ตัวแบบการถดถอยในตัวผสมสำหรับข้อมูลตลาดหุ้นไทย

นายอภิชา สุทธิชยาพิพัฒน์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์และวิทยาการคณนา
ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2566

MIXTURE AUTOREGRESSIVE MODELS FOR THAI STOCK MARKET DATA

Mr. Apicha Suthichayapipat

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Applied Mathematics and

Computational Science

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn University

Academic Year 2023

Thesis Title	MIXTURE AUTOREGRESSIVE MODELS FOR THAI STOCK MARKET DATA
By	Mr. Apicha Suthichayapipat
Field of Study	Applied Mathematics and Computational Science
Thesis Advisor	Assistant Professor Jiraphan Suntornchost, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

.... Dean of the Faculty of Science
(Professor Pranut Potiyaraj, Ph.D.)

THESIS COMMITTEE

.... Chairman
(Associate Professor Krung Sinapiromsaran, Ph.D.)

.... Thesis Advisor
(Assistant Professor Jiraphan Suntornchost, Ph.D.)

.... Examiner
(Monchai Kooakachai, Ph.D.)

.... External Examiner
(Assistant Professor Thidaporn Supapakorn, Ph.D.)

อภิชา สุทธิชยาพิพัฒน์ : ตัวแบบการถดถอยในตัวผสมสำหรับข้อมูลตลาดหุ้นไทย. (MIXTURE AUTOREGRESSIVE MODELS FOR THAI STOCK MARKET DATA) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : ผศ.ดร. จิราพรรณ สุนทรโชติ, 117 หน้า.

งานวิจัยนี้ศึกษาแบบจำลองการถดถอยในตัว ซึ่งเป็นหนึ่งในแบบจำลองที่ได้รับความนิยมมากที่สุดในการคาดการณ์ข้อมูลในอดีตของอนุกรมเวลา แบบจำลองการถดถอยในตัวอยู่ภายใต้การแจกแจงแบบปกติ ซึ่งแบบจำลองการถดถอยนี้ไม่เหมาะสมกับข้อมูลบางชุดโดยเฉพาะข้อมูลทางการเงิน แนวคิดของการแจกแจงแบบผสมยังนำไปสู่กลุ่มของตัวแบบการถดถอยในตัวผสมที่มีแบบจำลองการถดถอยในแต่ละชุดที่ต่างกัน ซึ่งพิจารณาเป็นตัวแบบการถดถอยในตัวผสม และตัวแบบการถดถอยในตัวผสมแบบเวกเตอร์ภายใต้การแจกแจงแบบปกติ และแบบที ในการศึกษานี้เราสร้างการประมาณค่าพารามิเตอร์ด้วยขั้นตอนวิธีค่าคาดหมายสูงสุด และตรวจสอบประสิทธิภาพในกลุ่มของตัวแบบการถดถอยในตัวผสมโดยใช้การประมาณค่าพารามิเตอร์ที่พัฒนาขึ้นด้วยขั้นตอนวิธีค่าคาดหมายสูงสุดเปรียบเทียบกับการประมาณภาวะน่าจะเป็นสูงสุด ซึ่งพิจารณาหุ้นที่มีความน่าเชื่อถือจาก 2 กลุ่มที่แตกต่างกันในตลาดหุ้นไทยซึ่งหุ้นในกลุ่มของพลังงาน และอิเล็กทรอนิกส์ในแต่ละกลุ่มมี 3 หุ้น เกณฑ์ในการเลือกแบบจำลองนั้นคือการใช้ค่าสถิติ เอไอซี, บีไอซี, เอชคิวไอซี และ ค่าเฉลี่ยความผิดพลาดกำลังสอง จากผลลัพธ์แสดงให้เห็นว่าตัวแบบจำลองที่เหมาะสมสำหรับข้อมูลตลาดหุ้นไทยมาจากกลุ่มของตัวแบบถดถอยในตัวผสม และการประมาณค่าที่แม่นยำสำหรับข้อมูลตลาดหุ้นไทยมาจากการประมาณค่าพารามิเตอร์ด้วยขั้นตอนวิธีค่าคาดหมายสูงสุด

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์และวิทยาการคณนา

ลายมือชื่อนิสิต อภิชา สุทธิชยาพิพัฒน์

ลายมือชื่อ อ.ที่ปรึกษาหลัก จิราพรรณ

ปีการศึกษา 2566

6370097423 : MAJOR APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE

KEYWORDS : AUTOREGRESSIVE MODEL, MULTIMODAL, MIXTURE MODEL, THAI STOCK MARKET

APICHA SUTHICHAYAPIPAT : MIXTURE AUTOREGRESSIVE MODELS FOR THAI STOCK MARKET DATA. ADVISOR : ASSISTANT PROFESSOR JIRAPHAN SUNTORNCHOST, PH.D., 117 pp.

The autoregressive (AR) model is one of the most widely used time series forecasting models. The standard AR model was established using the normal distribution, which is violated in some datasets, notably financial data. Therefore, alternative distributions are proposed in the literature, such as the concept of mixture distributions. This concept is also applied to time series modeling in the family of mixture autoregressive models that combine different autoregressive components. Specifically, we consider both the univariate mixture autoregressive model and the multivariate mixture autoregressive model based on the normal and t distributions. In this study, we construct the EM algorithm to estimate parameters and investigate the performance of this method compared with the MLE. The analysis focuses on top stocks from two different sectors in the market, namely energy and utility and electronic components, with each sector comprising three stocks. The fitted models are compared with the family of mixture autoregressive models by using AIC, HQIC, BIC, and MSE of predictions. The results indicate that the EM algorithm is preferred for Thai stock market data.

Department	: Mathematics and	Student's Signature	อภิชา สุทธิชยพิพัฒน์
	Computer Science	Advisor's Signature	จิราพรรณ
	Computational Science		
Academic Year :	2023		

ACKNOWLEDGEMENTS

This thesis has become a reality through the generous support and assistance of numerous individuals, to whom I am deeply grateful. My sincere thanks go out to each and every one of them.

Foremost, I would like to express my gratitude to my advisor, Assistant Professor Jiraphan Suntornchost, for guiding me in advancing my research project and providing continuous encouragement throughout my study. When I faced with challenges in understanding the theory or developing research ideas, she was always there to provide motivation and cheer me up. This thesis would not have been completed without her support. Thanks to the committee for giving suggestions to improve the project.

I would also like to express my gratitude to my father and mother, who encouraged me to pursue higher education and always listened to all of my problems.

CONTENTS

Page

CHAPTER

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A time series is an ordered sequence taken at successive equally spaced time intervals and has been widely used in various applications, particularly in finance, actuarial science, economics, methodology, and medicine. One of the most popular time series models is the family of autoregressive (AR) models, which have the unimodal Gaussian distribution as the underlying distribution, defined as follows.

Let y_t represent the value of the series at time t. We designate the process y_t as an autoregressive model with, p, order of autoregressive, denoted as $AR(p)$, if it can be expressed as a weighted linear combination of order p most recent past values of itself. Specifically,

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t, \tag{1.1}$$

where white noise, ϵ_t, is assumed to be normally distributed.

Even though the univariate $AR(p)$ models have been widely applied in many different displines, the unimodity assumption of the autoregressive models might not be applicable in some situations. For example, consider the Canadian lynx dataset and the time series of common stock closing prices for International Business Machines (IBM). For more details, refer to Wong and Li (2000). Therefore, alternative distributions for multimodal data have been investigated in literature. One such concept for multimodal data is to apply a mixture of different distributions. The finite mixture distribution is a class of probability distributions particularly useful for modelling data that contains relatively distinct subgroups or clusters of observations. In particular, the distribution of Φ_Y is considered as a finite mixture of Φ_i for $i = 1, 2, 3, \ldots, K$ if it can be written as

$$\Phi_Y = \alpha_1\Phi_1 + \alpha_2\Phi_2 + \cdots + \alpha_K\Phi_K,$$

where $\sum_{i=1}^{K} \alpha_i = 1$. Consequently, the density function ϕ_Y corresponding to Φ_Y can be written as

$$\phi_Y = \alpha_1\phi_1 + \alpha_2\phi_2 + \cdots + \alpha_K\phi_K,$$

where ϕ_i represents the probability density function corresponding to the distribution Φ_i for $i = 1, 2, \ldots, K$.

The concept of the mixture distribution was initially introduced into the time series context by Le et al. in 1996 [1], when the authors created the class of Gaussian Mixture Transition Distribution (GMTD) time series models.

In 2000, Wong and Li [2] generalized the concepts of the model to introduce a new class of mixture model for time series data, known as the Mixture Autoregressive (MAR) model. Moreover, they demonstrated that the MAR model outperforms other existing time series models, as evidenced by its application to datasets such as the International Business Machines stock prices and the Canadian lynx data.

Since then, the concepts of mixture time series models have attracted the attention of researchers. The models were then extended to more general models. For example, Fong et al. (2007) [3] extended the univariate MAR model to multivariate time series data by introducing the mixture vector autoregressive model. Subsequently, the model has been extended in various aspects. For instance, Lanne and Saikkonen [4] expanded the model by incorporating GARCH errors and applied it to the U.S. short-term interest rate. Meitz, Virolainen, and Savi [5] extended the model to a mixture of autoregressive and a model based on the Student's t-distribution. They developed the "uGMAR" R-package, which provides tools for estimating and analysing the Gaussian mixture autoregressive model. In 2022, Virolainen and Savi [6] developed the "gmvarkit" R-package, which offers tools for estimating and analyzing the Gaussian mixture vector autoregressive model.

In this study, we will construct appropriate mixture autoregressive models for Thai stock data and multivariate mixture vector autoregressive models to studies the correlation between different stock markets by using the maximum log-likelihood method and EM algorithm to estimate parameters in Chapter 4 and choose the best model by AIC, BIC, and HQIC. The thesis is organized as follows. In Chapter 2, we introduce the distributions that we consider in this study, discuss some of the parameter estimation, model selection, model diagnostic. At the end of the chapter, we introduce some basic ideas in time series analysis such as mean, covariance, correlation functions, and the concept of stationarity. Also, we take an overview of the basic linear time series models, which consists of the stationary time series model, which are AR, MA, ARMA, and the VAR models, and non-stationary time series models which are ARIMA models. In Chapter 3, we introduce the family of univariate mixture autoregressive models, which have univariate mixture autoregressive (MAR), t mixture autoregressive (TMAR) and describe the properties of each model. We construct the EM algorithm for estimating parameters in the univariate mixture autoregressive and the multivariate mixture vector autoregressive model. We investigate the performance of our method by comparing it with the maximum log-likelihood estimator and conducting simulation studies to test the accuracy of the estimation. After that, we apply the EM algorithm to analyze Thai stock market data. In Chapter 4, we introduce the multivariate mixture autoregressive (MVAR) models and the multivariate t mixture autoregressive (TMVAR) models and describe their properties. We investigate the performance of the parameter estimation of each method to the MVAR and TMVAR model. We fitted the models to our data set, obtained the best model for each dataset by using AIC, BIC, HQIC and mean square error (MSE), and the corresponding regression coefficients. We summarized the main results and conclusion in Chapter 5.

CHAPTER II

PRELIMINARY

In this chapter, we introduce the basic knowledge used in this thesis. The distributions considered in this study are discussed in Section 2.1. In Section 2.2, we discuss parameter estimation, the maximum likelihood estimator and the EM algorithm for efficient model parameter estimation. Section 2.3 covers the selection of the best candidate model for each model using selection criteria. In Section 2.4, we evaluate how well the models fit the data through model diagnostics. Fundamental concepts time series and time series model are discussed in Sections 2.5.

2.1 Distribution

In this study, we examine both the normal distribution and the Student's t distribution in the context of univariate and multivariate time series models. Additionally, we explore the finite mixed distribution, which is applied to time series models, resulting in a mixture model.

2.1.1 Normal distribution

Definition 1. Let X be a continuous random variable. We say that X follow a normal distribution with mean μ, variance σ^2 if the probability density function follows

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{2.1}$$

2.1.2 t distribution

Definition 2. Let X be a continuous random variable. We say that X follow a t distribution with v degree of freedom if the probability density function follows

$$f_v(x) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\Gamma(\frac{v}{2})}\Big(1 + \frac{x^2}{v}\Big)^{-\frac{v+1}{2}}, \tag{2.2}$$

where $\Gamma(\alpha)$ is the gamma function,

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx. \tag{2.3}$$

2.1.3 Finite mixture distribution

A mixture distribution is the probability distribution of a random variable derived from a collection of other random variables, which arises in various contexts in the literature and naturally occurs when a statistical population contains two or more subpopulations. The finite mixture distribution is a class of probability distributions particularly useful for modeling data that contains relatively distinct subgroups or clusters of observations.

Definition 3. The distribution Φ_Y is considered as a finite mixture of Φ_i for $i = 1, 2, \dots, K$ component if it can be written as

$$\Phi_Y = \alpha_1\Phi_1 + \alpha_2\Phi_2 + \cdots + \alpha_K\Phi_K, \tag{2.4}$$

where $\alpha_i > 0$ and $\sum_{i=1}^{K} \alpha_i = 1$. Consequently, the corresponding density function ϕ_Y to Φ_Y can be written as

$$\phi_Y = \alpha_1\phi_1 + \alpha_2\phi_2 + \cdots + \alpha_K\phi_K, \tag{2.5}$$

where ϕ_i is the corresponding probability density function to the distribution Φ_i for $i = 1, 2, \dots, K$.

2.2 Parameter estimation

Parameter estimation is the process of computing a model's parameter values from measured data. In this section, we introduce the basic concepts of the maximum likelihood estimator and the Expectation-Maximization algorithm that we investigated in our study.

2.2.1 The maximum likelihood estimator

Definition 4. Let $f(y|\theta)$ denote the joint probability distribution function of the sample $Y = (y_1, y_2, \ldots, y_n)$. Then, given that $Y = y$ is observed, the function of θ defined by

$$L(\theta|y) = f(y|\theta) \tag{2.6}$$

is called the likelihood function.

Definition 5. For each sample point y, let $\hat{\theta}(y)$ be a parameter value at which $L(\theta|y)$ attains its maximum as a function of θ, with y held fixed. The maximum likelihood estimator (MLE) of the parameter θ based on a sample Y is $\hat{\theta}(y)$.

Now that we have a likelihood function in the definition above, we set the partial derivative with respect to the parameter to zero, and then we will obtain the parameter estimate. The mixture components may not be solved in general [7] because, given the log likelihood, it is hard to derive the parameter estimation in closed form. Subsequently, the estimation of the parameters needed to be performed using a numerical method [5], and the expectation-maximization algorithm is an approach for performing maximum likelihood estimation in the presence of latent variables, addressing situations where the complete data likelihood is challenging to maximize directly due to the unobservable nature of certain variables.

2.2.2 The Expectation-Maximization Algorithm

The Expectation-Maximization algorithm, or EM algorithm, is an iterative method commonly used to obtain the maximum likelihood estimate when direct calculation is not applicable. In particular, in mixed effect models and mixture models that involve latent variables, say $Z = (Z_1, \ldots, Z_n)$, in addition of unknown parameters. To perform the EM algorithm, each iteration consists of two steps: the expectation step (E-step) and the maximization step (M-step). The parameter of the mixture model are $\boldsymbol{\theta} = \{\phi_1, \ldots, \phi_K, \sigma_1, \ldots, \sigma_K, \alpha_1, \ldots, \alpha_K\}$ and the observations $Y = (y_1, \ldots, y_n)$ are considered as incomplete data which have unobserved random variables $Z = (Z_1, \ldots, Z_K)$

and α_k, $k = 1, 2, \ldots, K$ as mixture proportions. The probability distribution of y_t is defined as

$$P(y_t|\boldsymbol{\theta}) = \sum_{k=1}^{K} \alpha_k P(y_t|\boldsymbol{\theta}, Z_t = k), \tag{2.7}$$

the conditional log likelihood of mixture distribution is given by

$$l(\boldsymbol{\theta}) = \sum_{t=p+1}^{n} \log \left(\sum_{k=1}^{K} \alpha_k P(y_t|\boldsymbol{\theta}, Z_t = k) \right). \tag{2.8}$$

To obtain MLE, the two steps of the EM algorithm are performed iteratively until convergence, as follows:

E step: suppose that $\boldsymbol{\theta}$ is known and the missing data Z is replaced by the conditional expectation $\tau_{t,k}$ of k^{th} component which is defined as

$$\tau_{t,k} = P(Z_t = k|\theta, y_t) = \frac{P(y_t|Z_t = k)P(Z_t = k)}{P(y_t)}. \tag{2.9}$$

M step: we evaluate $\tau_{t,k}$ from E step to the conditional log likelihood and then find the estimates of parameters $\boldsymbol{\theta}$ can be obtained by maximizing the log likelihood function $l(\boldsymbol{\theta})$ with respect to the parameters $\boldsymbol{\theta}$.

2.3 Model selection Criteria

The three variable criteria used in this study are the Akaike information criterion(AIC) which is defined as

$$AIC = -2l + 2K, \tag{2.10}$$

where K represents the number of estimated parameters, and l denotes the maximized log-likelihood, calculated from the conditional probability density function as defined in equation (2.8). Secondly, the Bayesian information criterion(BIC) is given by

$$BIC = -2l + \log(S)K, \tag{2.11}$$

where S is the sample size, K represents the number of estimated parameters, and l denotes the maximized log-likelihood, calculated from the conditional probability density function which defined in equation (2.8). Thirdly, the Hannan-Quinn information criterion is defined as

$$HQIC = -2l + 2K\log(\log(S)), \tag{2.12}$$

where S is the sample size, K represents the number of estimated parameters, and l denotes the maximized log-likelihood, calculated from the conditional probability density function which defined in equation (2.8), and the best model chosen from the smallest value of each criterion.

2.4 Model diagnostics

All statistical models are collections of assumptions about the data generation process, and estimation is useless if these assumptions do not hold true for the data. As previously said, selecting a decent model is far more crucial than selecting a good prior.

2.4.1 Residual analysis

Definition 6. The residual for each observation is the difference between predicted values and actual data which is defined as

$$\hat{e}_t = y_t - \hat{y}_t, \tag{2.13}$$

where $\hat{e}_t$ is the residual for each observation, actual observation y_t, and predicted $\hat{y}_t$.

If the model is correctly specified and the parameter estimates are reasonably close to the true values, then the residuals should have nearly the properties of white noise.

They should behave similarly to independently distributed, identically distributed normal variables with zero means and common standard deviations. The plot of residuals over time is examined as a diagnostic check. If the model is suitable, we expect the plot to suggest a rectangular scatter around a zero horizontal level with no trends whatsoever.

Some tools for diagnostics include residual analysis, diagnostic plots, and checks for the normality of residuals, such as residual plots, Q-Q plots, histograms of residuals, and normality tests. However, the empirical counterparts of error terms e_{kt} cannot be calculated because the process generating each observation is unknown. Therefore, residual-based diagnostics are unavailable. Building on Kalliovirta's work [8], the quantile residuals are placed within a general framework. Computational tests are then derived to detect autocorrelation, conditional heteroscedasticity, and non-normality in quantile residuals.

$$R_t = \Phi^{-1}(F(y_t|\mathcal{F}_{t-1})), \tag{2.14}$$

where $t = 1, 2, 3, \ldots, n$, $\Phi^{-1}(\cdot)$ is the standard normal quantile function and $F(\cdot|\mathcal{F}_{t-1})$ is the conditional cumulative distribution function.

2.4.2 Normality test

A quantile-quantile (Q-Q) plot is an effective tool for assessing normality. It is a plot of the quantile residuals of two distributions against each other or a graphical representation based on quantile estimations. The pattern of points in the plot is utilized to compare the two distributions. The most crucial stage in creating a Q-Q plot is the calculation or estimation of the quantiles to be plotted. All quantiles are uniquely defined and can be obtained by inverting the cumulative distribution function (CDF) if one or both axes of a Q-Q plot are based on a theoretical distribution with a continuous CDF. Additionally, the Shapiro-Wilk normality test and Kolmogorov-Smirnov test, when applied to the residuals, yield a test statistic.

Definition 7. The Shapiro-Wilk goodness of fit test is a statistical test used to determine if a random sample, $Y = (y_1, \ldots, y_t)$ is drawn from a normal distribution with mean μ

and variance σ^2 which the test following hypothesis:

H_0 : The random sample was drawn from a normal distribution

H_a : The random sample does not follow normal distribution

Collect a sample of data to test the normality assumption, and then sort the data in ascending order. Use the sorted data and the sample size to calculate the test statistic, W, which is given by the formula

$$W = \frac{(\sum_{t=1}^{n} a_t y_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}, \tag{2.15}$$

where y_t is the t^{th} ordered observation, $\bar{y}$ is the sample mean, and a_t is the coefficient. Next, compare the calculated test statistic (W) to the critical value from the Shapiro-Wilk tables or use statistical software to obtain the p-value associated with the test statistic. If the p-value is less than the chosen significance level (commonly 0.05), indicate that the test rejects the null hypothesis and concludes that the data does not follow a normal distribution. If the p-value is greater than the significance level, it means that the test accepts the null hypothesis.

Definition 8. The Kolmogorov-Smirnov test is based on the empirical distribution function. Given t order data points $Y = (y_1, \dots, y_t)$ which is defined by

H_0 : The sample follow a specified distribution

H_a : The sample does not follow a specified distribution

Collect a sample of data to test the distribution and calculate the test statistic D which is the maximum absolute difference between the observed CDF of the sample and the expected CDF of the reference distribution

$$D = \max\Big(|F_n(y) - F(y)|\Big), \tag{2.16}$$

where $F_n(y)$ is is the empirical distribution function of the sample and $F(y)$ is the cumulative distribution function of the reference distribution. Compare the calculated test statistic (D) to the critical value from the Kolmogorov-Smirnov table or use statistical software to obtain the p-value associated with the test statistic. If the p-value is less than the chosen significance level (commonly 0.05), indicate that the test rejects the null hypothesis and If the p-value is greater than the significance level, the test fails to reject the null hypothesis.

2.5 Time series

A time series is a sequence of data points measured at successive points in time or over successive periods. These data points are often collected, recorded, or observed in sequential order. Time series analysis in statistics involves examining and modelling the patterns, trends, and dependencies within the data to make predictions or understand the underlying structure.

2.5.1 Means, Variances, and Covariances

For a stochastic process $\{Y_t : t = 0, \pm1, \pm2, \pm3, \ldots\}$, the mean function is defined by

$$\mu_t = E(Y_t) \text{ for all } t, \tag{2.17}$$

where μ_t is the expected value of the process at time t.

The autocovariance function, $\gamma_{t,k}$, is given by

$$\begin{aligned}\gamma_{t,k} &= Cov(Y_t, Y_k) \text{ for all } t \text{ and } k\\ &= E[(Y_t - \mu_t)(Y_k - \mu_k)]\\ &= E(Y_tY_k) - \mu_t\mu_k. \end{aligned} \tag{2.18}$$

The correlation between a time series and a lagged version of itself. In simpler terms, it quantifies the relationship between observations at different time points within the same time series. The autocorrelation function, $\rho_{t,k}$, is given by

$$\begin{aligned}\rho_{t,k} &= Corr(Y_t, Y_k) \text{ for all } t \text{ and } k \\ &= \frac{Cov(Y_t, Y_k)}{\sqrt{Var(Y_t)Var(Y_k)}} \\ &= \frac{\gamma_{t,k}}{\sqrt{(\gamma_{t,t})(\gamma_{k,k})}}. \qquad (2.19)\end{aligned}$$

2.5.2 Stationarity

A stationary time series is one whose statistical properties, such as mean, variance, and autocorrelation, remain constant over time. In other words, the behaviour of the time series does not exhibit systematic changes or trends, and it is considered to be in a stable and consistent state. There are two main two types, strictly stationary, in which the entire probability distribution of the time series remains unchanged over time, and weakly or second-order stationary, in which the mean, variance, and autocorrelation structure of the time series remain constant over time, though individual observations may not be identically distributed.

Definition 9. A process $\{Y_t\}$ is said to be strictly stationary if the joint distribution of $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ is the same as the joint distribution of $Y_{t_1-k}, Y_{t_{2-k}}, \dots, Y_{t_{n-k}}$ for all choices of time points $t_1, t_2, \dots, t_n$ and all choices of time lag k. Strictly stationary can be written as

$$P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n) = P(X_{t_1+k} \leq x_1, \dots, X_{t_n+k} \leq x_n). \qquad (2.20)$$

Definition 10. A time series $\{Y_t\}$ is said to be weakly stationary or second-order stationary as a stochastic process if

1. The mean function of the process does not depend on time as

$$E(Y_t) = \mu \text{ for all } t, \qquad (2.21)$$

2. The variance of the process is constant, which is defined as

$$Var(Y_t) = \sigma^2 \text{ for all } t, \tag{2.22}$$

3. The covariance between Y_t and Y_{t-k} depends on time k which is defined by

$$\gamma(k) = Cov(Y_t, Y_{t-k}), \tag{2.23}$$

this is called the autocovariance function.

2.5.3 Time series models

Stationary time series models are statistical models that are based on the assumption that the underlying time series data is stationary. Stationarity is an important assumption in many time series models since it simplifies analysis and makes making solid predictions easier. The stationary time series models such as the autoregressive (AR), moving average (MA), and autoregressive moving average (ARMA) models are the most fundamental stationary models in time series analysis. There is also a non-stationary model, including the autoregressive integrated moving average (ARIMA) model, which is a generalization of the autoregressive moving average (ARMA) model, and the random walk.

2.5.3.1 Autoregressive processes

The Autoregressive (AR) model is a type of time series model that expresses the current observation in terms of its past values. Autoregressive process Y_t is a linear combination of the observation at p previous time in the past which denoted as AR(p) is given by

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \cdots + \phi_p Y_{t-p} + e_t, \tag{2.24}$$

where Y_t is the value of the time series at time t, $\phi_1, \phi_2, \ldots, \phi_p$ are the autoregressive coefficients and e_t is the error term or white noise at time t.

We call such a series a Autoregressive of order p as AR(p) with AR characteristic polynomial

$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2 - \phi_3 x^3 - \cdots - \phi_p x^p, \tag{2.25}$$

and corresponding AR characteristic equation

$$1 - \phi_1 x - \phi_2 x^2 - \phi_3 x^3 - \cdots - \phi_p x^p = 0. \tag{2.26}$$

Assuming that e_t is independent of $Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-4}, \ldots, Y_{t-p}$ the stationary solution to equation (2.26) occurs if and only if the absolute value (modulus) of the p roots of the AR characteristic equation exceeds 1. Other relationships between polynomial roots and coefficients can be used to demonstrate that the following two inequalities are required for stationarity: That is, for the roots to have a modulus greater than one, it is necessary but not sufficient that both

$$\begin{cases} \phi_1 + \phi_2 + \phi_3 + \cdots + \phi_p < 1 \\ \text{and} \;\; |\phi_p| < 1. \end{cases} \tag{2.27}$$

2.5.3.2 Moving average processes

The Moving Average (MA) model is a time series model used to explain the relationship between an observation and a residual error from a moving average process. It is often denoted as MA(q), where q represents the order of the moving average.

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - \cdots - \theta_q e_{t-q}, \tag{2.28}$$

where Y_t is the value of the time series at time t, $\theta_1, \ldots, \theta_q$ are the parameters of the model, representing the past error terms, and e_t is the error term or white noise at time

t. The variance of MA(q) process is given by

$$VAR(Y_t) = VAR(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}) \tag{2.29}$$

$$= \sigma^2(1 + \theta_1^2 + \cdots + \theta_q^2). \tag{2.30}$$

The autocorrelation of general MA(q) process is defined as

$$\rho_k = \begin{cases} \frac{-\theta_k+\theta_1\theta_{k+1}+\theta_2\theta_{k+2}+\ldots++\theta_q\theta_{q-k}}{1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2} & \\ 0 & \text{for } k > q, \end{cases} \tag{2.31}$$

where the numerator of ρ_q is simply θ_q. The autocorrelation function cuts off after lag q, meaning it becomes zero. Its shape can take on almost any form for the earlier lags.

2.5.3.3 Autoregressive Moving average model

We assume that the mixed series between autoregressive and moving average, obtained a time series model that is autoregressive moving model which defined as

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}, \tag{2.32}$$

where Y_t is a mixed autoregressive moving average process of orders p and q, respectively, which we call ARMA(p, q).

2.5.3.4 Vector autoregressive model

A vector autoregressive (VAR) model is a multivariate time series model comprising a system of n equations with n distinct, stationary response variables represented as linear functions of lagged responses and other terms. The VAR models are also characterized by their degree p which is denoted VAR(p) is written as

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \Phi_3 Y_{t-3} + \cdots + \Phi_p Y_{t-p} + e_{n,t}, \tag{2.33}$$

where $\Phi_1, \Phi_2, \ldots, \Phi_p$ are $(n \times n)$ coefficient matrices for lags 1 through p, Y_t is a vector of endogenous variables at time t, n is a dimension vector, $e_t \overset{\text{iid}}{\sim} N(0, \sigma^2)$, mean $E[e_t] = 0$, and covariance matrix $E[e_t e_t'] = \Omega$. The error term is related to the covariance matrix, which is a $k \times k$ positive semi definite matrix labeled Ω.

Data points often exhibit non-stationary characteristics, such as varying means, variances, and covariances across time. Non-stationary behaviours may manifest as trends, cycles, random walks, or combinations of the three. In this study, we introduce non-stationary time series, particularly the autoregressive integrated moving average (ARIMA) model.

2.5.3.5 Autoregressive integrated Moving average model

An autoregressive integrated moving average (ARIMA) model is a generalisation of an autoregressive moving average (ARMA) model, the order of which is commonly expressed by the notation ARIMA(p, d, q), where p is the autoregressive component order, d is the difference $D_t = \nabla^d Y_t$ is a stationary ARMA process, and q is the moving-average process order. With $D_t = Y_t - Y_{t-1}$, we have

$$D_t = \phi_1 D_{t-1} + \phi_2 D_{t-2} + \ldots + \phi_p D_{t-p} + e_t + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}, \tag{2.34}$$

where $\phi_1, \ldots, \phi_p$ are the autoregressive coefficients, the $\theta_1, \ldots, \theta_q$ are the parameter of the moving average coefficients, and the e_t are the error terms, which are generally assumed to be independent and identically distributed to the normal distribution.

CHAPTER III

THE FAMILY OF UNIVARIATE MIXTURE AUTOREGRESSIVE MODELS

In this chapter, we introduce the family of univariate mixture autoregressive models and discuss their specifications. This family includes the mixture autoregressive model and the t mixture autoregressive model, both of which are estimated using maximum likelihood estimation and the EM algorithm. Initially, we evaluate the performance of the family of univariate mixture autoregressive models on individual stock markets. We consider the top stocks from two different sectors, with each sector comprising three stocks. These stocks include BANPU, ESSO, and BCP from the energy and utility sectors, and HANA, TEAM, and KCE from the electronic components sector. Subsequently, we compare the performance of parameter estimation using information criteria and mean square error.

3.1 Mixture autoregressive model

The mixture autoregressive (MAR) model is a mixture of different autoregressive components. Specifically, the time series $\{y_t\}_{t\geq 1}$ is said to be the K-component MAR model, ϕ_{ki} is the coefficients for the k^{th} component, where $i = 1, 2, \ldots, p_k$, denoted as MAR$(K; p_1, p_2, p_3, \ldots, p_K)$, if it satisfies

$$F(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \Phi\Big(\frac{y_t - \phi_{k0} - \phi_{k1}y_{t-1} - \cdots - \phi_{kp_k}y_{t-p_k}}{\sigma_k}\Big), \tag{3.1}$$

where $F(y_t|\mathcal{F}_{t-1})$ is the cumulative distribution function of data y_t given the past information $y_{t-1}, y_{t-2}, y_{t-3}, \ldots,$ y_1, $\mathcal{F}_t$ is the information set up to time t, the function $\Phi(\cdot)$ represents the cumulative distribution function of the standard normal distribution and $\alpha_1 + \alpha_2 + \cdots + \alpha_K = 1$, $1 > \alpha_k > 0$, $k = 1, 2, 3, \ldots, K$.

The mixture autoregressive model's conditional mean and variance are provided as

$$E(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \mu_{kt},$$

where $\mu_{kt} = \phi_{k0} + \phi_{k1} y_{t-1} + \cdots - \phi_{kp_k} y_{t-p_k}$ and

$$\mathrm{Var}(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \mu_{kt}^2 + \sum_{k=1}^{K} \alpha_k \sigma_k^2 - \left(\sum_{k=1}^{K} \alpha_k \mu_{kt}\right)^2,$$

respectively.

In the case of the mixture autoregressive model, the maximum likelihood method is the parameter estimation method used in this study. Specifically, given a time series $\mathbf{y} = (y_1, y_2, \ldots, y_t)$, the likelihood function for the mixture autoregressive model is the product of conditional density

$$L(\phi, \sigma, \alpha|\mathbf{y}) = \prod_{t=p+1}^{n} \sum_{k=1}^{K} \frac{\alpha_k}{\sigma_k} \phi\Big(\frac{y_t - \sum_{i=1}^{p_k} \phi_{ki} y_{t-i}}{\sigma_k}\Big), \tag{3.2}$$

the log likelihood function of the mixture autoregressive model can be written as

$$l(\phi, \sigma, \alpha) = \sum_{t=p+1}^{n} \log\Big[\sum_{k=1}^{K} \frac{\alpha_k}{\sigma_k} \phi\Big(\frac{y_t - \sum_{i=1}^{p_k} \phi_{ki} y_{t-i}}{\sigma_k}\Big)\Big]. \tag{3.3}$$

Some parameters of the mixture autoregressive model may not be solved in general [7]. Consequently, the estimation of these parameters must be performed using a numerical method [5].

3.1.1 Simulation study for the MAR model

In this section, we study the performance of parameter estimation using the maximum likelihood estimation procedure implemented in the "uGMAR"[5]. We examine the correctness in choosing the number of components, K, and the, p, order of the autoregressive models. Furthermore, we examine the accuracy of parameter estimates. It's important to note that, under the restrictions of the package, the orders of autoregressive components for different components are assumed to be the same. Therefore, the models considered in this study are denoted as $(K; p)$, where K is the number of components and p is the common order of autoregressive components. The two models investigated in this study are the MAR$(2; 2)$, where K component is 2 and order p is 2, and the MAR$(3; 2)$ model, where K component is 3 and order p is 2.

In the first experiment, we generated a time length of 1000 data points from the MAR$(2; 2)$ model, where the coefficients $\alpha_1, \alpha_2, \phi_{10}, \phi_{11}, \phi_{12}, \sigma_1, \phi_{20}, \phi_{21}, \phi_{22}, \sigma_2$ are 0.65, 0.35, 0.02, 1.25, -0.26, 0.02, 0.1, 1.26, -0.32, 0.06, respectively. The data are fitted to the mixture autoregressive model for $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$ to assess the accuracy of the model selection. The corresponding AICs, HQICs, and BICs are obtained, and the model with the smallest values of the criterion statistics is selected to match the generated model.

Table 3.1: Criteria for the Simulation of the MAR(2;2) model

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	3329.921	3335.516	3344.642	MAR(3:1)	3337.778	3358.293	3391.752
MAR(1:2)	3296.972	3304.431	3316.595	MAR(3:2)	3263.439	3288.545	3331.120
MAR(1:3)	3290.448	3299.770	3314.971	MAR(3:3)	3271.536	3303.231	3354.917
MAR(1:4)	3288.045	3299.229	3317.467	MAR(3:4)	3278.134	3315.416	3376.208
MAR(2:1)	3331.997	3345.052	3366.344	MAR(4:1)	3340.433	3368.408	3414.034
MAR(2:2)	3262.900	3279.683	3307.052	MAR(4:2)	3278.041	3313.470	3371.250
MAR(2:3)	3273.840	3294.348	3327.792	MAR(4:3)	3277.415	3320.297	3390.224
MAR(2:4)	3274.082	3298.316	3337.831	MAR(4:4)	3282.092	3332.423	3414.493

From Table 3.1, The MAR$(K; p)$ model, whose K component is equal to 1, such as MAR(1;p), in the first four lines, is the original autoregressive model with order p, while the other K components represent the MAR models with multiple components. The best candidate model, determined by the smallest corresponding criterion, is the MAR(2;2) model. This result highlights the accuracy of the criterion in selecting the optimal model. Table 3.2 presents the parameter estimation of the MAR(2;2) models compared with the true values of the parameters we generated.

Table 3.2: Parameter estimates for MAR(2;2) models

	α_1	σ_1	ϕ_{10}	ϕ_{11}	ϕ_{12}
True value	0.65	0.02	0.02	1.25	-0.26
Mean of estimates	0.67	0.02	0.03	1.13	-0.15
Empirical standard error	0.02	0.00	0.01	0.12	0.11
Theoretical standard error	0.06	0.00	0.01	0.15	0.15
	α_2	σ_2	ϕ_{20}	ϕ_{21}	ϕ_{22}
True value	0.35	0.06	0.10	1.26	-0.32
Mean of estimates	0.33	0.06	0.12	1.23	-0.30
Empirical standard error	0.02	0.00	0.02	0.03	0.02
Theoretical standard error	0.06	0.00	0.06	0.08	0.09

From Table 3.2 show that the mean of the estimation using Maximum Likelihood Estimation (MLE) is quite close to the true value we generated. The theoretical standard error and the empirical standard error are close, except in the parameters of α_1 and α_2. Therefore, the fitting of the data is correct to choose the model, and the parameter estimate is quite accurate to the true value.

In the second experiment, we generated a time series with 1000 data points from MAR(3; 2) model, which the coefficients $\alpha_1, \alpha_2, \alpha_3, \phi_{10}, \phi_{11}, \phi_{12}, \sigma_1, \phi_{20}, \phi_{21}, \phi_{22}, \sigma_2, \phi_{30}, \phi_{31}, \phi_{32}, \sigma_3$ are 0.08, 0.59, 0.33, 0.04, 1.24, -0.26, 0.03, 0.03, 1.44, -0.46, 0.01, 0.11, 1.25, -0.32, 0.08, respectively. The data are fitted to the mixture autoregressive model for $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$ to assess the accuracy of the model selection. The corresponding AICs, HQICs, and BICs are obtained, and the model with the smallest criterion is then selected to match the generated model.

Table 3.3: Criteria for the Simulation of the MAR(3;2) model

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	3366.566	3372.161	3381.286	MAR(3:1)	3368.609	3389.124	3217.291
MAR(1:2)	3197.668	3205.127	3422.583	MAR(3:2)	3177.452	3195.612	3252.414
MAR(1:3)	3195.831	3205.153	3220.355	MAR(3:3)	3190.232	3221.927	3273.613
MAR(1:4)	3193.981	3205.166	3223.404	MAR(3:4)	3185.144	3222.427	3283.219
MAR(2:1)	3369.439	3382.494	3403.787	MAR(4:1)	3375.655	3403.630	3449.257
MAR(2:2)	3178.829	3209.839	3222.981	MAR(4:2)	3190.324	3225.754	3283.533
MAR(2:3)	3180.831	3201.340	3234.784	MAR(4:3)	3184.771	3227.653	3297.581
MAR(2:4)	3188.115	3212.348	3251.863	MAR(4:4)	3183.734	3227.783	3309.853

From Table 3.3, the best candidate model, corresponding to AIC and HQIC, is the MAR(3;2) model, but in the case of BIC, the best model is MAR(3;1). However, the 2 out of 3 criteria lead to the selection of the MAR(3;2) model as the best match, aligning with the model we generated. Tables 3.4 shows the parameter estimation of MAR(3;2) models comparing with the true values of parameters.

Table 3.4: Parameter estimates for MAR(3;2) models

	α_1	σ_1	ϕ_{10}	ϕ_{11}	ϕ_{12}
True value	0.08	0.03	0.04	1.24	-0.26
Mean of estimates	0.10	0.02	0.04	1.17	-0.19
Empirical standard error	0.02	0.01	0.00	0.07	0.07
Theoretical standard error	0.16	0.01	0.05	0.18	0.20
	α_2	σ_2	ϕ_{20}	ϕ_{21}	ϕ_{22}
True value	0.59	0.01	0.02	1.44	-0.46
Mean of estimates	0.67	0.01	0.07	1.02	-0.23
Empirical standard error	0.08	0.00	0.05	0.42	0.23
Theoretical standard error	0.15	0.01	0.01	0.23	0.24
	α_3	σ_3	ϕ_{30}	ϕ_{31}	ϕ_{32}
True value	0.33	0.08	0.11	1.25	-0.32
Mean of estimates	0.23	0.09	0.26	1.15	-0.30
Empirical standard error	0.10	0.01	0.15	0.10	0.02
Theoretical standard error	0.15	0.03	0.13	0.16	0.20

From Table 3.4, the parameter estimation for the MAR$(3;2)$ model, the mean of the estimates using Maximum Likelihood Estimation (MLE) is quite close to the true values used to generate the data. The theoretical standard error and the empirical standard error are close, except in the parameters of α_1, α_2 and α_3. Therefore, we observe a small bias in this simulation study, with the theoretical standard errors being smaller than the empirical standard errors. In general, the empirical and theoretical standard errors for the parameters are close, suggesting that the accuracy of the estimates is reasonably good. Next, we will apply the model to the Thai stock data in the next section.

3.1.2 Mixture autoregressive model for Thai stock markets

In this section, we investigate the performance of the mixture autoregressive model on individual stock markets using the daily closing prices of the top stock from the energy and utility, and electronic component sectors over the five-year period from August 1st, 2017, to August 1st, 2022 (1214 observations). In particular, BANPU, ESSO, and BCP from the energy and utility sector, HANA, TEAM, and KCE from the electronic components sector. The histogram of each dataset is shown in Figure 3.1, in which we can see that the histogram has a multimodal. In the tables below, we show the AIC, HQIC, and BIC values of each candidate model of the MAR model for each stock. Subsequently, we provide an analysis of the best MAR model.

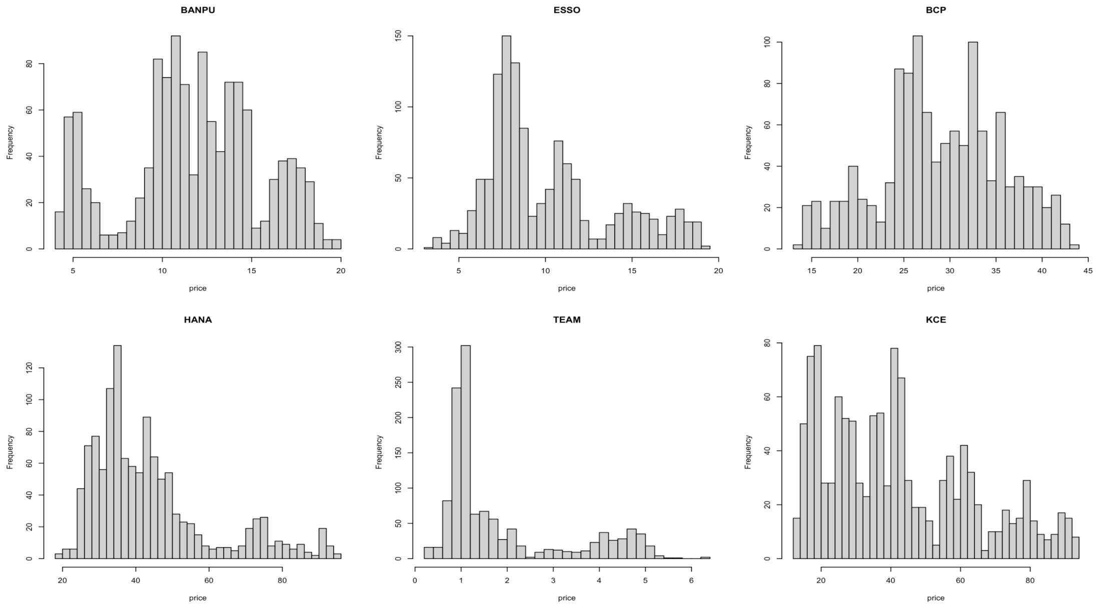


Figure 3.1: The histogram of stock market dataset

From Figure 3.1, the histogram reveals that the data exhibits various modes, including bimodal, trimodal, and multimodal patterns. To begin the analysis, we fit the BANPU stock from the energy and utility sectors with the MAR$(K;p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$ due to calculation complexity, and the histograms suggest less than four peaks. The criteria values for each model are provided in Table 3.5.

Table 3.5: Criteria for the candidate MAR model for BANPU

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	291.96387	297.72516	307.26643	MAR(3:1)	26.40063	47.52535	82.51000
MAR(1:2)	293.92742	301.60821	314.32753	MAR(3:2)	66.93240	93.81516	138.33278
MAR(1:3)	296.50757	306.10740	322.00358	MAR(3:3)	83.92518	116.56458	170.61161
MAR(1:4)	299.07955	310.59795	329.66981	MAR(3:4)	99.62837	138.02301	201.59588
MAR(2:1)	60.79663	74.23964	96.50259	MAR(4:1)	17.12187	45.92831	93.63465
MAR(2:2)	89.71769	106.99947	135.61794	MAR(4:2)	58.82009	95.30384	155.72061
MAR(2:3)	104.73908	125.85869	160.83030	MAR(4:3)	76.42850	120.58768	193.71014
MAR(2:4)	118.70635	143.66286	184.98523	MAR(4:4)	98.11703	149.94979	235.77317

From Table 3.5, the MAR($K;p$) model, whose K component is equal to 1, such as MAR(1;p), in the first four lines, is the original autoregressive model with order p, while the other K components represent the mixture autoregressive models with multiple components. All the criterion values for multiple components are smaller than those for the single component, confirming the motivation of the mixture distribution in the stock dataset. Among these models, the one with the smallest AIC and HQIC is the MAR(4:1), while the smallest BIC corresponds to the MAR(3;1). Based on the 2 out of 3 criteria, the MAR(4;1) is identified as the best model. Therefore, the optimal model for BANPU is the MAR(4;1) model.

Next, the diagnostic check involves quantile residuals, which are utilized for computationally simple tests aimed at detecting autocorrelation, quantile residual plots, Q-Q plots, and histograms of quantile residuals are shown in Figure 3.2 and Figure 3.3, respectively. The normality test of the quantile residuals is presented in Table 3.6.

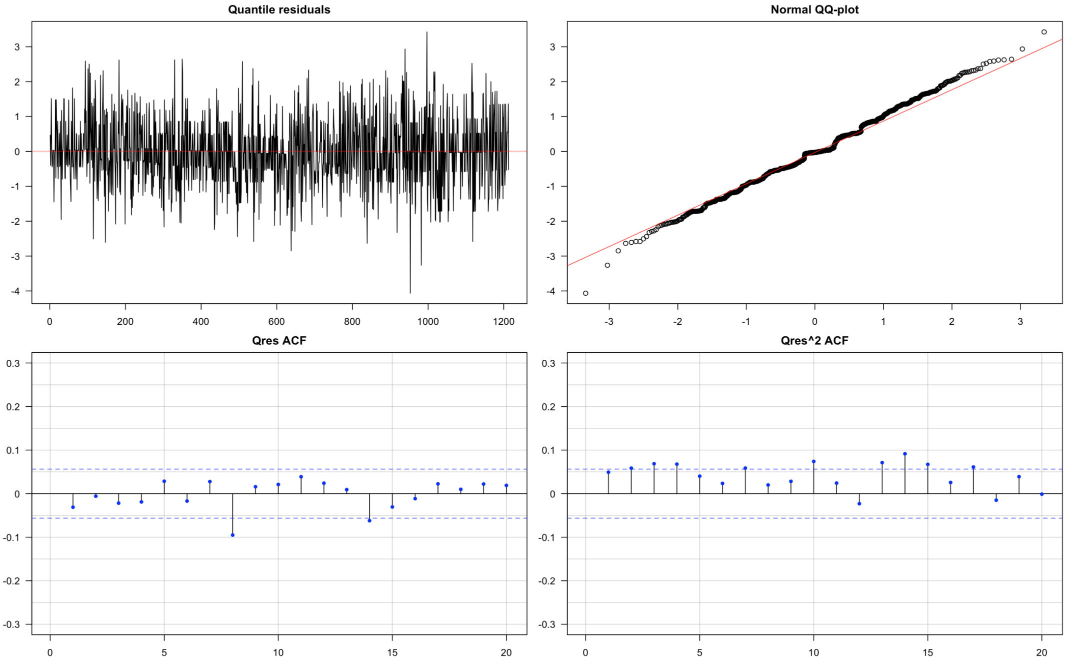


Figure 3.2: Quantile residual plot of MAR(4:1) for BANPU

From Figure 3.2, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. The autocorrelation plot of quantile residuals shows a spike at lags 8 and 14, although it is not highly significant. Additionally, Figure 3.3 includes a histogram of the quantile residuals. The results from normality tests, specifically the Shapiro-Wilk test and Kolmogorov-Smirnov test, are presented in Table 3.6.

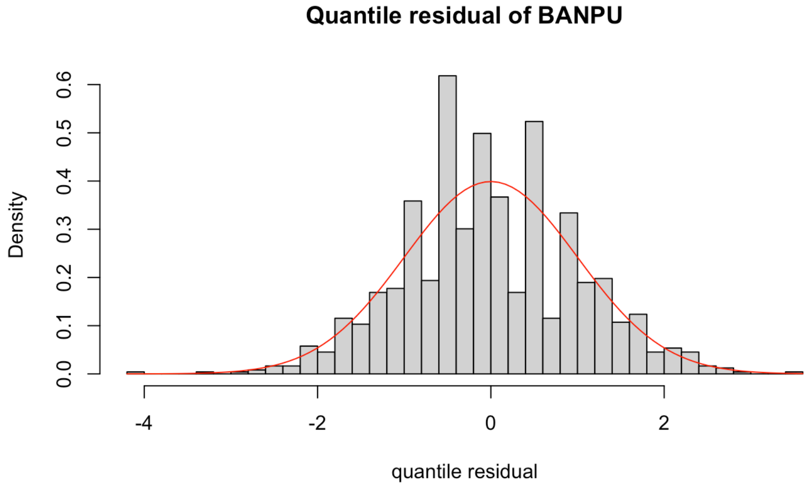


Figure 3.3: Histogram of quantile residual of MAR(4:1) for BANPU

Table 3.6: Normality test of MAR(4:1) for BANPU

Test	Statistic	p-value
Shapiro-Wilk	0.9961	0.0033
Kolmogorov-Smirnov	0.0672	0.0000

In the normality test of the quantile residuals, it is evident that the histogram does not fit the normal curve. Both normality tests, the Shapiro-Wilk test and the Kolmogorov-Smirnov test, reveal p-values less than 0.05. Consequently, the distribution of the given data does not conform to a normal distribution.

Next, we analyze ESSO stock data in the energy and utility sectors by fitting it with the MAR$(K;p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$. The criteria values for each model are presented in Table 3.7.

Table 3.7: Criteria for the candidate MAR model for ESSO

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	277.218070	282.97936	292.52063	MAR(3:1)	14.670105	35.79483	70.77948
MAR(1:2)	277.753155	285.43394	298.15326	MAR(3:2)	24.144946	51.02771	95.54533
MAR(1:3)	280.475325	290.07515	305.97133	MAR(3:3)	34.464661	67.10406	121.15109
MAR(1:4)	264.807048	276.32544	295.39730	MAR(3:4)	38.488671	76.88331	140.45618
MAR(2:1)	42.200631	55.64364	77.90659	MAR(4:1)	6.457981	35.26442	82.97076
MAR(2:2)	55.935559	73.21733	101.83580	MAR(4:2)	23.035381	59.51913	119.93590
MAR(2:3)	59.550001	80.66961	115.64122	MAR(4:3)	27.350684	71.50987	144.63232
MAR(2:4)	59.184602	84.14112	125.46348	MAR(4:4)	28.840646	80.67341	166.49679

From Table 3.7, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. However, the MAR(4;1) model exhibits the smallest AIC and HQIC value, while the BIC criteria favours the MAR (3;1) model. Therefore, considering the three criteria, two out of three indicate that the MAR (4;1) model is the best model for ESSO stock data.

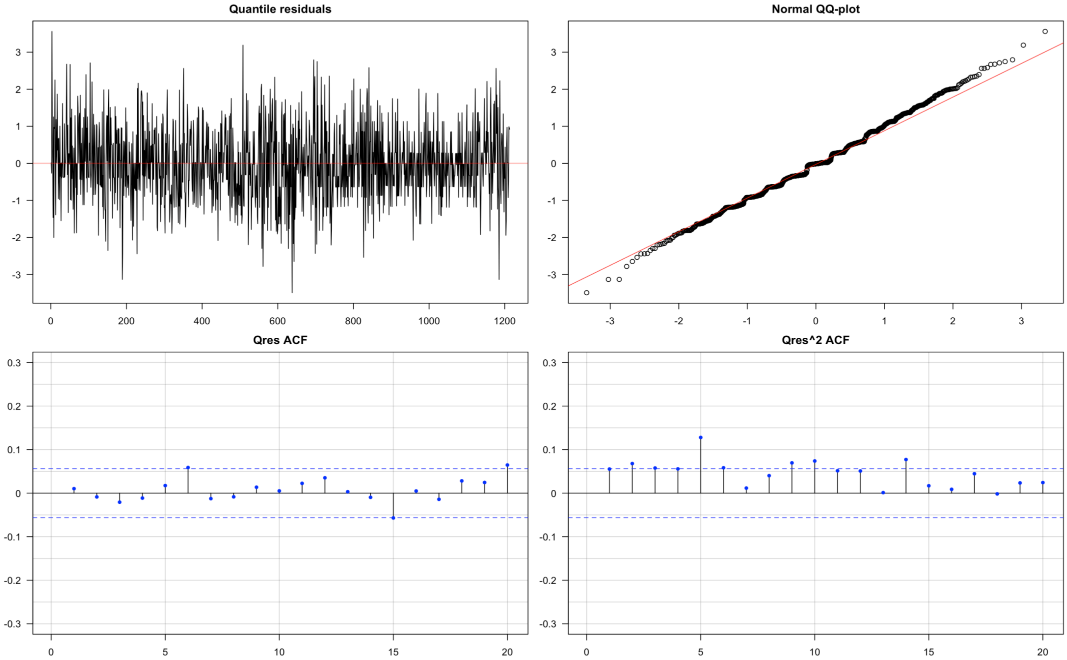


Figure 3.4: Quantile residual plot of MAR(4:1) for ESSO

From Figure 3.4, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. The autocorrelation plot of quantile residuals shows a spike at lags 6 and 20, although it is not highly significant. Additionally, Figure 3.5 includes a histogram of the quantile residuals. The results from normality tests, specifically the Shapiro-Wilk test and Kolmogorov-Smirnov test, are presented in Table 3.8.

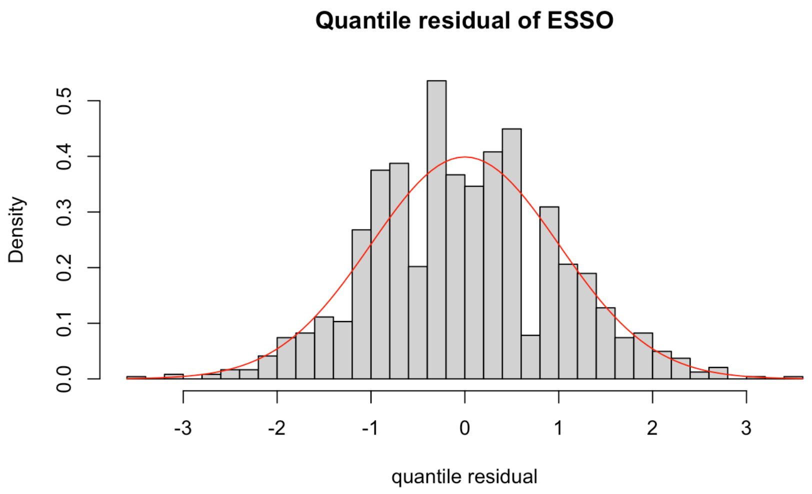


Figure 3.5: Histogram residual of MAR(4:1) for ESSO

Table 3.8: Normality test of MAR(4:1) for ESSO

Test	Statistic	p-value
Shapiro-Wilk	0.9965	0.0081
Kolmogorov-Smirnov	0.0561	0.0010

In the normality test of the quantile residuals for ESSO, it is evident that the histogram does not fit the normal curve. Both normality tests, the Shapiro-Wilk test and the Kolmogorov-Smirnov test, yield p-values less than 0.05. Consequently, the distribution of the given data does not follow a normal distribution.

Next, we analyze BCP stock data in the energy and utility sectors by fitting it with the MAR($K;p$) model, where we explore values of $K = 1,2,3,4$ and $p = 1,2,3,4$. The criteria values for each model are presented in Table 3.9.

Table 3.9: Criteria for the candidate MAR model for BCP

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	2001.458	2007.219	2016.760	MAR(3:1)	1876.064	1897.189	1932.174
MAR(1:2)	2002.554	2010.235	2022.954	MAR(3:2)	1893.043	1919.926	1964.443
MAR(1:3)	2003.298	2012.898	2028.794	MAR(3:3)	1876.633	1909.273	1963.320
MAR(1:4)	2003.598	2015.117	2034.188	MAR(3:4)	1887.430	1925.824	1989.397
MAR(2:1)	1882.570	1896.013	1918.276	MAR(4:1)	1868.379	1897.185	1944.892
MAR(2:2)	1901.797	1919.078	1947.697	MAR(4:2)	1880.222	1916.706	1977.122
MAR(2:3)	1886.544	1907.664	1942.636	MAR(4:3)	1867.738	1911.898	1985.020
MAR(2:4)	1904.845	1929.802	1971.124	MAR(4:4)	1875.381	1927.214	2013.037

From Table 3.9, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. However, the MAR(4;3) model exhibits the smallest AIC value, while the HQIC and BIC criteria favour the MAR(2;1) model. Therefore, considering the three criteria, two out of three indicate that the MAR(2;1) model is the best model for BCP stock data.

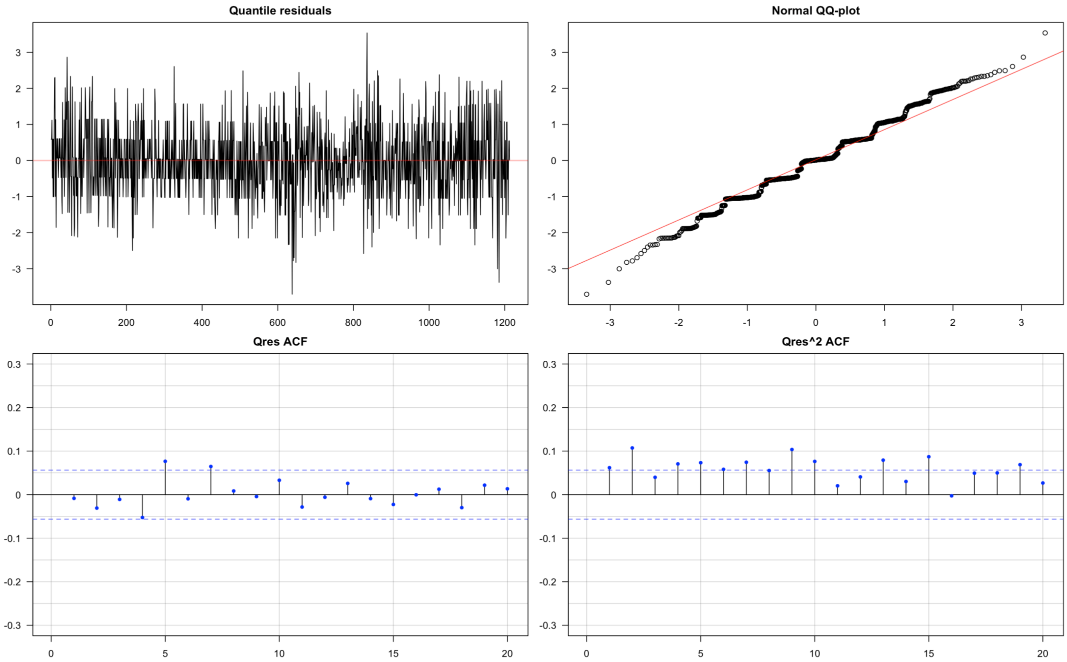


Figure 3.6: Quantile residual plot of MAR(2:1) for BCP

From Figure 3.6, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. The autocorrelation plot of quantile residuals shows a spike at lags 5 and 7, although it is not highly significant. Additionally, Figure 3.7 includes a histogram of the quantile residuals. The results from normality tests, specifically the Shapiro-Wilk test and Kolmogorov-Smirnov test, are presented in Table 3.10.

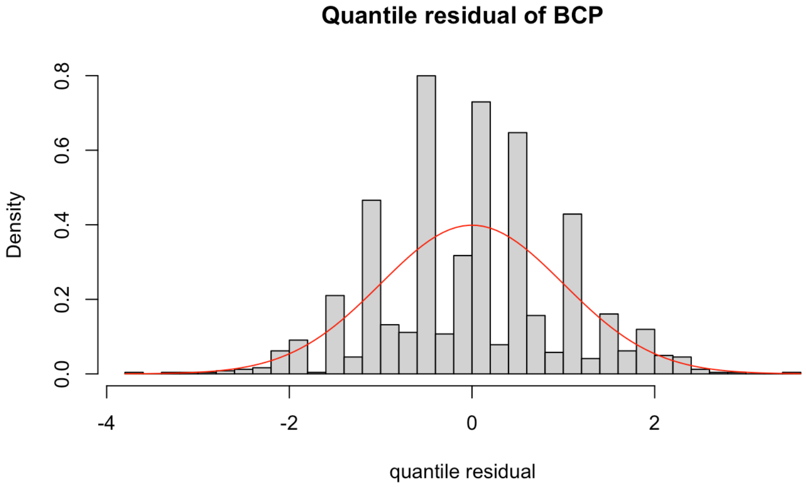


Figure 3.7: Histogram residual of MAR(2:1) for BCP

Table 3.10: Normality test of MAR(2:1) for BCP

Test	Statistic	p-value
Shapiro-Wilk	0.9907	0.0000
Kolmogorov-Smirnov	0.0683	0.0000

In the normality test of the quantile residuals for BCP, it is evident that the histogram does not fit the normal curve. Both normality tests, the Shapiro-Wilk test and the Kolmogorov-Smirnov test, reveal p-values less than 0.05. Consequently, the distribution of the given data does not conform to a normal distribution.

Next, we analyze HANA stock data in the electronic components sectors by fitting it with the MAR$(K;p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$. The criteria values for each model are presented in Table 3.11.

Table 3.11: Criteria for the candidate MAR model for HANA

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	4037.957	4043.718	4053.260	MAR(3:1)	3671.415	3692.540	3727.525
MAR(1:2)	4029.229	4036.910	4049.630	MAR(3:2)	3667.780	3694.663	3739.180
MAR(1:3)	4026.492	4036.092	4051.988	MAR(3:3)	3664.830	3697.470	3751.517
MAR(1:4)	4023.619	4035.137	4054.209	MAR(3:4)	3683.869	3722.264	3785.837
MAR(2:1)	3709.368	3722.811	3745.074	MAR(4:1)	3668.354	3697.161	3744.867
MAR(2:2)	3716.446	3733.727	3762.346	MAR(4:2)	3667.833	3704.317	3764.733
MAR(2:3)	3714.777	3735.896	3770.868	MAR(4:3)	3656.219	3700.378	3773.500
MAR(2:4)	3713.376	3738.333	3779.655	MAR(4:4)	3686.830	3738.663	3824.487

From Table 3.11, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. However, the MAR(4;3) model exhibits the smallest AIC value, while the HQIC and BIC criteria favour the MAR(3;1) model. Therefore, considering the three criteria, two out of three indicate that the MAR(3;1) model is the best model for HANA stock data.

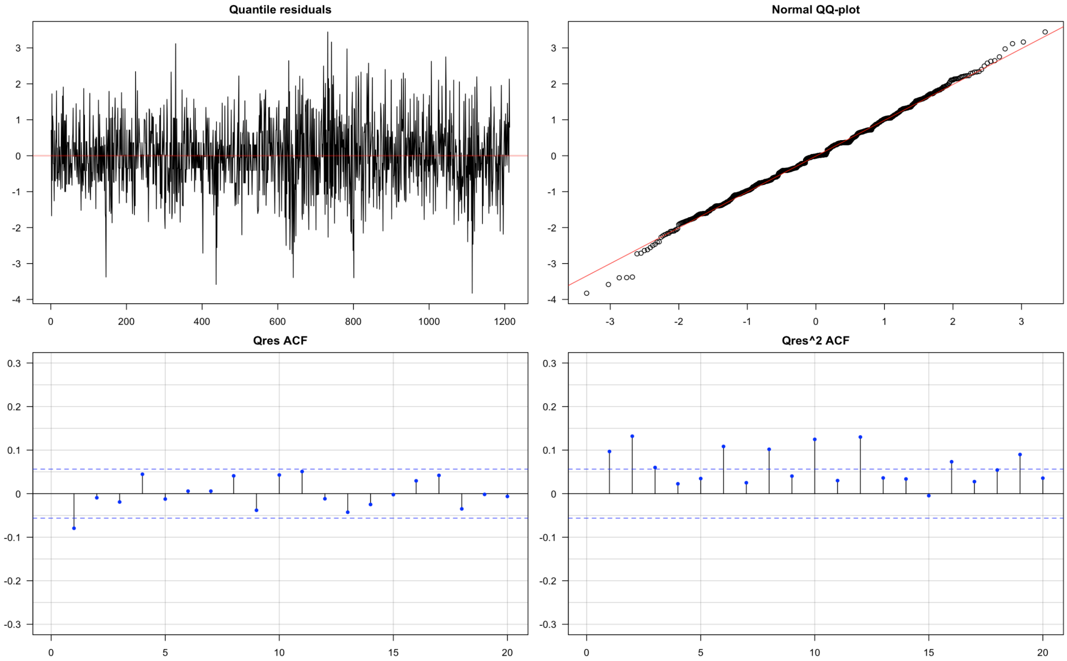


Figure 3.8: Quantile residual plot of MAR(3:1) for HANA

From Figure 3.8, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. The autocorrelation plot of quantile residuals shows a spike at lags 1, although it is not highly significant. Additionally, Figure 3.9 includes a histogram of the quantile residuals. The results from normality tests, specifically the Shapiro-Wilk test and Kolmogorov-Smirnov test, are presented in Table 3.12.

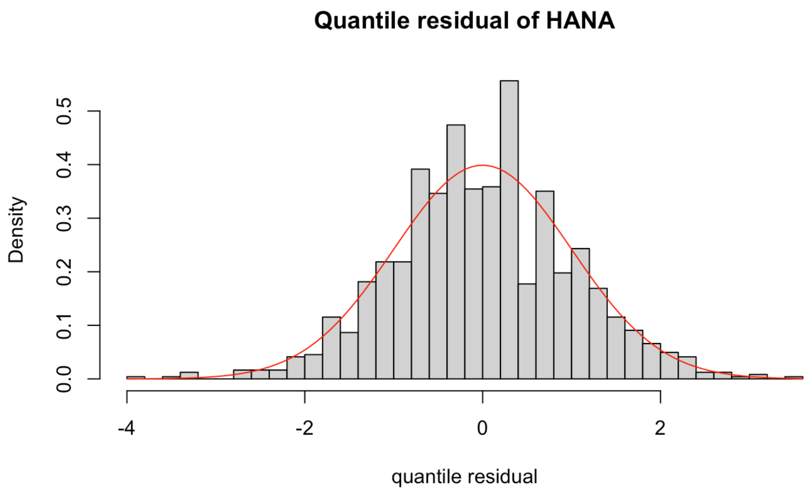


Figure 3.9: Histogram residual of MAR(3:1) for HANA

Table 3.12: Normality test of MAR(3:1) for HANA

Test	Statistic	p-value
Shapiro-Wilk	0.9967	0.0112
Kolmogorov-Smirnov	0.0468	0.0098

In the normality test of the quantile residuals for HANA, it is evident that the histogram does not conform to a normal distribution curve. Both normality tests, the Shapiro-Wilk test and the Kolmogorov-Smirnov test, yield p-values less than 0.05. Consequently, the distribution of the given data does not follow a normal distribution.

Next, we analyze TEAM stock data in the electronic components sectors by fitting it with the MAR$(K;p)$ model, where we explore values of $K = 1,2,3,4$ and $p = 1,2,3,4$. The criteria values for each model are presented in Table 3.13.

Table 3.13: Criteria for the candidate MAR model for TEAM

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	-1996.893	-1991.132	-1981.590	MAR(3:1)	-4194.116	-4172.991	-4138.006
MAR(1:2)	-1997.430	-1989.749	-1977.029	MAR(3:2)	-4163.073	-4136.191	-4091.673
MAR(1:3)	-2008.362	-1998.762	-1982.866	MAR(3:3)	-4062.829	-4030.189	-3976.142
MAR(1:4)	-2013.303	-2001.785	-1982.713	MAR(3:4)	-4021.485	-3983.090	-3919.517
MAR(2:1)	-3961.356	-3947.913	-3925.650	MAR(4:1)	-4228.198	-4199.391	-4151.685
MAR(2:2)	-3899.877	-3882.596	-3853.977	MAR(4:2)	-4211.877	-4175.393	-4114.976
MAR(2:3)	-3781.215	-3760.096	-3725.124	MAR(4:3)	-4155.016	-4110.857	-4037.734
MAR(2:4)	-3708.141	-3683.185	-3641.862	MAR(4:4)	-4114.097	-4062.264	-3976.441

From Table 3.13 shows that multiple components have smaller AIC, HQIC, and BIC values than the single component. Among these, the MAR(4;1) model exhibits the smallest AIC, HQIC, and BIC criteria. Therefore, considering the three criteria, it indicates that the MAR(4;1) model is the best fit for TEAM stock data.

The diagnostics check are the quantile residuals, which are used to obtain computationally simple tests aimed at detecting autocorrelation, quantile residual plot and Q-Q plot and the histogram of quantile residual are shown in Figure 3.10 and Figure 3.11, respectively. The normality test of quantile residuals is shown in Table 3.14.

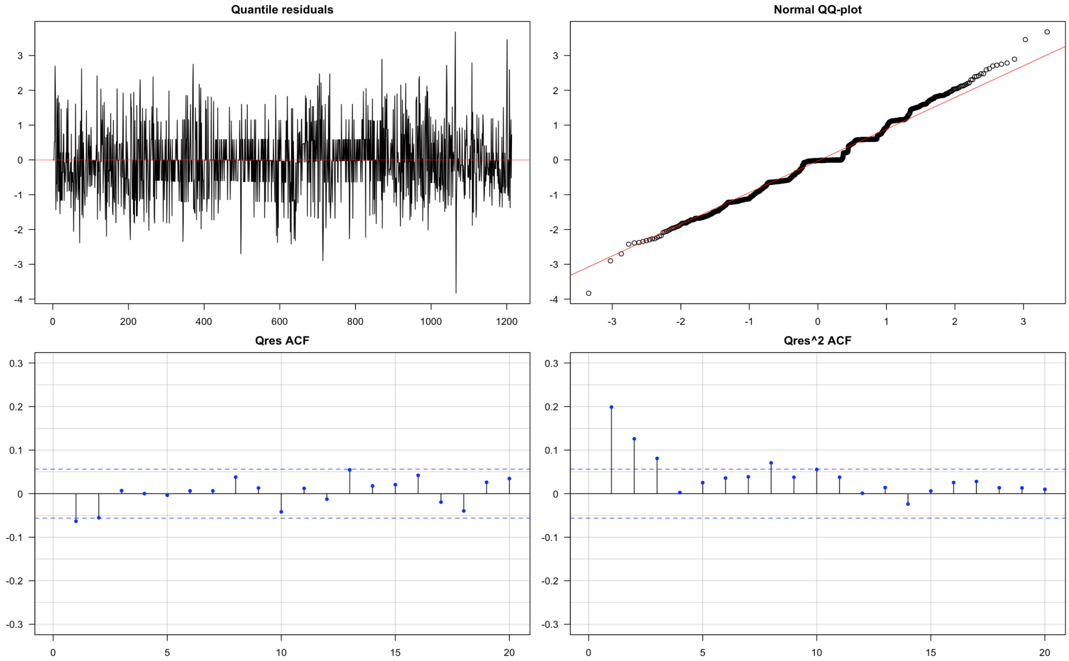


Figure 3.10: Quantile residual plot of MAR(4:1) for TEAM

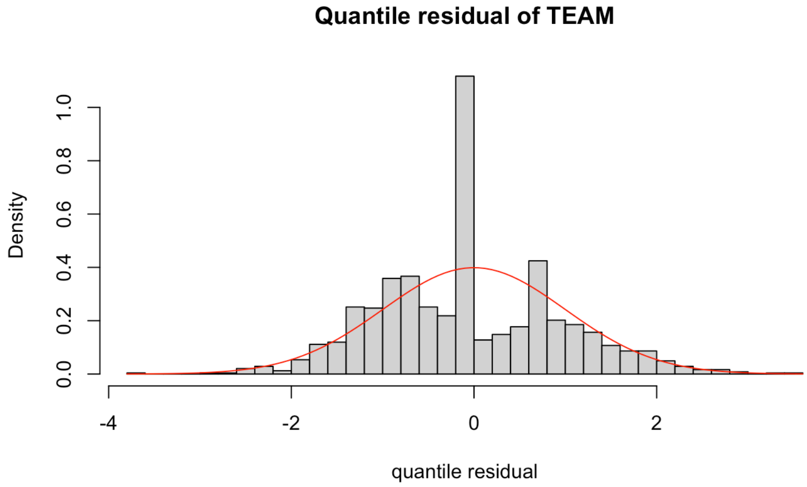


Figure 3.11: Histogram residual of MAR(4:1) for TEAM

Table 3.14: Normality test of MAR(4:1) for TEAM

Test	Statistic	p-value
Shapiro-Wilk	0.9882	0.0000
Kolmogorov-Smirnov	0.1158	0.0000

In the normality test of the quantile residuals for TEAM, it is evident that the histogram does not conform to a normal distribution curve. Both normality tests, the Shapiro-Wilk test and the Kolmogorov-Smirnov test, reveal p-values less than 0.05. Consequently, the distribution of the given data does not conform to a normal distribution.

Finally, we analyze KCE stock data in the electronic components sectors by fitting it with the MAR$(K;p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$. The criteria values for each model are presented in Table 3.15.

Table 3.15: Criteria for the candidate MAR model for KCE

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MAR(1:1)	3869.682	3875.443	3884.985	MAR(3:1)	3512.307	3533.432	3568.416
MAR(1:2)	3869.460	3877.140	3889.860	MAR(3:2)	3539.025	3565.908	3610.425
MAR(1:3)	3867.575	3877.175	3893.071	MAR(3:3)	3562.900	3595.540	3649.587
MAR(1:4)	3865.001	3876.520	3895.591	MAR(3:4)	3592.518	3630.912	3694.485
MAR(2:1)	3539.375	3552.818	3575.081	MAR(4:1)	3512.342	3541.149	3588.855
MAR(2:2)	3576.907	3594.188	3622.807	MAR(4:2)	3539.380	3575.863	3636.280
MAR(2:3)	3593.877	3614.996	3649.968	MAR(4:3)	3564.405	3608.564	3681.686
MAR(2:4)	3623.873	3648.829	3690.152	MAR(4:4)	3593.855	3645.687	3731.511

From Table 3.15 shows that multiple components have smaller AIC, HQIC, and BIC values than the single component. Among these, the MAR(3;1) model exhibits the smallest AIC, HQIC, and BIC criteria. Therefore, considering the three criteria, it indicates that the MAR(3;1) model is the best fit for KCE stock data.

Next, the diagnostic check involves quantile residuals, which are utilized for computationally simple tests aimed at detecting autocorrelation, quantile residual plot, Q-Q plot in Figure 3.12, the histogram of quantile residual and normality of the quantile residuals are shown in Figure 3.13 and Table 3.16, respectively.

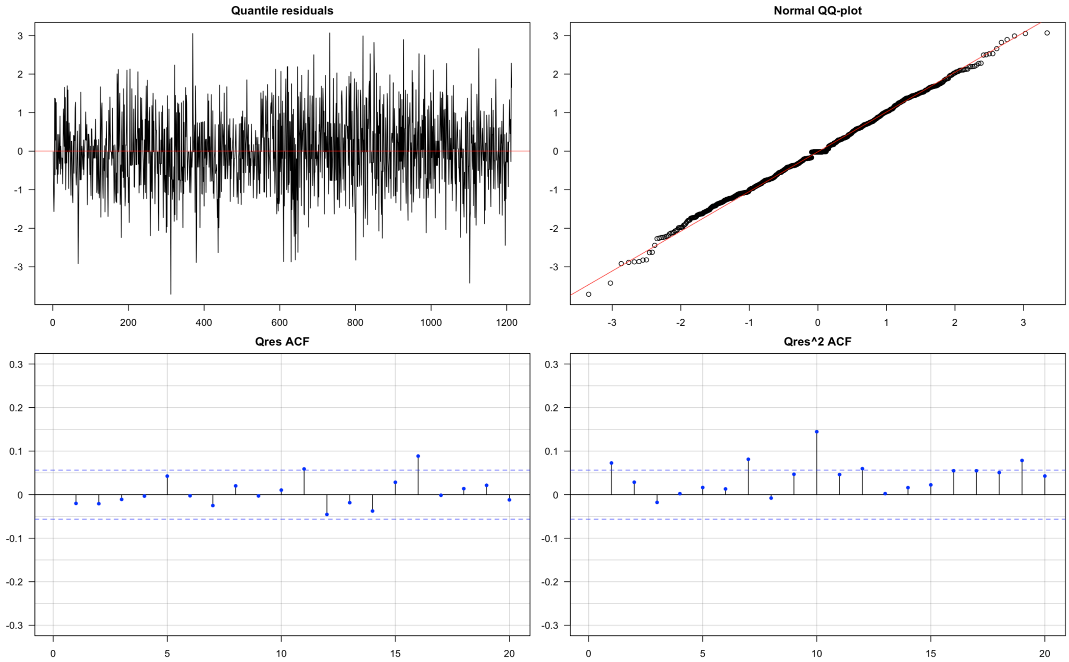


Figure 3.12: Quantile residual plot of MAR(3:1) for KCE

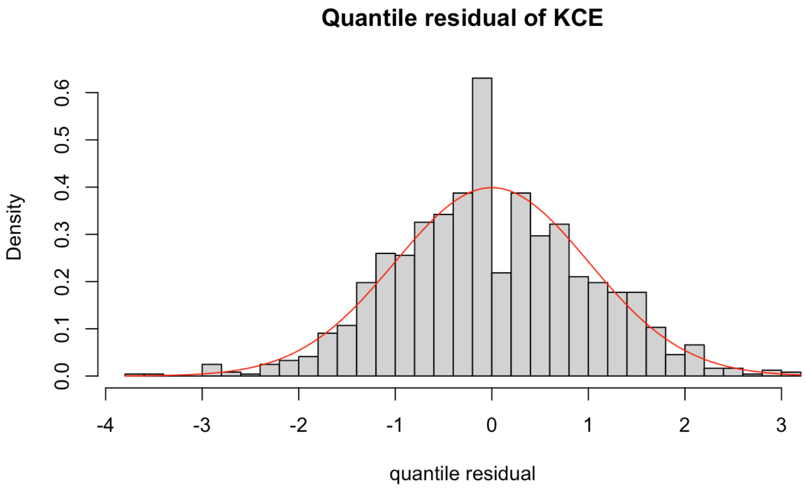


Figure 3.13: Histogram residual of MAR(3:1) for KCE

Table 3.16: Normality test of MAR(3:1) for KCE

Test	Statistic	p-value
Shapiro-Wilk	0.9976	0.0715
Kolmogorov-Smirnov	0.0501	0.0045

In the normality test of the quantile residuals for KCE, it is evident that the histogram does not conform to a normal distribution curve. In the normality tests, the p-values of the Shapiro-Wilk test are greater than 0.05, but the Kolmogorov-Smirnov test reveals p-values less than 0.05. Consequently, the distribution of the given data does not follow a normal distribution.

As a result of the mixture autoregressive (MAR) model for the 6 stock datasets, each table of criteria for the candidates shows that the best model for Thai stock data is the multiple component model, which exhibits the smallest criteria values. However, during the diagnostic modeling process, it became apparent that almost all the residuals and histogram of residuals do not adhere to a normal distribution. As an alternative, we considered the t distribution, which is known for its heavier tails compared to the normal distribution. Therefore, we have opted for the t mixture autoregressive (TMAR) model as an alternative.

3.2 t Mixture autoregressive model

The student t mixture autoregressive (TMAR) model is a collection of various autoregressive components that are extended from (3.1) by using the student t distribution. The heavy tails of component distributions can be adjusted, making this model more flexible than the mixture autoregressive model. Specifically, the time series $\{y_t\}_{t\geq 1}$ is said to be the K component TMAR model, denoted as TMAR($K; p_1, p_2, p_3, \ldots, p_K$), if it satisfies

$$F(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k F_{v_k}\Big(\frac{y_t - \phi_{k0} - \phi_{k1}y_{t-1} - \phi_{k2}y_{t-2} - \cdots - \phi_{kp_k}y_{t-p_k}}{\sigma_k}\Big), \tag{3.4}$$

where $F(y_t|\mathcal{F}_{t-1})$ is the conditional cumulative distribution function of y_t given the past information $y_{t-1}, y_{t-2}, y_{t-3}, \ldots,\ y_1$, $\mathcal{F}_t$ is the information set up to time t, $F_{v_k}(\cdot)$ is the cumulative distribution function of the standardized t distribution with v_k degrees of freedom for the k^{th} component, the mixing proportion $\alpha_k > 0$, $k = 1, 2, 3, \ldots, K$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_K = 1$ and assume that the error term of autoregressive e_t is follow t distribution. The probability distribution function of a standardized t distribution with the unit variance is

$$f_v(x) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi(v-2)}\Gamma(\frac{v}{2})}\Big(1 + \frac{x^2}{v-2}\Big)^{-\frac{v+1}{2}}, \tag{3.5}$$

where $\Gamma(\cdot)$ is the gamma function, and $2 < v < \infty$.

The conditional mean and conditional variance of the student t mixture autoregressive model similar to the mixture autoregressive model which is given as

$$E(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \mu_{kt},$$

where $\mu_{kt} = \phi_{k0} + \phi_{k1}y_{t-1} + \phi_{k2}y_{t-2} + \cdots - \phi_{kp_k}y_{t-p_k}$ and

$$\mathrm{Var}(y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \mu_{kt}^2 + \sum_{k=1}^{K} \alpha_k \sigma_k^2 - \left(\sum_{k=1}^{K} \alpha_k \mu_{kt}\right)^2,$$

respectively.

3.2.1 Parameter estimation

In this section, we discuss the method to estimate the parameters that we developed in this study, which is the EM algorithm, and compare it with the maximum likelihood function.

3.2.1.1 Parameter estimation by maximum likelihood function

In the case of the t mixture autoregressive model, the maximum likelihood method is the parameter estimation method used in this study to compare with the EM algorithm. Specifically, given a time series $\mathbf{y} = (y_1, y_2, \ldots, y_t)$, the likelihood function for the t mixture autoregressive model is the product of conditional density

$$L(\phi, \sigma, \alpha, v|\mathbf{y}) = \prod_{t=p+1}^{n} \sum_{k=1}^{K} \frac{\alpha_k}{\sigma_k} f_{v_k}\Big(\frac{y_t - \sum_{i=1}^{p_k} \phi_{ki} y_{t-i}}{\sigma_k}\Big), \tag{3.6}$$

the log likelihood of the t mixture autoregressive model can be written as

$$l(\phi, \sigma, \alpha, v) = \sum_{t=p+1}^{n} \log \Big[\sum_{k=1}^{K} \frac{\alpha_k}{\sigma_k} f_{v_k} \Big(\frac{y_t - \sum_{i=1}^{p_k} \phi_{ki} y_{t-i}}{\sigma_k} \Big) \Big], \tag{3.7}$$

where f_{v_k} is the probability distribution function of a standardized t distribution. Some parameters of the t mixture autoregressive model may not be solved in general [7]. Consequently, the estimation of these parameters must be performed using a numerical method [5].

3.2.1.2 Parameter estimation by the EM algorithm

In this section, parameter estimation is conducted using the EM algorithm, and the log-likelihood is constructed using the normal scale mixture model. Assume that we have observations $\mathbf{y} = (y_1, y_2, \ldots, y_t)$ generated from the TMAR model. Let $Z = (Z_1, Z_2, \ldots, Z_t)$ be a $K \times n$ unobservable random matrix, where $Z_t = (Z_{kt})$ for $t = 1, 2, 3, \ldots, n$, is a K-dimensional column indicator vector showing the origin of the k^{th} observation, that is, $Z_{kt} = 1$, if the observations y_t is generated from the k^{th} component of the TMAR model and $Z_{kt} = 0$ otherwise. Analogous to the formulation of Z, we consider another missing random matrix, $W = (W_1, W_2, \ldots, W_t)$, where $W_t = (W_{kt})$ for $t = 1, 2, 3, \ldots, n$ is also a K-dimensional vector. Given $Z_{kt} = 1$, the conditional distribution of W_{kt} is $W_{kt}|Z_{kt} = 1 \sim \text{gamma}(\frac{v_k}{2}, \frac{v_k - 2}{2})$, and $W_1, \ldots, W_n$ are distributed independently. The conditional loglikelihood function for t mixture autoregressive model is

$$l = l_1(\alpha) + l_2(v) + l_3(\theta), \tag{3.8}$$

where

$$
\begin{aligned}
l_1(\alpha) &= \sum_{k=1}^{K}\sum_{t=p+1}^{n} Z_{kt}\log(\alpha_k), && (3.9)\\
l_2(v) &= \sum_{k=1}^{K}\sum_{t=p+1}^{n} Z_{kt}\Big[-\log\Big\{\Gamma\Big(\frac{1}{2}v_k\Big)\Big\} + \Big(\frac{1}{2}v_k\Big)\log\Big(\frac{v_k-2}{2}\Big)\\
&\quad + \Big(\frac{v_k}{2}\Big)(\log W_{kt} - W_{kt}) + W_{kt} - \log(W_{kt})\Big], && (3.10)\\
l_3(\theta) &= \sum_{k=1}^{K}\sum_{t=p+1}^{n} Z_{kt}\Big(-\frac{1}{2}\{\log(2\pi) + \log\sigma_k^2 - \log W_{kt}\} - \frac{e_{kt}^2 W_{kt}}{2\sigma_k^2}\Big), && (3.11)
\end{aligned}
$$

and $e_{kt} = y_t - \phi_{k0} - \phi_{k1}y_{t-1} - \cdots - \phi_{kp_k}y_{t-p_k}$.

The parameters are estimated by iteratively maximum likelihood through the Expectation-Maximization(EM) procedure [9], which involves two main steps: the Expectation(E-step) and the Maximization(M-step). These steps are repeated iteratively until the algorithm converges. An illustration of the EM algorithm is presented in Figure 3.15.

The Expectation step. Assume that α, θ, and v are known. The unobserved random variable Z, the missing data W and $\log W$ in the loglikelihood are replaced by their expectations conditional on the parameters and the observed data $\mathbf{y}$. Let τ_{kt} be the conditional expectation of the k^{th} component of unobserved data Z. Let η_{kt} be the conditional expectation of the k^{th} component of missing data W. The Expectation step equations are

$$
\begin{aligned}
\tau_{kt} = E(Z_{kt}|y_t) &= \frac{\alpha_k\sigma_k^{-1}f_{vk}(\delta_{kt})}{\sum_{j=1}^{g}\alpha_j\sigma_j^{-1}f_{vj}(\delta_{jt})} \quad (k=1,\ldots,K), && (3.12)\\
\eta_{kt} = E(W_{kt}|y_t, z_{kt}=1) &= \frac{v_k+1}{\delta_{kt}^2 + v_k - 2} \quad (k=1,\ldots,K), && (3.13)\\
E(\log W_{kt}|y_t, z_{kt}=1) &= \log\eta_{kt} + \Big\{\psi\Big(\frac{v_k+1}{2}\Big) - \log\Big(\frac{v_k+1}{2}\Big)\Big\}, && (3.14)
\end{aligned}
$$

where $\delta_{kt}^2 = \frac{e_{kt}^2}{\sigma_k^2}$ and $\psi(s) = \frac{d\log\{\Gamma(s)\}}{ds}$ is the digamma function.

The Maximization step. Suppose that the unobserved random variable, Z, and the missing random variable, W, is known. By maximizing the log-likelihood function (3.8), the estimates of our model are obtained through the first derivatives with respect to the parameters α_k, v_k, ϕ_{ki}, and σ_k are

$$\frac{\partial l_1}{\partial \alpha_k} = \sum_{t=p+1}^{n} \Big(\frac{Z_{kt}}{\alpha_k} - \frac{Z_{gt}}{\alpha_K} \Big), \tag{3.15}$$

$$\frac{\partial l_2}{\partial v_k} = \sum_{t=p+1}^{n} Z_{kt} \Big[-\frac{1}{2}\psi\Big(\frac{1}{2}v_k\Big) + \frac{1}{2}\log\Big(\frac{v_k-2}{2}\Big) + \frac{1}{2}\log\Big(\frac{v_k}{v_k-2}\Big) + \frac{1}{2}\{\log(W_{kt}) - W_{kt}\} \Big], \tag{3.16}$$

$$\frac{\partial l_3}{\partial \phi_{ki}} = \sum_{t=p+1}^{n} \frac{Z_{kt} W_{kt} u(y_t, i) e_{kt}}{\sigma_k^2}, \tag{3.17}$$

$$\frac{\partial l_3}{\partial \sigma_k} = \sum_{t=p+1}^{n} \frac{Z_{kt}}{\sigma_k} \Big(\frac{W_{kt} e_{kt}^2}{\sigma_k^2} - 1 \Big), \tag{3.18}$$

where $u(y_t, i) = y_{t-i}$ for $i > 0$, and $u(y_t, i) = 1$ for $i = 0$.

Next, we substitute the conditional expectations of Z_{kt}, W_{kt}, and $\log(W_{kt})$ in (3.15) to (3.18) and these equations are set to zero to find the optimal values. The estimates of the mixing proportions α are

$$\hat{\alpha}_k = \frac{\sum_{t=p+1}^{n} \tau_{kt}}{n-p}. \tag{3.19}$$

The estimate of ϕ_{ki} are obtained by solving the system of equations

$$\sum_{t=p+1}^{n} \tau_{kt}\eta_{kt} y_t u(y_t, i) = \sum_{j=0}^{p_k} \phi_{kj} \sum_{t=p+1}^{n} \tau_{kt}\eta_{kt} y_t u(y_t, i) u(y_t, j), \tag{3.20}$$

where $u(y_t, i) = y_{t-i}$ for $i > 0$, and $u(y_t, i) = 1$ for $i = 0$. We can rewritten $\hat{e}_{kt} = y_t - \hat{\phi}_{k0} - \hat{\phi}_{k1} y_{t-1} \cdots - \hat{\phi}_{kp_k} y_{t-p_k}$, the estimate of σ is

$$\hat{\sigma}_k = \Big(\frac{\sum_{t=p+1}^{n} \tau_{kt}\eta_{kt}\hat{e}_{kt}^2}{\sum_{t=p+1}^{n} \tau_{kt}} \Big)^{\frac{1}{2}}. \tag{3.21}$$

The estimate of the degree of freedom must satisfy the equations

$$\begin{aligned}\left(\frac{v_k}{v_k-2}\right) - \psi\left(\frac{v_k}{2}\right) + \psi\left(\frac{v_k^{(m)}+1}{2}\right) + \log\left(\frac{v_k-2}{2}\right) - \log\left(\frac{v_k^{(m)}+1}{2}\right)\\ +\frac{1}{\sum_{t=p+1}^{n}\tau_{kt}^{(m)}}\sum_{t=p+1}^{n}\tau_{kt}^{(m)}\Big(\log(\eta_{kt}^{(m)})-\eta_{kt}^{(m)}\Big)=0,\end{aligned} \tag{3.22}$$

where, $v_k^{(m)}$ represents the estimated v_k in the m^th iteration of the EM algorithm. This estimation is employed to obtain a numerical solution using the Newton-Raphson method following

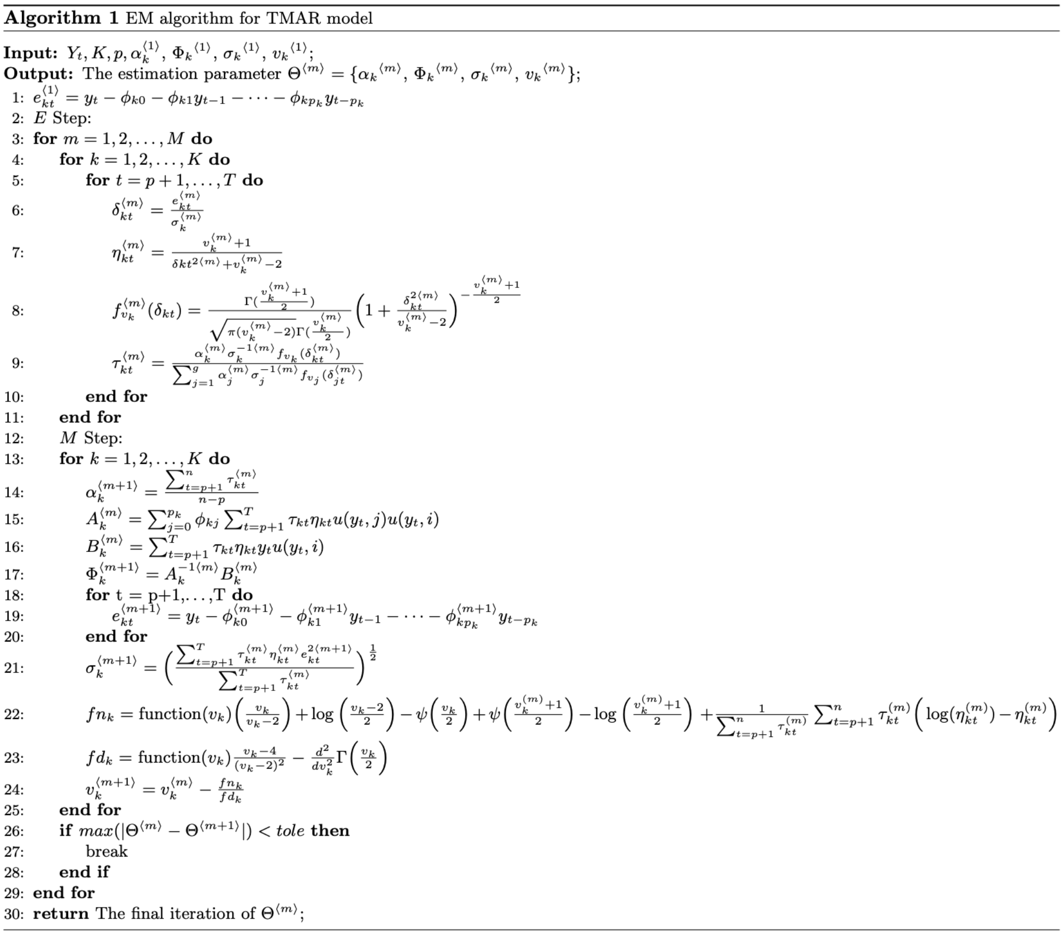
$$v_k^n = v_k^0 - \frac{f(v_k^0)}{f'(v_k^0)}, \tag{3.23}$$

where

$$f'(v_k^0) = \frac{v_k-4}{(v_k-2)^2 - \frac{1}{2}\psi'(\frac{v_k}{2})}. \tag{3.24}$$

In practice, it is feasible that the estimated values of v_k are fewer than two. To prevent this, we impose the condition $v_k > 2$ during EM estimation.

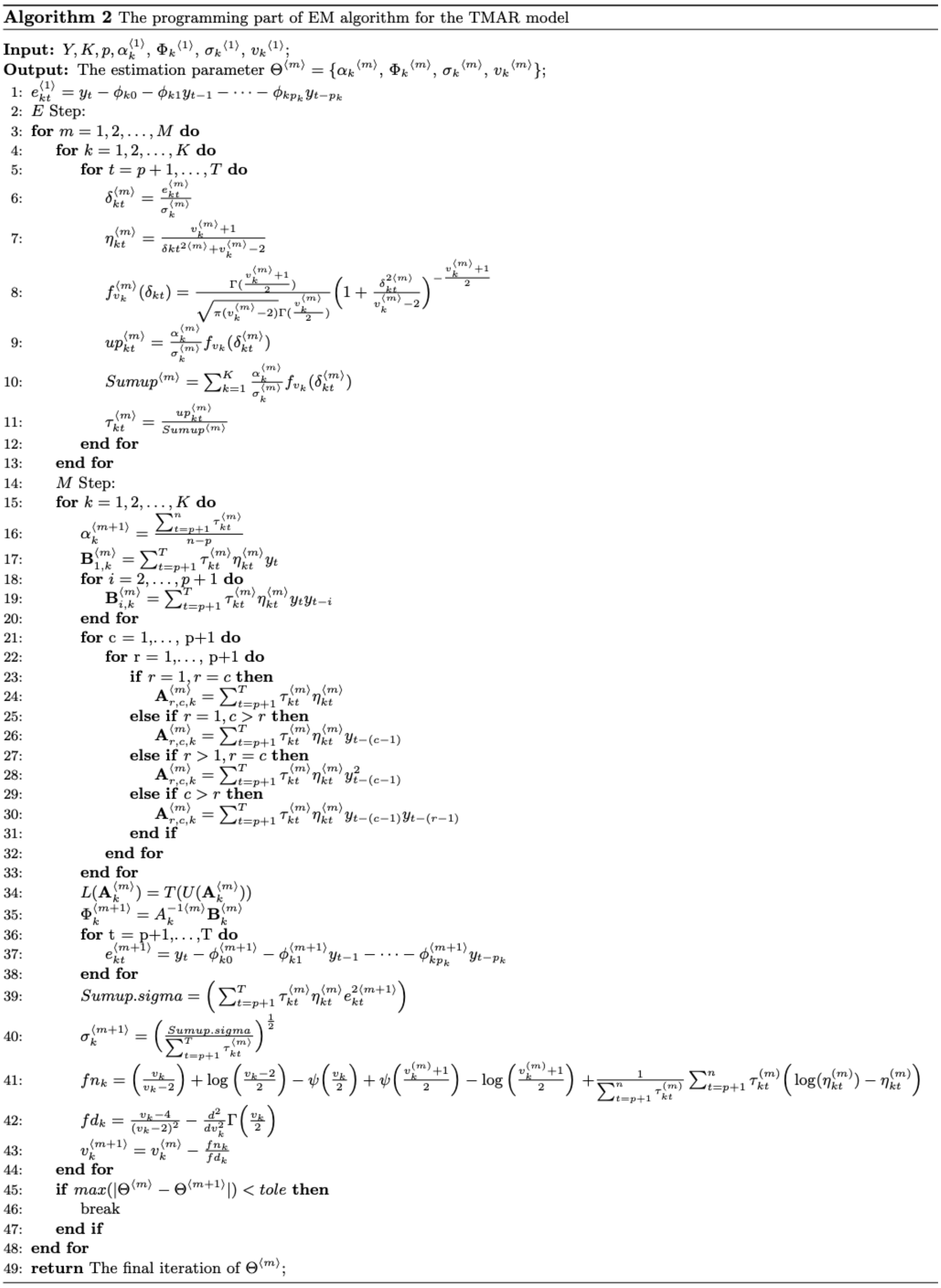
The EM algorithm for the TMAR model, which the K is a number of components, order of autoregressive p, weight of probability distribution α_k, degree of freedom v_k, autoregressive coefficient $\Phi = \phi_{k0}, \phi_{k1}, \ldots, \phi_{kp_k}$, standard deviation σ_k. The expectation step and maximization step are performed repeatedly to derive the parameters of the probability distribution, as detailed in Figure 3.15.



```
Algorithm 1 EM algorithm for TMAR model
Input: $Y_t, K, p, \alpha_k^{\langle 1 \rangle}, \Phi_k{}^{\langle 1 \rangle}, \sigma_k{}^{\langle 1 \rangle}, v_k{}^{\langle 1 \rangle}$;
Output: The estimation parameter $\Theta^{\langle m \rangle} = \{\alpha_k{}^{\langle m \rangle}, \Phi_k{}^{\langle m \rangle}, \sigma_k{}^{\langle m \rangle}, v_k{}^{\langle m \rangle}\}$;
$e_{kt}^{\langle 1 \rangle} = y_t - \phi_{k0} - \phi_{k1} y_{t-1} - \cdots - \phi_{kp_k} y_{t-p_k}$
$E$ Step:
for $m = 1, 2, \ldots, M$ do
    for $k = 1, 2, \ldots, K$ do
        for $t = p + 1, \ldots, T$ do
            $\delta_{kt}^{\langle m \rangle} = \frac{e_{kt}^{\langle m \rangle}}{\sigma_k^{\langle m \rangle}}$
            $\eta_{kt}^{\langle m \rangle} = \frac{v_k^{\langle m \rangle}+1}{\delta kt^{2\langle m \rangle}+v_k^{\langle m \rangle}-2}$
            $f_{v_k}^{\langle m \rangle}(\delta_{kt}) = \frac{\Gamma(\frac{v_k^{\langle m \rangle}+1}{2})}{\sqrt{\pi(v_k^{\langle m \rangle}-2)}\Gamma(\frac{v_k^{\langle m \rangle}}{2})} \left(1 + \frac{\delta_{kt}^{2\langle m \rangle}}{v_k^{\langle m \rangle}-2}\right)^{-\frac{v_k^{\langle m \rangle}+1}{2}}$
            $\tau_{kt}^{\langle m \rangle} = \frac{\alpha_k^{\langle m \rangle} \sigma_k^{-1\langle m \rangle} f_{v_k}(\delta_{kt}^{\langle m \rangle})}{\sum_{j=1}^{g} \alpha_j^{\langle m \rangle} \sigma_j^{-1\langle m \rangle} f_{v_j}(\delta_{jt}^{\langle m \rangle})}$
        end for
    end for
    $M$ Step:
    for $k = 1, 2, \ldots, K$ do
        $\alpha_k^{\langle m+1 \rangle} = \frac{\sum_{t=p+1}^{n} \tau_{kt}^{\langle m \rangle}}{n-p}$
        $A_k^{\langle m \rangle} = \sum_{j=0}^{p_k} \phi_{kj} \sum_{t=p+1}^{T} \tau_{kt}\eta_{kt} u(y_t, j) u(y_t, i)$
        $B_k^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}\eta_{kt} y_t u(y_t, i)$
        $\Phi_k^{\langle m+1 \rangle} = A_k^{-1\langle m \rangle} B_k^{\langle m \rangle}$
        for t = p+1,...,T do
            $e_{kt}^{\langle m+1 \rangle} = y_t - \phi_{k0}^{\langle m+1 \rangle} - \phi_{k1}^{\langle m+1 \rangle} y_{t-1} - \cdots - \phi_{kp_k}^{\langle m+1 \rangle} y_{t-p_k}$
        end for
        $\sigma_k^{\langle m+1 \rangle} = \left(\frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} e_{kt}^{2\langle m+1 \rangle}}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}\right)^{\frac{1}{2}}$
        $fn_k = \text{function}(v_k)\left(\frac{v_k}{v_k-2}\right) + \log\left(\frac{v_k-2}{2}\right) - \psi\left(\frac{v_k}{2}\right) + \psi\left(\frac{v_k^{(m)}+1}{2}\right) - \log\left(\frac{v_k^{(m)}+1}{2}\right) + \frac{1}{\sum_{t=p+1}^{n} \tau_{kt}^{(m)}} \sum_{t=p+1}^{n} \tau_{kt}^{(m)} \left(\log(\eta_{kt}^{(m)}) - \eta_{kt}^{(m)}\right)$
        $fd_k = \text{function}(v_k)\frac{v_k-4}{(v_k-2)^2} - \frac{d^2}{dv_k^2}\Gamma\left(\frac{v_k}{2}\right)$
        $v_k^{\langle m+1 \rangle} = v_k^{\langle m \rangle} - \frac{fn_k}{fd_k}$
    end for
    if $max(|\Theta^{\langle m \rangle} - \Theta^{\langle m+1 \rangle}|) < tole$ then
        break
    end if
end for
return The final iteration of $\Theta^{\langle m \rangle}$;
```

Figure 3.14: The EM algorithm for the TMAR model

The EM algorithm for the mixture autoregressive model based on the t distribution that we mention and develop in Section 3.2.1.2 has the following steps in the programme:



Algorithm 2 The programming part of EM algorithm for the TMAR model

```
Input: $Y, K, p, \alpha_k^{\langle 1 \rangle}, {\Phi_k}^{\langle 1 \rangle}, {\sigma_k}^{\langle 1 \rangle}, {v_k}^{\langle 1 \rangle}$;
Output: The estimation parameter $\Theta^{\langle m \rangle} = \{{\alpha_k}^{\langle m \rangle}, {\Phi_k}^{\langle m \rangle}, {\sigma_k}^{\langle m \rangle}, {v_k}^{\langle m \rangle}\}$;
$e_{kt}^{\langle 1 \rangle} = y_t - \phi_{k0} - \phi_{k1} y_{t-1} - \cdots - \phi_{kp_k} y_{t-p_k}$
$E$ Step:
for $m = 1, 2, \ldots, M$ do
    for $k = 1, 2, \ldots, K$ do
        for $t = p + 1, \ldots, T$ do
            $\delta_{kt}^{\langle m \rangle} = \frac{e_{kt}^{\langle m \rangle}}{\sigma_k^{\langle m \rangle}}$
            $\eta_{kt}^{\langle m \rangle} = \frac{v_k^{\langle m \rangle}+1}{\delta kt^{2\langle m \rangle} + v_k^{\langle m \rangle} - 2}$
            $f_{v_k}^{\langle m \rangle}(\delta_{kt}) = \frac{\Gamma(\frac{v_k^{\langle m \rangle}+1}{2})}{\sqrt{\pi(v_k^{\langle m \rangle}-2)}\Gamma(\frac{v_k^{\langle m \rangle}}{2})} \left(1 + \frac{\delta_{kt}^{2\langle m \rangle}}{v_k^{\langle m \rangle}-2}\right)^{-\frac{v_k^{\langle m \rangle}+1}{2}}$
            $up_{kt}^{\langle m \rangle} = \frac{\alpha_k^{\langle m \rangle}}{\sigma_k^{\langle m \rangle}} f_{v_k}(\delta_{kt}^{\langle m \rangle})$
            $Sumup^{\langle m \rangle} = \sum_{k=1}^{K} \frac{\alpha_k^{\langle m \rangle}}{\sigma_k^{\langle m \rangle}} f_{v_k}(\delta_{kt}^{\langle m \rangle})$
            $\tau_{kt}^{\langle m \rangle} = \frac{up_{kt}^{\langle m \rangle}}{Sumup^{\langle m \rangle}}$
        end for
    end for
    $M$ Step:
    for $k = 1, 2, \ldots, K$ do
        $\alpha_k^{\langle m+1 \rangle} = \frac{\sum_{t=p+1}^{n} \tau_{kt}^{\langle m \rangle}}{n-p}$
        $\mathbf{B}_{1,k}^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} y_t$
        for $i = 2, \ldots, p + 1$ do
            $\mathbf{B}_{i,k}^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} y_t y_{t-i}$
        end for
        for c = 1,. . . , p+1 do
            for r = 1,. . . , p+1 do
                if $r = 1, r = c$ then
                    $\mathbf{A}_{r,c,k}^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle}$
                else if $r = 1, c > r$ then
                    $\mathbf{A}_{r,c,k}^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} y_{t-(c-1)}$
                else if $r > 1, r = c$ then
                    $\mathbf{A}_{r,c,k}^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} y_{t-(c-1)}^2$
                else if $c > r$ then
                    $\mathbf{A}_{r,c,k}^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} y_{t-(c-1)} y_{t-(r-1)}$
                end if
            end for
        end for
        $L(\mathbf{A}_k^{\langle m \rangle}) = T(U(\mathbf{A}_k^{\langle m \rangle}))$
        $\Phi_k^{\langle m+1 \rangle} = A_k^{-1\langle m \rangle} \mathbf{B}_k^{\langle m \rangle}$
        for t = p+1,. . . ,T do
            $e_{kt}^{\langle m+1 \rangle} = y_t - \phi_{k0}^{\langle m+1 \rangle} - \phi_{k1}^{\langle m+1 \rangle} y_{t-1} - \cdots - \phi_{kp_k}^{\langle m+1 \rangle} y_{t-p_k}$
        end for
        $Sumup.sigma = \left(\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} e_{kt}^{2\langle m+1 \rangle}\right)$
        $\sigma_k^{\langle m+1 \rangle} = \left(\frac{Sumup.sigma}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}\right)^{\frac{1}{2}}$
        $fn_k = \left(\frac{v_k}{v_k-2}\right) + \log\left(\frac{v_k-2}{2}\right) - \psi\left(\frac{v_k}{2}\right) + \psi\left(\frac{v_k^{\langle m \rangle}+1}{2}\right) - \log\left(\frac{v_k^{\langle m \rangle}+1}{2}\right) + \frac{1}{\sum_{t=p+1}^{n} \tau_{kt}^{\langle m \rangle}} \sum_{t=p+1}^{n} \tau_{kt}^{\langle m \rangle} \left(\log(\eta_{kt}^{\langle m \rangle}) - \eta_{kt}^{\langle m \rangle}\right)$
        $fd_k = \frac{v_k-4}{(v_k-2)^2} - \frac{d^2}{dv_k^2}\Gamma\left(\frac{v_k}{2}\right)$
        $v_k^{\langle m+1 \rangle} = v_k^{\langle m \rangle} - \frac{fn_k}{fd_k}$
    end for
    if $max(|\Theta^{\langle m \rangle} - \Theta^{\langle m+1 \rangle}|) < tole$ then
        break
    end if
end for
return The final iteration of $\Theta^{\langle m \rangle}$;
```

Figure 3.15: The programming part of the EM algorithm for the TMAR model

Next, the calculation example of the TMAR(2;2) model from the programme TMAR_EM in the part of simulation which parameter $\alpha_1, \phi_{10}, \phi_{11}, \phi_{12}, \sigma_1, v_1, \alpha_2$, ϕ_{20}, ϕ_{21}, ϕ_{22}, σ_2 are 0.6, 0, 0.33, -0.36, 1.2, 8.68, 0.40, 0.00, -0.21, -0.1, 1.50, 6.37, respectively. The data length that we generate is 10 data points, and M is the iteration of the EM algorithm.

In the first step, we generate time series data from 10 data points.

	y
1	0.443
2	-0.410
3	1.116
4	-0.091
5	-1.150
6	0.293
7	-0.382
8	-2.521
9	-0.898
10	0.779

and the initial value for the EM algorithm is

$$\alpha = [\alpha_1, \alpha_2] = [0.791, 0.209]$$

$$\Phi = \begin{bmatrix} \phi_{10} & \phi_{20} \\ \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{bmatrix} = \begin{bmatrix} 0.443 & -2.261 \\ -0.187 & -0.405 \\ -0.478 & -0.996 \end{bmatrix}$$

$$\sigma = [\sigma_1, \sigma_2] = [0.856, 0.013]$$

$$v = [v_1, v_2] = [3.157, 3.634]$$

In the initial step of the computation in the Expectation step, we calculate $e_{kt}^{\langle 1\rangle}$.

$$
\begin{aligned}
e_{kt}^{\langle 1\rangle} &= y_t - \phi_{k0} - \phi_{k1}y_{t-1} - \phi_{k2}y_{t-2} \\
&= \begin{bmatrix}
0.000 & 0.000 \\
0.000 & 0.000 \\
0.809 & 3.653 \\
-0.521 & 2.214 \\
-1.076 & 2.186 \\
-0.408 & 1.998 \\
-1.320 & 0.852 \\
-2.895 & -0.123 \\
-1.994 & -0.038 \\
-1.037 & 0.165
\end{bmatrix}
\end{aligned}
$$

In the Expectation step ($m = 1$), we need to compute δ_{kt}, η_{kt}, and τ_{kt}. Begin with $K = 1, 2$ and $t = p + 1, \ldots, T$. In this example, we illustrate the case of $t = p + 1 = 3$ to T and for all k.

$$
\delta_{kt}^{\langle m\rangle} = \frac{e_{kt}^{\langle m\rangle}}{\sigma_k^{\langle m\rangle}} = \begin{bmatrix}
NA & NA \\
NA & NA \\
0.945 & 271.781 \\
-0.608 & 164.760 \\
-1.256 & 162.684 \\
-0.477 & 148.636 \\
-1.542 & 63.413 \\
-3.381 & -9.159 \\
-2.329 & -2.855 \\
-1.212 & 12.300
\end{bmatrix},
$$

$$\eta_{kt}^{\langle m\rangle} = \frac{v_k^{\langle m\rangle}+1}{\delta kt^{2\langle m\rangle} + v_k^{\langle m\rangle} - 2} = \begin{bmatrix} NA & NA \\ NA & NA \\ 0.884 & 0.068 \\ 0.491 & 0.143 \\ 0.052 & 0.795 \\ 1.296 & 0.351 \\ 0.078 & 0.785 \\ 0.550 & 0.354 \\ 0.798 & 0.002 \\ 1.007 & 0.000 \end{bmatrix},$$

$$\mathrm{up}_{kt}^{\langle m\rangle} = \frac{\alpha_k^{\langle m\rangle}}{\sigma_k^{\langle m\rangle}} f_{v_k}(\delta_{kt}^{\langle m\rangle}) = \begin{bmatrix} NA & NA \\ NA & NA \\ 0.211 & 0.000 \\ 0.005 & 0.000 \\ 0.000 & 0.000 \\ 2.380 & 0.000 \\ 0.000 & 0.000 \\ 0.010 & 0.000 \\ 0.110 & 0.000 \\ 0.480 & 0.000 \end{bmatrix},$$

where $f_v(x) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi(v-2)}\Gamma(\frac{v}{2})}\left(1+\frac{x^2}{v-2}\right)^{-\frac{v+1}{2}}$,

$$\text{Sumup}^{\langle m\rangle} = \sum_{k=1}^{K} \text{up}_{kt}^{\langle m\rangle} = \begin{bmatrix} 0.00 \\ 0.00 \\ 0.21 \\ 0.01 \\ 0.00 \\ 2.38 \\ 0.00 \\ 0.01 \\ 0.11 \\ 0.48 \end{bmatrix},$$

$$\tau_{kt}^{\langle m\rangle} = \frac{\text{up}_{kt}^{\langle m\rangle}}{\text{Sumup}^{\langle m\rangle}} = \begin{bmatrix} NA & NA \\ NA & NA \\ 1.00 & 0.00 \\ 1.00 & 0.00 \\ 1.00 & 0.00 \\ 1.00 & 0.00 \\ 1.00 & 0.00 \\ 0.81 & 0.19 \\ 0.10 & 0.90 \\ 1.00 & 0.00 \end{bmatrix}.$$

In the Maximization Step: we need to estimate $\alpha_k^{\langle m+1\rangle}, \Phi_k^{\langle m+1\rangle}, \sigma_k^{\langle m+1\rangle}$, and $v_k^{\langle m+1\rangle}$. In this example, we illustrate the case of $t = p + 1 = 3$ to T and for all k,

$$\begin{aligned}\alpha_k^{\langle m+1\rangle} &= \frac{\sum_{t=p+1}^{n} \tau_{kt}^{\langle m\rangle}}{n-p} \\ &= [\alpha_1, \alpha_2] \\ &= [0.916, 0.084].\end{aligned}$$

When estimating $\Phi_k^{\langle m+1\rangle}$, we need to construct the matrix equation $A\Phi_k = B$, where

$$\begin{aligned}A_k^{\langle m\rangle} &= \sum_{j=0}^{p_k} \phi_{kj} \sum_{t=p+1}^{T} \tau_{kt}\eta_{kt} u(y_t, j) u(y_t, i) \\ &= \begin{bmatrix} 2.732 & 1.694 & -0.139 \\ 1.694 & 2.867 & -2.937 \\ -0.139 & -2.937 & 4.690 \end{bmatrix},\end{aligned}$$

$$\begin{aligned}B_k^{\langle m\rangle} &= \sum_{t=p+1}^{T} \tau_{kt}\eta_{kt} y_t u(y_t, i) \\ &= \begin{bmatrix} 1.564 \\ 0.038 \\ 1.312 \end{bmatrix},\end{aligned}$$

$$\begin{aligned}\Phi_k^{\langle m+1\rangle} &= A_k^{-1\langle m\rangle} B_k^{\langle m\rangle} \\ &= \begin{bmatrix} 0.226 & -2.358 \\ -0.175 & -0.413 \\ -0.334 & -1.097 \end{bmatrix}.\end{aligned}$$

Before estimating $\sigma_k^{\langle m+1\rangle}$, we update the term $\phi_{kp}^{\langle m+1\rangle}$ to $e_{kt}^{\langle m+1\rangle}$.

$$
\begin{aligned}
e_{kt}^{\langle m+1\rangle} &= y_t - \phi_{k0}^{\langle m+1\rangle} - \phi_{k1}^{\langle m+1\rangle} y_{t-1} - \phi_{k2}^{\langle m+1\rangle} y_{t-2} \\
&= \begin{bmatrix} 0.000 & 0.000 \\ 0.000 & 0.000 \\ 0.966 & 3.791 \\ -0.259 & 2.278 \\ -1.020 & 2.395 \\ -0.165 & 2.076 \\ -0.941 & 0.835 \\ -2.717 & -0.000 \\ -1.692 & 0.000 \\ -0.446 & -0.000 \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\text{Sumup.sigma} &= \Big(\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle} \eta_{kt}^{\langle m\rangle} e_{kt}^{2\langle m+1\rangle} \Big) \\
&= [6.0256, 0.0000],
\end{aligned}
$$

$$
\begin{aligned}
\sigma_k^{\langle m+1\rangle} &= \Big(\frac{\text{Sumup.sigma}}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}} \Big)^{\frac{1}{2}} \\
&= [1.001, 0.000].
\end{aligned}
$$

The estimate of degree of freedom must satisfy the equations fn_k, which is defined as

$$
\begin{aligned}
fn_k &= \Big(\frac{v_k}{v_k-2}\Big) - \psi\Big(\frac{v_k}{2}\Big) + \psi\Big(\frac{v_k^{(m)}+1}{2}\Big) + \log\Big(\frac{v_k-2}{2}\Big) - \log\Big(\frac{v_k^{(m)}+1}{2}\Big) \\
&\quad + \frac{1}{\sum_{t=p+1}^{n}\tau_{kt}^{(m)}} \sum_{t=p+1}^{n} \tau_{kt}^{(m)} \Big(\log(\eta_{kt}^{(m)}) - \eta_{kt}^{(m)}\Big), \\
fd_k &= \frac{v_k-4}{(v_k-2)^2} - \frac{d^2}{dv_k^2}\Gamma\Big(\frac{v_k}{2}\Big),
\end{aligned}
$$

where, $v_k^{(m)}$ represents the estimated v_k in the m^{th} iteration of the EM algorithm. This estimation is employed to obtain a numerical solution using the Newton-Raphson method following

$$
\begin{aligned}
v_k^{\langle m+1\rangle} &= v_k^{\langle m\rangle} - \frac{fn_k}{fd_k} \\
&= [3.288, 3.574].
\end{aligned}
$$

For now, we complete the m^{th} iteration, where $m = 1$. We then repeat the Expectation step and Maximization step until the parameter estimate $\Theta^{\langle m\rangle}$ convergence by using $max(|\Theta^{\langle m\rangle} - \Theta^{\langle m+1\rangle}|) < tol$, where tol is the tolerance with a default value of 1×10^{-6} in this study.

3.2.2 Simulation study for the TMAR model

In this section, we examine the performance of parameter estimation using two different methods. First, we employ the EM algorithm in Section 3.2.1.2 that we develop and compare it with the maximum likelihood estimation procedure implemented in the package “uGMAR”[5] in Section 3.2.1.1. Furthermore, we examine the accuracy of parameter estimates. It’s important to note that, under the restrictions of the package, the orders of autoregressive components for different components are assumed to be the same. Therefore, the models considered in this study are denoted as TMAR($K; p$), where K is the number of components and p is the common order of autoregressive components. Two models investigated in this study are the TMAR($2; 2$), where K component is 2 and order p is 2, and the TMAR($3; 3$) model, where K component is 3 and order p is 3. Comparing the parameter estimate between the EM algorithm and the MLE. In the part of the EM algorithm using the parameter from the MLE to be an initial value and random uniform to an initial for the degree of freedom v_k. The best candidate models with the smallest corresponding criterion are the TMAR (2;2) and the TMAR(3;3) models.

In the first experiment, we generated a time length of 1000 data points from the TMAR$(2;2)$ model, where the coefficients $\alpha_1, \phi_{10}, \phi_{11}, \phi_{12}, \sigma_1, v_1, \alpha_2$, ϕ_{20}, ϕ_{21}, ϕ_{22}, σ_2 are 0.6, 0, 0.33, -0.36, 1.2, 8.68, 0.40, 0.00, -0.21, -0.1, 1.50, 6.37, respectively. The data are fitted to the mixture autoregressive model for $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$ to assess the accuracy of the model selection. The corresponding AICs, HQICs, and BICs are obtained, and the model with the smallest criterion is then selected to match the generated model. Table 3.17 presents the parameter estimation for the TMAR(2;2) models, comparing the exact values of parameters to the mean of estimates for each method, along with the error of each method.

Table 3.17: Parameter estimates using EM algorithm of TMAR(2;2) models

	α_1	ϕ_{10}	ϕ_{11}	ϕ_{12}	σ_1	v_1
Exact value	0.60	0.00	0.33	-0.36	1.20	8.68
Estimate of EM	0.76	-0.007	0.10	-0.23	1.17	9.43
Estimate of MLE	0.79	-0.013	0.08	-0.26	1.32	19076.05
Error of EM	0.16	0.007	0.23	0.13	0.03	0.75
Error of MLE	0.19	0.013	0.25	0.10	0.12	19069.68
	α_2	ϕ_{20}	ϕ_{21}	ϕ_{22}	σ_2	v_2
Exact value	0.40	0.00	-0.21	-0.10	1.50	6.37
Estimate of EM	0.24	0.073	-0.03	-0.33	0.61	8.806
Estimate of MLE	0.21	0.107	-0.07	-0.43	1.27	19076.05
Error of EM	0.16	0.073	0.18	0.23	0.89	2.437
Error of MLE	0.19	0.107	0.14	0.33	0.23	19069.68

From Table 3.17, the results of the parameter estimates for TMAR(2:2) models are presented. In the part α, the parameter estimates are quite close to the exact values.

The performance of the EM algorithm is notably good in 5 out of the 8 parameters based on the autoregressive coefficients ϕ_{kp} and standard deviation σ_k. In addition, we also calculated the mean square error(MSE) of the parameter estimates obtained from the EM algorithm and the maximum likelihood estimation (MLE), resulting in values of 1.242664 and 1.281587, respectively. Therefore, based on the parameter errors and the MSE values, it is evident that the EM algorithm outperforms MLE.

In the second experiment, we generated a time length of 1000 data points from the TMAR$(3;3)$ model with parameters $\alpha_1, \phi_{10}, \phi_{11}, \phi_{12}, \phi_{13}, \sigma_1, v_1, \alpha_2, \phi_{20}, \phi_{21}, \phi_{22}$, ϕ_{23}, σ_2, $v_2, \alpha_3, \phi_{30}, \phi_{31}, \phi_{32}, \phi_{33}, \sigma_3, v_3$ are 0.30, 0.00, 0.50, 0.24, 0.00, 2.00, 4.00, 0.30, 0.0, -0.90, 0.0, 0.0, 1.000, 6.000, 0.400, 0.0, 1.5, -0.740, 0.120, 0.500, 10.000, respectively. The data are fitted to the mixture autoregressive model for $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$ to assess the accuracy of the model selection. The corresponding AICs, HQICs, and BICs are obtained, and the model with the smallest criterion is then selected to match the generated model. Table 3.18 presents the parameter estimation for the TMAR(3;3) models, comparing the exact values of parameters to the mean of estimates for each method, along with the error of each method.

Table 3.18: Parameter estimates using EM algorithm of TMAR(3;3) models

	α_1	ϕ_{10}	ϕ_{11}	ϕ_{12}	ϕ_{13}	σ_1	v_1
Exact value	0.30	0.0	0.50	0.24	0.0	2.0	4.00
Estimate of EM	0.455	0.008	0.642	0.161	-0.138	0.925	8.919
Estimate of MLE	0.535	0.071	0.690	-0.198	-0.013	1.844	8710.071
Error of EM	0.155	0.008	0.142	0.079	0.138	1.075	4.919
Error of MLE	0.235	0.071	0.190	0.438	0.013	0.156	8706.071
	α_2	ϕ_{20}	ϕ_{21}	ϕ_{22}	ϕ_{23}	σ_2	v_2
Exact value	0.30	0.0	-0.90	0.0	0.0	1.000	6.000
Estimate of EM	0.327	0.006	0.164	0.224	-0.010	0.914	8.787
Estimate of MLE	0.316	0.018	0.061	-0.145	-0.003	2.227	15686.806
Error of EM	0.027	0.006	1.064	0.224	0.110	0.086	2.787
Error of MLE	0.016	0.018	0.961	0.145	0.003	1.227	15680.806
	α_3	ϕ_{30}	ϕ_{31}	ϕ_{32}	ϕ_{33}	σ_3	v_3
Exact value	0.400	0.0	1.5	-0.740	0.120	0.500	10.000
Estimate of EM	0.217	0.064	0.076	0.173	0.001	0.544	8.166
Estimate of MLE	0.148	-0.058	-0.023	-0.161	-0.009	2.028	23903.901
Error of EM	0.183	0.064	1.424	0.913	0.119	0.044	1.834
Error of MLE	0.252	0.058	1.523	0.579	0.129	1.528	23893.901

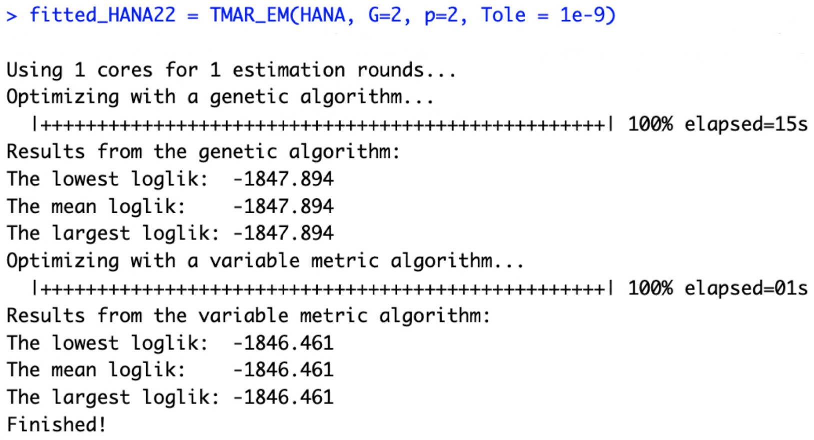
Table 3.18 displays the parameter estimates for the TMAR(3:3) models. In the part α_k, the error from the EM algorithm is smaller than the error from the maximum likelihood estimation (MLE) in 2 out of 3 parameters. Additionally, the performance of the EM algorithm exceeds that of the MLE in 6 out of 12 parameters for the autoregressive coefficient ϕ_{kp}. In terms of the standard deviation σ_k, the EM algorithm outperforms MLE in 2 out of 3 parameters. Furthermore, the parameter estimates from the EM algorithm exhibit better performance in terms of degree of freedom (v_k). In addition, we also calculated the mean square error (MSE) of the parameter estimates obtained from

the EM algorithm and Maximum Likelihood Estimation (MLE), resulting in values of 0.924623 and 1.08726, respectively. Therefore, based on the parameter errors and the MSE values, it is evident that the EM algorithm outperforms MLE.

3.2.3 *t* mixture autoregressive model for Thai stock markets

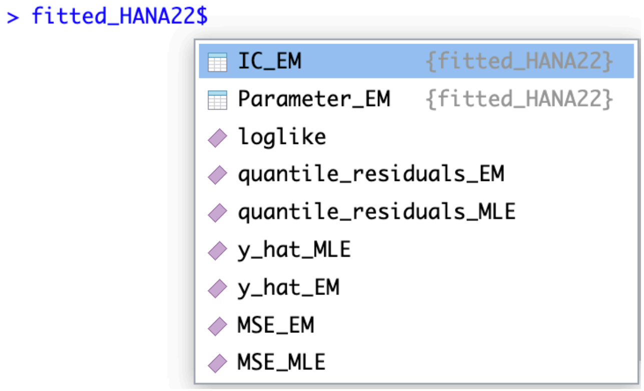
In part of the programme following the algorithm, we mention in Figure 3.15 and developing in function: TMAR_EM (data, K, p, tol) which inputs the data, number of components K, autoregressive order p, and tol is the tolerance, with default being 1×10^{-6}. The initial value of the parameter for this function is obtained by maximum likelihood estimation.

The following code fits a TMAR model with autoregressive order $p = 2$ and $K = 2$ mixture components, sets the tolerance with 1×10^{-9} to the HANA stock in the electronic components sector, and returns the list of elements, such as the information criteria IC, the estimation of parameters, the loglikelihood, the quantile residual of MLE and EM, the prediction of MLE and EM, and the mean square error of MLE and EM in Figure 3.17. For example, Figure 3.18 returns the parameter estimate from the function TMAR_EM.



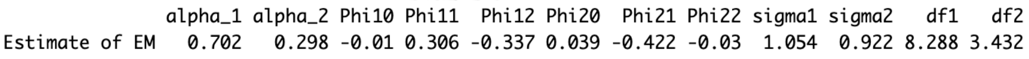
```
> fitted_HANA22 = TMAR_EM(HANA, G=2, p=2, Tole = 1e-9)

Using 1 cores for 1 estimation rounds...
Optimizing with a genetic algorithm...
  |++++++++++++++++++++++++++++++++++++++++++++++++++| 100% elapsed=15s
Results from the genetic algorithm:
The lowest loglik:  -1847.894
The mean loglik:    -1847.894
The largest loglik: -1847.894
Optimizing with a variable metric algorithm...
  |++++++++++++++++++++++++++++++++++++++++++++++++++| 100% elapsed=01s
Results from the variable metric algorithm:
The lowest loglik:  -1846.461
The mean loglik:    -1846.461
The largest loglik: -1846.461
Finished!
```

Figure 3.16: The EM function for HANA



```
> fitted_HANA22$
IC_EM                       {fitted_HANA22}
Parameter_EM                {fitted_HANA22}
loglike
quantile_residuals_EM
quantile_residuals_MLE
y_hat_MLE
y_hat_EM
MSE_EM
MSE_MLE
```

Figure 3.17: The list of elements in the EM function



```
                alpha_1 alpha_2 Phi10 Phi11  Phi12 Phi20  Phi21 Phi22 sigma1 sigma2   df1   df2
Estimate of EM    0.702   0.298 -0.01 0.306 -0.337 0.039 -0.422 -0.03  1.054  0.922 8.288 3.432
```

Figure 3.18: The parameter estimate from EM function

In this section, we investigate the performance of the mixture autoregressive model on individual stock markets using the daily closing prices of the top stock from the energy and utility, and electronic components sectors over the five-year period from August 1st, 2017, to August 1st, 2022 (1214 observations). In particular, BANPU, ESSO, and BCP from the energy and utility sector, HANA, TEAM, and KCE from the electronic components sector. In the table below, we show the AIC, HQIC, and BIC of each candidate model of the TMAR model using the EM algorithm to estimate parameters for each stock, and we will show the analysis of the best model.

To begin the analysis, we fit the BANPU stock from the energy and utility sectors with the TMAR$(K;p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$. The criteria values for each model are provided in Table 3.19.

Table 3.19: Criteria for the candidate TMAR model with using EM for BANPU

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMAR(1:1)	1308.7010	1320.22362	1339.30615	TMAR(3:1)	-427.7792	-397.05236	-346.16560
TMAR(1:2)	1188.0849	1201.52786	1223.79082	TMAR(3:2)	218.5742	255.06239	315.49041
TMAR(1:3)	1052.9590	1068.32246	1093.76584	TMAR(3:3)	1038.2250	1080.47444	1150.44373
TMAR(1:4)	907.9109	925.19473	953.81853	TMAR(3:4)	-603.1273	-555.11654	-475.60598
TMAR(2:1)	446.1429	467.26758	502.25222	TMAR(4:1)	-173.8239	-133.49488	-66.70601
TMAR(2:2)	2073.0176	2097.98320	2139.32869	TMAR(4:2)	-60.4234	-12.41266	67.09790
TMAR(2:3)	1379.9607	1408.76715	1456.47349	TMAR(4:3)	405.1463	460.83872	553.07097
TMAR(2:4)	306.9051	339.55238	393.61956	TMAR(4:4)	-703.0540	-639.67982	-534.72588

From Table 3.19, The TMAR($K;p$) model, whose K component is equal to 1, such as MAR(1;p), in the first four lines, is the original autoregressive model with order p, while the other K components represent the t mixture autoregressive models with multiple components. All the criterion values for multiple components are smaller than those for the single component, confirming the motivation of the mixture distribution in the stock dataset. Among these models, the one with the smallest AIC, HQIC, and BIC is the MAR(4:4). Therefore, the optimal model for BANPU is the MAR(4;1) model.

Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting quantile residual plots, Q-Q plots, and test t distribution of the quantile residuals by using Kolmogorov Smirnov test is presented in Figure 3.19 and Table 3.20.

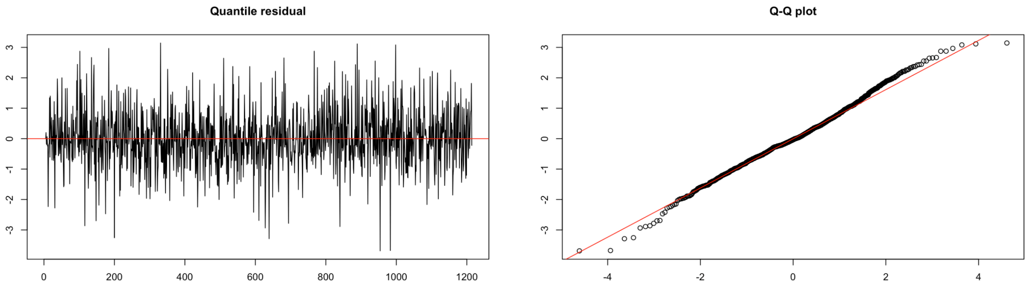


Figure 3.19: Quantile residual plot of TMAR(4:4) for BANPU

Table 3.20: *t* distribution test of TMAR(4:4) for BANPU

Test	Statistic	p-value
Kolmogorov-Smirnov	0.055	0.001283

From Figure 3.19, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. Additionally, Table 3.20 presents the *t* distribution test of the quantile residuals. Specifically the Kolmogorov-Smirnov test for the quantile residuals of BANPU reveals p-values less than 0.05, indicating that the distribution of the given data does not conform to a *t* distribution. The summary of the family of univariate mixture autoregressive models for BANPU is presented in Table 3.21, which includes criteria such as AIC, HQIC, and BIC, as well as the mean square error (MSE).

Table 3.21: The best of each candidate model for BANPU

Model	AIC	HQIC	BIC	MSE
TMAR_{EM}(4;4)	-703.0540	-639.67982	-534.72588	0.5633856
TMAR_{MLE}(4;1)	-83.07111	-46.58295	13.84507	1.0105105
MAR_{MLE}(4;1)	17.12187	45.92831	93.63465	1.0005787

From Table 3.21, the summary of the family of univariate mixture autoregressive models, which includes the MAR model, the TMAR model with parameter estimates by MLE, and the TMAR model with parameter estimates using the EM algorithm, reveals that the best candidate model for BANPU is the TMAR model estimated with the EM algorithm. This conclusion is based on the smallest criterion and MSE values.

Next, we analyze ESSO stock data in the energy and utility sectors by fitting it with the TMAR($K;p$) model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$. The criteria values for each model are presented in Table 3.22.

Table 3.22: Criteria for the candidate TMAR model with using EM for ESSO

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMAR(1:1)	1001.7015	1013.2241	1032.3066	TMAR(3:1)	-1152.7885	-1122.0617	-1071.1749
TMAR(1:2)	851.3204	864.7634	887.0264	TMAR(3:2)	747.2210	783.7092	844.1372
TMAR(1:3)	660.1353	675.4987	700.9421	TMAR(3:3)	1867.1625	1909.4119	1979.3812
TMAR(1:4)	520.3834	537.6673	566.2911	TMAR(3:4)	1242.6832	1290.6939	1370.2045
TMAR(2:1)	993.0267	1014.1515	1049.1361	TMAR(4:1)	-861.8696	-821.5406	-754.7517
TMAR(2:2)	901.8843	926.8499	968.1954	TMAR(4:2)	490.6107	538.6214	618.1320
TMAR(2:3)	-1064.1187	-1035.3123	-987.6059	TMAR(4:3)	-553.7844	-498.0919	-405.8597
TMAR(2:4)	961.2732	993.9205	1047.9877	TMAR(4:4)	532.6991	596.0732	701.0272

From Table 3.22, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. The TMAR(3:1) model exhibits the lowest AIC, HQIC, and BIC values, indicating that it is the best model for ESSO stock data. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting quantile residual plots, Q-Q plots, and test t distribution of the quantile residuals by using Kolmogorov Smirnov test is presented in Figure 3.20 and Table 3.23.

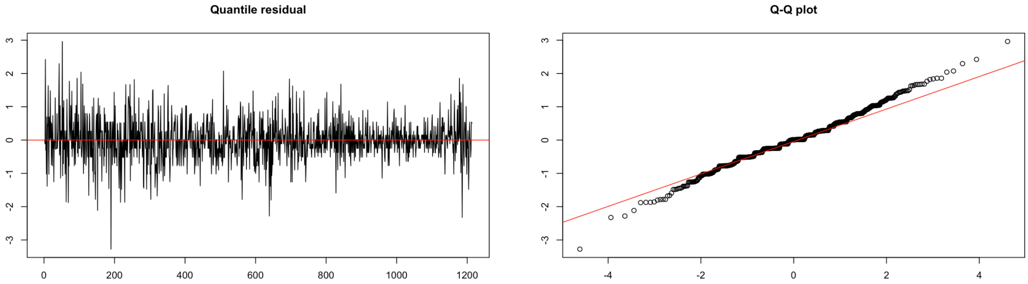


Figure 3.20: Quantile residual plot of TMAR(3:1) for ESSO

Table 3.23: t distribution test of TMAR(3:1) for ESSO

Test	Statistic	p-value
Kolmogorov-Smirnov	0.17115	2.2e-16

From Figure 3.20, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. Additionally, Table 3.23 presents the t distribution test of the quantile residuals. Specifically the Kolmogorov-Smirnov test for the quantile residuals of BANPU reveals p-values less than 0.05, indicating that the distribution of the given data does not conform to a t distribution. The summary of the family of univariate mixture autoregressive models for ESSO is presented in Table 3.24, which includes criteria such as AIC, HQIC, and BIC, as well as the mean square error (MSE).

Table 3.24: The best of each candidate model for ESSO

Model	AIC	HQIC	BIC	MSE
TMAR_{EM}(3;1)	-1152.7885	-1122.0617	-1071.1749	0.3804303
TMAR_{MLE}(4;1)	-23.73528	12.75288	73.18090	0.9937875
MAR_{MLE}(4;1)	6.45798	35.26442	82.97076	0.9946433

From Table 3.24, the summary of the family of univariate mixture autoregressive models reveals that the best candidate model for ESSO is the TMAR model estimated with the EM algorithm. This conclusion is based on the smallest criterion values and MSE values.

Next, we analyze BCP stock data in the energy and utility sectors by fitting it with the TMAR$(K;p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$. The criteria values for each model are presented in Table 3.25.

Table 3.25: Criteria for the candidate TMAR model with using EM for BCP

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMAR(1:1)	1605.9925	1617.5151	1636.5976	TMAR(3:1)	2999.2762	3030.0031	3080.8899
TMAR(1:2)	1461.5116	1474.9547	1497.2176	TMAR(3:2)	549.7641	586.2522	646.6803
TMAR(1:3)	-491.9944	-476.6310	-451.1876	TMAR(3:3)	2033.7533	2076.0028	2145.9721
TMAR(1:4)	-335.1214	-317.8375	-289.2137	TMAR(3:4)	1254.1858	1302.1965	1381.7071
TMAR(2:1)	2536.4205	2557.5452	2592.5299	TMAR(4:1)	590.5399	630.8689	697.6578
TMAR(2:2)	2812.8611	2837.8267	2879.1722	TMAR(4:2)	2274.9069	2322.9177	2402.4282
TMAR(2:3)	1402.3748	1431.1813	1478.8876	TMAR(4:3)	1133.206	1188.899	1281.131
TMAR(2:4)	3590.3761	3623.0234	3677.0906	TMAR(4:4)	1150.7145	1214.0886	1319.0426

From Table 3.25, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. The TMAR(3:2) model exhibits the lowest AIC, HQIC, and BIC values, indicating that it is the best model for BCP stock data. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting quantile residual plots, Q-Q plots, and test t distribution of the quantile residuals by using Kolmogorov Smirnov test is presented in Figure 3.21 and Table 3.26.

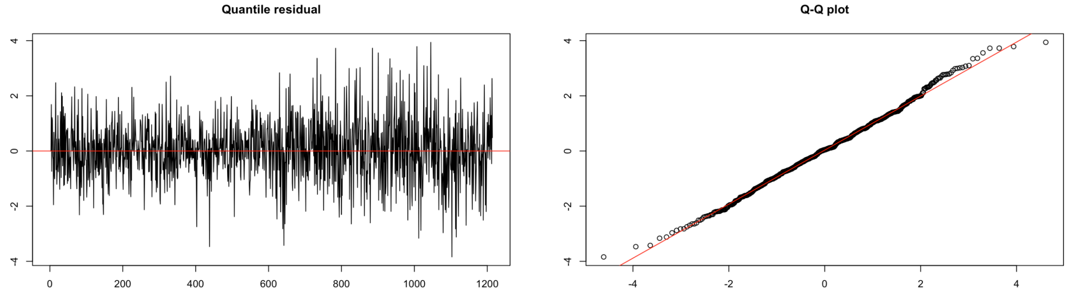


Figure 3.21: Quantile residual plot of TMAR(3:2) for BCP

Table 3.26: t distribution test of TMAR(3:2) for BCP

Test	Statistic	p-value
Kolmogorov-Smirnov	0.032672	0.1497

From Figure 3.21, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. Additionally, Table 3.26 presents the t distribution test of the quantile residuals. Specifically the Kolmogorov-Smirnov test for the quantile residuals of BANPU reveals p-values greater than 0.05, indicating that the distribution of the given data conform to a t distribution. The summary of the family of univariate mixture autoregressive models for BCP is presented in Table 3.27, which includes criteria such as AIC, HQIC, and BIC, as well as the mean square error (MSE).

Table 3.27: The best of each candidate model for BCP

Model	AIC	HQIC	BIC	MSE
TMAR$_{\text{EM}}$(3:2)	549.7641	586.2522	646.6803	0.6821307
TMAR$_{\text{MLE}}$(3:1)	1710.122	1737.008	1781.534	0.9942205
MAR$_{\text{MLE}}$(2:1)	1882.570	1896.013	1918.276	0.9951634

From Table 3.27, the summary of the family of univariate mixture autoregressive models reveals that the best candidate model for BCP is the TMAR(3:2) model estimated with the EM algorithm. This conclusion is based on the smallest criterion values and MSE values.

Next, we analyze HANA stock data in the electronic components sector by fitting it with the TMAR($K;p$) model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$, and the criteria values for each model are presented in Table 3.28.

Table 3.28: Criteria for the candidate TMAR model with using EM for HANA

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMAR(1:1)	4951.816	4963.339	4982.422	TMAR(3:1)	2744.888	2775.615	2826.502
TMAR(1:2)	4842.794	4856.237	4878.500	TMAR(3:2)	4004.121	4040.610	4101.038
TMAR(1:3)	4698.406	4713.769	4739.212	TMAR(3:3)	2734.783	2777.033	2847.002
TMAR(1:4)	4549.531	4566.815	4595.439	TMAR(3:4)	5689.577	5737.588	5817.098
TMAR(2:1)	4564.979	4586.104	4621.089	TMAR(4:1)	2699.12	2739.449	2806.238
TMAR(2:2)	1576.191	1601.156	1642.502	TMAR(4:2)	3802.404	3850.415	3929.926
TMAR(2:3)	4349.171	4377.978	4425.684	TMAR(4:3)	2528.586	2584.278	2676.510
TMAR(2:4)	4361.067	4393.714	4447.781	TMAR(4:4)	3998.370	4061.745	4166.699

From Table 3.28, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. The TMAR(2:2) model exhibits the lowest AIC, HQIC, and BIC values, indicating that it is the best model for HANA stock data. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting quantile residual plots, Q-Q plots, and test t distribution of the quantile residuals by using Kolmogorov Smirnov test is presented in Figure 3.22 and Table 3.29.

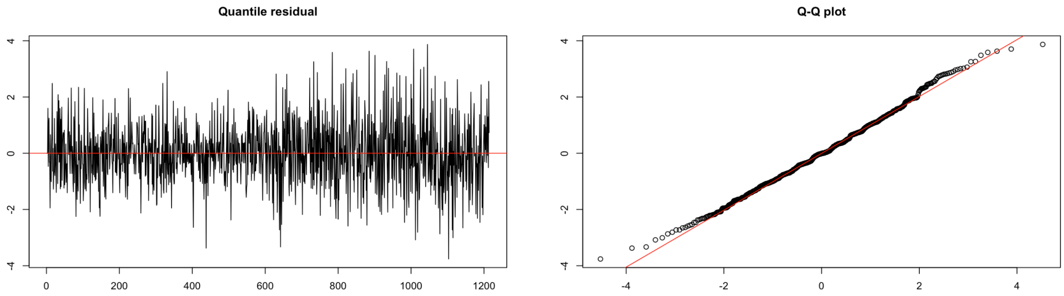


Figure 3.22: Quantile residual plot of TMAR(2:2) for HANA

Table 3.29: t distribution test of TMAR(2:2) for HANA

Test	Statistic	p-value
Kolmogorov-Smirnov	0.030268	0.216

From Figure 3.22, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. Additionally, Table 3.29 presents the t distribution test of the quantile residuals. Specifically the Kolmogorov-Smirnov test for the quantile residuals of BANPU reveals p-values greater than 0.05, indicating that the distribution of the given data conform to a t distribution. The summary of the family of univariate mixture autoregressive models for HANA is presented in Table 3.30, which includes criteria such as AIC, HQIC, and BIC, as well as the mean square error (MSE).

Table 3.30: The best of each candidate model for HANA

Model	AIC	HQIC	BIC	MSE
$\text{TMAR}_{\text{EM}}(2;2)$	1576.191	1601.156	1642.502	1.186730
$\text{TMAR}_{\text{MLE}}(4;1)$	3649.325	3685.814	3746.242	0.6950017
$\text{MAR}_{\text{MLE}}(3;1)$	3671.415	3692.540	3727.525	1.0053008

From Table 3.30, the summary of the family of univariate mixture autoregressive models reveals that the best candidate model for HANA, based on the criteria, is the TMAR(3:2) model estimated with the EM algorithm. However, the MSE values indicate that the best model is the TMAR(4:1) model estimated with the MLE. Therefore, both criteria and MSE lead to the same result, indicating that the TMAR model is preferred over the MAR model.

Next, we analyze TEAM stock data in the electronic components sector by fitting it with the TMAR($K;p$) model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$, and the criteria values for each model are presented in Table 3.31.

Table 3.31: Criteria for the candidate TMAR model with using EM for TEAM

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMAR(1:1)	6714.593	6726.116	6745.198	TMAR(3:1)	6445.419	6466.544	6501.529
TMAR(1:2)	5088.825	5102.268	5124.531	TMAR(3:2)	737.3544	773.8425	834.2705
TMAR(1:3)	4149.540	4164.903	4190.346	TMAR(3:3)	924.5772	966.8266	1036.7959
TMAR(1:4)	3495.515	3512.799	3541.423	TMAR(3:4)	2225.2418	2273.2526	2352.7631
TMAR(2:1)	6445.419	6466.544	6501.529	TMAR(4:1)	3417.119	3457.448	3524.237
TMAR(2:2)	5265.5946	5290.5602	5331.9057	TMAR(4:2)	326.1886	374.1993	453.7099
TMAR(2:3)	8632.9837	8661.7902	8709.4965	TMAR(4:3)	-1175.9961	-1120.3036	-1028.0714
TMAR(2:4)	1945.7314	1978.3787	2032.4459	TMAR(4:4)	-396.8455	-333.4713	-228.5174

From Table 3.31, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. The TMAR(4:3) model exhibits the lowest AIC, HQIC, and BIC values, indicating that it is the best model for TEAM stock data. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting quantile residual plots, Q-Q plots, and test t distribution of the quantile residuals by using Kolmogorov Smirnov test is presented in Figure 3.23 and Table 3.32.

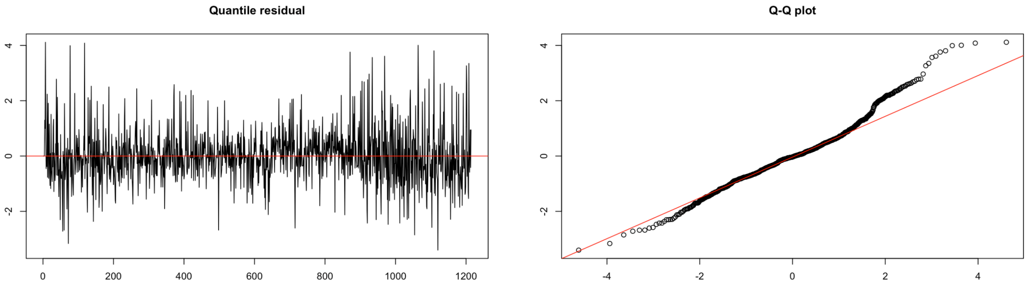


Figure 3.23: Quantile residual plot of TMAR(4:3) for TEAM

Table 3.32: t distribution test of TMAR(4:3) for TEAM

Test	Statistic	p-value
Kolmogorov-Smirnov	0.0857	3.466e-08

From Figure 3.23, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. Additionally, Table 3.32 presents the *t* distribution test of the quantile residuals. Specifically the Kolmogorov-Smirnov test for the quantile residuals of BANPU reveals p-values less than 0.05, indicating that the distribution of the given data does not conform to a *t* distribution. The summary of the family of univariate mixture autoregressive models for TEAM is presented in Table 3.33, which includes criteria such as AIC, HQIC, and BIC, as well as the mean square error (MSE).

Table 3.33: The best of each candidate model for TEAM

Model	AIC	HQIC	BIC	MSE
$\text{TMAR}_{\text{EM}}(4;3)$	-1175.9961	-1120.3036	-1028.0714	0.6678313
$\text{TMAR}_{\text{MLE}}(4;4)$	-4922.316	-4862.804	-4764.266	0.9977746
$\text{MAR}_{\text{MLE}}(4;1)$	-4228.198	-4199.391	-4151.685	0.9929871

From Table 3.33, the summary of the family of univariate mixture autoregressive models reveals that the best candidate model for TEAM, based on the criteria, is the TMAR(4:4) model estimated with the MLE. However, the MSE values indicate that the best model is the TMAR(4:3) model estimated with the EM algorithm. Therefore, both criteria and MSE lead to the same result, indicating that the TMAR model is preferred over the MAR model.

Next, we analyze KCE stock data in the electronic components sector by fitting it with the TMAR$(K; p)$ model, where we explore values of $K = 1, 2, 3, 4$ and $p = 1, 2, 3, 4$, and the criteria values for each model are presented in Table 3.34.

Table 3.34: Criteria for the candidate TMAR model with using EM for KCE

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMAR(1:1)	5560.454	5571.977	5591.059	TMAR(3:1)	3873.866	3904.593	3955.48
TMAR(1:2)	5341.030	5354.473	5376.736	TMAR(3:2)	3356.733	3393.221	3453.649
TMAR(1:3)	5129.536	5144.899	5170.343	TMAR(3:3)	3077.612	3119.861	3189.830
TMAR(1:4)	4953.420	4970.704	4999.328	TMAR(3:4)	3411.2	3459.211	3538.722
TMAR(2:1)	2881.607	2902.731	2937.716	TMAR(4:1)	2339.492	2379.821	2446.61
TMAR(2:2)	3081.329	3106.294	3147.640	TMAR(4:2)	3295.044	3343.055	3422.566
TMAR(2:3)	3090.927	3119.733	3167.440	TMAR(4:3)	1884.784	1940.477	2032.709
TMAR(2:4)	4959.662	4992.309	5046.376	TMAR(4:4)	2648.045	2711.420	2816.374

From Table 3.34, the multiple components have smaller AIC, HQIC, and BIC values than the single component model. The TMAR(4:3) model exhibits the lowest AIC, HQIC, and BIC values, indicating that it is the best model for KCE stock data. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting quantile residual plots, Q-Q plots, and test t distribution of the quantile residuals by using Kolmogorov Smirnov test is presented in Figure 3.24 and Table 3.35.

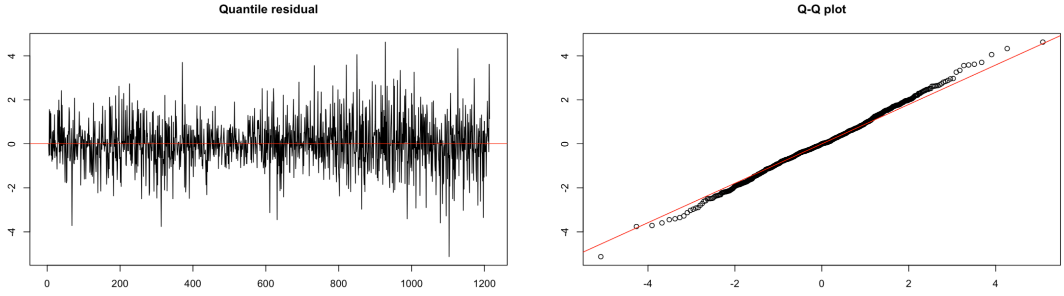


Figure 3.24: Quantile residual plot of TMAR(4:3) for KCE

Table 3.35: *t* distribution test of TMAR(4:3) for KCE

Test	Statistic	p-value
Kolmogorov-Smirnov	0.035523	0.0934

From Figure 3.24, In the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, there are some outliers. Additionally, Table 3.35 presents the t distribution test of the quantile residuals. Specifically the Kolmogorov-Smirnov test for the quantile residuals of BANPU reveals p-values greater than 0.05, indicating that the distribution of the given data conform to a t distribution. The summary of the family of univariate mixture autoregressive models for KCE is presented in Table 3.36, which includes criteria such as AIC, HQIC, and BIC, as well as the mean square error (MSE).

Table 3.36: The best of each candidate model for KCE

Model	AIC	HQIC	BIC	MSE
$\text{TMAR}_{\text{EM}}(4{:}3)$	1884.784	1940.477	2032.709	1.198544
$\text{TMAR}_{\text{MLE}}(1{:}1)$	3522.908	3530.589	3543.311	0.9684834
$\text{MAR}_{\text{MLE}}(3{:}1)$	3512.307	3533.432	3568.416	0.9971298

From Table 3.36, the summary of the family of univariate mixture autoregressive models reveals that the best candidate model for KCE, based on the criteria, is the TMAR(4:3) model estimated with the EM algorithm. However, the MSE values indicate that the best model is the TMAR(1:1) model estimated with the MLE. Therefore, both criteria and MSE lead to the same result, indicating that the TMAR model is preferred over the MAR model.

From two different sectors in the stock market, namely energy and utility, and electronics components, each sector having three stocks, we selected the best model using the selection criteria outlined in Section 2.3, which include the AIC, HQIC, and BIC criteria. We then assessed the validity of the chosen model through model diagnostics, as discussed in Section 2.4. The comparison involved the best candidate models in each sector, considering both MAR model and the TMAR model. We used maximum likelihood estimation and the EM algorithm to estimate parameters for each dataset, selecting the best model based on mean square error (MSE), as presented in Table 3.37.

Table 3.37: The best of each candidate model for the Thai stock market

	MLE	MSE	MLE	MSE	EM algorithm	MSE
BANPU	MAR(4;1)	1.0005787	TMAR(4;1)	1.0105105	TMAR(3;4)	0.5633856
ESSO	MAR(4;1)	0.9946433	TMAR(4;1)	0.9937875	TMAR(3;1)	0.3804303
BCP	MAR(2;1)	0.9951634	TMAR(3;1)	0.9942205	TMAR(3;2)	0.6821307
HANA	MAR(3;1)	1.0053008	TMAR(4;1)	0.6950017	TMAR(2;2)	1.186730
TEAM	MAR(4;1)	0.9929871	TMAR(4;4)	0.9977746	TMAR(4;3)	0.6678313
KCE	MAR(3;1)	0.9971298	TMAR(1;1)	0.9684834	TMAR(4;3)	1.198544

Table 3.37 presents the results of the TMAR models, considering different parameter estimation methods based on mean square error (MSE). For BANPU, ESSO, BCP, and TEAM, almost all instances of the EM algorithm yield significantly lower MSE values compared to maximum likelihood estimation. Conversely, HANA and KCE prefer parameter estimates from maximum likelihood estimation. This table indicates that in 4 out of 6 stock datasets, the EM algorithm outperforms in parameter estimation.

In this chapter, we introduce the family of univariate mixture autoregressive models, comprising the mixture autoregressive (MAR) model, the t mixture autoregressive (TMAR) model, and employ both the EM algorithm and maximum likelihood estimation (MLE) for parameter estimation. We conduct a simulation study to test the accuracy of the model and then apply it to Thai stock market data. For the mixture autoregressive (MAR) model, all criteria for each stock indicate that the multiple component model is better than the single component model. However, almost the entire residual of the model does not follow a normal distribution. Consequently, the alternative distribution, the t mixture autoregressive model, is considered, which is suitable for data exhibiting heavy tails, such as stock market data. In Table 3.37, the t mixture autoregressive model, utilizing both the maximum likelihood estimator and EM algorithm developed in Section 3.2.1.2 to estimate parameters, is preferred over the mixture autoregressive model. The next chapter introduces into the multivariate mixture autoregressive model, employing multiple variables to forecast potential outcomes, examining the correlation within each dataset, and applying the model to various stock sectors.

CHAPTER IV

THE FAMILY OF MULTIVARIATE MIXTURE AUTOREGRESSIVE MODELS

In this chapter, we introduce the family of multivariate mixture autoregressive models and discuss their specifications. We construct the EM algorithm for estimating parameters in the multivariate mixture vector autoregressive model. The simulation study of the mixture vector autoregressive (MVAR) model and the t mixture vector autoregressive (TMVAR) model and investigates the performance of the EM algorithm that we develop compare with the maximum likelihood estimation. Initially, we consider the top stocks from two different sectors, with each sector comprising of three stocks. The energy and utility sectors include BANPU, ESSO, and BCP, and the electronic components sector includes HANA, TEAM, and KCE. Subsequently, we compare the performance of parameter estimation using information criteria and the mean square error.

4.1 Mixture vector autoregressive model

The n dimensional vector time series Y_t is said to be the mixture vector autoregressive model denoted as MVAR(n:$K; p_1, p_2, p_3, \ldots, p_k$) if the distribution function of Y_t given pass information can be written as

$$F(Y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k \Phi(\Omega_k^{-1/2}(Y_t - \Theta_{k0} - \Theta_{k1}Y_{t-1} - \cdots - \Theta_{kp_k}Y_{t-p_k})), \tag{4.1}$$

where $F(Y_t|\mathcal{F}_{t-1})$ is the cumulative distribution function of Y_t given the past information $Y_{t-1}, Y_{t-2}, Y_{t-3}, \ldots,$ Y_1, $\Phi(\cdot)$ represents the cumulative distribution function of the multivariate Gaussian distribution with mean zero and variance-covariance matrix equal to identity matrix, Θ_{k0} is an n dimension vector, $\Theta_{k1}, \ldots, \Theta_{kp_k}$ are $n \times n$ coefficient matrices and Ω_k is the $n \times n$ variance covariance matrix for k^{th} component, the mixing proportion

$\alpha_k > 0$, $k = 1, 2, 3, \ldots, K$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_K = 1$.

In this study, we will construct an appropriate mixture autoregressive model for Thai stock data and a suitable multivariate mixture vector autoregressive model to study correlation between different stock markets by using the EM algorithm to estimate parameters.

4.1.1 Parameter estimation

In this section, we discuss the method to estimate the parameters that we developed in this study, which is the EM algorithm, and compare it with the maximum likelihood function.

4.1.1.1 Parameter estimation by the maximum likelihood function

In the case of the mixture vector autoregressive model, the maximum likelihood method is the parameter estimation method used in this study to compare with the EM algorithm. Specifically, given a time series $\mathbf{Y} = (Y_1, Y_2, Y_3, \ldots, Y_t)$, the likelihood function for the mixture vector autoregressive model is the product of conditional density

$$L(\tilde{\Theta}, \Omega, \alpha|\mathbf{Y}) = \prod_{t=p+1}^{T} \sum_{k=1}^{K} \frac{\alpha_k}{(2\pi)^{\frac{n}{2}} |\Omega_k|^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}(Y_t - \mu_{kt})^T \Omega_k^{-1} (Y_t - \mu_{kt})\Big), \qquad (4.2)$$

where Ω_k is a $n \times n$ variance covariance matrix for k^{th} component, and $\mu_{kt} = Y_t - \tilde{\Theta}_k X_{kt}$ is a $n \times n$ autoregressive matrices, $X_{kt} = (1, Y_{t-1}^T, Y_{t-2}^T)$, and $\tilde{\Theta}_k = [\Theta_{k0}, \Theta_{k1}, \ldots, \Theta_{kp_k}]$. The maximum likelihood function to estimate $\hat{\Upsilon}$ which is defined as

$$\hat{\Upsilon} = \arg\max_{\Upsilon} \ell(\Upsilon|\mathbf{Y}), \qquad (4.3)$$

where $\Upsilon = (\alpha_k, \tilde{\Theta}_k, \Omega_k)$.

The estimation of parameters requires the use of a numerical technique [6]. In the context of the mixture components of the mixture vector autoregressive model, finding a general solution may not be feasible [7]. The alternative to the parameter estimate is the

EM algorithm.

4.1.1.2 Parameter estimation by the EM algorithm

The parameter estimation is conducted using the EM algorithm, and the log-likelihood is constructed using the normal scale mixture model. Assume that the n dimension vectors of observations Y_T are generated from MVAR$(n, K; p)$ model for $t = 1, 2, \ldots, T$ and let $Z_t = (Z_{1t}, Z_{2t}, Z_{3t}, \ldots, Z_{kt})^T$, where

$$Z_{it} = \begin{cases} 1 & \text{if } \text{ if } Y_t \text{ comes from the } i^{th} \text{ component component; } 1 \leq i \leq K, \\ 0 & \text{otherwise.} \end{cases}$$

and the conditional log likelihood function of the mixture vector autoregressive model at time t is

$$l_t = \sum_{k=1}^{K} Z_{kt} \log(\alpha_k) - \frac{1}{2} \sum_{k=1}^{K} Z_{kt} \log |\Omega_k| - \frac{1}{2} \sum_{k=1}^{K} Z_{kt} (\mathbf{e_{kt}}^T \Omega_k^{-1} \mathbf{e_{kt}}), \tag{4.4}$$

where

$$\tilde{\Theta}_k = [\Theta_{k0}, \Theta_{k1}, \Theta_{k2}, \ldots, \Theta_{kp_k}], \tag{4.5}$$

$$X_{kt} = (1, Y_{t-1}^T, Y_{t-2}^T, \ldots, Y_{t-p_k}^T)^T, \tag{4.6}$$

$$\mathbf{e_{kt}} = Y_t - \tilde{\Theta}_k X_{kt}. \tag{4.7}$$

The log likelihood function of the mixture vector autoregressive model is given by

$$l = \sum_{t=p+1}^{T} \Big\{ \sum_{k=1}^{K} Z_{kt} \log(\alpha_k) - \frac{1}{2} \sum_{k=1}^{K} Z_{kt} \log |\Omega_k| - \frac{1}{2} \sum_{k=1}^{K} Z_{kt} (\mathbf{e_{kt}}^T \Omega_k^{-1} \mathbf{e_{kt}}) \Big\}. \tag{4.8}$$

The parameters are estimated by iteratively maximizing the log-likelihood through the Expectation-Maximization(EM) procedure [3], which involves two main steps: the Expectation(E-step) and the Maximization(M-step). These steps are repeated iteratively until the algorithm converges. An illustration of the EM algorithm is provide in Figure

4.1.

The Expectation step. Assume that the parameters $\alpha_k, \tilde{\Theta}_k, \Omega_k$ is known. The unobserved random variable Z are replaced by their expectations, conditional over the parameters and the observed data $Y_1, \ldots, Y_T$. Let τ_{kt} be the conditional expectation of the k^{th} component of Z_t which defined as

$$\tau_{kt} = \frac{\alpha_k |\Omega_k|^{-\frac{1}{2}} exp(-\frac{1}{2}\mathbf{e_{kt}}^T \Omega_k^{-1} \mathbf{e_{kt}})}{\sum_{k=1}^{K} \alpha_k |\Omega_k|^{-\frac{1}{2}} exp(-\frac{1}{2}\mathbf{e_{kt}}^T \Omega_k^{-1} \mathbf{e_{kt}})}. \tag{4.9}$$

The Maximization step. Suppose that the unobserved random variable is actually known. By maximizing the loglikelihood function (4.8), the estimates of our model are obtained through the first derivatives with respect to the parameters $\alpha_k, \tilde{\Theta}_k$, and Ω_k are

$$\frac{\partial l}{\partial \alpha_k} = \sum_{t=p+1}^{T} \Big(\frac{Z_{kt}}{\alpha_k} - \frac{Z_{Kt}}{\alpha_k} \Big), \tag{4.10}$$

$$\frac{\partial l}{\partial \tilde{\Theta}_k} = \Omega^{-1} \Big(\sum_{t=p+1}^{T} Z_{kt} Y_t X_{kt}^T - \sum_{t=p+1}^{T} Z_{kt} X_{kt} X_{kt}^T \Big), \tag{4.11}$$

$$\frac{\partial l}{\partial \Omega_k} = \frac{1}{2} \Big\{ \Omega^{-1} \sum_{t=p+1}^{T} (Z_{kt} \mathbf{e_{kt}} \mathbf{e_{kt}}^T) - \sum_{t=p+1}^{T} Z_{kt}, \tag{4.12}$$

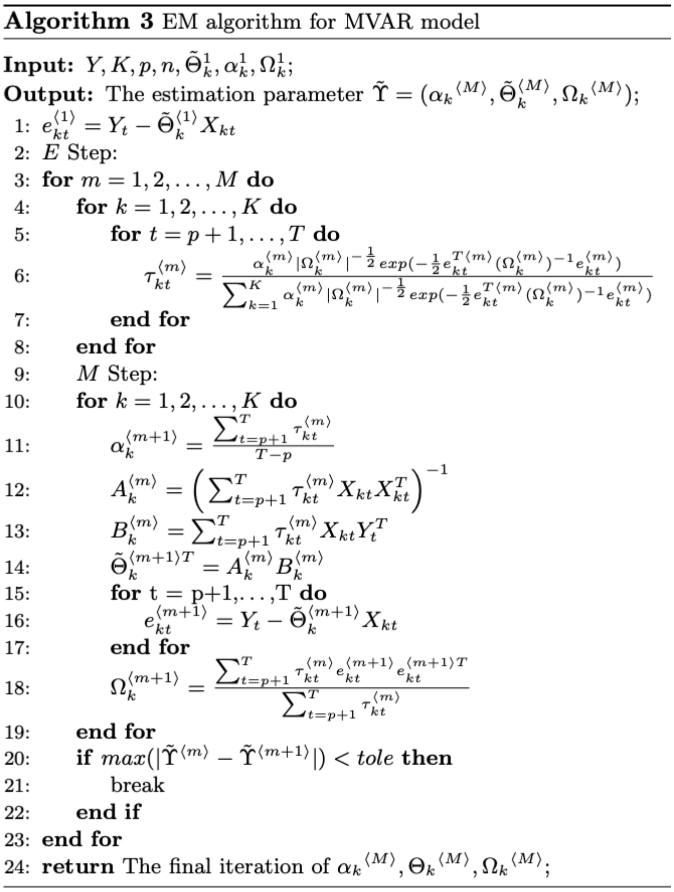
where $\mathbf{e_{kt}} = Y_t - \Theta_{k0} - \Theta_{k1} Y_{t-1} - \Theta_{k2} Y_{t-2} - \cdots - \Theta_{kp_k} Y_{t-p_k}$. Subsequently, we substitute the conditional expectation of Z_{kt} in (4.10) to (4.12) and setting equation to zero. The estimates of the parameter are

$$\hat{\alpha}_k = \frac{1}{T-p} \sum_{t=p+1}^{T} \tau_{kt}, \tag{4.13}$$

$$\hat{\tilde{\Theta}}_k^T = \Big(\sum_{t=p+1}^{T} \tau_{kt} X_{kt} X_{kt}^T \Big)^{-1} \Big(\sum_{t=p+1}^{T} \tau_{kt} X_{kt} Y_t^T \Big), \tag{4.14}$$

$$\hat{\Omega}_k = \frac{\sum_{t=p+1}^{T} \tau_{kt} \widehat{\mathbf{e}}_{\mathbf{kt}} \widehat{\mathbf{e}}_{\mathbf{kt}}^T}{\sum_{t=p+1}^{T} \tau_{kt}}, \tag{4.15}$$

where $k = 1, 2, \ldots, K$.

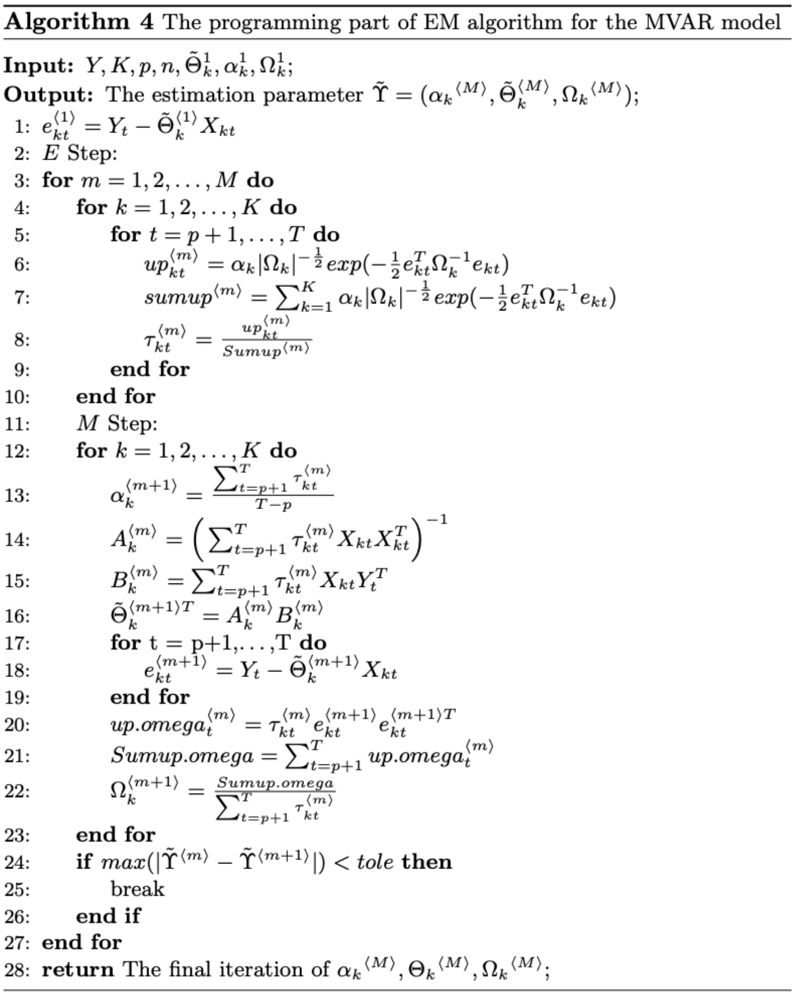


Algorithm 3 EM algorithm for MVAR model

```
Input: $Y, K, p, n, \tilde{\Theta}_k^1, \alpha_k^1, \Omega_k^1$;
Output: The estimation parameter $\tilde{\Upsilon} = (\alpha_k{}^{\langle M \rangle}, \tilde{\Theta}_k^{\langle M \rangle}, \Omega_k{}^{\langle M \rangle})$;
$e_{kt}^{\langle 1 \rangle} = Y_t - \tilde{\Theta}_k^{\langle 1 \rangle} X_{kt}$
$E$ Step:
for $m = 1, 2, \ldots, M$ do
    for $k = 1, 2, \ldots, K$ do
        for $t = p + 1, \ldots, T$ do
            $\tau_{kt}^{\langle m \rangle} = \frac{\alpha_k^{\langle m \rangle} |\Omega_k^{\langle m \rangle}|^{-\frac{1}{2}} exp(-\frac{1}{2} e_{kt}^{T\langle m \rangle} (\Omega_k^{\langle m \rangle})^{-1} e_{kt}^{\langle m \rangle})}{\sum_{k=1}^{K} \alpha_k^{\langle m \rangle} |\Omega_k^{\langle m \rangle}|^{-\frac{1}{2}} exp(-\frac{1}{2} e_{kt}^{T\langle m \rangle} (\Omega_k^{\langle m \rangle})^{-1} e_{kt}^{\langle m \rangle})}$
        end for
    end for
    $M$ Step:
    for $k = 1, 2, \ldots, K$ do
        $\alpha_k^{\langle m+1 \rangle} = \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}{T-p}$
        $A_k^{\langle m \rangle} = \left( \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} X_{kt} X_{kt}^T \right)^{-1}$
        $B_k^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} X_{kt} Y_t^T$
        $\tilde{\Theta}_k^{\langle m+1 \rangle T} = A_k^{\langle m \rangle} B_k^{\langle m \rangle}$
        for t = p+1,...,T do
            $e_{kt}^{\langle m+1 \rangle} = Y_t - \tilde{\Theta}_k^{\langle m+1 \rangle} X_{kt}$
        end for
        $\Omega_k^{\langle m+1 \rangle} = \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} e_{kt}^{\langle m+1 \rangle} e_{kt}^{\langle m+1 \rangle T}}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}$
    end for
    if $max(|\tilde{\Upsilon}^{\langle m \rangle} - \tilde{\Upsilon}^{\langle m+1 \rangle}|) < tole$ then
        break
    end if
end for
return The final iteration of $\alpha_k{}^{\langle M \rangle}, \Theta_k{}^{\langle M \rangle}, \Omega_k{}^{\langle M \rangle}$;
```

Figure 4.1: The EM algorithm for the MVAR model

The EM algorithm for the multivariate mixture autoregressive model that we mention and develop in Section 4.1.1.2 has the following steps in the programme, are show in Figure 4.2:



Algorithm 4 The programming part of EM algorithm for the MVAR model

```
Input: $Y, K, p, n, \tilde{\Theta}_k^1, \alpha_k^1, \Omega_k^1$;
Output: The estimation parameter $\tilde{\Upsilon} = (\alpha_k{}^{\langle M \rangle}, \tilde{\Theta}_k^{\langle M \rangle}, \Omega_k{}^{\langle M \rangle})$;
$e_{kt}^{\langle 1 \rangle} = Y_t - \tilde{\Theta}_k^{\langle 1 \rangle} X_{kt}$
$E$ Step:
for $m = 1, 2, \ldots, M$ do
    for $k = 1, 2, \ldots, K$ do
        for $t = p + 1, \ldots, T$ do
            $up_{kt}^{\langle m \rangle} = \alpha_k |\Omega_k|^{-\frac{1}{2}} exp(-\frac{1}{2} e_{kt}^T \Omega_k^{-1} e_{kt})$
            $sumup^{\langle m \rangle} = \sum_{k=1}^{K} \alpha_k |\Omega_k|^{-\frac{1}{2}} exp(-\frac{1}{2} e_{kt}^T \Omega_k^{-1} e_{kt})$
            $\tau_{kt}^{\langle m \rangle} = \frac{up_{kt}^{\langle m \rangle}}{Sumup^{\langle m \rangle}}$
        end for
    end for
    $M$ Step:
    for $k = 1, 2, \ldots, K$ do
        $\alpha_k^{\langle m+1 \rangle} = \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}{T-p}$
        $A_k^{\langle m \rangle} = \left( \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} X_{kt} X_{kt}^T \right)^{-1}$
        $B_k^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} X_{kt} Y_t^T$
        $\tilde{\Theta}_k^{\langle m+1 \rangle T} = A_k^{\langle m \rangle} B_k^{\langle m \rangle}$
        for t = p+1,...,T do
            $e_{kt}^{\langle m+1 \rangle} = Y_t - \tilde{\Theta}_k^{\langle m+1 \rangle} X_{kt}$
        end for
        $up.omega_t^{\langle m \rangle} = \tau_{kt}^{\langle m \rangle} e_{kt}^{\langle m+1 \rangle} e_{kt}^{\langle m+1 \rangle T}$
        $Sumup.omega = \sum_{t=p+1}^{T} up.omega_t^{\langle m \rangle}$
        $\Omega_k^{\langle m+1 \rangle} = \frac{Sumup.omega}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}$
    end for
    if $max(|\tilde{\Upsilon}^{\langle m \rangle} - \tilde{\Upsilon}^{\langle m+1 \rangle}|) < tole$ then
        break
    end if
end for
return The final iteration of $\alpha_k{}^{\langle M \rangle}, \Theta_k{}^{\langle M \rangle}, \Omega_k{}^{\langle M \rangle}$;
```

Figure 4.2: The programming part of the EM algorithm for the MVAR model

In this section, we will show the calculation example by using the MVAR_EM program for simulation. The model considered in this example is the MVAR(3:2;2) model with generate length of each time series is 10 data points, a dimension vector of 3, K components, and an autoregressive order of 2. The parameter M represents the number of iterations in the EM algorithm.

In the initial step, we generate time series data comprising 10 data points. For the purpose of this illustration, the calculation example is presented with m iterations limited to 1.

	$Y_{T,1}$	$Y_{T,2}$	$Y_{T,3}$
1	0.65	0.00	-0.11
2	1.55	-1.36	-0.14
3	2.96	-0.03	-0.33
4	2.09	0.43	0.48
5	3.17	-0.11	-0.24
6	-1.86	1.22	1.34
7	-0.44	1.08	0.53
8	0.68	1.64	0.47
9	1.04	0.43	1.24
10	-0.57	1.49	1.64

and the initial values for the EM algorithm are

$$
\begin{aligned}
\alpha &= [\alpha_1, \alpha_2] \\
&= [0.50, 0.50], \\
\tilde{\Theta}_1 &= \begin{bmatrix} \Theta_{10} & \Theta_{11} & \Theta_{12} \end{bmatrix} \\
&= \begin{bmatrix} -0.648 & -0.537 & 0.212 & 0.319 & 0.279 & 0.108 & -0.580 \\ 0.114 & -0.083 & 1.054 & 0.607 & 0.638 & -1.026 & 0.097 \\ 0.048 & 0.106 & 1.425 & -0.978 & 0.619 & 0.251 & -0.849 \end{bmatrix}, \\
\tilde{\Theta}_2 &= \begin{bmatrix} \Theta_{20} & \Theta_{21} & \Theta_{22} \end{bmatrix} \\
&= \begin{bmatrix} 0.15 & 0.514 & -0.363 & 0.019 & 0.259 & -0.398 & -0.767 \\ 0.43 & -0.050 & 0.256 & 0.209 & -0.010 & 0.356 & -0.500 \\ 0.45 & 0.092 & 0.459 & 0.292 & -0.205 & 0.233 & -0.030 \end{bmatrix}, \\
\Omega_1 &= \begin{bmatrix} 1.772 & 0.495 & 0.100 \\ 0.495 & 1.560 & 1.151 \\ 0.100 & 1.151 & 0.972 \end{bmatrix}, \\
\Omega_2 &= \begin{bmatrix} 5.538 & -1.188 & -1.242 \\ -1.188 & 0.364 & 0.341 \\ -1.242 & 0.341 & 0.330 \end{bmatrix}.
\end{aligned}
$$

In the initial step of the computation in the Expectation step, we calculate $\mathbf{e_{kt}}^{\langle 1 \rangle} =$

$Y_t - \tilde{\Theta}_k X_{kt}$, where $X_{kt} = (1, Y_{t-1}^T, Y_{t-2}^T)$:

$$\mathbf{e_{1t}}^{\langle 1\rangle} = \begin{bmatrix} NA & NA & 4.536 & 4.077 & 3.675 & 0.236 & -2.488 & 1.857 & 1.865 & 0.052 \\ NA & NA & 1.100 & -1.571 & -2.693 & 0.694 & -3.402 & 2.345 & -0.307 & 1.463 \\ NA & NA & 0.758 & -0.902 & -2.765 & -0.112 & -1.884 & 1.434 & -0.310 & 1.649 \end{bmatrix}.$$

In the Expectation step we have to compute $\tau_{kt} = \frac{\alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}})}{\sum_{k=1}^{K}\alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}})}$. Begin with $K = 1, 2$ and $t = p+1, \ldots, T$. In this example, we illustrate the case of $t = p + 1 = 3$ to T and for all k.

$$\text{up}_{kt}^{\langle m\rangle} = \alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}}) = \begin{bmatrix} NA & NA \\ NA & NA \\ 0.000 & 18.745 \\ 0.000 & 7.925 \\ 0.000 & 6.314 \\ 0.084 & 4.718 \\ 0.006 & 25.551 \\ 0.121 & 0.000 \\ 0.229 & 0.011 \\ 0.110 & 15.770 \end{bmatrix},$$

$$\begin{aligned} \text{sumup}^{\langle m\rangle} &= \sum_{k=1}^{K}\alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}}) \\ &= (0.000, 0.000, 0.306, 0.838, 0.487, 0.500, 0.338, 0.129, 0.104, 0.419), \end{aligned}$$

$$\tau_{kt}^{\langle m \rangle} = \frac{\text{up}_{kt}^{\langle m \rangle}}{\text{sumup}^{\langle m \rangle}} = \begin{bmatrix} NA & NA \\ NA & NA \\ 0.000 & 1.000 \\ 0.000 & 1.000 \\ 0.000 & 1.000 \\ 0.018 & 0.982 \\ 0.000 & 1.000 \\ 1.000 & 0.000 \\ 0.955 & 0.045 \\ 0.007 & 0.993 \end{bmatrix}.$$

In the Maximization Step: we need to estimate $\alpha_k^{\langle m+1 \rangle}, \tilde{\Theta}_1^{\langle m+1 \rangle T}$, and $\Omega_k^{\langle m+1 \rangle}$. In this example, we illustrate the case of $t = p + 1 = 3$ to T and for all k.

$$\begin{aligned} \alpha_k^{\langle m+1 \rangle} &= \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}{T - p} \\ &= [\alpha_1, \alpha_2] \\ &= [0.5620.438]. \end{aligned}$$

When estimating $\tilde{\Theta}_k^{\langle m+1 \rangle T}$, we need to construct the matrix equation $A_k^{\langle m \rangle T} \tilde{\Theta}_k^{\langle m+1 \rangle T} = B_k^{\langle m+1 \rangle T}$, where

$$\begin{aligned} A_k^{\langle m \rangle} &= \Big(\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} X_{kt} X_{kt}^T \Big)^{-1}, \\ B_k^{\langle m \rangle} &= \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} X_{kt} Y_t^T, \end{aligned}$$

$$
\begin{aligned}
\tilde{\Theta}_1^{\langle m+1\rangle T} &= A_1^{\langle m\rangle} B_1^{\langle m\rangle} \\
&= \begin{bmatrix} -0.012 & 0.028 & 0.034 & 0.030 & -0.001 & 0.029 & 0.015 \\ -0.020 & 0.046 & 0.056 & 0.049 & -0.001 & 0.048 & 0.024 \\ -0.019 & 0.043 & 0.052 & 0.046 & -0.001 & 0.045 & 0.022 \end{bmatrix}.
\end{aligned}
$$

Before estimating $\Omega_k^{\langle m+1\rangle}$, we update the term $\tilde{\Theta}_1^{\langle m+1\rangle T}$ to $\mathbf{e_{1t}}^{\langle m+1\rangle}$.

$$
\begin{aligned}
\mathbf{e_{1t}}^{\langle m+1\rangle} &= Y_t - \tilde{\Theta}_k X_{kt} \\
&= \begin{bmatrix} NA & NA & 3.294 & 2.422 & 3.494 & -1.530 & -0.111 & 1.006 & 1.371 & -0.246 \\ NA & NA & 0.508 & 0.972 & 0.427 & 1.759 & 1.618 & 2.185 & 0.975 & 2.034 \\ NA & NA & 0.170 & 0.983 & 0.263 & 1.845 & 1.029 & 0.976 & 1.745 & 2.145 \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\text{Sumup.omega} &= \sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle} e_{kt}^{\langle m+1\rangle} e_{kt}^{\langle m+1\rangle T} \\
&= \begin{bmatrix} 39.216 & 23.503 & 22.665 \\ 23.503 & 44.052 & 28.440 \\ 22.665 & 28.440 & 23.798 \end{bmatrix}, \\
\Omega_k^{\langle m+1\rangle} &= \left(\frac{\text{Sumup.omega}}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}} \right)^{\frac{1}{2}} \\
&= \begin{bmatrix} 1.366 & 1.214 & 1.057 \\ 21.214 & 2.096 & 1.376 \\ 1.057 & 1.376 & 1.159 \end{bmatrix}.
\end{aligned}
$$

For now, we complete the m^{th} iteration, where $m = 1$. We then repeat the Expectation(E) step and Maximization(M) step until the parameter estimate $\tilde{\Upsilon} = (\alpha_k{}^{\langle M\rangle}, \tilde{\Theta}_k^{\langle M\rangle}, \Omega_k{}^{\langle M\rangle})$ convergence by using $max(|\tilde{\Upsilon}^{\langle m\rangle} - \tilde{\Upsilon}^{\langle m+1\rangle}|) < tol$, where tol is the tolerance with a default value of 1×10^{-6} in this study.

4.1.2 Simulation study for the MVAR model

In this section, we examine the performance of parameter estimation using two different methods. First, we employ the EM algorithm in Section 4.1.1.2 that we develop and compare it with the maximum likelihood estimation procedure implemented in the package "uGMAR" [6] in Section 4.1.1.1. Furthermore, we examine the accuracy of parameter estimates. It's important to note that, under the restrictions of the package, the orders of autoregressive components for different components are assumed to be the same. Therefore, the models considered in this study are denoted as MVAR(n:K; p), where n is the dimensional vector, K is the number of components and p is the common order of autoregressive components. The model investigated in this study are the MVAR(3:2; 2), where K component is 2 and order p is 2, and n dimensional vector is 3. Comparing the parameter estimate between the EM algorithm and the MLE. In the part of the EM algorithm using the parameter from the MLE to be an initial value. The best candidate models with the smallest corresponding criterion is the MVAR(3:2; 2).

For the experiments, we generate a time series from the MVAR(3:2;2) which the dimension, n, is 3, the number of component, K, is 2, order of autoregressive model is 2 with a time length of 1000 data points and simulation 1000 replications. In comparing parameter estimates obtained with the maximum likelihood estimation(MLE) and the Expectation-Maximization(EM) algorithm, the EM algorithm utilized parameters from the MLE as an initial value. The best candidate models, identified based on the smallest corresponding criterion, were determined to be MVAR(3:2;2) models, which were correctly chosen. Table 4.1 presents parameter estimates for both methods, comparing them with the exact values.

Table 4.1: Parameter estimates for the MVAR(3:2;2) model using the EM algorithm, when $K = 1$.

	ϕ_{10}	ϕ_{20}	ϕ_{30}						
Exact value	0.00	0.00	0.00						
Estimate of EM	0.004	-0.014	-0.011						
Estimate of MLE	-0.648	0.114	0.048						
Error of EM	0.004	0.014	0.011						
Error of MLE	0.648	0.114	0.048						
	$p = 1$								
	$\phi_{11,1}$	$\phi_{21,1}$	$\phi_{31,1}$	$\phi_{12,1}$	$\phi_{22,1}$	$\phi_{32,1}$	$\phi_{13,1}$	$\phi_{23,1}$	$\phi_{33,1}$
Exact value	0.500	0.100	0.000	0.000	0.100	0.200	0.000	0.300	0.30
Estimate of EM	-0.009	0.030	0.026	-0.011	0.037	0.032	-0.010	0.033	0.03
Estimate of MLE	-0.537	-0.083	0.106	0.212	1.054	1.425	0.319	0.607	-0.98
Error of EM	0.509	0.070	0.026	0.011	0.063	0.168	0.010	0.267	0.27
Error of MLE	1.037	0.183	0.106	0.212	0.954	1.225	0.319	0.307	1.28
	$p = 2$								
	$\phi_{11,2}$	$\phi_{21,2}$	$\phi_{31,2}$	$\phi_{12,2}$	$\phi_{22,2}$	$\phi_{32,2}$	$\phi_{13,2}$	$\phi_{23,2}$	$\phi_{33,2}$
Exact value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Estimate of EM	0.000	-0.001	-0.001	-0.009	0.032	0.027	-0.005	0.016	0.01
Estimate of MLE	0.279	0.638	0.619	0.108	-1.026	0.251	-0.580	0.097	-0.85
Error of EM	0.000	0.001	0.001	0.009	0.032	0.027	0.005	0.016	0.01
Error of MLE	0.279	0.638	0.619	0.108	1.026	0.251	0.580	0.097	0.85
	Ω_{11}	Ω_{21}	Ω_{31}	Ω_{22}	Ω_{32}	Ω_{33}	α_1		
Exact value	2.250	0.000	0.000	1.000	0.500	0.740	0.400		
Estimate of EM	2.718	1.201	1.282	2.445	1.554	1.292	0.519		
Estimate of MLE	1.772	0.495	0.100	1.560	1.151	0.972	0.981		
Error of EM	0.468	1.201	1.282	1.445	1.054	0.552	0.119		
Error of MLE	0.478	0.495	0.100	0.560	0.651	0.232	0.581		

Table 4.2: Parameter estimates for the MVAR(3:2;2) model using the EM algorithm, when $K = 2$.

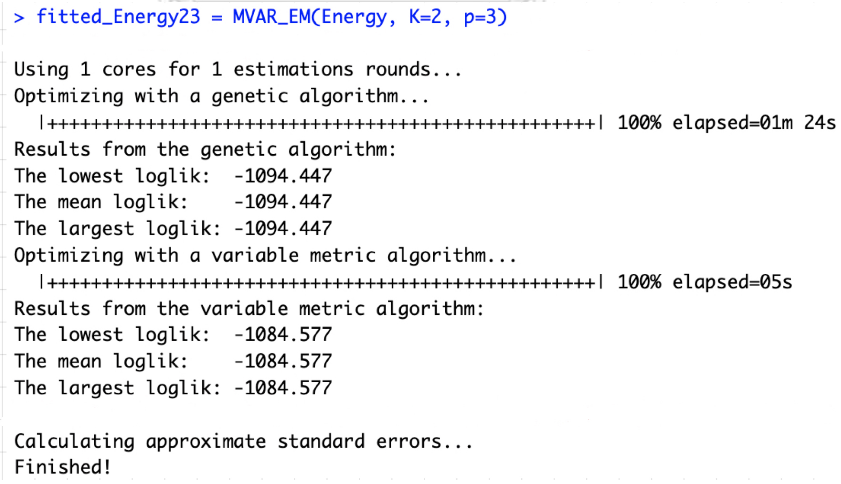
	ϕ_{10}	ϕ_{20}	ϕ_{30}						
Exact value	2.000	1.00	0.00						
Estimate of EM	0.016	0.01	0.011						
Estimate of MLE	0.150	0.43	0.450						
Error of EM	1.984	0.99	0.011						
Error of MLE	1.850	0.57	0.450						
	$p = 1$								
	$\phi_{11,1}$	$\phi_{21,1}$	$\phi_{31,1}$	$\phi_{12,1}$	$\phi_{22,1}$	$\phi_{32,1}$	$\phi_{13,1}$	$\phi_{23,1}$	$\phi_{33,1}$
Exact value	0.700	0.000	0.900	0.100	-0.400	0.000	0.000	0.100	0.800
Estimate of EM	-0.037	-0.023	-0.026	-0.045	-0.029	-0.032	-0.040	-0.025	-0.028
Estimate of MLE	0.514	-0.050	0.092	-0.363	0.256	0.459	0.019	0.209	0.292
Error of EM	0.737	0.023	0.926	0.145	0.371	0.032	0.040	0.125	0.828
Error of MLE	0.186	0.050	0.808	0.463	0.656	0.459	0.019	0.109	0.508
	$p = 2$								
	$\phi_{11,2}$	$\phi_{21,2}$	$\phi_{31,2}$	$\phi_{12,2}$	$\phi_{22,2}$	$\phi_{32,2}$	$\phi_{13,2}$	$\phi_{23,2}$	$\phi_{33,2}$
Exact value	-0.200	0.000	0.000	0.000	0.100	0.000	0.000	0.100	0.000
Estimate of EM	0.001	0.001	0.001	-0.039	-0.025	-0.027	-0.020	-0.012	-0.014
Estimate of MLE	0.259	-0.010	-0.205	-0.398	0.356	0.233	-0.767	-0.500	-0.030
Error of EM	0.201	0.001	0.001	0.039	0.125	0.027	0.020	0.112	0.014
Error of MLE	0.459	0.010	0.205	0.398	0.256	0.233	0.767	0.600	0.030
	Ω_{11}	Ω_{21}	Ω_{31}	Ω_{22}	Ω_{32}	Ω_{33}	α_2		
Exact value	0.260	0.030	0.000	0.090	0.000	0.810	0.600		
Estimate of EM	3.164	-0.376	-0.575	0.654	0.744	0.867	0.481		
Estimate of MLE	5.538	-1.188	-1.242	0.364	0.341	0.330	0.019		
Error of EM	2.904	0.406	0.575	0.564	0.744	0.057	0.119		
Error of MLE	5.278	1.218	1.242	0.274	0.341	0.480	0.581		

The results of the parameter estimates for MVAR(3:2;2) models are presented in Table 4.1. In the part of α, the parameter estimates are quite close to the exact values. The performance of the EM algorithm is good in 39 out of the 54 parameters base on the autoregressive coefficients $\phi_{n \times n,p}$ and standard deviation Ω_k. Furthermore, we compute the mean square error (MSE) for the parameter estimates obtained from both the EM algorithm and the maximum likelihood estimation(MLE). The resulting MSE values are 0.7405123 for the EM algorithm and 1.720199 for MLE. Consequently, judging from the parameter errors and MSE values, it is evident that the EM algorithm outperforms MLE.

4.1.3 Mixture vector autoregressive model for Thai stock market data

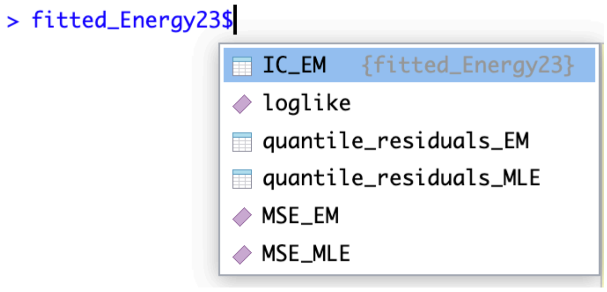
In this section, we are up to analysing a dataset, which we refer to as the sector dataset. We apply the MVAR model to analyze the sector dataset including three stocks in the energy and utility sectors, as well as the electronic sector. The goal is to explore the correlation within each dataset. In the program, following the algorithm outlined in Figure 4.2, we develop a function called MVAR_EM($\mathbf{y_{T\times n}}$, K, p, tol). This function takes input parameters such as the data ($\mathbf{y_{T\times n}}$, where T is the length of the data points and n is the number of dimensional time series), the number of components (K), autoregressive order (p), and tolerance (tol, with a default value of 1×10^{-6}). The initial parameter values for this function are obtained through maximum likelihood estimation.

The following code fits an MVAR model to the energy data, using two components and an autoregressive order of 3 mixture components, as illustrated in Figure 4.3. The MVAR_EM function returning a list of elements including the information criteria (IC), log-likelihood, quantile residuals for both Maximum Likelihood Estimation(MLE) and Expectation-Maximization(EM), as well as mean square error for MLE and EM, as illustrate in Figure 4.4. For instance, Figure 4.5 displays some elements from the MVAR_EM function, which is the information criteria.



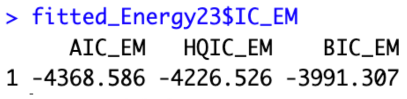
```
> fitted_Energy23 = MVAR_EM(Energy, K=2, p=3)

Using 1 cores for 1 estimations rounds...
Optimizing with a genetic algorithm...
  |++++++++++++++++++++++++++++++++++++++++++++++++++| 100% elapsed=01m 24s
Results from the genetic algorithm:
The lowest loglik:  -1094.447
The mean loglik:    -1094.447
The largest loglik: -1094.447
Optimizing with a variable metric algorithm...
  |++++++++++++++++++++++++++++++++++++++++++++++++++| 100% elapsed=05s
Results from the variable metric algorithm:
The lowest loglik:  -1084.577
The mean loglik:    -1084.577
The largest loglik: -1084.577

Calculating approximate standard errors...
Finished!
```

Figure 4.3: The MVAR_EM function for Energy sector



```
> fitted_Energy23$
```

Figure 4.4: The list of elements in the MVAR_EM function



```
> fitted_Energy23$IC_EM
     AIC_EM    HQIC_EM     BIC_EM
1 -4368.586 -4226.526 -3991.307
```

Figure 4.5: The information criteria from EM function

We analyze the energy sector data, which includes three stocks: BANPU, ESSO, and BCP. The stock plots for this sector, displayed in Figure 4.6, exhibit a similar pattern. We then examine the corresponding correlation values, which are presented in Table 4.3.

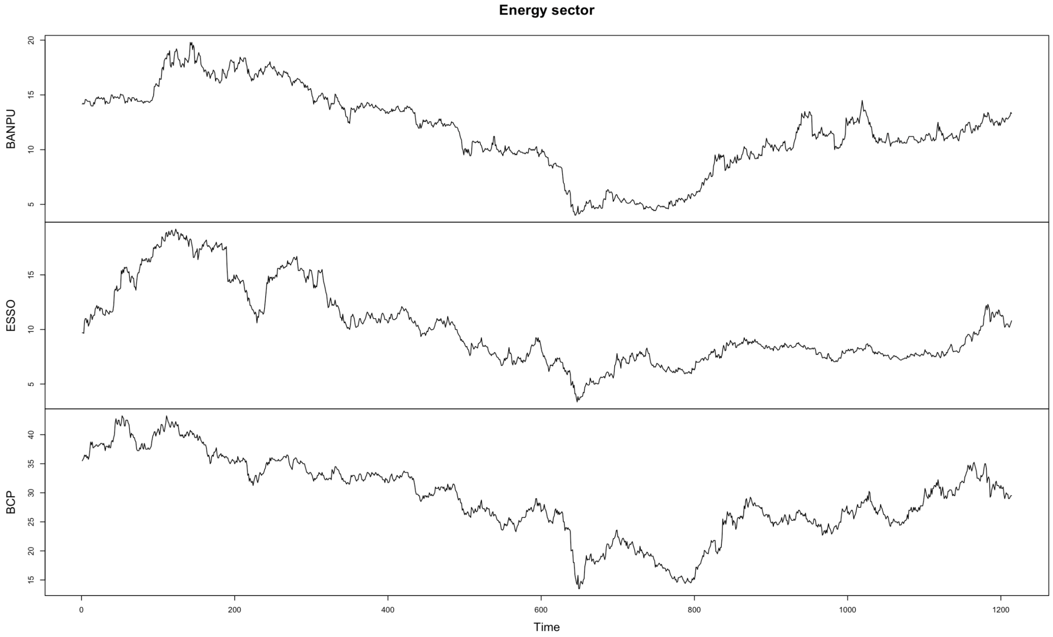


Figure 4.6: The plot of Energy sector

Table 4.3: The correlation between the stock in Energy sector

	BANPU	ESSO	BCP
BANPU	1.00	0.87	0.91
ESSO	0.87	1.00	0.88
BCP	0.91	0.88	1.00

From Figure 4.6, the plot shows that the data exhibits similar patterns and displays a correlation. To initiate the analysis, we apply the MVAR(n:K; p) model to the Energy sector. We explore values of K from 1 to 4, p from 1 to 4, and since the dataset includes 3 stocks, the number of dimension, n, is set to 3. The criteria values for each model are present in Table 4.4.

Table 4.4: The criteria for candidate MVAR models applied to Energy sector data using the EM algorithm

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MVAR(1:1)	-4539.33	-4502.85	-4442.43	MVAR(3:1)	-5098.47	-4989.02	-4807.77
MVAR(1:2)	-4514.08	-4460.32	-4371.30	MVAR(3:2)	-5165.11	-5003.83	-4736.78
MVAR(1:3)	-4497.82	-4426.79	-4309.18	MVAR(3:3)	-5053.58	-4840.49	-4487.66
MVAR(1:4)	-4472.31	-4384.01	-4237.82	MVAR(3:4)	-4672.74	-4407.85	-3969.28
MVAR(2:1)	-5145.72	-5072.75	-4951.92	MVAR(4:1)	-5225.11	-5079.18	-4837.51
MVAR(2:2)	-2996.60	-2889.08	-2711.04	MVAR(4:2)	-5120.87	-4905.83	-4549.76
MVAR(2:3)	-3230.99	-3088.93	-2853.71	MVAR(4:3)	-5033.49	-4749.37	-4278.93
MVAR(2:4)	-4888.88	-4712.29	-4419.91	MVAR(4:4)	-4221.64	-3868.45	-3283.69

From Table 4.4, the MVAR(n:K; p) model, where the number of dimension vector, n, and the number of component K is equal to 1, MVAR(n:1; p), represents the original vector autoregressive model with order p in the first four lines while the other K components represent the MVAR models with multiple components. However, the MVAR(3:2;1) model exhibits the smallest BIC value while the AIC and HQIC criteria favor the MVAR(3:4;1) model. Therefore, considering the three criteria, two out of three indicate that the MVAR(3:4;1) model is the best model. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting autocorrelation, quantile residual plots, Q-Q plots, and test normality test of the quantile residuals by using Kolmogorov Smirnov test is present in Figure 4.7 and Table 4.5.

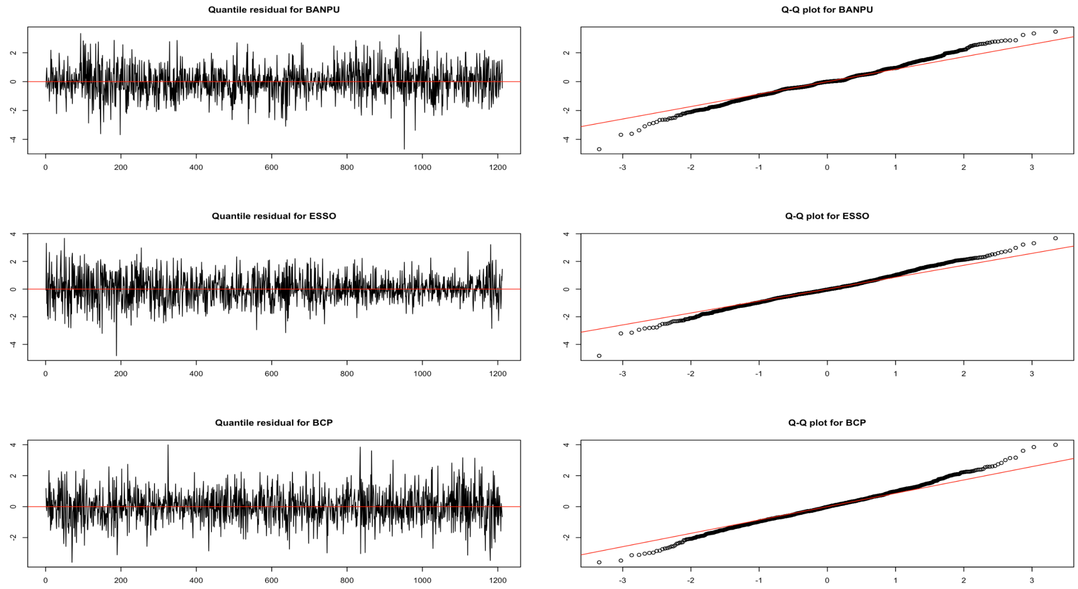


Figure 4.7: Quantile residual plot of MVAR(3:4:1) for Energy sector

Table 4.5: Normality test of MVAR(3:4:1) for Energy stock

Stock	Statistic	p-value
BANPU	0.05438	0.0015
ESSO	0.032046	0.1658
BCP	0.021262	0.6437

From Figure 4.7, in the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, some outliers are observed. Furthermore, in Figure 4.5, the Kolmogorov-Smirnov test for the quantile residuals of ESSO and BCP reveals p-values greater than 0.05, indicating that the distribution of the given data conforms to a normal distribution, while BANPU does not conform to a normal distribution.

Finally, we analyze the energy sector data, which includes three stocks: HANA, TEAM, and KCE. The stock plots for this sector, displayed in Figure 4.8, exhibit a similar pattern. We then examine the corresponding correlation values, which are presented in Table 4.6.

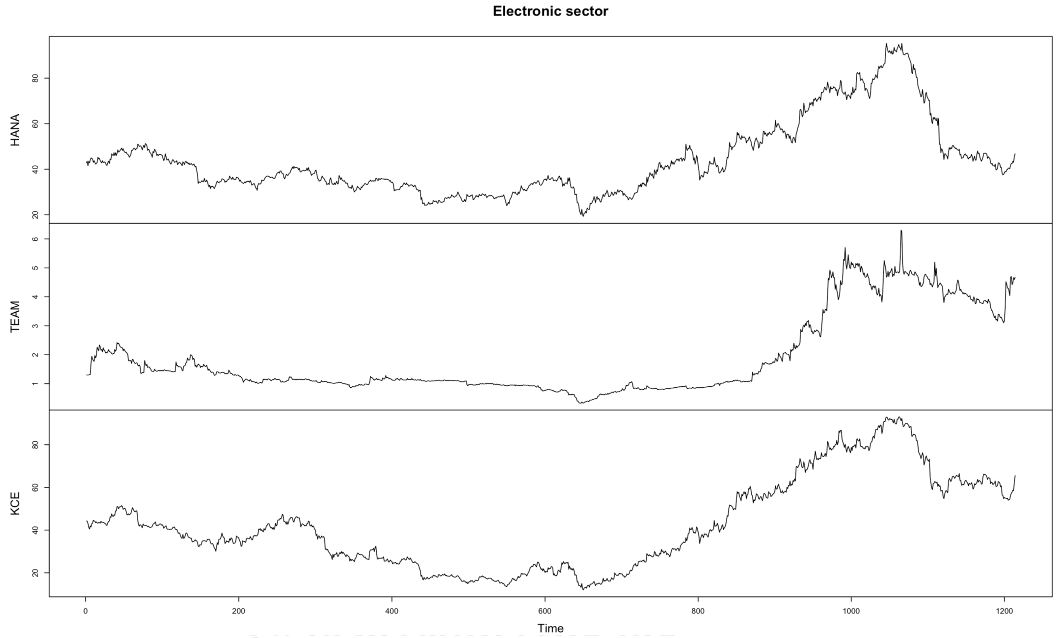


Figure 4.8: The plot of Electronic sector

Table 4.6: The correlation between the stock in Electronic sector

	HANA	TEAM	KCE
HANA	1.00	0.81	0.94
TEAM	0.81	1.00	0.88
KCE	0.94	0.88	1.00

From Figure 4.8, the plot shows that the data exhibits similar patterns and displays a correlation. To initiate the analysis, we apply the MVAR(n:K; p) model to the Energy sector. We explore values of K from 1 to 4, p from 1 to 4, and since the dataset includes

3 stocks, the n dimension is set to 3. The criteria values for each model are show in Table 4.7.

Table 4.7: The criteria for candidate MVAR models applied to Electronic sector data using the EM algorithm

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
MVAR(1:1)	-1380.55	-1344.06	-1283.65	MVAR(3:1)	-3560.35	-3450.90	-3269.65
MVAR(1:2)	-1375.10	-1321.34	-1232.32	MVAR(3:2)	-3539.36	-3378.09	-3111.03
MVAR(1:3)	-1387.67	-1316.64	-1199.03	MVAR(3:3)	-3506.46	-3293.37	-2940.54
MVAR(1:4)	-1367.90	-1279.60	-1133.41	MVAR(3:4)	-3446.61	-3181.72	-2743.15
MVAR(2:1)	-3367.23	-3294.27	-3173.43	MVAR(4:1)	-3694.24	-3548.30	-3306.64
MVAR(2:2)	-3352.07	-3244.55	-3066.51	MVAR(4:2)	-3747.04	-3532.00	-3175.93
MVAR(2:3)	-3177.15	-3035.09	-2799.87	MVAR(4:3)	-1867.85	-1583.73	-1113.29
MVAR(2:4)	-3282.16	-3105.57	-2813.19	MVAR(4:4)	-1718.10	-1364.92	-780.15

From Table 4.7, the MVAR(n:K; p) model, where the number of dimension vector, n, and the number of component K is equal to 1, MVAR(n:1; p), represents the original vector autoregressive model with order p in the first four lines while the other K components represent the MVAR models with multiple components. However, the MVAR(3:4;2) model exhibits the smallest AIC value while the HQIC and BIC criteria favor the MVAR(3:4;1) model. Therefore, considering the three criteria, two out of three indicate that the MVAR(3:4;1) model is the best model for Electronic sector. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting autocorrelation, quantile residual plots, Q-Q plots, and test normality test of the quantile residuals by using Kolmogorov Smirnov test is present in Figure 4.9 and Table 4.8.

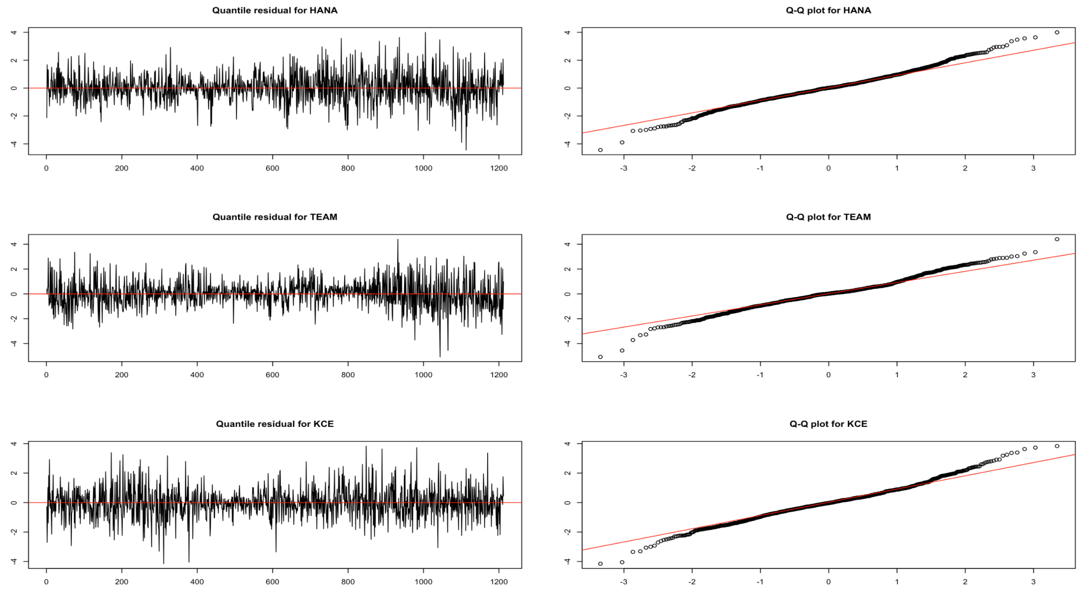


Figure 4.9: Quantile residual plot of MVAR(3:4:1) for Electronic sector

Table 4.8: Normality test of MVAR(3:4:1) for Electronic stock

Stock	Statistic	p-value
HANA	0.032871	0.1457
TEAM	0.045641	0.0128
KCE	0.05033	0.0043

From Figure 4.9, in the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, some outliers are observed. Furthermore, Figure 4.8, the Kolmogorov-Smirnov test for the quantile residuals of TEAM and KCE reveals p-values less than 0.05, indicating that the distribution of the given data does not conforms to a normal distribution, while HANA conform to a normal distribution.

4.2 t Mixture vector autoregressive model

The n dimensional vector time series Y_t is said to be the mixture vector autoregressive model denoted as TMVAR$(n : K; p_1, p_2, p_3, \ldots, p_k)$ if the distribution function of Y_t

given pass information can be written as

$$F(Y_t|\mathcal{F}_{t-1}) = \sum_{k=1}^{K} \alpha_k F_{v_k}(\Omega_k^{-1/2}(Y_t - \Theta_{k0} - \Theta_{k1}Y_{t-1} - \cdots - \Theta_{kp_k}Y_{t-p_k})), \tag{4.16}$$

where $F(Y_t|\mathcal{F}_{t-1})$ is the cumulative distribution function of Y_t given the past information $Y_{t-1}, Y_{t-2}, Y_{t-3}, \ldots, Y_1$, $F_{v_k}(\cdot)$ is the cumulative distribution function of the multivariate standardized t distribution with v_k degrees of freedom, Θ_{k0} is an n dimension vector, $\Theta_{k1}, \ldots, \Theta_{kp_k}$ are $n \times n$ coefficient matrices and Ω_k is the $n \times n$ variance covariance matrix for k^{th} component, the mixing proportion $\alpha_k > 0$, $k = 1, 2, 3, \ldots, K$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_K = 1$. The probability distribution function of a multivariate standardized t - distribution with unit variance is

$$f_v(\mathbf{X}) = \frac{\Gamma(\frac{v+n}{2})}{\pi^{\frac{n}{2}}(v-2)^{\frac{n}{2}}\Gamma(\frac{v}{2})}\Big(1 + \frac{1}{v-2}\mathbf{X}^T\mathbf{X}\Big)^{-\frac{v+n}{2}}, \tag{4.17}$$

where $\mathbf{X} = (X_1, \ldots, X_n)^T$ is a real random vector, $2 < v < \infty$, and $\Gamma(\cdot)$ is the gamma function.

4.2.1 Parameter estimation

In this section, we discuss the method we developed in this study to estimate parameters, namely the EM algorithm, and compare it with the maximum likelihood function.

4.2.1.1 Parameter estimation by maximum likelihood function

In the case of the t mixture vector autoregressive model, the maximum likelihood method is the parameter estimation method used in this study to compare with the EM algorithm. Specifically, given a time series $\mathbf{Y} = (Y_1, Y_2, Y_3, \ldots, Y_t)$, the likelihood function for the t mixture vector autoregressive model is the product of conditional density

$$L(\tilde{\Theta},\Omega,\alpha,v|Y_t)=\prod_{t=p+1}^{T}\sum_{k=1}^{K}\frac{\alpha_k|\Omega_k|^{-\frac{1}{2}}\Gamma(\frac{v+n}{2})}{\pi^{\frac{n}{2}}(v-2)^{\frac{n}{2}}\Gamma(\frac{v}{2})}\Big(1+\frac{1}{v-2}(Y_t-\mu_{kt})^T\Omega_k^{-1}(Y_t-\mu_{kt})\Big)^{-\frac{v+n}{2}}, \tag{4.18}$$

where Ω_k is a $n \times n$ variance covariance matrix for k^{th} component, and $\mu_{kt} = Y_t - \tilde{\Theta}_k X_{kt}$ is a $n \times n$ autoregressive matrices, $X_{kt} = (1, Y_{t-1}^T, Y_{t-2}^T)$, and $\tilde{\Theta}_k = [\Theta_{k0}, \Theta_{k1}, \ldots, \Theta_{kp_k}]$. The maximum likelihood estimate $\hat{\Theta}$ which is defined as

$$\hat{\Upsilon} = \arg\max_{\Upsilon} \ell(\Upsilon|\mathbf{Y}), \tag{4.19}$$

where $\Upsilon = (\tilde{\Theta}, \Omega, \alpha, v)$.

The estimation of parameters requires the use of a numerical technique [6]. In the context of the mixture components of the t mixture vector autoregressive model, finding a general solution may not be feasible [7]. The alternative to the parameter estimate is the EM algorithm.

4.2.1.2 Parameter estimation by the EM algorithm

The parameter estimation method used in this study is the EM algorithm. Assume that the n dimension vectors of observations Y_T are generated from TMVAR$(n, K; p)$ model for $t = 1, 2, \ldots, T$ and let $Z_t = (Z_{t,1}, Z_{t,2}, \ldots, Z_{kt})^T$, where

$$Z_{it} = \begin{cases} 1 & \text{if if } Y_t \text{ comes from the } i^{th} \text{ component component; } 1 \le i \le K, \\ 0 & \text{otherwise,} \end{cases}$$

and we consider another missing random variable matrix, $W = (W_1, W_2, \ldots, W_t)$, where $W_t = (W_{kt})$ for $t = 1, 2, 3, \ldots, n$ is also a K-dimensional vector. Given $Z_{kt} = 1$, the conditional distribution of W_{kt} is $W_{kt}|Z_{kt} = 1 \sim \text{gamma}(\frac{v_k}{2}, \frac{v_k - 2}{2})$, and $W_1, \ldots, W_n$ are distributed independently. The conditional loglikelihood function of the TMVAR model is

$$l = l_1(\alpha) + l_2(v) + l_3(\theta), \tag{4.20}$$

where

$$\begin{aligned}
l_1(\alpha) &= \sum_{k=1}^{K}\sum_{t=p+1}^{n} Z_{kt}\log(\alpha_k), && (4.21)\\
l_2(v) &= \sum_{k=1}^{K}\sum_{t=p+1}^{n} Z_{kt}\Big[-\log\Big\{\Gamma\Big(\frac{v_k}{2}\Big)\Big\} + \Big(\frac{v_k}{2}\Big)\log\Big(\frac{v_k-2}{2}\Big) \\
&\quad + \Big(\frac{v_k}{2}\Big)(\log W_{kt} - W_{kt}) + W_{kt} - \log(W_{kt})\Big], && (4.22)\\
l_3(\theta) &= \sum_{k=1}^{K}\sum_{t=p+1}^{n} Z_{kt}\Big(-\frac{1}{2}\{\log(2\pi) + \log\Omega_k - \log W_{kt}\} - \frac{{\mathbf{e_{kt}}}^T\mathbf{e_{kt}}W_{kt}}{2\Omega_k}\Big), && (4.23)
\end{aligned}$$

where $\mathbf{e_{kt}} = Y_t - \tilde{\Theta}_k X_{kt}$ and $X_{kt} = (1, Y_{t-1}^T, Y_{t-2}^T, \ldots, Y_{t-p_k}^T)^T$.

The parameters are estimated by iteratively maximizing the log-likelihood through the Expectation-Maximization(EM) procedure [3], which involves two main steps: the Expectation(E-step) and the Maximization(M-step). These steps are repeated iteratively until the algorithm converges. An illustration of the EM algorithm is provide in Figure 4.10.

The Expectation step. Assume that the parameters $\tilde{\Theta}, \Omega, \alpha, v$ is known. The unobserved random variable Z, the missing data W, and $\log W$ in the loglikelihood are replaced by their expectations, conditional over the parameters and the observed data $Y_1, \ldots, Y_T$. Let τ_{kt} be the conditional expectation of the k^{th} component of unobserved data Z. And then let η_{kt} be the conditional expectation of the k^{th} component of missing data W which defined as

$$\begin{aligned}
\tau_{kt} &= \frac{\alpha_k|\Omega_k|^{-\frac{1}{2}} f_{v_k}(\mathbf{e_{kt}}\Omega_k^{-1})}{\sum_{k=1}^{K}\alpha_k|\Omega_k|^{-\frac{1}{2}} f_{v_k}(e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}})}, && (4.24)\\
\eta_{kt} &= \frac{v_k+1}{e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}} + v_k - 2}. && (4.25)
\end{aligned}$$

The Maximization step. Suppose that the unobserved random variable, Z, and the missing random variable, W, is actually known. By maximizing the loglikelihood function (4.20), the estimates of our model are obtained through the first derivatives with respect to the parameters $\alpha_k, \tilde{\Theta}_k$, Ω_k and v_k are

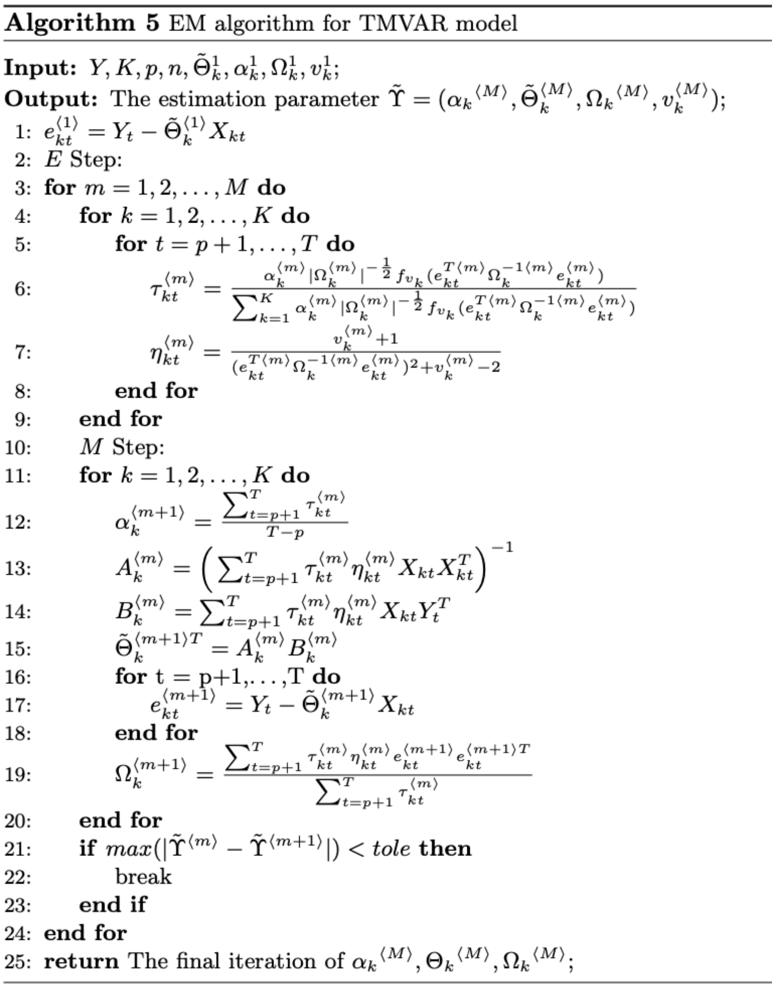
$$\hat{\alpha}_k = \frac{1}{T-p}\sum_{t=p+1}^{T}\tau_{kt}, \tag{4.26}$$

$$\hat{\tilde{\Theta}}_k^T = \Big(\sum_{t=p+1}^{T}\tau_{kt}\eta_{kt}X_{kt}X_{kt}^T\Big)^{-1}\Big(\sum_{t=p+1}^{T}\tau_{kt}\eta_{kt}X_{kt}Y_t^T\Big), \tag{4.27}$$

$$\hat{\Omega}_k = \frac{\sum_{t=p+1}^{T}\tau_{kt}\eta_{kt}\widehat{\mathbf{e}}_{\mathbf{kt}}\widehat{\mathbf{e}}_{\mathbf{kt}}^T}{\sum_{t=p+1}^{T}\tau_{kt}}, \tag{4.28}$$

where $k = 1, 2, \ldots, K$. The estimate of degree of freedom must satisfy the equations

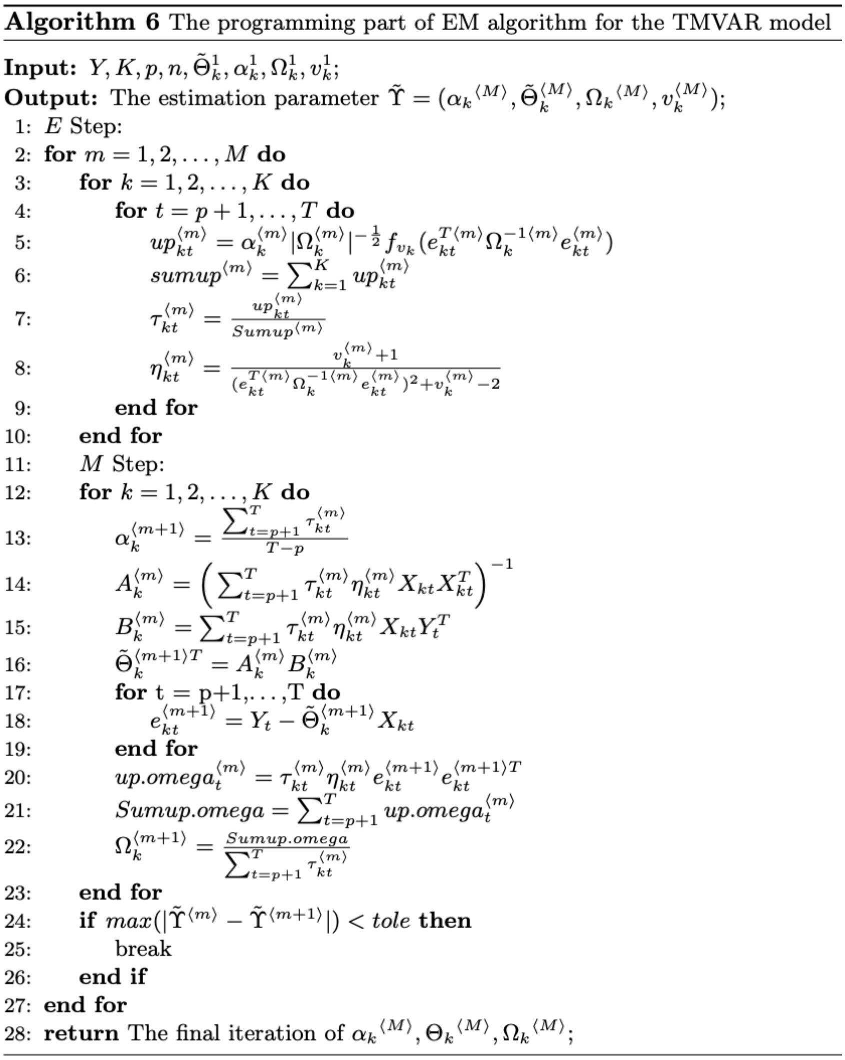
$$\begin{aligned}\Big(\frac{v_k}{v_k-2}\Big) + \log\Big(\frac{v_k-2}{2}\Big) - \psi\Big(\frac{v_k}{2}\Big) + \psi\Big(\frac{v_k^{(m)}+1}{2}\Big) - \log\Big(\frac{v_k^{(m)}+1}{2}\Big)\\ + \frac{1}{\sum_{t=p+1}^{n}\tau_{kt}^{(m)}}\sum_{t=p+1}^{n}\tau_{kt}^{(m)}\Big(\log(\eta_{kt}^{(m)}) - \eta_{kt}^{(m)}\Big) = 0.\end{aligned} \tag{4.29}$$



```
Algorithm 5 EM algorithm for TMVAR model
Input: $Y, K, p, n, \tilde{\Theta}_k^1, \alpha_k^1, \Omega_k^1, v_k^1$;
Output: The estimation parameter $\tilde{\Upsilon} = ({\alpha_k}^{\langle M\rangle}, \tilde{\Theta}_k^{\langle M\rangle}, {\Omega_k}^{\langle M\rangle}, v_k^{\langle M\rangle})$;
$e_{kt}^{\langle 1\rangle} = Y_t - \tilde{\Theta}_k^{\langle 1\rangle} X_{kt}$
$E$ Step:
for $m = 1, 2, \ldots, M$ do
    for $k = 1, 2, \ldots, K$ do
        for $t = p + 1, \ldots, T$ do
            $\tau_{kt}^{\langle m\rangle} = \frac{\alpha_k^{\langle m\rangle} |\Omega_k^{\langle m\rangle}|^{-\frac{1}{2}} f_{v_k}(e_{kt}^{T\langle m\rangle} \Omega_k^{-1\langle m\rangle} e_{kt}^{\langle m\rangle})}{\sum_{k=1}^{K} \alpha_k^{\langle m\rangle} |\Omega_k^{\langle m\rangle}|^{-\frac{1}{2}} f_{v_k}(e_{kt}^{T\langle m\rangle} \Omega_k^{-1\langle m\rangle} e_{kt}^{\langle m\rangle})}$
            $\eta_{kt}^{\langle m\rangle} = \frac{v_k^{\langle m\rangle}+1}{(e_{kt}^{T\langle m\rangle} \Omega_k^{-1\langle m\rangle} e_{kt}^{\langle m\rangle})^2 + v_k^{\langle m\rangle} - 2}$
        end for
    end for
    $M$ Step:
    for $k = 1, 2, \ldots, K$ do
        $\alpha_k^{\langle m+1\rangle} = \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}}{T-p}$
        $A_k^{\langle m\rangle} = \left( \sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle} \eta_{kt}^{\langle m\rangle} X_{kt} X_{kt}^T \right)^{-1}$
        $B_k^{\langle m\rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle} \eta_{kt}^{\langle m\rangle} X_{kt} Y_t^T$
        $\tilde{\Theta}_k^{\langle m+1\rangle T} = A_k^{\langle m\rangle} B_k^{\langle m\rangle}$
        for t = p+1,...,T do
            $e_{kt}^{\langle m+1\rangle} = Y_t - \tilde{\Theta}_k^{\langle m+1\rangle} X_{kt}$
        end for
        $\Omega_k^{\langle m+1\rangle} = \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle} \eta_{kt}^{\langle m\rangle} e_{kt}^{\langle m+1\rangle} e_{kt}^{\langle m+1\rangle T}}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}}$
    end for
    if $max(|\tilde{\Upsilon}^{\langle m\rangle} - \tilde{\Upsilon}^{\langle m+1\rangle}|) < tole$ then
        break
    end if
end for
return The final iteration of ${\alpha_k}^{\langle M\rangle}, {\Theta_k}^{\langle M\rangle}, {\Omega_k}^{\langle M\rangle}$;
```

Figure 4.10: The EM algorithm for the TMVAR model

The EM algorithm for the multivariate mixture autoregressive model that we mention and develop in Section 4.2.1.2 has the following steps in the programme, are show in Figure 4.11:



Algorithm 6 The programming part of EM algorithm for the TMVAR model

```
Input: $Y, K, p, n, \tilde{\Theta}_k^1, \alpha_k^1, \Omega_k^1, v_k^1$;
Output: The estimation parameter $\tilde{\Upsilon} = (\alpha_k{}^{\langle M \rangle}, \tilde{\Theta}_k^{\langle M \rangle}, \Omega_k{}^{\langle M \rangle}, v_k^{\langle M \rangle})$;
$E$ Step:
for $m = 1, 2, \ldots, M$ do
    for $k = 1, 2, \ldots, K$ do
        for $t = p+1, \ldots, T$ do
            $up_{kt}^{\langle m \rangle} = \alpha_k^{\langle m \rangle} |\Omega_k^{\langle m \rangle}|^{-\frac{1}{2}} f_{v_k}(e_{kt}^{T\langle m \rangle} \Omega_k^{-1\langle m \rangle} e_{kt}^{\langle m \rangle})$
            $sumup^{\langle m \rangle} = \sum_{k=1}^{K} up_{kt}^{\langle m \rangle}$
            $\tau_{kt}^{\langle m \rangle} = \frac{up_{kt}^{\langle m \rangle}}{Sumup^{\langle m \rangle}}$
            $\eta_{kt}^{\langle m \rangle} = \frac{v_k^{\langle m \rangle}+1}{(e_{kt}^{T\langle m \rangle} \Omega_k^{-1\langle m \rangle} e_{kt}^{\langle m \rangle})^2 + v_k^{\langle m \rangle} - 2}$
        end for
    end for
    $M$ Step:
    for $k = 1, 2, \ldots, K$ do
        $\alpha_k^{\langle m+1 \rangle} = \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}{T-p}$
        $A_k^{\langle m \rangle} = \left( \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} X_{kt} X_{kt}^T \right)^{-1}$
        $B_k^{\langle m \rangle} = \sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} X_{kt} Y_t^T$
        $\tilde{\Theta}_k^{\langle m+1 \rangle T} = A_k^{\langle m \rangle} B_k^{\langle m \rangle}$
        for t = p+1,...,T do
            $e_{kt}^{\langle m+1 \rangle} = Y_t - \tilde{\Theta}_k^{\langle m+1 \rangle} X_{kt}$
        end for
        $up.omega_t^{\langle m \rangle} = \tau_{kt}^{\langle m \rangle} \eta_{kt}^{\langle m \rangle} e_{kt}^{\langle m+1 \rangle} e_{kt}^{\langle m+1 \rangle T}$
        $Sumup.omega = \sum_{t=p+1}^{T} up.omega_t^{\langle m \rangle}$
        $\Omega_k^{\langle m+1 \rangle} = \frac{Sumup.omega}{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m \rangle}}$
    end for
    if $max(|\tilde{\Upsilon}^{\langle m \rangle} - \tilde{\Upsilon}^{\langle m+1 \rangle}|) < tole$ then
        break
    end if
end for
return The final iteration of $\alpha_k{}^{\langle M \rangle}, \Theta_k{}^{\langle M \rangle}, \Omega_k{}^{\langle M \rangle}$;
```

Figure 4.11: The programming part of the EM algorithm for the TMVAR model

In this section, we will show the calculation example by using the TMVAR_EM program for simulation. The model consider in this example is the TMVAR(2:2;1) model with generate length of each time series is 10 data points, a dimension vector of 2, the number of component, K, is 2, and the order of autoregressive is 1. The parameter M represents the number of iterations in the EM algorithm.

In the initial step, we generate time series data comprising 10 data points. For the purpose of this illustration, the calculation example is presented with m iterations limited to 1.

	$Y_{T,1}$	$Y_{T,2}$
1	-0.13	2.87
2	-0.44	2.30
3	-1.01	2.27
4	0.48	2.53
5	1.85	2.06
6	2.09	1.99
7	1.66	2.26
8	0.11	2.25
9	0.91	2.35
10	-0.13	1.93

and the initial values for the EM algorithm are

$$\begin{aligned}
\alpha &= [\alpha_1, \alpha_2] \\
&= [0.7, 0.3], \\
\tilde{\Theta}_1 &= \begin{bmatrix} \Theta_{10} & \Theta_{11} \end{bmatrix} \\
&= \begin{bmatrix} 1.568 & 0.224 & -1.509 \\ 0.938 & 0.108 & -0.492 \end{bmatrix}, \\
\tilde{\Theta}_2 &= \begin{bmatrix} \Theta_{20} & \Theta_{21} \end{bmatrix} \\
&= \begin{bmatrix} 0.130 & -0.279 & 0.497 \\ 1.247 & 0.030 & -0.495 \end{bmatrix}, \\
\Omega_1 &= \begin{bmatrix} 0.327 & 0.016 \\ 0.016 & 0.009 \end{bmatrix}, \\
\Omega_2 &= \begin{bmatrix} 0.097 - 0.074 \\ -0.0740.057 \end{bmatrix}, \\
v &= [v_1, v_2] \\
&= [7.568, 9.33].
\end{aligned}$$

In the initial step of the computation in the Expectation step, we calculate $\mathbf{e_{kt}}^{\langle 1 \rangle} = Y_t - \tilde{\Theta}_k X_{kt}$ where $X_{kt} = (1, Y_{t-1}^T, Y_{t-2}^T)$:

$$\mathbf{e_{1t}}^{\langle 1\rangle} = \begin{bmatrix} NA & 0.40 & 1.38 & -0.85 & 0.53 & -0.46 & 1.11 & -0.94 & 0.02 & -0.21 \\ NA & 0.38 & -0.17 & -0.22 & -0.40 & -0.52 & -0.74 & -0.67 & -0.37 & -0.25 \end{bmatrix}.$$

In the Expectation step we have to compute $\tau_{kt} = \frac{\alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}})}{\sum_{k=1}^K \alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}})}$. Begin with $K = 1, 2$ and $t = p + 1, \ldots, T$. In this example, we illustrate the case of $t = p + 1 = 3$ to T and for all k.

$$\text{up}_{kt}^{\langle m\rangle} = \alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}}) = \begin{bmatrix} NA & NA \\ 0.06 & 0.00 \\ 0.15 & 0.00 \\ 1.21 & 0.11 \\ 0.02 & 3.41 \\ 0.01 & 0.67 \\ 0.00 & 0.00 \\ 0.00 & 0.00 \\ 0.05 & 0.00 \\ 0.74 & 0.00 \end{bmatrix},$$

$$\eta_{kt} = \frac{v_k+1}{(e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}})^2+v_k-2} = \begin{bmatrix} NA & NA \\ 0.41 & 0.02 \\ 0.54 & 0.00 \\ 0.86 & 0.33 \\ 0.28 & 0.71 \\ 0.18 & 0.52 \\ 0.03 & 0.00 \\ 0.08 & 0.01 \\ 0.39 & 0.10 \\ 0.78 & 0.06 \end{bmatrix},$$

$$\begin{aligned} \text{sumup}^{\langle m\rangle} &= \sum_{k=1}^{K} \alpha_k|\Omega_k|^{-\frac{1}{2}}exp(-\frac{1}{2}e_{kt}^T\Omega_k^{-1}\mathbf{e_{kt}}) \\ &= (0.000, 0.061, 0.153, 1.316, 3.425, 0.675, 0.000, 0.001, 0.053, 0.737), \end{aligned}$$

$$\tau_{kt}^{\langle m\rangle} = \frac{\text{up}_{kt}^{\langle m\rangle}}{\text{sumup}^{\langle m\rangle}} = \begin{bmatrix} NA & NA \\ 1.00 & 0.00 \\ 1.00 & 0.00 \\ 0.92 & 0.08 \\ 0.01 & 0.99 \\ 0.01 & 0.99 \\ 1.00 & 0.00 \\ 0.99 & 0.01 \\ 0.95 & 0.04 \\ 1.00 & 0.00 \end{bmatrix}.$$

In the Maximization Step: we need to estimate $\alpha_k^{\langle m+1\rangle}, \tilde{\Theta}_1^{\langle m+1\rangle T}$, and $\Omega_k^{\langle m+1\rangle}$. In this example, we illustrate the case of $t = p+1 = 3$ to T and for all k.

$$
\begin{aligned}
\alpha_k^{\langle m+1\rangle} &= \frac{\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}}{T-p} \\
&= [\alpha_1, \alpha_2] \\
&= [0.848, 0.152].
\end{aligned}
$$

When estimating $\tilde{\Theta}_k^{\langle m+1\rangle T}$, we need to construct the matrix equation $A_k^{\langle m\rangle T}\tilde{\Theta}_k^{\langle m+1\rangle T} = B_k^{\langle m+1\rangle T}$, where

$$
\begin{aligned}
A_k^{\langle m\rangle} &= \Big(\sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}\eta_{kt}^{\langle m\rangle} X_{kt}X_{kt}^T\Big)^{-1}, \\
B_k^{\langle m\rangle} &= \sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}\eta_{kt}^{\langle m\rangle} X_{kt}Y_t^T,
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\Theta}_1^{\langle m+1\rangle T} &= A_1^{\langle m\rangle} B_1^{\langle m\rangle} \\
&= \begin{bmatrix} 0.42 & -0.02 & 0.24 \\ 0.52 & 0.02 & 0.23 \end{bmatrix}.
\end{aligned}
$$

Before estimating $\Omega_k^{\langle m+1\rangle}$, we update the term $\tilde{\Theta}_1^{\langle m+1\rangle T}$ to $\mathbf{e_{1t}}^{\langle m+1\rangle}$.

$$
\begin{aligned}
\mathbf{e_{1t}}^{\langle m+1\rangle} &= Y_t - \tilde{\Theta}_k X_{kt} \\
&= \begin{bmatrix} NA & -0.13 & 1.10 & 0.03 & 0.55 & 0.53 & 1.76 & 0.60 & 0.59 & 0.43 \\ NA & 0.10 & -0.34 & 0.07 & -0.44 & -0.17 & -0.54 & -0.11 & -0.20 & -0.05 \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\text{Sumup.omega} &= \sum_{t=p+1}^{T} \tau_{kt}^{\langle m\rangle}\eta_{kt}^{\langle m\rangle} e_{kt}^{\langle m+1\rangle} e_{kt}^{\langle m+1\rangle T} \\
&= \begin{bmatrix} 4.965-0.188 \\ -0.1880.136 \end{bmatrix}, \\
\Omega_k^{\langle m+1\rangle} &= \Big(\frac{\text{Sumup.omega}}{\sum_{t=p+1}^{T}\tau_{kt}^{\langle m\rangle}}\Big)^{\frac{1}{2}} \\
&= \begin{bmatrix} 0.327 & 0.016 \\ 0.016 & 0.009 \end{bmatrix}.
\end{aligned}
$$

The estimate of degree of freedom must satisfy the equations fn_k, which is defined as

$$
\begin{aligned}
fn_k &= \Big(\frac{v_k}{v_k-2}\Big) + \log\Big(\frac{v_k-2}{2}\Big) - \psi\Big(\frac{v_k}{2}\Big) + \psi\Big(\frac{v_k^{(m)}+1}{2}\Big) - \log\Big(\frac{v_k^{(m)}+1}{2}\Big) \\
&\quad + \frac{1}{\sum_{t=p+1}^{n}\tau_{kt}^{(m)}} \sum_{t=p+1}^{n} \tau_{kt}^{(m)}\Big(\log(\eta_{kt}^{(m)}) - \eta_{kt}^{(m)}\Big), \\
fd_k &= \frac{v_k-4}{(v_k-2)^2} - \frac{d^2}{dv_k^2}\Gamma\Big(\frac{v_k}{2}\Big),
\end{aligned}
$$

where, $v_k^{(m)}$ represents the estimated v_k in the m^{th} iteration of the EM algorithm. This estimation is employed to obtain a numerical solution using the Newton-Raphson method following

$$
\begin{aligned}
v_k^{\langle m+1\rangle} &= v_k^{\langle m\rangle} - \frac{fn_k}{fd_k} \\
&= [4.223, 5.674].
\end{aligned}
$$

For now, we complete the m^{th} iteration, where $m = 1$. We then repeat the Expectation step and Maximization step until the parameter estimate $\tilde{\Upsilon} = ({\alpha_k}^{\langle M\rangle}, \tilde{\Theta}_k^{\langle M\rangle}, {\Omega_k}^{\langle M\rangle})$ convergence by using $max(|\tilde{\Upsilon}^{\langle m\rangle} - \tilde{\Upsilon}^{\langle m+1\rangle}|) < tol$.

4.2.2 Simulation study for TMVAR model

In this section, we study the performance of parameter estimation using the maximum likelihood estimation procedure implemented in the “gmvarkit”[6] in Section 4.1.1.1. We examine the correctness in choosing the number of components, K, and the, p, order of the autoregressive models. Furthermore, we examine the accuracy of parameter estimates. It’s important to note that, under the restrictions of the package, the orders of autoregressive components for different components are assumed to be the same. Therefore, the models considered in this study are denoted as TMVAR(n:K; p), where K is the number of components and p is the common order of autoregressive components. The model investigated in this study are the MAR(2:2; 1), where n dimensional vector is 2, K component is 2, and order p is 1.

For the experiments, we generate a time series from the TMVAR(2:2;1) which the dimension is 2, the number of component is 2, order of autoregressive model is 1 with a time length of 1000 data points and simulation 1000 replications. In comparing parameter estimates between the Expectation-Maximization(EM) with the Maximum Likelihood Estimation(MLE). The EM algorithm utilized parameters from MLE as an initial value. The best candidate models, identified based on the smallest corresponding criterion, were determined to be TMVAR(2:2;1) models, which were correctly chosen. Table 4.9 presents the parameter estimation for TMVAR(2:2;1) models, comparing the exact values of parameters to the mean of estimates for each method, along with the error of each method, respectively.

Table 4.9: Parameter estimates for the TMVAR(2:2;1) model using the EM algorithm.

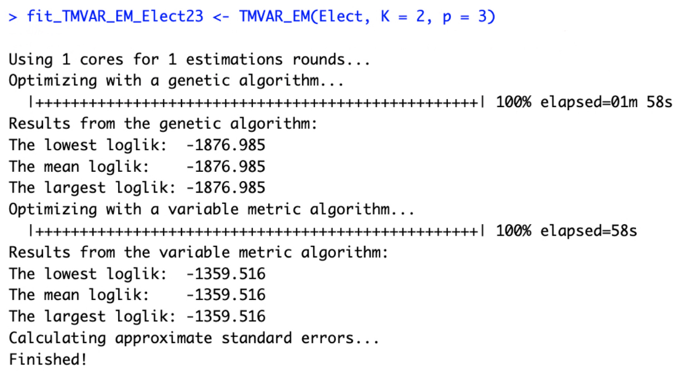
	$\phi_{10,1}$	$\phi_{20,1}$	$\phi_{11,1}$	$\phi_{21,1}$	$\phi_{12,1}$	$\phi_{22,1}$
Exact value	0.545	0.116	0.331	0.054	-0.042	0.709
Estimate of EM	0.588	0.141	0.298	0.049	-0.040	0.67
Estimate of MLE	0.493	0.073	0.275	0.035	-0.041	0.760
Error of EM	0.043	0.025	0.033	0.005	0.002	0.031
Error of MLE	0.052	0.043	0.056	0.019	0.001	0.051
	$\phi_{10,2}$	$\phi_{20,2}$	$\phi_{11,2}$	$\phi_{21,2}$	$\phi_{12,2}$	$\phi_{22,2}$
Exact value	1.598	0.483	0.126	-0.031	-0.613	0.723
Estimate of EM	1.218	0.453	0.107	0.003	-0.399	0.613
Estimate of MLE	1.198	0.289	0.012	-0.016	-0.470	0.818
Error of EM	0.380	0.030	0.019	0.034	0.214	0.110
Error of MLE	0.400	0.194	0.114	0.015	0.143	0.095
	$\Omega_{11,1}$	$\Omega_{21,1}$	$\Omega_{22,1}$	$\Omega_{11,2}$	$\Omega_{21,2}$	$\Omega_{22,2}$
Exact value	0.418	0.002	0.041	1.212	-0.036	0.138
Estimate of EM	0.318	0.003	0.033	0.748	-0.013	0.094
Estimate of MLE	0.291	0.001	0.031	0.595	-0.001	0.047
Error of EM	0.100	0.001	0.008	0.464	0.023	0.044
Error of MLE	0.127	0.001	0.010	0.617	0.035	0.091
	α_1	α_2	υ_1	υ_2		
Exact value	0.834	0.166	7.568	9.33		
Estimate of EM	0.787	0.129	6.932	8.546		
Estimate of MLE	0.834	0.166	972.902	13307.56		
Error of EM	0.047	0.037	0.636	0.784		
Error of MLE	0	0	965.334	13298.23		

From Table 4.9, the results of the parameter estimates for the TMVAR (2:2;1) models compare between the Maximum Likelihood Estimation and the EM algorithms. In the part α, the parameter estimates are quite close to the exact values by using the maximum likelihood estimation. In the other parameter, the performance of the EM algorithm is good in 15 out of the 22 parameters base on the autoregressive coefficients ϕ_{ki}, standard deviation Ω_k, and degree of freedom v_k. Furthermore, we compute the mean square error (MSE) for the parameter estimates obtained from both the EM algorithm and the maximum likelihood estimation(MLE). The resulting MSE values are 0.891 for the EM algorithm and 1.046 for the MLE. Consequently, judging from the parameter errors and MSE values, it is evident that the EM algorithm outperforms MLE.

4.2.3 t Mixture vector autoregressive model for Thai stock market data

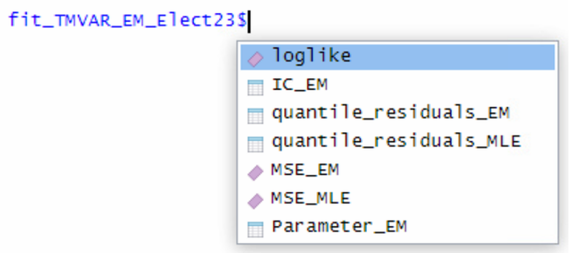
In this section, we are up to analysing a dataset, which we refer to as the sector dataset. We apply the TMVAR model to analyze the sector dataset including three stocks in the energy and utility sectors, as well as the electronic sector that we mention in Section 4.1.3. The goal is to explore the correlation within each dataset. In the program, following the algorithm outlined in Figure 4.11, we develop a function called TMVAR_EM($\mathbf{y_{T\times n}}$, K, p, tol). This function takes input parameters such as the data $\mathbf{y_{T\times n}}$, where T is the length of the data points and n is the number of dimensional time series, the number of components (K), autoregressive order (p), and tolerance (tol, with a default value of 1×10^{-6}). The initial parameter values for this function are obtained through maximum likelihood estimation.

The following code fits an TMVAR model to the energy data, using two components and an autoregressive order of 3 mixture components, as illustrated in Figure 4.12. The TMVAR_EM function returning a list of elements including the information criteria (IC), log-likelihood, quantile residuals for both Maximum Likelihood Estimation(MLE) and Expectation-Maximization(EM), as well as mean square error for MLE and EM, as illustrate in Figure 4.13. For instance, Figure 4.14 displays some elements from the TMVAR_EM function, which is the information criteria.



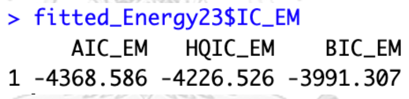
```
> fit_TMVAR_EM_Elect23 <- TMVAR_EM(Elect, K = 2, p = 3)

Using 1 cores for 1 estimations rounds...
Optimizing with a genetic algorithm...
  |++++++++++++++++++++++++++++++++++++++++++++++++++| 100% elapsed=01m 58s
Results from the genetic algorithm:
The lowest loglik:  -1876.985
The mean loglik:    -1876.985
The largest loglik: -1876.985
Optimizing with a variable metric algorithm...
  |++++++++++++++++++++++++++++++++++++++++++++++++++| 100% elapsed=58s
Results from the variable metric algorithm:
The lowest loglik:  -1359.516
The mean loglik:    -1359.516
The largest loglik: -1359.516
Calculating approximate standard errors...
Finished!
```

Figure 4.12: The TMVAR_EM function for Electronic sector



```
fit_TMVAR_EM_Elect23$
```

Figure 4.13: The list of elements in the TMVAR_EM function



```
> fitted_Energy23$IC_EM
    AIC_EM    HQIC_EM     BIC_EM
1 -4368.586 -4226.526 -3991.307
```

Figure 4.14: The information criteria from EM function

We analyze the energy sector data, which includes three stocks: BANPU, ESSO, and BCP that mentioned in Section 4.1.3. The stock plots for this sector show in Figure 4.6, exhibit a similar pattern, and the corresponding correlation values, which are presented in Table 4.3.

Since the data exhibits similar patterns and displays a correlation. We apply the TMVAR(n:K; p) model to the Energy sector and explore values of K from 1 to 4, p from 1 to 4, and since the dataset includes 3 stocks, the number of dimension, n, is set to 3. The criteria values for each model are present in Table 4.10.

Table 4.10: The criteria for candidate TMVAR models applied to Energy sector data using the EM algorithm

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMVAR(1:1)	-4211.82	-4173.42	-4109.82	TMVAR(3:1)	-3758.02	-3642.81	-3452.02
TMVAR(1:2)	-4198.04	-4142.36	-4050.16	TMVAR(3:2)	-4415.12	-4248.08	-3971.49
TMVAR(1:3)	-4192.60	-4119.65	-3998.86	TMVAR(3:3)	-4376.17	-4157.32	-3794.96
TMVAR(1:4)	-4179.38	-4089.16	-3939.79	TMVAR(3:4)	-4326.34	-4055.69	-3607.59
TMVAR(2:1)	-4508.71	-4431.90	-4304.70	TMVAR(4:1)	-4428.57	-4274.95	-4020.56
TMVAR(2:2)	-4472.65	-4361.29	-4176.90	TMVAR(4:2)	-4356.69	-4133.97	-3765.18
TMVAR(2:3)	-4451.27	-4305.37	-4063.79	TMVAR(4:3)	-3063.24	-2771.44	-2288.29
TMVAR(2:4)	-4418.94	-4238.51	-3939.77	TMVAR(4:4)	-3869.609	-3508.44	-2911.27

From Table 4.10, the TMVAR(n:K;p) model, where the number of dimension vector, n, and the number of component K is equal to 1 ,TMVAR(n:1;p), represents the original vector autoregressive model with order p in the first four lines while the other K components represent the TMVAR models with multiple components. However, the TMVAR(3:2;1) model exhibits the smallest AIC, HQIC, and BIC value. Therefore, considering the three criteria, indicate that the TMVAR(3:2;1) model is the best model. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting autocorrelation, quantile residual plots, Q-Q plots, and test normality test of the quantile residuals by using Kolmogorov Smirnov test is present in Figure 4.15 and Table 4.11.

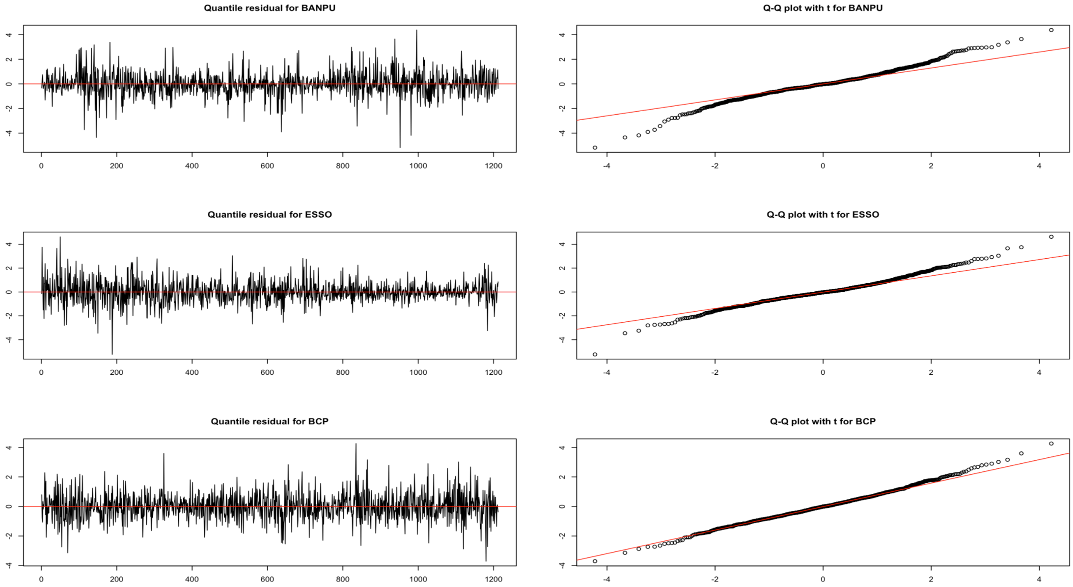


Figure 4.15: Quantile residual plot of TMVAR(3:2:1) for Energy sector

Table 4.11: t distribution test of TMVAR(3:4:1) for Energy stock

Stock	Statistic	p-value
BANPU	0.081831	1.785e-07
ESSO	0.081713	1.87e-07
BCP	0.057935	0.0005855

From Figure 4.15, in the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, some outliers are observed. Furthermore, in Figure 4.11, the Kolmogorov-Smirnov test for the quantile residuals of BANPU, ESSO, and BCP reveals p-values less than 0.05, indicating that the distribution of the given data does not conform to a t distribution.

Finally, we analyze the energy sector data, which includes three stocks: HANA, TEAM, and KCE that mentioned in Section 4.1.3. The stock plots for this sector show in Figure 4.8, exhibit a similar pattern. We then examine the corresponding correlation values, which are presented in Table 4.6. Since the data exhibits similar patterns and displays a correlation. We apply the TMVAR(n:K; p) model to the Electronic sector and explore values of K from 1 to 4, p from 1 to 4, and since the dataset includes 3 stocks, the number of dimension, n, is set to 3. The criteria values for each model are present in Table 4.12.

Table 4.12: The criteria for candidate TMVAR models applied to Electronic sector data using the EM algorithm

Models	AIC	HQIC	BIC	Models	AIC	HQIC	BIC
TMVAR(1:1)	-834.8752	-796.4713	-732.8747	TMVAR(3:1)	-1356.374	-1241.162	-1050.373
TMVAR(1:2)	-839.6012	-783.9222	-691.724	TMVAR(3:2)	-1396.63	-798.506	-1197.59
TMVAR(1:3)	-868.9479	-795.9981	-675.209	TMVAR(3:3)	-680.1413	-693.8274	-733.4976
TMVAR(1:4)	-859.7261	-769.5096	-620.1413	TMVAR(3:4)	-1204.38	-1186.249	-998.632
TMVAR(2:1)	-1396.864	-1320.056	-1192.863	TMVAR(4:1)	-1316.079	-1162.463	-908.0765
TMVAR(2:2)	-1380.917	-1269.56	-1085.164	TMVAR(4:2)	-1055.824	-967.664	-1129.342
TMVAR(2:3)	-1385.094	-1239.195	-997.6177	TMVAR(4:3)	-1172.976	-884.726	-1014.760
TMVAR(2:4)	-1353.939	-1173.507	-874.7699	TMVAR(4:4)	-659.4661	-769.5096	-850.1413

From Table 4.12, the TMVAR(n:K; p) model, where the number of dimension vector, n, and the number of component K is equal to 1 ,TMVAR(n:1; p), represents the original vector autoregressive model with order p in the first four lines while the other K components represent the TMVAR models with multiple components. However, the TMVAR(3:2;1) model exhibits the smallest AIC, HQIC, and BIC value. Therefore, considering the three criteria, indicate that the TMVAR(3:2;1) model is the best model. Next, the diagnostic check involves quantile residuals, which are used to perform computationally simple tests aimed at detecting autocorrelation, quantile residual plots, Q-Q plots, and test normality test of the quantile residuals by using Kolmogorov Smirnov test is present in Figure 4.16 and Table 4.13.

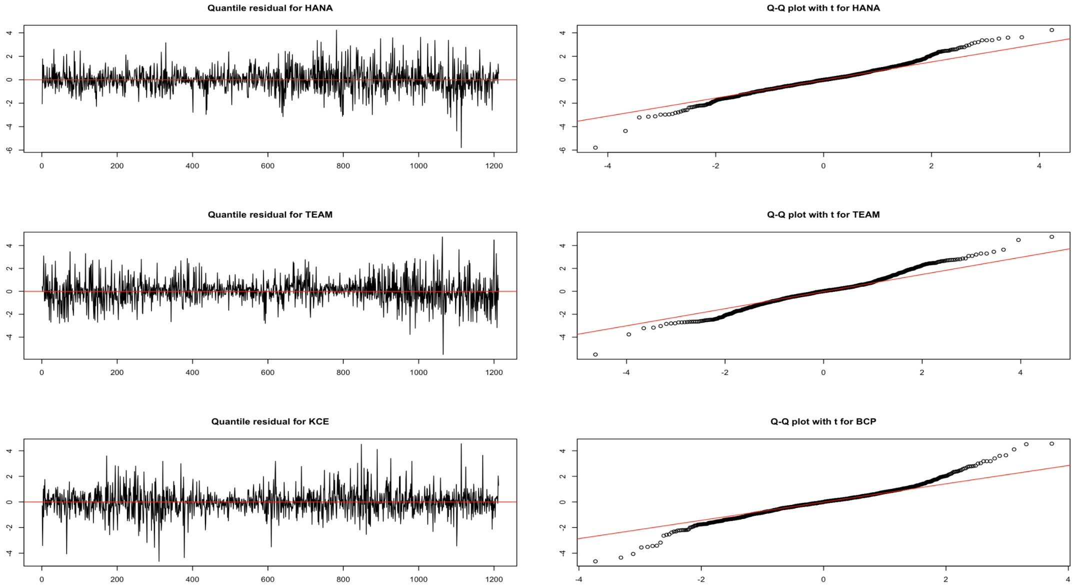


Figure 4.16: Quantile residual plot of TMVAR(3:2:1) for Electronic sector

Table 4.13: Normality test of TMVAR(3:2:1) for Electronic stock

Stock	Statistic	p-value
HANA	0.076517	1.372e-06
TEAM	0.23319	2.2e-16
KCE	0.072948	5.001e-06

From Figure 4.16, in the quantile residual analysis, the quantile residual plot is randomly dispersed around 0. While part of the Q-Q plot follows a diagonal line, some outliers are observed. Furthermore, in Figure 4.13, the Kolmogorov-Smirnov test for the quantile residuals of HANA, TEAM, and KCE reveals p-values less than 0.05, indicating that the distribution of the given data does not conform to a *t* distribution.

From two different sector dataset, we select the best model using the AIC, HQIC, and BIC that we mention in section 2.3. We assessed the validity of the best model, which using the EM algorithm to estimate parameter, through model diagnostic and compare the mean square error(MSE) for each stock data, as show in Table 4.14.

Table 4.14: The mean square error best candidate of the multivariate mixture autoregressive model for the Thai stock market

	Model	MSE	Model	MSE
BANPU	MVAR(3:4;1)	1.092	TMVAR(3:2;1)	0.893
ESSO		1.049		0.806
BCP		1.067		0.809
HANA	MVAR(3:4;1)	1.080	TMVAR(3:2;1)	1.075
TEAM		1.049		1.136
KCE		1.067		1.03

From Table 4.14, when comparing the mean square error (MSE) of the mixture vector autoregressive model with the t mixture vector autoregressive model, the mean square error (MSE) from the multivariate mixture autoregressive model based on the *t* distribution outperforms, which suitable for data exhibiting heavy tails such as stock market data.

In this chapter, we construct the family of multivariate mixture autoregressive models, which includes the mixture vector autoregressive (MVAR) model, the t mixture vector autoregressive (TMVAR) model using maximum likelihood estimation (MLE) to estimate parameters, and the t mixture autoregressive model using the EM algorithm for parameter estimation. We conduct a simulation study to test the accuracy of the method which is the EM algorithm is preferred over the MLE and then apply it to Thai stock market data. For the MVAR model, all criteria for each stock indicate that the multiple component model is better than the single component model. However, almost the entire residual of the model does not follow a normal distribution. Consequently, the alternative distribution, the t mixture autoregressive model, is considered, which is suitable for data exhibiting heavy tails, such as stock market data. In Table 4.14, the TMVAR model, utilizing the EM algorithm developed in Section 4.2.1.2 to estimate parameters, is preferred over the mixture autoregressive model.

CHAPTER V

CONCLUSIONS AND FUTURE WORK

In this chapter, we discuss and conclude this thesis, encompassing the family of univariate mixture autoregressive models and multivariate mixture autoregressive models based on normal and t distributions. The chapter examines the performance of the EM algorithm we constructed, compares it with the MLE, and explores the application of each model to Thai stock market data.

5.1 Conclusions

In Chapter I, we delved into the background of time series models and the concept of mixture distributions. In 1996, Le et al. introduced the class of mixture Gaussian transition distribution(GMTD) models [1]. Wong and Li [2] later extended these concepts, introducing a new class of the mixture models known as the mixture autoregressive(MAR) model in 2000. Following this, Meitz, Virolainen, and Savi [5] further extended the model to a mixture autoregressive model based on Student's t distribution. They developed the "uGMAR" R-package, which provides tools for estimating and analyzing the mixture autoregressive model base on normal distribution.

In Chapter II, we studied the time series and stochastic processes, exploring concepts such as stationarity, time series models including both stationary and non-stationary models, the vector of time series models, parameter estimation employing the maximum likelihood estimator and the Expectation-Maximization algorithm, model diagnostics, model selection criteria, and distributions, encompassing the normal distribution, the t distribution, and the finite mixture distribution.

In Chapter III, we introduced the family of the univariate mixture autoregressive model, including the mixture autoregressive(MAR) model, the t mixture autoregressive(TMAR) model, and the t mixture autoregressive model using the EM algorithm to estimate parameters, simulated a simulation study and test the accuracy of the model, and then applied it to the Thai stock market data. In the mixture autoregressive(MAR) model. The AIC, HQIC, and BIC values for each stock suggest that the multiple component model is better than the single component model, but almost the entire residual of the model does not follow a normal distribution. The t mixture autoregressive model is the alternative model, which is satisfied for data that has a heavy tail, such as stock market data. In which the first two models in Section 3.1 and 3.2 use maximum likelihood to estimate the parameter. After that, we developed the program for the TMAR model using the EM algorithm to estimate parameters in Section 3.2.1.2. The summary of the family of univariate mixture autoregressive models is shown in Table 3.37. The t-mixture autoregressive model, in which the maximum likelihood estimator and EM algorithm that we developed for the TMAR model are used to estimate the parameter, outperforms the mixture autoregressive model.

In Chapter IV, we introduced a family of multivariate mixture autoregressive models, including the mixture vector autoregressive (MVAR) model and the t mixture vector autoregressive (TMVAR) model, using the EM algorithm that we constructed to estimate parameters, simulate a simulation study, test the accuracy of the model, and then apply it to Thai stock market data. In the multivariate mixture autoregressive models, the information criteria for each stock suggest that the multiple component model is better than the single component model, but almost the entire residual of the model does not follow a normal distribution. The t mixture autoregressive model is the alternative model, which is satisfied for data that has a heavy tail, such as stock market data. The summary of the family of multivariate mixture autoregressive models is shown in Table 4.14. The t mixture vector autoregressive model, in which using the EM algorithm are used to estimate the parameter, outperforms the mixture vector autoregressive model.

5.2 Future work

Some directions of future work can be done such as developing the program based on the independence order of the autoregressive model. Another direction is to extend the mixture model of time series, which does not have a constant variance.

REFERENCES

[1] N. D. Le, R. D. Martin, and A. E. Raftery, "Modeling flat stretches, bursts outliers in time series using mixture transition distribution models," *Journal of the American Statistical Association*, vol. 91, no. 436, pp. 1504–1515, 1996.

[2] C. S. Wong and W. K. Li, "On a mixture autoregressive model," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 62, no. 1, pp. 95–115, 2000.

[3] P. W. Fong, W. K. Li, C. Yau, and C. S. Wong, "On a mixture vector autoregressive model," *Canadian Journal of Statistics*, vol. 35, no. 1, pp. 135–150, 2007.

[4] M. Lanne and P. Saikkonen, "Modeling the us short-term interest rate by mixture autoregressive processes," *Journal of Financial Econometrics*, vol. 1, no. 1, pp. 96–125, 2003.

[5] S. Virolainen and M. S. Virolainen, "Package 'ugmar'," 2023.

[6] S. Virolainen and M. S. Virolainen, "Package 'gmvarkit'," 2023.

[7] G. J. McLachlan and D. Peel, "Mixtures of factor analyzers," in *Proceedings of the seventeenth international conference on machine learning*, pp. 599–606, 2000.

[8] L. Kalliovirta, "Misspecification tests based on quantile residuals," *The Econometrics Journal*, vol. 15, no. 2, pp. 358–393, 2012.

[9] C. S. Wong, W.-S. Chan, and P. Kam, "A student t-mixture autoregressive model with applications to heavy-tailed financial data," *Biometrika*, vol. 96, no. 3, pp. 751–760, 2009.

Author Profile

Mr. Apicha Suthichayapipat ID 637 00974 23

Applied Mathematics and Computational Science

Department of Mathematics and Computer Science

Faculty of Science

Chulalongkorn university