequencey matrix###

length(e1$x1)###entry of the frequencey matrix###
```

```
length(f1$x1)###entry of the frequencey matrix###

s2=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

######

c2=subset(b,x1==3)

d2=subset(c2,x4==1)

e2=subset(c2,x4==2)

f2=subset(c2,x4==3)

length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s3=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))

F11=matrix(c(length(d$x1),length(e$x1),length(f$x1),

length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),

length(e2$x1),length(f2$x1)),ncol=3,byrow=T)

P11=round(matrix(c(length(d$x1)/s1,length(e$x1)/s1,

length(f$x1)/s1,

length(d1$x1)/s2,length(e1$x1)/s2,length(f1$x1)/s2,

length(d2$x1)/s3,length(e2$x1)/s3,length(f2$x1)/s3),

ncol=3,byrow=T),3)
```

```
##############################P21############
c=subset(b,x1==1)
d=subset(c,x5==1)
e=subset(c,x5==2)
f=subset(c,x5==3)
length(d$x1)###entry of the frequencey matrix###
length(e$x1)###entry of the frequencey matrix###
length(f$x1)###entry of the frequencey matrix###
s11=sum(length(d$x1)+length(e$x1)+length(f$x1))
c1=subset(b,x1==2)
d1=subset(c1,x5==1)
e1=subset(c1,x5==2)
f1=subset(c1,x5==3)
length(d1$x1)###entry of the frequencey matrix###
length(e1$x1)###entry of the frequencey matrix###
length(f1$x1)###entry of the frequencey matrix###
s12=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))
###########
c2=subset(b,x1==3)
```

```
d2=subset(c2,x5==1)

e2=subset(c2,x5==2)

f2=subset(c2,x5==3)

length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s13=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))

F21=matrix(c(length(d$x1),length(e$x1),length(f$x1),

length(d1$x1),length(e1$x1),length(f1$x1),

length(d2$x1),length(e2$x1),length(f2$x1)),ncol=3,byrow=T)

P21=round(matrix(c(length(d$x1)/s11,length(e$x1)/s11,

length(f$x1)/s11,

length(d1$x1)/s12,length(e1$x1)/s12,length(f1$x1)/s12,

length(d2$x1)/s13,length(e2$x1)/s13,length(f2$x1)/s13),

ncol=3,byrow=T),3)

###########################P31###########

c=subset(b,x1==1)

d=subset(c,x6==1)

e=subset(c,x6==2)
```

```
f=subset(c,x6==3)

length(d$x1)###entry of the frequencey matrix###

length(e$x1)###entry of the frequencey matrix###

length(f$x1)###entry of the frequencey matrix###

s1_1=sum(length(d$x1)+length(e$x1)+length(f$x1))

############

c1=subset(b,x1==2)

d1=subset(c1,x6==1)

e1=subset(c1,x6==2)

f1=subset(c1,x6==3)

length(d1$x1)###entry of the frequencey matrix###

length(e1$x1)###entry of the frequencey matrix###

length(f1$x1)###entry of the frequencey matrix###

s1_2=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

###########

c2=subset(b,x1==3)

d2=subset(c2,x6==1)

e2=subset(c2,x6==2)

f2=subset(c2,x6==3)
```

```
length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s1_3=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))

######

F31=matrix(c(length(d$x1),length(e$x1),length(f$x1),

length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),

length(e2$x1),length(f2$x1)),ncol=3,byrow=T)

P31=round(matrix(c(length(d$x1)/s1_1,length(e$x1)/s1_1,

length(f$x1)/s1_1,

length(d1$x1)/s1_2,length(e1$x1)/s1_2,length(f1$x1)/s1_2,

length(d2$x1)/s1_3,length(e2$x1)/s1_3,length(f2$x1)/s1_3),

ncol=3,byrow=T),3)

#####################P22##########

c=subset(b,x2==1)

d=subset(c,x5==1)

e=subset(c,x5==2)

f=subset(c,x5==3)

length(d$x5)###entry of the frequencey matrix###
```

```
length(e$x5)###entry of the frequencey matrix###

length(f$x5)###entry of the frequencey matrix###

s_1=sum(length(d$x1)+length(e$x1)+length(f$x1))

#######

c1=subset(b,x2==2)

d1=subset(c1,x5==1)

e1=subset(c1,x5==2)

f1=subset(c1,x5==3)

length(d1$x5)###entry of the frequencey matrix###

length(e1$x5)###entry of the frequencey matrix###

length(f1$x5)###entry of the frequencey matrix###

s_2=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

###############

c2=subset(b,x2==3)

d2=subset(c2,x5==1)

e2=subset(c2,x5==2)

f2=subset(c2,x5==3)

length(d2$x5)###entry of the frequencey matrix###

length(e2$x5)###entry of the frequencey matrix###
```

```
length(f2$x5)###entry of the frequencey matrix###

s_3=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))

######

F22=matrix(c(length(d$x1),length(e$x1),length(f$x1),

length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),

length(e2$x1),length(f2$x1)),ncol=3,byrow=T)

P22=round(matrix(c(length(d$x1)/s_1,length(e$x1)/s_1,

length(f$x1)/s_1,

length(d1$x1)/s_2,length(e1$x1)/s_2,length(f1$x1)/s_2,

length(d2$x1)/s_3,length(e2$x1)/s_3,length(f2$x1)/s_3),

ncol=3,byrow=T),3)

###########################P12################

c=subset(b,x2==1)

d=subset(c,x4==1)

e=subset(c,x4==2)

f=subset(c,x4==3)

length(d$x1)###entry of the frequencey matrix###

length(e$x1)###entry of the frequencey matrix###

length(f$x1)###entry of the frequencey matrix###
```

```
s1_=sum(length(d$x1)+length(e$x1)+length(f$x1))

###

c1=subset(b,x2==2)

d1=subset(c1,x4==1)

e1=subset(c1,x4==2)

f1=subset(c1,x4==3)

length(d1$x1)###entry of the frequencey matrix###

length(e1$x1)###entry of the frequencey matrix###

length(f1$x1)###entry of the frequencey matrix###

s2_=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

##############

c2=subset(b,x2==3)

d2=subset(c2,x4==1)

e2=subset(c2,x4==2)

f2=subset(c2,x4==3)

length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s3_=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))
```

```
#####
F12=matrix(c(length(d$x1),length(e$x1),length(f$x1),
length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),
length(e2$x1),length(f2$x1)),ncol=3,byrow=T)
P12=round(matrix(c(length(d$x1)/s1_,length(e$x1)/s1_,
length(f$x1)/s1_,
length(d1$x1)/s2_,length(e1$x1)/s2_,length(f1$x1)/s2_,
length(d2$x1)/s3_,length(e2$x1)/s3_,length(f2$x1)/s3_),
ncol=3,byrow=T),3)
#####################################P32##########
c=subset(b,x2==1)
d=subset(c,x6==1)
e=subset(c,x6==2)
f=subset(c,x6==3)
length(d$x1)###entry of the frequencey matrix###
length(e$x1)###entry of the frequencey matrix###
length(f$x1)###entry of the frequencey matrix###
s_11=sum(length(d$x1)+length(e$x1)+length(f$x1))
##########
```

```
c1=subset(b,x2==2)

d1=subset(c1,x6==1)

e1=subset(c1,x6==2)

f1=subset(c1,x6==3)

length(d1$x1)###entry of the frequencey matrix###

length(e1$x1)###entry of the frequencey matrix###

length(f1$x1)###entry of the frequencey matrix###

s_22=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

####

c2=subset(b,x2==3)

d2=subset(c2,x6==1)

e2=subset(c2,x6==2)

f2=subset(c2,x6==3)

length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s_33=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))

#########

F32=matrix(c(length(d$x1),length(e$x1),length(f$x1),
```

```
length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),
length(e2$x1),length(f2$x1)),ncol=3,byrow=T)
P32=round(matrix(c(length(d$x1)/s_11,length(e$x1)/s_11,
length(f$x1)/s_11,
length(d1$x1)/s_22,length(e1$x1)/s_22,length(f1$x1)/s_22,
length(d2$x1)/s_33,length(e2$x1)/s_33,length(f2$x1)/s_33),
ncol=3,byrow=T),3)
#############P33############
c=subset(b,x3==1)
d=subset(c,x6==1)
e=subset(c,x6==2)
f=subset(c,x6==3)
length(d$x1)###entry of the frequencey matrix###
length(e$x1)###entry of the frequencey matrix###
length(f$x1)###entry of the frequencey matrix###
s__1=sum(length(d$x1)+length(e$x1)+length(f$x1))
######
c1=subset(b,x3==2)
d1=subset(c1,x6==1)
```

```
e1=subset(c1,x6==2)

f1=subset(c1,x6==3)

length(d1$x1)###entry of the frequencey matrix###

length(e1$x1)###entry of the frequencey matrix###

length(f1$x1)###entry of the frequencey matrix###

s__2=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

#####

c2=subset(b,x3==3)

d2=subset(c2,x6==1)

e2=subset(c2,x6==2)

f2=subset(c2,x6==3)

length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s__3=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))

####

F33=matrix(c(length(d$x1),length(e$x1),length(f$x1),

length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),

length(e2$x1),length(f2$x1)),ncol=3,byrow=T)
```

```
P33=round(matrix(c(length(d$x1)/s__1,length(e$x1)/s__1,
length(f$x1)/s__1,
length(d1$x1)/s__2,length(e1$x1)/s__2,length(f1$x1)/s__2,
length(d2$x1)/s__3,length(e2$x1)/s__3,length(f2$x1)/s__3),
ncol=3,byrow=T),3)
##########################P13##########
c=subset(b,x3==1)
d=subset(c,x4==1)
e=subset(c,x4==2)
f=subset(c,x4==3)
length(d$x1)###entry of the frequencey matrix###
length(e$x1)###entry of the frequencey matrix###
length(f$x1)###entry of the frequencey matrix###
s1__=sum(length(d$x1)+length(e$x1)+length(f$x1))
#####
c1=subset(b,x3==2)
d1=subset(c1,x4==1)
e1=subset(c1,x4==2)
f1=subset(c1,x4==3)
```

```
length(d1$x1)###entry of the frequencey matrix###
length(e1$x1)###entry of the frequencey matrix###
length(f1$x1)###entry of the frequencey matrix###
s2__=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))
###
c2=subset(b,x3==3)
d2=subset(c2,x4==1)
e2=subset(c2,x4==2)
f2=subset(c2,x4==3)
length(d2$x1)###entry of the frequencey matrix###
length(e2$x1)###entry of the frequencey matrix###
length(f2$x1)###entry of the frequencey matrix###
s3__=sum(length(d2$x1)+length(e2$x1)+length(f2$x1))
F13=matrix(c(length(d$x1),length(e$x1),length(f$x1),
length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),
length(e2$x1),length(f2$x1)),ncol=3,byrow=T)
P13=round(matrix(c(length(d$x1)/s1__,length(e$x1)/s1__,
length(f$x1)/s1__,
length(d1$x1)/s2__,length(e1$x1)/s2__,length(f1$x1)/s2__,
```

```
length(d2$x1)/s3__,length(e2$x1)/s3__,length(f2$x1)/s3__),

ncol=3,byrow=T),3)

####################23############

c=subset(b,x3==1)

d=subset(c,x5==1)

e=subset(c,x5==2)

f=subset(c,x5==3)

length(d$x1)###entry of the frequencey matrix###

length(e$x1)###entry of the frequencey matrix###

length(f$x1)###entry of the frequencey matrix###

s_1_=sum(length(d$x1)+length(e$x1)+length(f$x1))

##########

c1=subset(b,x3==2)

d1=subset(c1,x5==1)

e1=subset(c1,x5==2)

f1=subset(c1,x5==3)

length(d1$x1)###entry of the frequencey matrix###

length(e1$x1)###entry of the frequencey matrix###

length(f1$x1)###entry of the frequencey matrix###
```

```
s_2_=sum(length(d1$x1)+length(e1$x1)+length(f1$x1))

#########

c2=subset(b,x3==3)

d2=subset(c2,x5==1)

e2=subset(c2,x5==2)

f2=subset(c2,x5==3)

length(d2$x1)###entry of the frequencey matrix###

length(e2$x1)###entry of the frequencey matrix###

length(f2$x1)###entry of the frequencey matrix###

s_3_=sum(length(d$x1)+length(e$x1)+length(f$x1))

F23=matrix(c(length(d$x1),length(e$x1),length(f$x1),

length(d1$x1),length(e1$x1),length(f1$x1),length(d2$x1),

length(e2$x1),length(f2$x1)),ncol=3,byrow=T)

P23=round(matrix(c(length(d$x1)/s_1_,length(e$x1)/s_1_,

length(f$x1)/s_1_,

length(d1$x1)/s_2_,length(e1$x1)/s_2_,length(f1$x1)/s_2_,

length(d2$x1)/s_3_,length(e2$x1)/s_3_,length(f2$x1)/s_3_),

ncol=3,byrow=T),3)

#########################Frequency Matrixes####################
```

```
##############Xt##########
xt1=subset(b,x1==1)
length(xt1$x1)
xt2=subset(b,x1==2)
length(xt2$x1)
xt3=subset(b,x1==3)
length(xt3$x1)
Xt1=t(cbind(length(xt1$x1)/124,length(xt2$x1)/124,length(xt3$x1)/124))
Xt1=round(Xt1,3)
##
xt1=subset(b,x2==1)
length(xt1$x2)
xt2=subset(b,x2==2)
length(xt2$x2)
xt3=subset(b,x2==3)
length(xt3$x2)
Xt2=t(cbind(length(xt1$x1)/124,length(xt2$x1)/124,length(xt3$x1)/124))
Xt2=round(Xt2,3)
######
```

```
xt1=subset(b,x3==1)

length(xt1$x3)

xt2=subset(b,x3==2)

length(xt2$x3)

xt3=subset(b,x3==3)

length(xt3$x3)

Xt3=t(cbind(length(xt1$x1)/124,length(xt2$x1)/124,length(xt3$x1)/124))

Xt3=round(Xt3,3)

Xt=(matrix(c(Xt1,Xt2,Xt3),ncol=3,byrow=T))

Xt=round(Xt,3)

###################Matrix B#################

c1=P11%*%Xt1

c2=P12%*%Xt2

c3=P13%*%Xt3

B1=round(cbind(c1,c2,c3),3)

c11=P21%*%Xt1

c22=P22%*%Xt2

c33=P23%*%Xt3

B2=round(cbind(c11,c22,c33),3)
```

```
##########

c111=P31%*%Xt1

c222=P32%*%Xt2

c333=P33%*%Xt3

B3=round(cbind(c111,c222,c333),3)

#############Zt+1####the negative correlation bit############

e=matrix(c(1,1,1,1,1,1,1,1,1),ncol=3,byrow=T)

Z=round(1/2*(e-(Xt)),3)

Zt1=round(matrix(c(0.214,0.448,0.338),ncol=1,byrow=F),3)

Zt2=round(matrix(c(0.326,0.330,0.342),ncol=1,byrow=F),3)

Zt3=round(matrix(c(0.314,0.338, 0.347),ncol=1,byrow=F),3)

###########Matrix B for negative part##########

c1=P11%*%Zt1

c2=P12%*%Zt2

c3=P13%*%Zt3

B11=round(cbind(c1,c2,c3),3)

c11=P21%*%Zt1

c22=P22%*%Zt2

c33=P23%*%Zt3
```

```
B22=round(cbind(c11,c22,c33),3)

##########

c111=P31%*%Zt1

c222=P32%*%Zt2

c333=P33%*%Zt3

B33=round(cbind(c111,c222,c333),3)

################end######################

\textbf{2. Lingo Commands for Optimisation

}

 min=w;

w>=0.573-0.451*x-0.335*y-0.345*z;

w>=-0.573+0.451*x+0.335*y+0.345*z;

w>=0.105-0.479*x-0.335*y-0.349*z;

w>=-0.105+0.479*x+0.335*y+0.349*z;

w>=0.323-0.420*x-0.337*y-0.342*z;

w>=-0.323+0.420*x+0.337*y+0.342*z;

w>=0.214-0.274*x1-0.332*y1-0.327*z1;

w>=-0.214+0.274*x1+0.332*y1+0.327*z1;

w>=0.448-0.261*x1-0.331*y1-0.325*z1;
```

```
w>=-0.448+0.261*x1+0.331*y1+0.325*z1;

w>=0.338-0.290*x1-0.331*y1-0.329*z1;

w>=-0.338+0.290*x1+0.331*y1+0.329*z1;

w>=0;

@bnd(0,x,1);

@bnd(0,y,1);

@bnd(0,z,1);

@bnd(0,x1,1);

@bnd(0,y1,1);

@bnd(0,z1,1);

x+y+z+x1+y1+z1=1;

#################

min=w;

w>=0.347-0.348*x-0.335*y-0.335*z;

w>=-0.347+0.348*x+0.335*y+0.335*z;

w>=0.339-0.336*x-0.337*y-0.334*z;

w>=-0.339+0.336*x+0.337*y+0.334*z;

w>=0.315-0.3151*x-0.331*y-0.276*z;

w>=-0.315+0.3151*x+0.331*y+0.276*z;
```

```
w>=0.326-0.327*x1-0.332*y1-0.332*z1;

w>=-0.326+0.327*x1+0.332*y1+0.332*z1;

w>=0.330-0.333*x1-0.332*y1-0.333*z1;

w>=-0.330+0.333*x1+0.332*y1+0.333*z1;

w>=0.342-0.342*x1-0.334*y1-0.275*z1;

w>=-0.342+0.342*x1+0.334*y1+0.275*z1;

w>=0;

@bnd(0,x,1);

@bnd(0,y,1);

@bnd(0,z,1);

@bnd(0,x1,1);

@bnd(0,y1,1);

@bnd(0,z1,1);

x+y+z+x1+y1+z1=1;

#############

min=w;

w>=0.371-0.329*x-0.336*y-0.337*z;

w>=-0.371+0.329*x+0.336*y+0.337*z;

w>=0.323-0.388*x-0.335*y-0.334*z;
```

```
w>=-0.323+0.388*x+0.335*y+0.334*z;

w>=0.306-0.367*x-0.332*y-0.335*z;

w>=-0.306+0.367*x+0.332*y+0.335*z;

w>=0.314-0.336*x1-0.331*y1-0.331*z1;

w>=-0.314+0.336*x1+0.331*y1+0.331*z1;

w>=0.338-0.306*x1-0.332*y1-0.333*z1;

w>=-0.338+0.306*x1+0.332*y1+0.333*z1;

w>=0.347-0.317*x1-0.333*y1-0.331*z1;

w>=-0.347+0.317*x1+0.333*y1+0.331*z1;

w>=0;

@bnd(0,x,1);

@bnd(0,y,1);

@bnd(0,z,1);

@bnd(0,x1,1);

@bnd(0,y1,1);

@bnd(0,z1,1);

x+y+z+x1+y1+z1=1;

###########End###
```