A Maximum Likelihood Approach towards Aggregating Mixture Models: A Financial Time Series Analysis

Seuk-Yen Phoong1* and Seuk-Wai Phoong2

EM algorithm is an iterative procedure of maximum likelihood estimation which essential in finding parameter estimates. The procedures in EM algorithm will iterated until it converges to a local maximum and increase the accuracy of the log-likelihood results in each of the iterations. The EM algorithm is used in this study to analysis the financial time series data since most of the financial data involves latent variables. The correlations between returns of stock market price, gold price and exchange rate are examined in this study since there is the symbol of economic growth of a country but with different volatility. A two-component normal mixture model is used to examine the time series data via EM algorithm. Results found that all variables are correlated and shows positive relationship between each other.

JEL Codes: C58

1. Introduction

Exchange rate play essential role as a symbol of development of economy for a country. Also, exchange rate is regarded as the rate of currency for a country exchanged to another foreign currency. Fluctuations in exchange rate influence the development of a country and the country’s relative level of trade. Therefore, the smaller scale for exchange rate denotes better.

Exchange rate is a financial asset price that confront with remarkable changes every day. According to Hopper (1997), there are five factors that influence the changes of exchange rate, including interest and inflation rates, trade balance, exchange rate, market speculation, central bank actions and foreign investment. The first factor explained that the interest rates, inflation and exchange rates are highly correlated. This indicates that higher value of the interest rates tend to increase the exchange rate because it will lead to attract the foreign capital and increase demand by investors. Secondly, the trade balance is measured in order to identify the fluctuation of exchange rate. If the exports are higher than imports, a positive trade balance or trade surplus exits. This indicates that the demand of exchange rate increases because foreign buyers exchange more currency in order to buy goods. Meanwhile if the imports are higher than exports, a negative trade balance or trade deficit exits and this denotes that the currency for that country could lead to devaluation.

Thirdly, the market speculators such as financial experts, economists and technical analysts often affect the market because the foreign traders often influence by the opinion of the speculators. This is an important factor that leads to short-term fluctuations of the

1Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia.Email: phoong@fsmt.upsi.edu.my, *Corresponding Author
2Department of Operations and Management Information Systems, Faculty of Business and Accountancy, University of Malaya, 50603 Kuala Lumpur, Malaysia.Email: phoongs@um.edu.my
exchange rate. Fourthly, the central bank manage the interest rates and use market intervention in administer the economy growth for a country. The last factor that influences the exchange rate is foreign investment. Investors make investment decisions based on the economy growth of a country. When the investment brings into a country, the demand of exchange rate for that country increases in order to buy the local exchange rate to buy goods and this leads to increase the value of exchange rate.

Alternation of exchange rate may effect on the movement of investment price such as import or export price. This reveals that exchange rate for a nation’s will devaluate with the export price meanwhile the import price will increased with the depreciation of exchange rate. The pros and cons of devaluation of exchange rate include creating jobs opportunity, increase aggregate demand, national output and gross domestic product (GDP) (Melander, 2009). This is because investors will increase their investment since the export price and exchange rate is low. Otherwise, raising the exchange rate may reduce the excessive aggregate demand and control the inflation rates. Hence, the fluctuation of exchange rate can effect on inflation or economy growth occurred.

Stock market is a public entity which listed a company shares at an agreed price. Stock market also known as equity market or stock exchanges which vital for a company to raise value and trade the stocks or bonds. This is because of rising stock market price denotes that a company is tends to increase the business investment. In addition, stock market can be categorizes as bull market and bear market. According to Tramontana et al. (2013), bull market is referred to the stock price increase with the growing of investor confidence meanwhile bear market denotes that the market price is depreciates and the economy is in recession.

The alternation of stock market price is affected by factors such as company news and performance. The development of a company influences the stock market price. This is because the rising of company earnings tend to increase the stock market price and vice versa. Moreover, the stock market price for a company also influence by the company news because the competitor companies may spread bad news to affect a company stock market price when competing the same market (Tangjitprom, 2012).

In addition, economic factors such as inflation, deflation, interest rates, economic policy and world events also act as the reason that influences the fluctuation of stock market price. In economics, inflation means the prices of goods and service rise in value and this leads to a higher interest rates to slow down the inflation with devaluate of stock price. While for deflation, the stock market price also confront with devaluate because the prices of goods and service is decreasing and this tend to decrease the interest rate.

Economy policy is vital for a company in which the policy from government tends to affect the interest rate, inflation and stock price. Moreover, the world events such as terrorist attacks and disaster also play role in depreciating the stock market price because the investors and traders will stop trading in that country. Another consequence from the world events is the devaluation of exchange rate.

Gold is a precious metal with characteristic such as good conductor of heat and unaffected by corrosive reagents. Gold is widely used to make jewellery, coins, electric products, foods and others. In addition, gold price is normally measure by troy ounce and by grams. According to the article at Chard (1964), gold prices are increased during first
quarter of 2008 although there is Global Economic Recession happened. The effect of rise in value for gold price leads to increase the demand of gold and depreciated the value of US dollar.

There are four factors which can affect the gold prices such as the value of US dollar, inflation rates, international politics, demand and supply of gold. The gold price is rise in value when there is a downturn of US dollar since central banks will invest other assets such as gold to reduce their risk when depreciates of US dollar. Secondly, gold price will increase when the inflation rates is high. Thirdly, gold is bought during the period of international political tension and this situation often leads to the rising of gold prices. Lastly, gold price is increasing when the demand for gold is higher than the supply. This can be supported by referred to the study of Phoong et al. (2013) that the demand of oil and gold would be increased as the global economy was in the recession period.

According to Phoong, et al. (2014), the gold price is founded to have a positive relationship with the stock returns which means that the rise in value of gold price will lead to a higher stock returns. The data that collected and being analyzed are from Malaysia, Singapore, Thailand and Indonesia.

Motivation of this paper is to investigate the tendency of economy in Malaysia since there are some unusual events that occurred such as Malaysia fixed and float the exchange rate until year 2005 and global economic recession happened in 2008. Hence, the situation of economic growth in Malaysia before the economic crisis and after crisis is drawn the interest in study the relation between it. This is an interesting issue that either Malaysia enables to confront this financial collapse which is known as the worst depression after World War II by economist. This is because Malaysia stock index fall sharply that is 13.3 per cent on the year 1998 after Malaysia government announced the exchange rate control policy. This regulation is wanted to keep the investors and funds during financial crisis 1997. Therefore, a two-component model is used in the present paper to examine the relationship between gold price, exchange rate and stock market price in Malaysia.

In addition, economic and financial time series data always exhibit nonlinear properties. For nonlinear data, the data elements are not organized in a sequential fashion. The main difference between linear and nonlinear data is the way to organize the data elements. For linear data, the elements are organized sequentially meanwhile the data elements for nonlinear data can be constructed by attaching a data element to several other data elements. Multiple probability distributions, structural change and jumps as well as breaks might occur passing the time trend. Thus, normal mixture model which is able to modelling large heterogeneous data, capture the skewness and kurtosis and is a flexible model family (Schnatter, 2006) is employ in this paper to examine the economic relationship between returns of gold price, exchange rate and stock market before and after the crisis.

The structure of the present paper is as follow. Section 2 discussed the previous studies. Section 3 described the methodology of the model. Then, the findings are collected and discussed in section 4. Lastly, the paper is concluded in section 5.
2. Literature Review

Maximum likelihood estimation is a popular statistical tool since it presents findings with consistent, asymptotic normal and optimal properties (Lepage, 2012). In another work, Bekti, Rahayu and Sutikno (2013) applied maximum likelihood estimation in examined the spatial relationship among dependent variable and independent variable on Spatial Durbin model. According to Avdis and Wachter (2013), maximum likelihood estimation provides reliable findings when the observation is finite.

Moreover, maximum likelihood approach can be used to compute the parameters of the models. Li and Zha (2006) studied maximum likelihood estimation to fit finite mixture model in order to estimate the parameters of the models. Hence, a mixture of Poisson distributions is applied to measure the parameters of document classification and word clustering.

Ateya (2012) studied that maximum likelihood estimation is suitable in examined real time series data. Ateya (2012) applied maximum likelihood estimation to fit the finite mixture model in modeling the censored samples. Then, a comparison between the models among generalized exponential distributions and Weibull distributions in analyzed the real data. Results showed that generalized exponential distributions fits the time series data better than Weibull distributions. Furthermore, the analysis of maximum likelihood estimation also showed asymptotic properties.

Instead of obtaining the maximum likelihood estimation directly by maximizing the likelihood function, the maximum likelihood estimation can be obtained through an alternative method based on an iterative procedure which is EM algorithm. These algorithm is an iterative procedure which yield a class of asymptotically efficient estimators for the cointegrating vectors, which are defined in terms of the maximum number of iterations. The iteration is done until it reaches convergence (Dempster et al., 1977).

In year 1997, Malaysia confronted with Asian financial crisis. This leads to government announce currency control policy to rescue the continuous declining economic which is fixed the currency on 3.80 ringgit per US dollar in year 1998. Briefly explanation of the previous studies on Malaysia situation during financial crisis 1997 can referred to Caramazza and Aziz (1998), Hernández and Montiel (2001), Khor (2005), Talib (2004).

The government of Malaysia is then abandoned the pegged exchange rate policy in July 21 2005. Since the exchange rate unable to recover to pre-crisis highs, Malaysia government announced a floating system which mainly used to stabilize the currency. The floating system is useful because the currency is then increased steadily from July 2005 until July 2008. However, there is another unusual event occurred which is Global Financial Crisis in year 2007-2008. This financial crisis results the stock market depreciated around the world, in addition to the crumble of large financial institutions.

3. The Methodology and Model

The nonlinear financial time series data that adopted in the present paper includes returns of nominal exchange rate, gold price and stock market price in Malaysia. Returns data are
used in this research. The EM algorithm is used to investigate the significant relationship between the variables. However, an important step need to progress before the analysis which is transformation of non-stationary data to stationary data. Most of the financial time series data exhibits non-stationary which may leads to robust findings obtained. This transformation process is undergone by taking the natural logarithm or first differencing of the raw data.

After the differencing process, the analysis involves a normal mixture model is fit by EM algorithm in which to investigate the relationship between returns of gold price, exchange rate and stock market price. This analysis is based on the time series model before and after the Global Economic Recession in 2008.

Expectation-Maximization (EM) algorithm is one of the iterative algorithms of maximum likelihood estimation to find parameter estimates. According to Phoong and Mohd Tahir (2015), maximum likelihood estimation is popular in modelling the nonlinear financial time series data. With the advent of the expectation maximization (EM) algorithm, maximum likelihood estimation has been largely applied to fit the mixture distributions (Dempster et al., 1977). This is because the finite mixture model is a special statistical tool to analyze the incomplete data and EM algorithm can be easily applied it. Besides that, a fixed number of components in the finite mixture model are commonly calculated by using the expectation maximization algorithm within a maximum likelihood framework.

Instead of obtaining the maximum likelihood estimation directly by maximizing the likelihood function, the maximum likelihood estimation can be obtained through an alternative method based on an iterative procedure such as Newton-Raphson algorithm or EM algorithm. These algorithms are iterative procedures which yield a class of asymptotically efficient estimators for the co-integrating vectors, which are defined in terms of the maximum number of iterations. The iteration is done until it reaches convergence. In this study, the EM algorithm is being chosen to analysis the time series data because most of the financial time series data may involves latent variables or parameters. The main reason of application of EM algorithm rather than Newton-Raphson algorithm in this study is that the model formulated by EM algorithm may involves latent or unknown variables in addition to unknown parameters and known data observations.

Dempster et al. (1977) studied the usefulness of the EM algorithm in the finite mixture model to model the heterogeneous data and indicated the expectation and maximization for each of the iterations. Iterations of the EM algorithm consist of two processes which are E-step and M-step. Expectation step (E-step) uses the current parameters to estimate missing or hidden data while Maximization step (M-step) is used to maximize the expectation computed in the E-step. EM algorithm increases the accuracy of the log-likelihood result in each of the iterations since these steps are iterated until it converges to a local maximum. In addition, Muthen and Shedden (1999) have used an EM algorithm to fit the finite mixture model to calculate the three latent classes of heavy drink trajectory for young adults.

EM algorithm is an iterative algorithm which used to estimate the marginal likelihood of maximum likelihood estimation by iterated the E-step and M-step. Thus, the marginal likelihood with unknown parameters is given as
\[ L(\theta;x) = p(x|\theta) = \sum_m p(x,m|\theta) \]  

(1)

where \( x \) represents the observed variables, \( \theta \) denotes unknown parameters and \( m \) denotes missing data.

According to Leisch (2004), EM algorithm can be divided into two steps which are E-step and M-step. E-step is known as expectation step in which use current parameters estimate to create an expectation of log-likelihood function. Hence, the general formula to compute E-step is

\[ Q(\theta|\theta^{(k)}) = E_{m|x,\theta^{(k)}}[\log L(\theta;x,m)] \]  

(2)

where \( x \) is observed data, \( \theta^{(k)} \) is parameters estimate at \( k \) iteration, \( m \) denotes the latent variables, \( E_{m|x,\theta^{(k)}} \) denotes the expected value of function and \( \log L(\theta;x,m) \) represents the log likelihood function.

However, M-step or maximization step is used to estimate the parameters maximizing based on the expected log-likelihood in E-step. Thus, the formula for M-step is

\[ \theta^{(k+1)} = \arg \max_{\theta} Q(\theta|\theta^{(k)}) \]  

\[ = \arg \max_{\theta} E_{m|x,\theta^{(k)}}[\log L(\theta;x,m)] \]  

(3)

The iteration of E-step and M-step is repeated until the findings reaches converges.

EM algorithm performs better when the missing values is small and dimensionality of data is low. However, EM algorithm can slow down the E-step when its dimensionality of data is high.

4. The Findings

The findings related to the investigation of the correlation between the financial time series data are explained in this section. The time series data that adopted is first being examined whether it is a non-stationary or stationary data. This is the essential steps for most of the financial time series data because a non-stationary data shows the time series with trends, or with seasonality. These characteristics of time series data may affect the value of the data at different times. Meanwhile, a stationary time series data is a white noise series where the data is look much the same at any period of time.

In order to transform the non-stationary data into stationary form, a differencing process or natural logarithm is applied to transform the data into stationary form. The transformation using natural logarithm is effective in stabilized the variance of a data sets while the transformation using differencing, is useful in stabilized the mean of a data sets by eliminating the changes in the level of a time series, and so removing the trend and seasonality of the time series. In addition, first differencing process or technique brings means of the changes between one observation and the next. The financial time series data that being differenced is called returns data.
Then, the relationship among the exchange rate, gold price and stock market price in Malaysia is analyzed by using EM algorithm. There are four important steps about the statistical analysis which used to explain the correlation of the data sets.

Step 1: Identify the hypothesis
This step plays crucial role to identify the hypothesis of the analysis. The objective to generate the hypothesis, include guide us to ensure the right things to be tested and communicated the findings against the hypotheses that we need. For this study, the hypotheses are as follows:

\[ H_0: \text{There is no correlation between the stock market price, gold price and exchange rate.} \]
\[ H_1: \text{There is a correlation between the stock market price, gold price and exchange rate.} \]

Step 2: Find the critical value and test value
In this step, the critical value and test value based on the analysis of the time series data are obtained and being compared.

Based on the analysis, the iteration process is looped and repeated until the model successfully converged. Hence, the findings that obtained are information after optimize process, parameter estimates for model and graph of mixture model. The information that observed from optimize process includes the total parameters that involve in iterations and Pearson statistic.

<table>
<thead>
<tr>
<th>Table 1: Information after Optimize Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total parameters</td>
</tr>
<tr>
<td>Mean Function Parameters</td>
</tr>
<tr>
<td>Scale Parameters</td>
</tr>
<tr>
<td>Mixing Parameters</td>
</tr>
<tr>
<td>Pearson Statistic</td>
</tr>
</tbody>
</table>

By referring to Table 1, the total parameters that involved in the optimization process are five which includes two mean functions, two scale parameters and a mixing parameter. Mean function parameter is used to explain the mean value that obtained for those component meanwhile scale parameter is adopted to describe the variance for each component. The mixing parameter is used to depict the probability density function or weight for the time series data. In this study, the probability density function that obtained is 1. The findings reveal the probability density function or weight between the data sets before and after crisis. Lastly, the Pearson statistic that obtained for the financial time series data is 142.0 which are roughly equal or equal to the sample size, indicates that the model is not over-dispersed.

<table>
<thead>
<tr>
<th>Table 2: Parameter Estimates for Gold, Exchange rate and Stock Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

As referred to Table 2, the Component column reveals that there are four parameters (same as referred to Table 1), include two mean parameters and two scale parameters. In this study, the component 1 denotes the data sets before crisis. Meanwhile, the
component 2 represents the financial time series data after crisis. For the values that displays at the column named “Gold”, “Exchange Rate” and “Stock Price”, bring means of the values for that parameter and component.

The component 1 in Table 2 is the model before Global Economic Recession meanwhile component 2 is the model after the crisis. For gold price and exchange rate, there are low mean low variance in component 1 and high mean high variance in component 2. For stock price, there are low mean high variance for component 1 and high mean low variance for component 2.

The variance for both component 1 and component 2 for all three variables include gold price, exchange rate and stock returns are small. This indicates that the data sets are close to the mean. Thus, the findings that obtained can be concluded as valid, reliable and significant.

Based on the findings above, it can be concluded that gold price has positive relationship with stock price in Malaysia. Moreover, there is a positive relationship between exchange rate and stock price. Also, the findings above depicts that gold price is alter with the fluctuation of exchange rate. Gold prices remain its values although there are wars, crisis period or changes of government. Thus, it can be concluded that gold price and exchange rate in Malaysia is not strongly related with each other.

Exchange rate in Malaysia is fixed and float before Global Economic Recession in year 2008 and rise in value after the crisis. Apart from that, both gold price and stock price in Malaysia shows the mean after crisis are greater than before crisis. This indicates that gold and stock price is rise in value after the economic recession in year 2008. Thus, Table 4 is providing to determine whether these three variables are correlated.

For all three variables, the variances for both component 1 and component 2 are small. This illustrated that both components are reliable and significant for returns of gold price, exchange rate and stock price.

Table 3: Mixing Parameter for Gold, Exchange rate and Stock Price

<table>
<thead>
<tr>
<th>Component</th>
<th>Gold</th>
<th>Exchange rate</th>
<th>Stock Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1697</td>
<td>0.5994</td>
<td>0.8353</td>
</tr>
<tr>
<td>2</td>
<td>0.8303</td>
<td>0.4006</td>
<td>0.1647</td>
</tr>
</tbody>
</table>

By referring to Table 3 and Figure 1, the information about the normal mixture model is explained. Table 3 lists the mixing parameter that obtained for each variable along with the component. Note that component 1 is the model before crisis and component 2 is the model after crisis. Mixing parameter or weight is used to describe the probability density function for each variable. The general equation for probability density function is

\[ f(x) = \pi \left( \mu_1, \sigma_1^2 \right) + \left( 1 - \pi \right) \left( \mu_2, \sigma_2^2 \right) \]

where \( \pi \) denotes weight for the model, \( \mu \) is the mean and \( \sigma^2 \) is variance for those model.

Hence, it can be expressed as the probability density function for gold in Malaysia is \( 0.1697(0.0030, 0.0003) + 0.8303(0.0104, 0.0018) \). For exchange rate in Malaysia, the function is \( 0.5994(0.0115, 0.0004) + 0.4006(-0.0012, 0.0031) \). Lastly, the probability density function for stock price in Malaysia is \( 0.8353(-0.0012, 0.0002) + 0.1647(-0.0001, \).
4.635E-8). Apart from that, it also can be concluded that the component 2 is more prevalent as compared to component 1 for gold price. Otherwise, component 1 for both exchange rate and stock price in Malaysia is more prevalent as compared to component 2. Moreover, Figure 1 shows the mixture models for returns of gold price, exchange rate and stock price in Malaysia. Each graph in Figure 1 contains three normal models lines that are component 1, component 2 and mixture model. The blue line represents the mixture model, and the red line denotes the component 1 or first normal model. Meanwhile, the green line reveals the component 2 or the second normal model.

**Figure 1: Mixture Model for variables**

- **Gold**
  - Blue line: Mixture model
  - Red line: Normal model 1
  - Green line: Normal model 2

- **Exchange Rate**
  - Blue line: Mixture model
  - Red line: Normal model 1
  - Green line: Normal model 2
In order to determine whether the gold price, exchange rate and stock price in Malaysia are correlated with each other, a correlation table is provided. Table 4 lists the estimates for each component along with the p-value that obtained. Based on the findings in Table 4, it can be concluded that the gold price, exchange rate and stock price are correlated since the p-value that obtained are greater than the estimate value for both component 1 and component 2. For component 1, the p-value that obtained is 0.0507 which is greater than the estimate value, 0.0068. While for component 2, the p-value is 0.0827 which is larger than the estimate value, 0.0027.

### Table 4: Correlation table

<table>
<thead>
<tr>
<th>Component</th>
<th>Estimate</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00684</td>
<td>0.0507</td>
</tr>
<tr>
<td>2</td>
<td>0.0027</td>
<td>0.0827</td>
</tr>
</tbody>
</table>

Step 3: Decision making
Decision making is a process of selecting the plausible alternative among others. Due to Robbins and Mary (2005), this process is defines as the selection of a preferred course of action from two or more alternatives. Moreover, decision making plays vital role in reveal the links or connection between the variables.

In this study, the decision that can be made based on the findings obtained is that there is not enough evidence to reject the null hypothesis since both components shows a larger p-value as compared to estimate value.

Step 4: Make conclusion
In this study, there is not enough evidence to reject the null hypothesis, $H_0$ so it can be concluded that there is a correlation between the stock price, gold price and exchange rate in Malaysia.

### 5. Summary and Conclusions

In this paper, a normal mixture model is established to determine the interaction among the non-stationary time series data via EM algorithm. In relation to the results, gold price, exchange rate and stock market price in Malaysia are correlated. The precious metal and
market prices are appreciated after the crisis and this indicates that the economy in Malaysia is getting growth and develops.

The contribution for this paper is to employ and investigate the relationship between the nonlinear models, include gold price, exchange rate and stock returns in Malaysia. Additionally, the causality between variables based on the effects of Global Economic Crisis in year 2007 is taking into account to determine whether the variables are correlated or not.

The suggestions for the future study are studying other financial time series data such as oil price which act as one of the affective time series data on economy. While the statistical methods such as Newton-Raphson algorithm, Bayesian method, and ARIMA can be use instead of EM algorithm in studying the relationship between the financial time series data. Apart from that, the comparison method such as EM algorithm and Newton-Raphson algorithm on this time series data can be used to determine the most plausible method in analyzing the data sets.

References


