การเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยเชิงเส้นด้วยแนวคิดเชิงแลตทิซ

นายสุรสิทธิ์ ฤทธิ์สมิตชัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2551

A COMPARISON ON SELECTION CRITERIA FOR LINEAR REGRESSION MODELS BASED ON LATTICE CONCEPT

Mr. Surasit Ritsmitchai

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Statistics Program in Statistics

Department of Statistics

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2008

หัวข้อวิทยานิพนธ์	การเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยเชิงเส้นด้วยแนวคิดเชิงแลตทิซ
โดย	นายสุรสิทธิ์ ฤทธิ์สมิตชัย
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.ธีระพร วีระถาวร

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

....คณบดีคณะพาณิชยศาสตร์และการบัญชี

(รองศาสตราจารย์ ดร.อรรณพ ตันละมัย)

คณะกรรมการสอบวิทยานิพนธ์

....ประธานกรรมการ

(รองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา)

ธีระ ธ..อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(รองศาสตราจารย์ ดร.ธีระพร วีระถาวร)

.... กรรมการ

(รองศาสตราจารย์ ผกาวดี ศิริรังษี)

มนต์ทิพย์ เทียนสุวรรณ.. กรรมการภายนอกมหาวิทยาลัย

(รองศาสตราจารย์ ดร. มนต์ทิพย์ เทียนสุวรรณ)

สุรสิทธิ์ ฤทธิ์สมิตชัย : การเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยเชิงเส้นด้วยแนวคิดเชิงแลตทิซ (A COMPARISON ON SELECTION CRITERIA FOR LINEAR REGRESSION MODELS BASED ON LATTICE CONCEPT). อ. ที่ปรึกษาวิทยานิพนธ์หลัก : รศ.ดร. ธีระพร วีระถาวร, 108 หน้า

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยเชิงเส้น (linear regression models) โดยทำการเปรียบเทียบภายในระแนงตัวแบบ (lattice) เกณฑ์ที่ใช้ในการคัดเลือกตัวแบบมี 2 เกณฑ์ คือ ค่าผลรวมความคลาดเคลื่อนกำลังสอง (Residual Sum of Squares (RSS)) และค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (Mean Square Prediction Error (MSPE)) ซึ่งได้ศึกษาในกรณีที่จำนวนตัวแปรอิสระเท่ากับ 3 และ 4 ความคลาดเคลื่อนมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับศูนย์ และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1, 2, 3 และ 5 โดยใช้ขนาดตัวอย่าง 20, 35 และ 50 ข้อมูลที่ใช้วิจัยจำลองขึ้นในคอมพิวเตอร์โดยใช้เทคนิคมอนติคาร์โล ทำซ้ำ 500 ครั้ง

ผลการวิจัยพบว่า ในกรณีที่ความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่มีค่า $\rho<0.55$ เมื่อใช้เกณฑ์การคัดเลือกตัวแบบเป็นเกณฑ์ RSS หรือ MSPE การคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมด (all possible models) พร้อมกัน และการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบให้ผลการคัดเลือกตัวแบบเหมือนกันคือ ตัวแบบเต็มรูป (full model) ส่วนกรณีที่ความสัมพันธ์ระหว่างตัวแปรอิสระบางคู่มีค่า $\rho \geq 0.55$ การคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบมีโอกาสเลือกตัวแบบลดรูป (reduced model) มากกว่าการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน นอกจากนั้นการใช้เกณฑ์ MSPE มีโอกาสเลือกได้ตัวแบบลดรูปมากกว่าการใช้เกณฑ์ RSS โดยที่เกณฑ์ RSS แปรผันตามกับขนาดตัวอย่างและส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน ส่วนเกณฑ์ MSPE แปรผันตามกับส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่านั้น อย่างไรก็ตามระดับความสัมพันธ์ของตัวแปรอิสระไม่มีผลต่อค่าของเกณฑ์ทั้งสอง

ภาควิชา.............สถิติ....................	ลายมือชื่อนิสิต.......สุรสิทธิ์ ฤทธิ์สมิตชัย..........
สาขาวิชา...........สถิติ....................	ลายมือชื่อ อ. ที่ปรึกษาวิทยานิพนธ์หลัก..............
ปีการศึกษา.........2551.................	

4982249426: MAJOR STATISTICS

KEYWORDS: MOELS SELECTION CRITERION/ LATTICE

SURASIT RITSMITCHAI: A COMPARISON ON SELECTION CRITERIA FOR LINEAR REGRESSION MODELS BASED ON LATTICE CONCEPT. ADVISOR: ASSOC.PROF. THEERAPORN VERATHAWORN, Ph.D., 108 pp.

The objective of this research is to compare two criterions of linear regression models selection by using the lattice method. The two criterions are Residual Sum of Squares (RSS) and Mean Square Prediction Error (MSPE) of which the number of independent variables in this study equal to 3 and 4. The distribution of error is zero mean normal distribution and standard deviation equal to 1, 2, 3 and 5. The sample sizes are 20, 35 and 50. The data in this research are obtained from computer employing Monte Carlo technique for 500 times.

The research shows that when all of correlation coefficients between variables in the model less than 0.55, for using RSS or MSPE criterions, the model which is selected by concurrently comparing all possible models and using the lattice method is the full model. In case of the model which has some two independent variables with correlation coefficient exceeding 0.55, the models selection by using the lattice method is more probable to select reduced model than the models selection by concurrently comparing all possible models. Moreover, MSPE is more probable to select reduced model than RSS. The RSS variation depends on sample sizes and standard deviation, but the MSPE variation just depends only on standard deviation. However, the level of correlation among the independent variables does not affect both criterions.

Department: ....Statistics.........
Field of Study: ....Statistics.........
Academic Year: ......2008.............

Student's Signature. Surasit Ritsmitchai
Advisor's Signature.......................

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร.ธีระพร วีระถาวร อาจารย์ที่ปรึกษาวิทยานิพนธ์เป็นอย่างสูงที่กรุณาสละเวลาในการให้ความรู้ คำแนะนำและคำปรึกษา ตลอดจนแก้ไขข้อบกพร่องต่างๆ ด้วยความเมตตาและเอาใจใส่อย่างดียิ่งจนทำให้วิทยานิพนธ์ ฉบับนี้สำเร็จลุล่วงได้ด้วยดี

ผู้วิจัยขอกราบขอบพระคุณรองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา ประธานกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ผกาวดี ศิริรังษี กรรมการสอบวิทยานิพนธ์ และรองศาสตราจารย์ ดร.มนต์ทิพย์ เทียนสุวรรณ กรรมการสอบวิทยานิพนธ์จากมหาวิทยาลัยมหิดล ที่กรุณาให้คำแนะนำ ตรวจสอบ และแก้ไขวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น ตลอดจนครูอาจารย์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ให้แก่ผู้วิจัยตั้งแต่การศึกษาขั้นต้นจนถึงปัจจุบัน

สุดท้ายนี้ ผู้วิจัยขอกราบขอบพระคุณ คุณพ่อ คุณแม่ที่ส่งเสริมและสนับสนุนด้านการเรียนของผู้วิจัยและเป็นกำลังใจให้เสมอมาจนสำเร็จการศึกษา และขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่เป็นกำลังใจและให้ความช่วยเหลือเสมอมา

สารบัญ

หน้า

สารบัญตาราง

สารบัญภาพ

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การวิเคราะห์ความถดถอยเชิงเส้น (linear regression analysis) เป็นวิธีการทางสถิติที่ใช้วิเคราะห์ความสัมพันธ์หรือหาสาเหตุที่มีผลกระทบต่อตัวแปรที่สนใจ ซึ่งเป็นเทคนิคที่นิยมใช้กันอย่างกว้างขวางในงานวิจัยด้านวิทยาศาสตร์ การแพทย์ สังคมศาสตร์ หรือในด้านธุรกิจ เป็นต้น การวิเคราะห์ความถดถอยนี้จะมีการแบ่งตัวแปรออกเป็นสองส่วนคือ ตัวแปรตาม และตัวแปรอิสระ ซึ่งเป็นตัวแปรที่คาดว่าจะสามารถอธิบายอิทธิพล หรือการเปลี่ยนแปลงของตัวแปรตามได้ โดยผลลัพธ์ที่ได้คือสมการการถดถอยที่ประ กอบไปด้วยกลุ่มของตัวแปรอิสระที่ถูกคัดเลือกโดยเทคนิคการคัดเลือกตัวแปร ซึ่งการเลือกเทคนิคการคัดเลือกตัวแปรที่แตกต่างกัน หรือการพิจารณาระแนงตัวแบบ (lattice) ที่แตกต่างกันในแต่ละลูกโซ่ ก็อาจจะให้ผลลัพธ์ที่แตกต่างกันได้

(ธีระพร วีระถาวร, 2541: 201) พิจารณาระแนงตัวแบบต่อไปนี้

$$
\begin{tikzcd}[column sep=small]
 & y=\beta_0'+\beta_1(x_1-\bar{x}_{.1})+\beta_2(x_2-\bar{x}_{.2})+\varepsilon \quad (M1) \arrow[dl, "\beta_1=0"'] \arrow[dr, "\beta_2=0"] & \\
y=\beta_0'+\beta_2(x_2-\bar{x}_{.2})+\varepsilon \quad (M2) \arrow[dr, "\beta_2=0"'] & & y=\beta_0'+\beta_1(x_1-\bar{x}_{.1})+\varepsilon \quad (M3) \arrow[dl, "\beta_1=0"] \\
 & y=\beta_0'+\varepsilon \quad (M4) &
\end{tikzcd}
$$

เมื่อ Mi หมายถึงตัวแบบความถดถอยเชิงเส้นตัวแบบที่ i

จากระแนงตัวแบบข้างต้น เมื่อพิจารณาการเปรียบเทียบตัวแบบภายในลูกโซ่เดียวกัน (เป็นการเปรียบเทียบตัวแบบที่ซับซ้อนกว่าหรือตัวแบบที่มีตัวแปรอิสระมากกว่า ว่าให้ผลไม่แตกต่างจากตัวแบบง่ายขึ้น หรือตัวแบบที่มีจำนวนตัวแปรอิสระลดลงหรือไม่) สมมติว่า ตัวแบบ M2 และตัวแบบ M3 คือตัวแบบที่ได้จากการเปรียบเทียบตัวแบบภายในลูกโซ่ทางซ้าย และลูกโซ่ทางข วา ตามลำดับ

ปัญหาที่น่าสนใจคือ ตัวแบบ M2 และตัวแบบ M3 ซึ่งเป็นตัวแบบความถดถอยเชิงเส้นที่มาจากต่างลูกโซ่กัน ตัวแบบใดเหมาะสมที่สุดในแง่ของการศึกษาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ และเกณฑ์หรือมาตรวัดใดที่สามารถใช้เปรียบเทียบตัวแบบที่มาจากต่างลูกโซ่กันได้

การคัดเลือกตัวแปรอิสระ (variable selection) ที่เป็นขั้นตอนหนึ่งของการวิเคราะห์ความถดถอยนั้น มีผู้วิเคราะห์หรือผู้วิจัยบางส่วนเข้าใจในวัตถุประสงค์ของขั้นตอนนี้ผิดพลาดไป ในขั้นตอนของการคัดเลือกตัวแปรอิสระนั้น เป็นขั้นตอนของการลดจำนวนของตัวแปรอิสระทั้ง หมดที่มีอยู่ให้น้อยลง เพื่อที่จะให้ปัญหาความรุนแรงของความสัมพันธ์ร่วมระหว่างตัวแปรอิสระเบาบางลงไปได้ ไม่ได้ทำไปเพื่อวัตถุประสงค์ของการค้นหาสมการถดถอยสำหรับการพยากรณ์ ซึ่งผลลัพธ์ที่ได้จะมาจากการพิจารณาตัวแบบที่อยู่ในลูกโซ่เดียวกัน กล่าวคือเป็นการเปรียบเทียบ ตัวแบบที่ติดกลุ่มกัน (nested model) ดังนั้นหากผู้วิจัยต้องการที่จะหาสมการถดถอยที่ดีที่สุดในแง่ของการศึกษาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ จึงจำเป็นที่จะต้องพิจารณาการคัดเลือกตัวแบบ (model selection) ด้วย

การคัดเลือกตัวแบบความถดถอยนั้น มีจุดประส งค์หลักในการหาสมการถดถอยที่ดีที่สุดเพื่อใช้ในการพยากรณ์ตัวแปรตามโดยการใช้เกณฑ์ที่สามารถวัดประสิทธิภาพของแต่ละตัวแบบได้มาเปรียบเทียบกันโดยตัวแบบที่นำมาเปรียบเทียบกันนั้นจะเป็นตัวแบบติดกลุ่ม หรือไม่ติดกลุ่มก็ได้ ดังนั้นจากคำถามข้างต้นจึงต้องทำการคัดเลือกตัวแบบความถดถอยโดยเลือกเกณฑ์ที่เหมาะสมมาใช้เปรียบเทียบ เพื่อหาสมการถดถอยที่ดีที่สุดภายใต้ข้อมูลชุดเดียวกัน

ในปัจจุบันเกณฑ์ที่ใช้คัดเลือกตัวแบบความถดถอยมีมากมาย เช่น ค่าเฉลี่ยความผิดพลาดกำลังสอง(Mean Square Error), สัมประสิทธิ์การตัดสินใจ (R^2), Akaike's Information Criterion: AIC (1974), Sawa's Bayesian Information Criterion: BIC (1978) เป็นต้น นอกจากนี้ก็ยังมีการพัฒนาเกณฑ์เหล่านี้ให้ซับซ้อนมากยิ่งขึ้น เกิดเป็นเกณฑ์ต่างๆอีกมากมาย คำถามหนึ่งที่เกิดขึ้นคือเกณฑ์เหล่านี้อธิบายสมการถดถอยที่นำมาเปรียบเทียบอย่างไร ค่ าที่คำนวณได้มีความหมายอย่างไร คำถามเหล่านี้เกิดขึ้นเนื่องจากสูตรการคำนวณของเกณฑ์เหล่านั้นดูยุ่งยาก และซับซ้อน ยากแก่การอธิบายในเชิงความหมาย นอกจากนี้ เป็นที่น่าสงสัยว่าบางเกณฑ์ จะนำมาใช้เปรียบเทียบได้จริงหรือไม่ เนื่องจากเป็นค่าที่ได้มาจากพื้นฐานที่ไม่เหมื อนกัน ในแง่ของตัวแปรอิสระในตัวแบบที่แตกต่างกัน หรือมาตรวัดของตัวแปรที่แตกต่างกัน เช่น ค่าเฉลี่ยความผิดพลาดกำลังสอง (Mean Square Error) หรือ สัมประสิทธิ์การตัดสินใจ (R^2) เป็นต้น

(Jiri Andel, 1993: 335-340) ในปี ค.ศ. 1940 โฮเทลลิ่ง ได้เสนอแนวคิดการเปรียบเ ทียบตัวแบบ M2 และ M3 ด้วยการเปรียบเทียบค่าสัมประสิทธ์สหสัมพันธ์ (correlation coefficient) ระหว่าง

Y กับ X_1 และ Y กับ X_2 โดยทำการสร้างตัวสถิติ เพื่อใช้ทดสอบสมมติฐานเชิงสถิติ เปรียบเทียบตัวแบบ M2 และ M3 ว่าตัวแบบใดที่เหมาะสมกว่ากัน ซึ่งพิจารณาจากผลรวมความคลา ดเคลื่อนกำลังสอง (residual sum of squares: RSS) ในปี ค.ศ. 1955 ฮีลลี่ ได้แสดงว่าวิธีการทดสอบของโฮเทลลิ่งนั้นเป็นการทดสอบเดียวกันกับการทดสอบสมมติฐานเชิงสถิติของสัมประสิทธิ์ความถดถอย ต่อมาในปี ค.ศ. 1959 วิลเลี่ยมส์ ได้นำแนวความคิดนี้มาประยุกต์สร้างตัวสถิติที่ใช้ทดสอบเปรียบเทียบตัวแบบที่มีตัวแปรอิสระมากกว่า 1 ตัวแปร และได้กล่าวว่า ในกรณีทั่วไปการทดสอบสมมติฐานเชิงสถิติของ โฮเทลลิ่ง ไม่ได้เหมือนกับการทดสอบสมมติฐานเชิงสถิติของสัมประสิทธิ์ความถดถอย

ในปี ค.ศ. 1967 เคนดอลล์ และสจวต ได้นำวิธีทดสอบดังกล่าวมาใช้เปรี ยบเทียบ ตัวแบบ M2 และ M3 โดยพิจารณาเบื้องต้นจาก

$$RSS_1 = \sum_{i=1}^{n}(y_i - \bar{y})^2 - \frac{\left[\sum_{i=1}^{n}(x_{i1} - \bar{x}_{.1})(y_i - \bar{y})\right]^2}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_{.1})^2}$$

$$RSS_2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 - \frac{\left[\sum_{i=1}^{n}(x_{i2} - \bar{x}_{.2})(y_i - \bar{y})\right]^2}{\sum_{i=1}^{n}(x_{i2} - \bar{x}_{.2})^2}$$

ซึ่งคือผลรวมความคลาดเคลื่อนกำลังสอง จากการประมาณด้วยตัวแบบ M2 และ M3 ตามลำดับ

จากงานวิจัยดังกล่าว จึงเป็นที่น่าสังเกตว่าค่าผลรวมกำลังสองของความคลาดเคลื่อนจ าก การประมาณน่าจะเป็นเกณฑ์ที่ใช้เปรียบเทียบตัวแบบได้ ถ้าพิจารณาโครงสร้างของค่า RSS ข้างต้นจะเห็นว่า ประกอบไปด้วยสองส่วนคือ ส่วนที่เป็นค่าที่คำนวณจากค่า y เพียงอย่างเดียว และ ส่วนที่สองเป็นค่าที่พิจารณาร่วมกันระหว่างตัวแปรอิสระที่นำเข้ามาในตัวแบบความถดถอยกั บตัวแปรตาม Y อาจกล่าวได้ว่า ค่าผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) เป็นค่าความแตกต่างระหว่างค่าสังเกต y กับค่าที่คำนวณหลังจากทำการประมาณด้วยตัวแบบความถดถอยที่ประกอบไปด้วยตัวแปรอิสระต่างๆ

ในกรณีทั่วไป ตัวแบบ $y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1,2,\ldots,n$ $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ เป็นอิสระซึ่งกันและกัน และมีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ^2 เหมือนกัน

กำหนดให้

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix}$$ มีค่าลำดับชั้น (rank) เท่ากับ k+1

และ $\bar{x}_j = \dfrac{1}{n}\sum_{i=1}^{n} x_{ij}, \quad j = 1,2,\ldots,k$

เพราะฉะนั้นตัวแบบข้างต้น สามารถเขียนใหม่ได้เป็น

$$y_i = \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \cdots + \beta_k(x_{ik} - \bar{x}_k) + \varepsilon_i, \quad i = 1,2,\ldots,n$$

เมื่อ $\beta_0 = \beta_0' + \beta_1\bar{x}_1 + \beta_2\bar{x}_2 + \cdots + \beta_k\bar{x}_k$

ซึ่งเมื่อนำมาเขียนให้อยู่ในรูปของเมทริกซ์ จะได้ว่า

$$\underset{\sim}{y} = (\underset{\sim}{1}, H)\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

$$\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \underset{\sim}{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

และ $$H = \begin{pmatrix} x_{11} - \bar{x}_1 & \ldots & x_{1k} - \bar{x}_k \\ \cdots & \cdots & \cdots \\ x_{n1} - \bar{x}_1 & \ldots & x_{nk} - \bar{x}_k \end{pmatrix}$$

ในกรณีนี้ค่าผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) คือ

$$RSS = (\underset{\sim}{y}'\underset{\sim}{y} - n\bar{y}^2) - \underset{\sim}{y}' H(H'H)^{-1}H'\underset{\sim}{y}$$

ในกรณีสมก ารถดถอยซึ่งต้องการเปรียบเทียบ มีจำนวนตัวแปรอิสระมากกว่า 1 ตัวแปร หรือตัวแปรอิสระของแต่ละสมการมีจำนวนไม่เท่ากัน อาจกล่าวได้ว่าความซับซ้อนของตัวแบบแตกต่างกัน การเปรียบเทียบความเหมาะสมระหว่างสองสมการถดถอยใดๆ จึงควรคำนึงถึงจำนวนตัวแปรอิสระด้วย เนื่องจากยิ่งตัวแปรอิสระมาก ความคลาดเคลื่อนยิ่งมีโอกาสมีค่าสูงขึ้นด้วย กล่าวคือความผิดพลาดในการประมาณมีค่ามากขึ้น นั่นเอง

ในปัจจุบันเกณฑ์ที่ใช้เปรียบเทียบตัวแบบ เช่น Akaike’s Information Criterion (AIC), Sawa’s Bayesian Information Criterion (BIC) และเกณฑ์อื่นๆที่พัฒนาขึ้นจากเกณฑ์เหล่านี้ มีพื้นฐานมาจากค่าผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) แต่เนื่องจาก

เหตุผลของความซับซ้อนที่แตกต่างกันของตัวแบบ จึงมีการปรับเกณฑ์ให้ สามารถวัดความซับซ้อนของตัวแบบด้วย

ในทฤษฎีการตัดสินใจ นอกจากการพิจารณาความไม่เอนเอียงของตัวสถิติแล้ว การเลือกตัวประมาณที่ดี ยังต้องอาศัยองค์ประกอบอื่นๆร่วมด้วย เช่น ฟังก์ชันความสูญเสีย ฟังก์ชันการเสี่ยง เป็นต้น ในกรณีที่ความแปรปรวนของตัวประมาณที่ไม่เอนเอี ยงมีค่ามาก ฟังก์ชันความสูญเสียและฟังก์ชันการเสี่ยง เป็นองค์ประกอบที่สำคัญของเกณฑ์การเลือกตัวประมาณที่ดี

โดยทั่วไปแล้วฟังก์ชันความสูญเสียจะอยู่ในรูปของ

$$L(\theta, D) = (D - q(\theta))^2$$

โดยที่ D เป็นตัวประมาณ และ q เป็นฟังก์ชันของพารามิเตอร์ θ เรียกฟังก์ชันความสูญเสียนี้ว่า "*ฟังก์ชันความสูญเสียแบบความผิดพลาด กำลังสอง*" (square error loss) ในกรณีนี้ฟังก์ชันการเสี่ยงคือ

$$MSPE = R(\theta, D) = E[L(\theta, D)] = E[(D - q(\theta))^2] = Var(D) + bias(D, \theta)^2$$

ซึ่งเรียกว่า "*ค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง*" (mean square prediction error)

ถ้าพิจารณาว่าการคัดเลือกตัวแบบความถดถอยเชิงเส้นที่ดีที่สุดเป็นการเลือกตัวประมาณที่ดีและเหมาะสมที่สุด สำหรับค่าสังเกต $\underset{\sim}{y}$ นั่นคือ สามารถพิจารณาค่าเฉลี่ยของควา มผิดพลาดจากการพยากรณ์กำลังสองได้ ซึ่ง (Toby Dylan Hocking, 2006) ได้กล่าวถึงความสัมพันธ์ระหว่างค่า MSPE กับ RSS ว่า $E(MSPE) = E(RSS) + 2p\sigma^2 - n\sigma^2$ ซึ่งเทอม $2p\sigma^2$ เป็นมาตรวัดความซับซ้อนของตัวแบบ โดยที่ ค่าประมาณของ MSPE คือ $\widehat{MSPE} = RSS - ns^2 + 2ps^2$ ซึ่งสามารถนำค่าประมาณนี้เป็นเกณฑ์เปรียบเทียบสมการถดถอยเชิงเส้นที่มีจำนวนตัวแปรอิสระไม่เท่ากันได้

1.2 วัตถุประสงค์การวิจัย

1. เพื่อศึกษาเกณฑ์การคัดเลือกสมการถดถอยเชิงเส้นที่มีประสิทธิภาพและน่าเชื่อถือ ซึ่งนำไปสู่การได้ตัวแบบที่ถูกต้องและเหมาะสมที่สุดภายใต้ข้อมูลชุดเดียวกัน
2. เพื่อศึกษาเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยเชิงเส้น 2 เกณฑ์ ระหว่างค่าผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) และ

ค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (mean square prediction error: MSPE) ว่ามีความสอดคล้อง หรือแตกต่างกันอย่างไร

1.3 ข้อตกลงเบื้องต้น

1. ตัวแบบความถดถอยเชิงเส้นที่ศึกษามีรูปแบบทั่วไปดังนี้

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ $\underset{\sim}{y}$ คือเวกเตอร์ของตัวแปรตามขนาด $n \times 1$

X คือเมทริกซ์ของตัวแปรอิสระขนาด $n \times p$

$\underset{\sim}{\beta}$ คือเวกเตอร์ของสัมประสิทธิ์ความถดถอยขนาด $p \times 1$

$\underset{\sim}{\varepsilon}$ คือเวกเตอร์ของความคลาดเคลื่อนสุ่มขนาด $n \times 1$

n คือขนาดตัวอย่าง

และ p คือจำนวนพารามิเตอร์ในตัวแบบ

โดยที่ $\underset{\sim}{\varepsilon} \sim N_n(\underset{\sim}{0}, \sigma^2 I_n)$, $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$, $\mathrm{cov}(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

2. ตัวแปรอิสระแต่ละตัวเป็นค่าคงที่
3. ความคลาดเคลื่อน (ε_i) เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ $(N(0,\sigma^2))$ เหมือนกันและเป็นอิสระซึ่งกันและกัน (i.i.d.)
4. ตัวประมาณค่าของพารามิเตอร์สัมประสิทธิ์ความถดถอย $\underset{\sim}{\beta}$ ของตัวแบบความถดถอยเชิงเส้นข้างต้น คือ ตัวประมาณค่ากำลังสองน้อยสุด (least squares estimator (L.S.E.))

$$\hat{\underset{\sim}{\beta}} = (X'X)^{-1} X' \underset{\sim}{y}$$

1.4 คำจำกัดความที่ใช้ในการวิจัย

1. (ธีระพร วีระถาวร, 2541: 49) ระแนงตัวแบบหรือแลตทิซ (lattice) หมายถึง เซตที่เป็นคู่อันดับได้บางส่วน (partially ordered set) โดยที่ทุกคู่ของสมาชิกจะมีส่วนร่วม (union) ซึ่งกรณีนี้ตัวแบบซับซ้อน (complicated model) บรรจุตัวแบบทั้งสองตัวแบบ และทุกคู่ของสมาชิกจะมีส่วนร่วม (intersection) ซึ่งตัวแบบง่ายขึ้น (simpler model) จะบรรจุอยู่ในตัวแบบทั้งสอง
2. (พจนา แววสวัสดิ์, 2543: 26) ตัวแบบติดกลุ่ม (nested model) หมายถึง ตัวแบบ 2 ตัวแบบจะติดกัน ถ้าในแต่ละพจน์ของตัวแบบแรกเป็นส่วนหนึ่งของตัวแบบที่สอง ซึ่งตัวแบบที่สองจะมีพจน์มากกว่าตัวแบบแรกอย่างน้อย 1 เทอม ตัวแบบที่สองที่มีความซับซ้อน

มากกว่าตัวแบบแรกเรียกว่าตัวแบบที่ซับซ้อน (complicated model) และตัวแบบแรกที่เป็นตัวแบบอย่างง่ายของตัวแบบสองเรียกว่าตัวแบบง่ายขึ้น (simpler model)

3. (บุญจิรา มากอ้น, 2545: 5) ตัวแบบไม่ติดกลุ่ม (non-nested model) หมายถึง ตัวแบบ 2 ตัวแบบจะไม่ติดกลุ่มกัน ถ้าตัวแบบหนึ่งไม่สามารถลดรูปเป็นตัวแบบหนึ่งได้ ด้วยการสมมติให้พารามิเตอร์บางตัวมีค่าเป็นศูนย์

1.5 ขอบเขตของการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ต่างๆที่ต้องการศึกษาดังนี้

1. จำนวนตัวแปรอิสระที่ใช้ในการวิจัยมี 2 ระดับคือ 3 และ 4
2. ขนาดตัวอย่างที่ใช้ในการวิจัยมี 3 ขนาดคือ 20, 35 และ 50 ตามลำดับ
3. ความคลาดเคลื่อนเป็นกลุ่มตัวอย่างที่สุ่มจากประชากรที่มีการแจกแจงแบบปกติ ซึ่งมีพารามิเตอร์ $\mu = 0$ และ $\sigma = 1,2,3$ และ 5 ตามลำดับ
4. สร้างตัวแปรอิสระที่มีความสัมพันธ์กันทุกตัวแปรในระดับต่างๆที่เป็นไปได้ทั้งหมดตั้งแต่ 0.05 ถึง 0.95 ที่ทำให้ $\sum$ เป็นเมทริกซ์ความแปรปรวนร่วม (covariance matrix) ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระเป็นเมทริกซ์บวกแน่นอน (positive definite matrix)
5. กำหนดระดับนัยสำคัญของการทดสอบสมมติฐาน(α) เท่ากับ 0.05
6. ข้อมูลที่ใช้ในการศึกษาวิจัยครั้งนี้ใช้โปรแกรมคอมพิวเตอร์จำลองด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 500 รอบในแต่ละสถานการณ์

1.6 เกณฑ์การตัดสินใจ

เกณฑ์การตัดสินใจว่าตัวแบบใดมีความถูกต้องมากที่สุดจะพิจารณาจากเกณฑ์ผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) และค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (mean squared prediction error: MSPE) ซึ่งมีสูตรดังนี้

$$RSS = (\underset{\sim}{y}'\underset{\sim}{y} - n\bar{y}^2) - \underset{\sim}{y}'H(H'H)^{-1}H'\underset{\sim}{y}$$

เมื่อตัวแบบความถดถอยอยู่ในรูป

$$\underset{\sim}{y} = (\underset{\sim}{1}, H)\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

โดยที่ $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \underset{\sim}{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

และ $H = \begin{pmatrix} x_{11} - \bar{x}_1 & \ldots & x_{1k} - \bar{x}_k \\ \cdots & \cdots & \cdots \\ x_{n1} - \bar{x}_1 & \ldots & x_{nk} - \bar{x}_k \end{pmatrix}$

และ $M\hat{S}PE = RSS - ns^2 + 2ps^2$

โดยที่ s^2 คือตัวประมาณค่าไม่เอนเอียงสำหรับ σ^2

และ p คือจำนวนพารามิเตอร์ในตัวแบบความถดถอย

นอกจากนี้พิจารณาค่าอัตราส่วนผลต่างของร้อยละของผลรวมความคลาดเคลื่อนกำลังสอง (RDRSS) และ อัตราส่วนผลต่างของร้ อยละของค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (RDMSPE) เพื่อพิจารณาว่าแต่ละเกณฑ์ที่ใช้ ตัวแบบที่เลือกได้จะดีกว่าตัวแบบอื่นกี่เปอร์เซ็นต์ โดยมีสูตรดังนี้

$$RDRSS_k = \frac{(RSS_k - RSS_{\min})}{RSS_k} \times 100$$

$$RD\,M\hat{S}PE_k = \frac{(M\hat{S}PE_k - M\hat{S}PE_{\min})}{M\hat{S}PE_k} \times 100$$

เมื่อ RSS_k แทนค่าผลรวมความคลาดเคลื่อนกำลังสองของตัวแบบที่ k ที่สนใจเปรียบเทียบ

$RSS_{\min}$ แทนค่าผลรวมความคลาดเคลื่อนกำลังสองที่มีค่าต่ำที่สุด

$RDRSS_k$ แทนค่าอัตราส่วนผลต่างของร้อยละของผลรวมความคลาดเคลื่อนกำลังส องของตัวแบบที่ k

$M\hat{S}PE_k$ แทนค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสองของตัวแบบที่ k ที่สนใจเปรียบเทียบ

$M\hat{S}PE_{\min}$ แทนค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพ ยากรณ์กำลังสองที่มีค่าต่ำที่สุด

และ $RD\,M\hat{S}PE_k$ แทนค่าอัตราส่วนผลต่างของร้อยละของค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสองของตัวแบบที่ k

1.7 วิธีดำเนินการวิจัย

1. กำหนดลักษณะการแจกแจงของความคลาดเคลื่อน ขน าดตัวอย่าง และจำนวนตัวแปรอิสระตามที่กำหนดไว้ในขอบเขตการวิจัย

2. สร้างข้อมูลตัวแปรอิสระ (X) ให้มีระดับความสัมพันธ์ตามที่กำหนดไว้ในขอบเขตการวิจัย และสร้างข้อมูลของตัวแปรตาม ($\underset{\sim}{y}$) จากรูปแบบความสัมพันธ์ $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ โดยกำหนดให้ $\underset{\sim}{\beta}' = (1,1,1,\ldots,1)_{1\times p}$

3. ประมาณค่าสัมประสิทธิ์ความถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุด

4. ทำการคัดเลือกตัวแบบโดยใช้หลักการกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) เพื่อหาตัวแบบที่ยอมรับได้ในแต่ละลูกโซ่ โดยทำการเปรียบเทียบตัวแบบเชิงซ้อน (complicated model) กับตัวแบบง่ายขึ้น (simpler model) ด้วยการทดสอบเอฟบางส่วน (partial F-test)

5. คำนวณหาค่าเฉลี่ยผลรวมความคลาดเคลื่อนกำลังสอง ($R\bar{S}S$) และค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง ($M\bar{S}PE$) จากการทำซ้ำจำนวน 500 รอบของตัวแบบที่ได้จากข้อ 4 เพื่อเปรียบเทียบหาตัวแบบที่เหมาะสมที่สุดภายใต้แลตทิซ พร้อมทั้งหาค่าอัตราส่วนผลต่างของร้อยละของผลรวมความคลาดเคลื่อนกำลังสอง ($RDRSS$) และค่าอัตราส่วนผลต่างของร้อยละของค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง ($RD\,M\hat{S}PE$) เพื่อวัดว่าแต่ละเกณฑ์ที่ใช้ ตัวแบบที่เลือกได้จะดีกว่าตัวแบบอื่นกี่เปอร์เซ็นต์

6. สรุปผลการวิจัยจากการใช้เกณฑ์ทั้งสอง ในแต่ละสถานการณ์

7. พิจารณาเปรียบเทียบผลที่ได้จากการใช้เกณฑ์ ที่แตกต่างกัน ว่ามีความสอดคล้อง หรือแตกต่างกันอย่างไร

1.8 ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นแนวทางในการพิจารณาเปรียบเทียบอิทธิพลที่แท้จริงของตัวแปรอิสระ (ปัจจัยที่มีผลกระทบจริงต่อเรื่องที่สนใจศึกษา) ของแต่ละตัวแบบความถดถอยเชิงเส้นที่มีผลต่อตัวแปรตาม

2. เป็นแนวทางในการศึกษาวิธีการคัดเลือกตัวแบบความถดถอยที่ดีที่สุดวิธีอื่นๆ สำหรับการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณต่อไป

บทที่ 2

แนวคิดและทฤษฎี

2.1 การวิเคราะห์ความถดถอยพหุคูณเชิงเส้น (Multiple Linear Regression Analysis)

(วลัยทิพย์ บุญญาดิศัย, 2549: 13) ตัวแบบที่ใช้สำหรับการวิจัยครั้งนี้เป็นตัวแบบความถดถอยเชิงเส้น นั่นคือ เป็นเชิงเส้นในพารามิเตอร์ ซึ่งมีรูปแบบทั่วไปดังนี้

$$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i, \qquad i = 1,2,\ldots,n$$

เมื่อ k เป็นจำนวนตัวแปรอิสระ

ส่วนข้อตกลงเบื้องต้นของการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณจะเป็นดังนี้

1. ε_i มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และมีค่าความแปรปรวนคงที่ เท่ากับ σ^2 นั่นคือ $\varepsilon_i \sim N(0,\sigma^2)$ โดยที่ ε_i และ ε_j สำหรับ $i \neq j$ มีการแจกแจงที่เหมือนกันและเป็นอิสระต่อกัน ซึ่งจะทำให้ $\mathrm{cov}(\varepsilon_i, \varepsilon_j) = 0$ สำหรับ $i \neq j$

2. เนื่องจากตัวแบบการถดถอยที่ใช้ในการพิจารณาเป็นตัวแบบเชิงเส้น และค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติ ดังนั้นในการประมาณค่าพารามิเตอร์ของสัมประสิทธิ์ความถดถอยด้วยวิธีกำลังสองน้อยสุด (least square estimation) ซึ่งจะได้ว่า

ตัวประมาณกำลังสองน้อยสุด (Least Square Estimator: $\underset{\sim}{\hat{\beta}}$) คือ

$$\underset{\sim}{\hat{\beta}} = (X'X)^{-1} X' \underset{\sim}{y}$$

2.2 การประมาณค่าสัมประสิทธิ์ความถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุด

(วลัยทิพย์ บุญญาดิศัย, 2549: 14-17) วิธีการประมาณค่าสัมประสิทธิ์นี้มีรากฐานมาจากทฤษฎีการประมาณเชิงเส้น โดยมีหลักการในการประมาณค่าสัมประสิทธิ์ คือ ทำให้ผลรวมกำลังสองของความคลาดเคลื่อน (Sum Square of Error: SSE) มีค่าน้อยสุด ซึ่งแสดงรายละเอียดดังนี้

<u>นิยามที่ 2.2.1</u>

กำหนดให้ $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ โดยที่ $\underset{\sim}{\varepsilon} \sim N_n(\underset{\sim}{0}, \sigma^2 I_n)$ จะได้ตัวประมาณกำลังสองน้อยสุดของ $\underset{\sim}{\beta}$ คือ $\underset{\sim}{\hat{\beta}}$ ที่จะทำให้ผลรวมกำลังสองของความคลาดเคลื่อนมีค่าน้อยที่สุด โดยที่ค่าประมาณ

ของ $\underset{\sim}{\hat{\beta}}$ คือ $\underset{\sim}{\hat{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$ และเมทริกซ์ความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\hat{\beta}}$ คือ $\operatorname{cov}(\underset{\sim}{\hat{\beta}}) = \sigma^2 (X'X)^{-1}$

ดังนั้นจากนิยามและจากตัวแบบทั่วไปจะได้ผลรวมกำลังสองของความคลาดเคลื่อนอยู่ในรูปแบบดังนี้

$$\begin{aligned} SS(\underset{\sim}{\beta}) &= \underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = (\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta}) \\ &= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\beta}'X\underset{\sim}{y} - \underset{\sim}{y}'X\underset{\sim}{\beta} + \underset{\sim}{\beta}'X'X\underset{\sim}{\beta} \\ &= \underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{\beta}'X\underset{\sim}{y} + \underset{\sim}{\beta}'X'X\underset{\sim}{\beta} \qquad (2.2.1) \end{aligned}$$

เนื่องจากวิธีกำลังสองน้อยสุดมีหลักการที่ทำให้ผลรวมความคลาดเคลื่อนกำลังสองมีค่าต่ำที่สุด ดังนั้นเราจะได้ค่าประมาณสัมประสิทธิ์ความถดถอยพหุคูณจากการหาอนุพันธ์อันดับที่ 1 ของสมการที่ (2.2.1) เทียบกับ $\underset{\sim}{\beta}$ แล้วให้เท่ากับศูนย์ ซึ่งผลดังกล่าวอยู่ในรูปแบบดังนี้

$$\frac{\partial}{\partial \underset{\sim}{\beta}} SS(\underset{\sim}{\beta}) = -2X'\underset{\sim}{y} - 2X'X\underset{\sim}{\hat{\beta}} = \underset{\sim}{0}$$

กล่าวคือ $X'X\underset{\sim}{\hat{\beta}} = X'\underset{\sim}{y}$

ซึ่งเราเรียกสมการดังกล่าวว่า สมการปกติ (normal equation) และจะได้ตัวประมาณสัมประสิทธิ์ความถดถอยพหุคูณในรูปของ

$$\underset{\sim}{\hat{\beta}} = (X'X)^{-1}(X'\underset{\sim}{y}) \qquad (2.2.2)$$

ตัวประมาณในสมการที่ (2.2.2) มีคุณสมบัติเป็นตัวประมาณที่ไม่เอนเอียงของ $\underset{\sim}{\beta}$ กล่าวคือ

$$\begin{aligned} E(\underset{\sim}{\hat{\beta}}) &= E[(X'X)^{-1}X'\underset{\sim}{y}] \\ &= E[(X'X)^{-1}X'(X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon})] \\ &= E[(X'X)^{-1}X'X(\underset{\sim}{\beta}) + (X'X)^{-1}X'\underset{\sim}{\varepsilon}] \\ &= E[\underset{\sim}{\beta}] + (X'X)^{-1}X'E(\underset{\sim}{\varepsilon}) \\ &= \underset{\sim}{\beta} \end{aligned}$$

เราควรสังเกตว่าการหาตัวประมาณค่าของตัวประมาณ $\underset{\sim}{\beta}$ โดยวิธีกำลังสองน้อยสุดนี้เราไม่จำเป็นต้องทราบลักษณะการแจกแจงความน่าจะเป็นของความคลาดเคลื่อน $\underset{\sim}{\varepsilon}$ นอกจากนั้นตัวประมาณ $\underset{\sim}{\hat{\beta}}$ ยังเป็นตัวประมาณค่าด้วยวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator) ของ $\underset{\sim}{\beta}$ ด้วย รวมทั้งเป็นตัวประมาณที่ไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator

(BLUE)) คงเส้นคงวา (consistant) และพอเพียง (sufficient) ด้วย แต่ในการประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยสุดมีข้อสมมติฐานที่สำคัญข้อหนึ่งคือตัวแปรอิสระแต่ละตัวต้องไม่มีความสัมพันธ์กับตัวแปรอิสระตัวอื่นซึ่งในทางปฏิบัติเกิดขึ้นได้น้อยมาก เนื่องจากตัวแปรบางตัวที่นำมาศึกษาอาจเป็นฟังก์ชันของตัวแปรอิสระตัวอื่น หรือที่เรียกว่า ตัวแปรอิสระมีพหุสัมพันธ์ (multicollinearity) เมื่อตัวแปรอิสระมีพหุสัมพันธ์กันสูงจะทำให้เมทริกซ์ $X'X$ เกิดเงื่อนไขที่ไม่ดี (ill-condition) อาจมีผลทำให้การประมาณ $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยสุดไม่ได้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำสุด เราจึงควรตรวจสอบประสิทธิภาพของตัวประมาณที่ได้โดยเราพิจารณาประสิทธิภาพของตัวประมาณค่า $\underset{\sim}{\beta}$ สองส่วน กล่าวคือ เมทริกซ์ความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\hat{\beta}}$ และค่าเฉลี่ยของกำลังสองระยะทางจาก $\underset{\sim}{\hat{\beta}}$ ไปยัง $\underset{\sim}{\beta}$ ซึ่งเราสามารถเขียนเมทริกซ์ความแปรปรวนร่วมของ $\underset{\sim}{\hat{\beta}}$ ในรูปฟังก์ชันของ $X'X$ และ σ^2 ดังต่อไปนี้

(2.2.3) $$\operatorname{cov}(\underset{\sim}{\hat{\beta}}) = \sigma^2 (X'X)^{-1}$$

ให้ L_1 คือระยะทางจาก $\underset{\sim}{\hat{\beta}}$ ไปยัง $\underset{\sim}{\beta}$ ดังนั้น

(2.2.4) $$L_1^{\,2} = (\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})'(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})$$

และเราจะได้ค่าเฉลี่ยของกำลังสองระยะทางจาก $\underset{\sim}{\hat{\beta}}$ ไปยัง $\underset{\sim}{\beta}$ ในรูปของ

(2.2.5) $$E(L_1^{\,2}) = \sigma^2 trace(X'X)^{-1}$$

$$E(L_1^{\,2}) = E[(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})'(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})]$$

$$= E[\underset{\sim}{\hat{\beta}}'\,\underset{\sim}{\hat{\beta}}] - \underset{\sim}{\beta}'\,\underset{\sim}{\beta}$$

(2.2.6) $$E[\underset{\sim}{\hat{\beta}}'\,\underset{\sim}{\hat{\beta}}] = \underset{\sim}{\beta}'\,\underset{\sim}{\beta} + \sigma^2 trace(X'X)^{-1}$$

เมื่อ $\underset{\sim}{\varepsilon}$ มีการแจกแจงแบบปกติจะได้ว่า

(2.2.7) $$Var(L_1^{\,2}) = 2\sigma^4 trace(X'X)^{-1}$$

จากสมการที่ (2.2.3),(2.2.5) และ (2.2.7) เราจะเห็นได้ว่า $\operatorname{cov}(\underset{\sim}{\hat{\beta}}), E(L_1^{\,2})$ และ $Var(L_1^{\,2})$ ต่างก็เป็นฟังก์ชันของเมทริกซ์ $X'X$ ดังนั้นเพื่อความสะดวกในการทำความเข้าใจเราจึงแปลงเมทริกซ์ $X'X$ ให้อยู่ในรูปของค่าเฉพาะ (eigenvalue) ของเมทริกซ์ $X'X$ โดยใช้ทฤษฎีที่

สำคัญข้อหนึ่งคือ ถ้า λ_i เป็นค่าเฉพาะของเมทริกซ์ $X'X$ แล้ว

$\sum_{i=1}^{p} \lambda_i = trace(X'X); i = 1,2,\ldots,p$ เมื่อ p เป็นจำนวนพารามิเตอร์ในตัวแบบ

กำหนดให้ค่าเฉพาะของเมทริกซ์ $X'X$ มีค่าเป็น

$(\lambda_{\max} = \lambda_1) \geq \lambda_2 \geq \ldots\ldots \geq (\lambda_p = \lambda_{\min}); \quad \lambda_1, \lambda_2, \ldots\ldots, \lambda_p > 0$

จากสมการ (2.2.5) เราสามารถเขียนค่าเฉลี่ยกำลังสองของระยะทางจาก $\hat{\underset{\sim}{\beta}}$ ไปยัง $\underset{\sim}{\beta}$ ในรูปฟังก์ชันของค่าเฉพาะได้ดังนี้

(2.2.8) $$E(L_1^{\,2}) = \sigma^2 \sum_{i=1}^{p} \left(\frac{1}{\lambda_i}\right)$$

และจากสมการที่ (2.2.7) เราสามารถเขียนค่าความแปรปรวนกำลังสองของระยะทางจาก $\hat{\underset{\sim}{\beta}}$ ไปยัง $\underset{\sim}{\beta}$ ในรูปของฟังก์ชันของค่าเฉพาะได้ดังนี้

(2.2.9) $$Var(L_1^{\,2}) = 2\sigma^4 \sum_{i=1}^{p} \left(\frac{1}{\lambda_i}\right)^2$$

ในกรณีที่ตัวแปรอิสระมีสภาพไม่เหมาะสม กล่าวคือ เกิดความสัมพันธ์ระหว่างตัวแปรอิสระในอัตราที่สูงจะทำให้ $|X'X|$ มีค่าเล็กลงเข้าใกล้ศูนย์ เนื่องจาก $|X'X|$ มีค่าเท่ากับผลคูณของค่าเฉพาะของเมทริกซ์ $X'X$ จึงส่งผลให้ค่าเฉพาะบางค่าต่ำมาก ดังนั้นจากสมการที่ (2.2.8) และ (2.2.9) เราจะเห็นได้ว่า $E(L_1^{\,2})$ และ $Var(L_1^{\,2})$ จึงมีค่าสูงขึ้นตามไปด้วย นอกจากนี้การเกิดความสัมพันธ์ระหว่างตัวแปรอิสระส่งผลให้ความแปรปรวนของค่าประมาณสัมประสิทธิ์ความถดถอยมีค่ามากและเกิดความสัมพันธ์กันสูงระหว่างสัมประสิทธิ์ความถดถอยที่ใช้ประมาณค่า

<u>การแก้ไขปัญหาตัวแปรอิสระมีพหุสัมพันธ์กัน</u>

1. ไม่ต้องแก้อะไรทั้งสิ้น ทั้งนี้อาจเป็นเพราะพหุสัมพันธ์ที่เกิดขึ้นนั้นไม่สูงนัก หรือผู้วิเคราะห์อาจสนใจเพียงเพื่อให้ได้สมการถดถอยสำหรับการประมาณค่าหรือพยากรณ์ โดยคำนึงถึงเพียงว่าค่า R^2 มีค่าสูงพอที่จะนำไปใช้ประมาณได้

2. สามารถแก้ไขด้วยการเก็บข้อมูลเพิ่มขึ้น (ถ้าเป็นไปได้) โดยพยายามหาข้อมูลตัวแปรอิสระใหม่ที่ไม่สัมพันธ์กับตัวแปรอิสระตัวอื่นๆ หรือถ้าเป็นไปไม่ได้ในกรณีดังกล่าวอาจหาข้อมูลโดยเพิ่มขนาดตัวอย่างก็ได้ ทั้งนี้จะช่วยให้ค่าความแปรปรวนของค่าประมาณของสัมประสิทธิ์ความถดถอยมีค่าลดลง

3. การแปลงข้อมูลตัวแปรอิสระที่สงสัยว่าจะก่อปัญหาพหุสัมพันธ์กับตัวแปรอิสระอื่นๆ

4. ตัดตัวแปรอิสระที่ก่อให้เกิดปัญหาออก

5. แก้ไขด้วยวิธีการหาองค์ประกอบหลัก (Principal Components)

6. แก้ไขด้วยการประมาณและวิเคราะห์ด้วยวิธีการวิเคราะห์ความถดถอยแบบริดจ์ (Ridge Regression Analysis)

2.3 การแปลงเมทริกซ์ X เข้าสู่ศูนย์กลาง

(วลัยทิพย์ บุญญาดิศัย, 2549: 17-18) การแปลงเมทริกซ์เข้าสู่ศูนย์กลาง เป็นวิธีการหนึ่งในการแก้ไขปัญหาหารเกิดความสัมพันธ์ร่วมเช้งพหุของตัวแปรอิสระ ซึ่งเป็นวิธีที่ใช้เมทริกซ์สหสัมพันธ์มาแทนที่เมทริกซ์ $X'X$ โดยพิจารณาจากการกำหนดให้ข้อมูลมีรูปแบบดังตารางนั่นคือ

x_0		x_1	x_2	…	x_k	Y
1		x_{11}	x_{21}	…	x_{k1}	y_1
1		x_{12}	x_{22}	…	x_{k2}	y_2
	…					
1		x_{1n}	x_{2n}	…	x_{kn}	y_n
ผลบวกของคอลัมน์		$\sum x_{1i}$	$\sum x_{2i}$	…	$\sum x_{ki}$	$\sum y_i$
ค่าเฉลี่ยของคอลัมน์		$\bar{x}_1$	$\bar{x}_2$	…	$\vec{x}_k$	$\bar{y}$

จากตัวแบบ

$$y = \beta_0' + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

สามารถเขียนในรูปแบบของ

$$y = \{\beta_0' + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \ldots + \beta_k \bar{x}_k\} + \beta_1 (x_1 - \bar{x}_1) + \beta_2 (x_2 - \bar{x}_2) + \ldots + \beta_k (x_k - \bar{x}_k) + \varepsilon$$

เมื่อ $\bar{y}, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ เป็นค่าจำนวนจริงที่ได้มาจากข้อมูล และถ้าเราแทน

$\beta_0 = \{\beta_0' + \beta_1 (x_1 - \bar{x}_1) + \beta_2 (x_2 - \bar{x}_2) + \ldots + \beta_k (x_k - \bar{x}_k)\}$ จะได้ว่า

$$y = \beta_0 + \beta_1 (x_1 - \bar{x}_1) + \beta_2 (x_2 - \bar{x}_2) + \ldots + \beta_k (x_k - \bar{x}_k) + \varepsilon$$

ซึ่ง β_0 สามารถประมาณได้จาก $\hat{\beta}_0 = \dfrac{1}{n} \sum_{i=1}^{n} y_i = \bar{y}$

2.4 เมทริกซ์ความแปรปรวนร่วม

(ธีระพร วีระถาวร, 2541: 123)

<u>นิยามที่ 2.4.1</u>

ถ้า $\underset{\sim(k\times1)}{x}$ เป็นเวกเตอร์สุ่ม ดังนั้น เมทริกซ์ความแปรปรวนร่วมของ $\underset{\sim(k\times1)}{x}$ คือ

$$\operatorname{cov}(\underset{\sim}{x}) = (\operatorname{cov}(x_i, x_j))_{k\times k} = E[(\underset{\sim}{x} - E(\underset{\sim}{x}))(\underset{\sim}{x} - E(\underset{\sim}{x}))']$$

<u>นิยามที่ 2.4.2</u>

เราเรียกเมทริกซ์สมมาตร $A_{(k\times k)}$ ว่าไม่เป็นลบแน่นอน (Non-negative definite or positive semi-definite) ถ้า $\underset{\sim}{\alpha}' A \underset{\sim}{\alpha} \geq 0, \forall \underset{\sim}{\alpha} \in \Re^p$ และ เรียกว่าเป็นบวกแน่นอน (positive semi-definite) ถ้า $\underset{\sim}{\alpha}' A \underset{\sim}{\alpha} > 0, \forall \underset{\sim}{\alpha} \neq \underset{\sim}{0}$

<u>บทตั้งที่ 2.4.1</u>

เมทริกซ์ $\Sigma_{(k\times k)}$ เป็นเมทริกซ์ความแปรปรวนร่วมก็ต่อเมื่อไม่เป็นลบแน่นอน

<u>หมายเหตุ</u>

ถ้า A เป็นเมทริกซ์บวกแน่นอน จะได้ว่า A เป็นเมทริหซ์ไม่เป็นลบแน่นอน และเงื่อนไขที่เพียงพอและจำเป็นสำหรับ $A = (a_{ij})_{k\times k}$ จะเป็นบวกแน่นอน คือ

$a_{11} > 0$, $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$, …, และ $|A| > 0$ (n leading principal minors เป็นบวก)

2.5 การพิจารณาความเหมาะสมของตัวแบบ

เกณฑ์การตัดสินใจว่าตัวแบบใดมีความถูกต้องมากที่สุดจะพิจารณาจากเกณฑ์ผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) และค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (mean squared prediction error: MSPE) โดยสามารถคำนวณได้ดังนี้

$$RSS = \frac{1}{500}\sum_{j=1}^{500}[(\underset{\sim}{y}'\underset{\sim}{y} - n\bar{y}^2) - \underset{\sim}{y}' H(H'H)^{-1}H'\underset{\sim}{y}]$$

$$MSPE = \frac{1}{500}\sum_{j=1}^{500}[RSS - ns^2 + 2ps^2]$$

เมื่อตัวแบบความถดถอยอยู่ในรูป

$$\underset{\sim}{y} = (\underset{\sim}{1}, H)\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

โดยที่ $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \underset{\sim}{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix}, \quad \underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

และ $H = \begin{pmatrix} x_{11} - \bar{x}_1 & \ldots & x_{1k} - \bar{x}_k \\ \cdots & \cdots & \cdots \\ x_{n1} - \bar{x}_1 & \ldots & x_{nk} - \bar{x}_k \end{pmatrix}$

เมื่อ n คือขนาดตัวอย่าง

s^2 คือตัวประมาณไม่เอนเอียงสำหรับ σ^2

p คือจำนวนพารามิเตอร์ในตัวแบบความถดถอย

RSS คือผลรวมความคลาดเคลื่อนกำลังสอง เฉลี่ยจากการทำซ้ำ 500 รอบ

และ $MSPE$ คือค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง เฉลี่ยจากการทำซ้ำ 500 รอบ

การตัดสินใจว่าตัวแบบความถดถอยเชิงเส้นหนึ่งๆ ในตัวแบบที่เป็นไปได้ทั้งหมด ตัวแบบเต็มรูป หรือตัวแบบลดรูปใด มีความเหมาะสมมากกว่ากัน โดยการใช้เกณฑ์ผลรวมความคลาดเคลื่อนกำลังสอง และค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง จะมีหลักการพิจารณาเหมือนกัน คือเลือกตัวแบบที่ให้ค่าเกณฑ์ต่ำที่สุด

2.6 การกำจัดตัวแปรแบบถอยหลัง

(วลัยทิพย์ บุญญาดิศัย, 2549: 22-23) วิธีการนี้เริ่มต้นด้วยสมการถดถอยด้วยตัวแบบเต็มรูป (full model) คือ ประกอบด้วยตัวแปรอิสระทุกตัวที่ใช้พิจารณาแล้วจะคัดเลือกตัวแปรอิสระออกจากสมการครั้งละ 1 ตัวแปร โดยเลือกตัวแปรที่มีความสัมพันธ์กับ y น้อยที่สุด (เมื่อตัวแปรอื่นๆคงที่) และค่าสัมประสิทธิ์ความถดถอยของตัวแปรนั้นไม่มีนัยสำคัญ จากนั้นคำนวณหาสมการถดถอยสำหรับตัวแปรอิสระที่เหลือ และคัดตัวแปรอิสระที่มีนัยสำคัญน้อยที่สุดออก จนกระทั่งไม่มีตัวแปรอิสระใดถูกคัดออกแล้วตัวแปรอิสระที่เหลืออยู่จะอยู่ในสมการถดถอยทุกตัว ซึ่งเป็นอันสิ้นสุดของวิธีการนี้ และสามารถแสดงขั้นตอนได้ดังนี้

1. สร้างสมการถดถอยเต็มรูปที่รวมเอาตัวแปรอิสระที่ควรพิจารณาทั้งหมด สมมติว่ามี k ตัว สมการจะเป็นดังนี้

$$\underset{\sim}{\hat{y}} = \hat{\beta}_0' + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_k X_k \;\; ; \; k = \text{จำนวนตัวแปรอิสระ}$$

2. คำนวณหาค่าเอฟบางส่วน (Partial F) ของตัวแปรอิสระทุกตัว เสมือนว่าตัวแปรอิสระนั้นเข้าสู่สมการเป็นตัวสุดท้าย

3. ในจำนวนเอฟบางส่วนนี้เลือกค่าเอฟบางส่วนที่น้อยที่สุด สมมติได้ F_j แล้วนำค่า F_j ไปเปรียบเทียบกับ $F_{\alpha(1,n-k-1)}$ ที่กำหนดจากตาราง ถ้าพบว่า

ก) $F_j < F_{\alpha(1,n-k-1)}$ ให้กำจัดตัวแปรอิสระ X_j ออกจกสมการถดถอย และตั้งสมการใหม่โดยไม่รวม X_j ในสมการ จากนั้นกลับไปทำในขั้นตอนที่ 2

ข) $F_j > F_{\alpha(1,n-k-1)}$ จะหยุดกระบวนการคัดเลือกตัวแปรอิสระ และได้สมการถดถอยที่เหมาะสม

2.7 การทดสอบเอฟบางส่วน (Partial F-test)

(วลัยทิพย์ บุญญาดิศัย, 2549: 23-26) การทดสอบเอฟบางส่วนเป็นการทดสอบที่ใช้ตรวจสอบนัยสำคัญของ β_j เพื่อตัดสินใจว่าตัวแปรอิสระใดควรอยู่ในสมการหรือตัวแปรอิสระใดไม่ควรอยู่ในสมการ โดยที่ β_j จะปรากฏอยู่ ณ ตำแหน่งใดในแบบจำลองก็ได้ แต่ในทางปฏิบัติจะทำโดยถือว่าตัวแปรอิสระนั้นเข้าสู่สมการเป็นตัวสุดท้าย

ให้ $\underset{\sim}{y} = \beta_0' + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_q X_q + \underset{\sim}{\varepsilon}$ (1)

จากสมการ (1) เราสามารถหาค่าตัวประมาณ $\underset{1\,\sim}{\hat{\beta}} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_q)'$ โดยวิธีกำลังสองน้อยที่สุด และค่าผลบวกกำลังสองที่เกี่ยวข้องได้ดังนี้

ก) $\underset{1\,\sim}{\hat{\beta}} = ({}_1X'X)^{-1}\,{}_1X'\underset{\sim}{y}$ เมื่อ ${}_1X$ คือ เมทริกซ์ขนาด $n \times (q+1)$ ซึ่งรวมเทอมของค่าคงที่

ข) $SSR_1 = \underset{1\,\sim}{\hat{\beta}'}\,{}_1X'\underset{\sim}{y}$

ค) $SSE_1 = \underset{\sim}{y}'\underset{\sim}{y} - \underset{1\,\sim}{\hat{\beta}'}\,{}_1X'\underset{\sim}{y}$ และ $MSE_1 = \sigma_1^{\,2} = \dfrac{1}{n-(q+1)}(\underset{\sim}{y}'\underset{\sim}{y} - \underset{1\,\sim}{\hat{\beta}'}\,{}_1X'\underset{\sim}{y})$

เมื่อ q = จำนวนพารามิเตอร์สัมประสิทธิ์ความถดถอยของสมการ 1

$\underset{1\,\sim}{\hat{\beta}'}$ = สัมประสิทธิ์ ความถดถอยของสมการ (1)

SSR_1 = ผลรวมกำลังสองของความถดถอย (Sum Squares of Regression) ของสมการ (1)

MSE_1 = ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Squares Error) ของสมการ (1)

ให้ $\underset{\sim}{y} = \beta_0' + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_q X_q + \beta_{q+1} X_{q+1} + \ldots + \beta_k X_k + \underset{\sim}{\varepsilon}$ (2)

เป็นสมการแสดงความสัมพันธ์ของ $\underset{\sim}{y}$ กับ X โดยที่ $k > q$

จากสมการ (2) สามารถหาค่าตัวประมาณ $\underset{2\,\sim}{\hat{\beta}} = (\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_q, \hat{\beta}_{q+1}, \ldots, \hat{\beta}_k)'$ โดยวิธีกำลังสองน้อยที่สุด และค่าผลบวกกำลังสองที่เกี่ยวข้องได้ดังนี้

ก) $\underset{2\,\sim}{\hat{\beta}} = ({}_2X'X)^{-1}\,{}_2X'\underset{\sim}{y}$ เมื่อ ${}_2X$ คือ เมทริกซ์ขนาด $n \times (k+1)$ ซึ่งรวมเทอมของค่าคงที่

ข) $SSR_2 = \underset{2\,\sim}{\hat{\beta}'}\,{}_2X'\underset{\sim}{y}$

ค) $SSE_2 = \underset{\sim}{y'}\underset{\sim}{y} - \underset{2\,\sim}{\hat{\beta}'}\,{}_2X'\underset{\sim}{y}$ และ $MSE_2 = {\sigma_2}^2 = \dfrac{1}{n-(k+1)}(\underset{\sim}{y'}\underset{\sim}{y} - \underset{2\,\sim}{\hat{\beta}'}\,{}_2X'\underset{\sim}{y})$

จากผลลัพธ์ข้างต้นจะพบว่า Extra Sum Squares of Regression คือ SSR

เมื่อ $ESSR = SSR_2 - SSR_1 = \underset{2\,\sim}{\hat{\beta}'}\,{}_2X'\underset{\sim}{y} - \underset{1\,\sim}{\hat{\beta}'}\,{}_1X'\underset{\sim}{y}$

ซึ่ง Extra Sum Squares นี้เป็นค่าผลรวมกำลังสองของตัวแปรอิสระ $X_{q+1}, X_{q+2}, \ldots, X_k$ ที่เพิ่มขึ้นจากสมการ (1)

จาก Distribution of Quadratic Form เราสามารถพิสูจน์ได้ว่า

$$*\frac{ESSR}{\sigma^2} \sim \chi^2_{(k-q)} \text{ และ } \frac{SSR_2}{\sigma^2} \sim \chi^2_{(n-(k+1))}$$

ดังนั้น ค่าทดสอบเอฟบางส่วน $= \dfrac{ESSR/(k-q)\sigma^2}{SSE_2/(n-k-1)\sigma^2}$

$$= \frac{ESSR/(k-q)}{\hat{\sigma}_2^2}$$

จะมีการแจกแจงแบบเอฟ ณ ระดับขั้นความเสรี $(k-q, n-k-1)$ และจะปฏิเสธสมมติฐาน

$H_0 : \beta_{q+1} = \beta_{q+2} = \ldots = \beta_k = 0$

$H_1 : \beta_j$ ไม่เท่ากับศูนย์ทั้งหมด ; $j = q+1, q+2, \ldots, k$

ณ ระดับนัยสำคัญ α เมื่อค่าทดสอบเอฟบางส่วนมากกว่า $F_\alpha(k-q, n-q-1)$

จากความรู้ในเรื่องการทดสอบเอฟบางส่วนนี้ เราสามารถนำมาประยุกต์ใช้กับวิธีการหาสมการถดถอยที่ดีที่สุดได้ เช่น วิธีการถดถอยขั้นบันได วิธีกำจัดตัวแปรแบบถอยหลัง เป็นต้น โดยการนำไปใช้ทำได้ดังนี้

จากสมการ $\underset{\sim}{y} = \beta_0' + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \underset{\sim}{\varepsilon}$ เราจะหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวแบบ (Mean Squares Error (MSE)) และผลบวกกำลังสองของความถดถอย (Sum Squares of Regression) ของเฉพาะ β_j ได้ดังนี้

$SS(\beta_j \mid \beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k) = SS(\beta_1, \beta_2, \ldots, \beta_k) - SS(\beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)$

เมื่อ $SS(\beta_j \mid \beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)$ คือ ผลรวมกำลังสองของ X_j เมื่อ $X's$ อื่นๆ เข้าสู่แบบจำลองแล้ว

ดังนั้นค่าสถิติเอฟบางส่วน คือ

$$F_c = \frac{SS(\beta_j \mid \beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)}{\hat{\sigma}^2}$$
$$= \frac{SS(\beta_1, \beta_2, \ldots, \beta_k) - SS(\beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)}{\hat{\sigma}^2}$$

โดยจะปฏิเสธ $H_0 : \beta_j = 0$ ณ ระดับนัยสำคัญ α เมื่อ $F_c > F_\alpha(1, n-k-1)$

การทดสอบเอฟบางส่วน ใช้สำหรับตรวจสอบนัยสำคัญของ β_j เพื่อตัดสินใจว่าตัวแปรอิสระใดควรคงไว้ ตัวแปรอิสระใดควรตัดทิ้งจากสมการถดถอยโดยที่ β_j ปรากฏอยู่ ณ ตำแหน่งในสมการถดถอยใดก็ได้ แต่ในทางปฏิบัติจะต้องยึดหลักเกณฑ์ของ Extra Sum Square ไว้เป็นแนวทางเสมอ กล่าวคือ จะต้องคำนวณหาค่า $SS(\beta_j \mid \beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)$ ได้เฉพาะเมื่อจัดให้ X_j เข้าสู่สมการเป็นตัวสุดท้าย นั่นคือ $SS(\beta_j \mid \beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)$ คำนวณหาได้จากผลต่างระหว่างผลรวมกำลังสองของความถดถอย

จากสมการ

$$\underset{\sim}{y} = \beta_0' + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_{j-1} X_{j-1} + \beta_{j+1} X_{j+1} + \ldots + \beta_k X_k + \underset{\sim}{\varepsilon}$$

และ

$$\underset{\sim}{y} = \beta_0' + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_{j-1} X_{j-1} + \beta_{j+1} X_{j+1} + \ldots + \beta_k X_k + \beta_j X_j + \underset{\sim}{\varepsilon}$$

และจากที่กล่าวมาแล้วข้างต้นสามารถสรุปได้ดังนี้

$$Partial\,F = \frac{S_1 - S_2}{(\underset{\sim}{y}'\underset{\sim}{y} - S_1)/(n-k-1)}$$

เมื่อ S_1 เท่ากับ $\hat{\beta}' X' \underset{\sim}{y}$ ของสมการถดถอยที่รวม X_j อยู่ด้วย

S_2 เท่ากับ $\hat{\beta}' X' \underset{\sim}{y}$ ของสมการถดถอยที่ไม่รวม X_j อยู่ด้วย

n คือขนาดตัวอย่าง

และ k คือจำนวนตัวแปรอิสระ

บทที่ 3

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้ได้ทำการจำลองข้อมูลในคอมพิวเตอร์ ด้วยเทคนิคมอนติคาร์โล (Monte Carlo simulation technique) ในการประมวลผลและวิเคราะห์ข้อมูล ซึ่งขั้นตอนของเทคนิคการจำลองมอนติคาร์โล ขอบเขตการวิจัย และขั้นตอนดำเนินการวิจัยมีรายละเอียดดังต่อไปนี้

3.1 การจำลองข้อมูลด้วยวิธีมอนติคาร์โล

(วลัยทิพย์ บุญญาดิศัย, 2549: 27-28) วิธีมอนติคาร์โลเป็นเทคนิคในการจำลองแบบทางคณิตศาสตร์ที่นิยมใช้กันอย่างแพร่หลาย โดยมีการจำลองตัวเลขสุ่ม (random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษาซึ่งยังไม่แน่ใจผลลัพธ์ที่เกิดขึ้น ซึ่งการวิจัยครั้งนี้จะใช้เทคนิคมอนติคาร์โลในการสร้างข้อมูลที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา โดยขั้นตอนสำคัญของการจำลองข้อมูลด้วยวิธีมอนติคาร์โลมี 3 ขั้นตอนดังนี้

<u>ขั้นตอนที่ 1</u> การสร้างตัวเลขสุ่ม (generate random number)

การใช้ตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีมอนติคาร์โล ทั้งนี้ก็เพราะว่าหลักการของวิธีมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหา ซึ่งลักษณะของตัวเลขสุ่มที่ดีจะมีการแจกแจงแบบสม่ำเสมอ (Uniform distribution) ในช่วง (0,1) และเป็นอิสระซึ่งกันและกัน จากนั้นนำเลขสุ่มที่ได้ไปสร้างตัวแปรสุ่มตามลักษณะการแจกแจงที่ต้องการศึกษา เพื่อเป็นข้อมูลสำหรับปัญหานั้นๆ

<u>ขั้นตอนที่ 2</u> การประยุกต์ปัญหาที่ต้องการศึกษาโดยใช้ตัวเลขสุ่ม

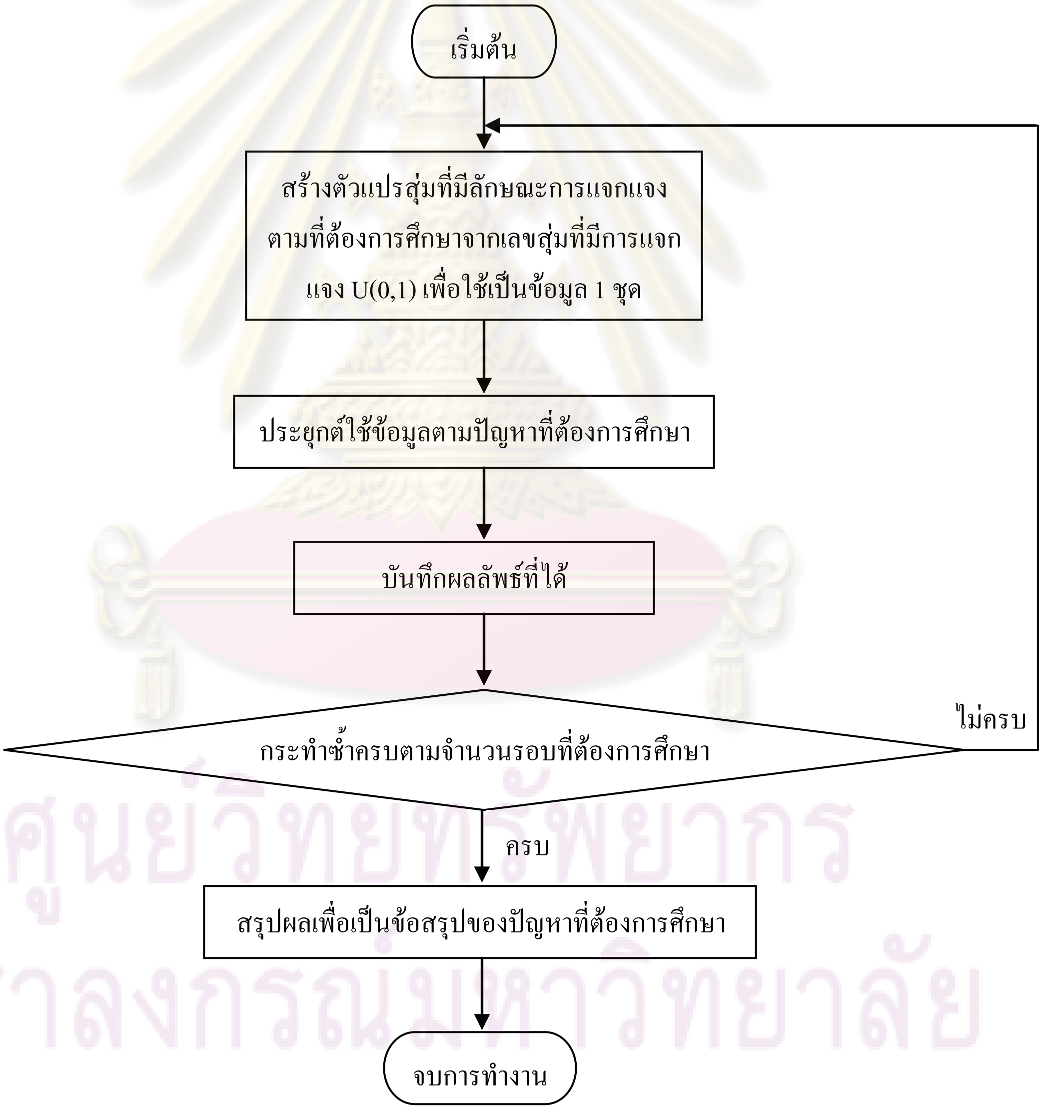
ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา ซึ่งเป็นขั้นตอนที่ใช้เลขสุ่มในการหาคำตอบตามสูตรหรือการคำนวณในปัญหาที่ศึกษา บางปัญหาอาจใช้ตัวเลขสุ่มโดยตรง แต่บางปัญหาอาจใช้ตัวเลขสุ่มเพียงบางขั้นตอนของปัญหาเท่านั้น

<u>ขั้นตอนที่ 3</u> การทดลองกระทำ

เมื่อประยุกต์ปัญหาที่ต้องการศึกษาโดยใช้ตัวเลขสุ่มแล้ว ขั้นตอนต่อไปก็คือ การทดลองโดยใช้กระบวนการสุ่ม (random process) มาทดลองกระทำในลักษณะที่ซ้ำๆกัน (replication)

จำนวนหลายครั้งเพื่อหาคำตอบของปัญหาที่ต้องการศึกษา ซึ่งการทดลองกระทำซ้ำๆกันนั้นจะเป็นการช่วยลดความไม่แน่นอนของคำตอบได้

จากหลักการของวิธีมอนติคาร์โลจะเห็นว่า การใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหา เป็นวิธีการที่จะนำไปสู่แนวคิดทางทฤษฎีที่เกี่ยวข้องกับการคำนวณ โดยเฉพาะทฤษฎีความน่าจะเป็นที่จะนำไปสู่การอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริง เพราะไม่มีผลกระทบจากปัจจัยอื่นๆเข้ามาเกี่ยวข้อง ในการทดลองเมื่อกระทำซ้ำๆกันนั้นเป็นจำนวนมากแล้ว ความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่างๆในแต่ละครั้งจะหมดไป (counter balance) จากขั้นตอนของวิธีมอนติคาร์โล สามารถเขียนผังงานได้ดังนี้



รูปที่ 3.1 แผนผังแสดงขั้นตอนของวิธีมอนติคาร์โล

3.2 ขอบเขตการวิจัย

ในการวิจัยครั้งนี้กำหนดสถานการณ์ต่างๆที่ต้องการศึกษาดังนี้

1. จำนวนตัวแปรอิสระที่ใช้ในการวิจัยมี 2 ระดับคือ 3 และ 4
2. ขนาดตัวอย่างที่ใช้ในการวิจัยมี 3 ขนาดคือ 20, 35 และ 50 ตามลำดับ
3. ความคลาดเคลื่อนเป็นกลุ่มตัวอย่างที่สุ่มจากประชากรที่มีการแจกแจงแบบปกติ ซึ่งมีพารามิเตอร์ $\mu = 0$ และ $\sigma = 1,2,3$ และ 5 ตามลำดับ
4. สร้างตัวแปรอิสระที่มีความสัมพันธ์กันทุกตัวแปรในระดับต่างๆที่เป็นไปได้ทั้งหมด ตั้งแต่ 0.05 ถึง 0.95 ที่ทำให้ $\sum$ เป็นเมทริกซ์ความแปรปรวนร่วม (covariance matrix) ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระเป็นเมทริกซ์บวกแน่นอน (positive definite matrix)
5. กำหนดระดับนัยสำคัญของการทดสอบสมมติฐาน(α) เท่ากับ 0.05
6. ข้อมูลที่ใช้ในการศึกษาวิจัยครั้งนี้ใช้โปรแกรมคอมพิวเตอร์จำลองด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 500 รอบในแต่ละสถานการณ์

3.3 ขั้นตอนการวิจัย

ในการวิจัยครั้งนี้จะพิจารณาผลการคัดเลือกตัวแบบความถดถอยเชิงเส้น 2 ตัวแบบ คือตัวแบบความถดถอยที่มีจำนวนตัวแปรอิสระ 3 ตัวแปร และตัวแบบความถดถอยที่มีจำนวนตัวแปรอิสระ 4 ตัวแปร โดยในแต่ละตัวแบบจะพิจารณาการคัดเลือกตัวแบบ 2 วิธี คือ การเปรียบเทียบตัวแบบที่เป็นไปได้ทั้งหมด (All Possible Models) พร้อมกัน และการเปรียบเทียบตัวแบบภายในระแนงตัวแบบ (Lattice) โดยการใช้เกณฑ์ผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) และค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (mean squared prediction error: MSPE) ซึ่งในแต่ละสถานการณ์ที่แตกต่างกัน จะมีขั้นตอนการทดลองดังต่อไปนี้

1. สร้างข้อมูลของความคลาดเคลื่อนที่มีการแจกแจงแบบปกติ โดยมีพารามิเตอร์ตามที่กำหนด
2. สร้างข้อมูลของตัวแปรอิสระ (X) ให้มีระดับความสัมพันธ์กันตามที่กำหนดไว้ และสร้างข้อมูลของตัวแปรตาม $(\underset{\sim}{y})$ จากรูปแบบความสัมพันธ์ $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ โดยกำหนดให้ $\underset{\sim}{\beta}' = (1,1,1\ldots1)_{1\times p}$

3. ประมาณค่าสัมประสิทธิ์ความถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุด (Least square estimation) และคำนวณค่า RSS และ MSPE ของทุกตัวแบบที่เป็นไปได้ กล่าวคือ ตัวแบบเต็มรูป และตัวแบบลดรูปที่เป็นไปได้ทั้งหมด
4. ทำการประมวลผลซ้ำ 500 รอบ โดยในแต่ละรอบให้ทำการคัดเลือกตัวแบบที่ยอมรับได้ภายในลูกโซ่เดียวกัน โดยใช้หลักการของวิธีการกำจัดตัวแปรแบบถอยหลัง และนับความถี่ว่าใน 500 รอบ ตัวแบบใดได้รับเลือกจากวิธีการกำจัดตัวแปรแบบถอยหลังมากที่สุด ทำเช่นเดียวกันกับทุกลูกโซ่ที่เป็นไปได้ทั้งหมด
5. แต่ละตัวแบบที่เป็นไปได้ คำนวณค่าเฉลี่ย RSS และ MSPE 500 รอบ เพื่อเป็นตัวแทนค่า RSS และ MSPE ของแต่ละตัวแบบ
6. พิจารณาเลือกตัวแบบความถดถอยจากตัวแบบที่เป็นไปได้ทั้งหมดด้วยวิธีการเปรียบเทียบตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน โดยที่ทั้งเกณฑ์ RSS และ MSPE จะพิจารณาเลือกตัวแบบที่ให้ค่าต่ำสุดเหมือนกัน
7. พิจารณาเลือกตัวแบบความถดถอยจากตัวแบบที่เป็นไปได้ทั้งหมดด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ โดยนำตัวแบบที่ได้รับเลือกจากวิธีการกำจัดตัวแปรแบบถอยหลังที่มีความถี่มากที่สุดของแต่ละลูกโซ่ มาเปรียบเทียบกันด้วยค่าเฉลี่ย RSS และ MSPE ที่คำนวณจาก 500 รอบ และพิจารณาเลือกตัวแบบที่ให้ค่าต่ำสุด
8. สรุปผลการวิจัยในแต่ละสถานการณ์

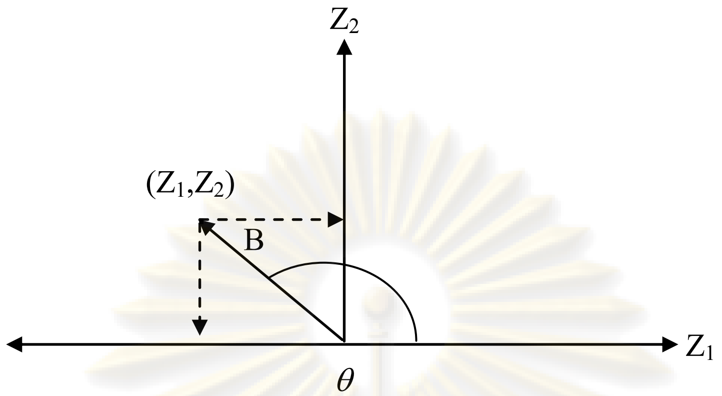
<u>สำหรับรายละเอียดของแต่ละขั้นตอนมีดังนี้</u>

1. การสร้างการแจกแจงของค่าความคลาดเคลื่อนตามที่ต้องการศึกษา

(วลัยทิพย์ บุญญาดิศัย, 2549: 30-31) การสร้างค่าความคลาดเคลื่อนให้มีลักษณะการแจกแจงแบบปกติจะใช้เลขสุ่ม (random number) ซึ่งมีการแจกแจงแบบสม่ำเสมอ (uniform distribution) ในช่วง (0,1) เป็นพื้นฐาน

<u>การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติ</u>

การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติใช้วิธีของบอกซ์ (Box) และมุลเลอร์ (Muller) (1958) ซึ่งผลิตเลขสุ่มที่มีการแจกแจงแบบปกติมาตรฐานที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 พร้อมกัน 2 ค่า โดยใช้ตัวผลิต (generator) Z_1 และ Z_2 ดังรูปต่อไปนี้



จากรูปจะได้ว่า

(1) $Z_1 = B\cos(\theta)$

(2) $Z_2 = B\sin(\theta)$

โดยที่ $B^2=Z_1^2+Z_2^2$ มีการแจกแจงแบบไคกำลังสอง (chi-square distribution) ด้วยระดับความเสรีเท่ากับ 2 ซึ่งเทียบเท่ากับการแจกแจงชี้กำลัง (exponential distribution) ที่มีค่าเฉลี่ยเท่ากับ 2 เมื่อใช้วิธีการแปลงผกผัน (inverse transformation) สร้างเลขสุ่มที่มีการแจกแจงดังต่อไปนี้

(3) $B = (-2\ln(R))^{1/2}$

โดยที่ R เป็นเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ

จากการสมมาตรของการแจกแจงปกติ เราจะได้ว่า θ มีการแจกแจงสม่ำเสมอระหว่าง 0 ถึง 2π เรเดียน และรัศมี B กับ θ เป็นอิสระกัน จากสมการที่ (1), (2) และ (3) เราสามารถสร้างเลขสุ่มที่มีการแจกแจงแบบปกติมาตรฐานจากเลขสุ่ม 2 ชุด คือ R_1 และ R_2 กล่าวคือ

$$Z_1 = (-2\ln(R_1))^{1/2}\cos(2\pi R_2)$$
$$Z_2 = (-2\ln(R_2))^{1/2}\sin(2\pi R_2)$$

ซึ่ง R_1 และ R_2 เป็นเลขสุ่ม เมื่อเราได้เลขสุ่มที่มีการแจกแจงแบบปกติมาตรฐานแล้ว เราจะทำการแปลงค่าเลขสุ่มดังกล่าวโดยอาศัยสมการ

$$NORMAL_1 = \mu+\sigma Z_1$$
$$NORMAL_2 = \mu+\sigma Z_2$$

ซึ่งเราจะได้ว่า $NORMAL_1$ และ $NORMAL_2$ มีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และค่าความแปรปรวนเท่ากับ σ^2 ($NORMAL_i \sim N(\mu,\sigma^2)$; i=1,2)

2. การสร้างข้อมูลให้มีความสัมพันธ์เชิงเส้น

(วลัยทิพย์ บุญญาดิศัย, 2549: 32-34) ในการวิจัยครั้งนี้สร้างข้อมูลของตัวแปรอิสระ $X = [\underset{\sim}{1}\ X_1^{**}]$ โดยที่ $X_1^{**} = (\underset{\sim 1}{x}^*, \underset{\sim 2}{x}^*, \ldots, \underset{\sim n}{x}^*)'$ ซึ่ง $x^{**}_{i(1\times k)}, i = 1,2,\ldots,n$ เป็นข้อมูลชุดที่ i ที่มีตัวแปรอิสระ k ตัว ซึ่งเราสามารถสร้างข้อมูลในแต่ละชุด k ตัวให้มีการแจกแจงปกติหลายตัวแปร (multivariate normal distribution) ที่มีค่าพารามิเตอร์ $\underset{\sim}{\mu} = \underset{\sim}{0}$ และ $\sum$ เป็นเมทริกซ์ความแปรปรวนร่วม (covariance matrix) ขนาด $k \times k$ และเป็นเมทริกซ์บวกแน่นอน (positive definite matrix) ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระได้เป็น $\underset{\sim i}{x}^* \sim N_k(\underset{\sim}{\mu}, \sum_{k\times k})$

ในปี ค.ศ. 1972 บารร์ (Barr) และเซิลสาค (Slesak) เสนอวิธีการสร้างข้อมูลที่มีการแจกแจงปกติของหลายตัวแปรดังนี้

กำหนดให้ $\underset{\sim i}{x}^{*'} = (X_{i1}, X_{i2}, \ldots, X_{ik})'$ เป็นเวกเตอร์ของตัวแปรที่มีการแจกแจงแบบปกติที่มาความสัมพันธ์กันโดยมีเวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\mu} = (\mu_1, \mu_2, \ldots, \mu_k)'$ เราสามารถเขียนเมทริกซ์ความแปรปรวนร่วมได้โดย

$$\sum = E[(\underset{\sim i}{x}^{*'} - \mu)(\underset{\sim i}{x}^{*'} - \mu)'] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{bmatrix}$$

เมื่อ $\sigma_{ij} = \mathrm{cov}(x_i, x_j) = \sigma_{ji} = \mathrm{cov}(x_j, x_i)$ สำหรับ $i \neq j$

และ $\sigma_{ii} = Var(x_i), i = 1,2,\ldots,k, j = 1,2,\ldots,k$

เนื่องจาก $\sum$ เป็นเมทริกซ์บวกแน่นอน ดังนั้นเขียน $\sum$ ได้เป็น

$$\sum = C'C$$

เมื่อ C เป็นเมทริกซ์สามเหลี่ยมล่าง (lower triangle matrix) กล่าวคือ

$$C = \begin{bmatrix} c_{11} & 0 & 0 & \cdots & 0 \\ c_{21} & c_{22} & 0 & \cdots & 0 \\ c_{31} & c_{32} & c_{32} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \cdots & c_{kk} \end{bmatrix}$$

ดังนั้นในขั้นเริ่มต้นเราจะคำนวณหาเมทริกซ์แบบสามเหลี่ยมล่าง (lower triangle matrix) C ที่ทำให้ $\sum = C'C$ หลังจากนั้นทำตามขั้นตอนเพื่อสร้างข้อมูลที่มีการแจกแจงปกติของหลายตัวแปรดังนี้

1. สร้างเวกเตอร์ $\underset{\sim}{z}$ ที่ประกอบด้วยตัวแปรที่มีการแจกแจงแบบปกติมาตรฐานที่เป็นอิสระซึ่งกันและกัน k ตัว
2. คำนวณ $\underset{\sim i}{x}^{*'} = \underset{\sim}{\mu} + C\underset{\sim}{z}$ เนื่องจาก $\sum$ เป็นเมทริกซ์สมมาตร เราจึงใช้ Cholesky Factorization ในการคำนวณหาเมทริกซ์ C ได้ดังนี้

$$c_{ij} = \frac{\left(\sigma_{ij} - \sum_{l=1}^{j-1} c_{il}c_{jl}\right)}{\left(\sigma_{jj} - \sum_{l=1}^{j-1} c_{jl}^{2}\right)^{\frac{1}{2}}}$$

เมื่อ $\quad \sum_{l=1}^{j-1} c_{il}c_{jl} = 0 \qquad ,1 \le j \le i \le l$

นั่นคือ $\quad c_{i1} = \dfrac{\sigma_{i1}}{\sqrt{\sigma_{11}}} \qquad ,1 \le i \le l,$

$$c_{ii} = \left(\sigma_{ii} - \sum_{l=1}^{i-1} c_{il}^{2}\right)^{\frac{1}{2}} \qquad ,1 \le i \le l,$$

$$c_{ij} = \frac{\left(\sigma_{ij} - \sum_{l=1}^{j-1} c_{il}c_{jl}\right)}{c_{jj}} \qquad ,1 < j \le i \le l$$

และ $\quad c_{ij} = 0 \qquad ,1 \le i < j \le l$

เราสามารถแสดงการคำนวณได้ว่า

$$\begin{aligned}\sum &= E[(C\underset{\sim}{z})(C\underset{\sim}{z})'] \\ &= E[C\underset{\sim}{z}\underset{\sim}{z}'C'] \\ &= CE[\underset{\sim}{z}\underset{\sim}{z}']C' \\ &= CC'\end{aligned}$$

เมื่อการแจกแจงของค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติ ผู้วิจัยได้ทำการสร้างข้อมูลของตัวแปรอิสระ X โดยให้มีความสัมพันธ์ตามที่กำหนดดังที่เคยกล่าวมาแล้วหลังจากนั้นจึงสร้างตัวแปรตาม $\underset{\sim}{y}$ ให้มีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระ X ให้มีการแจกแจงความคลาดเคลื่อนที่กำหนดตามรูปแบบของความสัมพันธ์ดังนี้คือ $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ เมื่อ $\underset{\sim}{\varepsilon}$ เป็นความคลาดเคลื่อนที่มีรูปแบบการแจกแจงดังที่ได้กล่าวมาแล้วในตอนต้น สำหรับการสร้าง $\underset{\sim}{y}$ นั้นจะเริ่มจาก

การกำหนดขนาดตัวอย่าง และจำนวนตัวแปรอิสระที่ต้องการศึกษา เมื่อกำหนดพารามิเตอร์ $\underset{\sim}{\beta}' = (1,1,1\ldots1)_{1\times p}$ ค่าเฉลี่ย ค่าความแปรปรวน และลักษณะการแจกแจงของความคลาดเคลื่อน รวมทั้งค่าคงที่ของตัวแปรอิสระ X แล้วจึงสร้างตัวแปรตาม $\underset{\sim}{y}$ ตามรูปแบบดังกล่าว

3. การประมาณค่าสัมประสิทธิ์ความถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุด

(วลัยทิพย์ บุญญาดิศัย, 2549: 34) ตัวประมาณค่าของพารามิเตอร์สัมประสิทธิ์ความถดถอย $\underset{\sim}{\beta}$ ของตัวแบบความถดถอยเชิงเส้นในการวิจัยครั้งนี้ คือ ตัวประมาณกำลังสองน้อยสุด (least square estimator (L.S.E))

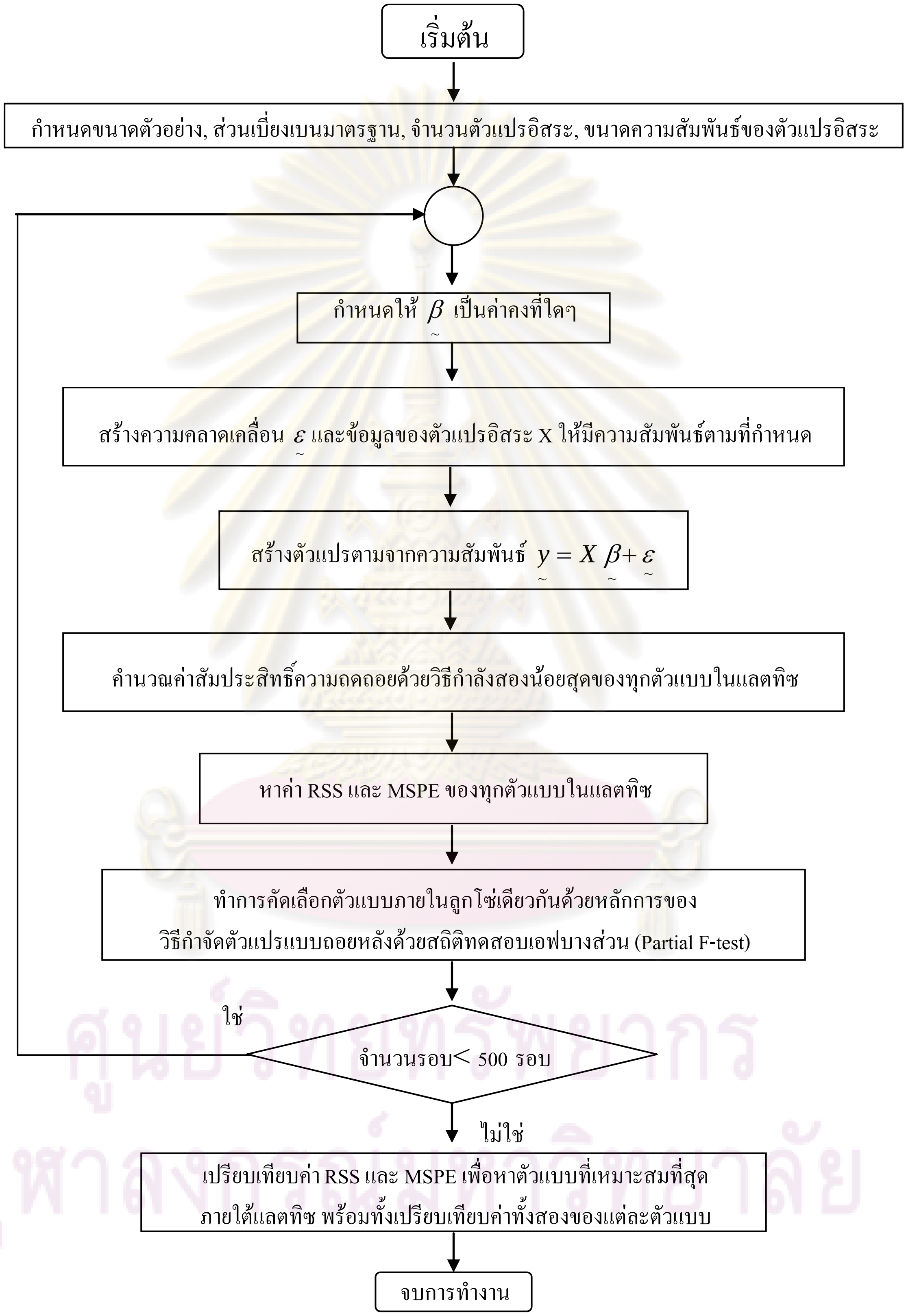
$$\underset{\sim}{\hat{\beta}} = (X'X)^{-1}X'\underset{\sim}{y}$$

4. การคัดเลือกตัวแบบที่ยอมรับได้ในลูกโซ่

(วลัยทิพย์ บุญญาดิศัย, 2549: 34-35) ทำการคัดเลือกตัวแบบที่ยอมรับได้ภายในลูกโซ่เดียวกันด้วยหลักการของวิธีกำจัดตัวแปรแบบถอยหลัง โดยการเปรียบเทียบตัวแบบเชิงซ้อน (complicated model) กับตัวแบบง่ายขึ้น (simpler model) ด้วยการทดสอบเอฟบางส่วน (Partial F-test)

$$F_c = \frac{SS(\beta_j \mid \beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)}{s^2}$$
$$= \frac{SS(\beta_1, \beta_2, \ldots, \beta_{p-1}) - SS(\beta_1, \beta_2, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_k)}{s^2}$$

โดยจะปฏิเสธ $H_0 : \beta_j = 0$ ณ ระดับนัยสำคัญ α เมื่อ $F_c > F_\alpha(1, n-k-1)$



<u>รูปที่ 3.2</u> <u>แผนผังแสดงขั้นตอนการดำเนินการวิจัย</u>

บทที่ 4

ผลการวิจัย

งานวิจัยครั้งนี้เป็นการศึกษาการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ (Lattice) ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบตัวแบบ มี 2 เกณฑ์ คือ ค่าผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) และค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (mean square prediction error: MSPE)

ผู้วิจัยเสนอผลการวิจัยโดยแบ่งเป็น 3 ส่วนคือ

4.1 ความแตกต่างของผลลัพธ์ที่ได้จากการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบตัวแบบที่เป็นไปได้ทั้งหมด (All Possible Models)พร้อมกัน และวิธีการเปรียบเทียบตัวแบบภายในระแนงตัวแบบ (Lattice)

4.2 ปัจจัยทีมีผลต่อค่า RSS และ MSPE

4.3 การคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ (Lattice) โดยใช้เกณฑ์ RSS และ MSPE

การนำเสนอผลการวิจัยจะใช้สัญลักษณ์แทนความหมายต่างๆดังนี้

n หมายถึง ขนาดตัวอย่าง

σ (sigma) หมายถึง ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน

RSS หมายถึง ค่าผลรวมความคลาดเคลื่อนกำลังสอง

RDRSS หมายถึง ค่าอัตราส่วนผลต่างของร้อยละของผลรวมความคลาดเคลื่อนกำลังสอง

MSPE หมายถึง ค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง

RDMSPE หมายถึง ค่าอัตราส่วนผลต่างของร้อยละของค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง

<u>ระดับความสัมพันธ์</u>

ระดับต่ำ หมายถึง $\rho = 0.05, 0.15$

ระดับกลาง หมายถึง $\rho = 0.25, 0.35, 0.45$

ระดับสูง หมายถึง $\rho = 0.55, 0.65, 0.75, 0.85, 0.95$

<u>กรณีที่จำนวนตัวแปรอิสระเท่ากับ 3</u>

พิจารณาตัวแบบความถดถอยเชิงเส้น $y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i, \qquad i = 1,2,\ldots,n$

จะได้ว่าตัวแบบความถดถอยที่เป็นไปได้ทั้งหมดมี 8 ตัวแบบ คือ

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 1 (M1) หรือเรียกว่า *<u>ตัวแบบเต็มรูป</u>* (Full Model)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ ---ตัวแบบที่ 2 (M2)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 3 (M3)

$y_i = \beta_0' + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 4 (M4)

$y_i = \beta_0' + \beta_1 x_{i1} + \varepsilon_i$ ---ตัวแบบที่ 5 (M5)

$y_i = \beta_0' + \beta_2 x_{i2} + \varepsilon_i$ ---ตัวแบบที่ 6 (M6)

$y_i = \beta_0' + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 7 (M7)

$y_i = \beta_0' + \varepsilon_i$ ---ตัวแบบที่ 8 (M8)

<u>ตัวแบบลดรูป</u> (Reduced Model)

และมีจำนวนลูกโซ่ที่เป็นไปได้ 6 ลูกโซ่ คือ

ลูกโซ่ที่ 1 ประกอบด้วยตัวแบบ M1⟶M2⟶M5⟶M8

ลูกโซ่ที่ 2 ประกอบด้วยตัวแบบ M1⟶M2⟶M6⟶M8

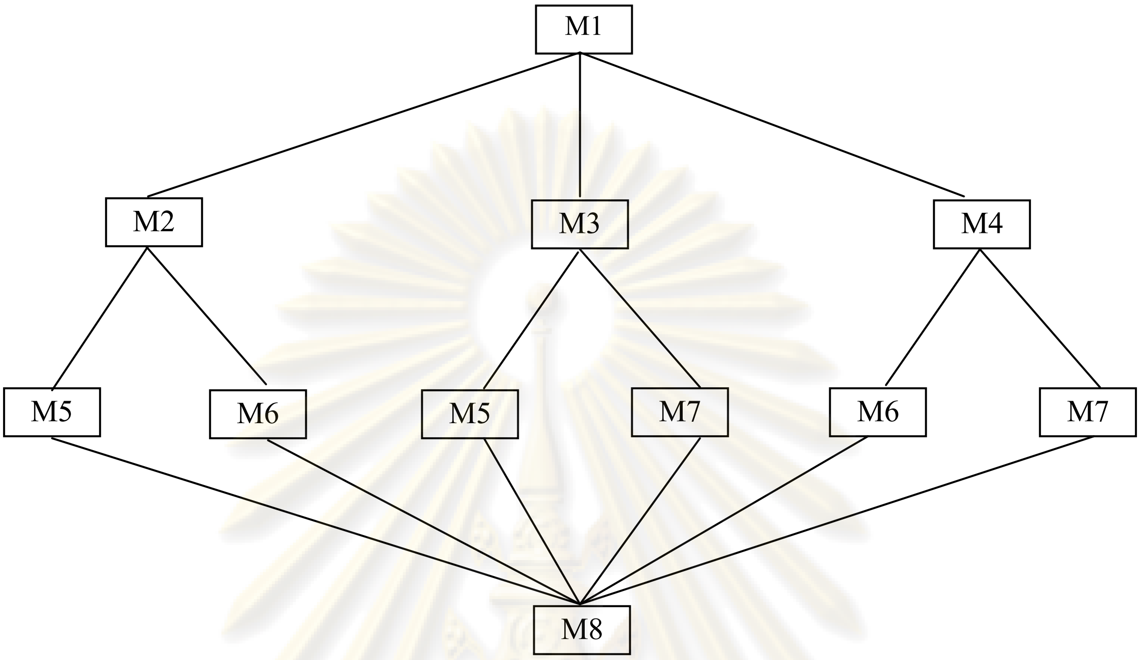
ลูกโซ่ที่ 3 ประกอบด้วยตัวแบบ M1⟶M3⟶M5⟶M8

ลูกโซ่ที่ 4 ประกอบด้วยตัวแบบ M1⟶M3⟶M7⟶M8

ลูกโซ่ที่ 5 ประกอบด้วยตัวแบบ M1⟶M4⟶M6⟶M8

ลูกโซ่ที่ 6 ประกอบด้วยตัวแบบ M1⟶M4⟶M7⟶M8

โดยเขียนเป็นระแนงตัวแบบได้ดังนี้



รูปที่ 4.1 แผนภาพระแนงตัวแบบความถดถอยที่มีตัวแปรอิสระ 3 ตัว

กรณีที่จำนวนตัวแปรอิสระเท่ากับ 4

พิจารณาตัวแบบความถดถอยเชิงเส้น

$$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i, \qquad i = 1,2,\ldots,n$$

จะได้ว่าตัวแบบความถดถอยที่เป็นไปได้ทั้งหมดมี 16 ตัวแบบ คือ

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 1 (M1)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 2 (M2)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 3 (M3)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 4 (M4)

$y_i = \beta_0' + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 5 (M5)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ ---ตัวแบบที่ 6 (M6)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 7 (M7)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 8 (M8)

$y_i = \beta_0' + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 9 (M9)

$y_i = \beta_0' + \beta_2 x_{i2} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 10 (M10)

$y_i = \beta_0' + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 11 (M11)

$y_i = \beta_0' + \beta_1 x_{i1} + \varepsilon_i$ ---ตัวแบบที่ 12 (M12)

$y_i = \beta_0' + \beta_2 x_{i2} + \varepsilon_i$ ---ตัวแบบที่ 13 (M13)

$y_i = \beta_0' + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 14 (M14)

$y_i = \beta_0' + \beta_4 x_{i4} + \varepsilon_i$ ---ตัวแบบที่ 15 (M15)

$y_i = \beta_0' + \varepsilon_i$ ---ตัวแบบที่ 16 (M16)

และมีจำนวนลูกโซ่ที่เป็นไปได้ 24 ลูกโซ่ ดังนี้

ลูกโซ่ที่ 1 ประกอบด้วยตัวแบบ M1⟶M2⟶M6⟶M12⟶M16

ลูกโซ่ที่ 2 ประกอบด้วยตัวแบบ M1⟶M2⟶M6⟶M13⟶M16

ลูกโซ่ที่ 3 ประกอบด้วยตัวแบบ M1⟶M2⟶M7⟶M12⟶M16

ลูกโซ่ที่ 4 ประกอบด้วยตัวแบบ M1⟶M2⟶M7⟶M14⟶M16

ลูกโซ่ที่ 5 ประกอบด้วยตัวแบบ M1⟶M2⟶M9⟶M13⟶M16

ลูกโซ่ที่ 6 ประกอบด้วยตัวแบบ M1⟶M2⟶M9⟶M14⟶M16

ลูกโซ่ที่ 7 ประกอบด้วยตัวแบบ M1⟶M3⟶M6⟶M12⟶M16

ลูกโซ่ที่ 8 ประกอบด้วยตัวแบบ M1⟶M3⟶M6⟶M13⟶M16

ลูกโซ่ที่ 9 ประกอบด้วยตัวแบบ M1⟶M3⟶M8⟶M12⟶M16

ลูกโซ่ที่ 10 ประกอบด้วยตัวแบบ M1⟶M3⟶M8⟶M15⟶M16

ลูกโซ่ที่ 11 ประกอบด้วยตัวแบบ M1⟶M3⟶M10⟶M13⟶M16

ลูกโซ่ที่ 12 ประกอบด้วยตัวแบบ M1⟶M3⟶M10⟶M15⟶M16

ลูกโซ่ที่ 13 ประกอบด้วยตัวแบบ M1⟶M4⟶M7⟶M12⟶M16

ลูกโซ่ที่ 14 ประกอบด้วยตัวแบบ M1⟶M4⟶M7⟶M14⟶M16

ลูกโซ่ที่ 15 ประกอบด้วยตัวแบบ M1⟶M4⟶M8⟶M12⟶M16

ลูกโซ่ที่ 16 ประกอบด้วยตัวแบบ M1⟶M4⟶M8⟶M15⟶M16

ลูกโซ่ที่ 17 ประกอบด้วยตัวแบบ M1⟶M4⟶M11⟶M14⟶M16

ลูกโซ่ที่ 18 ประกอบด้วยตัวแบบ M1⟶M4⟶M11⟶M15⟶M16

ลูกโซ่ที่ 19 ประกอบด้วยตัวแบบ M1⟶M5⟶M9⟶M13⟶M16

ลูกโซ่ที่ 20 ประกอบด้วยตัวแบบ M1⟶M5⟶M9⟶M14⟶M16

ลูกโซ่ที่ 21 ประกอบด้วยตัวแบบ M1⟶M5⟶M10⟶M13⟶M16

ลูกโซ่ที่ 22 ประกอบด้วยตัวแบบ M1⟶M5⟶M10⟶M15⟶M16

ลูกโซ่ที่ 23 ประกอบด้วยตัวแบบ M1⟶M5⟶M11⟶M14⟶M16

ลูกโซ่ที่ 24 ประกอบด้วยตัวแบบ M1⟶M5⟶M11⟶M15⟶M16

4.1 ความแตกต่างของผลลัพธ์ที่ได้จากการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน และวิธีการเปรียบเทียบตัวแบบภายในระแนงตัวแบบ

ในหัวข้อนี้ผู้วิจัยจะวิเคราะห์เฉพาะกรณีที่มีตัวแปรอิสระ 3 ตัวแปร และทำการจำลองข้อมูลในคอมพิวเตอร์ ตามสถานการณ์ต่างๆ ดังนี้

1. ขนาดตัวอย่าง (n) ที่ใช้มี 3 ระดับคือ 20, 35 และ 50
2. ความคลาดเคลื่อนสุ่มจากประชากรมีการแจกแจงแบบปกติโดยที่ $\mu = 0$ และ $\sigma = 1,2,3$ และ 5
3. ความสัมพันธ์ระหว่างตัวแปรอิสระคู่ใดๆ มีค่าต่ำสุดเท่ากับ 0.05 โดยเพิ่มความสัมพันธ์ทีละ 0.10 และมีค่าสูงสุดเท่ากับ 0.95 และทำให้ $\sum$ เป็นเมทริกซ์ความแปรปรวนร่วม (covariance matrix) ที่แสดงความสัมพันธ์ระหว่างตัวแปรอิสระเป็นเมทริกซ์บวกแน่นอน (positive definite matrix)

หมายเหตุ

1) จากสถานการณ์ข้างต้นได้ว่ากรณีตัวแบบที่มีตัวแปรอิสระ 3 ตัว จะมีการจำลองข้อมูลที่มีสถานการณ์แตกต่างกันถึง 12,000 แบบ แต่เนื่องจากบางกรณี ความสัมพันธ์ระหว่างตัวแปรอิสระ 3 ตัวไม่เป็นเมทริกซ์บวกแน่นอน จึงทำให้มีสถานการณ์ที่เป็นไปได้อยู่เพียง **9,804 แบบ**

2) รูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัว ที่เป็นไปได้ (โดยไม่คำนึงถึงลำดับ) มี 10 รูปแบบ คือ **(ต่ำ-ต่ำ-ต่ำ), (ต่ำ-ต่ำ-กลาง), (กลาง-กลาง-ต่ำ), (ต่ำ-ต่ำ-สูง), (กลาง-กลาง-กลาง), (กลาง-กลาง-สูง), (ต่ำ-กลาง-สูง), (สูง-สูง-ต่ำ), (สูง-สูง-กลาง)** และ **(สูง-สูง-สูง)** โดยที่สถานการณ์ซึ่งแตกต่างกัน 9804 แบบ สามารถจัดกลุ่มตามรูปแบบความสัมพันธ์นี้ ได้ดังนี้

รูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัว	จำนวน
ต่ำ-ต่ำ-ต่ำ	96
ต่ำ-ต่ำ-กลาง	432
กลาง-กลาง-ต่ำ	648
กลาง-กลาง-กลาง	324
ต่ำ-ต่ำ-สูง	720
กลาง-กลาง-สูง	1620
ต่ำ-กลาง-สูง	2088
สูง-สูง-ต่ำ	756
สูง-สูง-กลาง	1800
สูง-สูง-สูง	1320
Total	**9804**

แต่ละรูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัว จะแบ่งสถานการณ์ตามขนาดตัวอย่าง และค่าความแปรปรวน ออกเป็นจำนวนเท่าๆกันด้วย เช่น รูปแบบความสัมพันธ์แบบ (กลาง-กลาง-สูง) ซึ่งมีจำนวนเท่ากับ 1,620 แบบ ซึ่งแบ่งออกตามสถานการณ์ต่างๆได้ดังนี้

		σ				
		1	**2**	**3**	**5**	**Total**
n	**20**	135	135	135	135	540
	35	135	135	135	135	540
	50	135	135	135	135	540
Total		405	405	405	405	1,620

ในหัวข้อนี้จะนำเสนอผลการวิจัยออกเป็น 4 ส่วน โดยจัดอยู่ในรูปแบบตารางดังนี้

4.1.1 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ RSS โดยเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน

4.1.2 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ RSS โดยเปรียบเทียบภายในระแนงตัวแบบ

4.1.3 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ MSPE โดยเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน

4.1.4 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ MSPE โดยเปรียบเทียบภายในระแนงตัวแบบ

4.1.1 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ RSS โดยเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน

ตารางที่ 4.1.1 ตารางผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ RSS โดยเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน

รูปแบบ ความสัมพันธ์ของ ตัวแปรอิสระ 3 ตัว	ตัวแบบที่ 1		Total
	จำนวน	%	
ต่ำ-ต่ำ-ต่ำ	96	100.0	96
ต่ำ-ต่ำ-กลาง	432	100.0	432
กลาง-กลาง-ต่ำ	648	100.0	648
กลาง-กลาง-กลาง	324	100.0	324
ต่ำ-ต่ำ-สูง	720	100.0	720
กลาง-กลาง-สูง	1620	100.0	1620
ต่ำ-กลาง-สูง	2088	100.0	2088
สูง-สูง-ต่ำ	756	100.0	756
สูง-สูง-กลาง	1800	100.0	1800
สูง-สูง-สูง	1320	100.0	1320
Total	9804		9804

4.1.2 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ RSS โดยเปรียบเทียบภายในระแนงตัวแบบ

<u>ตารางที่ 4.1.2</u> ตารางผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ RSS โดยเปรียบเทียบภายในระแนงตัวแบบ

รูปแบบ ความสัมพันธ์ของตัวแปรอิสระ 3 ตัว	ตัวแบบที่ 1		ตัวแบบที่ 2		ตัวแบบที่ 3		ตัวแบบที่ 4		ตัวแบบที่ 5		ตัวแบบที่ 6		ตัวแบบที่ 7		Total
	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	
ต่ำ-ต่ำ-ต่ำ	96	100.0	-		-		-		-		-		-		96
ต่ำ-ต่ำ-กลาง	432	100.0	-		-		-		-		-		-		432
กลาง-กลาง-ต่ำ	648	100.0	-		-		-		-		-		-		648
กลาง-กลาง-กลาง	324	100.0	-		-		-		-		-		-		324
ต่ำ-ต่ำ-สูง	720	100.0	-		-		-		-		-		-		720
กลาง-กลาง-สูง	1620	100.0	-		-		-		-		-		-		1620
*ต่ำ-กลาง-สูง	1944	93.1	44	2.1	54	2.6	46	2.2	-		-		-		2088
*สูง-สูง-ต่ำ	648	85.7	38	5.0	34	4.5	36	4.8	-		-		-		756
*สูง-สูง-กลาง	1708	94.9	28	1.6	31	1.7	33	1.8	-		-		-		1800
*สูง-สูง-สูง	1167	88.4	55	4.2	53	4.0	41	3.1	-		1	0.1	3	0.2	1320
Total	9307		165		172		156		-		1		3		9804

*รูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัวที่ผลการคัดเลือกตัวแบบมีทั้งตัวแบบเต็มรูปและตัวแบบลดรูป

4.1.3 ผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ MSPE โดยเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน

ตารางที่ 4.1.3 ตารางผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ MSPE โดยเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน

รูปแบบ ความสัมพันธ์ของตัวแปรอิสระ 3 ตัว	ตัวแบบที่ 1		ตัวแบบที่ 2		ตัวแบบที่ 3		ตัวแบบที่ 4		ตัวแบบที่ 5		ตัวแบบที่ 6		ตัวแบบที่ 7		Total
	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	
ต่ำ-ต่ำ-ต่ำ	96	100.0	-		-		-		-		-		-		96
ต่ำ-ต่ำ-กลาง	432	100.0	-		-		-		-		-		-		432
กลาง-กลาง-ต่ำ	648	100.0	-		-		-		-		-		-		648
กลาง-กลาง-กลาง	324	100.0	-		-		-		-		-		-		324
ต่ำ-ต่ำ-สูง	720	100.0	-		-		-		-		-		-		720
*กลาง-กลาง-สูง	1597	98.6	8	0.5	8	0.5	7	0.4	-		-		-		1620
*ต่ำ-กลาง-สูง	1896	90.8	66	3.2	71	3.4	55	2.6	-		-		-		2088
*สูง-สูง-ต่ำ	644	85.2	34	4.5	34	4.5	26	3.4	8	1.1	4	0.5	6	0.8	756
*สูง-สูง-กลาง	1686	93.7	32	1.8	26	1.4	20	1.1	12	0.7	12	0.7	12	0.7	1800
*สูง-สูง-สูง	1146	86.8	49	3.7	54	4.1	55	4.2	4	0.3	4	0.3	8	0.6	1320
Total	9189		189		193		163.		24		20		26		9804

---

*รูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัวที่ผลการคัดเลือกตัวแบบมีทั้งตัวแบบเต็มรูปและตัวแบบลดรูป

ตารางที่ 4.1.4 ตารางผลการคัดเลือกตัวแบบที่ใช้เกณฑ์ MSPE โดยเปรียบเทียบภายในระแนงตัวแบบ

รูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัว	ตัวแบบที่ 1		ตัวแบบที่ 2		ตัวแบบที่ 3		ตัวแบบที่ 4		ตัวแบบที่ 5		ตัวแบบที่ 6		ตัวแบบที่ 7		Total
	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	จำนวน	%	
ต่ำ-ต่ำ-ต่ำ	96	100.0	-		-		-		-		-		-		96
ต่ำ-ต่ำ-กลาง	432	100.0	-		-		-		-		-		-		432
กลาง-กลาง-ต่ำ	648	100.0	-		-		-		-		-		-		648
กลาง-กลาง-กลาง	324	100.0	-		-		-		-		-		-		324
*ต่ำ-ต่ำ-สูง	620	86.1	37	5.1	30	4.2	33	4.6	-		-		-		720
*กลาง-กลาง-สูง	1350	83.3	99	6.1	81	5.0	90	5.6	-		-		-		1620
*ต่ำ-กลาง-สูง	1715	82.1	135	6.5	109	5.2	129	6.2	-		-		-		2088
*สูง-สูง-ต่ำ	530	70.1	38	5.0	32	4.2	42	5.6	36	4.8	42	5.6	36	4.8	756
*สูง-สูง-กลาง	1364	75.8	108	6.0	102	5.7	112	6.2	34	1.9	40	2.2	40	2.2	1800
*สูง-สูง-สูง	784	59.4	126	9.5	123	9.3	139	10.5	40	3.0	56	4.2	52	3.9	1320
Total	7863		543		477		545		110		138		128		9804

*รูปแบบความสัมพันธ์ของตัวแปรอิสระ 3 ตัวที่ผลการคัดเลือกตัวแบบมีทั้งตัวแบบเต็มรูปและตัวแบบลดรูป

จากตารางที่ 4.1.1-4.1.4 สรุปได้ว่า

1. เมื่อทุกคู่ของความสัมพันธ์ระหว่างตัวแปรอิสระ 3 ตัว มีค่าไม่เกิน 0.45 การใช้เกณฑ์ RSS และ MSPE จะให้ผลการคัดเลือกตัวแบบเหมือนกัน ไม่ว่าจะใช้วิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน หรือวิธีเปรียบเทียบภายในระแนงตัวแบบ โดยตัวแบบที่เลือกได้คือ ตัวแบบเต็มรูป
2. เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระ 3 ตัวอยู่ในระดับสูงมากกว่า 1 คู่ หรืออยู่ในรูปแบบความสัมพันธ์แบบ ต่ำ-กลาง-สูง การคัดเลือกตัวแบบด้วยวีธีเปรียบเทียบภายในระแนงตัวแบบ โดยการใช้เกณฑ์ RSS โอกาสที่ตัวแบบที่เลือกได้จะเป็นตัวแบบลดรูปจะมากกว่าการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน
3. เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระ 3 ตัวบางคู่อยู่ในระดับสูงหรือมีค่าตั้งแต่ 0.55 ขึ้นไปการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบ โดยการใช้เกณฑ์ MSPE โอกาสที่ตัวแบบที่เลือกได้จะเป็นตัวแบบลดรูปจะมากกว่า การคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน
4. เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระ 3 ตัวบางคู่มีค่าตั้งแต่ 0.55 ขึ้นไป การคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบ โอกาสที่ตัวแบบที่เลือกได้จากการใช้เกณฑ์ MSPE จะเป็นตัวแบบลดรูปจะมากกว่า ตัวแบบที่เลือกได้จากการใช้เกณฑ์ RSS

จากผลสรุปเห็นได้ชัดว่าเมื่อความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับที่ไม่สูงนัก ตัวแปรอิสระที่อยู่ในตัวแบบความถดถอยที่กำลังพิจารณาอยู่นั้นก็น่าจะมีความเหมาะสมทุกตัว ดังนั้นเกณฑ์การคัดเลือกตัวแบบที่ดี ผลลัพธ์ที่ได้ก็ควรเลือกตัวแบบเต็มรูป หรือตัวแบบความถดถอยที่ไม่มีการตัดตัวแปรอิสระออกเลย ซึ่งจากผลสรุป การใช้เกณฑ์ RSS และ MSPE ก็ได้ผลตามนั้น แต่เมื่อตัวแปรอิสระมีความสัมพันธ์กันมากขึ้น การตัดตัวแปรอิสระบางตัวออกจากตัวแบบ หรือการเลือกตัวแบบลดรูป อาจเหมาะสมกว่าการเลือกตัวแบบเต็มรูป ดังนั้นเมื่อต้องทำการเปรียบเทียบกันระหว่างตัวแบบเต็มรูปกับตัวแบบลดรูป หรือระหว่างตัวแบบลดรูปด้วยกันเอง การใช้เกณฑ์ MSPE ที่พิจารณาความซับซ้อนของตัวแบบ (จำนวนตัวแปรอิสระ) ด้วย ก็น่าจะมีความเหมาะสมมากกว่าการใช้เกณฑ์ RSS จากผลสรุปที่ได้ก็เป็นไปตามนั้นเช่นกัน คือ เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระบางคู่อยู่ในระดับสูงการใช้เกณฑ์ MSPE มีโอกาสที่จะเลือกตัวแบบลดรูปมากกว่าการใช้เกณฑ์ RSS

นอกจากนี้ในบางครั้งตัวแบบเต็มรูปอาจจะมีค่าเกณฑ์การคัดเลือกที่ไม่แตกต่างจากค่าเกณฑ์การคัดเลือกของตัวแบบลดรูปบางตัวแบบก็ได้ กล่าวคือตัวแบบลดรูปนั้นให้ผลไม่แตกต่าง

จากตัวแบบเต็มรูปอย่างมีนัยสำคัญ ซึ่งถ้าเป็นไปตามนี้ การคัดเลือกตัวแบบที่ดีก็ควรเลือกตัวแบบลดรูป เพราะว่าจะมีความผิดพลาดจากการประมาณน้อยลงเนื่องจากมีตัวแปรอิสระน้อยลง ดังนั้น การเปรียบเทียบตัวแบบภายในระแนงตัวแบบ ที่ในขั้นตอนการเปรียบเทียบมีการทดสอบสมมติฐานหานัยสำคัญของความแตกต่างระหว่างค่าผลรวมกำลังสองของการประมาณของตัวแบบซับซ้อนและตัวแบบง่ายขึ้น น่าจะมีโอกาสเลือกตัวแบบลดรูปมากกว่าวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน ซึ่งจากผลสรุปการใช้เกณฑ์ RSS และ MSPE ก็ได้ผลตามนั้นเช่นกัน

4.2 ปัจจัยที่มีผลต่อค่า RSS และ MSPE

ในหัวข้อนี้จะศึกษาการเปลี่ยนแปลงของค่า RSS และ MSPE เมื่อขนาดตัวอย่างเพิ่มขึ้นและส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเพิ่มขึ้น ซึ่งเลือกศึกษาเฉพาะค่า RSS และ MSPE ของ *ตัวแบบเต็มรูปเท่านั้น* โดยแยกวิเคราะห์ตามจำนวนตัวแปรอิสระ 3 ตัวแปร และ 4 ตัวแปร

การนำเสนอผลการวิจัยในหัวข้อนี้จะแสดงผลในรูปกราฟความสัมพันธ์ และตาราง โดยพิจารณาที่ระดับความสัมพันธ์ 3 รูปแบบ ดังนี้

- ความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่เท่ากับ 0.05
- ความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่เท่ากับ 0.45
- ความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่เท่ากับ 0.95

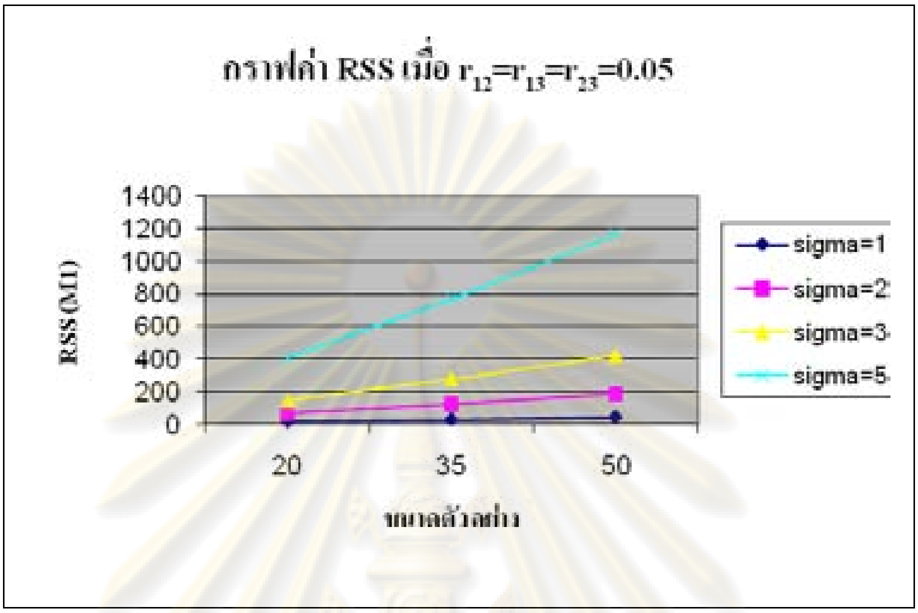
ความสัมพันธ์ระหว่างค่า RSS กับขนาดตัวอย่างและส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 ตัวแปร

กำหนดให้

r_{12} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_1 กับ X_2

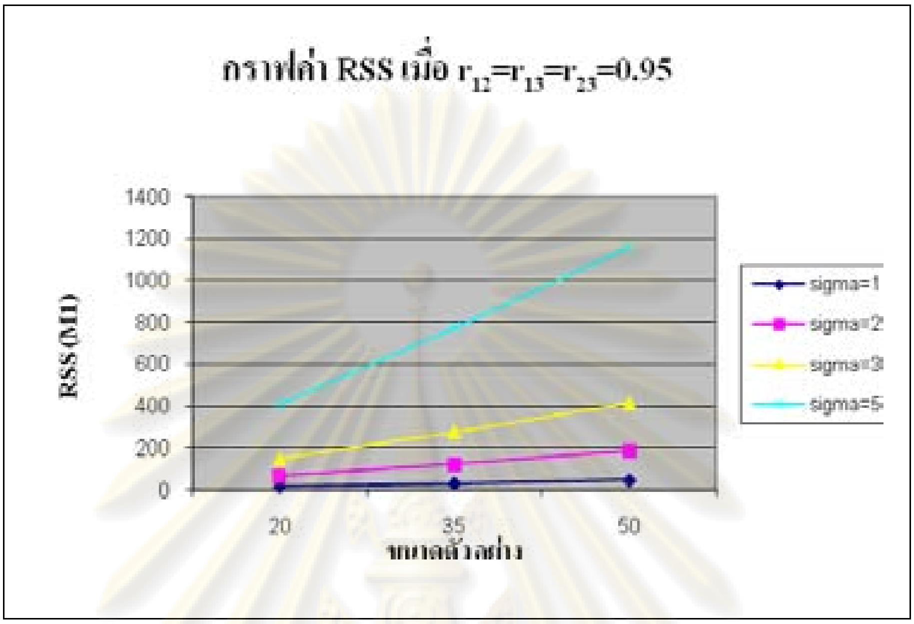
r_{13} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_1 กับ X_3

r_{23} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_2 กับ X_3

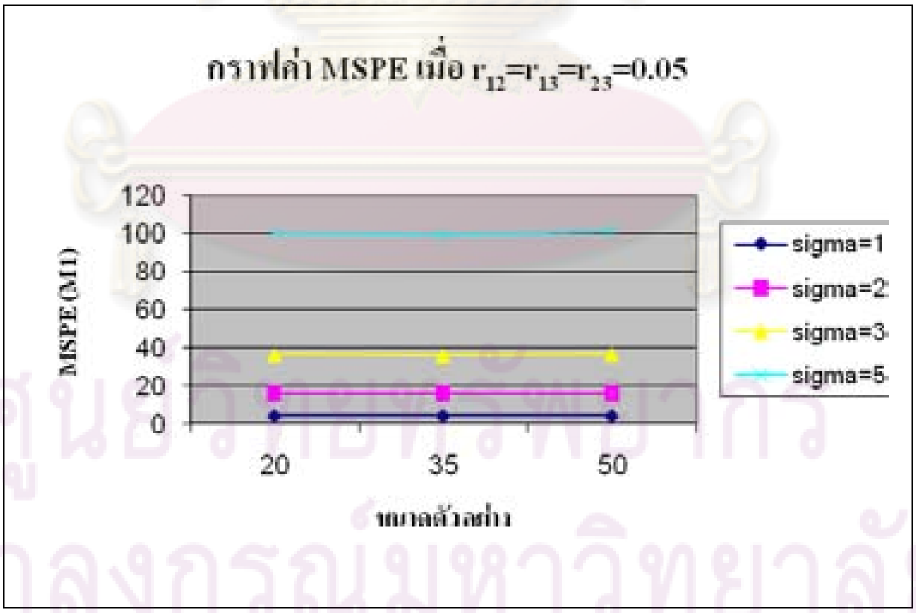


<u>**รูปที่ 4.2.1.1** กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 3 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ</u>

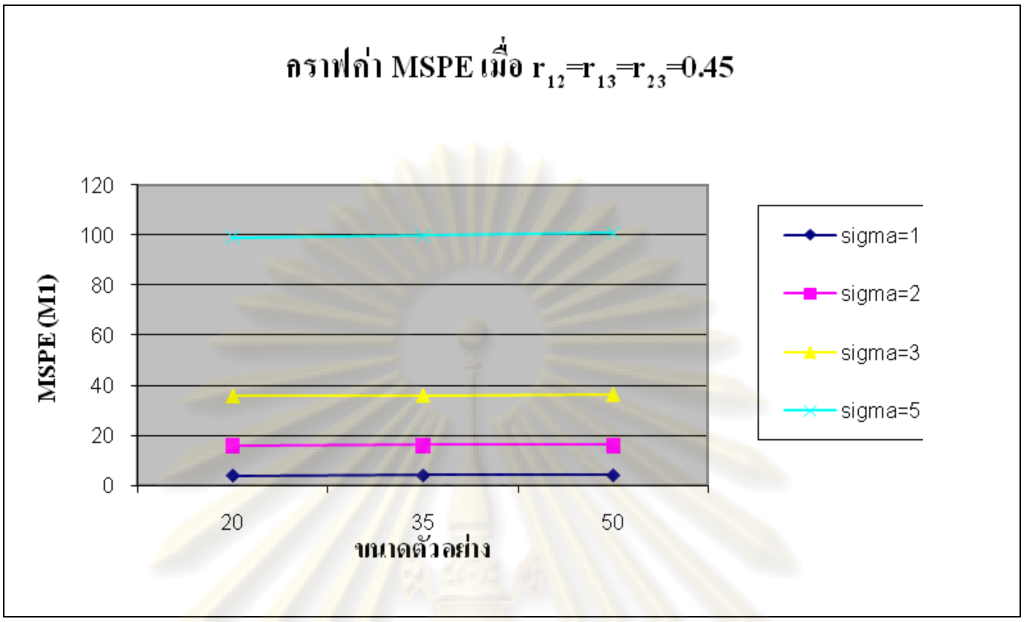
<u>**รูปที่ 4.2.1.2** กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 3 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับกลาง</u>



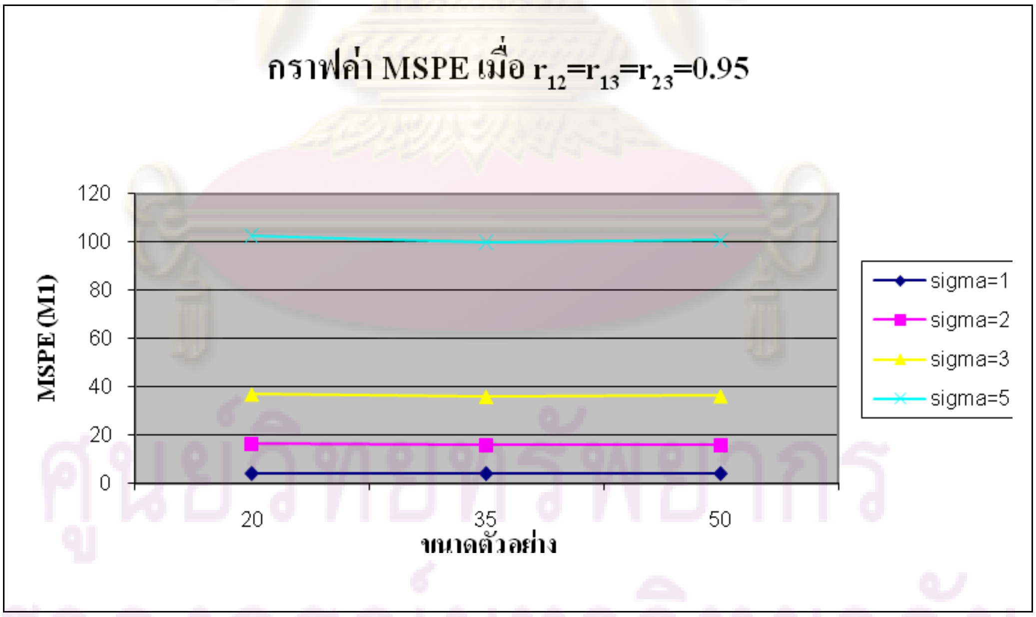
<u>**รูปที่ 4.2.1.3** กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 3 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง</u>



<u>**รูปที่ 4.2.1.4** กราฟความสัมพันธ์ระหว่างค่า MSPE ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 3 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ</u>



<u>**รูปที่ 4.2.1.5** กราฟความสัมพันธ์ระหว่างค่า MSPE ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 3 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับกลาง</u>



<u>**รูปที่ 4.2.1.6** กราฟความสัมพันธ์ระหว่างค่า MSPE ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 3 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง</u>

ความสัมพันธ์ระหว่างค่า RSS กับขนาดตัวอย่างและส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เมื่อจำนวนตัวแปรอิสระเท่ากับ 4 ตัวแปร

กำหนดให้

r_{12} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_1 กับ X_2

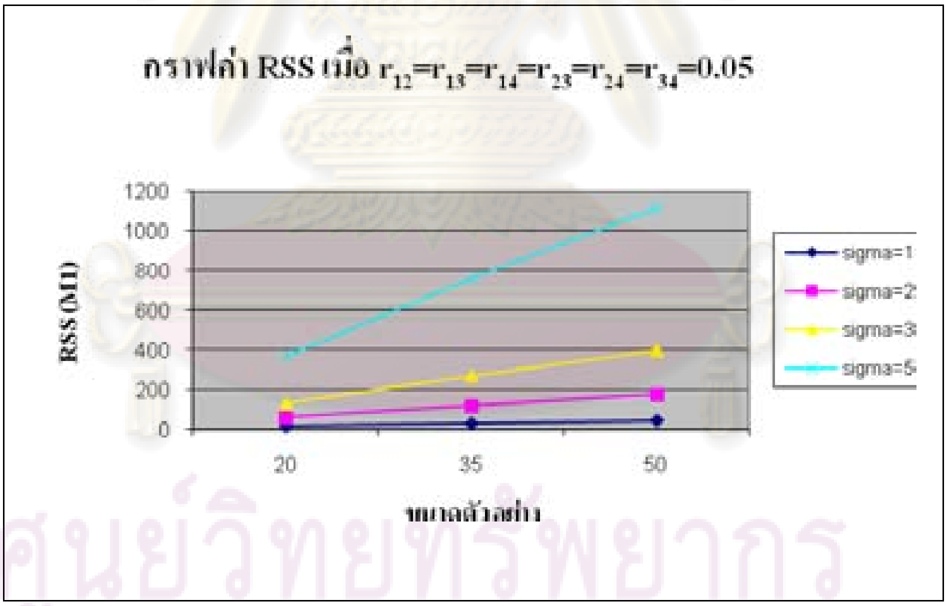
r_{13} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_1 กับ X_3

r_{14} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_1 กับ X_4

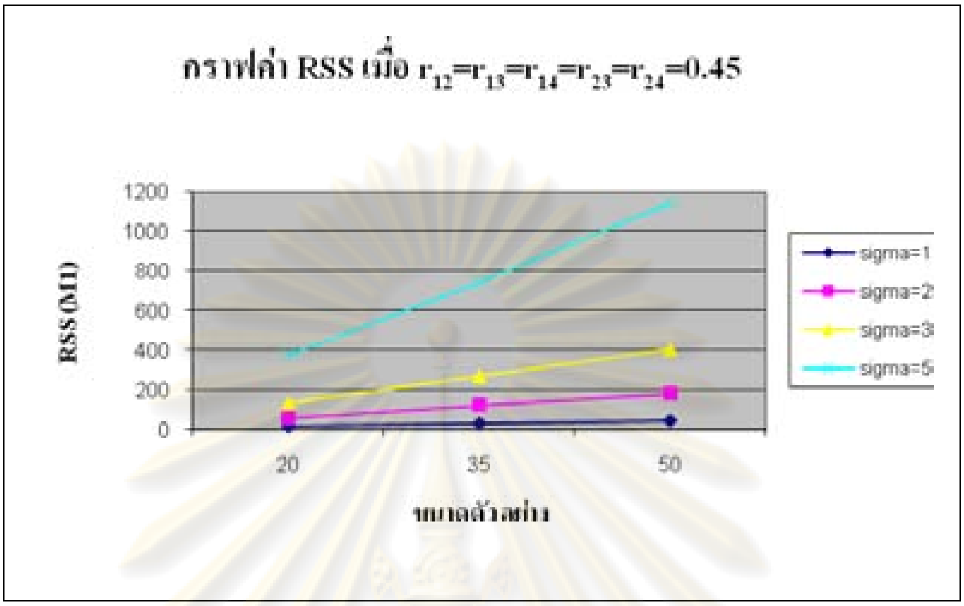
r_{23} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_2 กับ X_3

r_{24} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_2 กับ X_4

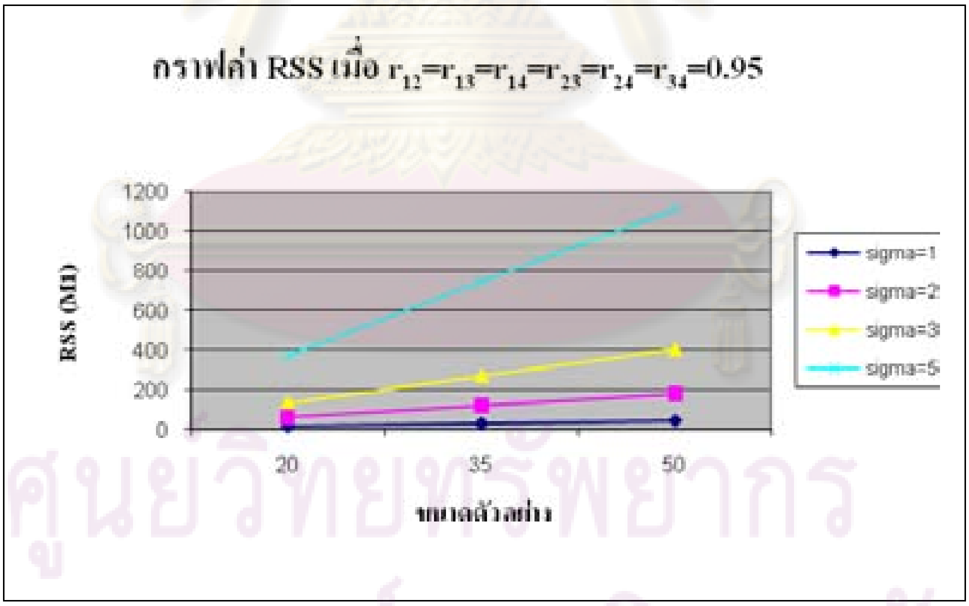
r_{34} หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระว่างตัวแปร X_3 กับ X_4



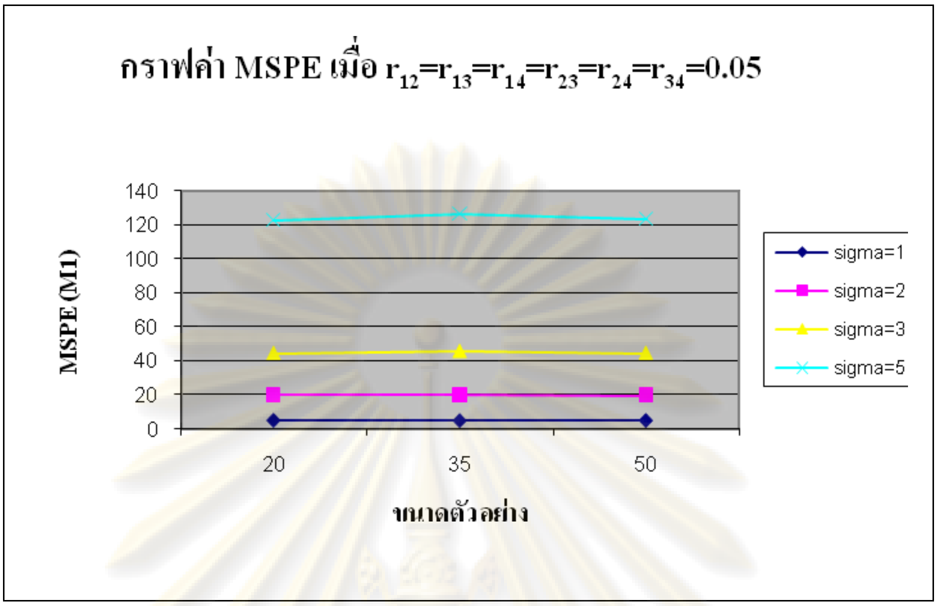
<u>**รูปที่ 4.2.2.1** กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 4 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ</u>



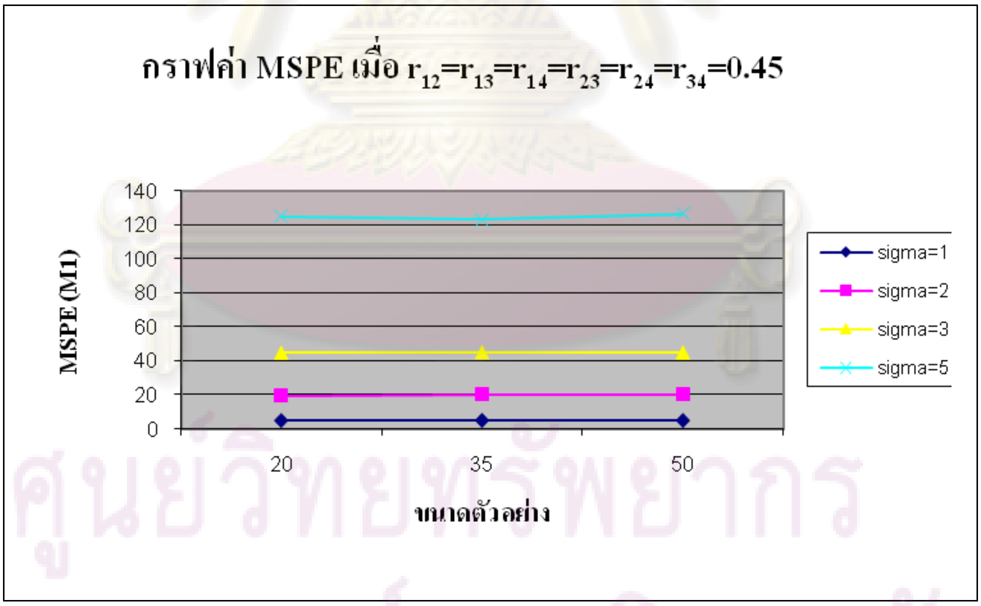
<u>**รูปที่ 4.2.2.2** กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 4 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับกลาง</u>



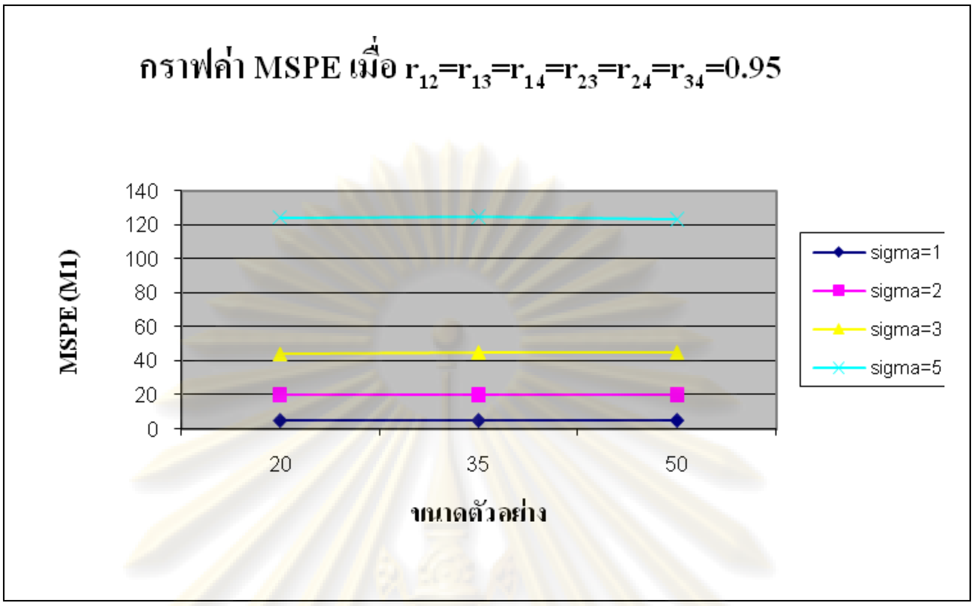
<u>**รูปที่ 4.2.2.3** กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 4 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง</u>



รูปที่ 4.2.2.4 กราฟความสัมพันธ์ระหว่างค่า MSPE ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 4 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับต่ำ



รูปที่ 4.2.2.5 กราฟความสัมพันธ์ระหว่างค่า MSPE ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 4 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับกลาง



รูปที่ 4.2.2.6 กราฟความสัมพันธ์ระหว่างค่า RSS ของตัวแบบเต็มรูป และขนาดตัวอย่าง เมื่อมีตัวแปรอิสระ 4 ตัวแปร โดยที่ความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง

จาก 4.2.1 และ 4.2.2 สามารถสรุปผลในกรณีที่มีตัวแปรอิสระ 3 ตัวแปร และกรณีที่มีตัวแปรอิสระ 4 ตัวแปรได้เหมือนกันคือ

ค่า RSS จะแปรผันตามกับขนาดตัวอย่าง และส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน เนื่องจากค่า RSS เป็นค่าที่ใช้วัดความแตกต่างระหว่างค่าของตัวแปรตามก่อนการประมาณและหลังการประมาณ หรือก็คือเป็นการรวมความผิดพลาดระหว่างค่าประมาณกับค่าจริงของทุกๆหน่วยตัวอย่างที่พิจารณา ดังนั้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้น ความผิดพลาดดังกล่าวจึงมีโอกาสเพิ่มมากขึ้นด้วย เป็นผลให้ค่า RSS แปรผันตามกับขนาดตัวอย่าง

ในขณะที่ ค่า MSPE แปรผันตามกับส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเท่านั้น โดยที่เมื่อขนาดตัวอย่างเพิ่มมากขึ้น ค่า MSPE มีแนวโน้มคงที่ ซึ่งถ้าพิจารณาจากสูตรการคำนวณค่า MSPE ($MSPE = RSS - ns^2 + 2ps^2$) และจากจำนวนพารามิเตอร์ในตัวแบบ (p) ที่มีจำนวนน้อยกว่าขนาดตัวอย่าง (n) หลายเท่า จะเห็นว่าเมื่อขนาดตัวอย่างเพิ่มขึ้น ค่า RSS ที่เพิ่มขึ้น จะถูกลบออกด้วยเทอมที่ขึ้นกับขนาดตัวอย่างอยู่เสมอ จึงเป็นผลให้ค่า MSPE มีแนวโน้มคงที่เมื่อขนาดตัวอย่างเพิ่มมากขึ้น

ข้อสังเกต *ถ้าจำนวนพารามิเตอร์ในตัวแบบเป็นครึ่งหนึ่งของขนาดตัวอย่าง* $(n = 2p)$ *จะได้ว่า* $MSPE = RSS$ *หรือกล่าวได้ว่าการพิจารณาตัวแบบความถดถอยที่ใช้จำนวนตัวอย่างไม่มากพอต่อจำนวนตัวแปรอิสระที่มี หรือจำนวนตัวแปรอิสระเพิ่มมากขึ้น ในขณะที่ใช้ขนาดตัวอย่างเท่าเดิม จะทำให้ ความสามารถในการเปรียบเทียบความซับซ้อนของตัวแบบของค่า MSPE น้อยลงหรือก็คือทำให้ค่า MSPE มีแนวโน้มเพิ่มมากขึ้น*

ส่วนระดับความสัมพันธ์ของตัวแปรอิสระไม่มีผลต่อค่า RSS และ MSPE เนื่องจากการที่ค่า RSS และ MSPE จะเพิ่มหรือลดนั้นขึ้นอยู่กับว่าค่าประมาณตัวแปรตามต่างจากค่าจริงของตัวแปรตามมากหรือน้อย ถึงแม้ว่าความสัมพันธ์ระหว่างตัวแปรอิสระจะอยู่ในระดับสูง แต่ถ้าตัวแปรอิสระเหล่านั้นมีความสัมพันธ์กับตัวแปรตามที่กำลังศึกษาอยู่ในระดับหนึ่ง ก็ย่อมได้ค่าประมาณที่ไม่แตกต่างจากค่าจริงมากนัก ดังนั้นระดับความสัมพันธ์ของตัวแปรอิสระจึงไม่มีผลต่อค่า RSS และ MSPE

ต่อไปเพื่อศึกษาอิทธิพลของจำนวนตัวแปรอิสระที่มีต่อค่า RSS และ MSPE ผู้วิจัยได้นำผลลัพธ์ของกรณีที่มีจำนวนตัวแปรอิสระเท่ากับ 3 และกรณีที่มีจำนวนตัวแปรอิสระเท่ากับ 4 มาเปรียบเทียบกัน และแสดงผลในรูปตารางดังนี้

ตารางที่ 4.2 ตารางเปรียบเทียบค่า RSS และ MSPE เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 และ 4

ความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่	n	σ	RSS		MSPE	
			ตัวแปรอิสระ3 ตัวแปร	ตัวแปรอิสระ4 ตัวแปร	ตัวแปรอิสระ3 ตัวแปร	ตัวแปรอิสระ4 ตัวแปร
0.05	**20**	**1**	16.0729	15.3074	4.0182	5.1024
		2	64.5689	59.9723	16.1422	19.9907
		3	145.2801	133.3801	36.3200	44.4600
		5	403.5559	368.2027	100.8889	122.7342
	35	**1**	30.8844	29.8992	3.9850	4.9832
		2	123.5378	119.3750	15.9403	19.8958
		3	278.1752	273.5832	35.8935	45.5972
		5	772.7090	758.8581	99.7043	126.4764
	50	**1**	45.8118	45.1282	3.9836	5.0142
		2	183.2472	177.2428	15.9345	19.6936
		3	420.0877	401.0278	36.5293	44.5586
		5	1166.9104	1113.9000	101.4704	123.7667
0.45	**20**	**1**	15.7564	14.7570	3.9391	4.9190
		2	63.4177	58.7202	15.8544	19.5734
		3	142.6899	134.4859	35.6724	44.8286
		5	396.3608	375.4382	99.0902	125.1461
	35	**1**	30.9796	29.9497	3.9973	4.9916
		2	123.9187	121.9805	15.9895	20.3300
		3	278.4539	271.2640	35.9295	45.2106
		5	773.4831	738.2142	99.8042	123.0357
	50	**1**	46.3235	44.2611	4.0281	4.9179
		2	185.2941	181.8459	16.1125	20.2051
		3	417.6293	405.4216	36.3155	45.0468
		5	1160.0814	1138.5770	100.8766	126.5085

ตารางที่ 4.2(ต่อ) ตารางเปรียบเทียบค่า RSS และ MSPE เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 และ 4

ความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่	n	σ	RSS		MSPE	
			ตัวแปรอิสระ3 ตัวแปร	ตัวแปรอิสระ4 ตัวแปร	ตัวแปรอิสระ3 ตัวแปร	ตัวแปรอิสระ4 ตัวแปร
0.95	**20**	**1**	16.2480	14.9930	4.0620	4.9862
		2	65.6508	61.3847	16.4127	20.2070
		3	147.7144	135.5991	36.9286	44.2149
		5	410.3178	373.5144	102.5794	124.5048
	35	**1**	30.8232	29.9372	3.9771	4.9895
		2	123.2929	121.2494	15.9087	20.2082
		3	278.7333	271.4034	35.9655	45.2339
		5	774.2593	749.7894	99.9044	124.9649
	50	**1**	46.1316	45.1640	4.0114	5.0182
		2	184.5267	179.6973	16.0458	19.9663
		3	417.1347	407.4939	36.2725	45.2771
		5	1158.7075	1110.3750	100.7571	123.3750

จากตารางที่ 4.2 สรุปได้ว่า ณ ระดับความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่ ขนาดตัวอย่างและส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนหนึ่งๆ เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น จะได้ว่าค่า RSS ลดลง แต่ในทางกลับกันค่า MSPE เพิ่มขึ้น ที่เป็นเช่นนี้เนื่องจากการทดลองข้างต้นเป็นการจำลองสถานการณ์เพื่อเปรียบเทียบตัวแบบที่ติดกลุ่มกัน ดังนั้นตัวแบบที่มีจำนวนตัวแปรอิสระ 3 ตัวแปรจึงเสมือนเป็นตัวแบบลดรูปของตัวแบบที่มีจำนวนตัวแปรอิสระ 4 ตัวแปร เพราะฉะนั้นจึงทำให้ตัวแบบที่มีตัวแปรอิสระ 4 ตัวแปร มีค่าผลรวมกำลังสองของความถดถอยมากกว่าตัวแบบที่มีตัวแปรอิสระ 3 ตัวแปร เป็นผลให้เมื่อตัวแปรอิสระเพิ่มมากขึ้น ความแตกต่างระหว่างค่าของตัวแปรตามก่อนการประมาณและหลังการประมาณจะลดน้อยลง กล่าวคือ ค่า RSS จะลดลงนั่นเอง แต่ในทางกลับกันเหตุผลที่ค่า MSPE เพิ่มขึ้นนั้น ตามที่ได้กล่าวมาแล้วข้างต้น เมื่อจำนวนตัวแปรอิสระเพิ่มมากขึ้น ในขณะที่ใช้ขนาดตัวอย่างเท่าเดิม จะทำให้ค่า MSPE มีแนวโน้มเพิ่มขึ้น

4.3 การคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ (Lattice) โดยใช้เกณฑ์ RSS และ MSPE

ในหัวข้อนี้จะศึกษาการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบระหว่างการใช้เกณฑ์ RSS และ MSPE ว่ามีลักษณะการคัดเลือกอย่างไร โดยนำเสนอในรูปตารางแยกวิเคราะห์ตามจำนวนตัวแปรอิสระ 3 ตัวแปร และ 4 ตัวแปร ในแต่ละตารางจะแสดงตัวแบบที่ถูกเลือกโดยวิธีการเปรียบเทียบภายในตัวแบบจากการใช้เกณฑ์ RSS และ MSPE และระบุค่า RDRSS และ RDMSPE ของทุกตัวแบบที่เป็นไปได้ เพื่อใช้ในการเปรียบเทียบกับผลลัพธ์ที่ได้จากวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้พร้อมกัน

4.3.1 ผลการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ (Lattice) โดยใช้เกณฑ์ RSS และ MSPE เมื่อมีตัวแปรอิสระ 3 ตัวแปร

จากหัวข้อ 4.12 และ 4.14 พบว่าผลการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ โดยการใช้เกณฑ์ RSS และ MSPE จะมีความแตกต่างกัน เมื่อมีความสัมพันธ์ระหว่างตัวแปรอิสระบางคู่อยู่ในระดับสูง ดังนั้นผู้วิจัยจึงทำการวิเคราะห์ผลเฉพาะกรณีที่มีความสัมพันธ์ระหว่างตัวแปรอิสระบางคู่อยู่ในระดับสูง ซึ่งแบ่งออกเป็นกรณีย่อยๆ ดังนี้

- r_{12}= 0.95, r_{13}= 0.25, r_{23}= 0.15
- r_{12}= 0.95, r_{13}= 0.45, r_{23}= 0.15
- r_{12}= 0.85, r_{13}= 0.85, r_{23}= 0.45
- r_{12}= 0.85, r_{13}= 0.85, r_{23}= 0.65

52

ตารางที่ 4.3.1.1 ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 3 ตัวแปร เมื่อ r_{12}=0.95 r_{13}=0.25 r_{23}=0.15

n	σ	ตัวแบบที่ถูกเลือก			ตัวแบบ							
		RSS	MSPE		1	2	3	4	5	6	7	8
20	1	M1	M3	RDRSS	0	49.39	13.80	12.70	52.09	60.82	81.05	87.15
		15.65	3.39	RDMSPE	0	77.29	12.32	7.55	77.00	83.90	94.15	96.24
	2	M1	M4	RDRSS	0	49.10	13.58	12.93	51.96	60.71	81.69	87.72
		62.47	13.11	RDMSPE	0	77.06	11.40	8.59	76.89	83.82	94.40	96.44
	3	M1	M4	RDRSS	0	49.05	12.85	13.11	51.89	61.38	80.75	87.19
		144.90	31.05	RDMSPE	0	77.01	8.24	9.36	76.82	84.27	94.04	96.26
	5	M1	M3	RDRSS	0	49.59	12.33	14.20	52.43	62.11	81.01	87.38
		391.05	85.31	RDMSPE	0	77.45	5.89	13.95	77.32	84.75	94.14	96.32
35	1	M1	M4	RDRSS	0	47.81	11.99	10.70	49.57	58.69	80.33	86.44
		30.94	3.64	RDMSPE	0	86.84	35.76	29.99	86.87	90.92	96.84	97.96
	2	M1	M4	RDRSS	0	48.36	10.42	10.82	49.94	59.24	79.67	86.19
		124.38	14.64	RDMSPE	0	87.11	28.63	30.59	87.06	91.12	96.71	97.91
	3	M1	M3	RDRSS	0	47.40	10.85	11.53	49.04	59.02	80.25	86.34
		277.90	32.81	RDMSPE	0	86.64	30.69	33.77	86.59	91.04	96.82	97.94
	5	M1	M3	RDRSS	0	48.31	10.58	10.79	49.88	59.42	79.87	86.38
		780.32	91.54	RDMSPE	0	87.08	29.43	30.41	87.03	91.19	96.75	97.94
50	1	M1	M1	RDRSS	0	47.08	10.76	9.29	48.24	57.88	79.39	85.80
		45.57	3.96	RDMSPE	0	90.68	47.01	40.41	90.67	93.67	97.74	98.55
	2	M1	M1	RDRSS	0	47.45	10.11	10.10	48.52	58.58	80.06	86.09
		173.30	15.07	RDMSPE	0	90.81	44.22	44.20	90.77	93.85	97.83	98.58
	3	M1	M1	RDRSS	0	48.29	9.99	9.78	49.33	58.77	79.91	86.16
		399.44	34.73	RDMSPE	0	91.10	43.70	42.73	91.07	93.90	97.81	98.59
	5	M1	M1	RDRSS	0	47.16	9.87	9.90	48.26	58.39	79.65	85.94
		1098.76	95.54	RDMSPE	0	90.71	43.14	43.31	90.68	93.80	97.78	98.57

ตารางที่ 4.3.1.2 ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 3 ตัวแปร เมื่อ r_{12}=0.95 r_{13}=0.45 r_{23}=0.15

n	σ	ตัวแบบที่ถูกเลือก			ตัวแบบ							
		RSS	MSPE		1	2	3	4	5	6	7	8
20	1	M3	M3	RDRSS	0	6.79	6.14	6.15	34.55	67.67	80.40	88.30
		15.72	2.61	RDMSPE	23.83	3.72	0	0.06	63.92	90.90	95.36	97.43
	2	M3	M3	RDRSS	0	6.50	5.85	5.75	33.12	66.27	80.57	88.20
		62.41	10.41	RDMSPE	25.59	4.35	0.58	0	62.43	90.53	95.51	97.47
	3	M3	M3	RDRSS	0	6.41	5.69	5.51	32.42	66.21	80.58	88.08
		141.15	23.80	RDMSPE	26.66	5.21	1.04	0	61.78	90.64	95.58	97.48
	5	M2	M3	RDRSS	0	6.45	5.73	5.68	32.46	66.54	80.41	88.14
		400.78	68.32	RDMSPE	25.90	4.48	0.31	0	61.47	90.68	95.49	97.46
35	1	M3	M3	RDRSS	0	3.70	3.13	3.10	29.90	65.04	78.98	86.98
		31.79	2.77	RDMSPE	25.21	6.29	0.35	0	77.37	94.81	97.43	98.54
	2	M3	M3	RDRSS	0	3.64	3.10	3.19	29.33	64.63	79.35	87.06
		124.65	10.84	RDMSPE	25.23	5.68	0	1.06	76.75	94.72	97.49	98.55
	3	M3	M3	RDRSS	0	3.81	3.09	2.96	29.47	64.60	78.99	86.84
		284.54	24.57	RDMSPE	26.39	8.77	1.50	0	77.27	94.79	97.47	98.55
	5	M3	M3	RDRSS	0	3.81	3.25	3.31	30.18	65.27	79.50	87.18
		801.25	71.80	RDMSPE	23.96	5.76	0	0.69	77.31	94.78	97.47	98.54
50	1	M3	M4	RDRSS	0	2.81	2.13	2.09	28.04	63.96	78.97	86.71
		47.12	2.817	RDMSPE	25.50	10.48	0.68	0	83.38	96.35	98.27	99.00
	2	M3	M3	RDRSS	0	2.92	2.31	2.31	28.47	64.67	79.39	87.03
		184.36	11.16	RDMSPE	22.82	8.74	0	0.07	83.14	96.33	98.26	98.99
	3	M3	M3	RDRSS	0	2.68	2.09	2.13	28.15	63.86	78.55	86.53
		422.20	24.53	RDMSPE	25.44	8.72	0	0.63	83.45	96.33	98.23	98.98
	5	M3	M3	RDRSS	0	2.88	2.28	2.24	27.72	63.74	78.88	86.61
		1185.50	70.18	RDMSPE	23.61	9.21	0.61	0	82.68	96.22	98.22	98.97

54

ตารางที่ 4.3.1.3 ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 3 ตัวแปร เมื่อ r_{12}=0.85 r_{13}=0.85 r_{23}=0.45

n	σ	ตัวแบบที่ถูกเลือก			ตัวแบบ							
		RSS	MSPE		1	2	3	4	5	6	7	8
20	1	M4	M5	RDRSS	0	6.71	6.61	6.36	12.01	71.15	70.52	90.10
		16.97	1.62	RDMSPE	45.42	30.73	30.30	29.27	0	94.47	94.30	98.48
	2	M4	M5	RDRSS	0	6.60	6.58	5.84	11.88	70.69	71.55	90.23
		67.66	6.50	RDMSPE	46.08	31.09	31.02	27.91	0	94.41	94.64	98.52
	3	M4	M5	RDRSS	0	6.91	6.91	6.32	12.09	70.17	69.50	89.54
		154.93	16.26	RDMSPE	44.98	30.94	30.96	28.51	0	94.15	93.96	98.37
	5	M4	M5	RDRSS	0	5.82	5.86	4.76	11.14	70.17	70.51	89.76
		420.10	38.64	RDMSPE	49.87	32.91	33.06	28.38	0	94.67	94.76	98.55
35	1	M3	M5	RDRSS	0	4.03	3.94	3.66	6.95	68.84	67.70	88.92
		32.53	1.72	RDMSPE	42.13	29.89	29.24	27.14	0	96.62	96.44	99.06
	2	M4	M5	RDRSS	0	4.22	4.21	3.38	7.12	68.47	68.70	88.89
		128.20	6.96	RDMSPE	40.56	29.35	29.26	22.91	0	96.47	96.50	99.03
	3	M4	M5	RDRSS	0	4.25	4.15	3.46	7.22	68.96	68.36	89.04
		284.86	14.76	RDMSPE	39.66	28.51	27.75	22.40	0	96.50	96.40	99.03
	5	M4	M5	RDRSS	0	3.63	3.55	3.361	6.64	69.25	67.90	88.91
		807.32	41.96	RDMSPE	44.92	30.46	29.90	28.44	0	96.84	96.64	99.11
50	1	M4	M5	RDRSS	0	3.13	3.19	2.53	5.20	68.13	68.56	88.80
		46.60	1.69	RDMSPE	36.90	27.59	28.18	20.98	0	97.43	97.48	99.30
	2	M4	M5	RDRSS	0	3.11	3.07	2.51	5.10	68.64	68.48	88.65
		186.92	6.84	RDMSPE	38.21	28.91	28.52	22.36	0	97.54	97.53	99.31
	3	M3	M5	RDRSS	0	2.95	2.87	2.33	4.98	68.43	68.13	88.74
		421.44	14.55	RDMSPE	39.70	29.02	28.22	22.15	0	97.58	97.55	99.33
	5	M3	M5	RDRSS	0	3.23	3.22	2.77	5.18	68.35	68.00	88.45
		1184.43	42.76	RDMSPE	37.15	28.93	28.83	24.04	0	97.47	97.43	99.28

55

<u>ตารางที่ 4.3.1.4</u> ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 3 ตัวแปร เมื่อ r_{12}=0.85 r_{13}=0.85 r_{23}=0.65

n	σ	ตัวแบบที่ถูกเลือก			ตัวแบบ							
		RSS	MSPE		1	2	3	4	5	6	7	8
20	1	M1	M1	RDRSS	0	25.21	25.47	16.62	36.74	63.62	64.63	90.63
		14.04	3.51	RDMSPE	0	45.89	46.45	22.92	56.95	85.70	86.32	97.38
	2	M1	M1	RDRSS	0	24.74	24.48	17.19	36.12	63.73	63.52	90.06
		56.26	14.06	RDMSPE	0	44.90	44.35	24.83	55.79	85.77	85.64	97.20
	3	M1	M1	RDRSS	0	25.40	25.22	16.75	37.03	64.05	64.12	90.40
		129.15	32.29	RDMSPE	0	46.30	45.91	23.35	57.49	85.97	86.01	97.31
	5	M1	M1	RDRSS	0	25.08	24.98	16.32	36.66	64.39	63.97	90.38
		357.90	89.48	RDMSPE	0	45.61	45.40	21.89	56.81	86.18	85.92	97.30
35	1	M1	M1	RDRSS	0	22.50	22.99	13.00	32.92	60.81	60.80	89.31
		30.96	3.99	RDMSPE	0	63.63	64.46	39.70	73.70	91.69	91.68	98.44
	2	M1	M1	RDRSS	0	23.30	23.54	13.11	33.89	61.12	61.41	89.28
		123.37	15.92	RDMSPE	0	64.96	65.35	40.10	74.83	91.79	91.89	98.44
	3	M1	M1	RDRSS	0	23.13	22.75	13.58	33.27	61.24	61.34	89.48
		271.69	35.06	RDMSPE	0	64.69	64.05	41.79	74.12	91.83	91.87	98.47
	5	M1	M1	RDRSS	0	23.59	23.37	13.56	34.20	61.70	61.63	89.68
		753.39	97.21	RDMSPE	0	65.43	65.07	41.71	75.17	91.99	91.97	98.50
50	1	M1	M1	RDRSS	0	21.45	22.04	13.39	31.55	60.72	60.83	89.31
		44.75	3.89	RDMSPE	0	72.52	73.35	56.10	81.14	94.38	94.40	98.95
	2	M1	M1	RDRSS	0	22.56	22.51	12.50	32.35	60.85	61.21	89.42
		177.58	15.44	RDMSPE	0	74.02	73.96	53.34	81.81	94.41	94.49	98.97
	3	M1	M1	RDRSS	0	22.29	22.19	12.77	32.01	60.62	60.90	89.16
		404.24	35.15	RDMSPE	0	73.68	73.54	54.21	81.53	94.35	94.42	98.94
	5	M1	M1	RDRSS	0	21.29	22.29	13.14	31.45	60.35	61.30	89.14
		1125.59	97.88	RDMSPE	0	72.31	73.68	55.34	81.05	94.29	94.51	98.94

4.3.2 ผลการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ (Lattice) โดยใช้เกณฑ์ RSS และ MSPE เมื่อมีตัวแปรอิสระ 4 ตัวแปร

จากหัวข้อ 4.3.1 พบว่าเมื่อขนาดตัวอย่างต่างๆกัน ผลการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ โดยการใช้เกณฑ์ RSS และ MSPE จะมีความคล้ายคลึงกัน ดังนั้นผู้วิจัยจึงทำการวิเคราะห์ผลเฉพาะที่ขนาดตัวอย่างเท่ากับ 20 เท่านั้น โดยพิจารณาที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระต่างๆ แบ่งออกเป็นกรณีย่อยๆ ดังนี้

- $r_{12}= 0.15$, $r_{13}= 0.25$, $r_{14}= 0.35$, $r_{23}= 0.15$, $r_{24}= 0.25$, $r_{34}= 0.35$
- $r_{12}= 0.45$, $r_{13}= 0.55$, $r_{14}= 0.85$, $r_{23}= 0.45$, $r_{24}= 0.55$, $r_{34}= 0.85$
- $r_{12}= 0.75$, $r_{13}= 0.85$, $r_{14}= 0.95$, $r_{23}= 0.75$, $r_{24}= 0.85$, $r_{34}= 0.95$

57

<u>ตารางที่ 4.3.2.1</u> ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 4 ตัวแปร เมื่อ $r_{12}=0.15$ $r_{13}=0.25$ $r_{14}=0.35$ $r_{23}=0.15$ $r_{24}=0.25$ $r_{34}=0.35$

N=20	σ							
	1		2		3		5	
	RSS	**MSPE**	**RSS**	**MSPE**	**RSS**	**MSPE**	**RSS**	**MSPE**
ตัวแบบที่เลือกได้	M1	M1	M1	M1	M1	M1	M1	M1
min	14.80	4.93	60.21	20.07	131.88	43.96	362.01	120.67
ตัวแบบ	RDRSS	RDMSPE	RDRSS	RDMSPE	RDRSS	RDMSPE	RDRSS	RDMSPE
1	0	0	0	0	0	0	0	0
2	47.05	69.37	46.19	68.51	46.95	69.28	46.51	68.83
3	49.35	71.62	48.32	70.64	47.64	69.97	49.75	71.99
4	51.00	73.14	50.77	72.92	51.67	73.73	52.97	74.86
5	50.21	72.42	49.58	71.84	48.21	70.52	49.30	71.57
6	72.49	87.66	71.86	87.28	71.73	87.20	72.78	87.84
7	70.87	86.66	70.83	86.64	71.13	86.83	71.73	87.20
8	69.56	85.83	68.88	85.38	69.03	85.48	70.90	86.68
9	73.03	87.98	72.45	87.64	72.34	87.57	72.76	87.82
10	70.62	86.51	70.32	86.31	69.22	85.61	70.54	86.46
11	69.73	85.94	69.36	85.70	69.22	85.61	70.35	86.34
12	82.99	93.08	82.59	92.87	82.55	92.86	83.53	93.34
13	85.05	94.07	84.94	94.02	84.42	93.77	85.24	94.16
14	83.07	93.11	82.77	92.96	82.80	92.98	83.44	93.30
15	80.24	91.65	79.86	91.45	79.61	91.32	80.90	92.01
16	90.13	96.27	89.93	96.18	89.76	96.11	90.46	96.41

58

ตารางที่ 4.3.2.2 ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 4 ตัวแปร เมื่อ r_{12}=0.45 r_{13}=0.55 r_{14}=0.85 r_{23}=0.45 r_{24}=0.55 r_{34}=0.85

N=20	σ							
	1		2		3		5	
	RSS	MSPE	RSS	MSPE	RSS	MSPE	RSS	MSPE
ตัวแบบที่เลือกได้	M1	M10	M1	M10	M1	M10	M1	M10
min	14.23	3.99	57.74	15.66	125.37	34.44	357.10	92.02
ตัวแบบ	RDRSS	RDMSPE	RDRSS	RDMSPE	RDRSS	RDMSPE	RDRSS	RDMSPE
1	0	2.46	0	0.73	0	1.18	0	0
2	11.12	0	11.58	0	11.46	0	11.97	0.78
3	21.97	32.49	19.30	24.64	21.45	30.38	18.52	21.98
4	44.62	67.67	44.65	67.13	45.93	68.61	46.33	68.65
5	20.45	28.87	20.02	26.51	20.34	27.65	19.84	25.52
6	66.30	84.01	65.86	83.42	65.47	83.22	65.83	83.28
7	53.22	73.00	54.20	73.53	55.07	74.52	55.53	74.66
8	51.88	71.60	51.63	70.82	52.96	72.37	52.30	71.34
9	65.71	83.60	65.91	83.46	65.47	83.22	65.89	83.32
10	28.82	31.04	26.91	23.89	28.32	28.67	26.35	21.45
11	51.53	71.22	51.82	71.03	52.45	71.84	52.44	71.49
12	80.67	92.08	80.88	92.05	80.99	92.15	81.02	92.07
13	87.11	95.14	86.76	94.90	87.19	95.11	87.11	95.02
14	80.47	91.98	80.93	92.08	80.57	91.93	81.05	92.08
15	55.26	72.18	55.27	71.69	56.14	72.86	55.66	71.96
16	93.60	97.75	93.65	97.73	93.84	97.81	93.80	97.77

59

ตารางที่ 4.3.2.3 ตารางค่า RSS และ MSPE สำหรับตัวแบบที่มีตัวแปรอิสระ 4 ตัวแปร เมื่อ r_{12}=0.75 r_{13}=0.85 r_{14}=0.95 r_{23}=0.75 r_{24}=0.85 r_{34}=0.95

N=20	σ							
	1		2		3		5	
	RSS	MSPE	RSS	MSPE	RSS	MSPE	RSS	MSPE
ตัวแบบที่เลือกได้	M2	M15	M2	M15	M2	M15	M2	M15
min	14.67	2.25	60.53	8.80	135.11	22.74	384.87	61.17
ตัวแบบ	RDRSS	RDMSPE	RDRSS	RDMSPE	RDRSS	RDMSPE	RDRSS	RDMSPE
1	0	22.28	0	30.02	0	21.88	0	27.07
2	7.12	6.34	6.77	14.43	6.81	4.62	6.74	10.73
3	10.22	17.46	8.95	21.80	9.85	15.81	9.84	21.35
4	19.15	40.69	16.58	41.50	17.69	37.24	18.49	43.04
5	10.61	18.71	8.84	21.44	10.16	16.83	9.57	20.51
6	41.67	66.83	41.40	69.83	41.39	66.31	40.86	67.91
7	44.79	70.49	44.49	73.13	43.19	68.50	44.36	71.86
8	24.05	32.42	22.32	34.10	23.04	28.86	23.18	34.03
9	41.90	67.11	41.27	69.68	42.38	67.54	40.85	67.90
10	16.14	0	14.28	0	16.23	0	14.99	0
11	23.78	31.59	22.07	33.32	22.62	27.45	23.51	35.02
12	67.97	87.40	68.13	88.74	67.59	87.10	67.73	88.04
13	79.46	93.19	79.36	93.83	78.45	92.72	78.50	93.22
14	68.46	87.69	67.67	88.49	67.88	87.28	67.67	88.00
15	28.08	19.99	26.59	21.08	27.69	17.67	27.40	21.79
16	94.86	98.58	94.85	98.72	94.66	98.51	94.75	98.64

จากตารางที่ 4.3.1.1-4.3.1.4 และตารางที่ 4.3.2.1-4.3.2.3 พบว่า การใช้เกณฑ์ RSS และ MSPE ในการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบ สำหรับตัวแบบที่มีตัวแปรอิสระ 3 หรือ 4 ตัวแปร จะมีลักษณะการเลือกตัวแบบอยู่ 2 ลักษณะคือ

1. เมื่อพิจารณาค่า RDRSS หรือ RDMSPE แล้วปรากฏว่าตัวแบบที่ให้ค่า RDRSS=0 หรือ RDMSPE=0 (ตัวแบบที่ให้ค่า RSS หรือ MSPE ต่ำที่สุด โดยวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้พร้อมกัน) มีค่าต่างจากตัวแบบอื่นๆมาก <u>ตัวแบบที่เลือกได้จากการเปรียบเทียบภายในระแนงตัวแบบอาจเป็นตัวแบบที่ให้ค่า RDRSS=0 หรือ RDMSPE=0 หรือ ตัวแบบอื่นๆที่มีค่า RDRSS หรือ RDMSPE น้อยเป็นอันดับที่สอง</u>
2. เมื่อตัวแบบลดรูปบางตัวแบบมีค่า RDRSS หรือ RDMSPE ใกล้เคียง 0 หรือเมื่อมีกลุ่มของตัวแบบลดรูปบางกลุ่มที่มีค่า RDRSS หรือ RDMSPE ใกล้เคียงกันและมีค่าไม่มากนัก <u>ตัวแบบที่เลือกได้จากการใช้วิธีเปรียบเทียบภายในระแนงตัวแบบอาจจะเป็นตัวแบบหนึ่งในกลุ่มตัวแบบลดรูปนั้น และไม่จำเป็นที่จะต้องเป็นตัวแบบที่มีค่า RDRSS หรือ RDMSPE น้อยเป็นอันดับที่สองก็ได้</u>

บทที่ 5

สรุปผลการวิจัย และข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองเพื่อศึกษาหาแนวทางการคัดเลือกตัวแบบความถดถอยที่เหมาะสมที่สุด ซึ่งได้พิจารณาการใช้เกณฑ์การคัดเลือกตัวแบบ 2 เกณฑ์ คือ เกณฑ์ค่าผลรวมความคลาดเคลื่อนกำลังสอง (residual sum of squares: RSS) และเกณฑ์ค่าประมาณค่าเฉลี่ยของความผิดพลาดจากการพยากรณ์กำลังสอง (mean square prediction error: MSPE) นอกจากนั้นยังพิจารณาวิธีการเปรียบเทียบตัวแบบความถดถอย 2 วิธี คือ วิธีการเปรียบเทียบตัวแบบความถดถอยที่เป็นไปได้ทั้งหมดพร้อมกัน และวิธีการเปรียบเทียบตัวแบบความถดถอยภายในระแนงตัวแบบ โดยในบทนี้จะนำเสนอสรุปผลการวิจัยแบ่งออกเป็น 3 ส่วนดังนี้

5.1 สรุปผลการวิจัย

5.2 ข้อเสนอแนะในการเลือกใช้เกณฑ์การคัดเลือกตัวแบบและวิธีการเปรียบเทียบตัวแบบ

5.3 ข้อเสนอแนะด้านการวิจัย

5.1 สรุปผลการวิจัย

จากผลการวิจัยในบทที่ 4 เราสรุปผลได้ดังนี้

5.1.1 สรุปความแตกต่างของผลลัพธ์ที่ได้จากการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน และวิธีการเปรียบเทียบตัวแบบภายในระแนงตัวแบบ

ในบทที่ 4 หัวข้อที่ 4.1 สรุปได้ว่า

1. เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระทุกคู่ในตัวแบบความถดถอยมีค่าไม่สูงนัก (มีค่าไม่เกิน 0.45) กล่าวคือตัวแปรอิสระที่อยู่ในตัวแบบความถดถอยที่กำลังพิจารณาอยู่นั้น มีความเหมาะสมในการอธิบายตัวแปรตามที่กำลังศึกษาอยู่ในระดับหนึ่ง ตัวแบบความถดถอยที่มีความเหมาะสมจึงควรเป็นตัวแบบเต็มรูป ซึ่งการใช้เกณฑ์ RSS และ MSPE จะให้ผลการคัดเลือกตัวแบบตามนั้น ไม่ว่าจะใช้วิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน หรือวิธีเปรียบเทียบภายในระแนงตัวแบบ
2. เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระบางคู่เริ่มเข้าสู่ระดับสูง (มีค่าตั้งแต่ 0.55 ขึ้นไป) การตัดตัวแปรอิสระบางตัวออกจากตัวแบบความถดถอย อาจจะมีความเหมาะสมกว่าการเลือกตัวแบบเต็มรูป ซึ่งการใช้เกณฑ์ RSS และ MSPE ในการคัดเลือกตัวแบบ และการคัดเลือก

ตัวแบบด้วยวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน และวิธีเปรียบเทียบตัวแบบภายในระแนงตัวแบบ จะเริ่มมีความแตกต่างของผลการคัดเลือกตัวแบบ โดยการใช้เกณฑ์ MSPE จะมีโอกาสเลือกตัวแบบลดรูปมากกว่าการใช้เกณฑ์ RSS และการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบจะมีโอกาสเลือกตัวแบบลดรูปมากกว่าวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้พร้อมกัน

5.1.2 สรุปปัจจัยที่มีผลต่อค่า RSS และ MSPE

1. ปัจจัยที่มีผลต่อค่า RSS

1.1 ขนาดตัวอย่าง

เมื่อขนาดตัวอย่างเพิ่มมากขึ้น ค่า RSS มีแนวโน้มเพิ่มขึ้น เนื่องจากค่า RSS เป็นค่าที่ใช้วัดความแตกต่างระหว่างค่าของตัวแปรตามก่อนการประมาณและหลังการประมาณ ซึ่งเป็นการรวมความผิดพลาดระหว่างค่าประมาณกับค่าจริงของทุกๆหน่วยตัวอย่างที่พิจารณา ดังนั้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้น ความผิดพลาดดังกล่าวจึงมีโอกาสเพิ่มมากขึ้นด้วย

1.2 ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน

เมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเพิ่มมากขึ้น ค่า RSS มีแนวโน้มเพิ่มขึ้นด้วย เนื่องจากข้อมูลมีการกระจายมากขึ้นหรือก็คือข้อมูลมีความแตกต่างกันมากขึ้น ดังนั้นโอกาสที่ค่าประมาณจะคลาดเคลื่อนไปจากค่าจริงมากก็จะสูงขึ้น

1.3 ความสัมพันธ์ระหว่างตัวแปรอิสระ

ความสัมพันธ์ของตัวแปรอิสระไม่มีผลต่อค่า RSS เนื่องจากค่าประมาณตัวแปรตามไม่ขึ้นกับระดับความสัมพันธ์ของตัวแปรอิสระ ดังนั้นระดับความสัมพันธ์ระหว่างตัวแปรอิสระจึงไม่มีผลต่อความแตกต่างระหว่างค่าจริงกับค่าประมาณของตัวแปรตาม นั่นคือความสัมพันธ์ของตัวแปรอิสระไม่มีผลต่อค่า RSS

1.4 จำนวนตัวแปรอิสระ

เมื่อตัวแปรอิสระในตัวแบบมีจำนวนลดลง ค่า RSS มีแนวโน้มเพิ่มขึ้น เนื่องจากเมื่อตัวแปรอิสระลดลง ก็จะทำให้ผลรวมกำลังสองของความถดถอยลดลงด้วย นั่นคือทำให้ความแตกต่างระหว่างค่าของตัวแปรตามก่อนการประมาณและหลังการประมาณเพิ่มมากขึ้น

2. ปัจจัยที่มีผลต่อค่า MSPE

2.1 ขนาดตัวอย่าง

เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่า MSPE มีแนวโน้มคงที่ เนื่องจากในสูตรการคำนวณค่า MSPE จะมีส่วนของค่า RSS ที่ถูกการปรับค่าให้เล็กลงด้วยเทอมที่ขึ้นกับขนาดตัวอย่างเสมอ ซึ่งเป็นผลให้ค่า MSPE มีค่าน้อยกว่าค่า RSS ด้วย

2.2 ส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน

เมื่อส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนเพิ่มมากขึ้น ค่า MSPE มีแนวโน้มเพิ่มขึ้นด้วย เนื่องจากข้อมูลมีการกระจายมากขึ้นหรือก็คือข้อมูลมีความแตกต่างกันมากขึ้น ดังนั้นโอกาสที่ค่าประมาณจะคลาดเคลื่อนไปจากค่าจริงมากก็จะสูงขึ้น

2.3 ความสัมพันธ์ระหว่างตัวแปรอิสระ

ในทำนองเดียวกันกับค่า RSS ความสัมพันธ์ระหว่างตัวแปรอิสระก็ไม่มีผลต่อค่า MSPE

2.4 จำนวนตัวแปรอิสระ

เมื่อจำนวนตัวแปรอิสระเพิ่มมากขึ้น ถ้าใช้ขนาดตัวอย่างเท่าเดิม จะทำให้ค่า MSPE มีแนวโน้มเพิ่มขึ้น เนื่องจากในสูตรการคำนวณค่า MSPE นอกจากมีการปรับค่า RSS ลงด้วยเทอมที่ขึ้นกับขนาดตัวอย่างแล้ว ยังมีการปรับค่าเพิ่มขึ้นด้วยเทอมที่ขึ้นกับจำนวนพารามิเตอร์ด้วย ดังนั้นถ้าตัวแปรอิสระมีจำนวนใกล้เคียงกับขนาดตัวอย่าง จะทำให้ค่า MSPE มีแนวโน้มเข้าใกล้ค่า RSS มากขึ้น หรือก็คือค่า MSPE มีแนวโน้มเพิ่มขึ้นนั่นเอง

5.1.3 สรุปผลการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ (Lattice) โดยการใช้เกณฑ์ RSS และ MSPE

จากบทที่ 4 หัวข้อที่ 4.3 เราพบลักษณะของการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ โดยการใช้เกณฑ์ RSS และ MSPE 2 ลักษณะ คือ

1. ถ้าพิจารณาค่า RDRSS หรือ RDMSPE แล้วปรากฏว่าตัวแบบที่ให้ค่า RDRSS=0 หรือ RDMSPE=0 (ตัวแบบที่ให้ค่า RSS หรือ MSPE ต่ำที่สุด โดยวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้พร้อมกัน) มีค่าต่างจากตัวแบบอื่นๆมาก <u>ตัวแบบที่เลือกได้จากการเปรียบเทียบภายในระแนงตัวแบบอาจเป็นตัวแบบที่ให้ค่า RDRSS=0 หรือ RDMSPE=0 หรือ ตัวแบบอื่นๆที่มีค่า RDRSS หรือ RDMSPE น้อยเป็นอันดับที่สอง</u>

2. ถ้าตัวแบบลดรูปบางตัวแบบมีค่า RDRSS หรือ RDMSPE ใกล้เคียง 0 หรือเมื่อมีกลุ่มของตัวแบบลดรูปบางกลุ่มที่มีค่า RDRSS หรือ RDMSPE ใกล้เคียงกันและมีค่าไม่มากนัก <u>ตัวแบบที่เลือกได้จากการใช้วิธีเปรียบเทียบภายในระแนงตัวแบบอาจจะเป็นตัวแบบหนึ่งในกลุ่มตัวแบบลดรูปนั้น และไม่จำเป็นที่จะต้องเป็นตัวแบบที่มีค่า RDRSS หรือ RDMSPE น้อยเป็นอันดับที่สองก็ได้</u>

จากทั้งสองลักษณะเราสรุปได้ว่า การใช้ระแนงตัวแบบในการเปรียบเทียบตัวแบบ จะเปิดโอกาสให้ตัวแบบที่เลือกได้เป็นตัวแบบที่ไม่ได้มีค่า RSS หรือ MSPE ต่ำที่สุดในบรรดาตัวแบบที่เป็นไปได้ทั้งหมด หรือก็คือการใช้ระแนงตัวแบบจะช่วยพิจารณาความเหมาะสมของความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ เพื่อค้นหาเซตของตัวแปรอิสระที่ดีที่สุด โดยไม่พิจารณาเฉพาะในแง่ของการพยากรณ์เท่านั้น

นอกจากนี้การใช้เกณฑ์ MSPE จะช่วยเพิ่มประสิทธิภาพในการเปรียบเทียบตัวแบบที่มีจำนวนตัวแปรอิสระแตกต่างกันด้วย

5.2 ข้อเสนอแนะในการเลือกใช้เกณฑ์การคัดเลือกตัวแบบและวิธีการเปรียบเทียบตัวแบบ

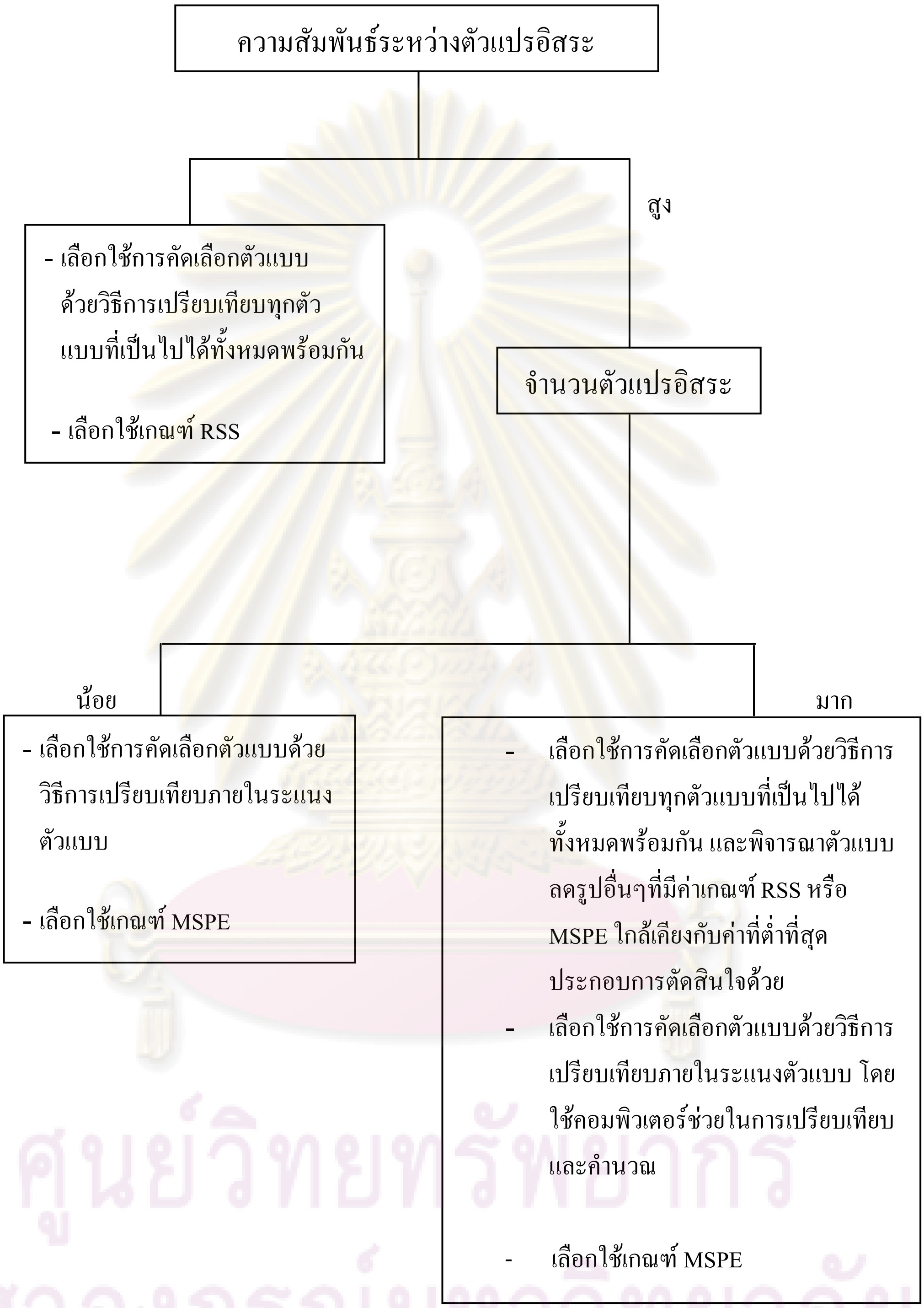
จากผลการวิจัยทั้งหมดสามารถสรุปเป็นแนวทางในการเลือกใช้เกณฑ์การคัดเลือกตัวแบบและวิธีการเปรียบเทียบตัวแบบ ให้เหมาะสมกับแต่ละสถานการณ์ได้ดังนี้

1. เมื่อตัวแบบความถดถอยที่พิจารณามีตัวแปรอิสระเริ่มต้นเป็นจำนวนมาก การคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบ จะมีความยุ่งยากและซับซ้อนมากและเสียเวลาในการเปรียบเทียบนาน เนื่องจากจำนวนลูกโซ่ที่เป็นไปได้ และความยาวของแต่ละลูกโซ่จะแปรผันตามจำนวนตัวแปรอิสระเริ่มต้น การคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกันจึงมีความเหมาะสมมากกว่า
2. เมื่อตัวแปรอิสระในตัวแบบความถดถอยมีความสัมพันธ์กันในระดับต่ำ การคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน จะสามารถทำได้ง่ายกว่าการเปรียบเทียบภายในระแนงตัวแบบ และให้ผลที่ไม่แตกต่างกัน
3. เมื่อตัวแปรอิสระในตัวแบบความถดถอยมีความสัมพันธ์กันในระดับต่ำ การเลือกใช้เกณฑ์ RSS จะมีความสะดวกกว่าการใช้เกณฑ์ MSPE เนื่องจากการใช้โปรแกรมสำเร็จรูปทางสถิติทั่วไป จะมีการแสดงค่า RSS อยู่แล้ว ในขณะที่ไม่มีการคำนวณค่า MSPE นอกจากนั้นเมื่อความสัมพันธ์ระหว่างตัวแปรอิสระมีค่าไม่สูงนัก การคัดเลือกตัวแบบความถดถอยที่ใช้เกณฑ์ RSS ก็ให้ผลไม่แตกต่างกับการใช้เกณฑ์ MSPE

4. เมื่อตัวแปรอิสระในตัวแบบความถดถอยมีความสัมพันธ์กันสูง และมีความจำเป็นในการวิเคราะห์หาตัวแบบที่ดีที่สุดในแง่ของการพยากรณ์ และในแง่ของความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ การเลือกใช้เทคนิคการคัดเลือกตัวแปรเพียงอย่างเดียวอาจไม่เพียงพอ เนื่องจากเป็นเพียงการลดปัญหาความสัมพันธ์ระหว่างตัวแปรอิสระเท่านั้น และการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมด ก็อาจให้ผลที่ไม่ดีเพียงพอ ดังนั้นการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบจึงเป็นทางเลือกหนึ่งที่เหมาะสม และเมื่อใช้เกณฑ์ MSPE แทนการใช้เกณฑ์ RSS ในการคัดเลือกตัวแบบร่วมด้วยแล้ว จะช่วยเพิ่มประสิทธิภาพในการคัดเลือกได้ดียิ่งขึ้น
5. สำหรับตัวแบบความถดถอยที่จำนวนพารามิเตอร์เป็นครึ่งหนึ่งของขนาดตัวอย่าง การใช้เกณฑ์ MSPE จะให้ผลเหมือนกับการใช้เกณฑ์ RSS

จากข้อเสนอแนะข้างต้น ในกรณีที่ตัวแบบความถดถอยมีจำนวนตัวแปรอิสระเริ่มต้นมาก และตัวแปรอิสระเหล่านั้นมีความสัมพันธ์กันสูง การคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบจะมีความเหมาะสมในแง่ทฤษฎี แต่ไม่เหมาะสมในทางปฏิบัติ เนื่องจากใช้เวลานานในการเปรียบเทียบ ดังนั้นการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกันจึงเหมาะสมกว่าในทางปฏิบัติ ดังนั้นเมื่อจำเป็นต้องใช้การเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน จึงควรจะพิจารณาตัวแบบลดรูปอื่นๆที่มีค่าเกณฑ์ RSS หรือ MSPE ใกล้เคียงกับค่าที่ต่ำที่สุดด้วย เนื่องจากการคัดเลือกตัวแบบด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบอาจได้ผลการคัดเลือกเป็นตัวแบบลดรูปที่มีค่าเกณฑ์ที่พิจารณาไม่ต่างจากค่าเกณฑ์ที่ต่ำที่สุดมากนัก โดยการพิจารณาว่าตัวแบบลดรูปใดจะมีความเหมาะสมกว่าตัวแบบเต็มรูปหรือไม่ สำหรับตัวแบบที่ไม่ติดกลุ่มกันจำเป็นต้องเลือกตัวแบบที่ค่าเกณฑ์ RSS หรือ MSPE ต่ำสุดเช่นเดิม แต่สำหรับตัวแบบที่ติดกลุ่มกันสามารถพิจารณาได้ด้วยตัวสถิติเอฟ

นอกจากการแก้ไขปัญหาด้วยวิธีดังกล่าว การใช้คอมพิวเตอร์อาจเป็นอีกหนึ่งทางเลือก อย่างไรก็ตามโปรแกรมสำเร็จรูปทางสถิติทั่วไปยังไม่รองรับวิธีการเปรียบเทียบภายในระแนงตัวแบบ ดังนั้นผู้วิจัยจึงใส่รายละเอียดของโปรแกรมที่ใช้ในงานวิจัยนี้ในภาคผนวก เพื่อให้ผู้สนใจได้ใช้เป็นแนวทางในการพัฒนาโปรแกรมให้ดียิ่งขึ้นต่อไป



รูปที่ 5.2 แผนผังแนวทางเลือกใช้เกณฑ์การคัดเลือกตัวแบบ และวิธีการเปรียบเทียบตัวแบบ

5.3 ข้อเสนอแนะด้านการวิจัย

เพื่อเป็นแนวทางให้ผู้สนใจได้ศึกษาเพิ่มเติม โดยทำการศึกษาในกรณีต่างๆดังนี้

ก) ศึกษาเพิ่มเติมเมื่อความคลาดเคลื่อนไม่ได้มีการแจกแจงแบบปกติ

ข) ศึกษาเพิ่มเติมสำหรับตัวแบบความถดถอยอื่นๆที่ไม่เป็นเชิงเส้น เช่น ตัวแบบความถดถอยพหุนาม เป็นต้น

ค) ศึกษาเพิ่มเติมในกรณีที่เมทริกซ์ X ไม่เป็นค่าคงที่ แต่เป็นตัวแปรสุ่ม

ง) ใช้วิธีการประมาณค่าสัมประสิทธิ์ความถดถอยหลายตัวแปรแบบอื่น สำหรับข้อมูลที่มีความสัมพันธ์ระหว่างตัวแปรอิสระอยู่ในระดับสูง เช่นการวิเคราะห์ความถดถอยแบบริดจ์ เป็นต้น

จ) ศึกษาเพิ่มเติมสำหรับเกณฑ์การคัดเลือกตัวแบบอื่นๆที่มีความซับซ้อนมากขึ้น เช่น Akaike's Information Criterion (AIC), Sawa's Bayesian Information Criterion (BIC)

รายการอ้างอิง

ภาษาไทย

ธีระพร วีระถาวร. **ตัวแบบเชิงเส้น ทฤษฎีและการประยุกต์**. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: วิทยพัฒน์, 2541.

บุญจิรา มากอ้น. **การเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบการถดถอยแบบไม่ติดกลุ่ม**. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย. 2545.

พจนา แววสวัสดิ์. **การเปรียบเทียบเกณฑ์การคัดเลือกตัวแบบความถดถอยพหุนามแบบติดกลุ่ม**. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย. 2543.

วลัยทิพย์ บุญญาติศัย. **การคัดเลือกสมการถดถอยเชิงเส้นที่ดีที่สุดภายใต้แลตทิซ**. วิทยานิพนธ์ปริญญามหาบัณฑิต, ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย. 2549.

ภาษาอังกฤษ

Jiri, A. On non-nested regression models. Comment.Math.Univ.Carolinae, 34, 1993: 335-340.

Toby, D. Model Selection Criteria. 2006. Available from E-mail: tdhock@ocf.berkeley.edu.

บรรณานุกรม

ภาษาไทย

ธีระพร วีระถาวร. **การอนุมานเชิงสถิติขั้นกลาง โครงสร้างและความหมาย**. พิมพ์ครั้งที่ 1. กรุงเทพมหานคร: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2536.

สุพล ดุรงค์วัฒนา. **ตัวแบบและการวิเคราะห์ความถดถอยสำหรับการวิจัยขั้นสูง**. กรุงเทพมหานคร: ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย, 2548.

ภาษาอังกฤษ

Bovas A. and Johannes L. Introduction to regression Modeling. CA: Thomson Brooks/Cole. 2006.

G.A.F. SEBER. Linear Regression Analysis. New York: John Wiley&Sons. 1997.

John, O.R. Applied Regression Analysis: A Research Tool. CA: Wadsworth. 1998.

Raymond H. Myers. Classical and Modern Regression with Applications. Massachuseetts: PWS-KENT. 1990

ภาคผนวก

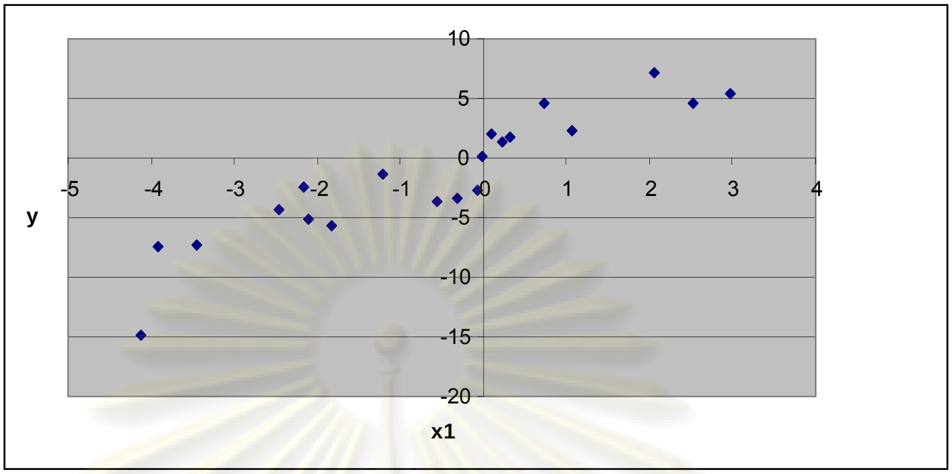
ภาคผนวก ก

ตัวอย่างการวิเคราะห์หาตัวแบบความถดถอยที่ดีที่สุด โดยวิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน และวิธีการเปรียบเทียบภายในระแนงตัวแบบ

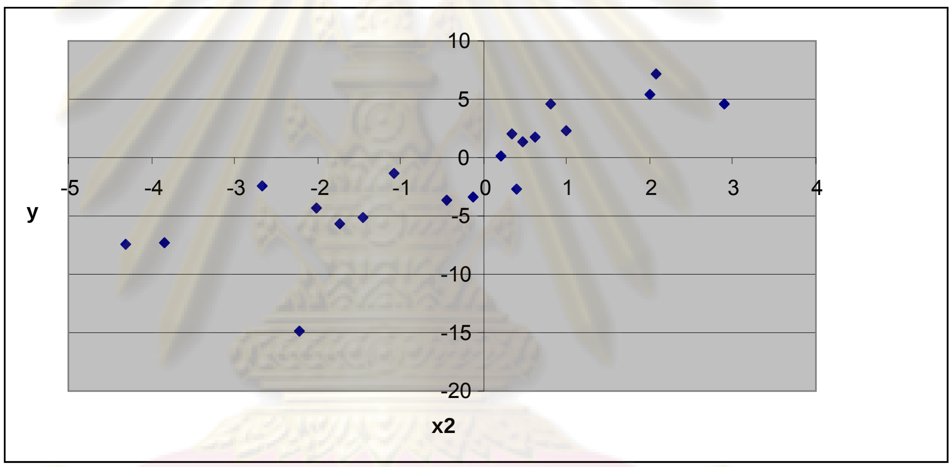
พิจารณาข้อมูลชุดหนึ่ง ซึ่งประกอบด้วยตัวแปรอิสระ 3 ตัว และตัวแปรตาม 1 ตัว ดังนี้

y	x_1	x_2	x_3
2.286	1.066	0.996	0.438
7.172	2.066	2.071	0.662
-5.131	-2.104	-1.448	-2.620
4.659	0.724	0.802	0.006
-2.656	-0.074	0.402	-1.302
-4.310	-2.455	-2.015	-1.799
-3.605	-0.550	-0.440	-0.347
-3.442	-0.321	-0.130	-1.106
-2.430	-2.167	-2.673	0.822
-7.428	-3.908	-4.303	0.064
-5.613	-1.819	-1.736	-0.796
0.130	-0.011	0.211	-0.484
-7.270	-3.447	-3.833	0.175
4.563	2.526	2.907	-0.078
-1.295	-1.201	-1.072	-0.519
-14.867	-4.128	-2.216	-6.544
1.395	0.222	0.467	-0.921
5.423	2.975	2.010	3.488
1.990	0.106	0.340	-0.514
1.718	0.324	0.620	-0.913

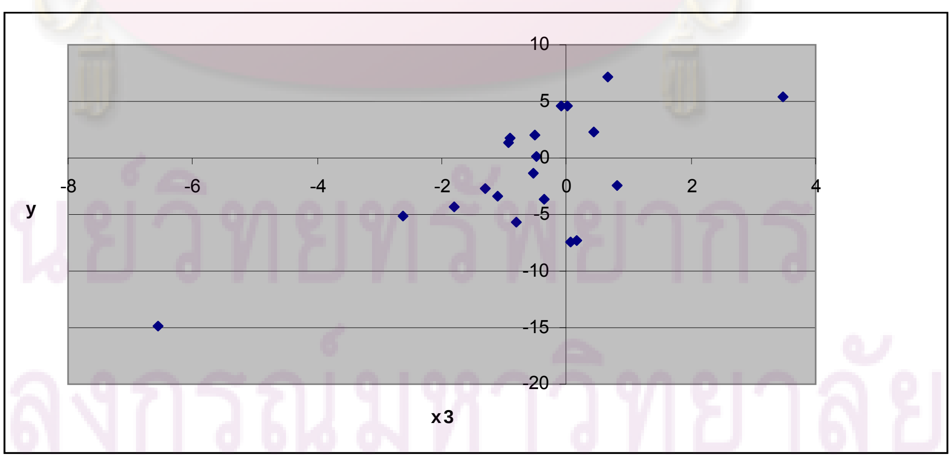
จากข้อมูลข้างต้นทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัว กับตัวแปรตาม โดยพิจารณาจากกราฟความสัมพันธ์ดังนี้



รูปที่ 1 กราฟความสัมพันธ์ระหว่าง y และ x_1



รูปที่ 2 กราฟความสัมพันธ์ระหว่าง y และ x_2



รูปที่ 3 กราฟความสัมพันธ์ระหว่าง y และ x_3

จากกราฟความสัมพันธ์ ตัวแปรอิสระทั้ง 3 ตัวแปร (x_1, x_2, และ x_3) น่าจะมีความสัมพันธ์กับตัวแปรตาม (y) ในลักษณะเชิงเส้น ดังนั้นถ้าต้องการสร้างตัวแบบความถดถอยเพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรอิสระ x_1, x_2, และ x_3 และตัวแปรตาม (y) จึงควรพิจารณาตัวแบบความถดถอยเชิงเส้น

พิจารณาการคัดเลือกตัวแบบความถดถอยด้วยวิธีการเปรียบเทียบภายในระแนงตัวแบบ โดยใช้เกณฑ์ RSS และ MSPE ในการเปรียบเทียบ ซึ่งมีขั้นตอนดังนี้

<u>ขั้นตอนที่ 1</u> พิจารณาระแนงตัวแบบ

พิจารณาตัวแบบความถดถอยเชิงเส้น $y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i , \qquad i = 1,2,\ldots,n$
จะได้ว่าตัวแบบความถดถอยที่เป็นไปได้ทั้งหมดมี 8 ตัวแบบ คือ

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 1 (M1) หรือเรียกว่า <u>*ตัวแบบเต็มรูป*</u> (Full Model)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$ ---ตัวแบบที่ 2 (M2)

$y_i = \beta_0' + \beta_1 x_{i1} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 3 (M3)

$y_i = \beta_0' + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 4 (M4)

$y_i = \beta_0' + \beta_1 x_{i1} + \varepsilon_i$ ---ตัวแบบที่ 5 (M5)

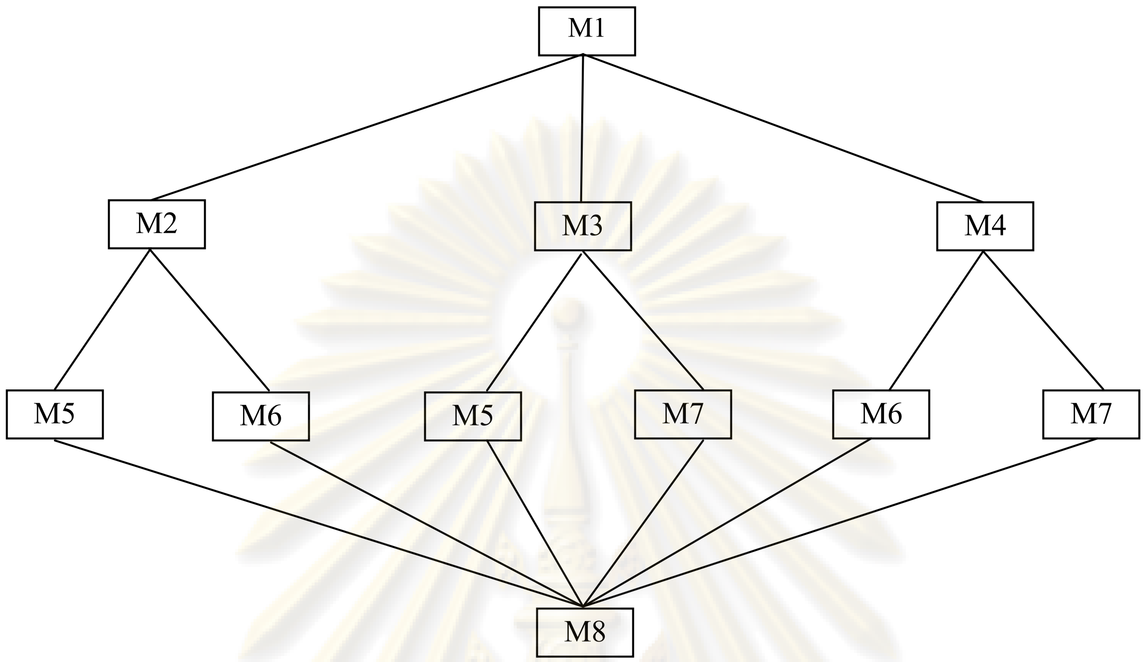
$y_i = \beta_0' + \beta_2 x_{i2} + \varepsilon_i$ ---ตัวแบบที่ 6 (M6)

$y_i = \beta_0' + \beta_3 x_{i3} + \varepsilon_i$ ---ตัวแบบที่ 7 (M7)

$y_i = \beta_0' + \varepsilon_i$ ---ตัวแบบที่ 8 (M8)

<u>*ตัวแบบลดรูป*</u> (Reduced Model)

โดยเขียนเป็นระแนงตัวแบบได้ดังนี้



จากระแนงตัวแบบข้างต้นมีจำนวนลูกโซ่ที่เป็นไปได้ 6 ลูกโซ่ คือ

ลูกโซ่ที่ 1 ประกอบด้วยตัวแบบ M1⟶M2⟶M5⟶M8

ลูกโซ่ที่ 2 ประกอบด้วยตัวแบบ M1⟶M2⟶M6⟶M8

ลูกโซ่ที่ 3 ประกอบด้วยตัวแบบ M1⟶M3⟶M5⟶M8

ลูกโซ่ที่ 4 ประกอบด้วยตัวแบบ M1⟶M3⟶M7⟶M8

ลูกโซ่ที่ 5 ประกอบด้วยตัวแบบ M1⟶M4⟶M6⟶M8

ลูกโซ่ที่ 6 ประกอบด้วยตัวแบบ M1⟶M4⟶M7⟶M8

<u>ขั้นตอนที่ 2</u> ปรับค่าข้อมูลตัวแปรอิสระแต่ละตัวด้วยค่าเฉลี่ย

จากตัวแบบความถดถอย $y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ ทำการปรับค่าตัวแปรอิสระแต่ละตัวด้วยการลบด้วยค่าเฉลี่ย ซึ่งเขียนเป็นตัวแบบความถดถอยใหม่ได้เป็น

$$y_i = \beta_0 + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_{.2}) + \beta_3(x_{i3} - \bar{x}_3) + \varepsilon_i$$

เมื่อ $\beta_0 = \beta_0' + \beta_1\bar{x}_1 + \beta_2\bar{x}_2 + \beta_3\bar{x}_3$ และ $\bar{x}_j = \dfrac{1}{20}\sum_{i=1}^{20} x_{ij},\ \ j = 1,2,3$

ซึ่งเมื่อนำมาเขียนให้อยู่ในรูปของเมทริกซ์ จะได้ว่า

$$\underset{\sim}{y} = [\underset{\sim}{1}, H]\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

$$\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_{20} \end{pmatrix}, \quad \underset{\sim}{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_3 \end{pmatrix}, \quad \underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{20} \end{pmatrix}$$

และ $$H = \begin{pmatrix} x_{1,1} - \bar{x}_1 & \dots & x_{1,3} - \bar{x}_3 \\ \cdots & \cdots & \cdots \\ x_{20,1} - \bar{x}_1 & \dots & x_{20,3} - \bar{x}_3 \end{pmatrix}$$

<u>ขั้นตอนที่ 3</u> ประมาณค่าสัมประสิทธิ์ความถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุด เพื่อคำนวณหาค่า RSS และ MSPE ของทุกตัวแบบในระแนงตัวแบบ

ทำการคำนวณค่า RSS และ MSPE ด้วยสูตรการคำนวณดังนี้

$$- RSS = (\underset{\sim}{y}'\underset{\sim}{y} - 20\bar{y}^2) - \underset{\sim}{y}' H(H'H)^{-1}H'\underset{\sim}{y}$$

$$- MSPE = RSS - 20s^2 + 2ps^2$$

เมื่อ $$s^2 = \frac{RSS(M1)}{20-3-1} = \frac{59.994}{16} = 3.750$$

และ p คือจำนวนพารามิเตอร์ในตัวแบบความถดถอย

จากข้อมูลข้างต้นได้ผลการคำนวณค่า RSS และ MSPE ดังนี้

MODEL	p	SSR	RSS	MSPE
M1	4	481.238	**59.994**	14.999
M2	3	477.926	63.306	10.811
M3	3	479.645	61.587	9.092
M4	3	480.651	60.580	**8.085**
M5	2	455.685	85.546	25.552
M6	2	369.204	172.028	112.033
M7	2	264.631	276.601	216.607
M8	1	0	541.232	473.738

โดยการเปรียบเทียบเบื้องต้น เมื่อพิจารณาค่า RSS และ MSPE ของทั้ง 8 ตัวแบบ พบว่าตัวแบบที่ให้ค่า RSS ต่ำที่สุด คือตัวแบบเต็มรูป ($y_i = \beta_0' + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$) ขณะที่ตัวแบบที่ให้ค่า MSPE ต่ำที่สุดคือตัวแบบลดรูปที่ไม่มีตัวแปรอิสระ x_1 ($y_i = \beta_0' + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$)

<u>ขั้นตอนที่ 4</u> ทำการเปรียบเทียบตัวแบบภายในแต่ละลูกโซ่ ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง

ในที่นี้จะแสดงการคำนวณแค่การเปรียบเทียบภายในลูกโซ่ที่ 1

ลูกโซ่ที่ 1 ประกอบไปด้วยตัวแบบ M1⟶M2⟶M5⟶M8

พิจารณาเปรียบเทียบ M1 กับ M2 โดยมีสมมติฐานทางสถิติคือ

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0$$

กำหนดระดับนัยสำคัญ $\alpha = 0.05$

คำนวณค่าสถิติเอฟ $F_c = \dfrac{SSR(M1) - SSR(M2)}{s^2} = \dfrac{481.238 - 477.926}{3.750} = 0.883$

จะได้ว่า $F_c = 0.883 < 4.494 = F_{0.05}(1,16)$

สรุปว่า ยอมรับ H_0 นั่นคือ M1 ไม่ทัดเทียม M2 อย่างมีนัยสำคัญ

ต่อไปพิจารณาว่า M5 จะมีความเหมาะสมกว่า M2 หรือไม่

พิจารณาเปรียบเทียบ M2 กับ M5 โดยมีสมมติฐานทางสถิติคือ

$$H_0' : \beta_2 = 0$$

$$H_1' : \beta_2 \neq 0$$

คำนวณค่าสถิติเอฟ $F_c = \dfrac{SSR(M2) - SSR(M5)}{s^2} = \dfrac{477.926 - 455.685}{3.750} = 5.931$

จะได้ว่า $F_c = 5.931 > 0.494 = F_{1-0.05}(1,16)$

สรุปว่า ปฏิเสธ H_0' หรือก็คือ ยอมรับ H_1' นั่นคือ M5 ไม่ทัดเทียม M2 อย่างมีนัยสำคัญ

เพราะฉะนั้นผลการคัดเลือกตัวแบบภายในลูกโซ่ที่ 1 ด้วยวิธีการกำจัดตัวแปรแบบถอยหลัง ตัวแบบที่มีความเหมาะสมที่สุดคือ M2

พิจารณาในทำนองเดียวกันกับลูกโซ่ที่เหลือ

<u>ขั้นตอนที่ 5</u> นำตัวแบบที่ได้รับเลือกของแต่ละลูกโซ่ในขั้นตอนที่ 4 มาเปรียบเทียบกันโดยใช้เกณฑ์ RSS หรือ MSPE

ผลลัพธ์ของการเปรียบเทียบในขั้นตอนที่ 4 นำมาสรุปรวมเป็นตารางได้ดังนี้

ลูกโซ่	**ตัวแบบที่เลือกได้จากแต่ละลูกโซ่**	**RSS**	**MSPE**
1	M2	63.306	10.811
2	M2	63.306	10.811
3	M3	61.587	9.092
4	M3	61.587	9.092
5	**M4**	60.580	8.085
6	**M4**	60.580	8.085

จากตารางข้างต้นพบว่า ทุกๆลูกโซ่ตัวแบบที่เลือกได้เป็นตัวแบบที่มีตัวแปรอิสระ 2 ตัวแปร โดยลูกโซ่ที่ 1 และ 2 ตัวแบบที่เลือกได้คือ M2 ลูกโซ่ที่ 3 และ 4 ตัวแบบที่เลือกได้คือ M3 และลูกโซ่ที่ 5 และ 6 ตัวแบบที่เลือกได้คือ M4

เมื่อพิจารณาเปรียบเทียบค่า RSS และ MSPE พบว่าตัวแบบที่ให้ค่า RSS และ MSPE ต่ำที่สุดคือตัวแบบลดรูปที่ไม่มีตัวแปรอิสระ x_1 ($y_i = \beta_0' + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$)

สรุปผลการคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบ การใช้เกณฑ์ RSS ในการเปรียบเทียบจะให้ผลที่แตกต่างจากวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน โดยที่วิธีเปรียบเทียบภายในระแนงตัวแบบ ตัวแบบที่เลือกได้คือ ตัวแบบลดรูปที่ไม่มีตัวแปรอิสระ x_1 ในขณะที่วิธีการเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน ตัวแบบที่เลือกได้คือตัวแบบเต็มรูป ส่วนการใช้เกณฑ์ MSPE การคัดเลือกตัวแบบด้วยวิธีเปรียบเทียบภายในระแนงตัวแบบ และวิธีเปรียบเทียบจากทุกตัวแบบที่เป็นไปได้ทั้งหมดพร้อมกัน ได้ผลการคัดเลือกตัวแบบเหมือนกันคือ ตัวแบบลดรูปที่ไม่มีตัวแปรอิสระ x_1

ภาคผนวก ข

รายละเอียดของขั้นตอนการเขียนโปรแกรมทั้งหมดที่ใช้ในงานวิจัย

ตัวแปรอิสระ 3 ตัวแปร

```
clc;

n = 20;
sigma = 3;
N=500;
r12 = 0.85;
r13 = 0.85;
r23 = 0.85;

rho = [1 r12 r13; r12 1 r23; r13 r23 1];

L = chol(rho);
for i = 1:N
z1 = normrnd(0,sigma,n,1);
z2 = normrnd(0,sigma,n,1);
z3 = normrnd(0,sigma,n,1);
x1 = L(1,1)*z1;
x2 = L(1,2)*z1+L(2,2)*z2;
x3 = L(1,3)*z1+L(2,3)*z2+L(3,3)*z3;
err = normrnd(0,sigma,n,1);
b=[1 1 1 1]';
X=[ones(n,1) x1 x2 x3];
y=(X*b)+err;

h1 = x1-mean(x1);
h2 = x2-mean(x2);
h3 = x3-mean(x3);
H_123 = [h1 h2 h3];
H_12  = [h1 h2];
H_13  = [h1 h3];
H_23  = [h2 h3];
H_1   = [h1];
H_2   = [h2];
H_3   = [h3];

SST = (y'*y-n*mean(y)^2);
rss_123 = SST-y'*H_123*inv(H_123'*H_123)*H_123'*y;
rss_12  = SST-y'*H_12*inv(H_12'*H_12)*H_12'*y;
rss_13  = SST-y'*H_13*inv(H_13'*H_13)*H_13'*y;
rss_23  = SST-y'*H_23*inv(H_23'*H_23)*H_23'*y;
rss_1   = SST-y'*H_1*inv(H_1'*H_1)*H_1'*y;
rss_2   = SST-y'*H_2*inv(H_2'*H_2)*H_2'*y;
rss_3   = SST-y'*H_3*inv(H_3'*H_3)*H_3'*y;
rss_0   = SST;

EstMSE = rss_123/(n-3-1);
mspe_123 = rss_123+(2*4-n)*EstMSE;
mspe_12  = rss_12+(2*3-n)*EstMSE;
mspe_13  = rss_13+(2*3-n)*EstMSE;
mspe_23  = rss_23+(2*3-n)*EstMSE;
mspe_1   = rss_1+(2*2-n)*EstMSE;
mspe_2   = rss_2+(2*2-n)*EstMSE;
mspe_3   = rss_3+(2*2-n)*EstMSE;
mspe_0   = rss_0+(2*1-n)*EstMSE;
```

```
SSR_123 = SST-rss_123;
SSR_12  = SST-rss_12;
SSR_13  = SST-rss_13;
SSR_23  = SST-rss_23;
SSR_1   = SST-rss_1;
SSR_2   = SST-rss_2;
SSR_3   = SST-rss_3;
SSR_0   = SST-rss_0;

ModelResult_RSS(i,1)= rss_123;
ModelResult_RSS(i,2)= rss_12;
ModelResult_RSS(i,3)= rss_13;
ModelResult_RSS(i,4)= rss_23;
ModelResult_RSS(i,5)= rss_1;
ModelResult_RSS(i,6)= rss_2;
ModelResult_RSS(i,7)= rss_3;
ModelResult_RSS(i,8)= rss_0;
ModelResult_MSPE(i,1)= mspe_123;
ModelResult_MSPE(i,2)= mspe_12;
ModelResult_MSPE(i,3)= mspe_13;
ModelResult_MSPE(i,4)= mspe_23;
ModelResult_MSPE(i,5)= mspe_1;
ModelResult_MSPE(i,6)= mspe_2;
ModelResult_MSPE(i,7)= mspe_3;
ModelResult_MSPE(i,8)= mspe_0;

alpha = 0.05;

if (SSR_123-SSR_12)/EstMSE > finv(1-alpha,1,n-3-1)
    C1258 = 1;
    RSS1258 = rss_123;
    MSPE1258 = mspe_123;
elseif (SSR_12-SSR_1)/EstMSE > finv(1-alpha,1,n-3-1)
    C1258 = 2;
    RSS1258 = rss_12;
    MSPE1258 = mspe_12;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-3-1)
    C1258 = 5;
    RSS1258 = rss_1;
    MSPE1258 = mspe_1;
else
    C1258 = 8;
    RSS1258 = rss_0;
    MSPE1258 = mspe_0;
end

Chain1258Result(i,1) = C1258;
Chain1258Result(i,2) = RSS1258;
Chain1258Result(i,3) = MSPE1258;

if (SSR_123-SSR_12)/EstMSE > finv(1-alpha,1,n-3-1)
    C1268 = 1;
    RSS1268 = rss_123;
    MSPE1268 = mspe_123;
elseif (SSR_12-SSR_2)/EstMSE > finv(1-alpha,1,n-3-1)
    C1268 = 2;
    RSS1268 = rss_12;
    MSPE1268 = mspe_12;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-3-1)
    C1268 = 6;
    RSS1268 = rss_2;
    MSPE1268 = mspe_2;
```

```
else
    C1268 = 8;
    RSS1268 = rss_0;
    MSPE1268 = mspe_0;
end

Chain1268Result(i,1) = C1268;
Chain1268Result(i,2) = RSS1268;
Chain1268Result(i,3) = MSPE1268;

if (SSR_123-SSR_13)/EstMSE > finv(1-alpha,1,n-3-1)
    C1358 = 1;
    RSS1358 = rss_123;
    MSPE1358 = mspe_123;
elseif (SSR_13-SSR_1)/EstMSE > finv(1-alpha,1,n-3-1)
    C1358 = 3;
    RSS1358 = rss_13;
    MSPE1358 = mspe_13;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-3-1)
    C1358 = 5;
    RSS1358 = rss_1;
    MSPE1358 = mspe_1;
else
    C1358 = 8;
    RSS1358 = rss_0;
    MSPE1358 = mspe_0;
end

Chain1358Result(i,1) = C1358;
Chain1358Result(i,2) = RSS1358;
Chain1358Result(i,3) = MSPE1358;

if (SSR_123-SSR_13)/EstMSE > finv(1-alpha,1,n-3-1)
    C1378 = 1;
    RSS1378 = rss_123;
    MSPE1378 = mspe_123;
elseif (SSR_13-SSR_3)/EstMSE > finv(1-alpha,1,n-3-1)
    C1378 = 3;
    RSS1378 = rss_13;
    MSPE1378 = mspe_13;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-3-1)
    C1378 = 7;
    RSS1378 = rss_3;
    MSPE1378 = mspe_3;
else
    C1378 = 8;
    RSS1378 = rss_0;
    MSPE1378 = mspe_0;
end

Chain1378Result(i,1) = C1378;
Chain1378Result(i,2) = RSS1378;
Chain1378Result(i,3) = MSPE1378;

if (SSR_123-SSR_23)/EstMSE > finv(1-alpha,1,n-3-1)
    C1468 = 1;
    RSS1468 = rss_123;
    MSPE1468 = mspe_123;
elseif (SSR_23-SSR_2)/EstMSE > finv(1-alpha,1,n-3-1)
    C1468 = 4;
    RSS1468 = rss_23;
    MSPE1468 = mspe_23;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-3-1)
```

```
    C1468 = 6;
    RSS1468 = rss_2;
    MSPE1468 = mspe_2;
else
    C1468 = 8;
    RSS1468 = rss_0;
    MSPE1468 = mspe_0;
end

Chain1468Result(i,1) = C1468;
Chain1468Result(i,2) = RSS1468;
Chain1468Result(i,3) = MSPE1468;

if (SSR_123-SSR_23)/EstMSE > finv(1-alpha,1,n-3-1)
    C1478 = 1;
    RSS1478 = rss_123;
    MSPE1478 = mspe_123;
elseif (SSR_23-SSR_3)/EstMSE > finv(1-alpha,1,n-3-1)
    C1478 = 4;
    RSS1478 = rss_23;
    MSPE1478 = mspe_23;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-3-1)
    C1478 = 7;
    RSS1478 = rss_3;
    MSPE1478 = mspe_3;
else
    C1478 = 8;
    RSS1478 = rss_0;
    MSPE1478 = mspe_0;
end

Chain1478Result(i,1) = C1478;
Chain1478Result(i,2) = RSS1478;
Chain1478Result(i,3) = MSPE1478;

end

u1 = mode(Chain1258Result);
u2 = mode(Chain1268Result);
u3 = mode(Chain1358Result);
u4 = mode(Chain1378Result);
u5 = mode(Chain1468Result);
u6 = mode(Chain1478Result);

fq1258 = 0;
fq1268 = 0;
fq1358 = 0;
fq1378 = 0;
fq1468 = 0;
fq1478 = 0;

SumRss1258 =0;
SumRss1268 =0;
SumRss1358 =0;
SumRss1378 =0;
SumRss1468 =0;
SumRss1478 =0;

SumMspe1258 =0;
SumMspe1268 =0;
SumMspe1358 =0;
SumMspe1378 =0;
SumMspe1468 =0;
```

```
SumMspe1478 =0;

for k=1:N
    if Chain1258Result(k,1) == u1(1,1)
        fq1258 = fq1258+1;
        SumRss1258 = SumRss1258+Chain1258Result(k,2);
        SumMspe1258 = SumMspe1258+Chain1258Result(k,3);
    else
    end

    if Chain1268Result(k,1) == u2(1,1)
        fq1268 = fq1268+1;
        SumRss1268 = SumRss1268+Chain1268Result(k,2);
        SumMspe1268 = SumMspe1268+Chain1268Result(k,3);
    else
    end

    if Chain1358Result(k,1) == u3(1,1)
        fq1358 = fq1358+1;
        SumRss1358 = SumRss1358+Chain1358Result(k,2);
        SumMspe1358 = SumMspe1358+Chain1358Result(k,3);
    else
    end

    if Chain1378Result(k,1) == u4(1,1)
        fq1378 = fq1378+1;
        SumRss1378 = SumRss1378+Chain1378Result(k,2);
        SumMspe1378 = SumMspe1378+Chain1378Result(k,3);
    else
    end

    if Chain1468Result(k,1) == u5(1,1)
        fq1468 = fq1468+1;
        SumRss1468 = SumRss1468+Chain1468Result(k,2);
        SumMspe1468 = SumMspe1468+Chain1468Result(k,3);
    else
    end

    if Chain1478Result(k,1) == u6(1,1)
        fq1478 = fq1478+1;
        SumRss1478 = SumRss1478+Chain1478Result(k,2);
        SumMspe1478 = SumMspe1478+Chain1478Result(k,3);
    else
    end
end

ChainResult(1,1) = u1(1,1);
ChainResult(2,1) = u2(1,1);
ChainResult(3,1) = u3(1,1);
ChainResult(4,1) = u4(1,1);
ChainResult(5,1) = u5(1,1);
ChainResult(6,1) = u6(1,1);
ChainResult(1,2) = fq1258;
ChainResult(2,2) = fq1268;
ChainResult(3,2) = fq1358;
ChainResult(4,2) = fq1378;
ChainResult(5,2) = fq1468;
ChainResult(6,2) = fq1478;
ChainResult(1,3) = SumRss1258/fq1258;
ChainResult(2,3) = SumRss1268/fq1268;
ChainResult(3,3) = SumRss1358/fq1358;
ChainResult(4,3) = SumRss1378/fq1378;
ChainResult(5,3) = SumRss1468/fq1468;
```

```
ChainResult(6,3) = SumRss1478/fq1478;
ChainResult(1,4) = SumMspe1258/fq1258;
ChainResult(2,4) = SumMspe1268/fq1268;
ChainResult(3,4) = SumMspe1358/fq1358;
ChainResult(4,4) = SumMspe1378/fq1378;
ChainResult(5,4) = SumMspe1468/fq1468;
ChainResult(6,4) = SumMspe1478/fq1478;

v = min(ChainResult);
MinRssChain = v(1,3);
MinMspeChain = v(1,4);

if v(1,3) == ChainResult(1,3)
    ChainSelModel_RSS = u1(1,1);
elseif v(1,3) == ChainResult(2,3)
    ChainSelModel_RSS = u2(1,1);
elseif v(1,3) == ChainResult(3,3)
    ChainSelModel_RSS = u3(1,1);
elseif v(1,3) == ChainResult(4,3)
    ChainSelModel_RSS = u4(1,1);
elseif v(1,3) == ChainResult(5,3)
    ChainSelModel_RSS = u5(1,1);
else
    ChainSelModel_RSS = u6(1,1);
end

if v(1,4) == ChainResult(1,4)
    ChainSelModel_MSPE = u1(1,1);
elseif v(1,4) == ChainResult(2,4)
    ChainSelModel_MSPE = u2(1,1);
elseif v(1,4) == ChainResult(3,4)
    ChainSelModel_MSPE = u3(1,1);
elseif v(1,4) == ChainResult(4,4)
    ChainSelModel_MSPE = u4(1,1);
elseif v(1,4) == ChainResult(5,4)
    ChainSelModel_MSPE = u5(1,1);
else
    ChainSelModel_MSPE = u6(1,1);
end

a = mean(ModelResult_RSS);
Mrss1 = a(1,1);
Mrss2 = a(1,2);
Mrss3 = a(1,3);
Mrss4 = a(1,4);
Mrss5 = a(1,5);
Mrss6 = a(1,6);
Mrss7 = a(1,7);
Mrss8 = a(1,8);
b = mean(ModelResult_MSPE);
Mmspe1 = b(1,1);
Mmspe2 = b(1,2);
Mmspe3 = b(1,3);
Mmspe4 = b(1,4);
Mmspe5 = b(1,5);
Mmspe6 = b(1,6);
Mmspe7 = b(1,7);
Mmspe8 = b(1,8);

if min(a) == Mrss1
    DirSelModel_RSS = 1;
    MinRSS = Mrss1;
elseif min(a) == Mrss2
```

```
        DirSelModel_RSS = 2;
        MinRSS = Mrss2;
    elseif min(a) == Mrss3
        DirSelModel_RSS = 3;
        MinRSS = Mrss3;
    elseif min(a) == Mrss4
        DirSelModel_RSS = 4;
        MinRSS = Mrss4;
    elseif min(a) == Mrss5
        DirSelModel_RSS = 5;
        MinRSS = Mrss5;
    elseif min(a) == Mrss6
        DirSelModel_RSS = 6;
        MinRSS = Mrss6;
    elseif min(a) == Mrss7
        DirSelModel_RSS = 7;
        MinRSS = Mrss7;
    else
        DirSelModel_RSS = 8;
        MinRSS = Mrss8;
    end

    if min(b) == Mmspe1
        DirSelModel_MSPE = 1;
        MinMSPE = Mmspe1;
    elseif min(b) == Mmspe2
        DirSelModel_MSPE = 2;
        MinMSPE = Mmspe2;
    elseif min(b) == Mmspe3
        DirSelModel_MSPE = 3;
        MinMSPE = Mmspe3;
    elseif min(b) == Mmspe4
        DirSelModel_MSPE = 4;
        MinMSPE = Mmspe4;
    elseif min(b) == Mmspe5
        DirSelModel_MSPE = 5;
        MinMSPE = Mmspe5;
    elseif min(b) == Mmspe6
        DirSelModel_MSPE = 6;
        MinMSPE = Mmspe6;
    elseif min(b) == Mmspe7
        DirSelModel_MSPE = 7;
        MinMSPE = Mmspe7;
    else
        DirSelModel_MSPE = 8;
        MinMSPE = Mmspe8;
    end

    FinalResult(1,1) = n;
    FinalResult(2,1) = sigma;
    FinalResult(3,1) = r12;
    FinalResult(4,1) = r13;
    FinalResult(5,1) = r23;
    FinalResult(1,2)= ChainSelModel_RSS;
    FinalResult(2,2)= MinRssChain;
    FinalResult(1,3)= ChainSelModel_MSPE;
    FinalResult(2,3)= MinMspeChain;
    FinalResult(1,4)= ((Mrss1-MinRSS)/MinRSS)*100;
    FinalResult(1,5)= ((Mrss2-MinRSS)/MinRSS)*100;
    FinalResult(1,6)= ((Mrss3-MinRSS)/MinRSS)*100;
    FinalResult(1,7)= ((Mrss4-MinRSS)/MinRSS)*100;
    FinalResult(1,8)= ((Mrss5-MinRSS)/MinRSS)*100;
    FinalResult(1,9)= ((Mrss6-MinRSS)/MinRSS)*100;
```

```
FinalResult(1,10)= ((Mrss7-MinRSS)/MinRSS)*100;
FinalResult(1,11)= ((Mrss8-MinRSS)/MinRSS)*100;
FinalResult(2,4)= ((Mmspe1-MinMSPE)/MinMSPE)*100;
FinalResult(2,5)= ((Mmspe2-MinMSPE)/MinMSPE)*100;
FinalResult(2,6)= ((Mmspe3-MinMSPE)/MinMSPE)*100;
FinalResult(2,7)= ((Mmspe4-MinMSPE)/MinMSPE)*100;
FinalResult(2,8)= ((Mmspe5-MinMSPE)/MinMSPE)*100;
FinalResult(2,9)= ((Mmspe6-MinMSPE)/MinMSPE)*100;
FinalResult(2,10)= ((Mmspe7-MinMSPE)/MinMSPE)*100;
FinalResult(2,11)= ((Mmspe8-MinMSPE)/MinMSPE)*100;
```

ตัวแปรอิสระ 4 ตัวแปร

```
clc;

n = 20;
sigma = 1;
N=500;
r12 = 0.15;
r13 = 0.25;
r14 = 0.35;
r23 = 0.15;
r24 = 0.25;
r34 = 0.35;

rho = [1 r12 r13 r14; r12 1 r23 r24; r13 r23 1 r34; r14 r24 r34 1];
L = chol(rho);
for i = 1:N
z1 = normrnd(0,sigma,n,1);
z2 = normrnd(0,sigma,n,1);
z3 = normrnd(0,sigma,n,1);
z4 = normrnd(0,sigma,n,1);

x1 = L(1,1)*z1;
x2 = L(1,2)*z1+L(2,2)*z2;
x3 = L(1,3)*z1+L(2,3)*z2+L(3,3)*z3;
x4 = L(1,4)*z1+L(2,4)*z2+L(3,4)*z3+L(4,4)*z4;
err = normrnd(0,sigma,n,1);

b=[1 1 1 1 1]';
X=[ones(n,1) x1 x2 x3 x4];
y=(X*b)+err;

h1 = x1-mean(x1);
h2 = x2-mean(x2);
h3 = x3-mean(x3);
h4 = x4-mean(x4);
H_1234 = [h1 h2 h3 h4];
H_123  = [h1 h2 h3];
H_124  = [h1 h2 h4];
H_134  = [h1 h3 h4];
H_234  = [h2 h3 h4];
H_12   = [h1 h2];
H_13   = [h1 h3];
H_14   = [h1 h4];
H_23   = [h2 h3];
H_24   = [h2 h4];
H_34   = [h3 h4];
H_1    = [h1];
H_2    = [h2];
```

```
H_3    = [h3];
H_4    = [h4];

SST = (y'*y-n*mean(y)^2);
rss_1234 = SST-y'*H_1234*inv(H_1234'*H_1234)*H_1234'*y;
rss_123  = SST-y'*H_123*inv(H_123'*H_123)*H_123'*y;
rss_124  = SST-y'*H_124*inv(H_124'*H_124)*H_124'*y;
rss_134  = SST-y'*H_134*inv(H_134'*H_134)*H_134'*y;
rss_234  = SST-y'*H_234*inv(H_234'*H_234)*H_234'*y;
rss_12   = SST-y'*H_12*inv(H_12'*H_12)*H_12'*y;
rss_13   = SST-y'*H_13*inv(H_13'*H_13)*H_13'*y;
rss_14   = SST-y'*H_14*inv(H_14'*H_14)*H_14'*y;
rss_23   = SST-y'*H_23*inv(H_23'*H_23)*H_23'*y;
rss_24   = SST-y'*H_24*inv(H_24'*H_24)*H_24'*y;
rss_34   = SST-y'*H_34*inv(H_34'*H_34)*H_34'*y;
rss_1    = SST-y'*H_1*inv(H_1'*H_1)*H_1'*y;
rss_2    = SST-y'*H_2*inv(H_2'*H_2)*H_2'*y;
rss_3    = SST-y'*H_3*inv(H_3'*H_3)*H_3'*y;
rss_4    = SST-y'*H_4*inv(H_4'*H_4)*H_4'*y;
rss_0    = SST;

EstMSE = rss_1234/(n-4-1);
mspe_1234 = rss_1234+(2*5-n)*EstMSE;
mspe_123  = rss_123+(2*4-n)*EstMSE;
mspe_124  = rss_124+(2*4-n)*EstMSE;
mspe_134  = rss_134+(2*4-n)*EstMSE;
mspe_234  = rss_234+(2*4-n)*EstMSE;
mspe_12  = rss_12+(2*3-n)*EstMSE;
mspe_13  = rss_13+(2*3-n)*EstMSE;
mspe_14  = rss_14+(2*3-n)*EstMSE;
mspe_23  = rss_23+(2*3-n)*EstMSE;
mspe_24  = rss_24+(2*3-n)*EstMSE;
mspe_34  = rss_34+(2*3-n)*EstMSE;
mspe_1   = rss_1+(2*2-n)*EstMSE;
mspe_2   = rss_2+(2*2-n)*EstMSE;
mspe_3   = rss_3+(2*2-n)*EstMSE;
mspe_4   = rss_4+(2*2-n)*EstMSE;
mspe_0   = rss_0+(2*1-n)*EstMSE;

SSR_1234 = SST-rss_1234;
SSR_123  = SST-rss_123;
SSR_124  = SST-rss_124;
SSR_134  = SST-rss_134;
SSR_234  = SST-rss_234;
SSR_12   = SST-rss_12;
SSR_13   = SST-rss_13;
SSR_14   = SST-rss_14;
SSR_23   = SST-rss_23;
SSR_24   = SST-rss_24;
SSR_34   = SST-rss_34;
SSR_1    = SST-rss_1;
SSR_2    = SST-rss_2;
SSR_3    = SST-rss_3;
SSR_4    = SST-rss_4;
SSR_0    = SST-rss_0;

ModelResult_RSS(i,1)= rss_1234;
ModelResult_RSS(i,2)= rss_123;
ModelResult_RSS(i,3)= rss_124;
ModelResult_RSS(i,4)= rss_134;
ModelResult_RSS(i,5)= rss_234;
ModelResult_RSS(i,6)= rss_12;
ModelResult_RSS(i,7)= rss_13;
```

```
ModelResult_RSS(i,8)= rss_14;
ModelResult_RSS(i,9)= rss_23;
ModelResult_RSS(i,10)= rss_24;
ModelResult_RSS(i,11)= rss_34;
ModelResult_RSS(i,12)= rss_1;
ModelResult_RSS(i,13)= rss_2;
ModelResult_RSS(i,14)= rss_3;
ModelResult_RSS(i,15)= rss_4;
ModelResult_RSS(i,16)= rss_0;
ModelResult_MSPE(i,1)= mspe_1234;
ModelResult_MSPE(i,2)= mspe_123;
ModelResult_MSPE(i,3)= mspe_124;
ModelResult_MSPE(i,4)= mspe_134;
ModelResult_MSPE(i,5)= mspe_234;
ModelResult_MSPE(i,6)= mspe_12;
ModelResult_MSPE(i,7)= mspe_13;
ModelResult_MSPE(i,8)= mspe_14;
ModelResult_MSPE(i,9)= mspe_23;
ModelResult_MSPE(i,10)= mspe_24;
ModelResult_MSPE(i,11)= mspe_34;
ModelResult_MSPE(i,12)= mspe_1;
ModelResult_MSPE(i,13)= mspe_2;
ModelResult_MSPE(i,14)= mspe_3;
ModelResult_MSPE(i,15)= mspe_4;
ModelResult_MSPE(i,16)= mspe_0;

alpha = 0.05;

if (SSR_1234-SSR_123)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_12_16 = 1;
    RSS1_2_6_12_16 = rss_1234;
    MSPE1_2_6_12_16 = mspe_1234;
elseif (SSR_123-SSR_12)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_12_16 = 2;
    RSS1_2_6_12_16 = rss_123;
    MSPE1_2_6_12_16 = mspe_123;
elseif (SSR_12-SSR_1)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_12_16 = 6;
    RSS1_2_6_12_16 = rss_12;
    MSPE1_2_6_12_16 = mspe_12;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_12_16 = 12;
    RSS1_2_6_12_16 = rss_1;
    MSPE1_2_6_12_16 = mspe_1;
else
    C1_2_6_12_16 = 16;
    RSS1_2_6_12_16 = rss_0;
    MSPE1_2_6_12_16 = mspe_0;
end

Chain_1_2_6_12_16_Result(i,1) = C1_2_6_12_16;
Chain_1_2_6_12_16_Result(i,2) = RSS1_2_6_12_16;
Chain_1_2_6_12_16_Result(i,3) = MSPE1_2_6_12_16;

if (SSR_1234-SSR_123)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_13_16 = 1;
    RSS1_2_6_13_16 = rss_1234;
    MSPE1_2_6_13_16 = mspe_1234;
elseif (SSR_123-SSR_12)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_13_16 = 2;
    RSS1_2_6_13_16 = rss_123;
    MSPE1_2_6_13_16 = mspe_123;
elseif (SSR_12-SSR_2)/EstMSE > finv(1-alpha,1,n-4-1)
```

```
    C1_2_6_13_16 = 6;
    RSS1_2_6_13_16 = rss_12;
    MSPE1_2_6_13_16 = mspe_12;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_6_13_16 = 13;
    RSS1_2_6_13_16 = rss_2;
    MSPE1_2_6_13_16 = mspe_2;
else
    C1_2_6_13_16 = 16;
    RSS1_2_6_13_16 = rss_0;
    MSPE1_2_6_13_16 = mspe_0;
end

Chain_1_2_6_13_16_Result(i,1) = C1_2_6_13_16;
Chain_1_2_6_13_16_Result(i,2) = RSS1_2_6_13_16;
Chain_1_2_6_13_16_Result(i,3) = MSPE1_2_6_13_16;


if (SSR_1234-SSR_123)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_12_16 = 1;
    RSS1_2_7_12_16 = rss_1234;
    MSPE1_2_7_12_16 = mspe_1234;
elseif (SSR_123-SSR_13)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_12_16 = 2;
    RSS1_2_7_12_16 = rss_123;
    MSPE1_2_7_12_16 = mspe_123;
elseif (SSR_13-SSR_1)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_12_16 = 7;
    RSS1_2_7_12_16 = rss_13;
    MSPE1_2_7_12_16 = mspe_13;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_12_16 = 12;
    RSS1_2_7_12_16 = rss_1;
    MSPE1_2_7_12_16 = mspe_1;
else
    C1_2_7_12_16 = 16;
    RSS1_2_7_12_16 = rss_0;
    MSPE1_2_7_12_16 = mspe_0;
end

Chain_1_2_7_12_16_Result(i,1) = C1_2_7_12_16;
Chain_1_2_7_12_16_Result(i,2) = RSS1_2_7_12_16;
Chain_1_2_7_12_16_Result(i,3) = MSPE1_2_7_12_16;

if (SSR_1234-SSR_123)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_14_16 = 1;
    RSS1_2_7_14_16 = rss_1234;
    MSPE1_2_7_14_16 = mspe_1234;
elseif (SSR_123-SSR_13)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_14_16 = 2;
    RSS1_2_7_14_16 = rss_123;
    MSPE1_2_7_14_16 = mspe_123;
elseif (SSR_13-SSR_3)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_14_16 = 7;
    RSS1_2_7_14_16 = rss_13;
    MSPE1_2_7_14_16 = mspe_13;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_7_14_16 = 14;
    RSS1_2_7_14_16 = rss_3;
    MSPE1_2_7_14_16 = mspe_3;
else
    C1_2_7_14_16 = 16;
    RSS1_2_7_14_16 = rss_0;
```

```
    MSPE1_2_7_14_16 = mspe_0;
end

Chain_1_2_7_14_16_Result(i,1) = C1_2_7_14_16;
Chain_1_2_7_14_16_Result(i,2) = RSS1_2_7_14_16;
Chain_1_2_7_14_16_Result(i,3) = MSPE1_2_7_14_16;

if (SSR_1234-SSR_123)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_13_16 = 1;
    RSS1_2_9_13_16 = rss_1234;
    MSPE1_2_9_13_16 = mspe_1234;
elseif (SSR_123-SSR_23)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_13_16 = 2;
    RSS1_2_9_13_16 = rss_123;
    MSPE1_2_9_13_16 = mspe_123;
elseif (SSR_23-SSR_2)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_13_16 = 9;
    RSS1_2_9_13_16 = rss_23;
    MSPE1_2_9_13_16 = mspe_23;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_13_16 = 13;
    RSS1_2_9_13_16 = rss_2;
    MSPE1_2_9_13_16 = mspe_2;
else
    C1_2_9_13_16 = 16;
    RSS1_2_9_13_16 = rss_0;
    MSPE1_2_9_13_16 = mspe_0;
end

Chain_1_2_9_13_16_Result(i,1) = C1_2_9_13_16;
Chain_1_2_9_13_16_Result(i,2) = RSS1_2_9_13_16;
Chain_1_2_9_13_16_Result(i,3) = MSPE1_2_9_13_16;

if (SSR_1234-SSR_123)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_14_16 = 1;
    RSS1_2_9_14_16 = rss_1234;
    MSPE1_2_9_14_16 = mspe_1234;
elseif (SSR_123-SSR_23)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_14_16 = 2;
    RSS1_2_9_14_16 = rss_123;
    MSPE1_2_9_14_16 = mspe_123;
elseif (SSR_23-SSR_3)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_14_16 = 9;
    RSS1_2_9_14_16 = rss_23;
    MSPE1_2_9_14_16 = mspe_23;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_2_9_14_16 = 14;
    RSS1_2_9_14_16 = rss_3;
    MSPE1_2_9_14_16 = mspe_3;
else
    C1_2_9_14_16 = 16;
    RSS1_2_9_14_16 = rss_0;
    MSPE1_2_9_14_16 = mspe_0;
end

Chain_1_2_9_14_16_Result(i,1) = C1_2_9_14_16;
Chain_1_2_9_14_16_Result(i,2) = RSS1_2_9_14_16;
Chain_1_2_9_14_16_Result(i,3) = MSPE1_2_9_14_16;

if (SSR_1234-SSR_124)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_12_16 = 1;
    RSS1_3_6_12_16 = rss_1234;
    MSPE1_3_6_12_16 = mspe_1234;
```

```
elseif (SSR_124-SSR_12)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_12_16 = 3;
    RSS1_3_6_12_16 = rss_124;
    MSPE1_3_6_12_16 = mspe_124;
elseif (SSR_12-SSR_1)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_12_16 = 6;
    RSS1_3_6_12_16 = rss_12;
    MSPE1_3_6_12_16 = mspe_12;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_12_16 = 12;
    RSS1_3_6_12_16 = rss_1;
    MSPE1_3_6_12_16 = mspe_1;
else
    C1_3_6_12_16 = 16;
    RSS1_3_6_12_16 = rss_0;
    MSPE1_3_6_12_16 = mspe_0;
end

Chain_1_3_6_12_16_Result(i,1) = C1_3_6_12_16;
Chain_1_3_6_12_16_Result(i,2) = RSS1_3_6_12_16;
Chain_1_3_6_12_16_Result(i,3) = MSPE1_3_6_12_16;

if (SSR_1234-SSR_124)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_13_16 = 1;
    RSS1_3_6_13_16 = rss_1234;
    MSPE1_3_6_13_16 = mspe_1234;
elseif (SSR_124-SSR_12)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_13_16 = 3;
    RSS1_3_6_13_16 = rss_124;
    MSPE1_3_6_13_16 = mspe_124;
elseif (SSR_12-SSR_2)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_13_16 = 6;
    RSS1_3_6_13_16 = rss_12;
    MSPE1_3_6_13_16 = mspe_12;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_6_13_16 = 13;
    RSS1_3_6_13_16 = rss_2;
    MSPE1_3_6_13_16 = mspe_2;
else
    C1_3_6_13_16 = 16;
    RSS1_3_6_13_16 = rss_0;
    MSPE1_3_6_13_16 = mspe_0;
end

Chain_1_3_6_13_16_Result(i,1) = C1_3_6_13_16;
Chain_1_3_6_13_16_Result(i,2) = RSS1_3_6_13_16;
Chain_1_3_6_13_16_Result(i,3) = MSPE1_3_6_13_16;

if (SSR_1234-SSR_124)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_12_16 = 1;
    RSS1_3_8_12_16 = rss_1234;
    MSPE1_3_8_12_16 = mspe_1234;
elseif (SSR_124-SSR_14)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_12_16 = 3;
    RSS1_3_8_12_16 = rss_124;
    MSPE1_3_8_12_16 = mspe_124;
elseif (SSR_14-SSR_1)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_12_16 = 8;
    RSS1_3_8_12_16 = rss_14;
    MSPE1_3_8_12_16 = mspe_14;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_12_16 = 12;
    RSS1_3_8_12_16 = rss_1;
```

```
    MSPE1_3_8_12_16 = mspe_1;
else
    C1_3_8_12_16 = 16;
    RSS1_3_8_12_16 = rss_0;
    MSPE1_3_8_12_16 = mspe_0;
end

Chain_1_3_8_12_16_Result(i,1) = C1_3_8_12_16;
Chain_1_3_8_12_16_Result(i,2) = RSS1_3_8_12_16;
Chain_1_3_8_12_16_Result(i,3) = MSPE1_3_8_12_16;

if (SSR_1234-SSR_124)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_15_16 = 1;
    RSS1_3_8_15_16 = rss_1234;
    MSPE1_3_8_15_16 = mspe_1234;
elseif (SSR_124-SSR_14)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_15_16 = 3;
    RSS1_3_8_15_16 = rss_124;
    MSPE1_3_8_15_16 = mspe_124;
elseif (SSR_14-SSR_4)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_15_16 = 8;
    RSS1_3_8_15_16 = rss_14;
    MSPE1_3_8_15_16 = mspe_14;
elseif (SSR_4-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_8_15_16 = 15;
    RSS1_3_8_15_16 = rss_4;
    MSPE1_3_8_15_16 = mspe_4;
else
    C1_3_8_15_16 = 16;
    RSS1_3_8_15_16 = rss_0;
    MSPE1_3_8_15_16 = mspe_0;
end

Chain_1_3_8_15_16_Result(i,1) = C1_3_8_15_16;
Chain_1_3_8_15_16_Result(i,2) = RSS1_3_8_15_16;
Chain_1_3_8_15_16_Result(i,3) = MSPE1_3_8_15_16;

if (SSR_1234-SSR_124)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_13_16 = 1;
    RSS1_3_10_13_16 = rss_1234;
    MSPE1_3_10_13_16 = mspe_1234;
elseif (SSR_124-SSR_24)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_13_16 = 3;
    RSS1_3_10_13_16 = rss_124;
    MSPE1_3_10_13_16 = mspe_124;
elseif (SSR_24-SSR_2)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_13_16 = 10;
    RSS1_3_10_13_16 = rss_24;
    MSPE1_3_10_13_16 = mspe_24;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_13_16 = 13;
    RSS1_3_10_13_16 = rss_2;
    MSPE1_3_10_13_16 = mspe_2;
else
    C1_3_10_13_16 = 16;
    RSS1_3_10_13_16 = rss_0;
    MSPE1_3_10_13_16 = mspe_0;
end

Chain_1_3_10_13_16_Result(i,1) = C1_3_10_13_16;
Chain_1_3_10_13_16_Result(i,2) = RSS1_3_10_13_16;
Chain_1_3_10_13_16_Result(i,3) = MSPE1_3_10_13_16;
```

```
if (SSR_1234-SSR_124)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_15_16 = 1;
    RSS1_3_10_15_16 = rss_1234;
    MSPE1_3_10_15_16 = mspe_1234;
elseif (SSR_124-SSR_24)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_15_16 = 3;
    RSS1_3_10_15_16 = rss_124;
    MSPE1_3_10_15_16 = mspe_124;
elseif (SSR_24-SSR_4)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_15_16 = 10;
    RSS1_3_10_15_16 = rss_24;
    MSPE1_3_10_15_16 = mspe_24;
elseif (SSR_4-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_3_10_15_16 = 15;
    RSS1_3_10_15_16 = rss_4;
    MSPE1_3_10_15_16 = mspe_4;
else
    C1_3_10_15_16 = 16;
    RSS1_3_10_15_16 = rss_0;
    MSPE1_3_10_15_16 = mspe_0;
end

Chain_1_3_10_15_16_Result(i,1) = C1_3_10_15_16;
Chain_1_3_10_15_16_Result(i,2) = RSS1_3_10_15_16;
Chain_1_3_10_15_16_Result(i,3) = MSPE1_3_10_15_16;

if (SSR_1234-SSR_134)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_12_16 = 1;
    RSS1_4_7_12_16 = rss_1234;
    MSPE1_4_7_12_16 = mspe_1234;
elseif (SSR_134-SSR_13)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_12_16 = 4;
    RSS1_4_7_12_16 = rss_134;
    MSPE1_4_7_12_16 = mspe_134;
elseif (SSR_13-SSR_1)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_12_16 = 7;
    RSS1_4_7_12_16 = rss_13;
    MSPE1_4_7_12_16 = mspe_13;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_12_16 = 12;
    RSS1_4_7_12_16 = rss_1;
    MSPE1_4_7_12_16 = mspe_1;
else
    C1_4_7_12_16 = 16;
    RSS1_4_7_12_16 = rss_0;
    MSPE1_4_7_12_16 = mspe_0;
end

Chain_1_4_7_12_16_Result(i,1) = C1_4_7_12_16;
Chain_1_4_7_12_16_Result(i,2) = RSS1_4_7_12_16;
Chain_1_4_7_12_16_Result(i,3) = MSPE1_4_7_12_16;

if (SSR_1234-SSR_134)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_14_16 = 1;
    RSS1_4_7_14_16 = rss_1234;
    MSPE1_4_7_14_16 = mspe_1234;
elseif (SSR_134-SSR_13)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_14_16 = 4;
    RSS1_4_7_14_16 = rss_134;
    MSPE1_4_7_14_16 = mspe_134;
elseif (SSR_13-SSR_3)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_14_16 = 7;
    RSS1_4_7_14_16 = rss_13;
```

```
    MSPE1_4_7_14_16 = mspe_13;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_7_14_16 = 14;
    RSS1_4_7_14_16 = rss_3;
    MSPE1_4_7_14_16 = mspe_3;
else
    C1_4_7_14_16 = 16;
    RSS1_4_7_14_16 = rss_0;
    MSPE1_4_7_14_16 = mspe_0;
end

Chain_1_4_7_14_16_Result(i,1) = C1_4_7_14_16;
Chain_1_4_7_14_16_Result(i,2) = RSS1_4_7_14_16;
Chain_1_4_7_14_16_Result(i,3) = MSPE1_4_7_14_16;

if (SSR_1234-SSR_134)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_12_16 = 1;
    RSS1_4_8_12_16 = rss_1234;
    MSPE1_4_8_12_16 = mspe_1234;
elseif (SSR_134-SSR_14)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_12_16 = 4;
    RSS1_4_8_12_16 = rss_134;
    MSPE1_4_8_12_16 = mspe_134;
elseif (SSR_14-SSR_1)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_12_16 = 8;
    RSS1_4_8_12_16 = rss_14;
    MSPE1_4_8_12_16 = mspe_14;
elseif (SSR_1-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_12_16 = 12;
    RSS1_4_8_12_16 = rss_1;
    MSPE1_4_8_12_16 = mspe_1;
else
    C1_4_8_12_16 = 16;
    RSS1_4_8_12_16 = rss_0;
    MSPE1_4_8_12_16 = mspe_0;
end

Chain_1_4_8_12_16_Result(i,1) = C1_4_8_12_16;
Chain_1_4_8_12_16_Result(i,2) = RSS1_4_8_12_16;
Chain_1_4_8_12_16_Result(i,3) = MSPE1_4_8_12_16;

if (SSR_1234-SSR_134)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_15_16 = 1;
    RSS1_4_8_15_16 = rss_1234;
    MSPE1_4_8_15_16 = mspe_1234;
elseif (SSR_134-SSR_14)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_15_16 = 4;
    RSS1_4_8_15_16 = rss_134;
    MSPE1_4_8_15_16 = mspe_134;
elseif (SSR_14-SSR_4)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_15_16 = 8;
    RSS1_4_8_15_16 = rss_14;
    MSPE1_4_8_15_16 = mspe_14;
elseif (SSR_4-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_8_15_16 = 15;
    RSS1_4_8_15_16 = rss_4;
    MSPE1_4_8_15_16 = mspe_4;
else
    C1_4_8_15_16 = 16;
    RSS1_4_8_15_16 = rss_0;
    MSPE1_4_8_15_16 = mspe_0;
end
```

```
Chain_1_4_8_15_16_Result(i,1) = C1_4_8_15_16;
Chain_1_4_8_15_16_Result(i,2) = RSS1_4_8_15_16;
Chain_1_4_8_15_16_Result(i,3) = MSPE1_4_8_15_16;

if (SSR_1234-SSR_134)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_14_16 = 1;
    RSS1_4_11_14_16 = rss_1234;
    MSPE1_4_11_14_16 = mspe_1234;
elseif (SSR_134-SSR_34)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_14_16 = 4;
    RSS1_4_11_14_16 = rss_134;
    MSPE1_4_11_14_16 = mspe_134;
elseif (SSR_34-SSR_3)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_14_16 = 11;
    RSS1_4_11_14_16 = rss_34;
    MSPE1_4_11_14_16 = mspe_34;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_14_16 = 14;
    RSS1_4_11_14_16 = rss_3;
    MSPE1_4_11_14_16 = mspe_3;
else
    C1_4_11_14_16 = 16;
    RSS1_4_11_14_16 = rss_0;
    MSPE1_4_11_14_16 = mspe_0;
end

Chain_1_4_11_14_16_Result(i,1) = C1_4_11_14_16;
Chain_1_4_11_14_16_Result(i,2) = RSS1_4_11_14_16;
Chain_1_4_11_14_16_Result(i,3) = MSPE1_4_11_14_16;

if (SSR_1234-SSR_134)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_15_16 = 1;
    RSS1_4_11_15_16 = rss_1234;
    MSPE1_4_11_15_16 = mspe_1234;
elseif (SSR_134-SSR_34)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_15_16 = 4;
    RSS1_4_11_15_16 = rss_134;
    MSPE1_4_11_15_16 = mspe_134;
elseif (SSR_34-SSR_4)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_15_16 = 11;
    RSS1_4_11_15_16 = rss_34;
    MSPE1_4_11_15_16 = mspe_34;
elseif (SSR_4-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_4_11_15_16 = 15;
    RSS1_4_11_15_16 = rss_4;
    MSPE1_4_11_15_16 = mspe_4;
else
    C1_4_11_15_16 = 16;
    RSS1_4_11_15_16 = rss_0;
    MSPE1_4_11_15_16 = mspe_0;
end

Chain_1_4_11_15_16_Result(i,1) = C1_4_11_15_16;
Chain_1_4_11_15_16_Result(i,2) = RSS1_4_11_15_16;
Chain_1_4_11_15_16_Result(i,3) = MSPE1_4_11_15_16;

if (SSR_1234-SSR_234)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_13_16 = 1;
    RSS1_5_9_13_16 = rss_1234;
    MSPE1_5_9_13_16 = mspe_1234;
elseif (SSR_234-SSR_23)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_13_16 = 5;
    RSS1_5_9_13_16 = rss_234;
```

```
    MSPE1_5_9_13_16 = mspe_234;
elseif (SSR_23-SSR_2)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_13_16 = 9;
    RSS1_5_9_13_16 = rss_23;
    MSPE1_5_9_13_16 = mspe_23;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_13_16 = 13;
    RSS1_5_9_13_16 = rss_2;
    MSPE1_5_9_13_16 = mspe_2;
else
    C1_5_9_13_16 = 16;
    RSS1_5_9_13_16 = rss_0;
    MSPE1_5_9_13_16 = mspe_0;
end

Chain_1_5_9_13_16_Result(i,1) = C1_5_9_13_16;
Chain_1_5_9_13_16_Result(i,2) = RSS1_5_9_13_16;
Chain_1_5_9_13_16_Result(i,3) = MSPE1_5_9_13_16;

if (SSR_1234-SSR_234)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_14_16 = 1;
    RSS1_5_9_14_16 = rss_1234;
    MSPE1_5_9_14_16 = mspe_1234;
elseif (SSR_234-SSR_23)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_14_16 = 5;
    RSS1_5_9_14_16 = rss_234;
    MSPE1_5_9_14_16 = mspe_234;
elseif (SSR_23-SSR_3)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_14_16 = 9;
    RSS1_5_9_14_16 = rss_23;
    MSPE1_5_9_14_16 = mspe_23;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_9_14_16 = 14;
    RSS1_5_9_14_16 = rss_3;
    MSPE1_5_9_14_16 = mspe_3;
else
    C1_5_9_14_16 = 16;
    RSS1_5_9_14_16 = rss_0;
    MSPE1_5_9_14_16 = mspe_0;
end

Chain_1_5_9_14_16_Result(i,1) = C1_5_9_14_16;
Chain_1_5_9_14_16_Result(i,2) = RSS1_5_9_14_16;
Chain_1_5_9_14_16_Result(i,3) = MSPE1_5_9_14_16;

if (SSR_1234-SSR_234)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_13_16 = 1;
    RSS1_5_10_13_16 = rss_1234;
    MSPE1_5_10_13_16 = mspe_1234;
elseif (SSR_234-SSR_24)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_13_16 = 5;
    RSS1_5_10_13_16 = rss_234;
    MSPE1_5_10_13_16 = mspe_234;
elseif (SSR_24-SSR_2)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_13_16 = 10;
    RSS1_5_10_13_16 = rss_24;
    MSPE1_5_10_13_16 = mspe_24;
elseif (SSR_2-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_13_16 = 13;
    RSS1_5_10_13_16 = rss_2;
    MSPE1_5_10_13_16 = mspe_2;
else
    C1_5_10_13_16 = 16;
```

```
    RSS1_5_10_13_16 = rss_0;
    MSPE1_5_10_13_16 = mspe_0;
end

Chain_1_5_10_13_16_Result(i,1) = C1_5_10_13_16;
Chain_1_5_10_13_16_Result(i,2) = RSS1_5_10_13_16;
Chain_1_5_10_13_16_Result(i,3) = MSPE1_5_10_13_16;

if (SSR_1234-SSR_234)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_15_16 = 1;
    RSS1_5_10_15_16 = rss_1234;
    MSPE1_5_10_15_16 = mspe_1234;
elseif (SSR_234-SSR_24)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_15_16 = 5;
    RSS1_5_10_15_16 = rss_234;
    MSPE1_5_10_15_16 = mspe_234;
elseif (SSR_24-SSR_4)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_15_16 = 10;
    RSS1_5_10_15_16 = rss_24;
    MSPE1_5_10_15_16 = mspe_24;
elseif (SSR_4-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_10_15_16 = 15;
    RSS1_5_10_15_16 = rss_4;
    MSPE1_5_10_15_16 = mspe_4;
else
    C1_5_10_15_16 = 16;
    RSS1_5_10_15_16 = rss_0;
    MSPE1_5_10_15_16 = mspe_0;
end

Chain_1_5_10_15_16_Result(i,1) = C1_5_10_15_16;
Chain_1_5_10_15_16_Result(i,2) = RSS1_5_10_15_16;
Chain_1_5_10_15_16_Result(i,3) = MSPE1_5_10_15_16;

if (SSR_1234-SSR_234)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_14_16 = 1;
    RSS1_5_11_14_16 = rss_1234;
    MSPE1_5_11_14_16 = mspe_1234;
elseif (SSR_234-SSR_34)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_14_16 = 5;
    RSS1_5_11_14_16 = rss_234;
    MSPE1_5_11_14_16 = mspe_234;
elseif (SSR_34-SSR_3)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_14_16 = 11;
    RSS1_5_11_14_16 = rss_34;
    MSPE1_5_11_14_16 = mspe_34;
elseif (SSR_3-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_14_16 = 14;
    RSS1_5_11_14_16 = rss_3;
    MSPE1_5_11_14_16 = mspe_3;
else
    C1_5_11_14_16 = 16;
    RSS1_5_11_14_16 = rss_0;
    MSPE1_5_11_14_16 = mspe_0;
end

Chain_1_5_11_14_16_Result(i,1) = C1_5_11_14_16;
Chain_1_5_11_14_16_Result(i,2) = RSS1_5_11_14_16;
Chain_1_5_11_14_16_Result(i,3) = MSPE1_5_11_14_16;

if (SSR_1234-SSR_234)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_15_16 = 1;
    RSS1_5_11_15_16 = rss_1234;
```

```
    MSPE1_5_11_15_16 = mspe_1234;
elseif (SSR_234-SSR_34)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_15_16 = 5;
    RSS1_5_11_15_16 = rss_234;
    MSPE1_5_11_15_16 = mspe_234;
elseif (SSR_34-SSR_4)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_15_16 = 11;
    RSS1_5_11_15_16 = rss_34;
    MSPE1_5_11_15_16 = mspe_34;
elseif (SSR_4-SSR_0)/EstMSE > finv(1-alpha,1,n-4-1)
    C1_5_11_15_16 = 15;
    RSS1_5_11_15_16 = rss_4;
    MSPE1_5_11_15_16 = mspe_4;
else
    C1_5_11_15_16 = 16;
    RSS1_5_11_15_16 = rss_0;
    MSPE1_5_11_15_16 = mspe_0;
end

Chain_1_5_11_15_16_Result(i,1) = C1_5_11_15_16;
Chain_1_5_11_15_16_Result(i,2) = RSS1_5_11_15_16;
Chain_1_5_11_15_16_Result(i,3) = MSPE1_5_11_15_16;

end

u1 = mode(Chain_1_2_6_12_16_Result);
u2 = mode(Chain_1_2_6_13_16_Result);
u3 = mode(Chain_1_2_7_12_16_Result);
u4 = mode(Chain_1_2_7_14_16_Result);
u5 = mode(Chain_1_2_9_13_16_Result);
u6 = mode(Chain_1_2_9_14_16_Result);
u7 = mode(Chain_1_3_6_12_16_Result);
u8 = mode(Chain_1_3_6_13_16_Result);
u9 = mode(Chain_1_3_8_12_16_Result);
u10 = mode(Chain_1_3_8_15_16_Result);
u11 = mode(Chain_1_3_10_13_16_Result);
u12 = mode(Chain_1_3_10_15_16_Result);
u13 = mode(Chain_1_4_7_12_16_Result);
u14 = mode(Chain_1_4_7_14_16_Result);
u15 = mode(Chain_1_4_8_12_16_Result);
u16 = mode(Chain_1_4_8_15_16_Result);
u17 = mode(Chain_1_4_11_14_16_Result);
u18 = mode(Chain_1_4_11_15_16_Result);
u19 = mode(Chain_1_5_9_13_16_Result);
u20 = mode(Chain_1_5_9_14_16_Result);
u21 = mode(Chain_1_5_10_13_16_Result);
u22 = mode(Chain_1_5_10_15_16_Result);
u23 = mode(Chain_1_5_11_14_16_Result);
u24 = mode(Chain_1_5_11_15_16_Result);

fq_1_2_6_12_16 = 0;
fq_1_2_6_13_16 = 0;
fq_1_2_7_12_16 = 0;
fq_1_2_7_14_16 = 0;
fq_1_2_9_13_16 = 0;
fq_1_2_9_14_16 = 0;
fq_1_3_6_12_16 = 0;
fq_1_3_6_13_16 = 0;
fq_1_3_8_12_16 = 0;
fq_1_3_8_15_16 = 0;
fq_1_3_10_13_16 = 0;
fq_1_3_10_15_16 = 0;
fq_1_4_7_12_16 = 0;
```

```
fq_1_4_7_14_16 = 0;
fq_1_4_8_12_16 = 0;
fq_1_4_8_15_16 = 0;
fq_1_4_11_14_16 = 0;
fq_1_4_11_15_16 = 0;
fq_1_5_9_13_16 = 0;
fq_1_5_9_14_16 = 0;
fq_1_5_10_13_16 = 0;
fq_1_5_10_15_16 = 0;
fq_1_5_11_14_16 = 0;
fq_1_5_11_15_16 = 0;
SumRss_1_2_6_12_16 = 0;
SumRss_1_2_6_13_16 = 0;
SumRss_1_2_7_12_16 = 0;
SumRss_1_2_7_14_16 = 0;
SumRss_1_2_9_13_16 = 0;
SumRss_1_2_9_14_16 = 0;
SumRss_1_3_6_12_16 = 0;
SumRss_1_3_6_13_16 = 0;
SumRss_1_3_8_12_16 = 0;
SumRss_1_3_8_15_16 = 0;
SumRss_1_3_10_13_16 = 0;
SumRss_1_3_10_15_16 = 0;
SumRss_1_4_7_12_16 = 0;
SumRss_1_4_7_14_16 = 0;
SumRss_1_4_8_12_16 = 0;
SumRss_1_4_8_15_16 = 0;
SumRss_1_4_11_14_16 = 0;
SumRss_1_4_11_15_16 = 0;
SumRss_1_5_9_13_16 = 0;
SumRss_1_5_9_14_16 = 0;
SumRss_1_5_10_13_16 = 0;
SumRss_1_5_10_15_16 = 0;
SumRss_1_5_11_14_16 = 0;
SumRss_1_5_11_15_16 = 0;
SumMspe_1_2_6_12_16 = 0;
SumMspe_1_2_6_13_16 = 0;
SumMspe_1_2_7_12_16 = 0;
SumMspe_1_2_7_14_16 = 0;
SumMspe_1_2_9_13_16 = 0;
SumMspe_1_2_9_14_16 = 0;
SumMspe_1_3_6_12_16 = 0;
SumMspe_1_3_6_13_16 = 0;
SumMspe_1_3_8_12_16 = 0;
SumMspe_1_3_8_15_16 = 0;
SumMspe_1_3_10_13_16 = 0;
SumMspe_1_3_10_15_16 = 0;
SumMspe_1_4_7_12_16 = 0;
SumMspe_1_4_7_14_16 = 0;
SumMspe_1_4_8_12_16 = 0;
SumMspe_1_4_8_15_16 = 0;
SumMspe_1_4_11_14_16 = 0;
SumMspe_1_4_11_15_16 = 0;
SumMspe_1_5_9_13_16 = 0;
SumMspe_1_5_9_14_16 = 0;
SumMspe_1_5_10_13_16 = 0;
SumMspe_1_5_10_15_16 = 0;
SumMspe_1_5_11_14_16 = 0;
SumMspe_1_5_11_15_16 = 0;

for k=1:N
    if Chain_1_2_6_12_16_Result(k,1) == u1(1,1)
        fq_1_2_6_12_16 = fq_1_2_6_12_16+1;
```

```
    SumRss_1_2_6_12_16 = SumRss_1_2_6_12_16+Chain_1_2_6_12_16_Result(k,2);
    SumMspe_1_2_6_12_16 = SumMspe_1_2_6_12_16+Chain_1_2_6_12_16_Result(k,3);
else
end

if Chain_1_2_6_13_16_Result(k,1) == u2(1,1)
    fq_1_2_6_13_16 = fq_1_2_6_13_16+1;
    SumRss_1_2_6_13_16 = SumRss_1_2_6_13_16+Chain_1_2_6_13_16_Result(k,2);
    SumMspe_1_2_6_13_16 = SumMspe_1_2_6_13_16+Chain_1_2_6_13_16_Result(k,3);
else
end

if Chain_1_2_7_12_16_Result(k,1) == u3(1,1)
    fq_1_2_7_12_16 = fq_1_2_7_12_16+1;
    SumRss_1_2_7_12_16 = SumRss_1_2_7_12_16+Chain_1_2_7_12_16_Result(k,2);
    SumMspe_1_2_7_12_16 = SumMspe_1_2_7_12_16+Chain_1_2_7_12_16_Result(k,3);
else
end

if Chain_1_2_7_14_16_Result(k,1) == u4(1,1)
    fq_1_2_7_14_16 = fq_1_2_7_14_16+1;
    SumRss_1_2_7_14_16 = SumRss_1_2_7_14_16+Chain_1_2_7_14_16_Result(k,2);
    SumMspe_1_2_7_14_16 = SumMspe_1_2_7_14_16+Chain_1_2_7_14_16_Result(k,3);
else
end

if Chain_1_2_9_13_16_Result(k,1) == u5(1,1)
    fq_1_2_9_13_16 = fq_1_2_9_13_16+1;
    SumRss_1_2_9_13_16 = SumRss_1_2_9_13_16+Chain_1_2_9_13_16_Result(k,2);
    SumMspe_1_2_9_13_16 = SumMspe_1_2_9_13_16+Chain_1_2_9_13_16_Result(k,3);
else
end

if Chain_1_2_9_14_16_Result(k,1) == u6(1,1)
    fq_1_2_9_14_16 = fq_1_2_9_14_16+1;
    SumRss_1_2_9_14_16 = SumRss_1_2_9_14_16+Chain_1_2_9_14_16_Result(k,2);
    SumMspe_1_2_9_14_16 = SumMspe_1_2_9_14_16+Chain_1_2_9_14_16_Result(k,3);
else
end

 if Chain_1_3_6_12_16_Result(k,1) == u7(1,1)
    fq_1_3_6_12_16 = fq_1_3_6_12_16+1;
    SumRss_1_3_6_12_16 = SumRss_1_3_6_12_16+Chain_1_3_6_12_16_Result(k,2);
    SumMspe_1_3_6_12_16 = SumMspe_1_3_6_12_16+Chain_1_3_6_12_16_Result(k,3);
else
end

if Chain_1_3_6_13_16_Result(k,1) == u8(1,1)
    fq_1_3_6_13_16 = fq_1_3_6_13_16+1;
    SumRss_1_3_6_13_16 = SumRss_1_3_6_13_16+Chain_1_3_6_13_16_Result(k,2);
    SumMspe_1_3_6_13_16 = SumMspe_1_3_6_13_16+Chain_1_3_6_13_16_Result(k,3);
else
end

if Chain_1_3_8_12_16_Result(k,1) == u9(1,1)
    fq_1_3_8_12_16 = fq_1_3_8_12_16+1;
    SumRss_1_3_8_12_16 = SumRss_1_3_8_12_16+Chain_1_3_8_12_16_Result(k,2);
    SumMspe_1_3_8_12_16 = SumMspe_1_3_8_12_16+Chain_1_3_8_12_16_Result(k,3);
else
end

if Chain_1_3_8_15_16_Result(k,1) == u10(1,1)
    fq_1_3_8_15_16 = fq_1_3_8_15_16+1;
```

```
            SumRss_1_3_8_15_16 = SumRss_1_3_8_15_16+Chain_1_3_8_15_16_Result(k,2);
            SumMspe_1_3_8_15_16 = SumMspe_1_3_8_15_16+Chain_1_3_8_15_16_Result(k,3);
        else
        end

        if Chain_1_3_10_13_16_Result(k,1) == u11(1,1)
            fq_1_3_10_13_16 = fq_1_3_10_13_16+1;
            SumRss_1_3_10_13_16 =
SumRss_1_3_10_13_16+Chain_1_3_10_13_16_Result(k,2);
            SumMspe_1_3_10_13_16 =
SumMspe_1_3_10_13_16+Chain_1_3_10_13_16_Result(k,3);
        else
        end

        if Chain_1_3_10_15_16_Result(k,1) == u12(1,1)
            fq_1_3_10_15_16 = fq_1_3_10_15_16+1;
            SumRss_1_3_10_15_16 =
SumRss_1_3_10_15_16+Chain_1_3_10_15_16_Result(k,2);
            SumMspe_1_3_10_15_16 =
SumMspe_1_3_10_15_16+Chain_1_3_10_15_16_Result(k,3);
        else
        end

         if Chain_1_4_7_12_16_Result(k,1) == u13(1,1)
            fq_1_4_7_12_16 = fq_1_4_7_12_16+1;
            SumRss_1_4_7_12_16 = SumRss_1_4_7_12_16+Chain_1_4_7_12_16_Result(k,2);
            SumMspe_1_4_7_12_16 = SumMspe_1_4_7_12_16+Chain_1_4_7_12_16_Result(k,3);
        else
        end

        if Chain_1_4_7_14_16_Result(k,1) == u14(1,1)
            fq_1_4_7_14_16 = fq_1_4_7_14_16+1;
            SumRss_1_4_7_14_16 = SumRss_1_4_7_14_16+Chain_1_4_7_14_16_Result(k,2);
            SumMspe_1_4_7_14_16 = SumMspe_1_4_7_14_16+Chain_1_4_7_14_16_Result(k,3);
        else
        end

        if Chain_1_4_8_12_16_Result(k,1) == u15(1,1)
            fq_1_4_8_12_16 = fq_1_4_8_12_16+1;
            SumRss_1_4_8_12_16 = SumRss_1_4_8_12_16+Chain_1_4_8_12_16_Result(k,2);
            SumMspe_1_4_8_12_16 = SumMspe_1_4_8_12_16+Chain_1_4_8_12_16_Result(k,3);
        else
        end

        if Chain_1_4_8_15_16_Result(k,1) == u16(1,1)
            fq_1_4_8_15_16 = fq_1_4_8_15_16+1;
            SumRss_1_4_8_15_16 = SumRss_1_4_8_15_16+Chain_1_4_8_15_16_Result(k,2);
            SumMspe_1_4_8_15_16 = SumMspe_1_4_8_15_16+Chain_1_4_8_15_16_Result(k,3);
        else
        end

        if Chain_1_4_11_14_16_Result(k,1) == u17(1,1)
            fq_1_4_11_14_16 = fq_1_4_11_14_16+1;
            SumRss_1_4_11_14_16 =
SumRss_1_4_11_14_16+Chain_1_4_11_14_16_Result(k,2);
            SumMspe_1_4_11_14_16 =
SumMspe_1_4_11_14_16+Chain_1_4_11_14_16_Result(k,3);
        else
        end

        if Chain_1_4_11_15_16_Result(k,1) == u18(1,1)
            fq_1_4_11_15_16 = fq_1_4_11_15_16+1;
```

```
        SumRss_1_4_11_15_16 =
SumRss_1_4_11_15_16+Chain_1_4_11_15_16_Result(k,2);
        SumMspe_1_4_11_15_16 =
SumMspe_1_4_11_15_16+Chain_1_4_11_15_16_Result(k,3);
    else
    end

     if Chain_1_5_9_13_16_Result(k,1) == u19(1,1)
        fq_1_5_9_13_16 = fq_1_5_9_13_16+1;
        SumRss_1_5_9_13_16 = SumRss_1_5_9_13_16+Chain_1_5_9_13_16_Result(k,2);
        SumMspe_1_5_9_13_16 = SumMspe_1_5_9_13_16+Chain_1_5_9_13_16_Result(k,3);
    else
    end

    if Chain_1_5_9_14_16_Result(k,1) == u20(1,1)
        fq_1_5_9_14_16 = fq_1_5_9_14_16+1;
        SumRss_1_5_9_14_16 = SumRss_1_5_9_14_16+Chain_1_5_9_14_16_Result(k,2);
        SumMspe_1_5_9_14_16 = SumMspe_1_5_9_14_16+Chain_1_5_9_14_16_Result(k,3);
    else
    end

    if Chain_1_5_10_13_16_Result(k,1) == u21(1,1)
        fq_1_5_10_13_16 = fq_1_5_10_13_16+1;
        SumRss_1_5_10_13_16 =
SumRss_1_5_10_13_16+Chain_1_5_10_13_16_Result(k,2);
        SumMspe_1_5_10_13_16 =
SumMspe_1_5_10_13_16+Chain_1_5_10_13_16_Result(k,3);
    else
    end

    if Chain_1_5_10_15_16_Result(k,1) == u22(1,1)
        fq_1_5_10_15_16 = fq_1_5_10_15_16+1;
        SumRss_1_5_10_15_16 =
SumRss_1_5_10_15_16+Chain_1_5_10_15_16_Result(k,2);
        SumMspe_1_5_10_15_16 =
SumMspe_1_5_10_15_16+Chain_1_5_10_15_16_Result(k,3);
    else
    end

    if Chain_1_5_11_14_16_Result(k,1) == u23(1,1)
        fq_1_5_11_14_16 = fq_1_5_11_14_16+1;
        SumRss_1_5_11_14_16 =
SumRss_1_5_11_14_16+Chain_1_5_11_14_16_Result(k,2);
        SumMspe_1_5_11_14_16 =
SumMspe_1_5_11_14_16+Chain_1_5_11_14_16_Result(k,3);
    else
    end

    if Chain_1_5_11_15_16_Result(k,1) == u24(1,1)
        fq_1_5_11_15_16 = fq_1_5_11_15_16+1;
        SumRss_1_5_11_15_16 =
SumRss_1_5_11_15_16+Chain_1_5_11_15_16_Result(k,2);
        SumMspe_1_5_11_15_16 =
SumMspe_1_5_11_15_16+Chain_1_5_11_15_16_Result(k,3);
    else
    end
end

ChainResult(1,1) = u1(1,1);
ChainResult(2,1) = u2(1,1);
ChainResult(3,1) = u3(1,1);
ChainResult(4,1) = u4(1,1);
ChainResult(5,1) = u5(1,1);
```

```
ChainResult(6,1) = u6(1,1);
ChainResult(7,1) = u7(1,1);
ChainResult(8,1) = u8(1,1);
ChainResult(9,1) = u9(1,1);
ChainResult(10,1) = u10(1,1);
ChainResult(11,1) = u11(1,1);
ChainResult(12,1) = u12(1,1);
ChainResult(13,1) = u13(1,1);
ChainResult(14,1) = u14(1,1);
ChainResult(15,1) = u15(1,1);
ChainResult(16,1) = u16(1,1);
ChainResult(17,1) = u17(1,1);
ChainResult(18,1) = u18(1,1);
ChainResult(19,1) = u19(1,1);
ChainResult(20,1) = u20(1,1);
ChainResult(21,1) = u21(1,1);
ChainResult(22,1) = u22(1,1);
ChainResult(23,1) = u23(1,1);
ChainResult(24,1) = u24(1,1);
ChainResult(1,2) = fq_1_2_6_12_16;
ChainResult(2,2) = fq_1_2_6_13_16;
ChainResult(3,2) = fq_1_2_7_12_16;
ChainResult(4,2) = fq_1_2_7_14_16;
ChainResult(5,2) = fq_1_2_9_13_16;
ChainResult(6,2) = fq_1_2_9_14_16;
ChainResult(7,2) = fq_1_3_6_12_16;
ChainResult(8,2) = fq_1_3_6_13_16;
ChainResult(9,2) = fq_1_3_8_12_16;
ChainResult(10,2) = fq_1_3_8_15_16;
ChainResult(11,2) = fq_1_3_10_13_16;
ChainResult(12,2) = fq_1_3_10_15_16;
ChainResult(13,2) = fq_1_4_7_12_16;
ChainResult(14,2) = fq_1_4_7_14_16;
ChainResult(15,2) = fq_1_4_8_12_16;
ChainResult(16,2) = fq_1_4_8_15_16;
ChainResult(17,2) = fq_1_4_11_14_16;
ChainResult(18,2) = fq_1_4_11_15_16;
ChainResult(19,2) = fq_1_5_9_13_16;
ChainResult(20,2) = fq_1_5_9_14_16;
ChainResult(21,2) = fq_1_5_10_13_16;
ChainResult(22,2) = fq_1_5_10_15_16;
ChainResult(23,2) = fq_1_5_11_14_16;
ChainResult(24,2) = fq_1_5_11_15_16;
ChainResult(1,3) = SumRss_1_2_6_12_16/fq_1_2_6_12_16;
ChainResult(2,3) = SumRss_1_2_6_13_16/fq_1_2_6_13_16;
ChainResult(3,3) = SumRss_1_2_7_12_16/fq_1_2_7_12_16;
ChainResult(4,3) = SumRss_1_2_7_14_16/fq_1_2_7_14_16;
ChainResult(5,3) = SumRss_1_2_9_13_16/fq_1_2_9_13_16;
ChainResult(6,3) = SumRss_1_2_9_14_16/fq_1_2_9_14_16;
ChainResult(7,3) = SumRss_1_3_6_12_16/fq_1_3_6_12_16;
ChainResult(8,3) = SumRss_1_3_6_13_16/fq_1_3_6_13_16;
ChainResult(9,3) = SumRss_1_3_8_12_16/fq_1_3_8_12_16;
ChainResult(10,3) = SumRss_1_3_8_15_16/fq_1_3_8_15_16;
ChainResult(11,3) = SumRss_1_3_10_13_16/fq_1_3_10_13_16;
ChainResult(12,3) = SumRss_1_3_10_15_16/fq_1_3_10_15_16;
ChainResult(13,3) = SumRss_1_4_7_12_16/fq_1_4_7_12_16;
ChainResult(14,3) = SumRss_1_4_7_14_16/fq_1_4_7_14_16;
ChainResult(15,3) = SumRss_1_4_8_12_16/fq_1_4_8_12_16;
ChainResult(16,3) = SumRss_1_4_8_15_16/fq_1_4_8_15_16;
ChainResult(17,3) = SumRss_1_4_11_14_16/fq_1_4_11_14_16;
ChainResult(18,3) = SumRss_1_4_11_15_16/fq_1_4_11_15_16;
ChainResult(19,3) = SumRss_1_5_9_13_16/fq_1_5_9_13_16;
ChainResult(20,3) = SumRss_1_5_9_14_16/fq_1_5_9_14_16;
```

```
ChainResult(21,3) = SumRss_1_5_10_13_16/fq_1_5_10_13_16;
ChainResult(22,3) = SumRss_1_5_10_15_16/fq_1_5_10_15_16;
ChainResult(23,3) = SumRss_1_5_11_14_16/fq_1_5_11_14_16;
ChainResult(24,3) = SumRss_1_5_11_15_16/fq_1_5_11_15_16;
ChainResult(1,4) = SumMspe_1_2_6_12_16/fq_1_2_6_12_16;
ChainResult(2,4) = SumMspe_1_2_6_13_16/fq_1_2_6_13_16;
ChainResult(3,4) = SumMspe_1_2_7_12_16/fq_1_2_7_12_16;
ChainResult(4,4) = SumMspe_1_2_7_14_16/fq_1_2_7_14_16;
ChainResult(5,4) = SumMspe_1_2_9_13_16/fq_1_2_9_13_16;
ChainResult(6,4) = SumMspe_1_2_9_14_16/fq_1_2_9_14_16;
ChainResult(7,4) = SumMspe_1_3_6_12_16/fq_1_3_6_12_16;
ChainResult(8,4) = SumMspe_1_3_6_13_16/fq_1_3_6_13_16;
ChainResult(9,4) = SumMspe_1_3_8_12_16/fq_1_3_8_12_16;
ChainResult(10,4) = SumMspe_1_3_8_15_16/fq_1_3_8_15_16;
ChainResult(11,4) = SumMspe_1_3_10_13_16/fq_1_3_10_13_16;
ChainResult(12,4) = SumMspe_1_3_10_15_16/fq_1_3_10_15_16;
ChainResult(13,4) = SumMspe_1_4_7_12_16/fq_1_4_7_12_16;
ChainResult(14,4) = SumMspe_1_4_7_14_16/fq_1_4_7_14_16;
ChainResult(15,4) = SumMspe_1_4_8_12_16/fq_1_4_8_12_16;
ChainResult(16,4) = SumMspe_1_4_8_15_16/fq_1_4_8_15_16;
ChainResult(17,4) = SumMspe_1_4_11_14_16/fq_1_4_11_14_16;
ChainResult(18,4) = SumMspe_1_4_11_15_16/fq_1_4_11_15_16;
ChainResult(19,4) = SumMspe_1_5_9_13_16/fq_1_5_9_13_16;
ChainResult(20,4) = SumMspe_1_5_9_14_16/fq_1_5_9_14_16;
ChainResult(21,4) = SumMspe_1_5_10_13_16/fq_1_5_10_13_16;
ChainResult(22,4) = SumMspe_1_5_10_15_16/fq_1_5_10_15_16;
ChainResult(23,4) = SumMspe_1_5_11_14_16/fq_1_5_11_14_16;
ChainResult(24,4) = SumMspe_1_5_11_15_16/fq_1_5_11_15_16;

v = min(ChainResult);
MinRssChain = v(1,3);
MinMspeChain = v(1,4);

if v(1,3) == ChainResult(1,3)
    ChainSelModel_RSS = u1(1,1);
elseif v(1,3) == ChainResult(2,3)
    ChainSelModel_RSS = u2(1,1);
elseif v(1,3) == ChainResult(3,3)
    ChainSelModel_RSS = u3(1,1);
elseif v(1,3) == ChainResult(4,3)
    ChainSelModel_RSS = u4(1,1);
elseif v(1,3) == ChainResult(5,3)
    ChainSelModel_RSS = u5(1,1);
elseif v(1,3) == ChainResult(6,3);
    ChainSelModel_RSS = u6(1,1);
elseif v(1,3) == ChainResult(7,3)
    ChainSelModel_RSS = u7(1,1);
elseif v(1,3) == ChainResult(8,3)
    ChainSelModel_RSS = u8(1,1);
elseif v(1,3) == ChainResult(9,3)
    ChainSelModel_RSS = u9(1,1);
elseif v(1,3) == ChainResult(10,3)
    ChainSelModel_RSS = u10(1,1);
elseif v(1,3) == ChainResult(11,3);
    ChainSelModel_RSS = u11(1,1);
elseif v(1,3) == ChainResult(12,3)
    ChainSelModel_RSS = u12(1,1);
elseif v(1,3) == ChainResult(13,3)
    ChainSelModel_RSS = u13(1,1);
elseif v(1,3) == ChainResult(14,3)
    ChainSelModel_RSS = u14(1,1);
elseif v(1,3) == ChainResult(15,3)
    ChainSelModel_RSS = u15(1,1);
```

```
elseif v(1,3) == ChainResult(16,3);
    ChainSelModel_RSS = u16(1,1);
elseif v(1,3) == ChainResult(17,3)
    ChainSelModel_RSS = u17(1,1);
elseif v(1,3) == ChainResult(18,3)
    ChainSelModel_RSS = u18(1,1);
elseif v(1,3) == ChainResult(19,3)
    ChainSelModel_RSS = u19(1,1);
elseif v(1,3) == ChainResult(20,3)
    ChainSelModel_RSS = u20(1,1);
elseif v(1,3) == ChainResult(21,3);
    ChainSelModel_RSS = u21(1,1);
elseif v(1,3) == ChainResult(22,3)
    ChainSelModel_RSS = u22(1,1);
elseif v(1,3) == ChainResult(23,3)
    ChainSelModel_RSS = u23(1,1);
else
    ChainSelModel_RSS = u24(1,1);
end

if v(1,4) == ChainResult(1,4)
    ChainSelModel_MSPE = u1(1,1);
elseif v(1,4) == ChainResult(2,4)
    ChainSelModel_MSPE = u2(1,1);
elseif v(1,4) == ChainResult(3,4)
    ChainSelModel_MSPE = u3(1,1);
elseif v(1,4) == ChainResult(4,4)
    ChainSelModel_MSPE = u4(1,1);
elseif v(1,4) == ChainResult(5,4)
    ChainSelModel_MSPE = u5(1,1);
elseif v(1,4) == ChainResult(6,4);
    ChainSelModel_MSPE = u6(1,1);
elseif v(1,4) == ChainResult(7,4)
    ChainSelModel_MSPE = u7(1,1);
elseif v(1,4) == ChainResult(8,4)
    ChainSelModel_MSPE = u8(1,1);
elseif v(1,4) == ChainResult(9,4)
    ChainSelModel_MSPE = u9(1,1);
elseif v(1,4) == ChainResult(10,4)
    ChainSelModel_MSPE = u10(1,1);
elseif v(1,4) == ChainResult(11,4);
    ChainSelModel_MSPE = u11(1,1);
elseif v(1,4) == ChainResult(12,4)
    ChainSelModel_MSPE = u12(1,1);
elseif v(1,4) == ChainResult(13,4)
    ChainSelModel_MSPE = u13(1,1);
elseif v(1,4) == ChainResult(14,4)
    ChainSelModel_MSPE = u14(1,1);
elseif v(1,4) == ChainResult(15,4)
    ChainSelModel_MSPE = u15(1,1);
elseif v(1,4) == ChainResult(16,4);
    ChainSelModel_MSPE = u16(1,1);
elseif v(1,4) == ChainResult(17,4)
    ChainSelModel_MSPE = u17(1,1);
elseif v(1,4) == ChainResult(18,4)
    ChainSelModel_MSPE = u18(1,1);
elseif v(1,4) == ChainResult(19,4)
    ChainSelModel_MSPE = u19(1,1);
elseif v(1,4) == ChainResult(20,4)
    ChainSelModel_MSPE = u20(1,1);
elseif v(1,4) == ChainResult(21,4);
    ChainSelModel_MSPE = u21(1,1);
elseif v(1,4) == ChainResult(22,4)
```

```
    ChainSelModel_MSPE = u22(1,1);
elseif v(1,4) == ChainResult(23,4)
    ChainSelModel_MSPE = u23(1,1);
else
    ChainSelModel_MSPE = u24(1,1);
end

a = mean(ModelResult_RSS);
Mrss1 = a(1,1);
Mrss2 = a(1,2);
Mrss3 = a(1,3);
Mrss4 = a(1,4);
Mrss5 = a(1,5);
Mrss6 = a(1,6);
Mrss7 = a(1,7);
Mrss8 = a(1,8);
Mrss9 = a(1,9);
Mrss10 = a(1,10);
Mrss11 = a(1,11);
Mrss12 = a(1,12);
Mrss13 = a(1,13);
Mrss14 = a(1,14);
Mrss15 = a(1,15);
Mrss16 = a(1,16);

b = mean(ModelResult_MSPE);
Mmspe1 = b(1,1);
Mmspe2 = b(1,2);
Mmspe3 = b(1,3);
Mmspe4 = b(1,4);
Mmspe5 = b(1,5);
Mmspe6 = b(1,6);
Mmspe7 = b(1,7);
Mmspe8 = b(1,8);
Mmspe9 = b(1,9);
Mmspe10 = b(1,10);
Mmspe11 = b(1,11);
Mmspe12 = b(1,12);
Mmspe13 = b(1,13);
Mmspe14 = b(1,14);
Mmspe15 = b(1,15);
Mmspe16 = b(1,16);

if min(a) == Mrss1
    DirSelModel_RSS = 1;
    MinRSS = Mrss1;
elseif min(a) == Mrss2
    DirSelModel_RSS = 2;
    MinRSS = Mrss2;
elseif min(a) == Mrss3
    DirSelModel_RSS = 3;
    MinRSS = Mrss3;
elseif min(a) == Mrss4
    DirSelModel_RSS = 4;
    MinRSS = Mrss4;
elseif min(a) == Mrss5
    DirSelModel_RSS = 5;
    MinRSS = Mrss5;
elseif min(a) == Mrss6
    DirSelModel_RSS = 6;
    MinRSS = Mrss6;
elseif min(a) == Mrss7
    DirSelModel_RSS = 7;
```

```
        MinRSS = Mrss7;
    elseif min(a) == Mrss8
        DirSelModel_RSS = 8;
        MinRSS = Mrss8;
    elseif min(a) == Mrss9
        DirSelModel_RSS = 9;
        MinRSS = Mrss9;
    elseif min(a) == Mrss10
        DirSelModel_RSS = 10;
        MinRSS = Mrss10;
    elseif min(a) == Mrss11
        DirSelModel_RSS = 11;
        MinRSS = Mrss11;
    elseif min(a) == Mrss12
        DirSelModel_RSS = 12;
        MinRSS = Mrss12;
    elseif min(a) == Mrss13
        DirSelModel_RSS = 13;
        MinRSS = Mrss13;
    elseif min(a) == Mrss14
        DirSelModel_RSS = 14;
        MinRSS = Mrss14;
    elseif min(a) == Mrss15
        DirSelModel_RSS = 15;
        MinRSS = Mrss15;
    else
        DirSelModel_RSS = 16;
        MinRSS = Mrss16;
    end

    if min(b) == Mmspe1
        DirSelModel_MSPE = 1;
        MinMSPE = Mmspe1;
    elseif min(b) == Mmspe2
        DirSelModel_MSPE = 2;
        MinMSPE = Mmspe2;
    elseif min(b) == Mmspe3
        DirSelModel_MSPE = 3;
        MinMSPE = Mmspe3;
    elseif min(b) == Mmspe4
        DirSelModel_MSPE = 4;
        MinMSPE = Mmspe4;
    elseif min(b) == Mmspe5
        DirSelModel_MSPE = 5;
        MinMSPE = Mmspe5;
    elseif min(b) == Mmspe6
        DirSelModel_MSPE = 6;
        MinMSPE = Mmspe6;
    elseif min(b) == Mmspe7
        DirSelModel_MSPE = 7;
        MinMSPE = Mmspe7;
    elseif min(b) == Mmspe8
        DirSelModel_MSPE = 8;
        MinMSPE = Mmspe8;
    elseif min(b) == Mmspe9
        DirSelModel_MSPE = 9;
        MinMSPE = Mmspe9;
    elseif min(b) == Mmspe10
        DirSelModel_MSPE = 10;
        MinMSPE = Mmspe10;
    elseif min(b) == Mmspe11
        DirSelModel_MSPE = 11;
        MinMSPE = Mmspe11;
```

```
elseif min(b) == Mmspe12
    DirSelModel_MSPE = 12;
    MinMSPE = Mmspe12;
elseif min(b) == Mmspe13
    DirSelModel_MSPE = 13;
    MinMSPE = Mmspe13;
elseif min(b) == Mmspe14
    DirSelModel_MSPE = 14;
    MinMSPE = Mmspe14;
elseif min(b) == Mmspe15
    DirSelModel_MSPE = 15;
    MinMSPE = Mmspe15;
else
    DirSelModel_MSPE = 16;
    MinMSPE = Mmspe16;
end

FinalResult(1,1) = n;
FinalResult(2,1) = sigma;
FinalResult(3,1) = r12;
FinalResult(4,1) = r13;
FinalResult(5,1) = r14;
FinalResult(6,1) = r23;
FinalResult(7,1) = r24;
FinalResult(8,1) = r34;
FinalResult(1,2)= ChainSelModel_RSS;
FinalResult(2,2)= MinRssChain;
FinalResult(1,3)= ChainSelModel_MSPE;
FinalResult(2,3)= MinMspeChain;
FinalResult(1,4) = ((Mrss1-MinRSS)/Mrss1)*100;
FinalResult(2,4) = ((Mrss2-MinRSS)/Mrss2)*100;
FinalResult(3,4)= ((Mrss3-MinRSS)/Mrss3)*100;
FinalResult(4,4)= ((Mrss4-MinRSS)/Mrss4)*100;
FinalResult(5,4)= ((Mrss5-MinRSS)/Mrss5)*100;
FinalResult(6,4)= ((Mrss6-MinRSS)/Mrss6)*100;
FinalResult(7,4)= ((Mrss7-MinRSS)/Mrss7)*100;
FinalResult(8,4)= ((Mrss8-MinRSS)/Mrss8)*100;
FinalResult(9,4) = ((Mrss9-MinRSS)/Mrss9)*100;
FinalResult(10,4) = ((Mrss10-MinRSS)/Mrss10)*100;
FinalResult(11,4)= ((Mrss11-MinRSS)/Mrss11)*100;
FinalResult(12,4)= ((Mrss12-MinRSS)/Mrss12)*100;
FinalResult(13,4)= ((Mrss13-MinRSS)/Mrss13)*100;
FinalResult(14,4)= ((Mrss14-MinRSS)/Mrss14)*100;
FinalResult(15,4)= ((Mrss15-MinRSS)/Mrss15)*100;
FinalResult(16,4)= ((Mrss16-MinRSS)/Mrss16)*100;
FinalResult(1,5)= ((Mmspe1-MinMSPE)/Mmspe1)*100;
FinalResult(2,5)= ((Mmspe2-MinMSPE)/Mmspe2)*100;
FinalResult(3,5)= ((Mmspe3-MinMSPE)/Mmspe3)*100;
FinalResult(4,5)= ((Mmspe4-MinMSPE)/Mmspe4)*100;
FinalResult(5,5)= ((Mmspe5-MinMSPE)/Mmspe5)*100;
FinalResult(6,5)= ((Mmspe6-MinMSPE)/Mmspe6)*100;
FinalResult(7,5)= ((Mmspe7-MinMSPE)/Mmspe7)*100;
FinalResult(8,5)= ((Mmspe8-MinMSPE)/Mmspe8)*100;
FinalResult(9,5)= ((Mmspe9-MinMSPE)/Mmspe9)*100;
FinalResult(10,5)= ((Mmspe10-MinMSPE)/Mmspe10)*100;
FinalResult(11,5)= ((Mmspe11-MinMSPE)/Mmspe11)*100;
FinalResult(12,5)= ((Mmspe12-MinMSPE)/Mmspe12)*100;
FinalResult(13,5)= ((Mmspe13-MinMSPE)/Mmspe13)*100;
FinalResult(14,5)= ((Mmspe14-MinMSPE)/Mmspe14)*100;
FinalResult(15,5)= ((Mmspe15-MinMSPE)/Mmspe15)*100;
FinalResult(16,5)= ((Mmspe16-MinMSPE)/Mmspe16)*100;
```

ประวัติผู้เขียนวิทยานิพนธ์

นายสุรสิทธิ์ ฤทธิ์สมิตชัย เกิดเมื่อวันที่ 9 ธันวาคม พ.ศ. 2526 สำเร็จการศึกษาปริญญาตรีจากคณะวิทยาศาสตร์ ภาควิชาคณิตศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2548 และได้เข้าศึกษาต่อในระดับปริญญาโท ที่ภาคสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปี พ.ศ. 2549