

**ANALYSIS OF THE EXCHANGE RATE OF THE THAI
BAHT AGAINST THE CHINESE YUAN USING A
SUPPORT VECTOR MACHINE MODEL**

Phakkhaphon Sawatkamon



**A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Applied Mathematics
Suranaree University of Technology**

Academic Year 2019

การวิเคราะห์อัตราแลกเปลี่ยนระหว่างเงินบาทไทยกับเงินหยวนจีน
ด้วยแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาคณิตศาสตร์ประยุกต์

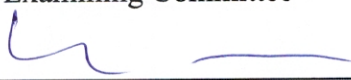
มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2562

**ANALYSIS OF EXCHANGE RATE OF THE THAI BAHT
AGAINST THE CHINESE YUAN USING A SUPPORT
VECTOR MACHINE MODEL**

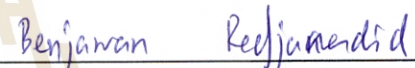
Suranaree University of Technology has approved this thesis submitted
in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee



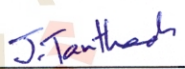
(Assoc. Prof. Dr. Eckart Schulz)

Chairperson




(Asst. Prof. Dr. Benjawan Rodjanadid)

Member (Thesis Advisor)




(Asst. Prof. Dr. Jessada Tanthanuch)

Member (Thesis Co-advisor)




(Dr. Panu Yimmunag)

Member




(Dr. Amornrat Suriyawichitseranee)

Member



(Assoc. Prof. Flt. Lt. Dr. Kontron Chamniprasart)



(Assoc. Prof. Dr. Worawat Meevasana)

Vice Rector for Academic Affairs
and Internationalization

Dean of Institute of Science

กัศพล สวัสดิ์กมล : การวิเคราะห์อัตราแลกเปลี่ยนระหว่างเงินบาทไทยกับเงินหยวนจีนด้วย
แบบจำลองซัพพอร์ตเวกเตอร์แมชชีน (ANALYSIS OF THE EXCHANGE RATE OF
THE THAI BAHT AGAINST THE CHINESE YUAN USING A SUPPORT VECTOR
MACHINE MODEL) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.เบญจวรรณ โจนดิษฐ์,
61 หน้า.

คำสำคัญ : ซัพพอร์ตเวกเตอร์แมชชีน/การถดถอยเชิงเส้นพหุคูณ/อัตราแลกเปลี่ยนสกุลเงินบาทไทย
ต่อสกุลเงินหยวนจีน

งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลกระทบต่ออัตราแลกเปลี่ยนเงินบาท
ไทยต่อสกุลเงินหยวนจีน โดยใช้แบบจำลองการถดถอยเชิงเส้นพหุคูณ และหาแบบจำลองที่
เหมาะสมใช้ในการพยากรณ์อัตราแลกเปลี่ยนเงินบาทไทยต่อสกุลเงินหยวนจีนด้วยวิธีซัพพอร์ต
เวกเตอร์แมชชีน โดยคำนวณผ่าน โปรแกรม R (รุ่น 3.6.3)

ผลการศึกษาพบว่า ปัจจัยที่ส่งผลต่ออัตราแลกเปลี่ยนสกุลเงินบาทไทยต่อสกุลเงินหยวนจีน
อย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญร้อยละ 5 มีดังนี้ ปัจจัยที่ส่งผลในทิศทางเดียวกัน ได้แก่
อัตราดอกเบี้ยนโยบาย และ มูลค่าการนำเข้า ไทย-จีน สำหรับปัจจัยที่ส่งผลต่ออัตราแลกเปลี่ยนใน
ทิศทางตรงกันข้าม ได้แก่ เงินสำรองระหว่างประเทศ ทั้งนี้แบบจำลองสำหรับการพยากรณ์อัตรา
แลกเปลี่ยนสร้างด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนได้ ให้ผลพยากรณ์อัตราแลกเปลี่ยนสกุลเงินบาท
ไทยต่อสกุลเงินหยวนจีน ตั้งแต่เดือนมกราคม พ.ศ. 2562 จนถึง เดือนมิถุนายน พ.ศ. 2562 มี
ความคลาดเคลื่อนราคากำลังสองเฉลี่ยเท่ากับ 0.1435

สาขาวิชาคณิตศาสตร์

ปีการศึกษา 2562

ลายมือชื่อนักศึกษา Suw

ลายมือชื่ออาจารย์ที่ปรึกษา Benjawan

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม J. Tantham

PHAKKHAPHON SAWATKAMON : ANALYSIS OF THE EXCHANGE
RATE OF THE THAI BAHT AGAINST THE CHINESE YUAN USING
A SUPPORT VECTOR MACHINE MODEL. THESIS ADVISOR :
ASST. PROF. BENJAWAN RODJANADID, Ph.D. 61 PP.

SUPPORT VECTOR MACHINE/MULTIPLE LINEAR REGRESSION/THE
EXCHANGE RATE THAI BAHT AGAINST CHINESE YUAN

The objectives of this research were to study the factors that affect the exchange rate of the Thai Baht against the Chinese Yuan by using multiple linear regression and to find an appropriate model for forecasting the exchange rate of the Thai Baht against the Chinese Yuan by using Support Vector Machine (SVM), work computations by R statistical software.

The result showed that factors which affect the interest rate at a statistically significant level of at least 5% were Policy Interest Rate and the Value of Imports from China, both in the positive direction, and the Foreign Exchange Reserves in the negative direction. The SVM model could predict the exchange rate of the Thai Baht against the Chinese Yuan during the period January 2019 until June 2019 with a Root Mean Square Error of 0.1435.

School of Mathematics

Academic Year 2019

Student's Signature _____

Advisor's Signature Benjawan

Co-Advisor's Signature S. Tanthad

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my thesis advisor Asst. Prof. Dr. Benjawan Rodjanadid and co-advisor Asst. Prof. Dr. Jessada Tanthanuch for giving me to opportunity to do this research. They were kindly guiding, encouraging, and explaining to me whenever I did not understand some details in this thesis work.

In addition, I am grateful for professional support received from Assoc. Prof. Dr. Eckart Schulz, Dr. Panu Yimmunag, Dr. Amornrat Suriyawichitseranee and all professors in the School of Mathematics, Institute of Science, Suranaree University of Technology.

Finally, I most gratefully acknowledge my family and my friends for their good support and encouragement.

Phakkhaphon Sawatkamon

มหาวิทยาลัยเทคโนโลยีสุรนารี

CONTENTS

	Page
ABSTRACT IN THAI	I
ABSTRACT IN ENGLISH	II
ACKNOWLEDGEMENTS	III
CONTENTS	IV
LIST OF TABLES	VIII
LIST OF FIGURES	IX
CHAPTER	
I INTRODUCTION	1
1.1 Research Objective	3
1.2 Scope Limitations	3
1.3 Research Procedure	4
II LITERATURE REVIEW	5
2.1 Multiple Linear Regression Analysis	5
2.1.1 The Assumption of the Classical Linear Regression	5
2.1.2 Estimation of the Model Parameters	6
2.1.3 Test for Significance of Regression	9
2.2 Linear Classification Problem	10
2.2.1 Classification Problem	10
2.2.2 Support Vector Classification (SVC) for Linearly Separable Problems	10
2.2.3 Maximal Margin Method	11

CONTENTS (Continued)

	Page
2.2.4 Properties of the Maximal Margin Method in Linear Classification	13
2.3 The Dual Problem in Linear Classification	14
2.4 Support Vector	16
2.5 Linear C-Support Vector Classification	16
2.5.1 Maximal Margin Method	16
2.5.2 Properties of the Maximal Margin Method in Linear C-Support Vector Classification	18
2.6 The Dual Problem in Linear C-Support Vector Classification	18
2.7 Linear Regression	20
2.7.1 Regression Problem	20
2.7.2 Hard $\bar{\epsilon}$ -Band Hyperplane	21
2.7.3 Hard $\bar{\epsilon}$ -Band Hyperplane and Linear Classification	22
2.7.4 Optimization Problem of Constructing a Hard $\bar{\epsilon}$ -Band Hyperplane	23
2.8 Linear Hard ϵ -Band Support Vector Regression	25
2.8.1 Primal Problem	25
2.8.2 Relationship between the Primal and Dual Problems	25
2.9 Linear ϵ -Support Vector Regression	28
2.9.1 Primal Problem	28
2.9.2 Relationship between Dual Problem and Primal Problem	29
2.10 Kernel	31
2.10.1 Commonly Used Kernels	31

CONTENTS (Continued)

	Page
2.11 Root Mean Square Error	32
2.12 Related Researches	32
III RESEARCH METHODOLOGY	35
3.1 Collect Training Data and Analyze Descriptive Statistics	35
3.2 Program R and Packages Installation	36
3.3 Analyzing Independent Factors that Affect the Exchange Rate . . .	36
3.3.1 Study Relationship between Dependent and Independent Variables	36
3.3.2 Estimating Parameter of Model	37
3.3.3 Test for Significance of Regression	37
3.4 Creating a Forecasting Model	38
3.5 Predicting	38
3.6 Accuracy Measurement of Predicting Model	38
IV RESULTS AND DISCUSSION	40
4.1 Descriptive Statistics of Training Data	40
4.2 Factors that Affect the Exchange Rate	41
4.3 Creating a Forecasting Model	42
4.4 Predicting	45
4.5 Accuracy Measurement of Predicting Model	45
V CONCLUSION AND RECOMMENDATION	46
REFERENCES	48
APPENDICES	
APPENDIX A ANALYZING DATA BY USING PROGRAM R	52

CONTENTS (Continued)

	Page
A.1 Import Data and Install Packages.	53
A.2 Analyzing Descriptive Statistics	53
A.3 Multiple Linear Regression Analysis	53
A.4 Support Vector Machine	54
A.5 Predicting	55
A.6 Performance of model	55
APPENDIX B ROOT MEAN SQUARE ERROR VALUES OF MODEL BY USING REGRESSION ANALYSIS . . .	56
B.1 Choose one independent variable from 5 variables.	57
B.2 Choose two independent variables from 5 variables.	58
B.3 Choose three independent variables from 5 vari- ables.	58
B.4 Choose four independent variables from 5 vari- ables.	59
B.5 Choose five independent variables from 5 vari- ables.	60
CURRICULUM VITAE	61

LIST OF TABLES

Table		Page
1.1	The list of government spending (share as a percentage of GDP) of some countries.	1
3.1	Shows Training data.	35
4.1	Training data Descriptive statistics.	40
4.2	Coefficient, t-statistic values(t^*), and standard error of independent variables (5% of a significant level).	41
4.3	Parameter, SVM Type, and Kernel Type of a predicting model.	43
4.4	The appropriate parameter of predicting model.	43
4.5	Predicted values from the SVR-THC and the MLR models.	45
B.1	One independent variable.	57
B.2	Two independent variables.	58
B.3	Three independent variables.	58
B.4	Four independent variables.	59
B.5	Five independent variables.	60

LIST OF FIGURES

Figure		Page
2.1	Linearly separable problem.	11
2.2	Optimal separating line with fixed normal direction.	12
2.3	Separating line with maximal margin.	12
2.4	Geometric interpretation of Theorem 2.12.	16
2.5	A regression problem in \mathbb{R}	21
2.6	A linear regression problem in \mathbb{R}	21
2.7	A hard $\bar{\epsilon}$ -band hyperplane (line) in \mathbb{R}	22
4.1	Showing the root mean square error depending on epsilon and penalty parameters.	44
4.2	The actual and predicted value of training data.	44

CHAPTER I

INTRODUCTION

A mixed economy is an economy that uses both market signals and government directives to allocate goods and resources (Schiller and Gebhardt, 2016). Most countries have mixed economies with varying degrees of private enterprise and state intervention. The list of government spending (share as a percent of GDP) of some countries is shown in Table 1.1 (Pettinger, 2011).

Table 1.1 The list of government spending (share as a percentage of GDP) of some countries.

Country	Share (%)
Sweden	52.5
United Kingdom	47.3
Germany	43.7
United States	38.9
Japan	37.1
Russia	34.1
China	20.8
Thailand	17.7

The goodness of a mixed economy is that the economy is largely driven by private investment and enterprise, but government can intervene to reduce fluctuations in the economic cycle, for example, reduce inflation, boost economic

growth (fiscal policy) and have money exchange controls. Exchange controls are government-imposed limitations on the purchase and/or sale of currencies. These controls allow countries to better stabilize their economies by limiting in-flows and out-flows of currency, which can create exchange rate volatility (Kenton, 2019). Exchange rate is the price of one country's currency expressed in terms of another's (the domestic price of a foreign currency) (Schiller and Gebhardt, 2016). In the foreign exchange market and international finance, a world currency, or global currency, is accepted for trade throughout the world. The most popular are the U.S. Dollar, the Euro, the British Pound and the Japanese Yen. However, because of the U.S. economic crisis in 2008 and the Eurozone debt crisis in 2010, the International Monetary Fund (IMF) awarded the Chinese yuan the status of a reserve currency on November 30, 2015 and later added the yuan to its Special Drawing Rights basket on October 1, 2016. ("How the Yuan Could Become a Global Currency", n.d). The Renminbi (RMB) is the official currency of China, whereas the yuan is the basic unit of the renminbi. The Chinese authorities announced the RMB internationalization policy in 2009 to promote the use of RMB as an international currency (Bank of Thailand, n.d.). China exports a lot of commodities which makes China have more economic clout in relation to many countries. The IMF has claimed that China has been the world's second largest economic by nominal GDP since 2010-2018 (Congressional Research Service, 2019). Trading Economics (2019) claimed that China is the main trading partner of Thailand (12 percent of total exports and 15 percent of total imports-updated on October of 2019). Thus, the currency fluctuations of the yuan have a very high impact on trade and investment. An understanding of the currency exchange rate of Thai bath (THB) against Chinese yuan (CNY) provides a good opportunity in international business planning.

This research focuses on the study of the forecasting model for THB-CNY exchange rate. Many models of the exchange rate forecasting are reviewed. The applications of support vector machine are considered and applied to analyze the historical THB-CNY exchange rate. The model obtained is examined and compared to the existing forecasting models. The results can be used in the real-life business and be extended to other forecasting model.

1.1 Research Objective

- 1) To study the factors that affect the exchange rate of the Thai Baht against the Chinese Yuan by Multiple Linear Regression.
- 2) To find an appropriate model for forecasting the exchange rate of Thai Baht against Chinese Yuan by using Support Vector Machine.

1.2 Scope Limitations

- 1) The dependent variable is the exchange rate of the Thai Baht against the Chinese Yuan, available on <https://www.bot.or.th/Thai/Statistics/Pages/default.aspx>.
- 2) The independent variable factors in this study consist of the Foreign Exchange Reserves(FER), the Policy Interest Rate (PR), the Nominal Effective Exchange Rate (NEER), the Value of Imports from China(IM) and the Value of Exports from China(EV), available on <https://www.bot.or.th/-Thai/Statistics/Pages/default.aspx>, <http://www2.ops3.moc.go.th/>.
- 3) The secondary data series(Time Series Data) used in this study are the

monthly data from January 2009 until December 2018, a total of 120 months.

- 4) Statistical calculations are based on Program R Version 3.6.3, working on Microsoft Window 10.

1.3 Research Procedure

The research work proceeded as follows:

- 1) Study the application of Multiple Linear regression.
- 2) Study the application of Support Vector Machine.
- 3) Study the program R software(Version3.6.3).
- 4) Collect the Training data, and descriptive statistics.
- 5) Analyze the independent factors that affect the exchange rate of the Thai Baht against the Chinese Yuan by using Multiple Linear Regression.
- 6) Construct the model for forecasting the exchange rate of the Thai Baht against the Chinese Yuan by using Support Vector Machine.
- 7) Use the obtained model to predict the exchange rate of the Thai Baht against the Chinese Yuan from January 2019 until June 2019.
- 8) Measure the accuracy for the predictive model by using Root Mean Square Error.

CHAPTER II

LITERATURE REVIEW

In this chapter, the knowledge of primary mathematics regarding Multiple Linear Regression, Support Vector Machines, and related research works are reviewed.

For the content of Multiple Linear Regression, we follow from Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining(2012).

2.1 Multiple Linear Regression Analysis

Regression analysis is a statistical technique for investigating and modelling the relationship between variables. Multiple linear regression is a type of a linear regression model with one dependent variable and two or more independent variables. The general form of the multiple linear regression with k predictor variables is as the follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (2.1)$$

where Y is the response(or dependent) variable, X_i are the predictor(or independent) variable, β_j for $j = 1, 2, 3, \dots, k$ are the regression coefficients(or parameters) and ε is the random error component.

2.1.1 The Assumption of the Classical Linear Regression

We now make some basic assumption on the model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i ; \quad i = 1, 2, 3, \dots, n \quad (2.2)$$

- 1) ε_i is a random variable with zero mean and variance σ^2 (unknown), that is $E(\varepsilon_i) = 0$, $\text{Var}(\varepsilon_i) = \sigma^2$.
- 2) ε_i and ε_j are uncorrelated for $i \neq j$ and $i, j \in \{1, 2, 3, \dots, n\}$, that is $\text{Cov}[\varepsilon_i, \varepsilon_j] = 0$.
- 3) ε_i is a normally distributed random variable with mean zero and variance σ^2 , that is $\varepsilon_i \sim N(0, \sigma^2)$.
- 4) There is no a perfect linear relationship between the independent variables. That is, there is no multicollinearity.

2.1.2 Estimation of the Model Parameters

Consider the linear model in equation (2.2), we can rewrite it in the matrix form as follows;

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

It can be rewritten as;

$$Y = X\beta + \varepsilon \tag{2.3}$$

where Y is an $n \times 1$ column vector of observations.

X is an $n \times (k + 1)$ matrix of the levels of the regressor variables.

β is an $(k + 1) \times 1$ vector of the regression coefficients.

ε is an $n \times 1$ vector of random errors.

Definition 2.1 [Residual] The difference between the observed value Y_i and the corresponding fitted value \hat{Y}_i is the residual.

$$\varepsilon_i = Y_i - \hat{Y}_i \tag{2.4}$$

Least - Square Estimation of the Regression Coefficient

The method of least squares is used to estimate the regression coefficient in equation (2.1). Now, we assume $b_0, b_1, b_2, \dots, b_k$ are the estimators of $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, respectively, and suppose that $n > k$. The least squares function is

$$S(\beta_0, \beta_1, \dots, \beta_k) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij}]^2 \quad (2.5)$$

The function S must be minimized with respect to $\beta_0, \beta_1, \dots, \beta_k$. The least - squares estimators of $\beta_0, \beta_1, \dots, \beta_k$ must satisfy

$$\left. \frac{\partial S}{\partial \beta_0} \right|_{b_0, b_1, \dots, b_k} = -2 \sum_{i=1}^n \left[Y_i - b_0 - \sum_{j=1}^k b_j X_{ij} \right] = 0 \quad (2.6)$$

and

$$\left. \frac{\partial S}{\partial \beta_j} \right|_{b_0, b_1, \dots, b_k} = -2 \sum_{i=1}^n \left[Y_i - b_0 - \sum_{j=1}^k b_j X_{ij} \right] X_{ij} = 0, \quad j = 0, 1, 2, \dots, k \quad (2.7)$$

From equation (2.6) and (2.7), we obtain the least square normal equations

$$\begin{aligned} nb_0 + b_1 \sum_{i=1}^n X_{i1} + b_2 \sum_{i=1}^n X_{i2} + b_3 \sum_{i=1}^n X_{i3} + \dots + b_k \sum_{i=1}^n X_{ik} &= \sum_{i=1}^n Y_i \\ b_0 \sum_{i=1}^n X_{i1} + b_1 \sum_{i=1}^n X_{i1}^2 + b_2 \sum_{i=1}^n X_{i1}X_{i2} + \dots + b_k \sum_{i=1}^n X_{i1}X_{ik} &= \sum_{i=1}^n X_{i1}Y_i \end{aligned} \quad (2.8)$$

⋮

$$b_0 \sum_{i=1}^n X_{ik} + b_1 \sum_{i=1}^n X_{ik}X_{i1} + b_2 \sum_{i=1}^n X_{ik}X_{i2} + \dots + b_k \sum_{i=1}^n X_{ik}^2 = \sum_{i=1}^n X_{ik}Y_i$$

Therefore, there are $k + 1$ normal equations, one for each of the unknown regression coefficients, and the solution to the normal equation will be the least square estimators $b_0, b_1, b_2, \dots, b_k$.

From the least square normal equations (2.8) as follows;

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & x_{21} & \cdots & X_{n1} \\ X_{12} & X_{22} & \cdots & X_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1k} & X_{2k} & \cdots & X_{nk} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n X_{i1} & \cdots & \sum_{i=1}^n X_{ik} \\ \sum_{i=1}^n X_{i1} & \sum_{i=1}^n X_{i1}^2 & \cdots & \sum_{i=1}^n X_{i1}X_{ik} \\ \sum_{i=1}^n X_{i2} & \sum_{i=1}^n X_{i1}X_{i2} & \cdots & \sum_{i=1}^n X_{i2}X_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n X_{ik} & \sum_{i=1}^n X_{i1}X_{ik} & \cdots & \sum_{i=1}^n X_{ik}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

It can be rewritten as;

$$X^T Y = (X^T X) b \quad (2.9)$$

where $b = [b_0 \ b_1 \ b_2 \ \cdots \ b_k]^T$

Multiplying both sides of equation (2.9) by $(X^T X)^{-1}$, we get

$$(X^T X)^{-1} X^T Y = (X^T X)^{-1} (X^T X) b$$

$$(X^T X)^{-1} X^T Y = I b$$

where I is $(k+1) \times (k+1)$ identity matrix.

Therefore, $b = (X^T X)^{-1} X^T Y$.

Theorem 2.2 The estimator $b = (X^T X)^{-1} X^T Y$ is an unbiased estimator of β .

In addition,

$$\text{Var}[b] = (X^T X)^{-1} \sigma^2$$

Proof In the first step, we will show that b is an unbiased estimator;

$$\begin{aligned} E[b] &= E[(X^T X)^{-1} X^T Y] \\ &= (X^T X)^{-1} X^T E[Y] \\ &= (X^T X)^{-1} X^T E[X\beta + \varepsilon] \end{aligned}$$

from the assumption of the Classical Linear Regression, we have $E[\varepsilon] = 0$.

Therefore, $E[b] = (X^T X)^{-1} X^T X \beta = \beta$. This completes the proof of the unbiasedness of b .

Next, the variance of b can be computed directly:

$$\begin{aligned} \text{Var}[b] &= \text{Var}[(X^T X)^{-1} X^T Y] \\ &= (X^T X)^{-1} X^T \text{Var}[Y] (X^T X)^{-1} X^T \\ &= (X^T X)^{-1} X^T X (X^T X)^{-1} \sigma^2 \\ &= (X^T X)^{-1} \sigma^2 \end{aligned}$$

Theorem 2.3 The unbiased estimator of the variance in the multiple linear regression is given by

$$\sigma^2 = \frac{Y^T (I - X(X^T X)^{-1} X^T) Y}{n - p} = \frac{1}{n - p} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

where $p = k + 1$

2.1.3 Test for Significance of Regression

The test for Significance of regression is a test to determine if there is a linear relationship between the response Y and any of the regressor variables $X_1, X_2, X_3, \dots, X_n$.

Assume that the estimator parameter $\beta_0, \beta_1, \dots, \beta_k$ are b_0, b_1, \dots, b_k , respectively.

The statement for the hypotheses are:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

$$H_1 : \beta_i \neq 0 \text{ for at least one } i$$

The statistic that can be used for testing H_0 versus H_1 is the t test statistic in the following form:

$$t^* = \frac{b_i}{\text{se}(b_i)} \sim t_{\frac{\alpha}{2}, n-(k+1)} \quad \forall i = 1, 2, 3, \dots, k$$

where $\text{se}(b_i) = \sqrt{\sigma^2 C_{ii}}$, C_{ii} is an element on the main diagonal of $(X^T X)^{-1}$ and we reject H_0 if $t_{\frac{\alpha}{2}, n-(k+1)} < |t^*|$.

2.2 Linear Classification Problem

For the content of Support Vector Machine, we follow Deng, Tain, and Zhang(2013).

2.2.1 Classification Problem

Classification Problem: Given a training set

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$$

where $x_i \in \mathbb{R}^n, y_i \in \mathbb{Y} = \{-1, 1\}, i = 1, 2, 3, \dots, m$, find a real function $g(x)$ in \mathbb{R}^n , to derive the value of y for any x by the decision function

$$y_i = \text{sgn}(g(x_i)).$$

Thus the solving of a classification problem is to find a criterion to separate the \mathbb{R}^n space into two regions according to the training set T .

Note that in the training set T , $(x_i, y_i) \in \mathbb{R}^n \times \mathbb{Y}$ is called *training point* for all $i = 1, 2, 3, \dots, m$. The vector $x_i \in \mathbb{R}^n$ is called *input*, its components are called *features* and $y_i \in \mathbb{Y}$ is the *label* or *output*.

2.2.2 Support Vector Classification (SVC) for Linearly Separable Problems

The linear classification problem is the problem whether the training set can be separated by a hyperplane correctly. The definition is follows.

A hyperplane \mathbb{H} in \mathbb{R}^n is the set of points that satisfy a linear equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where the vector $u = (a_1, a_2, \dots, a_n)$ is not zero.

Definition 2.4 (Linearly separable problem)

Consider the training set $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} \in (\mathbb{R}^n \times \mathbb{Y})^m$, where $x_i \in \mathbb{R}^n, y_i \in \mathbb{Y} = \{-1, 1\}, i = 1, 2, 3, \dots, m$. If there exist $w \in \mathbb{R}^n, b \in \mathbb{R}$ and a positive number ε such that for any subscripts i with $y_i = 1$, we have $(w \cdot x_i) + b \geq \varepsilon$, and for any subscripts i with $y_i = -1$, we have $(w \cdot x_i) + b \leq -\varepsilon$, we say the training set and its corresponding classification problem are linearly separable.

2.2.3 Maximal Margin Method

Derivation of the maximal margin method

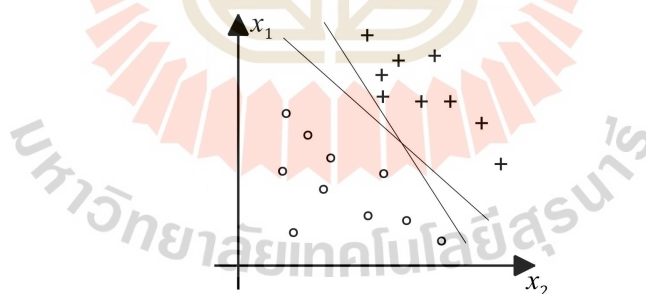


Figure 2.1 Linearly separable problem.

Consider the separable problem in \mathbb{R}^2 shown in the Figure 2.1, and try to find a suitable straight line to separate the \mathbb{R}^2 space into two regions. We want to select the best separating line among the straight lines which is able to separate all of the positive inputs “+” and negative inputs “o” correctly.

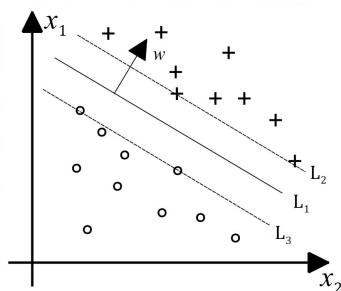


Figure 2.2 Optimal separating line with fixed normal direction.

In Figure 2.2, L_1 is one of the straight lines with the given w (w is normal vector), separating all of the positive and negative inputs correctly. However such a line is not unique, any line obtained parallel moving L_1 before approaching any input is a candidates. The two lines L_2, L_3 in the extreme cases are called *support lines*. The straight line between L_2 and L_3 is the best separating line.

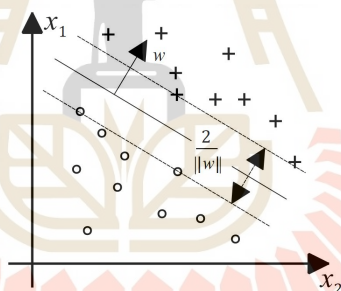


Figure 2.3 Separating line with maximal margin.

We focus on how the best normal direction w is select In Figure 2.3. The distance between the two support lines is called “*margin*”. Thus we select the normal direction which makes the margin maximal.

The idea of maximal margin leads to the following optimization problem for w and b ;

$$\max_{w,b} \frac{2}{\|w\|}, \quad (2.10)$$

subject to

$$(w \cdot x_i) + b \geq 1, \quad \forall i : y_i = 1, \quad (2.11)$$

$$(w \cdot x_i) + b \leq -1, \quad \forall i : y_i = -1, \quad (2.12)$$

or

$$\min_{w,b} \frac{1}{2} \|w\|^2, \quad (2.13)$$

subject to

$$y_i((w \cdot x_i) + b) \geq 1, \quad i = 1, 2, 3, \dots, m. \quad (2.14)$$

The above optimization problem is derived from maximizing the margin between the two support lines in two-dimensional space \mathbb{R}^2 . For solving classification problem in n -dimensional space \mathbb{R}^n , we should maximize the margin between the two support hyperplanes. It is called *the principle of maximal margin*.

2.2.4 Properties of the Maximal Margin Method in Linear Classification

Theorem 2.5 For a linearly separable problem, there exist a solution (w^*, b^*) to the optimization problem (2.13) \sim (2.14) and the solution satisfies:

- 1) $w^* \neq 0$;
- 2) there exist a $j \in \{i \mid y_i = 1\}$ such that

$$(w^* \cdot x_j) + b^* = 1; \quad (2.15)$$

- 3) there exist a $k \in \{i \mid y_i = -1\}$ such that

$$(w^* \cdot x_k) + b^* = -1. \quad (2.16)$$

Theorem 2.6 For a linearly separable problem, the solution to the optimization problem (2.13) \sim (2.14) is unique.

2.3 The Dual Problem in Linear Classification

By the theory of Lagrange multipliers, one can formulate a related optimization problem, called *the dual problem*.

To derive the dual problem of the primal problem (2.13) \sim (2.14), we introduce the Lagrange function;

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i (y_i ((w \cdot x_i) + b) - 1),$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$ is the *Lagrange multiplier vector*.

Theorem 2.7 (Optimization problem)

$$\max_{\alpha} -\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i y_j (x_i \cdot x_j) \alpha_i \alpha_j + \sum_{j=1}^m \alpha_j, \quad (2.17)$$

subject to

$$\sum_{i=1}^m y_i \alpha_i = 0, \quad (2.18)$$

$$\alpha_i \geq 0, \quad i = 1, 2, 3, \dots, m, \quad (2.19)$$

or

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i y_j (x_i \cdot x_j) \alpha_i \alpha_j - \sum_{j=1}^m \alpha_j, \quad (2.20)$$

subject to

$$\sum_{i=1}^m y_i \alpha_i = 0, \quad (2.21)$$

$$\alpha_i \geq 0, \quad i = 1, 2, 3, \dots, m, \quad (2.22)$$

is the dual problem of the primal problem (2.13) \sim (2.14).

Theorem 2.8 For a linearly separable problem, the dual problem (2.17) \sim (2.19) or (2.20) \sim (2.22) has a solution.

Theorem 2.9 Consider the linearly separable problem. For any solution to the dual problem (2.20) \sim (2.22), $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$, there must be a nonzero component α_j^* . Furthermore, given arbitrary nonzero component α_j^* of α^* , the unique solution to the primal problem (2.13) \sim (2.14) can be obtained by

$$w^* = \sum_{i=1}^m \alpha_i^* y_i x_i ,$$

$$b^* = y_j - \sum_{i=1}^m \alpha_i^* y_i (x_i \cdot x_j).$$

Algorithm 2.10 (Linear separable support vector classification)

Step 1) Input the training set $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, where $x_i \in \mathbb{R}^n$, $y_i \in \mathbb{Y} = \{-1, 1\}$, $i = 1, 2, 3, \dots, m$;

Step 2) Construct and solve the convex quadratic programming (2.20) \sim (2.22), to obtain a solution $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$;

Step 3) Compute $w^* = \sum_{i=1}^m \alpha_i^* y_i x_i$. Choose a positive component of α^* , α_j^* , then compute b^*

$$b^* = y_j - \sum_{i=1}^m \alpha_i^* y_i (x_i \cdot x_j);$$

Step 4) Construct the separating hyperplane $(w^* \cdot x) + b^* = 0$, obtaining the decision function

$$f(x) = \text{sgn}(g(x)) ,$$

where

$$g(x) = (w^* \cdot x) + b^* = \sum_{i=1}^m y_i \alpha_i^* (x_i \cdot x) + b^*. \quad (2.23)$$

2.4 Support Vector

Definition 2.11 (Support vector)

Suppose that $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$ is a solution to the dual problem obtained using Algorithm (2.10). The input x_i , associated with the training point (x_i, y_i) , is said to be a *support vector* if the corresponding component α_i^* of α^* is nonzero and otherwise it is a *nonsupport vector*.

The next theorem characterizes the support vectors.

Theorem 2.12 Suppose that the linearly separable problems are solved using Algorithm (2.10) and that the $g(x)$ defined by (2.23). Then

- 1) support vector x_i satisfies $y_i g(x_i) = y_i((w^* \cdot x_i) + b^*) = 1$.
- 2) nonsupport vector x_i satisfies $y_i g(x_i) = y_i((w^* \cdot x_i) + b^*) > 1$.

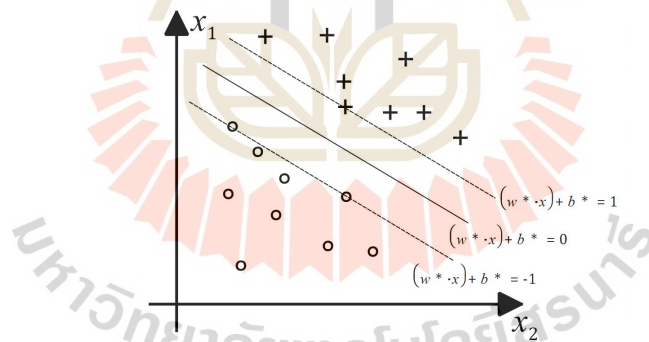


Figure 2.4 Geometric interpretation of Theorem 2.12.

2.5 Linear C-Support Vector Classification

2.5.1 Maximal Margin Method

For a general classification problem, which may be linearly nonseparable problem, it is possible that any hyperplane is unable to separate all of the positive

and negative input correctly. If we still want to use a hyperplane as a separator, we need to adopt the following two strategies: On one hand, in order to relax the requirement to separate all of the inputs correctly, allow the existence of training points that violate the constrains $y_i((w \cdot x_i) + b) \geq 1$ by the introducing slack variables

$$\xi_i \geq 0, \quad i = 1, 2, 3, \dots, m,$$

then yielding constraints

$$y_i((w \cdot x_i) + b) \geq 1 - \xi_i, \quad i = 1, 2, 3, \dots, m.$$

On the other hand in order to make the above violation as small as possible and avoid making ξ too large, one superimposes a penalty upon them in the objective function, For instance, we can add a term

$$\sum_{i=1}^m \xi_i$$

to the objective function. We obtain the primal problem (2.13) ~ (2.14) as

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i, \quad (2.24)$$

subject to

$$y_i((w \cdot x_i) + b) \geq 1 - \xi_i, \quad i = 1, 2, 3, \dots, m, \quad (2.25)$$

$$\xi_i \geq 0, \quad i = 1, 2, 3, \dots, m, \quad (2.26)$$

where $\xi = (\xi_1, \xi_2, \dots, \xi_m)^T$, and $C > 0$ is a penalty parameter.

Algorithm 2.13 (Linear maximal margin method)

Solve the primal problem (2.24) ~ (2.26), obtaining a solution (w^*, b^*, ξ^*) . Then construct the separating hyperplane and the corresponding decision function $f(x) = \text{sgn}(g(x))$, where $g(x) = (w^* \cdot x) + b^*$.

2.5.2 Properties of the Maximal Margin Method in Linear C-Support Vector Classification

Theorem 2.14 There exists a solution to the primal problem (2.24) \sim (2.26), with respect to (w, b) .

Theorem 2.15 The solution w^* of the primal problem (2.24) \sim (2.26) with respect to w is unique.

2.6 The Dual Problem in Linear C-Support Vector Classification

To derive the dual problem of the primal problem (2.24) \sim (2.26), we introduce the Lagrange function;

$$L(w, b, \xi, \alpha, \beta) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i - \sum_{i=1}^m \alpha_i (y_i ((w \cdot x_i) + b) - 1 + \xi_i) - \sum_{i=1}^m \beta_i \xi_i,$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$ and $\beta = (\beta_1, \beta_2, \dots, \beta_m)^T$ are *Lagrange multiplier vectors*.

Then we have the following theorem.

Theorem 2.16 (Optimization problem in Linear C-support vector)

$$\max_{\alpha, \beta} -\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) + \sum_{j=1}^m \alpha_j, \quad (2.27)$$

subject to

$$\sum_{i=1}^m y_i \alpha_i = 0, \quad (2.28)$$

$$C - \alpha_i - \beta_i = 0, \quad i = 1, 2, 3, \dots, m, \quad (2.29)$$

$$\alpha_i \geq 0, \quad i = 1, 2, 3, \dots, m, \quad (2.30)$$

$$\beta_i \geq 0, \quad i = 1, 2, 3, \dots, m, \quad (2.31)$$

is the dual problem of the primal problem (2.24) ~ (2.26).

Theorem 2.17 The dual problem (2.27) ~ (2.31) has a solutions.

The dual problem (2.27) ~ (2.31) can be simplified to a problem only for a single variable α by eliminating the variable β and then rewriting as a minimization problem:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m y_i y_j (x_i \cdot x_j) \alpha_i \alpha_j - \sum_{j=1}^m \alpha_j, \quad (2.32)$$

subject to

$$\sum_{i=1}^m y_i \alpha_i = 0, \quad (2.33)$$

$$0 \leq \alpha_i \leq C, i = 1, 2, 3, \dots, m. \quad (2.34)$$

Theorem 2.18 Suppose that $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$ is any solution to the convex quadratic program (2.32) ~ (2.34). If there exist a component of α^* , α_j^* , such that $\alpha_j^* \in (0, C)$, then a solution (w^*, b^*) to the primal problem (2.24) ~ (2.26) with respect to (w, b) , can be obtained by

$$w^* = \sum_{i=1}^m \alpha_i^* y_i x_i,$$

$$b^* = y_j - \sum_{i=1}^m \alpha_i^* y_i (x_i \cdot x_j).$$

Algorithm 2.19 (Linear C-support vector classification)

Step 1) Input the training set $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$, where $x_i \in \mathbb{R}^n$, $y_i \in \mathbb{Y} = \{-1, 1\}$, $i = 1, 2, 3, \dots, m$;

Step 2) Choose an appropriate penalty parameter $C > 0$;

Step 3) Construct and solve the convex quadratic program (2.32) ~ (2.34) and we obtain a solution $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_m^*)^T$;

Step 4) Compute b^* : Choose a positive component of α^* , $\alpha_j^* \in (0, C)$, and compute

$$b^* = y_j - \sum_{i=1}^m \alpha_i^* y_i (x_i \cdot x_j);$$

Step 5) Construct the decision function

$$f(x) = \text{sgn}\left(\sum_{i=1}^m y_i \alpha_i^* (x_i \cdot x) + b^*\right).$$

2.7 Linear Regression

2.7.1 Regression Problem

Given a training set

$$\tilde{T} = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\},$$

where $x_i \in \mathbb{R}^n$, $y_i \in \mathbb{Y} = \mathbb{R}$, $i = 1, 2, 3, \dots, m$, the regression problem consists of finding $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ in \mathbb{R}^n , to derive the value of y for any x_i by the function:

$$y_i = g(x_i).$$

Particularly, when the function $g(x)$ is restricted to be a linear function

$$y = g(x) = (w \cdot x) + b,$$

the corresponding problem is defined as the linear regression problem.

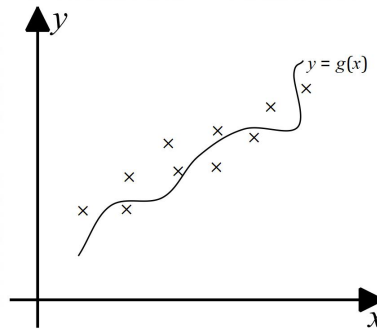


Figure 2.5 A regression problem in \mathbb{R} .

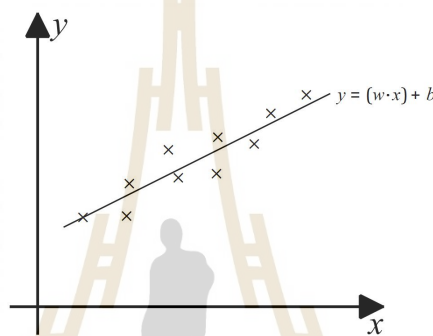


Figure 2.6 A linear regression problem in \mathbb{R} .

2.7.2 Hard $\bar{\epsilon}$ -Band Hyperplane

In order to solve the regression problem, firstly we will introduce the definitions of $\bar{\epsilon}$ -band and hard $\bar{\epsilon}$ -band hyperplane.

Definition 2.20 ($\bar{\epsilon}$ -band of a hyperplane)

For given $\bar{\epsilon} > 0$, the $\bar{\epsilon}$ -band of a hyperplane $y = (w \cdot x) + b$ is the set

$$\{(x, y) | (w \cdot x) + b - \bar{\epsilon} < y < (w \cdot x) + b + \bar{\epsilon}\}$$

.

Definition 2.21 (Hard $\bar{\epsilon}$ -band of a hyperplane)

For given $\bar{\epsilon} > 0$ and a training set \tilde{T} , we say that a hyperplane $y = (w \cdot x) + b$ is the *hard $\bar{\epsilon}$ -band hyperplane* for a training set \tilde{T} , if all the training points are inside

its $\bar{\varepsilon}$ -band, i.e. the hyperplane $y = (w \cdot x) + b$ satisfies that

$$-\bar{\varepsilon} < y_i - ((w \cdot x_i) + b) < \bar{\varepsilon}, \quad i = 1, 2, \dots, m.$$

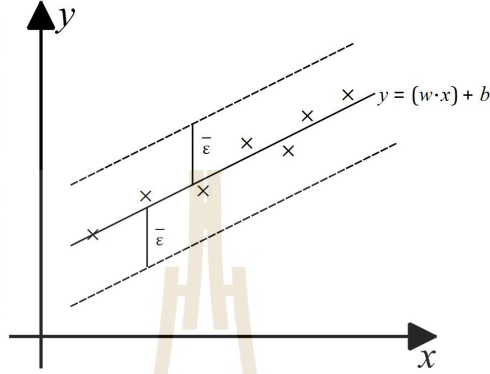


Figure 2.7 A hard $\bar{\varepsilon}$ -band hyperplane (line) in \mathbb{R} .

Consider the hard $\bar{\varepsilon}$ -band hyperplane for any training set \tilde{T} . When $\bar{\varepsilon}$ is large enough, there always exists a hard $\bar{\varepsilon}$ -band hyperplane, since the number of the training points is finite. And the value of $\bar{\varepsilon}$ corresponding to a hard $\bar{\varepsilon}$ -band hyperplane should be not too small, it should be larger than the optimal value ε_{inf} of the following optimization problem:

$$\min_{w, b, \bar{\varepsilon}} \bar{\varepsilon}, \quad (2.35)$$

subject to

$$-\bar{\varepsilon} \leq y_i - ((w \cdot x_i) + b) \leq \bar{\varepsilon}, \quad i = 1, 2, \dots, m. \quad (2.36)$$

2.7.3 Hard $\bar{\varepsilon}$ -Band Hyperplane and Linear Classification

In this section, we construct a hard $\bar{\varepsilon}$ -band hyperplane using the classification method. We construct two classes based on the training set \tilde{T} by adding

and subtracting $\bar{\varepsilon}$ to y of every training points, then we obtain two sets of the positive and negative points respectively:

$$D^+ = \{(x_i^T, y_i + \bar{\varepsilon})^T, i = 1, 2, 3, \dots, m\}, \quad (2.37)$$

$$D^- = \{(x_i^T, y_i - \bar{\varepsilon})^T, i = 1, 2, 3, \dots, m\}. \quad (2.38)$$

Then, the training set for classification is

$$\{((x_1^T, y_1 + \bar{\varepsilon})^T, 1), \dots, ((x_m^T, y_m + \bar{\varepsilon})^T, 1), ((x_1^T, y_1 - \bar{\varepsilon})^T, -1), \dots, ((x_m^T, y_m - \bar{\varepsilon})^T, -1)\},$$

where $(x_i^T, y_i + \bar{\varepsilon})^T$ or $(x_i^T, y_i - \bar{\varepsilon})^T$ represents the input, and the last component, 1 or -1 , represents the output. The problem of constructing a hard $\bar{\varepsilon}$ -band hyperplane is equivalent to linearly separating the above training sets .

Theorem 2.22 For a given training set \tilde{T} and $\bar{\varepsilon} > 0$, a hyperplane $y = (w \cdot x) + b$ is a hard $\bar{\varepsilon}$ -band hyperplane if and only if the set D^+ and D^- defined by (2.37) ~ (2.38) are located on different sides of this hyperplane respectively, and neither D^+ nor D^- touch this hyperplane.

2.7.4 Optimization Problem of Constructing a Hard $\bar{\varepsilon}$ -Band Hyperplane

In this section, we try to construct a hard $\bar{\varepsilon}$ -band hyperplane. For the case $\bar{\varepsilon} > \varepsilon_{\text{inf}}$, where ε_{inf} is the optimal value of the problem (2.35) ~ (2.36). We can derive the optimization problem of constructing a hard $\bar{\varepsilon}$ -band hyperplane, according to the maximal margin method for a linearly separable problem.

The classification problem is in \mathbb{R}^{n+1} . Assume that the hyperplane is $(w \cdot x) + \eta y + b = 0$, where the normal vector is $(w^T, \eta)^T$, $w \in \mathbb{R}^n$ corresponds to x , $\eta \in \mathbb{R}$ corresponds to y .

Similar to the classification problem (2.10) ~ (2.12), we get following programming problem with respect to (w, η, b)

$$\min_{w, \eta, b} \frac{1}{2} \|w\|^2 + \frac{1}{2} \eta^2, \quad (2.39)$$

subject to

$$(w \cdot x_i) + \eta(y_i + \bar{\varepsilon}) + b \geq 1 \quad ; i = 1, 2, \dots, m, \quad (2.40)$$

$$(w \cdot x_i) + \eta(y_i - \bar{\varepsilon}) + b \leq -1 \quad ; i = 1, 2, \dots, m. \quad (2.41)$$

Then the separating hyperplane is

$$(\bar{w} \cdot x) + \bar{\eta}y + \bar{b} = 0, \quad (2.42)$$

where $(\bar{w}, \bar{\eta}, \bar{b})$ is the solution to the problem (2.40) ~ (2.42), the regression problem is

$$y = (w^* \cdot x) + b^*, \quad (2.43)$$

where

$$w^* = -\frac{\bar{w}}{\bar{\eta}}, \quad b^* = -\frac{\bar{b}}{\bar{\eta}}. \quad (2.44)$$

Theorem 2.23 Suppose that $(\bar{w}, \bar{\eta}, \bar{b})$ is the solution to the problem (2.40) ~ (2.42) then $\bar{\eta} \neq 0$. Furthermore, let

$$\varepsilon = \bar{\varepsilon} - \frac{1}{\bar{\eta}},$$

then

1) ε satisfies

$$\varepsilon_{\inf} \leq \varepsilon < \bar{\varepsilon}_i,$$

where ε_{\inf} is the optimal value of the problem (2.35) ~ (2.36);

2) $(w^*, b^*) = (-\frac{\bar{w}}{\bar{\eta}}, -\frac{\bar{b}}{\bar{\eta}})$ is the solution to the following problem:

$$\min_{w,b} \frac{1}{2} \|w\|^2, \quad (2.45)$$

subject to

$$(w \cdot x_i) + b - y_i \leq \varepsilon \quad ; i = 1, 2, \dots, m, \quad (2.46)$$

$$y_i - (w \cdot x_i) - b \leq \varepsilon \quad ; i = 1, 2, \dots, m. \quad (2.47)$$

2.8 Linear Hard ε -Band Support Vector Regression

2.8.1 Primal Problem

The Primal problem is the problem (2.45) ~ (2.47)

Theorem 2.24 Suppose that ε_{inf} , is the optimal value of the following problem

$$\min_{w,b,\varepsilon} \varepsilon, \quad (2.48)$$

subject to

$$-\varepsilon \leq y_i - ((w \cdot x_i) + b) \leq \varepsilon, \quad i = 1, 2, 3, \dots, m, \quad (2.49)$$

If $\varepsilon > \varepsilon_{\text{inf}}$, then the primal problem (2.45) ~ (2.47) has solution, and the solution with respect to w is unique.

2.8.2 Relationship between the Primal and Dual Problems

In order to derive the dual problem, we introduce the Lagrange function

$$L(w, b, \alpha^{(*)}) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i (\varepsilon + y_i - (w \cdot x_i) - b) - \sum_{i=1}^m \alpha_i^* (\varepsilon - y_i + (w \cdot x_i) + b),$$

where $\alpha^{(*)} = (\alpha_1, \alpha_1^*, \alpha_2, \alpha_2^*, \dots, \alpha_m, \alpha_m^*)^T \in \mathbb{R}^{2m}$ is the *Lagrange multiplier vector*,

and $(*)$ is a shorthand implying both the vector with and without asterisks.

Theorem 2.25

$$\max_{\alpha^{(*)} \in \mathbb{R}^{2m}} -\frac{1}{2} \sum_{i,j=1}^m (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) - \varepsilon \sum_{i=1}^m (\alpha_i^* + \alpha_i) + \sum_{i=1}^m y_i (\alpha_i^* - \alpha_i) , \quad (2.50)$$

subject to

$$\sum_{i=1}^m (\alpha_i^* - \alpha_i) = 0 , \quad (2.51)$$

$$\alpha_i^{(*)} \geq 0, \quad i = 1, 2, \dots, m , \quad (2.52)$$

is the dual problem of the primal problem (2.45) ~ (2.47).

Theorem 2.26 If $\varepsilon > \varepsilon_{\text{inf}}$, then the dual problem (2.50) ~ (2.52) has a solution, where ε_{inf} is the optimal value of the problem (2.48) ~ (2.49).

Algorithm 2.27 (Linear hard $\bar{\varepsilon}$ -band support vector regression)

Step 1) Input the training set \tilde{T} ;

Step 2) Choose the parameter $\varepsilon > 0$;

Step 3) Construct and solve the convex quadratic programming

$$\min_{\alpha^{(*)} \in \mathbb{R}^{2m}} \frac{1}{2} \sum_{i,j=1}^m (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) + \varepsilon \sum_{i=1}^m (\alpha_i^* + \alpha_i) - \sum_{i=1}^m y_i (\alpha_i^* - \alpha_i) , \quad (2.53)$$

subject to

$$\sum_{i=1}^m (\alpha_i^* - \alpha_i) = 0 , \quad (2.54)$$

$$\alpha_i^{(*)} \geq 0, \quad i = 1, 2, \dots, m , \quad (2.55)$$

and we obtain the solution $\bar{\alpha}^{(*)} = (\bar{\alpha}_1, \bar{\alpha}_1^*, \bar{\alpha}_2, \bar{\alpha}_2^*, \dots, \bar{\alpha}_m, \bar{\alpha}_m^*)^T$;

Step 4) Compute

$$\bar{w} = \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i) x_i .$$

Choose a positive component of $\bar{\alpha}^{(*)}$, $\bar{\alpha}_j > 0$, then compute $\bar{b} = y_j - (\bar{w} \cdot x_j) + \varepsilon$;
 or choose a positive component of $\bar{\alpha}^{(*)}$, $\bar{\alpha}_k^* > 0$, then compute $\bar{b} = y_k - (\bar{w} \cdot x_k) - \varepsilon$;

Step 5) Construct the regression function

$$y = g(x) = (\bar{w} \cdot x) + \bar{b} = \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i)(x_i \cdot x) + \bar{b}.$$

Definition 2.28 Suppose that $\bar{\alpha}^{(*)} = (\bar{\alpha}_1, \bar{\alpha}_1^*, \bar{\alpha}_2, \bar{\alpha}_2^*, \dots, \bar{\alpha}_m, \bar{\alpha}_m^*)^T$ is a solution to the dual problem (2.53) ~ (2.55) obtain by Algorithm 2.8.1 The input (x_i, y_i) is said to be a *support vector* if the corresponding component $\bar{\alpha}_i$ or $\bar{\alpha}_i^*$ is nonzero, otherwise it is the *nonsupport vector*.

Theorem 2.29 Suppose that $\bar{\alpha}^{(*)} = (\bar{\alpha}_1, \bar{\alpha}_1^*, \bar{\alpha}_2, \bar{\alpha}_2^*, \dots, \bar{\alpha}_m, \bar{\alpha}_m^*)^T$ is the solution to problem (2.53) ~ (2.55) for $i = 1, 2, 3, \dots, m$, there exists only one nonzero component between $\bar{\alpha}_i$ and $\bar{\alpha}_i^*$.

Theorem 2.30 Suppose that $\bar{\alpha}^{(*)} = (\bar{\alpha}_1, \bar{\alpha}_1^*, \bar{\alpha}_2, \bar{\alpha}_2^*, \dots, \bar{\alpha}_m, \bar{\alpha}_m^*)^T$ is a solution to problem (2.53) ~ (2.55) solved by Algorithm 2.4.5.2.3 and $y = (\bar{w} \cdot x) + \bar{b}$ is the regression function obtained by Algorithm 2.4.5.2.3. If $\varepsilon > \varepsilon_{\text{inf}}$, where ε_{inf} is the optimal value of the problem (2.48) ~ (2.49), then

- 1) All support vector are on the boundary of the ε - band of hyperplane $y = (\bar{w} \cdot x) + \bar{b}$;
- 2) All nonsupport vector are inside or on the boundary of the ε - band of hyperplane $y = (\bar{w} \cdot x) + \bar{b}$.

2.9 Linear ε -Support Vector Regression

2.9.1 Primal Problem

Similar to support vector classification, the primal problem of slack variable $\xi^{(*)} = (\xi_1, \xi_1^*, \xi_2, \xi_2^*, \dots, \xi_m, \xi_m^*)^T$ and penalty parameter C , the primal problem of linear ε -support vector regression can be written as:

$$\min_{w, b, \xi^{(*)}} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*), \quad (2.56)$$

subject to

$$((w \cdot x_i) + b) - y_i \leq \varepsilon + \xi_i, \quad i = 1, 2, 3, \dots, m, \quad (2.57)$$

$$y_i - ((w \cdot x_i) + b) \leq \varepsilon + \xi_i^*, \quad i = 1, 2, 3, \dots, m, \quad (2.58)$$

$$\xi_i^{(*)} \geq 0, \quad i = 1, 2, 3, \dots, m, \quad (2.59)$$

where $(*)$ is a shorthand implying both the vector with and without asterisks.

This problem is a convex quadratic programming.

After obtaining the solution to the primal problem (2.56) \sim (2.59), $(\bar{w}, \bar{b}, \bar{\xi}^{(*)})$, then construct the regression function:

$$y = g(x) = (\bar{w} \cdot x) + \bar{b}.$$

Theorem 2.31 There exist solution to the primal problem (2.56) \sim (2.59) with respect to (w, b) and the solution with respect to w is unique.

2.9.2 Relationship between Dual Problem and Primal Problem

In order to derive the dual problem of the primal problem (2.56) ~ (2.59), we introduce the Lagrange function

$$L(w, b, \xi^{(*)}, \alpha^{(*)}, \eta^{(*)}) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) - \sum_{i=1}^m (\eta_i \xi_i + \eta_i^* \xi_i^*) - \sum_{i=1}^m \alpha_i (\varepsilon + \xi_i + y_i - (w \cdot x_i) - b) - \sum_{i=1}^m \alpha_i^* (\varepsilon + \xi_i^* - y_i + (w \cdot x_i) + b),$$

where $\alpha^{(*)} = (\alpha_1, \alpha_1^*, \alpha_2, \alpha_2^*, \dots, \alpha_m, \alpha_m^*)^T$, $\eta^{(*)} = (\eta_1, \eta_1^*, \eta_2, \eta_2^*, \dots, \eta_m, \eta_m^*)^T$ are the *Lagrange multiplier vectors*.

Theorem 2.32

$$\max_{\alpha^{(*)}, \eta^{(*)} \in \mathbb{R}^{2m}} -\frac{1}{2} \sum_{i,j=1}^m (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) - \varepsilon \sum_{i=1}^m (\alpha_i^* + \alpha_i) + \sum_{i=1}^m y_i (\alpha_i^* - \alpha_i), \quad (2.60)$$

subject to

$$\sum_{i=1}^m (\alpha_i^* - \alpha_i) = 0, \quad (2.61)$$

$$C - \alpha_i^{(*)} - \eta_i^{(*)} = 0, \quad i = 1, 2, 3, \dots, m, \quad (2.62)$$

$$\alpha_i^{(*)} \geq 0, \eta_i^{(*)} \geq 0, \quad i = 1, 2, \dots, m, \quad (2.63)$$

is the dual problem of the primal problem (2.56) ~ (2.59).

Theorem 2.33 Dual problem (2.60) ~ (2.63) has solutions.

Theorem 2.34 Suppose that $\bar{\alpha}^{(*)} = (\bar{\alpha}_1, \bar{\alpha}_1^*, \bar{\alpha}_2, \bar{\alpha}_2^*, \dots, \bar{\alpha}_m, \bar{\alpha}_m^*)^T$ is any solution to the problem

$$\min_{\alpha^{(*)} \in \mathbb{R}^{2m}} \frac{1}{2} \sum_{i,j=1}^m (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) + \varepsilon \sum_{i=1}^m (\alpha_i^* + \alpha_i) - \sum_{i=1}^m y_i(\alpha_i^* - \alpha_i) , \quad (2.64)$$

subject to

$$\sum_{i=1}^m (\alpha_i^* - \alpha_i) = 0 , \quad (2.65)$$

$$0 \leq \alpha_i^{(*)} \leq C, \quad i = 1, 2, \dots, m , \quad (2.66)$$

If there exists a component of $\bar{\alpha}^{(*)}$, $\bar{\alpha}_j \in (0, C)$ or $\bar{\alpha}_k^* \in (0, C)$, then a solution (\bar{w}, \bar{b}) to the primal problem (48) ~ (51) with respect to (w, b) can be obtained by

$$\begin{aligned} \bar{w} &= \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i) x_i , \\ \bar{b} &= y_j - \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i)(x_i \cdot x_j) + \varepsilon , \end{aligned}$$

or

$$\bar{b} = y_k - \sum_{i=1}^m (\bar{\alpha}_i^* - \bar{\alpha}_i)(x_i \cdot x_k) - \varepsilon .$$

Algorithm 2.35 (Linear ε - support vector regression)

Step 1) Input the training set \tilde{T} ;

Step 2) Choose an appropriate parameter ε the penalty parameter $C > 0$;

Step 3) Construct and solve the convex quadratics programming (2.64) ~ (2.66);

Step 4) Compute \bar{b} : choose a component of $\bar{\alpha}^{(*)}$ in the interval $(0, C)$.

If the component $\bar{\alpha}_j \in (0, C)$, compute

$$\bar{b} = y_j - \sum_{i=1}^m (\alpha_i^* - \alpha_i)(x_i \cdot x_j) + \varepsilon ;$$

If the component $\bar{\alpha}_k^* \in (0, C)$, compute

$$\bar{b} = y_k - \sum_{i=1}^m (\alpha_i^* - \alpha_i)(x_i \cdot x_k) - \varepsilon ;$$

Step 5) Construct the decision function

$$y = g(x) = \sum_{i=1}^m (\alpha_i^* - \alpha_i)(x_i \cdot x) + \bar{b}.$$

2.10 Kernel

In case of nonlinear problem, we can use the kernel function to solve problems.

Definition 2.36 (Kernel function)

A function $K(x, y)$ is called a *kernel* on $\mathbb{R}^n \times \mathbb{R}^n$ if there exists a map Φ from the space \mathbb{R}^n to Hilbert space \mathcal{H} ,

$$\begin{aligned} \Phi : \mathbb{R}^n &\rightarrow \mathcal{H} \\ x &\mapsto \Phi(x) \end{aligned}$$

such that

$$K(x, y) = (\Phi(x) \cdot \Phi(y))$$

where (\cdot) denotes the inner product of space \mathcal{H} .

2.10.1 Commonly Used Kernels

- Polynomial kernel

$$K(x, y) = ((x \cdot y) + 1)^d \tag{2.67}$$

where d is the degree of the polynomial.

- Gaussian kernel

$$K(x, y) = e^{-\frac{\|x-y\|^2}{2\sigma^2}}. \tag{2.68}$$

- **Gaussian radial basis function kernel**

$$K(x, y) = e^{-\gamma \|x-y\|^2}. \quad (2.69)$$

2.11 Root Mean Square Error

Root Mean Square Error is a commonly used measure of the difference between the observed value Y_i and the corresponding fitted value \hat{Y}_i ,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N \varepsilon_i^2}{N}} \quad (2.70)$$

where ε_i is the residual and N is a number of observations.

2.12 Related Researches

Boonyanam and Mungmaiphol (2019) analyzed the factors that affect the exchange rate between the Thai Baht and Japanese Yen. They found that, in the long term, the inflation differential between Thailand and Japan (DINF), the monthly import of Thailand (RXM) and the Thai Business Sentiment (BSI), Correlate positively with the exchange rate (Yen v.s. Bath). On the other hand, the real effective exchange rate (REER) and the manufacturing production index (MPI) correlate negatively.

Chanaim, Rungruang, Srichaikul and Sriboonchitta (2019) used support vector machine to investigate the predictability of stock index movement direction with SVM by forecasting the daily movement direction of the SET 50 index over the period 5 April, 2000 to 22 August, 2018. The experiment a results show

that SVM with autoregressive lag $p = 10$ and training data equal 37 have accuracy(ACC) 92.56%.

Grigoryan (2016) studied the financial time series prediction problem to predict stock market, by support vector machines (SVM) with independent component analysis (ICA) method (or SVM-ICA) and SVM method. He used the root mean square error (RMSE) measure to evaluate the performance of the model and found that the SVM-ICA method is better than the SVM method for nonstationary time series.

Chih-Feng Chao and Ming-Huwi Horng (2015) used the firefly-based SVM (firefly-SVM) for exploring binary and multiclass classification. Based on the result of the current experiments on the binary classification of ten data sets of the UCI database and multiclass classification of ultrasonic supraspinatus images, the experimental results advocate the use of fireflySVM to classify the pattern classifications for maximum accuracy.

Madge (2015) studied forecasting the stock price, using the daily closing prices for 34 technology stocks to calculate price volatility and momentum for individual stocks and for the overall sector. The result shows that SVM model is better suitable for the long-run than the short-run.

B.Yu, Han, Yan and Ning (2014) used the support vector machine for forecasting Dry Bulk Freight Index. The BDI data for 2005 to 2012 are presented to test the model. The result shows that the proposed method has higher accuracy and can be used to forecast the short-term trend of the BDI.

K. Okasha (2014) used the support vector machines in financial time series, the autoregressive integrated moving average (ARIMA), the artificial neural networks (ANN) to forecast the A1-Quds index of the Palestinian stock exchange market time series data. The result shows that the SVM model is better than the

ARIMA and ANN models.

Mu, Wu and Zhang (2014) used the support vector machine (SVM), least squares support vector machine (LSSVM) and partial least squares (PLS) method, to predict house prices in a Boston suburb. The result shows that the SVM method is better in prediction than LSSVM and PLS method.



CHAPTER III

RESEARCH METHODOLOGY

This chapter presents the process used in this research. The process consists of 6 parts, as follows:

3.1 Collect Training Data and Analyze Descriptive Statistics

This research used monthly data (secondary time series data) from January 2009 until December 2018, a total of 120 months. This section will analyze the descriptive statistics of the training data.

Table 3.1 Shows Training data.

variables	Type	Unit
Exchange Rate of the Thai Baht against the Chinese Yuan	Dependent	Baht/Yuan
Policy Interest rate	independent	Percent/Year
Value of Imports from China	independent	Billion Baht
Value of Exports to China	independent	Billion Baht
Foreign Exchange Reserves	independent	Ten Billion Dollar
Nominal Effective Exchange Rate	independent	-

The data in Table 3.1 was saved in the CSV file (data.csv), CSV file is available for working with program R, the detail of data.csv will be imported and

program R will analyze descriptive statistics shown in Appendix A.

3.2 Program R and Packages Installation

Program R is a free software environment for statistical computing and graphics. This research installed program R (Version 3.6.3) and installed its packages. The complete installed package in program R shown in Appendix A.

3.3 Analyzing Independent Factors that Affect the Exchange Rate

This research used the Multiple Linear Regression to analyze the independent factors that affect the exchange rate of the Thai Baht against the Chinese Yuan. This section divided into 3 parts, as follows:

3.3.1 Study Relationship between Dependent and Independent Variables

To construct the Multiple Linear Regression Model. The model will be obtained for analyzing the independent factors that affect the exchange rate is the following:

$$EXC = \beta_0 + \beta_1(IV) + \beta_2(EV) + \beta_3(PR) + \beta_4(FER) + \beta_5(NEER) \quad (3.1)$$

Where

EXC is the exchange rate of the Thai Baht against the Chinese Yuan.

IV is Value of Import from China.

EV is Value of Exports to China.

PR is Policy Interest Rate.

FER is Foreign Exchange Reserves.

NEER is Nominal Effective Exchange Reserves.

$\beta_0, \beta_1, \dots, \beta_5$ are parameters or regression coefficient.

3.3.2 Estimating Parameter of Model

This research used the least square method to estimate the parameters in the model (3.1).

3.3.3 Test for Significance of Regression

The test for Significance of regression is a test to determine if there is a linear relationship between the dependent and any of the regressor variables $X_1, X_2, X_3, \dots, X_n$.

Assume that the estimator parameter $\beta_0, \beta_1, \dots, \beta_5$ are b_0, b_1, \dots, b_5 respectively.

The statement for the hypotheses are:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_5 = 0$$

$$H_1 : \beta_i \neq 0 \text{ for at least one } i$$

The statistic that can be used for testing H_0 versus H_1 is the t test statistic in the following form:

$$t^* = \frac{b_i}{\text{se}(b_i)} \sim t_{\frac{\alpha}{2}, n-(k+1)} \quad \forall i = 1, 2, 3, 4, 5$$

where $\text{se}(b_i) = \sqrt{\sigma^2 C_{ii}}$, C_{ii} is an element on the main diagonal of $(X^T X)^{-1}$ and we reject H_0 if $t_{\frac{\alpha}{2}, n-(k+1)} < |t^*|$.

This research used 5% of a significant level ($\alpha = 0.05$), and number of training data equal to 120 ($n=120$). Since $\frac{\alpha}{2} = 0.025$ and $n - (k + 1) = 116$ from the t-Distribution table, then the value of $t_{0.0025, 116}$ is approximately 1.980.

3.4 Creating a Forecasting Model

The purpose of this research was to use the factors, obtained from the regression analysis, to construct an appropriate model for forecasting the exchange rate by using Support Vector Machine (SVM).

Support Vector Machine is the most popular tool for classification and regression and for predicting problem in SVM. We use Support Vector Regression (SVR) to solve this problem; SVM can efficiently perform a non-linear classification and regression by using the kernel function.

Four kernel functions are commonly used.

- ⊙ The linear kernel
- ⊙ The polynomial kernel
- ⊙ The radial basis function kernel

3.5 Predicting

Use the obtained model predicting the exchange rate of the Thai Baht against the Chinese Yuan from January 2019 until June 2019, a total of 6 months.

3.6 Accuracy Measurement of Predicting Model

In data mining, when developing a predictive model, as in linear regression, SVM, SVR, etc, it is important to quantify how well the model is suited to future observations.

The frequently used accuracy measure of forecasting model are:

- ⊙ Root Mean Square Error (RMSE)

- ⊙ Mean Absolute Error (MAE)
- ⊙ Mean Square Error (MSE)

This research use the Root Mean Square Error (RMSE) to measure the efficiency of predicting model.



CHAPTER IV

RESULTS AND DISCUSSION

This chapter presents the results from the process in Chapter III Research Methodology. The aim is to show the consequence computed in program R (Version 3.6.3), the independent factors that affect the exchange rate, the best model, and the predicted values of the exchange rate of the Thai Baht against the Chinese Yuan from January 2019 until June 2019, a total of 6 months.

4.1 Descriptive Statistics of Training Data

Descriptive statistics is a summary statistic that quantitatively describes or summarizes features from the collection of data and help to understand the training data that is essential in part of Machine Learning. This part shows the result of some descriptive statistics of the training data (data.csv).

Table 4.1 Training data Descriptive statistics.

variables	Scale	Max	Min	Mean	Range	Variance
EXC	ratio	5.6483	4.4938	5.03307583	1.1545	0.07102295
PR	ratio	3.50	1.25	1.94791667	2.25	0.41060487
IV	ratio	16.0832168	3.2892282	10.4436373	12.7939886	8.6547762
FER	ratio	2.15614700	1.107222	1.71045380	1.048925	0.05149714
EV	ratio	13.5195499	4.9475996	9.9501433	8.5719503	2.5706560

Table 4.1 (Continued) Training data Descriptive statistics.

variables	Scale	Max	Min	Mean	Range	Variance
NEER	ratio	17.32	9.255	10.50624	2.477	3.304180

4.2 Factors that Affect the Exchange Rate

The next step is on analysis of those factors that affect the Thai Baht versus Chinese Yuan exchange rate. Each of the above data were modeled, either individually or on groups, by using Multiple Linear Regression. Those factors that resulted in the lowest root mean square error where singled out to be used in the model. The result showed that Policy Interest Rate, Value of Imports from China, and Foreign Exchange Reserves affect the exchange rate at a larger than 5% statistically significant level. The estimation coefficients by the least square method are shown in Table 4.2 and the root mean square errors of all models shown in Appendix(B).

Table 4.2 Coefficient, t-statistic values(t^*), and standard error of independent variables (5% of a significant level).

variables	coefficient	Std.Error	t-statistic(t^*)
Intercept	5.5927	0.1250	47.6050
PR	0.0596	0.0286	2.0840
IV	0.1020	0.0085	12.0040
FER	-1.2280	0.1124	-10.9230

Then the obtained model is the following:

$$\text{EXC} = 5.5927 + 0.0596(\text{PR}) + 0.1020(\text{IV}) - 1.2280(\text{FER}) \quad (4.1)$$

Where

EXC is the exchange rate of the Thai Baht against the Chinese Yuan.

PR is Policy Interest Rate.

IV is Value of Imports from China.

FER is Foreign Exchange Reserves.

Finally, from Table 4.2, it is concluded that Policy Interest Rate and Value of Imports from China effected the exchange rate of the Thai Baht against the Chinese Yuan in the same direction. Whereas, Foreign Exchange Reserves effected in the opposite direction.

4.3 Creating a Forecasting Model

Use factors that affected the exchange rate from regression analysis to construct a predicting model. The new training data consists of:

Output Data

- The exchange rate of the Thai Baht against the Chinese Yuan

Input Data

- Policy Interest Rate
- Value of Imports from China
- Foreign Exchange Reserves

After that, we used the new training data to construct a predicting model by using the Support Vector Machine Calculated with program R and use the “e107” package.

The penalty parameter(or C parameter) in program R is written as **Cost**. The gamma parameter is set to the default value equal to $\frac{1}{m}$, m is the dimension in Euclidian space of the input data vector. The **Linear- ϵ -support vector**

Table 4.3 Parameter, SVM Type, and Kernel Type of a predicting model.

Parameter	
SVM-Type	Eps-regression
SVM-kernel	Radial basis function
Cost	1
gamma	0.33
epsilon	0.1

regression is one type in SVR, such that works with the epsilon insensitive loss function.

From Table 4.3, the obtained predicting model is a Support Vector Regression(SVR) with radial basis function. Afterward, creating the best model by using program R to compute the proper epsilon and the penalty parameters are commonly used intervals $[0, 1]$ and $[0, 100]$, respectively.

In this research, we used 100 values of the penalty parameters(i.e., $C = 1, 2, 3, \dots, 100$) and 10 values of the epsilon parameters(i.e., $\epsilon = 0, 0.1, 0.2, \dots, 1.0$) to find the proper parameter for predicting model.

Table 4.4 The appropriate parameter of predicting model.

Parameter	
Cost	17
epsilon	0.1

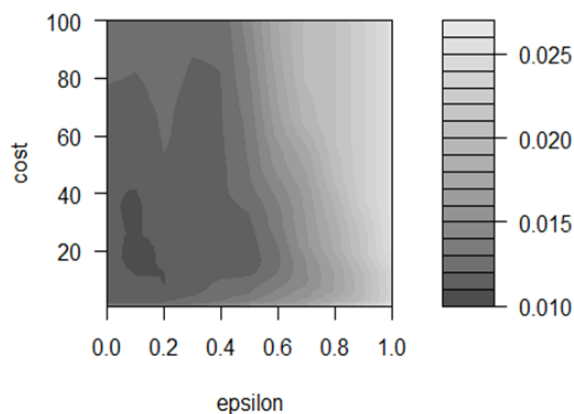


Figure 4.1 Showing the root mean square error depending on epsilon and penalty parameters.

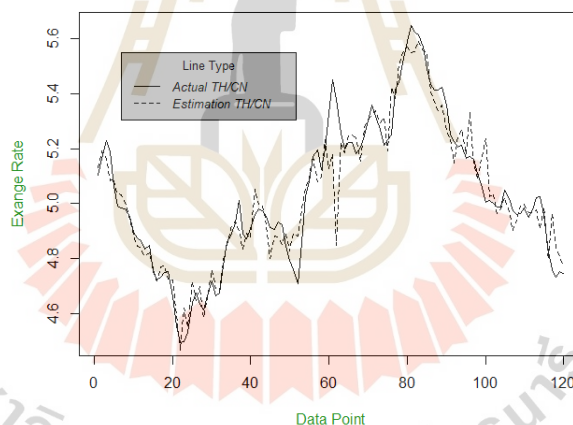


Figure 4.2 The actual and predicted value of training data.

Therefore, the suitable model for predicting the exchange rate is Support Vector Regression with a radial basis function; the gamma parameter is 0.33; the epsilon parameter is 0.1; the penalty parameter is 17 and has Root Mean Square Error equal to 0.0861, the obtained appropriate model in this study is called SVR-THC model.

4.4 Predicting

The result predicts the value of the exchange rate of the Thai Baht against the Chinese Yuan from January 2019 until June 2019 by using the SVR-THC model shown in Table 4.5.

Table 4.5 Predicted values from the SVR-THC and the MLR models.

Month	Actual Value	Predicted Value of SVR-THC model	Predicted Value of MLR model
Jan.2019	4.6807	4.7392	5.1314
Feb.2019	4.6411	4.6065	4.5411
Mar.2019	4.7231	4.7434	4.7270
Apr.2019	4.7390	4.7934	4.9027
May.2019	4.6292	4.7894	4.9382
Jun.2019	4.5066	4.8056	4.7798

Where MLR is Multiple Linear Regression.

4.5 Accuracy Measurement of Predicting Model

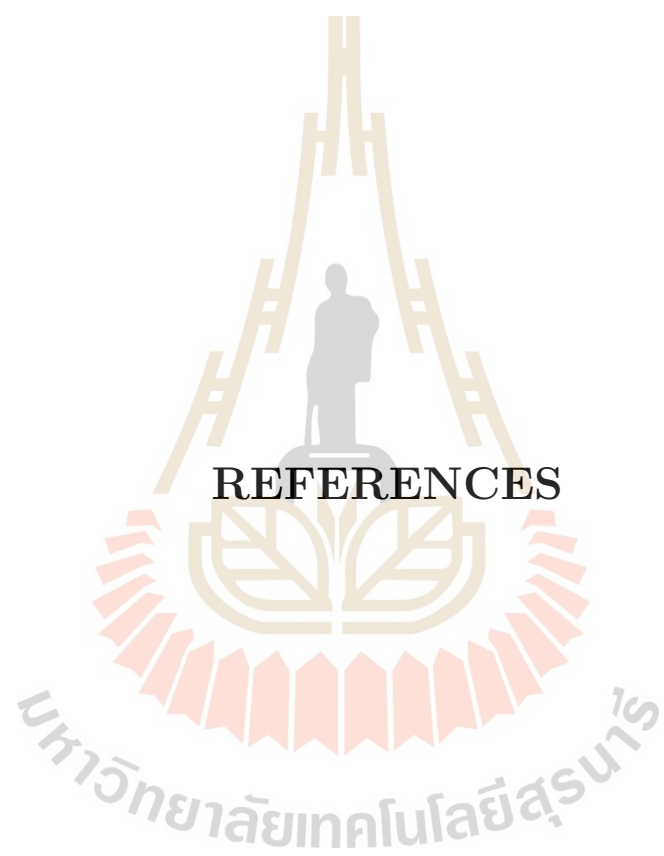
Computing the root mean square error of the above 6 months predictions shows an error of 0.1435 of the SVR-THC Model. On the other hand the Multiple Linear Regression Model shows a substantially larger error of 0.2614.

CHAPTER V

CONCLUSION AND RECOMMENDATION

The purposes of this research were to study the factors that affected the exchange rate of the Thai Baht against the Chinese Yuan by using Multiple Linear Regression, and find an appropriate model for forecasting the exchange rate of the Thai Baht against the Chinese Yuan by using Support Vector Machine(SVM) calculated by Program R software version 3.6.3. The analysis was divided into two steps. The first step analyzes the factors that affect the exchange rate and the result showed that Policy Interest Rate and Value of Imports from China affected the exchange rate of the Thai Baht against the Chinese Yuan in the same direction, whereas Foreign Exchange Reserves in the opposite direction, which were both calculated by 5% statistically significant level. The second step used the factors in the first step to construct an appropriate model for predicting the exchange rate by using Support Vector Machine and the result found that the proper model is Support Vector Regression(SVR) with radial basis function; the gamma parameter is 0.33; the epsilon parameter is 0.1; the penalty parameter is 17. The obtained appropriate model in this study is called the SVR-THC model.

The SVR-THC model predicted the exchange rate of the Thai Baht against the Chinese Yuan from January 2019 until June 2019 with Root Mean Square Error of 0.1435 only. Whereas, the general forecasting of the exchange rate by using Multiple Linear Regression model on the same data has a Root Mean Square Error equal to 0.2614. Therefore, the SVR-THC model is more accurate predictions than the Multiple Linear Regression model.



REFERENCES

REFERENCES

- Bank of Thailand, **RMB Transaction** [On-line]. Available: [https:// www.bot.or.th/English/FinancialMarkets/ForeignExchangeMarket/ LocalCurrencyMarkets/RMBTransaction/Pages/default.aspx](https://www.bot.or.th/English/FinancialMarkets/ForeignExchangeMarket/LocalCurrencyMarkets/RMBTransaction/Pages/default.aspx) .
- Boonyanam, N., and Mungmaiphol, D. (2019). Exchange Rate Forecasting Between Thai Baht And Japanese Yen. **Veridian E-Journal**. 11(3).
- Bo, Y., Yu, B., Ning, G., and Han, Q. (2014). Forecasting Dry Bulk Freight Index with Improved SVM. **Mathematical Problems in Engineering**. (2014,June). doi: 10.1155/2014/460684.
- Chanaim, S., Sriboonchitta, S., Rungruang, C., and Srichaikul, W. (2019). Prediction the Direction of SET50 Index Using Support Vector Machines. **Thai Journal of Mathematics**. (2019): 153-165.
- Chao, C. F., Horng, M. H. (2015). The Construction of Support Vector Machine Classifier Using the Firefly Algorithm. **Computational Intelligence and Neuroscience**. (2015,January). doi: 10.1155/2015/212719.
- Congressional Research Service (2019). **China's Economic Rise: History Trends, Challenges, and Implications for the United States** [On-line]. Available: <https://fas.org/sgp/crs/row/RL33534.pdf> .
- Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining (2012). **Introduction to linear regression analysis**. John Wiley & Sons: Canada.
- Deng, N., Tain, Y., and Zhang, C. (2013). **Support Vector Machines Op-**

timization Based Theory Algorithms and Extensions. Boca Raton: CRC Press.

Gebhardt, K., and Schiller, B. R. (2016). **The Micro Economy Today** (14th ed.). New York: McGraw-Hill.

Grigoyan, H. (2016). A Stock Market Prediction Method Based on Support Vector Machines(SVM) and Independent Component Analysis(ICA). **Database Systems Journal**. 7 (1): 12-21.

Kenton, W. (2019). **Exchange Controls** [On-line]. Available: <https://www.investopedia.com/terms/e/exchangecontrol.asp> .

Madge, S. (2015). **Predicting Stock Price Direction using Support Vector Machines** [On-line]. Available: http://www.cs.princeton.edu/sites/default/files/uploads/saahil_madge.pdf .

Mu, J., Wu, F., and Zhang, A. (2014). Housing Value Forecasting Based on Machine Learning Methods. **Abstract and Applied Analysis**. (2014, August). doi: 10.1155/2014,648047.

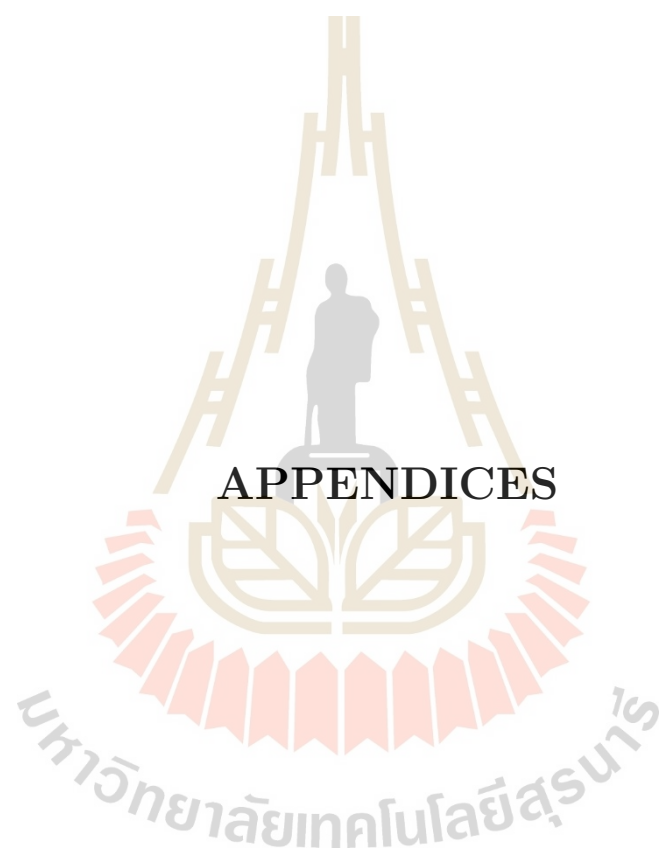
Okasha, K. (2014). Using Support Vector Machines in Financial Time Series Forecasting **International Journal of Statistics and Applications**. 4(1): 28-39.

Pettinger, T. (2011). **List of Government Spending as % of GDP** [On-line]. Available: <https://www.economicshelp.org/blog/3963/economics/list-of-government-spending-as-of-gdp/> .

Trading Economics (2019). **Thailand Balance of Trade** [On-line]. Available: <https://tradingeconomics.com/thailand/balance-of-trade> .

V. Vapnik. (1995). **The Nature of Statistical Learning Theory**. Springer-Verlag: New York .





APPENDICES

The logo of Sakon Nakhon Rajabhat University is a large, faint watermark in the background. It features a central figure of a person standing on a pedestal, surrounded by a circular emblem with a lotus flower. The emblem is topped by a tall, tiered structure resembling a stupa or a traditional Thai architectural element. The entire logo is rendered in a light beige or gold color.

APPENDIX A
ANALYZING DATA BY USING PROGRAM R

มหาวิทยาลัยเทคโนโลยีสุรนารี

The appendix A shown R commands using in this thesis.

A.1 Import Data and Install Packages.

In this step, load and install program R, install packages, and load Training data into program R are following:

- Download and Install Program R

Download Program R Version 3.6.3 available on <https://cran.r-project.org/src/base/R-3/>

- Installation Packages

```
> install.packages("e1071") ( for Support Vector Machine)
```

```
> install.packages("pastecs") ( for Descriptive Statistics)
```

```
> install.packages("CARS") ( for Regression Analysis)
```

- Reading a CSV file in R

```
> Trainingdata = read.csv(file="data.csv")
```

A.2 Analyzing Descriptive Statistics

Descriptive Statistics of training data used library "pastec", as follows:

```
> library("pastec")
```

```
> descriptivestatistic = stat.desc(Trainingdata)
```

```
> descriptivestatistic
```

A.3 Multiple Linear Regression Analysis

Analyzing the independent factors that affected the exchange rate of the Thai Baht against the Chinese Yuan is the following:

```
> library("CARS")
```

```
> model = lm(EXC ~ PR + IV + FER + NEER + EV, data = Trainingdata)
```

A.4 Support Vector Machine

The first step used the new training data to construct a model by using Support Vector Machine is following:

```
> library("e1071")
> newtrainingdata = read.csv(file="newdata.csv")
> modelsvm = svm(EXC ~ PR + IV + FER, data = newtrainingdata)
> print(modelsvm)
```

The second step is to find the proper parameter of the svm model, using 100 values of the penalty parameter(i.e., 1, 2, 3, ..., 100) and 10 values of the epsilon parameter(i.e., 0, 0.1, 0.2, ..., 1.0) by setting Cross-Validation = 3.

```
> bestmodel = tune(svm, EXC ~ PR+IV+FER, data = newtrainingdata,
ranges=list(epsilon=seq(0,1,0.1),
cost=1:100), tunecontrol = tune.control("cross", cross=3))
> print(bestmodel)
> plot(bestmodel)
```

The obtained model(bestmodel) is called the SVR-THC model.

```
> SVR-THC = svm(EXC ~ PR+IV+FER, data = newtrainingdata, kernel =
"radial", epsilon = "0.1", gamma = "0.33", cost = "17")
```

A.5 Predicting

Used independent data that affect the exchange rate to predict the exchange rate of the Thai Baht against the Chinese Yuan from January 2019 until June 2019, a total of 6 months (datapredict.csv).

```
> predictexc = predict(bestmodel,newdata = datapredict)
```

A.6 Performance of model

Used Root Mean Square Error to measure the efficiency of predicting model.

```
> actualvalue = c(4.6807, 4.6411, 4.7231, 4.7390, 4.6292, 4.5066)
```

```
> residual = predictexc - actualvalue
```

```
> sumsquare = sum(residual2)
```

```
> meansumsquare = sumsquare/6
```

```
> rmse = sqrt(meansumsquare)
```

The logo of Sakon Nakhon Rajabhat University is a large, faint watermark in the background. It features a central figure of a person standing on a pedestal, surrounded by a circular emblem with a sunburst design. The text 'มหาวิทยาลัยเทคโนโลยีสุรนารี' is written in Thai script around the bottom of the emblem.

APPENDIX B
ROOT MEAN SQUARE ERROR VALUES OF
MODEL BY USING REGRESSION ANALYSIS

มหาวิทยาลัยเทคโนโลยีสุรนารี

This part shows the Root Mean Square Error of all models from Regression Analysis.

List of Symbols

x_1 is Value of Imports from China.

x_2 is Value of Exports to China.

x_3 is Policy Interest Rate.

x_4 is Foreign Exchange Rate.

x_5 is Nominal Effective Exchange Reserves.

RMSE is Root Mean Square Error.

The bold letter is an independent variable that does not affect the exchange rate of the Thai Baht against the Chinese Yuan.

B.1 Choose one independent variable from 5 variables.

Table B.1 One independent variable.

Independent Variable	RMSE
x_1	0.259
x_2	0.272
x_3	0.260
x_4	0.267
x_5	0.272

B.2 Choose two independent variables from 5 variables.

Table B.2 Two independent variables.

Independent Variable	RMSE
x_1x_2	0.248
x_1x_3	0.250
x_1x_4	0.180
x_1x_5	0.253
x_2x_3	0.261
x_2x_4	0.249
x_2x_5	0.273
x_3x_4	0.260
x_3x_5	0.262
x_4x_5	0.252

B.3 Choose three independent variables from 5 variables.

Table B.3 Three independent variables.

Independent Variable	RMSE
$x_1x_2x_3$	0.243
$x_1\mathbf{x}_2x_3$	0.180

B.3 (Continued) Three independent variables.

Independent Variable	RMSE
$x_1x_2x_5$	0.242
$x_1x_3x_4$	0.179
$x_1x_3x_5$	0.236
$x_1x_3x_5$	0.236
$x_1x_4x_5$	0.178
$x_2x_3x_4$	0.248
$x_2x_3x_5$	0.262
$x_2x_4x_5$	0.240
$x_3x_4x_5$	0.254

B.4 Choose four independent variables from 5 variables.

Table B.4 Four independent variables.

Independent Variable	RMSE
$x_1x_2x_3x_4$	0.180
$x_1x_2x_3x_5$	0.230
$x_1x_2x_4x_5$	0.179
$x_1x_3x_4x_5$	0.179

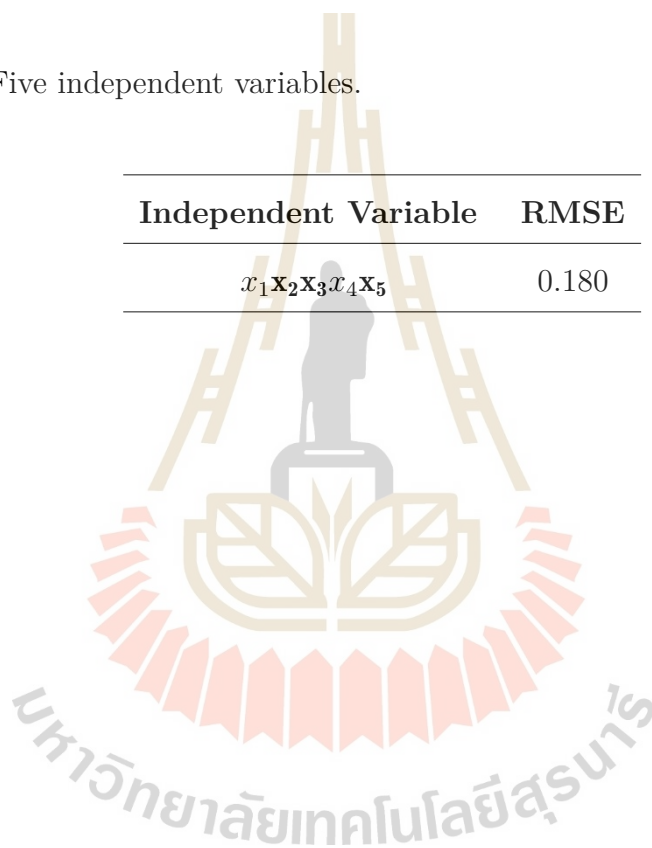
B.4 (Continued) Four independent variables.

Independent Variable	RMSE
$x_1x_2x_3x_5$	0.242

B.5 Choose five independent variables from 5 variables.

Table B.5 Five independent variables.

Independent Variable	RMSE
$x_1x_2x_3x_4x_5$	0.180



CURRICULUM VITAE

NAME : Phakkhaphon Sawatkamon

GENDER : Male

EDUCATION BACKGROUND:

- Bachelor of Science (Mathematics), Suranaree University of Technology, Thailand, 2017

SCHOLARSHIP:

- His Majesty the King's 7th Cycle Birthday Anniversary Suranaree University of Technology Scholarship
- Outstanding Study Performance Scholarship for Graduate Student of Suranaree University of Technology.

CONFERENCE:

- 5th National Conference on Quality Management and Technology Innovation Proceeding, The Eastern University of Management and Technology, Ubon Ratchathani, May 23rd, 2020.

EXPERIENCE:

- Special teacher in Advance Mathematics at Surathampitak School(Intensive Program), Term 1/2019 and Term 2/2019
- Special teacher in Mathematics at Surathampitak School(Mini English Program), Term 2/2019