การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์ เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ

นางสาวจิตติมา ผสมญาติ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาสถิติศาสตรมหาบัณฑิต
สาขาวิชาสถิติ ภาควิชาสถิติ
คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2546
ISBN 974-17-4282-7
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF BAYESIAN SELECTION METHODS FOR BEST REGRESSION EQUATION WITH CONJUGATE NORMAL PRIOR DISTRIBUTION

Miss Jittima Phasomyard

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Statistics

Department of Statistics
Faculty of Commerce and Accountancy
Chulalongkorn University

Academic Year 2003
ISBN 974-17-4282-7

หัวข้อวิทยานิพนธ์	การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ
โดย	นางสาวจิตติมา ผสมญาติ
สาขาวิชา	สถิติ
อาจารย์ที่ปรึกษา	รองศาตราจารย์ ดร. ธีระพร วีระถาวร

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.....คณบดีคณะพาณิชยศาสตร์และการบัญชี
(ผู้ช่วยศาสตราจารย์ ดร. ดนุชา คุณพนิชกิจ)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาตราจารย์ ดร. กัลยา วานิชย์บัญชา)

..... อาจารย์ที่ปรึกษา
(รองศาตราจารย์ ดร. ธีระพร วีระถาวร)

..... กรรมการ
(รองศาตราจารย์ ดร. สุพล ดุรงค์วัฒนา)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จิตติมา ผสมญาติ : การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ (A COMPARISON OF BAYESIAN SELECTION METHODS FOR BEST REGRESSION EQUATION WITH CONJUGATE NORMAL PRIOR DISTRIBUTION.) อ. ที่ปรึกษา : รศ.ดร. ธีระพร วีระถาวร, 165 หน้า. ISBN 974-17-4282-7

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ โดยจะเปรียบเทียบวิธีการคัดเลือกสมการถดถอย 3 วิธี ได้แก่ วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (Model Uncertainty via Simultaneous Variable and Transformation Selection)(BMA_{SVT}) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (Optimal Predictive Model Selection: median probability model)(OPM) และวิธีการถดถอยแบบขั้นบันได (Stepwise Regression method)(SR) เกณฑ์ที่ใช้ในการตัดสินใจ คือ เกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Average of Mean Square Error)(AMSE) และเกณฑ์ที่ใช้ประกอบการตัดสินใจ คือ เกณฑ์ค่าอัตราส่วนผลต่างของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Ratio of Different Average Mean Square Error) (RDAMSE) การแจกแจงของค่าคลาดเคลื่อนที่ศึกษา คือ การแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐาน (σ) เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ ขนาดตัวอย่างที่ศึกษา (n) คือ 15 30 50 และ100 จำนวนตัวแปรอิสระที่ศึกษาคือ 3 5 8 10 12 และ 15 ค่าคงที่สำหรับวิธี BMA_{SVT} และวิธี OPM $(\sigma_\beta / \tau, c)$ ที่ศึกษามี 4 ระดับ คือ (1,5) (1,10) (10,100) และ (10,500) ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองแบบด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 500 รอบ ในแต่ละสถานการณ์ ซึ่งผลการวิจัยได้ข้อสรุปดังนี้

การเปรียบเทียบค่า AMSE ของทั้ง 3 วิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุกสถานการณ์ วิธี OPM จะให้ค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อยโดยเฉพาะอย่างยิ่งเมื่อค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าต่ำ ๆ ส่วนวิธี SR มีค่า AMSE แตกต่างจากวิธี BMA_{SVT} และวิธี OPM อย่างชัดเจนในทุกสถานการณ์ ปัจจัยที่มีผลต่อค่า AMSE ของทุกวิธี คือ ขนาดตัวอย่าง ค่าส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อนและจำนวนตัวแปรอิสระ โดยที่ค่า AMSE จะแปรผันตามค่าส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อนและจำนวนตัวแปรอิสระ แต่จะแปรผกผันกับขนาดตัวอย่าง

นอกจากนั้นค่า AMSE ของ 2 วิธีซึ่งเป็นวิธีการภายใต้แนวทางของเบส์ คือ วิธี BMA_{SVT} และวิธี OPM จะแปรผันตามค่าคงที่ $(\sigma_\beta / \tau, c)$ โดยที่ค่า AMSE ของวิธี SR ไม่เปลี่ยนแปลงเนื่องจากไม่ได้นำค่าคงที่ $(\sigma_\beta / \tau, c)$ มาพิจารณา

ภาควิชา สถิติ ลายมือชื่อนิสิต .....

สาขาวิชา สถิติ ลายมือชื่ออาจารย์ที่ปรึกษา .....

ปีการศึกษา 2546

4482197026 : MAJOR STATISTICS
KEY WORD : BAYESIAN MODEL AVERAGING / MARKOV CHAIN MONTE CARLO MODEL COMPOSITION / CHANGE-POINT TRANSFORMATION / OPTIMAL PREDICTIVE MODEL SELECTION / STEPWISE REGRESSION

JITTIMA PHASOMYARD : A COMPARISON OF BAYESIAN SELECTION METHODS FOR BEST REGRESSION EQUATION WITH CONJUGATE NORMAL PRIOR DISTRIBUTION.)
THESIS ADVISOR : ASSOC. PROF. THEERAPORN VERATHAWORN, 165 pp. ISBN 974-17-4282-7

The objective of this research is to compare the Bayesian selection methods for best regression equation with conjugate normal prior distribution. The three Bayesian selection methods for best regression equation in this comparison composed of Bayesian Model Averaging method using Markov Chain Monte Carlo model composition via Simultaneous Variable and Transformation Selection (BMA_{SVT}) , Optimal Predictive Model Selection (OPM) and Stepwise Regression method (SR). The average of mean square error (AMSE) and the ratio of different average mean square error (RDAMSE) were used as the criteria in this project. The details of the data were represented as follows. In this study, the distribution of random errors are normal distribution with mean equal to 0 and standard deviation equal to 0.25 0.50 and 2.50, respectively. The sizes of the samples are varied, which composed of 15 , 30 , 50 and 100 samples. The numbers of independent variables in regression model are 3 , 5 , 8 , 10 , 12 , and 15. The Bayesian hyperparameters $(\sigma_\beta / \tau, c)$ for BMA_{SVT} and OPM are (1,5) (1,10) (10,100) and (10,500), respectively. Using the Monte Carlo simulation technique with 500 repetitions for each case generated all data.

The analyzed results of data were demonstrated as follow. The comparisons of the AMSE from three methods ranging from minimum to maximum were BMA_{SVT}, OPM and SR for all cases. The OPM gave AMSE slightly higher than BMA_{SVT} especially when the Bayesian hyperparameters $(\sigma_\beta / \tau, c)$ had low values. The SR gave AMSE clearly different from BMA_{SVT} and OPM in all cases. The factors that affected AMSE of all methods are sample size, the standard deviation of random errors and the number of independent variables. The AMSE of all method were proportionate to the standard deviation of random errors and the number of independent variables although they were inversely proportionate to sample size.

Furthermore, the AMSE of two methods under Bayesian Approach, BMA_{SVT} and OPM were proportionate to the Bayesian hyperparameters $(\sigma_\beta / \tau, c)$ and whereas the AMSE of SR was constant because we did not consider the Bayesian hyperparameters $(\sigma_\beta / \tau, c)$.

Department	Statistics	Student's signature…….....
Field of study	Statistics	Advisor's signature…….....
Academic year	2003	

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี จากความอนุเคราะห์ของบุคคลหลายฝ่ายด้วยกัน ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.ธีระพร วีระถาวร อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่กรุณาให้คำแนะนำ ปรึกษา ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี

ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ ดร.กัลยา วานิชย์บัญชา และรองศาสตราจารย์ ดร.สุพล ดุรงค์วัฒนา ในฐานะประธานกรรมการและกรรมการสอบวิทยานิพนธ์ ที่กรุณาตรวจแก้วิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น ทั้งนี้ผู้วิจัยขอกราบขอบพระคุณคณาจารย์ประจำภาควิชาสถิติ ที่ให้โอกาสทางการศึกษา และประสิทธิประสาทความรู้ให้แก่ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

ท้ายนี้ผู้วิจัยใคร่ขอกราบขอบพระคุณ บิดา มารดา ซึ่งให้การสนับสนุน และเพื่อน ๆ ที่ให้กำลังใจแก่ผู้วิจัยเสมอมาจนกระทั่งสำเร็จการศึกษา และเนื่องจากการวิจัยครั้งนี้ได้รับทุนอุดหนุนการวิจัยของบัณฑิตวิทยาลัย ผู้วิจัยจึงขอกราบขอบพระคุณบัณฑิตวิทยาลัยมา ณ ที่นี้ด้วย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....ง

บทคัดย่อภาษาอังกฤษ.....จ

กิตติกรรมประกาศ.....ฉ

สารบัญ.....ช

สารบัญตาราง.....ญ

สารบัญภาพ.....ฐ

บทที่ 1 บทนำ.....1

1.1 ความเป็นมาและความสำคัญของปัญหา.....1

1.2 วัตถุประสงค์ของการวิจัย.....5

1.3 ข้อตกลงเบื้องต้น.....5

1.4 สมมติฐานของการวิจัย.....5

1.5 ขอบเขตของการวิจัย.....5

1.6 ขั้นตอนในการศึกษาวิจัย.....6

1.7 ประโยชน์ที่คาดว่าจะได้รับ.....7

1.8 เกณฑ์ที่ใช้ในการตัดสินใจ.....7

บทที่ 2 ทฤษฎีและสถิติที่เกี่ยวข้อง.....8

2.1 การวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ.....8

2.2 ความน่าจะเป็นที่มีเงื่อนไข ทฤษฎีบทของเบส์ และความเป็นอิสระ.....9

2.3 วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ.....9

2.3.1 วิธีการเฉลี่ยตัวแบบของเบส์.....10

2.3.1.1 ความน่าจะเป็นก่อนของตัวแบบ.....11

2.3.1.2 ความน่าจะเป็นภายหลังขอบของตัวแบบ.....12

2.3.1.3 ขั้นตอนการคัดเลือกตัวแปรของวิธี BMA_{MC3}.....14

2.3.2 ขั้นตอนของการคัดเลือกตัวแปรและการแปลงไปพร้อม ๆ กันของวิธี BMA_{SVT}.....16

2.4 วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด.....16

หน้า

2.4.1 การค้นหาตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ.....17

2.4.2 การคัดเลือกตัวแบบที่เหมาะสมที่สุด.....18

2.5 วิธีการถดถอยแบบขั้นบันได.....19

บทที่ 3 วิธีการดำเนินการวิจัย..... 21

3.1 การหาข้อสรุปโดยใช้การจำลองข้อมูลด้วยวิธีมอนติคาร์โล.....21

3.2 แผนการทดลอง.....22

3.3 ขั้นตอนในการศึกษาวิจัย.....23

บทที่ 4 ผลการวิจัย.....37

4.1 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ (σ_β / τ , c) มีค่าเท่ากับ (1,5).....41

4.2 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ (σ_β / τ , c) มีค่าเท่ากับ (1,10).....63

4.3 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ (σ_β / τ , c) มีค่าเท่ากับ (10, 100).....85

4.4 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวน

หน้า

ตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ (σ_β / τ, c) มีค่าเท่ากับ (10, 500).....107

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....129

5.1 สรุปผลการวิจัย.....130

5.1.1 ผลการเปรียบเทียบค่าเฉลี่ยของค่าคลาดเคลื่อนกำสังสองเฉลี่ย.....130

5.1.2 ปัจจัยที่มีผลต่อค่าเฉลี่ยของค่าคลาดเคลื่อนกำสังสองเฉลี่ยของแต่ละวิธี.....131

5.1.3 ผลสรุปการเลือกวิธีการสร้างตัวแบบการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ.....132

5.2 ข้อเสนอแนะ.....135

รายการอ้างอิง.....136

ภาคผนวก.....138

ประวัติผู้เขียนวิทยานิพนธ์.....165

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตารางที่ หน้า

4.1 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....42

4.2 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....45

4.3 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....49

4.4 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับดัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....52

4.5 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับดัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....55

4.6 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....58

4.7 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....64

4.8 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....67

ตารางที่ หน้า

4.9 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....70

4.10 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....73

4.11 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....76

4.12 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....79

4.13 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....86

4.14 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....89

4.15 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....92

4.16 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....95

ตารางที่ หน้า

4.17 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....98

4.18 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....101

4.19 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$.....108

4.20 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$.....111

4.21 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$.....114

4.22 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$.....117

4.23 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$.....120

4.24 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$.....123

สารบัญภาพ

รูปที่ หน้า

3.1 แสดงผังงานสำหรับขั้นตอนของวิธีมอนติคาร์โล..... 22

3.2 แสดงผังงานสำหรับขั้นตอนในการดำเนินการวิจัย.....24

3.3 แสดงผังงานสำหรับขั้นตอนของวิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมตัวแปรอิสระ.....28

3.4 แสดงผังงานสำหรับขั้นตอนของวิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด.....31

3.5 แสดงผังงานสำหรับขั้นตอนของวิธีการถดถอยแบบนบันได.....34

4.1 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....43

4.2 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....46

4.3 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....50

4.4 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....53

4.5 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....56

4.6 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$.....59

4.7 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....65

รูปที่ หน้า

4.8 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....68

4.9 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....71

4.10 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....74

4.11 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....77

4.12 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำมวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$.....80

4.13 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....87

4.14 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....90

4.15 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$.....93

รูปที่ หน้า

4.16 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,100).....:96

4.17 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,100).....99

4.18 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,100).....102

4.19 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,500).....109

4.20 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,500).....112

4.21 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,500).....115

4.22 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c)$ = (10,500).....118

รูปที่ หน้า

4.23 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) =$ (10,500).....121

4.24 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) =$ (10,500).....124

5.1 แสดงแผนผังผลสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุค.....133

5.2 แสดงแผนผังผลสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคในเชิงปฏิบัติ.....134

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันการวิจัยสาขาต่าง ๆ ไม่ว่าจะเป็นด้านวิทยาศาสตร์ ศึกษาศาสตร์ สังคมศาสตร์ อุตสาหกรรม เกษตรกรรม และทางด้านธุรกิจ จำเป็นต้องอาศัยระเบียบวิธีวิจัยทางสถิติเข้ามาช่วยในการค้นคว้าและดำเนินงานวิจัยอย่างเป็นระบบ เพื่อให้งานวิจัยที่ได้มีความถูกต้องและมีความน่าเชื่อถือ การใช้ระเบียบวิธีวิจัยทางสถิติช่วยในการวิเคราะห์เพื่อหาคำตอบสำหรับการวิจัยนั้น ๆ ผู้วิจัยนิยมใช้วิธีการวิเคราะห์ความถดถอย (regression analysis) ช่วยในการค้นหาคำตอบเพื่ออธิบายความสัมพันธ์ระหว่างปัญหาที่สนใจศึกษาและสาเหตุของปัญหานั้น ๆ หรือเพื่อคาดคะเนเหตุการณ์ล่วงหน้า หรือเรียกอีกนัยหนึ่งว่าการพยากรณ์

วิธีการวิเคราะห์ความถดถอยที่นิยมใช้กันอย่างแพร่หลายคือ การวิเคราะห์ความถดถอยเชิงเส้น (linear regression analysis) ซึ่งมีตัวแบบในรูปทั่วไปดังนี้

(1.1) $$\underset{\sim}{y} = X\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

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

X เป็นเมทริกซ์ของตัวแปรอิสระขนาด $n\times(p+1)$

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่าขนาด $(p+1)\times 1$

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

n เป็นขนาดตัวอย่าง

และ p เป็นจำนวนตัวแปรอิสระ

โดยทั่วไปการพิจารณาตัวแปรอิสระเพื่อใช้ในตัวแบบถดถอยเชิงเส้นในการพยากรณ์นั้น ถ้าใช้ตัวแปรอิสระเพียงตัวเดียวในตัวแบบถดถอยเชิงเส้นจะเรียกว่า การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย(simple linear regression) ซึ่งอาจจะทำให้การพยากรณ์คลาดเคลื่อนจากค่าที่แท้จริงได้ เนื่องจากตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรตามอาจมีมากกว่า 1 ตัวแปร ดังนั้น ผู้วิจัยจะต้องใช้ตัวแปรอิสระมากกว่า 1 ตัวแปรในการวิเคราะห์ความถดถอยหรือเรียกว่า การวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ (multiple linear regression) ซึ่งจะช่วยให้การพยากรณ์แม่นยำและถูกต้องมากขึ้น

บางครั้งการใช้ตัวแปรอิสระมากเกินไปในตัวแบบถดถอยอาจไม่ให้ผลดีเสมอไป ในทางกลับกันอาจจะเพิ่มความคลาดเคลื่อนในการพยากรณ์ให้สูงกว่าการใช้ตัวแปรอิสระเพียงตัวเดียว

ในตัวแบบถดถอยก็เป็นได้ ดังนั้น การคัดเลือกตัวแปรอิสระในการวิเคราะห์ความถดถอยเชิงเส้นจึงเป็นสิ่งสำคัญเพื่อให้ได้ตัวแบบถดถอยเชิงเส้นที่ดีที่สุดเพื่อใช้ในการพยากรณ์ หรือในบางครั้งเมื่อผู้วิจัยนำข้อมูลมาวิเคราะห์ความถดถอยเชิงเส้น จนกระทั่งภายหลังได้ตัวแบบถดถอยเชิงเส้น (linear regression model) เพื่อนำไปใช้ในการพยากรณ์นั้น ผู้วิจัยส่วนใหญ่ไม่ได้ทำการตรวจสอบความเหมาะสมของตัวแบบถดถอยเชิงเส้นที่ได้ว่ามีความเหมาะสมที่จะนำไปใช้พยากรณ์หรือไม่ เพราะเหตุนี้การคัดเลือกตัวแบบที่ดีที่สุดเพื่อใช้ในการพยากรณ์ จึงสามารถจำแนกได้เป็น 2 แนวทาง คือ [1]

1) ใช้ค่าสถิติในการพิจารณาความเหมาะสมของตัวแบบทุกรูปแบบที่เป็นไปได้ หรือเรียกว่าวิธีพิจารณาทุกรูปแบบ (all possible regression) ค่าสถิติที่นิยมใช้ในวิธีนี้ได้แก่ ค่าสัมประสิทธิ์การกำหนด (R^2) ค่าสัมประสิทธิ์การกำหนดที่ปรับแล้ว (R^2_{adj}) และค่าสถิติของมอลโลวส์ (Mallow Cp) เป็นต้น

2) ใช้วิธีการคัดเลือกตัวแปรอิสระ (variable selection) โดยการเพิ่ม และ/หรือ ลดตัวแปรอิสระจากตัวแบบการถดถอย ซึ่งวิธีที่นิยมใช้กันอย่างแพร่หลายได้แก่ วิธีการคัดเลือกตัวแปรแบบไปข้างหน้า (forward selection) วิธีการกำจัดตัวแปรแบบถอยหลัง (backward elimination) และวิธีการถดถอยแบบขั้นบันได (stepwise regression) เป็นต้น

การใช้เกณฑ์ในการคัดเลือกตัวแบบ 2 แนวทางข้างต้นเป็นวิธีการที่ใช้เพียงข้อมูลปัจจุบันในการประมาณค่าพารามิเตอร์ ซึ่งอาจทำให้เกิดความคลาดเคลื่อนของการประมาณระดับหนึ่ง ดังนั้น วิธีที่จะลดความคลาดเคลื่อนให้ต่ำลงจึงควรใช้ข้อมูลในอดีตของพารามิเตอร์มาช่วยในการประเมินค่า กล่าวคือ ใช้การแจกแจงก่อน (prior distribution) ของพารามิเตอร์มาพิจารณาด้วย ซึ่งแนวความคิดนี้เรียกว่า การวิเคราะห์เชิงเบส์ (Bayesian approach) การนำแนวทางของเบส์มาใช้ในการหาเกณฑ์หรือวิธีการคัดเลือกตัวแบบการถดถอยเชิงเส้นที่ดีที่สุด จึงเป็นแนวทางที่นักสถิติให้ความสนใจกันมากในปัจจุบัน เนื่องจากมีแนวโน้มทำให้ได้วิธีการคัดเลือกตัวแบบที่มีประสิทธิภาพและเหมาะสมกว่าวิธีการภายใต้แนวทางอื่น ๆ

ในปี ค.ศ. 1997 ราฟเทอร์รี (Raftery) เมดิแกน (Madigan) และโฮเอ็ททิง (Hoeting) ได้เสนอวิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA)) สำหรับตัวแบบการถดถอยเชิงเส้น ซึ่งวิธีนี้มีการคำนึงถึงหลักการเกี่ยวกับความไม่แน่นอนของตัวแบบ (model uncertainty) วิธีการเฉลี่ยตัวแบบของเบส์จะพิจารณาค่าความน่าจะเป็นภายหลัง (posterior probability) สำหรับทุก ๆ ตัวแบบที่เราสนใจ และนำตัวแบบทุกตัวแบบที่เราสนใจมาเฉลี่ยกันโดยใช้ความน่าจะเป็นภายหลังของแต่ละตัวแบบเป็นตัวถ่วงน้ำหนักเพื่อหาค่าพยากรณ์ที่เหมาะสม

---

[1] ทรงศิริ แต้สมบัติ, การวิเคราะห์การถดถอย (กรุงเทพมหานคร : มหาวิทยาลัยเกษตรศาสตร์, 2541) หน้า 205.

สำหรับการหาปริภูมิตัวแบบที่จะนำมาใช้ในวิธีการเฉลี่ยตัวแบบของเบส์นั้น ราฟเทอร์รี เมดิแกน และโฮเอ็ททิง (Raftery, Madigan and Hoeting, 1997) ได้เสนอไว้ 2 แนวทางดังนี้คือ

1. วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA))โดยใช้การค้นหาปริภูมิตัวแบบด้วยวิธีออกส์แคมวินโดว์ (Occam's Window)(BMA $_{OCC}$)

2. วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA))โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) (BMA $_{MC3}$)

ในปี ค.ศ. 1999 โฮเอ็ททิง (Hoeting) ราฟเทอร์รี (Raftery) และเมดิแกน (Madigan) ได้เสนอวิธีสำหรับพิจารณาตัวแปรอิสระพร้อมกับคัดเลือกการแปลงด้วยค่าความน่าจะเป็นภายหลัง ซึ่งพิจารณาทุก ๆ ตัวแบบที่เป็นไปได้โดยใช้ตัวแบบเปลี่ยนจุด (change-point model) หรือการแปลงเปลี่ยนจุด(change-point transformation) ทำให้ได้ตัวแบบที่เหมาะสมมากกว่าการเลือกกำลังการแปลงของบอกซ์และทิดเวล (the standard Box-Tidwell power transformations) ราฟเทอร์รี เมดิแกน และโฮเอ็ททิง (Raftery, Madigan and Hoeting, 1999) เสนอการวิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการคัดเลือกตัวแปรและการแปลงไปพร้อม ๆ กัน ซึ่งทำให้ค่าพยากรณ์เหมาะสมขึ้นทั้งในเรื่องของคะแนนการพยากรณ์ทั้งหมด (overall predictive score) และความครอบคลุมของช่วงการพยากรณ์ (the coverage of prediction intervals)

ในปี ค.ศ. 2002 บาร์บีรี (Barbieri) และเบอร์เกอร์ (Berger) ได้เสนอวิธีการคัดเลือกตัวแบบของเบส์สำหรับการคัดเลือกระหว่างกลุ่มของตัวแบบเชิงเส้นปกติ (selection among normal linear models) ตัวแบบที่เหมาะสมที่สุดอาจไม่ใช่ตัวแบบที่มีค่าความน่าจะเป็นภายหลังสูงสุด (the model with highest posterior probability) แต่กลับเป็นตัวแบบที่ค่าความน่าจะเป็นภายหลังกลาง ๆ (the median probability model) และเกณฑ์ที่ใช้ในการเลือกตัวแบบที่เหมาะสมที่สุดจากกลุ่มของตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ คือความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสอง (square error loss) ซึ่งจะเลือกตัวแบบที่มีค่าดังกล่าวต่ำสุด

ในปี ค.ศ. 2002 นิทัสน์ สุขสุวรรณ เปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางของเบส์ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ โดยเปรียบเทียบวิธีการคัดเลือกสมการถดถอย 5 วิธีได้แก่ วิธีการคัดเลือกตัวแบบที่ดีที่สุดโดยใช้เกณฑ์ข้อสนเทศของเบส์ (BIC) วิธีการคัดเลือกตัวแปรของเบส์ (BVS) วิธีการเฉลี่ยตัวแบบของเบส์โดยใช้การค้นหาปริภูมิตัวแบบด้วยวิธีออกส์แคมวินโดว์ (BMA $_{OCC}$) วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (BMA $_{MC3}$) และวิธีการถดถอยแบบขั้นบันได

(SR) เกณฑ์ที่ใช้ในการตัดสินใจคือเกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ผลการวิจัยได้ข้อสรุปดังนี้

การเปรียบเทียบค่า AMSE ของทั้ง 5 วิธีเรียงลำดับจากน้อยไปมากได้แก่ วิธี BMA_{MC3} BMA_{OCC} BVS BIC และ SR ตามลำดับสำหรับทุกสถานการณ์ วิธี BMA_{OCC} จะได้ค่าAMSE สูงกว่าวิธี BMA_{MC3} เพียงเล็กน้อยโดยเฉพาะอย่างยิ่งเมื่อระดับนัยสำคัญลดลง สำหรับวิธี BVS จะให้ค่า AMSE ใกล้เคียงกับวิธี BMA_{MC3} และวิธี BMA_{OCC} ก็ต่อเมื่อค่าคงที่ $(\sigma_\beta/\tau, c)$ มีค่าต่ำ ๆ ส่วนวิธี BIC และ SR มีค่า AMSE แตกต่างจากวิธี BMA_{MC3} และ BMA_{OCC} อย่างชัดเจนในทุกสถานการณ์

ในการวิจัยครั้งนี้ผู้วิจัยได้สนใจทำการศึกษาเปรียบเทียบวิธีการที่ใช้ในการคัดเลือกตัวแบบการถดถอยเชิงเส้นพหุคูณ 3 วิธี คือ

(1) วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA)) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (Model Uncertainty via Simultaneous Variable and Transformation Selection)(BMA_{SVT})

(2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (Optimal Predictive Model Selection: median probability model)(OPM)

(3) วิธีการถดถอยแบบขั้นบันได (Stepwise Regression method)(SR)

โดยสองวิธีแรกเป็นวิธีการภายใต้แนวทางของเบส์ ส่วนวิธีที่ 3 เป็นวิธีการพื้นฐานซึ่งนิยมใช้กันอย่างแพร่หลายและมีประสิทธิภาพดีในการคัดเลือกสมการถดถอยที่ดีที่สุดเมื่อการวิเคราะห์ความถดถอยเป็นไปตามข้อตกลงเบื้องต้น[1] ซึ่งผู้วิจัยได้นำมาเปรียบเทียบกับวิธีการภายใต้แนวทางของเบส์เพื่อศึกษาดูแนวโน้มว่าวิธีการคัดเลือกตัวแบบภายใต้แนวทางของเบส์จะมีความเหมาะสมกว่าวิธีการพื้นฐานหรือไม่เมื่อใช้หลักเกณฑ์คู่สังยุคแบบปกติ ดังนั้น ถ้าผลการวิจัยพบว่าวิธีการใดมีประสิทธิภาพในการพยากรณ์มากที่สุด ก็ควรจะนำไปใช้ในการคัดเลือกตัวแบบการถดถอยเชิงเส้นพหุคูณที่ดีที่สุดต่อไป

[1] จะเด็จ สวรรค์ตรานนท์, " การเปรียบเทียบวิธีที่ใช้สำหรับการเลือกสมการถดถอยที่ดีที่สุด ", (วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2530), หน้า 212

1.2 วัตถุประสงค์ของการวิจัย

1) เพื่อศึกษาการคัดเลือกสมการถดถอยที่ดีที่สุดในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ 3 วิธีดังกล่าวข้างต้น
2) เพื่อเปรียบเทียบความถูกต้องของการพยากรณ์จากตัวแบบที่ได้จากวิธีการต่าง ๆ

1.3 ข้อตกลงเบื้องต้น

1) รูปแบบทั่วไปของสมการถดถอยเชิงเส้นพหุคูณมีรูปแบบดังสมการ (1.1)
2) ตัวแปรอิสระแต่ละตัวถือว่าเป็นค่าคงที่
3) ความคลาดเคลื่อนสุ่มเป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และส่วนเบี่ยงเบนมาตรฐานเป็น σ เหมือนกันและเป็นอิสระซึ่งกันและกัน

1.4 สมมุติฐานของการวิจัย

1) วิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางของเบส์ ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณนั้น จะให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากกว่าวิธีการพื้นฐาน
2) วิธีการเฉลี่ยตัวแบบของเบส์เมื่อพิจารณาการแปลงค่าตัวแปรอิสระที่เหมาะสม จะให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากกว่าวิธีอื่น ๆ

1.5 ขอบเขตของการวิจัย

1) ตัวแบบของการถดถอยเชิงเส้นพหุคูณที่สนใจศึกษามีรูปแบบดังสมการ (1.1) ซึ่งรูปแบบการถดถอยนอกจากจะเป็นเชิงเส้นในพารามิเตอร์แล้ว ยังเป็นเชิงเส้นในตัวแปรอิสระด้วย
2) ตัวแบบการถดถอยในการวิจัยครั้งนี้เป็นตัวแบบติดกลุ่ม(nested models)
3) กำหนดขนาดตัวอย่างที่ศึกษา คือ 15 30 50 และ 100 ซึ่งในกรณีขนาดตัวอย่างเท่ากับ 15 ผู้วิจัยได้ทำการศึกษาเฉพาะกรณีจำนวนตัวแปรอิสระเท่ากับ 3 และ 5 เท่านั้น และในกรณีขนาดตัวอย่างเท่ากับ 30 ผู้วิจัยได้ทำการศึกษาเฉพาะกรณีจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 และ12 เท่านั้น
4) จำนวนตัวแปรอิสระที่ศึกษาคือ 3 5 8 10 12 และ 15 โดยตัวแปรอิสระแต่ละตัวแปรสร้างจากการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

5) ในการวิจัยครั้งนี้จะศึกษาเมื่อความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ
6) ในการวิจัยครั้งนี้จะศึกษากรณีการแจกแจงก่อนของเวกเตอร์พารามิเตอร์สัมประสิทธิ์ความถดถอย $\underset{\sim}{\beta}$ เมื่อทราบค่า σ^2 แบบคู่สังยุคปกติ
7) ค่าคงที่สำหรับวิธี BMA_{SVT} และวิธี OPM มี 4 ระดับ คือ (1,5) (1,10) (10,100) และ (10,500)
8) ข้อมูลที่ใช้ในการศึกษาวิจัยครั้งนี้ใช้โปรแกรมคอมพิวเตอร์จำลองด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 500 รอบในแต่ละสถานการณ์

1.6 ขั้นตอนในการศึกษาวิจัย

1) กำหนดลักษณะการแจกแจงของค่าคลาดเคลื่อน ขนาดตัวอย่าง จำนวนตัวแปรอิสระ ระดับนัยสำคัญและค่าคงที่สำหรับวิธี BMA_{SVT}และวิธี OPM
2) สร้างข้อมูลตัวแปรตามจากตัวแปรอิสระและความคลาดเคลื่อนที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรอิสระ
3) สร้างตัวแบบโดยวิธีทั้ง 3 วิธี คือ

   (1) วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA)) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (Model Uncertainty via Simultaneous Variable and Transformation Selection)(BMA_{SVT})

   (2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (Optimal Predictive Model Selection: median probability model)(OPM)

   (3) วิธีการถดถอยแบบขั้นบันได (Stepwise Regression method)(SR)
4) หาค่าคลาดเคลื่อนกำลังสองเฉลี่ยของแต่ละตัวแบบ
5) สรุปผลในรูปของตารางและรูปภาพ

1.7 ประโยชน์ที่คาดว่าจะได้รับ

1) เป็นแนวทางในการคัดเลือกตัวแบบที่เหมาะสมเพื่อใช้ในการพยากรณ์สำหรับการวิเคราะห์ความถดถอยเชิงพหุคูณ
2) เป็นแนวทางในการศึกษาวิธีการคัดเลือกตัวแบบการถดถอยที่ดีที่สุดวิธีอื่น ๆ ภายใต้แนวทางของเบส์ สำหรับการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณต่อไป

1.8 เกณฑ์ที่ใช้ในการตัดสินใจ

เกณฑ์ที่ใช้ในการตัดสินใจว่าวิธีการคัดเลือกตัวแบบการถดถอยวิธีใดจะมีประสิทธิภาพมากที่สุดพิจารณาจากเกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Average of Mean Square Error (AMSE)) และเกณฑ์ที่ใช้ประกอบการตัดสินใจจะใช้เกณฑ์ค่าอัตราส่วนผลต่างของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Ratio of Different Average Mean Square Error (RDAMSE)) ซึ่งมีสูตรดังนี้

เพราะว่า $$MSE_j = \frac{\sum_{j=1}^{n} (y_i - \hat{y}_i)^2}{n-p}$$

ดังนั้น $$AMSE = \frac{\sum_{j=1}^{500} MSE_j}{500}$$

เมื่อ y_i เป็นค่าสังเกตที่ i

$\hat{y}_i$ เป็นค่าพยากรณ์ที่ j

p เป็นจำนวนพารามิเตอร์ในตัวแบบการถดถอย

n เป็นขนาดตัวอย่าง

MSE_j เป็นค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการทำซ้ำรอบที่ j

และ $AMSE$ เป็นค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากการทำซ้ำ 500 รอบ

ส่วน $$RDAMSE_i = \frac{(AMSE_i - AMSE_{\min})}{AMSE_{\min}} x100\%$$

เมื่อ $AMSE_i$ เป็นค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากวิธีที่ i

และ $AMSE_{\min}$ เป็นค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่มีค่าต่ำสุดจากทุกวิธี

บทที่ 2

ทฤษฎีและสถิติที่เกี่ยวข้อง

การศึกษาเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์ในการวิจัยครั้งนี้เป็นการพิจารณาหาวิธีการคัดเลือกสมการถดถอยที่ให้ค่าพยากรณ์ถูกต้องและแม่นยำมากที่สุด โดยวิธีการคัดเลือกสมการถดถอยที่นำมาศึกษามีดังนี้

1) วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA)) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (Model Uncertainty via Simultaneous Variable and Transformation Selection) (BMA_{SVT})

2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (Optimal Predictive Model Selection: median probability model (OPM))

3) วิธีการถดถอยแบบขั้นบันได (Stepwise Regression method (SR))

การนำเสนอทฤษฎีและสถิติที่เกี่ยวข้องจะประกอบด้วยแนวคิดและทฤษฎีที่เกี่ยวข้องกับการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ ความน่าจะเป็นที่มีเงื่อนไข ทฤษฎีบทของเบส์ ความเป็นอิสระและทฤษฎีของวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีข้างต้น โดยมีรายละเอียดดังนี้

2.1 การวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis)[1]

ตัวแบบและข้อตกลงเบื้องต้นที่ใช้ในการวิจัยครั้งนี้ คือ ตัวแบบความถดถอยเชิงเส้นพหุคูณซึ่งมีรูปแบบทั่วไปดังนี้

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_j x_{ip} + \varepsilon_i \quad ; \quad i = 1, 2, \ldots, n$$

เมื่อ β_j เป็นสัมประสิทธิ์การถดถอยที่ j, $j = 1, 2, \ldots, p$

ข้อตกลงเบื้องต้นของการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณมีดังนี้

1. ε_i มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวนเป็น σ^2 คือ ε_i จ.ม.อ. (i.i.d.) $N(0, \sigma^2)$ กล่าวคือ ε_i และ ε_j สำหรับ $i \neq j$ มีการแจกแจงที่เหมือนกันและเป็นอิสระต่อกัน ซึ่งจะทำให้ $Cov(\varepsilon_i, \varepsilon_j) = 0$ สำหรับ $i \neq j$

---

[1] นิทัสน์ สุขสุวรรณ, "การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางของเบส์ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ" (วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2545), หน้า 11

1) รูปแบบการถดถอยเป็นแบบเชิงเส้นตรงของพารามิเตอร์
2) ตัวแปรอิสระแต่ละตัวเป็นค่าคงที่และไม่มีพหุสัมพันธ์กัน

เนื่องจากตัวแบบการถดถอยที่ใช้ในการพิจารณาเป็นตัวแบบเชิงเส้น ดังนั้นในการประมาณค่าพารามิเตอร์เราสามารถใช้วิธีกำลังสองน้อยสุดเพื่อประมาณค่าเวกเตอร์พารามิเตอร์สัมประสิทธิ์การถดถอย $\underset{\sim}{\beta}$ ของตัวแบบความถดถอยเชิงเส้นพหุคูณซึ่งมีรูปแบบตัวประมาณดังนี้

ตัวประมาณค่ากำลังสองน้อยสุด (Least Square estimator: $\underset{\sim}{\hat{\beta}}$) อยู่ในรูปของ

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

เมื่อ X มีค่าลำดับชั้นเป็นหนึ่ง

2.2 ความน่าจะเป็นที่มีเงื่อนไข (Conditional Probability) **ทฤษฎีบทของเบส์** (Bayes' Theorem) **และความเป็นอิสระ** (Independence) ดังแสดงรายละเอียดในภาคผนวก

2.3 วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA)) **โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ** (Markov Chain Monte Carlo model composition (MC^3)) **เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ** (Model Uncertainty via Simultaneous Variable and Transformation Selection) (BMA_{SVT})

ในปี ค.ศ. 1999 โฮเอ็ททิง (Hoeting) ราฟเทอร์รี (Raftery) และเมดิแกน (Madigan) ได้เสนอวิธีสำหรับการคัดเลือกตัวแปรอิสระตามความน่าจะเป็นภายหลัง (posterior probability) และการพิจารณาการแปลงที่เหมาะสมสำหรับตัวแปรอิสระไปพร้อมๆกัน วิธีการนี้จะพิจารณาทุกๆตัวแบบที่เป็นไปได้โดยใช้ตัวแบบเปลี่ยนจุด (change-point model) หรือการแปลงเปลี่ยนจุด (change-point transformation) ทำให้ได้ตัวแบบที่เหมาะสมมากกว่าการเลือกกำลังการแปลงของบอกซ์และทิดเวล (the standard Box-Tidwell power transformations) ราฟเทอร์รี เมดิแกน และโฮเอ็ททิง (Raftery, Madigan and Hoeting, 1999) เสนอการวิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการคัดเลือกตัวแปรและการแปลงไปพร้อม ๆ กัน ซึ่งทำให้ค่าพยากรณ์เหมาะสมขึ้นทั้งในเรื่องของคะแนนการพยากรณ์ทั้งหมด (overall predictive score) และความครอบคลุมของช่วงการพยากรณ์ (the coverage of prediction intervals)

การนำเสนอแนวคิดและทฤษฎีของวิธี BMA_{SVT} จำเป็นต้องกล่าวถึงวิธีการเฉลี่ยตัวแบบของเบส์ ขั้นตอนการคัดเลือกตัวแปรของวิธี BMA_{MC3} ก่อนที่จะกล่าวถึงขั้นตอนของการคัดเลือกตัวแปรและการแปลงไปพร้อม ๆ กันของวิธี BMA_{SVT} โดยมีรายละเอียดดังนี้

2.3.1 วิธีการเฉลี่ยตัวแบบของเบส์ (BMA)

วิธี BMA สำหรับตัวแบบการถดถอยเชิงเส้น เสนอโดยราฟเทอร์รี เมดิแกน และโฮเอ็ททิง(Raftery Madigan and Hoeting, 1997) เป็นวิธีการพิจารณาหาค่าความน่าจะเป็นภายหลัง สำหรับทุก ๆ ตัวแบบที่เราสนใจ และนำตัวแบบทุกตัวแบบที่เราสนใจมาเฉลี่ยกันโดยใช้ความน่าจะเป็นภายหลังของแต่ละตัวแบบเป็นตัวถ่วงน้ำหนัก เพื่อหาค่าพยากรณ์ที่เหมาะสม

วิธีการเฉลี่ยตัวแบบของเบส์นี้เป็นการพิจารณาโดยคำนึงถึงความไม่แน่นอนของตัวแบบ (model uncertainty) และมีแนวคิดว่าการคัดเลือกตัวแบบที่ดีที่สุดเพียงตัวแบบเดียวจากทุกตัวแบบที่เป็นไปได้ ถือเป็นการละเลยตัวแบบอื่น ๆ ซึ่งบ่อยครั้งเราจะพบว่ามีตัวแบบหลายรูปแบบที่มีประสิทธิภาพในการพยากรณ์ใกล้เคียงกัน ดังนั้นเมื่อพิจารณาถึงหลักการเกี่ยวกับความไม่แน่นอนของตัวแบบแล้ว วิธีการเฉลี่ยตัวแบบของเบส์น่าจะทำให้ได้ค่าพยากรณ์ที่มีประสิทธิภาพมากกว่าค่าพยากรณ์ที่ได้จากตัวแบบเดียว

การแจกแจงความน่าจะเป็นภายหลังของสิ่งที่สนใจ เมื่อมีข้อมูลจะเป็นดังนี้

$$p\left(\Delta \mid D\right) = \sum_{k=1}^{K} p\left(\Delta \mid M_k, D\right) \cdot p\left(M_k \mid D\right) \tag{2.1}$$

เมื่อ Δ เป็นปริมาณของสิ่งที่สนใจ เช่น ค่าพยากรณ์ที่สนใจ

D เป็นข้อมูลของตัวแบบที่สนใจ

และ $M_1, \ldots, M_K$ เป็นตัวแบบที่พิจารณา

สมการ (2.1) เป็นการเฉลี่ยการแจกแจงความน่าจะเป็นภายหลัง ถ่วงน้ำหนักด้วยความน่าจะเป็นภายหลังของแต่ละตัวแบบ โดยเรียกแนวคิดนี้ว่า "การเฉลี่ยตัวแบบของเบส์ “

ความน่าจะเป็นภายหลังของตัวแบบ M_k คือ

$$p\left(M_k \mid D\right) = \frac{p\left(D \mid M_k\right) \cdot p\left(M_k\right)}{\sum_{i=1}^{K} p\left(D \mid M_k\right) \cdot p\left(M_i\right)} \tag{2.2}$$

เมื่อ

$$p\left(D \mid M_k\right) = \int p\left(D \mid \underset{\sim}{\theta}_k, M_k\right) p\left(\underset{\sim}{\theta}_k \mid M_k\right) d\,\underset{\sim}{\theta}_k \tag{2.3}$$

ซึ่งเป็นความควรจะเป็นขอบ (marginal likelihood) ของตัวแบบ M_k

โดยที่ $\underset{\sim k}{\theta}$ เป็นเวกเตอร์ของพารามิเตอร์ สำหรับการวิเคราะห์ความถดถอย หมายถึง $\underset{\sim}{\beta}$ และ σ^2

$p\left(\underset{\sim k}{\theta} \mid M_k\right)$ เป็นความหนาแน่นก่อนของ $\underset{\sim k}{\theta}$ ภายใต้ตัวแบบ M_k

$p\left(D \mid \underset{\sim k}{\theta}, M_k\right)$ เป็นฟังก์ชันความควรจะเป็น

และ $p\left(M_k\right)$ เป็นความน่าจะเป็นก่อน (prior probability) สำหรับตัวแบบ M_k

ค่าเฉลี่ยและความแปรปรวนภายหลังของสิ่งที่สนใจ เป็นดังนี้

(2.4) $$E\left[\Delta \mid D\right] = \sum_{k=0}^{K} \hat{\Delta}_k \; p\left(M_k \mid D\right)$$

(2.5) และ $$Var\left[\Delta \mid D\right] = \sum_{k=0}^{K}\left(Var\left[\Delta \mid D, M_k\right] + \hat{\Delta}_k^2\right) \bullet p\left(M_k \mid D\right) - E\left[\Delta \mid D\right]^2$$

เมื่อ $\hat{\Delta}_k = E\left[\Delta \mid D, M_k\right]$

จากสมการ (2.4) เราสามารถสรุปได้ว่า ถ้าสิ่งที่เราสนใจ Δ เป็นค่าพยากรณ์ $\hat{y}$ ค่าคาดหมายของค่าพยากรณ์ $\hat{y}$ เมื่อมีข้อมูล คือ การเฉลี่ยค่าพยากรณ์ของแต่ละตัวแบบถ่วงน้ำหนักด้วยความน่าจะเป็นภายหลังของแต่ละตัวแบบ M_k จะทำให้เราได้ค่าพยากรณ์ $\hat{y}$ ตามที่ต้องการ

นอกจากนี้ในงานวิจัยของเมดิแกนและราฟเทอร์รี (Madigan and Raftery, 1994) ยังแสดงให้เห็นว่าการเฉลี่ยตัวแบบจะทำให้การพยากรณ์มีประสิทธิภาพสูงกว่าการใช้ตัวแบบเดียวในการพยากรณ์ โดยใช้กฎของคะแนนลอการิทึมและแนวคิดเกี่ยวกับรูปแบบข้อสนเทศของคูลล์แบ็ค-ไลเบอร์ (Kullback-leibler, 1951) จะได้ว่า

$$-E\left[\log\left\{\sum_{k=1}^{K} p\left(\Delta \mid M_k, D\right) \bullet p\left(M_k \mid D\right)\right\}\right] \leq -E\left[\log\left\{p\left(\Delta \mid M_j, D\right)\right\}\right],\ j = 1, \ldots, K$$

ซึ่งเป็นการยืนยันว่าวิธีการเฉลี่ยตัวแบบจะมีประสิทธิภาพในการพยากรณ์สูงกว่าการใช้ตัวแบบเดียวในการพยากรณ์

2.3.1.1 **ความน่าจะเป็นก่อนของตัวแบบ**

ความน่าจะเป็นก่อนของแต่ละตัวแบบจะอยู่ในรูปแบบดังนี้

(2.6) $$p(M_i) = \prod_{j=1}^{k} p_j^{\delta_{ij}} \left(1-p_j\right)^{1-\delta_{ij}}$$

เมื่อ p_j เป็นความน่าจะเป็นที่ตัวแปรอิสระที่ j จะรวมอยู่ในตัวแบบ M_j

δ_{ij} เป็นตัวบ่งชี้ที่มีค่า 0 หรือ 1

โดยที่ δ_{ij} เป็น 0 เมื่อตัวแปรอิสระที่ j ไม่อยู่ในตัวแบบ M_j

และ δ_{ij} เป็น 1 เมื่อตัวแปรอิสระที่ j อยู่ในตัวแบบ M_j

ในการวิจัยครั้งนี้จะกำหนด p_j สำหรับทุก ๆ ตัวแปรอิสระเป็น $\left(\frac{1}{2}\right)$ หรือตัวแปรอิสระทุกตัวจะมีโอกาสอยู่ในตัวแบบหรือไม่อยู่ในตัวแบบเท่า ๆ กัน

2.3.1.2 **ความน่าจะเป็นภายหลังขอบ** (marginal posterior probability) **ของตัวแบบ**

เนื่องจากความน่าจะเป็นภายหลังขอบของตัวแบบ *M* ใด ๆ จะแปรผันตามฟังก์ชันความควรจะเป็นคูณกับการแจกแจงก่อนของพารามิเตอร์ กล่าวคือ

$$p(D \mid M) \propto likelihood \times prior$$

การวิจัยครั้งนี้ผู้วิจัยจะกำหนดการแจกแจงก่อนคู่สังยุค (conjugate prior distribution) สำหรับ $\underset{\sim}{\beta}$ เมื่อ σ^2 ทราบค่าเป็นการแจกแจงปกติหลายตัวแปร (multivariate normal distribution) กล่าวคือเป็นหลักเกณฑ์สังยุคแบบปกติ (conjugate normal prior)

$$p(\underset{\sim}{\beta}) \propto \exp\left\{-\frac{1}{2\sigma^2}(A^{1/2}\underset{\sim}{\overline{\beta}} - A^{1/2}\underset{\sim}{\beta})'(A^{1/2}\underset{\sim}{\overline{\beta}} - A^{1/2}\underset{\sim}{\beta})\right\}$$

เมื่อ $\underset{\sim}{\overline{\beta}}$ คือ เวกเตอร์ค่าเฉลี่ยก่อน (prior mean)

A คือ เมทริกซ์บวกแน่นอน (positive definite matrix) ซึ่งกำหนดให้ $A^{1/2}$เป็นเมทริกซ์สมมาตร (symmetric matrix) ที่มีคุณสมบัติ $A^{1/2}\ A^{1/2} = A$

และ $A = \sigma^2 \underset{\sim}{\sum}_{\beta}^{-1}$ เมื่อ $\underset{\sim}{\sum}_{\beta}^{-1}$คือเมทริกซ์ความแปรปรวนร่วมก่อน (prior covariance matrix) ของ $\underset{\sim}{\beta}$

การแจกแจงความน่าจะเป็นภายหลังสำหรับตัวแบบความถดถอย *M* คือ

$$p(D \mid M) \propto \text{likelihood} \times \text{prior}$$

$$p(D \mid M) \propto l\left(\underset{\sim}{\beta} \middle| \sigma, \underset{\sim}{y}\right) \times p\left(\underset{\sim}{\beta}\right)$$

$$p(D \mid M) \propto \exp\left[-\frac{1}{2\sigma^2}\{(A^{1/2}\underset{\sim}{\bar{\beta}} - A^{1/2}\underset{\sim}{\beta})'(A^{1/2}\underset{\sim}{\bar{\beta}} - A^{1/2}\underset{\sim}{\beta}) + (\underset{\sim}{y} - X\underset{\sim}{\beta})'(\underset{\sim}{y} - X\underset{\sim}{\beta})\}\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}\left\{\begin{pmatrix} A^{1/2}\underset{\sim}{\bar{\beta}} - A^{1/2}\underset{\sim}{\beta} \\ \underset{\sim}{y} - X\underset{\sim}{\beta} \end{pmatrix}'\begin{pmatrix} A^{1/2}\underset{\sim}{\bar{\beta}} - A^{1/2}\underset{\sim}{\beta} \\ \underset{\sim}{y} - X\underset{\sim}{\beta} \end{pmatrix}\right\}\right]$$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}(\underset{\sim}{w} - G\underset{\sim}{\beta})'(\underset{\sim}{w} - G\underset{\sim}{\beta})\right\}$$

เมื่อ $\underset{\sim}{w}_{\{(n+p+1)\times 1\}} = \begin{pmatrix} A^{1/2}\underset{\sim}{\bar{\beta}} \\ \underset{\sim}{y} \end{pmatrix}$

$$G_{\{(n+p+1)\times 1\}} = \begin{pmatrix} A^{1/2} \\ X \end{pmatrix}$$

และ $(\underset{\sim}{w} - G\underset{\sim}{\beta})'(\underset{\sim}{w} - G\underset{\sim}{\beta}) = (\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})'G'G(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}}) + (\underset{\sim}{w} - G\underset{\sim}{\hat{\beta}})'(\underset{\sim}{w} - G\underset{\sim}{\hat{\beta}})$

โดยที่ $\underset{\sim}{\hat{\beta}} = (G'G)^{-1}G'\underset{\sim}{w}$

$$= \left\{(A^{1/2}X')\begin{pmatrix} A^{1/2} \\ X \end{pmatrix}\right\}^{-1}(A^{1/2}X')\begin{pmatrix} A^{1/2}\underset{\sim}{\bar{\beta}} \\ \underset{\sim}{y} \end{pmatrix}$$

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

$$= (A + X'X)^{-1}(A\underset{\sim}{\bar{\beta}} + X'X\underset{\sim}{b})$$

จะได้ว่า $p(D \mid M) \propto \exp\left\{-\frac{1}{2\sigma^2}(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})'G'G(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})\right\}$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})'(A + X'X)(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})\right\}$$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})'(\underset{\sim}{\bar{\Sigma}}_{\beta}^{-1} + X'X/\sigma^2)(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})\right\}$$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})'\underset{\sim}{\hat{\Sigma}}_{\beta}^{-1}(\underset{\sim}{\beta} - \underset{\sim}{\hat{\beta}})\right\}$$

เมื่อ $\underset{\sim}{\hat{\Sigma}}_{\beta}^{-1} = \underset{\sim}{\bar{\Sigma}}_{\beta}^{-1} + X'X/\sigma^2$

และ $\underset{\sim}{\overline{\overline{\beta}}} = (A + X'X)^{-1}(A\overline{\beta} + X'X\underset{\sim}{b})$

$$= \left[\overline{\sum}_{\underset{\sim}{\beta}}^{-1} + X'X/\sigma^2\right]^{-1}\left[\overline{\sum}_{\underset{\sim}{\beta}}^{-1}\underset{\sim}{\overline{\beta}} + (X'X/\sigma^2)\underset{\sim}{b}\right]$$

ดังนั้น การแจกแจงภายหลังสำหรับ $\underset{\sim}{\beta}$ คือ การแจกแจงปกติหลายตัวแปรด้วยเวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\overline{\overline{\beta}}}$ และเมทริกซ์ความแปรปรวนร่วม $\underset{\sim}{\overline{\overline{\sum}}}_{\beta}$ ซึ่งก็คือ $N_p(\underset{\sim}{\overline{\overline{\beta}}}, \underset{\sim}{\overline{\overline{\sum}}}_{\beta})$

2.3.1.3 ขั้นตอนการคัดเลือกตัวแปรของวิธี BMA_{MC3}

การคัดเลือกตัวแปรของวิธี BMA_{MC3} มีขั้นตอนเหมือนวิธีการคัดเลือกตัวแปรของเบส์เซียนแต่วิธี BMA_{MC3}จะนำตัวแบบในทุกรูปแบบที่ปรากฏอยู่ในปริภูมิตัวแบบมาเฉลี่ยกันโดยใช้ความน่าจะเป็นภายหลังของแต่ละตัวแบบเป็นตัวถ่วงน้ำหนัก โดยรายละเอียดของขั้นตอนการคัดเลือกตัวแปรของเบส์เซียน[2]มีดังนี้

ขั้นตอนแรกของวิธีการนี้เริ่มต้นจากสมการถดถอยแบบเต็มรูป คือ ประกอบด้วยตัวแปรอิสระครบทุกตัว แล้วทำการสุ่มแบบกิบส์ เพื่อสร้างลำดับ

(2.7) $$\underset{\sim}{\delta}^1, \underset{\sim}{\delta}^2, \cdots, \underset{\sim}{\delta}^j, \cdots$$

ลำดับในสมการ (2.7) จะลู่เข้าสู่การแจกแจงของ $\underset{\sim}{\delta} \sim f(\underset{\sim}{\delta} \mid \underset{\sim}{y})$ ซึ่งเป็นลำดับที่มีความน่าจะเป็นค่อนข้างสูง เพราะบรรจุข้อมูลที่สอดคล้องกับการคัดเลือกตัวแปร โดย $\underset{\sim}{\delta}^j$ ใด ๆ ที่มีความน่าจะเป็นสูงสุด ทำให้ง่ายต่อการตัดสินใจในการเลือกตัวแบบที่ดีที่สุด

เราใช้วิธีการสุ่มแบบกิบส์เพื่อสร้างเวกเตอร์เป็นลำดับแบบกิบส์ ดังนี้

(2.8) $$\underset{\sim}{\beta}^0, \sigma^0, \underset{\sim}{\delta}^0, \underset{\sim}{\beta}^1, \sigma^1, \underset{\sim}{\delta}^1, \ldots, \underset{\sim}{\beta}^j, \sigma^j, \underset{\sim}{\delta}^j, \ldots$$

เมื่อ $\underset{\sim}{\beta}^0, \sigma^0$ เป็นค่าเริ่มต้นที่ได้จากการประมาณค่าด้วยวิธีกำลังสองน้อยสุดของสมการ (1.1) และ $\underset{\sim}{\delta}^0$ ถูกกำหนดค่าเริ่มต้นให้เป็น $\underset{\sim}{\delta}^0 \equiv (1,1,\ldots,1)'$

ในแต่ละรอบขอบของการสุ่มค่าของลำดับ $\underset{\sim}{\beta}^j, \sigma^j, \underset{\sim}{\delta}^j$ ได้จากการสร้างค่าที่สอดคล้องกับขั้นตอนต่อไปนี้

---

[2] นิทัสน์ สุขสุวรรณ, "การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางของเบส์ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ" (วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2545), หน้า 24-26

1) สุ่มเวกเตอร์ของสัมประสิทธิ์การถดถอย $\underset{\sim}{\beta}^{j}$ จากการแจกแจง

(2.9) $$f\left(\underset{\sim}{\beta}\,\middle|\,\underset{\sim}{y},\sigma^{j-1},\underset{\sim}{\delta}^{j-1}\right)=N\,p\left(\mathrm{A}_{\delta\,j-1}\left(\sigma^{j-1}\right)^{-2}X'X\,\underset{\sim LS}{\beta},\mathrm{A}_{\delta\,j-1}\right)$$

เมื่อ $\mathrm{A}_{\delta\,j-1}=\left(\left(\sigma^{j-1}\right)^{-2}X'X+D^{-1}_{\delta\,j-1}R^{-1}D^{-1}_{\delta\,j-1}\right)^{-1}$

และ $D^{-1}_{\delta\,j-1}=diag\left[\left(a_1\tau_1\right)^{-1},\left(a_2\tau_2\right)^{-1},\ldots,\left(a_k\tau_k\right)^{-1}\right]$

2) สุ่มค่าความแปรปรวน σ^{j} จากการแจกแจง

(2.10) $$f\left(\sigma^{j}\,\middle|\,\underset{\sim}{y},\underset{\sim}{\beta}^{j},\underset{\sim}{\delta}^{j-1}\right)=IG\left(\frac{n+\upsilon_{\delta\,j-1}}{2},\frac{\left|\underset{\sim}{y}-X\,\underset{\sim}{\beta}^{j}\right|^{2}+\upsilon_{\delta\,j-1}\;\lambda_{\delta\,j-1}}{2}\right)$$

3) ค่าของเวกเตอร์ $\underset{\sim}{\delta}^{j}$ ได้จากการแยกองค์ประกอบ (component wise) โดยการสุ่มอย่างต่อเนื่องกัน (sampling consecutively) จากการแจกแจงแบบมีเงื่อนไขดังนี้

(2.11) $$\delta_i^{j}\;\sim\;f\left(\delta_i^{j}\,\middle|\,\underset{\sim}{y},\underset{\sim}{\beta}^{j},\sigma^{j},\delta_i^{j}\right)=f\left(\delta_i^{j}\,\middle|\,\underset{\sim}{\beta}^{j},\sigma^{j},\delta_i^{j}\right)$$

เมื่อ $\delta_{-i}^{j}=\left(\delta_1^{j},\ldots,\delta_{i-1}^{j},\delta_{i+1}^{j},\ldots,\delta_k^{j}\right)$

จะเห็นได้ว่าการแจกแจงในสมการ (1.12) จะไม่ขึ้นอยู่กับค่าของ $\underset{\sim}{y}$

ในสมการ (1.12) δ_i^{j} แต่ละตัวจะมีการแจกแจงแบบแบร์นูลลี (Bernoulli distribution) ด้วยความน่าจะเป็น

(2.12) $$p\left(\delta_i^{j}=1\,\middle|\,\underset{\sim}{\beta}^{j},\sigma^{j},\delta_{-i}^{j}\right)=\frac{a}{a+b}$$

(2.13) เมื่อ $a=f\left(\underset{\sim}{\beta}^{j}\,\middle|\,\delta_{-i}^{j},\delta_i^{j}=1\right)\times f\left(\sigma^{j}\,\middle|\,\delta_{-i}^{j},\delta_i^{j}=1\right)\times f\left(\delta_{-i}^{j},\delta_i^{j}=1\right)$

(2.14) และ $b=f\left(\underset{\sim}{\beta}^{j}\,\middle|\,\delta_{-i}^{j},\delta_i^{j}=0\right)\times f\left(\sigma^{j}\,\middle|\,\delta_{-i}^{j},\delta_i^{j}=0\right)\times f\left(\delta_{-i}^{j},\delta_i^{j}=0\right)$

ความยาวของลำดับในสมการ (2.7) จะเพิ่มขึ้นเรื่อย ๆ ทำให้การแจกแจงของค่าที่แท้จริงของ $\underset{\sim}{\delta}$ จะลู่เข้าสู่การแจกแจงภายหลัง $\underset{\sim}{\delta}\;\sim\;f(\underset{\sim}{\delta}\mid\underset{\sim}{y})$ การลู่เข้านี้จะเป็นไปอย่างรวดเร็ว ถ้า $\underset{\sim}{\delta}\;\sim\;f(\underset{\sim}{\delta}\mid\underset{\sim}{y})$ เป็นค่าสูงสุด ตัวแบบที่มีน้ำหนักมาก ๆ มีจำนวนไม่มาก ตัวแบบเหล่านี้จะมีความแม่นยำสูงเมื่อ $\underset{\sim}{\delta}\;\sim\;f(\underset{\sim}{\delta}\mid\underset{\sim}{y})$ บรรจุข้อมูลส่วนใหญ่ของการคัดเลือกตัวแปร

2.3.2 ขั้นตอนของการคัดเลือกตัวแปรและการแปลงไปพร้อม ๆ กันของวิธี BMA_{SVT}

วิธีการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method (BMA)) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3)) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (Model Uncertainty via Simultaneous Variable and Transformation Selection (BMA_{SVT})) เสนอโดย ราฟเทอร์รี่ เมดิแกน และโฮเอ็ททิง (Raftery Madigan and Hoeting, 1999) ซึ่งสามารถเพิ่มประสิทธิภาพของวิธีการเฉลี่ยตัวแบบของเบส์ให้มากขึ้น ซึ่งวิธีการดังกล่าวมีขั้นตอนดังนี้

1) นำตัวแปรอิสระเข้าสู่กระบวนการ ACE (Alternating Conditional Expectation Algorithm)

2) พิจารณารูปกราฟแสดงความสัมพันธ์ระหว่างตัวแปรอิสระที่ทำการแปลงและตัวแปรอิสระไม่ได้ทำการแปลง

3) ถ้าตัวแปรตามในกรณีที่ทำการแปลงและตัวแปรตามดั้งเดิมมีความสัมพันธ์กันแบบไม่เป็นเชิงเส้น (non-linear of feature) แสดงว่าควรทำการแปลงตัวแปรตามด้วย หลังจากนั้นนำข้อมูลเข้าสู่กระบวนการเฉลี่ยตัวแบบของเบส์ (Bayesian Model Averaging method) โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (Markov Chain Monte Carlo model composition (MC^3))ด้วยกำลังการแปลงที่เหมาะสม 4 กรณีคือ p = (-1,0,0.5,1) แต่ถ้าหากความสัมพันธ์ดังกล่าวเป็นเชิงเส้น ไม่จำเป็นต้องทำการแปลงค่าตัวแปรตาม

4) สำหรับตัวแปรอิสระใด ๆ ถ้าตัวแปรอิสระในกรณีที่ทำการแปลงและตัวแปรอิสระเดิมมีความสัมพันธ์กันแบบไม่เป็นเชิงเส้น (non-linear of feature) แสดงว่าควรทำการแปลงตัวแปรอิสระนั้น ๆ โดยแต่ละตัวแปรอิสระ จะทำการพิจารณากราฟของการแปลงที่แนะนำโดยกระบวนการ ACE และพิจารณาการแปลงที่เหมาะสมด้วยปัจจัยเบส์ (Bayes factor) เพื่อตัดสินใจว่าตัวแปรอิสระนั้น ๆ สมควรทำการแปลงหรือไม่

5) นำตัวแปรอิสระทุกตัวและตัวแปรอิสระที่ควรทำการแปลงเข้าสู่กระบวนการ MC^3 พร้อมกันเพื่อหาตัวแบบที่เหมาะสมที่สุดในการพยากรณ์

2.4 วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (Optimal Predictive Model Selection)

วิธีการนี้เสนอโดยบาร์บีรีและเบอเกอร์ (Barbieri and Berger, 2002) ซึ่งมีแนวคิดว่าตัวแบบที่เหมาะสมที่สุดในการพยากรณ์อาจไม่ใช่ตัวแบบที่มีความน่าจะเป็นภายหลังสูงสุด (the model with highest posterior probability) แต่กลับเป็นตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ (median probability model) ซึ่งใกล้เคียงกับวิธีการเฉลี่ยตัวแบบของเบส์เมื่อพิจารณาความไม่แน่นอนของตัวแบบ และเกณฑ์ที่ใช้ในการเลือกตัวแบบที่เหมาะสมที่สุดจากกลุ่มของตัวแบบที่มี

ความน่าจะเป็นภายหลังกลาง ๆ คือความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสอง (square error loss) โดยจะเลือกตัวแบบที่มีค่าดังกล่าวต่ำสุด

2.4.1 การค้นหาตัวแบบที่มีความน่าจะเป็นภายหลังกลางๆ

การค้นหาตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ กระทำโดยใช้แผนการค้นหาตัวแบบด้วยเทคนิคมาคอฟเชนมอนติคาร์โล (ordinary MCMC model search schemes) แผนการนี้ได้พัฒนาลูกโซ่มาคอฟให้เคลื่อนไประหว่างตัวแบบด้วยค่าความน่าจะเป็นภายหลังของตัวแบบนั้น ๆ ซึ่งคำนวณค่าความน่าจะเป็นภายหลังได้จากสัดส่วนของจำนวนครั้งที่ตัวแบบอยู่ในลูกโซ่ การค้นหาตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ จะพิจารณาค่าความน่าจะเป็นภายหลังสำหรับตัวแปรอิสระใด ๆ เป็นเกณฑ์ในการคัดเลือก กล่าวคือ เมื่อจบสิ้นกระบวนการ MCMC เราจะทำการเลือกตัวแปรอิสระใด ๆ ที่มีค่าสัดส่วนของจำนวนครั้งที่ตัวแปรอิสระดังกล่าวอยู่ในตัวแบบมากกว่า 1/2

หากกำหนดให้ สมการ (1.1) เป็นสมการของตัวแบบเต็ม เราจะได้สมการของตัวแบบย่อย (submodels) ที่นำมาพิจารณาเป็นดังสมการ 2.1

(2.15) $$M_{\underset{\sim}{l}} \; : \; \underset{\sim}{y} = X_{\underset{\sim}{l}} \underset{\sim}{\beta}_{\underset{\sim}{l}} + \underset{\sim}{\varepsilon}$$

เมื่อ $\underset{\sim}{l} = (l_1, l_2, \ldots, l_p)$ เป็นดัชนีของตัวแบบซึ่งบอกว่าตัวแปรอิสระใดควรอยู่ในตัวแบบ

หรือ $l_j = \begin{cases} 1, & X_j \text{ อยู่ในตัวแบบ} \\ 0, & \text{อื่น ๆ} \end{cases}$

ซึ่งสมมูลกับ

$$l_j = \begin{cases} 1, & \beta_j \neq 0 \\ 0, & \beta_j = 0 \end{cases}$$

และค่าพยากรณ์ที่ได้จะอยู่ในรูป

(2.16) $$\underset{\sim}{y}^* = X^* \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

เมื่อ $X^* = \left(\mathrm{x}_1^* , \ldots , \mathrm{x}_p^* \right)$ เป็นเวกเตอร์ของตัวแปรอิสระที่ใช้ในการพยากรณ์

ค่าพยากรณ์ที่ได้จะมีค่าความสูญเสียอันเกิดจากการพยากรณ์ $\underset{\sim}{y}^*$ ด้วย $\hat{\underset{\sim}{y}}^*$ เป็น

(2.17) $$L(\hat{y}^{*}, y^{*}) = (\hat{y}^{*} - y^{*})^2$$

นิยามที่ 1

ผลรวมของความน่าจะเป็นภายหลังสำหรับตัวแปรอิสระที่ j (posterior inclusion probability for variable j)

(2.18) $$p_j = \sum_{l: l_j=1} P\left(M_l \middle| y\right)$$

ดังนั้น ตัวแปรอิสระที่ j จะอยู่ในตัวแบบก็ต่อเมื่อ $p_j \geq \frac{1}{2}$ และจะได้ตัวแบบที่มีความน่าจะเป็นกลาง ๆ (median probability model: M_l^*) เป็นตัวแบบที่ประกอบด้วยตัวแปรอิสระที่มี p_j อย่างน้อยเท่ากับ 1/2 หรือ

$$l_j^* = \begin{cases} 1, & p_j \geq \frac{1}{2} \\ 0, & \text{อื่น ๆ} \end{cases}$$

2.4.2 การคัดเลือกตัวแบบที่เหมาะสมที่สุด

ภายหลังการค้นหาตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ เราจะทำการคัดเลือกตัวแบบที่เหมาะสมที่สุดและเกณฑ์ที่ใช้ในการคัดเลือกตัวแบบที่เหมาะสมที่สุดจากกลุ่มของตัวแบบที่มีความน่าจะเป็นภายหลังกลาง ๆ คือ ความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสอง (square error loss) โดยจะเลือกตัวแบบที่มีค่าดังกล่าวต่ำสุด ซึ่งรายละเอียดของการคัดเลือกตัวแบบที่เหมาะสมที่สุดมีดังนี้

สมมติว่า $Q = X'X$ เป็นเมทริกซ์ทแยงมุม เราจะได้ค่าเฉลี่ยภายหลัง $\tilde{\beta}_l$ (the posterior means) เป็นดังสมการ 2.19

(2.19) $$\tilde{\beta}_l = H_l' \tilde{\beta}$$

เมื่อ $\tilde{\beta}$ เป็นค่าเฉลี่ยภายหลังของตัวแบบเต็ม

และ H_l เป็นเมทริกซ์ขนาด $p \times p_l$ ของคู่ลำดับ (j, k) ที่มีค่าเป็น 1ถ้า $l_j = 1$ และ $k = \sum_{r-1}^{i} l_r$ หรือมีค่า เป็น 0 ในกรณีอื่น ๆ

สำหรับการวิจัยครั้งนี้จะกำหนดหลักเกณฑ์สังยุคแบบปกติ (Independent conjugate normal priors) กล่าวคือ ในตัวแบบเต็มสมมติให้ $\pi\left(\beta \middle| \sigma\right) \sim N_p\left(\mu, \sigma^2 \Lambda\right)$ ซึ่งเป็น

การแจกแจงแบบปกติ p ตัวแปร มีค่าเฉลี่ยเท่ากับ $\underset{\sim}{\mu}$ ความแปรปรวนแนวทแยงเท่ากับ $\sigma^2\Lambda$ โดยกำหนดให้ $\Lambda = n(X'X)^{-1}$ หรือ $\Lambda = c(X'X)^{-1}$ ด้วยการเลือกค่า c ที่เหมาะสมซึ่งทำให้หลักเกณฑ์ของ $\underset{\sim}{\beta}_{\underset{\sim}{l}}$ ในตัวแบบย่อยเป็น $N_{p_{\underset{\sim}{l}}}\left(\mathrm{H}'_{\underset{\sim}{l}}\underset{\sim}{\mu}, \sigma^2 \mathrm{H}'_{\underset{\sim}{l}}\Lambda \mathrm{H}_{\underset{\sim}{l}} \right)$ เมื่อ $p_{\underset{\sim}{l}}$ คือจำนวนมิติของ $\underset{\sim}{\beta}_{\underset{\sim}{l}}$ ในตัวแบบย่อยและจากการใช้หลักเกณฑ์ดังกล่าวจะทำให้สมการ 2.19 เป็นจริงสำหรับทุกหลักเกณฑ์ใด ๆ ของ σ^2

ดังนั้น ถ้าQ เป็นเมทริกซ์ทแยงมุมซึ่งประกอบด้วยค่าในแนวทแยง $q_j > 0$ และสมการ (2.19) เป็นจริงแล้วเราสามารถหาค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสองของตัวแบบย่อยใด ๆ ได้ดังสมการ

(2.20) $$R(M_{\underset{\sim}{l}}) = \sum_{j=1}^{p} \tilde{\beta}_j^2 q_j \ (l_j - p_j)^2$$

เมื่อ p_j เป็นผลรวมของความน่าจะเป็นภายหลังสำหรับตัวแปรอิสระที่ j ดังสมการ (2.18)

2.5 วิธีการถดถอยแบบขั้นบันได (Stepwise Regression method (SR))

วิธีการนี้เป็นวิธีการคัดเลือกตัวแปรอิสระแบบเลือกเข้า โดยในแต่ละขั้นตอนของการเลือกตัวแปรอิสระเข้าสู่สมการถดถอยจะตรวจสอบตัวแปรอิสระที่มีในสมการถดถอยก่อนที่ตัวแปรอิสระตัวล่าสุดจะเพิ่มเข้าในสมการโดยวิธีการกำจัดตัวแปรอิสระแบบถอยหลังเสียก่อนซึ่งสามารถแสดงขั้นตอนของวิธีการถดถอยแบบขั้นบันไดได้ดังนี้

1) สร้างสมการถดถอยโดยเริ่มจากไม่มีตัวแปรอิสระอยู่ในสมการ

2) เลือกตัวแปรอิสระที่มีสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) กับตัวแปรตาม (y) สูงสุดเข้าสู่สมการเป็นตัวแปรแรก และทดสอบว่าตัวแปรอิสระดังกล่าวมีนัยสำคัญหรือไม่ หากตัวแปรอิสระดังกล่าวไม่มีนัยสำคัญ จะได้สมการถดถอยที่เหมาะสมคือ $y = \bar{y}$ แต่ถ้าตัวแปรอิสระดังกล่าวมีนัยสำคัญจะทำในขั้นตอนต่อไป

3) คำนวณค่าสถิติเอฟบางส่วนสำหรับตัวแปรอิสระที่ไม่อยู่ในสมการ

4) ค่าสถิติเอฟบางส่วนที่มากที่สุด (F_U) จะถูกเปรียบเทียบกับค่า F ณ ระดับนัยสำคัญที่กำหนด (F_O)

- ถ้า $F_U < F_O$ จะไม่นำตัวแปรอิสระที่มีค่าสถิติเอฟบางส่วนมากที่สุดนั้นเข้าสู่สมการถดถอย

- ถ้า $F_u > F_o$ จะนำตัวแปรอิสระที่มีค่าสถิติเอฟบางส่วนมากที่สุดนั้นเข้าสู่สมการถดถอย และคำนวณค่าสัมประสิทธิ์ของสมการถดถอยเมื่อนำตัวแปรอิสระนั้นเข้าสู่สมการด้วยวิธีกำลังสองน้อยสุด

5) คำนวณค่าสถิติเอฟบางส่วนของทุกตัวแปรอิสระที่อยู่ในสมการถดถอย

6) ค่าสถิติเอฟบางส่วนที่น้อยที่สุด (F_L) จะถูกเปรียบเทียบกับค่า F ณ ระดับนัยสำคัญที่กำหนด (F_o)

- ถ้า $F_L < F_o$ จะตัดตัวแปรอิสระที่มีค่าสถิติเอฟบางส่วนน้อยที่สุดนั้นออกจากสมการถดถอย และคำนวณค่าสัมประสิทธิ์ของสมการถดถอย เมื่อตัดตัวแปรอิสระนั้นแล้ว
- ถ้า $F_L > F_o$ จะไม่ตัดตัวแปรอิสระที่มีค่าสถิติที่มีค่าสถิติเอฟบางส่วนน้อยที่สุดในนั้นออกจากสมการถดถอย

7) ถ้าไม่มีตัวแปรใดเข้าและออกจากสมการถดถอยแล้ว จะได้สมการถดถอยที่เหมาะสมแต่ถ้ายังมีตัวแปรอิสระใดที่เป็นไปตามเงื่อนไขของการเข้าหรือออกจากสมการ ให้กลับไปทำในขั้นตอนที่ 3

จากขั้นตอนของวิธีการถดถอยแบบขั้นบันไดจะเห็นได้ว่าต้องอาศัยวิธีการทดสอบเอฟบางส่วน และการหาสัมประสิทธิ์สหสัมพันธ์ ซึ่งได้กล่าวถึงรายละเอียดไว้ในภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

3.1 การหาข้อสรุปโดยใช้การจำลองของข้อมูลด้วยวิธีมอนติคาร์โล [1] (Simulation by Monte Carlo Method)

วิธีมอนติคาร์โลเป็นเทคนิคในการจำลองตัวแบบทางคณิตศาสตร์ ซึ่งหลักการเบื้องต้นนั้นต้องจำลองตัวเลขสุ่ม (random number) เพื่อช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา โดยในการศึกษาวิจัยนี้จะใช้เทคนิคมอนติคาร์โลในการสร้างข้อมูลที่มีสภาพการแจกแจงตามที่ต้องการศึกษา ขั้นตอนที่สำคัญของการจำลองข้อมูลด้วยวิธีมอนติคาร์โลมี 3 ขั้นตอนคือ

ขั้นตอนที่ 1 การสร้างตัวเลขสุ่ม

การสร้างตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีมอนติคาร์โล ทั้งนี้ก็เพราะว่าหลักการของมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหา ลักษณะของตัวเลขสุ่มที่ดีจะมีการแจกแจงสม่ำเสมอ (Uniform Distribution) ในช่วง $[0,1]$ และเป็นอิสระซึ่งกันและกัน

ขั้นตอนที่ 2 การประยุกต์ของปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม

ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจใช้ตัวเลขสุ่มโดยตรง แต่บางปัญหาอาจใช้ตัวเลขสุ่มเพียงบางขั้นตอนของปัญหาเท่านั้น

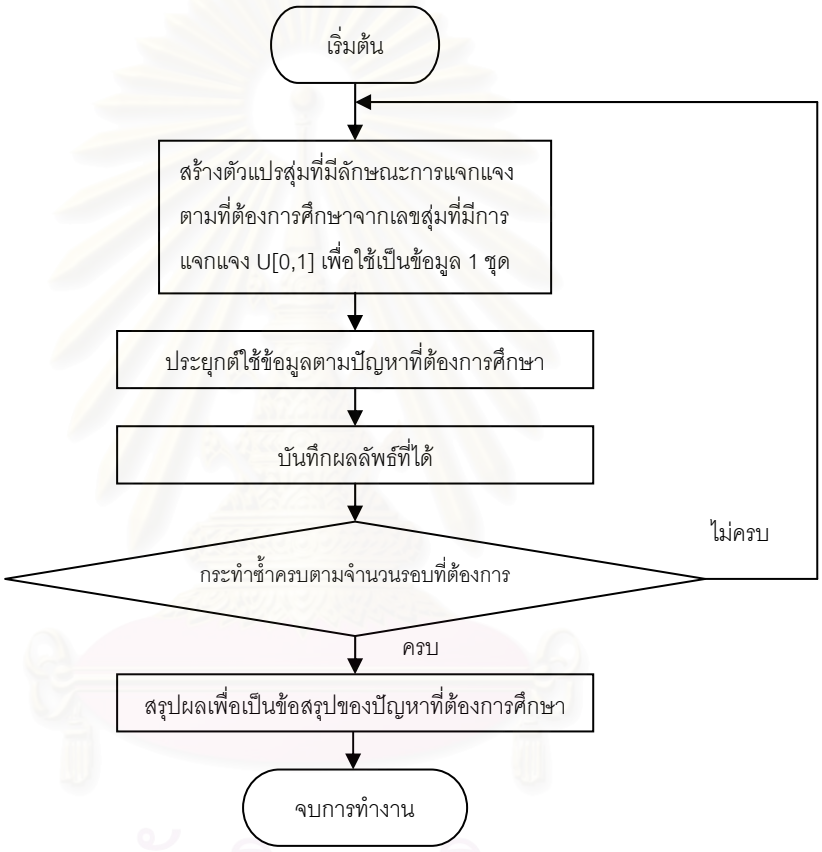
ขั้นตอนที่ 3 การทดลองกระทำ

เมื่อประยุกต์ปัญหาเพื่อใช้กับตัวเลขสุ่มได้แล้วขั้นตอนต่อไปก็คือการทดลองโดยใช้กระบวนการสุ่ม (random process) มากระทำในลักษณะที่ซ้ำ ๆ กัน (replication) เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

---

[1] จะเด็จ สวรรค์ตรานนท์,"การเปรียบเทียบวิธีที่ใช้สำหรับการคัดเลือกสมการถดถอยที่ดีที่สุด,"(วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย,2531),หน้า36.

จากหลักการของวิธีมอนติคาร์โลจะเห็นได้ว่า การใช้เลขสุ่มเพื่อเป็นพื้นฐานในการหาคำตอบของปัญหา เป็นวิธีการที่จะนำไปสู่แนวคิดทางทฤษฎีที่เกี่ยวข้องกับการคำนวณ โดยเฉพาะทฤษฎีความน่าจะเป็น ที่จะนำไปสู่การอ้างอิงผลสรุปในสถานการณ์ของข้อมูลจริง เพราะไม่มีผลกระทบจากปัจจัยอื่น ๆ เข้ามาเกี่ยวข้อง ในการทดลองเมื่อกระทำซ้ำ ๆ กันเป็นจำนวนมากแล้วความคลาดเคลื่อนอย่างสุ่มที่เกิดขึ้นในการวิเคราะห์หาค่าต่าง ๆ ในแต่ละครั้งจะหมดไป (counter balance) จากขั้นตอนของวิธีมอนติคาร์โล สามารถเขียนผังงานได้ดังนี้



รูปที่ 3.1 แสดงผังงานสำหรับขั้นตอนของวิธีมอนติคาร์โล

3.2 แผนการทดลอง

ผู้วิจัยได้กำหนดสถานการณ์ต่าง ๆ สำหรับการวิจัยครั้งนี้ไว้ดังนี้

1) เลือกตัวอย่างสุ่มเพื่อใช้เป็นค่าความคลาดเคลื่อนสุ่มจากประชากรที่มีการแจกแจงเดียวกันโดยการวิจัยครั้งนี้สนใจศึกษาเฉพาะตัวอย่างสุ่มที่มาจากการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ

2) กำหนดขนาดตัวอย่างที่ศึกษา คือ 15 30 50 และ 100 ซึ่งในกรณีขนาดตัวอย่างเท่ากับ 15 ผู้วิจัยได้ทำการศึกษาเฉพาะกรณีจำนวนตัวแปรอิสระเท่ากับ 3 และ 5 เท่านั้น และในกรณีขนาดตัวอย่างเท่ากับ 30 ผู้วิจัยได้ทำการศึกษาเฉพาะกรณีจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 และ12 เท่านั้น

3) กำหนดจำนวนตัวแปรอิสระที่ศึกษา คือ 3 5 8 10 12 และ15 ตามลำดับ

4) ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM มี 4 ระดับ คือ (σ_β / τ , c) มีค่าเป็น (1,5) (1,10) (10,100) และ (10,500)

3.3 ขั้นตอนในการศึกษาวิจัย

ขั้นตอนในการดำเนินการวิจัยมีดังนี้

1) กำหนดลักษณะการแจกแจงของค่าคลาดเคลื่อน ขนาดตัวอย่าง จำนวนตัวแปรอิสระ และค่าคงที่สำหรับวิธี BMA_{SVT} และวิธี OPM

2) สร้างข้อมูลตัวแปรตามจากตัวแปรอิสระและความคลาดเคลื่อนที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรอิสระ

3) สร้างตัวแบบโดยวิธีทั้ง 3 วิธี คือ

(1) วิธีการเฉลี่ยตัวแบบของเบส์ (BMA)โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (BMA_{SVT})

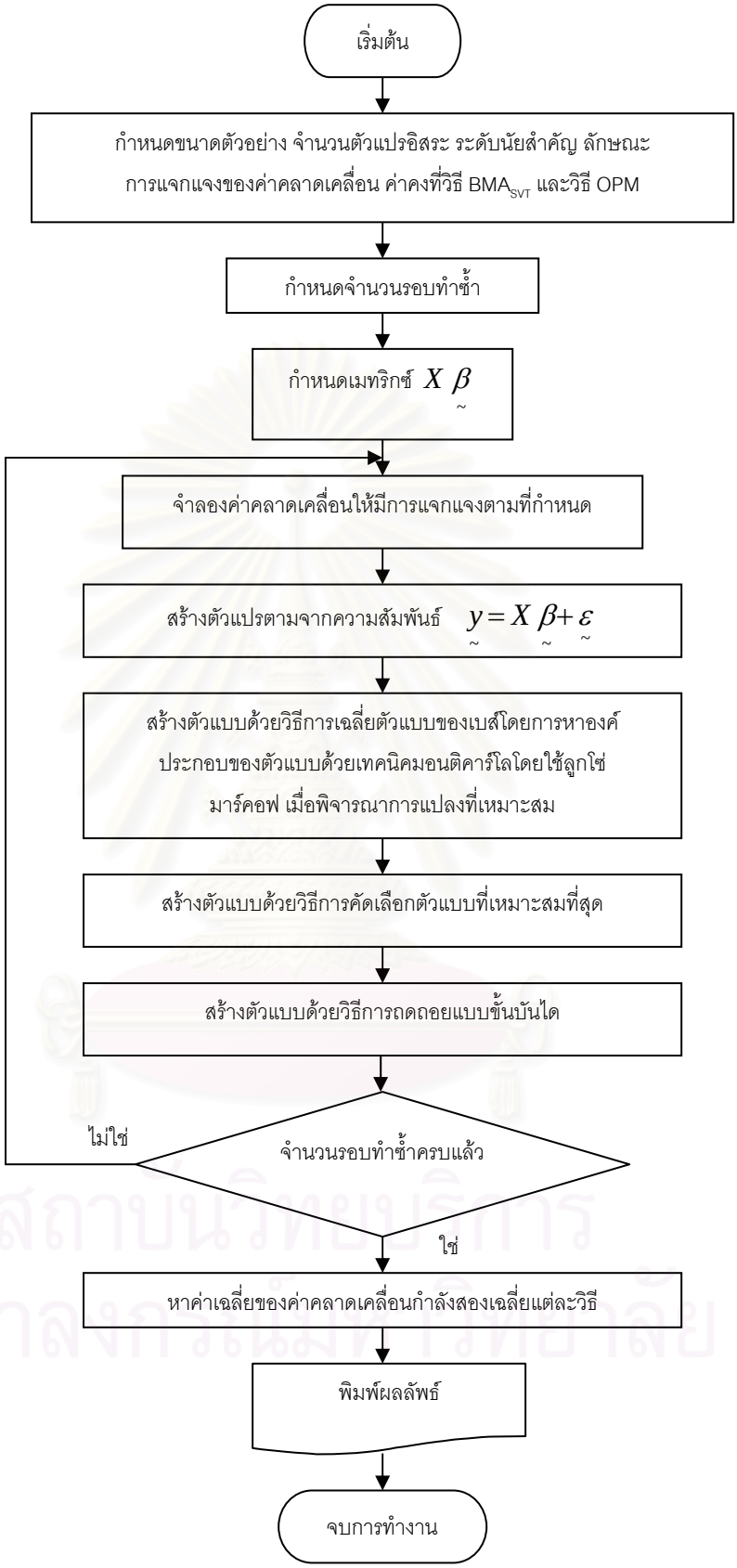
(2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (OPM)

(3) วิธีการถดถอยแบบขั้นบันได (SR)

4) คำนวณค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากการทำซ้ำจำนวน 500 รอบ ของตัวแบบที่ได้จากวิธีการคัดเลือกตัวแบบการถดถอยทั้ง 3 วิธี

5) ทำการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากการทำซ้ำจำนวน500 รอบ ของตัวแบบที่ได้จากวิธีการคัดเลือกตัวแบบการถดถอยทั้ง3วิธี โดยสรุปผลในรูปของตารางและรูปภาพ

ผังงานแสดงขั้นตอนในการดำเนินการวิจัย ดังแสดงรูปที่ 3.2



รูปที่ 3.2 แสดงผังงานของขั้นตอนในการดำเนินการวิจัย

สำหรับรายละเอียดของแต่ละขั้นตอนเป็นดังนี้

ขั้นตอนที่ 1

กำหนดลักษณะการแจกแจงของค่าคลาดเคลื่อน ขนาดตัวอย่าง (n) จำนวนตัวแปรอิสระ ระดับนัยสำคัญ ค่าคงที่สำหรับวิธี BMA_{SVT} และวิธี OPM โดยจะกำหนดตามแผนการทดลองที่กล่าวมาแล้วข้างต้น

ขั้นตอนที่ 2

การสร้างข้อมูลในการวิจัยครั้งนี้ประกอบด้วย ข้อมูลตัวแปรอิสระ ข้อมูลค่าคลาดเคลื่อนที่มีลักษณะการแจกแจงที่ต้องการศึกษา ข้อมูลตัวแปรตาม โดยมีรายละเอียดดังนี้

1) <u>ข้อมูลตัวแปรอิสระ</u>

เนื่องจากการวิจัยครั้งนี้ศึกษาเฉพาะการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณในกรณีที่ตัวแปรเป็นไปตามข้อตกลงเบื้องต้น กล่าวคือ ตัวแปรอิสระเป็นค่าคงที่ ไม่มีพหุสัมพันธ์กัน กำลังสูงสุดของตัวแปรอิสระเป็น 1 และไม่มีอันตรกิริยาระหว่างตัวแปรอิสระ นั่นคือ การวิจัยครั้งนี้ศึกษาเฉพาะการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณในรูปทั่วไป ไม่รวมถึงตัวแบบความถดถอยพหุนาม ดังนั้นในการสร้างข้อมูลตัวแปรอิสระจากการแจกแจงพหุแบบปกติ (multivariate normal distribution) ที่มีเวกเตอร์ค่าเฉลี่ยเท่ากับ 0 และเมทริกซ์ความแปรปรวนร่วมเป็นเมทริกซ์เอกลักษณ์ (identity matrix) ซึ่งจะทำให้ได้ตัวแปรอิสระแต่ละตัวที่มีค่าเฉลี่ยเป็น 0 ความแปรปรวนเป็น 1 และตัวแปรอิสระทุกตัวที่ได้จะไม่มีพหุสัมพันธ์กัน โดยสร้างจากฟังก์ชัน rmvnorm(n , mean , cov , sd ,rho , d) ในโปรแกรม S-plus 2000 ซึ่งรายละเอียดสามารถดูได้จากภาคผนวก จากนั้นนำข้อมูลตัวแปรอิสระที่ได้เก็บไว้ในแฟ้มข้อมูลสำหรับกรณีจำนวนตัวแปรอิสระในระดับต่าง ๆ เพื่อเรียกใช้ในการประมวลผลต่อไป

2) <u>ข้อมูลค่าคลาดเคลื่อน</u>

การวิจัยครั้งนี้สนใจศึกษาเฉพาะกรณีที่ค่าคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนร่วมเท่ากับ σ^2 เหมือนกันและเป็นอิสระซึ่งกันและกัน โดยสร้างจากฟังก์ชัน rnorm(n , mean , sd) ในโปรแกรม S-plus 2000 ซึ่งรายละเอียดสามารถดูได้จากภาคผนวก

3) <u>ข้อมูลตัวแปรตาม</u>

เมื่อมีข้อมูลตัวแปรอิสระ และข้อมูลค่าคลาดเคลื่อนที่มีลักษณะการแจกแจงตาม

ที่ต้องการศึกษาแล้ว จากนั้นจะทำการสร้างข้อมูลตัวแปรตามให้มีความสัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรอิสระ ซึ่งมีรูปแบบความสัมพันธ์ดังนี้

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

เมื่อ $\underset{\sim}{y}$ เป็นเวกเตอร์ของตัวแปรตาม

X เป็นเมทริกซ์ของตัวแปรอิสระ

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์ที่กำหนดขึ้น ซึ่งในการวิจัยครั้งนี้กำหนดให้ $\underset{\sim}{\beta}' = (1,1,...,1)_{1\times p}$ นั่นคือ กำหนดให้ตัวแปรอิสระทุกตัวมีอิทธิพลเท่ากันในตัวแบบเริ่มต้น

และ $\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ของค่าคลาดเคลื่อนที่กำหนด ซึ่ง $\underset{\sim}{\varepsilon} \sim N(0, \sigma^2 I_n)$

ขั้นตอนที่ 3

กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูป (full model) ในการคัดเลือกตัวแบบที่เหมาะสม

ขั้นตอนที่ 4

หลังจากที่มีข้อมูลพร้อมแล้วทำการสร้างตัวแบบที่เหมาะสมจากทั้ง 3 วิธี ดังต่อไปนี้

1) วิธีการเฉลี่ยตัวแบบของเบส์ (BMA)โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ (MC^3) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (BMA_{SVT}) ใช้โปรแกรมย่อย SVT(Y , X , num.its , M0.var , M0.out , outs.list , PI , K , IP , CC , n , numx , sd)

2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (OPM) ใช้โปรแกรมย่อย OPM(Y , X , num.its , M0.var , M0.out , outs.list , PI , K , IP , CC , n , numx , sd , p)

3) วิธีการถดถอยแบบขั้นบันได (SR)ใช้โปรแกรมย่อย Stepwise(X , Y , intercept=T , Tolerance=1.e-07, method="ex" , nbest=3)

ซึ่งโปรแกรมย่อยทั้ง 3 โปรแกรมเป็นโปรแกรมย่อยที่สร้างขึ้นโดยใช้โปรแกรมS-plus 2000 สำหรับรายละเอียดของโปรแกรมสามารถดูได้จากภาคผนวก

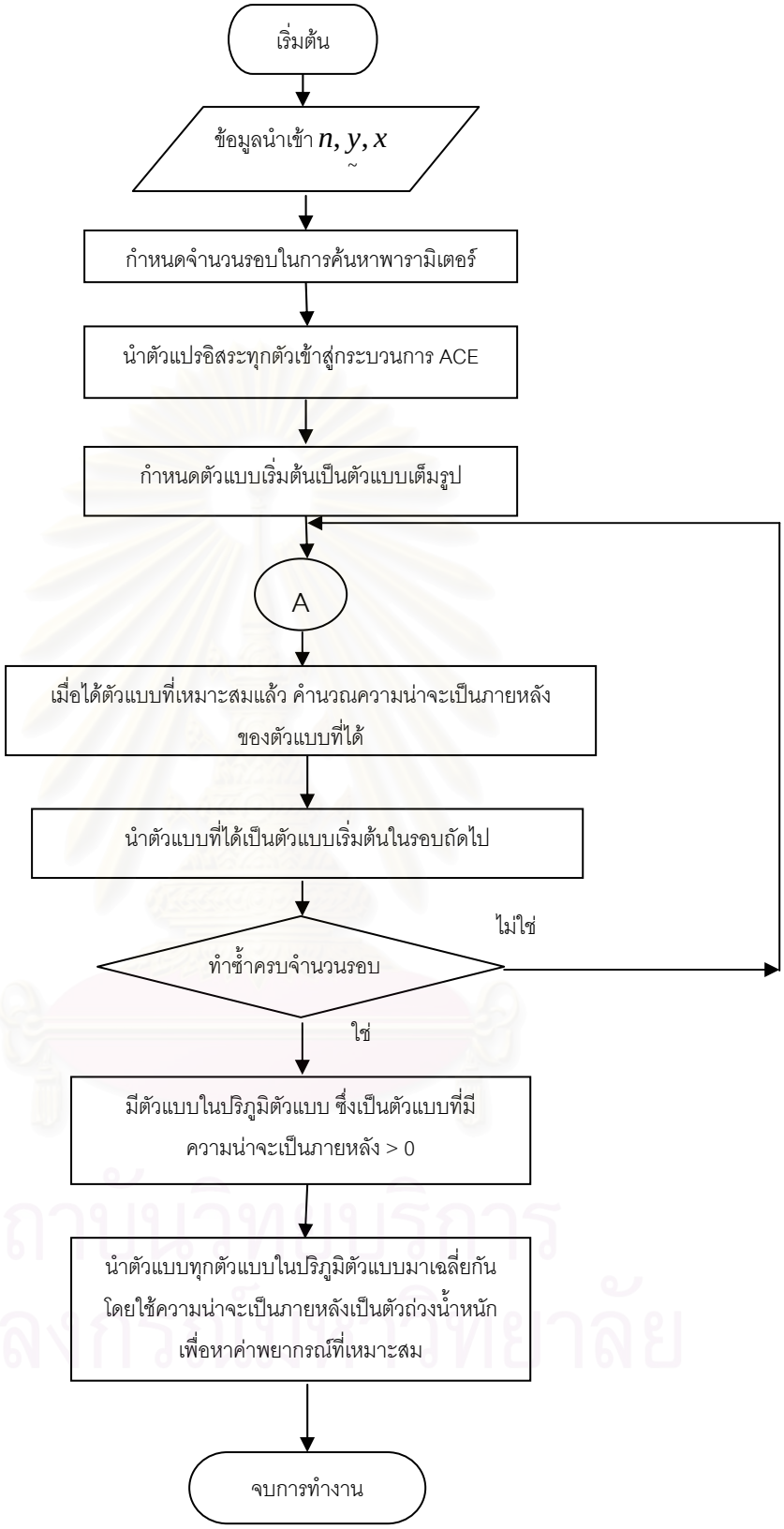
วิธีการสร้างตัวแบบในแต่ละวิธีมีรายละเอียดดังนี้

1) วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (BMA_{SVT})

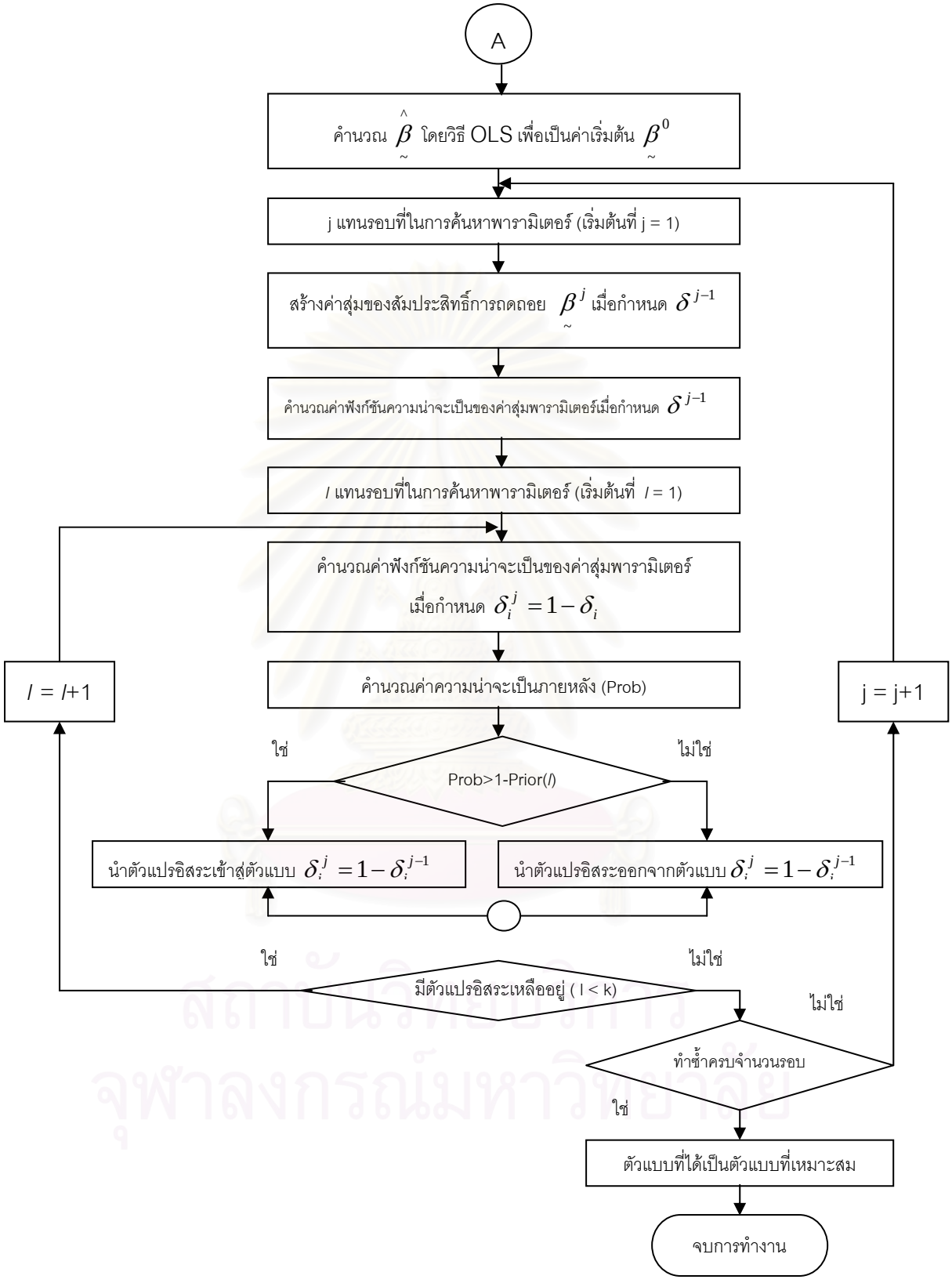
วิธีการนี้เริ่มต้นจากพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระโดยนำเข้าสู่กระบวนการ ACE (Alternating Conditional Expectation Algorithm) จากนั้นนำตัวแปรอิสระทุกตัวและตัวแปรอิสระที่ควรทำการแปลงเข้าสู่กระบวนการ MC^3 พร้อมกันโดยกำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูป คือ มีตัวแปรอิสระครบทุกตัวแปรในสมการถดถอย กำหนดจำนวนรอบทำซ้ำและค่าคงที่ในการค้นหาพารามิเตอร์ จากนั้นทำการวนซ้ำเพื่อคำนวณความน่าจะเป็นภายหลังของแต่ละตัวแบบ และทำการคัดเลือกตัวแปรด้วยวิธีการเช่นเดียวกับวิธีการคัดเลือกตัวแปรของเบส์ กล่าวคือ เริ่มต้นด้วยการกำหนดความน่าจะเป็นก่อน และค่าคงที่ของการคัดเลือกตัวแปรแบบเบส์ กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูปและคำนวณค่าประมาณของสัมประสิทธิ์การถดถอยของตัวแบบเต็มรูปโดยวิธีกำลังสองน้อยสุด เพื่อเป็นค่าเริ่มต้นที่ใช้ในการสุ่มพารามิเตอร์ $\underset{\sim}{\beta}$ หลังจากนั้นทำการสุ่มค่าพารามิเตอร์ $\underset{\sim}{\beta}$ จากการแจกแจงแบบคู่สังยุคปกติ (conjugate normal distribution) และคำนวณหาฟังก์ชันความน่าจะเป็นของค่าที่สุ่มได้ เพื่อนำไปหาค่าความน่าจะเป็นแบบมีเงื่อนไขภายหลังจากทราบข้อมูลแล้ว

เมื่อทำซ้ำครบจำนวนรอบที่กำหนดไว้แล้วจะได้ตัวแบบที่มีค่าความน่าจะเป็นภายหลังมากกว่าศูนย์อยู่ในปริภูมิตัวแบบ นำตัวแบบทุกตัวแบบในปริภูมิตัวแบบมาเฉลี่ยกันโดยใช้ความน่าจะเป็นภายหลังของแต่ละตัวแบบเป็นตัวถ่วงน้ำหนัก เพื่อหาค่าพยากรณ์ที่เหมาะสม

ผังงานแสดงขั้นตอนของวิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ ดังแสดงในรูปที่ 3.3



รูปที่ 3.3 แสดงผังงานสำหรับขั้นตอนของวิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



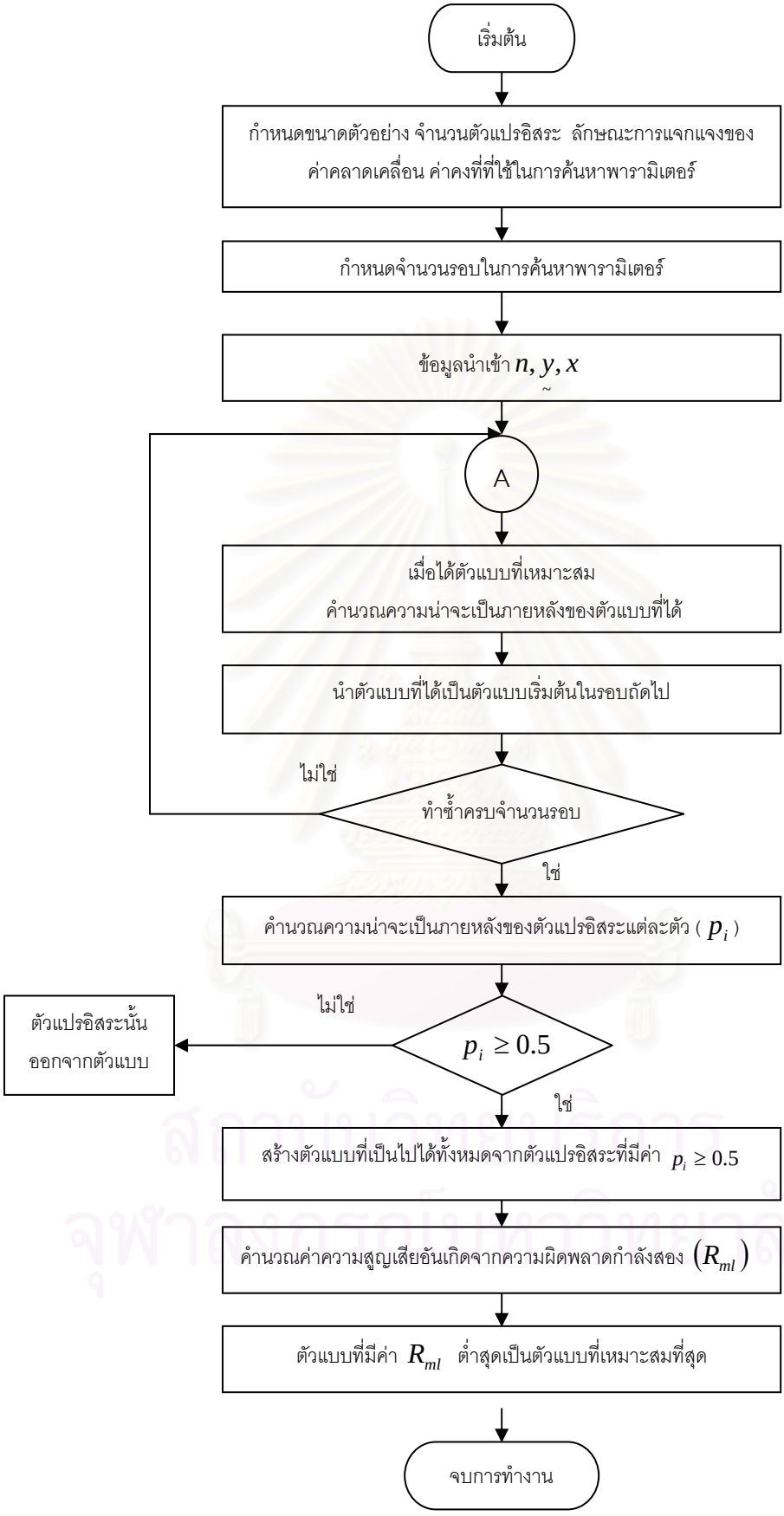
รูปที่ 3.3 แสดงผังงานสำหรับขั้นตอนของวิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (ต่อ)

2) **วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด** (OPM)

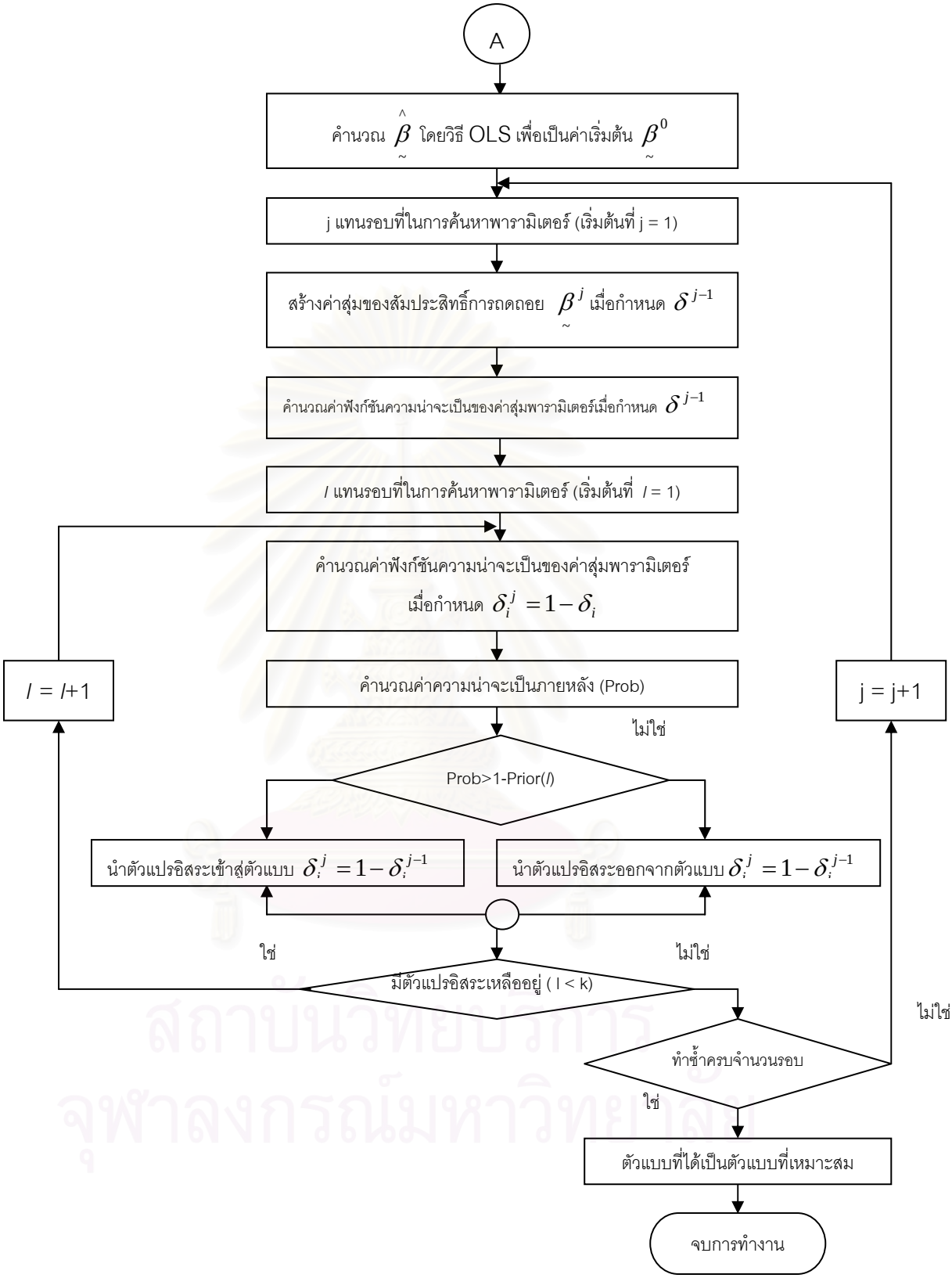
วิธีการนี้เริ่มต้นจากกำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูป กำหนดจำนวนรอบทำซ้ำและค่าคงที่ในการค้นหาพารามิเตอร์ จากนั้นทำการวนซ้ำเพื่อคำนวณความน่าจะเป็นภายหลังของแต่ละตัวแบบ หลังจากนั้นทำการคัดเลือกตัวแปรอิสระที่มีความน่าจะเป็นภายหลังมากกว่าหรือเท่ากับ 0.5 และทำการกำหนดตัวแบบที่เป็นไปได้ทั้งหมดจากตัวแปรอิสระดังกล่าวซึ่งตัวแบบที่สร้างเป็นตัวแบบติดกลุ่ม จากนั้นทำการคัดเลือกตัวแบบที่เหมาะสมที่สุดจากตัวแบบที่มีค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสองของตัวแบบต่ำสุด (square error loss)

ผังงานแสดงขั้นตอนของวิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด ดังแสดงในรูปที่ 3.4

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.4 แสดงผังงานสำหรับขั้นตอนของวิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด



รูปที่ 3.4 แสดงผังงานสำหรับขั้นตอนของวิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (ต่อ)

3) วิธีการถดถอยแบบขั้นบันได (SR)[2]

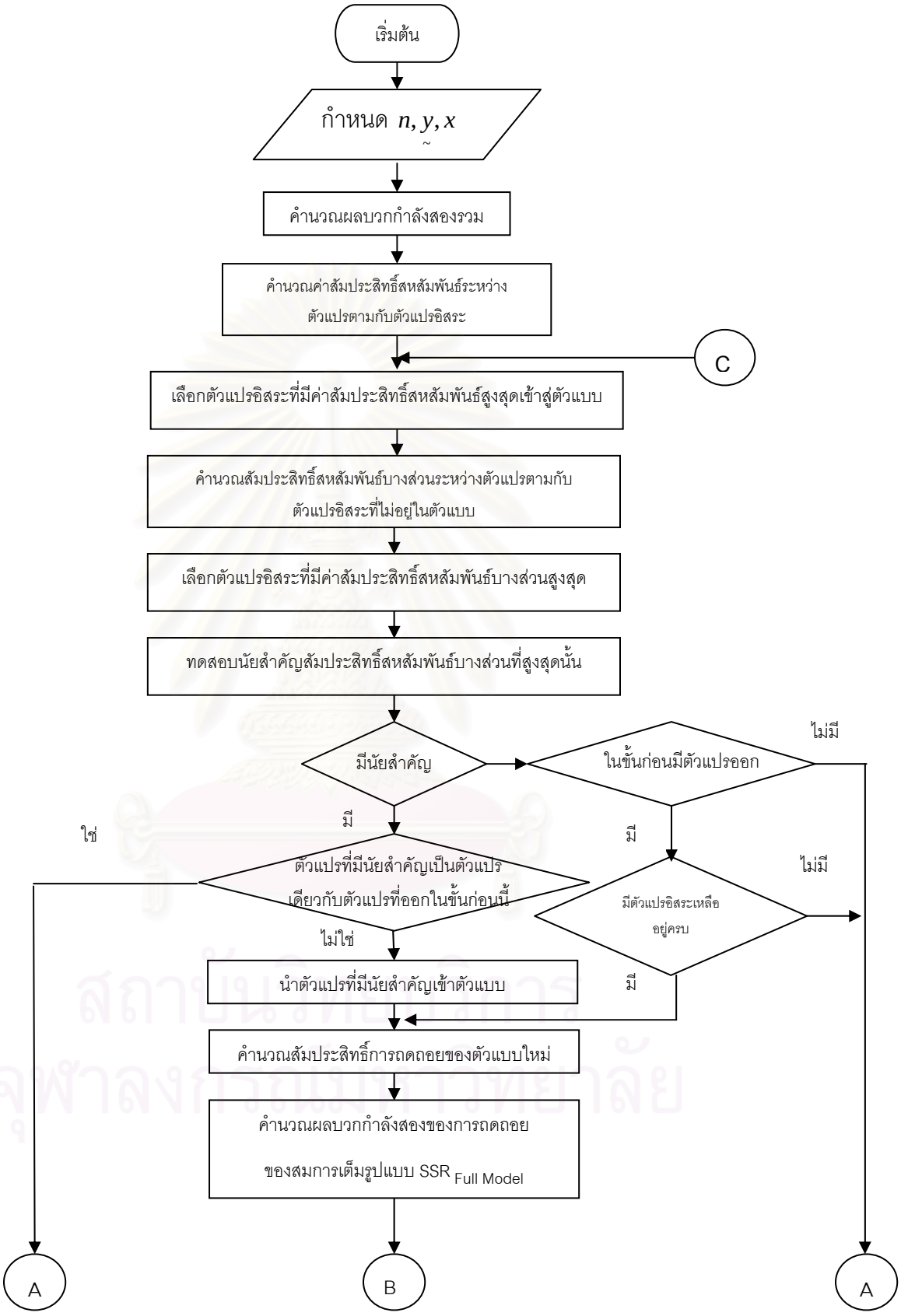
วิธีการนี้จะเริ่มจากการคัดเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (r_{xy}) มากที่สุดเข้าสู่สมการถดถอย (X_i) เข้าสู่สมการถดถอย โดยที่ $r_{xy}^2 = 1 - \frac{SSE}{SST}$ จากนั้นจะหาสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรตามกับตัวแปรอิสระแต่ละตัวที่ยังไม่อยู่ในสมการถดถอยโดยถือว่าได้รวมตัวแปรอิสระ X_i ไว้ในตัวแบบแล้ว และเลือกตัวแปรอิสระที่มีค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนสูงที่สุดและมีนัยสำคัญเข้าสู่ตัวแบบ ขั้นตอนต่อไปเป็นการคัดเลือกตัวแปรอิสระออก โดยจะตัดตัวแปรอิสระที่มีค่าสถิติเอฟบางส่วนน้อยที่สุด และไม่มีนัยสำคัญออกจากตัวแบบ จะทำซ้ำการคัดเลือกตัวแปรอิสระเข้าและออกจากตัวแบบจนกระทั่งเงื่อนไขการหยุดข้อใดข้อหนึ่งต่อไปนี้เป็นจริง

1) ไม่มีตัวแปรอิสระที่สามารถเข้าหรือออกจากตัวแบบในขั้นถัดกัน
2) ตัวแปรอิสระที่เข้าและออกจากตัวแบบในขั้นถัดกันเป็นตัวแปรเดียวกัน
3) ในขั้นตอนการออกไม่มีตัวแปรอิสระเหลืออยู่ในตัวแบบ
4) ในขั้นตอนการเข้าไม่เหลือตัวแปรอิสระที่ไม่อยู่ในสมการถดถอย

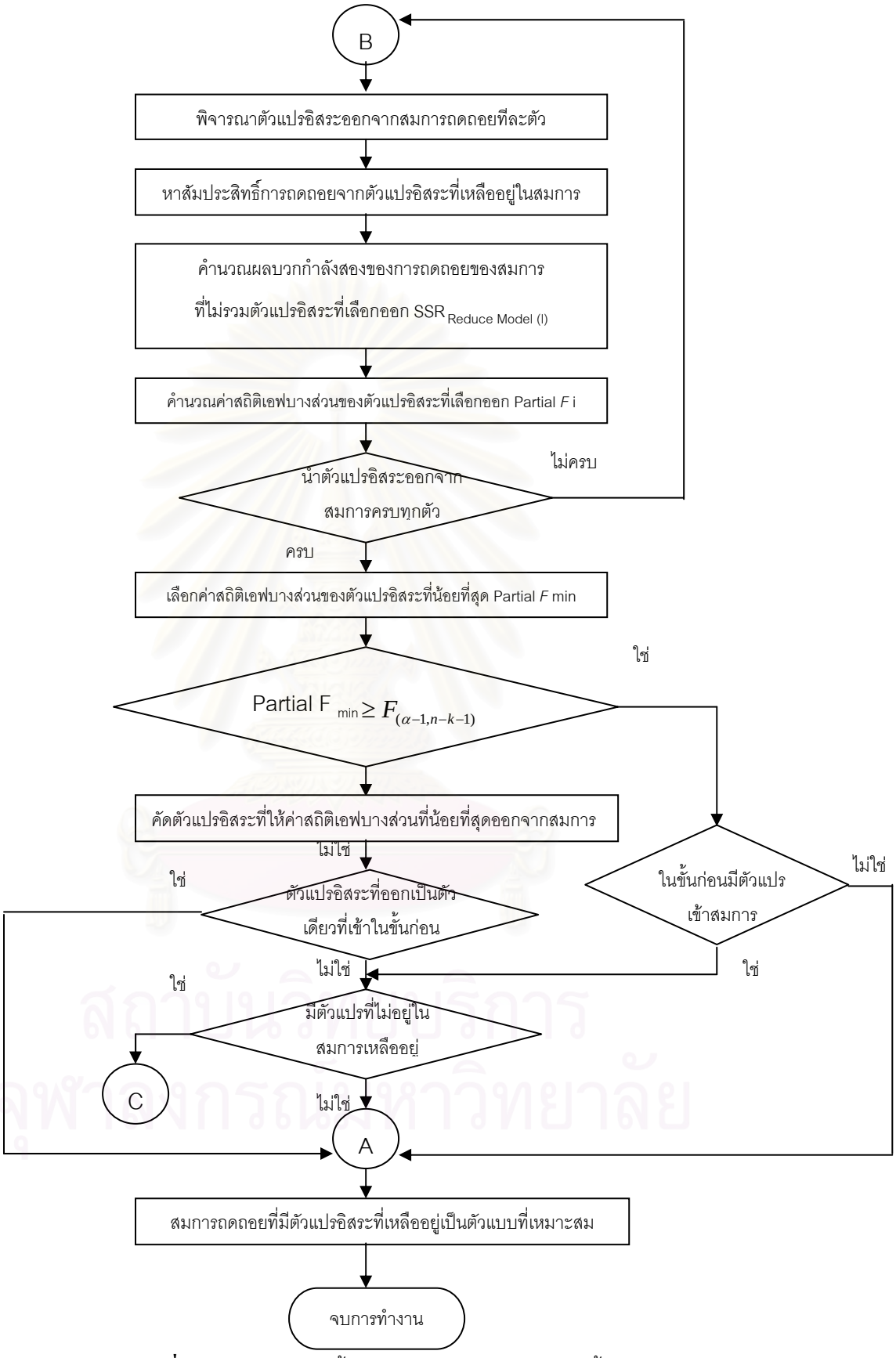
ผังงานแสดงขั้นตอนของวิธีการถดถอยแบบขั้นบันได ดังแสดงในรูปที่ 3.5

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

[2] นิทัสน์ สุขสุวรรณ, "การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางของเบส์ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ" (วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2545), หน้า 61



รูปที่ 3.5 แสดงผังงานขั้นตอนของวิธีการถดถอยแบบขั้นบันได



รูปที่ 3.5 แสดงผังงานขั้นตอนของวิธีการถดถอยแบบขั้นบันได (ต่อ)

ขั้นตอนที่ 5

คำนวณหาค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) จากการทำซ้ำจำนวน 500 รอบ ของทั้ง 5 วิธีเพื่อใช้เป็นเกณฑ์ในการตัดสินใจ และคำนวณค่าส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อนกำลังสองเฉลี่ย รวมถึงค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RDAMSE) ของทั้ง 3 วิธีดังกล่าวข้างต้น เพื่อเป็นเกณฑ์ประกอบการตัดสินใจ

ขั้นตอนที่ 6

ทำการเปรียบเทียบค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) สำหรับตัวแบบที่ได้จากการคัดเลือกตัวแบบการถดถอยทั้ง 3 วิธี โดยสรุปผลในรูปของตารางและรูปภาพเพื่อแสดงการเปรียบเทียบและศึกษาแนวโน้มของแต่ละวิธี

สำหรับรายละเอียดของโปรแกรมทั้งหมดสามารถดูได้จากภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการวิจัย

งานวิจัยครั้งนี้เป็นการศึกษาวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ เพื่อเปรียบเทียบเทียบความถูกต้องของการพยากรณ์ที่ได้จากการคัดเลือกตัวแบบการถดถอยซึ่งวิธีการคัดเลือกสมการถดถอยที่นำมาใช้ในการสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณมี 3 วิธีดังนี้

1) วิธีการเฉลี่ยตัวแบบของเบส์ (BMA)โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟ(MC^3) เมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (BMA_{SVT})

2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (OPM)

3) วิธีการถดถอยแบบขั้นบันได (SR)

โดยเกณฑ์ที่ใช้ในการตัดสินใจว่าวิธีการคัดเลือกตัวแบบการถดถอยวิธีใดจะมีความถูกต้องและแม่นยำมากที่สุดพิจารณาจากเกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Average of Mean Square Error (AMSE)) และเกณฑ์ที่ใช้ประกอบการตัดสินใจจะใช้เกณฑ์ค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Ratio of Different Average Mean Square Error (RDAMSE)) ซึ่งวิธีใดให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดจะเป็นวิธีที่ดีที่สุด

ขอบเขตของการวิจัย จะเป็นการศึกษาเฉพาะตัวแบบการถดถอยเชิงเส้นพหุคูณที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 ซึ่งใช้สถานการณ์ในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 กำหนดขนาดตัวอย่างเท่ากับ 15 30 50 และ 100 และค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM คือ (σ_β / τ , c) เท่ากับ (1,5) (1,10) (10,100) และ (10,500) ซึ่งการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์ในงานวิจัยครั้งนี้พิจารณาจากเกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย(AMSE) และเกณฑ์ที่ใช้ประกอบการตัดสินใจจะใช้เกณฑ์ค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย(RDAMSE) จากการทำซ้ำจำนวน 500 รอบในแต่ละสถานการณ์

การนำเสนอผลการวิจัย ผู้วิจัยได้ใช้สัญลักษณ์ต่าง ๆ ในตารางและการสรุปผล โดยมีความหมายดังนี้

1) σ แทน ค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่ม

2) n แทน ขนาดตัวอย่าง

ค่าที่แสดงในตารางผลการวิจัย สำหรับแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เลขเรียงลงมาได้แก่ ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ค่าส่วนเบี่ยงเบนมาตรฐานของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยซึ่งแสดงในวงเล็บ และค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RDAMSE) ตามลำดับ

การนำเสนอผลการวิจัย ของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติทั้ง 3 วิธีนั้นประกอบด้วยตารางและรูปภาพ โดยแบ่งการนำเสนอเป็น 4 ตอน (ตอนที่ 4.1 – 4.4) ซึ่งใช้ระดับค่าคงที่ที่กำหนดลักษณะการกระจายของการแจกแจงของพารามิเตอร์สัมประสิทธิ์การถดถอยสำหรับวิธี BMA_{SVT} และวิธี OPM ที่มี 4 ระดับเป็นเกณฑ์ในการแบ่งตอนของการนำเสนอผลการวิจัย โดยมีลำดับในการนำเสนอดังนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การนำเสนอผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติทั้ง 4 ตอน (ตอนที่ 4.1 – 4.4) กำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ

ตอนที่ 4.1 **ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (1,5)** ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตารางที่ 4.1 – 4.6 และรูปที่ 4.1 – 4.6

ตอนที่ 4.2 **ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (1,10)** ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตารางที่ 4.7–4.12 และรูปที่ 4.7 – 4.12

ตอนที่ 4.3 **ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (10,100)** ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตารางที่ 4.13 – 4.18 และรูปที่ 4.13 – 4.18

ตอนที่ 4.4 **ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (10,500)** ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์ แสดงในตารางที่ 4.19 – 4.24 และรูปที่ 4.19 – 4.24

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปแบบการนำเสนอผลการวิจัยในทุกตอนจะเริ่มจากผลการวิจัยที่ประกอบด้วยตารางและรูปภาพสำหรับแต่ละตอนซึ่งเมื่อเปลี่ยนระดับของตัวแปรอิสระจะมีการอธิบายผลการวิจัยที่ได้ และทำการอธิบายผลการวิจัยทั้งหมดของตอนนั้น ๆ เมื่อนำเสนอตารางและรูปภาพของผลการวิจัยในตอนนั้นครบแล้ว (6 ตาราง และ 6 รูปภาพ) ในตอนท้ายของบท หลังจากที่นำเสนอผลการวิจัยครบทุกตอนแล้วจะมีการอธิบายสรุปผลการวิจัยทั้งหมดอีกครั้งหนึ่ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4.1 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ (σ_β / τ , c) มีค่าเท่ากับ (1,5) ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตารางที่ 4.1 – 4.6 และรูปที่ 4.1 – 4.6

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.1 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

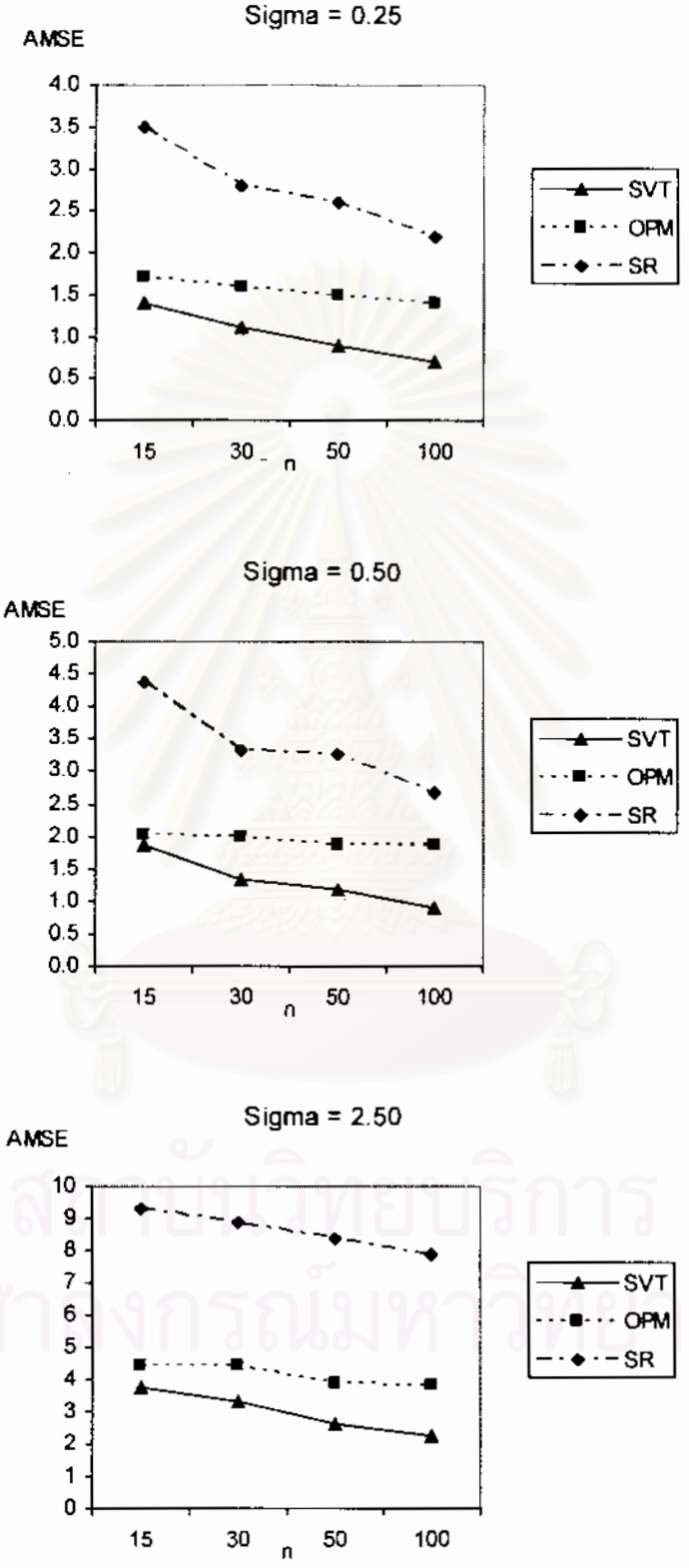
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	15	3.5397	1.6613	1.3862
		(0.3205)	(0.1384)	(0.1078)
		155.3528	19.8456	0.0000
	30	2.7610	1.6205	1.0793
		(0.2563)	(0.1278)	(0.0960)
		155.8140	50.1436	0.0000
	50	2.5781	1.5404	0.9035
		(0.2381)	(0.1192)	(0.0808)
		185.3459	70.4925	0.0000
	100	2.2167	1.4339	0.7423
		(0.2105)	(0.1171)	(0.0597)
		198.6259	93.1699	0.0000
0.50	15	4.3851	2.0204	1.8469
		(0.3808)	(0.2013)	(0.1503)
		137.4303	9.3941	0.0000
	30	3.3277	1.9846	1.2272
		(0.3168)	(0.1879)	(0.1086)
		171.1619	61.7177	0.0000
	50	3.2659	1.8623	1.1834
		(0.2694)	(0.1635)	(0.0926)
		175.9760	57.3686	0.0000
	100	2.6644	1.8450	0.9103
		(0.2328)	(0.1576)	(0.0735)
		192.6947	102.6804	0.0000
2.50	15	9.2892	4.4254	3.7546
		(0.9138)	(0.4594)	(0.3514)
		147.4085	17.8661	0.0000
	30	8.8458	4.4254	3.3321
		(0.7665)	(0.3890)	(0.3112)
		165.4722	32.8111	0.0000
	50	8.3773	3.8741	2.6227
		(0.7257)	(0.3478)	(0.2299)
		219.4151	47.7142	0.0000
	100	7.8477	3.7953	2.2511
		(0.7143)	(0.3201)	(0.1422)
		248.6162	68.5976	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัวเรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.1 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,5) (ตารางที่ 4.1 และรูปที่ 4.1) พบว่า วิธี BMA_{SVT} ให้ค่าAMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่าAMSE สูงสุด โดยวิธี OPM มีค่าAMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่าAMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่าAMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่าAMSE <u>แปรผันตาม</u>ค่า σ เพราะค่าAMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่าAMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น และค่า AMSE <u>แปรผกผัน</u>กับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธีBMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.50 วิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 50 ทั้งนี้อาจเป็นเพราะวิธี BMA_{SVT}ให้ค่า AMSE ลดลงเพียงเล็กน้อย และเมื่อ σ มีค่าเพิ่มขึ้นจะส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มขึ้น ดังนั้นค่า RDAMSE จึงมีแนวโน้มเพิ่มขึ้นเมื่อ σ มีค่าเพิ่มขึ้น

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.2 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, C) = (1,5)$

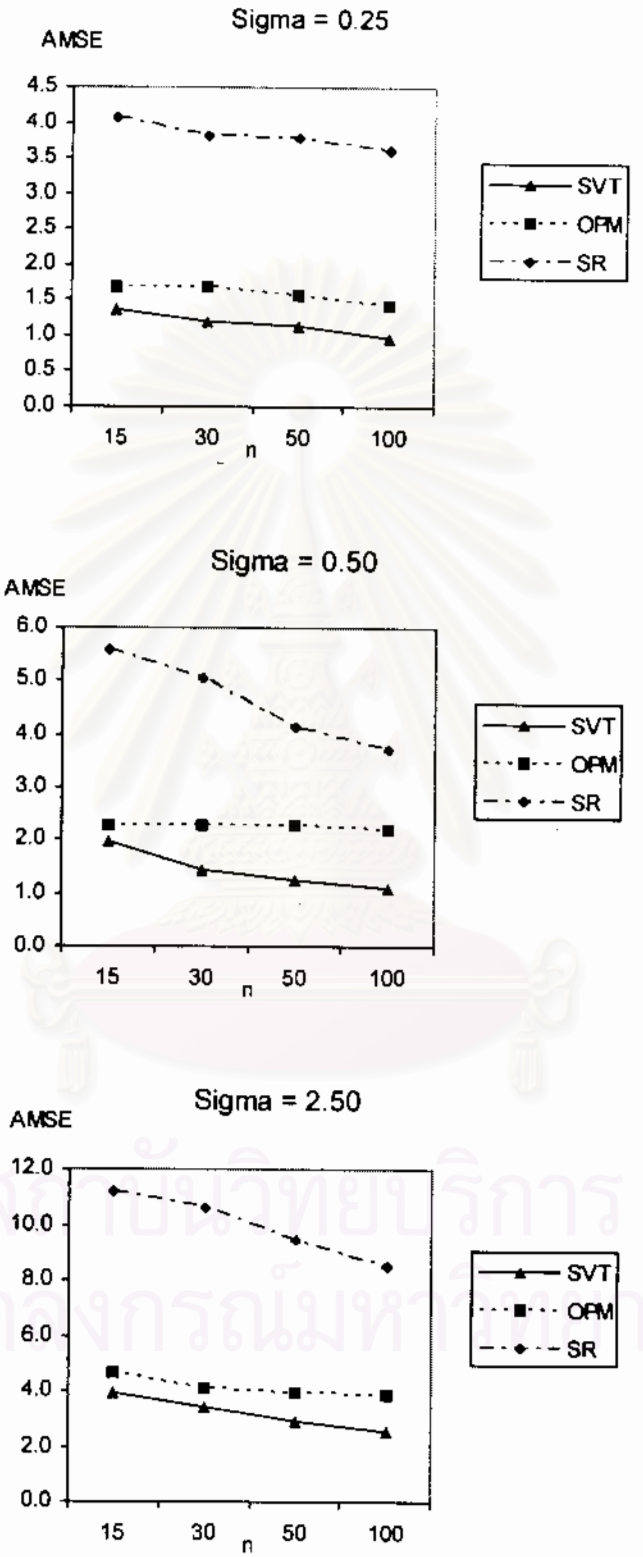
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	15	4.0798	1.6758	1.3649
		(0.3964)	(0.1371)	(0.1063)
		198.9083	22.7782	0.0000
	30	3.8236	1.6703	1.2043
		(0.3691)	(0.1268)	(0.1048)
		217.4956	38.6947	0.0000
	50	3.8027	1.5618	1.1289
		(0.3677)	(0.1223)	(0.0805)
		236.8500	38.3471	0.0000
	100	3.6347	1.4158	0.9684
		(0.3424)	(0.1101)	(0.0618)
		275.3304	46.1999	0.0000
0.50	15	5.5953	2.2849	1.9639
		(0.4810)	(0.2131)	(0.1632)
		184.9076	16.3450	0.0000
	30	5.3521	2.2772	1.4576
		(0.4612)	(0.2051)	(0.1041)
		224.3515	46.1993	0.0000
	50	4.1375	2.2745	1.2512
		(0.3861)	(0.1938)	(0.0979)
		230.6825	81.7855	0.0000
	100	3.7412	0.1888	1.0875
		(0.3371)	(0.0788)	(0.0862)
		244.0184	101.2689	0.0000
2.50	15	11.1879	4.6893	3.9527
		(0.9531)	(0.4149)	(0.2868)
		183.0445	18.6354	0.0000
	30	10.6167	4.1022	3.4475
		(0.8545)	(0.3542)	(0.2688)
		207.9536	18.9906	0.0000
	50	9.4775	3.9503	2.9163
		(0.8108)	(0.3227)	(0.2234)
		224.9837	35.4559	0.0000
	100	8.5101	3.8316	2.5374
		(0.7056)	(0.3153)	(0.2042)
		235.3866	51.0050	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.2 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ (σ_β / τ, c) เหมือนกันคือ (1,5) (ตารางที่ 4.2 และรูปที่ 4.2) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 3 ตัวแปร เป็น 5 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น เนื่องจากในการวิจัยครั้งนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูป เมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้น ทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มสูงขึ้น โดยเฉพาะอย่างยิ่งวิธี SR ซึ่งใช้หลักการในการคัดเลือกตัวแปรเข้าออกจากตัวแบบผ่านการทดสอบสมมติฐานเท่านั้น ในขณะที่วิธีการภายใต้แนวทางของเบส์ทั้ง 2 วิธี คือ วิธี BMA_{SVT} และ OPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการพิจารณาถึงความเหมาะสมของตัวแบบด้วย เช่น มีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT} จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น เพราะนอกจากวิธีนี้จะมีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบแล้วยังมีการพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระด้วย ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธีBMA_{SVT}ทำให้ค่า AMSEมีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.25 วิธี OPM ให้ค่า RDAMSE ลดลงเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 50 ทั้งนี้อาจเป็นเพราะวิธี BMA_{SVT}ให้ค่าAMSE ลดลงเพียงเล็กน้อย และเมื่อ σ มีค่าเพิ่มขึ้นจาก 0.25 เป็น

0.50 และ 2.50 ตามลำดับ จะส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มขึ้น ดังนั้นค่า RDAMSE จึงมีแนวโน้มเพิ่มขึ้นเมื่อ σ มีค่าเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้นจาก 3 ตัวแปรเป็น 5 ตัวแปร วิธี SR ให้ค่า AMSE เพิ่มขึ้น ดังนั้นค่า RDAMSE ของวิธี SR จึงมีแนวโน้มเพิ่มขึ้นเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.3 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, C) = (1,5)$

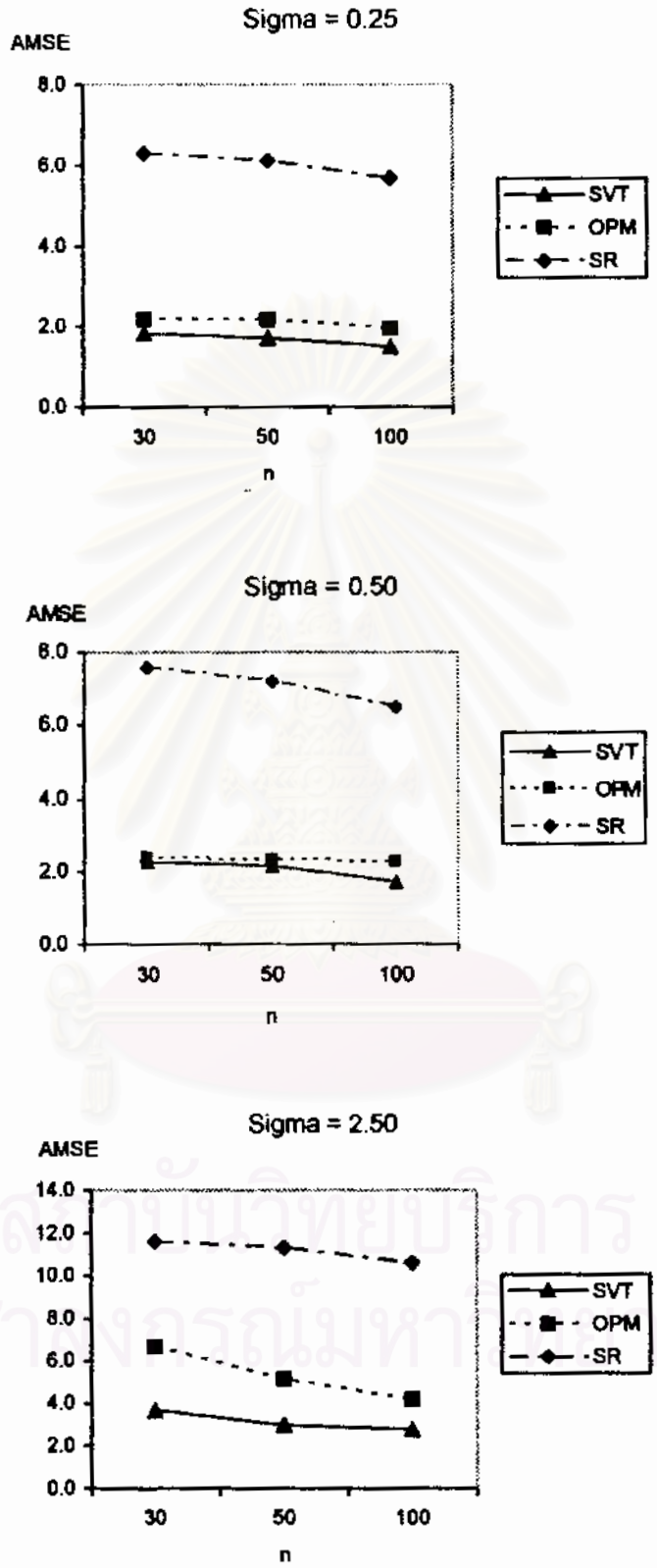
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	6.2869	2.1785	1.8430
		(0.6279)	(0.1865)	(0.1448)
		241.1231	18.2040	0.0000
	50	6.1168	2.1674	1.7176
		(0.5557)	(0.1737)	(0.1319)
		256.1248	26.1993	0.0000
	100	5.6928	1.9612	1.5128
		(0.5392)	(0.1580)	(0.0646)
		276.3088	29.6404	0.0000
0.50	30	7.5850	2.3759	2.2849
		(0.6886)	(0.2021)	(0.1739)
		231.9620	3.9827	0.0000
	50	7.1875	2.3193	2.1607
		(0.7025)	(0.1948)	(0.1864)
		232.6468	7.3402	0.0000
	100	6.5047	2.2469	1.7237
		(0.6271)	(0.1769)	(0.1485)
		277.3684	30.3533	0.0000
2.50	30	11.5936	6.6973	3.6788
		(1.1661)	(0.5754)	(0.3157)
		215.1462	82.0512	0.0000
	50	11.2964	5.1587	2.9642
		(0.9661)	(0.4137)	(0.2357)
		281.0944	74.0335	0.0000
	100	10.5906	4.1587	2.7642
		(0.9584)	(0.4037)	(0.1822)
		283.1344	50.4486	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.3 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,5) (ตารางที่ 4.3 และรูปที่ 4.3) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และการเพิ่มจำนวนตัวแปรอิสระเป็น 8 ตัวแปรนั้น จะทำให้ค่า AMSE มีแนวโน้มสูงขึ้นกว่ากรณีตัวแปรอิสระ 5 ตัวแปรเล็กน้อย ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 5 ตัวแปรเป็น 8 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในกรณีที่ตัวแปรอิสระเป็น 5 ตัวแปร ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSEมีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 2.50 วิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น ทั้งนี้อาจเป็นเพราะวิธี OPM ให้ค่า AMSE ลดลงอย่างรวดเร็วและวิธี BMA_{SVT} ลดลงเล็กน้อย

เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้นจาก 5 ตัวแปรเป็น 8 ตัวแปร วิธี SR ให้ค่า AMSE เพิ่มขึ้น ดังนั้นค่า RDAMSE ของวิธี SR จึงมีแนวโน้มเพิ่มขึ้นเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนวิธี OPM นั้นค่า RDAMSE มีแนวโน้มลดลงทั้งนี้อาจเป็นเพราะวิธีOPM ให้ค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเมื่อเพิ่มจำนวนตัวแปรอิสระเป็น 8 ตัว (ยกเว้นในกรณี σ เท่ากับ 2.50)

ตารางที่ 4.4 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

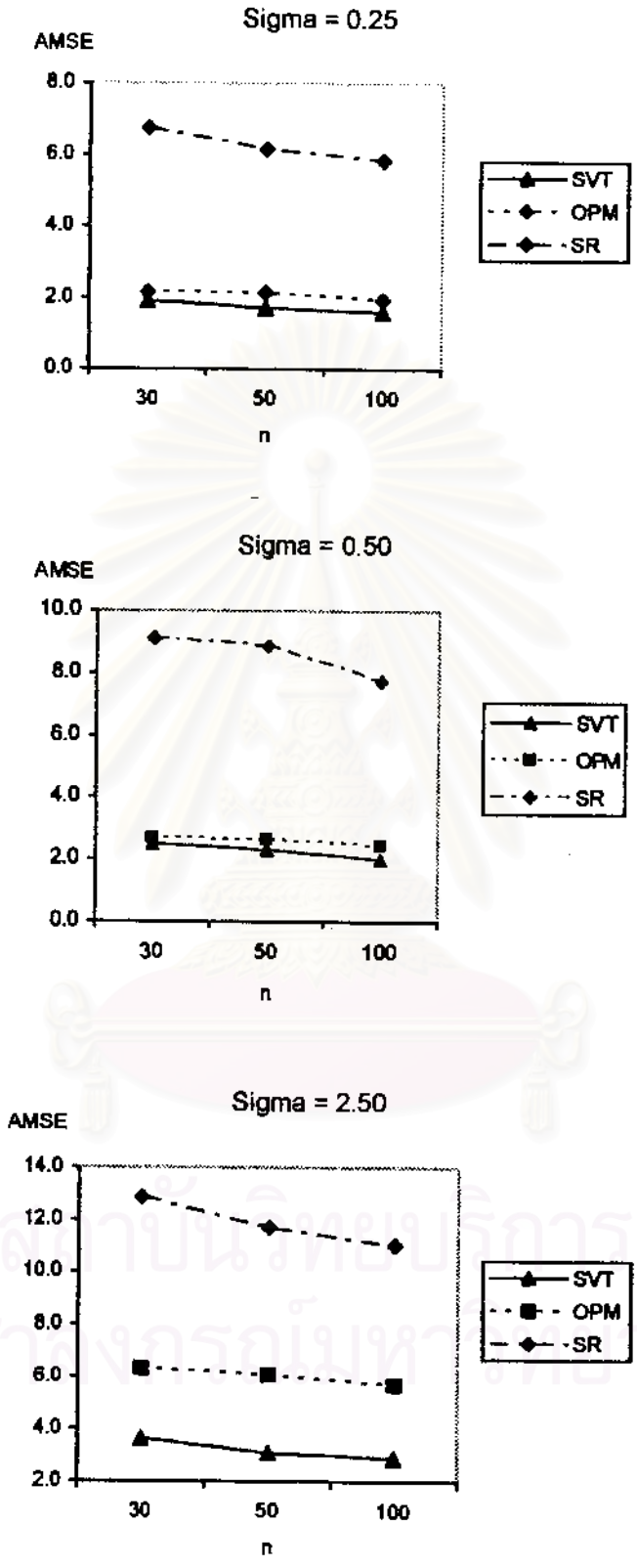
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	6.7617	2.1671	1.9228
		(0.6390)	(0.2189)	(0.1846)
		251.6590	12.7054	0.0000
	50	6.1753	2.1463	1.7079
		(0.5708)	(0.1985)	(0.1635)
		261.5727	25.6690	0.0000
	100	5.8338	1.9509	1.5873
		(0.5009)	(0.1881)	(0.1470)
		267.5298	22.9068	0.0000
0.50	30	9.1478	2.6895	2.5026
		(0.8355)	(0.2524)	(0.1929)
		265.5318	7.4682	0.0000
	50	8.8717	2.6541	2.3367
		(0.8485)	(0.2357)	(0.2285)
		279.6679	13.5833	0.0000
	100	7.7219	2.4352	2.0110
		(0.4205)	(0.2115)	(0.1847)
		283.9831	21.0940	0.0000
2.50	30	12.8407	6.3416	3.6529
		(1.1250)	(0.6448)	(0.3121)
		251.5207	73.6045	0.0000
	50	11.7042	6.0829	3.1001
		(1.1158)	(0.5759)	(0.2614)
		277.5427	96.2163	0.0000
	100	11.0116	5.6872	2.8870
		(0.9708)	(0.5211)	(0.2645)
		281.4202	96.9934	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.4 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,5) (ตารางที่ 4.4 และรูปที่ 4.4) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่าAMSE สูงสุด โดยวิธี OPM มีค่าAMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และการเพิ่มจำนวนตัวแปรอิสระเป็น 10 ตัวแปรนั้น จะทำให้ค่า AMSE มีแนวโน้มสูงขึ้นกว่ากรณีตัวแปรอิสระ 8 ตัวแปรเล็กน้อย ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่าAMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 8 ตัวแปรเป็น 10 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในกรณีที่ตัวแปรอิสระเป็น 5 ตัวแปร ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSEมีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้นจาก 8 ตัวแปรเป็น 10 ตัวแปร วิธี SR ให้ค่า AMSE เพิ่มขึ้น ดังนั้นค่า RDAMSE ของวิธี SR จึงมีแนวโน้มเพิ่มขึ้นเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนวิธี OPM นั้นค่า RDAMSE มีแนวโน้มลดลงทั้งนี้อาจเป็นเพราะวิธีOPM ให้ค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเมื่อเพิ่มจำนวนตัวแปรอิสระเป็น 8 ตัว (ยกเว้นในกรณี σ เท่ากับ 2.50)

ตารางที่ 4.5 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

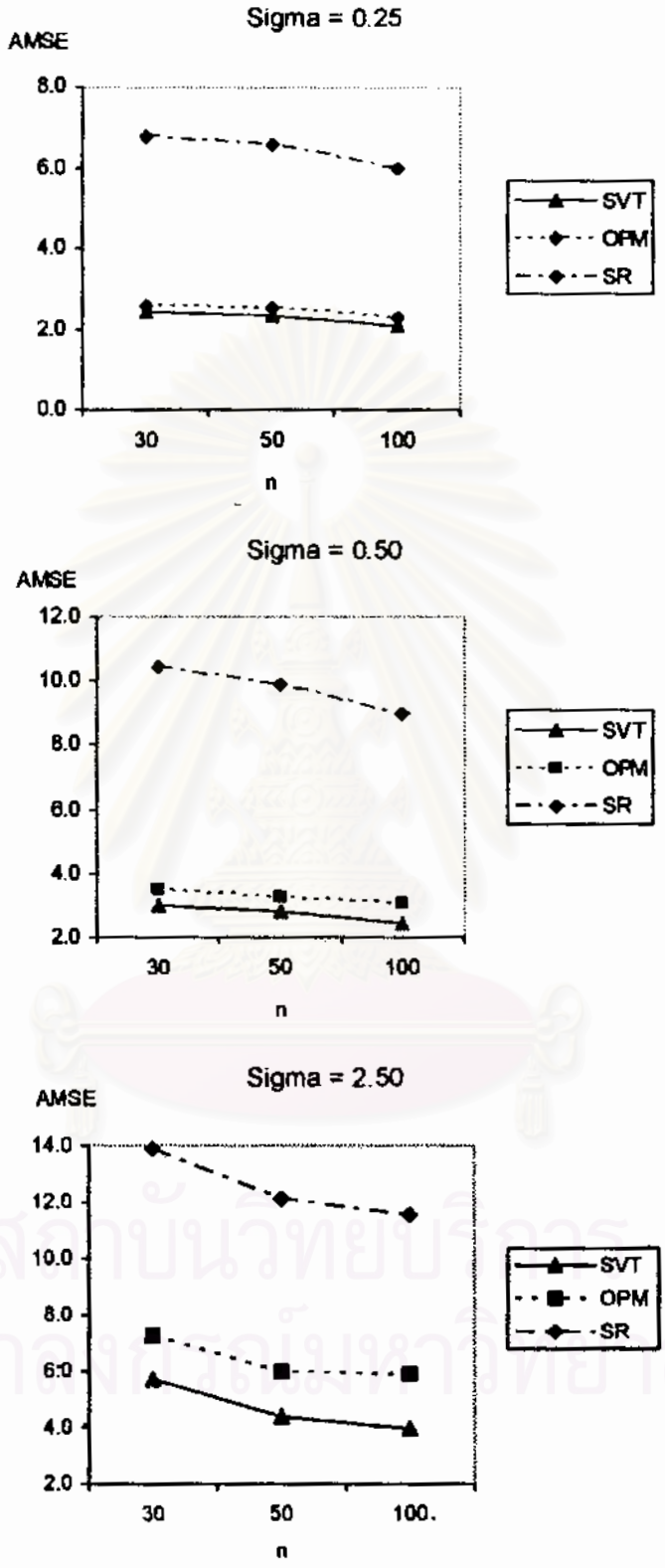
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	6.7606	2.5512	2.4018
		(0.6696)	(0.2388)	(0.2216)
		181.4806	6.2203	0.0000
	50	6.5587	2.5111	2.3139
		(0.6442)	(0.2365)	(0.2173)
		183.4478	8.5224	0.0000
	100	5.9911	2.2631	2.0722
		(0.5559)	(0.2047)	(0.1897)
		189.1178	9.2124	0.0000
0.50	30	10.4180	3.4897	3.0101
		(0.9860)	(0.2958)	(0.2646)
		246.1015	15.9330	0.0000
	50	9.8757	3.2170	2.8145
		(0.9400)	(0.2837)	(0.2581)
		250.8865	14.3009	0.0000
	100	8.9512	3.0265	2.4281
		(0.8635)	(0.2692)	(0.2077)
		268.6504	24.6448	0.0000
2.50	30	13.8762	7.2581	5.7083
		(1.1187)	(0.5383)	(0.4125)
		143.0881	27.1499	0.0000
	50	12.1164	5.9923	4.3932
		(1.0177)	(0.4903)	(0.3912)
		175.7990	36.3994	0.0000
	100	11.5373	5.8910	3.9610
		(1.0034)	(0.4081)	(0.3325)
		191.2724	48.7251	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.5 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ (σ_β / τ , c) = (1,5)

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ (σ_β / τ, c) เหมือนกันคือ (1,5) (ตารางที่ 4.5 และรูปที่ 4.5) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และการเพิ่มจำนวนตัวแปรอิสระเป็น 12 ตัวแปรนั้น จะทำให้ค่า AMSE มีแนวโน้มสูงขึ้นกว่ากรณีตัวแปรอิสระ 10 ตัวแปรเล็กน้อย ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่าAMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่าAMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 10 ตัวแปรเป็น 12 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในกรณีที่ตัวแปรอิสระเป็น 5 ตัวแปร ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSEมีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSE เพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.50 วิธี OPM ให้ค่า RDAMSE ลดลงเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 50 ทั้งนี้อาจเป็นเพราะวิธี BMA_{SVT}ให้ค่า AMSE ลดลงเพียงเล็กน้อย

เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้นจาก 10 ตัวแปรเป็น 12 ตัวแปร วิธี SR ให้ค่า AMSE เพิ่มขึ้นเพียงเล็กน้อย ดังนั้นค่า RDAMSE ของวิธี SR จึงมีแนวโน้มลดลง ส่วนวิธี OPM นั้นให้ผลสรุปเช่นเดียวกับวิธี SR

ตารางที่ 4.6 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

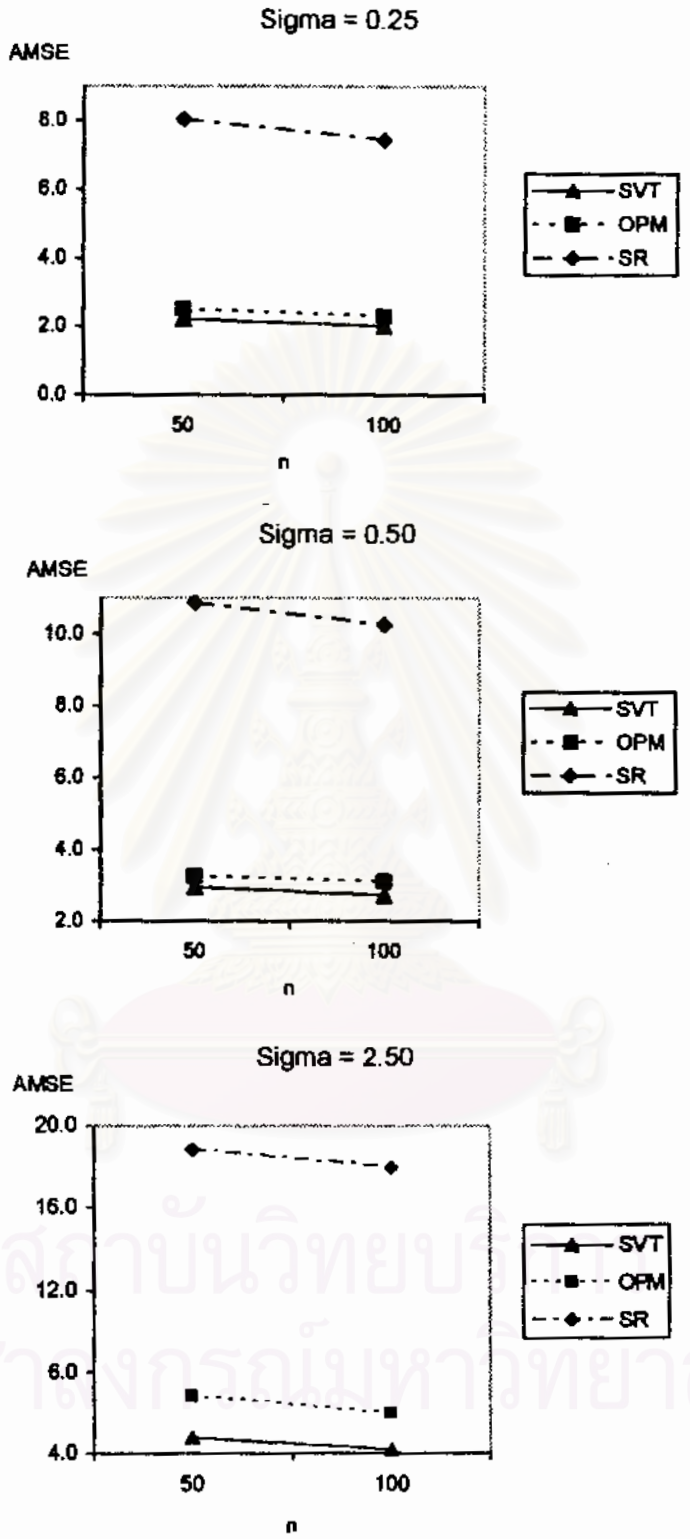
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	50	8.0429	2.5064	2.2120
		(0.7770)	(0.2402)	(0.2028)
		263.6031	13.3092	0.0000
	100	7.4425	2.3057	2.0035
		(0.6902)	(0.2230)	(0.1921)
		271.4749	15.0836	0.0000
0.50	50	10.8595	3.2544	2.9545
		(1.0832)	(0.3071)	(0.2754)
		267.5579	10.1506	0.0000
	100	10.2588	3.1036	2.7083
		(0.9659)	(0.2723)	(0.2379)
		278.7911	14.5959	0.0000
2.50	50	18.8468	6.8236	4.8215
		(1.3040)	(0.6230)	(0.3697)
		290.8908	41.5244	0.0000
	100	17.9775	5.9301	4.1823
		(1.2233)	(0.5156)	(0.3321)
		329.8472	41.7904	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.6 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,5)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ (σ_β / τ, c) เหมือนกันคือ (1,5) (ตารางที่ 4.4 และรูปที่ 4.4) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และการเพิ่มจำนวนตัวแปรอิสระเป็น 15 ตัวแปรนั้น จะทำให้ค่า AMSE มีแนวโน้มสูงขึ้นกว่ากรณีตัวแปรอิสระ 12 ตัวแปรเล็กน้อย ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่าAMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 12 ตัวแปรเป็น 15 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในกรณีจำนวนตัวแปรอิสระเท่ากับ 5 ตัวแปร ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 50 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น

เมื่อจำนวนตัวแปรอิสระเพิ่มขึ้นจาก 12 ตัวแปรเป็น 15 ตัวแปร วิธี SR ให้ค่า AMSE เพิ่มขึ้นอย่างชัดเจน ดังนั้นค่า RDAMSE ของวิธี SR จึงมีแนวโน้มเพิ่มขึ้นเมื่อจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนวิธี OPM นั้นเมื่อค่า σ เท่ากับ 0.25 ค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเล็กน้อยและเมื่อ σ เท่ากับ 0.50 และ 2.50 ค่า RDAMSE มีแนวโน้มลดลงเล็กน้อย

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่(σ_β / τ , c) เหมือนกันคือ (1,5) (ตารางที่ 4.1 – 4.6 และรูปที่ 4.1 – 4.6) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุดโดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกัน กล่าวคือค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เนื่องจากในการวิจัยครั้งนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูปเมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้นทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มสูงขึ้นโดยเฉพาะอย่างยิ่งวิธี SR ซึ่งใช้หลักการในการคัดเลือกตัวแปรเข้าออกจากตัวแบบผ่านการทดสอบสมมติฐานเท่านั้น ในขณะที่วิธีการภาย ใต้แนวทางของเบส์ทั้ง 2 วิธี คือ วิธี BMA_{SVT} และ OPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการพิจารณาถึงความเหมาะสมของตัวแบบด้วย เช่น มีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT} จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น เพราะนอกจากวิธีนี้จะมีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบแล้วยังมีการพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระด้วย ค่า AMSE แปรผกผันกับขนาดตัวอย่างเพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

ผลการวิจัยของวิธี BMA_{SVT} จะให้ค่า RDAMSE น้อยที่สุด รองลงมาคือวิธี OPM และวิธี SR ตามลำดับ โดยส่วนใหญ่ค่า RDAMSE ของวิธี SR เมื่อขนาดตัวอย่างและจำนวนตัวแปรอิสระเพิ่มขึ้น ส่วนค่า RDAMSE ของวิธี OPM นั้นมีแนวโน้มลดลงเมื่อขนาดตัวอย่างและจำนวนตัวแปรอิสระเพิ่มขึ้นเนื่องจากวิธีนี้เป็นวิธีการภายใต้แนวทางของเบส์ ซึ่งการพิจารณาถึงความเหมาะสมของตัวแบบด้วยความน่าจะเป็นภายหลังของตัวแบบเพื่อคัดเลือกตัวแปรอิสระที่มีค่าความน่าจะเป็นภายหลังกลาง ๆ และหาสมการถดถอยที่ดีที่สุดจากตัวแบบที่ให้ค่าความสูญเสียอันเกิดจากความผิดพลาดยกกำลังสอง (square error loss) ต่ำสุด ดังนั้นค่า AMSE ของวิธี OPM จึงไม่เพิ่มขึ้นอย่างรวดเร็วเมื่อตัวอย่างและจำนวนตัวแปรอิสระเพิ่มขึ้นเหมือนวิธี SR ซึ่งเป็นวิธีพื้นฐาน

สรุปตอนที่ 4.1 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ สำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,5)

เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1, 5) พบว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุก ๆ กรณี โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย และวิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

จากผลการวิจัยในตอนที่ 4.1 การเปลี่ยนแปลงของค่า AMSE และค่า RDAMSE มีลักษณะดังนี้

1) ค่า AMSE

(1) แปรผันตามส่วนเบี่ยงเบนมาตรฐาน รองลงมาคือจำนวนตัวแปรอิสระ

(2) แปรผกผันกับขนาดตัวอย่าง

2) ค่า RDAMSE

(1) ค่า RDAMSE ของวิธี SR แปรผันตามขนาดตัวอย่าง รองลงมาคือจำนวนตัวแปรอิสระ

(2) ค่า RDAMSE ของวิธี OPM แปรผันตามขนาดตัวอย่างและแปรผกผันกับจำนวนตัวแปรอิสระ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4.2 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี QPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (1,10) ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตารางที่ 4.7 – 4.12 และรูปที่ 4.7 – 4.12

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.7 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

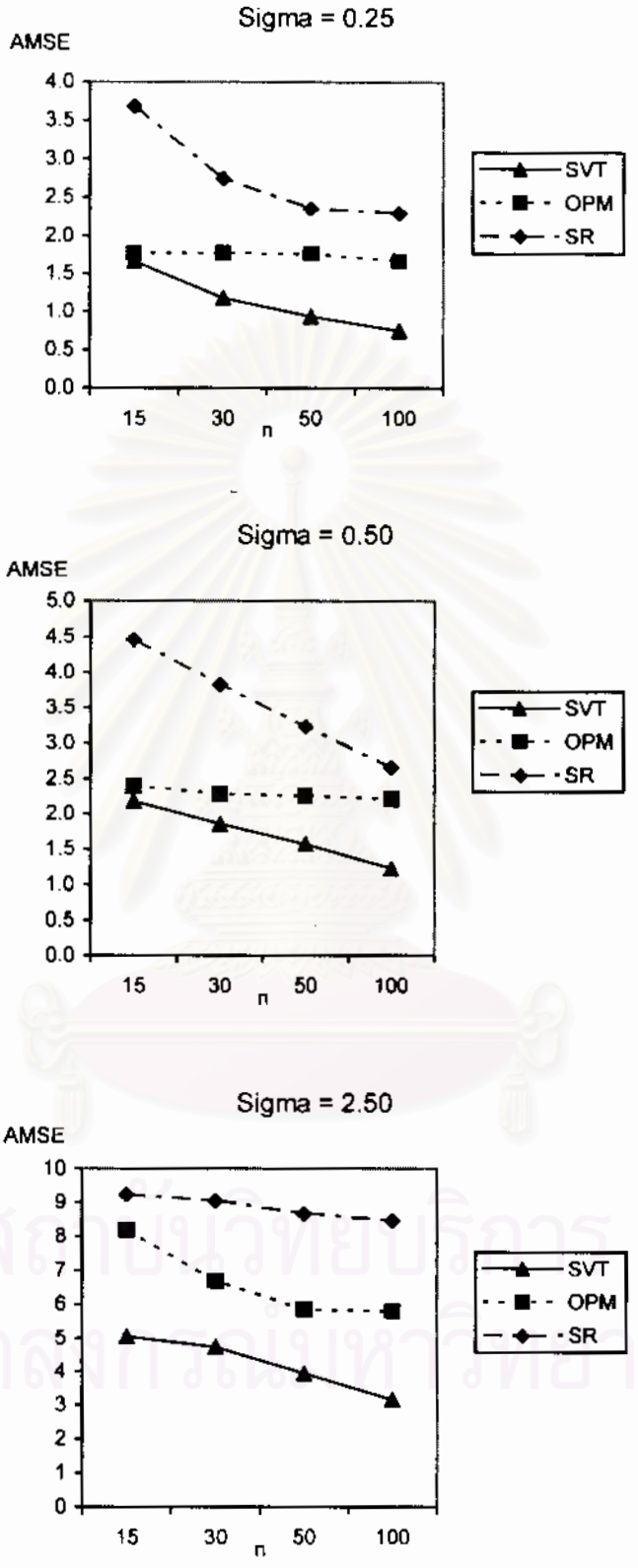
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	15	3.6894	1.7701	1.6562
		(0.3755)	(0.1384)	(0.1078)
		122.7630	6.8772	0.0000
	30	2.7396	1.7643	1.1793
		(0.2547)	(0.1278)	(0.0960)
		132.3073	49.6057	0.0000
	50	2.3545	1.7609	0.9368
		(0.2049)	(0.1192)	(0.0808)
		151.3343	87.9697	0.0000
	100	2.2929	1.6599	0.7501
		(0.1996)	(0.1171)	(0.0597)
		205.6792	121.2905	0.0000
0.50	15	4.4604	2.3927	2.1783
		(0.3826)	(0.2543)	(0.1503)
		104.7652	9.8425	0.0000
	30	3.8260	2.2838	1.8634
		(0.3130)	(0.2069)	(0.1086)
		105.3236	22.5609	0.0000
	50	3.2261	2.2475	1.5767
		(0.2687)	(0.2021)	(0.0926)
		104.6109	42.5446	0.0000
	100	2.6483	2.2135	1.2250
		(0.2172)	(0.1576)	(0.0735)
		116.1859	80.6924	0.0000
2.50	15	9.2421	8.1896	5.0460
		(0.9114)	(0.9183)	(0.4345)
		83.1570	62.2989	0.0000
	30	9.0484	6.6835	4.7303
		(0.8948)	(0.6369)	(0.3803)
		91.2860	41.2913	0.0000
	50	8.6672	5.8439	3.9521
		(0.8405)	(0.5540)	(0.3286)
		119.3062	47.8682	0.0000
	100	8.4512	5.7763	3.1627
		(0.8188)	(0.5065)	(0.3039)
		167.2147	82.6383	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



<u>รูปภาพที่ 4.7</u> แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วย<u>จำนวนตัวแปรอิสระเท่ากับ 3</u> เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ (σ_β / τ , C) = (1.10)

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$เหมือนกันคือ (1,10) (ตารางที่ 4.7 และรูปที่ 4.7) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่าAMSE สูงสุด วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) ส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อยเพราะวิธีนี้มีหลักการเฉลี่ยตัวแบบด้วยความน่าจะเป็นหลังของตัวแบบ แต่จะส่งผลต่อวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.1และตารางที่ 4.7) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่าAMSE <u>แปรผันตาม</u>ค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น และค่า AMSE <u>แปรผกผัน</u>กับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้ ดังนั้นการเพิ่มขนาดตัวอย่างจะทำให้ค่าAMSEของทุกวิธีมีแนวโน้มลดลง

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่าAMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 2.50 วิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 15 เป็น 30 ทั้งนี้อาจเป็นเพราะวิธี BMA_{SVT}ให้ค่า AMSE ลดลงเพียงเล็กน้อยเมื่อเทียบกับวิธี OPM

ตารางที่ 4.8 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

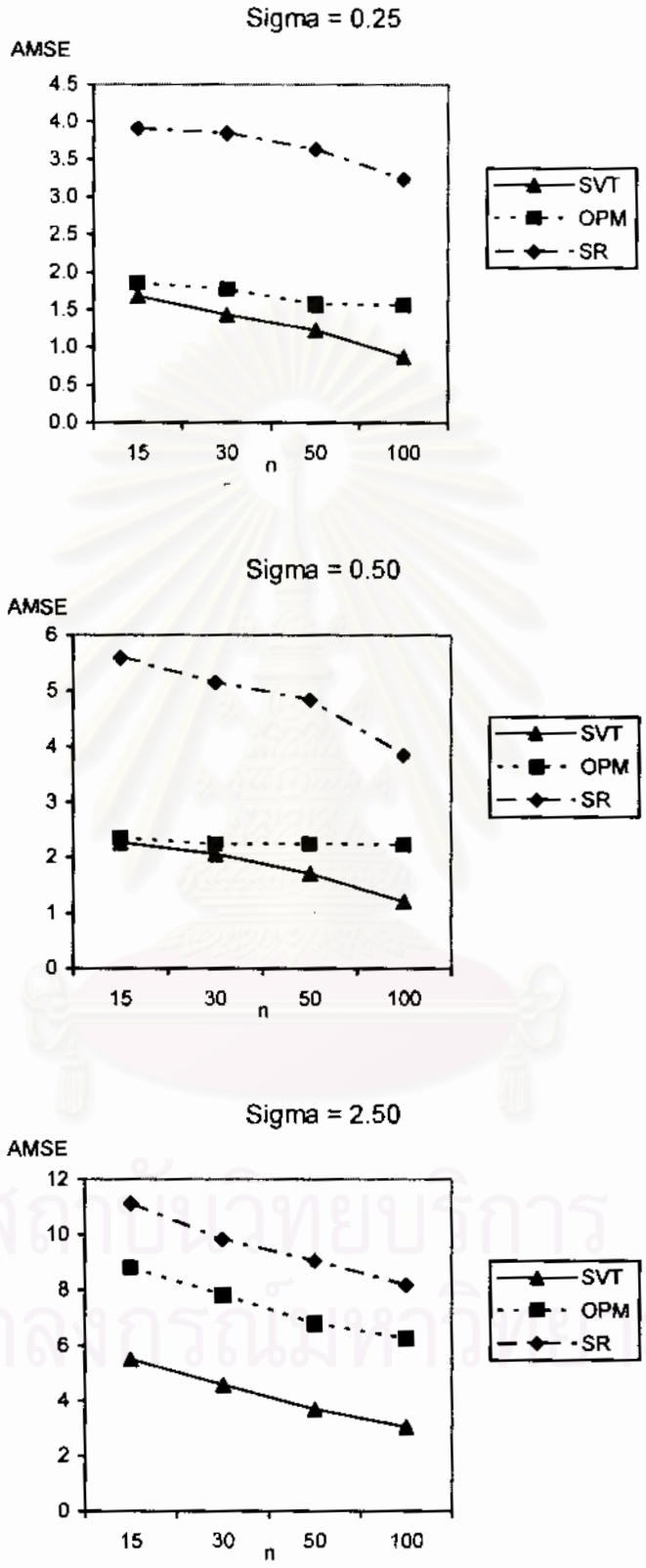
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	15	3.9101	1.8564	1.6829
		(0.3537)	(0.1584)	(0.1428)
		132.3430	10.3096	0.0000
	30	3.8459	1.7716	1.4291
		(0.3479)	(0.1472)	(0.1239)
		169.1134	23.9661	0.0000
	50	3.6333	1.5718	1.2276
		(0.3568)	(0.1387)	(0.0943)
		195.9677	28.0384	0.0000
	100	3.2404	1.5599	0.8690
		(0.2922)	(0.1271)	(0.0697)
		272.8884	79.5052	0.0000
0.50	15	5.5933	2.3481	2.2742
		(0.4943)	(0.2442)	(0.1984)
		145.9458	3.2495	0.0000
	30	5.1512	2.2438	2.0615
		(0.4816)	(0.2108)	(0.1915)
		149.8763	8.8430	0.0000
	50	4.8411	2.2409	1.7147
		(0.4730)	(0.2068)	(0.1574)
		182.3293	30.6876	0.0000
	100	3.8454	2.2272	1.2115
		(0.3929)	(0.2009)	(0.0954)
		217.4082	83.8382	0.0000
2.50	15	11.1243	8.8099	5.4914
		(1.1232)	(0.8526)	(0.5083)
		102.5768	60.4309	0.0000
	30	9.8160	7.8176	4.5586
		(0.9741)	(0.7159)	(0.4497)
		115.3293	71.4912	0.0000
	50	9.0548	6.7856	3.6712
		(0.8851)	(0.5995)	(0.3276)
		146.6441	84.8333	0.0000
	100	8.1707	6.2474	3.0415
		(0.8113)	(0.5562)	(0.2657)
		168.6405	105.4052	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.8 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,10) (ตารางที่ 4.8 และรูปที่ 4.8) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) ส่งผลกระทบต่อวิธี BMA_{SVT} และวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.2และตารางที่ 4.8) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 3 ตัวแปรเป็น 5 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยง เบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น

ตารางที่ 4.9 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

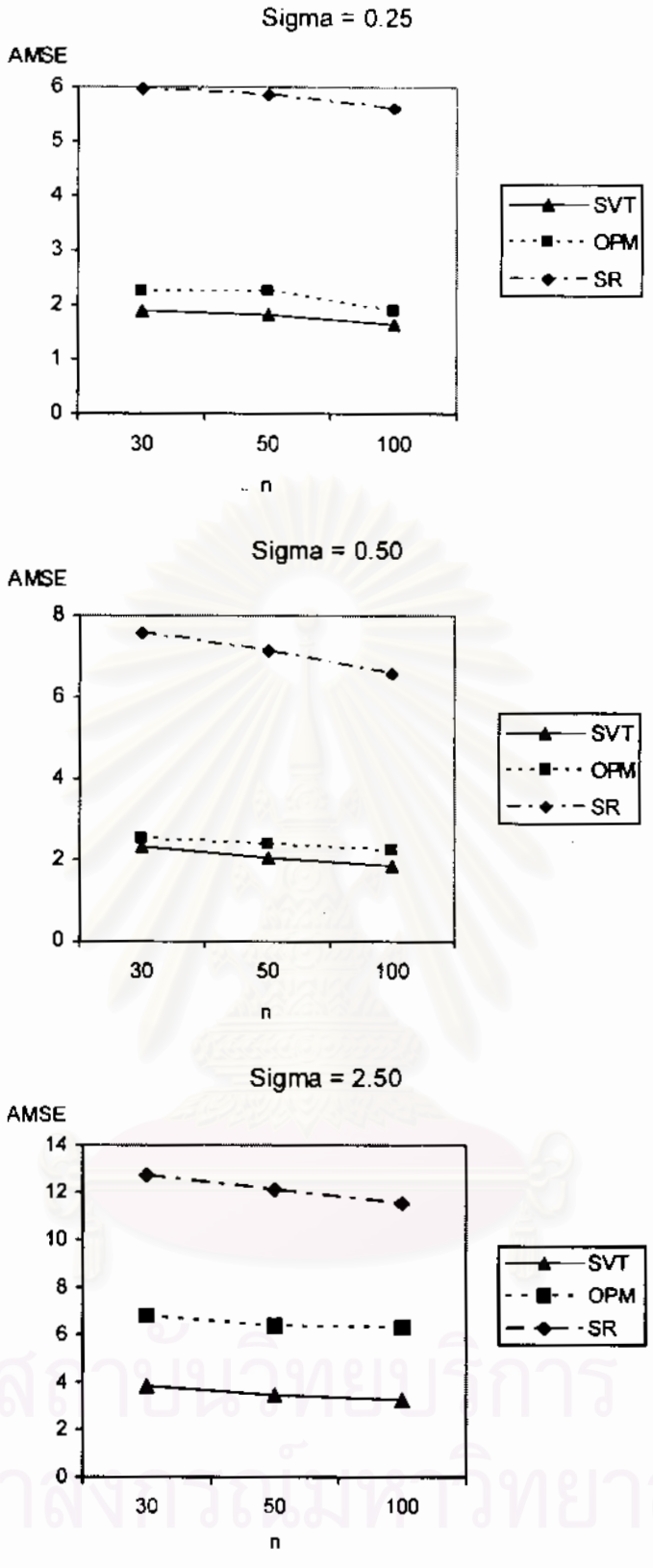
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	5.9639	2.2633	1.9060
		(0.5723)	(0.1165)	(0.1048)
		212.9014	18.7461	0.0000
	50	5.8647	2.2624	1.8121
		(0.5396)	(0.1137)	(0.0819)
		223.6411	24.8496	0.0000
	100	5.5910	1.9029	1.6304
		(0.5179)	(0.0980)	(0.0646)
		242.9220	16.7137	0.0000
0.50	30	7.5808	2.5307	2.3425
		(0.7232)	(0.2519)	(0.2136)
		223.6201	8.0342	0.0000
	50	7.1632	2.3749	2.0496
		(0.7004)	(0.2303)	(0.1819)
		249.4926	15.8714	0.0000
	100	6.5862	2.2535	1.8419
		(0.6366)	(0.2103)	(0.1303)
		257.5764	22.3465	0.0000
2.50	30	12.7082	6.7956	3.8279
		(1.1232)	(0.6278)	(0.3820)
		231.9888	77.5281	0.0000
	50	12.1147	6.3870	3.4472
		(1.0542)	(0.6050)	(0.3246)
		251.4359	85.2808	0.0000
	100	11.5243	6.3236	3.2506
		(1.0065)	(0.6161)	(0.2653)
		254.5284	94.5364	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.9 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,10) (ตารางที่ 4.9 และรูปที่ 4.9) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อย ซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.3และตารางที่ 4.9) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่าAMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่าσ นั้นจะทำให้ค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 5 ตัวแปรเป็น 8 ตัวแปรจะทำให้ค่าAMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 (วิธี OPM จะให้ค่า AMSE ลดลงในกรณี σ เท่ากับ 2.50) ค่าAMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยง เบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่าAMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.25 วิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 50 เป็น 100 ทั้งนี้อาจเป็นเพราะวิธี OPM ให้ค่าAMSE ลดลงมากเมื่อเทียบกับวิธี BMA_{SVT}

ตารางที่ 4.10 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

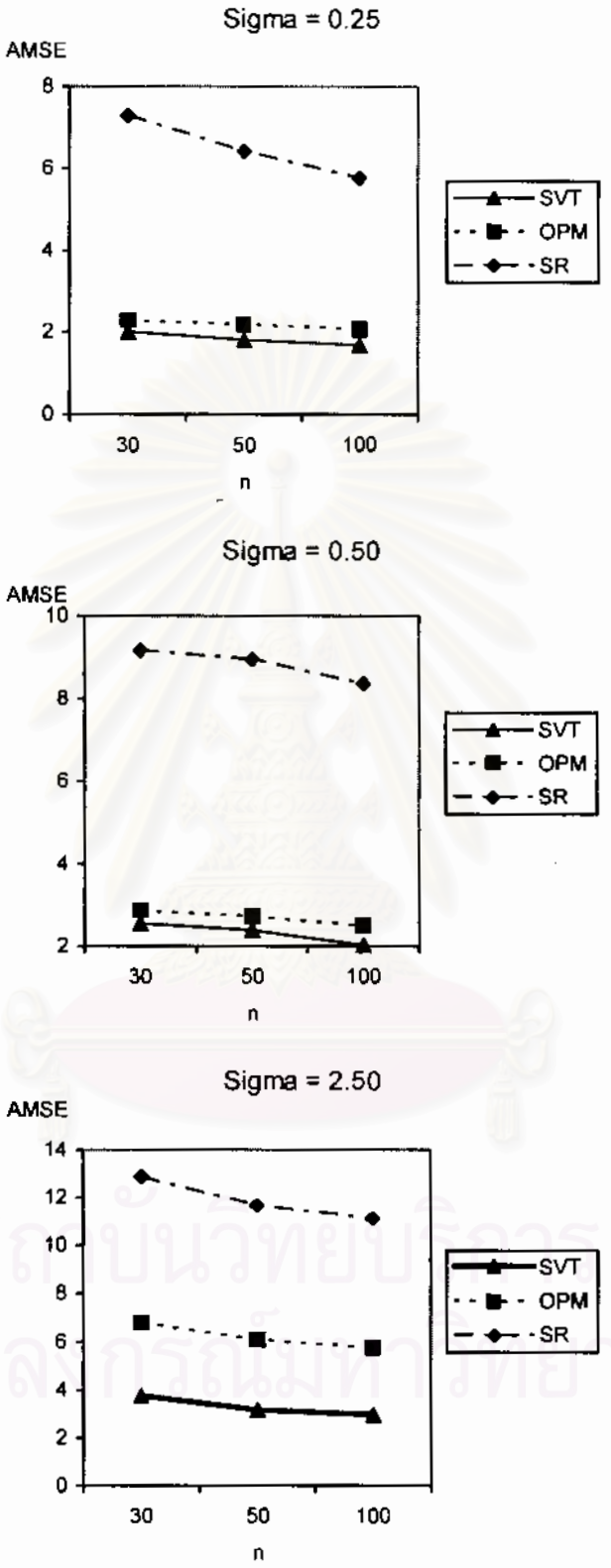
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	7.2937	2.2697	2.0052
		(0.7220)	(0.2122)	(0.1840)
		263.7393	13.1907	0.0000
	50	6.4131	2.1778	1.8099
		(0.6096)	(0.1998)	(0.1684)
		254.3345	20.3271	0.0000
	100	5.7591	2.0672	1.6772
		(0.5616)	(0.1841)	(0.1431)
		243.3759	23.2530	0.0000
0.50	30	9.1666	2.8723	2.5569
		(0.8817)	(0.2606)	(0.1929)
		258.5044	12.3352	0.0000
	50	8.9499	2.7247	2.3885
		(0.8978)	(0.2349)	(0.0864)
		274.7080	14.0758	0.0000
	100	8.3595	2.4995	2.0237
		(0.8068)	(0.2107)	(0.0785)
		313.0800	23.5114	0.0000
2.50	30	12.8496	6.7943	3.7412
		(1.2804)	(0.5995)	(0.3237)
		243.4620	81.6075	0.0000
	50	11.6818	6.0909	3.1367
		(1.1548)	(0.5923)	(0.3023)
		272.4232	94.1818	0.0000
	100	11.1229	5.7462	2.9627
		(1.0347)	(0.5102)	(0.2543)
		275.4312	93.9514	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.10 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ (σ_β / τ, c) เหมือนกันคือ (1,10) (ตารางที่ 4.10 และรูปที่ 4.10) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่าAMSE สูงสุด โดยวิธี OPM มีค่าAMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อย ซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.4 และตารางที่ 4.10) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่าAMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่าAMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 8 ตัวแปรเป็น 10 ตัวแปรจะทำให้ค่าAMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน

ตารางที่ 4.11 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

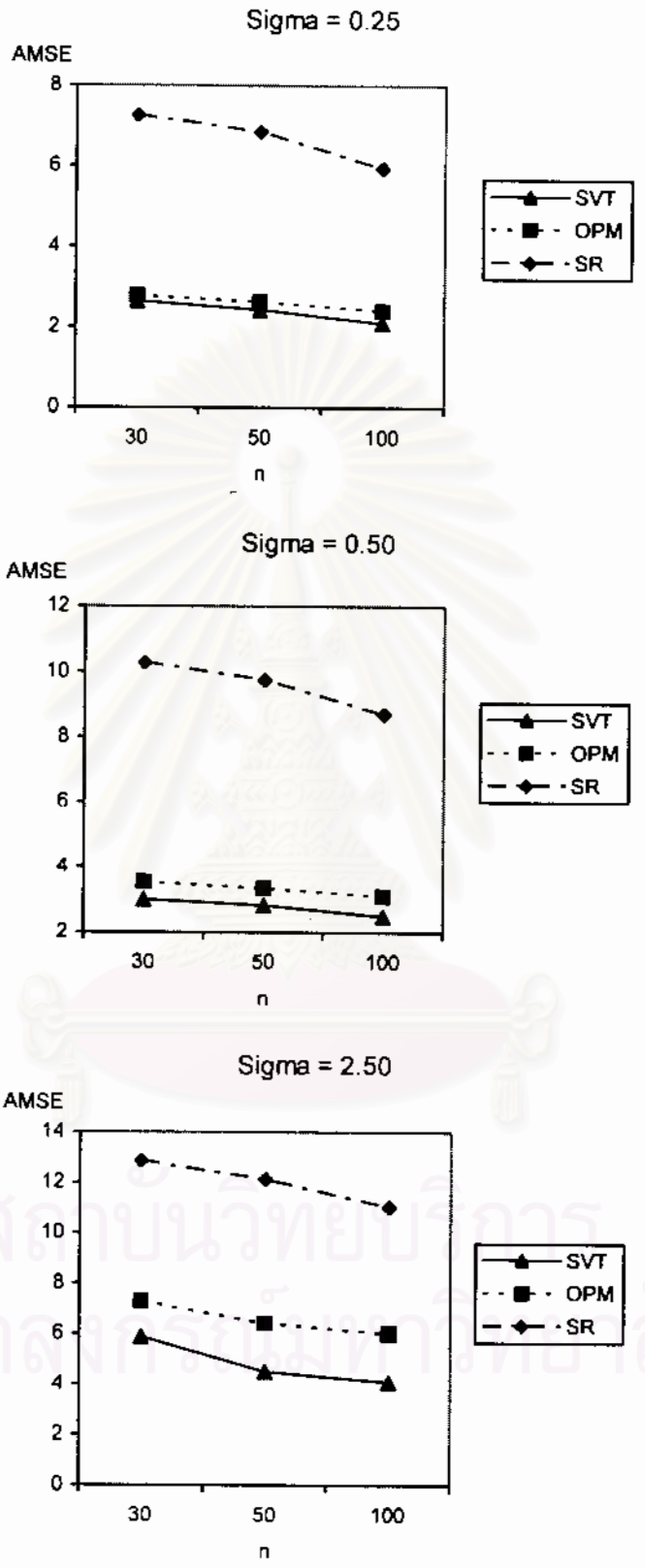
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	7.2648	2.7722	2.6548
		(0.7168)	(0.2487)	(0.2337)
		173.6477	4.4222	0.0000
	50	6.8563	2.6227	2.4339
		(0.6002)	(0.2365)	(0.2173)
		185.8088	7.7571	0.0000
	100	5.9394	2.4032	2.1021
		(0.5338)	(0.2268)	(0.1901)
		183.4186	14.3922	0.0000
0.50	30	10.2895	3.5422	3.0147
		(0.9762)	(0.2906)	(0.2829)
		241.3109	17.4976	0.0000
	50	9.7479	3.3491	2.8292
		(0.9509)	(0.2749)	(0.2341)
		244.5462	18.3762	0.0000
	100	8.6933	3.1105	2.4914
		(0.8254)	(0.2607)	(0.1643)
		248.9323	24.8495	0.0000
2.50	30	12.8479	7.2593	5.8793
		(1.2781)	(0.6261)	(0.3593)
		118.5277	23.4722	0.0000
	50	12.1263	6.4103	4.5098
		(1.0771)	(0.5891)	(0.3070)
		168.8878	42.1416	0.0000
	100	11.0152	6.0119	4.0911
		(1.0072)	(0.5731)	(0.2974)
		169.2479	46.9507	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงด้วยสถิ 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.11 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,10) (ตารางที่ 4.11 และรูปที่ 4.11) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) ส่งผลกระทบต่อวิธี BMA_{SVT} และวิธี OPM เพียงเล็กน้อย ซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.5 และตารางที่ 4.11) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่าAMSE <u>แปรผันตาม</u>ค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยัง<u>แปรผันตาม</u>จำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 10 ตัว แปรเป็น 12 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE <u>แปรผกผัน</u>กับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่าRDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.25 วิธี SR ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 50 เป็น 100 ทั้งนี้อาจเป็นเพราะวิธี SRให้ค่าAMSE ลดลงมากเมื่อเทียบกับวิธี BMA_{SVT}

ตารางที่ 4.12 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

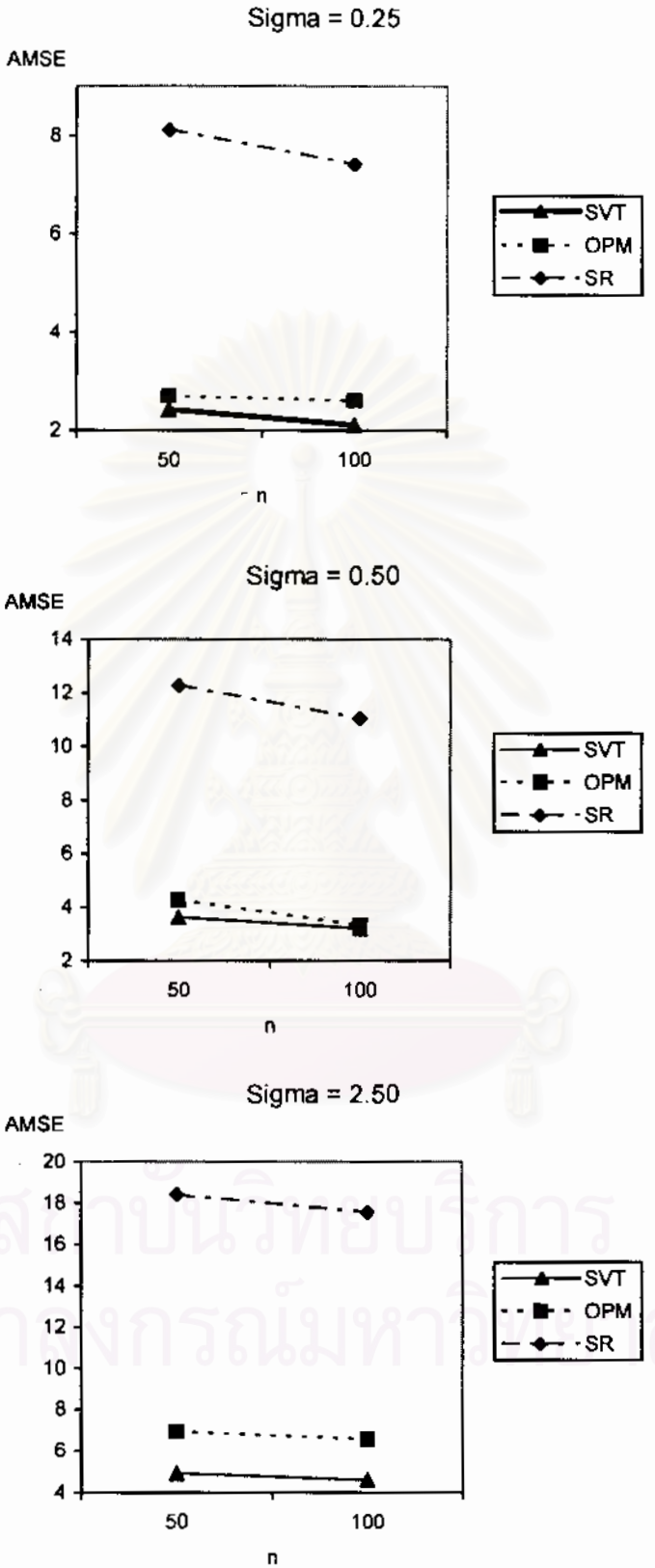
σ	n	วิธีการ		
		SR	**OPM**	**BMA_{SVT}**
0.25	**50**	8.1068	2.6964	2.4159
		(0.8023)	(0.2537)	(0.2009)
		235.5602	11.6106	**0.0000**
	100	7.4229	2.5900	2.1051
		(0.6659)	(0.2514)	(0.1943)
		252.6151	23.0345	**0.0000**
0.50	**50**	12.2733	4.2586	3.6279
		(1.1074)	(0.3702)	(0.3379)
		238.3032	17.3847	**0.0000**
	100	11.0542	3.3151	3.1983
		(1.0006)	(0.3098)	(0.2637)
		245.6274	3.6519	**0.0000**
2.50	**50**	18.4067	6.9405	4.9272
		(1.8689)	(0.6402)	(0.4291)
		273.5732	40.8609	**0.0000**
	100	17.5353	6.5456	4.6082
		(1.5452)	(0.6032)	(0.3925)
		280.5238	42.0424	**0.0000**

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.12 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,10) (ตารางที่ 4.12 และรูปที่ 4.12) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่าAMSE สูงสุด โดยวิธี OPM มีค่าAMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่าAMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อย ซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.6 และตารางที่ 4.12) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์ส้มประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 12 ตัวแปรเป็น 15 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่าAMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธีBMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ สำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่(σ_β / τ, c) เหมือนกันคือ (1,10) (ตารางที่ 4.7 – 4.12 และรูปที่ 4.7 – 4.12) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อ σ มีค่าเพิ่มขึ้น ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่าคงที่ σ_β / τ และ c จะมีผลต่อวิธี BMA_{SVT} และ OPM เท่านั้น โดยค่า AMSE จะ เพิ่มขึ้นเมื่อค่าคงที่ σ_β / τ และ c ที่สูงขึ้น เนื่องจากค่าคงที่ σ_β / τ และ c เป็นค่าที่กำหนดลักษณะการกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยที่สุ่มได้ ค่าคงที่ σ_β / τ และ c ที่สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น ทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้ค่า AMSE มีค่าสูงขึ้น และจากผลการวิจัยตอนที่ 4.2 นี้ ค่าคงที่ σ_β / τ และ c เปลี่ยนแปลงจากตอนที่ 4.1 เล็กน้อย (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) จึงส่งผลให้ค่า AMSE ของวิธี BMA_{SVT} และวิธี OPM มีการเปลี่ยนแปลงไม่มากนัก

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เนื่องจากในการวิจัยครั้งนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูปเมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้นทำให้ค่าAMSE ของทุกวิธีมีแนวโน้มสูงขึ้นโดยเฉพาะอย่างยิ่งวิธี SR ซึ่งใช้หลักการในการคัดเลือกตัวแปรเข้าออกจากตัวแบบผ่านการทดสอบสมมติฐานเท่านั้น ในขณะที่วิธีการภาย ใต้แนวทางของเบส์ทั้ง 2 วิธี คือ วิธี BMA_{SVT} และ OPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการพิจารณาถึงความเหมาะสมของตัวแบบด้วย เช่น มีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT} จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น เพราะนอกจากวิธีนี้จะมีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบแล้ว ยังมีการพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระด้วย ค่าAMSE แปรผกผันกับขนาดตัวอย่าง กล่าวคือ การเพิ่มขนาดตัวอย่างส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มลดลงเพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

ผลการวิจัยของวิธี BMA_{SVT} จะให้ค่า RDAMSE น้อยที่สุด รองลงมาคือวิธี OPM และวิธี SR ตามลำดับ โดยส่วนใหญ่ค่า RDAMSE ของวิธี SR และ OPM มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

<u>สรุปตอนที่ 4.2</u> ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ สำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (1,10)

เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (1,10)$ พบว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุก ๆ กรณี โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย และวิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

เนื่องจากการวิจัยตอนที่ 4.2 มีการเปลี่ยนแปลงค่าคงที่ c จากตอนที่ 4.1 เล็กน้อย (ตอนที่ 4.1 กำหนดให้ค่าคงที่ c เท่ากับ 5 และตอนที่ 4.2 กำหนดให้ค่าคงที่ c เท่ากับ10) จึงส่งผลให้ค่า AMSE ของวิธี BMA_{SVT} และวิธี OPM มีการเปลี่ยนแปลงไม่มากนัก

จากผลการวิจัยในตอนที่ 4.2 การเปลี่ยนแปลงของค่า AMSE และค่า RDAMSE มีลักษณะดังนี้

1) <u>ค่า AMSE</u>
   (1) แปรผันตามส่วนเบี่ยงเบนมาตรฐาน จำนวนตัวแปรอิสระและค่าคงที่ σ_β / τ และ c (เฉพาะวิธี BMA_{SVT}และ OPM เท่านั้น) ตามลำดับ
   (2) แปรผกผันกับขนาดตัวอย่าง
2) <u>ค่า RDAMSE</u> แปรผันตามขนาดตัวอย่าง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4.3 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (10,100) ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตารางที่ 4.13 – 4.18 และรูปที่ 4.13 – 4.18

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.13 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ (σ_β / τ , c) = (10,100)

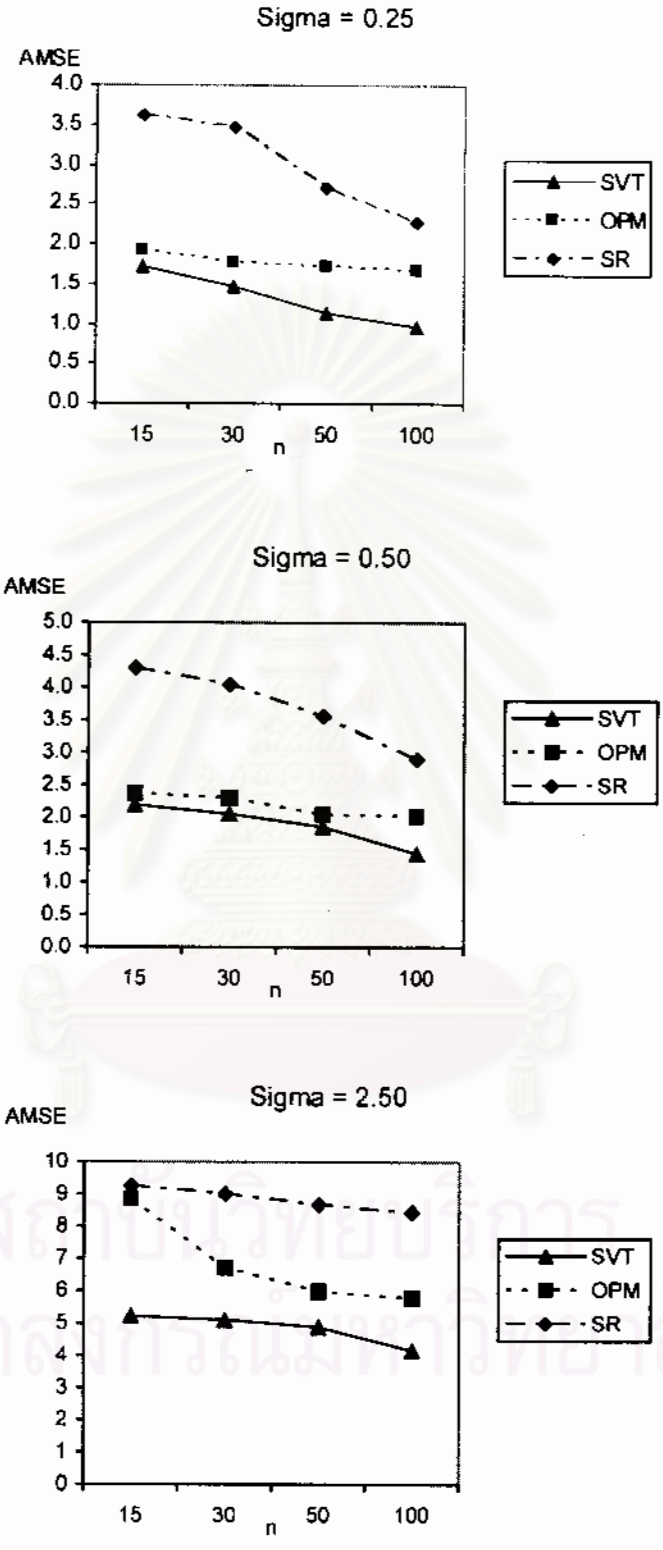
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	15	3.6208	1.9237	1.7292
		(0.3575)	(0.1446)	(0.1356)
		109.3916	11.2480	0.0000
	30	3.4653	1.7786	1.4689
		(0.3377)	(0.1242)	(0.1124)
		135.9112	21.0838	0.0000
	50	2.6983	1.7329	1.1502
		(0.2644)	(0.1196)	(0.1012)
		134.5940	50.6608	0.0000
	100	2.2940	1.6604	0.9485
		(0.2184)	(0.1186)	(0.0813)
		141.8556	75.0554	0.0000
0.50	15	4.3033	2.3661	2.1857
		(0.3504)	(0.1957)	(0.1785)
		96.8843	8.2536	0.0000
	30	4.0477	2.3031	2.0505
		(0.3496)	(0.1904)	(0.1545)
		97.4006	12.3189	0.0000
	50	3.5601	2.0400	1.8498
		(0.2679)	(0.1600)	(0.1218)
		97.8646	10.2822	0.0000
	100	2.8932	2.0234	1.4422
		(0.2069)	(0.1548)	(0.1072)
		100.6102	40.2995	0.0000
2.50	15	9.2683	8.4498	5.2171
		(0.9105)	(0.6275)	(0.4843)
		77.6523	61.9635	0.0000
	30	9.0102	6.7312	5.1086
		(0.8972)	(0.6083)	(0.4567)
		76.3732	31.7621	0.0000
	50	8.6844	5.9898	4.8940
		(0.8451)	(0.5292)	(0.3825)
		77.4499	22.3907	0.0000
	100	8.4321	5.7862	4.1649
		(0.7989)	(0.4468)	(0.3474)
		102.4562	38.9277	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.13 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,100) (ตารางที่ 4.13 และรูปที่ 4.13) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ $(\sigma_\beta / \tau, c)$ (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ 10 เท่าของตอนที่ 4.2) ส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.7 และตารางที่ 4.13) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง และเป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} ชัดเจนมากกว่าผลการวิจัยตอนที่ 4.1 และ 4.2 เมื่อ σ มีค่าเท่ากับ 2.50

ค่า AMSE <u>แปรผันตาม</u>ค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น และค่า AMSE <u>แปรผกผัน</u>กับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้ ดังนั้นการเพิ่มขนาดตัวอย่างส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มลดลง

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.50 วิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 50 ทั้งนี้อาจเป็นเพราะวิธีBMA_{SVT}ให้ค่า AMSE ลดลงเพียงเล็กน้อยเมื่อเทียบกับวิธี OPM

ตารางที่ 4.14 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, C) = (10,100)$

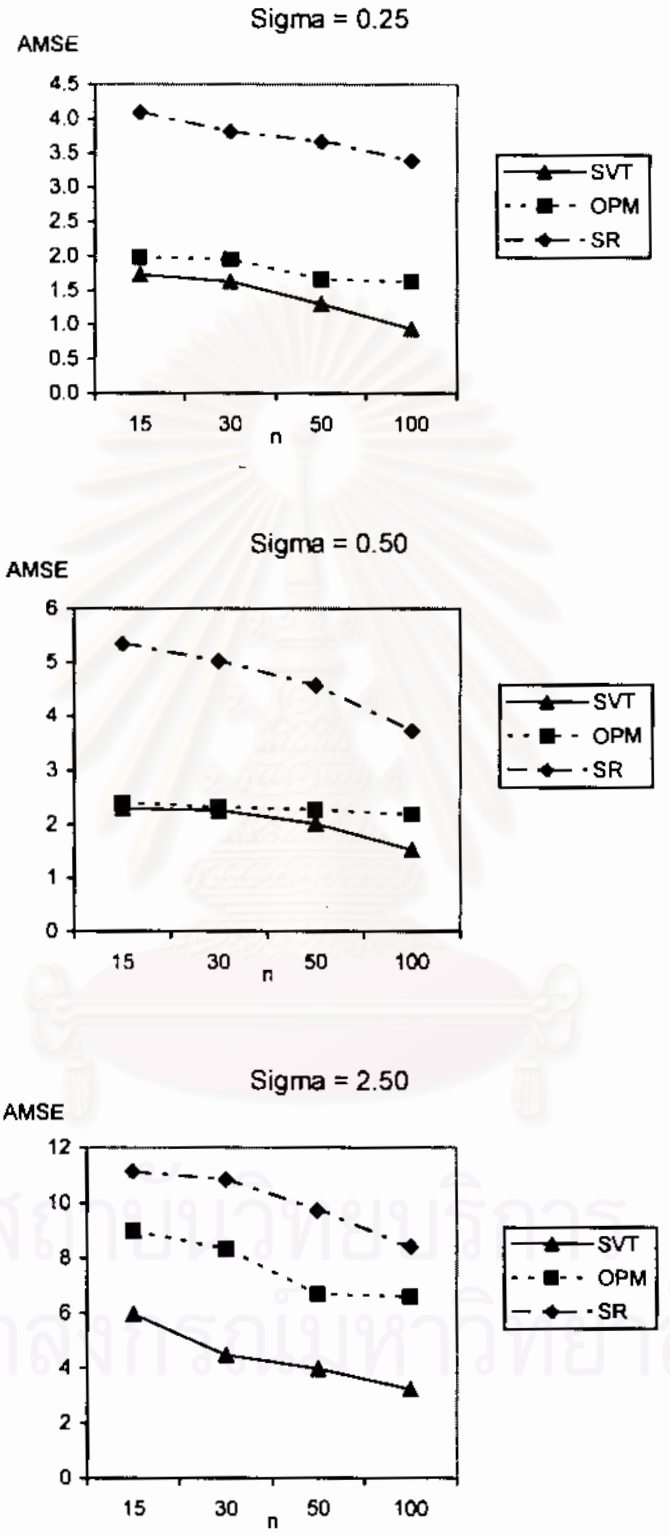
σ	n	วิธีการ		
		SR	**OPM**	**BMA_{SVT}**
0.25	15	4.0921	1.9793	1.7312
		(0.3932)	(0.1511)	(0.1476)
		136.3736	14.3311	0.0000
	30	3.8085	1.9549	1.6266
		(0.3754)	(0.1489)	(0.1380)
		134.1387	20.1832	0.0000
	50	3.6714	1.6555	1.3004
		(0.3632)	(0.1386)	(0.1068)
		182.3285	27.3070	0.0000
	100	3.3931	1.6170	0.9416
		(0.3174)	(0.1250)	(0.0775)
		260.3547	71.7290	0.0000
0.50	15	5.3451	2.3968	2.2995
		(0.5022)	(0.2191)	(0.1952)
		132.4462	4.2314	0.0000
	30	5.0153	2.3180	2.2646
		(0.4734)	(0.1918)	(0.1931)
		121.4652	2.3580	0.0000
	50	4.5663	2.2711	2.0081
		(0.4336)	(0.1838)	(0.1508)
		127.3941	13.0970	0.0000
	100	3.7318	2.1841	1.5237
		(0.2869)	(0.1743)	(0.1036)
		144.9170	43.3418	0.0000
2.50	15	11.1307	8.9098	5.9504
		(1.1027)	(0.8646)	(0.5706)
		87.0580	49.7345	0.0000
	30	10.8464	8.3353	4.4741
		(1.0437)	(0.6970)	(0.4008)
		142.4264	86.3012	0.0000
	50	9.7264	6.6849	3.9584
		(0.9549)	(0.6013)	(0.3498)
		145.7154	68.8788	0.0000
	100	8.3907	6.5721	3.2246
		(0.8349)	(0.5932)	(0.2611)
		160.2090	103.8113	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.14 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,100) (ตารางที่ 4.14 และรูปที่ 4.14) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่$(\sigma_\beta / \tau, c)$ (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ 10 เท่าของตอนที่ 4.2) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อยซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.8และตารางที่ 4.14) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 3 ตัวแปรเป็น 5 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยง เบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธีBMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน

ตารางที่ 4.15 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

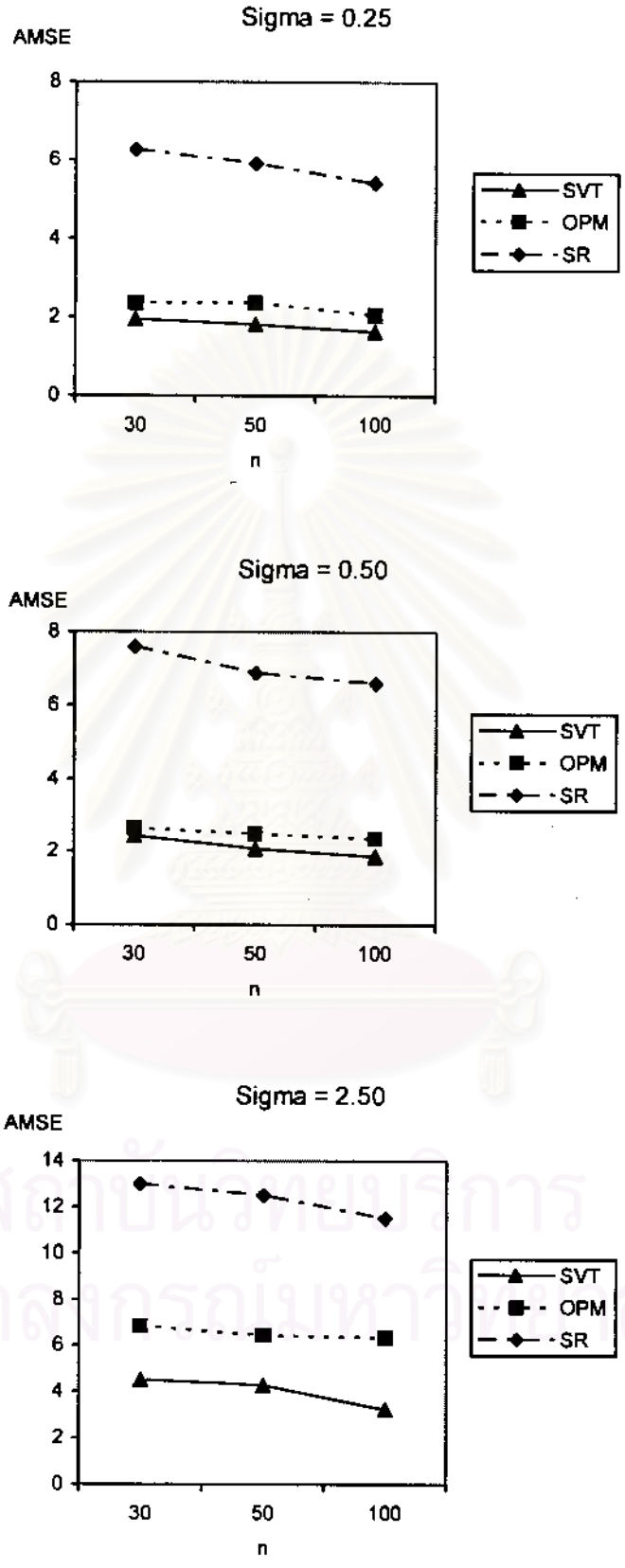
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	6.2668	2.3633	1.9460
		(0.5764)	(0.1687)	(0.1479)
		222.0349	21.4440	0.0000
	50	5.9160	2.3584	1.8150
		(0.5851)	(0.1556)	(0.1253)
		225.9504	29.9394	0.0000
	100	5.4073	2.0623	1.6347
		(0.5250)	(0.1215)	(0.1047)
		230.7824	26.1577	0.0000
0.50	30	7.5891	2.6320	2.4346
		(0.6733)	(0.1897)	(0.1695)
		211.7186	8.1081	0.0000
	50	6.8660	2.4712	2.0727
		(0.6742)	(0.1753)	(0.1204)
		231.2587	19.2261	0.0000
	100	6.5892	2.3541	1.8721
		(0.5746)	(0.1692)	(0.1188)
		251.9684	25.7465	0.0000
2.50	30	12.9712	6.8446	4.5175
		(1.1136)	(0.6291)	(0.3877)
		187.1323	51.5130	0.0000
	50	12.5017	6.4267	4.2735
		(1.0716)	(0.5649)	(0.3671)
		192.5401	50.3849	0.0000
	100	11.5134	6.3501	3.2289
		(1.0007)	(0.5546)	(0.2847)
		256.5734	96.6645	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.15 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ (σ_β / τ, c) เหมือนกันคือ (10,100) (ตารางที่ 4.15 และรูปที่ 4.15) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ ชัดเจนขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่(σ_β / τ, c) (ตอนที่ 4.3 กำหนดให้ค่าคงที่ (σ_β / τ, c) มีค่าเท่ากับ 10 เท่าของตอนที่ 4.2) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อยซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.9 และตารางที่ 4.15) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง และเป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} ชัดเจนมากกว่าผลการวิจัยตอนที่ 4.1 และ 4.2 เมื่อ σ มีค่าเท่ากับ 2.50

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่าAMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 5 ตัวแปรเป็น 8 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น

ตารางที่ 4.16 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

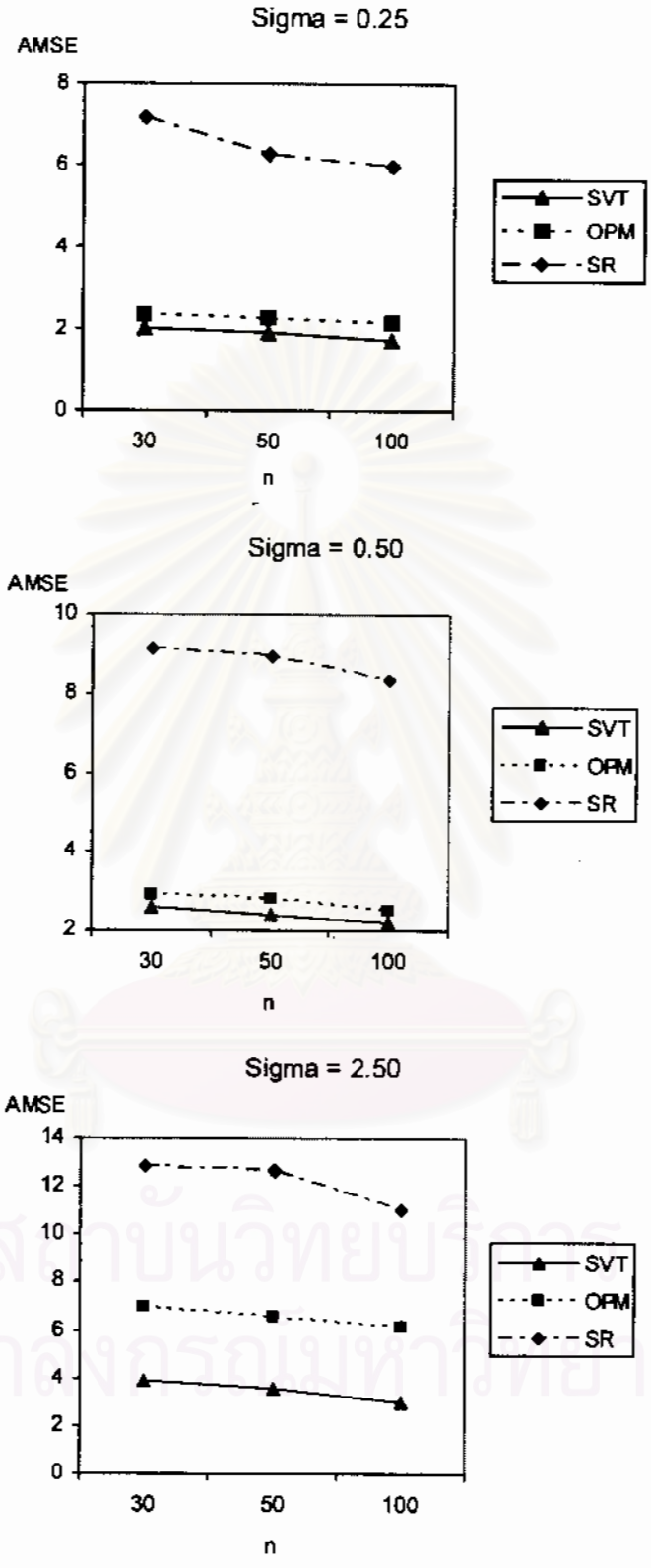
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	7.1731	2.3636	2.0127
		(0.7134)	(0.2163)	(0.1805)
		256.3919	17.4343	0.0000
	50	6.2773	2.2522	1.9082
		(0.5943)	(0.1912)	(0.1682)
		228.9645	18.0275	0.0000
	100	5.9778	2.1521	1.7198
		(0.5483)	(0.1820)	(0.1413)
		247.5869	25.1366	0.0000
0.50	30	9.4188	2.9328	2.6137
		(0.8619)	(0.2665)	(0.1910)
		249.2637	12.2087	0.0000
	50	8.9530	2.8007	2.4076
		(0.8549)	(0.2328)	(0.1715)
		271.8641	16.3275	0.0000
	100	8.3233	2.5215	2.1931
		(0.7753)	(0.2115)	(0.1213)
		279.5221	14.5638	0.0000
2.50	30	12.8127	7.0001	3.9012
		(1.2621)	(0.6710)	(0.3511)
		228.4297	79.4345	0.0000
	50	12.6754	6.5706	3.5644
		(1.1351)	(0.5863)	(0.2938)
		255.6110	84.3396	0.0000
	100	11.0226	6.0987	3.0101
		(1.0581)	(0.5111)	(0.2526)
		266.1872	102.6079	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองศ์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.16 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,100) (ตารางที่ 4.16 และรูปที่ 4.16) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ ชัดเจนมากขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่$(\sigma_\beta / \tau, c)$ (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ 10 เท่าของตอนที่ 4.2) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อยซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.10 และตารางที่ 4.16) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์ส้มประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง และเป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} ชัดเจนมากกว่าผลการวิจัยตอนที่ 4.1 และ 4.2

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 8 ตัว แปรเป็น 10 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.50 วิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 50 เป็น 100 ทั้งนี้อาจเป็นเพราะวิธี BMA_{SVT}ให้ค่า AMSE ลดลงเพียงเล็กน้อยเมื่อเทียบกับวิธี OPM

ตารางที่ 4.17 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

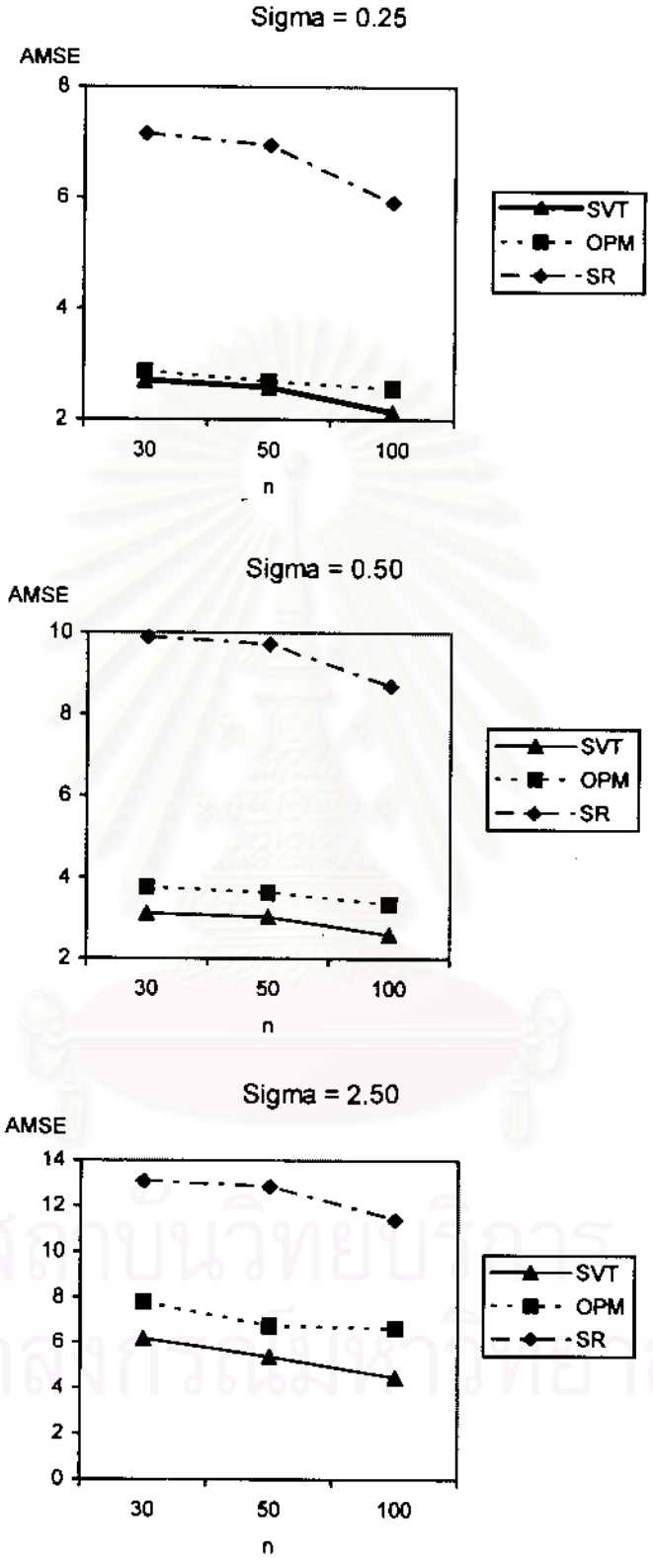
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	7.1604	2.8560	2.7012
		(0.7102)	(0.2415)	(0.2336)
		165.0822	5.7308	0.0000
	50	6.9352	2.6752	2.5828
		(0.6053)	(0.2316)	(0.2124)
		168.5148	3.5775	0.0000
	100	5.9071	2.5580	2.1258
		(0.5587)	(0.2210)	(0.1810)
		177.8766	20.3312	0.0000
0.50	30	9.8937	3.7488	3.1052
		(0.9897)	(0.2975)	(0.2834)
		218.6172	20.7265	0.0000
	50	9.7307	3.6214	3.0167
		(0.9223)	(0.2762)	(0.2354)
		222.5611	20.0451	0.0000
	100	8.7028	3.3384	2.5849
		(0.8431)	(0.2638)	(0.1621)
		236.6784	29.1501	0.0000
2.50	30	13.0857	7.7564	6.1674
		(1.2935)	(0.6279)	(0.3557)
		112.1753	25.7645	0.0000
	50	12.8682	6.7411	5.3737
		(1.1269)	(0.5898)	(0.3011)
		139.4663	25.4462	0.0000
	100	11.3963	6.6147	4.4790
		(1.0087)	(0.5752)	(0.2948)
		154.4385	47.6825	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.17 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,100) (ตารางที่ 4.17 และรูปที่ 4.17) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่าAMSE สูงสุด โดยวิธี OPM มีค่าAMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่าAMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่$(\sigma_\beta / \tau, c)$ (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ 10 เท่าของตอนที่ 4.2) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM มากพอควร ซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.11 และตารางที่ 4.17) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง และเป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} ชัดเจนมากกว่าผลการวิจัยตอนที่ 4.1 และ 4.2

ค่า AMSE <u>แปรผันตาม</u>ค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยัง<u>แปรผันตาม</u>จำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 10 ตัวแปรเป็น 12 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่าAMSE <u>แปรผกผัน</u>กับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน

ตารางที่ 4.18 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

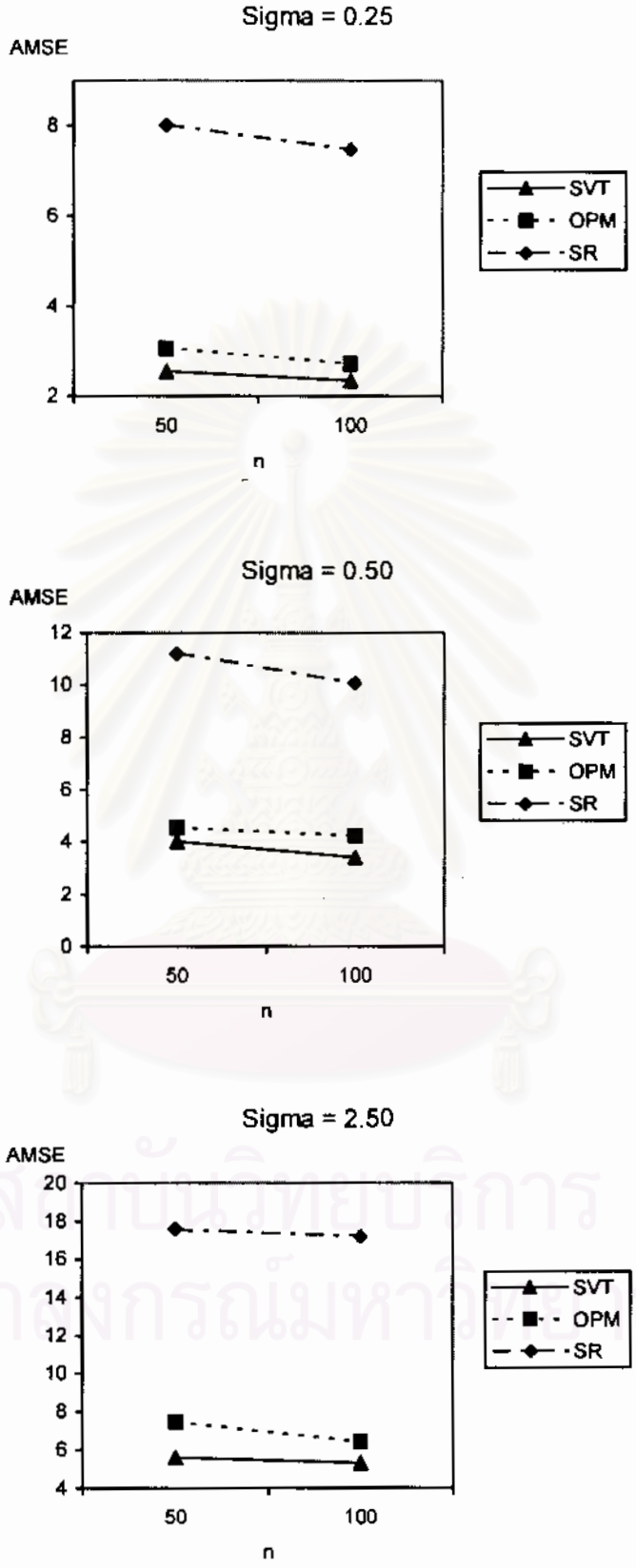
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	50	8.0144	3.0640	2.5364
		(0.7717)	(0.2509)	(0.2031)
		215.9754	20.8011	0.0000
	100	7.4655	2.7103	2.3351
		(0.6722)	(0.2413)	(0.1930)
		219.7079	16.0678	0.0000
0.50	50	11.2100	4.5323	4.0091
		(1.0961)	(0.3745)	(0.3374)
		179.6139	13.0503	0.0000
	100	10.0615	4.2198	3.4104
		(0.9697)	(0.3063)	(0.2679)
		195.0240	23.7333	0.0000
2.50	50	17.5690	7.4389	5.6143
		(1.6491)	(0.6452)	(0.4249)
		212.9330	32.4992	0.0000
	100	17.1906	6.4091	5.3158
		(1.6405)	(0.6046)	(0.3927)
		223.3869	20.5670	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.18 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,100)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ (σ_β / τ, c) เหมือนกันคือ (10,100) (ตารางที่ 4.18 และรูปที่ 4.18) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่(σ_β / τ, c) (ตอนที่ 4.3 กำหนดให้ค่าคงที่ (σ_β / τ, c) มีค่าเท่ากับ 10 เท่าของตอนที่ 4.2) ส่งผลกระทบต่อวิธี BMA_{SVT}และวิธี OPM เพียงเล็กน้อยซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี(ตารางที่ 4.12 และตารางที่ 4.18) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง และเป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} ชัดเจนมากกว่าผลการวิจัยตอนที่ 4.1 และ 4.2

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น และค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 50 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 100 ยังคงให้ผลสรุปเหมือนกัน

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่(σ_β / τ, c) เหมือนกันคือ (10,100) (ตารางที่ 4.13 – 4.18 และรูปที่ 4.13 – 4.18) พบว่า วิธี BMA_{SVT} ให้ค่าAMSE ต่ำสุด รองลงมาคือ วิธี OPM และวิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่าσ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่าσ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

ค่าคงที่ σ_β / τ และ c จะมีผลต่อวิธี BMA_{SVT} และ OPM เท่านั้น โดยค่า AMSE จะเพิ่มขึ้นเมื่อค่าคงที่ σ_β / τ และ c ที่สูงขึ้น เนื่องจากค่าคงที่ σ_β / τ และ c เป็นค่าที่กำหนดลักษณะการกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยที่สุ่มได้ ค่าคงที่ σ_β / τ และ c ที่สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น ทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้ค่า AMSE มีค่าสูงขึ้น และจากผลการวิจัยตอนที่ 4.3 นี้ ค่าคงที่σ_β / τ และ c เปลี่ยนแปลงจากตอนที่ 4.1 และ 4.2 มากพอควร จึงส่งผลให้ค่า AMSE ของวิธี BMA_{SVT} และวิธี OPM มีการเปลี่ยนแปลงมากขึ้น โดยเฉพาะวิธี OPM และเป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} ชัดเจนมากกว่าผลการวิจัยตอนที่ 4.1 และ 4.2 โดยเฉพาะเมื่อค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่ามาก($\sigma \geq 0.50$)

ค่า AMSE <u>แปรผันตาม</u>ค่าσ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2) ดังนั้นการเพิ่มค่าσ นั้นจะทำให้ค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอก จากนี้ค่า AMSE ยัง<u>แปรผันตาม</u>จำนวนตัวแปรอิสระด้วย เนื่องจากในการวิจัยครั้งนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูปเมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้นทำให้ค่าAMSE ของทุกวิธีมีแนวโน้มสูงขึ้นโดยเฉพาะอย่างยิ่งวิธี SR ซึ่งใช้หลักการในการคัดเลือกตัวแปรเข้าออกจากตัวแบบผ่านการทดสอบสมมติฐานเท่านั้น ในขณะที่วิธีการภายใต้แนวทางของเบส์ทั้ง 2 วิธี คือ วิธี BMA_{SVT} และ OPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการพิจารณาถึงความเหมาะสมของตัวแบบด้วย เช่น มีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT} จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น เพราะนอกจากวิธีนี้จะมีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบแล้วยังมีการ

พิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระด้วย ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

ผลการวิจัยของวิธี BMA_{SVT} จะให้ค่า RDAMSE น้อยที่สุด รองลงมาคือวิธี OPM และวิธี SR ตามลำดับ โดยส่วนใหญ่ค่า RDAMSE ของวิธี SR และ OPM มีแนวโน้มเพิ่มขึ้นเมื่อค่า ขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

<u>สรุปตอนที่ 4.3</u> ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ สำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,100)

เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,100) พบว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุก ๆ กรณี โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย และวิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} ชัดเจนมากขึ้นเมื่อเปรียบเทียบกับผลการวิจัยตอนที่ 4.1 และ 4.2 โดยเฉพาะเมื่อค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่ามาก($\sigma \geq 0.50$)

จากผลการวิจัยในตอนที่ 4.3 การเปลี่ยนแปลงของค่า AMSE และค่า RDAMSE มีลักษณะดังนี้

1) <u>ค่า AMSE</u>
   (1) แปรผันตามส่วนเบี่ยงเบนมาตรฐาน จำนวนตัวแปรอิสระและค่าคงที่ σ_β / τ และ c (เฉพาะวิธี BMA_{SVT}และ OPM เท่านั้น) ตามลำดับ
   (2) แปรผกผันกับขนาดตัวอย่าง
2) <u>ค่า RDAMSE</u> แปรผันตามขนาดตัวอย่าง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4.4 ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ซึ่งกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 และศึกษาในกรณีที่ความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่ค่าคงที่ของวิธี BMA_{SVT} และวิธี QPM จะศึกษาเมื่อ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (10,500) ซึ่งผลการศึกษาที่ได้ในแต่ละสถานการณ์แสดงในตาราง 4.19 – 4.24 และรูปที่ 4.19 – 4.24

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.19 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, C) = (10,500)$

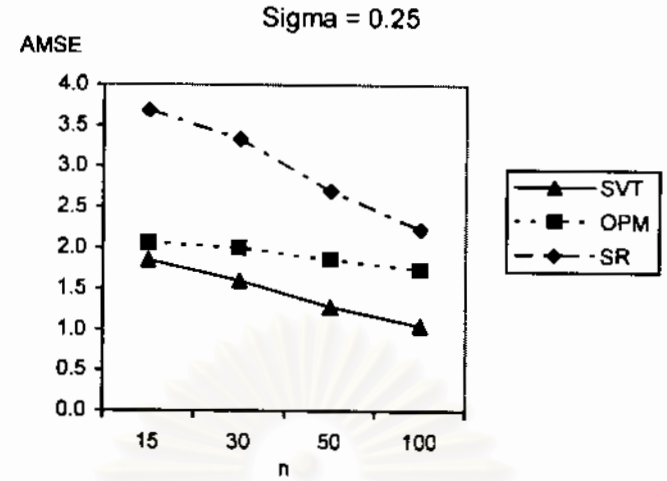
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	15	3.6859	2.0603	1.8456
		(0.3755)	(0.1604)	(0.1241)
		99.7128	11.6331	0.0000
	30	3.3278	2.0057	1.6037
		(0.2547)	(0.1217)	(0.1131)
		107.5076	25.0677	0.0000
	50	2.7028	1.8649	1.2806
		(0.2049)	(0.1177)	(0.0969)
		111.0573	45.6270	0.0000
	100	2.2294	1.7418	1.0488
		(0.1996)	(0.1089)	(0.0809)
		112.5667	66.0755	0.0000
0.50	15	4.4039	2.5713	2.2190
		(0.3826)	(0.2183)	(0.1649)
		98.4633	15.8765	0.0000
	30	3.8916	2.3083	1.8265
		(0.3130)	(0.1851)	(0.1358)
		113.0632	26.3783	0.0000
	50	3.5092	2.2446	1.5934
		(0.2687)	(0.1732)	(0.1139)
		120.2335	40.8686	0.0000
	100	2.7057	2.0234	1.2034
		(0.2172)	(0.1534)	(0.0988)
		124.8380	68.1403	0.0000
2.50	15	9.2899	8.5452	5.2317
		(0.9114)	(0.8101)	(0.4253)
		77.5694	63.3351	0.0000
	30	9.1392	6.7916	5.1103
		(0.8948)	(0.6318)	(0.4174)
		78.8388	32.9002	0.0000
	50	8.9969	6.0776	4.9842
		(0.8405)	(0.5472)	(0.3831)
		80.5084	21.9373	0.0000
	100	8.7487	5.7862	4.1721
		(0.8188)	(0.5167)	(0.3549)
		109.6954	38.6880	0.0000

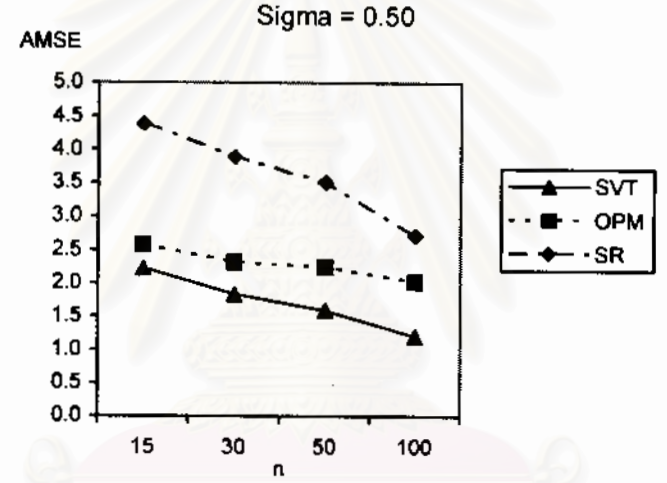
ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

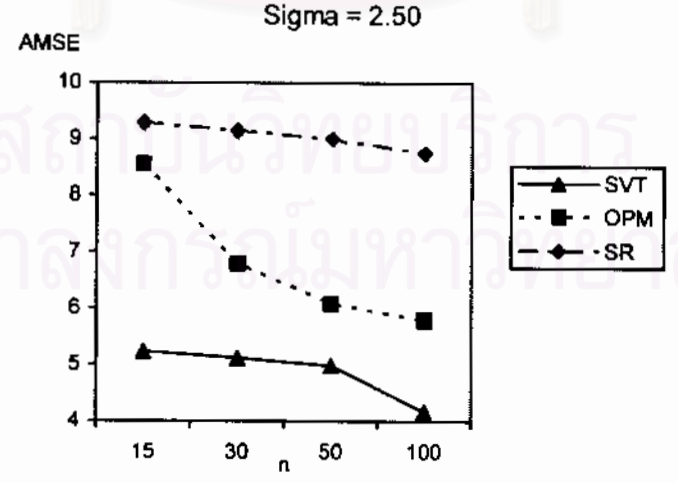
SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ







สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.19 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$เหมือนกันคือ (10,500) (ตารางที่ 4.19 และรูปที่ 4.19) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่าAMSEเป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,100)$ และตอนที่ 4.4 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,500)$) ส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี (ตารางที่ 4.13และตารางที่ 4.19) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้นจึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จึงส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นและค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้ ดังนั้นการเพิ่มขนาดตัวอย่างจึงส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มลดลง

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธีBMA_{SVT} ทำให้ค่าAMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 2.50 นั้นวิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 50 ทั้งนี้อาจเป็นเพราะวิธี OPM ให้ค่าAMSE ลดลงมากเมื่อเทียบกับวิธี BMA_{SVT}

ตารางที่ 4.20 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

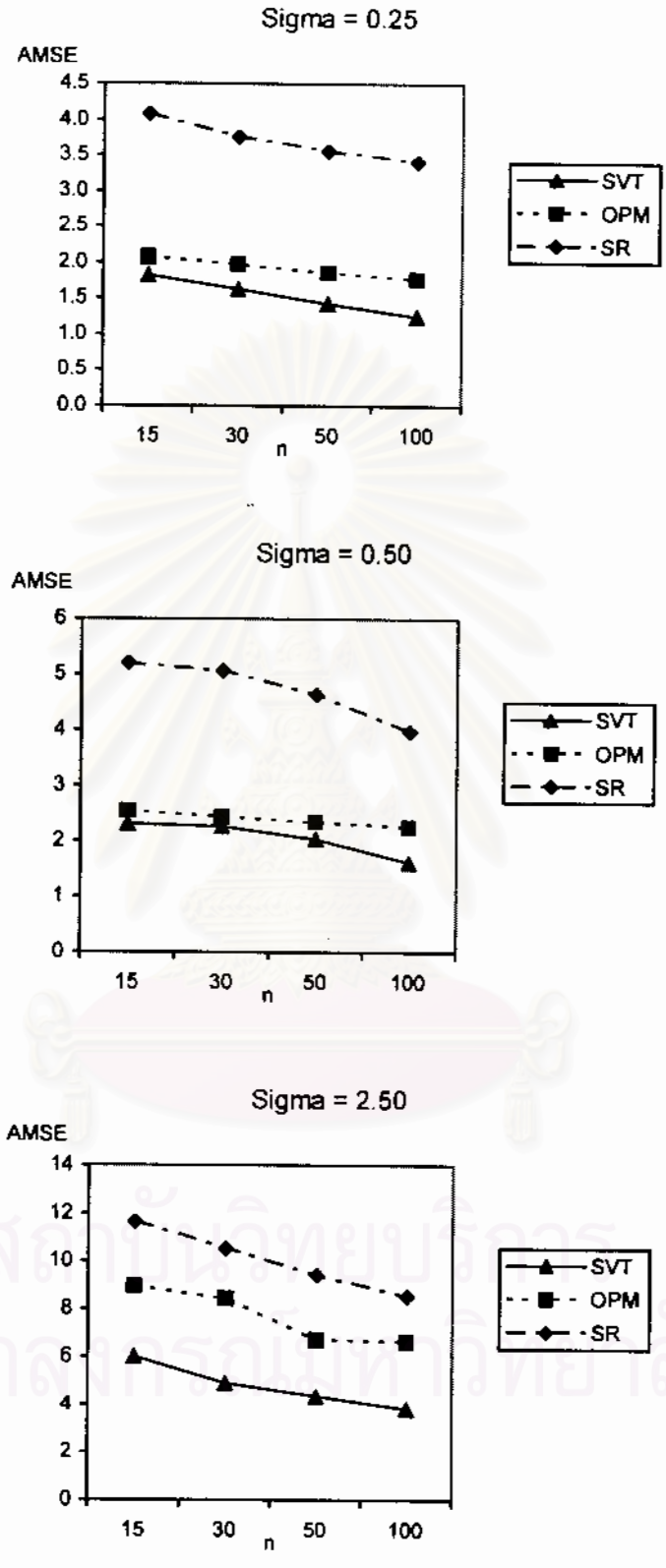
σ	n	วิธีการ		
		SR	**OPM**	**BMA_{SVT}**
0.25	15	4.0765	2.0685	1.8243
		(0.3537)	(0.1623)	(0.1455)
		123.4556	13.3860	0.0000
	30	3.7636	1.9741	1.6275
		(0.3479)	(0.1472)	(0.1227)
		131.2504	21.2965	0.0000
	50	3.5661	1.8581	1.4293
		(0.3568)	(0.1457)	(0.1075)
		149.4998	30.0007	0.0000
	100	3.4106	1.7667	1.2469
		(0.2922)	(0.1365)	(0.0836)
		173.5263	41.6874	0.0000
0.50	15	5.1961	2.5418	2.3063
		(0.4943)	(0.2110)	(0.1948)
		125.3003	10.2112	0.0000
	30	5.0699	2.4431	2.2655
		(0.4816)	(0.2049)	(0.1920)
		123.7872	7.8393	0.7888
	50	4.6362	2.3407	2.0351
		(0.4730)	(0.1940)	(0.1522)
		127.8119	15.0165	0.0000
	100	3.9811	2.2554	1.6215
		(0.3929)	(0.1928)	(0.1025)
		145.5196	39.0934	0.0000
2.50	15	11.6749	8.9223	6.0021
		(1.1232)	(0.8591)	(0.5166)
		94.5136	48.6530	0.0000
	30	10.5261	8.4179	4.9129
		(0.9741)	(0.7501)	(0.4628)
		114.2543	71.3428	0.0000
	50	9.4021	6.7225	4.3643
		(0.8851)	(0.5816)	(0.3810)
		115.4320	54.0339	0.0000
	100	8.5261	6.6292	3.8220
		(0.8113)	(0.57277)	(0.2779)
		123.0795	73.4485	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.20 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 5 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,500) (ตารางที่ 4.20 และรูปที่ 4.20) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,100)$ และตอนที่ 4.4 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,500)$) ส่งผลกระทบต่อวิธี BMA_{SVT} และวิธี OPM เพียงเล็กน้อย ซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี (ตารางที่ 4.14 และตารางที่ 4.20) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 3 ตัวแปรเป็น 5 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่าAMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 15 วิธี BMA_{SVT} ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 30 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.50 และ 2.50 นั้นวิธี OPM ให้ ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 15 เป็น30 และจาก 30 เป็น 50 ตามลำดับ ทั้งนี้อาจเป็นเพราะวิธี OPM ให้ค่า AMSE ลดลงมากเมื่อเทียบกับวิธี BMA_{SVT}

ตารางที่ 4.21 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

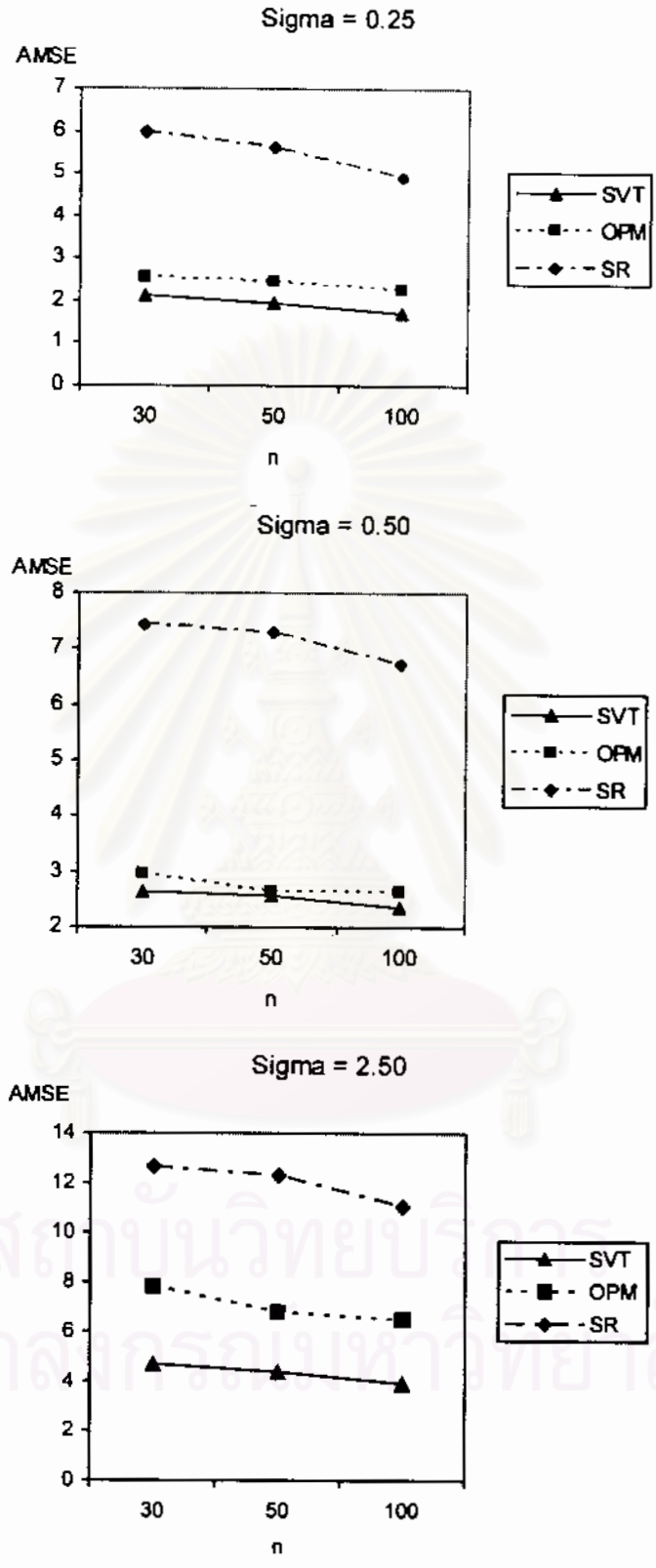
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	5.9791	2.5604	2.1087
		(0.5723)	(0.1697)	(0.1429)
		183.5444	21.4208	0.0000
	50	5.6016	2.4547	1.9723
		(0.5396)	(0.1483)	(0.1265)
		184.0136	24.4588	0.0000
	100	4.9026	2.2546	1.7092
		(0.5179)	(0.1274)	(0.1072)
		186.8359	31.9097	0.0000
0.50	30	7.4391	2.9382	2.6468
		(0.7232)	(0.2713)	(0.2201)
		181.0601	11.0095	0.0000
	50	7.2986	2.6266	2.5722
		(0.7004)	(0.2464)	(0.2037)
		183.7493	2.1149	0.0000
	100	6.7157	2.6150	2.3638
		(0.6366)	(0.2341)	(0.1842)
		184.1061	10.6270	0.0000
2.50	30	12.6755	7.8360	4.6837
		(1.1232)	(0.6821)	(0.3992)
		170.6301	67.3036	0.0000
	50	12.3249	6.7741	4.3666
		(1.0542)	(0.6130)	(0.3451)
		182.2539	55.1344	0.0000
	100	11.0691	6.4884	3.8997
		(1.0065)	(0.6093)	(0.2845)
		183.8449	66.3820	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.21 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 8 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$เหมือนกันคือ (10,500) (ตารางที่ 4.21 และรูปที่ 4.21) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (10,100) และตอนที่ 4.4 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c)$ มีค่าเท่ากับ (10,500)) ส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี (ตารางที่ 4.15 และตารางที่ 4.21) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจาก เมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE <u>แปรผันตาม</u>ค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยัง<u>แปรผันตาม</u>จำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 5 ตัวแปรเป็น 8 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE <u>แปรผกผัน</u>กับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธีBMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSEเพิ่มขึ้น แต่ในกรณี σ เท่ากับ 0.50 และ 2.50 นั้นวิธี OPM ให้ค่า RDAMSE ลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นจาก 30 เป็น 50 ทั้งนี้อาจเป็นเพราะวิธี OPM ให้ค่า AMSE ลดลงมากเมื่อเทียบกับวิธี BMA_{SVT}

ตารางที่ 4.22 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ (σ_β / τ , C) = (10,500)

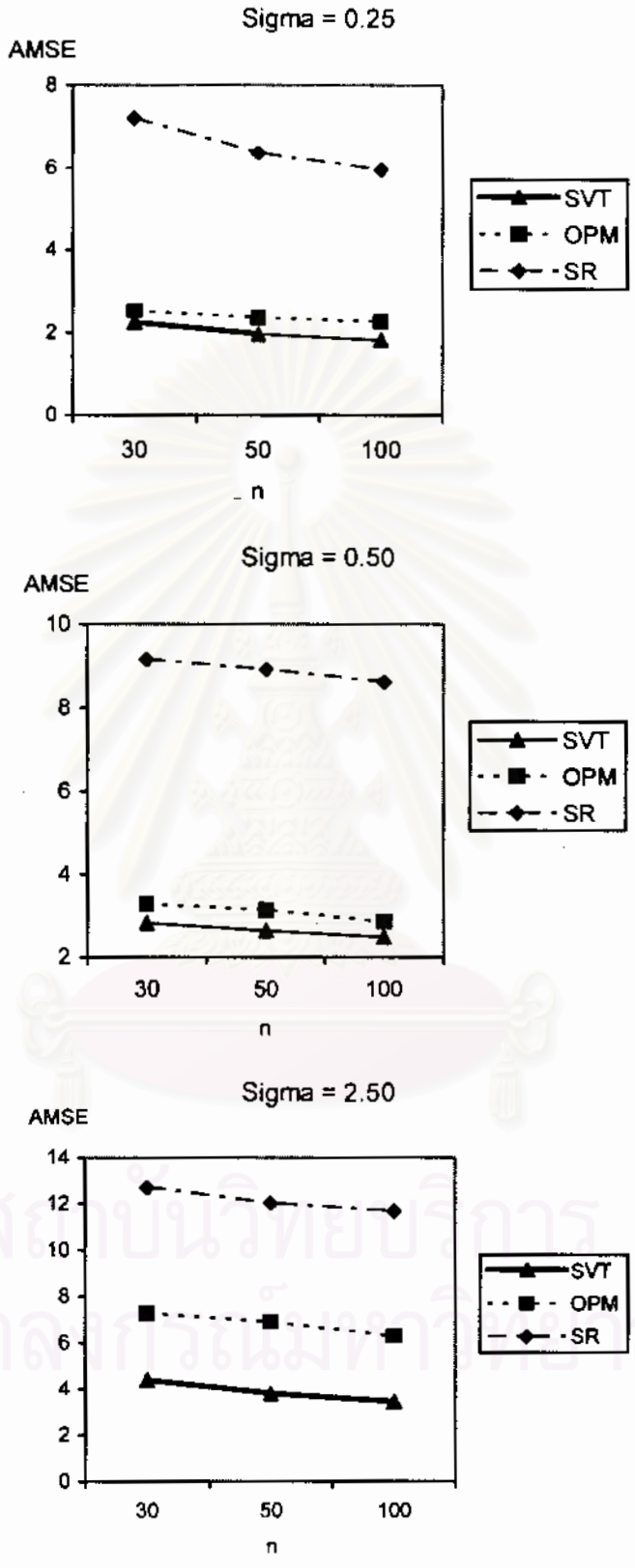
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	30	7.1976	2.5249	2.2549
		(0.7220)	(0.2219)	(0.1925)
		219.1982	11.9739	0.0000
	50	6.3590	2.3620	1.9587
		(0.6096)	(0.2075)	(0.1628)
		224.6541	20.5902	0.0000
	100	5.9508	2.2667	1.8202
		(0.5616)	(0.1989)	(0.1522)
		226.9311	24.5303	0.0000
0.50	30	9.1478	3.2727	2.8186
		(0.8817)	(0.2789)	(0.2092)
		224.5512	16.1108	0.0000
	50	8.9242	3.1357	2.6374
		(0.8978)	(0.2420)	(0.1913)
		238.3711	18.8936	0.0000
	100	8.6188	2.8528	2.4822
		(0.8068)	(0.2298)	(0.1728)
		247.2242	14.9303	0.0000
2.50	30	12.7253	7.2648	4.3766
		(1.2804)	(0.6452)	(0.3271)
		190.7577	65.9919	0.0000
	50	12.0198	6.8844	3.8010
		(1.1548)	(0.5942)	(0.3021)
		216.2273	81.1208	0.0000
	100	11.6819	6.2762	3.4253
		(1.0347)	(0.5356)	(0.2534)
		241.0475	83.2307	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปภาพที่ 4.22 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 10 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,500) (ตารางที่ 4.22 และรูปที่ 4.22) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,100)$ และตอนที่ 4.4 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,500)$) ส่งผลกระทบต่อวิธี BMA_{SVT} และวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี (ตารางที่ 4.16 และตารางที่ 4.22) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจาก เมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 8 ตัวแปรเป็น 10 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSE เพิ่มขึ้น

ตารางที่ 4.23 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

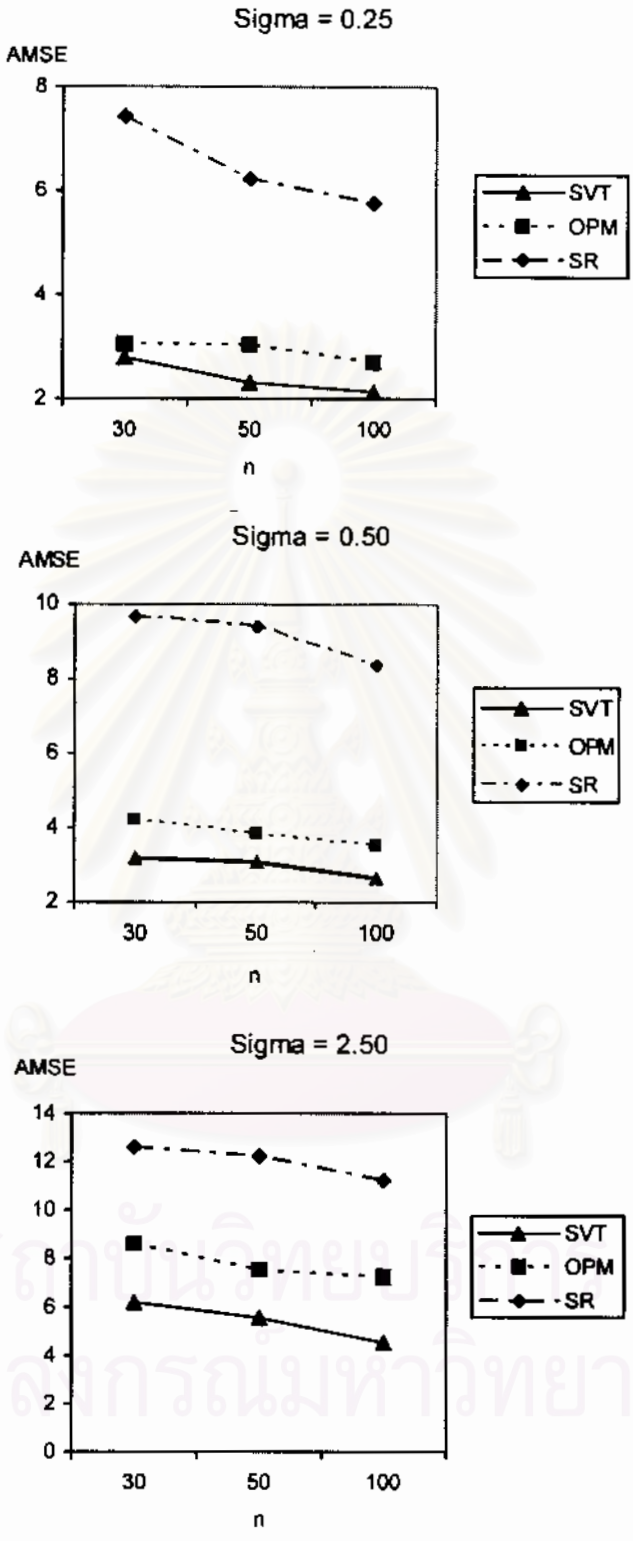
σ	n	วิธีการ		
		SR	**OPM**	**BMA_{SVT}**
0.25	**30**	7.4254	3.0562	2.8016
		(0.7168)	(0.2530)	(0.2412)
		165.0414	9.0877	0.0000
	50	6.2257	3.0451	2.3147
		(0.6002)	(0.2496)	(0.2213)
		168.9636	31.5548	0.0000
	100	5.7556	2.6961	2.1294
		(0.5338)	(0.2272)	(0.1897)
		170.2921	26.6131	0.0000
0.50	**30**	9.6903	4.2201	3.2033
		(0.9762)	(0.3635)	(0.2819)
		202.5099	31.7423	0.0000
	50	9.4065	3.8226	3.0791
		(0.9509)	(0.2951)	(0.2689)
		205.4951	24.1467	0.0000
	100	8.4044	3.5381	2.5637
		(0.8254)	(0.2723)	(0.1929)
		215.5160	32.8265	0.0000
2.50	**30**	12.6013	8.5946	6.1783
		(1.2781)	(0.7467)	(0.5656)
		103.9606	39.1095	0.0000
	50	12.2183	7.5157	5.5457
		(1.0771)	(0.6726)	(0.5127)
		120.3202	35.5230	0.0000
	100	11.2172	7.2276	4.5296
		(1.0072)	(0.5916)	(0.3467)
		147.6422	59.5638	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.23 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 12 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,500) (ตารางที่ 4.23 และรูปที่ 4.23) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,100)$ และตอนที่ 4.4 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,500)$) ส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี (ตารางที่ 4.17 และตารางที่ 4.23) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 10 ตัวแปรเป็น 12 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ขนาดตัวอย่างเท่ากับ 30 วิธี BMA_{SVT} ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 50 และ 100 ยังคงให้ผลสรุปเหมือนกัน และค่า RDAMSE ของวิธี SR มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธีBMA_{SVT} ทำให้ค่า AMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่า RDAMSE เพิ่มขึ้น

ตารางที่ 4.24 การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

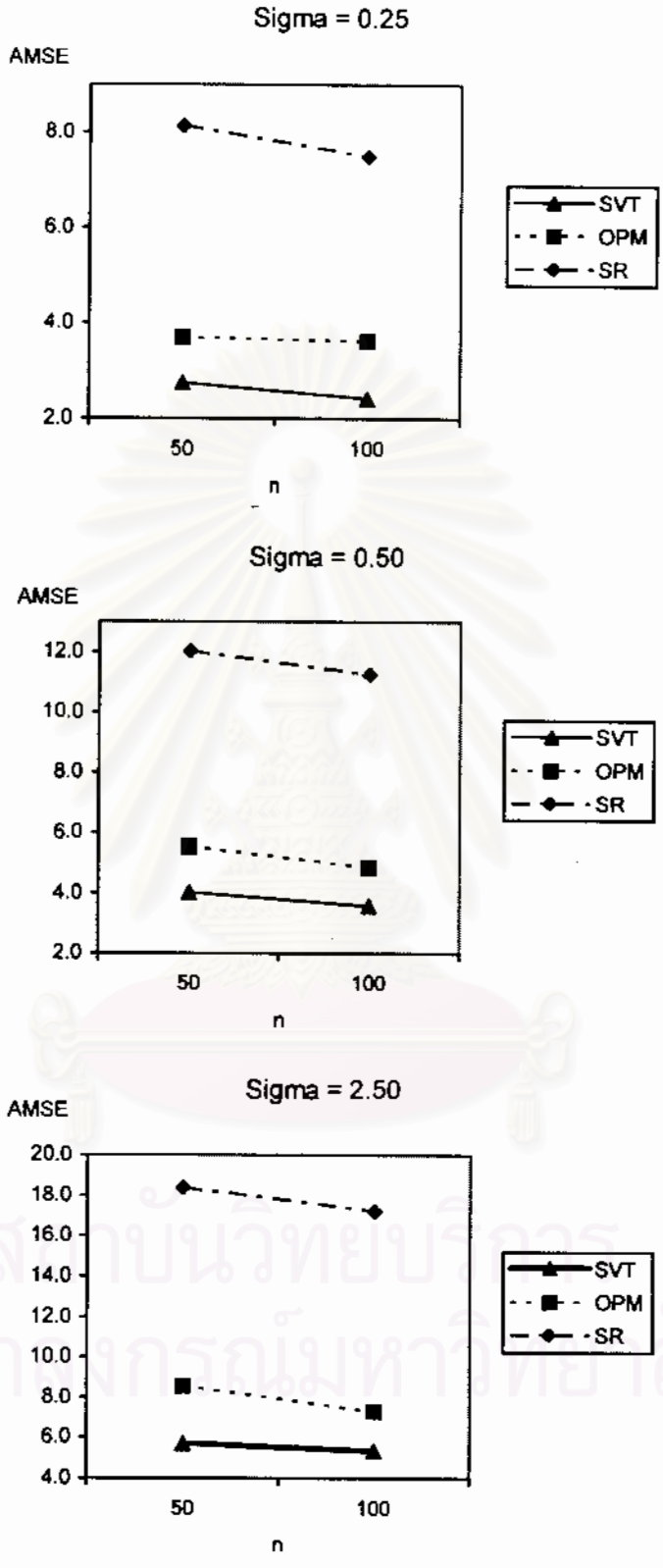
σ	n	วิธีการ		
		SR	OPM	BMA_{SVT}
0.25	50	8.1402	3.6871	2.7380
		(0.8023)	(0.2709)	(0.2484)
		197.3046	34.6640	0.0000
	100	7.4894	3.6182	2.4182
		(0.6659)	(0.2592)	(0.1959)
		209.7097	49.6237	0.0000
0.50	50	12.2058	5.5427	4.0134
		(1.1074)	(0.4820)	(0.3412)
		204.1262	38.1048	0.0000
	100	11.2512	4.8467	3.5731
		(1.0006)	(0.4026)	(0.2919)
		214.8862	35.6441	0.0000
2.50	50	18.3867	8.5312	5.6951
		(1.8689)	(0.7433)	(0.4738)
		222.8512	49.7989	0.0000
	100	17.2153	7.2481	5.3312
		(1.5452)	(0.6527)	(0.4024)
		222.9160	35.9563	0.0000

ค่าที่แสดงในแต่ละกรณีของแต่ละวิธีจะแสดงตัวเลข 3 ตัว เรียงลงมาได้แก่ 1. AMSE 2. S. D. ของ AMSE 3. RDAMSE

SR แทน วิธีการถดถอยแบบขั้นบันได

OPM แทน วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด

BMA_{SVT} แทน วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โลโดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ



รูปภาพที่ 4.24 แสดงการเปรียบเทียบค่า AMSE ของวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 24 เมื่อวิธี BMA_{SVT} และวิธี OPM มีค่าคงที่เหมือนกันคือ $(\sigma_\beta / \tau, c) = (10,500)$

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการถดถอยเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,500) (ตารางที่ 4.24 และรูปที่ 4.24) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

การเพิ่มขึ้นของค่าคงที่ c (ตอนที่ 4.3 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,100)$ และตอนที่ 4.4 กำหนดให้ค่าคงที่ $(\sigma_\beta / \tau, c) = (10,500)$) ส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM ค่อนข้างมากซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE ของทั้งสองวิธี (ตารางที่ 4.18 และตารางที่ 4.24) สาเหตุที่ทำให้ค่า AMSE มีค่าเพิ่มขึ้นเนื่องมาจากเมื่อค่าคงที่ σ_β / τ และ c สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น จึงทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง

ค่า AMSE แปรผันตามค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2) ดังนั้นการเพิ่มค่า σ จะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยังแปรผันตามจำนวนตัวแปรอิสระด้วย เมื่อเพิ่มจำนวนตัวแปรอิสระจาก 12 ตัว แปรเป็น 15 ตัวแปรจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นด้วยเหตุผลดังที่กล่าวไว้ในตอนที่ 4.1 ค่า AMSE แปรผกผันกับขนาดตัวอย่าง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

เมื่อ σ เท่ากับ 0.25 0.50 และ2.50 ขนาดตัวอย่างเท่ากับ 50 วิธีBMA_{SVT}ให้ค่า RDAMSE ต่ำสุด รองลงมาคือวิธี OPM และ SR ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มขึ้นเป็น 100 ยังคงให้ผลสรุปเหมือนกัน

จากผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติสำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT} และวิธี OPM มีค่าคงที่(σ_β / τ, c) เหมือนกันคือ (10,500) (ตารางที่ 4.19 – 4.24 และรูปที่ 4.19 – 4.24) พบว่า วิธี BMA_{SVT} ให้ค่า AMSE ต่ำสุด รองลงมาคือ วิธี OPM และ วิธี SR ให้ค่า AMSE สูงสุด โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนการเปลี่ยนค่า σ จะมีการสรุปผลการวิจัยไม่แตกต่างกันโดยค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้นเมื่อเพิ่มค่า σ ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2)

เมื่อค่าคงที่ σ_β / τ และ c เพิ่มขึ้น จะส่งผลให้ค่า AMSE เพิ่มขึ้น เนื่องจากค่าคงที่ σ_β / τ และ c เป็นค่าที่กำหนดลักษณะการกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยที่สุ่มได้ ค่าคงที่ σ_β / τ และ c ที่สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น ทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้ค่า AMSE มีค่าสูงขึ้น การเปลี่ยนแปลงของค่าคงที่นี้จะส่งผลกระทบต่อวิธี BMA_{SVT} และ วิธี OPM เพียงเล็กน้อยซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE เมื่อค่าคงที่ σ_β / τ และ c เพิ่มขึ้น แต่อย่างไรก็ตามการเปลี่ยนแปลงของค่า AMSE ของวิธี BMA_{SVT} และวิธี OPM นั้น ยังคงทำให้ค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุก ๆ กรณี จึงทำให้ข้อสรุปต่าง ๆ ยังคงเหมือนกับผลการวิจัยตอนที่ 4.1 ทุกประการ แต่เป็นที่สังเกตว่าค่า AMSE ของวิธี OPM จะสูงกว่าค่า AMSE ของวิธี BMA_{SVT} อย่างชัดเจน โดยเฉพาะเมื่อค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่ามาก($\sigma \geq 0.50$)

ค่า AMSE <u>แปรผันตาม</u>ค่า σ เพราะค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน (σ^2) ดังนั้นการเพิ่มค่า σ นั้นจะทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มเพิ่มสูงขึ้น นอกจากนี้ค่า AMSE ยัง<u>แปรผันตาม</u>จำนวนตัวแปรอิสระด้วย เนื่องจากในการวิจัยครั้งนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูปเมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้นทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มสูงขึ้นโดยเฉพาะอย่างยิ่งวิธี SR ซึ่งใช้หลักการในการคัดเลือกตัวแปรเข้าออกจากตัวแบบผ่านการทดสอบสมมติฐานเท่านั้น ในขณะที่วิธีการภายใต้แนวทางของเบส์ทั้ง 2 วิธี คือ วิธี BMA_{SVT} และ OPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการพิจารณาถึงความเหมาะสมของตัวแบบด้วย เช่น มีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT} จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น เพราะนอกจากวิธีนี้จะมีการคำนึงถึงความน่าจะเป็นภายหลังของตัวแบบแล้วยังมีการ

พิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระด้วย ค่าAMSE แปรผกผันกับขนาดตัวอย่างการเพิ่มขนาดตัวอย่างส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มลดลงเพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้

ผลการวิจัยของวิธี BMA_{SVT} จะให้ค่า RDAMSE น้อยที่สุด รองลงมาคือวิธี OPM และวิธี SR ตามลำดับ โดยส่วนใหญ่ค่า RDAMSE ของวิธี SR และ OPM มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เนื่องจากเมื่อขนาดตัวอย่างเพิ่มขึ้นจะส่งผลต่อการคัดเลือกสมการถดถอยของวิธี BMA_{SVT} ทำให้ค่าAMSE มีค่าลดลงมากกว่าวิธีอื่นจึงส่งผลทำให้ค่าRDAMSEเพิ่มขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

<u>สรุปตอนที่ 4.4</u> ผลการวิจัยของการเปรียบเทียบวิธีการคัดเลือกสมการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ สำหรับตัวแบบการถดถอยที่ประกอบด้วยจำนวนตัวแปรอิสระเท่ากับ 3 5 8 10 12 และ 15 เมื่อ σ เท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ โดยที่วิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,500)

เมื่อวิธี BMA_{SVT}และวิธี OPM มีค่าคงที่ $(\sigma_\beta / \tau, c)$ เหมือนกันคือ (10,500) พบว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุก ๆ กรณี โดยวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อย วิธี BMA_{SVT} มีค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนมากยิ่งขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และวิธี OPM มีค่า AMSE สูงกว่าวิธี BMA_{SVT} ชัดเจนมากขึ้นเมื่อเปรียบเทียบกับผลการวิจัยตอนที่ 4.3 โดยเฉพาะเมื่อค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่ามาก($\sigma \geq 0.50$)

จากผลการวิจัยในตอนที่ 4.4 การเปลี่ยนแปลงของค่า AMSE และค่า RDAMSE มีลักษณะดังนี้

1) <u>ค่า AMSE</u>
   (1) แปรผันตามส่วนเบี่ยงเบนมาตรฐาน จำนวนตัวแปรอิสระและค่าคงที่ σ_β / τ และ c (เฉพาะวิธี BMA_{SVT}และ OPM เท่านั้น) ตามลำดับ
   (2) แปรผกผันกับขนาดตัวอย่าง
2) <u>ค่า RDAMSE</u> แปรผันตามขนาดตัวอย่าง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเพื่อศึกษาการคัดเลือกสมการถดถอยที่ดีที่สุดในกาวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ และเพื่อเปรียบเทียบความถูกต้องของการพยากรณ์จากตัวแบบที่ได้จากวิธีการต่าง ๆ ซึ่งวิธีการคัดเลือกสมการถดถอยที่นำมาใช้ในการสร้างตัวแบบการถดถอยเชิงเส้นพหุคูณมี 3 วิธีดังนี้

1) วิธีการเฉลี่ยตัวแบบของเบส์โดยการหาองค์ประกอบของตัวแบบด้วยเทคนิคมอนติคาร์โล โดยใช้ลูกโซ่มาร์คอฟเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ(BMA_{SVT})
2) วิธีการคัดเลือกตัวแบบที่เหมาะสมที่สุด (OPM)
3) วิธีการถดถอยแบบขั้นบันได(SR)

โดยสองวิธีแรกเป็นวิธีการภายใต้แนวทางของเบส์ ส่วนวิธีที่ 3 เป็นวิธีการพื้นฐานซึ่งนิยมใช้กันอย่างแพร่หลายและมีประสิทธิภาพดีในการคัดเลือกสมการถดถอยที่ดีที่สุดเมื่อการวิเคราะห์ความถดถอยเป็นไปตามข้อตกลงเบื้องต้น ซึ่งผู้วิจัยได้นำมาเปรียบเทียบกับวิธีการภายใต้แนวทางของเบส์เพื่อศึกษาดูแนวโน้มว่าวิธีการคัดเลือกตัวแบบภายใต้แนวทางของเบส์ จะมีความเหมาะสมมากกว่าวิธีการพื้นฐานหรือไม่เมื่อใช้หลักเกณฑ์คู่สังยุคแบบปกติ ดังนั้นถ้าผลการศึกษาวิจัยพบว่าวิธีการใดมีประสิทธิภาพในการพยากรณ์มากที่สุดก็ควรจะนำไปใช้ในการคัดเลือกตัวแบบการถดถอยเชิงเส้นพหุคูณที่ดีที่สุดต่อไป ซึ่งมีการกำหนดสถานการณ์ต่าง ๆ ในการวิจัยครั้งนี้ไว้ดังนี้

1) การแจกแจงของค่าคลาดเคลื่อนสุ่มเป็นแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 0.50 และ 2.50 ตามลำดับ
2) กำหนดขนาดตัวอย่างที่ศึกษา คือ 15 30 50 และ100
3) จำนวนตัวแปรอิสระที่ศึกษาคือ 3 5 8 10 12 และ 15 ตามลำดับ
4) ค่าคงที่สำหรับวิธี BMA_{SVT} และวิธี OPM มี 4 ระดับ คือ (1,5) (1,10) (10,100) และ (10,500)

เกณฑ์ที่ใช้ในการตัดสินใจว่าวิธีการคัดเลือกตัวแบบการถดถอยวิธีใดจะมีประสิทธิภาพมากที่สุดจะพิจารณาจากเกณฑ์ของค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Average of Mean Square Error (AMSE)) และเกณฑ์ที่ใช้ประกอบการตัดสินใจจะใช้เกณฑ์ค่าอัตราส่วนผลต่างของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Ratio of Different Average Mean Square Error (RDAMSE))

ซึ่งวิธีใดให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่าอัตราส่วนผลต่างของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดจะเป็นวิธีที่ดีที่สุด ผลการวิจัยได้ข้อสรุปดังนี้

5.1 สรุปผลการวิจัย

5.1.1 ผลการเปรียบเทียบค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย

จากการเปรียบเทียบค่า AMSE ของทั้ง 3 วิธี พบว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่วิธี BMA_{SVT} OPM และ SR ตามลำดับ สำหรับทุกระดับขนาดตัวอย่าง ค่าส่วนเบี่ยงเบนมาตรฐาน จำนวนตัวแปรอิสระ และค่าคงที่วิธี BMA_{SVT} และวิธี OPM แสดงว่าวิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากที่สุดในการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ วิธี OPM นั้นมีความสามารถในการคัดเลือกสมการถดถอยใกล้เคียงกับวิธี BMA_{SVT} ซึ่งจะสังเกตได้จากค่า AMSE ของวิธี OPM จะสูงกว่าวิธี BMA_{SVT} เพียงเล็กน้อยเท่านั้น โดยเฉพาะอย่างยิ่งเมื่อค่าคงที่ของวิธี BMA_{SVT} และวิธี OPM (σ_β / τ , c) มีค่าต่ำ ๆ ส่วนวิธี SR มีค่า AMSE สูงกว่าวิธี BMA_{SVT} และวิธี OPM อย่างชัดเจนในทุกสถานการณ์ ซึ่งจากผลวิจัยที่ได้นี้ นอกจากจะสรุปได้ว่าวิธี BMA_{SVT} มีความสามารถในการคัดเลือกสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติได้ดีที่สุดแล้ว ยังสามารถสรุปได้ว่าวิธีการภายใต้แนวทางของเบส์มีความสามารถในการคัดเลือกสมการถดถอยเชิงเส้นพหุคูณดีกว่าวิธีการพื้นฐาน ซึ่งในที่นี้คือวิธีการถดถอยแบบขั้นบันได โดยจะเห็นได้จากวิธีการภายใต้แนวทางของเบส์จะให้ค่า AMSE ต่ำกว่าวิธีการพื้นฐานอย่างชัดเจนทุกสถานการณ์ ส่วนข้อสรุปเกี่ยวกับหลักการความไม่แน่นอนของตัวแบบและการแปลงที่เหมาะสมของตัวแปรอิสระจะพบว่าวิธี BMA_{SVT} ซึ่งคำนึงถึงหลักดังกล่าวจะให้ค่าพยากรณ์ที่ได้มีความถูกต้องและแม่นยำมากยิ่งขึ้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

5.1.2 ปัจจัยที่มีผลต่อค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของแต่ละวิธี

1) ขนาดตัวอย่าง (n)

เมื่อขนาดตัวอย่างเพิ่ม ค่าAMSE ของทุกวิธีมีแนวโน้มลดลง เพราะขนาดตัวอย่างที่เพิ่มขึ้นจะช่วยลดความเบี่ยงเบนที่ไม่ทราบสาเหตุลงได้ และเมื่อขนาดตัวอย่างเพิ่มขึ้นวิธี BMA_{SVT} และวิธี OPM จะมีค่าAMSE ต่ำกว่าวิธี SRอย่างชัดเจน

2) ค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่ม (σ)

เมื่อค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนสุ่มมีค่า เพิ่มขึ้น ค่า AMSEของทุกวิธีมีแนวโน้มเพิ่มขึ้น ทั้งนี้เนื่องจากค่า AMSE เป็นตัวประมาณค่าความแปรปรวนของค่าคลาดเคลื่อน(σ^2)

3) จำนวนตัวแปรอิสระ

เมื่อจำนวนตัวแปรอิสระในตัวแบบการถดถอยเพิ่มขึ้น จะส่งผลให้ค่า AMSE ของทุกวิธีมีแนวโน้มสูงขึ้น เนื่องจากในการวิจัยครั้งนี้ได้กำหนดตัวแบบเริ่มต้นเป็นตัวแบบเต็มรูป เมื่อมีจำนวนตัวแปรอิสระเพิ่มขึ้นโอกาสที่จะได้ตัวแบบที่ไม่เหมาะสมก็จะมีมากขึ้น ทำให้ค่า AMSE ของทุกวิธีมีแนวโน้มสูงขึ้น โดยเฉพาะอย่างยิ่งวิธีการภายใต้แนวทางของเบส์ทั้ง 2 วิธีคือ วิธี BMA_{SVT} และวิธี OPM ค่า AMSE จะมีอัตราการเพิ่มขึ้นน้อยกว่าวิธี SR เพราะวิธีการภายใต้แนวทางของเบส์จะมีการพิจารณาถึงความน่าจะเป็นภายหลังของตัวแบบ โดยเฉพาะอย่างยิ่งวิธี BMA_{SVT} จะได้รับผลกระทบน้อยมากจากการเพิ่มจำนวนตัวแปรอิสระ ซึ่งสังเกตได้จากค่า AMSE ที่เพิ่มขึ้นเพียงเล็กน้อยเท่านั้น

4) ค่าคงที่ σ_β / τ และ $\mathbf{C}$

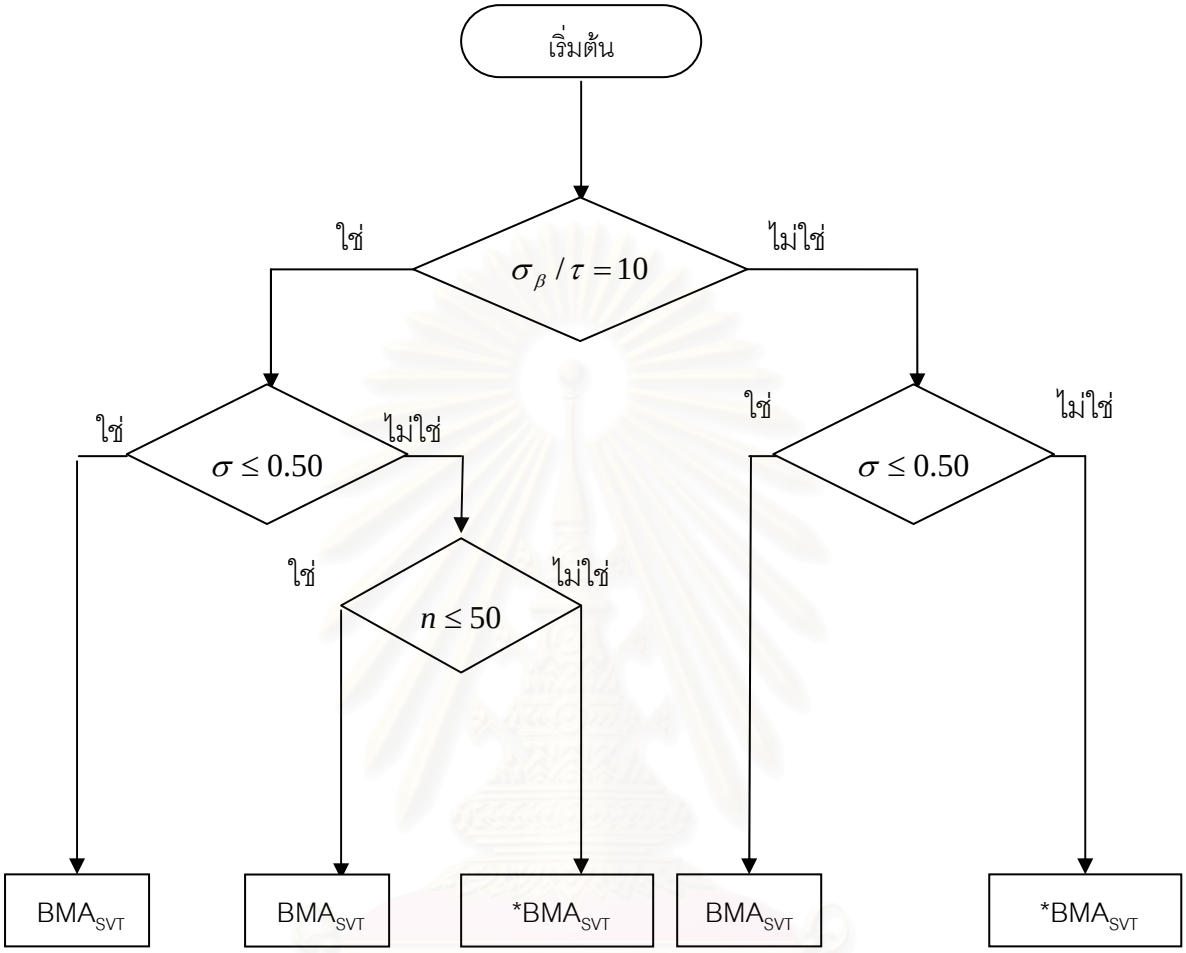
ค่าคงที่ σ_β / τ และ $\mathbf{C}$ จะมีผลต่อค่า AMSE ของวิธี BMA_{SVT} และวิธี OPMเท่านั้น โดยค่า AMSE จะเพิ่มขึ้นเมื่อ ค่าคงที่ σ_β / τ และ $\mathbf{C}$ สูงขึ้น เนื่องจาก ค่าคงที่ σ_β / τ และ $\mathbf{C}$ เป็นค่าที่กำหนดลักษณะการกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยที่สุ่มได้ ค่าคงที่ σ_β / τ และ $\mathbf{C}$ ที่สูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้น ทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้ค่า AMSE มีค่าสูงขึ้น การเปลี่ยนแปลงของค่าคงที่นี้

จะส่งผลกระทบต่อวิธี BMA_{SVT} เพียงเล็กน้อย แต่จะส่งผลต่อวิธี OPM มากพอควรซึ่งสังเกตได้จากอัตราการเพิ่มขึ้นของค่า AMSE เมื่อ ค่าคงที่ σ_β/τ และ $\mathbf{C}$ เพิ่มขึ้น

5.1.3 ผลสรุปการเลือกวิธีการสร้างตัวแบบถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ

ผลการสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์ในการวิจัยครั้งนี้ พบว่าการสร้างสมการถดถอยด้วยวิธี BMA_{SVT} จะส่งผลให้ได้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากที่สุด ส่วนวิธีการสร้างสมการถดถอยที่มีประสิทธิภาพในการพยากรณ์รองลงมา คือ วิธี OPM และวิธี SR ตามลำดับ สำหรับทุกระดับขนาดตัวอย่าง ค่าส่วนเบี่ยงเบนมาตรฐาน จำนวนตัวแปรอิสระ และค่าคงที่วิธี BMA_{SVT} และวิธี OPM (ค่าคงที่ σ_β/τ และ $\mathbf{C}$) ซึ่งวิธี OPM จะมีประสิทธิภาพใกล้เคียงกับวิธี BMA_{SVT} ค่อนข้างมากในกรณีที่ ค่าคงที่ σ_β/τ และ $\mathbf{C}$ มีค่าต่ำ ๆ และจากผลการวิจัยเปรียบเทียบค่า AMSE ของทั้ง 3 วิธี แม้ว่าค่า AMSE ของแต่ละวิธีเรียงลำดับจากน้อยไปมาก ได้แก่ วิธี BMA_{SVT} OPM และ SR ตามลำดับ ในทุกกรณีก็ตาม แต่พบว่ามีบางกรณีที่วิธี BMA_{SVT} มีค่า ต่ำกว่าวิธีอื่น ๆ ค่อนข้างมาก ซึ่งแสดงให้เห็นว่าในกรณีดังกล่าวเราควรเลือกวิธี BMA_{SVT} ในการสร้างสมการถดถอยเพื่อให้ได้ค่าพยากรณ์ที่ถูกต้องและแม่นยำมากที่สุด ส่วนในกรณีอื่น ๆ นั้นแม้ว่าวิธี BMA_{SVT}จะให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ แต่ในการเลือกใช้วิธีการสร้างสมการถดถอย เราอาจคำนึงถึงปัจจัยด้านอื่น ๆ ด้วย เช่น ระยะเวลาในการคำนวณ ซึ่งวิธี BMA_{SVT} นั้นใช้ระยะเวลาในการคำนวณค่อนข้างมากเมื่อเทียบกับวิธี OPM โดยแผนผังแสดงข้อสรุปของกรณีต่าง ๆ ที่วิธี BMA_{SVT} จะให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างชัดเจนแสดงดังรูปที่ 5.1

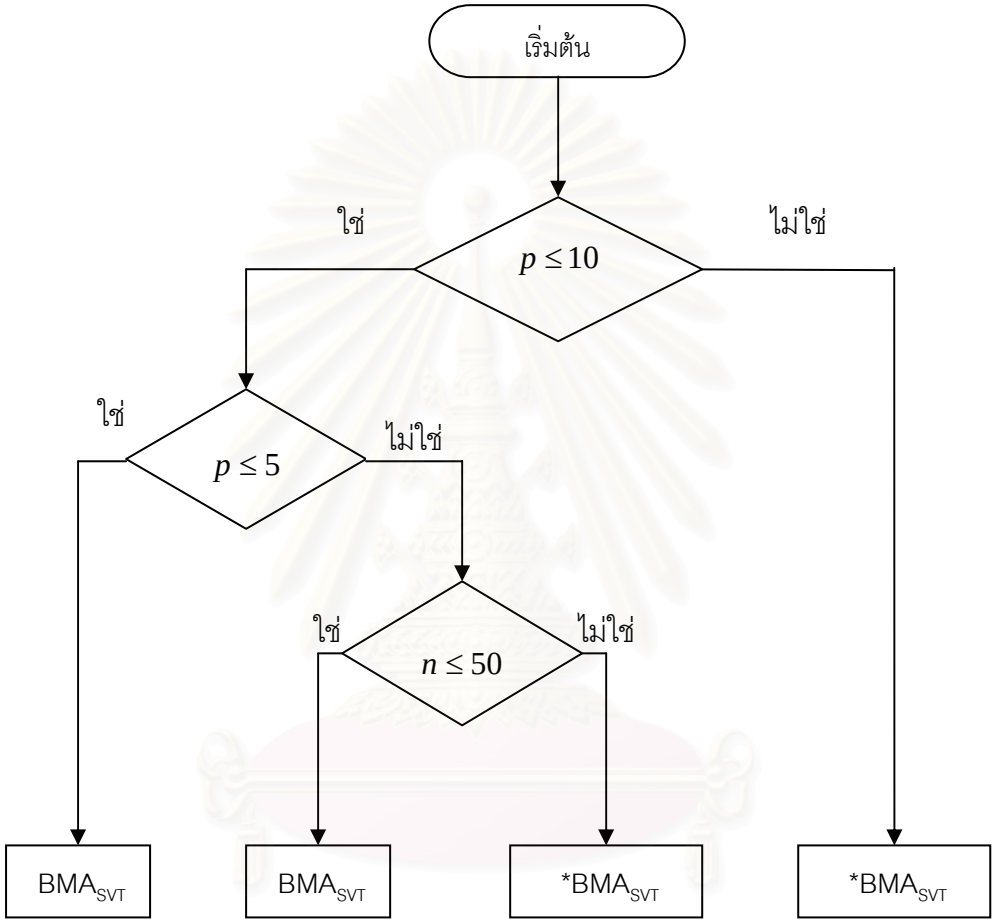
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 5.1 แสดงแผนผังผลสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์ เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ

หมายเหตุ : * หมายถึงกรณีที่วิธี BMA_{SVT} ให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ ค่อนข้างมาก

ในทางปฏิบัตินอกจากผู้วิจัยจะต้องคำนึงถึงระยะเวลาในการคำนวณแล้วขนาดตัวอย่าง (n) ยังเป็นอีกปัจจัยหนึ่งที่ผู้วิจัยมักคำนึงถึงก่อนเสมอ ด้วยเหตุนี้ผู้วิจัยจึงขอเสนอแนวทางการพิจารณาเลือกใช้วิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติ โดยแผนผังแสดงข้อสรุปของกรณีต่าง ๆ ที่วิธี BMA_{SVT} จะให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ ดังแสดงในรูปที่ 5.2



รูปที่ 5.2 แสดงแผนผังผลสรุปการเลือกวิธีการสร้างสมการถดถอยที่ดีที่สุดเชิงเบส์เมื่อใช้การแจกแจงก่อนแบบคู่สังยุคปกติในเชิงปฏิบัติ

หมายเหตุ : * หมายถึงกรณีที่วิธี BMA_{SVT} ให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ ค่อนข้างมาก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

5.2 ข้อเสนอแนะ

1) การวิจัยครั้งนี้ศึกษาเฉพาะการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณไม่รวมถึงตัวแบบความถดถอยพหุนาม ซึ่งเป็นกรณีเฉพาะที่มักพบได้บ่อยครั้งในการวิเคราะห์ความถดถอย ดังนั้นในงานวิจัยแบบเต็มรูป ควรมีการศึกษาเปรียบเทียบวิธีการคัดเลือกสมการถดถอยสำหรับตัวแบบความถดถอยพหุนามด้วย ทั้งนี้เพื่อศึกษาว่าการนำวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธีไปใช้กับการวิเคราะห์ความถดถอยพหุนามจะให้ผลสรุปเหมือนหรือแตกต่างกับการนำไปใช้การวิเคราะห์ความถดถอยเชิงเส้นพหุคูณดังเช่นในงานวิจัยครั้งนี้อย่างไร

2) วิธีการคัดเลือกสมการถดถอยภายใต้แนวทางของเบส์จะได้รับผลกระทบจากการแจกแจงก่อนค่อนข้างมาก ในการวิจัยครั้งนี้ใช้การแจกแจงก่อนแบบคู่สังยุคปกติ ดังนั้นถ้ามีการปรับการแจกแจงก่อนหรือข้อมูลเพิ่มเติมให้เหมาะสมกับปัญหาที่จะศึกษา เช่น การแจกแจงก่อนแบบแกมมาซึ่งมีลักษณะเบ้ขวา เป็นต้น แล้วทำการศึกษาว่าวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี จะให้ผลสรุปเหมือนหรือต่างกันเมื่อปรับข้อมูลเพิ่มเติม

3) วิธี BMA_{SVT} ซึ่งคำนึงถึงหลักการความไม่แน่นอนของตัวแบบเมื่อพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระโดยเลือกใช้วิธีการแปลงเปลี่ยนจุด (change-point transformation) ช่วยให้ได้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากขึ้น หากมีการพิจารณาข้อมูลที่มีค่าผิดปกติ (outlier identification) ควบคู่ไปกับการพิจารณาการแปลงที่เหมาะสมของตัวแปรอิสระ (transformation selection) อาจทำให้วิธีการเฉลี่ยตัวแบบของเบส์มีประสิทธิภาพมากยิ่งขึ้น

4) เนื่องจากในการวิจัยครั้งนี้วิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากที่สุดอย่างคงเส้นคงวาและกรณีที่ค่าคงที่(σ_β/τ และ $\mathbf{C}$) ของวิธี BMA_{SVT}และวิธี OPM มีค่าต่ำ ๆ วิธีOPM จะให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำใกล้เคียงกับวิธี BMA_{SVT}มาก ส่วนกรณีที่ค่าคงที่ (σ_β/τ และ $\mathbf{C}$) ของวิธี BMA_{SVT}และวิธี OPM มีค่าสูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้นทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้วิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากกว่าวิธี OPM ชัดเจนยิ่งขึ้น ดังนั้นในการพิจารณาว่าวิธีการใดจะมีประสิทธิภาพมากที่สุดสำหรับกรณีที่วิธีBMA_{SVT}ให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างคงเส้นคงวาจึงควรใช้การวิเคราะห์ความแปรปรวนของการทดลองแบบสุ่มโดยสมบูรณ์ภายในกลุ่ม (Randomized Completely Block Design: RCBD) เพราะการวิเคราะห์ดังกล่าวจะทำให้ได้วิธีการคัดเลือกสมการถดถอยที่มีประสิทธิภาพมากที่สุดอย่างมีนัยสำคัญทางสถิติ (ดังแสดงตัวอย่างการวิเคราะห์ในหัวข้อที่ 7 ของภาคผนวก)

รายการอ้างอิง

ภาษาไทย

จะเด็จ สวรรค์ตรานนท์.**การเปรียบเทียบวิธีที่ใช้สำหรับการคัดเลือกสมการถดถอยที่ดีที่สุด**. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2530.

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

ธีระพร วีระถาวร. **ความน่าจะเป็นกับการประยุกต์**. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: วิทยพัฒน์, 2539.

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

นิทัสน์ สุขสุวรรณ. **การเปรียบเทียบวิธีการคัดเลือกสมการถดถอยที่ดีที่สุดภายใต้แนวทางเบส์ในการวิเคราะห์ความถดถอยเชิงเส้นพหุคูณ**. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติบัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2545.

นพมาศ อัครจันทโชติ. **การเปรียบเทียบวิธีการสร้างตัวแบบในการวิเคราะห์ความถดถอยเชิงเส้นพหุนามกรณีที่มี 2 ตัวแปรซึ่งเกิดอันตรกิริยา**. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2539

นุชรินทร์ ทิพยวรรณากร. **การเปรียบเทียบค่าพยากรณ์ที่ได้จากตัวแบบที่คัดเลือกตัวแปรด้วยวิธีเบส์เซียน วิธีการกำจัดตัวแปรแบบถอยหลัง และวิธีการถดถอยแบบขั้นได ในการวิเคราะห์ความถดถอยเชิงเส้นพหุนามแบบลำดับขั้น** วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2540

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

วีรพา ฐานะปรัชญ์. **การเปรียบเทียบตัวประมาณกำลังสองน้อยสุดกับตัวประมาณเบส์สำหรับตัวแบบการถดถอยเชิงเส้นเชิงเดียว**. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2545

ภาษาอังกฤษ

Adrian E. Raftery, David Madigan, and Jennifer A. Hoeting. Bayesian Model Averaging for Linear Regression Models. **Journal of the American Statistical Association. Vol. 92, No.437 (1997),** 179-191

Barbieri, M. M. and Berger, J. O. Optimal predictive model selection. **Technical Report 02-02.** Duke University, Durham, 2002

Edward I. George and Robert E. McCulloch. Variable Selection via Gibbs Sampling. **Journal of the American Statistical Association. Vol. 88, No.423 (1993),** 881-889

Hoeting, J. A., Raftery, A. E., and Madigan, D. Bayesian Simultaneous Variable and Transformation Selection in Linear Regression." **Technical Report 99-05.** Colorado State University, Dept. of Statistics, 1999

Jose, B., and Smith A.F.M. **Bayesian Theory.** New York: John Wiley & Sons, 1994.

Kass, R.E. ,and Rafterry, A.E. Bayes factors. **Journal of the American Statistical Association. Vol. 90, (June1995),** 773-795

Raiffa, H. and Schlaifer, R. **Applied Statistical Decision Theory.** The MIT Press, Cambridge, MA, 1961

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

1. **ความน่าจะเป็นที่มีเงื่อนไข** (Conditional Probability)[1]

ความน่าจะเป็นที่มีเงื่อนไขเป็นการศึกษาความน่าจะเป็นเมื่อทราบข้อมูลเพิ่มเติมเกี่ยวกับเหตุการณ์ที่สนใจศึกษา

นิยามที่ 1

ให้ A และ B เป็นเหตุการณ์ซึ่ง $P(A)>0$ เราเรียก $P(B|A)$ ว่าเป็น "ความน่าจะเป็นที่มีเงื่อนไขของ B เมื่อกำหนด A (the conditional probability of B given A)" ถ้า

$$P(B|A)=\frac{P(B\cap A)}{P(A)} \quad , P(A)>0$$

ทฤษฎีบทที่ 1 ทฤษฎีการคูณ (Multiplication Theorem)

ถ้าในการทดลองชนิดหนึ่งเหตุการณ์ A และ B สามารถเกิดขึ้นได้พร้อมกัน จะได้ว่า

$$P(B\cap A)=P(B|A)\cdot P(A)$$

ทฤษฎีบทที่ 2

ถ้าในการทดลองชนิดหนึ่งซึ่งเหตุการณ์ A_1 , A_2 ,..., A_n สามารถเกิดขึ้นพร้อมกันได้ จะได้ว่า

$$P(A_1\cap A_2\cap\ldots\cap A_n)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1,A_2)\ldots P(A_n|A_1,A_2,\ldots,A_{n-1})$$

2. **ทฤษฎีบทของเบส์** (Bayes' Theorem)[2]

นิยามที่ 2

เหตุการณ์ B_1 ,B_2 , ... จะแทนผลแบ่งกั้นของปริภูมิตัวอย่าง S ถ้า

1) $B_i\cap B_j=\phi$, $\forall i\neq j$

2) $\bigcup_{i=1}^{\infty} B_i=S$

3) $P(B_i)>0$, $\forall i$

---

[1] ธีระพร วีระถาวร, ความน่าจะเป็นกับการประยุกต์, พิมพ์ครั้งที่ 2 (กรุงเทพมหานคร : วิทยพัฒน์, 2539), หน้า 100-106

[2] ธีระพร วีระถาวร, ความน่าจะเป็นกับการประยุกต์, พิมพ์ครั้งที่ 2 (กรุงเทพมหานคร : วิทยพัฒน์, 2539), หน้า 107-108

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

กล่าวคือเมื่อเราทำการทดลอง E จะได้ว่าเหตุการณ์ $B_i, i=1,2,...,n$ จะไม่เกิดขึ้นพร้อม ๆ กัน

ทฤษฎีบทที่ 3

ให้ $B_1, B_2, \ldots$ แทนผลแบ่งกั้นของปริภูมิตัวอย่าง S และถ้า A เป็นเหตุการณ์ ซึ่ง $P(A) > 0$ จะได้ว่า

$$P(B_k \mid A) = \frac{P(B_k)\, P(A \mid B_k)}{\sum_{i=1}^{\infty} P(B_i) \cdot P(A \mid B_i)} \quad , \quad \forall k = 1, 2, 3, \ldots$$

3. **ความเป็นอิสระ** (Independence)[3]

นิยามที่ 3

เหตุการณ์ 2 เหตุการณ์ A และ B เรียกว่า " เป็นอิสระกัน (independence)" ก็ต่อเมื่อ

$P(B \mid A) = P(B)$ และ $P(A \mid B) = P(A)$ กล่าวคือ

$$P(A \cap B) = P(A) \cdot P(B)$$

นิยามที่ 4

เหตุการณ์ 3 เหตุการณ์ A , B และ C เป็นอิสระซึ่งกันและกัน (mutually Independent) ก็ต่อเมื่อ

1) แต่ละคู่ของเหตุการณ์เป็นอิสระต่อกัน (pairwise independent) กล่าวคือ

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A \cap C) = P(A) \cdot P(C)$$
$$P(B \cap C) = P(B) \cdot P(C)$$

2) เหตุการณ์หนึ่งเหตุการณ์ใดเป็นอิสระจากสองเหตุการณ์ใด ๆ กล่าวคือ

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$

---

[3] ธีระพร วีระถาวร, ความน่าจะเป็นกับการประยุกต์, พิมพ์ครั้งที่ 2 (กรุงเทพมหานคร: วิทยพัฒน์, 2539), หน้า 114-116.

4. **การทดสอบเอฟบางส่วน** (Partial F-test)[4]

การทดสอบเอฟบางส่วนเป็นการทดสอบที่ใช้ตรวจสอบนัยสำคัญของ β_j เพื่อตัดสินใจว่าตัวแปรอิสระใดควรอยู่ในสมการหรือตัวแปรอิสระใดควรตัดออกจากสมการถดถอย โดยที่ β_j จะปรากฏอยู่ ณ ตำแหน่งใดแบบจำลองก็ได้ แต่ในทางปฏิบัติจะทำโดยถือว่าตัวแปรอิสระนั้นเข้าสู่สมการถดถอยเป็นตัวสุดท้าย

จากสมการถดถอยแสดงความสัมพันธ์ระหว่างตัวแปรตาม *(y)* กับตัวแปรอิสระ *(x_j)* คือ

(a) $$\underset{\sim}{y} = \beta_0 + \beta_1 \underset{\sim}{x}_1 + \beta_2 \underset{\sim}{x}_2 + \ldots + \beta_k \underset{\sim}{x}_k + \underset{\sim}{\varepsilon}$$

เราสามารถหาค่าตัวประมาณ $\underset{\sim}{\hat{\beta}}_1 = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k)'$ โดยวิธีกำลังสองน้อยสุด และค่าผลบวกกำลังสองที่เกี่ยวข้องเป็นดังนี้

1) $\underset{\sim}{\hat{\boldsymbol{\beta}}}_1 = ({}_1X'{}_1X)^{-1}\,{}_1X'\,\underset{\sim}{y}$

เมื่อ ${}_1X$ คือเมทริกซ์ของตัวแปรอิสระขนาด $n \times (k+1)$ ซึ่งรวมเทอมของค่าคงที่

2) $SSR_1 = \underset{\sim}{\hat{\boldsymbol{\beta}}}'\,{}_1X'\,\underset{\sim}{y}$

3) $SSE_1 = \underset{\sim}{y}'\,\underset{\sim}{y} - \underset{\sim}{\hat{\boldsymbol{\beta}}}'\,{}_1X'\,\underset{\sim}{y}$ และ

$$MSE_1 = \hat{\sigma}_1^2 = \frac{1}{n-(k+1)}\left(\underset{\sim}{y}'\,\underset{\sim}{y} - \underset{\sim}{\hat{\boldsymbol{\beta}}}'\,{}_1X'\,\underset{\sim}{y}\right)$$

กำหนดให้

(b) $$\underset{\sim}{y} = \beta_0 + \beta_1 \underset{\sim}{x}_1 + \beta_2 \underset{\sim}{x}_2 + \ldots + \beta_k \underset{\sim}{x}_k + \beta_{k+1} \underset{\sim}{x}_{k+1} + \ldots + \beta_p \underset{\sim}{x}_p + \underset{\sim}{\varepsilon}$$

เป็นสมการถดถอยแสดงความสัมพันธ์ระหว่างตัวแปรตาม *(y)* กับตัวแปรอิสระ *(x_j)* โดยที่ $p > k$ จากสมการ (b)สามารถหาค่าตัวประมาณ $\underset{\sim}{\hat{\beta}}_2 = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k, \hat{\beta}_{k+1}, \ldots, \hat{\beta}_p)$ โดยวิธีกำลังสองน้อยสุด และค่าผลบวกกำลังสองที่เกี่ยวข้องเป็นดังนี้

---

[4] นพมาศ อัครจันทโชติ, "การเปรียบเทียบวิธีการสร้างตัวแบบในการวิเคราะห์ความถดถอยพหุนามกรณีที่มี 2 ตัวแปรอิสระซึ่งเกิดอันตรกิริยา", (วิทยานิพนธ์ ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2539), หน้า 15-17.

1) $\underset{\sim}{\hat{\beta}}_2 = (X'_2X)^{-1}\,X'_2\underset{\sim}{y}$

เมื่อ X_2 คือเมทริกซ์ของตัวแปรอิสระขนาด $n \times (p+1)$ ซึ่งรวมเทอมของค่าคงที่

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

3) $SSE_2 = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\hat{\beta}}'_2X'\underset{\sim}{y}$ และ

$$MSE_2 = \hat{\sigma}^2_2$$

$$= \frac{1}{n-(p+1)}\left(\underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\hat{\beta}}'_2X'\underset{\sim}{y}\right)$$

จากผลลัพธ์ข้างต้นพบว่า Extra Sum of Squares Regression (ESSR) คือ

$$ESSR = SSR_2 - SSR_1$$

$$= \underset{\sim}{\hat{\beta}}'_2X'\underset{\sim}{y} - \underset{\sim}{\hat{\beta}}'_1X'\underset{\sim}{y}$$

ซึ่ง Extra Sum of Squares นี้เป็นค่าผลบวกกำลังสองของตัวแปรอิสระ x_{k+1} , x_{k+2} , . . . , x_p ที่เพิ่มขึ้นมาจากสมการ (b)

โดยอาศัยความรู้เรื่อง Distribution of Quadratic Form เราสามารถพิสูจน์ได้ว่า

$$\frac{ESSR}{\sigma^2} \sim \chi^2_{(p-k)}$$

$$\frac{SSE_2}{\sigma^2} \sim \chi^2_{(n-(p+1))}$$

ดังนั้น

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

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

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จะมีการแจกแจงเอฟบางส่วน ณ ระดับขั้นความเสรี $(p-k\ ,\ n-p-1)$[5] และเราจะปฏิเสธสมมติฐาน $H_0 : \beta_{k+1} = \beta_{k+2} = \ldots = \beta_p = 0$ ณ ระดับนัยสำคัญ α เมื่อค่าทดสอบเอฟบางส่วนมากกว่า $F_{1-\alpha, p-k, n-p-1}$

จากความรู้ในเรื่องการทดสอบเอฟบางส่วนนี้ เราสามารถนำมาประยุกต์ใช้กับวิธีการหาสมการถดถอยที่ดีที่สุด โดยวิธีการถดถอยแบบขั้นบันไดได้ดังนี้

จากสมการถดถอย $\underset{\sim}{y} = \beta_0 + \beta_1 \underset{\sim}{X}_1 + \beta_2 \underset{\sim}{X}_2 + \ldots + \beta_p \underset{\sim}{X}_p + \underset{\sim}{\varepsilon}$ เราจะหาค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squares Error (MSE)) และผลบวกกำลังสองของความถดถอย (Sum of Squares Regression) ของเฉพาะ X_j ได้จากสมการต่อไปนี้

$$MSE = \hat{\sigma}^2 = \frac{1}{(n-p-1)}\left(\underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\hat{\beta}}'\underset{\sim}{X}'\underset{\sim}{y}\right)$$

เมื่อ $SS\left(X_j \middle| X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p\right) = SS\left(X_1, X_2, \ldots, X_p\right) - SS\left(X_1, X_2, \ldots, X_{j-1}, X_{j+1}, \ldots, X_p\right)$

ดังนั้น

ค่าสถิติเอฟบางส่วนคือ

$$F_c = \frac{SS\left(X_j \middle| X_1, X_1, \ldots X_{j-1}, X_{j+1}, \ldots X_p\right)}{\hat{\sigma}^2}$$

โดยจะปฏิเสธสมมติฐาน $H_0 : \beta_j = 0$ ณ ระดับนัยสำคัญ α เมื่อ $F_c > F_{1-\alpha, p-k, n-p-1}$

5. **สัมประสิทธิ์สหสัมพันธ์** (Correlation Coefficient)[6]

สัมประสิทธิ์สหสัมพันธ์เป็นตัวสถิติที่ใช้วัดความสัมพันธ์เชิงเส้นของตัวแปรว่ามีความสัมพันธ์กันมากน้อยเพียงใด ซึ่งค่าสัมประสิทธิ์สหสัมพันธ์จะเป็นค่าที่บอกทั้งทิศทางและขนาดของสหสัมพันธ์ โดยจะมีค่าอยู่ระหว่าง -1 ถึง 1 ถ้าค่าสัมประสิทธิ์สหสัมพันธ์เข้าใกล้ -1 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์เชิงเส้นกันมากในทางตรงกันข้าม ถ้าเข้าใกล้ 1 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์เชิงเส้นกันมากในทิศทางเดียวกัน แต่ถ้าค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเป็นศูนย์หรือเข้าใกล้ศูนย์ แสดงว่าตัวแปรทั้งสองไม่มีความสัมพันธ์เชิงเส้นต่อกัน หรือมี

---

[5] $U = \frac{\chi_1^2/\upsilon_1}{\chi_2^2/\upsilon_2} \sim F_{(\upsilon_1, \upsilon_2)}$

[6] นพมาศ อัครจันทโชติ, "การเปรียบเทียบวิธีการสร้างตัวแบบในการวิเคราะห์ความถดถอยพหุนามกรณีที่มี 2 ตัวแปรอิสระซึ่งเกิดอันตรกิริยา", (วิทยานิพนธ์ ปริญญามหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2539), หน้า 17-18.

ความสัมพันธ์เชิงเส้นกันน้อย สำหรับสัมประสิทธิ์สหสัมพันธ์ที่ใช้วัดความสัมพันธ์ระหว่างตัวแปรสองตัวสามารถจำแนกได้ดังนี้

1) สัมประสิทธิ์สหสัมพันธ์เชิงเดียว (Simple Correlation Coefficient) เป็นตัวสถิติที่ใช้วัดความสัมพันธ์ระหว่าง 2 ตัวแปรใด ๆ ซึ่งมีสูตรดังนี้

$$r_{xy} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

เมื่อ n เป็นขนาดตัวอย่าง

y_i เป็นตัวแปรตามที่ i

และ x_i เป็นตัวแปรอิสระที่ i

2) สัมประสิทธิ์สหสัมพันธ์บางส่วน (Coefficient of Partial Correlation) เป็นตัวสถิติที่ใช้เป็นค่าวัดระดับและทิศทางของความสัมพันธ์เชิงเส้นระหว่างตัวแปร 2 ตัวแปร โดยที่ควบคุมให้ตัวแปรอื่น ๆ คงที่ เช่น กรณีตัวแปรอิสระ 3 ตัวแปร ได้แก่ x_1 , x_2 , x_3 และตัวแปรตาม y ถ้าต้องการหาสัมประสิทธิ์สหสัมพันธ์บางส่วนของตัวแปรตาม y กับตัวแปรอิสระ x_1 โดยควบคุมให้ x_2 และ x_3 คงที่ จะใช้สัญลักษณ์ $r_{y1.\ 23}$ สำหรับการคำนวณหาค่าของสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรคู่ใด ๆ นั้นหาได้จาก

$$r^2_{ij.1,2,3,\ldots,i-1,i+1,\ldots,j-1,j+1,\ldots k} = \left(\frac{a^2_{ij}}{a_{ii}a_{jj}}\right)$$

และ $$r_{ij\ .1,2,3,\ldots,i\ -1,\ i+1,\ldots\ ,j\ -1,\ j\ +1,\ldots k} = sign\sqrt{r^2}$$

$$R = \begin{bmatrix} 1 & a_{12} & \ldots & r_{1k} \\ a_{21} & 1 & \ldots & r_{2k} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ r_{k1} & r_{k2} & \ldots & 1 \end{bmatrix}_{k\ x\ k}$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

$$A = R^{-1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}_{k \times k}$$

เมื่อ $r_{ij \,.\, 1, 2, 3, \ldots, i-1, i+1, \ldots, j-1, j+1, \ldots, k}$ เป็นค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรที่ i และ j โดยตัวแปรอื่นคงที่

r_{ij} เป็นสัมประสิทธิ์สหสัมพันธ์เชิงเดียวระหว่างตัวแปรที่ i และ j

k เป็นจำนวนตัวแปรทั้งหมด

a_{ij} เป็นสมาชิกแถวที่ i แนวตั้งที่ j ของเมทริกซ์ผกผันของเมทริกซ์สหสัมพันธ์ *(R)*

R เป็นเมทริกซ์สหสัมพันธ์

และ A เป็นเมทริกซ์ผกผันของเมทริกซ์สหสัมพันธ์

ส่วนเครื่องหมายของสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระใด ๆ นั้นกำหนดตามเครื่องหมายของสัมประสิทธิ์การถดถอยของตัวแปรอิสระนั้น

6. รายละเอียดของโปรแกรมที่ใช้ในการวิจัย

ในการวิจัยครั้งนี้ได้ใช้โปรแกรม S-plus 2000 สำหรับการสร้างตัวแบบด้วยวิธี BMA_{SVT} OPM และ SR เนื่องจากในงานวิจัยของ ราฟเทอร์รี่ เมดิแกน และโฮเอ็ททิง (Raftery Madigan and Hoeting , 1997) ซึ่งเป็นผู้นำเสนอวิธีการเฉลี่ยตัวแบบของเบส์ได้มีการนำเสนออัลกอริทึมของวิธีการเฉลี่ยตัวแบบของเบส์โดยใช้โปรแกรม S-plus 2000 ดังนั้น เพื่อความสะดวกในการพัฒนาโปรแกรมของผู้วิจัยจึงได้ใช้โปรแกรม S-plus 2000 ในการเขียนโปรแกรมเพื่อสร้างตัวแบบด้วยวิธี BMA_{SVT} OPM และ SR ดังกล่าว

สำหรับรายละเอียดทั้งหมดของโปรแกรมที่ใช้ในการวิจัยมีดังนี้

ตารางแสดงลักษณะการทำงานของโปรแกรมที่ใช้ในการวิจัย

ลำดับที่	ชื่อโปรแกรม	การทำงานของโปรแกรม	ชื่อโปรแกรมย่อยหรือฟังก์ชันที่เรียกใช้
โปรแกรมหลัก	MAIN	- สร้างข้อมูลตัวแปรตาม - สร้างข้อมูลตัวแปรอิสระ - สร้างตัวแบบด้วยวิธี BMA_{SVT}, OPM และ SR - คำนวณค่า AMSE และค่าส่วนเบี่ยงเบนมาตรฐานของค่า AMSE จากการทำซ้ำ 500 รอบ	rnorm, rmvnorm, fy, SVT, OPM, stepwise, fmse, fsd
โปรแกรมย่อย			
1	SVT	ดำเนินการสร้างตัวแบบตามขั้นตอนของวิธี BMA_{SVT}	choose, flogpost,
2	OPM	ดำเนินการสร้างตัวแบบตามขั้นตอนของวิธี OPM	choose, flogpost, fr
3	stepwise	ดำเนินการสร้างตัวแบบตามขั้นตอนของวิธี SR	
4	MC3	ดำเนินการหาปริภูมิตัวแบบและคำนวณความน่าจะเป็นภายหลัง	
5	choose	สุ่มหาตัวแบบรูปแบบอื่น ๆ	
6	flogpost	คำนวณค่าล็อกของความน่าจะเป็นภายหลังของตัวแบบ	
7	fy	สร้างตัวแปรตาม	
8	fyp	คำนวณค่าประมาณของตัวแปรตาม	
9	fr	คำนวณค่าความสูญเสียอันเกิดจากความผิดพลาดกำลังสอง	
10	fmse	คำนวณค่า AMSE	
ฟังก์ชัน			
1	rnorm	ทำการสร้างตัวเลขสุ่มให้มีการแจกแจงแบบปกติ	
2	rmvnorm	ทำการสร้างตัวเลขสุ่มให้มีการแจกแจงแบบปกติหลายตัวแปร	
3	sum	หาผลรวมของข้อมูล	
4	min	หาค่าต่ำสุด	
5	ace	หาการแปลงที่เหมาะสม	
6	leaps	ทำการสร้างตัวแบบที่เป็นไปได้	

```
**********************************************************************************************************************
MC3<-function(all.y,all.x,num.its,M0.var,M0.out,outs.list,PI,K,ip,cc,n,sde)
{
  Ys <<-scale(all.y)
  Xs <<-scale(all.x)

  M0.var<<-M0.var
  M0.out<<-M0.out
  outs.list <<- outs.list
  PI<<-PI
  K<<-K
  flag <<- 1
  outcnt<<-sum(outs.list)
  ip<<-ip
  cc<<-cc
  sde<<-sde
  n<<-n

 big.list<<-matrix(0,1,4)
 big.list[1,1]<<-sum(2^((0:(length(M0.var)-1))[M0.var]))+1
 if (sum(M0.out)!=0)
    big.list[1,2]<<-sum(2^((0:(length(M0.out)-1))[M0.out]))+1
 else big.list[1,2]<<-1
 if (outcnt!=0) big.list[1,3]<<-(dim(Ys)[1]-sum(M0.out))*log(1-PI)+
            sum(M0.out)*log(PI)+flogpost(Ys,Xs,M0.var,
            sde,n,sum(M0.var),ip,cc)
 else big.list[1,3]<<-flogpost(Ys,Xs,M0.var,
                sde,n,sum(M0.var),ip,cc)

i_1
while (i<=num.its)
{
      if (flag==1)
      {
         if (sum(M0.var)!=0)
```

```
M0.1<<-sum(2^((0:(length(M0.var)-1))[M0.var]))+1
          else M0.1<<-1
         if (sum(M0.out)!=0)
M0.2<<-sum(2^((0:(length(M0.out)-1))[M0.out]))+1
         else M0.2<<-1
       }

M1 <- choose(M0.var,M0.out)
      if (sum(M1$var)!=0)
M1.1<<-sum(2^((0:(length(M0.var)-1))[M1$var]))+1
       else M1.1<-1
       if (sum(M1$out)!=0)
M1.2<<-sum(2^((0:(length(M0.out)-1))[M1$out]))+1
       else M1.2<-1

if (sum(big.list[,1]==M1.1 & big.list[,2]==M1.2)==0)
     {
        if (M1.1==1)
          {
          if (outcnt!=0) a<-(dim(Ys)[1]-sum(M1$out))*log(1-PI)+
              sum(M1$out)*log(PI)+ flogpost(Ys,Xs,0,sde,n,0,ip,cc)
          else a<-flogpost(Ys,Xs,0,sde,n,0,ip,cc)
          }
        else
         {
          if (outcnt!=0) a<-(dim(Ys)[1]-sum(M1$out))*log(1-PI)+
          sum(M1$out)*log(PI)+flogpost(Ys,Xs,M1$var,
          sde,n,sum(M1$var),ip,cc)
         else a<-flogpost(Ys,Xs,M1$var,
          sde,n,sum(M1$var),ip,cc)
          }
       big.list<<-rbind(big.list,c(M1.1,M1.2,a,0))
      }

      BF <- exp(big.list[big.list[,1]==M1.1 & big.list[,2]==M1.2,3]-
```

```
            big.list[big.list[,1]==M0.1 & big.list[,2]==M0.2,3])

      if (BF >= 1) flag <<- 1
      else flag <<- rbinom(1,1,BF)

       if (flag == 1)
        {
          M0.var <<- M1$var
          M0.out <<- M1$out
          M0.1 <<- M1.1
          M0.2 <<- M1.2
        }
      big.list[big.list[,1]==M0.1 & big.list[,2]==M0.2,4] <<-
           big.list[big.list[,1]==M0.1 & big.list[,2]==M0.2,4]+1
      i<-i+1
}

  var.vect<<-matrix(as.logical(rep(big.list[,1]-1,rep(length(M0.var),
           length(big.list[,1])))
           %/% 2^(0:(length(M0.var) - 1)) %% 2),ncol=length(M0.var),byrow=T)

    n.var<-length(M0.var)
    ndx<-1:n.var
    Xn<-rep("X",n.var)
    labs<-paste(Xn,ndx,sep="")

    dimnames(var.vect)<-list(c(1:length(var.vect[,1])),labs)

    postprob<<-matrix((exp(big.list[,3]))/(sum(exp(big.list[,3]))),ncol=1)
    dimnames(postprob)[2]_list(c("Post.Mod.Pr."))
    visits<<-matrix(big.list[,4],ncol=1)

    dimnames(visits)[2]<-list(c("#visits"))
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
 if (length(outs.list)!=0)
{out.vect<<-matrix(as.logical(rep(big.list[,2]-1,
                    rep(length(outs.list),length(big.list[,2])))
            %/% 2^(0:(length(outs.list) - 1)) %%
2),ncol=length(outs.list),byrow=T)
    dimnames(out.vect)<-list(c(1:length(out.vect[,1])),c(outs.list))
 model.matrix<<-cbind(var.vect,out.vect,postprob,visits)}
 else model.matrix<<-cbind(var.vect,postprob,visits)

 colno<-length(M0.var)+length(M0.out)+1
 model.matrix<<-model.matrix[order(-model.matrix[,colno]),]

 return(model.matrix)
}
*****************************************************************************************************************
choose<-function(M0.var,M0.out)
{
  var <- M0.var
  in.or.out <- sample(c(1:length(M0.var),rep(0,length(M0.out))),1)
  if (in.or.out == 0){
     out<-M0.out
     in.or.out2 <- sample(1:length(M0.out),1)
     out[in.or.out2]<-!M0.out[in.or.out2]
     }
  else {
      var[in.or.out]<- !M0.var[in.or.out]
      out <- M0.out
     }
return(var,out)
}
*****************************************************************************************************************
flogpost<-function(y,x,model.vect, sde,n,numx,ip,cc)
{
     x_x[,model.vect]
     xtx_t(x)%*%x
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
xty_t(x)%*%y
sig2_sde^2
bls_ginverse(xtx)%*%xty
s2x_matrix(nrow=numx,ncol=1)
tau_matrix(nrow=numx,ncol=1)
sigb_matrix(nrow=numx,ncol=1)
diffx_matrix(nrow=samplesize,ncol=numx)
diffx2_matrix(nrow=samplesize,ncol=numx)
s2x_cor(x,trim=0,na.method="fail", unbiased=T)
for (indexj in 1:numx)
{
      tau[indexj,1]_((sde/xtx[indexj,indexj])^0.5)/ip
      if (model.vect[indexj]!=F)
      {
            sigb[indexj,1]_(ip/tau[indexj,1])^0.5
      }
      else
      {
            sigb[indexj,1]_(ip/(cc*tau[indexj,1]))^0.5
      }
}
bigsigb_diag(rep(1,numx))
for (j in 1:numx)
{
      bigsigb[j,j]_sigb[j,1]/s2x[j,j]
}
b11_t(bls)
A_sig2*bigsigb
covb_bigsigb
bigb_rmvnorm(1, mean=b11, cov=covb,d=numx)
bigbbar_rep(sum(bigb)/(numx),numx)
b1_A+(xtx)
b2_(A%*%bigbbar)+(xtx%*%bls)
b2bar_ginverse(b1)%*%b2
bigsigb2b_bigsigb+((sig2^-1)*xtx)
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
    cov2b_diag(rep(1,numx))
    for (j in 1:numx)
    {
        cov2b[j,j]_bigsigb2b[j,j]
    }
    q_matrix(rep(numx,numx),nrow=1,ncol=numx)
    mpost_pmvnorm(q,mean=t(b2bar),cov=cov2b)
    lmpost_log(mpost,10)

    return(lmpost)
}
*****************************************************************************************************************
fyp_function(rowresult,resultsvt,newx,y,yp1)
{
    for (i in 1:rowresult)
    {
        numxtemp_0
        newnumx_ncol(resultsvt)-2
        for (j in 1:newnumx)
        {
            if (resultsvt[i,j]!=0)
            {
                numxtemp_numxtemp+1
            }
        }
        xtemp_matrix(,nrow=samplesize,ncol=numxtemp)
        sumx1_1
        for (indexj in 1:newnumx)
        {
            if (resultsvt[i,indexj]!=0)
            {
                xtemp[,sumx1]_newx[,indexj]
                sumx1_sumx1+1
            }
        }
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
            regmodel_glm(y~xtemp)
            yhat_predict(regmodel)
            yp1_yp1+resultsvt[i,newnumx+1]*yhat
      }
      yp_yp1

      return(yp)
}
*****************************************************************************************************************
fmse_function(y,yp,n,numx)
{
      dif_y-yp
      se_matrix(nrow=n,ncol=1)
      for (i in 1:n)
      {
            se[i]_(dif[i])^2
      }
      sse_sum(se)
      mse_sse/(n-numx-1)

      return(mse)
}
*****************************************************************************************************************
fy_function(samplesize,sde,x,numx)
{
      error_rnorm(samplesize,0,sde)
      ones_rep(1,samplesize)
      xones_cbind(ones,x)
      beta_rep(1,numx+1)
      y_(xones%*%beta)+error

      return(y)
}
*****************************************************************************************************************
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
fr_function(regwhich,xtemp,p.5,y,
rowwhich,numxtemp)
{
      q_t(xtemp)%*%xtemp
      bfull_solve(q)%*%(t(xtemp)%*%y)

r_matrix(nrow=rowwhich,ncol=numxtemp)
      sumr_matrix(nrow=numxtemp,ncol=1)
      for (j in 1:rowwhich)
      {
            for (k in 1:numxtemp)
            {
                  r[j,k]_((bfull[k,1]^2)*q[k,k])%*%((regwhich[j,k]-p.5[k,1])^2)
            }
            sumr[j,1]_sum(r[j,])
      }
      return(sumr)
}

**********************************************************************************************************************
                                                MAIN_PROGRAM
**********************************************************************************************************************

numx_3
sde_0.25
ip_1
cc_5
numloop_500
samplesize_15
x_rmvnorm(samplesize,mean=rep(0,numx),cov=diag(rep(1,numx)),d=numx)
msesr_matrix(nrow=numloop,ncol=1)
mseopm_matrix(nrow=numloop,ncol=1)
msesvt_matrix(nrow=numloop,ncol=1)
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
for (i in 1:numloop)
{
    y_fy(samplesize,sde,x,numx)
    ********************************************************************************
                                        SR
    ********************************************************************************
    resultsr_stepwise(x,y,intercept=T,tolerance=1.e-07,method="ex",
            nbest=3)
    regmodel_lm(y~x[,resultsr$which[3,]])
    msesr[i]_sum(regmodel$residuals^2)/regmodel$df
    ********************************************************************************
                                        OPM
    ********************************************************************************
    resultmc<-MC3(y,x,10000,rep(T,numx),NULL,NULL,0,0,ip,cc,n,sde)
    rowresult_nrow(resultmc)
    p_matrix(nrow=numx,ncol=1)
    numxtemp_0
    for (j in 1:numx)
    {   prob_matrix(nrow=rowresult,ncol=1)
        for (k in 1:rowresult)
        {
            prob[k,1]_resultmc[k,j]%*%resultmc[k,numx+1]
        }
        p[j,1]_sum(prob)
        if (p[j,1]>=0.5)
        {
            numxtemp_numxtemp+1
        }
    }
    xtemp_matrix(nrow=samplesize,ncol=numxtemp)
    p.5_matrix(nrow=numxtemp,ncol=1)
    sumx_1
    for (I in 1:numx)
    {
        if (p[I,1]>=0.5)
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
    {
        p.5[l,1]_p[l,1]
        xtemp[,sumx]_x[,l]
        sumx_sumx+1
    }
}
reg_leaps(xtemp, y,rep(1,samplesize), int=T, method="adjr2", keep.int=T,
  nbest=1, df=samplesize)
numxtemp_ncol(xtemp)
rowwhich_nrow(reg$which)
mr_fr(reg$which,xtemp,p.5,y,rowwhich,numxtemp)
minr_min(mr)
for (l in 1:rowwhich)
{
    if (mr[l,1]==minr)
    {
        regmodel_lm(y~xtemp[,reg$which[l,]])
        mseopm[i]_sum(regmodel$residuals^2)/regmodel$df
    }
}
*******************************************************************************
                                    SVT
*******************************************************************************
tr_ace(x,y)
tx_matrix(nrow=samplesize,ncol=numx)
for (i in 1:numx)
{
    tx[,i]_tr$tx[,i]
}
xandtx_cbind(x,tx)
newnumx_ncol(xandtx)
resultsvt_MC3(y,xandtx,10000,rep(T,newnumx),NULL,NULL,0,0,ip,cc,n,sde)
rowresult_nrow(resultsvt)
yp_rep(0,samplesize)
ysvt_fyp(rowresult,resultsvt,xandtx,y,yp)
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
        msesvt [i]_fmse(y,ysvt,samplesize,rowresult)
        print(samplesize)
        print(i)
}
amsesr_sum(msesr)/numloop
stdsr_matrix(nrow=numloop,ncol=1)
for (indexi in 1:numloop)
{
        stdsr[indexi]_(msesr[indexi]-amsesr)^2
}
stdamsesr_sqrt(sum(stdsr)/(numloop-1))
amseopm_sum(mseopm)/numloop
stdopm_matrix(nrow=numloop,ncol=1)
for (indexi in 1:numloop)
{
        stdopm[indexi]_(mseopm[indexi]-amseopm)^2
}
stdamseopm_sqrt(sum(stdopm)/(numloop-1))
amsesvt_sum(msesvt)/numloop
stdsvt_matrix(nrow=numloop,ncol=1)
for (indexi in 1:numloop)
{
        stdsvt[indexi]_(msesvt[indexi]-amsesvt)^2
}
stdamsesvt_sqrt(sum(stdsvt)/(numloop-1))
amsesr
stdamsesr
amseopm
stdamseopm
amsesvt
stdamsesvt

**********************************************************************************************************************

                                        END MAIN_PROGRAM

**********************************************************************************************************************
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

7. การวิเคราะห์ความแปรปรวนของการทดลองแบบสุ่มโดยสมบูรณ์ภายในกลุ่ม (Randomized Completely Block Design: RCBD)

ในกรณีที่ค่าคงที่(σ_β / τ และ $\mathbf{C}$) ของวิธี BMA_{SVT}และวิธี OPM มีค่าต่ำ ๆ วิธีOPM จะให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำใกล้เคียงกับวิธี BMA_{SVT} มาก ส่วนในกรณีที่ค่าคงที่ (σ_β / τ และ $\mathbf{C}$)ของวิธี BMA_{SVT}และวิธี OPM มีค่าสูงขึ้นจะทำให้การกระจายของพารามิเตอร์สัมประสิทธิ์การถดถอยมีการกระจายมากขึ้นทำให้ค่าที่สุ่มได้มีความแม่นยำลดลง จึงส่งผลให้วิธี BMA_{SVT} ให้ค่าพยากรณ์ที่มีความถูกต้องและแม่นยำมากกว่าวิธี OPM ชัดเจนยิ่งขึ้น ดังนั้นในการพิจารณาว่าวิธีการใดจะมีประสิทธิภาพมากที่สุดในกรณีที่วิธี BMA_{SVT} ให้ค่า AMSE ต่ำกว่าวิธีอื่น ๆ อย่างคงเส้นคงวาจึงควรใช้การวิเคราะห์ความแปรปรวนของการทดลองแบบสุ่มโดยสมบูรณ์ภายในกลุ่ม (Randomized Completely Block Design: RCBD) เพราะการวิเคราะห์ดังกล่าวจะทำให้ได้วิธีการคัดเลือกสมการถดถอยที่มีประสิทธิภาพมากที่สุดอย่างมีนัยสำคัญทางสถิติ โดยผู้วิจัยได้ยกตัวอย่างของการวิเคราะห์ความแปรปรวนของการทดลองแบบสุ่มโดยสมบูรณ์ภายในกลุ่มสำหรับผลการวิจัยตอนที่ 1 (ตารางที่4.1) และผลการวิจัยตอนที่ 2 (ตารางที่4.13) ดังนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการวิเคราะห์สำหรับตารางที่ 4.1

กรณี $\sigma = 0.25$

H_0: ไม่มีความแตกต่างกันระหว่างวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี

H_1: มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ

ระดับนัยสำคัญที่ 0.05

เขตปฏิเสธ จะปฏิเสธ H_0 ถ้า $F>F_{.95:3,\ 12}$ หรือ Sig. <.05

Tests of Between-Subjects Effects

Dependent Variable: AMSE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	44.788[a]	3	14.929	112.567	.000
METHOD	44.788	3	14.929	112.567	.000 (1)
Error	1.194	9	.133		
Total	45.982	12			

a. R Squared = .974 (Adjusted R Squared = .965)

Multiple Comparisons

Dependent Variable: AMSE

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	SR	OPM	1.209850*	.2575128	.001	.627315	1.792385
		SVT	1.746050*	.2575128	.000	1.163515	2.328585
	OPM	SR	-1.209850*	.2575128	.001	-1.792385	-.627315
		SVT	.536200	.2575128	(2) .067	-.046335	1.118735
	SVT	SR	-1.746050*	.2575128	.000	-2.328585	-1.163515
		OPM	-.536200	.2575128	.067	-1.118735	.046335
Bonferroni	SR	OPM	1.209850*	.2575128	.003	.454481	1.965219
		SVT	1.746050*	.2575128	.000	.990681	2.501419
	OPM	SR	-1.209850*	.2575128	.003	-1.965219	-.454481
		SVT	.536200	.2575128	(3) .201	-.219169	1.291569
	SVT	SR	-1.746050*	.2575128	.000	-2.501419	-.990681
		OPM	-.536200	.2575128	.201	-1.291569	.219169

Based on observed means.

*. The mean difference is significant at the .05 level.

จากผลการวิเคราะห์ (1) แสดงว่า มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ ณ ระดับนัยสำคัญ 0.05 และจาก (2), (3) แสดงว่าวิธี BMA_{SVT} และวิธี OPM ไม่มีความแตกต่างกัน ณ ระดับนัยสำคัญ 0.05 และวิธี SR มีความแตกต่างจากวิธีการคัดเลือกสมการถดถอยอื่น ๆ ณ ระดับนัยสำคัญ 0.05

กรณี $\sigma = 0.50$

H_0: ไม่มีความแตกต่างกันระหว่างวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี

H_1: มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ

ระดับนัยสำคัญที่ 0.05

เขตปฏิเสธ จะปฏิเสธ H_0 ถ้า $F>F_{.95:3,\ 12}$ หรือ Sig. <.05

Tests of Between-Subjects Effects

Dependent Variable: AMSE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	68.080[a]	3	22.693	100.768	.000
METHOD	68.080	3	22.693	100.768	.000 (1)
Error	2.027	9	.225		
Total	70.107	12			

a. R Squared = .971 (Adjusted R Squared = .961)

Multiple Comparisons

Dependent Variable: AMSE

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
LSD	SR	OPM	1.482700*	.3355612	.002	.723608	2.241792
		SVT	2.118825*	.3355612	.000	1.359733	2.877917
	OPM	SR	-1.482700*	.3355612	.002	-2.241792	-.723608
		SVT	.636125	.3355612	(2) .091	-.122967	1.395217
	SVT	SR	-2.118825*	.3355612	.000	-2.877917	-1.359733
		OPM	-.636125	.3355612	.091	-1.395217	.122967
Bonferroni	SR	OPM	1.482700*	.3355612	.005	.498390	2.467010
		SVT	2.118825*	.3355612	.000	1.134515	3.103135
	OPM	SR	-1.482700*	.3355612	.005	-2.467010	-.498390
		SVT	.636125	.3355612	(3) .272	-.348185	1.620435
	SVT	SR	-2.118825*	.3355612	.000	-3.103135	-1.134515
		OPM	-.636125	.3355612	.272	-1.620435	.348185

Based on observed means.

*. The mean difference is significant at the .05 level.

จากผลการวิเคราะห์ (1) แสดงว่า มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ ณ ระดับนัยสำคัญ 0.05 และจาก (2), (3) แสดงว่าวิธี BMA_{SVT} และวิธี OPM ไม่มีความแตกต่างกัน ณ ระดับนัยสำคัญ 0.05 และวิธี SRมีความแตกต่างจากวิธีการคัดเลือกสมการถดถอยอื่น ๆ ณ ระดับนัยสำคัญ 0.05

กรณี $\sigma = 2.50$

H_0: ไม่มีความแตกต่างกันระหว่างวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี

H_1: มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ

ระดับนัยสำคัญที่ 0.05

เขตปฏิเสธ จะปฏิเสธ H_0 ถ้า $F>F_{.95:3,\ 12}$ หรือ Sig. <.05

Tests of Between-Subjects Effects

Dependent Variable: AMSE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	68.080[a]	3	22.693	100.768	.000
METHOD	68.080	3	22.693	100.768	.000 (1)
Error	2.027	9	.225		
Total	70.107	12			

a. R Squared = .971 (Adjusted R Squared = .961)

Multiple Comparisons

Dependent Variable: AMSE

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	SR	OPM	1.482700*	.3355612	.002	.723608	2.241792
		SVT	2.118825*	.3355612	.000	1.359733	2.877917
	OPM	SR	-1.482700*	.3355612	.002	-2.241792	-.723608
		SVT	.636125	.3355612	(2) .091	-.122967	1.395217
	SVT	SR	-2.118825*	.3355612	.000	-2.877917	-1.359733
		OPM	-.636125	.3355612	.091	-1.395217	.122967
Bonferroni	SR	OPM	1.482700*	.3355612	.005	.498390	2.467010
		SVT	2.118825*	.3355612	.000	1.134515	3.103135
	OPM	SR	-1.482700*	.3355612	.005	-2.467010	-.498390
		SVT	.636125	.3355612	(3) .272	-.348185	1.620435
	SVT	SR	-2.118825*	.3355612	.000	-3.103135	-1.134515
		OPM	-.636125	.3355612	.272	-1.620435	.348185

Based on observed means.

*. The mean difference is significant at the .05 level.

จากผลการวิเคราะห์ (1) แสดงว่า มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ ณ ระดับนัยสำคัญ 0.05 และจาก (2), (3) แสดงว่าวิธี BMA_{SVT} และวิธี OPM ไม่มีความแตกต่างกัน ณ ระดับนัยสำคัญ 0.05 และวิธี SR มีความแตกต่างจากวิธีการคัดเลือกสมการถดถอยอื่น ๆ ณ ระดับนัยสำคัญ 0.05

ผลการวิเคราะห์สำหรับตารางที่ 4.13

กรณี $\sigma = 0.25$

H_0: ไม่มีความแตกต่างกันระหว่างวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี

H_1: มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ

ระดับนัยสำคัญที่ 0.05

เขตปฏิเสธ จะปฏิเสธ H_0 ถ้า $F>F_{.95:3,\ 12}$ หรือ Sig. <.05

Tests of Between-Subjects Effects

Dependent Variable: AMSE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	56.073[a]	3	18.691	106.250	.000
METHOD	56.073	3	18.691	106.250	.000
Error	1.583	9	.176		
Total	57.656	12			

(1)

a. R Squared = .973 (Adjusted R Squared = .963)

Multiple Comparisons

Dependent Variable: AMSE

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	SR	OPM	1.245700*	.2965762	.002	.574798	1.916602
		SVT	1.695400*	.2965762	.000	1.024498	2.366302
	OPM	SR	-1.245700*	.2965762	.002	-1.916602	-.574798
		SVT	.449700	.2965762	.164	-.221202	1.120602
	SVT	SR	-1.695400*	.2965762	.000	-2.366302	-1.024498
		OPM	-.449700	.2965762	.164	-1.120602	.221202
Bonferroni	SR	OPM	1.245700*	.2965762	.007	.375746	2.115654
		SVT	1.695400*	.2965762	.001	.825446	2.565354
	OPM	SR	-1.245700*	.2965762	.007	-2.115654	-.375746
		SVT	.449700	.2965762	.491	-.420254	1.319654
	SVT	SR	-1.695400*	.2965762	.001	-2.565354	-.825446
		OPM	-.449700	.2965762	.491	-1.319654	.420254

(2) (3)

Based on observed means.

*. The mean difference is significant at the .05 level.

จากผลการวิเคราะห์ (1) แสดงว่า มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ ณ ระดับนัยสำคัญ 0.05 และจาก (2), (3) แสดงว่าวิธี BMA_{SVT} และวิธีOPM ไม่มีความแตกต่างกัน ณ ระดับนัยสำคัญ 0.05 และวิธี SR มีความแตกต่างจากวิธีการคัดเลือกสมการถดถอยอื่น ๆ ณ ระดับนัยสำคัญ 0.05

กรณี $\sigma = 0.50$

H_0: ไม่มีความแตกต่างกันระหว่างวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี

H_1: มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ

ระดับนัยสำคัญที่ 0.05

เขตปฏิเสธ จะปฏิเสธ H_0 ถ้า $F > F_{.95:3,\ 12}$ หรือ Sig. <.05

Tests of Between-Subjects Effects

Dependent Variable: AMSE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	88.025[a]	3	29.342	168.811	.000
METHOD	88.025	3	29.342	168.811	.000
Error	1.564	9	.174		
Total	89.589	12			

a. R Squared = .983 (Adjusted R Squared = .977)

(1)

Multiple Comparisons

Dependent Variable: AMSE

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	SR	OPM	1.517925*	.2947991	.001	.851043	2.184807
		SVT	1.819025*	.2947991	.000	1.152143	2.485907
	OPM	SR	-1.517925*	.2947991	.001	-2.184807	-.851043
		SVT	.301100	.2947991	.334	-.365782	.967982
	SVT	SR	-1.819025*	.2947991	.000	-2.485907	-1.152143
		OPM	-.301100	.2947991	.334	-.967982	.365782
Bonferroni	SR	OPM	1.517925*	.2947991	.002	.653184	2.382666
		SVT	1.819025*	.2947991	.000	.954284	2.683766
	OPM	SR	-1.517925*	.2947991	.002	-2.382666	-.653184
		SVT	.301100	.2947991	1.000	-.563641	1.165841
	SVT	SR	-1.819025*	.2947991	.000	-2.683766	-.954284
		OPM	-.301100	.2947991	1.000	-1.165841	.563641

(2)

(3)

Based on observed means.

*. The mean difference is significant at the .05 level.

จากผลการวิเคราะห์ (1) แสดงว่า มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ ณ ระดับนัยสำคัญ 0.05 และจาก (2), (3) แสดงว่าวิธี BMA_{SVT} และวิธี OPM ไม่มีความแตกต่างกัน ณ ระดับนัยสำคัญ 0.05 และวิธี SRมีความแตกต่างจากวิธีการคัดเลือกสมการถดถอยอื่น ๆ ณ ระดับนัยสำคัญ 0.05

กรณี $\sigma = 2.50$

H_0: ไม่มีความแตกต่างกันระหว่างวิธีการคัดเลือกสมการถดถอยทั้ง 3 วิธี

H_1: มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ

ระดับนัยสำคัญที่ 0.05

เขตปฏิเสธ จะปฏิเสธ H_0 ถ้า $F > F_{.95:3,\ 12}$ หรือ Sig. <.05

Tests of Between-Subjects Effects

Dependent Variable: AMSE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Model	588.812[a]	3	196.271	322.837	.000
METHOD	588.812	3	196.271	322.837	.000 (1)
Error	5.472	9	.608		
Total	594.284	12			

a. R Squared = .991 (Adjusted R Squared = .988)

Multiple Comparisons

Dependent Variable: AMSE

	(I) Method	(J) Method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
LSD	SR	OPM	2.109500*	.5513417	.004	.862279	3.356721
		SVT	4.002600*	.5513417	.000	2.755379	5.249821
	OPM	SR	-2.109500*	.5513417	.004	-3.356721	-.862279
		SVT	1.893100*	.5513417	(2) .007	.645879	3.140321
	SVT	SR	-4.002600*	.5513417	.000	-5.249821	-2.755379
		OPM	-1.893100*	.5513417	.007	-3.140321	-.645879
Bonferroni	SR	OPM	2.109500*	.5513417	.012	.492236	3.726764
		SVT	4.002600*	.5513417	.000	2.385336	5.619864
	OPM	SR	-2.109500*	.5513417	.012	-3.726764	-.492236
		SVT	1.893100*	.5513417	(3) .022	.275836	3.510364
	SVT	SR	-4.002600*	.5513417	.000	-5.619864	-2.385336
		OPM	-1.893100*	.5513417	.022	-3.510364	-.275836

Based on observed means.

*. The mean difference is significant at the .05 level.

จากผลการวิเคราะห์ (1) แสดงว่า มีอย่างน้อย 1 วิธีการคัดเลือกสมการถดถอยที่แตกต่างจากวิธีอื่น ๆ ณ ระดับนัยสำคัญ 0.05 และจาก (2), (3) แสดงว่าวิธี BMA_{SVT} วิธี OPM และวิธี SR มีความแตกต่างกัน ณ ระดับนัยสำคัญ 0.05

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวจิตติมา ผสมญาติ เกิดเมื่อวันที่ 23 ธันวาคม 2520 ที่จังหวัดชลบุรี สำเร็จการศึกษาระดับปริญญาตรีวิทยาศาสตรบัณฑิต เกียรตินิยมอันดับสอง สาขาเศรษฐศาสตร์ จากคณะเศรษฐศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ เมื่อ พ.ศ.2541 และเข้าศึกษาต่อในระดับปริญญามหาบัณฑิต สาขาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ.2544

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย