การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียโดยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่

กรองทิพย์ เนียมถนอม

ดุษฎีนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต
สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา
วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา
สิงหาคม 2559

คณะกรรมการควบคุมดุษฎีนิพนธ์และคณะกรรมการสอบดุษฎีนิพนธ์ ได้พิจารณาดุษฎีนิพนธ์ของ กรองทิพย์ เนียมถนอม ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพาได้

คณะกรรมการควบคุมดุษฎีนิพนธ์

....อาจารย์ที่ปรึกษาหลัก

(รองศาสตราจารย์ ดร.เสรี ชัดแช้ม)

....อาจารย์ที่ปรึกษาร่วม

(ผู้ช่วยศาสตราจารย์ ดร.กิดาการ สายธนู)

คณะกรรมการสอบดุษฎีนิพนธ์

....ประธาน

(ผู้ช่วยศาสตราจารย์ ดร.กุศยา ปลั่งพงษ์พันธ์)

....กรรมการ

(รองศาสตราจารย์ ดร.เสรี ชัดแช้ม)

....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.กิดาการ สายธนู)

....กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.พูลพงศ์ สุขสว่าง)

วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา อนุมัติให้รับดุษฎีนิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาวิชาการวิจัยและสถิติทางวิทยาการปัญญา ของมหาวิทยาลัยบูรพา

....คณบดีวิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา

(ผู้ช่วยศาสตราจารย์ ดร.สุชาดา กรเพชรปาณี)

วันที่...1...เดือน...สิงหาคม...พ.ศ...2559...

ประกาศคุณูปการ

ดุษฎีนิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วย ความกรุณาจาก รองศาสตราจารย์ ดร.เสรี ชัดแช้ม อาจารย์ที่ปรึกษาหลัก และผู้ช่วยศาสตราจารย์ ดร.กิดาการ สายธนู อาจารย์ที่ปรึกษาร่วม ที่กรุณาให้คำแนะนำ คำปรึกษา ตลอดจนช่วยเหลือแก้ไขข้อบกพร่องต่าง ๆ เป็นอย่างดียิ่ง จนกระทั่งดุษฎีนิพนธ์ฉบับนี้เสร็จสมบูรณ์ ผู้วิจัยขอกราบขอบพระคุณด้วยความรู้สึกซาบซึ้งและสำนึกในพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ผู้วิจัยขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.สนิท สิทธิ รองผู้อำนวยการฝ่ายเทคโนโลยีสารสนเทศ สำนักบริหารและพัฒนาวิชาการ มหาวิทยาลัยแม่โจ้ ที่กรุณาตรวจสอบความถูกต้องของการเขียนโปรแกรมในการวิเคราะห์ผลงานวิจัย และขอขอบคุณส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ที่กรุณาให้คำแนะนำข้อมูลยางพารา

ผู้วิจัยขอกราบขอบพระคุณครูอาจารย์ทุกท่าน ที่ได้ให้โอกาสทางการศึกษาและประสิทธิ์ประสาทวิชาความรู้ให้แก่ผู้วิจัยจนกระทั่งสำเร็จการศึกษา

ผู้วิจัยขอระลึกถึงพระคุณ คุณพ่อ คุณแม่ ที่ให้ความสนับสนุนด้านการศึกษาและให้กำลังใจจนกระทั่งสำเร็จการศึกษา

ท้ายนี้คุณค่าและประโยชน์ของดุษฎีนิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นกตัญญูกตเวทิตาแด่บุพการี บูรพาจารย์ และผู้มีพระคุณทุกท่านทั้งในอดีตและปัจจุบัน ที่ทำให้ผู้วิจัยเป็นผู้มีการศึกษาและประสบความสำเร็จมาจนถึงทุกวันนี้

กรองทิพย์ เนียมถนอม

51810899: สาขาวิชา: การวิจัยและสถิติทางวิทยาการปัญญา;
ปร.ด. (การวิจัยและสถิติทางวิทยาการปัญญา)
คำสำคัญ: พหุสัมพันธ์/ ค่าผิดปกติ/ อัตสหสัมพันธ์/ การถดถอยที่มีความแกร่ง/ วิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่/ อุปสงค์การส่งออกยางพาราของไทย

กรองทิพย์ เนียมถนอม: การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียโดยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ (FORECASTING THAI RUBBER EXPORTING DEMAND TO MALAYSIA USING NEW PARAMETER ESTIMATION IN MULTIPLE LINEAR REGRESSION) คณะกรรมการควบคุมดุษฎีนิพนธ์: เสรี ชัดแช้ม, ค.ด., กิดาการ สายธนู, Ph.D. 180 หน้า. ปี พ.ศ. 2559.

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณและเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตสหสัมพันธ์ของค่าคลาดเคลื่อน จำลองข้อมูลด้วยเทคนิคมอนติคาร์โลภายใต้สถานการณ์ 162 สถานการณ์ และพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียด้วยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น โดยนำมาใช้กับข้อมูลทุติยภูมิแบบอนุกรมเวลาเป็นรายปีตั้งแต่ ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2556 ผลการวิจัยปรากฏว่า

1) ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตสหสัมพันธ์ของค่าคลาดเคลื่อน คือ $\hat{\mathbf{b}}_{NEW} = (\mathbf{Q}'\mathbf{WQ} + k\mathbf{I})^{-1}(\mathbf{Q}'\mathbf{WR} + k\mathbf{J})$

2) ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นใช้ได้ดีกว่าวิธีอื่นภายใต้สถานการณ์ 61 สถานการณ์ เมื่อข้อมูลมีค่าอัตสหสัมพันธ์ 0.3, 0.5 และ 0.9 กรณีข้อมูลมีค่าผิดปกติในตัวแปรตาม ขนาดตัวอย่างเท่ากับ 20 ข้อมูลมีค่าผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันระดับ 0.3 แต่ถ้าขนาดตัวอย่างเท่ากับ 50 และ 100 ข้อมูลมีค่าผิดปกติ 1% และ 5% ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันระดับ 0.3 และ 0.5 ในกรณีข้อมูลมีค่าผิดปกติในตัวแปรอิสระและตัวแปรตาม ขนาดตัวอย่างเท่ากับ 20 และ 50 ข้อมูลมีค่าผิดปกติ 1% และ 5% ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันระดับ 0.3 ขนาดตัวอย่างเท่ากับ 100 ข้อมูลผิดปกติ 1% และ 5% ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันระดับ 0.3 และ 0.5 และในกรณีข้อมูลมีค่าผิดปกติในตัวแปรอิสระ ไม่เหมาะสมที่จะใช้ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

3) สมการพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น คือ $z_y = 0.274685z_1 - 0.495796z_2 + 1.057818z_3$ เมื่อ z_y เป็นปริมาณการส่งออกยางพาราไทยไปมาเลเซียในรูปมาตรฐาน z_1 เป็นผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซียในรูปมาตรฐาน z_2 เป็นราคาส่งออกยางพาราไทยในรูปมาตรฐาน และ z_3 เป็นปริมาณการส่งออกยางพาราไทยไปมาเลเซียในปีที่ผ่านมาในรูปมาตรฐาน ค่า $R^2_{adj} = 0.848639$ เมื่อพยากรณ์ปริมาณการส่งออกยางพาราไทยไปมาเลเซีย ในปี พ.ศ. 2557 พบว่า มีปริมาณเฉลี่ย 661.565 พันตัน

51810899: MAJOR: RESEARCH AND STATISTICS IN COGNITIVE SCIENCE;
Ph.D. (RESEARCH AND STATISTICS IN COGNITIVE SCIENCE)
KEYWORDS: MULTICOLLINEARITY/ OUTLIERS/ AUTOCORRELATION/ ROBUST REGRESSION/ NEW PARAMETER ESTIMATION IN MULTIPLE LINEAR REGRESSION/ THE DEMAND FOR THAI RUBBER

KRONGTHIP NEAMTHANOM: FORECASTING THAI RUBBER EXPORTING DEMAND TO MALAYSIA USING NEW PARAMETER ESTIMATION IN MULTIPLE LINEAR REGRESSION. ADVISORY COMMITTEE: SEREE CHADCHAM, Ph.D., KIDAKAN SAITHANU, Ph.D. 180 P. 2016.

The purposes of this study were to develop and compared the estimator efficiency of linear multiple regression coefficients among the methods of Ordinary Least Square, Generalized Least Square, Ridge Regressions, Julmusik's method, and Generalized M-estimator Ridge if the problems of multicollinearity, outliers, and autocorrelation occurred, using Monte Carlo simulations under 162 situations, to predict Thai rubber exporting demand to Malaysia. The developed estimation method of coefficient of multiple linear regression was employed with the annual secondary time series data between 1999 and 2013.

The results were as follows:

1) The new estimator of linear multiple regression in the case of problems of multicollinearity, outliers, and autocorrelation was $\hat{\mathbf{b}}_{NEW} = (\mathbf{Q'WQ} + k\mathbf{I})^{-1}(\mathbf{Q'WR} + k\mathbf{J})$

2) The developed coefficient estimates of the multiple linear regressions were better than those of other methods under 61 situations where the autocorrelations were 0.3, 0.5, and 0.9. In the case of outliers in the dependent variable, when the sample size was 20 the data were 1% and 5% outliers and the multicollinearity was 0.3. When the sample sizes were 50 and 100 the data of 1% and 5% outliers should be used when the multicollinearity was 0.3 and 0.5. In the case of outliers in the independent and dependent variables, when the sample sizes were 20 and 50, the data of 1% and 5% outliers should be used when the multicollinearity was 0.3. When the sample size was 100 the data of 1% and 5% outliers should be used when the multicollinearity was 0.3 and 0.5, but should not be used in the presence of outliers among the independent variables.

3) The equation of Thai rubber exporting demand to Malaysia using the developed estimator of linear multiple regression coefficient was $z_y = 0.274685z_1 - 0.495796z_2 + 1.057818z_3$ where z_y is exporting volumes of the Thai rubber to Malaysia in standard form, z_1 is Malaysia's gross domestic product in standard form, z_2 is exporting price of Thai rubber in standard form, and z_3 is exporting volumes of Thai rubber to Malaysia of the preceding year in standard form. $R^2_{adj} = 0.848639$ can be predicted that the amount of Thai rubber exports to Malaysia in A.D. 2014 will be on average by 661.565 thousand tons.

สารบัญ

หน้า

บทคัดย่อภาษาไทย.......... ง
บทคัดย่อภาษาอังกฤษ.......... จ
สารบัญ.......... ฉ
สารบัญตาราง.......... ณ
สารบัญภาพ.......... ฎ
บทที่

1 บทนำ.......... 1
ความเป็นมาและความสำคัญของปัญหา.......... 1
วัตถุประสงค์ของการวิจัย.......... 6
กรอบแนวคิดในการวิจัย.......... 6
สมมติฐานของการวิจัย.......... 8
ประโยชน์ที่ได้รับจากการวิจัย.......... 9
ขอบเขตของการวิจัย.......... 9
นิยามศัพท์เฉพาะ.......... 11
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.......... 12
ตอนที่ 1 หลักการ แนวคิด และทฤษฎีที่.......... 12
1. วิธีการประมาณค่าพารามิเตอร์.......... 12
2. ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity).......... 24
3. ปัญหาการมีข้อมูลที่มีค่าผิดปกติ (Outliers).......... 27
4. ปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน (Autocorrelation).......... 29
ตอนที่ 2 การจำลองสถานการณ์ด้วยเทคนิคมอนติคาร์โล.......... 52
1. ความหมายของเทคนิคมอนติคาร์โล.......... 52
2. ความเป็นมาของเทคนิคมอนติคาร์โล.......... 53
3. ขั้นตอนของเทคนิคมอนติคาร์โล.......... 53
ตอนที่ 3 ความรู้เกี่ยวกับยางพารา.......... 55
1. ความเป็นมาของยางพาราในประเทศไทย.......... 55
2. ลักษณะทางพฤกษศาสตร์ของต้นยางพารา.......... 56
3. สภาพการทำสวนยางพาราของประเทศไทย.......... 56
4. การผลิต การส่งออก และการใช้ยางธรรมชาติในประเทศต่าง ๆ.......... 57
5. การผลิต การส่งออก และการใช้ยางธรรมชาติในประเทศไทย.......... 59
6. การส่งออกยางพาราของไทยไปประเทศมาเลเซีย.......... 61
ตอนที่ 4 ความรู้เกี่ยวกับทฤษฎีอุปสงค์.......... 61

สารบัญ (ต่อ)

บทที่ หน้า

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน....... 65

ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับอุปสงค์ของการส่งออกยางพาราของไทย.................... 76

3 วิธีดำเนินการวิจัย.. 82

ตอนที่ 1 การพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ในกรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน.. 82

ตอนที่ 2 การเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธี คือวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม... 88

ตอนที่ 3 การพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น...... 97

4 ผลการวิจัย.. 99

ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ กรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน... 99

ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธี คือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม... 100

ตอนที่ 3 ผลการพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น....... 124

5 สรุปและอภิปรายผล.. 127

สรุปผลการวิจัย.. 127

อภิปรายผล.. 133

ข้อเสนอแนะ.. 134

บรรณานุกรม.. 136

ภาคผนวก.. 143

สารบัญ (ต่อ)

บทที่ หน้า

ภาคผนวก ก การเขียนโปรแกรมชุดคำสั่งเพื่อประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี........ 144

ภาคผนวก ข ข้อมูลยางพารา และการทดสอบข้อสมมติของตัวแบบการถดถอย ค่าประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการประมาณค่า พารามิเตอร์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี........ 155

ภาคผนวก ค การเขียนโปรแกรมชุดคำสั่งเพื่อพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่........ 162

ภาคผนวก ง ตารางสถิติเดอร์บิน-วัตสัน........ 170

ภาคผนวก จ การพิสูจน์ค่าคาดหวัง ค่าความแปรปรวนร่วม และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าวิธีการถดถอยริดจ์........ 172

ภาคผนวก ฉ การปรับข้อมูลให้อยู่ในรูปมาตรฐาน........ 177

ประวัติย่อของผู้วิจัย........ 180

สารบัญตาราง

ตารางที่ หน้า

1 ปริมาณการผลิตยางธรรมชาติของประเทศต่างๆ ปี พ.ศ. 2550-2554 57
2 ปริมาณการส่งออกยางธรรมชาติของประเทศต่างๆ ปี พ.ศ. 2550–2554 58
3 ปริมาณการใช้ยางธรรมชาติของประเทศต่างๆ ปี พ.ศ. 2550–2554 59
4 ปริมาณการผลิตการส่งออก การใช้ในประเทศ และสต็อกของยางธรรมชาติของประเทศไทย ปี พ.ศ. 2542–2557 60
5 ตลาดส่งออกยางธรรมชาติที่สำคัญของไทยปี พ.ศ. 2553–2557 60
6 ปริมาณการส่งออกยางพาราของไทยไปมาเลเซีย ปี พ.ศ. 2553–2557 61
7 จำนวนค่าที่ผิดปกติในแต่ละขนาดตัวอย่าง จำแนกตามเปอร์เซนต์การเกิดค่าผิดปกติ 91
8 จำนวนค่าที่ผิดปกติและการกำหนดค่าผิดปกติ 91
9 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 106
10 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 108
11 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 110
12 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 112
13 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 114

สารบัญตาราง (ต่อ)

ตารางที่ หน้า

14 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100........ 116

15 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20........ 118

16 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50........ 120

17 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100........ 122

18 ค่าประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าสัมประสิทธิ์การกำหนดของวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่........ 126

สารบัญภาพ

ภาพที่ หน้า

1 กรอบแนวคิดในการวิจัยเรื่องการพัฒนาวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่.. 7

2 กรอบแนวทางของการวิจัยในภาพรวม.. 8

3 แผนผังของการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุแบบใหม่......... 87

4 แผนผังของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล.. 93

5 แผนผังการทำงานของวิธีกำลังสองน้อยที่สุดถ่วงน้ำหนักชนิดทำซ้ำ (IRLS)............................... 96

6 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20.. 107

7 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50.. 109

8 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100.. 111

9 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20.. 113

10 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50.. 115

11 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100.. 117

สารบัญภาพ (ต่อ)

ภาพที่ หน้า

12 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 119

13 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 121

14 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 6 วิธี กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 123

บทที่ 1
บทนำ

ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันวิธีการทางสถิติได้ถูกนำมาใช้ในการวิเคราะห์ข้อมูลกันอย่างแพร่หลาย การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติวิธีการหนึ่งที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไปโดยพิจารณารูปแบบความเป็นไปได้ของความสัมพันธ์ ที่มีเป้าหมายของการวิเคราะห์ คือ หาสมการทางคณิตศาสตร์ที่ใช้เป็นตัวแทนแสดงความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะประมาณค่าหรือพยากรณ์ค่าของตัวแปรหนึ่งจากค่าที่กำหนดให้ของตัวแปรอื่น ๆ ที่เกี่ยวข้องกัน ตัวแปรในตัวแบบการถดถอยจำแนกออกเป็น 2 ประเภท คือ ตัวแปรที่ต้องการประมาณหรือพยากรณ์เรียกว่าตัวแปรตาม (Dependent Variable) หรือตัวแปรตอบสนอง (Response Variable) และตัวแปรที่มีอิทธิพลต่อตัวแปรตามเรียกว่าตัวแปรอิสระ (Independent Variable) หรือตัวแปรอธิบาย (Explainable Variable) การวิเคราะห์การถดถอยเป็นวิธีการวิเคราะห์ข้อมูลเพื่ออธิบายความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ในการวิเคราะห์การถดถอยนั้นจะมีตัวแปรอิสระเป็นตัวแปรเชิงปริมาณเพียงอย่างเดียว หรืออาจจะมีตัวแปรอิสระบางตัวเป็นตัวแปรเชิงปริมาณและตัวแปรอิสระบางตัวเป็นตัวแปรเชิงกลุ่ม ในขณะที่ตัวแปรตามจะเป็นตัวแปรเชิงปริมาณ กรณีความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามเป็นแบบเชิงเส้น ถ้ามีตัวแปรอิสระเพียงตัวเดียว จะเรียกการวิเคราะห์การถดถอยเชิงเส้นเชิงเดียว (Simple Linear Regression Analysis) และถ้ามีตัวแปรอิสระหลายตัว จะเรียกการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis)

การวิเคราะห์การถดถอย จะต้องทำภายใต้ข้อสมมติ (Assumption) เกี่ยวกับค่าคลาดเคลื่อนสุ่ม (Random Error) ดังนั้นเมื่อได้ตัวแบบการถดถอยแล้วจำเป็นต้องตรวจสอบค่าคลาดเคลื่อนสุ่มว่าเป็นไปตามข้อสมมติหรือไม่ แต่เนื่องจากไม่ทราบค่าของค่าคลาดเคลื่อนสุ่มจริง ในการตรวจสอบจึงใช้ค่าตกค้าง (Residuals) ซึ่งเป็นค่าประมาณของค่าคลาดเคลื่อนสุ่มมาใช้ในการตรวจสอบแทน และผลการตรวจสอบจะถูกนำไปใช้ในการพิจารณาปรับหาตัวแบบการถดถอยใหม่ เพื่อให้ได้ตัวแบบที่เหมาะสม ตัวแบบที่นำมาใช้แสดงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ คือ

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon_i \; ; i = 1, 2, \ldots, n$$

เมื่อ	Y_i	แทน	ค่าสังเกตที่ i ของตัวแปรตามที่ทราบค่า
	$X_{i1}, X_{i2}, \ldots, X_{ip}$	แทน	ค่าสังเกตที่ i ของตัวแปรอิสระ p ตัวที่ทราบค่า
	$\beta_0, \beta_1, \beta_2, \ldots, \beta_p$	แทน	สัมประสิทธิ์การถดถอยที่ไม่ทราบค่า
	ε_i	แทน	ค่าคลาดเคลื่อนสุ่มของค่าสังเกตที่ i
	n	แทน	จำนวนค่าสังเกต
และ	p	แทน	จำนวนตัวแปรอิสระ

หรืออาจเขียนความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระได้ในรูปของเมทริกซ์ ดังนี้

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

เมื่อ $\mathbf{Y}$ แทน เวกเตอร์ของค่าสังเกตของตัวแปรตามที่ทราบค่า มีมิติ n x 1

$\mathbf{X}$ แทน เมทริกซ์ของค่าสังเกตของตัวแปรอิสระที่ทราบค่า มีมิติ n x (p+1)

$\boldsymbol{\beta}$ แทน เวกเตอร์ของสัมประสิทธิ์การถดถอยที่ไม่ทราบค่า มีมิติ (p+1) x 1

$\boldsymbol{\varepsilon}$ แทน เวกเตอร์ของค่าคลาดเคลื่อนสุ่ม มีมิติ n x 1

n แทน จำนวนค่าสังเกต

และ p แทน จำนวนตัวแปรอิสระ

เนื่องจาก $\beta_0, \beta_1, \beta_2, ..., \beta_p$ เป็นพารามิเตอร์ของตัวแบบการถดถอยที่ไม่ทราบค่า หากสามารถประมาณค่าพารามิเตอร์เหล่านี้ได้ใกล้เคียงกับค่าจริงแล้วจะทำให้การพยากรณ์ถูกต้องมากยิ่งขึ้น

ขั้นตอนในการวิเคราะห์การถดถอยมีหลายขั้นตอน ซึ่งขั้นตอนที่สำคัญ คือ การคัดเลือกตัวแปรอิสระที่มีอิทธิพลต่อการอธิบายความแปรผันของตัวแปรตาม และการประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอย การประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอยมีด้วยกันหลายวิธี วิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่นิยมใช้กันมากที่สุดคือ วิธีกำลังสองน้อยที่สุดสามัญ (Ordinary Least Square Method: OLS) ซึ่งจะมีตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณอยู่ในรูปของ $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ โดยจะเป็นวิธีที่ทำให้ได้ตัวประมาณที่มีคุณสมบัติไม่เอนเอียง และมีความแปรปรวนต่ำที่สุด (Minimum Variance Unbiased Estimator: MVUE) และจะเรียกตัวประมาณที่มีคุณสมบัติดังกล่าวนี้ว่าตัวประมาณค่าไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator: BLUE) ภายใต้ข้อสมมติ คือ

1. ε_i เป็นตัวแปรสุ่ม
2. ε_i มีการแจกแจงปกติ ที่มีค่าเฉลี่ยเท่ากับ 0 คือ $\mathrm{E}(\varepsilon_i) = 0$ และค่าความแปรปรวนคงที่ (σ^2) คือ $\mathrm{V}(\varepsilon_i) = \sigma^2$
3. ε_i และ ε_j เป็นอิสระซึ่งกันและกัน คือ $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$ โดยที่ $i \neq j$
4. Y_i และ Y_j จะต้องเป็นอิสระซึ่งกันและกัน และมีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + \varepsilon_i$ และค่าความแปรปรวนคงที่ (σ^2)
5. ตัวแปรอิสระแต่ละตัวมีความเป็นอิสระกัน

ปัญหาของการวิเคราะห์การถดถอยเชิงเส้นพหุคูณที่พบเนื่องมาจากข้อสมมติของตัวแบบการถดถอยไม่เป็นจริงได้แก่

1. ค่าคลาดเคลื่อนไม่ได้มีการแจกแจงปกติ
2. ค่าคลาดเคลื่อนไม่เป็นอิสระกัน
3. ความแปรปรวนของค่าคลาดเคลื่อนมีค่าไม่คงที่
4. รูปแบบของการถดถอยไม่ถูกต้อง
5. ตัวแปรอิสระบางตัวซึ่งสำคัญต่อตัวแปรตามไม่ได้ถูกนำมาวิเคราะห์ในการถดถอย
6. ตัวแปรอิสระในตัวแบบการถดถอยมีความสัมพันธ์กัน

7. ค่าของตัวแปรตาม หรือค่าของตัวแปรอิสระบางค่าเป็นค่าผิดปกติ (Outlier) หรือเป็นค่าที่มีอิทธิพล (Influential)

ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดสามัญหากข้อมูลประกอบด้วยค่าสูญหาย (Missing Value) ค่าผิดปกติ หรือเกิดปัญหาของค่าคลาดเคลื่อนไม่ได้มีการแจกแจงปกติ หรือเกิดการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) แล้วจะทำให้ผลลัพธ์ที่วิเคราะห์ได้เป็นไปในทางที่ไม่ถูกต้อง แต่ในบางครั้งการวิเคราะห์การถดถอยเชิงเส้นพหุคูณตัวแปรอิสระที่นำมาศึกษาอาจมีความสัมพันธ์กัน ทำให้เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอย (Variance-Covariance Matrix of the Regression Coefficient Estimator) มีค่าสูงซึ่งจะส่งผลให้การทดสอบสมมติฐานเกี่ยวกับค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ได้ขาดความแม่นยำ

ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรหรือตัวแปรอิสระมีความสัมพันธ์กัน เป็นปัญหาสำคัญที่พบโดยทั่วไป โดยเฉพาะข้อมูลทางด้านเศรษฐศาสตร์ สาเหตุของการเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์กันนั้นอาจเกิดขึ้นเนื่องมาจากการใช้ตัวแปรในอดีตมาเป็นตัวแปรอิสระ การเปลี่ยนแปลงของสภาพเศรษฐกิจจะมีผลกระทบกับปัจจัยต่าง ๆ ในทางเศรษฐศาสตร์ในลักษณะเดียวกันหรือตรงข้ามกัน การมีตัวแปรอิสระมากเกินไป หรือขนาดของข้อมูลตัวอย่างไม่เพียงพอ เมื่อตัวแปรอิสระมีความสัมพันธ์กันผลกระทบที่เกิดขึ้นคือจะทำให้ค่าประมาณสัมประสิทธิ์การถดถอยมีค่าแตกต่างไปจากความเป็นจริงและยังทำให้มีเครื่องหมายไม่ตรงกับข้อเท็จจริง (ทรงศิริ แต้สมบัติ, 2548, หน้า 182) มีวิธีการแก้ปัญหาโดยการตัดตัวแปรอิสระบางตัวที่มีความสัมพันธ์กันสูงกับตัวแปรอิสระตัวอื่นออกไปจากตัวแบบเชิงคณิตศาสตร์ โดยพิจารณาจากตัวแปรอิสระที่มีความสัมพันธ์กัน และตัดตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามน้อยกว่าออกไป เช่น การกำจัดตัวแปรแบบลดตัวแปรอิสระ (Backward Elimination) การคัดเลือกตัวแปรแบบเพิ่มตัวแปรอิสระ (Forward Selection) การถดถอยแบบทีละขั้นตอน (Stepwise Regression) และการวิเคราะห์ส่วนประกอบหลัก (Principle Component Analysis) สำหรับกรณีที่ไม่ต้องการตัดตัวแปรอิสระตัวใดออกจากตัวแบบอาจเนื่องมาจากตัวแปรอิสระทุกตัวมีผลต่อการเปลี่ยนแปลงของตัวแปรตามเท่า ๆ กันหรือกล่าวได้ว่าตัวแปรอิสระทุกตัวมีความสำคัญต่อตัวแปรตามเท่ากัน หรือมีความสัมพันธ์ระหว่างตัวแปรอิสระไม่ชัดเจนพอ ได้มีผู้เสนอศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ดังนี้

ในปี ค.ศ. 1970 Hoerl and Kennard ได้เสนอวิธีการแก้ปัญหาตัวแปรอิสระมีความสัมพันธ์กันโดยไม่ต้องตัดตัวแปรอิสระทิ้ง โดยเป็นการศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองน้อยกว่าวิธีกำลังสองน้อยที่สุดสามัญ ซึ่งเรียกว่าวิธีการถดถอยริดจ์ (Ridge Regression Method) หลักการของวิธีนี้จะเป็นการนำค่าคงที่ k ที่มีค่าอยู่ระหว่าง 0 และ 1 มาบวกกับสมาชิกในแนวทแยงมุมของเมทริกซ์ $\mathbf{X'X}$ เพื่อลดค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากวิธีกำลังสองน้อยที่สุดสามัญให้มีค่าน้อยลง จึงได้ตัวแบบของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ คือ

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X} + k\mathbf{I})^{-1}\mathbf{X'Y} \;; 0 < k < 1$$

แต่ตัวประมาณที่ได้เป็นตัวประมาณที่มีคุณสมบัติเอนเอียง และค่า k ที่ใช้มีรูปแบบไม่แน่นอน

ต่อมา ปีค.ศ. 1993 Liu ได้เสนอวิธีเจเนรัลไลซ์ ลู (Generalized Liu Method) ที่พัฒนามาจาก 2 วิธีการ คือ วิธีการถดถอยริดจ์ที่เสนอโดย Hoerl and Kennard (1970) และวิธีสไตน์ (Stein Estimation) ซึ่งตัวประมาณที่ได้เป็นตัวประมาณค่าที่มีคุณสมบัติเอนเอียงปี ค.ศ. 1995 Crouse, Jinand Hanumara ได้พัฒนาวิธีการถดถอยริดจ์ที่เสนอโดย Hoerl and Kennard เรียกว่าการถดถอยริดจ์ที่มีข้อมูลเบื้องต้น (Ridge Regression with Prior Information Method) วิธีการนี้เป็นวิธีที่นำข้อมูลในอดีตหรือข้อมูลที่มีประโยชน์มาเป็นส่วนประกอบเสริมวิธีการถดถอยริดจ์จึงทำให้ได้ตัวประมาณที่มีคุณสมบัติไม่เอนเอียง และมีสูตรในการคำนวณค่า k ในรูปที่แน่นอน ในปีเดียวกัน Akdeniz and Kaciranlasr ได้เสนอวิธีออลโมส อันไบแอส เจเนรัลไลซ์ ลู (Almost Unbiased Generalized Liu Method) ซึ่งเป็นวิธีที่พัฒนามาจากวิธีเจเนรัลไลซ์ ลู ที่เสนอโดย Liu วิธีออลโมส อันไบแอส เจเนรัลไลซ์ ลู เป็นวิธีที่ใช้หลักการของวิธีการถดถอยริดจ์ และวิธีการประมาณค่าแบบสไตน์ร่วมกัน ซึ่งตัวประมาณที่ได้จะเป็นตัวประมาณที่มีคุณสมบัติเอนเอียงน้อยกว่าวิธีเจเนรัลไลซ์ ลู ในปี ค.ศ. 2009 Muniz and Kibria ได้เสนอวิธีการประมาณค่า k ของการถดถอยริดจ์ โดยพัฒนามาจากวิธีของการถดถอยริดจ์ที่เสนอโดย Hoerland Kennard เพื่อทำให้มีค่าเฉลี่ยคลาดเคลื่อนกำลังสองของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณน้อยกว่าวิธีกำลังสองน้อยที่สุดสามัญ

การเก็บรวบรวมข้อมูลในบางครั้งอาจพบปัญหาข้อมูลที่เก็บรวบรวมมามีค่าผิดปกติเกิดขึ้น คือ มีค่าสังเกตบางค่าสูงหรือต่ำกว่าค่าสังเกตส่วนใหญ่ หากพบว่าค่าผิดปกติเหล่านั้นเกิดจากธรรมชาติของข้อมูล จะต้องนำค่าผิดปกติเหล่านั้นมารวมในการวิเคราะห์การถดถอยด้วย (ทรงศิริ แต้สมบัติ, 2548, หน้า 211) เนื่องจากข้อมูลที่มีค่าผิดปกติเหล่านั้นจะมีผลต่อการประมาณค่าพารามิเตอร์ของวีธีกำลังสองน้อยที่สุดสามัญ สำหรับการแก้ปัญหานี้มีนักสถิติหลายคนเสนอวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณที่สามารถช่วยลดอิทธิพลของข้อมูลที่มีค่าผิดปกติ โดยการให้น้ำหนักของค่าสังเกตที่ผิดปกติน้อยกว่าค่าสังเกตที่เป็นข้อมูลส่วนใหญ่ วิธีการเหล่านี้จัดอยู่ในการวิเคราะห์การถดถอยที่มีความแกร่ง (Robust Regression) โดยการวิเคราะห์การถดถอยที่มีความแกร่ง เป็นการวิเคราะห์การถดถอยที่ใช้หลักของการลดอิทธิพลของข้อมูลที่ผิดปกติลง โดยการสร้างสมการถดถอยสำหรับข้อมูลส่วนใหญ่ แล้วตรวจสอบข้อมูลที่ผิดปกติจากสมการถดถอยที่สร้างเป็นเกณฑ์ มีนักสถิติหลายคนได้ศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่มีความแกร่งหลายวิธี เช่น วิธีตัวประมาณค่าเอ็ม (M–Estimator Method) วิธีค่าสัมบูรณ์น้อยที่สุด (Least Absolute Value Method) วิธีตัวประมาณค่าเอ็มเอ็ม (MM–Estimator Method) วิธีค่ามัธยฐานกำลังสองน้อยที่สุด (Least Median Square Estimator: LMS) และวิธีตัวประมาณค่าแอลทีเอส (Least Trimmed Square Estimator: LTS) ในปี ค.ศ.1996 Simpson and Montgomery ได้พัฒนาวิธีประมาณค่าสัมประสิทธิ์การถดถอยจากวิธีตัวประมาณค่าเอ็มมาเป็นวิธีตัวประมาณค่าจีเอ็ม (Generalized M-Estimator Method: GM) แล้วรวมกับการประมาณโดยวิธีการถดถอยริดจ์ เรียกว่า วิธีริดจ์ชนิดตัวประมาณจีเอ็ม (Generalized M-Estimator Ridge Method: GMR) เพื่อประมาณค่าสัมประสิทธิ์การถดถอยในกรณีเมื่อเกิดปัญหาการมีความสัมพันธ์ระหว่างตัวแปรอิสระและการมีค่าผิดปกติเกิดขึ้นพร้อม

การประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดสามัญยังมีข้อสมมติที่สำคัญอีกอย่างหนึ่ง คือ ค่าคลาดเคลื่อนเป็นอิสระกัน ซึ่งในทางปฏิบัติข้อมูลทางด้านเศรษฐศาสตร์ที่เก็บรวบรวมมาเป็นข้อมูลอนุกรมเวลา ปัญหาที่พบบ่อยคือ ค่าคลาดเคลื่อนมีความสัมพันธ์กันหรือเรียกว่ามีอัตตสหสัมพันธ์ (Autocorrelation) ซึ่งอาจเกิดขึ้นจากสาเหตุการกำหนดฟังก์ชันความสัมพันธ์ในแบบจำลองผิดพลาด การละเลยตัวแปรอิสระบางตัวที่สำคัญ หรือกำหนดรูปแบบความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระไม่ถูกต้อง การแปลงข้อมูล หรือความผิดพลาดอันเกิดจากการวัด ซึ่งส่งผลให้การประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญไม่มีประสิทธิภาพและส่งผลให้ตัวประมาณไม่มีคุณสมบัติตัวประมาณค่าไม่เอนเอียงเชิงเส้นที่ดีที่สุด วิธีการแก้ปัญหาเมื่อเกิดอัตตสหสัมพันธ์ เช่น การเพิ่มตัวแปรอิสระ การเปลี่ยนรูปแบบความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระหรือใช้การแปลงข้อมูล และการแปลงแบบจำลองที่เกิดปัญหาไปเป็นแบบจำลองที่ไม่มีอัตตสหสัมพันธ์ โดยการประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดทั่วไปแทนวิธีกำลังสองน้อยที่สุดสามัญ ซึ่งมีข้อสมมติ ดังนี้

$$\mathrm{E}(\varepsilon) = 0 \text{ และ } \mathrm{Cov}(\varepsilon) = \mathrm{E}(\varepsilon\varepsilon') = \sigma^2\mathbf{U}$$

โดยที่ $\mathbf{U}$ แทน เมทริกซ์อัตตสหสัมพันธ์ของ ε_i มิติ $n \times n$ ที่มี rank = n และจะเรียกข้อสมมตินี้ว่า ข้อสมมติของวิธีกำลังสองน้อยที่สุดทั่วไป

วิธีกำลังสองน้อยที่สุดทั่วไปเป็นวิธีการที่นำเอาค่าสัมประสิทธิ์อัตตสหสัมพันธ์ของค่าคลาดเคลื่อน (Autocorrelation Coefficient: ρ) มาปรับข้อมูล ในกรณีที่ทราบค่า ρ ปกติจะได้จากข้อมูลที่มีมาก่อน (Prior Information) เรียกการประมาณค่าวิธีนี้ว่า วิธีกำลังสองน้อยที่สุดทั่วไป (Generalized Least Square: GLS) แต่โดยทั่วไปจะไม่ทราบค่า ρ จึงต้องมีการประมาณค่า ρ ขึ้นมาด้วยวิธีการที่เหมาะสม จะเรียกการประมาณค่าวิธีนี้ว่า วิธีตัวประมาณกำลังสองน้อยที่สุดทั่วไป (Estimated Generalized Least Square: EGLS) ซึ่งมีวิธีการประมาณค่า ρ หลายวิธี เช่น วิธี Generalized Difference วิธี Paris-Winsten วิธี Durbin-Watson วิธี Cochrance-Orcutt iterative วิธี Durbin’s Two-step วิธี Cochrance-Orcutt Two-step และวิธี Hildreth-Lu

จากปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีปัญหาอัตตสหสัมพันธ์ ล้วนเป็นปัญหาที่สำคัญทั้งสิ้นเมื่อต้องทำการพยากรณ์โดยใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ เนื่องจากปัญหาดังกล่าวจะส่งผลให้การประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญไม่มีประสิทธิภาพ และส่งผลให้ตัวประมาณไม่มีคุณสมบัติตัวประมาณค่าไม่เอนเอียงเชิงเส้นที่ดีที่สุด โดยเฉพาะการพยากรณ์ข้อมูลทางเศรษฐศาสตร์นั้น ตัวแปรอิสระที่ใช้ในการทำนายตัวแปรตามจะมีหลายตัว จึงอาจมีความสัมพันธ์ระหว่างตัวแปรอิสระเกิดขึ้นได้ ประกอบกับข้อมูลทางเศรษฐศาสตร์เป็นข้อมูลที่จัดเก็บตามช่วงเวลาเรียกว่าข้อมูลอนุกรมเวลา ทำให้ค่าสังเกตในอนุกรมเวลามีความสัมพันธ์กันได้ และอาจมีค่าสังเกตบางค่ามีค่าสูงหรือต่ำกว่าค่าสังเกตส่วนใหญ่ ในการพยากรณ์ข้อมูลทางเศรษฐศาสตร์จึงจำเป็นต้องตระหนักถึงวิธีการประมาณค่าที่เหมาะสมกับข้อมูลเป็นอย่างมาก ซึ่งในงานวิจัยที่ผ่านมาได้มีการศึกษาเพื่อหาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่สามารถแก้ปัญหาอย่างใดอย่างหนึ่งในกรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ

การมีข้อมูลที่มีค่าผิดปกติ และปัญหาการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน ส่วนในกรณีการเกิดปัญหา 2 อย่างร่วมกันนั้น ได้มีผู้ศึกษากรณีเกิดปัญหาการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนร่วมกับการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ และการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระร่วมกับการมีข้อมูลที่มีค่าผิดปกติเท่านั้น แต่ยังไม่มีผู้ศึกษากรณีการเกิดปัญหาทั้ง 3 อย่างพร้อม ๆ กัน ในการศึกษาครั้งนี้ ผู้วิจัยจึงต้องการจะพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเพื่อแก้ปัญหาทั้ง 3 อย่างพร้อมกัน คือ ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน ซึ่งจะนำไปสู่วิธีการประมาณค่าพารามิเตอร์แบบใหม่ที่สามารถลดขั้นตอนในการประมาณค่าสัมประสิทธิ์การถดถอยและได้ค่าประมาณสัมประสิทธิ์การถดถอยที่มีประสิทธิภาพ

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน
2. เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ภายใต้สถานการณ์ 162 สถานการณ์ ได้แก่ 1) ขนาดตัวอย่าง 3 ขนาด คือ 20, 50 และ 100 2) ระดับการมีพหุสัมพันธ์ 3 ระดับ คือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 3) ระดับการมีอัตตสหสัมพันธ์ 3 ระดับ คือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 4) ลักษณะของการเกิดค่าผิดปกติของข้อมูล 3 แบบ คือ ข้อมูลผิดปกติในตัวแปรอิสระ ข้อมูลผิดปกติในตัวแปรตาม และข้อมูลผิดปกติในตัวแปรอิสระและตัวแปรตาม 5) เปอร์เซนต์ในการมีค่าผิดปกติ 2 ระดับ คือ 1% และ 5% โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล
3. เพื่อพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

กรอบแนวคิดในการวิจัย

การพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เริ่มต้นจากการศึกษาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ปรากฏว่าค่าพารามิเตอร์ที่ได้จากวิธีประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญจะไม่มีประสิทธิภาพขาดความแม่นยำ เมื่อข้อมูลเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์กัน ข้อมูลมีค่าผิดปกติ และเกิดอัตตสหสัมพันธ์กันของค่าคลาดเคลื่อน Hoerl and Kennard (1970) ได้พัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีตัวแปรอิสระมีความสัมพันธ์กันโดยไม่ต้องตัดตัวแปรอิสระทิ้ง เรียกว่าวิธีการถดถอยริดจ์ ต่อมา Swindel (1976) ได้พัฒนาวิธีการถดถอยริดจ์ที่เสนอโดย Hoerl and Kennard เรียกว่าวิธีการถดถอยริดจ์ที่มีค่าเบื้องต้น ส่วนในการประมาณค่าพารามิเตอร์กรณีข้อมูลมีค่าผิดปกติ ได้มีนักสถิติพัฒนาวิธีการแก้ปัญหาขึ้นมาเรียกว่า การวิเคราะห์การถดถอย

ที่มีความแกร่ง เช่น Huber (1964) ได้เสนอ วิธีตัวประมาณเอ็ม ต่อมา Simpson and Montgomery (1996) ได้พัฒนาจากวิธีตัวประมาณเอ็มเป็นวิธีตัวประมาณจีเอ็ม และนำเอาวิธีตัวประมาณจีเอ็มรวมกับการประมาณโดยวิธีการถดถอยริดจ์ เรียกวิธีใหม่ว่า วิธีริดจ์ชนิดตัวประมาณจีเอ็ม เพื่อแก้ปัญหากรณีมีความสัมพันธ์ระหว่างตัวแปรอิสระ และข้อมูลที่มีค่าผิดปกติเกิดขึ้นพร้อมกัน ส่วนกรณีที่ค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์กัน การประมาณค่าพารามิเตอร์สามารถทำด้วยวิธีกำลังสองน้อยที่สุดทั่วไปที่เสนอไว้โดย Aitken (1934) (มนตรี พิริยะกุล, 2545, หน้า 165) จากการศึกษาปรากฏว่าข้อมูลทางเศรษฐศาสตร์ในปัจจุบันนั้น ส่วนใหญ่แล้วตัวแปรอิสระจะมีความสัมพันธ์กัน และข้อมูลมีค่าผิดปกติซึ่งเกิดจากธรรมชาติของข้อมูล นอกจากนี้ยังเป็นข้อมูลอนุกรมเวลา จึงมีความเป็นไปได้สูงที่ค่าคลาดเคลื่อนจะมีความสัมพันธ์กัน ดังนั้นในงานวิจัยนี้จึงเป็นการแก้ปัญหาทั้งสามอย่างพร้อม ๆ กัน กรอบแนวคิดในการวิจัย จึงเป็นการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน ซึ่งแสดงในภาพที่ 1 และภาพที่ 2 ดังนี้

การประมาณค่าพารามิเตอร์ ด้วย
วิธีกำลังสองน้อยที่สุดทั่วไป

$$\hat{\mathbf{b}}_{GLS} = (\mathbf{X}'\mathbf{U}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{U}^{-1}\mathbf{Y})$$

วิธีการถดถอยริดจ์

$$\hat{\mathbf{b}}_{RR} = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\mathbf{Y}\right)$$

$$\hat{\mathbf{b}}_{RJ} = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\mathbf{Y} + k\mathbf{J}\right)$$

วิธีของนันทวัน จุลมุสิก

$$\hat{\mathbf{b}}_{NT} = \left(\mathbf{H}'\mathbf{U}^{-1}\mathbf{H} + k\mathbf{I}\right)^{-1}\left(\mathbf{H}'\mathbf{U}^{-1}\mathbf{N} + k\mathbf{J}\right)$$

วิธีริดจ์ชนิดตัวประมาณจีเอ็ม

$$\hat{\mathbf{b}}_{GMR} = \left(\mathbf{X}'\mathbf{W}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\mathbf{W}\mathbf{Y}\right)$$

วิธีการประมาณค่าพารามิเตอร์
การถดถอยเชิงเส้นพหุคูณแบบใหม่

$$\hat{\mathbf{b}}_{NEW} = \left(\mathbf{Q}'\mathbf{W}\mathbf{Q} + k\mathbf{I}\right)^{-1}\left(\mathbf{Q}'\mathbf{W}\mathbf{R} + k\mathbf{J}\right)$$

การพัฒนาวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่
โดยอาศัยวิธีของ Generalized Least Square, Hoerl and Kennard
(1970), Swindel (1976), Simpson and Montgomery (1996)
และวิธีของนันทวัน จุลมุสิก (2552)

ภาพที่ 1 กรอบแนวคิดในการวิจัยเรื่องการพัฒนาวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่

พัฒนาวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

↓

เปรียบเทียบประสิทธิภาพระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธี คือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ภายใต้สถานการณ์ 162 สถานการณ์

↓

พยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

ภาพที่ 2 กรอบแนวทางของการวิจัยในภาพรวม

สมมติฐานของการวิจัย

1. วิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำกว่าวิธีประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ได้จากวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ภายใต้สถานการณ์ 162 สถานการณ์ ซึ่งเกิดจากเงื่อนไข 5 เงื่อนไข ดังนี้

1.1 ขนาดตัวอย่าง 3 ขนาด คือ ขนาดตัวอย่างเท่ากับ 20, 50 และ 100

1.2 ระดับการมีพหุสัมพันธ์ 3 ระดับ คือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9)

1.3 ระดับการมีอัตตสหสัมพันธ์ 3 ระดับคือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9)

1.4 ลักษณะของการเกิดค่าผิดปกติของข้อมูล 3 แบบ คือ ข้อมูลผิดปกติในตัวแปรอิสระ ข้อมูลผิดปกติในตัวแปรตาม และข้อมูลผิดปกติในตัวแปรอิสระและตัวแปรตาม

1.5 เปอร์เซนต์ในการมีค่าผิดปกติ 2 ระดับ คือ 1% และ 5%

2. การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้นมีค่าสัมประสิทธิ์การกำหนดสูงกว่าร้อยละ 70

ประโยชน์ที่ได้รับจากการวิจัย

1. ได้วิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นมาใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

2. ได้เกณฑ์ในการตัดสินใจเลือกใช้วิธีประมาณค่าที่เหมาะสมในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

3. เป็นแนวทางในการนำไปประยุกต์กับข้อมูลจริงเมื่อเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์กัน ข้อมูลมีค่าผิดปกติ และค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับหนึ่งในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

ขอบเขตของการวิจัย

การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น ภายใต้ขอบเขตของกรณีการเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน โดยมีตัวแปรอิสระ 3 ตัว และมีขอบเขตของการวิจัยตามวัตถุประสงค์ของการวิจัย ดังนี้

ขอบเขตของการวิจัยสำหรับวัตถุประสงค์การวิจัยข้อที่ 1 และข้อที่ 2

1. ตัวแบบทั่วไป (General Model) ของความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามเป็นเชิงเส้นในรูปเมทริกซ์ คือ

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{nx1} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}_{nx(p+1)} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1)x1} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{nx1}$$

2. กำหนดให้ตัวแปรอิสระ X และค่าคลาดเคลื่อน ε_i มีการแจกแจงปกติ (Normal Distribution) ซึ่งฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function :pdf.) ของ X คือ

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \qquad ; -\infty < x < \infty$$

โดยตัวแปรอิสระและค่าคลาดเคลื่อน มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

3. จำนวนตัวแปรอิสระในตัวแบบมี 3 ตัว กำหนดให้ตัวแปร X_1 และ X_2 มีความสัมพันธ์กัน และตัวแปร X_3 เป็นอิสระจากตัวแปร X_1 และ X_2

4. ขนาดตัวอย่างที่ใช้ในการศึกษามี 3 ขนาด คือ ขนาดตัวอย่างเท่ากับ 20, 50 และ 100

5. ค่าพารามิเตอร์หรือค่าสัมประสิทธิ์การถดถอยที่ใช้ มีค่าดังนี้ $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1,$ $\beta_3 = 1$

6. ศึกษากรณีที่ข้อมูลมีค่าผิดปกติเกิดขึ้นในระดับรุนแรง โดยกำหนดลักษณะของการเกิดค่าผิดปกติของข้อมูล 3 แบบ คือ ข้อมูลผิดปกติในตัวแปรอิสระ กำหนดให้เป็นตัวแปร X_3 ข้อมูลผิดปกติในตัวแปรตาม (Y) และข้อมูลผิดปกติในตัวแปรอิสระ (X_3) และตัวแปรตาม (Y)

7. ระดับความสัมพันธ์ระหว่างตัวแปรอิสระแบ่งเป็น 3 ระดับ คือ ระดับต่ำมีค่าเท่ากับ 0.3 ระดับกลางมีค่าเท่ากับ 0.5 และระดับสูงมีค่าเท่ากับ 0.9

8. อัตตสหสัมพันธ์ของค่าคลาดเคลื่อนแบ่งเป็น 3 ระดับ คือ ระดับต่ำมีค่าเท่ากับ 0.3 ระดับกลางมีค่าเท่ากับ 0.5 และระดับสูงมีค่าเท่ากับ 0.9

โดยระดับความสัมพันธ์ระหว่างตัวแปรอิสระและอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนในแต่ละชุด จะเป็นไปตามที่กำหนดไว้ดังต่อไปนี้ (0.3,0.3), (0.3,0.5), (0.3,0.9), (0.5,0.3), (0.5,0.5), (0.5,0.9), (0.9,0.3), (0.9,0.5) และ(0.9,0.9)

เมื่อค่าระดับความสัมพันธ์ตัวแรก หมายถึง ความสัมพันธ์ระหว่างตัวแปรอิสระ 2 ตัว คือ X_1 และ X_2 และค่าระดับความสัมพันธ์ตัวหลัง หมายถึงความสัมพันธ์ของค่าคลาดเคลื่อน

ในการวิจัยนี้จะศึกษาการมีพหุสัมพันธ์ของตัวแปรอิสระเฉพาะความสัมพันธ์ในทิศทางบวกเท่านั้น และศึกษาอัตตสหสัมพันธ์ของความคลาดเคลื่อนเฉพาะค่าบวก โดยจำลองข้อมูลให้ค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับหนึ่ง (AR(1))

9. กำหนดเปอร์เซนต์ในการมีค่าผิดปกติที่ใช้ในการศึกษา 2 ระดับ คือ 1% และ 5%

10. กำหนดจำนวนการทำซ้ำในแต่ละสถานการณ์เป็น 1,000 ครั้ง

11. การวิจัยนี้ใช้วิธีการจำลองให้มีสถานการณ์ตามที่กำหนด ด้วยโปรแกรม MATLAB เวอร์ชัน 8.7

12. การเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ว่าวิธีการใดเหมาะสมมากที่สุดจะใช้เกณฑ์ในการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณจากทุกวิธี โดยวิธีที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณน้อยที่สุด จะเป็นวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่มีประสิทธิภาพมากที่สุด

ขอบเขตของการวิจัยสำหรับวัตถุประสงค์การวิจัยข้อที่ 3

การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น เป็นการนำข้อมูลทุติยภูมิเกี่ยวกับยางพาราเก็บรวบรวมเป็นรายปี ตั้งแต่ปี พ.ศ. 2541 ถึง ปี พ.ศ. 2557 ได้แก่ ปริมาณการส่งออกยางพาราไทยไปประเทศมาเลเซีย ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย และราคาส่งออก

ยางพาราไทย โดยกำหนดให้ปัจจัยสำคัญในการศึกษาอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ได้แก่ ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ส่วนตัวแปรราคาได้ใช้ราคาส่งออกยางพาราของไทยไปประเทศมาเลเซีย และปริมาณการส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมา โดยยางพาราหรือยางธรรมชาติที่นำมาใช้ในการพยากรณ์ครั้งนี้ใช้ยางธรรมชาติชนิดน้ำยางข้น เนื่องจากประเทศมาเลเซียมีการนำเข้ายางธรรมชาติชนิดน้ำยางข้นในปริมาณมากที่สุดเมื่อเทียบกับชนิดอื่น ๆ ทุกปี โดยใช้ข้อมูลในช่วงปี พ.ศ. 2542 ถึงปี พ.ศ. 2556 ในการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น และพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยใช้ข้อมูล ปี พ.ศ. 2557

นิยามศัพท์เฉพาะ

พหุสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) หมายถึง เหตุการณ์ที่ตัวแปรอิสระมีความสัมพันธ์เชิงเส้นต่อกันในทิศทางบวก

ค่าผิดปกติ (Outliers) หมายถึง เหตุการณ์ที่ข้อมูลหรือค่าสังเกตมีค่าสูงหรือต่ำกว่าค่าสังเกตส่วนใหญ่โดยเกิดขึ้นในระดับรุนแรง

อัตตสหสัมพันธ์ (Autocorrelation) หมายถึง เหตุการณ์ที่ค่าความคลาดเคลื่อนมีความสัมพันธ์กันทางบวก และเป็นตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง

ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) หมายถึง ค่าที่ใช้วัดความมีประสิทธิภาพของตัวประมาณค่า ซึ่งคำนวณได้จากผลรวมของผลต่างของค่าจริงกับค่าประมาณยกกำลังสอง แล้วหารด้วยจำนวนข้อมูล

ยางธรรมชาติ (Natural Rubber) หมายถึง วัตถุดิบชั้นกลางซึ่งผ่านการแปรรูปชั้นต้นมาจากน้ำยางธรรมชาติ ซึ่งได้แก่ น้ำยางข้น ยางแผ่น ยางแท่ง ยางเครพ เป็นต้น

การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย (Forecasting the Demand for Thai Rubber Export) หมายถึง การทำนายปริมาณของยางพาราไทยชนิดน้ำยางข้นที่ประเทศมาเลเซียต้องการจะเสนอซื้อ ณ ระดับราคาต่าง ๆ ในเวลาหนึ่ง โดยมีอำนาจซื้อ (Purchasing power) สนับสนุน

อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย (The Demand for Thai Rubber Export) หมายถึง ปริมาณการส่งออกยางพาราไทยชนิดน้ำยางข้นไปประเทศมาเลเซีย

ผลิตภัณฑ์มวลรวมภายในประเทศ (Gross Domestic Product: GDP) หมายถึง มูลค่าตลาดของสินค้าและบริการขั้นสุดท้ายที่ผลิตในประเทศมาเลเซียในช่วงเวลาหนึ่ง ๆ โดยไม่คำนึงว่าผลผลิตนั้นจะผลิตขึ้นมาด้วยทรัพยากรของชาติใด

ราคาส่งออกยางพาราไทย (Exporting Price of Thai Rubber) หมายถึง ราคาส่งออกยางพาราที่แท้จริงชนิดน้ำยางข้นของไทยไปประเทศมาเลเซีย

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยนี้เป็นการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอย 5 วิธี คือวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ภายใต้สถานการณ์ 162 สถานการณ์ โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล และนำตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นไปพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ซึ่งผู้วิจัยได้ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้อง โดยนำเสนอผลการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้

ตอนที่ 1 หลักการ แนวคิด และทฤษฎีที่เกี่ยวข้อง

ตอนที่ 2 การจำลองสถานการณ์ด้วยเทคนิคมอนติคาร์โล

ตอนที่ 3 ความรู้เกี่ยวกับยางพารา

ตอนที่ 4 ความรู้เกี่ยวกับทฤษฎีอุปสงค์

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับอุปสงค์การส่งออกยางพาราของไทย

ตอนที่ 1 หลักการ แนวคิด และทฤษฎีที่เกี่ยวข้อง

1. วิธีการประมาณค่าพารามิเตอร์

1.1 วิธีกำลังสองน้อยที่สุดสามัญ (Ordinary Least Square Method)

ในปี ค.ศ. 1777-1855 คาร์ล เฟรดริก เกาส์ (Karl Friecrich Gauss) และปี ค.ศ. 1856-1922 อังเดร แอนดรีวิช (Andrei Andreevich Markov) ได้เสนอวิธีกำลังสองน้อยที่สุดสามัญ ในการประมาณค่าพารามิเตอร์โดยหลักการของวิธีนี้ คือ หาตัวประมาณของพารามิเตอร์ที่ทำให้ผลรวมค่าคลาดเคลื่อนกำลังสอง (Sum Square Errors: SSE) มีค่าน้อยที่สุด

การหาตัวประมาณโดยวิธีกำลังสองน้อยที่สุดสามัญ

จากตัวแบบการถดถอยเชิงเส้น $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

ให้ $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณของพารามิเตอร์ $\boldsymbol{\beta}$ ที่ทำให้ผลรวมค่าคลาดเคลื่อนกำลังสองมีค่าน้อยที่สุดได้โดยการหาอนุพันธ์ (Differentiate) เทียบกับ $\hat{\boldsymbol{\beta}}$ แล้วกำหนดให้เท่ากับ $\mathbf{0}$ นั่นคือ

$$SSE = \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}$$

$$= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

$$= (\mathbf{Y}' - \hat{\boldsymbol{\beta}}'\mathbf{X}')(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

$$= \mathbf{Y}'\mathbf{Y} - \mathbf{Y}'\mathbf{X}\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

$$= \mathbf{Y}'\mathbf{Y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{Y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

$$\frac{\partial SSE}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$$

$$-2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = 0$$

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$$

ดังนั้น $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

1.2 คุณสมบัติของตัวประมาณโดยวิธีกำลังสองน้อยที่สุดสามัญ

1.2.1 ความไม่เอนเอียง (Unbiasedness)

ถ้า $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณของพารามิเตอร์ $\boldsymbol{\beta}$ แล้ว $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณที่ไม่เอนเอียงของ $\boldsymbol{\beta}$ ก็ต่อเมื่อ $\mathrm{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$

พิสูจน์ จาก $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

$$\mathrm{E}(\hat{\boldsymbol{\beta}}) = \mathrm{E}(\boldsymbol{\beta}) + \mathrm{E}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{E}(\boldsymbol{\varepsilon})$$

$= \boldsymbol{\beta}$;จาก $\mathrm{E}(\boldsymbol{\varepsilon}) = 0$

ดังนั้น $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณที่ไม่เอนเอียงของ $\boldsymbol{\beta}$

1.2.2 ความแปรปรวนของ $\hat{\boldsymbol{\beta}} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

พิสูจน์ $\mathrm{V}(\hat{\boldsymbol{\beta}}) = \mathrm{E}(\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}}))(\hat{\boldsymbol{\beta}} - \mathrm{E}(\hat{\boldsymbol{\beta}}))'$

$$= \mathrm{E}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'$$

$$= \mathrm{E}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} - \boldsymbol{\beta})((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} - \boldsymbol{\beta})'$$

$$= \mathrm{E}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) - \boldsymbol{\beta})((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) - \boldsymbol{\beta})'$$

$$= \mathrm{E}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon})((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon})'$$

$$= \mathrm{E}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon})(\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1})$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

$$= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \qquad ;\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2$$

ดังนั้น $\mathrm{V}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

1.3 วิธีกำลังสองน้อยที่สุดทั่วไป (Generalized Least Square Method)

จากรูปแบบ $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ มีข้อสมมติว่า $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\mathbf{I_n}$ ซึ่งแสดงว่าไม่มีสหสัมพันธ์ในตัวเองและความแปรปรวนคงที่ แต่ถ้าทดสอบแล้วพบว่า ε ไม่เป็นไปตามข้อสมมติคือ $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\boldsymbol{\Omega}$ โดย $\boldsymbol{\Omega} \neq \mathbf{I_n}$ เมื่อ $\boldsymbol{\Omega}$ เป็นเมทริกซ์ขนาด n x n ที่มีคุณสมบัติเป็นเมทริกซ์สมมาตรที่เป็นบวกแน่นอน (Symmetric Positive-definite Matrix) และ σ^2 เป็นพารามิเตอร์ที่ไม่ทราบค่า ในกรณีที่มีข้อตกลงว่า $\boldsymbol{\varepsilon}$ มีการแจกแจงแบบปกติมีค่าเฉลี่ยเป็น 0 เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมเป็น $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2\boldsymbol{\Omega}$ นั่นคือ $\boldsymbol{\varepsilon} \sim \mathrm{N}(0, \sigma^2\boldsymbol{\Omega})$ แล้ว สามารถใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method) จะได้

$$\mathrm{L} = \frac{1}{(2\pi)^{\frac{n}{2}}\left|\boldsymbol{\Omega}\right|^{\frac{1}{2}}\sigma^n} e^{(-\frac{1}{2\sigma^2}\boldsymbol{\varepsilon}'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon})}$$

$$\begin{aligned} 1\mathrm{nL} &= -\frac{n}{2}\mathrm{In}2\pi - \frac{1}{2}1\mathrm{n}\left|\boldsymbol{\Omega}\right| - n1\mathrm{n}\sigma - \frac{1}{2\sigma^2}\boldsymbol{\varepsilon}'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon} \\ &= -\frac{n}{2}\mathrm{In}2\pi - \frac{1}{2}1\mathrm{n}\left|\boldsymbol{\Omega}\right| - n1\mathrm{n}\sigma - \frac{1}{2\sigma^2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Omega}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned}$$

ในที่นี้ต้องการหาค่า $\hat{\boldsymbol{\beta}}$ และ σ^2 เพื่อเป็นตัวประมาณค่าที่ทำให้ $\ln \mathrm{L}$ มีค่ามากที่สุดเท่าที่จะเป็นไปได้ ในกรณีนี้ต้องการหาค่า $\hat{\boldsymbol{\beta}}$ เมื่อ σ^2 คงที่ แล้วทำให้ $(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\boldsymbol{\Omega}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$ มีค่าต่ำที่สุด ถ้า $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณของ $\boldsymbol{\beta}$ จะได้ว่าต้องการหาค่าต่ำสุดของ $\mathrm{S}(\hat{\boldsymbol{\beta}})$

$$\begin{aligned} \mathrm{S}(\hat{\boldsymbol{\beta}}) &= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'\boldsymbol{\Omega}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{Y}'\boldsymbol{\Omega}^{-1}\mathbf{Y} - \mathbf{2}\hat{\boldsymbol{\beta}}'\mathbf{X}\boldsymbol{\Omega}^{-1}\mathbf{Y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned}$$

$$\frac{\partial S(\hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}} = -\mathbf{2X'\Omega^{-1}Y} + \mathbf{2X'\Omega^{-1}X\hat{\boldsymbol{\beta}}}$$

$$-\mathbf{2X'\Omega^{-1}Y} + \mathbf{2X'\Omega^{-1}X\hat{\boldsymbol{\beta}}} = 0$$

$$2\mathbf{X'\Omega^{-1}X\hat{\boldsymbol{\beta}}} = 2\mathbf{X'\Omega^{-1}Y}$$

จะได้ $\hat{\boldsymbol{\beta}} = (\mathbf{X'\Omega^{-1}X})^{-1}\mathbf{X'\Omega^{-1}Y}$ เป็นตัวประมาณโดยวิธีกำลังสองน้อยที่สุดทั่วไป (Generalized Least Square Estimator) หรือตัวประมาณแบบ GLS ซึ่งถ้า $\mathbf{\Omega}$ มีลักษณะเป็นเมทริกซ์เอกลักษณ์ (Identity Matrix) แล้ว $\hat{\boldsymbol{\beta}}$ จะเป็นตัวประมาณโดยวิธีกำลังสองน้อยที่สุดสามัญนั่นเอง

เนื่องจาก $\mathbf{\Omega}$ และ $\mathbf{\Omega}^{-1}$ มีคุณสมบัติเป็นเมทริกซ์สมมาตรที่เป็นบวกแน่นอนแล้ว จึงสามารถเขียน $\mathbf{\Omega}^{-1}$ อยู่ในรูป $\mathbf{\Omega}^{-1} = \mathbf{PP'}$ เมื่อ $\mathbf{P}$ เป็นเมทริกซ์ไม่เอกฐาน (Nonsingular Matrix)

จาก $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ เมื่อเอา $\mathbf{P}$ คูณข้างหน้าตลอดจะได้

$$\mathbf{PY} = (\mathbf{PX})\boldsymbol{\beta} + \mathbf{P}\boldsymbol{\varepsilon}$$

หรือ $\mathbf{Y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$

เมื่อ $\mathbf{Y}^* = \mathbf{PY},\ \mathbf{X}^* = \mathbf{PX},\ \boldsymbol{\varepsilon}^* = \mathbf{P}\boldsymbol{\varepsilon}$

และ $E(\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}) = E(\mathbf{P}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{P}') = \sigma^2\mathbf{P\Omega P'}$ แต่จาก $\mathbf{\Omega}^{-1} = \mathbf{PP'}$ จะได้ $\mathbf{\Omega} = \mathbf{P}^{-1}(\mathbf{P}')^{-1}$ และจะได้ $\mathbf{P\Omega P'} = \mathbf{I_n}$ ดังนั้น $E(\boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon}^{*\prime}) = \sigma^2\mathbf{I_n}$ และ $E(\boldsymbol{\varepsilon}^*) = E(\mathbf{P}\boldsymbol{\varepsilon}) = 0$ ซึ่งเป็นเงื่อนไขที่สอดคล้องกับข้อสมมติ ดังนั้นสามารถดำเนินการโดยวิธีกำลังสองน้อยที่สุดสามัญได้ คือ

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\mathbf{Y}^*$$

โดยวิธีนี้จะเป็นการเปลี่ยนแปลงข้อมูลจาก $\mathbf{Y}$ เป็น $\mathbf{PY}$ และ $\mathbf{X}$ เป็น $\mathbf{PX}$ แล้วดำเนินการโดยวิธีกำลังสองน้อยที่สุดสามัญได้ แต่โดยปกติการที่จะทราบ $\mathbf{P}$ นั้นต้องทราบ $\mathbf{\Omega}$ ซึ่งเป็นเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของความคลาดเคลื่อนก่อน

1.4 วิธีการถดถอยริดจ์ (Ridge Regression Method)

ปี ค.ศ. 1970 Hoerl and Kennard ได้เสนอวิธีการถดถอยริดจ์ในการแก้ปัญหาตัวแปรอิสระมีความสัมพันธ์เชิงเส้นต่อกัน วิธีนี้จะให้ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณ $\hat{\boldsymbol{\beta}}$ ลดลง เรียกว่าวิธีการถดถอยริดจ์ มีหลักการ คือ ถ้าตัวแปรอิสระมีพหุสัมพันธ์เชิงเส้นกันแล้ว ค่าดีเทอร์มิแนนต์ (Determinant) ของ $\mathbf{X'X}$ จะมีค่าเข้าใกล้ศูนย์ จึงมีผลทำให้ความแปรปรวนของตัวประมาณของสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสอง

น้อยที่สุดสามัญมีค่าสูงผิดปกติด้วย ดังนั้นจึงทำให้ $|\mathbf{X'X}|$ เพิ่มขึ้นโดยการบวกค่าคงที่ k ที่มากกว่า 0 เข้ากับสมาชิกทุกตัวบนเส้นทแยงมุมของเมทริกซ์ $\mathbf{X'X}$ ซึ่งจะทำให้ค่าเจาะจงมีค่ามากขึ้น และส่งผลให้ผลบวกกำลังสองของ $\hat{\boldsymbol{\beta}}$ มีค่าลดลง ทำให้ตัวประมาณค่าของสัมประสิทธิ์การถดถอยถูกต้องมากขึ้น

ตัวประมาณค่าของ $\boldsymbol{\beta}$ ซึ่งเป็นตัวประมาณที่เอนเอียง คือ

$$\hat{\boldsymbol{\beta}}_{\mathrm{R}} = \left(\mathbf{X'X}+k\mathbf{I}\right)^{-1}\mathbf{X'Y}\ ;k>0$$

ในระยะแรกไม่สามารถกำหนดค่า k ที่แน่นอนได้ จึงทดลองกำหนดค่า k ให้มีค่าระหว่าง 0 ถึง 1 จนได้ค่า k ที่เหมาะสมที่ทำให้ $MSE(\hat{\boldsymbol{\beta}}_{\mathrm{R}})$ มีค่าน้อยกว่า $MSE(\hat{\boldsymbol{\beta}}_{\mathrm{OLS}})$ ในปี ค.ศ. 1975 Hoerl, Kenard, and Baldwin ได้มีการพัฒนาการหาค่า k ให้เหมาะสมยิ่งขึ้น ซึ่งค่า k ที่เป็นที่นิยมใช้และได้ผลดีวิธีหนึ่ง คือ

$$k = \frac{p\hat{\sigma}^2}{\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}}$$

โดยที่ p คือ จำนวนตัวแปรอิสระในตัวแบบ

$\hat{\boldsymbol{\beta}}$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ $\boldsymbol{\beta}$

$\hat{\sigma}^2$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ σ^2

เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของ $\hat{\boldsymbol{\beta}}_{\mathbf{R}}$ แสดงได้ดังนี้

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}_R) = \sigma^2\left(\mathbf{X'X}+k\mathbf{I}\right)^{-1}\mathbf{X'X}\left(\mathbf{X'X}+k\mathbf{I}\right)^{-1}$$

และตัวประมาณของเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของ $\hat{\boldsymbol{\beta}}_R$ คือ

$$\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}}_{\mathbf{R}}) = \hat{\sigma}^2\left(\mathbf{X'X}+k\mathbf{I}\right)^{-1}\mathbf{X'X}\left(\mathbf{X'X}+k\mathbf{I}\right)^{-1}$$

โดยที่ $\hat{\sigma}^2 = \dfrac{SSE}{\mathrm{n - 2}}$ เมื่อ SSE คือผลรวมค่าคลาดเคลื่อนกำลังสอง (Sum Square Error)

ต่อมาในปี ค.ศ. 1976 Swindel ได้เสนอเทคนิคการนำสารสนเทศเบื้องต้น (Prior Information) มาใช้ร่วมกับการถดถอยริดจ์ เรียกว่าการถดถอยริดจ์ที่มีค่าเบื้องต้น ทำให้ได้ตัวประมาณของสัมประสิทธิ์การถดถอย ดังนี้

$$\hat{\boldsymbol{\beta}}_{RJ} = \left(\mathbf{X'X}+k\mathbf{I}\right)^{-1}\left(\mathbf{X'Y}+k\mathbf{J}\right)$$

โดยที่ $\mathbf{J}$ คือเวกเตอร์ของค่าประมาณเบื้องต้นของ $\boldsymbol{\beta}$

โดย Swindel แสดงให้เห็นว่า มีค่าคงที่ k ที่ทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ที่มีค่าเบื้องต้นมีค่าน้อยกว่าวิธีกำลังสองน้อยที่สุดสามัญ แต่ Swindel ไม่ได้เสนอแนวทางในการประมาณค่า k

จนกระทั่ง ในปี ค.ศ. 1995 Crouse, Jin and Hanumara ได้พัฒนาวิธีการของ Swindel คือสามารถหาค่าประมาณ k ที่แน่นอน ที่ทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ที่มีค่าเบื้องต้นมีค่าน้อยกว่าวิธีกำลังสองน้อยที่สุดสามัญ และเป็นตัวประมาณค่าสัมประสิทธิ์การถดถอยมีคุณสมบัติไม่เอนเอียง และแนะนำให้ใช้สารสนเทศเบื้องต้นจาก $\mathbf{J} = \bar{\mathbf{J}}$ ซึ่งทำให้ค่าเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของค่าสัมประสิทธิ์การถดถอยมีเครื่องหมายที่ถูกต้อง โดยกำหนดให้ $\bar{\mathbf{J}}$ มีค่าเป็น

$$\bar{\mathbf{J}} = \sum_{j=1}^{p} \frac{\hat{\beta}_j}{p}$$

โดยที่ p คือ จำนวนตัวแปรอิสระในตัวแบบ

$\hat{\beta}_j$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ β_j

ค่าประมาณ k มีค่าเป็น

$$k = \begin{cases} \dfrac{p\hat{\sigma}^2}{(\hat{\boldsymbol{\beta}}-\mathbf{J})'(\hat{\boldsymbol{\beta}}-\mathbf{J}) - \hat{\sigma}^2 tr(\mathbf{X}'\mathbf{X})^{-1}} & ;(\hat{\boldsymbol{\beta}}-\mathbf{J})'(\hat{\boldsymbol{\beta}}-\mathbf{J}) - \hat{\sigma}^2 tr(\mathbf{X}'\mathbf{X})^{-1} > 0 \\[2ex] \dfrac{p\hat{\sigma}^2}{(\hat{\boldsymbol{\beta}}-\mathbf{J})'(\hat{\boldsymbol{\beta}}-\mathbf{J})} & ;\text{elsewhere} \end{cases}$$

โดยที่ $\hat{\sigma}^2$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ σ^2

ในงานวิจัยนี้ใช้วิธีการถดถอยริดจ์ โดยคำนวณค่าประมาณค่า k ตามหลักการของ Hoerl, Kenard, and Baldwin (HKB)

1.5 วิธีริดจ์ชนิดตัวประมาณจีเอ็ม (Generalized M-Estimator Ridge Method)

ในปี ค.ศ. 1964 Huber ได้ศึกษาถึงตัวประมาณที่มีความแกร่ง (Robust Estimator) ที่เรียกว่าวิธีตัวประมาณเอ็มเป็นครั้งแรก เพื่อประมาณค่าสัมประสิทธิ์การถดถอยสำหรับกรณีข้อมูลที่มีค่าผิดปกติในตัวแปรตามหรือกรณีที่ค่าคลาดเคลื่อนมีค่าผิดปกติ แต่ถ้าข้อมูลที่นำมาวิเคราะห์มีค่าผิดปกติในตัวแปรอิสระและตัวแปรตามพร้อม ๆ กัน วิธีตัวประมาณเอ็มอาจจะให้ค่าประมาณสัมประสิทธิ์การถดถอยที่ไม่เหมาะสม ต่อมาในปี ค.ศ. 1996 Simpson and Montgomery ได้พัฒนาวิธีตัวประมาณเอ็มเป็นวิธีตัวประมาณจีเอ็มเพื่อทำการสร้างตัวถ่วงน้ำหนักซึ่งจะพิจารณาทั้งค่าผิดปกติที่เกิดขึ้นในตัวแปรอิสระและตัวแปรตาม และได้นำวิธีตัวประมาณจีเอ็มมารวมกับการประมาณโดยวิธีการถดถอยริดจ์เรียกวิธีดังกล่าวว่า วิธีริดจ์ชนิดตัวประมาณจีเอ็ม (Generalized M-Estimator Ridge Method) โดยตัวประมาณจีเอ็มมาจากหลักการพื้นฐานของการประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) นั่นคือหาตัวประมาณ $\hat{\boldsymbol{\beta}}$ ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นของค่าคลาดเคลื่อนมีค่าสูงสุด ดังนี้

$$\max \sum_{i=1}^{n} \ln\, f(\varepsilon_i) = \max \sum_{i=1}^{n} \ln\, f(Y_i - X_i'\boldsymbol{\beta}) \qquad (1)$$

เมื่อ $\sum_{i=1}^{n} \ln\, f(Y_i - X_i'\boldsymbol{\beta})$ เป็นฟังก์ชันภาวะน่าจะเป็น

โดยที่ ε_i เป็น ค่าคลาดเคลื่อนของค่าสังเกตที่ i ; $i = 1,\ 2,\ ...,\ n$

Y_i เป็น ค่าสังเกตที่ i ของตัวแปรตาม ; $i = 1,\ 2,\ ...,\ n$

X_i เป็น ค่าสังเกตที่ i ของตัวแปรอิสระ ; $i = 1,\ 2,\ ...,\ n$

หรือ $X_i' = \left[1\ X_{i1}\ X_{i2}\ X_{i3}\ ...\ X_{ip}\right]$

p เป็น จำนวนตัวแปรอิสระ

$\boldsymbol{\beta}$ เป็น เมทริกซ์ของสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ

n เป็น ขนาดตัวอย่าง

จากสมการ (1) สามารถหาค่า $\hat{\boldsymbol{\beta}}$ ได้โดยการหาอนุพันธ์ (Differentiate) ของ $\sum_{i=1}^{n} \ln\, f(Y_i - X_i'\boldsymbol{\beta})$ เทียบกับ $\boldsymbol{\beta}$ แล้วกำหนดให้เท่ากับศูนย์

$$\sum_{i=1}^{n} \mathbf{X_i} \frac{f'(Y_i - X_i'\hat{\boldsymbol{\beta}})}{f(Y_i - X_i'\hat{\boldsymbol{\beta}})} = 0$$

เมื่อ $f'(z) = \dfrac{df(z)}{dz}$

จากหลักการของวิธีภาวะน่าจะเป็นสูงสุดจำเป็นต้องทราบรูปแบบของฟังก์ชันความหนาแน่นความน่าจะเป็นของค่าคลาดเคลื่อน คือ $f(\varepsilon_i)$ เมื่อ ε_i คือ $Y_i - X_i'\boldsymbol{\beta}$ ถ้า ε_i มีการแจกแจงแบบปกติแล้ว $\boldsymbol{\beta}$ จะมีรูปแบบเดียวกับตัวประมาณจากวิธีกำลังสองน้อยที่สุดสามัญ แต่ในกรณีที่ไม่ทราบรูปแบบการแจกแจงของค่าคลาดเคลื่อนจะทำให้ไม่สามารถหา $f'(\varepsilon_i) / f(\varepsilon_i)$ และ $\hat{\boldsymbol{\beta}}$ ได้ ดังนั้น Huber จึงได้กำหนดฟังก์ชัน ψ ซึ่งเป็นตัวแทนที่ดีในการอธิบายลักษณะการแจกแจงของค่าคลาดเคลื่อน ซึ่งการกำหนดฟังก์ชัน ψ จะมีผลต่อความแกร่งและประสิทธิภาพของตัวประมาณจีเอ็ม หลักการของตัวประมาณจีเอ็ม คือ หาตัวประมาณของ $\boldsymbol{\beta}$ ที่ทำให้ผลรวมของฟังก์ชันความผิดพลาดมีค่าน้อยที่สุด โดยมีรูปแบบ คือ

$$\min \sum_{i=1}^{n} \rho\left(\frac{Y_i - X_i'\hat{\boldsymbol{\beta}}}{\pi_i s}\right)\pi_i \qquad (2)$$

เมื่อ ρ เป็น ฟังก์ชันของความผิดพลาดที่ถูกเลือกให้เหมาะสม

s เป็น ตัวประมาณที่แกร่งของสเกล

π_i เป็น ค่าถ่วงน้ำหนักซึ่งสามารถคำนวณจาก

$$\pi_i = median|u_i| / u_i \; ; u_i = \left(\frac{e_i}{s} \right)$$

จากสมการ (2) หา $\hat{\boldsymbol{\beta}}$ ที่เหมาะสมโดยการหาอนุพันธ์บางส่วนของ $\sum_{i=1}^{n} \rho \left(\frac{Y_i - X_i'\boldsymbol{\beta}}{\pi_i s} \right) \pi_i$ เทียบกับ $\boldsymbol{\beta}$ แล้วกำหนดให้เท่ากับศูนย์ จะได้

$$\frac{\partial}{\partial \boldsymbol{\beta}} \sum_{i=1}^{n} \rho \left[\frac{Y_i - X_i'\boldsymbol{\beta}}{\pi_i s} \right] \pi_i = \sum_{i=1}^{n} \rho' \left[\frac{Y_i - X_i'\boldsymbol{\beta}}{\pi_i s} \right] \pi_i \frac{\partial}{\partial \boldsymbol{\beta}} \left[\frac{Y_i - X_i'\boldsymbol{\beta}}{\pi_i s} \right]$$

$$= \sum_{i=1}^{n} \pi_i X_i \psi \left(\frac{e_i}{\pi_i s} \right) = \mathbf{0}$$

$$= \sum_{i=1}^{n} \pi_i x_{ij} \psi \left(\frac{e_i}{\pi_i s} \right) = \mathbf{0} \; ; j = 0, 1, 2, ..., p \qquad (3)$$

เมื่อ $\psi(z) = \frac{\partial}{\partial z} \rho(z)$ และ $\psi(z) = z$ แล้ววิธีตัวประมาณจีเอ็มคือวิธีกำลังสองน้อยที่สุดสามัญ โดยที่ $z = \left(\frac{e_i}{\pi_i s} \right)$ และ x_{ij} แทนค่าสังเกตที่ i ของตัวแปรอิสระที่ j ซึ่งมีตัวแปรอิสระ p ตัว

จากสมการ (3) พบว่าฟังก์ชัน ψ ไม่เป็นเชิงเส้น (Nonlinear) ดังนั้นวิธีการแก้สมการนี้จะอาศัยเทคนิคการทำซ้ำเพื่อหาค่าที่เหมาะสม ซึ่งวิธีที่ใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยส่วนใหญ่นิยมใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักชนิดทำซ้ำ (Iteratively Reweighted Least Square: IRLS) ของ Beaton and Tukey (Beaton & Tukey, 1974, unpaged as cited Montgomery, Peck and Vining, 2001, p. 387) ซึ่งคำนวณวิธี IRLS จะต้องมีการกำหนดค่าเริ่มต้นของ $\hat{\boldsymbol{\beta}}$ และคำนวณ s แล้วหาน้ำหนักถ่วงที่เหมาะสมของแต่ละค่าสังเกตเพื่อหาตัวประมาณค่า $\boldsymbol{\beta}$ จากวิธีกำลังสองถ่วงน้ำหนักน้อยที่สุด (Weight Least Square: WLS)

พิจารณาจากสมการ (3) จะได้

$$\sum_{i=1}^{n} \pi_i x_{ij} \psi \left(\frac{e_i}{\pi_i s} \right) = \mathbf{0}$$

$$\sum_{i=1}^{n} \pi_i x_{ij} \frac{\psi \left(\frac{e_i}{\pi_i s} \right)}{\left(\frac{e_i}{\pi_i s} \right)} \left(\frac{e_i}{\pi_i s} \right) = \mathbf{0}$$

$$\sum_{i=1}^{n} \pi_i x_{ij} w_i \left(\frac{e_i}{\pi_i s} \right) = \mathbf{0}$$

เนื่องจาก $e_i = Y_i - X_{\mathbf{i}}'\hat{\boldsymbol{\beta}}$ จะได้

$$\sum_{i=1}^{n} \pi_i x_{ij} w_i \left(\frac{Y_i - X_i'\hat{\boldsymbol{\beta}}}{\pi_i s} \right) = \mathbf{0} \tag{4}$$

เมื่อ $w_i = \dfrac{\psi\left(\dfrac{e_i}{\pi_i s}\right)}{\left(\dfrac{e_i}{\pi_i s}\right)} = \dfrac{\psi(z_i)}{z_i}$

ดังนั้น จากสมการ (4) สามารถเขียนอยู่ในรูปเมทริกซ์ได้ดังนี้

$$\mathbf{X}'\mathbf{W}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{0}$$

$$(\mathbf{X}'\mathbf{W}\mathbf{Y}) - (\mathbf{X}'\mathbf{W}\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{0}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{Y})$$

ดังนั้น ตัวประมาณของ $\boldsymbol{\beta}$ จากวิธีกำลังสองถ่วงน้ำหนักน้อยที่สุดอยู่ในรูปของ

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\,\mathbf{W}\,\mathbf{X})^{-1}(\mathbf{X}'\,\mathbf{W}\,\mathbf{Y}) \tag{5}$$

เมื่อ $\mathbf{W}$ เป็น เมทริกซ์ทแยงมุม มีมิติ n x n

w_i เป็น สมาชิกแนวทแยงมุมของเมทริกซ์ $\mathbf{W}$

และการคำนวณตัวประมาณที่แกร่งของสเกล (s) ซึ่งเป็นตัวประมาณที่ไม่เอนเอียงของส่วนเบี่ยงเบนมาตรฐาน ในที่นี้ใช้ตัวประมาณส่วนเบี่ยงเบนสัมบูรณ์ของค่ามัธยฐาน (Median Absolute Deviation: MAD) ซึ่งเสนอโดย Mosteller and Huber (Mosteller & Huber, 1977, unpaged as cited Montgomery, Peck & Vining, 2001, p. 387) และถูกปรับด้วยค่าคงที่ 1.4826 ทำให้ s เป็นตัวประมาณที่ไม่เอนเอียงของ σ เมื่อ n มีขนาดใหญ่ และการแจกแจงของค่าคลาดเคลื่อนเป็นแบบปกติ ดังนั้นจะได้ s มีรูปแบบ ดังนี้

$$s = 1.4826\, median\left| e_i - median(e_i) \right| \qquad ;\ i = 1,\ 2,\ ...,\ n$$

Simpson and Montgomery (1996) ได้นำวิธีตัวประมาณจีเอ็มเข้ากับการประมาณโดยวิธีการถดถอยริดจ์ ซึ่งรูปแบบการคำนวณค่าสัมประสิทธิ์การถดถอยโดยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม คือ

$$\hat{\boldsymbol{\beta}}_{\mathrm{GMR}} = (\mathbf{X}'\,\mathbf{W}^{(\mathbf{r})}\,\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\,\mathbf{W}^{(\mathbf{r})}\,\mathbf{Y} \qquad ;k > 0 \tag{6}$$

เมื่อ $\mathbf{W}^{(\mathbf{r})}$ แทน น้ำหนักรอบสุดท้ายจากเกณฑ์ความแกร่งของตัวประมาณแบบจีเอ็ม

จากสมการ (6) จะต้องเลือกเกณฑ์ความแกร่งหรือรูปแบบฟังก์ชัน ψ ที่เหมาะสม เพื่อให้ได้สัมประสิทธิ์การถดถอยที่แกร่งและมีประสิทธิภาพ ซึ่งมีอยู่ด้วยกันหลายเกณฑ์ เช่น เกณฑ์ความแกร่งของ Huber’s Function เกณฑ์ความแกร่งของ Ramsay’s Function เกณฑ์ความแกร่งของ Andrews Function เกณฑ์ความแกร่งของ Hample’s 17A Function เกณฑ์ความแกร่งของ Barya Function และเกณฑ์ความแกร่งของ Talwar Function

จากการวิจัยของ อดุลย์เดช กรงทอง (2550) ได้ศึกษาการเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม สรุปได้ว่า ภายใต้กรณีที่ศึกษาวิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อประมาณค่าคงที่ k ด้วยวิธี Hoerl-Kennard-Baldwin (HKB) ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำกว่าวิธี Hoerl-Kennard (HK) เป็นส่วนใหญ่ และเมื่อขนาดตัวอย่างเพิ่มขึ้น การประมาณค่า k ทั้งสองวิธีให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณใกล้เคียงกัน นอกจากนี้เมื่อพิจารณาทุกระดับความรุนแรงของค่าผิดปกติ เปอร์เซนต์ค่าผิดปกติ ระดับของความสัมพันธ์และทุกลักษณะของตัวแปรที่เกิดค่าผิดปกติ พบว่าส่วนใหญ่วิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อใช้เกณฑ์ความแกร่งของ Ramsay จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด เมื่อขนาดตัวอย่างเท่ากับ 20 และวิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อใช้ความแกร่งของ Huber และ Ramsay ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณใกล้เคียงกันและมีค่าต่ำสุดเป็นส่วนใหญ่

ในที่นี้ผู้วิจัยใช้การประมาณค่า k ของวิธีริดจ์ชนิดตัวประมาณจีเอ็มด้วยวิธี Hoerl-Kennard-Baldwin (HKB) การคำนวณค่า s^2 และ k จะใช้ตัวประมาณ $\hat{\boldsymbol{\beta}}$ ที่ได้จากกำลังสองน้อยที่สุดสามัญและใช้เกณฑ์ความแกร่งของ Ramsay ซึ่งมีรูปแบบของฟังก์ชัน ρ ฟังก์ชัน ψ และฟังก์ชัน w ดังต่อไปนี้

$$\rho(z) = a^{-2}\left[1 - e^{-a|z|}\left(1 + a|z|\right)\right] \qquad ;|z| < \infty$$

เมื่อ $\psi(z) = \frac{\partial}{\partial z}\rho(z)$ แล้ว

$$\psi(z) = ze^{-a|z|} \qquad ;|z| < \infty$$

และ $w(z) = \frac{\psi(z)}{z}$ แล้วจะได้ว่า

$$w(z) = e^{-a|z|} \qquad ;|z| < \infty$$

เมื่อ z เป็นความคลาดเคลื่อนมาตรฐาน $\left(z = \frac{y_i - \mathbf{x}_i'\boldsymbol{\beta}}{\pi_i s} = \frac{e_i}{\pi_i s} \right)$

และ a แทนค่าคงที่ที่มีการปรับค่า (Turning Constant) ให้มีค่าเท่ากับ 0.3 ซึ่งกำหนดมาเพื่อให้ตัวประมาณที่ได้มีประสิทธิภาพตามที่ต้องการ เมื่อเปรียบเทียบกับตัวประมาณจากวิธีกำลังสองน้อยที่สุดสามัญ (Montgomery, Peck & Vining, 2001, p. 388)

ในการประมาณค่าด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ต้องหาตัวประมาณแบบจีเอ็ม เมื่อใช้เกณฑ์ความแกร่ง โดยอาศัยเทคนิคการทำซ้ำเพื่อหาค่าที่เหมาะสม ในที่นี้จะใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักชนิดทำซ้ำ (Iteratively Reweighted Least Square: IRLS) ซึ่งมีขั้นตอนการคำนวณดังนี้

1. คำนวณค่าเริ่มต้นของ $\hat{\beta}$ จากวิธีกำลังสองน้อยที่สุดสามัญ

$$\hat{\beta}^{(0)} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

2. เข้าสู่การทำซ้ำรอบที่ 1 เพื่อหาค่า $\mathrm{e}_\mathrm{i}^{(1)}, \mathrm{s}^{(1)}$ และ $\pi_\mathrm{i}^{(1)}$

$$e_i^{(1)} = Y_i - X_i'\hat{\boldsymbol{\beta}}^{(0)} \qquad ;\ i = 1,\ 2,\ ...,\ n$$

$$s^{(1)} = 1.4826\, median\left|e_i^{(1)} - median(e_i^{(1)})\right|$$

$$\pi_i^{(1)} = median\left|u_i^{(1)}\right| / u_i^{(1)} \qquad ;u_i^{(1)} = (\frac{e_i^{(1)}}{s^{(1)}})$$

3. ทำซ้ำรอบที่ 1 เพื่อหาน้ำหนักของแต่ละค่าสังเกตตามเงื่อนไขของเกณฑ์ความแกร่งหรือรูปแบบฟังชัน ψ ที่ถูกเลือกจากสมการที่ (3) ถ้าเลือกเกณฑ์ความแกร่งที่เหมาะสม จะทำให้ค่าสัมประสิทธิ์การถดถอยที่ได้มีความแกร่งและมีประสิทธิภาพ

4. การทำซ้ำรอบที่ 1 เพื่อหาค่า $\hat{\boldsymbol{\beta}}^{(1)}$ จาก $\hat{\boldsymbol{\beta}}^{(1)} = (\mathbf{X}'\mathbf{W}^{(1)}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{(1)}\mathbf{Y}$

โดยที่ $\mathbf{W}^{(1)}$ คือเมทริกซ์ทแยงมุมของน้ำหนักในรอบที่ 1 ขนาด n x n

$$\mathbf{W}^{(1)} = \begin{bmatrix} w_1^{(1)} & 0 & \cdots & 0 \\ 0 & w_2^{(1)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n^{(1)} \end{bmatrix}_{n \times n}$$

5. หาค่าสัมบูรณ์ของผลต่างระหว่างค่าประมาณของสัมประสิทธิ์การถดถอยเริ่มต้นกับค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ 1 ของค่าสัมประสิทธิ์การถดถอยทุกค่า

6. ถ้าค่าสัมบูรณ์จากข้อที่ 5 มีอย่างน้อยหนึ่งค่ามากกว่า 0.001 (ค่าคลาดเคลื่อนจากค่าคงที่ที่ยอมให้เกิดขึ้นได้ในการคำนวณวิธี IRLS เพื่อให้ได้ค่า $\boldsymbol{\beta}$ ค่อนข้างคงที่) ให้กลับไปทำซ้ำรอบที่

r ในข้อที่ 7 ต่อไป แต่ถ้าค่าสัมบูรณ์จากข้อที่ 5 ของสัมประสิทธิ์การถดถอยทุกค่าไม่มากกว่า 0.001 นั่นคือจะได้ค่าสัมประสิทธิ์การถดถอยจากรอบที่ 1

7. เข้าสู่การทำซ้ำรอบที่ r $\quad r = 2,\ 3, ...$

$$e_i^{(r)} = Y_i - X_i'\hat{\boldsymbol{\beta}}^{(r-1)} \qquad ;\ i = 1,\ 2,\ ...,\ n$$

$$s^{(r)} = 1.4826\, median\left| e_i^{(r)} - median(e_i^{(r)}) \right|$$

$$\pi_i^{(r)} = median\left| u_i^{(r)} \right| / u_i^{(r)} \qquad ;u_i^{(r)} = (\frac{e_i^{(r)}}{s^{(r)}})$$

ตามเงื่อนไขของเกณฑ์ความแกร่งหรือรูปแบบฟังก์ชัน ψ ที่ถูกเลือก

$$w_i^{(r)} = \frac{\psi(z)}{z}$$

เมื่อ $w_i^{(r)}$ คือน้ำหนักของค่าสังเกตตัวที่ i ในการทำซ้ำรอบที่ r

8. คำนวณการประมาณค่าสัมประสิทธิ์การถดถอยในการทำซ้ำรอบที่ r

$$\hat{\boldsymbol{\beta}}^{(r)} = (\mathbf{X}'\mathbf{W}^{(r)}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{(r)}\mathbf{Y}$$

9. คำนวณค่าสัมบูรณ์ของผลต่างระหว่างค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ $r-1$ กับค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ r ของสัมประสิทธิ์การถดถอยทุกค่า

ถ้าค่าสัมบูรณ์ของผลต่างระหว่างค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ $r-1$ กับค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ r ของสัมประสิทธิ์การถดถอยทุกค่า มีอย่างน้อยหนึ่งค่ามากกว่า 0.001 ให้กลับไปทำซ้ำในข้อที่ 7 จนถึงข้อที่ 9 ต่อไปเรื่อย ๆ จนกระทั่งค่าสัมบูรณ์ของผลต่างระหว่างค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ $r-1$ กับค่าประมาณของสัมประสิทธิ์การถดถอยรอบที่ r ของสัมประสิทธิ์การถดถอยทุกค่าไม่มากกว่า 0.001 นั่นคือจะได้ค่าสัมประสิทธิ์การถดถอยจากรอบที่ r

จากนั้นนำมาใช้กับสมการที่ Simpson และ Montgomery ได้เสนอไว้คือ นำวิธีตัวประมาณจีเอ็มรวมเข้ากับการประมาณโดยวิธีการถดถอยริดจ์ ซึ่งตัวประมาณสัมประสิทธิ์การถดถอยโดยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม คือ

$$\hat{\boldsymbol{\beta}}_{GMR} = (\mathbf{X}'\mathbf{W}^{(r)}\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{W}^{(r)}\mathbf{Y} \qquad ;k > 0$$

เมื่อ $\mathbf{W}^{(r)}$ คือ น้ำหนักรอบสุดท้ายที่ได้จากเกณฑ์ความแกร่งของตัวประมาณแบบจีเอ็ม

วิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณ มีวิธีที่ทำการประมาณค่าได้หลายวิธี เช่น วิธีกำลังสองน้อยที่สุดสามัญ เป็นวิธีที่สามารถทำการประมาณค่าพารามิเตอร์ได้ดีเมื่อค่าคลาดเคลื่อนมีการแจกแจงปกติ ค่าคลาดเคลื่อน $\varepsilon_i, \varepsilon_j$ เป็นอิสระต่อกัน ความแปรปรวนคงที่ และ

ตัวแปรอิสระเป็นอิสระต่อกัน แต่ถ้าค่าคลาดเคลื่อนไม่เป็นอิสระต่อกันควรใช้วิธีกำลังสองน้อยที่สุดทั่วไปหรือวิธีบูตสแตรป (Bootstrap) ถ้าตัวแปรอิสระมีความสัมพันธ์กันควรใช้วิธีการถดถอยริดจ์ และถ้าตัวแปรอิสระมีความสัมพันธ์กันและข้อมูลมีค่าผิดปกติควรใช้วิธีริดจ์ชนิดตัวประมาณจีเอ็ม

2 ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity)

ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณด้วยตัวประมาณกำลังสองน้อยที่สุดสามัญ เมื่อเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์กัน การเกิดปัญหาการมีพหุสัมพันธ์จะมากหรือน้อยขึ้นอยู่กับระดับความสัมพันธ์ระหว่างตัวแปรอิสระ ถ้าตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กับตัวแปรอิสระตัวอื่นต่ำ ผลที่ทำให้เกิดปัญหาการมีพหุสัมพันธ์อาจไม่รุนแรง แต่ถ้าตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กับตัวแปรอิสระตัวอื่นสูง จะส่งผลให้เกิดปัญหาการมีพหุสัมพันธ์รุนแรงด้วยเช่นกัน โดยที่ผลของการเกิดปัญหาการมีพหุสัมพันธ์มีดังนี้

2.1 ทำให้ค่าส่วนเบี่ยงเบนมาตรฐานของตัวประมาณสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ $(\mathrm{S_b})$ มีค่าสูงกว่าปกติ ซึ่งมีผลทำให้การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธ์การถดถอยของตัวแปรอิสระแต่ละตัวผิดพลาด (กัลยา วานิชบัญชา, 2546, หน้า 334-335) นั่นคือ

$$\mathrm{H_0}: \beta_j = 0 \qquad ; j = 1,\ 2,\ ...,\ p$$
$$\mathrm{H_1}: \beta_j \neq 0$$

สถิติทดสอบ $\mathrm{t} = \dfrac{b_j}{S_{b_j}}$

ซึ่งเป็นการทดสอบว่าตัวแปรอิสระ X_j มีอิทธิพลต่อตัวแปรตาม Y หรือไม่ ในขณะที่ตัวแปรอิสระตัวอื่น ๆ มีค่าคงที่ ดังนั้นเมื่อค่าสถิติ t มีค่าต่ำกว่าปกติเนื่องมาจากค่าส่วนเบี่ยงเบนมาตรฐานของตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณ (S_b) มีค่าสูงกว่าปกติ จะส่งผลให้ไม่สามารถปฏิเสธสมมติฐานหลัก $\mathrm{H_0}$ ได้ ทำให้สรุปได้ว่าเมื่อมีตัวแปรอิสระอื่น ๆ อยู่ในตัวแบบ ตัวแปรอิสระ X_j ไม่มีอิทธิพลต่อตัวแปรตาม Y ทั้งที่ความจริง X_j อาจจะมีอิทธิพลต่อตัวแปรตาม Y

2.2 เมื่อตัวแปรอิสระอยู่ในรูปของค่ามาตรฐาน (Standardize) แล้วสมาชิกของ $\mathbf{X'X}$ จะเป็นค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย และค่าดีเทอร์มิแนนท์ จะอยู่ในช่วง (0,1) ถ้า det $(\mathbf{X'X}) = 0$ แสดงว่ามีความสัมพันธ์กันอย่างสมบูรณ์ระหว่างตัวแปรอิสระอย่างน้อย 1 คู่ นั่นคือ คอลัมน์คู่ใดคู่หนึ่งไม่เป็นอิสระเชิงเส้น (Linearly Dependent) แต่ถ้า det$(\mathbf{X'X}) = 1$ หมายถึงตัวแปรอิสระทุกตัวเป็นอิสระต่อกันอย่างสมบูรณ์ (Orthogonal) (Draper & Smith, 1998, pp. 369-370) ดังนั้นจึงมีนักสถิติแนะนำให้พิจารณาขนาดของ det($\mathbf{X'X}$) ถ้ามีค่าใกล้ศูนย์ แสดงว่าตัวแปรอิสระมีความสัมพันธ์กันสูง ซึ่งมีผลทำให้ค่าความแปรปรวนของค่าประมาณสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณมีค่ามากผิดปกติ

2.3 มัลลิกา บุนนาค (2551, หน้า 320) เสนอว่าไม่สามารถประมาณค่าสัมประสิทธิ์การถดถอยได้ ถ้าตัวแปรอิสระมีความสัมพันธ์กันอย่างสมบูรณ์ เช่น ถ้าตัวแปรอิสระ X_1 และ X_2 มีความสัมพันธ์กันอย่างสมบูรณ์ นั่นคือ $\mathrm{r_{12}} = 1$

ให้ $x_j = X_{ij} - \bar{X}_j$ โดยที่ $j = 1,\ 2$ และ $i = 1,\ 2,\ ...,\ n$

จาก

$$r_{12} = \frac{\sum_{i=1}^{n}(X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2)/(n-1)}{\sqrt{\mathrm{V}(X_1)\mathrm{V}(X_2)}}$$

$$r_{12} = \frac{\sum_{i=1}^{n}(X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2)/(n-1)}{\sqrt{\left(\frac{\sum_{i=1}^{n}(X_{i1} - \bar{X}_1)^2}{n-1}\right)\left(\frac{\sum_{i=1}^{n}(X_{i2} - \bar{X}_2)^2}{n-1}\right)}}$$

$$r_{12} = \frac{\sum_{i=1}^{n} x_1 x_2}{\sqrt{\sum_{i=1}^{n} x_1^2 \sum_{i=1}^{n} x_2^2}}$$

$$r_{12}^2 = \frac{(\sum_{i=1}^{n} x_1 x_2)^2}{\sum_{i=1}^{n} x_1^2 \sum_{i=1}^{n} x_2^2} = 1$$

จะได้ว่า $(\sum_{i=1}^{n} x_1 x_2)^2 = \sum_{i=1}^{n} x_1^2 \sum_{i=1}^{n} x_2^2$

ตัวแบบการถดถอยเชิงเส้นพหุคูณ คือ

$$y_i = b_1 x_{i1} + b_2 x_{i2} + \varepsilon_i \qquad \text{โดยที่ } y_i = Y_i - \bar{Y}_i$$

โดยใช้วิธีกำลังสองน้อยที่สุดสามัญ จะได้สมการปกติดังนี้

$$\sum_{i=1}^{n} x_{i1} y = b_1 \sum_{i=1}^{n} x_{ij}^2 + b_2 \sum_{i=1}^{n} x_{i1} x_{i2}$$

$$\sum_{i=1}^{n} x_{i2} y = b_2 \sum_{i=1}^{n} x_{i1} x_{i2} + b_1 \sum_{i=1}^{n} x_{i2}^2$$

ซึ่งเมื่อแก้สมการแล้วจะได้ตัวประมาณการถดถอย b_1 และ b_2 ดังนี้

$$b_1 = \frac{\sum_{i=1}^{n} x_1 y \sum_{i=1}^{n} x_2^2 - \sum_{i=1}^{n} x_2 y \sum_{i=1}^{n} x_1 x_2}{\sum_{i=1}^{n} x_1^2 \sum_{i=1}^{n} x_2^2 - (\sum_{i=1}^{n} x_1 x_2)^2}$$

$$b_2 = \frac{\sum_{i=1}^{n} x_2 y \sum_{i=1}^{n} x_1^2 - \sum_{i=1}^{n} x_1 y \sum_{i=1}^{n} x_1 x_2}{\sum_{i=1}^{n} x_1^2 \sum_{i=1}^{n} x_2^2 - (\sum_{i=1}^{n} x_1 x_2)^2}$$

เมื่อ $(\sum_{i=1}^{n} x_1 x_2)^2 = \sum_{i=1}^{n} x_1^2 \sum_{i=1}^{n} x_2^2$ จะทำให้ตัวส่วนเป็น 0 ซึ่งทำให้ไม่สามารถหาค่า b_1 และ b_2 ได้

2.4 วิรัชช พานิชวงศ์ (2549, หน้า 165-167) ได้เสนอวิธีการตรวจสอบปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ดังนี้

2.4.1 ทดสอบสัมประสิทธิ์การถดถอยแต่ละตัว ถ้าตัวแปรอิสระตัวใดเมื่ออยู่ในตัวแบบเพียงตัวเดียวแล้วมีอิทธิลต่อตัวแปรตาม แต่เมื่อเพิ่มตัวแปรอิสระอื่นเข้ามาในตัวแบบแล้วตัวแปรอิสระตัวนั้นกลับไม่มีอิทธิพลต่อตัวแปรตาม Y อย่างมีนัยสำคัญ ให้สงสัยว่าเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ

2.4.2 ตรวจสอบค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายจากเมทริกซ์สหสัมพันธ์ โดยพิจารณาดูว่าตัวแปรอิสระคู่ใดบ้างที่มีค่าขนาดของสัมประสิทธิ์สหสัมพันธ์มากกว่า 0.8 ไม่ว่าจะเป็นบวกหรือลบ ให้สงสัยว่าตัวแปรอิสระคู่นั้นอาจมีความสัมพันธ์กัน

2.4.3 ตรวจสอบจากค่า VIF (Variance Inflation Factor) ค่า VIF สามารถคำนวณได้จากสูตร

$$\mathrm{VIF} = \frac{1}{1 - r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}}$$

เมื่อ $r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}$ คือ สัมประสิทธิ์การกำหนดของสมการถดถอยเชิงเส้นพหุคูณ X_j กับตัวแปรอิสระอื่น ๆ ที่เหลือ นั่นคือ

$$\hat{X}_j = b_0 + b_1 X_1 + \ldots + b_{j-1} X_{j-1} + b_{j+1} X_{j+1} + \ldots + b_k X_k$$

และสัมประสิทธิ์การกำหนดสามารถหาได้จากสูตร

$$r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)} = \frac{SST - SSE}{SST}$$

เมื่อ SST คือ ผลรวมกำลังสองทั้งหมด (Total Sum of Square)

และ SSE คือ ผลรวมค่าคลาดเคลื่อนกำลังสอง (Error Sum of Square)

เมื่อพิจารณาค่า $r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}$ คือ ค่าสัดส่วนความแปรผันทั้งหมดของ X_j ที่อธิบายโดยตัวแปรอิสระอื่น ๆ ที่เหลือ ค่า $r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}$ มีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าค่า $r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}$ มีค่าเป็นศูนย์หมายถึง X_j ไม่มีความสัมพันธ์กับตัวแปรอิสระที่อยู่ในสมการซึ่งจะทำให้ค่า VIF มีค่าเท่ากับ 1 แต่ถ้าค่าของ $r^2_{X_j(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)}$ มีค่าเข้าใกล้ 1 หมายถึง X_j มีความสัมพันธ์กับ

ตัวแปรอิสระอื่น ๆ ที่อยู่ในสมการซึ่งจะทำให้ค่าของ VIF มีค่าเข้าใกล้ ∞ และ Neter, Wasserman and Kctner (2005) ได้กล่าวว่า ถ้าตัวแปรอิสระตัวใดในตัวแบบสมการถดถอยมีค่า VIF สูง คือ มากกว่าหรือเท่ากับ 10 เราสามารถสรุปได้ว่า ตัวแปรอิสระนั้นมีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ นั่นคือ เกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ

วิธีการแก้ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระได้แก่ การเพิ่มค่าสังเกต การเพิ่มขนาดตัวอย่าง การแปลงตัวแปรอิสระบางตัว การตัดตัวแปรอิสระบางตัวที่เป็นสาเหตุของการเกิดปัญหาออก โดยพิจารณาว่าตัวแปรอิสระตัวใดมีความสัมพันธ์กับ Y น้อยที่สุดเมื่อมีตัวแปรอิสระตัวอื่น ๆ อยู่ในตัวแบบการถดถอยให้ตัดตัวแปรอิสระตัวนั้นออก หรือใช้วิธีการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีอื่น ๆ เช่นวิธีการวิเคราะห์ปัจจัย (Factor Analysis) การวิเคราะห์ส่วนประกอบหลัก (Principal Component Analysis) วิธีการถดถอยริดจ์ เป็นต้น

3. ปัญหาการมีข้อมูลที่มีค่าผิดปกติ (Outliers)

การวิเคราะห์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดสามัญเมื่อข้อมูลหรือค่าสังเกตที่ถูกเก็บรวบรวมเพื่อนำมาวิเคราะห์การถดถอยมีค่าสูงกว่าหรือต่ำกว่าค่าสังเกตส่วนใหญ่มาก หรือเป็นค่าสังเกตที่ไม่ได้มาจากประชากรเดียวกับค่าสังเกตส่วนใหญ่ ซึ่งลักษณะข้อมูลดังกล่าวถือว่าเป็นค่าผิดปกติ โดยสาเหตุของการเกิดค่าผิดปกติอาจเกิดจากความคลาดเคลื่อนของการวัด การบันทึกข้อมูลผิดพลาด การใช้เครื่องมือที่มีคุณภาพต่ำ รวมถึงประสบการณ์ของผู้ปฏิบัติ หรือเกิดจากเหตุการณ์ที่ผิดปกติซึ่งไม่สามารถคาดการณ์ล่วงหน้าได้ เช่น การจราจล ภาวะสงคราม หรือภัยธรรมชาติ เป็นต้น ลักษณะของข้อมูลที่มีค่าผิดปกติมีโอกาสเกิดขึ้นได้ทั้งกับข้อมูลของตัวแปรตามและข้อมูลของตัวแปรอิสระ และอาจส่งผลให้การประมาณค่าสัมประสิทธิ์การถดถอยโดยวิธีกำลังสองน้อยที่สุดสามัญไม่เหมาะสม เนื่องจากวิธีกำลังสองน้อยที่สุดสามัญมีความไว (Sensitive) ต่อค่าที่ผิดปกติ ทำให้สมการที่ได้ถูกปรับทิศทางไปตามค่าที่ผิดปกติ ส่งผลให้การประมาณมีประสิทธิภาพต่ำ (Rousseeuw & Leroy, 2003)

ดังนั้นข้อมูลผิดปกติจึงเป็นค่าสังเกตที่มีอิทธิพล (Influential Observation) ซึ่งหมายถึงข้อมูลผิดปกติที่มีผลทำให้สมการถดถอยเบี่ยงเบนไปจากกรณีปกติหรือกรณีที่ไม่มีข้อมูลที่มีอิทธิพล นั่นคือข้อมูลที่มีอิทธิพลจะมีผลทำให้ค่าสัมประสิทธิ์การถดถอยเปลี่ยนไป เนื่องจากสมการถดถอยที่ได้จะถูกปรับทิศทางไปตามข้อมูลที่มีอิทธิพล และอาจทำให้ค่าประมาณของความแปรปรวนของ ε_i มีค่าสูงขึ้นกว่าปกติ ดังนั้นหากทราบว่าเกิดข้อมูลผิดปกติเนื่องจากตัวแปรอิสระและ/หรือตัวแปรตาม ควรพิจารณาต่อไปว่าข้อมูลผิดปกตินั้นเป็นข้อมูลที่มีอิทธิพลหรือไม่ แต่ข้อมูลที่ผิดปกติไม่จำเป็นต้องเป็นข้อมูลที่มีอิทธิพลเสมอไป การตรวจสอบว่าค่าสังเกตใด ๆ จะเป็นข้อมูลที่มีอิทธิพลหรือไม่นั้นจะตรวจสอบโดยใช้ค่าสถิติ ได้แก่ Cook’s Distance D, DEFIT และ DFBETA ในกรณีที่ตรวจสอบพบว่ามีข้อมูลที่ผิดปกติและมีอิทธิพลต่อการประมาณค่าสัมประสิทธิ์การถดถอย ต้องพิจารณาหาสาเหตุของการเกิดค่าผิดปกติ เช่น ถ้ามาจากความคลาดเคลื่อนของการวัด การบันทึกข้อมูลผิดพลาด การใช้เครื่องมือที่มีคุณภาพต่ำ รวมถึงประสบการณ์ของผู้ปฏิบัติ อาจตัดข้อมูลผิดปกตินั้นออกไป หรือ

ทำการปรับแก้ค่าให้ถูกต้อง จากนั้นจึงนำข้อมูลที่ได้ปรับแก้ค่าให้ถูกต้องแล้วไปทำการวิเคราะห์การถดถอยต่อไป

ในกรณีที่ข้อมูลผิดปกติเหล่านี้เกิดจากธรรมชาติของข้อมูล และผู้วิเคราะห์ไม่สามารถอธิบายถึงสาเหตุของการเกิดค่าผิดปกติได้ ผู้วิเคราะห์ต้องนำค่าผิดปกตินั้นมาร่วมในการวิเคราะห์การถดถอยด้วย (ทรงศิริ แต้สมบัติ, 2548, หน้า 211) แต่ควรทำการแก้ไขก่อนนำข้อมูลมาวิเคราะห์ หรือใช้วิธีประมาณค่าสัมประสิทธิ์การถดถอยที่ให้ค่าประมาณที่มีความแกร่งต่อการมีข้อมูลผิดปกติ วิธีการดังกล่าวเรียกว่า การวิเคราะห์การถดถอยที่มีความแกร่ง (Robust Regression) ได้แก่ วิธีประมาณ LMS (Least Median Squares Method) วิธีประมาณ LTS (Least Trimmed Squares Method) วิธีประมาณ M (M-Estimator Method) และวิธีประมาณ MM (MM-Estimator Method) ในปี ค.ศ. 1996 Simpson และ Montgomery ได้พัฒนาวิธีประมาณค่าสัมประสิทธ์การถดถอยจากวิธีตัวประมาณเอ็มที่แก้ปัญหากรณีข้อมูลที่มีค่าผิดปกติในตัวแปรตามหรือกรณีที่ค่าคลาดเคลื่อนมีค่าผิดปกติมาเป็นวิธีตัวประมาณจีเอ็ม (Generalized M-Eestimator Method) เพื่อให้สามารถแก้ปัญหากรณีข้อมูลมีค่าผิดปกติในตัวแปรอิสระและตัวแปรตามพร้อม ๆ กัน แล้วรวมกับการประมาณโดยวิธีการถดถอยริดจ์เรียกว่า วิธีริดจ์ชนิดตัวประมาณจีเอ็ม (Generalized M-Estimator RidgeMethod: GMR) ในกรณีเกิดปัญหาการมีความสัมพันธ์ระหว่างตัวแปรอิสระและการมีค่าผิดปกติเกิดขึ้นพร้อมกัน

การวิเคราะห์การถดถอยที่มีความแกร่ง เป็นการวิเคราะห์การถดถอยที่ใช้หลักการลดอิทธิพลของข้อมูลที่มีค่าผิดปกติ คุณลักษณะที่สำคัญของการวิเคราะห์การถดถอยที่มีความแกร่ง มีดังนี้

1. วิธีการวิเคราะห์การถดถอยที่มีความแกร่ง จะให้ผลได้ดีพอ ๆ กับวิธีกำลังสองน้อยที่สุดสามัญ เมื่อข้อมูลมีคุณสมบัติครบตามข้อสมมติ โดยเฉพาะเมื่อค่าคลาดเคลื่อนมีการแจกแจงปกติและไม่มีค่าผิดปกติ
2. ในกรณีที่ข้อมูลขาดคุณสมบัติเกี่ยวกับการแจกแจงปกติ ซึ่งอาจทำให้เกิดค่าผิดปกติที่เป็นค่าสังเกตที่มีอิทธิพลซึ่งมีผลต่อสมการถดถอย วิธีการวิเคราะห์การถดถอยที่มีความแกร่งจะให้ผลที่ดีกว่าวิธีกำลังสองน้อยที่สุดสามัญ

การใช้การถดถอยที่มีความแกร่ง จะได้ตัวประมาณพารามิเตอร์ที่มีคุณสมบัติ ดังนี้

ค่าเบรกดาวน์ (Breakdown) แสดงถึงเปอร์เซนต์ของค่าผิดปกติที่ปนอยู่ในข้อมูลทั้งหมดแล้วทำให้ค่าประมาณสัมประสิทธิ์การถดถอยที่ประมาณขึ้นมายังคงสามารถใช้ในการพยากรณ์ได้ โดยค่าเบรกดาวน์นั้นมีค่าสูงสุดประมาณ 50% และมีค่าต่ำสุดเท่ากับ 0% วิธีการประมาณค่าพารามิเตอร์แบบแกร่งที่ให้ค่าเบรกดาวน์สูง ได้แก่ วิธีประมาณในกลุ่ม LMS, LTS และเอ็มเอ็ม ส่วนวิธีที่ให้ค่าเบรกดาวน์ต่ำ ได้แก่ วิธีกำลังสองน้อยที่สุดสามัญ และวิธีเอ็ม

ความมีประสิทธิภาพ (Efficiency) วิธีการถดถอยที่มีความแกร่งจะมีประสิทธิภาพได้ดีใกล้เคียงกับวิธีกำลังสองน้อยที่สุดสามัญเมื่อข้อมูลไม่มีค่าผิดปกติหรือค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ

ประสิทธิภาพสัมพัทธ์ (Relative Efficiency: RE) เป็นค่าที่แสดงถึงประสิทธิภาพของวิธีการวิเคราะห์การถดถอยที่มีความแกร่งมีความสามารถเทียบเท่าวิธีกำลังสองน้อยที่สุดสามัญ เมื่อข้อมูลมีคุณสมบัติครบตามข้อสมมติ โดยเฉพาะเมื่อค่าคลาดเคลื่อนมีการแจกแจงแบบปกติ โดยค่า RE มากกว่า 1 แสดงว่าวิธีในกลุ่มวิธีการถดถอยที่มีความแกร่งเป็นวิธีที่เหมาะสมสำหรับใช้ในการพยากรณ์

4. ปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน (Serial Correlation or Autocorrelation)

4.1 ความหมายและสาเหตุการเกิดปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ปัญหาอัตตสหสัมพันธ์หรือสหสัมพันธ์ในตนเอง (Autocorrelation or Serial Correlation) ของค่าคลาดเคลื่อน คือ เหตุการณ์ที่ตัวแปรสุ่ม ε มีความสัมพันธ์ต่อกัน กล่าวคือ $E(\varepsilon_i\varepsilon_j) \neq 0\ ; i \neq j$ ซึ่งเป็นเหตุการณ์ที่ขัดแย้งกับข้อสมมติสำหรับการวิเคราะห์การถดถอยที่กำหนดว่า $E(\varepsilon_i\varepsilon_j) = 0\ ; i \neq j$ การคำนวณหาค่าสหสัมพันธ์ในตัวเองสามารถคำนวณหาได้จากสมาชิกของ ε ครั้งละ 2 ตัวจะคำนวณได้ทั้งหมด ${}^{n}\mathrm{C}_2 = \frac{n(n-1)}{2}$ ครั้ง เมื่อคำนวณหา $E(\varepsilon_i\varepsilon_j)\ ; i \neq j$ ครบ $\frac{n(n-1)}{2}$ ครั้งจะได้ค่าสหสัมพันธ์ซึ่งสามารถเขียนเป็นเมทริกซ์ $\mathrm{V}(\varepsilon)$ ถ้าให้ $\mathrm{V}(\varepsilon) \neq \sigma^2\mathbf{I}_n$ ตามข้อสมมติเดิม กล่าวคือ ถ้า $E(\varepsilon_i\varepsilon_j) \neq 0\ ; i \neq j$ จะได้เมทริกซ์ $\mathrm{V}(\varepsilon)$ ดังนี้

$$\mathrm{V}(\varepsilon) = \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1n} \\ \rho_{12} & 1 & \rho_{13} & \cdots & \rho_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{1n} & \rho_{2n} & \rho_{3n} & \cdots & 1 \end{bmatrix} \neq \sigma^2\mathrm{I}_n$$

เมื่อ $\sigma^2\rho_{ij} = E(\varepsilon_i\varepsilon_j)\ ; i \neq j$ เมื่อ $i = 1,2,\ldots,n\ ; j = 1,2,\ldots,n$

เมื่อพิจารณา $E(\varepsilon_i\varepsilon_j) \neq 0\ ; i \neq j$ แล้วจำนวนพารามิเตอร์ในเมทริกซ์ $\mathrm{V}(\varepsilon)$ จะมีจำนวนมากถึง $\frac{n(n-1)}{2}$ ตัว และ σ^2 อีกหนึ่งตัว จึงทำให้มีพารามิเตอร์ที่ไม่ทราบค่ามีจำนวนเท่ากับ $\frac{n(n-1)}{2}+1$ ตัว ซึ่งเป็นปัญหาสำคัญและทำให้เกิดความยุ่งยากแก่นักวิจัยเป็นอย่างมาก เพราะทำให้ต้องประมาณค่าพารามิเตอร์เหล่านี้ทั้งหมด แต่อย่างไรก็ตามความยุ่งยากนี้จะลดลงได้ ด้วยการกำหนดรูปแบบของสหสัมพันธ์ในตัวเอง โดยให้เป็นกระบวนการถดถอยในตัวเองลำดับที่ p (Autoregressive Process of Order p: AR(p)) กระบวนการค่าเฉลี่ยเคลื่อนที่ลำดับที่ q (Moving-Average Process of Order q: MA(q)) หรือ การผสมระหว่าง AR กับ MA คือ ถดถอยในตัวเองและค่าเฉลี่ยเคลื่อนที่ลำดับที่ p และ q (Autoregressive Moving-Average Process of Order (p,q): ARMA(p,q)) ซึ่งจะทำให้จำนวนพารามิเตอร์ในเมทริกซ์ $\mathrm{V}(\varepsilon)$ ลดจำนวนจาก

$\frac{n(n-1)}{2}+1$ ตัว เหลือเพียง p ตัว หรือ q ตัว หรือ p+q ตัว ซึ่งขึ้นอยู่กับว่าใช้รูปแบบ AR(p) หรือ MA(q) หรือ ARMA(p,q) ตามลำดับ

สหสัมพันธ์ในตัวเองเกิดขึ้นจากสาเหตุต่าง ๆ ดังนี้

1. การไม่ได้พิจารณาตัวแปรอิสระที่สำคัญ

โดยปกตินักวิจัยต้องกำหนดตัวแปรอิสระเพื่อใช้ในการอธิบายตัวแปร Y ที่สนใจให้ครบถ้วนที่สุดเท่าที่จะทำได้ เพราะถ้าสามารถกำหนดตัวแปรอิสระได้ครบถ้วนมากเพียงใดแล้ว จะทำให้ทราบธรรมชาติตลอดจนพฤติกรรมในอดีต ปัจจุบัน และอนาคตของตัวแปร Y ได้มากเพียงนั้น แต่อย่างไรก็ตามอาจเกิดความบกพร่องขึ้นได้ เพราะการกำหนดตัวแปรให้ได้ครบถ้วนนั้นเป็นเรื่องยาก จึงอาจเป็นไปได้ที่ตัวแปรอิสระบางตัวหรือหลายตัวไม่ได้ถูกพิจารณา นอกจากนี้อาจตัดตัวแปรอิสระบางตัวทิ้งไปด้วยเหตุจำเป็น เพราะไม่อาจวัดค่าสังเกตได้ ตัวแปรที่ไม่ได้พิจารณานี้จึงแสดงบทบาทในการควบคุมความเคลื่อนไหวของ Y อยู่นอกสมการ $Y = f\left(X's\right)$ โดยส่งอิทธิพลผ่านตัวแปรสุ่ม ε ส่งผลให้ตัวแปรสุ่ม ε ขาดความเป็นอิสระขึ้นได้ สหสัมพันธ์ในตัวเองลักษณะนี้จึงมิใช่ธรรมชาติแท้จริงของ ε หรือมิใช่สหสัมพันธ์แท้ ๆ ของ ε แต่แสดงให้เห็นเสมือนว่า ε มีสหสัมพันธ์ที่ เรียกว่า Quasi-Autocorrelation

2. การกำหนดความสัมพันธ์ทางคณิตศาสตร์ที่ผิดพลาด

การระบุความสัมพันธ์ทางคณิตศาสตร์ระหว่างตัวแปรอิสระกับตัวแปรตามผิดจากสภาพความสัมพันธ์ที่แท้จริงจะมีผลให้ตัวแปรสุ่ม ε เกิดความสัมพันธ์กันขึ้นได้ เช่น ความสัมพันธ์ที่แท้จริงของสมการขาดความเป็นอิสระตามธรรมชาติเดิมและเกิดปัญหาสหสัมพันธ์ในตัวเองขึ้น เช่นถ้า $Y = f\left(X's\right)$ คือ ความสัมพันธ์ชนิดพหุนามกำลังสอง แต่นักวิจัยกำหนดให้เป็นสมการเชิงเส้น ค่าคลาดเคลื่อนจึงเกิดขึ้นและค่าคลาดเคลื่อนเหล่านี้จะถูกนำไปรวมกันไว้ในตัวแปรสุ่ม ε ส่งผลให้ตัวแปรสุ่ม ε ขาดความเป็นอิสระตามธรรมชาติเดิมและเกิดปัญหาสหสัมพันธ์ในตัวเองขึ้น

3. การระบุข้อตกลงว่า $\mathrm{E}(\varepsilon_i\varepsilon_j)=0$ ทั้ง ๆ ที่ $\mathrm{E}\left(\varepsilon_i\varepsilon_j\right)\neq 0\ ;i\neq j$

โดยปกติตัวแปรเป็นจำนวนมากมีธรรมชาติของความเกี่ยวเนื่องกันภายในตัวเองในระหว่างช่วงเวลา เช่น ราคาสินค้าปัจจุบันสัมพันธ์ทางบวกกับราคาในอดีต กล่าวคือ การกำหนดราคาสินค้าผู้กำหนดราคานิยมเอาราคาในอดีตเป็นฐาน ซึ่งการกระทำเช่นนี้มีผลให้ตัวแปรราคามีสหสัมพันธ์ในตัวเอง ขณะเดียวกันตัวแปรอิสระอาจมีความสัมพันธ์ภายในได้ด้วยเหตุอื่น ๆ หลายประการเช่น การนัดหยุดงานของพนักงานมีผลให้ปริมาณการผลิต ยอดขาย สภาพคล่อง และอื่น ๆ กระทบกระเทือนเป็นลูกโซ่หลายช่วงเวลา ด้วยเหตุที่ตัวแปรอิสระมีธรรมชาติของสหสัมพันธ์ในตัวเองอยู่แล้วตามธรรมชาติ การกำหนดให้ $\mathrm{E}\left(\varepsilon_i\varepsilon_j\right)\neq 0\ ;i\neq j$ โดยไม่ได้ตรวจสอบจะส่งผลต่อประสิทธิภาพและความแม่นยำในการวิเคราะห์ข้อมูล

4.2 ผลกระทบของสหสัมพันธ์ในตัวเองต่อวิธีกำลังสองน้อยที่สุดสามัญ

เมื่อเกิดปัญหาสหสัมพันธ์ในตัวเองขึ้น คือ $\mathrm{V}(\varepsilon) \neq \sigma_\varepsilon^2 \mathbf{I}_n$ แต่ยังคงใช้วิธีกำลังสองน้อยที่สุดสามัญในการประมาณค่า $\boldsymbol{\beta}$ จะส่งผลให้ตัวประมาณที่ได้ขาดความแม่นยำโดยเฉพาะในแง่ของคุณภาพของ $\hat{\boldsymbol{\beta}}$ ดังนี้

4.2.1 $\hat{\boldsymbol{\beta}}$ ยังเป็นตัวประมาณค่าที่ไม่มีความเอนเอียงของ $\boldsymbol{\beta}$ แต่โดยทั่วไปจะมีประสิทธิภาพต่ำกว่าวิธีอื่น ๆ นั่นคือ

$$\begin{aligned} \mathrm{E}(\hat{\boldsymbol{\beta}}) &= \mathrm{E}\left[(\mathbf{X'X})^{-1}\mathbf{X'Y}\right] \\ &= \mathrm{E}\left[(\mathbf{X'X})^{-1}\mathbf{X'}(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})\right] \\ &= \mathrm{IE}(\boldsymbol{\beta}) + (\mathbf{X'X})^{-1}\mathbf{X'}\mathrm{E}(\boldsymbol{\varepsilon}) \\ &= \boldsymbol{\beta} \qquad\qquad ;\text{จาก } \mathrm{E}(\boldsymbol{\varepsilon})=0 \end{aligned}$$

ดังนั้น $\mathrm{E}(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$

4.2.2 ถ้าเกิดปัญหาสหสัมพันธ์ในตัวเองขึ้นโดยเฉพาะอย่างยิ่งคืออัตตสหสัมพันธ์ทางบวก ค่าประมาณของ σ^2 ซึ่งประมาณขึ้นโดยอาศัยค่าคลาดเคลื่อน คือ $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n}\varepsilon_i^2}{n-k-1}$; $\mathrm{i}=1,2,...,\mathrm{n}$ จะมีค่าต่ำกว่าความเป็นจริง ทั้งนี้เพราะในกรณีเช่นนี้ค่าจริงของ $\boldsymbol{\varepsilon}$ จะจับกลุ่มกันห่างไกลจากสมการเส้นตรงจริง ขณะที่ค่าประมาณของ $\boldsymbol{\varepsilon}$ จะจับกลุ่มกันใกล้สมการเส้นตรงที่ประมาณ

4.2.3 เมื่อเกิดปัญหาสหสัมพันธ์ในตัวเองขึ้น ค่าประมาณของ $\mathrm{V}(\hat{\boldsymbol{\beta}})$ คือ $\hat{\mathrm{V}}(\hat{\boldsymbol{\beta}})$ จะต่ำกว่าความเป็นจริง โดยเฉพาะอย่างยิ่ง เมื่อประมาณค่า β โดยอาศัยวิธีกำลังสองน้อยที่สุดสามัญ กล่าวคือ $\hat{\mathrm{V}}(\hat{\boldsymbol{\beta}})=\hat{\sigma}^2(\mathbf{X'X})^{-1}$ เหตุที่เป็นเช่นนี้ เพราะ $\hat{\sigma}^2$ มีค่าต่ำกว่าความเป็นจริง และในกรณีของสหสัมพันธ์ในตัวเองนั้น $\mathrm{E}(\varepsilon_i\varepsilon_j) \neq 0\ ; i \neq j$ แสดงว่า $Cov\left(\varepsilon_i\varepsilon_j\right)$ จะต้องมีบทบาทในการกำหนดค่า $\hat{\mathrm{V}}(\hat{\boldsymbol{\beta}})$ มีค่าต่ำกว่าความเป็นจริง ค่า $\hat{\mathrm{V}}(\hat{\boldsymbol{\beta}})$ ที่ต่ำ ๆ ในกรณีนี้หากไม่ระมัดระวังและไม่มีความเข้าใจในเรื่องเหล่านี้ อาจเข้าใจผิดคิดว่างานของตนมีคุณภาพสูงทั้ง ๆ ที่ความจริงไม่ได้เป็นเช่นนั้น

4.2.4 เมื่อเกิดปัญหาสหสัมพันธ์ในตัวเองขึ้น ตัวประมาณโดยวิธีกำลังสองน้อยที่สุดสามัญจะมีประสิทธิภาพต่ำ เพราะ $\hat{\mathrm{V}}(\hat{\boldsymbol{\beta}})$ ตามวิธีกำลังสองน้อยที่สุดสามัญ จะมีค่าสูงกว่า $\hat{\mathrm{V}}(\hat{\boldsymbol{\beta}})$ ที่ประมาณได้โดยอาศัยวิธีอื่นที่เหมาะสมกว่า

4.3 การกำหนดรูปแบบของสหสัมพันธ์ในตัวเอง

ถ้า $\mathrm{E}(\varepsilon_i\varepsilon_j) \neq \sigma^2\rho_{ij}; i \neq j$ จะต้องมีการประมาณค่าพารามิเตอร์จำนวนมาก เพื่อลดการประมาณค่าพารามิเตอร์จะกำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรสุ่ม ε_i กับ ε_j เพื่อความสะดวกและในการเชื่อมโยงระหว่างตัวแปรจะใช้ตัวห้อย (Subscript) t แทน i หรือ j ดังนี้

4.3.1 รูปแบบ AR(p) กำหนดให้ ε_t มีความสัมพันธ์กันดังนี้

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\ \varepsilon_{t-2} + ... + \phi_p\ \varepsilon_{t-p} + a_t$$

4.3.2 รูปแบบ MA(q) กำหนดให้ ε_t มีความสัมพันธ์กันดังนี้

$$\varepsilon_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots + \theta_q a_{t-q}$$

4.3.3 รูปแบบ ARMA(p,q) กำหนดให้ ε_t มีความสัมพันธ์กันดังนี้

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \ldots + \phi_p\varepsilon_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots\theta_q a_{t-q}$$

โดยมีข้อตกลงว่า $\mathrm{E}(a_t) = 0;\ \mathrm{E}(a_t^2) = \sigma_a^2;\ \mathrm{E}(a_t a_s) = 0;\ t \neq s$ โดยที่ ϕ และ θ เป็นพารามิเตอร์ที่ไม่ทราบค่า ในงานวิจัยนี้ใช้รูปแบบสหสัมพันธ์ในตัวเอง AR(1)

4.4 รูปแบบสหสัมพันธ์ถดถอยในตัวเองลำดับที่ 1 หรือ AR(1)

สำหรับAR(t) จะกำหนดให้ ε_t มีความสัมพันธ์กันดังนี้

$\varepsilon_t = \phi\varepsilon_{t-1} + a_t$ โดยที่ $-1 < \phi < 1$

ซึ่งสามารถกำหนดโครงสร้างของ ε_t ให้เป็นฟังชันของ a_t โดยอาศัยการแทนที่ซ้ำ ๆ เนื่องจากอิทธิพลของ ε ที่มีต่อกันและกันในลักษณะผลกระทบแบบลูกโซ่ตามวิธีมาร์คอฟ (Markov) ดังนี้

จากสมการ $\varepsilon_t = \phi\varepsilon_{t-1} + a_t$ (*)

$$\varepsilon_{t-1} = \phi\varepsilon_{t-2} + a_{t-1} \quad (1)$$

$$\varepsilon_{t-2} = \phi\varepsilon_{t-3} + a_{t-2} \quad (2)$$

$$\varepsilon_{t-3} = \phi\varepsilon_{t-4} + a_{t-3} \quad (3)$$

$$\vdots$$

$$\varepsilon_{t-r} = \phi\varepsilon_{t-(r+1)} + a_{t-r} \quad (r)$$

ดังนั้น จากสมการ (*) แทนที่ ε_{t-1} ด้วยสมการ (1) จะได้

$$\varepsilon_t = \phi(\phi\varepsilon_{t-2} + a_{t-1}) + a_t = \phi^2\varepsilon_{t-2} + (\phi a_{t-1} + a_t)$$

แทนที่ ε_{t-2} ด้วยสมการ (2) จะได้

$$\varepsilon_t = \phi^2(\phi\varepsilon_{t-3} + a_{t-2}) + (\phi a_{t-1} + a_t)$$

$$\varepsilon_t = \phi^3\varepsilon_{t-3} + (\phi^2 a_{t-2} + \phi a_{t-1} + a_t)$$

แทนที่ ε_{t-3} ด้วยสมการ (3) จะได้

$$\varepsilon_t = \phi^3(\phi\varepsilon_{t-4} + a_{t-3}) + (\phi^2 a_{t-2} + \phi a_{t-1} + a_t)$$

$$\varepsilon_t = \phi^4\varepsilon_{t-4} + (\phi^3 a_{t-3} + \phi^2 a_{t-2} + \phi a_{t-1} + a_t)$$

แทนที่ ε_{t-r} ด้วยสมการ (r) จะได้

$$\varepsilon_t = \phi^{(r+1)}\varepsilon_{t-(r+1)} + (a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + ... + \phi^r a_{t-r})$$

เนื่องจาก $-1 < \phi < 1$ และ $r \to \infty$ จะมีผลทำให้ $\phi^r \to 0$

ดังนั้น $\varepsilon_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + ...$

$$= \sum_{r=0}^{\infty} \phi^r a_{t-r}$$

นั่นคือ โดยอาศัยกระบวนการมาร์คอฟ (Markov Process) ตัวแปรสุ่ม ε_t สามารถเสนอในรูปฟังก์ชันของค่าช่วงเวลาย้อนหลัง (Lagged Value) ของ a_t ได้ ซึ่งถ้าพิจารณาแล้ว ε_t จะได้รับอิทธิพลจากค่าคลาดเคลื่อน a_t ในช่วงเวลาปัจจุบันหรือใกล้ปัจจุบันมากที่สุดในทำนองเดียวกัน a_t ที่อยู่ห่างจากปัจจุบันมาก ๆ ก็จะยิ่งมีอิทธิพลลดลง และถ้า ϕ ยิ่งมีค่าน้อยมากเท่าไรแล้ว a_t ในอดีตย่อมมีอิทธิพลต่อ ε_t ลดลงมากขึ้นเท่านั้น

คุณลักษณะทางประชากรของ ε_t เมื่อเกิดสหสัมพันธ์ในตัวเองใน AR(1)

1. ค่าคาดหมายของ ε_t

$$\varepsilon_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + ...$$

$$\mathrm{E}(\varepsilon_t) = \mathrm{E}(a_t) + \phi\mathrm{E}(a_{t-1}) + \phi^2\mathrm{E}(a_{t-2}) + \phi^3\mathrm{E}(a_{t-3}) + ...$$

แต่ $\mathrm{E}(a_t) = 0\ ; t = 1,\ 2,\ ...,\ n$

ดังนั้น $\mathrm{E}(\varepsilon_t) = 0\ ; t = 1,\ 2,\ ...,\ n$

แสดงว่าแม้ว่าเกิดปัญหาสหสัมพันธ์ในตัวเองใน AR(1) ε_t ยังมีคุณสมบัติสอดคล้องกับข้อสมมุติเดิมที่ว่า $\mathrm{E}(\varepsilon_t) = 0$

2. ความแปรปรวนของ ε_t

$$\varepsilon_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + ...$$

$$= \sum_{r=0}^{\infty} \phi^r a_{t-r}$$

จะพบว่า $\mathrm{E}(\varepsilon_t^2) = \mathrm{E}\{\sum_{r=0}^{\infty} \phi^r a_{t-r}\}^2$

$$= \mathrm{E}\{a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \phi^3 a_{t-3} + ...\}^2$$

$$= \mathrm{E}[\{a_t^2 + \phi^2 a_{t-1}^2 + \phi^4 a_{t-2}^2 + ...\} + \{2\phi a_t a_{t-1} + 2\phi^2 a_t a_{t-2} + ...\}]$$

แต่ $\mathrm{E}(a_t^2) = \sigma_a^2$ และ $(a_t a_{t-s}) = 0\ ; s = 1, 2, 3, ...$

ดังนั้น $\mathrm{E}(\varepsilon_t^2) = (\sigma_a^2 + \phi^2\sigma_a^2 + \phi^4\sigma_a^2 + \phi^6\sigma_a^2 + ...) + (0 + 0 + 0 + ... + 0)$

$$= \sigma_a^2(1 + \phi^2 + \phi^4 + \phi^6 + ...)$$

จากอนุกรมเรขาคณิต ถ้า $\mathrm{n} \to \infty$ และ $|\phi| < 1$ จะได้

$$1+\phi^2+\phi^4+\phi^6+...=\frac{1}{(1-\phi^2)}$$

นั่นคือ $\mathrm{E}(\varepsilon_t^2)=\dfrac{\sigma_a^2}{(1-\phi^2)}$ และจะได้ $\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2=\dfrac{\sigma_a^2}{(1-\phi^2)}$

แสดงว่าแม้ว่าจะเกิดปัญหาสหสัมพันธ์ในตัวเองใน AR(1) แต่ไม่เกิดปัญหาความแปรปรวนไม่คงที่เพราะ $\mathrm{Var}(\varepsilon_t)$ มีค่าคงที่เท่ากับ $\dfrac{\sigma_a^2}{(1-\phi^2)}$ เสมอในทุกค่าของ t

3. ความแปรปรวนร่วม ของ ε_t

ในที่นี้จะพิสูจน์เฉพาะ $\mathrm{E}(\varepsilon_t\varepsilon_{t-1})$ และ $\mathrm{E}(\varepsilon_t\varepsilon_{t-2})$ แล้วขยายความสู่กรณีทั่วไป คือ $\mathrm{E}(\varepsilon_t\varepsilon_{t-s})$; $s=1,2,3,...,n$

3.1 $$\begin{aligned}\mathrm{E}(\varepsilon_t\varepsilon_{t-1}) &= \mathrm{E}[\{a_t+\phi a_{t-1}+\phi^2 a_{t-2}+\phi^3 a_{t-3}+...\}\\ &\quad \{a_{t-1}+\phi a_{t-2}+\phi^2 a_{t-3}+\phi^3 a_{t-4}+...\}]\\ &= \mathrm{E}[a_t(a_{t-1}+\phi a_{t-2}+\phi^2 a_{t-3}+...)+\phi(a_{t-1}+\phi a_{t-2}+\phi^2 a_{t-3}+...)^2]\\ &= (0+0+0+...+0)+\phi\mathrm{E}[a_{t-1}+\phi a_{t-2}+\phi^2 a_{t-3}+...]^2\\ &= \phi\mathrm{E}[(a_{t-1}^2+\phi^2 a_{t-2}^2+\phi^4 a_{t-3}^2+...)+2(\phi a_{t-1}a_{t-2}+\phi^2 a_{t-2}a_{t-3}+...)]\\ &= \phi[\sigma_a^2+\phi^2\sigma_a^2+\phi^4\sigma_a^2+...)+2(0+0+0+...+0)]\\ &= \phi\sigma_a^2(1+\phi^2+\phi^4+...)\\ &= \phi\frac{\sigma_a^2}{(1-\phi^2)} = \phi\sigma_\varepsilon^2 \neq 0\end{aligned}$$

3.2 $$\begin{aligned}\mathrm{E}[\varepsilon_t\varepsilon_{t-2}] &= \mathrm{E}[\{a_t+\phi a_{t-1}+\phi^2 a_{t-2}+\phi^3 a_{t-3}+...\}\\ &\quad \{a_{t-2}+\phi a_{t-3}+\phi^2 a_{t-4}+\phi^3 a_{t-5}+...\}]\\ &= \mathrm{E}[a_t(a_{t-2}+\phi a_{t-3}+\phi^2 a_{t-4}+...)+\phi a_{t-1}(a_{t-2}+\phi a_{t-3}+\phi^2 a_{t-4}+...)\\ &\quad +\phi^2(a_{t-2}+\phi a_{t-3}+\phi^2 a_{t-4}+...)^2]\\ &= (0+0+0+...+0)+(0+0+0+...+0)\\ &\quad +\phi^2\mathrm{E}[a_{t-2}+\phi a_{t-3}+\phi^2 a_{t-4}+...]^2\\ &= \phi^2\mathrm{E}[(a_{t-2}^2+\phi^2 a_{t-3}^2+\phi^4 a_{t-4}^2+...)+2(\phi a_{t-2}a_{t-3}+\phi^2 a_{t-2}a_{t-4}+...)]\\ &= \phi^2[(\sigma_a^2+\phi^2\sigma_a^2+\phi^4\sigma_a^2+...)+2(0+0+0+...+0)]\\ &= \phi^2\sigma_a^2(1+\phi^2+\phi^4+...)\\ &= \phi^2\frac{\sigma_a^2}{(1-\phi^2)} = \phi^2\sigma_\varepsilon^2 \neq 0\end{aligned}$$

จะเห็นได้ว่า $\mathrm{E}[\varepsilon_t\varepsilon_{t-1}] = \mathrm{E}[\varepsilon_{t-1}\varepsilon_t] = \phi\sigma_\varepsilon^2, \mathrm{E}[\varepsilon_t\varepsilon_{t-2}] = \mathrm{E}(\varepsilon_{t-2}\varepsilon_t) = \phi^2\sigma_\varepsilon^2$ ดังนั้น โดยอาศัยการพิสูจน์ในทำนองเดียวกัน จะได้ความแปรปรวนร่วมระหว่าง ε_t กับ ε_{t-s} ดังนี้

$\mathrm{E}(\varepsilon_t\varepsilon_{t-s}) = \phi^s\sigma_\varepsilon^2 \neq 0$; $s = 1,2,3,\ldots$ เมื่อ $\sigma_\varepsilon^2 = \dfrac{\sigma_a^2}{(1-\phi^2)}$

ในกรณีตัวอย่าง เช่น

$$t = 2, s = 1 \ \mathrm{E}[\varepsilon_2\varepsilon_1] = \mathrm{E}[\varepsilon_1\varepsilon_2] = \phi\sigma_\varepsilon^2 = \phi\sigma_a^2 / (1-\phi^2)$$

$$t = 3, s = 1 \ \mathrm{E}[\varepsilon_3\varepsilon_2] = \mathrm{E}[\varepsilon_2\varepsilon_3] = \phi\sigma_\varepsilon^2 = \phi\sigma_a^2 / (1-\phi^2)$$

$$\mathrm{t} = 3, \mathrm{s} = 2 \ \mathrm{E}[\varepsilon_3\varepsilon_1] = \mathrm{E}[\varepsilon_1\varepsilon_3] = \phi^2\sigma_\varepsilon^2 = \phi^2\sigma_a^2 / (1-\phi^2)$$

$$t = 4, s = 1 \ \mathrm{E}[\varepsilon_4\varepsilon_3] = \mathrm{E}[\varepsilon_3\varepsilon_4] = \phi\sigma_\varepsilon^2 = \phi\sigma_a^2 / (1-\phi^2)$$

$$t = 4, s = 2 \ \mathrm{E}[\varepsilon_4\varepsilon_2] = \mathrm{E}[\varepsilon_2\varepsilon_4] = \phi^2\sigma_\varepsilon^2 = \phi^2\sigma_a^2 / (1-\phi^2)$$

$$t = 4, s = 3 \ \mathrm{E}[\varepsilon_4\varepsilon_1] = \mathrm{E}[\varepsilon_1\varepsilon_4] = \phi^3\sigma_\varepsilon^2 = \phi^3\sigma_a^2 / (1-\phi^2)$$

ดังนั้น จะได้เมทริกซ์ $\mathrm{V}[\varepsilon]$ ดังนี้

$$\mathrm{V}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \phi^2 & \cdots & \phi^{n-2} \\ \phi^2 & \phi & 1 & \phi & \cdots & \phi^{n-3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \phi^{n-4} & \cdots & 1 \end{bmatrix} = \sigma_\varepsilon^2 \mathbf{V} \neq \sigma_\varepsilon^2 \mathbf{I}_n$$

เมื่อ $\sigma_\varepsilon^2 = \dfrac{\sigma_a^2}{(1-\phi^2)}$ จะได้

$$\mathrm{V}(\boldsymbol{\varepsilon}) = \frac{\sigma_a^2}{(1-\phi^2)} \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \phi^2 & \cdots & \phi^{n-2} \\ \phi^2 & \phi & 1 & \phi & \cdots & \phi^{n-3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \phi^{n-4} & \cdots & 1 \end{bmatrix} = \sigma_a^2 \mathbf{U}$$

และพบว่า

$$\mathbf{V}^{-1} = \frac{1}{(1-\phi^2)} \begin{bmatrix} 1 & -\phi & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\phi & 1+\phi^2 & -\phi & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix}$$

และ $\mathbf{U}^{-1} = (1-\phi^2)\,\mathbf{V}^{-1}$ เนื่องจาก $\mathbf{U} \neq \mathbf{I}_n$ จึงไม่อาจใช้วิธีกำลังสองน้อยที่สุดสามัญประมาณค่า $\boldsymbol{\beta}$ ได้อย่างเหมาะสม

4.5 การทดสอบสหสัมพันธ์ในตัวเอง

การทดสอบว่าข้อมูลที่มีอยู่เกิดปัญหาสหสัมพันธ์ในตัวเองหรือไม่นั้น สามารถกระทำได้หลายวิธี ในที่นี้จะเสนอ 2 แบบ คือ การทดสอบสหสัมพันธ์ในตัวเองสำหรับกรณีทั่วไป ซึ่งอาจเป็น AR, MA หรือ ARMA และการทดสอบสหสัมพันธ์ในตัวเองเฉพาะ AR(1)

4.5.1 วิธีทดสอบสหสัมพันธ์ในตัวเองกรณีทั่วไป

การตรวจสอบวิธีนี้สามารถใช้กับ AR(p), MA(q) และ ARMA(p,q) ได้เพียงแต่ระบุดีกรี p และ q ตามที่พิจารณาเห็นสมควรแก่สถานการณ์ลงไปให้ชัดเจน ซึ่งจะแก้ปัญหาได้ เช่น

AR(p) คือ $\varepsilon_t = \sum_{i=1}^{p} \phi_i \varepsilon_{t-i} + a_t$

MA(q) คือ $\varepsilon_t = a_t + \sum_{i=1}^{q} \theta_i a_{t-i}$

ARMA(p,q) คือ $\varepsilon_t = \sum_{i=1}^{p} \phi_i \varepsilon_{i-1} + a_i + \sum_{i=1}^{q} \theta_i a_{t-i}$

ในที่นี้จะกล่าวถึงวิธีปฏิบัติเฉพาะ AR(1) เท่านั้น ส่วน AR(p); $p = 2, 3, \ldots$ หรือ MA(q) และ ARMA(p,q) จะปฏิบัติเช่นเดียวกัน ดังนี้

1. จากสมการ $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ ให้ประมาณค่า $\boldsymbol{\beta}$ โดยวิธีกำลังสองน้อยที่สุดสามัญและคำนวณค่าคลาดเคลื่อนได้ $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ แล้วนำเมทริกซ์ $\hat{\boldsymbol{\varepsilon}}$ มาจำแนกเป็น 2 กลุ่มคือ $\hat{\boldsymbol{\varepsilon}}_t = (\hat{\varepsilon}_2, \hat{\varepsilon}_3, \ldots \hat{\varepsilon}_n)'$ และ $\hat{\boldsymbol{\varepsilon}}_{t-1} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots \hat{\varepsilon}_{n-1})'$ โดยเมทริกซ์ $\hat{\boldsymbol{\varepsilon}}_t$ ทำหน้าที่ของเมทริกซ์ $\mathbf{Y}$ และเมทริกซ์ $\hat{\boldsymbol{\varepsilon}}_{t-1}$ ทำหน้าที่ของเมทริกซ์ $\mathbf{X}$

2. วิเคราะห์สมการถดถอย $\hat{\boldsymbol{\varepsilon}}_t = \phi \hat{\boldsymbol{\varepsilon}}_{t-1} + \hat{a}_t$ ตามวิธีกำลังสองน้อยที่สุดสามัญได้

$$\hat{\phi} = \frac{\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2} \quad , \hat{\sigma}_\varepsilon^2 = \frac{1}{n-1} \left(\sum_{t=1}^{n-1} \hat{\varepsilon}_t^2 - \hat{\phi} \sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} \right)$$

$$\hat{\mathrm{V}}(\hat{\phi}) = \frac{\hat{\sigma}_{\varepsilon}^{2}}{\sum_{t=2}^{n} \hat{\varepsilon}_{t-2}^{2}} \text{ และ } S_{\hat{\phi}} = \frac{\hat{\sigma}_{\varepsilon}}{\sqrt{\sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^{2}}}$$

3. ทดสอบสมมติฐาน $\mathrm{H}_0 : \phi = 0$

$\mathrm{H}_1 : \phi \neq 0$

โดยจะปฏิเสธสมมติฐานว่าง ณ ระดับนัยสำคัญ α เมื่อ $|t_c| > t_{n-1,1-\frac{\alpha}{2}}$ เมื่อ $t_c = \frac{\hat{\phi}}{S_{\hat{\phi}}}$ ซึ่งในกรณีนี้ ถ้าพบว่า ϕ ไม่มีนัยสำคัญทางสถิติ จะเป็นการยืนยันว่าข้อมูลที่มีอยู่ไม่เกิดปัญหาสหสัมพันธ์ในตัวเองในรูปแบบ AR(1)

สำหรับการทดสอบปัญหาสหสัมพันธ์ในตัวเองในรูปแบบ AR, MA และ ARMA ดีกรีอื่น ๆ ให้ใช้การทดสอบเอฟ (F-test) หรือยังคงใช้การทดสอบที (t-test) ได้เช่นใน AR(2) คือ

$$\hat{\boldsymbol{\varepsilon}}_t = \phi_1 \hat{\boldsymbol{\varepsilon}}_{t-1} + \phi_2 \hat{\boldsymbol{\varepsilon}}_{t-2} + \hat{a}_t$$

จำแนกเมทริกซ์ $\hat{\boldsymbol{\varepsilon}}$ ออกเป็น 3 กลุ่ม คือ $\hat{\boldsymbol{\varepsilon}}_{\mathrm{t}} = (\hat{\varepsilon}_3, \hat{\varepsilon}_4, ..., \hat{\varepsilon}_n)'$, $\hat{\boldsymbol{\varepsilon}}_{\mathrm{t}-1} = (\hat{\varepsilon}_2, \hat{\varepsilon}_3, ..., \hat{\varepsilon}_{n-1})'$ และ $\hat{\boldsymbol{\varepsilon}}_{\mathrm{t}-2} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_{n-2})'$

จากนั้นทำการวิเคราะห์สมการถดถอยตามวิธีการถดถอยพหุคูณ แล้วทดสอบสมมติฐาน

$\mathrm{H}_0 : \phi_1 = 0,\ \phi_2 = 0$

$\mathrm{H}_0 : \phi_1 \neq 0,\ \phi_2 \neq 0$

4.5.2 วิธีทดสอบสหสัมพันธ์ในตัวเอง AR(1)

การทดสอบของเดอร์บินวัตสัน (Durbin-Watson Test: DW)

วิธีทดสอบของเดอร์บิน และวัตสันเป็นเครื่องมือทดสอบสหสัมพันธ์ในตัวเองใน AR(1) ที่มีผู้นิยมใช้กันอย่างกว้างขวาง เดอร์บินและวัตสันเสนอตัวทดสอบสำหรับสมมติฐานต่อไปนี้ ดังนี้

$\mathrm{H}_0 :\ \phi = 0$

$\mathrm{H}_1 :\ \phi > 0$

ปฏิเสธสมมติฐานว่าง ณ ระดับนัยสำคัญ α ถ้า $d < d_L$

ยอมรับสมมติฐานว่าง ถ้า $d > d_U$ และไม่อาจตัดสินใจได้ ถ้า $d_L < d < d_U$

$\mathrm{H}_0 :\ \phi = 0$

$\mathrm{H}_1 :\ \phi < 0$

ปฏิเสธสมมติฐานว่าง ณ ระดับนัยสำคัญ α ถ้า $d > 4 - d_L$

ยอมรับสมมติฐานว่าง ถ้า $d < 4 - d_U$ และไม่อาจตัดสินใจได้ ถ้า $4 - d_U < d < 4 - d_L$

โดยที่ $d = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$

เมื่อ $\hat{\varepsilon}_t$ คือ ค่าคลาดเคลื่อนโดยวิธีกำลังสองน้อยที่สุดสามัญ $(\hat{\varepsilon}_t = Y_t - \hat{Y}_t)$

พิจารณา d ในรูปเมทริกซ์ จะได้

$$d = \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{A}\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}} = \frac{\boldsymbol{\varepsilon}'\mathbf{M}\mathbf{A}\mathbf{M}\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}}$$

พิสูจน์

$$d = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$$

$$= \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t^2 + \hat{\varepsilon}_{t-1}^2 - 2\hat{\varepsilon}_t\hat{\varepsilon}_{t-1})}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = \frac{\sum_{t=2}^{n}\hat{\varepsilon}_t^2 + \sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2 - 2\sum_{t=2}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}$$

$$= \frac{a}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}$$

พิจารณาเทอมเศษจะพบว่าสามารถจัดรูปเป็นรูปแบบกำลังสอง (Quadratic Form) ได้ดังนี้

$$a = (\hat{\varepsilon}_2^2 + \hat{\varepsilon}_3^2 + \ldots + \hat{\varepsilon}_n^2) + (\hat{\varepsilon}_1^2 + \hat{\varepsilon}_2^2 + \ldots + \hat{\varepsilon}_{n-1}^2) - 2(\hat{\varepsilon}_1\hat{\varepsilon}_2 + \hat{\varepsilon}_2\hat{\varepsilon}_3 + \hat{\varepsilon}_3\hat{\varepsilon}_4 + \ldots\hat{\varepsilon}_{n-1}\hat{\varepsilon}_n)$$

$$= \{\hat{\varepsilon}_1^2 + (2\hat{\varepsilon}_2^2 + 2\hat{\varepsilon}_3^2 + \ldots + 2\hat{\varepsilon}_{n-1}^2) + \hat{\varepsilon}_n^2\} - 2\hat{\varepsilon}_1\hat{\varepsilon}_2 - 2\hat{\varepsilon}_2\hat{\varepsilon}_3 - 2\hat{\varepsilon}_3\hat{\varepsilon}_4 - \ldots - 2\hat{\varepsilon}_{n-1}\hat{\varepsilon}_n$$

$$a = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n)\begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \hat{\varepsilon}_n \end{bmatrix} = \hat{\boldsymbol{\varepsilon}}'\mathbf{A}\hat{\boldsymbol{\varepsilon}}$$

ดังนั้น $d = \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{A}\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}$

แต่เนื่องจาก $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$

$$= [\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}] - \mathbf{X}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})]$$

$$= \mathbf{X\beta} + \boldsymbol{\varepsilon} - \mathbf{X}[\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}]$$

$$= \boldsymbol{\varepsilon} - \mathrm{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$$

$$= [\mathrm{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\boldsymbol{\varepsilon}$$

$$= \mathbf{M}\boldsymbol{\varepsilon}$$

ดังนั้น $$d = \frac{\hat{\boldsymbol{\varepsilon}}'\mathbf{A}\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}} = \frac{\boldsymbol{\varepsilon}'\mathbf{M}'\mathbf{A}\mathbf{M}\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}'\mathbf{M}'\mathbf{M}\boldsymbol{\varepsilon}} = \frac{\boldsymbol{\varepsilon}'\mathbf{M}\mathbf{A}\mathbf{M}\boldsymbol{\varepsilon}}{\boldsymbol{\varepsilon}'\mathbf{M}\boldsymbol{\varepsilon}}$$

โดยที่ $\mathbf{M} = \mathrm{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

และ $$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

ข้อสังเกต

1. $d = \dfrac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}$ สามารถพิสูจน์ได้ว่า $d \cong 2(1-\phi)$ ซึ่งแสดงว่า ถ้าไม่มีสหสัมพันธ์ในตัวเอง คือ $\phi = 0$ จะพบว่า $d \cong 2$ และถ้า d มีค่าต่ำใกล้ 0 แสดงว่าเกิดปัญหาอัตตสหสัมพันธ์เชิงบวกและถ้า d มีค่าใกล้ 4 แสดงว่าเกิดปัญหาอัตตสหสัมพันธ์เชิงลบ

2. ช่วงการตัดสินใจ (Indecision Interval) คือ ช่วงที่ $d_L < d < d_U$ และ $4 - d_U < d < 4 - d_L$ เป็นช่วงที่ไม่อาจสรุปผลใด ๆ ได้ นอกจากต้องเพิ่มขนาดตัวอย่างซึ่งไม่ใช่สิ่งที่กระทำได้ง่าย โดยเฉพาะการวิจัยที่อาศัยข้อมูลอนุกรมเวลา

3. วิธีทดสอบเดอร์บินและวัตสันไม่ได้กำหนดการแจกแจงค่าตัวอย่างของ d ไว้ ทั้งนี้เพราะค่าของ d ขึ้นอยู่กับลักษณะของเมทริกซ์ $\mathbf{X}$ ซึ่งถ้าเมทริกซ์ $\mathbf{X}$ มีลักษณะเปลี่ยนไป $f(d)$ จะเปลี่ยนแปลงไป ทำให้ไม่อาจกำหนดรูปทั่วไปของ $f(d)$ ได้

4. วิธีทดสอบเดอร์บินและวัตสันใช้ได้เฉพาะ AR(1) เท่านั้น ไม่อาจใช้กับ AR(p), p=2,3,... ได้ นอกจากนี้ถ้ามีตัวแปรช่วงเวลาย้อนหลัง (Lagged Variable) ปนอยู่ในระหว่างตัวแปรอิสระ วิธีทดสอบเดอร์บินและวัตสัน จะมีอำนาจการทดสอบต่ำกว่าวิธีอื่น ๆ

4.6 การแก้ไขปัญหาช่วงการตัดสินใจ

วิธีที่ 1 ให้ใช้ d_U และ $4-d_U$ เป็นเขตวิกฤต กล่าวคือ

ปฏิเสธ $H_0:\phi=0$

$H_1:\phi>0$ เมื่อ $d<d_U$

และปฏิเสธ $H_0:\phi=0$

$H_1:\phi<0$ เมื่อ $d>4-d_U$

วิธีนี้เสนอโดยแฮนแนนและเทอร์เรล (Hannan, E.J. & Terrell, R.D., 1968) หลักการของวิธีนี้คือ การรวมเอาช่วงการตัดสินใจเข้าเป็นเขตวิกฤต วิธีนี้แม้จะทำให้ระดับของความน่าจะเป็นของความผิดพลาดประเภทที่ 1 (Probability of type I error) สูงกว่าที่กำหนดไว้เดิม แต่ก็ยังใช้ได้ผลดีโดยเฉพาะอย่างยิ่งเมื่อค่าของตัวแปรอิสระมีค่าความแปรปรวนไม่มากนัก

วิธีที่ 2 ประมาณการแจกแจงของ d ด้วยการแจกแจงของ $a+bd_U$

วิธีนี้เรียกว่าการประมาณของ $a+bd_U$ ของเดอร์บินวัตสัน (Dubin-Watson’s $a+bd_U$ Approximation) ซึ่งเสนอโดยเดอร์บินและวัตสัน ในปี 1971 มีวิธีการดังนี้

1. คำนวณหาค่าของ $\mathrm{E}(d)$ และ $\mathrm{V}(d)$

$E(d) = \dfrac{P}{n-k-1}$ เมื่อ $P=2(n-1)-\mathrm{tr}[\mathbf{X}'\mathbf{A}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]$ และ $\mathbf{A}$ คือ เมทริกซ์ของรูปแบบกำลังสอง

เมื่อ $\mathrm{V}(d)=\dfrac{2}{(n-k-1)(n-k+1)}[\mathrm{Q}-\mathrm{PE}(d)]$

และ $\mathrm{Q}=2(3n-4)-2\mathrm{tr}[\mathbf{X}'\mathbf{A}^2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]+\mathrm{tr}[\mathbf{X}'\mathbf{A}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]^2$

2. แก้สมการหาค่าของ a และ b จากสมการ $\mathrm{E}(d)=a+b\mathrm{E}(d_U)$ และ $\mathrm{V}(d)=b^2\mathrm{V}(d_U)$

เมื่อพิจารณาสมการคู่นี้ พบว่าสามารถแก้สมการหาค่าของ a และ b ได้โดยง่าย ดังนี้

$$b = \sqrt{\frac{\mathrm{V}(d)}{\mathrm{V}(d_U)}}$$

$$a = \mathrm{E}(d)-b\mathrm{E}(d_U) = \mathrm{E}(d)-\mathrm{E}(d_U)\sqrt{\frac{\mathrm{V}(d)}{\mathrm{V}(d_U)}}$$

ทั้งนี้หากทราบค่าของ $\mathrm{E}(d)$ และ $\mathrm{V}(d)$ ที่ได้จากข้อที่ 1 ส่วน $\mathrm{E}(d_U)$ และ $\mathrm{V}(d_U)$ สามารถหาค่าได้จากสมการต่อไปนี้

$$\mathrm{E}(d_U)=2+\frac{2\sum_{j=1}^{k}\cos(j\pi/n)}{n-k-1}$$

$$\mathrm{V}(d_U)=\frac{4[n-2-2\sum_{j=1}^{k}\cos^2(j\pi/n)-2\dfrac{\left\{\sum_{j=1}^{k}\cos(j\pi/\mathrm{n})\right\}^2}{n-k-1}}{(n-k-1)(n-k+1)}$$

3. ปฏิเสธ $\mathrm{H}_0:\phi=0$

$\mathrm{H}_1:\phi>0$ ณ ระดับนัยสำคัญ α เมื่อ $d<a+bd_U^0$

เมื่อ d_U^0 คือ ค่าวิกฤตของ d_U จากตารางในภาคผนวก ง

วิธีที่ 3 ใช้ $\mathrm{N}(0,1)$ หรือ การทดสอบแซท ($\mathrm{Z-test}$)

ในปี ค.ศ. 1972 Blattberg แนะนำให้ใช้การทดสอบแซทตัดสินใจสำหรับการทดสอบสมมติฐาน

$\mathrm{H}_0:\phi=0$

$\mathrm{H}_1:\phi>0$ หรือ $\phi<0$ โดยให้ตัดสินใจดังนี้

ปฏิเสธ $\mathrm{H}_0:\phi=0$

$\mathrm{H}_1:\phi>0$ ณ ระดับนัยสำคัญ α เมื่อ $\mathrm{Z_c}>\mathrm{Z}_{1-\alpha}$

ปฏิเสธ $\mathrm{H}_0:\phi=0$

$\mathrm{H}_1:\phi<0$ ณ ระดับนัยสำคัญ α เมื่อ $\mathrm{Z_c}<\mathrm{Z}_\alpha$

โดยที่ $\mathrm{Z_c}=\dfrac{d-\mathrm{E}(d)}{\sqrt{\mathrm{V}(d)}}$

และ $E(d) = \dfrac{P}{n-k-1}$ เมื่อ $P=2(n-1)-\mathrm{tr}[\mathbf{X'\,A\,X}(\mathbf{X'X})^{-1}]$

$\mathrm{V}(d)=\dfrac{2}{(n-k-1)(n-k+1)}[\mathrm{Q}-\mathrm{PE}(d)]$ เมื่อ d คือ ค่าสถิติของเดอร์บินและวัตสัน และ $\mathrm{Q}=2(3n-4)-2\mathrm{tr}[\mathbf{X'\,A^2\,X}(\mathbf{X'X})^{-1}]+\mathrm{tr}[\mathbf{X'\,A\,X}(\mathbf{X'X})^{-1}]^2$

การตัดสินใจตามวิธีนี้นอกจากช่วยแก้ปัญหาเรื่องช่วงการตัดสินใจแล้วยังสามารถใช้ทดสอบได้ในทุกระดับนัยสำคัญ ขณะที่วิธีทดสอบเดอร์บินและวัตสันใช้ทดสอบได้เพียง 3 ระดับคือ 5%, 2.5% และ 1%

4.7 การประมาณค่าพารามิเตอร์เมื่อมีสหสัมพันธ์ในตัวเองในรูปแบบ AR(1)

4.7.1 การประมาณค่าพารามิเตอร์ $\boldsymbol{\beta}$ โดยวิธีกำลังสองน้อยที่สุดทั่วไป (Generalized Least Square: GLS)

จาก $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ เมื่อ $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')=\sigma_\varepsilon^2\,\mathbf{V}=\dfrac{\sigma_a^2}{(1-\phi^2)}\mathbf{V}=\sigma_a^2\,\mathbf{U}\neq\sigma_\varepsilon^2\mathbf{I}_{\mathrm{n}}$

$$\mathbf{U}=\frac{1}{(1-\phi^2)}\begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \phi^2 & \cdots & \phi^{n-2} \\ \phi^2 & \phi & 1 & \phi & \cdots & \phi^{n-3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \phi^{n-4} & \cdots & 1 \end{bmatrix}$$

$$\mathbf{U}^{-1}=\begin{bmatrix} 1 & -\phi & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\phi & 1+\phi^2 & -\phi & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix}$$

ถ้าทราบค่า ϕ จะสามารถประมาณค่า $\boldsymbol{\beta}$ ได้ดังนี้

$$\hat{\boldsymbol{\beta}}=(\mathbf{X}'\,\mathbf{U}^{-1}\,\mathbf{X})^{-1}\,\mathbf{X}'\,\mathbf{U}^{-1}\,\mathbf{Y}\ ,\ \hat{\sigma}^2=(\mathbf{Y}'\,\mathbf{U}^{-1}\,\mathbf{Y}-\hat{\boldsymbol{\beta}}'\,\mathbf{X}'\,\mathbf{U}^{-1}\,\mathbf{Y})/(n-k-1)$$

หรืออาศัยเทคนิคการแปลงรูปโดยใช้เมทริกซ์ $\mathbf{P}$ ที่ $\mathbf{P}'\,\mathbf{P}=\mathbf{U}^{-1}$ จะได้เมทริกซ์ $\mathbf{P}$ สำหรับกรณีนี้ คือ

$$\mathbf{P}=\begin{bmatrix} \sqrt{1-\phi^2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\phi & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\phi & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\phi & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix}$$

ดังนั้น

$$\mathbf{P}'\mathbf{P} = \begin{bmatrix} 1 & -\phi & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\phi & 1+\phi^2 & -\phi & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\phi & 1+\phi^2 & -\phi & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix} = \mathbf{U}^{-1}$$

โดยอาศัยเทคนิคการแปลงรูป สามารถนำเมทริกซ์ $\mathbf{P}$ คูณด้านหน้าสมการ $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ จะได้ $\mathbf{P}\mathbf{Y} = (\mathbf{P}\mathbf{X})\boldsymbol{\beta} + \mathbf{P}\boldsymbol{\varepsilon}$ หรือ $\mathbf{Y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$ ซึ่งในกรณีนี้จะพบว่า $\mathrm{V}(\mathbf{P}\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}_n$ โดยไม่ขัดกับข้อสมมติเดิมของการวิเคราะห์การถดถอย

จากสมการ $\mathbf{P}\mathbf{Y} = (\mathbf{P}\mathbf{X})\boldsymbol{\beta} + \mathbf{P}\boldsymbol{\varepsilon}$ สามารถประมาณค่า $\boldsymbol{\beta}$ ได้ดังนี้

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= [(\mathbf{P}\mathbf{X})'(\mathbf{P}\mathbf{X})]^{-1}[(\mathbf{P}\mathbf{X})'(\mathbf{P}\mathbf{Y})] \\ &= [\mathbf{X}'(\mathbf{P}'\mathbf{P})\mathbf{X}]^{-1}[\mathbf{X}'(\mathbf{P}'\mathbf{P})\mathbf{Y}] \\ &= [(1-\phi^2)\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}]^{-1}[(1-\phi^2)\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}] \\ &= (\mathbf{X}'\mathbf{U}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{U}^{-1}\mathbf{Y}) \end{aligned}$$

$$\begin{aligned} \hat{\sigma}^2 &= [(\mathbf{P}\mathbf{Y})'(\mathbf{P}\mathbf{Y}) - \hat{\boldsymbol{\beta}}'(\mathbf{P}\mathbf{X})'(\mathbf{P}\mathbf{Y})]/(n-k-1) \\ &= [(1-\phi^2)\mathbf{Y}'\mathbf{V}^{-1}\mathbf{Y} - (1-\phi^2)\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{V}^{-1}]/(n-k-1) \\ &= (\mathbf{Y}'\mathbf{U}^{-1}\mathbf{Y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{U}^{-1}\mathbf{Y})/(n-k-1) \end{aligned}$$

เมื่อนำเทคนิคการแปลงรูปไปวิเคราะห์ข้อมูลสามารถทำได้ ดังนี้

จากสมการ $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + \varepsilon_t$ (1)

เมื่อแทน t ด้วย t-1 จะได้

$$Y_{t-1} = \beta_0 + \beta_1 X_{1,t-1} + \beta_2 X_{2,t-1} + \ldots + \beta_k X_{k,t-1} + \varepsilon_{t-1} \quad (2)$$

คูณตลอดสมการที่ (2) ด้วย ϕ จะได้

$$\phi Y_{t-1} = \phi\beta_0 + \phi\beta_1 X_{1,t-1} + \phi\beta_2 X_{2,t-1} + \ldots + \phi\beta_k X_{k,t-1} + \phi\varepsilon_{t-1} \quad (3)$$

(1) – (3) จะได้

$$Y_t - \phi Y_{t-1} = \beta_0(1-\phi) + \beta_1(X_{1t} - \phi X_{1,t-1}) + \ldots + \beta_k(X_{kt} - \phi X_{k,t-1}) + (\varepsilon_t - \phi\varepsilon_{t-1}) \quad (4)$$

พิจารณาสมการที่ (4) จะพบว่า ค่าคลาดเคลื่อน คือ $\varepsilon_t - \phi\varepsilon_{t-1}$ นั่นคือ a_t ในสมการ AR (1) คือ $\varepsilon_t = \phi\varepsilon_{t-1} + a_t$ นั่นเอง แต่ $E(a_t^2) = \sigma_a^2$ ในทุกค่าของ t แสดงว่าค่าคลาดเคลื่อนในสมการที่ (4)

ไม่มีสหสัมพันธ์ในตัวเอง ดังนั้น เมื่อดำเนินการวิเคราะห์สมการถดถอยโดยใช้สมการที่ (4) สามารถแก้ปัญหาสหสัมพันธ์ในตัวเองในรูปแบบ AR (1) ได้

จากสมการที่ (4) คือ

$$Y_t - \phi Y_{t-1} = \beta_0(1-\phi) + \beta_1(X_{1t} - \phi X_{1,t-1}) + \ldots + \beta_k(X_{kt} - \phi X_{k,t-1}) + (\varepsilon_t - \phi\varepsilon_{t-1}) \ ; t = 2,3,\ldots n$$

เมื่อจัดให้อยู่ในรูปเมทริกซ์จะพบว่า

$$\begin{bmatrix} Y_2 - \phi Y_1 \\ Y_3 - \phi Y_2 \\ \vdots \\ Y_n - \phi Y_{n-1} \end{bmatrix} = \begin{bmatrix} 1-\phi & X_{12} - \phi X_{11} & \cdots & X_{k2} - \phi X_{k1} \\ 1-\phi & X_{13} - \phi X_{12} & \cdots & X_{k3} - \phi X_{k2} \\ \vdots & \vdots & & \vdots \\ 1-\phi & X_{1n} - \phi X_{1,n-1} & \cdots & X_{kn} - \phi X_{k,n-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_2 - \phi\varepsilon_1 \\ \varepsilon_3 - \phi\varepsilon_2 \\ \vdots \\ \varepsilon_n - \phi\varepsilon_{n-1} \end{bmatrix}$$

หรือ $\mathbf{Y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$ (5)

โดยที่ $\mathbf{Y}^*$ และ $\boldsymbol{\varepsilon}^*$ เป็นเมทริกซ์ขนาด $(n-1)\times 1$ ขณะที่ $\mathbf{X}^*$ เป็นเมทริกซ์ขนาด $(n-1)\times(k+1)$ และเมื่อพิจารณาเมทริกซ์ $\mathbf{Y}^*$ และเมทริกซ์ $\mathbf{X}^*$ จะพบว่า

$$\mathbf{Y}^* = \begin{bmatrix} Y_2 - \phi Y_1 \\ Y_3 - \phi Y_2 \\ Y_4 - \phi Y_3 \\ \vdots \\ Y_n - \phi Y_{n-1} \end{bmatrix} = \begin{bmatrix} -\phi & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\phi & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\phi & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\phi & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix} = \mathbf{P_1}\,\mathbf{Y}$$

ขณะเดียวกันเมทริกซ์ $\mathbf{X}^*$ คือ เมทริกซ์ $\mathbf{P_1}\,\mathbf{X}$ โดยที่ $\mathbf{P_1}$ คือ เมทริกซ์ขนาด $(n-1)\times n$

ด้วยเหตุนี้ สมการที่ (5) คือ $\mathbf{Y}^* = \mathbf{X}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}^*$ คือสมการที่เกิดจากการแปลงรูปโดยอาศัยเมทริกซ์ $\mathbf{P_1}$ กล่าวคือ แปลงรูปจากสมการ $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ เป็น $\mathbf{P_1}\,\mathbf{Y} = (\mathbf{P_1}\,\mathbf{X})\boldsymbol{\beta} + \mathbf{P_1}\,\boldsymbol{\varepsilon}$ สำหรับเมทริกซ์ $\mathbf{P_1}$ นั้นถ้าสังเกตให้ดีจะพบว่า $\mathbf{P_1}$ คือ เมทริกซ์ที่เกิดจากการตัดแถวที่ 1 ของเมทริกซ์ $\mathbf{P}$ ทิ้งไปนั่นเอง การประมาณค่าเมทริกซ์ $\boldsymbol{\beta}$ โดยอาศัยสมการที่ (4) หรือสมการที่ (5) จะแตกต่างจากเทคนิคการแปลงรูปเล็กน้อย เพราะสูญเสียความเป็นอิสระขององศาเสรี (Degree of Freedom) ไปเท่ากับ 1 เนื่องจากวิธีนี้ระดับความเป็นอิสระของค่าคลาดเคลื่อนมีเพียง $(n-1)-k-1$ ขณะที่วิธีกำลังสองน้อยที่สุดทั่วไปมีระดับความเป็นอิสระของค่าคลาดเคลื่อนเท่ากับ $n-k-1$

4.7.2 การประมาณค่าพารามิเตอร์ $\boldsymbol{\beta}$ โดยวิธีตัวประมาณกำลังสองน้อยที่สุดทั่วไป (Estimated Generalized Least Square: EGLS)

ในกรณีของการประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยที่สุดทั่วไปนั้น ต้องใช้เมทริกซ์ $\mathbf{P}$ และ $\mathbf{U}^{-1}$ ซึ่งต้องทราบค่า ϕ จึงจะสามารถประมาณค่าเมทริกซ์ $\boldsymbol{\beta}$ ได้ แต่โดยทั่วไปมักจะไม่ทราบค่า ϕ ด้วยเหตุนี้จึงได้มีความพยายามที่จะประมาณค่าของ ϕ ขึ้นมาด้วยวิธีการที่เหมาะสม

แล้วแทนที่ ϕ ด้วย $\hat{\phi}$ เพราะเมทริกซ์ที่จำเป็นต้องใช้สำหรับวิธีกำลังสองน้อยที่สุดทั่วไป คือ $\hat{\mathbf{P}}$ และ $\hat{\mathbf{U}}^{-1}$ จากนั้นจึงจะสามารถประมาณค่าเมทริกซ์ $\boldsymbol{\beta}$ ได้

$$\hat{\boldsymbol{\beta}}=(\mathbf{X}'\,\hat{\mathbf{U}}^{-1}\,\mathbf{X})^{-1}\,\mathbf{X}'\,\hat{\mathbf{U}}^{-1}\,\mathbf{Y}\,,\ \hat{\sigma}^2=(\mathbf{Y}'\,\hat{\mathbf{U}}^{-1}\,\mathbf{Y}-\hat{\boldsymbol{\beta}}'\,\mathbf{X}'\,\hat{\mathbf{U}}^{-1}\,\mathbf{Y})/(n-k-1)$$

และ $\mathrm{SS}(\beta_0)=\mathrm{n}\overline{\mathrm{Y}}^{*2}$ เมื่อ $\overline{\mathrm{Y}}^{*}$ คือ ค่าเฉลี่ยจากสมาชิกของเมทริกซ์ $\mathbf{Y}^*$ เมื่อ $\mathbf{Y}^*=\hat{\mathbf{P}}\,\mathbf{Y}$

การประมาณค่า ϕ สามารถกระทำได้หลายวิธี ดังนี้

1. ประมาณค่า ϕ โดยใช้สหสัมพันธ์ r ระหว่างค่าคลาดเคลื่อนโดยวิธีกำลังสองน้อยที่สุดสามัญ

จากค่าคลาดเคลื่อนโดยวิธีกำลังสองน้อยที่สุดสามัญ $\hat{\boldsymbol{\varepsilon}}=\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}$ เมื่อ $\hat{\boldsymbol{\varepsilon}}=(\hat{\varepsilon}_1,\hat{\varepsilon}_2,...,\hat{\varepsilon}_n)'$ แยกเมทริกซ์ $\hat{\boldsymbol{\varepsilon}}$ ออกเป็น 2 กลุ่ม คือ $\hat{\boldsymbol{\varepsilon}}_{t-1}=(\hat{\varepsilon}_1,\hat{\varepsilon}_2,...,\hat{\varepsilon}_{n-1})'$ และ $\hat{\boldsymbol{\varepsilon}}_t=(\hat{\varepsilon}_2,\hat{\varepsilon}_3,...,\hat{\varepsilon}_n)'$ แล้ว คำนวณค่าสหสัมพันธ์เชิงเส้นระหว่าง $\hat{\boldsymbol{\varepsilon}}_t$ กับ $\hat{\boldsymbol{\varepsilon}}_{t-1}$

$$r=\frac{\sum_{t=2}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sqrt{\sum_{t=2}^{n}\hat{\varepsilon}_t^2\sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2}}\cong\frac{\sum_{t=2}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2}=r_1=\hat{\phi}$$

หรือใช้ r_1 แทน r เป็นค่าประมาณค่าของ ϕ

แล้วนำค่า r_1 ที่ได้มาแทนค่าลงในเมทริกซ์ $\hat{\mathbf{P}}_1$ ตามวิธีการที่เลือกใช้ ดังนี้

(1) วิธีของ Cochrane-Orcutt

ใช้ r_1 เป็นตัวประมาณค่าของ ϕ โดยแทนที่ ϕ ในเมทริกซ์ $\hat{\mathbf{P}}_1$ ด้วย r_1 (เรียกว่า Cochrane-OrcuttTwo-Step Procedure)

$$\hat{\mathbf{P}}_1=\begin{bmatrix}-r_1 & 1 & 0 & \cdots & 0 & 0 & 0\\ 0 & -r_1 & 1 & \cdots & 0 & 0 & 0\\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & -r_1 & 1 & 0\\ 0 & 0 & 0 & \cdots & 0 & -r_1 & 1\end{bmatrix}$$

แล้ววิเคราะห์สมการถดถอย $\hat{\mathbf{P}}_1\,\mathbf{Y}=(\hat{\mathbf{P}}_1\,\mathbf{X})\boldsymbol{\beta}+\hat{\mathbf{P}}_1\,\boldsymbol{\varepsilon}$ หรืออีกนัยหนึ่งคือ ให้แทนที่ ϕ ด้วย r_1 แล้ววิเคราะห์สมการถดถอย

$$(Y_t-r_1Y_{t-1})=\beta_0(1-r_1)+\beta_1(X_{1t}-r_1X_{1,t-1})+...+\beta_k(X_{kt}-r_1X_{k,t-1})+(\varepsilon_t-r_1\varepsilon_{t-1})$$

เมื่อ $t=2,3,...,n$

(2) วิธีของ Prais-Winsten

ใช้ r_1 เป็นตัวประมาณค่าของ ϕ โดยแทนที่ ϕ ในเมทริกซ์ $\hat{\mathbf{P}}$ ด้วย r_1

$$\hat{\mathbf{P}} = \begin{bmatrix} \sqrt{1-r_1^2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -r_1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -r_1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -r_1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -r_1 & 1 \end{bmatrix}$$

แล้ววิเคราะห์สมการถดถอย $\hat{\mathbf{P}}\mathbf{Y} = (\hat{\mathbf{P}}\mathbf{X})\boldsymbol{\beta} + \hat{\mathbf{P}}\boldsymbol{\varepsilon}$

สำหรับวิธีที่ 1 นั้น ยังสามารถใช้สัมประสิทธิ์การถดถอยจากสมการ $\hat{\varepsilon}_t = \phi\hat{\varepsilon}_{t-1} + \hat{a}_t$ เป็นตัวประมาณค่าของ ϕ ได้ นั่นคือ

$$\hat{\phi} = \frac{\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2}$$

และเมื่อ $n \to \infty$ สัมประสิทธิ์สหสัมพันธ์ r_1 และสัมประสิทธิ์การถดถอย ϕ จะมีค่าเท่ากัน

2. วิธีค่าสถิติของเดอร์บินและวัตสัน (Durbin-Watson Statistics) คือ $\hat{\phi} \cong 1 - \dfrac{d}{2}$

จากค่าสถิติของเดอร์บินและวัตสัน คือ

$$d = \frac{\sum_{t=2}^{n} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

$$= \frac{\sum_{t=2}^{n} (\hat{\varepsilon}_t^2 + \hat{\varepsilon}_{t-1}^2 - 2\hat{\varepsilon}_t \hat{\varepsilon}_{t-1})}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2} = \frac{\sum_{t=2}^{n} \hat{\varepsilon}_t^2 + \sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2 - 2\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}$$

แต่เมื่อ $\mathrm{n} \to \infty$ จะพบว่า $\sum_{t=2}^{n} \hat{\varepsilon}_t^2$, $\sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2$, $\sum_{t=1}^{n} \hat{\varepsilon}_t^2$ จะมีค่าใกล้เคียงกัน

นั่นคือ เมื่อ $\mathrm{n} \to \infty$ แล้ว จะมีผลให้

$$d \cong \frac{2\sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2 - 2\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2} = 2 - 2\hat{\phi}$$

เมื่อ $\hat{\phi}$ คือ ค่าประมาณสัมประสิทธิ์การถดถอย จากสมการ $\hat{\varepsilon}_t = \phi\hat{\varepsilon}_{t-1} + \hat{a}_t$

นั่นคือ เมื่อ $n \to \infty$ จะได้ว่า $d \cong 2 - 2\hat{\phi}$ หรือ $\hat{\phi} \cong 1 - \frac{d}{2}$

เมื่อ d คือ ค่าสถิติของเดอร์บินและวัตสัน

วิธีนี้เหมาะสำหรับกรณีของกลุ่มตัวอย่างขนาดใหญ่ เพราะถ้า n มีขนาดเล็ก ค่าของ $\hat{\phi}$ จะไม่ค่อยแม่นยำ

3. วิธีธีลและนาการ์ (Theil-Nager Modification)

ในปี ค.ศ. 1961 ธีลและนาการ์ (Theil, H and Nager, A. L.) ได้เสนอค่าประมาณของ ϕ โดยปรับปรุงวิธีที่ 2 ให้เหมาะสมกับสถานการณ์ที่ค่าตัวแปรอิสระ $X's$ ต่าง ๆ มีค่าใกล้เคียงกัน ดังนี้

$$\phi^* = \frac{n^2(1-\frac{d}{2}) + (k+1)^2}{n^2 - (k+1)^2}$$

เมื่อ n แทน จำนวนข้อมูล

d แทน ค่าสถิติของเดอร์บินและวัตสัน

k แทน จำนวนตัวแปรอิสระ

การประมาณค่า ϕ ตามวิธีนี้จะดีที่สุดเมื่อค่า $\phi < 0.4$

ในการวิจัยนี้ใช้วิธีการประมาณค่า ϕ ด้วยวิธีการของเพรส วินสเทน (Prais-Winsten)

4.8 การพยากรณ์

4.8.1 การพยากรณ์ในกรณีวิธีกำลังสองน้อยที่สุดทั่วไป

จากสมการของความสัมพันธ์ที่แท้จริง (True Relationship)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon_i \qquad ; i = 1, 2, 3, \ldots, n$$

หรือ

$$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

โดยที่ $\mathrm{E}(\boldsymbol{\varepsilon}) = 0$ และ $\mathrm{E}(\boldsymbol{\varepsilon\varepsilon}') = \phi \neq \sigma^2 \mathbf{I}_n$

สมมติว่าทราบความความสัมพันธ์ที่แท้จริง และทราบค่าของ $X's$ ในอนาคตคือ ทราบว่า $X_0 = (1, X_{01}, X_{02}, \ldots, X_{0p})$ เราจะพยากรณ์ค่าในอนาคตของ Y ได้ดังนี้คือ

$$Y_0 = X_0\beta + \varepsilon_0$$

หรือ $Y_0 = \beta_0 + \beta_1 X_{01} + \beta_2 X_{02} + \ldots + \beta_p X_{0p} + \varepsilon_0$

โดยที่ ε_0 คือ Prediction Disturbance และ

$$\mathrm{E}(\varepsilon_0) = 0,\ \mathrm{E}(\varepsilon_o^2) = \sigma_0^2$$

และความผันแปรร่วมระหว่าง Prediction Disturbance กับ Disturbance ใน sample Period คือ

$$\mathrm{E}(\varepsilon_0, \boldsymbol{\varepsilon}) = \mathrm{E}(\varepsilon_0)\begin{bmatrix}\varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n\end{bmatrix} = \mathrm{E}\begin{bmatrix}\varepsilon_0\varepsilon_1 \\ \varepsilon_0\varepsilon_2 \\ \vdots \\ \varepsilon_0\varepsilon_n\end{bmatrix} = \mathbf{W}$$

ปัญหาของนักวิจัยคือ การหาตัวพยากรณ์ไม่เอนเอียงเชิงเส้นดีที่สุด (Best Linear Unbiased Predictor: BLUP) ของ Y_0

การหาตัวพยากรณ์ไม่เอนเอียงเชิงเส้นดีที่สุด อาศัยเทคนิคเดียวกันกับที่ใช้หา BLUE โดยดำเนินการเป็น 2 ระยะ ดังนี้คือ

1. สมมติการรวมเชิงเส้น (Linear Combination) ใด ๆ ของ $Y_1, Y_2, ..., Y_n$ ขึ้นมาชุดหนึ่ง แล้วตรวจสอบดูว่าการรวมเชิงเส้นดังกล่าวเป็นตัวทำนายค่าที่ไม่เอนเอียง (Unbiased Predictor) ของ Y_0 หรือไม่ และถ้าเป็นจะต้องเป็นภายใต้เงื่อนไขใด

2. การรวมเชิงเส้นดังกล่าวเป็นตัวทำนายค่าที่ดีที่สุดของ Y หรือไม่ โดยพิจารณาจากค่าความแปรปรวนต่ำที่สุด และถ้าเป็นตัวทำนายค่าที่ดีที่สุดจะต้องเป็นภายใต้เงื่อนไขใด

ขั้นที่ 1 สมมติว่า $\mathbf{P} = c_1Y_1 + c_2Y_2 + ... + c_nY_n = \mathbf{C}'\mathbf{Y}$ เป็นตัวทำนายค่าที่ไม่เอนเอียงของ Y_0 โดยที่ $\mathbf{C} = (c_1, c_2, ..., c_n)$ เป็นเวกเตอร์ของค่าคงที่ใด ๆ ที่เราต้องการหาโครงสร้างที่ทำให้ $\mathbf{P}$ เป็น BLUP ของ Y_0

เมื่อกำหนดให้ $\mathbf{P} = \mathbf{C}'\mathbf{Y}$ เป็นตัวทำนายค่าที่ไม่เอนเอียงของ Y_0 ก็แสดงว่า $\mathrm{E}(\mathbf{P}) = \mathbf{Y}_0$ หรือ $\mathrm{E}(\mathbf{P} - \mathrm{Y}_0) = 0$ แต่เมื่อทดลองตรวจสอบ $\mathrm{E}(\mathrm{P} - \mathrm{Y}_0)$ กลับพบว่า

$$\begin{aligned}\mathrm{E}(\mathbf{P} - \mathbf{Y}_0) &= \mathrm{E}(\mathbf{C}'\mathbf{Y} - \mathrm{Y}_0) \\ &= \mathrm{E}[(\mathbf{C}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) - (\mathrm{X}_0\beta + \varepsilon_0)] \\ &= \mathrm{E}[(\mathbf{C}'\mathbf{X} - \mathrm{X}_0)\beta + \mathbf{C}'\boldsymbol{\varepsilon} - \varepsilon_0)] \\ &= (\mathbf{C}'\mathbf{X} - \mathrm{X}_0)\boldsymbol{\beta} \neq 0\end{aligned}$$

หรือ $$\mathrm{E}(\mathbf{P}) = \mathrm{Y}_0 + (\mathbf{C}'\mathbf{X} - \mathrm{X}_0)\boldsymbol{\beta} \neq \mathrm{Y}_0$$

แสดงว่า $\mathrm{E}(\mathbf{P} - \mathrm{Y}_0)$ จะเท่ากับ 0 หรือ $\mathbf{P}$ จะเป็นตัวทำนายที่ไม่เอนเอียงของ Y_0 ได้ก็ต่อเมื่อ $(\mathbf{C}'\mathbf{X} - \mathrm{X}_0)\boldsymbol{\beta} = 0$ หรือ $(\mathbf{C}'\mathbf{X} - \mathrm{X}_0) = 0$ เท่านั้น

นั่นคือ $(\mathbf{C}'\mathbf{X}-\mathrm{X}_0)=0$ คือเงื่อนไขที่ทำให้ $\mathbf{P}=\mathbf{C}'\mathbf{Y}$ เป็นตัวทำนายค่าที่ไม่เอนเอียงของ Y_0

ขั้นที่ 2 พิจารณา $\mathrm{V}(\mathbf{P})$ พบว่า

$$\mathrm{V}(\mathbf{P})=\mathrm{E}[\mathbf{P}-\mathrm{E}(\mathbf{P})]^2$$

$$=\mathrm{E}[\mathbf{P}-\ \{\mathrm{Y}_0+(\mathbf{C}'\mathbf{X}-\mathrm{X}_0)\boldsymbol{\beta}\}]^2$$

แต่ $(\mathbf{C}'\mathbf{X}-\mathrm{X}_0)=0$

ดังนั้น $\mathrm{V}(\mathbf{P})=\mathrm{E}(\mathbf{P}-\mathrm{Y}_0)^2$

$$=\mathrm{E}[(\mathbf{P}-\mathrm{Y}_0)(\mathbf{P}-\mathrm{Y}_0)']$$

พิจารณา $\mathbf{P}-\mathrm{Y}_o$ พบว่า

$$\mathbf{P}-\mathrm{Y}_0=\ \mathbf{C}'\mathbf{Y}-\mathrm{Y}_0$$

$$=\ \mathbf{C}'(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})-\mathbf{X_0}\boldsymbol{\beta}-\varepsilon_0$$

$$=\ (\mathbf{C}'\mathbf{X}-\mathbf{X}_0)\boldsymbol{\beta}+\mathbf{C}'\boldsymbol{\varepsilon}-\varepsilon_0$$

แต่ $(\mathbf{C}'\mathbf{X}-\mathbf{X}_0)=0$

ดังนั้น $\mathbf{P}-\mathrm{Y}_0=\ \mathbf{C}'\boldsymbol{\varepsilon}-\varepsilon_0$

นั่นคือ $\mathrm{V}(\mathbf{P})=\mathrm{E}[(\mathbf{C}'\boldsymbol{\varepsilon}-\varepsilon_0)(\mathbf{C}'\boldsymbol{\varepsilon}-\varepsilon_0)']$

$$=\mathrm{E}[(\mathbf{C}'\boldsymbol{\varepsilon}-\varepsilon_0)(\boldsymbol{\varepsilon}'\mathbf{C}-\varepsilon_0)]$$

$$=\mathrm{E}[(\mathbf{C}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{C}-\mathbf{C}'\boldsymbol{\varepsilon}\varepsilon_0-\varepsilon_0\boldsymbol{\varepsilon}'\mathbf{C}+\varepsilon_0^2]$$

$$=\mathrm{E}[(\mathbf{C}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{C}-2\mathbf{C}'\boldsymbol{\varepsilon}\varepsilon_0\ +\varepsilon_0^2]$$

$$=\mathbf{C}'\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\mathrm{C}-2\mathbf{C}'\,\mathrm{E}(\boldsymbol{\varepsilon}\varepsilon_0)+\mathrm{E}(\varepsilon_0^2)$$

$$=\mathbf{C}'\phi\mathbf{C}-2\mathbf{C}'\mathbf{W}+\sigma_0^2$$

นั่นคือ $\mathrm{V}(\mathbf{P})=\mathbf{C}'\phi\mathbf{C}-2\mathbf{C}'\mathbf{W}+\sigma_0^2$

ผลการพิสูจน์ทั้งสองขั้นตอนนี้ ชี้ให้เห็นว่า $\mathbf{P} = \mathbf{C}'\mathbf{Y}$ จะเป็นตัวทำนายค่าที่ไม่เอนเอียงได้ภายใต้เงื่อนไขใด และมีผลให้ $\mathrm{V}(\mathbf{P})$ มีรูปแบบอย่างไร แต่ยังไม่ได้หาค่า $\mathbf{C}$ ที่เหมาะสมตามเงื่อนไขคือทำให้ $\mathrm{V}(\mathbf{P})$ มีค่าต่ำสุด

ขั้นที่ 3 หาค่าที่เหมาะสมของ $\mathbf{C}$ โดยวิธีตัวคูณลากรองจ์ (Lagrange Multiplier Method)

จากเงื่อนไข $(\mathbf{C}'\mathbf{X} - \mathrm{X}_0) = 0$ และฟังก์ชันวัตถุประสงค์ คือ $\mathrm{V}(\mathbf{P}) = \mathbf{C}'\phi\mathbf{C} - 2\mathbf{C}'\mathbf{W} + \sigma_0^2$ จะพบว่า

$\mathrm{F} = \mathrm{V}(\mathbf{P}) - 2(\mathbf{C}'\mathbf{X} - \mathrm{X}_0)\,\boldsymbol{\lambda}$ เมื่อ $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, ..., \lambda_p)'$ คือ เวกเตอร์ตัวคูณลากรองจ์ (Lagrange Multipliervector)

$$= \mathbf{C}'\phi\mathbf{C} - 2\mathbf{C}'\mathbf{W} + \sigma_0^2 - 2(\mathbf{C}'\mathbf{X} - \mathrm{X}_0)\boldsymbol{\lambda}$$

ค่าที่เหมาะสมของ $\mathrm{C} = (c_1, c_2, ..., c_n)'$ คือ ค่าที่ได้จากระบบสมการ

$$\frac{\partial \mathrm{F}}{\partial \mathbf{C}_{\mathrm{i}}} = 0\ ;\quad \text{i=1,2,...,n} \text{ และ } \frac{\partial \mathrm{F}}{\partial \boldsymbol{\lambda}_{\mathrm{i}}} = 0\ ;\quad \text{i=0,2,...,p}$$

หรือ $\frac{\partial \mathrm{F}}{\partial \mathbf{C}_{\mathrm{i}}} = 0$, $\frac{\partial \mathrm{F}}{\partial \boldsymbol{\lambda}_{\mathrm{i}}} = 0$ เมื่อ $\mathbf{C} = (c_1, c_2, ..., c_n)'$, $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, ..., \lambda_p)'$

และพบว่า

$$\frac{\partial \mathrm{F}}{\partial \mathbf{C}_{\mathrm{i}}} = 0 \Rightarrow 2\phi\mathbf{C} - 2\mathbf{W} - 2\mathbf{X}\boldsymbol{\lambda} = 0 \qquad \text{.......................... (1)}$$

$$\frac{\partial \mathrm{F}}{\partial \boldsymbol{\lambda}_{\mathrm{i}}} = 0 \Rightarrow \mathbf{C}'\mathbf{X} - \mathbf{X}_0 = 0 \qquad \text{.......................... (2)}$$

จาก (1) พบว่า $\phi\mathbf{C} - \mathbf{X}\boldsymbol{\lambda} = \mathbf{W}$ .... (3)

จาก (2) พบว่า $\mathbf{X}'\mathbf{C} - \mathbf{X}_0' = 0$

หรือ $\mathbf{X}'\mathbf{C} - 0\boldsymbol{\lambda} = \mathbf{X}_0'$ ....(4)

จาก (3) และ (4) จัดรูปเมทริกซ์ ได้ดังนี้ คือ

$$\begin{bmatrix} \phi & \mathrm{X} \\ \mathrm{X} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{C} \\ -\lambda \end{bmatrix} = \begin{bmatrix} \mathrm{W} \\ \mathrm{X}_o' \end{bmatrix} \qquad \text{.......................... (5)}$$

โดยอาศัยความรู้เรื่องการผกผันของเมทริกซ์แบ่งส่วน (Inverse of Partitioned Matrix) พบว่า

$$\hat{\mathrm{C}} = \phi^{-1}[\mathrm{I}_n - \mathbf{X}(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}'\phi^{-1}]\mathbf{W} + \phi^{-1}\mathbf{X}(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}_o'$$

คือค่าของเวกเตอร์ $\mathbf{C}$ ที่เหมาะสมที่มีผลให้ $\mathbf{P} = \mathbf{C}'\mathbf{Y}$ เป็น BLUP

นั่นคือ $\hat{\mathbf{P}} = \hat{\mathbf{C}}'\mathbf{Y}$

$$= [\ \phi^{-1}\{\mathrm{I}_n - \mathbf{X}(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}'\phi^{-1}\}\mathbf{W}]'\mathbf{Y} + \{\phi^{-1}\mathbf{X}(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}_o'\}'\mathbf{Y}$$

$$= [\ \mathbf{W}'\{\mathrm{I}_n - \phi^{-1}\mathbf{X}(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}'\}\phi^{-1}]\mathbf{Y} + \{\mathbf{X}_0(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}'\phi^{-1}\}\mathbf{Y}$$

$$= \mathbf{W}'\phi^{-1}\mathbf{Y} - \mathbf{W}'\phi^{-1}\mathbf{X}(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}'\phi^{-1}\mathbf{Y} + \mathbf{X}_0(\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}'\phi^{-1}\mathbf{Y}$$

แต่ $\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}'\phi^{-1}\mathbf{X})^{-1}\mathbf{X}\phi^{-1}\mathbf{Y}$

ดังนั้น $\hat{\mathbf{P}} = \mathbf{W}'\phi^{-1}\mathbf{Y} - \mathbf{W}'\phi^{-1}\mathbf{X}\hat{\boldsymbol{\beta}}_{GLS} + \mathbf{X}_0\hat{\boldsymbol{\beta}}_{GLS}$

$$= \mathbf{X}_0\hat{\boldsymbol{\beta}}_{GLS} + \mathbf{W}'\phi^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{GLS})$$

$$= \mathbf{X}_0\hat{\boldsymbol{\beta}}_{GLS} + \mathbf{W}'\phi^{-1}\hat{\boldsymbol{\varepsilon}}$$ โดยที่ $\hat{\boldsymbol{\varepsilon}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{GLS}$

นั่นคือ BLUP ของ Y_0 คือ $\hat{\mathbf{P}}$ โดยที่ $\hat{\mathbf{P}} = \mathbf{X_0}\hat{\boldsymbol{\beta}}_{GLS} + \boldsymbol{\Omega}'\phi^{-1}\hat{\boldsymbol{\varepsilon}}$ ซึ่งเป็นรูปที่ใช้ได้กับทุกกรณี

4.8.2 การพยากรณ์ในกรณีมีอัตตสหสัมพันธ์ในตัวเองในรูปแบบ AR(1)

จากกรณีการเกิดอัตตสหสัมพันธ์ในตัวเองในรูปแบบ AR(1) เราทราบว่า

$$\mathrm{E}(\varepsilon_t\varepsilon_{t-s}) = \begin{cases} \phi^s\sigma_\varepsilon^2 & ;\ \mathrm{s} \neq \mathrm{t} \\ \sigma_\varepsilon^2 = \dfrac{\sigma_a^2}{1-\phi^2} & ;\ \mathrm{s} = \mathrm{t} \end{cases}$$ หรือสามารถเสนอ $\mathrm{V}(\boldsymbol{\varepsilon})$ ได้ดังนี้

$$\mathrm{V}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \phi^2 & \cdots & \phi^{n-2} \\ \phi^2 & \phi & 1 & \phi & \cdots & \phi^{n-3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \phi^{n-4} & \cdots & 1 \end{bmatrix}$$

พิจารณาสูตร $\hat{\mathrm{Y}}_0 = \mathbf{X_0}\hat{\boldsymbol{\beta}}_{GLS} + \mathbf{W}'\phi^{-1}\hat{\boldsymbol{\varepsilon}}$ สมมติว่าเราต้องการพยากรณ์ค่า Y ใน period ที่ 1 นอก Sample Period จะพบว่า

$$\hat{\mathrm{Y}}_{n+1} = \mathbf{X}_{n+1}\hat{\boldsymbol{\beta}}_{GLS} + \mathbf{W}'\phi^{-1}\hat{\boldsymbol{\varepsilon}}$$

โดยในที่นี้จะพบว่า

$$\mathbf{X}_{n+1} = (1, X_{n+1,1}, X_{n+1,2}, ..., X_{n+1,p}),\ \ \hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_p)'$$

$$\mathbf{X}_{n+1}\hat{\boldsymbol{\beta}}_{GLS} = \hat{\beta}_0 + \hat{\beta}_1 X_{n+1,1} + \hat{\beta}_2 X_{n+1,2} + ... + \hat{\beta}_p X_{n+1,p}$$

$$\mathbf{W} = \mathrm{E}\begin{bmatrix} \varepsilon_1\varepsilon_{n+1} \\ \varepsilon_2\varepsilon_{n+1} \\ \varepsilon_3\varepsilon_{n+1} \\ \vdots \\ \varepsilon_n\varepsilon_{n+1} \end{bmatrix} = \mathrm{E}\begin{bmatrix} \varepsilon_{n+1}\varepsilon_{n+1-n} \\ \varepsilon_{n+1}\varepsilon_{n+1-(n-1)} \\ \varepsilon_{n+1}\varepsilon_{n+1-(n-2)} \\ \vdots \\ \varepsilon_{n+1}\varepsilon_{n+1-1} \end{bmatrix} = \begin{bmatrix} \phi^n\sigma_\varepsilon^2 \\ \phi^{n-1}\sigma_\varepsilon^2 \\ \phi^{n-2}\sigma_\varepsilon^2 \\ \vdots \\ \phi\sigma_\varepsilon^2 \end{bmatrix} = \phi\sigma_\varepsilon^2\begin{bmatrix} \phi^{n-1} \\ \phi^{n-2} \\ \phi^{n-3} \\ \vdots \\ \phi \\ 1 \end{bmatrix}$$

แสดงว่า $\mathbf{W}$ สำหรับการพยากรณ์ในอนาคต 1 วาระ มีค่าเท่ากับ ϕ ของสดมภ์สุดท้ายของเมทริกซ์ $\mathrm{V}(\boldsymbol{\varepsilon})$ และสามารถพิสูจน์ได้ว่า $\mathbf{W}$ สำหรับการพยากรณ์ไปในอนาคต r วาระ มีค่าเท่ากับ ϕ^r เท่ากับสดมภ์สุดท้ายของ $\mathrm{V}(\boldsymbol{\varepsilon})$ และพบว่าในการพยากรณ์ Y_{n+1} นั้น $\mathbf{W}'\phi^{-1}\hat{\boldsymbol{\varepsilon}}$ คือ

$$\phi\sigma_\varepsilon^2(\phi^{n-1},\phi^{n-2},\phi^{n-3},...,\phi,1)\frac{1}{\sigma_\varepsilon^2}\frac{1}{1-\phi^2}\begin{bmatrix} 1 & -\phi & 0 & \cdots & 0 & 0 & 0 \\ -\phi & 1+\phi^2 & \phi & \cdots & 0 & 0 & 0 \\ 0 & -\phi & 1+\phi^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\phi & 1+\phi^2 & -\phi \\ 0 & 0 & 0 & \cdots & 0 & -\phi & 1 \end{bmatrix}\begin{bmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \hat{\varepsilon}_3 \\ \vdots \\ \hat{\varepsilon}_n \end{bmatrix}$$

$$= \frac{\phi\sigma_\varepsilon^2}{(1-\phi^2)\sigma_\varepsilon^2}(0,0,0,...,1-\phi^2)\begin{bmatrix} \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \hat{\varepsilon}_3 \\ \vdots \\ \hat{\varepsilon}_n \end{bmatrix}$$

$$= \frac{\phi}{(1-\phi^2)}(1-\phi^2)\hat{\varepsilon}_n$$

$$= \phi\hat{\varepsilon}_n$$

ดังนั้น $\hat{\mathrm{Y}}_{n+1} = \mathbf{X}_{n+1}\hat{\boldsymbol{\beta}}_{GLS} + \phi\hat{\varepsilon}_n$ และในกรณีที่ไม่ทราบค่า ϕ ให้ใช้ค่าของ $\hat{\phi}$ แทน

ดังนั้น $\hat{\mathrm{Y}}_{n+1} = \mathbf{X}_{n+1}\hat{\boldsymbol{\beta}}_{GLS} + \phi\hat{\varepsilon}_n$ คือ BLUP ของ Y_{n+1}

หรือ $\hat{\mathrm{Y}}_{n+1} = \hat{\beta}_0 + \hat{\beta}_1 X_{n+1,1} + \hat{\beta}_2 X_{n+1,2} + ... + \hat{\beta}_k X_{n+1,k} + \phi\hat{\varepsilon}_n$

สำหรับการพยากรณ์ค่าของ Y ในวาระใด ๆ สมมติเป็นวาระที่ r คือ $\hat{\mathrm{Y}}_{n+r}$ นอก sample Period คือ

$$\hat{\mathrm{Y}}_{n+r} = \mathbf{X}_{n+r}\hat{\boldsymbol{\beta}}_{GLS} + \phi^r\hat{\varepsilon}_n \quad ; \mathrm{r} = 1,2,3,...$$ คือ BLUP ของ Y_{n+r}

ตอนที่ 2 การจำลองสถานการณ์ด้วยเทคนิคมอนติคาร์โล

1. ความหมายของเทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โล เป็นส่วนหนึ่งของกระบวนการของชุดคำสั่งในการคำนวณที่ทำซ้ำ ๆ ของชุดตัวเลขเชิงสุ่ม เพื่อหาผลลัพธ์ของชุดตัวเลขนั้น ๆ เทคนิคมอนติคาร์โลมักใช้ในงานที่มีลักษณะการจำลองทางกายภาพ และทางคณิตศาสตร์ และเป็นการใช้ตัวเลขที่สุ่มมาด้วยเครื่องคอมพิวเตอร์ ซึ่งเรียกว่าตัวเลขสุ่มเทียม (Pseudo–Random Number)

เทคนิคมอนติคาร์โล เป็นรูปแบบการสร้างค่าเชิงสุ่มเพื่อการศึกษาในเชิงสถิติ ซึ่งเป็นค่าเชิงสุ่มที่มักจะได้จากการใช้เครื่องคอมพิวเตอร์ในการสร้างขึ้นมา เป็นเทคนิคที่เกี่ยวข้องกับชุดข้อมูลที่สร้างขึ้นมาหลากหลายชุดเพื่อใช้ในการศึกษาคุณลักษณะของตัวประมาณ เช่น ในกรณีที่นักสถิติที่ต้องการพัฒนาทฤษฎีใหม่ ๆ จำเป็นที่จะต้องทดสอบทฤษฎีนั้น ๆ ด้วยข้อมูล ซึ่งแทนที่นักสถิติจะเก็บข้อมูลจากข้อมูลที่เกิดขึ้นในสถานการณ์จริงมาเพื่อทดสอบ และต่อมาสถานการณ์นั้น ๆ หรือสภาพแวดล้อมนั้นอาจจะเปลี่ยนไปในเวลาต่อมา นอกจากนั้นการเก็บข้อมูลจริงจำนวนมาก ๆ เพื่อทดสอบทฤษฎีจำเป็นต้องใช้ต้นทุนสูงมาก ดังนั้นเพื่อทดสอบทฤษฎีจึงใช้วิธีการสร้างชุดข้อมูลตามสถานการณ์นั้น ๆ โดยใช้เทคนิคมอนติคาร์โล

2. ความเป็นมาของเทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โล เป็นสาขาหนึ่งของคณิตศาสตร์ที่ใช้คอมพิวเตอร์ช่วยในการจำลองสถานการณ์ (Simulation) โดยอาศัยตัวเลขสุ่ม (Random Number) มาสร้างตัวแปรให้เหมือนกับสถานการณ์จริงและมีการทดลองซ้ำหลาย ๆ ครั้งเพื่อให้ค่าที่แน่นอนที่จะใช้เป็นข้อสรุปหรืออธิบายปรากฏการณ์ต่าง ๆ ในสถานการณ์จริง หรือช่วยหาคำตอบในเรื่องราวต่าง ๆ ที่ยังไม่แน่ใจในผลที่เกิดขึ้น

เทคนิคมอนติคาร์โล ได้มีการใช้มานานแล้ว เพียงแต่ไม่ได้เรียกว่า มอนติคาร์โล ต่อมาในราวปี ค.ศ. 1753 จอร์ส หลุยส์ เลคเลอร์ (Georges Lois Lectere & Comte de Buffon) ได้นำมาพัฒนาในทฤษฎีความน่าจะเป็นโดยการทดลองหาค่า π โดยการโยนเข็มที่มีความยาว k หน่วยอย่างสุ่มลงบนพื้นราบที่มีเส้นขนานอยู่ โดยให้ระยะห่างระหว่างเส้นขนานแต่ละเส้นห่างกัน d หน่วย และกำหนดให้ d มากกว่า k จะได้ความน่าจะเป็นที่เข็มจะตัดเส้นขนาน $P=2k/\pi d$ ซึ่งถ้าความน่าจะเป็นเป็นค่าสุ่มก็จะหาค่า π ได้ ต่อมาในปี ค.ศ. 1908 กอสเซท (Gooset) ได้ศึกษาการแจกแจงความถี่ของความสูงของนักโทษอาชญากรรมจำนวน 3,000 คน โดยเทียบกับการแจกแจงความถี่ของกลุ่มตัวอย่างที่สุ่มมาครั้งละ 4 คน จำนวน 750 กลุ่มตัวอย่าง ผลการศึกษาพบว่า การแจกแจงความถี่ทั้งสองลักษณะเหมือนกัน กอสเซทได้ตั้งชื่อการแจกแจงความถี่ที่ค้นพบนี้ว่า การแจกแจงที (t–distribution) ซึ่งถือว่าเป็นจุดเริ่มต้นของเทคนิคมอนติคาร์โล (Monte Carlo Technique)

เทคนิคมอนติคาร์โล ได้รับการพัฒนาอย่างจริงจังในราวปี ค.ศ. 1944 ช่วงสงครามโลกครั้งที่ 2 อูลาม และออน นิวแมน (Ulam & Von Neumann) เป็นผู้ตั้งชื่อ มอนติ คาร์โล ซึ่งเป็นรหัสลับของงานที่ทำในลอส อาลามาส (Los Alamas) และได้นำเทคนิคนี้มาหาผลของการแพร่อย่างสุ่มของนิวตรอนในวัสดุเชื้อเพลิงที่เป็นการทดลองทางคณิตศาสตร์เพื่อหาผลของคำตอบก่อนที่จะทำการทดลองจริงซึ่งเป็นการหลีกเหลี่ยงอันตรายและช่วยประหยัดค่าใช้จ่ายก่อนที่จะทดลองจริง

หลังจากนั้นเทคนิคมอนติคาร์โล จึงมีการนำมาใช้อย่างกว้างขวางทั้งทางด้านฟิสิกส์ คณิตศาสตร์ สถิติ และการวิจัย นับได้ว่าเทคนิคมอนติคาร์โล มีประโยชน์อย่างมากในการขยายความรู้เชิงทฤษฎี

3. ขั้นตอนของเทคนิคมอนติคาร์โล

หลักการที่สำคัญของเทคนิคมอนติคาร์โล คือ การนำเอาตัวเลขสุ่มมาประยุกต์แก้ปัญหาต่าง ๆ มีขั้นตอนที่สำคัญ ดังนี้ (ศิริรัตน์ วงศ์ประกรณ์กุล, 2539, หน้า 240–246)

3.1 สร้างตัวเลขสุ่ม (Generate Random Number) ระยะแรกทำได้โดยอาศัยเครื่องมือทางกายภาพ เช่น ล้อรูเล็ต ลูกเต๋า ไพ่ กระดาษเขียนหมายเลข เป็นต้น แต่ได้ตัวเลขสุ่มไม่มาก ต่อมาจึงหันมาใช้เครื่องมืออิเล็กทรอนิกส์ เช่น เครื่องสร้างตัวเลขสุ่มที่สร้างขึ้นโดยบริษัท แรนด์ (Reand) โดยสร้างตัวเลขสุ่มจากเครื่องกำเนิดพัลส์ (Pulse) ที่สามารถสร้างตัวเลขสุ่มได้หนึ่งล้านตัว

การสร้างหรือเลือกตัวเลขสุ่มดังกล่าวกับเครื่องคอมพิวเตอร์ยังมีปัญหา 2 ประการ คือ เป็นการยากที่จะทำให้เครื่องคอมพิวเตอร์สามารถเรียกใช้ได้เมื่อมีความต้องการ และยากที่จะทำให้เครื่องมือดังกล่าวสร้างตัวเลขสุ่มชุดเดิม เมื่อต้องการเปรียบเทียบวิธีการต่าง ๆ ภายใต้เงื่อนไขของระบบเลขสุ่มชุดเดียวกัน หรือถ้าจะเก็บเลขสุ่มเหล่านี้ไว้ในหน่วยความจำหรือจานแม่เหล็กก็จะทำให้สูญเสียหน่วยความจำหรือเสียเวลาในการค้นหา ฉะนั้นการสร้างเลขสุ่มในคอมพิวเตอร์จึงนิยมสร้างตัวเลขสุ่มเทียมโดยอาศัยสูตรทางคณิตศาสตร์

ในการผลิตเลขสุ่มชุดตัวเลขที่ผลิตขึ้นมาจะต้องมีคุณสมบัติทางสถิติที่สำคัญ 2 ประการ คือ ความสม่ำเสมอ (Uniformity) และความเป็นอิสระ (Independence) ดังนั้นตัวเลขสุ่มแต่ละตัวจะต้องถูกเลือกอย่างอิสระ และมีการแจกแจงสม่ำเสมออยู่ในช่วง 0 ถึง 1 ซึ่งวิธีการผลิตตัวเลขสุ่มแบบวิธีสมภาคเชิงเส้น (Linear Congruential Method) จะผลิตชุดตัวเลขสุ่มจำนวนเต็ม x_1, x_2,... มีค่าอยู่ระหว่าง 0 ถึง m–1 จากสมการตัวผลิต x_i=(ax_{i-1} + c) mod M; I = 1, 2,... ตัวเลขจำนวนเต็ม x_1, x_2, ... จะมีการแจกแจงสม่ำเสมอในช่วง 0 ถึง M–1 (U (0, M–1)) ดังนั้นตัวเลขสุ่ม R_1, R_2, ... ที่มีการแจกแจงสม่ำเสมอในช่วง 0 ถึง 1 (U(0,1)) จะผลิตได้จากสมการ $R_i = x_i / M$; i=1, 2, ...

กำหนดให้ a เป็นค่าคงที่ c เป็นค่าส่วนที่เพิ่ม (Increment) x_0 เป็นตัวเลขนำ M เป็นตัวคูณจำนวนคงที่ (Modulus) และ mod หมายความว่า (ax_{i-1} + c) หารด้วย M จนกระทั่งเหลือเศษน้อยกว่าค่า M เลขที่เหลือ จึงเป็นเลขสุ่มของตัวถัดไปคือ x_i

ถ้ากำหนดค่า $c \neq 0$ จะเรียกตัวผลิตนี้ว่าวิธีสมภาคแบบผสม (Mixed Congruential Method) แต่ถ้ากำหนด c=0 เรียกตัวผลิตนี้ว่าวิธีสมภาคแบบผลคูณ (Multiplicative Congruential Method) การกำหนดค่า c, a, M และ x_0 มีความสำคัญมากเนื่องจากค่าเหล่านี้มีผลโดยตรงต่อค่าทางสถิติและความยาวของชุดตัวเลขสุ่ม และจะทำให้ตัวเลขสุ่มแต่ละตัวมีความสม่ำเสมอ มีความเป็นอิสระและความเป็นวัฏจักรยาว

ตั้งแต่ยุคเริ่มแรกจนกระทั่งปัจจุบันภาษาคอมพิวเตอร์ที่ใช้เพื่อสร้างตัวเลขสุ่มในเครื่องคอมพิวเตอร์มีหลายภาษา เช่น ภาษาเบสิค (Basic) ภาษาฟอร์แทน (FORTRAN) ภาษาฟอกโปร (FoxPro) เป็นต้น นอกจากนั้นในปัจจุบันนี้มีความเจริญก้าวหน้าทางด้านเทคโนโลยีที่ได้เกิดขึ้นอย่างรวดเร็ว จนกระทั่งมีนักการศึกษาหลายท่านได้สร้างโปรแกรมสำเร็จรูป เพื่อสร้างตารางเลขสุ่มขึ้นมาอย่างมากมาย โดยแต่ละโปรแกรมที่สร้างขึ้นจะตอบรับการใช้งานที่คล้ายกันบ้างและแตกต่าง

กันบ้าง เช่น SAS, baseSim, C++Sim, JavaSim, Simulacion4, SimlAr ฯลฯ ซึ่งโปรแกรมสำเร็จรูปต่าง ๆ ได้พยายามสร้างชุดตัวเลขสุ่มให้มีคุณสมบัติ ดังนี้

1. ตัวเลขสุ่มจะต้องมีลักษณะการกระจายความน่าจะเป็นแบบสม่ำเสมอ
2. ตัวเลขสุ่มที่ได้ต้องเป็นอิสระกัน
3. อนุกรมของตัวเลขสุ่มที่ได้ต้องสามารถสร้างซ้ำเดิมได้
4. อนุกรมตัวเลขสุ่มที่ได้ต้องไม่ซ้ำเดิมในช่วงที่ต้องการใช้ตัวเลขสุ่ม
5. ต้องใช้เวลาน้อยในการสร้างตัวเลขสุ่ม
6. ต้องใช้หน่วยความจำในคอมพิวเตอร์น้อย

3.2 นำตัวเลขสุ่มมาประยุกต์ใช้กับการแก้ปัญหาต่าง ๆ เป็นการนำตัวเลขสุ่มไปสร้างตัวแปรตามลักษณะการแจกแจงของปัญหาที่จะศึกษาเพื่อเป็นข้อมูลของปัญหานั้น เช่น การสร้างตัวเลขสุ่มแล้วนำตัวเลขสุ่มไปสร้างเป็นคะแนนการสอบของผู้เรียน ซึ่งบางตัวแปรของปัญหาที่จะศึกษาไม่ได้สร้างจากตัวเลขสุ่มโดยตรงแต่อาจใช้ตัวเลขสุ่มเป็นพื้นฐานได้

3.3 ทำการทดลองซ้ำ ๆ หลายๆ ครั้ง การศึกษาด้วยเทคนิคมอนติคาร์โล ต้องมีการทดลองซ้ำหลายๆ ครั้งเพื่อลดความคลาดเคลื่อนของคำตอบที่จะได้ และสามารถสรุปเป็นความน่าจะเป็นของการเกิดเหตุการณ์ในปัญหานั้นๆ

3.4 จุดเด่นของเทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โล เป็นการใช้ตัวเลขสุ่มเป็นพื้นฐานการสร้างตัวแปรของปัญหาโดยอาศัยทฤษฎี สูตร หรือกฎเกณฑ์ต่าง ๆ ที่มีอยู่ และมีการทดลองซ้ำหลาย ๆ ครั้ง เพื่อลดความคลาดเคลื่อนต่าง ๆ ซึ่งนับว่ามีประโยชน์ที่สำคัญ ดังนี้

3.4.1 สามารถควบคุมตัวแปรแทรกซ้อน และสามารถสังเกตได้อย่างสมบรูณ์ และทำการทดลองซ้ำในสภาพแวดล้อมเดิมหลาย ๆ ครั้งได้ ส่วนในการทดลองจริงนั้นไม่สามารถทำได้เนื่องจากเมื่อเวลาเปลี่ยนไปสภาพแวดล้อมก็เปลี่ยนไปด้วย

3.4.2 ถ้ามีสูตรหรือกฎเกณฑ์ต่าง ๆ ที่ถูกต้องรองรับการสร้างตัวแปรของปัญหาในการทดลองแล้วจะให้ผลที่ถูกต้องแม่นยำกว่าเมื่อใช้ทดลองในสถานการณ์จริงเพราะสามารถลดตัวแปรแทรกซ้อนได้

3.4.3 สิ้นเปลืองเวลา แรงงาน และค่าใช้จ่ายน้อยกว่า เมื่อเทียบกับการทดลองในสถานการณ์จริง

ตอนที่ 3 ความรู้เกี่ยวกับยางพารา

1. ความเป็นมาของยางธรรมชาติในประเทศไทย

ยางพาราเป็นยางที่ได้จากต้นไม้ชนิดหนึ่งเรียกว่า ต้นยางพารา (Para Rubber) หรือเรียกตามภาษาพฤกษศาสตร์ว่า Heavea Brasiliensis เป็นเพราะว่าเมื่อประมาณ 100 ปีมาแล้ว ยางชนิดนี้จะทำการซื้อขายกันเฉพาะที่เมืองพารา ประเทศบราซิลเพียงแห่งเดียวเท่านั้น ถิ่นกำเนิดของต้นยางพาราอยู่ที่ทวีปอเมริกาใต้ ในป่าเขตร้อน ซึ่งเป็นพื้นที่ที่มีฝนตกชุกแถบอเมริกาใต้ ได้แก่ประเทศบราซิล เปรู โคลัมเบีย เวเนซุเอลา และโบลิเวีย ชาวพื้นเมืองคือชาวอินเดียนแดงในอเมริกาใต้

และอเมริกากลางรู้จักยางพารามานานแล้ว และได้นำมาใช้ประโยชน์กันมาหลายร้อยปีก่อนที่ชาวยุโรปจะพบโลกใหม่หรือทวีปอเมริกา ซึ่งเป็นถิ่นเดิมของต้นยางพารา ชาวอินเดียนแดงได้ใช้ยางทำลูกบอล ผ้ากันฝน ถุงเก็บน้ำปากแคบ เป็นต้น ในปี พ.ศ. 2520 Sir Henry Wickhan ได้นำเมล็ดยางพารามาปลูกที่เปรัค ประเทศมาเลเซีย จำนวน 9 ต้น และอีก 13 ต้น ปลูกที่สวนพฤกษาชาติ ประเทศสิงคโปร์ ซึ่งเป็นที่มาของต้นยางพาราในเอเชียและแอฟริกาในทุกวันนี้

การปลูกต้นยางพาราในประเทศไทยไม่มีบันทึกที่แน่นอน สันนิษฐานว่าราวปี พ.ศ. 2442-2444 (กลางรัชกาลที่ 5) พระยารัษฎานุประดิษฐ์ มหิศรภักดี (คอซิมบี้ ณ ระนอง) เจ้าเมืองตรัง ได้นำพันธุ์ยางพารามาจากรัฐเปรัค ประเทศมาเลเซีย มาทดลองปลูกครั้งแรกที่อำเภอกันตัง จังหวัดตรัง ต่อมามีราษฎรนำมาปลูกเป็นสวนยางในแถบจังหวัดตรังและนราธิวาส หลังจากนั้นจึงปลูกกันแพร่หลายไปทั่วทั้ง 14 จังหวัดของภาคใต้ ในจังหวัดจันทบุรี หลวงราชไมตรี (ปูม ปุณศรี) เป็นผู้นำไปปลูกเมื่อประมาณปี พ.ศ. 2454 นับเป็นการแพร่กระจายยางพาราสู่ภาคตะวันออก และสำหรับในภาคเหนือและภาคตะวันออกเฉียงเหนือนั้น กองการยาง กรมวิชาการเกษตร ได้ทดลองปลูกเป็นผลสำเร็จในระหว่างปี พ.ศ. 2521- 2523 และขยายไปสู่ภาคตะวันออกและภาคตะวันออกเฉียงเหนือ (เสาวณีย์ ก่อวุฒิกุลรังสี, 2547)

2. ลักษณะทางพฤกษศาสตร์ของต้นยางพารา

ยางพาราเป็นพืชยืนต้นขนาดใหญ่ มีอายุยืนยาวหลายปี ปัจจุบันมีต้นยางพาราที่มีอายุมากในลุ่มแม่น้ำอเมซอนเป็นจำนวนมาก ต้นที่เจริญเติบโตในประเทศบราซิลและในประเทศข้างเคียง ลำต้นวัดโดยรอบได้กว่า 3–5 เมตร ถ้าเป็นต้นที่สมบูรณ์และอยู่ในที่ระบายน้ำได้ดีจะมีความสูงถึง 40 เมตร แต่ต้นที่ปลูกในทวีปเอเชียจะเล็กลงมาก โดยลำต้นของต้นที่ปลูกด้วยเมล็ดจะโตประมาณ 1–2 เมตร และถ้าเป็นต้นติดตาลำต้นจะโตไม่เกิน 1 เมตร ส่วนความสูงก็ลดลงเหลือประมาณ 15–20 เมตรเท่านั้น ต้นยางมีเปลือกที่น้ำยางจะไหลออกได้หนาประมาณ 6.5–15 มิลลิเมตร ทรงต้นที่สมบูรณ์มักจะสูงชะลูด กิ่งแยกมักแยกตั้งขึ้นไปประมาณ 45 องศาจากลำต้น ใบมักจะรวมเป็นพุ่มที่ส่วนปลายของกิ่ง แต่ละก้านใบแยกออกเป็น 3 ใบ แต่ละใบใน 3 ใบ กว้างประมาณ 5–10 เซนติเมตร และยาวประมาณ 10–20 เซนติเมตร ในทางพฤกษศาสตร์ได้จัดให้ต้นยางพาราอยู่ในวงศ์ยูฟอร์เบียซีอี (Family Euphorb iaceae) ในสกุลฮีเวีย (Genus Hevea) ชนิดบราซิไลเอ็นซิส (Species brasiliensis) ต้นยางฮีเวียมีประมาณ 20 ชนิด แต่ปรากฏว่าฮีเวียบราซิไลเอ็นซิส (Hevea brasiliensis) เป็นชนิดที่ให้น้ำยางมากที่สุด และเนื้อยางก็มีคุณสมบัติทางวิทยาศาสตร์ดีกว่ายางชนิดอื่นๆ จึงปลูกกันแต่พันธุ์ฮีเวียบราซิไล เอ็นซิสเท่านั้น (เสาวณีย์ ก่อวุฒิกุลรังสี, 2547)

3. สภาพการทำสวนยางพาราของประเทศไทย

ในช่วงแรก ๆ ประมาณปี พ.ศ. 2443–2503 นั้นการปลูกยางพาราจะเป็นพันธุ์พื้นเมือง โดยปลูกร่วมกันไม้ผลและพืชผักชนิดอื่น ๆ ในบริเวณเชิงเขาและที่ราบเชิงเขา ซึ่งเรียกว่า “ป่ายาง” โดยเมล็ดพันธุ์จะนำมาจากประเทศมาเลเซีย ซึ่งขณะนั้นเป็นประเทศผู้ผลิตยางธรรมชาติเป็นอันดับ 1 ของโลก ต่อมาประเทศไทยได้มีการวิจัยและพัฒนาพันธุ์ให้มีผลผลิตต่อไร่สูง การทำสวนยางพารามีการเปลี่ยนแปลงรูปแบบจาก “ป่ายาง” ไปเป็นพืชเชิงเดียวที่มีเพียงต้นยางพาราอย่างเดียว โดยปลูกทดแทนพันธุ์พื้นเมืองเดิมด้วยยางพันธุ์ดีที่ให้ผลผลิตต่อไร่สูงกว่า และมีการบำรุงรักษามากขึ้น

มีการพัฒนาการทำสวนยางพาราให้เป็นพืชเศรษฐกิจที่สำคัญ โดยจัดตั้งหน่วยงานที่เกี่ยวข้องกับการทำสวนยางพารา เพื่อดำเนินการวิจัย ส่งเสริมและพัฒนา ทำให้พื้นที่ปลูกยางพาราขยายตัวอย่างต่อเนื่อง ผลผลิตยางพาราของไทยในช่วงต่อมาจึงเพิ่มขึ้นอย่างรวดเร็ว จนประเทศไทยก้าวขึ้นมาเป็นผู้ผลิตยางธรรมชาติอันดับหนึ่งของโลกตั้งแต่ปี พ.ศ. 2534

4. การผลิต การส่งออก และการใช้ยางธรรมชาติในประเทศต่าง ๆ

การผลิตของประเทศต่าง ๆ

ปริมาณการผลิตยางธรรมชาติของโลกในปี พ.ศ. 2554 ประเทศผู้ผลิตยางธรรมชาติมากที่สุด 3 อันดับแรก คือ ไทย อินโดนีเซีย และมาเลเซีย โดยมีปริมาณการผลิตรวม 7.483 ล้านตัน คิดเป็นร้อยละ 70.20 ของปริมาณการผลิตยางธรรมชาติทั้งหมดของโลก ประกอบด้วย ไทย 3.573 ล้านตัน คิดเป็นร้อยละ 33.52 อินโดนีเซีย 2.886 ล้านตัน คิดเป็นร้อยละ 27.08 และมาเลเซีย 1.024 ล้านตัน คิดเป็นร้อยละ 9.61 ส่วนประเทศอื่น ๆ ผลิตยางธรรมชาติได้น้อยกว่า 1 ล้านตันต่อปี ซึ่งแสดงดังตารางที่ 1

ตารางที่ 1 ปริมาณการผลิตยางธรรมชาติของประเทศต่าง ๆ ในปี พ.ศ. 2550–2554

หน่วย: 1,000 ตัน

ประเทศ	2550	2551	2552	2553	2554
1. ไทย	3,056	3,090	3,164	3,252	3,573
2. อินโดนีเซีย	2,755	2,751	2,440	2,736	2,886
3. มาเลเซีย	1,200	1,072	856	939	1,024
4. อินเดีย	811	881	820	851	885
5. เวียดนาม	606	660	711	755	776
6. จีน	590	560	644	665	685
7. โกตดิวัวร์	183	194	203	227	227
8. ศรีลังกา	118	129	137	153	160
9. ไลบีเรีย	121	84	60	62	76
10. ฟิลิปปินส์	101	103	98	99	101
11. อื่นๆ	349	604	557	662	266
รวม	9,890	10,128	9,690	10,401	10,659
อัตราการเปลี่ยนแปลง (ร้อยละ)	0.64	2.41	-4.32	7.34	2.48

ที่มา: Internatinal Rubber Study Group (2011)

การส่งออกของประเทศต่าง ๆ

ปริมาณการส่งออกยางธรรมชาติของโลกในปี พ.ศ. 2554 ประเทศผู้ส่งออกยางธรรมชาติมากที่สุด 4 อันดับแรกเป็นประเทศผู้ผลิตยางธรรมชาติรายใหญ่แถบอาเซียน คือ ไทย อินโดนีเซีย มาเลเซีย และเวียดนาม โดยมีปริมาณการส่งออกรวม 7.464 ล้านตัน คิดเป็นร้อยละ 92.09 ของปริมาณการส่งออกยางธรรมชาติทั้งหมดของโลก ประกอบด้วย ไทย 2.952 ล้านตัน คิดเป็นร้อยละ 36.42 อินโดนีเซีย 2.543 ล้านตัน คิดเป็นร้อยละ 31.38 มาเลเซีย 1.228 ล้านตัน คิดเป็นร้อยละ 15.15 และเวียดนาม 0.741 ล้านตัน คิดเป็นร้อยละ 9.14 ซึ่งแสดงดังตารางที่ 2

ตารางที่ 2 ปริมาณการส่งออกยางธรรมชาติของประเทศต่าง ๆ ในปี พ.ศ. 2550–2554

หน่วย: 1,000 ตัน

ประเทศ	2550	2551	2552	2553	2554
1. ไทย	2,704	2,675	2,726	2,866	2,952
2. อินโดนีเซีย	2,407	2,296	2,001	2,369	2,543
3. มาเลเซีย	1,213	1,155	1,088	1,245	1,228
4. เวียดนาม	716	658	731	782	741
5. โกตดิวัวร์	179	203	221	247	232
6. กัวเตมาลา	69	73	78	76	81
7. ไลบีเรีย	121	85	60	62	61
8. แคเมอรูน	62	64	59	58	58
9. ศรีลังกา	49	46	54	50	41
10. พม่า	62	55	51	67	68
11. อื่นๆ	182	231	244	189	100
รวม	7,764	7,541	7,313	8,011	8,105
อัตราการเปลี่ยนแปลง (ร้อยละ)	0.40	-2.88	-3.03	9.56	1.17

ที่มา: Internatinal Rubber Study Group (2011)

การใช้ยางธรรมชาติของประเทศต่าง ๆ

ปริมาณการใช้ยางธรรมชาติของโลกในช่วงปี พ.ศ. 2550 ถึงปี พ.ศ. 2554 มีอัตราการใช้ยางธรรมชาติอยู่ระหว่าง 9.329–10.778 ล้านตัน เมื่อพิจารณาเป็นรายประเทศพบว่าจีนเป็นประเทศที่มีการใช้ยางธรรมชาติมากที่สุดของโลกทุกปี โดยในปี พ.ศ. 2554 จีนมีปริมาณใช้ 3.610 ล้านตัน คิดเป็นร้อยละ 34.03 รองลงมาคือสหรัฐอเมริกา 0.991 ล้านตัน คิดเป็นร้อยละ 9.34 อินเดีย 0.962 ล้านตัน คิดเป็นร้อยละ 9.07 ญี่ปุ่น 0.783 ล้านตัน คิดเป็นร้อยละ 7.38 และไทย 0.488 ล้านตัน คิดเป็นร้อยละ 4.60 ซึ่งไทยมีปริมาณการใช้เพิ่มขึ้น จากที่เคยอยู่ในลำดับที่ 7 หรือ 8 ปรับขึ้นมาอยู่เป็นลำดับที่ 5 ของประเทศที่มีการใช้ยางมากที่สุดของโลกแทนที่มาเลเซีย ซึ่งแสดงดังตารางที่ 3

ตารางที่ 3 ปริมาณการใช้ยางธรรมชาติของประเทศต่าง ๆ ในปี พ.ศ. 2550–2554

หน่วย: 1,000 ตัน

ประเทศ	2550	2551	2552	2553	2554
1. จีน	2,843	2,947	3,384	3,646	3,610
2. สหรัฐอเมริกา	1,018	1,041	687	926	991
3. อินเดีย	851	881	904	944	962
4. ญี่ปุ่น	887	878	637	750	783
5. ไทย	374	398	399	459	488
6. อินโดนีเซีย	383	413	352	421	428
7. มาเลเซีย	450	469	470	458	423
8. เกาหลีใต้	377	358	330	384	405
9. บราซิล	345	357	279	374	369
10. เยอรมนี	282	247	17	291	307
11. อื่นๆ	2,368	2,186	1,870	2,125	1,842
รวม	10,178	10,175	9,329	10,778	10,608
อัตราการเปลี่ยนแปลง (ร้อยละ)	5.04	-0.03	-8.31	15.53	-1.58

ที่มา: Internatinal Rubber Study Group (2011)

5. การผลิต การส่งออก และการใช้ยางธรรมชาติในประเทศไทย

ยางพาราเป็นพืชเศรษฐกิจที่สำคัญของประเทศไทยมีการผลิตเป็นอันดับหนึ่งของโลก และสามารถสร้างรายได้ให้กับประเทศไทยปีละหลายหมื่นล้านบาท เนื่องจากยางธรรมชาติเป็นวัตถุดิบหลักในการผลิตผลิตภัณฑ์ที่จำเป็นได้หลากหลายชนิด ยางธรรมชาติมีหลายชนิดได้แก่ ยางแผ่นรมควัน ยางแท่ง และน้ำยางข้น สามารถนำไปใช้เป็นวัตถุดิบที่สำคัญในอุตสาหกรรมต่าง ๆ ทำให้ปริมาณการผลิตยางธรรมชาติของไทยปี พ.ศ.2557 มีปริมาณเพิ่มขึ้นถึง 153,547 ตัน หรือเพิ่มขึ้นร้อยละ 3.68 โดยมีปริมาณการส่งออกเพิ่มขึ้นถึง 105,708 ตันหรือเพิ่มขึ้นร้อยละ 2.88 และในขณะเดียวกันก็มีการใช้ภายในประเทศเพิ่มขึ้น 20,375 ตันหรือเพิ่มขึ้นร้อยละ 3.91 เช่นกัน ดังแสดงในตารางที่ 4 นอกจากประเทศไทยจะผลิตยางธรรมชาติมากเป็นอันดับหนึ่งของโลกแล้ว ประเทศไทยยังเป็นประเทศที่ส่งออกยางธรรมชาติมากที่สุดในโลกด้วย โดยมีปริมาณการส่งออกเพิ่มขึ้นทุกปี ในปี พ.ศ. 2557 ปริมาณการส่งออกยางของไทยมีทั้งสิ้น 3,770,649 ตัน เพิ่มขึ้นจากปี พ.ศ. 2553 ที่มีปริมาณการส่งออก 2,866,447 ตัน หรือเพิ่มขึ้นร้อยละ 31.54 โดยประเทศผู้ซื้อปลายทาง 3 ลำดับแรกคือ จีน มาเลเซีย และญี่ปุ่น ดังแสดงในตารางที่ 5

ตารางที่ 4 ปริมาณการผลิตการส่งออก การใช้ในประเทศ และสต็อกของยางธรรมชาติของประเทศไทย ปี พ.ศ. 2542–2557

หน่วย: ตัน

ปี พ.ศ.	ปริมาณการผลิต	ปริมาณส่งออก	ใช้ในประเทศ	สต็อก
2542	2,154,560	1,886,339	226,917	250,850
2543	2,346,487	2,166,153	242,549	188,635
2544	2,319,549	2,042,079	253,105	213,000
2545	2,615,104	2,354,416	278,355	196,680
2546	2,876,005	2,573,450	298,699	202,240
2547	2,984,293	2,637,096	318,649	232,560
2548	2,937,158	2,632,398	334,649	204,256
2549	3,136,993	2,771,673	320,885	249,895
2550	3,056,005	2,703,762	373,659	230,390
2551	3,089,751	2,675,283	397,595	251,721
2552	3,164,379	2,726,193	399,415	293,659
2553	3,252,135	2,866,447	458,637	227,252
2554	3,569,033	2,952,381	486,745	361,557
2555	3,778,010	3,121,332	505,052	516,675
2556	4,170,428	3,664,941	520,628	502,855
2557	4,323,975	3,770,649	541,003	516,756

ที่มา: สถาบันวิจัยยาง กรมวิชาการเกษตร (2559)

ตารางที่ 5 ตลาดส่งออกยางธรรมชาติที่สำคัญของไทยปี พ.ศ. 2553–2557

หน่วย:ตัน

ประเทศ	2553	2554	2555	2556	2557
1. จีน	1,128,553	1,274,188	1,630,322	2,075,776	2,142,199
2. มาเลเซีย	443,000	344,589	353,501	421,408	406,025
3. ญี่ปุ่น	346,302	333,669	269,418	281,091	256,578
4. เกาหลีใต้	171,530	186,634	181,403	183,466	188,675
5. สหรัฐอเมริกา	177,859	205,410	172,577	145,638	146,794
6. ยุโรป	268,693	223,938	179,302	205,498	231,053
7. อื่นๆ	330,510	383,953	334,809	352,064	399,325
รวม	2,866,447	2,952,381	3,121,332	3,664,941	3,770,649

ที่มา: สถาบันวิจัยยาง กรมวิชาการเกษตร (2559)

6. การส่งออกยางพาราของไทยไปประเทศมาเลเซีย

ประเทศไทยส่งยางพาราไปยังประเทศมาเลเซียเป็นลำดับ 2 รองจากประเทศจีน โดยผลิตภัณฑ์ที่ประเทศมาเลเซียซื้อ ได้แก่ น้ำยางข้น น้ำยางธรรมชาติอื่น ๆ ยางคอมปาวด์ ยางเครพ ยางแท่ง ยางแผ่นผึ่งแห้ง และยางแผ่นรมควัน จากข้อมูลพบว่าประเทศไทยส่งออกน้ำยางข้นไปยังประเทศมาเลเซียเป็นปริมาณที่มากที่สุดในทุกปี ในปี พ.ศ. 2557 ส่งออก 537,408 ตัน คิดเป็นร้อยละ 87.61 ของปริมาณการส่งออกยางพาราไปมาเลเซียในปีนั้นทั้งหมด ดังแสดงในตารางที่ 6

ตารางที่ 6 ปริมาณการส่งออกยางพาราของไทยไปมาเลเซีย ปี พ.ศ. 2553–2557

หน่วย:ตัน

ประเภท	2553	2554	2555	2556	2557
1. น้ำยางข้น	584,034	490,918	512,022	578,640	537,408
2. ยางคอมปาวด์	46,594	36,137	38,522	28,428	9,148
3. ยางเครพ	1,785	1,015	1,575	4,722	5,303
4. ยางแท่ง	2,018	1,818	3,364	7,286	8,515
5. ยางธรรมชาติอื่นๆ	29,020	17,601	24,369	54,522	45,356
6. ยางแผ่นผึ่งแห้ง	1,172	290	224	2,783	3,362
7. ยางแผ่นรมควัน	6,916	8,257	5,808	13,563	4,306
รวม	671,539	556,036	585,884	689,944	613,398

ที่มา: ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร (2559)

ดังนั้นในการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ครั้งนี้ผู้วิจัยจึงสนใจนำข้อมูลการส่งออกยางธรรมชาติชนิดน้ำยางข้นมาศึกษาเท่านั้น

ตอนที่ 4 ความรู้เกี่ยวกับทฤษฎีอุปสงค์

ทฤษฎีอุปสงค์

Marshall (cited in Nicholson, 1994, p. 123) กล่าวว่า อุปสงค์ของสินค้าชนิดใดชนิดหนึ่ง คือ ปริมาณต่าง ๆ ของสินค้าชนิดนั้นที่ผู้บริโภคต้องการจะเสนอซื้อ ณ ระดับราคาต่าง ๆ ในเวลาหนึ่ง โดยมีอำนาจซื้อ (Purchasing Power) สนับสนุน

ปัจจัยกำหนดอุปสงค์

อุปสงค์ในสินค้าชนิดใดชนิดหนึ่งจะมีมากหรือน้อยเพียงใดมิได้ขึ้นอยู่กับราคาของสินค้าชนิดนั้น ๆ เพียงอย่างเดียว หากแต่ยังขึ้นอยู่กับปัจจัยอื่น ๆ อีกหลายอย่างประกอบกัน เช่น รสนิยมของผู้ซื้อ รายได้ของผู้ซื้อ ราคาของสินค้าอย่างอื่นที่เกี่ยวข้อง การกระจายรายได้ และจำนวนผู้ซื้อในตลาด

จากปัจจัยกำหนดอุปสงค์ดังกล่าวสามารถเขียนฟังก์ชันอุปสงค์ (Demand Function) ได้ดังนี้ (Marshall cited in Nicholson, 1994, p. 124)

$$Q_x = f\left(P_x, I, P_y, E, \ldots\right)$$

โดยที่ Q_x แทน ปริมาณการซื้อของสินค้า X

P_x แทน ราคาของสินค้า X

I แทน รายได้ของผู้ซื้อหรือผลิตภัณฑ์มวลรวมในประเทศ

P_y แทน ราคาสินค้าอย่างอื่นที่เกี่ยวข้องกับสินค้า X เช่น สินค้าที่ใช้แทนหรือใช้ร่วมกัน

E แทน อัตราแลกเปลี่ยน

กฎว่าด้วยอุปสงค์ จะอธิบายความสัมพันธ์ระหว่างราคากับปริมาณการซื้อสินค้าชนิดหนึ่ง โดยสมมติให้ปัจจัยอื่น ๆ อยู่คงที่ ถ้าพิจารณาตามฟังก์ชันอุปสงค์ ก็คือการอธิบายความสัมพันธ์ระหว่าง P_x กับ Q_x เพียงสองตัวเท่านั้น โดยกฎว่าด้วยอุปสงค์จะอธิบายว่า เมื่อราคาของสินค้าชนิดหนึ่งสูงขึ้น ปริมาณการซื้อของสินค้าชนิดนั้นจะลดลง ในทางตรงกันข้ามถ้าราคาสินค้าชนิดนั้นลดลง ปริมาณการซื้อสินค้าขนิดนั้นจะเพิ่มขึ้น แสดงว่าปริมาณการซื้อของสินค้าชนิดไหนก็จะถูกกำหนดโดยราคาของสินค้าชนิดนั้นเพียงอย่างเดียว

อุปสงค์ต่อราคา

อุปสงค์ต่อราคา คือปริมาณสินค้าที่ผู้ซื้อต้องการซื้อ ณ ระดับราคาต่าง ๆ ณ เวลาใดเวลาหนึ่ง โดยกำหนดให้สิ่งอื่น ๆ อยู่คงที่ หรือกล่าวได้ว่าอุปสงค์ต่อราคาก็คือปริมาณการซื้อของสินค้าชนิดหนึ่ง ซึ่งจะมีจำนวนมากน้อยเพียงใดนั้น จะขึ้นอยู่กับราคาของสินค้าชนิดนั้น ๆ แสดงว่าราคาสินค้าเป็นตัวกำหนดปริมาณการซื้อ

ความยืดหยุ่นของอุปสงค์

ความยืดหยุ่นของอุปสงค์คือ อัตราการเปลี่ยนแปลงของปริมาณสินค้าที่ผู้ซื้อต้องการซื้อ ณ ขณะใดขณะหนึ่ง เพื่อตอบสนองต่ออัตราการเปลี่ยนแปลงของตัวแปรอื่น ๆ ที่เป็นตัวกำหนดปริมาณการซื้อ เนื่องจากในการอธิบายไม่สามารถที่จะแสดงให้เห็นพร้อม ๆ กันได้ว่าตัวแปรแต่ละตัวมีส่วนในการกำหนดปริมาณการซื้อมากน้อยเพียงใด ดังนั้น เวลาอธิบายจึงหยิบยกตัวแปรขึ้นมาอธิบายทีละตัวแปร โดยสมมติให้ตัวอื่น ๆ อยู่คงที่ไม่ให้เข้ามามีส่วนในการกำหนดปริมาณการซื้อด้วย ความยืดหยุ่นของอุปสงค์แบ่งออกเป็น 3 ประเภท คือ

1. ความยืดหยุ่นของอุปสงค์ต่อราคา
2. ความยืดหยุ่นของอุปสงค์ต่อรายได้
3. ความยืดหยุ่นของอุปสงค์ต่อราคาสินค้าชนิดอื่น

โดยอธิบายรายละเอียดของอุปสงค์แต่ละประเภทได้ ดังนี้

1. ความยืดหยุ่นของอุปสงค์ต่อราคา

ความยืดหยุ่นของอุปสงค์ต่อราคาคือ อัตราการเปลี่ยนแปลงของปริมาณสินค้าชนิดใดชนิดหนึ่งที่ผู้ซื้อต้องการซื้อ ซึ่งเป็นการตอบสนองต่ออัตราการเปลี่ยนแปลงของราคาสินค้าชนิดนั้น เนื่องจาก

สินค้าหรือบริการแต่ละชนิดมีความจำเป็นมากน้อยแตกต่างกัน หรือเป็นสินค้าที่สามารถใช้สินค้าชนิดอื่นทดแทนได้หรือไม่ได้ เป็นต้น

1.1 ประเภทของความยืดหยุ่นของอุปสงค์ต่อราคา การศึกษาเรื่องความยืดหยุ่นของอุปสงค์ต่อราคาในเบื้องต้น จะแบ่งลักษณะความยืดหยุ่นออกเป็น 5 ประเภทด้วยกันคือ

1.1.1 อุปสงค์ที่ไม่มีความยืดหยุ่น หรือมีความยืดหยุ่นเท่ากับศูนย์

1.1.2 อุปสงค์ที่มีความยืดหยุ่นน้อย คือมีความยืดหยุ่นน้อยกว่า 1 แต่มากกว่าศูนย์ โดยทั่วไปเรียกว่าอุปสงค์มีความยืดหยุ่นน้อย

1.1.3 อุปสงค์ที่มีความยืดหยุ่นเท่ากับหนึ่ง เส้นอุปสงค์ที่มีความยืดหยุ่นเท่ากับหนึ่งตลอดทั้งเส้นจะมีลักษณะเป็นโค้งแบบไฮเพอร์โบลามุมฉาก (Rectangular Hyperbola)

1.1.4 อุปสงค์มีความยืดหยุ่นมากกว่าหนึ่ง หรือน้อยกว่าอินฟินิตี้ โดยทั่วไปนิยมเรียกว่าอุปสงค์มีความยืดหยุ่นมาก

1.1.5 อุปสงค์มีความยืดหยุ่นอย่างสมบูรณ์ หรือมีความยืดหยุ่นเท่ากับอินฟินิตี้

1.2 ปัจจัยกำหนดค่าความยืดหยุ่นของอุปสงค์ต่อราคา

1.2.1 สินค้าจำเป็นหรือสินค้าฟุ่มเฟือย ถ้าเป็นสินค้าที่จำเป็นแก่การดำรงชีวิต ความยืดหยุ่นของอุปสงค์ต่อราคาจะมีน้อย ส่วนสินค้าฟุ่มเฟือยประเภทที่ราคาแพง อุปสงค์ต่อราคาจะมีความยืดหยุ่นมาก เนื่องจากอัตราการเปลี่ยนแปลงของปริมาณการซื้อมีมากกว่าอัตราการเปลี่ยนแปลงของราคา

1.2.2 สินค้าที่ใช้ทดแทนกันได้หรือทดแทนกันไม่ได้ สินค้าที่สามารถใช้สินค้าชนิดอื่นทดแทนได้ ความยืดหยุ่นของอุปสงค์ต่อราคาจะมีมาก ถ้าสินค้าชนิดใดหาสินค้าอย่างอื่นมาทดแทนไม่ได้เลย หรือพอทดแทนได้แต่ไม่ดีเท่าที่ควร กรณีนี้ความยืดหยุ่นของอุปสงค์ต่อราคาจะมีน้อย

1.2.3 ราคาสินค้าเมื่อเปรียบเทียบกับรายได้ของผู้บริโภค สินค้าชนิดใดที่เป็นสินค้าเล็ก ๆ น้อย ๆ ราคาถูก ๆ ใช้เงินจ่ายซื้อเพียงส่วนน้อยของรายได้ ความยืดหยุ่นของอุปสงค์ต่อราคาจึงมีน้อยมากด้วย ส่วนสินค้าที่มีราคาสูงซึ่งการซื้อจะต้องใช้เงินส่วนใหญ่ของรายได้ กรณีนี้ความยืดหยุ่นของอุปสงค์ต่อราคาจะมีมาก

1.2.4 สินค้าที่ใช้ได้ทนทานหรือบุบสลายง่าย สินค้าที่ทนทานหรืออายุการใช้งานยาวนานและสามารถซ่อมแซมได้ อุปสงค์ต่อราคาจะมีความยืดหยุ่นน้อย ส่วนสินค้าที่บุบสลายง่ายใช้งานได้ไม่นาน และราคาไม่ค่อยสูงมากนัก อุปสงค์ต่อราคาจะมีความยืดหยุ่นมาก

1.2.5 ระยะเวลา ความยืดหยุ่นของอุปสงค์จะมีมากน้อยเพียงใดขึ้นอยู่กับระยะเวลาด้วย ถ้าผู้บริโภคมีเวลาในการตัดสินใจนาน ปริมาณการซื้อของสินค้ามีโอกาสที่จะเปลี่ยนแปลงได้มากขึ้น ด้วยเหตุนี้ระยะเวลาจึงเข้ามามีส่วนทำให้ความยืดหยุ่นของอุปสงค์มากขึ้น

2. ความยืดหยุ่นของอุปสงค์ต่อรายได้

ความยืดหยุ่นของอุปสงค์ต่อรายได้ คืออัตราการเปลี่ยนแปลงของปริมาณสินค้าที่ผู้ซื้อต้องการซื้อ ณ เวลาใดเวลาหนึ่งที่กำหนด ซึ่งตอบสนองต่ออัตราการเปลี่ยนแปลงของรายได้ โดยสมมติให้สิ่งอื่น ๆ คงที่ ความยืดหยุ่นของอุปสงค์ต่อรายได้สามารถคำนวณได้ดังนี้

$$\text{ความยืดหยุ่นของอุปสงค์ต่อรายได้} = \frac{\text{อัตราการเปลี่ยนแปลงของปริมาณการซื้อ}}{\text{อัตราการเปลี่ยนแปลงของรายได้}}$$

3. ความยืดหยุ่นของอุปสงค์ต่อราคาสินค้าชนิดอื่น

ความยืดหยุ่นของอุปสงค์ต่อราคาสินค้าชนิดอื่น คืออัตราการเปลี่ยนแปลงปริมาณการซื้อของสินค้าชนิดใดชนิดหนึ่งที่ผู้บริโภคต้องการซื้อ ณ ขณะใดขณะหนึ่ง ซึ่งตอบสนองต่ออัตราการเปลี่ยนแปลงของราคาสินค้าชนิดอื่นที่เกี่ยวข้อง โดยสมมติให้สิ่งอื่น ๆ อยู่คงที่ สินค้าที่เกี่ยวข้องในที่นี้หมายถึงเป็นการเกี่ยวข้องในลักษณะที่ใช้ทดแทนกันได้อย่างหนึ่ง หรือเกี่ยวข้องกันเพราะต้องใช้ประกอบกันจะแยกกันใช้โดยลำพังไม่ได้

ความยืดหยุ่นของอุปสงค์ต่อราคาสินค้าชนิดอื่น

$$= \frac{\text{อัตราการเปลี่ยนแปลงปริมาณการซื้อของสินค้าชนิดอื่น}}{\text{อัตราการเปลี่ยนแปลงของราคาสินค้าชนิดอื่นที่เกี่ยวข้อง}}$$

ค่าความยืดหยุ่นของอุปสงค์ต่อราคาสินค้าชนิดอื่นจะมีลักษณะทำนองเดียวกับความยืดหยุ่นของอุปสงค์ต่อรายได้ คืออาจมีเครื่องหมายบวกหรือลบก็ได้ แต่เครื่องหมายเหล่านี้ไม่ใช่สิ่งที่จะบอกค่าของความยืดหยุ่นว่ามีมากหรือน้อยกว่าศูนย์ จะบอกเพียงลักษณะความสัมพันธ์ระหว่างสินค้า 2 ชนิดว่าเป็นสินค้าที่ใช้ทดแทนกันได้หรือเป็นสินค้าที่ใช้ประกอบกันเท่านั้น สำหรับค่าของตัวเลข (ไม่คำนึงถึงเครื่องหมาย) จะสามารถบอกให้ทราบได้ว่าสินค้าทั้ง 2 ชนิดมีความสัมพันธ์กันมากน้อยเพียงใด กล่าวคือถ้าค่าตัวเลขสูงก็แสดงว่าสินค้า 2 ชนิดมีความสัมพันธ์กันมาก แต่ถ้าค่าตัวเลขต่ำก็แสดงว่าสินค้า 2 ชนิดมีความสัมพันธ์กันน้อย และถ้าความยืดหยุ่นของอุปสงค์ต่อราคาสินค้าชนิดอื่นมีค่าเท่ากับศูนย์ แสดงว่าสินค้า 2 ชนิดนั้นเป็นอิสระต่อกันหรือไม่มีความสัมพันธ์กันเลย

ทฤษฎีอุปสงค์ของปัจจัยการผลิต

Shephard and Lemma (cited in Nicholson, 2000, pp. 147-148) กล่าวว่าผู้ผลิตไม่ได้ซื้อปัจจัยการผลิตไปบริโภคโดยตรง แต่ได้นำไปเป็นวัตถุดิบเพื่อผลิตสินค้าอื่น ในการแสวงหาจุดเหมาะสมในการผลิตของหน่วยผลิต ซึ่งสามารถแบ่งออกเป็น 3 ลักษณะ ตามเป้าหมายของการผลิตของหน่วยผลิต ดังต่อไปนี้

1. การผลิตเพื่อให้ได้ผลผลิตมากที่สุดภายใต้เงื่อนไขต้นทุนมีจำกัด
2. การผลิตโดยใช้ต้นทุนการผลิตน้อยที่สุดภายใต้เงื่อนไขให้ได้ปริมาณผลผลิตที่กำหนด
3. การผลิตให้ได้กำไรมากที่สุด

ในที่นี้จะอธิบายถึงทฤษฎีสมการอุปสงค์ที่มีผลต่อปัจจัยการผลิตตามเป้าหมายการผลิตโดยใช้ต้นทุนการผลิตน้อยที่สุด ภายใต้เงื่อนไขให้ได้ปริมาณผลผลิตตามที่กำหนดเป็นการวิเคราะห์พฤติกรรมของหน่วยผลิตที่พยายามเลือกซื้อปัจจัยการผลิตแต่ละชนิด ในจำนวนที่จะทำให้เสียต้นทุนน้อยที่สุด โดยให้ได้รับปริมาณผลผลิตในระดับที่หน่วยผลิตต้องการ ต้นทุนการผลิตน้อยที่สุดภายใต้เงื่อนไขให้ได้ปริมาณผลผลิตที่กำหนด โดยสร้างสมการเป้าหมายและสร้างสมการเงื่อนไข เมื่อสร้าง

สมการทั้งสองสมการขึ้นมาแล้ว นำมารวมกันเป็นสมการใหม่ โดยมีตัว λ เป็นตัวเชื่อม ซึ่งเรียกว่า ตัวคูณลากรานจ์ (Lagrange Multiplier)

กรณีที่สมการอุปสงค์อยู่ในรูปของเลขชี้กำลัง (Exponential) การหาค่าความยืดหยุ่นของอุปสงค์สามารถหาได้โดยการ Take Log เข้าไปในสมการ จากสมการอุปสงค์ของสินค้า

$$QX = AP\_X^{\alpha 1} GDPX^{\alpha 2} P\_Y^{\alpha 3} Q\_Z^{\alpha 4}$$

กำหนดให้
- QX แทน ปริมาณอุปสงค์ของสินค้า X
- P_X แทน ราคาสินค้า X
- $GDPX$ แทน รายได้ของประเทศที่ซื้อสินค้า X
- P_Y แทน ราคาสินค้า Y ซึ่งใช้แทนสินค้า X
- Q_Z แทน ปริมาณของสินค้า Z ซึ่งใช้วัตถุดิบจากสินค้า X
- A แทน ค่าคงที่
- $\alpha 1$ แทน ค่าความยืดหยุ่นของราคาสินค้า X
- $\alpha 2$ แทน ค่าความยืดหยุ่นของรายได้ของประเทศที่ซื้อสินค้า X
- $\alpha 3$ แทน ค่าความยืดหยุ่นของราคาสินค้า Y ซึ่งใช้แทนสินค้า X
- $\alpha 4$ แทน ค่าความยืดหยุ่นของปริมาณสินค้า Z ซึ่งใช้วัตถุดิบจากสินค้า X

จากสมการสามารถ Take Log เข้าไปทั้งสองข้าง จะได้

$$\ln QX = \alpha_0 + \alpha_1 \ln P\_X + \alpha_2 \ln GPAX + \alpha_3 \ln P\_Y + \alpha_{4\ln} \ln Q\_Z$$

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ผู้วิจัยได้ศึกษาผลงานวิจัยที่เกี่ยวข้องกับปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน มีรายละเอียด ดังนี้

เปรมวดี ชูไสว (2548) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณเมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ วิธีการเปรียบเทียบที่นำมาพิจารณา คือ วิธีกำลังสองน้อยที่สุด วิธีการถดถอยริดจ์ วิธีการถดถอยส่วนประกอบหลัก โดยเกณฑ์เปรียบเทียบ คือ อัตราส่วนของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของสัมประสิทธิ์การถดถอย โดยจำนวนตัวแปรอิสระที่ใช้เท่ากับ 3, 6 และ 9 มีส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อนเท่ากับ 1, 5 และ 10 ขนาดตัวอย่างเท่ากับ 5p, 10p, 15p, 20p, 25p และ 30p เมื่อ p คือจำนวนตัวแปรอิสระ โดยแบ่งระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็น น้อย (0.15-0.30) ปานกลาง (0.31-0.5,0.51-0.65) และมาก (0.66-0.85, 0.86-0.99) ผลที่ได้จากการวิจัยพบว่า มากกว่า 99% ของสถานการณ์สามารถสรุปได้ว่าในระดับความสัมพันธ์มาก ช่วง 0.86-0.99

นั้นจะทำให้เกิดปัญหาการมีพหุสัมพันธ์สูงจนส่งกระทบต่อการประมาณค่าด้วยตัวประมาณแบบวิธีกำลังสองน้อยที่สุด ดังนั้นวิธีการถดถอยส่วนประกอบหลักจึงเป็นวิธีที่มีประสิทธิภาพที่สุด ในทุกกรณีในส่วนระดับความสัมพันธ์อื่นนั้น จะต้องพิจารณาค่าส่วนเบี่ยงเบนของค่าคลาดเคลื่อนประกอบด้วย ถ้าส่วนเบี่ยงเบนของค่าคลาดเคลื่อนเท่ากับ 1 วิธีกำลังสองน้อยที่สุดจะมีประสิทธิภาพมากที่สุด และในส่วนของการเบี่ยงเบนค่าคลาดเคลื่อนเท่ากับ 5 และ 10 วิธีกำลังสองน้อยที่สุด หลังจากที่ทำการแก้ไขให้ X เป็นแกนตั้งฉากซึ่งกันและกันแล้ว จะเป็นวิธีที่ดีที่สุด ปัจจัยที่มีผลต่อความคลาดเคลื่อนนั้น พบว่าเมื่อขนาดตัวอย่างเพิ่มขึ้นค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย(AMSE) ที่ได้จะให้ค่าลดลง ตรงกันข้ามถ้าระดับความสัมพันธ์ระหว่างตัวแปรอิสระและส่วนเบี่ยงเบนของค่าคลาดเคลื่อนเพิ่มขึ้น ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ได้จะมีค่าเพิ่มขึ้น ยกเว้นกรณีวิธีการถดถอยส่วนประกอบหลักที่ระดับความสัมพันธ์ระหว่างตัวแปรอิสระช่วง 0.86-0.99 จะน้อยกว่าค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยในระดับอื่น

ปิยะนาฎ หาดารา (2549) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเมื่อข้อมูลมีค่าผิดปกติ 6 วิธี คือ วิธีกำลังสองน้อยที่สุด วิธีตัวประมาณแบบเอ็ม เมื่อใช้เกณฑ์ความแกร่งของ Huber, Ramsay และ Andrews วิธีบูสแตรป และวิธีแจ๊คไนฟ ศึกษาในกรณีที่จำนวนตัวแปรอิสระเท่ากับ 2 ขนาดตัวอย่างเท่ากับ 25, 50 และ 100 เปอร์เซนต์ค่าความผิดปกติเท่ากับ 1% และ 5% กำหนดระดับความรุนแรงของค่าผิดปกติแบ่งเป็น 2 ระดับ คือ ปานกลางและระดับรุนแรง เกณฑ์ในการเปรียบเทียบคือค่าเอนเอียง ค่าความแปรปรวน และค่าคลาดเคลื่อนกำลังสองเฉลี่ย การเปรียบเทียบกระทำภายใต้เงื่อนไขของข้อมูลที่มีค่าผิดปกติ ดังนี้ 1) กรณีที่ข้อมูลเกิดค่าผิดปกติที่ตัวแปรตาม Y 2) กรณีที่ข้อมูลเกิดค่าผิดปกติที่ตัวแปรอิสระ X_1 กับ X_2 3) กรณีที่ข้อมูลเกิดค่าผิดปกติที่ตัวแปรอิสระ X_1 และตัวแปรตาม Y และ 4) กรณีที่ข้อมูลเกิดค่าผิดปกติที่ตัวแปรอิสระ X_1, X_2 และตัวแปรตาม Y โดยข้อมูลที่ใช้ในการวิจัยได้มาจากการจำลองด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 1,000 รอบ ในแต่ละกรณี พบว่า กรณีที่เกิดค่าผิดปกติในตัวแปรตาม โดยทั่วไปวิธีตัวประมาณเอ็ม เมื่อใช้ความแกร่งของ Andrews ให้ค่าเอนเอียงค่าความแปรปรวนและค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธีอื่น ๆ กรณีข้อมูลเกิดค่าผิดปกติในตัวแปรอิสระ X_1 กับ X_2 และกรณีที่ข้อมูลเกิดค่าผิดปกติที่ตัวแปรอิสระ X_1 กับ X_2 และ ตัวแปรตาม โดยสถานการณ์ทั่วไปพบว่าการประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีตัวประมาณเอ็ม เมื่อใช้เกณฑ์ความแกร่งของ Andrews ให้ค่าเอนเอียงและค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธีอื่น ๆ แต่เมื่อใช้เกณฑ์ความแกร่งของ Huber ให้ค่าความแปรปรวนต่ำที่สุด และกรณีที่เกิดค่าผิดปกติที่ตัวแปรอิสระ X_1 กับตัวแปรตาม Y โดยทั่วไปพบว่าวิธีตัวประมาณเอ็มเมื่อใช้เกณฑ์ความแกร่งของ Andrews ให้ค่าเอนเอียงต่ำกว่าวิธีอื่น ๆ แต่เมื่อใช้เกณฑ์ความแกร่งของ Huber ให้ค่าความแปรปรวนและค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด

อุบลรัตน์ ภู่ระหงษ์ (2549) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของการถดถอยเชิงเส้นพหุคูณ โดยใช้วิธีประมาณ 5 วิธี คือวิธีกำลังสองน้อยที่สุดสามัญ วิธี Least Median Square วิธี Least Trimmed Square วิธีเอ็ม (M) และ วิธีเอ็มเอ็ม (MM) เกณฑ์ที่ใช้ในการเปรียบเทียบคือค่าประสิทธิภาพสัมพัทธ์ โดยกระทำภายใต้เงื่อนไขในกรณีศึกษาต่าง ๆ ดังนี้ ขนาดตัวอย่าง 3 ระดับ

คือ 30, 40 และ 80 จำนวนค่าผิดปกติ 3 ระดับ คือ 5%, 10% และ 15% ตำแหน่งการผิดปกติในตัวแปรอิสระและตัวแปรตาม 4 ระดับ คือ (2σ,2σ), (4σ,8σ), (8σ,4σ) และ (8σ,8σ) สรุปผลได้ดังนี้ กรณีไม่มีค่าผิดปกติเกิดขึ้นในข้อมูล พบว่าวิธีกำลังสองน้อยที่สุดเป็นวิธีที่ให้ค่าประสิทธิภาพสัมพัทธ์สูงที่สุด และเมื่อขนาดตัวอย่างใหญ่ขึ้น พบว่านอกจากวิธีกำลังสองน้อยที่สุดสามัญแล้ววิธีเอ็ม และวิธีเอ็มเอ็มยังเป็นวิธีที่เหมาะสมสำหรับใช้ประมาณค่าพารามิเตอร์ที่ดีอีกด้วย และในกรณีมีค่าผิดปกติเกิดขึ้นในข้อมูล ผลการวิจัยสามารถจำแนกตามตำแหน่งการผิดปกติของตัวแปรอิสระตัวแปรตาม และจำนวนค่าผิดปกติ ดังนี้ เมื่อตำแหน่งที่เกิดค่าผิดปกติ (2σ,2σ) พบว่าในทุกขนาดตัวอย่างและจำนวนค่าผิดปกติวิธีกำลังสองน้อยที่สุดให้ค่าประสิทธิภาพสัมพัทธ์สูงที่สุด และตำแหน่ง (4σ,8σ), (8σ,4σ) และ (8σ,8σ) พบว่าโดยรวมวิธีเอ็มเอ็ม ให้ค่าประสิทธิภาพสัมพัทธ์ต่ำที่สุด โดยค่าประสิทธิภาพสัมพัทธ์จะแปรผกผันตามจำนวนข้อมูลที่ผิดปกติ ตำแหน่งการเกิดค่าผิดปกติ และจำนวนตัวแปรอิสระ แต่จะแปรผันตามขนาดตัวอย่าง

กิตติพงษ์ ไตรทิพย์พานิชย์ (2550) ได้ศึกษาเปรียบเทียบความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติที่ใช้ตรวจสอบอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนในสถานการณ์ดังต่อไปนี้ 1) เปรียบเทียบสถิติ 3 วิธี คือ ตัวสถิติเดอร์บินวัตสัน บูทสแตรปเดอร์บินวัตสัน และ บูทสแตรปโรล ในการตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 1 2) เปรียบเทียบสถิติ 3 วิธี คือ ตัวสถิติวอลลิส บูทสแตรปวอลลิส และบูทสแตรปโรล และในการตรวจสอบอัตตสหสัมพันธ์ตำแหน่งที่ 4 เมื่อจำนวนตัวแปรอิสระเท่ากับ 1, 2 และ 5 และกำหนดขนาดตัวอย่างที่ทำการศึกษา คือ 15, 20, 50, 60, 90 และ 100 สำหรับในแต่ละสถานการณ์ ข้อมูลที่ใช้ในการวิจัยนี้ได้จากการทดลองด้วยเทคนิคมอนติคาร์โล โดยจำลองการทดลองด้วยเครื่องคอมพิวเตอร์ 1,000 ครั้ง สำหรับแต่ละสถานการณ์ที่ระดับนัยสำคัญ 0.05 ผลวิจัยสรุปได้ ดังนี้ 1.ความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1เมื่อมีอัตตสหสัมพันธ์ตำแหน่งที่ 1ตัวสถิติบูทสแตรปเดอร์บินวัตสัน และบูทสแตรปโรล สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ในทุกกรณีและเมื่อมีอัตตสหสัมพันธ์ตำแหน่งที่ 4 ตัวสถิติบูทสแตรปวอลลิส และบูทสแตรปโรล สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ในทุกกรณี 2. อำนาจการทดสอบ เมื่อมีอัตตสหสัมพันธ์ตำแหน่งที่ 1 โดยทั่วไปตัวสถิติบูทสแตรปเดอร์บินวัตสันให้อำนาจการทดสอบสูงที่สุด และเมื่อมีอัตตสหสัมพันธ์ตำแหน่งที่ 4 ตัวสถิติบูทสแตรปวอลลิส จะให้ค่าอำนาจการทดสอบที่สูงสุด

อดุลย์เดช กรงทอง (2550) ได้เปรียบเทียบวิธีประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระและข้อมูลมีค่าผิดปกติ ด้วยวีธีริดจ์ชนิดตัวประมาณจีเอ็ม โดยใช้ค่า k ที่ได้จากการประมาณด้วยวิธีโฮเอร์ล เคนนารด์ (Hoerl-Kennard: HK) และวิธีโฮเอร์ล เคนนารด์ บาลด์วิน (Hoerl-Kennard-Baldwin: HKB) ใช้เกณฑ์ความแกร่ง 6 เกณฑ์ คือ เกณฑ์ความแกร่งของ Huber, Ramsay, Andrews, Hample’s 17A, Barya และ Talwer กำหนดจำนวนตัวแปรอิสระที่ใช้ในการศึกษาเท่ากับ 3 ตัวแปร ขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ระดับความสัมพันธ์ของตัวแปรอิสระเป็น 3 ระดับ คือ ระดับต่ำ (0.1) ระดับปานกลาง (0.5) และระดับสูง

(0.9) กำหนดค่าผิดปกติ 2 ระดับ คือ ระดับปานกลาง และระดับรุนแรง เปอร์เซนต์การเกิดค่าผิดปกติเท่ากับ 1% และ 5% ของขนาดตัวอย่าง โดยแบ่งกรณีศึกษาเป็น 5 กรณี ดังนี้ ข้อมูลมีค่าผิดปกติในตัวแปร Y อย่างเดียว ข้อมูลมีค่าผิดปกติในตัวแปร X_3 อย่างเดียว ข้อมูลมีค่าผิดปกติในตัวแปร (X_1 ,X_2) (Y, X_3) (Y, X_1, X_2) ทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล และกระทำซ้ำ 1,000 รอบ ในแต่ละสถานการณ์การเปรียบเทียบพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ยผลการวิจัยพบว่าภายใต้กรณีที่ศึกษาวิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อประมาณค่าคงที่ k ด้วยวิธี HKB ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธี HK เป็นส่วนใหญ่ และเมื่อขนาดตัวอย่างเพิ่มขึ้น การประมาณค่า k ทั้งสองวิธีให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกัน นอกจากนี้เมื่อพิจารณาทุกระดับความรุนแรงของค่าผิดปกติ เปอร์เซนต์ค่าผิดปกติ ระดับของความสัมพันธ์ และทุกลักษณะของตัวแปรที่เกิดค่าผิดปกติ พบว่าส่วนใหญ่วิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อใช้เกณฑ์ความแกร่งของ Ramsay จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด เมื่อขนาดตัวอย่างเท่ากับ 20 และวิธีริดจ์ชนิดตัวประมาณจีเอ็มเมื่อใช้ความแกร่งของ Huber และ Ramsay ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยใกล้เคียงกันและมีค่าต่ำที่สุดเป็นส่วนใหญ่ เมื่อขนาดตัวอย่างเท่ากับ 50 และ 100 นอกจากนี้ยังพบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีแนวโน้มสูงขึ้นเมื่อระดับความรุนแรงของค่าผิดปกติและเปอร์เซนต์การเกิดค่าผิดปกติเพิ่มขึ้น แต่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีแนวโน้มลดลง เมื่อขนาดตัวอย่างเพิ่มขึ้น

วารุณี รวยดี (2551) ได้เปรียบเทียบวิธีการถ่วงน้ำหนักเฉพาะที่ในการวิเคราะห์การถดถอยสำหรับข้อมูลที่มีความสัมพันธ์กันและมีค่าผิดปกติ โดยทำการถ่วงน้ำหนักค่าสังเกต 3 วิธี ได้แก่ วิธี Lowess Bisquare วิธี Lowess Tri-cube และวิธี K-Nearest Neighbor (K-NN) เกณฑ์ที่ใช้ในการเปรียบเทียบคือค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ที่ได้จากการประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักการเปรียบเทียบกระทำภายใต้การจำลองสถานการณ์ และเปรียบเทียบในกรณีศึกษาจากข้อมูลจริง การจำลองสถานการณ์นั้นได้ใช้เทคนิควิธีของมอนติคาร์โล โดยกำหนดขนาดตัวอย่างในการศึกษาเท่ากับ 15, 30 และ 100 ระดับอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน 3 ระดับคือ ระดับต่ำ (0.2) ระดับปานกลาง (0.5) และระดับสูง (0.8) มีค่าผิดปกติปรากฏอยู่ร้อยละ 20 ของขนาดตัวอย่าง ซึ่งแบ่งการจำลองข้อมูลเป็น 2 กรณี คือ กรณีที่ 1 ตัวแปรอิสระและตัวแปรตามมีการแจกแจงปกติและมีความแปรปรวนเท่ากับ 1 ในแต่ละสถานการณ์กระทำซ้ำ 1,000 ครั้ง กรณีที่ 2 ตัวแปรอิสระและตัวแปรตามมีการแจกแจงปกติ โดยกำหนดให้ค่าความแปรปรวนเท่ากับค่าความแปรปรวนของข้อมูลจริง ข้อมูลจริงที่ใช้ในการศึกษาครั้งนี้ คือ อัตราแลกเปลี่ยนเงินบาทเทียบกับดอลลาร์สหรัฐ (X1) ราคาน้ำมันเบนซิน 95 (X2) และราคาทองคำแท่งขายออก (Y) ผลการวิจัยพบว่า การศึกษาจากการจำลองสถานการณ์ และจากกรณีศึกษาข้อมูลจริงนั้น ทุกสถานการณ์ของการจำลองข้อมูลนั้นวิธี K-NN ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ในขณะที่วิธี Lowess Tri-cube ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่สูงกว่าวิธี K-NN เพียงเล็กน้อยแต่วิธี K-NN เป็นวิธีที่ไม่มีความซับซ้อนมากนัก และเมื่อพิจารณาขนาดตัวอย่างและระดับอัตตสหสัมพันธ์ พบว่าการเพิ่มขนาดตัวอย่างมีผลทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยลดลง และการเพิ่มระดับอัตตสหสัมพันธ์ในตัวเองของค่าคลาดเคลื่อนนั้นมีผลทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยเพิ่มขึ้น สำหรับการศึกษากรณีศึกษาข้อมูล

จริงผลการศึกษาสอดคล้องกับการจำลองข้อมูลในกรณีที่ 2 สรุปได้ว่าวิธี K-NN เป็นวิธีที่เหมาะสมที่สุดในการถ่วงน้ำหนักเพื่อประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักในการถดถอยพหุคูณ เมื่อข้อมูลมีค่าสังเกตผิดปกติและค่าคลาดเคลื่อนเกิดอัตตสหสัมพันธ์กันในระดับต่ำหรือระดับปานกลาง แต่เมื่อข้อมูลมีค่าสังเกตผิดปกติและค่าคลาดเคลื่อนเกิดอัตตสหสัมพันธ์กันในระดับสูง ควรเลือกใช้วิธี Lowess Tri-cube และขนาดตัวอย่างที่ใช้ควรมีขนาดใหญ่

ณัฐพร ภักดี (2552) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณเมื่อตัวแปรอิสระมีพหุสัมพันธ์กันโดยวิธีการถดถอยริดจ์ โดยวิธีการถดถอยริดจ์ (RR) วิธีการถดถอยริดจ์แบบมีข้อจำกัด (RRR) และวิธีการถดถอยริดจ์แบบทางเลือก (ARR) และค่าประมาณ k คำนวณจากวิธีของโฮเอิร์น เคนนาร์ด และบาลด์วิน (HKB) โดยกำหนดระดับความสัมพันธ์ของตัวแปรอิสระเป็น 0.70 0.80 และ 0.90 และวิธีที่ดีที่สุด คือวิธีที่ให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (AMSE) ต่ำที่สุด ผลการวิจัยพบว่า กรณีตัวแปรอิสระ 3 ตัว เมื่อขนาดตัวอย่างเท่ากับ 30 ถ้าข้อจำกัดที่มีค่าเป็นลบ และข้อจำกัดมีค่าเป็น 0 ทุกระดับความสัมพันธ์ และความแปรปรวนของค่าคลาดเคลื่อนวิธีการถดถอยริดจ์จะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด และถ้าข้อจำกัดมีค่าเป็น 2 ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน หรือข้อจำกัดเป็นบวก ที่ระดับความสัมพันธ์เป็น 0.7 วิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ส่วนกรณีอื่น ๆ วิธีการถดถอยริดจ์จะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดที่ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน เมื่อขนาดตัวอย่างเท่ากับ 50 เมื่อข้อจำกัดมีค่าเป็น 2 ที่ระดับความสัมพันธ์ 0.7 และความแปรปรวนของค่าคลาดเคลื่อนมีค่า 0.1 และ 0.5 หรือที่ระดับความสัมพันธ์ 0.7 หรือ 0.8 และความแปรปรวนมีค่า 1 และ 5 วิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดสำหรับข้อจำกัดอื่น ๆ วิธีการถดถอยริดจ์จะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดที่ทุกระดับความสัมพันธ์ และความแปรปรวนของค่าคลาดเคลื่อน ส่วนในกรณีตัวแปรอิสระ 5 ตัว ถ้าขนาดตัวอย่างเท่ากับ 30 เมื่อข้อจำกัดมีค่าเป็น (2,2) ที่ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อนหรือเมื่อข้อจำกัดมีค่าเป็น (+,0) ที่ทุกระดับความสัมพันธ์แต่ความแปรปรวนของค่าคลาดเคลื่อนมากกว่า0.1 วิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด นอกจากนี้เมื่อข้อจำกัดเป็น (+,+) ที่ระดับความสัมพันธ์เป็น (0.7,0.7) และ (0.7,0.8) และความแปรปรวนของค่าคลาดเคลื่อนมีค่ามากกว่า 0.1 วิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดเช่นเดียวกัน สำหรับข้อจำกัดอื่น ๆ วิธีการถดถอยริดจ์จะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดที่ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน เมื่อขนาดตัวอย่างเท่ากับ 50 ที่ระดับความสัมพันธ์ของตัวแปรสองกลุ่มมีค่าเท่ากัน และข้อจำกัดมีค่าเดียวกัน วิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด เมื่อข้อจำกัดมีค่าเป็น (2,2) ที่ทุกระดับความสัมพันธ์และความแปรปรวนของความคลาดเคลื่อนหรือเมื่อข้อจำกัดมีค่าเป็น (+,+) เมื่อระดับความสัมพันธ์เป็น (0.7,0.7) และความแปรปรวนของค่าคลาดเคลื่อนมีค่ามากกว่า 0.1 ส่วนข้อจำกัดที่มีค่าเดียวกันอื่น ๆ วิธีการถดถอยริดจ์จะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดสำหรับทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน กรณีที่ข้อจำกัดมีค่าต่างกันวิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดเมื่อข้อจำกัดเป็น (+,0) ที่

ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อนมากกว่า 0.1 หรือเมื่อข้อจำกัดมีค่าเป็น (+,+) ที่ระดับความสัมพันธ์ (0.7,0.7) และทุกค่าความแปรปรวนของค่าคลาดเคลื่อน และที่ระดับความสัมพันธ์ (0.8,0.8) เมื่อความแปรปรวนของค่าคลาดเคลื่อนมีค่ามากกว่า 0.5 ส่วนข้อจำกัดที่มีค่าต่างกันอื่น ๆ วิธีการถดถอยริดจ์แบบมีข้อจำกัดจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดที่ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน เมื่อขนาดตัวอย่างเท่ากับ 50 และระดับความสัมพันธ์มีค่าต่างกัน ถ้าข้อจำกัดมีค่าเป็น (-,-) วิธีการถดถอยริดจ์ จะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดที่ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน ส่วนข้อจำกัดอื่น ๆ วิธีการถดถอยริดจ์แบบทางเลือกจะให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าต่ำสุดที่ทุกระดับความสัมพันธ์และความแปรปรวนของค่าคลาดเคลื่อน

ยุภาวดี สำราญฤทธิ์ (2552) ได้เปรียบเทียบวิธีการแก้ไขปัญหาอัตตสหสัมพันธ์อันดับหนึ่งของค่าคลาดเคลื่อนในสมการถดถอยเชิงเส้นอย่างง่าย โดยใช้ 3 วิธี ได้แก่ วิธี Cochrane-Orcutt วิธี Prais-Winsten และวิธี Hildreth-Lu และตรวจสอบความเหมาะสมของการพยากรณ์ที่ได้จากแต่ละวิธี โดยกำหนดขนาดตัวอย่างเป็น 10, 30 และ 100 และระดับสัมประสิทธิ์อัตตสหสัมพันธ์เป็น 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 และ 1.0 ข้อมูลที่ใช้ในการศึกษานี้เป็นข้อมูลที่ได้จากการจำลองด้วยเทคนิคมอนติคาร์โล รวม 30 กรณี แต่ละกรณีทำการจำลองซ้ำ 2,000 ครั้ง โดยใช้สถิติทดสอบเดอร์บิน-วัตสัน ในการตรวจสอบว่าค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์หรือไม่ โดยเกณฑ์ที่ใช้ในการเปรียบเทียบความสามารถในการแก้ปัญหา คือ วิธีที่มีจำนวนชุดข้อมูลที่แก้ปัญหาอัตตสหสัมพันธ์ได้มากที่สุดจะเป็นวิธีที่แก้ปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนที่ดีที่สุด ส่วนเกณฑ์ที่ใช้ในการเปรียบเทียบความเหมาะสมในการพยากรณ์ คือ การพิจารณาค่าคลาดเคลื่อนกำลังสองเฉลี่ย วิธีใดที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยโดยเฉลี่ยต่ำที่สุด จะเป็นวิธีที่เหมาะสมที่สุดในการพยากรณ์ ผลการวิจัยพบว่า เมื่อขนาดตัวอย่างมีขนาดเท่ากับ 10 ความสามารถในการแก้ปัญหาอัตตสหสัมพันธ์ของวิธี Cochrane-Orcutt ดีที่สุดที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ทุกระดับ ยกเว้นที่ระดับ 0.8 สามารถแก้ปัญหาได้มากกว่าร้อยละ 68.65 และวิธี Hildreth-Lu ดีที่สุดที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์เท่ากับ 0.8 สามารถแก้ปัญหาได้มากกว่าร้อยละ 69.65 เมื่อขนาดตัวอย่างมีขนาดเท่ากับ 30 ความสามารถในการแก้ปัญหาอัตตสหสัมพันธ์ของวิธี Cochrane-Orcutt ดีที่สุดที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์เท่ากับ 0.2 ถึง 1.0 สามารถแก้ปัญหาได้มากกว่าร้อยละ 88.00 และวิธี Prais-Winsten ดีที่สุดที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์เท่ากับ 0.1 สามารถแก้ปัญหาได้มากกว่าร้อยละ 99.6 เมื่อขนาดตัวอย่างมีขนาดเท่ากับ 100 ความสามารถในการแก้ปัญหาอัตตสหสัมพันธ์ของวิธี Cochrane-Orcutt ดีที่สุดที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ทุกระดับ ยกเว้นที่ระดับ 0.6 และ 0.9 สามารถแก้ปัญหาได้มากกว่าร้อยละ 90.05 และวิธี Prais-Winsten ดีที่สุดที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์เท่ากับ 0.1, 0.2, 0.3, 0.6, 0.9 และ 1.0 สามารถแก้ปัญหาได้มากกว่าร้อยละ 90.05 วิธี Cochrane-Orcutt เป็นวิธีที่เหมาะสมที่สุดในการพยากรณ์เมื่อขนาดตัวอย่าง 10 ที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ 0.2, 0.3 และ 0.8 ส่วนวิธี Prais-Winsten เป็นวิธีที่เหมาะสมที่สุดในการพยากรณ์เมื่อขนาดตัวอย่าง 10 ที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ 0.5, 0.7 และ 0.9 เมื่อขนาดตัวอย่าง 30 ที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ 0.1 ถึง 0.4 และวิธี Hildreth-Lu เป็นวิธีที่เหมาะสมที่สุดในการพยากรณ์ เมื่อขนาดตัวอย่าง 10 ที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ 0.1, 0.4, 0.6 และ 1.0

เมื่อขนาดตัวอย่าง 30 ที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ 0.5 ถึง 1.0 และเมื่อขนาดตัวอย่าง 100 ที่ระดับสัมประสิทธิ์อัตตสหสัมพันธ์ 0.1ถึง 1.0

วราฤทธิ์ พานิชกิจโกศลกุล และสุภารัตน์ นิวิศพงศ์ (2552) ได้เปรียบเทียบวิธีการพยากรณ์ข้อมูลล่วงหน้าหนึ่งคาบเวลาสำหรับตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง เมื่อข้อมูลมีค่าผิดปกติ ด้วยวิธีการพยากรณ์ 3 วิธี คือ วิธีการพยากรณ์โดยใช้วิธีกำลังสองน้อยที่สุดแบบค่าเฉลี่ยเวียนเกิด (RM) วิธีการพยากรณ์โดยใช้วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิด (RMD) และวิธีการพยากรณ์โดยใช้วิธีกำลังสองน้อยที่สุดแบบมัธยฐานเวียนเกิดปรับปรุง (IRMD) กำหนดขนาดตัวอย่าง 4 ระดับ คือ 25, 50, 100 และ 250 ความแปรปรวนของค่าคลาดเคลื่อนสุ่ม (σ_a^2) เท่ากับ 1 ร้อยละของค่าผิดปกติเท่ากับ 5 และ 10 ขนาดของค่าผิดปกติเท่ากับ $3\sigma_a$ และ $5\sigma_a$ ในการวิจัยครั้งนี้ใช้การจำลองแบบมอนติคาร์โล และทำการทำลองซ้ำ ๆ กัน 10,000 ครั้งในแต่ละสถานการณ์ เพื่อคำนวณค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าพยากรณ์ (PMSE) ผลการวิจัยสรุปได้ว่า วิธี RM ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าพยากรณ์ต่ำที่สุด เมื่อค่า ρ ไม่เข้าใกล้ 1 ในทุกกรณีที่ศึกษา วิธี RMD ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าพยากรณ์ต่ำกว่าวิธี RM เล็กน้อย เมื่อค่า ρ เข้าใกล้ 1 ในเกือบทุกกรณีที่ศึกษา และวิธี IRMDให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของค่าพยากรณ์ต่ำที่สุดเมื่อค่า ρ เข้าใกล้ 1 ในทุกกรณีที่ศึกษา

วริดา พลาศรี (2552) ได้ศึกษาเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระโดยเปรียบเทียบวิธีการถดถอยริดจ์ และวิธีการถดถอยบูตสแตรปริดจ์ เกณฑ์การเปรียบเทียบที่ใช้ สำหรับการประมาณค่าแบบจุด คือค่าคลาดเคลื่อนกำลังสองเฉลี่ยและสำหรับการประมาณค่าแบบช่วง คือค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น ซึ่งแบ่งเป็น 2 ขั้นตอนย่อย ขั้นแรกพิจารณาค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากแต่ละวิธีมีค่าไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ขั้นต่อไปทำการเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น โดยทำการศึกษาภายใต้เงื่อนไขของการแจกแจงของค่าคลาดเคลื่อนเป็นแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1, 5 และ 10 จำนวนตัวแปรอิสระที่ใช้เท่ากับ 2, 3, 4 และ 5 ขนาด ตัวอย่างเท่ากับ 15, 30, 50 และ 100 โดยแบ่งระดับความสัมพันธ์ระหว่างตัวแปรอิสระเป็นต่ำ (0.3) ปานกลาง (0.6) และสูง (0.9) วิธีการประมาณค่า k ซึ่งเป็นค่าพารามิเตอร์ริดจ์มี 3 วิธี คือวิธี KS วิธี New HKB และวิธี New LW และในการประมาณค่าแบบช่วงกำหนดค่าสัมประสิทธิ์ความเชื่อมั่นเท่ากับ 0.90, 0.95 และ 0.99 ตามลำดับ การวิจัยครั้งนี้ได้ทำการจำลองโดยใช้เทคนิคมอนติคาร์โลซึ่งกระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์ ด้วยโปรแกรม R เวอร์ชัน 2.8.1 ผลการวิจัยสามารถสรุปได้ดังนี้ 1. กรณีการประมาณค่าแบบจุด พบว่ามากกว่า 97% ของจำนวนสถานการณ์จำลองทั้งหมด วิธีการถดถอยบูตสแตรปริดจ์ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุด โดยที่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเพิ่มขึ้น เมื่อระดับความสัมพันธ์ ส่วนเบี่ยงเบนมาตรฐาน และจำนวนตัวแปรอิสระเพิ่มขึ้น และค่าคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น 2. กรณีการประมาณค่าแบบช่วง พบว่า จากจำนวนสถานการณ์จำลองทั้งหมด วิธีการถดถอยริดจ์ให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าที่กำหนดมากกว่าวิธีการถดถอยบูตสแตรปริดจ์ และ 66% ของสถานการณ์จำลองทั้งหมด วิธีการถดถอยบูตสแตรปริดจ์ให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่ำกว่าวิธีการถดถอยริดจ์

นันทวัน จุลมุสิก (2553) ได้พัฒนาสูตรเพื่อประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณ เมื่อค่าคลาดเคลื่อนมีอัตตสัมพันธ์และมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ซึ่งทำการแก้ปัญหาทั้งสองอย่างนี้ไปพร้อม ๆ กัน และพัฒนาเป็นตัวประมาณตัวใหม่ได้ ดังนี้

$$\hat{\beta}_{NN} = (H'V^{-1}H + kI)^{-1}(H'V^{-1}N + kJ)$$

และได้ศึกษาเปรียบเทียบตัวประมาณค่าสัมประสิทธิ์การถดถอยจากการศึกษากับข้อมูลทางด้านเศรษฐศาสตร์ โดยเปรียบเทียบกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จาก 4 วิธี คือ กำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ และวิธีของ Bayhan พบว่าค่าของสมาชิกในเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป และวิธีการถดถอยริดจ์มีค่ามากกว่าวิธีที่พัฒนาขึ้นใหม่ แต่สำหรับวิธีของ Bayhanนั้นจะให้ค่าใกล้เคียงกัน แต่อย่างไรก็ตามวิธีที่พัฒนาขึ้นใหม่นี้ให้ค่าของสมาชิกในเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอยมีเครื่องหมายที่ถูกต้อง

เบญจวรรณ ระหงษ์ (2554) ได้ศึกษาเปรียบเทียบการประมาณค่าสัมประสิทธ์การถดถอยเชิงเส้นพหุคูณ กรณีเกิดการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ โดยวิธีการถดถอยส่วนประกอบหลัก วิธีการถดถอยริดจ์ที่มีค่าเบื้องต้น และวิธีการถดถอยมูนิชไคเบรียริดจ์เกณฑ์ที่ใช้ในการเปรียบเทียบได้แก่ ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าเฉลี่ยความเอนเอียงของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ปัจจัยที่กำหนดในการศึกษามี 4 ปัจจัย คือระดับพหุสัมพันธ์ระหว่างตัวแปรอิสระ 4 ระดับ ขนาดตัวอย่าง 5 ขนาด ความแปรปรวนของค่าคลาดเคลื่อนสุ่ม 4 ระดับ และระดับความต่างของค่าคงที่และค่าสัมประสิทธิ์การถดถอย 6 แบบ ทำการศึกษาจำลองแบบโดยการทำซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ ผลการศึกษาภายใต้เกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ สรุปได้ว่า วิธีการถดถอยริดจ์ที่มีค่าเบื้องต้นจะมีประสิทธิภาพสูงสุด ที่ทุกระดับพหุสัมพันธ์ ทุกขนาดตัวอย่าง และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่าน้อย ส่วนวิธีการถดถอยส่วนประกอบหลักจะมีประสิทธิภาพสูงสุดบางกรณีเช่นที่ระดับพหุสัมพันธ์สูง ที่ขนาดตัวอย่างเล็ก และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่าน้อยและบางกรณีที่ระดับพหุสัมพันธ์สูงที่ขนาดตัวอย่างใหญ่ และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามากโดยที่วิธีการถดถอยมูนิชและไคเบรียริดจ์จะมีประสิทธิภาพสูงสุดบางกรณี เช่นที่ขนาดตัวอย่างเล็กและความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก และบางกรณีที่ระดับพหุสัมพันธ์ต่ำและขนาดตัวอย่างใหญ่ และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก ผลการศึกษาภายใต้เกณฑ์ค่าเฉลี่ยความเอนเอียงของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณสรุปได้ว่า วิธีการถดถอยริดจ์ที่มีค่าเบื้องต้นจะมีประสิทธิภาพสูงสุด ที่ทุกระดับพหุสัมพันธ์ ทุกขนาดตัวอย่าง และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่าน้อย ส่วนวิธีการถดถอยส่วนประกอบหลักที่มี 5 องค์ประกอบจะมีประสิทธิภาพสูงสุดบางกรณีเช่นที่ขนาดตัวอย่างเล็กและความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามากและบางกรณีที่ระดับพหุสัมพันธ์สูงและขนาดตัวอย่างเล็ก และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก โดยที่วิธีการถดถอยมูนิชและไคเบรียริดจ์จะมีประสิทธิภาพสูงสุดบางกรณี เช่นที่ขนาดตัวอย่างใหญ่ และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก และบางกรณีที่ระดับพหุสัมพันธ์ต่ำและ

ขนาดตัวอย่างเล็ก และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก ผลการศึกษาภายใต้เกณฑ์ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าเฉลี่ยความเอนเอียงของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณสรุปได้ว่า วิธีการถดถอยริดจ์ที่มีค่าเบื้องต้นจะมีประสิทธิภาพสูงสุดที่ทุกระดับความสัมพันธ์ทุกขนาดตัวอย่าง และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่าน้อย ส่วนวิธีการถดถอยมูนิชและไคเบรียริดจ์จะมีประสิทธิภาพสูงสุดบางกรณีที่ระดับพหุสัมพันธ์ต่ำและขนาดตัวอย่างเล็ก และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก และบางกรณีที่ระดับพหุสัมพันธ์สูงและขนาดตัวอย่างใหญ่ และความแปรปรวนของค่าคลาดเคลื่อนสุ่มมีค่ามาก

กิตติ วารีรัตน์ (2555) ได้ศึกษาการแก้ไขปัญหาการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์รุนแรงด้วยวิธีการวิเคราะห์ส่วนประกอบหลัก ได้ทำการศึกษาในขอบเขตของจำนวนตัวแปรอิสระ 2 และ 3 ตัว ที่มีการแจกแจงแบบปกติมีค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 2 และความแปรปรวนของค่าคลาดเคลื่อนเท่ากับ 2 โดยที่ค่าสัมประสิทธิ์การถดถอยมีค่าเท่ากับ 2 และมีขนาดตัวอย่างเท่ากับ 50, 100, 200 ซึ่งระดับความสัมพันธ์ของตัวแปรอิสระ 2 ตัว คือ 0.90, 0.91, 0.92,..., 0.99 และสำหรับตัวแปรอิสระ 3 ตัว จะพิจารณาค่าระดับความสัมพันธ์สูงสุดจากค่าระดับความสัมพันธ์ภายใต้ขอบเขตที่ทำการศึกษา ผลการศึกษากรณีตัวแปรอิสระ 2 ตัวพบว่า ระดับความสัมพันธ์ตั้งแต่ 0.95–0.99 จะเกิดปัญหาการมีพหุสัมพันธ์รุนแรงของทุก ๆ ขนาดตัวอย่างที่ทำการศึกษา ดังนั้น ระดับความสัมพันธ์ 0.95 คือจุดเปลี่ยนที่ทำให้เกิดปัญหาการมีพหุสัมพันธ์รุนแรง และกรณีตัวแปรอิสระ 3 ตัว พบว่าระดับความสัมพันธ์สูงสุดเท่ากับ 0.93 , 0.95, 0.97, 0.99 จะเกิดปัญหาการมีพหุสัมพันธ์รุนแรง เมื่อขนาดตัวอย่างเท่ากับ 50, 100 และระดับความสัมพันธ์สูงสุดเท่ากับ 0.91, 0.93, 0.95, 0.97, 0.99 จะเกิดปัญหาการมีพหุสัมพันธ์รุนแรง เมื่อขนาดตัวอย่างเท่ากับ 200 ดังนั้น ระดับความสัมพันธ์ 0.93 คือจุดเปลี่ยนที่ทำให้เกิดปัญหาการมีพหุสัมพันธ์รุนแรง เมื่อขนาดตัวอย่างเท่ากับ 50, 100 และ ระดับความสัมพันธ์ 0.95 คือจุดเปลี่ยนที่ทำให้เกิดปัญหาการมีพหุสัมพันธ์รุนแรง เมื่อขนาดตัวอย่างเท่ากับ 200 ในส่วนของการแก้ไขปัญหาการมีพหุสัมพันธ์ด้วยวิธีการวิเคราะห์ส่วนประกอบหลัก พบว่าค่าเฉลี่ยค่าคลาดเคลื่อนมาตรฐานก่อนการแก้ไขปัญหาการมีพหุสัมพันธ์รุนแรงมีค่ามากกว่าหลังการแก้ไขปัญหาการมีพหุสัมพันธ์รุนแรง แต่ค่าเฉลี่ยความเอนเอียงก่อนการแก้ไขปัญหาการมีพหุสัมพันธ์รุนแรงมีค่าน้อยกว่าหลังการแก้ไขปัญหาการมีพหุสัมพันธ์รุนแรง

เด่นนภา จุลเพชร (2555) ได้ประมาณค่าพารามิเตอร์ที่เหมาะสม โดยการเปรียบเทียบจากค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ สำหรับตัวแบบการถดถอยเชิงเส้นพหุคูณของวิธีกำลังสองน้อยสุดเชิงส่วน (PLS) และวิธีกำลังสองน้อยสุดด้วยวิธีการแปลงข้อมูลแบบกรามชมิดต์ (OLS_G) เมื่อเกิดพหุสัมพันธ์รุนแรงระหว่างตัวแปรอิสระ ซึ่งทำการศึกษาภายใต้ตัวแปรอิสระมีการแจกแจงแบบปกติค่าเฉลี่ยเท่ากับ 0 ค่าส่วนเบี่ยงเบนมาตรฐานของตัวแปรอิสระเท่ากับ 2 และค่าความแปรปรวนของค่าคลาดเคลื่อนเท่ากับ 10 โดยมีจำนวนตัวแปรอิสระ (p) เท่ากับ 2 และ 3 ตัว มีขนาดตัวอย่าง 50, 100 และ 200 มีค่าสัมประสิทธิ์การถดถอยเบื้องต้น $\beta_0 = \beta_1 = ... = \beta_p = 1$ และระดับความสัมพันธ์ระหว่างตัวแปรอิสระแบ่งเป็น 2 กรณี คือ กรณีที่มีตัวแปรอิสระ 2 ตัวแปร มีระดับความสัมพันธ์ (ρ) เป็น 0.9, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98 และ 0.99 ส่วนกรณีที่มีตัวแปรอิสระ 3 ตัวแปร มีระดับความสัมพันธ์ (ρ) เป็น 0.91, 0.93, 0.95, 0.97 และ 0.99 ซึ่ง

ในการศึกษาจะทำซ้ำจำนวน 1,000 รอบ ผลการศึกษาพบว่ากรณี p เท่ากับ 2 ตัวแปร จะเกิดปัญหาพหุสัมพันธ์รุนแรงที่ ρ ตั้งแต่ 0.95 ขึ้นไป และกรณี p เท่ากับ 3 ตัวแปร ที่ n เท่ากับ 50 และ 100 จะเกิดปัญหาพหุสัมพันธ์รุนแรงที่ ρ ตั้งแต่ 0.93 ขึ้นไป ส่วนที่ n เท่ากับ 200 จะเกิดปัญหาพหุสัมพันธ์รุนแรงที่ ρ ตั้งแต่ 0.95 ขึ้นไป และทั้งสองกรณีมีเปอร์เซ็นต์การเกิดปัญหาการมีพหุสัมพันธ์รุนแรงสูง ซึ่งส่งผลกระทบต่อการประมาณค่าสัมประสิทธิ์การถดถอย ดังนั้นกรณี p เท่ากับ 2 ตัวแปร ที่ n เท่ากับ 50, 100 และ 200 ตามลำดับ และที่ ρ เท่ากับ 0.9-0.94, 0.9-0.97 และ 0.9-0.98 ตามลำดับ วิธี OLS_G มีประสิทธิภาพดีกว่าวิธี PLS แต่ที่ ρ เท่ากับ 0.95-0.99, 0.98-0.99 และ 0.99 ตามลำดับ วิธี PLS มีประสิทธิภาพดีกว่าวิธี OLS_G และกรณี p เท่ากับ 3 ตัว วิธี PLS มีประสิทธิภาพดีกว่าวิธี OLS_G ทุกกรณีที่ทำการศึกษา โดยปัจจัยที่มีผลต่อค่าคลาดเคลื่อนกำลังสองเฉลี่ยของทั้งสองกรณีนั้นพบว่าเมื่อ ρ เพิ่มขึ้น ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธี PLS ไม่มีแนวโน้มที่แน่นอน ยกเว้นที่ p เท่ากับ 3 และ n เท่ากับ 200 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยจะมีค่าลดลงเล็กน้อย ในขณะที่วิธี OLS_G จะมีค่าเพิ่มขึ้น และถ้า n เพิ่มขึ้น ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของทั้งสองวิธีจะมีค่าลดลง

อรพรรณ ตันตระกูล (2555) ได้เปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่แกร่งสำหรับการถดถอยเชิงเส้นพหุคูณ เมื่อข้อมูลมีค่าผิดปกติ โดยเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยในตัวแบบการถดถอยเชิงเส้นพหุคูณด้วยวิธีการประมาณ 3 วิธี คือ วิธีกำลังสองน้อยที่สุดสามัญ (OLS) วิธี Least Trimmed Squares estimator (LTS) และวิธี Generalized M-estimator (GM) เมื่อข้อมูลมีค่าผิดปกติ โดยศึกษาข้อมูลจำลองด้วยเทคนิคมอนติคาร์โล กระทำซ้ำ 1,000 ครั้ง สำหรับแต่ละสถานการณ์ การจำลองข้อมูลกำหนดให้ข้อมูลมีตัวแปรอิสระ 2 ตัว ค่า β_0, β_1 และ β_2 เท่ากับ 1 และการแจกแจงของค่าคลาดเคลื่อน 2 แบบ คือ แบบทีและลอกนอร์มอล ตัวแปรอิสระมีระดับค่าผิดปกติ 2 ระดับ คือไม่รุนแรงและรุนแรง โดยแต่ละระดับจะมีอัตราส่วนค่าผิดปกติของตัวแปรอิสระเท่ากับ 0.10, 0.15 และ 0.20 ขนาดตัวอย่าง 4 ขนาดคือ 20, 50, 100 และ 200 เกณฑ์ที่ใช้การเปรียบเทียบคือค่าคลาดเคลื่อนกำลังสองเฉลี่ยผลการวิจัยพบว่า กรณีที่ไม่มีค่าผิดปกติในตัวแปรอิสระและตัวแปรตาม วิธี OLS ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดในทุกสถานการณ์ กรณีที่มีค่าผิดปกติในตัวแปรตามและกรณีที่มีค่าผิดปกติตัวแปรอิสระและตัวแปรตาม เมื่อค่าคลาดเคลื่อนมีการแจกแจงแบบทีที่องศาเสรีเท่ากับ 1 และ 4 โดยทั่วไปวิธี GM ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด แต่เมื่อองศาความอิสระเพิ่มขึ้นเท่ากับ 8 วิธี OLS ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด และเมื่อค่าคลาดเคลื่อนมีการแจกแจงแบบลอกนอร์มอล โดยทั่วไปวิธี LTS ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด โดยค่าคลาดเคลื่อนกำลังสองเฉลี่ยจะแปรผกผันกับขนาดตัวอย่าง อัตราส่วนค่าผิดปกติของตัวแปรอิสระ ระดับค่าผิดปกติในตัวแปรอิสระ องศาเสรีของค่าคลาดเคลื่อน และส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อนที่ขนาดตัวอย่างเท่ากับ 100 และ 200 สำหรับวิธี LTS และวิธี GM แต่จะแปรผันตามค่าส่วนเบี่ยงเบนมาตรฐานของค่าคลาดเคลื่อน

Bagheriet al. (2010) ได้ศึกษาวิธีการประมาณค่าพารามิเตอร์สำหรับรูปแบบการถดถอยเชิงเส้นในกรณีที่ข้อมูลมีค่าผิดปกติ โดยการเปรียบเทียบประสิทธิภาพกับวิธีกำลังสองน้อยที่สุดสามัญ และทำการตรวจสอบการประมาณค่าพารามิเตอร์ของวิธีต่าง ๆ ด้วยค่า Distance to Maximum Normal Curvature Direction (DMNCD) และค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ได้แก่ วิธี M

วิธี MM วิธี GM1 วิธี GM6 และวิธีปรับ GM6 กำหนดให้ค่าคลาดเคลื่อนมีการแจกแจงแบบปกติและการแจกแจงแบบหางหนา (Heavy Tailed Distribution) จากการศึกษาพบว่าวิธี GM1 วิธี GM6 และวิธีปรับ GM6 ให้ผลการประมาณค่าพารามิเตอร์ในรูปแบบการถดถอยเป็นที่น่าพอใจ เนื่องจากเป็นตัวประมาณที่มีความแกร่งต่อค่าอิทธิพล (Leverage) ทั้งในแกน X และแกน Y โดยวิธี GM6 ให้ค่า DMNCD และค่าคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุด ในบรรดาวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีอื่น ๆ นอกจากนี้ยังได้นำข้อมูลจริงมายืนยันผลการศึกษาโดยใช้ข้อมูลเนื้อเยื่อในปอดอักเสบ (Interstitial Lung Disease) พบว่าการประมาณค่าพารามิเตอร์ด้วยวิธี GM6 และวิธีปรับ GM6 มีประสิทธิภาพมากกว่าวิธีประมาณอื่น ๆ เนื่องจากวิธีดังกล่าวคำนึงถึงการถ่วงน้ำหนักข้อมูลตามค่าอิทธิพล ดังนั้นการประมาณด้วยวิธี GM6 และวิธีปรับ GM6 จึงเป็นทางเลือกที่ดีสำหรับการประมาณค่าพารามิเตอร์ที่มีความแกร่ง

Alma (2011) ได้เปรียบเทียบการประมาณค่าพารามิเตอร์จากรูปแบบการถดถอยเชิงเส้นภายใต้ข้อมูลที่มีค่าผิดปกติ และค่าอิทธิพล (Leverage) โดยใช้วิธีประมาณ 5 วิธี คือ วิธีกำลังสองน้อยที่สุดสามัญวิธี LTS วิธี M วิธี MM และวิธี S จากสถานการณ์ต่าง ๆ กัน และทำการตรวจสอบประสิทธิภาพของวิธีการประมาณด้วยค่าประสิทธิภาพสัมพัทธ์และศึกษาในส่วนของชุดข้อมูลที่มีสัดส่วนของค่าผิดปกติและค่าอิทธิพลที่แตกต่างกัน จากการศึกษาดังกล่าวพบว่า วิธี MM จะมีประสิทธิภาพต่ำเมื่อมีค่าอิทธิพลสูงและจำนวนตัวแปรอิสระน้อย ซึ่งเป็นการแสดงจุดอ่อนของวิธี MM ส่วนวิธี M และวิธี Sจะให้ผลที่น่าพอใจมากกว่าวิธี LTS และวิธี MM ทั้งนี้กล่าวได้ว่าวิธี M ที่ใช้เกณฑ์ความแกร่งของ Huber จะเป็นตัวประมาณที่ดี ถ้าลักษณะของข้อมูลมีค่าผิดปกติ แต่ไม่มีค่าอิทธิพลและวิธี S จะมีประสิทธิภาพดีกว่าวิธี MM และวิธี LTS สำหรับกรณีที่ชุดข้อมูลมีสัดส่วนของค่าผิดปกติต่ำในตัวแปรตาม

Drerny and Rashwan (2011) ได้ศึกษาวิธีแก้ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระด้วยแบบจำลองการถดถอยริดจ์ โดยศึกษาเปรียบเทียบทั้งหมด 4 วิธี ได้แก่วิธีกำลังสองน้อยที่สุดสามัญ (Ordinary Least Square: OLS) วิธีการถดถอยริดจ์สามัญ (Ordinary Ridge Regression: ORR) วิธีการถดถอยริดจ์ทั่วไป (Generlized Ridge Regression: GRR) วิธีการถดถอยริดจ์แบบระบุทิศทาง (Directed Ridge Regression: DRR) โดยวิธี ORR ใช้ค่า k ของ โฮเอิร์น เคนนาร์ด และบาลด์วิน(Hoerl, Kennard and Baldwin: HKB) กับ ลอแลส และแวง (Lawless and Wang: LW) โดยพิจารณาจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าสัมประสิทธิ์การกำหนด (R^2) พบว่าแบบจำลองริดจ์ทุกวิธีให้ค่าส่วนเบี่ยงเบนมาตรฐาน และค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธีกำลังสองน้อยที่สุดสามัญ โดยวิธีการถดถอยริดจ์ทั่วไปให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด และเมื่อพิจารณาจากค่าสัมประสิทธิ์การกำหนดพบว่าแบบจำลองการถดถอยริดจ์ทุกวิธีให้ค่าสัมประสิทธิ์การกำหนดสูงกว่าวิธีกำลังสองน้อยที่สุดสามัญ โดยวิธีการถดถอยริดจ์สามัญที่ใช้ค่า k ของโฮเอิร์น เคนนาร์ด และบาลด์วินมีค่าสูงสุด รองลงมา คือ วิธีการถดถอยริดจ์ทั่วไปกับวิธีการถดถอยริดจ์แบบระบุทิศทาง มีค่าสัมประสิทธิ์การกำหนดเท่ากัน

จากที่กล่าวมาชี้ให้เห็นว่าการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหามีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ และมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนนั้นได้มีผู้พัฒนาและศึกษาวิธีการประมาณค่าพารามิเตอร์กรณีมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ

วิธีการประมาณค่าพารามิเตอร์กรณีข้อมูลมีค่าผิดปกติ วิธีการประมาณค่าพารามิเตอร์กรณีค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์ วิธีการประมาณค่าพารามิเตอร์กรณีมีพหุสัมพันธ์ระหว่างตัวแปรอิสระและข้อมูลมีค่าผิดปกติ และวิธีการประมาณค่าพารามิเตอร์กรณีพหุสัมพันธ์ระหว่างตัวแปรอิสระและค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์ แต่ยังไม่มีวิธีการประมาณพารามิเตอร์เมื่อเกิดปัญหามีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ และมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน ซึ่งข้อมูลทางเศรษฐศาสตร์มีความเป็นไปได้สูงที่ข้อมูลจะเกิดปัญหาทั้ง 3 แบบพร้อม ๆ กัน วิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่จึงน่าจะสามารถแก้ปัญหาดังกล่าวได้อย่างมีประสิทธิภาพ

ตอนที่ 6 งานวิจัยที่เกี่ยวข้องกับอุปสงค์การส่งออกยางพาราของไทย

ผู้วิจัยได้ศึกษาผลงานวิจัยที่เกี่ยวข้องกับอุปสงค์การส่งออกยางพาราของไทยมีรายละเอียดดังนี้

ณรงค์ คงสังข์ (2549) ได้ศึกษาปัจจัยที่มีผลกระทบต่ออุปสงค์การนำเข้ายางพาราไทยของจีน ญี่ปุ่น และมาเลเซีย โดยมีวัตถุประสงค์เพื่อวิเคราะห์ถึงปัจจัยที่มีผลกระทบต่ออุปสงค์การนำเข้ายางพาราไทยของจีน ญี่ปุ่น และมาเลเซีย ในช่วงปี พ.ศ. 2520 ถึงปี พ.ศ. 2547 และพยากรณ์อุปสงค์การนำเข้ายางพาราไทยของจีน ญี่ปุ่น และมาเลเซียในปี พ.ศ. 2548 ถึงปี พ.ศ. 2557 ด้วยสมการถดถอยเชิงเส้นพหุคูณ ผลการศึกษาพบว่าปัจจัยที่มีผลกระทบต่ออุปสงค์การนำเข้ายางพาราไทยของจีน และมาเลเซีย มีความสัมพันธ์ในทิศทางเดียวกันกับผลิตภัณฑ์มวลรวมภายในประเทศทั้งสองประเทศ ปริมาณการนำเข้ายางพาราไทยในปีก่อนหน้า และมีความสัมพันธ์ในทิศทางตรงกันข้ามกับราคาส่งออกยางพาราของไทย นอกจากนี้การนำเข้ายางพาราของจีนยังมีความสัมพันธ์ในทิศทางเดียวกันกับปริมาณผลผลิตยานพาหนะและรถจักรยานยนต์ในจีน ส่วนปัจจัยที่มีผลกระทบต่อการนำเข้ายางพาราไทยของญี่ปุ่น มีความสัมพันธ์ในทิศทางเดียวกันกับผลิตภัณฑ์มวลรวมภายในประเทศของญี่ปุ่น ปริมาณการนำเข้ายางพาราไทยของญี่ปุ่นในปีก่อนหน้า และราคาส่งออกยางพาราของไทย การที่ราคายางพาราของไทยกับการนำเข้าของญี่ปุ่นมีความสัมพันธ์ในทิศทางเดียวกันอาจเป็นเพราะการนำเข้ายางพาราของญี่ปุ่นขึ้นกับการขยายตัวของเศรษฐกิจของญี่ปุ่น ประกอบกับราคายางพาราในตลาดโลกถูกกำหนดโดยประเทศผู้บริโภคยางพารารายใหญ่ของโลกอย่างสหรัฐอเมริกาและจีน ส่วนอัตราแลกเปลี่ยนของญี่ปุ่น เยนต่อดอลลาร์สหรัฐอเมริกามีความสัมพันธ์อย่างมีนัยสำคัญทางสถิติในทิศทางตรงข้ามกัน เมื่อทำการพยากรณ์ปริมาณการนำเข้ายางพาราไทยของจีน ญี่ปุ่น และมาเลเซีย ระหว่างปี พ.ศ. 2548 ถึงปี พ.ศ. 2557 พบว่า มีปริมาณเฉลี่ย 1595.81, 647.10 และ 786.29 พันตัน ตามลำดับ หรือมีอัตราการขยายตัวเฉลี่ยร้อยละ 16.2, 3.5 และ 14.4 ต่อปี ตามลำดับ

ทัศนีย์ ลักษณะ (2549) ได้วิเคราะห์การส่งออกยางธรรมชาติของไทย โดยมีวัตถุประสงค์เพื่อศึกษาปัจจัยที่กำหนดการส่งออกยางธรรมชาติของไทยไปยังประเทศคู่ค้าที่สำคัญ คือจีน ญี่ปุ่น และสหรัฐอเมริกาโดยใช้แบบจำลองสมการถดถอยพหุคูณ ในรูปกำลังสองน้อยที่สุดสามัญ โดยใช้ข้อมูลทุติยภูมิในช่วงปี 2534 ถึงปี พ.ศ. 2548 ผลการวิเคราะห์การส่งออกยางธรรมชาติไปจีนพบว่าราคาส่งออกยางแท่งของไทยไปจีน ราคายางแท่งของอินโดนีเซียไปจีน ราคายางสังเคราะห์มีผลกระทบต่อการส่งออกยางธรรมชาติของไทยอย่างมีนัยสำคัญทางสถิติ ณ ระดับความเชื่อมั่นร้อยละ 99 ทั้ง 3 ตัวแปร ผลการวิเคราะห์การส่งออกยางธรรมชาติของไทยไปญี่ปุ่นพบว่า ราคายางสังเคราะห์ ราคาส่งออกยางแท่ง

ของอินโดนีเซียไปญี่ปุ่น และราคาส่งออกยางแท่งของไทยไปญี่ปุ่นมีผลกระทบต่อการส่งออกยางธรรมชาติของไทยอย่างมีนัยสำคัญทางสถิติ ณ ระดับความเชื่อมั่นร้อยละ 99 ร้อยละ 99 และร้อยละ 90 ตามลำดับ ส่วนผลการวิเคราะห์การส่งออกยางธรรมชาติของไทยไปสหรัฐอเมริกาพบว่า ราคายางสังเคราะห์ ราคาส่งออกยางแท่งของไทยไปสหรัฐอเมริกา และราคาส่งออกยางแท่งอินโดนีเซียไปสหรัฐอเมริกา มีผลกระทบต่อการส่งออกยางธรรมชาติของไทยอย่างมีนัยสำคัญทางสถิติ ณ ระดับความเชื่อมั่นร้อยละ 99 ร้อยละ 95 และไม่มีนัยสำคัญทางสถิติ ตามลำดับ

พลพจน์ กิจชัยสวัสดิ์ (2549) ได้ศึกษาอุปสงค์การส่งออกยางธรรมชาติของไทยไปจีน โดยมีวัตถุประสงค์เพื่อศึกษาปัจจัยต่าง ๆ ที่มีผลต่อการเปลี่ยนแปลงของอุปสงค์การส่งออกยางธรรมชาติของไทยไปจีน โดยใช้ข้อมูลทุติยภูมิแบบอนุกรมเวลา เป็นรายเดือนระหว่างปี พ.ศ. 2544 ถึงปี พ.ศ. 2548 โดยการประมาณค่าแบบจำลองด้วยวิธีกำลังสองน้อยที่สุด แบบจำลองที่ใช้ คือ แบบจำลองการถดถอยพหุคูณ สมการไม่ใช่สมการเชิงเส้น โดยใช้สมการในรูปลอกสองชั้น (Double-log Model) เพื่อหาค่าความยืดหยุ่นของอุปสงค์ ผลการศึกษาพบว่าปัจจัยที่มีผลต่อการเปลี่ยนแปลงของอุปสงค์การส่งออกยางธรรมชาติของไทยไปจีนในทิศทางเดียวกันคือ ปริมาณการผลิตยางยานยนต์ทุกชนิดของจีน ราคา F.O.B. ยางแผ่นรมควันชั้น 3 ของไทย ผลิตภัณฑ์มวลรวมภายในประเทศที่แท้จริง (Real GDP) ของจีน ปริมาณการผลิตยางแผ่นรมควันชั้น 3 ของประเทศไทย อัตราแลกเปลี่ยนเงินตราระหว่างเงินเหรียญสหรัฐอเมริกากับเงินบาท ส่วนตัวแปรราคายางแผ่นรมควันชั้น 3 ณ ตลาดล่วงหน้าโตเกียวมีอิทธิพลต่อการเปลี่ยนแปลงของอุปสงค์การส่งออกยางธรรมชาติของไทยไปจีนในทิศทางตรงกันข้าม

สุภัทรา ไชยศร (2549) ได้ศึกษาการส่งออกยางแผ่นรมควันชั้น 3 ของไทยไปสหรัฐอเมริกา มีวัตถุประสงค์เพื่อวิเคราะห์ถึงปัจจัยที่มีอิทธิพลในการกำหนดอุปสงค์ต่อการส่งออกยางแผ่นรมควันชั้น 3 ของไทยไปสหรัฐอเมริกา ข้อมูลที่นำมาใช้ในการศึกษาเป็นข้อมูลทุติยภูมิแบบอนุกรมเวลาระหว่างปี พ.ศ. 2532 ถึงปี พ.ศ. 2548 การวิเคราะห์ใช้สมการถดถอยโดยกำหนดรูปแบบกึ่งลอก (Semi-log) ด้วยวิธีกำลังสองน้อยที่สุด ผลการวิเคราะห์พบว่า ปัจจัยที่มีผลกระทบต่อปริมาณการส่งออกยางแผ่นรมควันชั้น 3 ของไทยไปสหรัฐอเมริกา คือ ปัจจัยทางด้านรายได้ประชาชาติเบื้องต้นของสหรัฐอเมริกา และปริมาณการผลิตยางรถยนต์และยางรถบรรทุก

สัญชัย บริบูรณ์ (2550) ได้ศึกษาปัจจัยที่กำหนดการส่งออกยางพาราของไทยไปตลาดจีน มีวัตถุประสงค์เพื่อ 1) ศึกษาสถานการณ์การผลิต การส่งออก และสถานการณ์ราคายางของประเทศไทย 2) วิเคราะห์ปัจจัยที่กำหนดการส่งออกยางพาราของประเทศไทยไปตลาดจีน การศึกษาครั้งนี้ใช้ข้อมูลทุติยภูมิแบบอนุกรมเวลา ตั้งแต่ พ.ศ. 2529 ถึงปี พ.ศ. 2548 โดยใช้สมการถดถอยพหุคูณโดยการประมาณค่าด้วยวิธีกำลังสองน้อยที่สุด ผลการศึกษาพบว่า ประเทศไทยผลิตและส่งออกยางพารามากที่สุดในโลก โดยผลิตและส่งออกยางแท่งมากที่สุด ส่งออกไปตลาดจีนมากที่สุด ราคายางมีแนวโน้มเพิ่มขึ้นตลอดตั้งแต่ปี พ.ศ. 2544 ถึงปี พ.ศ. 2548 ส่วนปัจจัยที่มีผลต่อการส่งออกยางพาราของไทยไปตลาดจีนมากที่สุด คือ ราคาส่งออกยางแผ่นรมควันของไทย และปัจจัยที่กำหนดการส่งออกยางแท่งของไทยไปตลาดจีน พบว่าปริมาณการผลิตยางรถยนต์ของจีนมีผลต่อการส่งออกมากที่สุด ขณะที่ปัจจัยที่กำหนดการส่งออกยางแผ่นรมควันของไทยไปตลาดจีนพบว่าจำนวนประชากรมีผลต่อการส่งออกไปตลาดจีนมากที่สุด

ณิชกุล จันทร์มณี (2551) ได้ศึกษาปัจจัยที่ทำให้เกิดความผันผวนของราคายางพาราภายในประเทศไทย มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ทำให้เกิดความผันผวนของราคายางพาราภายในประเทศไทย ในการศึกษาใช้ข้อมูลทุติยภูมิแบบอนุกรมเวลา ตั้งแต่ปี พ.ศ. 2523 ถึงปี พ.ศ. 2550 โดยการศึกษาเชิงพรรณาและเชิงปริมาณ ได้ศึกษาพิจารณาถึงคุณสมบัติโดยทั่วไปและกลยุทธ์ทางการตลาด การศึกษาเชิงปริมาณได้สร้างแบบจำลองทางสถิติด้วยสมการถดถอยพหุคูณ และประมาณค่าด้วยกำลังสองน้อยที่สุด เพื่อวิเคราะห์ความสัมพันธ์ของปัจจัยที่มีอิทธิพลต่อความผันผวนของราคายางพาราภายในประเทศไทย ผลการศึกษาพบว่า ปัจจัยที่มีอิทธิพลต่อความผันผวนของราคายางพาราภายในประเทศไทย คือการบริโภคภายในประเทศ (IP) มีความสัมพันธ์ในทิศทางตรงกันข้ามกับราคายางที่เกษตรกรขายได้ (PP) และผลิตภัณฑ์มวลรวมภายในประเทศ (GDP) และมูลค่าการส่งออกของไทย (OP) มีความสัมพันธ์ในทิศทางเดียวกันกับราคายางที่เกษตรกรขายได้ (PP) อย่างมีนัยสำคัญที่ระดับความเชื่อมั่นร้อยละ 94 เนื่องจากราคายางพารามีราคาสูงขึ้น ทำให้ปริมาณความต้องการปลูกยางพาราภายในประเทศเพิ่มขึ้นด้วย เพราะความต้องการยางพาราในตลาดโลกมีแนวโน้มสูงขึ้น ดังนั้นรัฐบาลและหน่วยงานที่เกี่ยวข้องควรทำการค้นคว้าวิจัยถึงปัญหา มีการพัฒนา และมีการปรับปรุงคุณภาพของยางพารา

ธัญญารัตน์ ไชยเนตรไกรสิน (2551) ได้ศึกษาปัจจัยที่มีผลกระทบและพยากรณ์ปริมาณการส่งออกยางพาราของไทยไปญี่ปุ่น มีวัตถุประสงค์เพื่อศึกษาสภาพการผลิตและการตลาดของอุตสาหกรรมยางพาราของไทย วิเคราะห์ปัจจัยที่มีผลกระทบต่อการส่งออกยางพาราของไทยไปญี่ปุ่น โดยใช้ข้อมูลอนุกรมเวลารายปีตั้งแต่ปี พ.ศ. 2535 ถึงปี พ.ศ. 2550 และเพื่อพยากรณ์แนวโน้มปริมาณการส่งออกยางพาราของไทยไปญี่ปุ่นในช่วงเวลา 5 ปีข้างหน้า ระหว่างปี พ.ศ. 2551 ถึงปี พ.ศ. 2555 วิธีการศึกษาเป็นการนำข้อมูลที่รวบรวมได้มาวิเคราะห์โดยกำหนดรูปแบบจำลองการส่งออกเป็นแบบกึ่งลอก (Semi-log) และใช้วิธีกำลังสองน้อยที่สุดในการประมาณค่าสัมประสิทธิ์ของความสัมพันธ์ในแบบจำลอง ผลการวิเคราะห์ปัจจัยที่มีผลต่อปริมาณการส่งออกยางพาราของไทยไปญี่ปุ่นพบว่า ผลิตภัณฑ์มวลรวมประชาชาติของญี่ปุ่น อัตราแลกเปลี่ยนเงินตราสกุลเงินบาทไทยต่อสกุลเงินเยนญี่ปุ่น และราคาส่งออกยางแผ่นรมควันชั้น 3 ของไทย มีผลกระทบต่อปริมาณการส่งออกยางพาราของไทยไปญี่ปุ่นอย่างมีนัยสำคัญทางสถิติ ณ ระดับความเชื่อมั่นร้อยละ 99 สำหรับผลการพยากรณ์ปริมาณการส่งออกยางพาราของไทยไปญี่ปุ่นปรากฏว่ามีแนวโน้มเพิ่มขึ้นทุกปี

ศราวุธ อินแป้น (2551) ได้ศึกษาความได้เปรียบโดยการเปรียบเทียบในการส่งออกยางพาราไปสหรัฐอเมริกา โดยมีวัตถุประสงค์เพื่อศึกษาความได้เปรียบโดยการเปรียบเทียบในการส่งออกยางพาราของไทยไปสหรัฐอเมริกา ซึ่งเป็นตลาดส่งออกหลัก โดยทำการศึกษาเปรียบเทียบกับประเทศผู้แข่งขันทางการค้าที่สำคัญ คืออินโดนีเซียและมาเลเซีย ช่วงระยะเวลาที่ทำการศึกษาระหว่างปี พ.ศ. 2531 ถึงปี พ.ศ. 2550 โดยใช้ข้อมูลทุติยภูมิที่รวบรวมได้จากหน่วยงานราชการและเอกชนต่าง ๆ และใช้ค่าดัชนีความได้เปรียบโดยเปรียบเทียบที่ปรากฏ (RCA) เป็นตัวชี้วัดการศึกษา ผลการศึกษาพบว่าไทยมีความได้เปรียบโดยเปรียบเทียบในการส่งออกยางพารามากกว่ามาเลเซีย แต่น้อยกว่าอินโดนีเซียตลอดช่วงที่ทำการศึกษา สาเหตุหลักเนื่องจากค่าแรงของมาเลเซียสูง ส่วนอินโดนีเซียมีความได้เปรียบโดยเปรียบเทียบมากกว่าประเทศอื่นที่ทำการศึกษาในครั้งนี้ และมีแนวโน้มความได้เปรียบโดยเปรียบเทียบเพิ่มขึ้นเรื่อย ๆ เนื่องจากแนวโน้มความต้องการยางแท่งของสหรัฐอเมริกาเพิ่มขึ้น

สมลักษณ์ หอมสิน (2551) ได้ศึกษาการผลิต และการส่งออกยางพาราในประเทศไทย โดยมีวัตถุประสงค์เพื่อศึกษาการผลิต การตลาด และการส่งออกของยางพาราในประเทศไทยโดยใช้ข้อมูลทุติยภูมิ ใช้สถิติพรรณาเพื่ออธิบายถึงสภาพทั่วไป การผลิต การตลาด และการส่งออกของยางพารา ผลการศึกษาพบว่า ปัจจุบันประเทศไทยเป็นผู้ผลิต และส่งออกยางพาราเป็นอันดับ 1 ของโลก การเคลื่อนไหวของราคายางในประเทศไทยขึ้นอยู่กับความต้องการยางพาราในตลาดโลก และสามารถนำข้อมูลในการศึกษาครั้งนี้มาใช้เป็นแนวทางในการวางแผนการซื้อขายที่กำลังจะเกิดขึ้นในอนาคตได้อย่างมีประสิทธิภาพ

จุฑารัตน์ พรหมทัต (2553) ได้วิเคราะห์ขีดความสามารถในการแข่งขันเพื่อการส่งออกยางพาราไปจีน โดยมีวัตถุประสงค์เพื่อวิเคราะห์ปัจจัยที่มีผลต่อความได้เปรียบเชิงแข่งขันของการส่งออกยางพารา โดยใช้แบบจำลองความได้เปรียบเชิงแข่งขันระบบเพชรของ ไมเคิล. อี. พอร์ทเตอร์ และวิเคราะห์ปัจจัยที่มีผลต่ออุปสงค์การนำเข้ายางแผ่นรมควัน ยางแท่ง และน้ำยางข้นของจีนจากประเทศไทย อินโดนีเซีย มาเลเซีย และเวียดนาม โดยการวิเคราะห์ข้อมูลแบบผสมตัดขวางและอนุกรมเวลา (Panel Data) ผลการวิเคราะห์พบว่า ประเทศไทยมีความมีความได้เปรียบเชิงแข่งขันของการส่งออกยางแผ่นรมควัน และน้ำยางข้น ในด้านปัจจัยการผลิต เช่น ความเพียงพอของวัตถุดิบในการผลิต ความเชี่ยวชาญในการผลิตของแรงงาน ส่วนด้านอุตสาหกรรมที่สนับสนุนและเกี่ยวข้อง พบว่าไทยเริ่มมีการพัฒนานำยางแผ่นรมควันและน้ำยางข้นมาแปรรูปเป็นผลิตภัณฑ์มากขึ้น สำหรับยางแท่งมีข้อได้เปรียบในการแข่งขันภายในอุตสาหกรรมเพื่อการส่งออก ทำให้มีการพัฒนาคุณภาพด้านผลิตภัณฑ์เพิ่มขึ้น ส่วนข้อเสียเปรียบของการส่งออกยางแผ่นรมควัน ยางแท่ง และน้ำยางข้น คือระบบขนส่งสินค้าที่มีต้นทุนขนส่งราคาสูง การบริหารงานเป็นแบบครอบครัว ทำให้การบริหารไม่เป็นระบบ โดยเฉพาะในส่วนของยางแท่ง ยังมีข้อเสียเปรียบในด้านคุณภาพในการผลิต และมีต้นทุนการผลิตที่สูงกว่าประเทศคู่แข่ง อีกทั้งยังขาดเทคโนโลยีในการผลิต ในส่วนของผลการวิเคราะห์ปัจจัยที่มีผลต่ออุปสงค์การนำเข้ายางแผ่นรมควันของจีน พบว่า ราคานำเข้าที่แท้จริง และสัดส่วนของรายได้ประชาชาติเบื้องต้นที่แท้จริงของจีนกับประเทศผู้ส่งออก มีผลต่อการนำเข้ายางแผ่นรมควันของจีน ซึ่งถือได้ว่ายางแผ่นรมควันเป็นสินค้าสามัญในตลาดประเทศจีน สำหรับยางแท่งพบว่า ราคานำเข้าที่แท้จริง และอัตราแลกเปลี่ยนเงินตราระหว่างประเทศผู้ซื้อและผู้ขาย มีผลต่อการนำเข้ายางแท่งของจีน และน้ำยางข้น ปัจจัยที่มีผลต่อการนำเข้าในตลาดประเทศจีน คืออัตราแลกเปลี่ยนเงินตราระหว่างประเทศและสัดส่วนของรายได้ประชาชาติเบื้องต้นที่แท้จริงของจีนกับประเทศผู้ส่งออก แสดงว่าน้ำยางข้นที่จีนนำเข้าจากประเทศผู้ส่งออกต่าง ๆ เป็นสินค้าสามัญ ดังนั้นประเทศไทยจำเป็นต้องพัฒนาศักยภาพในการผลิตและการส่งออกยางพาราแต่ละประเภทโดยเฉพาะยางแท่ง และพัฒนาการแข่งขันเพื่อการส่งออกยางแผ่นรมควันและน้ำยางข้นที่มีความสามารถอยู่แล้วให้ดีขึ้น โดยควรส่งเสริมด้านการวิจัยและพัฒนาคุณภาพยางพาราและผลิตภัณฑ์ยางพาราเพื่อสร้างความได้เปรียบในการแข่งขันเพื่อการส่งออกยางพาราไปจีนให้เพิ่มขึ้น

อิศรา ชูบำรุง (2554) ได้วิเคราะห์ความสามารถในการแข่งขันการส่งออกยางพาราธรรมชาติของประเทศไทย มีวัตถุประสงค์เพื่อ 1) ศึกษาสภาพทั่วไปของการผลิตและการส่งออกยางพาราของประเทศไทย 2) วิเคราะห์ความได้เปรียบโดยเปรียบเทียบในการส่งออกยางพาราธรรมชาติของประเทศไทยไปตลาดโลก 3) วิเคราะห์ส่วนแบ่งการตลาดของการส่งออกยางพาราธรรมชาติของประเทศไทยไปตลาดโลก 4) วิเคราะห์ความสัมพันธ์ของความได้เปรียบโดยเปรียบเทียบและส่วนแบ่ง

การตลาดของการส่งออกยางพาราธรรมชาติของประเทศไทยไปตลาดโลก โดยการศึกษาในครั้งนี้ใช้ข้อมูลการผลิตและส่งออกในช่วงปี พ.ศ.2548 ถึงปี พ.ศ. 2552 วิเคราะห์ข้อมูลโดยคำนวณค่าความได้เปรียบโดยเปรียบเทียบที่ปรากฏและส่วนแบ่งการตลาด ผลการศึกษาพบว่า 1) ในปีพ.ศ. 2552 ประเทศไทยมีพื้นที่เพาะปลูกยางพาราสูงเป็นอันดับสองรองจากอินโดนีเซีย และสามารถผลิตยางพาราธรรมชาติได้มากที่สุดถึง 3,164 พันตัน รองลงมาได้แก่ประเทศอินโดนีเซีย มาเลเซีย และเวียดนาม ประเทศไทยส่งออกยางพาราธรรมชาติสูงที่สุดมีมูลค่ารวม 146,263 ล้านบาท 2) ประเทศไทยมีความได้เปรียบโดยเปรียบเทียบของการส่งออกเฉลี่ยสูงสุดในสินค้ายางแผ่นรมควัน และน้ำยางข้น ประเทศอินโดนีเซียมีความได้เปรียบโดยเปรียบเทียบของสูงสุดการส่งออกในสินค้ายางแท่ง และยางชนิดอื่น ๆ ประเทศเวียดนามได้เปรียบโดยเปรียบเทียบสูงสุดในสินค้ายางผสม 3) ประเทศไทยมีส่วนแบ่งการตลาดเฉลี่ยมากที่สุดในสินค้ายางแผ่นรมควัน น้ำยางข้น และยางผสม ประเทศอินโดนีเซียมีส่วนแบ่งการตลาดเฉลี่ยมากที่สุดในสินค้ายางแท่ง และยางชนิดอื่น ๆ 4) ความได้เปรียบโดยเปรียบเทียบของการส่งออกยางพาราธรรมชาติเฉลี่ยของประเทศไทยมีค่าสูงเป็นอันดับ 2 รองจากประเทศอินโดนีเซีย ในขณะที่มีส่วนแบ่งการตลาดสูงที่สุด

พรรณรวีย์ จันทรมาศ (2556) ได้ศึกษาความสัมพันธ์ระหว่างมูลค่าการส่งออกยางพารากับรายได้ในภาคเกษตรของประเทศไทย มีวัตถุประสงค์เพื่อแสดงให้เห็นถึงความสัมพันธ์ของตัวแปรอิสระ ซึ่งได้แก่ มูลค่าการส่งออกยางพารา มูลค่าการส่งออกข้าว และระดับอัตราแลกเปลี่ยนของประเทศไทยว่ามีผลอย่างไรต่อตัวแปรตาม คือรายได้ในภาคเกษตรของประเทศไทย โดยใช้ข้อมูลทุติยภูมิรายปีตั้งแต่ปี พ.ศ. 2539 ถึงปี พ.ศ. 2553 รวมระยะเวลา 15 ปี โดยวิธีทางเศรษฐมิติ และวิเคราะห์สมการถดถอยพหุคูณด้วยวิธีประมาณค่าแบบกำลังสองน้อยที่สุดสามัญผลการศึกษาพบว่า มูลค่าการส่งออกยางพาราของประเทศไทย มีความสัมพันธ์กับรายได้ในภาคเกษตรของประเทศไทย ในทิศทางเดียวกันสอดคล้องกับสมมติฐานที่ตั้งไว้อย่างมีนัยสำคัญทางสถิติ เช่นเดียวกับมูลค่าการส่งออกข้าวของประเทศไทย ส่วนอัตราแลกเปลี่ยนไม่มีความสัมพันธ์กับการเปลี่ยนแปลงของรายได้ในภาคเกษตรของประเทศไทย

ศิริประภา สุขสำโรง (2557) ได้ศึกษาความสัมพันธ์ของปัจจัยภายในและปัจจัยภายนอกประเทศที่มีต่อการส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ไปจีน มีวัตถุประสงค์เพื่อหาความสัมพันธ์ของปัจจัยภายในและปัจจัยภายนอกประเทศที่มีต่อการส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ไปจีน โดยใช้ข้อมูลทุติยภูมิรายเดือนตั้งแต่เดือนกรกฎาคม ปี พ.ศ. 2554 ถึงเดือนมิถุนายน ปี พ.ศ. 2557 รวม 36 เดือน การวิจัยครั้งนี้ใช้การวิเคราะห์ข้อมูลด้วยเครื่องมือทางเศรษฐมิติแบบสมการถดถอยพหุคูณ และใช้วิธีการประมาณค่าสัมประสิทธิ์แบบกำลังสองน้อยที่สุดสามัญ ผลการวิจัย พบว่า ปัจจัยในแบบจำลองนี้สามารถอธิบายการเปลี่ยนแปลงปริมาณการส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ไปจีนได้ที่ร้อยละ 77.31% ทั้งนี้ปัจจัยที่มีผลมากที่สุดได้แก่ ราคาส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ของไทย ถ้าราคาส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ของไทย ลดลงร้อยละ 1 จะมีผลทำให้ปริมาณการส่งออกเพิ่มขึ้นร้อยละ 4.653 ปัจจัยที่มีผลต่อปริมาณการส่งออกรองลงมา คือ ราคายางแผ่นรมควันชั้น 3 ณ ตลาดล่วงหน้าโตเกียว ปริมาณการผลิตรถยนต์นั่งของจีน ปริมาณการผลิตยางแผ่นรมควันของไทย ราคาส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ของเวียดนาม อัตราแลกเปลี่ยนเงินตราต่างประเทศ และผลิตภัณฑ์มวลรวมในประเทศจีน ข้อสรุปของการวิจัยครั้งนี้เสนอแนะให้ภาครัฐ กลุ่มนักวิชาการ ผู้ประกอบการผลิต ผู้ส่งออกและ

ผู้เกี่ยวข้องอื่น ๆ ในทุกภาคส่วน ร่วมกันกำหนดนโยบายควบคุมปริมาณการผลิตทั้งอุตสาหกรรมภายในและส่งออก เพิ่มการศึกษาในเรื่องการแปรรูปผลิตภัณฑ์ยาง เผยแพร่ข้อมูลข่าวสาร สร้างความเข้าใจถึงสภาวการณ์สภาพปัญหาที่เกิดขึ้นและผลกระทบที่จะเกิดส่งผ่านมาถึงตลาดยางพาราไทยให้ผู้ส่งออกและเกษตรกรเข้าใจ และมีความพร้อมในการวางแผนแก้ปัญหาและรับมือกับผลกระทบที่เกิดขึ้น

ศศิธร สว่างวงศ์ (2558) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อการส่งออกยางพาราของไทยไปสู่จีน มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อการส่งออกยางพาราของประเทศไทยไปสู่จีน โดยใช้ข้อมูลทุติยภูมิประเภทอนุกรมเวลา ระหว่างปี พ.ศ. 2540 ถึงปี พ.ศ. 2556 รวมระยะเวลา 17 ปี เพื่อหาสัมประสิทธิ์ของแบบจำลอง โดยใช้วิธีประมาณค่าสัมประสิทธิ์ด้วยวิธีกำลังสองน้อยที่สุดสามัญจากการใช้โปรแกรมอีวิว (Eviews) ผลการศึกษาพบว่า ราคาส่งออกยางพาราไปสู่จีน (P) ส่งผลต่อปริมาณการส่งออกยางพาราของประเทศไทยไปสู่จีน (Q) ในทิศทางตรงกันข้าม และผลิตภัณฑ์มวลรวมในประเทศของจีน (GDP) อัตราแลกเปลี่ยนของไทยกับจีน (EX) อัตราเงินเฟ้อของจีน (IF) ส่งผลต่อปริมาณการส่งออกยางพาราของไทยไปสู่จีน (Q) ในทิศทางเดียวกัน

จากที่กล่าวมาข้างต้นผู้วิจัยจะพยากรณ์อุปสงค์การส่งออกยางพาราของไทยเฉพาะในกรณีส่งออกยางธรรมชาติชนิดน้ำยางข้นไปประเทศมาเลเซีย โดยใช้การพิจารณาจากปัจจัยที่มีผลกระทบต่ออุปสงค์การส่งออกยางพาราไทยไปมาเลเซีย ได้แก่ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ปริมาณการส่งออกยางพาราไทยไปมาเลเซียในปีที่ผ่านมา และราคาส่งออกยางพาราของไทย โดยการศึกษาครั้งนี้จะใช้ปริมาณและราคาของน้ำยางข้น

บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอย 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ภายใต้สถานการณ์ 162 สถานการณ์ โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล และพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นโดยแบ่งออกเป็น 3 ขั้นตอนดังนี้

ตอนที่ 1 การพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ตอนที่ 2 การเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม

ตอนที่ 3 การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

ตอนที่ 1 การพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ขั้นตอนนี้เป็นการดำเนินการตามวัตถุประสงค์ของงานวิจัยที่ต้องการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณขึ้นมาใหม่ กรณีการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน ผู้วิจัยมีวิธีการดำเนินการดังนี้

1. ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ในสมการถดถอยเชิงเส้นพหุคูณกรณีมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เพื่อรวบรวมแนวคิดและวิธีการต่าง ๆ ในการนำไปพัฒนาวิธีการประมาณค่าพารามิเตอร์ของการถดถอยเชิงเส้นพหุคูณแบบใหม่ โดยจำแนกได้ดังนี้

1.1 ศึกษากรณีข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ

ปี ค.ศ. 1970 Hoerl and Kennard ได้เสนอวิธีการถดถอยริดจ์ในการแก้ปัญหาตัวแปรอิสระมีความสัมพันธ์เชิงเส้นต่อกัน วิธีนี้จะให้ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้น

พหุคูณที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณ $\hat{\boldsymbol{\beta}}$ ลดลง เรียกว่าวิธีการถดถอยริดจ์ มีหลักการคือถ้าตัวแปรอิสระมีพหุสัมพันธ์เชิงเส้นกันแล้ว ค่าดีเทอร์มิแนนต์ (Determinant) ของ $\mathbf{X'X}$ จะมีค่าเข้าใกล้ศูนย์จึงมีผลทำให้ความแปรปรวนของตัวประมาณของสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญมีค่าสูงผิดปกติด้วย ดังนั้นจะทำให้ $|\mathbf{X'X}|$ เพิ่มขึ้นได้โดยการบวกค่าคงที่ k ที่มากกว่า 0 เข้ากับสมาชิกทุกตัวบนเส้นทแยงมุมของเมทริกซ์ $(\mathbf{X'X})$ ซึ่งจะทำให้ค่าเจาะจงมีค่ามากขึ้น และส่งผลให้ผลบวกกำลังสองของ $\hat{\boldsymbol{\beta}}$ มีค่าลดลง ทำให้ตัวประมาณของสัมประสิทธิ์การถดถอยถูกต้องมากขึ้น

ตัวประมาณค่า $\boldsymbol{\beta}$ ซึ่งเป็นตัวประมาณที่เอนเอียง คือ

$$\hat{\boldsymbol{\beta}}_{RR} = (\mathbf{X'X} + k\mathbf{I})^{-1}\mathbf{X'Y} \;; k > 0$$

โดยในระยะแรกไม่สามารถกำหนดค่า k ที่แน่นอนได้

ต่อมาในปี ค.ศ. 1975 Hoerl, Kennard and Baldwin ได้มีการพัฒนาการหาค่า k ให้เหมาะสมยิ่งขึ้น เป็นที่นิยมใช้และได้ผลดีวิธีหนึ่ง โดยกำหนดให้

$$k = \frac{p\hat{\sigma}^2}{\hat{\boldsymbol{\beta}}'\hat{\boldsymbol{\beta}}}$$

โดยที่ p คือ จำนวนตัวแปรอิสระในตัวแบบ

$\hat{\boldsymbol{\beta}}$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ β

$\hat{\sigma}^2$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ σ^2

$$\hat{\sigma}^2 = \frac{\left|\mathbf{Y'Y} - \hat{\boldsymbol{\beta}}'\mathbf{X'Y}\right|}{n-p}$$

ในปี ค.ศ. 1976 Swindel ได้เสนอเทคนิคการนำสารสนเทศเบื้องต้น (Prior Information) มาใช้ร่วมกับการถดถอยริดจ์ เรียกว่าวิธีการถดถอยริดจ์ที่มีค่าเบื้องต้น ทำให้ได้ตัวประมาณของสัมประสิทธิ์การถดถอย ดังนี้

$$\hat{\boldsymbol{\beta}}_{RJ} = (\mathbf{X'X} + k\mathbf{I})^{-1}(\mathbf{X'Y} + k\mathbf{J})$$

โดยที่ $\mathbf{J}$ คือเวกเตอร์ของค่าประมาณเบื้องต้นของค่าเฉลี่ยสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณจากวิธีกำลังสองน้อยที่สุดสามัญ

$$\mathbf{J} = \left(\sum_{j=1}^{p} \frac{\hat{\boldsymbol{\beta}}_{OLS_j}}{p} \right) 1_{\text{p x 1}}$$

โดย Swindel แสดงให้เห็นว่า มีค่าคงที่ k ที่ทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ที่มีค่าเบื้องต้นมีค่าน้อยกว่าวิธีกำลังสองน้อยที่สุดสามัญ แต่ Swindel ไม่ได้เสนอแนวทางในการประมาณค่า k

จนกระทั่ง ในปี ค.ศ. 1995 Crouse, Jin and Hanumara ได้พัฒนาวิธีการของ Swindel คือสามารถหาค่าประมาณ k ที่แน่นอน ที่ทำให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ที่มีค่าเบื้องต้นมีค่าน้อยกว่าวิธีกำลังสองน้อยที่สุดสามัญ และเป็นตัวประมาณค่าสัมประสิทธิ์การถดถอยที่มีคุณสมบัติไม่เอนเอียง

โดยค่าประมาณ k แยกเป็น 2 กรณีคือ

กรณีทราบค่า σ^2

$$k = \begin{cases} \dfrac{p\sigma^2}{(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})'(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) - \sigma^2 tr(\mathbf{X}'\mathbf{X})^{-1}} & ; (\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})'(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) - \sigma^2 tr(\mathbf{X}'\mathbf{X})^{-1} > 0 \\ \dfrac{p\sigma^2}{(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})'(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})} & ; \text{elsewhere} \end{cases}$$

กรณีไม่ทราบค่า σ^2

$$k = \begin{cases} \dfrac{p\hat{\sigma}^2}{(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})'(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) - \hat{\sigma}^2 tr(\mathbf{X}'\mathbf{X})^{-1}} & ; (\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})'(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) - \hat{\sigma}^2 tr(\mathbf{X}'\mathbf{X})^{-1} > 0 \\ \dfrac{p\hat{\sigma}^2}{(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})'(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J})} & ; \text{elsewhere} \end{cases}$$

1.2 กรณีเกิดปัญหาข้อมูลมีค่าผิดปกติ

ในปี ค.ศ. 1964 Huber ได้ศึกษาถึงตัวประมาณที่มีความแกร่ง (Robust Estimator) ที่เรียกว่าวิธีตัวประมาณเอ็มเป็นครั้งแรกเพื่อประมาณค่าสัมประสิทธิ์การถดถอยสำหรับกรณีข้อมูลที่มีค่าผิดปกติในตัวแปรตามหรือกรณีที่ค่าคลาดเคลื่อนมีค่าผิดปกติ ต่อมาในปี ค.ศ. 1996 Simpson and Montgomery ได้พัฒนาวิธีตัวประมาณเอ็มเป็นวิธีตัวประมาณจีเอ็ม (Generalized M-Estimator Method) เพื่อประมาณค่าสัมประสิทธิ์การถดถอยสำหรับกรณีข้อมูลมีค่าผิดปกติในตัวแปรอิสระและตัวแปรตามพร้อม ๆ กัน โดยตัวประมาณจีเอ็มมาจากหลักการพื้นฐานของการ

ประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) นั่นคือหาตัวประมาณ $\hat{\boldsymbol{\beta}}$ ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นของค่าคลาดเคลื่อนมีค่าสูงสุด ดังนี้

$$\max \sum_{i=1}^{n} \ln f(\boldsymbol{\varepsilon}_i) = \max \sum_{i=1}^{n} \ln f(Y_i - \mathbf{X_i}'\boldsymbol{\beta})$$

โดยวิธีการที่ใช้ในการแก้สมการจะอาศัยเทคนิคการทำซ้ำเพื่อหาค่าที่เหมาะสม ซึ่งวิธีที่ใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยส่วนใหญ่นิยมใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักชนิดทำซ้ำ (Iteratively Reweighted Least Square: IRLS) ของ Beaton and Huber ในปี ค.ศ. 1964 ซึ่งคำนวณวิธี IRLS จะต้องมีการกำหนดค่าเริ่มต้นของ $\hat{\boldsymbol{\beta}}$ และคำนวณ s แล้วหาน้ำหนักถ่วงที่เหมาะสมของแต่ละค่าสังเกตเพื่อหาตัวประมาณค่า $\boldsymbol{\beta}$ จากวีธีกำลังสองถ่วงน้ำหนักน้อยที่สุด (Weight Least Square: WLS)

ได้สูตรการประมาณค่า $\hat{\boldsymbol{\beta}}$ จากวิธี WLS อยู่ในรูปของ

$$\hat{\boldsymbol{\beta}}_{GM} = (\mathbf{X'WX})^{-1}(\mathbf{X'WY})$$

เมื่อ $\mathbf{W}$ เป็น เมทริกซ์ทแยงมุมที่ได้จากการทำวิธีกำลังสองถ่วงน้ำหนักน้อยที่สุดชนิดทำซ้ำ มีมิติ $n \times n$

w_i เป็น สมาชิกแนวทแยงมุมของเมทริกซ์ $\mathbf{W}$

และการคำนวณตัวประมาณที่แกร่งของสเกล (s) ซึ่งเป็นตัวประมาณที่ไม่เอนเอียงของส่วนเบี่ยงเบนมาตรฐาน ในที่นี้ใช้ตัวประมาณส่วนเบี่ยงเบนสัมบูรณ์ของค่ามัธยฐาน (Median Absolute Deviation: MAD) ซึ่งเสนอโดย Mosteller and Huber ในปี ค.ศ. 1977 และถูกปรับด้วยค่าคงที่ 1.4826 ทำให้ s เป็นตัวประมาณที่ไม่เอนเอียง เมื่อ n มีขนาดใหญ่ และการแจกแจงของค่าคลาดเคลื่อนเป็นปกติ ดังนั้นจะได้ s มีรูปแบบดังนี้

$$s = 1.4826 median\left|e_i - median(e_i)\right| \; ; i = 1, 2, ..., n$$

1.3 กรณีเกิดปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

แก้ปัญหาโดยการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดทั่วไปซึ่งอาศัยเทคนิคการแปลงข้อมูล (Transformation Technique) ภายใต้ข้อสมมติเบื้องต้นของ GLS คือ $\mathrm{E}(\mathbf{e}) = 0, \mathrm{Cov}(\mathbf{e}) = \mathrm{E}(\mathbf{ee}') = \sigma^2\mathbf{U}$

ตัวประมาณที่ไม่เอนเองของ $\hat{b}$ มีสูตรดังนี้

$$\hat{\mathbf{b}}_{GLS} = (\mathbf{X}'\mathbf{U}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{U}^{-1}\mathbf{Y})$$

1.4 กรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระร่วมกับปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ในปี พ.ศ. 2552 นันทวัน จุลมุสิก ได้พัฒนาสูตรเพื่อประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณเมื่อค่าคลาดเคลื่อนมีอัตตสัมพันธ์และมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ซึ่งเป็นการแก้ปัญหาทั้งสองอย่างนี้ไปพร้อม ๆ กัน แล้วพัฒนาเป็นตัวประมาณตัวใหม่ได้ดังนี้

$$\hat{\mathbf{b}}_{NT} = (\mathbf{H}'\mathbf{U}^{-1}\mathbf{H} + k\mathbf{I})^{-1}(\mathbf{H}'\mathbf{U}^{-1}\mathbf{N} + k\mathbf{J})$$

และได้ศึกษาเปรียบเทียบตัวประมาณค่าสัมประสิทธิ์การถดถอยจากการศึกษากับข้อมูลทางด้านเศรษฐศาสตร์ โดยเปรียบเทียบกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จาก 4 วิธี คือ กำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ และวิธีของ Bayhan พบว่าค่าของสมาชิกในเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป และวิธีการถดถอยริดจ์มีค่ามากกว่าวิธีที่พัฒนาขึ้นมาใหม่ แต่สำหรับวิธีของ Bayhan นั้นจะให้ค่าใกล้เคียงกันแต่อย่างไรก็ตาม วิธีที่พัฒนาขึ้นใหม่นี้ให้ค่าของสมาชิกในเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอยมีเครื่องหมายที่ถูกต้อง

1.5 กรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระร่วมกับปัญหาการมีข้อมูลที่มีค่าผิดปกติ

ในปี ค.ศ. 1996 Simpson and Montgomery ได้พัฒนาวิธีตัวประมาณที่มีความแกร่ง เรียกว่าวิธีตัวประมาณจีเอ็มแทนวิธีตัวประมาณเอ็มที่ Huber (1964) ได้ศึกษาไว้ แล้วนำมารวมกับการประมาณโดยวิธีการถดถอยริดจ์เรียกวิธีดังกล่าวว่า วิธีริดจ์ชนิดตัวประมาณจีเอ็ม (Generalized M-Estimator Ridge Method)

ซึ่งรูปแบบการคำนวณค่าสัมประสิทธิ์การถดถอยโดยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม คือ

$$\hat{\boldsymbol{\beta}}_{GMR} = (\mathbf{X}'\mathbf{W}^{(\mathbf{r})}\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{W}^{(\mathbf{r})}\mathbf{Y} \; ; k > 0$$

เมื่อ $\mathbf{W}^{(\mathbf{r})}$ แทนน้ำหนักรอบสุดท้ายจากเกณฑ์ความแกร่งของตัวประมาณแบบจีเอ็ม

2. ศึกษาแนวคิดและทฤษฎีการอนุมานเชิงสถิติที่เกี่ยวข้องกับการหาค่าคาดหมายทางคณิตศาสตร์ ความแปรปรวน และความแปรปรวนร่วม เพื่อเป็นแนวทางในการพัฒนาวิธีการประมาณค่าพารามิเตอร์ของการถดถอยเชิงเส้นพหุคูณ

3. ศึกษาวิธีการได้มาของวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยใช้เกณฑ์ความแกร่งของ Ramsay และงานวิจัยอื่น ๆ

4. นำความรู้ที่ได้จากข้อที่ 1 ถึงข้อที่ 3 เป็นแนวทางในการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณขึ้นมาใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

สรุปเป็นแผนผังของการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่ ดังภาพที่ 3

ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการประมาณค่าพารามิเตอร์ในสมการถดถอยกรณีที่มีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

↓

ศึกษาแนวคิด หลักการ และทฤษฎีการอนุมานเชิงสถิติที่เกี่ยวข้องกับการหาค่าคาดหมายทางคณิตศาสตร์ ความแปรปรวน และความแปรปรวนร่วมของการถดถอยเชิงเส้นพหุคูณ

↓

ศึกษาวิธีการได้มาของวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณ ด้วยวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยใช้เกณฑ์ความแกร่งของ Ramsay และงานวิจัยอื่น ๆ

↓

พัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณขึ้นมาใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ภาพที่ 3 แผนผังของการพัฒนาวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่

ตอนที่ 2 การเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวันจุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม

การศึกษาในขั้นตอนนี้เพื่อเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลและใช้โปรแกรม MATLAB เวอร์ชัน 8.7 มีหลักการดังนี้

แผนการดำเนินการวิจัย

1. ตัวแบบทั่วไป (General Model) ของความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามเชิงเส้นในรูปเมทริกซ์ คือ

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

2. กำหนดให้ตัวแปรอิสระ X และค่าคลาดเคลื่อน ε_i มีการแจกแจงปกติ (Normal Distribution) ซึ่งฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function: pdf.) ของ X อยู่ในรูปของ

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2} \;; -\infty < x < \infty$$

โดยตัวแปรอิสระและความคลาดเคลื่อน มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1

3. จำนวนตัวแปรอิสระในตัวแบบมี 3 ตัว โดยกำหนดให้ตัวแปร X_1 และ X_2 เกิดปัญหาการมีพหุสัมพันธ์กัน และตัวแปร X_3 เป็นอิสระจากตัวแปร X_1 และ X_2

4. ขนาดตัวอย่างที่ใช้ในการศึกษามี 3 ขนาด คือขนาดตัวอย่างเท่ากับ 20, 50และ 100

5. ค่าพารามิเตอร์หรือค่าสัมประสิทธิ์การถดถอยที่ใช้มีค่าดังนี้ $\beta_0 = 1, \beta_1 = 1, \beta_2 = 1,$ $\beta_3 = 1$

6. ศึกษาในกรณีที่ข้อมูลมีค่าผิดปกติเกิดขึ้นในระดับรุนแรง โดยกำหนดลักษณะของการเกิดค่าผิดปกติของข้อมูล 3 แบบ ดังนี้ ข้อมูลผิดปกติในตัวแปรอิสระ กำหนดให้เป็นตัวแปร X_3 ข้อมูลผิดปกติในตัวแปรตาม (Y) และข้อมูลผิดปกติในตัวแปรอิสระ (X_3) และตัวแปรตาม (Y)

7. ระดับความสัมพันธ์ทั้งระหว่างตัวแปรอิสระแบ่งเป็น 3 ระดับ คือระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9)

8. อัตตสหสัมพันธ์ของค่าคลาดเคลื่อนแบ่งเป็น 3 ระดับ คือระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9)

โดยระดับความสัมพันธ์ระหว่างตัวแปรอิสระและอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนในแต่ละชุด จะเป็นไปตามที่กำหนดไว้ดังต่อไปนี้ (0.3,0.3), (0.3,0.5), (0.3,0.9), (0.5,0.3), (0.5,0.5), (0.5,0.9), (0.9,0.3), (0.9,0.5) และ (0.9,0.9)

เมื่อค่าระดับความสัมพันธ์ตัวแรก หมายถึง ความสัมพันธ์ระหว่างตัวแปรอิสระ 2 ตัว คือ X_1 และ X_2 และค่าระดับความสัมพันธ์ตัวหลัง หมายถึงความสัมพันธ์ของค่าคลาดเคลื่อน

การวิจัยนี้ ศึกษาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระโดยศึกษาเฉพาะความสัมพันธ์ในทิศทางบวกเท่านั้น ส่วนการศึกษาปัญหาอัตตสหสัมพันธ์จะศึกษาเฉพาะอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนทางบวก (Positive Autocorrelation) โดยการจำลองข้อมูลให้ค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์อันดับหนึ่ง (AR(1))

9. กำหนดเปอร์เซนต์ค่าผิดปกติที่ใช้ในการศึกษา 2 ระดับ คือ 1% และ 5%
10. กำหนดจำนวนการกระทำซ้ำในแต่ละสถานการณ์เป็น 1,000 ครั้ง

ขั้นตอนดำเนินการวิจัย

การดำเนินการวิจัยมีขั้นตอน ดังนี้

1. สุ่มตัวเลขเริ่มต้นในช่วง (0,1) ซึ่งมีการแจกแจงยูนิฟอร์ม เพื่อนำไปสร้างตัวแปรสุ่มที่มีการแจกแจงปกติ ด้วยวิธีบ็อกซ์ มูลเลอร์ (Box Muller)
2. สุ่มตัวอย่างจำนวน (n) เท่ากับ 20, 50 และ 100
3. สร้างข้อมูล X และ Y ที่มีความสัมพันธ์เชิงเส้นกัน โดยกำหนดค่าระดับความสัมพันธ์ทั้งระหว่างตัวแปรอิสระและอัตตสหสัมพันธ์ กำหนดค่าผิดปกติ

3.1 สร้างค่าคลาดเคลื่อน (ε_i) ให้มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ_u^2 เมื่ออัตตสหสัมพันธ์เป็น 0 แต่เมื่อมีอัตตสหสัมพันธ์อันดับหนึ่ง (AR(1)) รูปแบบของ ε_i จะเปลี่ยนแปลงไปโดย $\varepsilon_i = \rho\varepsilon_{i-1} + u_i$ โดย ε_i มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $\frac{\sigma_u^2}{(1-\rho^2)}$ จากขอบเขตของการวิจัย กำหนดให้ $\sigma_u^2 = 1$ โดยสร้างข้อมูล u_i มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ1

3.2 ปรับค่าความคลาดเคลื่อนให้มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9

3.3 สร้างตัวแปรอิสระ (X_{ij}) ตามระดับความสัมพันธ์ต่าง ๆ ที่กำหนดไว้ ในการสร้างตัวแปรอิสระนั้นผู้วิจัยได้ใช้วิธีจำลองของ Wichern and Churchill (1978)

$$X_{ij} = (1-\rho^2)^{1/2} Z_{ij} + \rho Z_{14} \,; i = 1,2,\ldots,n \;;\; j = 1,2,3$$

เมื่อ Z_{ij} เป็นตัวเลขสุ่มที่มีการแจกแจงปกติมาตรฐานมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 และเป็นอิสระต่อกัน

ρ เป็น สหสัมพันธ์ระหว่างตัวแปรอิสระ โดยกำหนดระดับความสัมพันธ์ของตัวแปรอิสระเท่ากับ 0.3, 0.5 และ 0.9

เช่น ขนาดตัวอย่างเท่ากับ 20 ระดับความสัมพันธ์ระหว่างตัวแปรอิสระเท่ากับ 0.3 จะได้ค่าตัวแปรอิสระ ดังนี้

$$X_{ij} = \left(1-0.3^2\right)^{1/2} Z_{ij} + 0.3Z_{i4}\ ; i=1,\ 2,\ 3,\ \ldots,\ 20 \quad ; j=1,\ 2,\ 3$$

โดยที่ $Z_{i1}, Z_{i2}, Z_{i3}, Z_{i4}$ เป็นค่าของตัวแปรอิสระที่สร้างขึ้นให้มีการแจกแจงแบบปกติมาตรฐานและเป็นอิสระต่อกัน

3.4 กำหนดค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเริ่มต้น (β)=1

3.5 สร้างตัวแปรตาม (Y_i) ตามตัวแปรอิสระที่สร้างขึ้น ณ ระดับความสัมพันธ์ที่กำหนด และค่าความคลาดเคลื่อนโดยใช้รูปแบบความสัมพันธ์เชิงเส้น

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i\ ; i = 1, 2, \ldots, n$$

3.6 สร้างข้อมูลตัวแปรอิสระและตัวแปรตามที่มีค่าผิดปกติ 3 กรณีดังนี้

กรณีที่ 1 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3

กรณีที่ 2 ข้อมูลมีค่าผิดปกติในตัวแปรตาม Y

กรณีที่ 3 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y

การสร้างข้อมูลของตัวแปรอิสระและตัวแปรตามศึกษาในกรณีความผิดปกติเกิดขึ้นในระดับรุนแรง ตามเงื่อนไขการตรวจสอบค่าผิดปกติด้วยวิธีแผนภาพกล่อง (Box Plot) มีรายละเอียดดังนี้

3.6.1 ตรวจสอบว่าข้อมูลมีค่าผิดปกติหรือไม่ โดยพิจารณาจากข้อมูลที่ค่าของตัวแปรตามและตัวแปรอิสระอยู่ในช่วง $\left[Q_1 - 1.5(IQR), Q_3 + 1.5(IQR)\right]$ เป็นข้อมูลที่ไม่มีค่าผิดปกติ

3.6.2 ปรับข้อมูลตัวแปรอิสระและตัวแปรตามให้เกิดค่าผิดปกติในกรณีต่าง ๆ ดังนี้ ข้อมูลผิดปกติในตัวแปรอิสระ กำหนดให้เป็นตัวแปร X_3 ข้อมูลผิดปกติในตัวแปรตาม (Y) และข้อมูลผิดปกติในตัวแปรอิสระ (X_3) และตัวแปรตาม (Y)

การปรับข้อมูลตัวแปรอิสระ X_3 และตัวแปรตาม Y จะปรับตามระดับค่าผิดปกติเกิดขึ้นในระดับรุนแรง ตามเงื่อนไขการตรวจสอบค่าผิดปกติด้วยวิธีแผนภาพกล่องโดยในแต่ละกรณีค่าผิดปกติกำหนดให้มีเปอร์เซนต์ค่าผิดปกติเท่ากับ 1% และ 5% ซึ่งจำนวนค่าผิดปกติของแต่ละขนาดตัวอย่างที่ศึกษา แสดงในตารางที่ 7 ดังนี้

ตารางที่ 7 จำนวนค่าที่ผิดปกติในแต่ละขนาดตัวอย่าง จำแนกตามเปอร์เซนต์การเกิดค่าผิดปกติ

เปอร์เซนต์ค่าผิดปกติ	ขนาดตัวอย่าง		
	20	50	100
1%	1	1	1
5%	2	3	5

มีขั้นตอนการสร้างข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระและตัวแปรตามดังนี้

ก. นำข้อมูลของตัวแปรอิสระและ/ หรือตัวแปรตามที่ต้องการสร้างให้มีค่าผิดปกติมาเรียงลำดับจากน้อยไปหามาก คือ $R_1, R_2, R_3, ..., R_{n-1}, R_n$ ตามลำดับ

เมื่อ R_1 แทน ค่าของตัวแปรอิสระและ/ หรือตัวแปรตามที่มีค่าน้อยที่สุด

R_n แทน ค่าของตัวแปรอิสระและ/ หรือตัวแปรตามที่มีค่ามากที่สุด

ข. แทนค่าของตัวแปรอิสระและ/ หรือตัวแปรตามที่มากที่สุดด้วยค่าค่าหนึ่งให้มีค่ามากกว่าเดิม หรือแทนค่าของตัวแปรอิสระและ/ หรือตัวแปรตามที่น้อยที่สุดด้วยค่าค่าหนึ่งให้มีค่าน้อยลงกว่าเดิม เพื่อให้ได้ค่าผิดปกติในระดับที่กำหนดและมีจำนวนตามที่ต้องการ แสดงในตารางที่ 8 ดังนี้

ตารางที่ 8 จำนวนค่าที่ผิดปกติและการกำหนดค่าผิดปกติ

จำนวนค่าผิดปกติ	กำหนดค่าผิดปกติ
1	$R_n = Q_3 + L(IQR)$
2	$R_1 = Q_1 - L(IQR)$ และ $R_n = Q_3 + L(IQR)$
3	$R_1 = Q_1 - L(IQR)$ $R_{n-1} = Q_3 + [0.95xL(IQR)]$ และ $R_n = Q_3 + L(IQR)$
5	$R_1 = Q_1 - L(IQR)$ $R_2 = Q_1 - [0.95xL(IQR)]$ $R_{n-2} = Q_3 + [0.9xL(IQR)]$ $R_{n-1} = Q_3 + [0.95xL(IQR)]$ และ $R_n = Q_3 + L(IQR)$

เมื่อ Q_1 แทน ค่าควอร์ไทล์ที่ 1 ของตัวแปรอิสระ/ ตัวแปรตาม

Q_3 แทน ค่าควอร์ไทล์ที่ 3 ของตัวแปรอิสระ/ ตัวแปรตาม

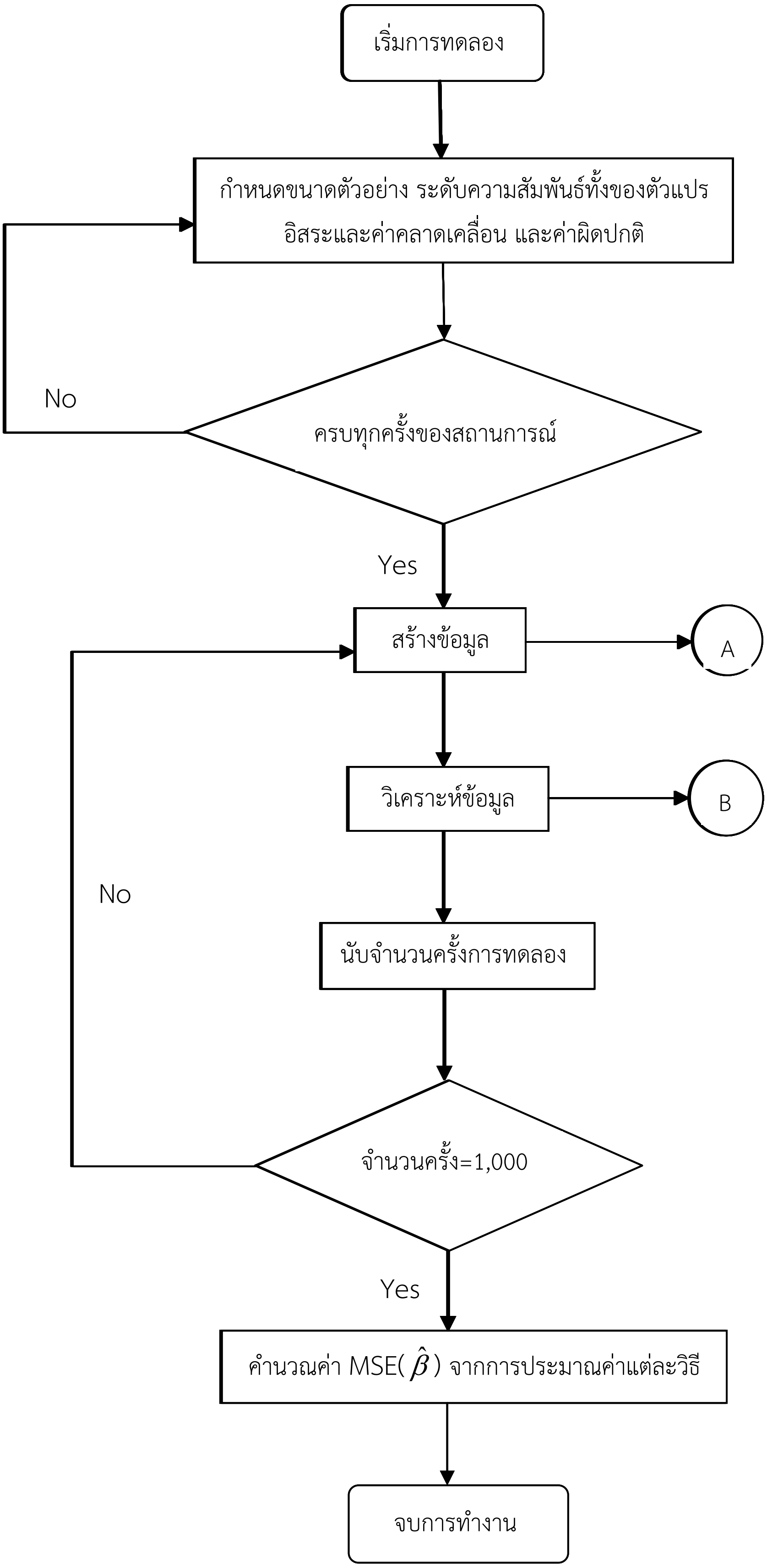
IQR แทน ผลต่างระหว่างค่าควอร์ไทล์ที่ 3 และค่าควอร์ไทล์ที่1 ของข้อมูล

กำหนดให้ L=4 สำหรับค่าผิดปกติในระดับรุนแรง เนื่องจากจะได้ค่าอยู่ระหว่างช่วง $[-\infty, Q_1 - 3(IQR)]$ หรือ $[Q_3 + 3(IQR), \infty]$ เป็นค่าผิดปกติในระดับรุนแรง ตามเงื่อนไขการตรวจสอบค่าผิดปกติด้วยวิธีแผนภาพกล่อง

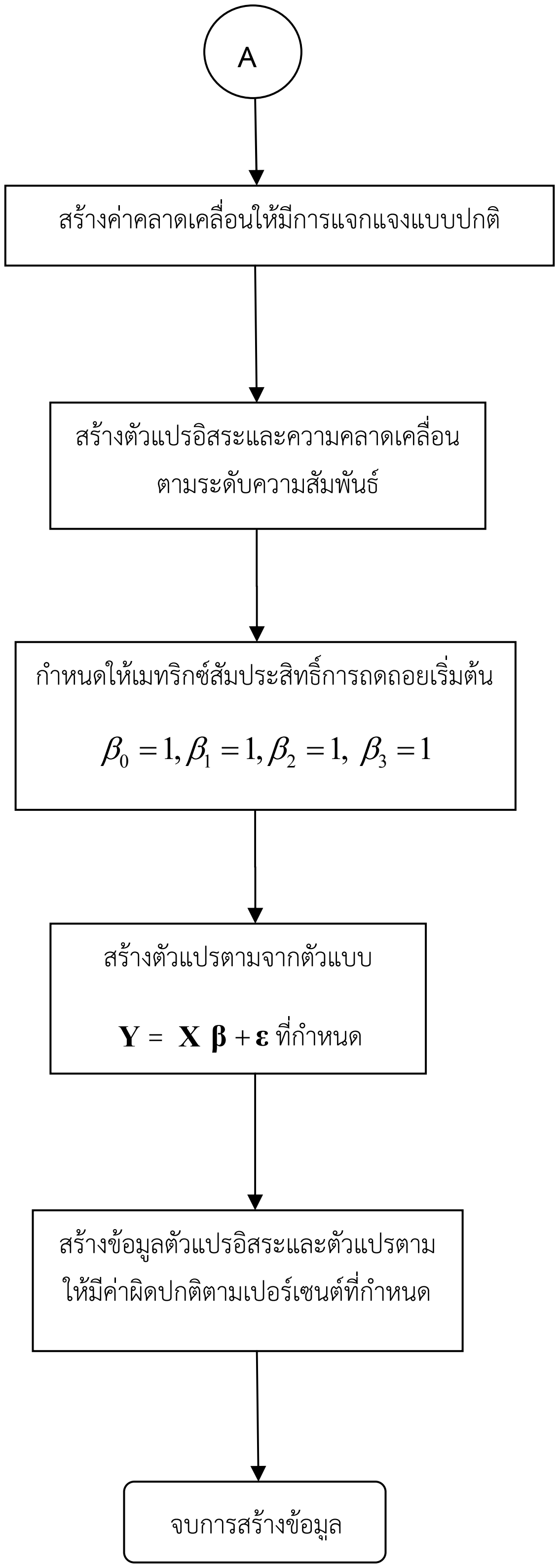
4. นำข้อมูลที่ได้ในข้อที่ 3. มาคำนวณหาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณของการถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น และวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยใช้เกณฑ์ความแกร่งของ Ramsay

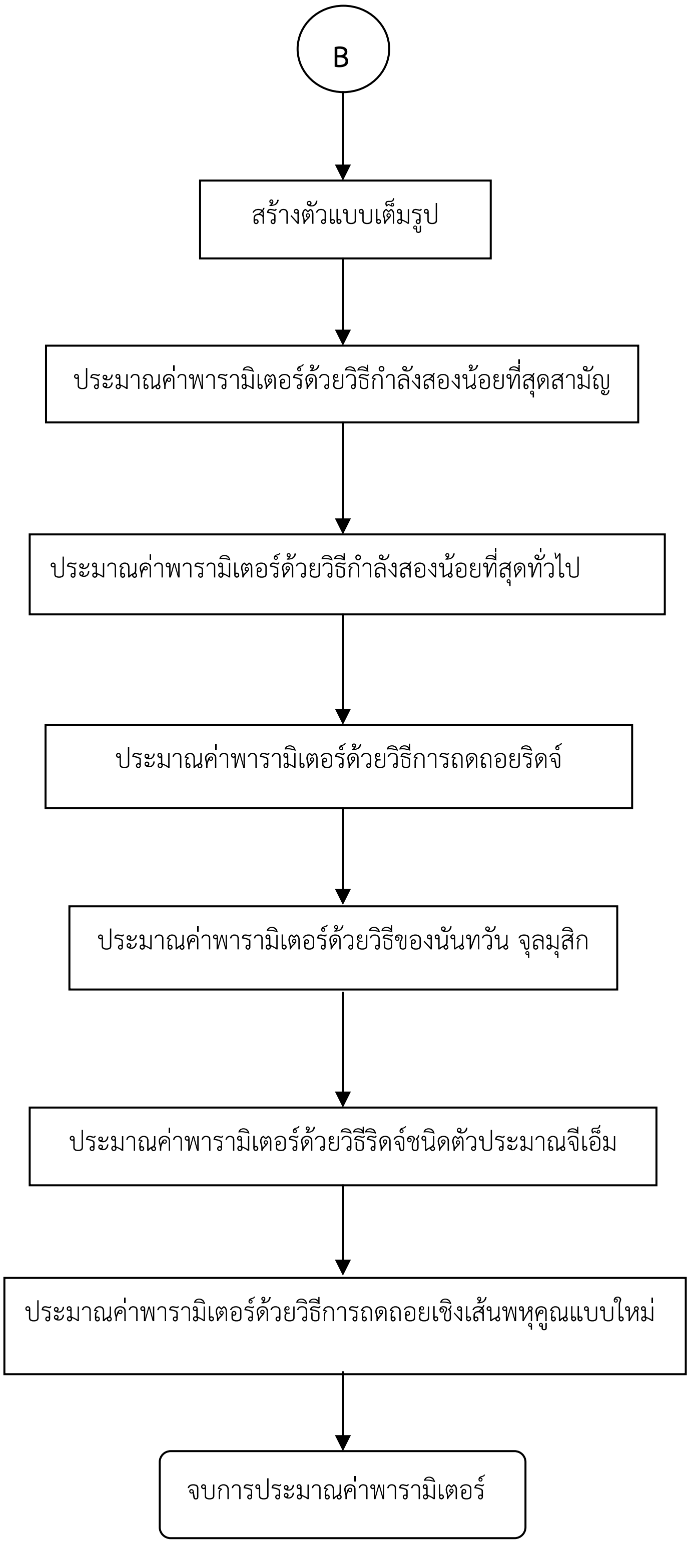
5. ทำซ้ำในแต่ละสถานการณ์จำนวน 1,000 ครั้ง โดยจะเก็บค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณที่คำนวณได้จากวิธีการประมาณแบบใหม่ที่พัฒนาขึ้น วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยใช้เกณฑ์ความแกร่งของ Ramsay เพื่อนำไปคำนวณหาค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณของแต่ละวิธี และแต่ละสถานการณ์

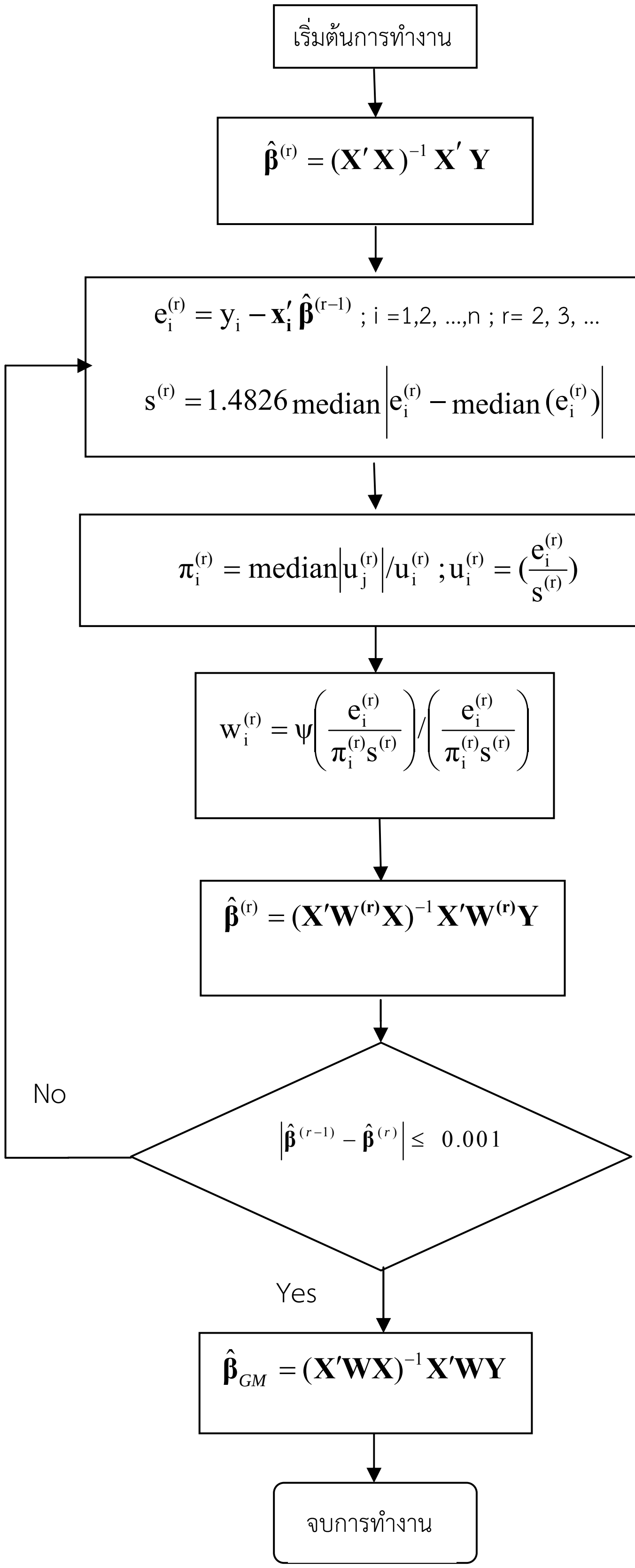
6. วิเคราะห์ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณของแต่ละวิธีและแต่ละสถานการณ์ เพื่อนำแต่ละวิธีและแต่ละสถานการณ์มาเปรียบเทียบกัน โดยวิธีที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่น้อยที่สุด จะเป็นวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่มีประสิทธิภาพมากที่สุด จึงสามารถสรุปวิธีการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ได้ดังภาพที่ 4



ภาพที่ 4 แผนผังของการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล


A
สร้างค่าคลาดเคลื่อนให้มีการแจกแจงแบบปกติ
สร้างตัวแปรอิสระและความคลาดเคลื่อน
ตามระดับความสัมพันธ์
กำหนดให้เมทริกซ์สัมประสิทธิ์การถดถอยเริ่มต้น
$\beta_0 = 1, \beta_1 = 1, \beta_2 = 1, \ \beta_3 = 1$
สร้างตัวแปรตามจากตัวแบบ
$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ ที่กำหนด
สร้างข้อมูลตัวแปรอิสระและตัวแปรตาม
ให้มีค่าผิดปกติตามเปอร์เซนต์ที่กำหนด
จบการสร้างข้อมูล


B
สร้างตัวแบบเต็มรูป
ประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดสามัญ
ประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดทั่วไป
ประมาณค่าพารามิเตอร์ด้วยวิธีการถดถอยริดจ์
ประมาณค่าพารามิเตอร์ด้วยวิธีของนันทวัน จุลมุสิก
ประมาณค่าพารามิเตอร์ด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม
ประมาณค่าพารามิเตอร์ด้วยวิธีการถดถอยเชิงเส้นพหุคูณแบบใหม่
จบการประมาณค่าพารามิเตอร์



ภาพที่ 5 แผนผังการทำงานของวิธีกำลังสองน้อยที่สุดถ่วงน้ำหนักชนิดทำซ้ำ (IRLS)

เกณฑ์การตัดสินใจ

ในการเปรียบเทียบประสิทธิภาพของวิธีประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณ ผู้วิจัยใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณทุกวิธี เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าที่พัฒนาขึ้นมาใหม่ และตัวประมาณค่าอีก 5 วิธี มีวิธีคำนวณ ดังนี้

$$\mathrm{MSE}(\hat{\boldsymbol{\beta}}) = \frac{\sum_{j=1}^{p}\left[\sum_{t=1}^{1000}\frac{(\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_{jt})^2}{1000}\right]}{p+1}$$

เมื่อ $\boldsymbol{\beta}_j$ แทน เวกเตอร์ของค่าพารามิเตอร์หรือค่าสัมประสิทธิ์การถดถอยที่แท้จริงตัวที่ $j;\ j = 0,1,2,3$

$\hat{\boldsymbol{\beta}}_{jt}$ แทน เวกเตอร์ของตัวประมาณค่าพารามิเตอร์หรือค่าประมาณสัมประสิทธิ์การถดถอยตัวที่ j ในการทำซ้ำรอบที่ $t;\ t = 1,2,3,\ldots,1000$

p แทน จำนวนตัวแปรอิสระ

หากวิธีการประมาณค่าสัมประสิทธิ์การถดถอยวิธีใดให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุดจะถือว่าวิธีนั้นมีประสิทธิภาพมากที่สุด

ตอนที่ 3 การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

ในขั้นตอนนี้นำวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่มาประยุกต์ใช้กับข้อมูลทุติยภูมิเพื่อพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยมีรายละเอียด ดังนี้

การเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลทุติยภูมิเกี่ยวกับยางพารา คือปริมาณการส่งออกยางพาราของไทยไปประเทศมาเลเซีย และราคาส่งออกยางพาราไทย ใช้ข้อมูลจากส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตรสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557 ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ใช้ข้อมูลจากสำนักงานคณะกรรมการเศรษฐกิจและสังคมภาคพื้นเอเชียและแปซิฟิค (Economic and Social Commission for Asia and Pacific) ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557

การวิเคราะห์ข้อมูล

วิธีการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ด้วยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่พัฒนาขึ้น มีวิธีดำเนินการ ดังนี้

1. รวบรวมข้อมูลทุติยภูมิปริมาณการนำเข้ายางพาราไทยของประเทศมาเลเซีย และราคาส่งออกยางพาราไทย จากส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557

2. รวบรวมข้อมูลทุติยภูมิเกี่ยวกับผลิตภัณฑ์มวลรวมภายในประเทศมาเลเซีย จากสำนักงานคณะกรรมการเศรษฐกิจและสังคมภาคพื้นเอเชียและแปซิฟิค ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557

3. นำข้อมูลในช่วง ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2556 มาทำการวิเคราะห์ข้อมูลพื้นฐาน และตรวจสอบข้อสมมติของตัวแบบ

4. ประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น โดยใช้ข้อมูลในช่วง ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2556

5. พยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ในปี พ.ศ. 2557 ด้วยค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ได้จากข้อที่ 4

บทที่ 4
ผลการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอย 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ภายใต้สถานการณ์ 162 สถานการณ์ โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล และพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น ผู้วิจัยได้เสนอผลการวิจัยออกเป็น 3 ตอน ดังนี้

ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม

ตอนที่ 3 ผลการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

จากการศึกษาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เมื่อใช้วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยประมาณค่า k ด้วยวิธีของ Hoerl-Kennard-Baldwin (HKB) และใช้เกณฑ์ความแกร่งของ Ramsay ทำให้สามารถพัฒนาเป็นตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ได้ ดังนี้

$$\hat{\mathbf{b}}_{NEW} = \left(\mathbf{Q}'\mathbf{W}\mathbf{Q} + k\mathbf{I}\right)^{-1}\left(\mathbf{Q}'\mathbf{W}\mathbf{R} + k\mathbf{J}\right)$$

เมื่อ $\mathbf{Q}$ เป็นเมทริกซ์ของตัวแปรอิสระที่อยู่ในรูปมาตรฐานและใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

R เป็นเวกเตอร์ของตัวแปรตามที่อยู่ในรูปมาตรฐานและใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

W เป็นเมทริกซ์ทแยงมุมที่ได้จากการทำวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักซ้ำหลายรอบ (IRLS) โดยใช้เกณฑ์ความแกร่งของ Ramsay ขนาด $n \times n$

และเมทริกซ์ความแปรปรวนความแปรปรวนร่วมของ $\hat{\mathbf{b}}_{NEW}$ แสดงได้ดังนี้

$$
\begin{aligned}
\mathrm{Cov}\left(\hat{\mathbf{b}}_{NEW}\right) &= \mathrm{Cov}\left(\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\left(\mathbf{Q}'\mathbf{W}\mathbf{R}+k\mathbf{J}\right)\right) \\
&= \mathrm{Cov}\left(\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\mathbf{R}+\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}k\mathbf{J}\right) \\
&= \left[\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\right]\mathrm{Var}(\mathbf{R})\left[\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}Q'\mathbf{W}\right]' \\
&= \sigma^2\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\mathbf{W}'\mathbf{Q}\left(\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\right)' \\
&= \sigma^2\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\mathbf{W}'\mathbf{Q}\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}
\end{aligned}
$$

จะได้ว่าตัวประมาณเมทริกความแปรปรวนและความแปรปรวนร่วมของ $\hat{\mathbf{b}}_{NEW}$ คือ

$$
\widehat{\mathrm{Cov}}\left(\hat{\mathbf{b}}_{NEW}\right) = \hat{\sigma}^2\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\mathbf{W}'\mathbf{Q}\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}
$$

โดยที่ $\hat{\sigma}^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม

ในการวิเคราะห์เปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ขั้นตอนนี้เพื่อทดสอบวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลและใช้โปรแกรม MATLAB เวอร์ชัน 8.7

ซึ่งการนำเสนอผลการวิจัยจะนำเสนอในรูปของตารางและกราฟ โดยใช้สัญลักษณ์ต่อไปนี้แทนความหมายต่าง ๆ ดังนี้

OLS หมายถึง ตัวประมาณที่คำนวณด้วยวิธีกำลังสองน้อยที่สุดสามัญ

GLS หมายถึง ตัวประมาณที่คำนวณด้วยวิธีกำลังสองน้อยที่สุดทั่วไป

RR หมายถึง ตัวประมาณที่คำนวณด้วยวิธีการถดถอยริดจ์ใช้ค่า k ของ Hoerl

		Kennard and Baldwin
NT	หมายถึง	ตัวประมาณที่คำนวณด้วยวิธีของนันทวัน จุลมุสิก
GMR	หมายถึง	ตัวประมาณที่คำนวณด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็มซึ่งใช้เกณฑ์ความแกร่งของ Ramsay
NEW	หมายถึง	ตัวประมาณที่คำนวณด้วยวิธีที่พัฒนาขึ้นซึ่งใช้ค่า k ของ Hoerl Kennard and Baldwin และเกณฑ์ความแกร่งของ Ramsay
MSE($\hat{\beta}$)	หมายถึง	ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ
n	หมายถึง	ขนาดตัวอย่าง
r	หมายถึง	ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_2
ρ	หมายถึง	ค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

ตัวประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณและเมทริกซ์ความแปรปรวนความแปรปรวนร่วมของแต่ละวิธีที่ใช้ในการเปรียบเทียบมีสูตร ดังนี้

วิธีกำลังสองน้อยที่สุดสามัญ

$$\hat{\mathbf{b}}_{OLS} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y}$$

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{OLS}\right) = \hat{\sigma}_1^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}$$

โดยที่ $\hat{\sigma}_1^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

วิธีกำลังสองน้อยที่สุดทั่วไป

$$\hat{\mathbf{b}}_{GLS} = \left(\mathbf{X}'\mathbf{U}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{X}'\mathbf{U}^{-1}\mathbf{Y}\right)$$

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{GLS}\right) = \hat{\sigma}_2^2\left(\mathbf{X}'\mathbf{U}^{-1}\mathbf{X}\right)^{-1}$$

โดยที่ $\hat{\sigma}_2^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

วิธีการถดถอยริดจ์

$$\hat{\mathbf{b}}_{RR} = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\mathbf{Y}\right)$$

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{RR}\right) = \hat{\sigma}_3^2\left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}\mathbf{X}'\mathbf{X}\left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}$$

โดยที่ $\hat{\sigma}_3^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

วิธีของนันทวัน จุลมุสิก

$$\hat{\mathbf{b}}_{NT} = \left(\mathbf{H}'\mathbf{U}^{-1}\mathbf{H} + k\mathbf{I}\right)^{-1}\left(\mathbf{H}'\mathbf{U}^{-1}\mathbf{N} + k\mathbf{J}\right)$$

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{NT}\right) = \hat{\sigma}_4^2\left(\mathbf{H}'\mathbf{U}^{-1}\mathbf{H} + k\mathbf{I}\right)^{-1}\mathbf{H}'\mathbf{U}^{-1}\mathbf{H}\left(\mathbf{H}'\mathbf{U}^{-1}\mathbf{H} + k\mathbf{I}\right)^{-1}$$

โดยที่ $\hat{\sigma}_4^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

วิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยใช้เกณฑ์ความแกร่งของ Ramsay

$$\hat{\mathbf{b}}_{GMR} = \left(\mathbf{X}'\mathbf{W}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\mathbf{W}\mathbf{Y}\right)$$

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{GMR}\right) = \hat{\sigma}_5^2\left(\mathbf{X}'\mathbf{W}\mathbf{X} + k\mathbf{I}\right)^{-1}\mathbf{X}'\mathbf{W}\mathbf{W}'\mathbf{X}\left(\mathbf{X}'\mathbf{W}\mathbf{X} + k\mathbf{I}\right)^{-1}$$

โดยที่ $\hat{\sigma}_5^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

วิธีที่พัฒนาขึ้น

$$\hat{\mathbf{b}}_{NEW} = \left(\mathbf{Q}'\mathbf{W}\mathbf{Q} + k\mathbf{I}\right)^{-1}\left(\mathbf{Q}'\mathbf{W}\mathbf{R} + k\mathbf{J}\right)$$

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{NEW}\right) = \hat{\sigma}_6^2\left(\mathbf{Q}'\mathbf{W}\mathbf{Q} + k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\mathbf{W}'\mathbf{Q}\left(\mathbf{Q}'\mathbf{W}\mathbf{Q} + k\mathbf{I}\right)^{-1}$$

โดยที่ $\hat{\sigma}_6^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

สัญลักษณ์ต่าง ๆ ใช้แทนความหมายดังต่อไปนี้

$\mathbf{Y}$ แทน เวกเตอร์ของตัวแปรตามที่ทราบค่า

$\mathbf{X}$ แทน เมทริกซ์ของตัวแปรอิสระที่ทราบค่า

$\mathbf{N}$ แทน เวกเตอร์ของตัวแปรตามที่อยู่ในรูปมาตรฐาน มีมิติ n x 1

ด้วยสูตร $n_i = \dfrac{y_i - \bar{y}}{s_y}$

$\mathbf{H}$ แทน เมทริกซ์ของตัวแปรอิสระที่อยู่ในรูปมาตรฐาน มีมิติ n x p

ด้วยสูตร $h_{i1} = \dfrac{x_{i1} - \bar{x}_1}{s_1}$

$$h_{i2} = \frac{x_{i2} - \bar{x}_2}{s_2}$$

$$h_{i3} = \frac{x_{i3} - \bar{x}_3}{s_3}$$

$\mathbf{U}^{-1}$ แทน อินเวอร์สของเมทริกซ์อัตตสหสัมพันธ์

$\mathbf{I}$ แทน เมทริกซ์เอกลักษณ์

k แทน ค่าประมาณที่คำนวณจากสูตร HKB $k = \dfrac{p\hat{\sigma}^2}{\hat{\mathbf{b}}'\hat{\mathbf{b}}}$

โดยที่ p คือ จำนวนตัวแปรอิสระในตัวแบบ

$\hat{\mathbf{b}}$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ $\boldsymbol{\beta}$

$\hat{\sigma}^2$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ σ^2

$\mathbf{J}$ แทน เวกเตอร์ของค่าประมาณเบื้องต้นของ $\mathbf{b}$ จากวิธีกำลังสองน้อยที่สุดสามัญ

ค่า $\mathbf{J}$ จะประมาณจาก $\bar{\mathbf{J}}$ ด้วยสูตร $\bar{\mathbf{J}} = \sum_{j=1}^{p} \dfrac{\hat{\mathbf{b}}_j}{p}$

โดยที่ p คือ จำนวนตัวแปรอิสระในตัวแบบ

$\hat{\mathbf{b}}_j$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ $\boldsymbol{\beta}_j$

$\mathbf{W}$ แทน เมทริกซ์ทแยงมุมที่ได้จากการทำวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักซ้ำหลายรอบ (IRLS) โดยใช้เกณฑ์ความแกร่งของ Ramsay

$\mathbf{R}$ แทน เวกเตอร์ของตัวแปรตามที่อยู่ในรูปมาตรฐานหลังจากใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

$\mathbf{Q}$ แทน เมทริกซ์ของตัวแปรอิสระที่อยู่ในรูปมาตรฐานหลังจากใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

โดยเทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทนนั้น เป็นการประมาณค่า ϕ โดยใช้สหสัมพันธ์ r ระหว่างค่าความคลาดเคลื่อนโดยวิธีกำลังสองน้อยที่สุด ทำการคำนวณโดยใช้สูตรดังนี้

$$r = \frac{\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sqrt{\sum_{t=2}^{n} \hat{\varepsilon}_t^2 \sum_{t=2}^{n} \hat{\varepsilon}_{t-1}^2}} \cong \frac{\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2} = r_1 = \hat{\phi}$$

$$\hat{\mathbf{P}} = \begin{bmatrix} \sqrt{1-r_1^2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -r_1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -r_1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -r_1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -r_1 & 1 \end{bmatrix}$$

สรุปได้เป็นเมทริกซ์

$$\mathbf{Y}^* = \hat{\mathbf{P}}\mathbf{Y} = \begin{bmatrix} y_1\sqrt{1-r_1^2} \\ y_2 - r_1 y_1 \\ y_3 - r_1 y_2 \\ \vdots \\ y_n - r_1 y_{n-1} \end{bmatrix}_{n \text{ x } 1}$$

$$\mathbf{X}^* = \hat{\mathbf{P}}\mathbf{X} = \begin{bmatrix} \sqrt{1-r_1^2} & x_{11}\sqrt{1-r_1^2} & x_{12}\sqrt{1-r_1^2} & \cdots & x_{1k}\sqrt{1-r_1^2} \\ 1-r_1 & x_{21} - r_1 x_{11} & x_{22} - r_1 x_{12} & \cdots & x_{2k} - r_1 x_{1k} \\ 1-r_1 & x_{31} - r_1 x_{21} & x_{32} - r_1 x_{22} & \cdots & x_{3k} - r_1 x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1-r_1 & x_{n1} - r_1 x_{(n-1)1} & x_{n2} - r_1 x_{(n-1)2} & \cdots & x_{nk} - r_1 x_{(n-1)k} \end{bmatrix}_{n \text{ x } (k+1)}$$

จะได้ว่า

$\mathbf{Y}^*$ แทน เวกเตอร์ของตัวแปรตามที่ใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

$\mathbf{X}^*$ แทน เมทริกซ์ของตัวแปรอิสระที่ใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

ดังนั้นจึงสรุปได้ว่า จากตัวแบบที่นำมาใช้แสดงความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ คือ

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + \varepsilon_i \quad ; i = 1, 2, 3, ..., n$$

เมื่อ Y_i แทน ค่าสังเกตที่ i ของตัวแปรตามที่ทราบค่า

$X_{i1}, X_{i2}, ..., X_{ip}$ แทน ค่าสังเกตที่ i ของตัวแปรอิสระ p ตัวที่ทราบค่า

$\beta_0, \beta_1, \beta_2, ..., \beta_p$ แทน สัมประสิทธิ์การถดถอยที่ไม่ทราบค่า

ε_i แทน ค่าคลาดเคลื่อนสุ่มของค่าสังเกตที่ i

n แทน จำนวนค่าสังเกต

และ p แทน จำนวนตัวแปรอิสระ

จะเปลี่ยนรูปเป็น

$$z_{yi} = \alpha_1 z_{i1} + \alpha_2 z_{i2} + ... + \alpha_p z_{ip} + z_{\varepsilon i} \qquad ; \text{i=1,2,3,...,n}$$

เมื่อ z_{yi} แทน ค่าสังเกตที่ i ของตัวแปรตามที่ทราบค่าที่อยู่ในรูปมาตรฐาน

$z_{i1}, z_{i2}, ..., z_{ip}$ แทน ค่าสังเกตที่ i ของตัวแปรอิสระ p ตัวที่ทราบค่าที่อยู่ในรูปมาตรฐาน

$\alpha_1, \alpha_2, ..., \alpha_p$ แทน สัมประสิทธิ์การถดถอยที่ไม่ทราบค่า

$z_{\varepsilon i}$ แทน ค่าคลาดเคลื่อนสุ่มมาตรฐานของค่าสังเกตที่ i

n แทน จำนวนค่าสังเกต

และ p แทน จำนวนตัวแปรอิสระ

จากตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น ผู้วิจัยทำการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี คือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก วิธีริดจ์ชนิดตัวประมาณจีเอ็ม และวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ ภายใต้สถานการณ์ 162 สถานการณ์ ได้แก่ 1) ขนาดตัวอย่าง 3 ขนาด คือ 20, 50 และ 100 2) ระดับการมีพหุสัมพันธ์ 3 ระดับ คือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 3) ระดับการมีอัตตสหสัมพันธ์ 3 ระดับ คือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 4) ลักษณะของการเกิดค่าผิดปกติของข้อมูล 3 แบบ คือ ข้อมูลผิดปกติในตัวแปรอิสระ ข้อมูลผิดปกติในตัวแปรตาม และข้อมูลผิดปกติในตัวแปรอิสระและตัวแปรตาม 5) เปอร์เซนต์ค่าผิดปกติ 2 ระดับ คือ 1% และ 5% ในแต่ละวิธีจะคำนวณค่าคลาดเคลื่อนกำลังสองเฉลี่ยของเวกเตอร์ตัวประมาณค่าสัมประสิทธิ์การถดถอย และจะนำแต่ละวิธีและแต่ละสถานการณ์มาเปรียบเทียบกัน โดยวิธีที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณน้อยที่สุด จะเป็นวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่มีประสิทธิภาพมากที่สุด ผู้วิจัยได้แบ่งการนำเสนอผลการวิจัยออกเป็น 3 กรณี ดังนี้

กรณีที่ 1 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3

กรณีที่ 2 ข้อมูลมีค่าผิดปกติในตัวแปรตาม Y

กรณีที่ 3 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y

สรุปผลการวิจัยได้ ดังนี้

กรณีที่ 1 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3

กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ได้ผลการวิจัยดังนี้

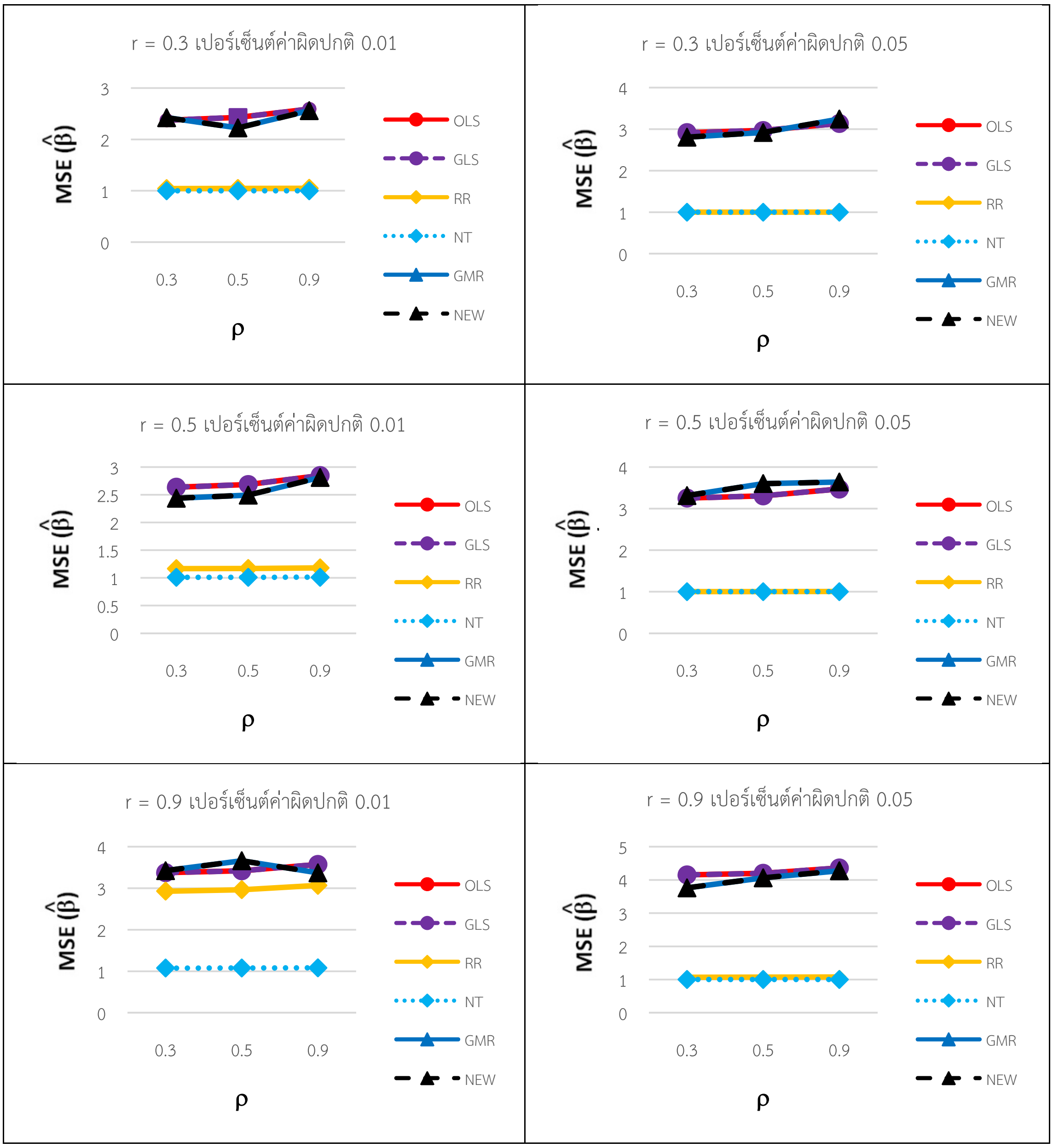
ตารางที่ 9 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	2.3739	2.3739	1.0418	**0.9998***	2.4244	2.4236
		5%	2.9163	2.9163	1.0010	**1.0001***	2.8068	2.8068
	0.5	1%	2.6372	2.6372	1.1654	**1.0099***	2.4371	2.4371
		5%	3.2547	3.2547	1.0029	**1.0001***	3.3141	3.3141
	0.9	1%	3.3738	3.3738	2.9296	**1.0776***	3.4229	3.4229
		5%	4.1530	4.1530	1.0718	**1.0025***	3.7622	3.7622
0.5	0.3	1%	2.4278	2.4278	1.0435	**1.0000***	2.2256	2.2256
		5%	2.9693	2.9693	1.0011	**1.0001***	2.9200	2.9200
	0.5	1%	2.6864	2.6864	1.1688	**1.0101***	2.4941	2.4941
		5%	3.3063	3.3063	1.0030	**1.0002***	3.6044	3.6044
	0.9	1%	3.4191	3.4191	2.9654	**1.0789***	3.6630	3.6630
		5%	4.2007	4.2007	1.0741	**1.0026***	4.0654	4.0654
0.9	0.3	1%	2.5918	2.5918	1.0484	**1.0002***	2.5598	2.5598
		5%	3.1383	3.1383	1.0012	**1.0001***	3.2451	3.2451
	0.5	1%	2.8478	2.8478	1.1793	**1.0106***	2.8128	2.8128
		5%	3.4729	3.4729	1.0034	**1.0002***	3.6427	3.6427
	0.9	1%	3.5727	3.5727	3.0695	**1.0823***	3.3710	3.3710
		5%	4.3594	4.3594	1.0798	**1.0028***	4.2780	4.2780

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 9 และภาพที่ 6 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 ค่าคลาดเคลื่อนกำลังสอง

เฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีของนันทวัน จุลมุสิกให้ค่าต่ำที่สุดในทุกสถานการณ์



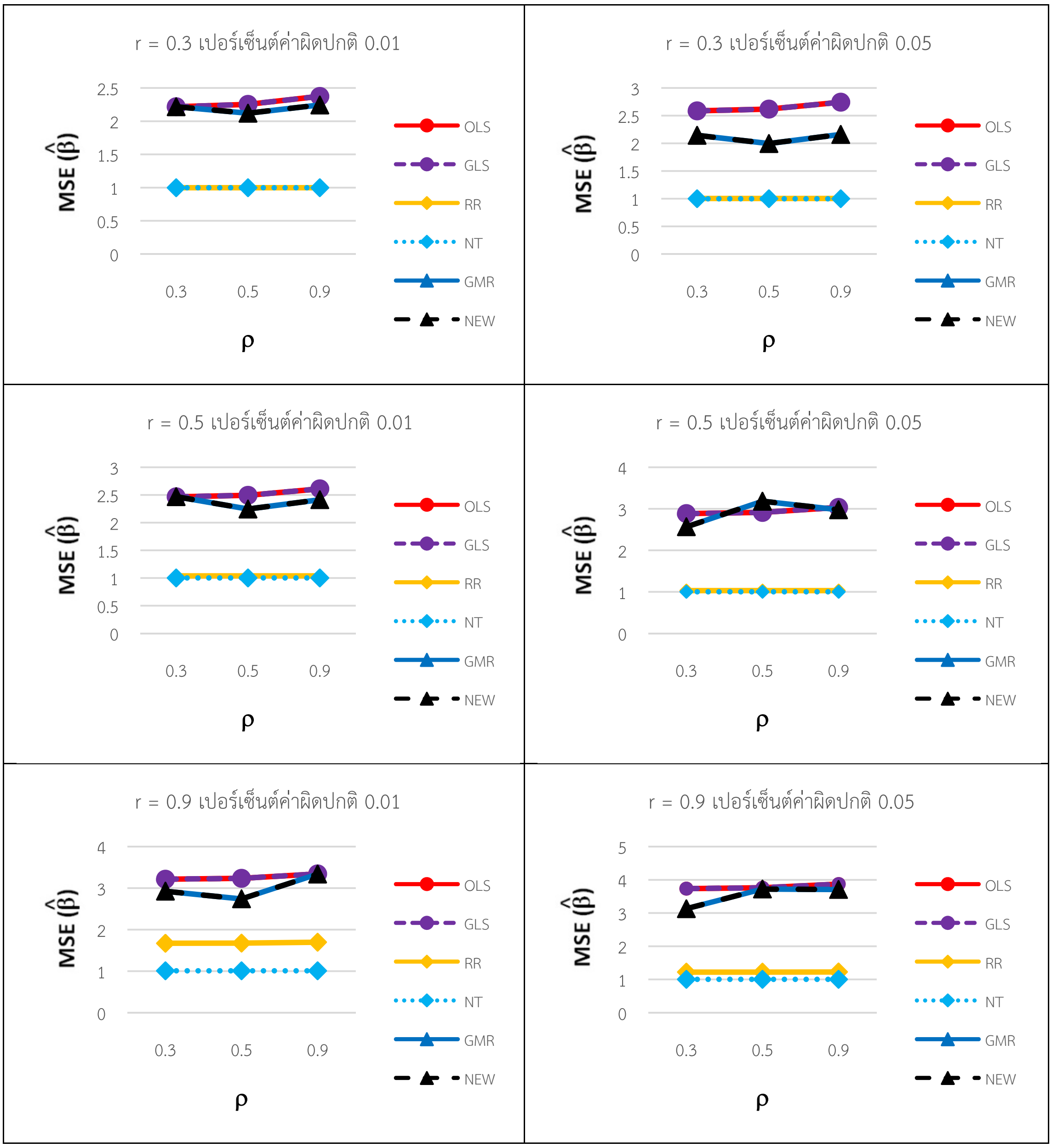
ภาพที่ 6 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20

ตารางที่ 10 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	2.2190	2.2190	**0.9982***	0.9991	2.2180	2.2179
		5%	2.5871	2.5871	1.0036	**0.9999***	2.1454	2.1454
	0.5	1%	2.4685	2.4685	1.0375	**1.0007***	2.4702	2.4702
		5%	2.8844	2.8844	1.0308	**1.0012***	2.5675	2.5675
	0.9	1%	3.2143	3.2143	1.6705	**1.0076***	2.9264	2.9264
		5%	3.7364	3.7364	1.2225	**1.0048***	3.1302	3.1302
0.5	0.3	1%	2.2533	2.2533	**0.9986***	0.9991	2.1205	2.1205
		5%	2.6187	2.6187	1.0036	**0.9999***	1.9971	1.9971
	0.5	1%	2.4951	2.4951	1.0378	**1.0007***	2.2462	2.2462
		5%	2.9134	2.9134	1.0307	**1.0012***	3.1838	3.1838
	0.9	1%	3.2360	3.2360	1.6751	**1.0076***	2.7416	2.7416
		5%	3.7606	3.7606	1.2227	**1.0048***	3.7233	3.7233
0.9	0.3	1%	2.3746	2.3746	0.9994	**0.9992***	2.2439	2.2439
		5%	2.7449	2.7449	1.0038	**0.9999***	2.1605	2.1605
	0.5	1%	2.6114	2.6114	1.0394	**1.0007***	2.4154	2.4154
		5%	3.0346	3.0346	1.0310	**1.0012***	2.9777	2.9777
	0.9	1%	3.3427	3.3427	1.6974	**1.0077***	3.3384	3.3384
		5%	3.8723	3.8723	1.2264	**1.0047***	3.7107	3.7107

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 10 และภาพที่ 7 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีของนันทวัน จุลมุสิกส่วนใหญ่จะให้ค่าต่ำที่สุดในทุกสถานการณ์ ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และมีอัตตสหสัมพันธ์เท่ากับ 0.3 และ 0.5 วิธีการถดถอยริดจ์จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



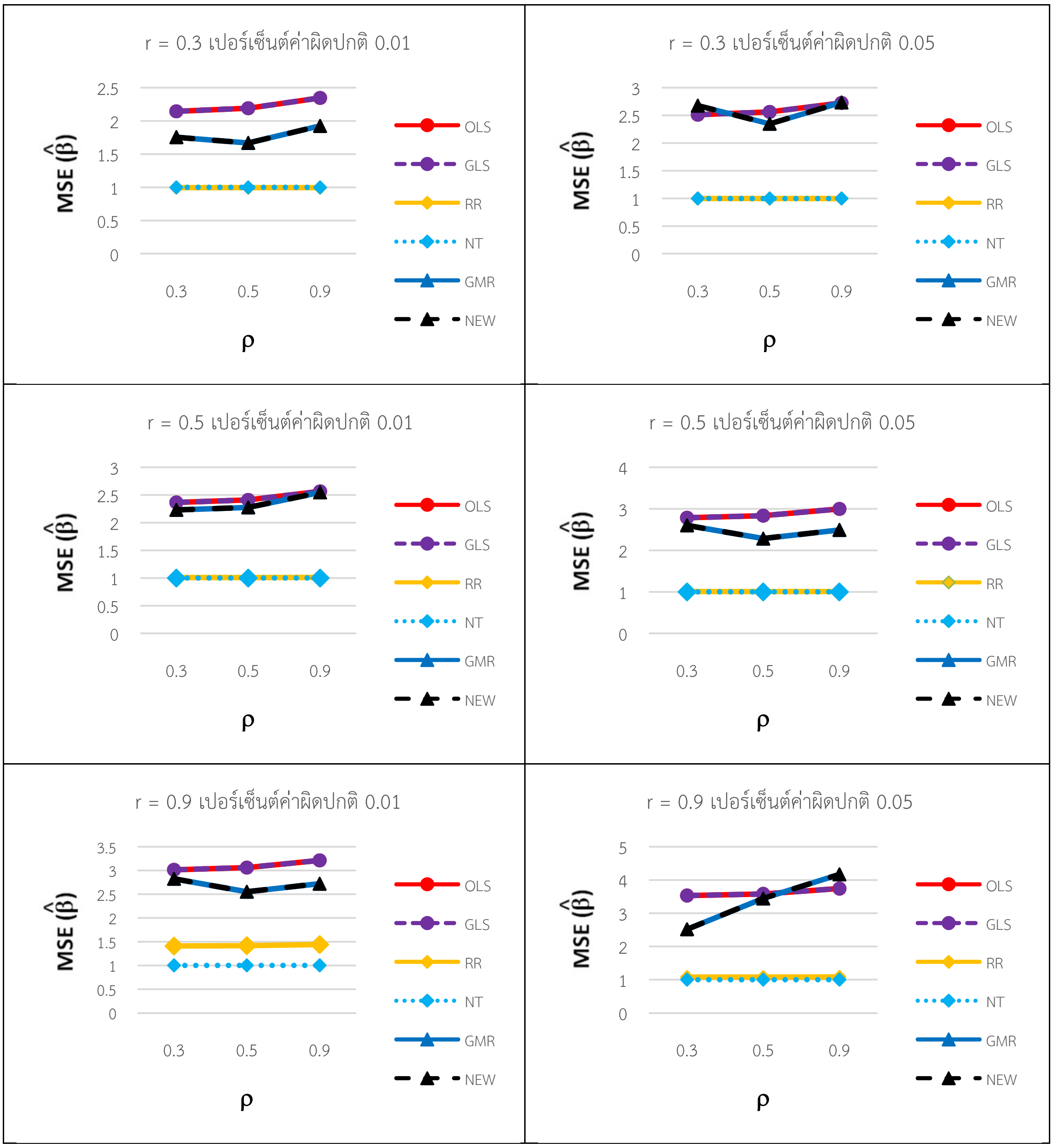
ภาพที่ 7 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50

ตารางที่ 11 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	2.1451	2.1451	**0.9954***	0.9997	1.7560	1.7560
		5%	2.5138	2.5138	**0.9993***	0.9999	2.6765	2.6765
	0.5	1%	2.3660	2.3660	1.0104	**1.0000***	2.2313	2.2313
		5%	2.7855	2.7855	1.0089	**1.0002***	2.6032	2.6032
	0.9	1%	3.0142	3.0142	1.4089	**1.0018***	2.8256	2.8256
		5%	3.5320	3.5320	1.0768	**1.0010***	2.5145	2.5145
0.5	0.3	1%	2.1911	2.1911	**0.9954***	0.9997	1.6692	1.6692
		5%	2.5635	2.5635	**0.9994***	0.9999	2.3489	2.3489
	0.5	1%	2.4118	2.4118	1.0106	**1.0000***	2.2760	2.2760
		5%	2.8353	2.8353	1.0090	**1.0002***	2.2846	2.2846
	0.9	1%	3.0593	3.0593	1.4161	**1.0018***	2.5533	2.5533
		5%	3.5809	3.5809	1.0778	**1.0010***	3.4436	3.4436
0.9	0.3	1%	2.3458	2.3458	**0.9958***	0.9997	1.9243	1.9243
		5%	2.7261	2.7261	**0.9994***	0.9994	2.7330	2.7330
	0.5	1%	2.5666	2.5666	1.0113	**1.0000***	2.5472	2.5472
		5%	2.9979	2.9979	1.0092	**1.0002***	2.4928	2.4928
	0.9	1%	3.2126	3.2126	1.4400	**1.0018***	2.7222	2.7222
		5%	3.7420	3.7420	1.0814	**1.0010***	4.1685	4.1685

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 11 และภาพที่ 8 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีของนันทวัน จุลมุสิกส่วนใหญ่จะให้ค่าต่ำที่สุดในทุกสถานการณ์ ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และมีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีการถดถอยริดจ์จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



ภาพที่ 8 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100

กรณีที่ 2 ข้อมูลมีค่าผิดปกติในตัวแปรตาม Y

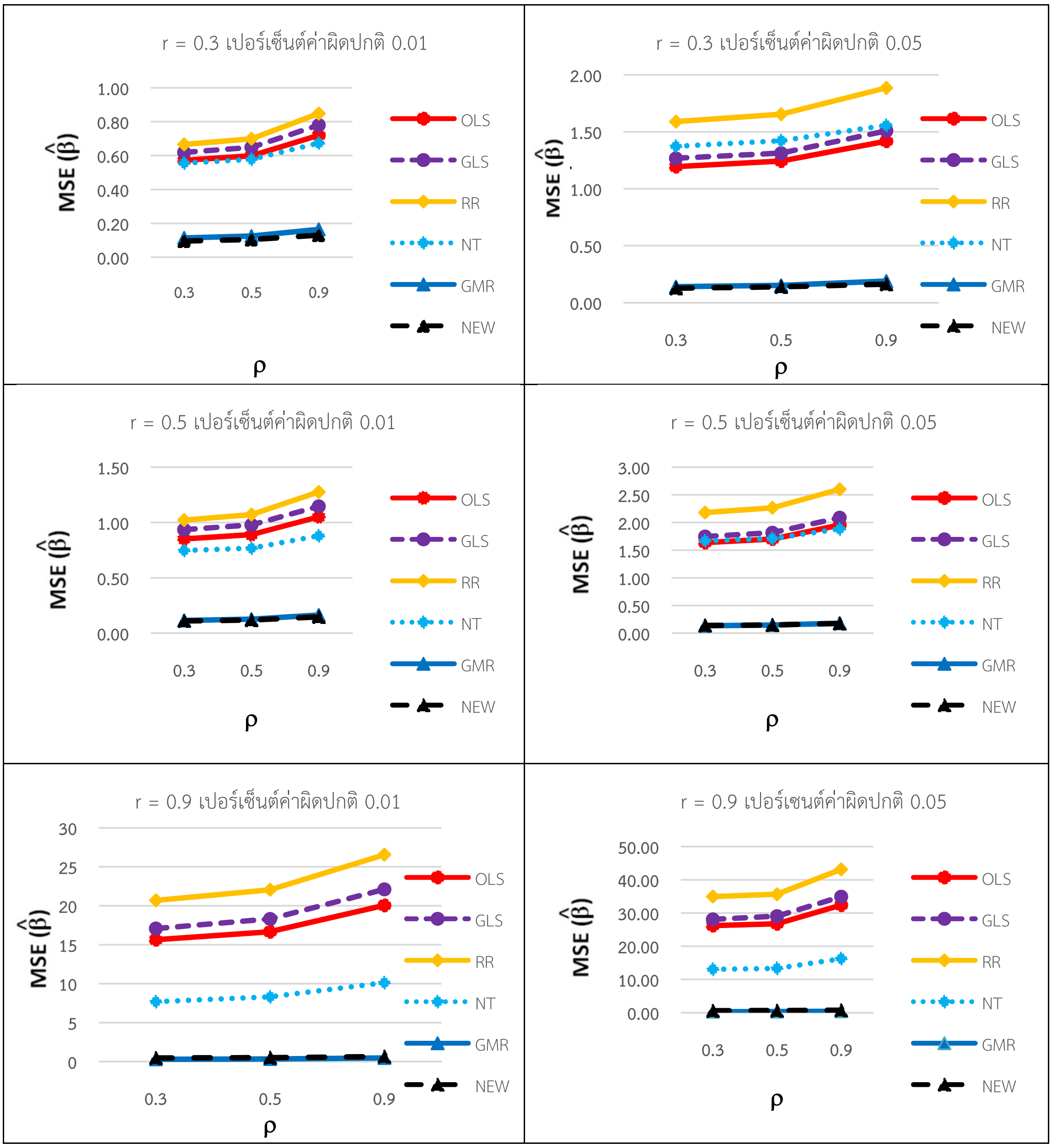
กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ได้ผลการวิจัย ดังนี้

ตารางที่ 12 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	0.5722	0.6196	0.6662	0.5559	0.1145	**0.0961***
		5%	1.1941	1.2669	1.5889	1.3722	0.1402	**0.1251***
	0.5	1%	0.8520	0.9353	1.0227	0.7474	0.1146	**0.1111***
		5%	1.6375	1.7469	2.1798	1.6702	**0.1350***	0.1397
	0.9	1%	15.6409	17.0743	20.6924	7.6830	**0.3021***	0.4604
		5%	26.2293	28.0961	34.9674	13.0552	**0.2909***	0.5699
0.5	0.3	1%	0.5992	0.6487	0.6993	0.5786	0.1257	**0.1051***
		5%	1.2431	1.3121	1.6539	1.4218	0.1517	**0.1377***
	0.5	1%	0.8908	0.9778	1.0717	0.7682	0.1260	**0.1184***
		5%	1.7021	1.8149	2.2657	1.7105	**0.1471***	0.1489
	0.9	1%	16.6615	18.3014	22.0499	8.3012	**0.3374***	0.5047
		5%	26.7444	29.0618	35.6539	13.3409	**0.3200***	0.6072
0.9	0.3	1%	0.7197	0.7798	0.8483	0.6743	0.1648	**0.1310***
		5%	1.4170	1.5081	1.8850	1.5568	0.1890	**0.1615***
	0.5	1%	1.0518	1.1459	1.2759	0.8800	0.1634	**0.1466***
		5%	1.9544	2.0866	2.6013	1.8842	0.1790	**0.1754***
	0.9	1%	20.0471	22.0935	26.5511	10.1223	**0.4744***	0.6446
		5%	32.3693	34.9197	43.1530	16.2601	**0.4288***	0.7313

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 12 และภาพที่ 9 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3 0.5 และ 0.9 กรณีค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3 และ 0.5 กรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.9 นอกนั้นวิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



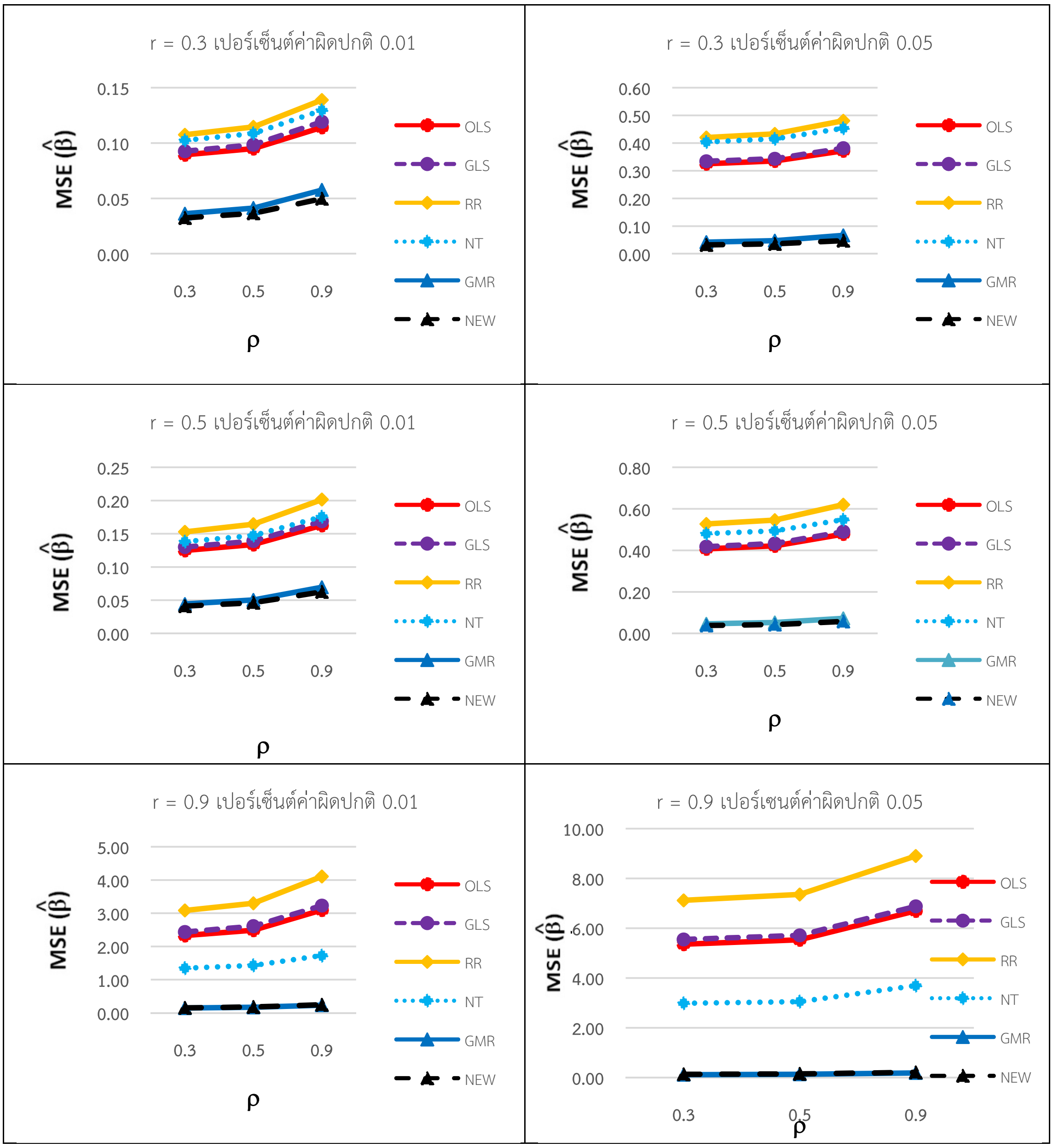
ภาพที่ 9 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20

ตารางที่ 13 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	0.0894	0.0927	0.1075	0.1025	0.0362	**0.0324***
		5%	0.3255	0.3345	0.4201	0.4041	0.0421	**0.0322***
	0.5	1%	0.1249	0.1301	0.1529	0.1383	0.0443	**0.0411***
		5%	0.4074	0.4182	0.5271	0.4807	0.0463	**0.0388***
	0.9	1%	2.3289	2.4372	3.0853	1.3470	**0.1461***	0.1532
		5%	5.3549	5.5460	7.1166	2.9831	**0.1136***	0.1325
0.5	0.3	1%	0.0950	0.0985	0.1146	0.1088	0.0413	**0.0367***
		5%	0.3357	0.3433	0.4332	0.4152	0.0475	**0.0359***
	0.5	1%	0.1337	0.1392	0.1644	0.1474	0.0504	**0.0462***
		5%	0.4216	0.4329	0.5457	0.4940	0.0527	**0.0429***
	0.9	1%	2.4902	2.6118	3.3000	1.4317	**0.1688***	0.1788
		5%	5.5325	5.7106	7.3528	3.0510	**0.1316***	0.1507
0.9	0.3	1%	0.1142	0.1192	0.1389	0.1296	0.0577	**0.0498***
		5%	0.3723	0.3817	0.4806	0.4537	0.0665	**0.0476***
	0.5	1%	0.1626	0.1689	0.2015	0.1757	0.0695	**0.0626***
		5%	0.4787	0.4901	0.6202	0.5481	0.0727	**0.0578***
	0.9	1%	3.0966	3.2286	4.1070	1.7248	**0.2330***	0.2456
		5%	6.6931	6.8770	8.8986	3.6942	**0.1821***	0.2077

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 13 และภาพที่ 10 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



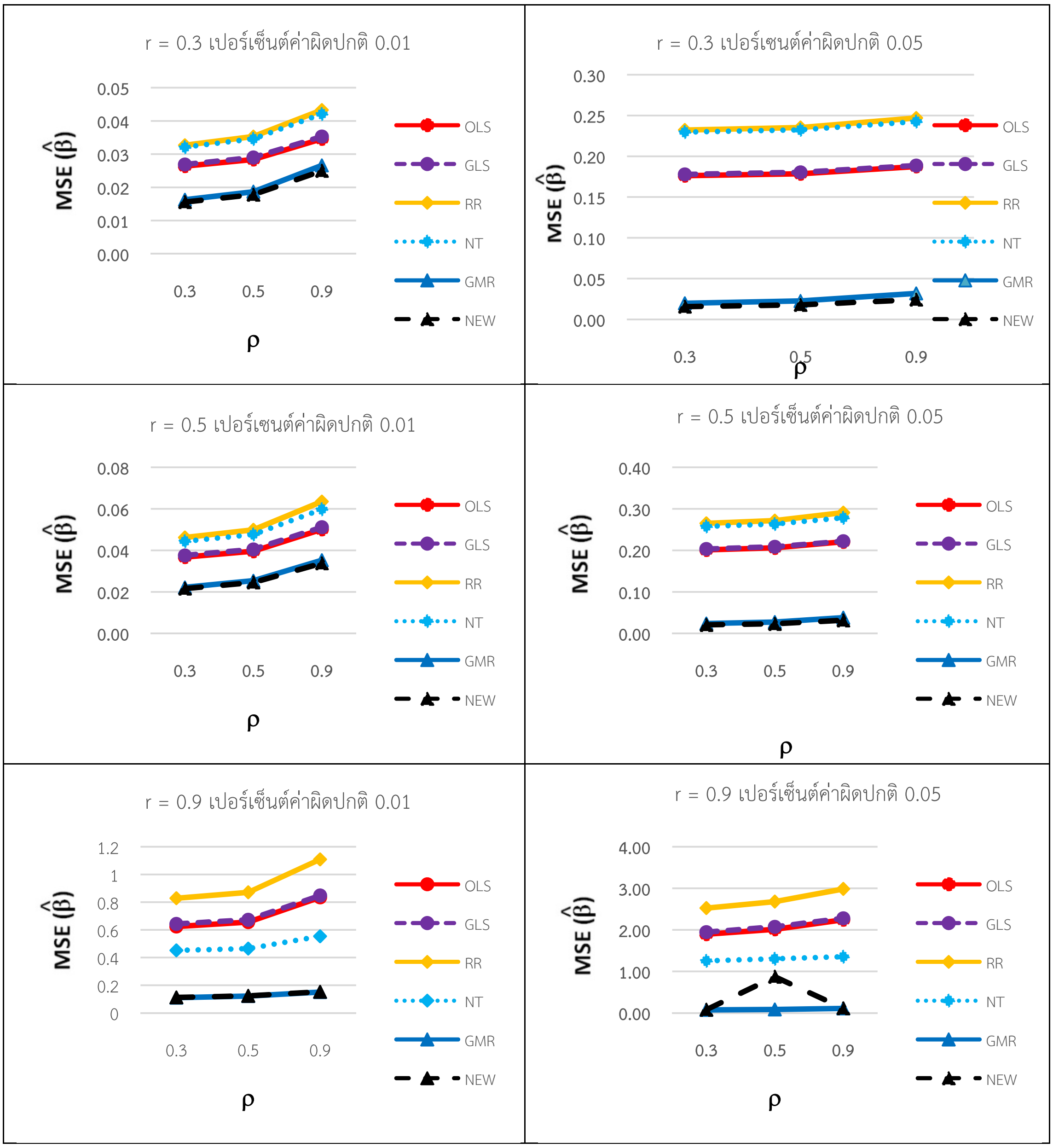
ภาพที่ 10 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50

ตารางที่ 14 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	0.0264	0.0269	0.0327	0.0321	0.0163	**0.0156***
		5%	0.1762	0.1778	0.2323	0.2297	0.0198	**0.0156***
	0.5	1%	0.0368	0.0376	0.0462	0.0444	0.0222	**0.0216***
		5%	0.2011	0.2034	0.2651	0.2577	0.0238	**0.0205***
	0.9	1%	0.6241	0.6427	0.8281	0.4518	**0.1098***	0.1124
		5%	1.8964	1.9464	2.5243	1.2533	**0.0725***	0.0767
0.5	0.3	1%	0.0284	0.0290	0.0353	0.0346	0.0187	**0.0178***
		5%	0.1786	0.1803	0.2354	0.2325	0.0226	**0.0177***
	0.5	1%	0.0396	0.0405	0.0499	0.0477	0.0254	**0.0246***
		5%	0.2062	0.2087	0.2718	0.2633	0.0272	**0.0231***
	0.9	1%	0.6565	0.6725	0.8711	0.4646	**0.1211***	0.1233
		5%	2.0147	2.0747	2.6817	1.3028	**0.0824***	0.8714
0.9	0.3	1%	0.0346	0.0353	0.0433	0.0421	0.0266	**0.0250***
		5%	0.1875	0.1887	0.2470	0.2426	0.0319	**0.0243***
	0.5	1%	0.0501	0.0512	0.0635	0.0598	0.0353	**0.0338***
		5%	0.2207	0.2222	0.2909	0.2785	0.0379	**0.0315***
	0.9	1%	0.8351	0.8480	1.1090	0.5536	**0.1504***	0.1523
		5%	2.2434	2.2797	2.9863	1.3524	**0.1069***	0.1115

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 14 และภาพที่ 11 ปรากฏในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยวิธีของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



ภาพที่ 11 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100

กรณีที่ 3 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y

กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ได้ผลการวิจัยดังนี้

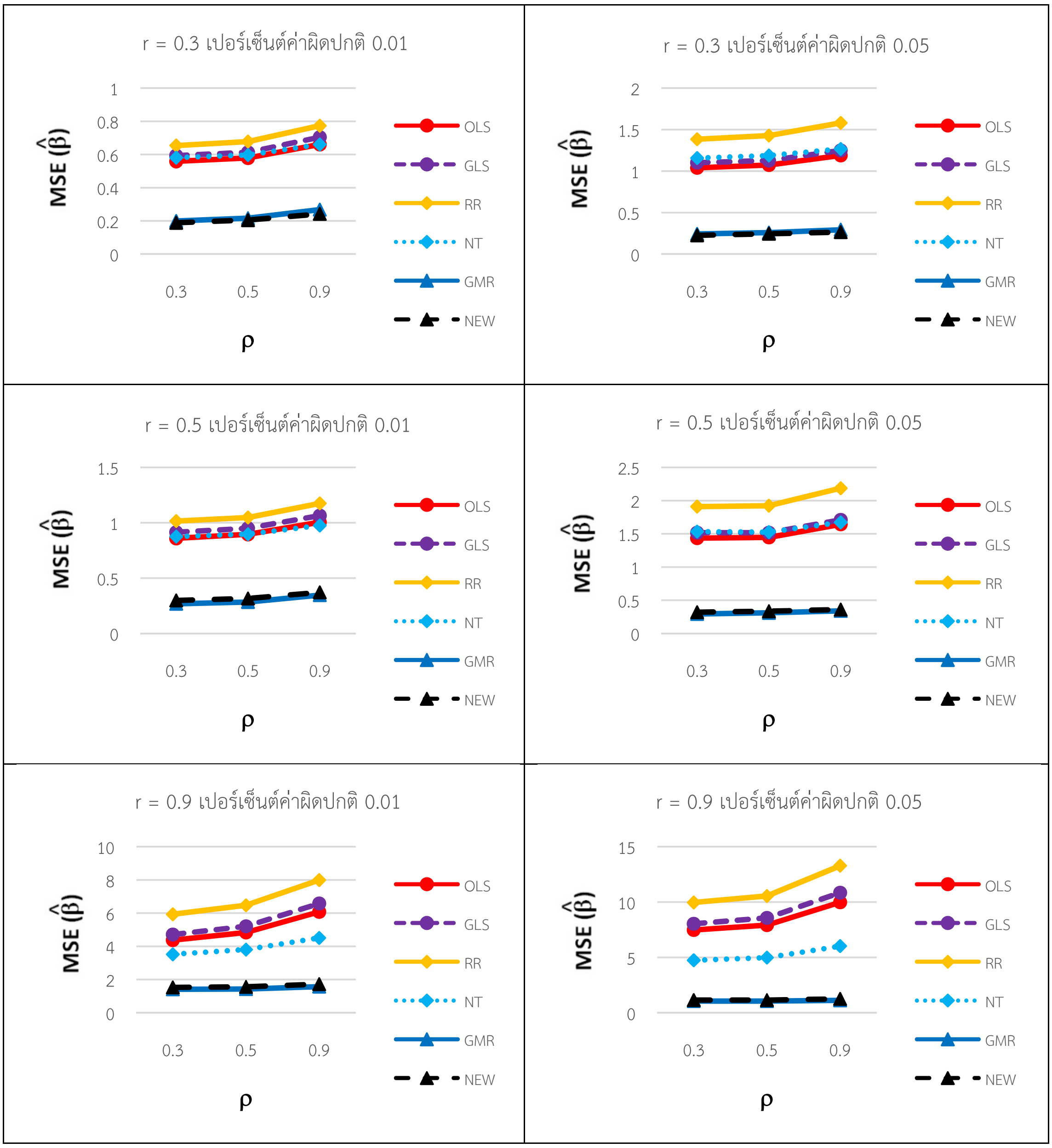
ตารางที่ 15 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	0.5586	0.5934	0.6543	0.5808	0.1996	**0.1888***
		5%	1.0403	1.1021	1.3837	1.1578	0.2431	**0.2259***
	0.5	1%	0.8592	0.9144	1.0154	0.8752	**0.2684***	0.2982
		5%	1.4356	1.5176	1.9100	1.5337	**0.2944***	0.3204
	0.9	1%	4.3773	4.7142	5.9331	3.5202	**1.4147***	1.5298
		5%	7.4863	8.0414	9.9636	4.7361	**1.0524***	1.1455
0.5	0.3	1%	0.5787	0.6130	0.6782	0.5970	0.2163	**0.2048***
		5%	1.0736	1.1268	1.4280	1.1882	0.2590	**0.2461***
	0.5	1%	0.8953	0.9500	1.0478	0.8970	**0.2848***	0.3157
		5%	1.4457	1.5175	1.9232	1.5266	**0.3107***	0.3346
	0.9	1%	4.8357	5.2090	6.4738	3.8062	**1.4367***	1.5631
		5%	7.9253	8.5744	10.5500	4.9774	**1.0588***	1.1435
0.9	0.3	1%	0.6607	0.7047	0.7733	0.6607	0.2682	**0.2428***
		5%	1.1886	1.2496	1.5804	1.2671	0.2939	**0.2655***
	0.5	1%	1.0050	1.0647	1.1754	0.9755	**0.3469***	0.3717
		5%	1.6434	1.7100	2.1866	1.6724	**0.3399***	0.3597
	0.9	1%	6.0753	6.5792	7.9943	4.5074	**1.5760***	1.7259
		5%	9.9858	10.8432	13.2865	6.0235	**1.1255***	1.2545

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 15 และภาพที่ 12 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้น

ให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 และ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



ภาพที่ 12 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนสำหรับขนาดตัวอย่างเท่ากับ 20

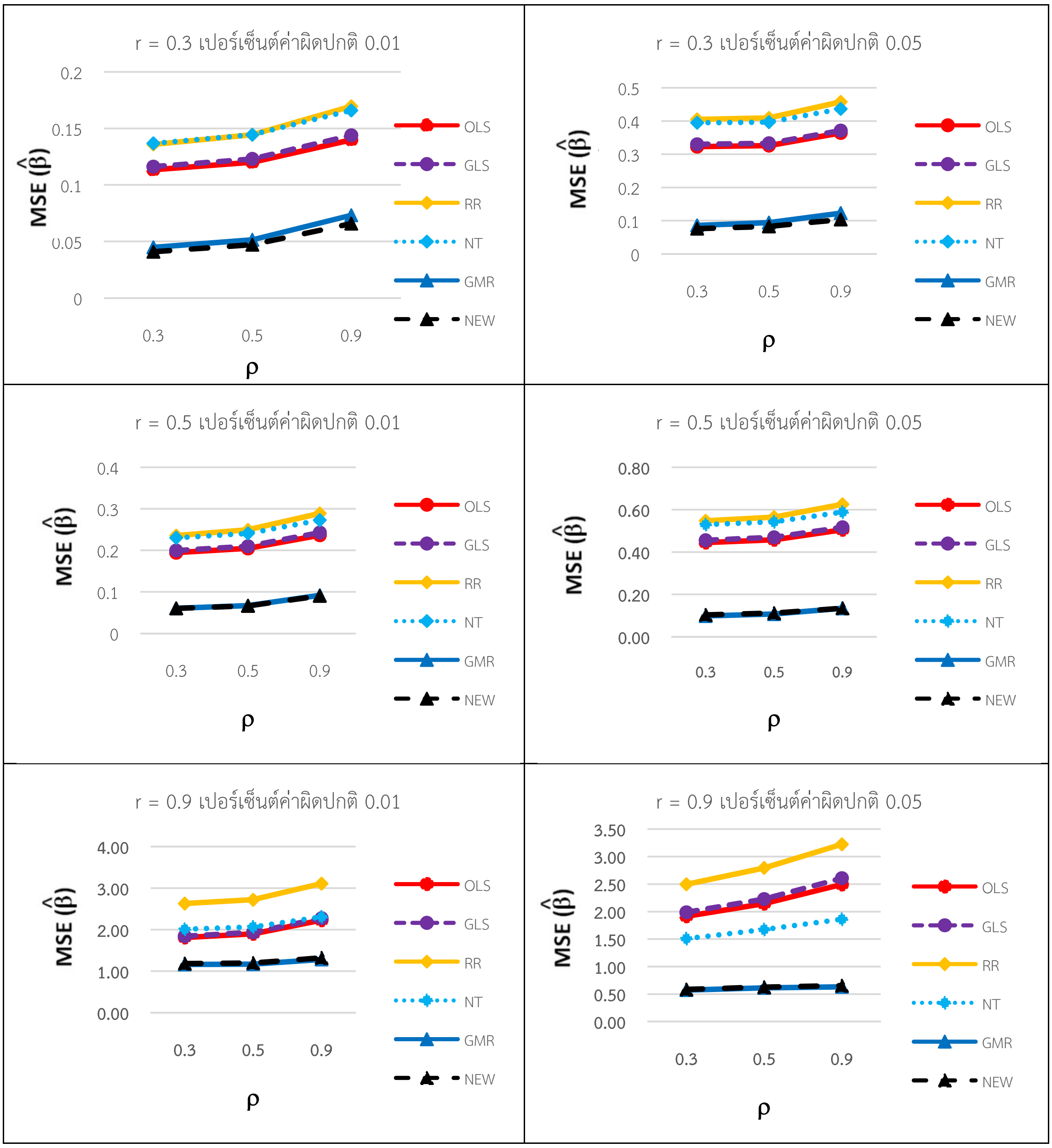
ตารางที่ 16 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนสำหรับขนาดตัวอย่างเท่ากับ 50

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	0.1134	0.1162	0.1359	0.1368	0.0448	**0.0409***
		5%	0.3229	0.3301	0.4050	0.3947	0.0857	**0.0761***
	0.5	1%	0.1944	0.1995	0.2355	0.2301	0.0604	**0.0601***
		5%	0.4446	0.4555	0.5476	0.5295	**0.0983***	0.1034
	0.9	1%	1.8121	1.8412	2.6283	2.0083	**1.1601***	1.1826
		5%	1.9139	1.9861	2.4959	1.5085	**0.5723***	0.5854
0.5	0.3	1%	0.1201	0.1230	0.1445	0.1444	0.0515	**0.0472***
		5%	0.3261	0.3329	0.4095	0.3967	0.0940	**0.0828***
	0.5	1%	0.2046	0.2097	0.2493	0.2409	0.0666	**0.0663***
		5%	0.4580	0.4694	0.5648	0.5424	**0.1071***	0.1118
	0.9	1%	1.8998	1.9326	2.7194	2.0637	**1.1664***	1.1952
		5%	2.1450	2.2266	2.7935	1.6744	**0.6123***	0.6258
0.9	0.3	1%	0.1399	0.1439	0.1696	0.1659	0.0732	**0.0658***
		5%	0.3640	0.3711	0.4579	0.4366	0.1227	**0.1036***
	0.5	1%	0.2364	0.2426	0.2889	0.2727	0.0915	**0.0903***
		5%	0.5054	0.5168	0.6254	0.5885	**0.1346***	0.1349
	0.9	1%	2.2277	2.2751	3.1084	2.3058	**1.2811***	1.3216
		5%	2.4932	2.6060	3.2231	1.8612	**0.6309***	0.6536

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 16 และภาพที่ 13 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 และค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 นอกนั้นวิธีริดจ์ชนิดตัวประมาณ

จีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด



ภาพที่ 13 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนสำหรับขนาดตัวอย่างเท่ากับ 50

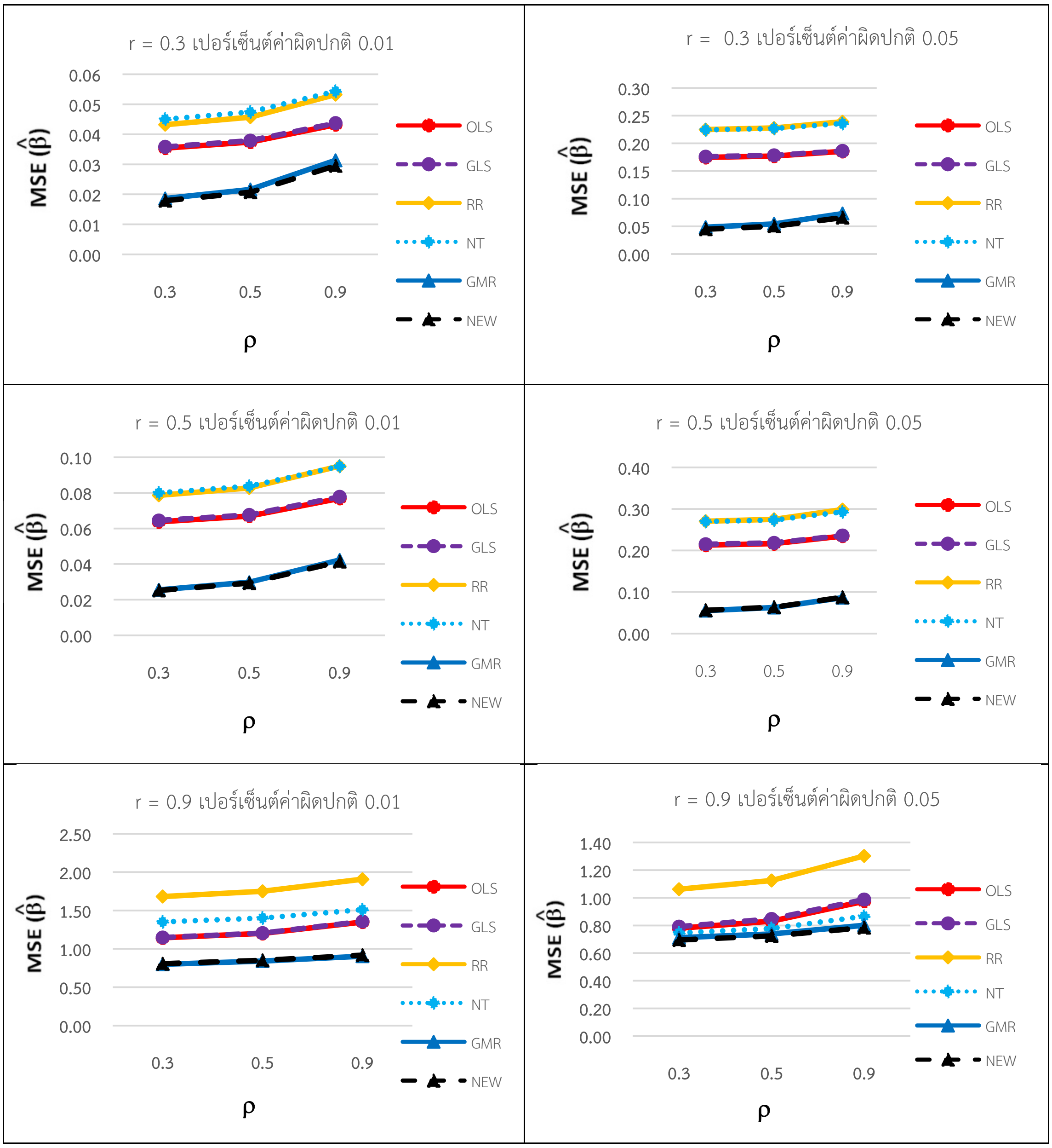
ตารางที่ 17 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100

ρ	r	เปอร์เซ็นต์ค่าผิดปกติ	วิธีประมาณ					
			OLS	GLS	RR	NT	GMR	NEW
0.3	0.3	1%	0.0354	0.0358	0.0432	0.0450	0.0186	**0.0179***
		5%	0.1745	0.1760	0.2248	0.2243	0.0485	**0.0449***
	0.5	1%	0.0637	0.0645	0.0786	0.0800	0.0255	**0.0252***
		5%	0.2128	0.2151	0.2704	0.2697	**0.0550***	0.0559
	0.9	1%	1.1427	1.1477	1.6821	1.3507	**0.7986***	0.8062
		5%	0.7800	0.7917	1.0622	0.7462	0.7115	**0.6947***
0.5	0.3	1%	0.0374	0.0379	0.0457	0.0474	0.0216	**0.0207***
		5%	0.1768	0.1783	0.2277	0.2267	0.0545	**0.0501***
	0.5	1%	0.0669	0.0676	0.0827	0.0837	0.0298	**0.0293***
		5%	0.2164	0.2184	0.2750	0.2733	**0.0626***	0.0634
	0.9	1%	1.2009	1.2054	1.7504	1.4016	**0.8416***	0.8498
		5%	0.8309	0.8473	1.1253	0.7783	0.7387	**0.7246***
0.9	0.3	1%	0.0432	0.0437	0.0532	0.0543	0.0313	**0.0295***
		5%	0.1855	0.1864	0.2390	0.2358	0.0731	**0.0659***
	0.5	1%	0.0769	0.0778	0.0950	0.0949	0.0423	**0.0414***
		5%	0.2347	0.2360	0.2983	0.2927	**0.0863***	0.0872
	0.9	1%	1.3490	1.3568	1.9072	1.5110	**0.9040***	0.9166
		5%	0.9747	0.9879	1.3032	0.8660	0.8019	**0.7845***

* ค่า MSE($\hat{\beta}$) น้อยที่สุด

จากตารางที่ 17 และภาพที่ 14 ปรากฏว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 และค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 และค่าเปอร์เซนต์ความผิดปกติ

5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3 0.5 และ 0.9 นอกนั้น วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด



ภาพที่ 14 การเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\beta}$ ทั้ง 6 วิธี เมื่อข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนสำหรับขนาดตัวอย่างเท่ากับ 100

สรุปได้ว่าสถานการณ์ทั้งหมดที่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นมีค่าต่ำกว่าวิธีประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม จำนวน 61 สถานการณ์ ที่สอดคล้องเป็นไปตามสมมติฐานการวิจัยข้อที่ 1 และขนาดตัวอย่าง ค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนมีอิทธิพลต่อค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ คือค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มมากขึ้น และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณมีแนวโน้มเพิ่มขึ้นเมื่อค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนมีค่าเพิ่มมากขึ้น

ตอนที่ 3 ผลการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

ในการนำวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่มาประยุกต์ใช้กับข้อมูลทุติยภูมิเพื่อพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยใช้ข้อมูลเกี่ยวกับยางพาราชนิดน้ำยางข้น ได้แก่ปริมาณการส่งออกยางพาราของไทยไปประเทศมาเลเซีย ราคาส่งออกยางพาราไทย ใช้ข้อมูลจากส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557 ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ใช้ข้อมูลจากสำนักงานคณะกรรมการเศรษฐกิจและสังคมภาคพื้นเอเชียและแปซิฟิค (Economic and Social Commission for Asia and Pacific) ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557

แบบจำลองอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย

จากการศึกษางานวิจัยที่เกี่ยวข้องกับอุปสงค์การส่งออกยางพาราของไทยสรุปได้ว่าปัจจัยสำคัญในการศึกษาอุปสงค์การส่งออกยางพาราไทยไปยังมาเลเซีย (Y) ได้แก่ ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย (X_1) ส่วนตัวแปรราคาได้ใช้ราคาส่งออกยางพาราของไทย (X_2) และปริมาณการส่งออกยางพาราไทยไปมาเลเซียในปีที่ผ่านมา (X_3) ซึ่งในแบบจำลองนี้ไม่ได้นำปัจจัยอัตราแลกเปลี่ยนของมาเลเซียมาพิจารณา เนื่องจากอัตราแลกเปลี่ยนของมาเลเซียได้กำหนดแบบคงที่กับค่าเงินดอลลาร์สหรัฐอเมริกา โดยยางพาราหรือยางธรรมชาติที่นำมาใช้ในการงานวิจัยนี้ใช้ยางธรรมชาติชนิดน้ำยางข้นเนื่องจากประเทศมาเลเซียมีการนำเข้าในปริมาณมากที่สุดเมื่อเทียบกับชนิดอื่น ๆ ทุกปี

ในการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ด้วยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้นภายใต้สมมติฐานที่ว่า อุปสงค์

การส่งออกยางพาราของไทยไปประเทศมาเลเซีย มีความสัมพันธ์ในทิศทางเดียวกับผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย และอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมา แต่มีความสัมพันธ์ในทิศทางตรงกันข้ามกับราคาส่งออกยางพาราของไทย โดยในการวิเคราะห์ข้อมูลนั้นจะใช้ข้อมูลในปี พ.ศ. 2542 ถึง ปี พ.ศ. 2556 ในการสร้างสมการพยากรณ์การถดถอยเชิงเส้นพหุคูณ และใช้ข้อมูล ปี พ.ศ. 2557 ในการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ใน ปี พ.ศ. 2557

สามารถเขียนเป็นสมการอุปสงค์การส่งออกยางพาราไทยไปยังมาเลเซีย ดังนี้

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

โดยที่ Y คือ ปริมาณการส่งออกยางพาราไทยไปมาเลเซีย (ตัน)
X_1 คือ ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย (ล้านริงกิต)
X_2 คือ ราคาส่งออกยางพาราไทย (บาทต่อตัน)
X_3 คือ ปริมาณการส่งออกยางพาราไทยไปมาเลเซียในปีที่ผ่านมา (ตัน)

นำข้อมูลที่เก็บรวบรวมได้มาทำการทดสอบข้อสมมติของตัวแบบการถดถอยพบว่า ค่าความคลาดเคลื่อนมีการแจกแจงปกติ ความแปรปรวนของค่าคลาดเคลื่อนมีค่าคงที่ ไม่เกิดปัญหาอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน ไม่มีข้อมูลที่มีค่าผิดปกติ แต่พบปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระเกิดขึ้นคือเกิดความสัมพันธ์กันระหว่างตัวแปร X_3 กับ X_1 และ X_3 กับ X_2 และตัวแปรอิสระทุกตัวมีความสัมพันธ์กับตัวแปรตาม (ภาคผนวก ข) จึงทำการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ด้วยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

เนื่องจากข้อมูลที่ใช้ในการวิจัยนี้มีค่าที่เป็นปริมาณสูงมาก จึงต้องปรับข้อมูลให้อยู่ในรูปมาตรฐาน ได้สมการอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ดังนี้

$$z_{yi} = \alpha_1 z_{i1} + \alpha_2 z_{i2} + \alpha_3 z_{i3}$$

เมื่อ z_y เป็น ปริมาณการส่งออกยางพาราไทยไปมาเลเซียในรูปมาตรฐาน
z_1 เป็น ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซียในรูปมาตรฐาน
z_2 เป็น ราคาส่งออกยางพาราไทยในรูปมาตรฐาน
และ z_3 เป็น ปริมาณการส่งออกยางพาราไทยไปมาเลเซียในปีที่ผ่านมาในรูปมาตรฐาน

ผลการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ด้วยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น โดยใช้ค่า k ของ Hoerl Kennard and Baldwin และเกณฑ์ความแกร่งของ Ramsay แสดงได้ดังตารางที่ 18 ดังนี้

ตารางที่ 18 ค่าประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย ค่าคลาดเคลื่อนกำลังสองเฉลี่ย และค่าสัมประสิทธิ์การกำหนดของวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่

$\hat{\alpha}$	$Cov(\hat{\alpha})$	MSE	R^2
$\begin{bmatrix} 0.274685 \\ -0.495796 \\ 1.057818 \end{bmatrix}$	$\begin{bmatrix} 0.038748 & -0.028177 & -0.007791 \\ -0.028177 & 0.045690 & -0.016016 \\ -0.007791 & -0.016016 & 0.028441 \end{bmatrix}$	0.139717	0.848639

จากตารางที่ 18 แสดงว่าในการประมาณค่าสัมประสิทธิ์การถดถอย $\hat{\alpha}$ ที่ได้จากการประมาณค่าด้วยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่มีสมการดังนี้

$$z_y = 0.274685z_1 - 0.495796z_2 + 1.057818z_3$$

โดยมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.139717 และมีค่าสัมประสิทธิ์การกำหนดเท่ากับ 0.848639 หรือกล่าวได้ว่าผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ราคาส่งออกยางพาราของไทย และอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมาสามารถพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีปัจจุบันได้ร้อยละ 84.86 สูงกว่าร้อยละ 70 ซึ่งสอดคล้องเป็นไปตามสมมติฐานการวิจัยข้อที่ 2

และจากค่าสัมประสิทธิ์การถดถอย ($\hat{\alpha}$) แสดงว่าอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย มีความสัมพันธ์ในทิศทางเดียวกับผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซียและอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมา แต่มีความสัมพันธ์ในทิศทางตรงกันข้ามกับราคาส่งออกยางพาราของไทย โดยอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมามีอิทธิพลต่ออุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียมากที่สุด รองลงมาคือราคาส่งออกยางพาราของไทย และผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซียตามลำดับ เมื่อพยากรณ์ปริมาณการนำเข้ายางพาราไทยของประเทศมาเลเซีย ในปี พ.ศ. 2557 พบว่ามีปริมาณเฉลี่ย 661.565 พันตัน หรือกล่าวได้ว่าในปี พ.ศ. 2557 อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียมีประมาณ 661.565 พันตัน

บทที่ 5
สรุปและอภิปรายผล

การวิจัยนี้มีวัตถุประสงค์เพื่อการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอย 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มภายใต้สถานการณ์ 162 สถานการณ์ โดยการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลและพยากรณ์อุปสงค์การส่งออกยางพาราไทยของไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น จากผลการศึกษาสามารถสรุปผลได้ดังนี้

สรุปผลการวิจัย

1. ผลการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน

จากการศึกษาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน เมื่อใช้วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็มโดยประมาณค่า k ด้วยวิธีของ Hoerl-Kennard-Baldwin (HKB) และใช้เกณฑ์ความแกร่งของ Ramsay ทำให้สามารถพัฒนาเป็นตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ ได้ดังนี้

$$\hat{\mathbf{b}}_{NEW} = \left(\mathbf{Q}'\mathbf{W}\mathbf{Q} + k\mathbf{I}\right)^{-1}\left(\mathbf{Q}'\mathbf{W}\mathbf{R} + k\mathbf{J}\right)$$

เมื่อ $\mathbf{Q}$ เป็น เมทริกซ์ของตัวแปรอิสระที่อยู่ในรูปมาตรฐานและใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

$\mathbf{R}$ เป็น เวกเตอร์ของตัวแปรตามที่อยู่ในรูปมาตรฐานและใช้เทคนิคการแปลงรูปด้วยวิธีเพรส วินสเทน

$\mathbf{W}$ เป็น เมทริกซ์ทแยงมุมที่ได้จากการทำวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักซ้ำหลายรอบ (IRLS) โดยใช้เกณฑ์ความแกร่งของ Ramsay ขนาด $n \times n$

k เป็น ค่าประมาณที่คำนวณจากสูตร HKB $k = \dfrac{p\hat{\sigma}^2}{\hat{\mathbf{b}}'\hat{\mathbf{b}}}$

$\mathbf{J}$ เป็น เวกเตอร์ของค่าประมาณเบื้องต้นของ $\mathbf{b}$ จากวิธีกำลังสองน้อยที่สุดสามัญ ค่า $\mathbf{J}$ จะประมาณจาก $\bar{\mathbf{J}}$ ด้วยสูตร $\bar{\mathbf{J}} = \sum_{j=1}^{p} \dfrac{\hat{\mathbf{b}}_j}{p}$

โดยที่ p คือ จำนวนตัวแปรอิสระในตัวแบบ

$\hat{\mathbf{b}}$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ $\boldsymbol{\beta}$

$\hat{\sigma}^2$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ σ^2

$\hat{\mathbf{b}}_j$ คือ ค่าประมาณที่ไม่เอนเอียงจากวิธีกำลังสองน้อยที่สุดสามัญของ $\boldsymbol{\beta}_j$

และเมทริกซ์ความแปรปรวนความแปรปรวนร่วมของ $\hat{\mathbf{b}}_{NEW}$ แสดงได้ดังนี้

$$\mathrm{Cov}\left(\hat{\mathbf{b}}_{NEW}\right) = \hat{\sigma}^2\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}\mathbf{Q}'\mathbf{W}\mathbf{W}'\mathbf{Q}\left(\mathbf{Q}'\mathbf{W}\mathbf{Q}+k\mathbf{I}\right)^{-1}$$

โดยที่ $\hat{\sigma}^2 = \dfrac{SSE}{n-2}$ เมื่อ SSE คือ ผลรวมค่าความคลาดเคลื่อนกำลังสอง

2. ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม

จากตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น ผู้วิจัยทำการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี คือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก วิธีริดจ์ชนิดตัวประมาณจีเอ็ม และวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ ภายใต้สถานการณ์ 162 สถานการณ์ ได้แก่ 1) ขนาดตัวอย่าง 3 ขนาดคือ 20, 50 และ 100 2) ระดับการมีพหุสัมพันธ์ 3 ระดับ คือระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 3) ระดับการมีอัตตสหสัมพันธ์ 3 ระดับ คือระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 4) ลักษณะของการเกิดค่าผิดปกติของข้อมูล 3 แบบ คือ ข้อมูลผิดปกติในตัวแปรอิสระ ข้อมูลผิดปกติในตัวแปรตาม และข้อมูลผิดปกติในตัวแปรอิสระและตัวแปรตาม 5) เปอร์เซนต์ค่าผิดปกติ 2 ระดับ คือ 1% และ 5% ในแต่ละวิธีจะคำนวณค่าคลาดเคลื่อนกำลังสองเฉลี่ยของเวกเตอร์ตัวประมาณค่าสัมประสิทธิ์การถดถอย และจะนำแต่ละวิธีและแต่ละสถานการณ์มาเปรียบเทียบกัน โดยวิธีที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณน้อยที่สุด จะเป็นวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่มีประสิทธิภาพมากที่สุด ผู้วิจัยได้แบ่งการนำเสนอผลการวิจัยออกเป็น 3 กรณีดังนี้

กรณีที่ 1 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3

กรณีที่ 2 ข้อมูลมีค่าผิดปกติในตัวแปรตาม Y

กรณีที่ 3 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y

สรุปผลการวิจัยได้ดังนี้

กรณีที่ 1 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3

กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ได้ผลการวิจัยดังนี้

1.1 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีของนันทวัน จุลมุสิกให้ค่าต่ำที่สุดในทุกสถานการณ์

1.2 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีของนันทวัน จุลมุสิกส่วนใหญ่จะให้ค่าต่ำที่สุดในทุกสถานการณ์ ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และมีอัตตสหสัมพันธ์เท่ากับ 0.3 และ 0.5 วิธีการถดถอยริดจ์จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

1.3 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีของนันทวัน จุลมุสิกส่วนใหญ่จะให้ค่าต่ำที่สุดในทุกสถานการณ์ ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และมีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีการถดถอยริดจ์จะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

กรณีที่ 2 ข้อมูลมีค่าผิดปกติในตัวแปรตาม Y

กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ได้ผลการวิจัยดังนี้

2.1 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ที่ค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่า

ความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 กรณีค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3 และ 0.5 กรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.9 นอกนั้นวิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

2.2 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 50 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ที่ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

2.3 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 100 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ที่ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

กรณีที่ 3 ข้อมูลมีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y

กรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20, 50 และ 100 ได้ผลการวิจัยดังนี้

3.1 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สำหรับขนาดตัวอย่างเท่ากับ 20 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ที่ค่าเปอร์เซนต์ความผิดปกติ 1%

และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 ยกเว้นกรณีค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 และ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 วิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

3.2 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนสำหรับขนาดตัวอย่างเท่ากับ 50 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 และค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 นอกนั้นวิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

3.3 ผลการเปรียบเทียบค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีที่ข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติในตัวแปรอิสระ X_3 และตัวแปรตาม Y และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนสำหรับขนาดตัวอย่างเท่ากับ 100 พบว่าค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีที่พัฒนาขึ้นให้ค่าต่ำที่สุดในสถานการณ์ที่ค่าเปอร์เซนต์ความผิดปกติ 1% และ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 และค่าเปอร์เซนต์ความผิดปกติ 1% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 และค่าเปอร์เซนต์ความผิดปกติ 5% ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 มีอัตตสหสัมพันธ์เท่ากับ 0.3, 0.5 และ 0.9 นอกนั้นวิธีริดจ์ชนิดตัวประมาณจีเอ็มจะให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณต่ำที่สุด

สรุปได้ว่าสถานการณ์ทั้งหมดที่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นมีค่าต่ำกว่าวิธีประมาณค่าสัมประสิทธิ์การถดถอยอีก 5 วิธีคือ วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม จำนวน 61 สถานการณ์ ที่สอดคล้องเป็นไปตามสมมติฐานการวิจัยข้อที่ 1 และขนาดตัวอย่าง ค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนมีอิทธิพลต่อค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ คือค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มมากขึ้น และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ

มีแนวโน้มเพิ่มขึ้นเมื่อค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนมีค่าเพิ่มมากขึ้น

3. ผลการพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

ในการนำวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่มาประยุกต์ใช้กับข้อมูลทุติยภูมิเพื่อพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยใช้ข้อมูลเกี่ยวกับยางพารา ได้แก่ปริมาณการส่งออกยางพาราของไทยไปประเทศมาเลเซีย และราคาส่งออกยางพาราไทย ใช้ข้อมูลจากสถิติยางประเทศไทยสถาบันวิจัยยาง กรมวิชาการเกษตร กระทรวงเกษตรและสหกรณ์ ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557 ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ใช้ข้อมูลจากสำนักงานคณะกรรมการเศรษฐกิจและสังคมภาคพื้นเอเชียและแปซิฟิค (Economic and Social Commission for Asia and Pacific) ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557 โดยในการวิเคราะห์ข้อมูลนั้นจะใช้ข้อมูลในปี พ.ศ. 2542 ถึง ปี พ.ศ. 2556 ในการสร้างสมการพยากรณ์การถดถอยเชิงเส้นพหุคูณ และใช้ข้อมูลปี พ.ศ. 2557 ในการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ใน ปี พ.ศ. 2557

ผลการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ด้วยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่พัฒนาขึ้นโดยประมาณค่า k ด้วยวิธีของ Hoerl-Kennard-Baldwin (HKB) และใช้เกณฑ์ความแกร่งของ Ramsay มีผลการประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย ค่าคลาดเคลื่อนกำลังสองเฉลี่ยและค่าสัมประสิทธิ์การกำหนด พบว่าการประมาณค่าสัมประสิทธิ์การถดถอย ได้ดังสมการ

$$z_y = 0.274685z_1 - 0.495796z_2 + 1.057818z_3$$

เมื่อ z_y เป็น ปริมาณการส่งออกยางพาราไทยไปมาเลเซียในรูปมาตรฐาน
z_1 เป็น ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซียในรูปมาตรฐาน
z_2 เป็น ราคาส่งออกยางพาราไทยในรูปมาตรฐาน
และ z_3 เป็น ปริมาณการส่งออกยางพาราไทยไปมาเลเซียในปีที่ผ่านมาในรูปมาตรฐาน

อธิบายได้ว่าอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย มีความสัมพันธ์ในทิศทางเดียวกับผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย และอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมา แต่มีความสัมพันธ์ในทิศทางตรงกันข้ามกับราคาส่งออกยางพาราของไทย โดยอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมามีอิทธิพลต่ออุปสงค์การส่งออกยางพาราของไทยไปยังประเทศมาเลเซียมากที่สุด รองลงมาคือราคาส่งออกยางพาราของไทย และผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ตามลำดับ

ค่าความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย คือ

$$Cov(\hat{b}) = \begin{bmatrix} 0.038748 & -0.028177 & -0.007791 \\ -0.028177 & 0.045690 & -0.016016 \\ -0.007791 & -0.016016 & 0.028441 \end{bmatrix}$$

มีค่าคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.139717 และค่าสัมประสิทธิ์การกำหนดเท่ากับ 0.848639 หรือกล่าวได้ว่าผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ราคาส่งออกยางพาราของไทย และอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีที่ผ่านมาสามารถพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียในปีปัจจุบันได้ร้อยละ 84.86 ซึ่งมีค่าสูงกว่าร้อยละ 70 จึงยอมรับสมมติฐานการวิจัยข้อที่ 2 เมื่อทำการพยากรณ์ในปี พ.ศ. 2557 พบว่าอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียมีประมาณ 661.565 พันตัน

อภิปรายผล

การวิจัยนี้ได้สิ่งที่ค้นพบซึ่งนำมาอภิปรายได้ดังนี้

1. ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนไปพร้อมกัน เมื่อพิจารณาค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณจากการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณโดยวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก วิธีริดจ์ชนิดตัวประมาณจีเอ็ม และวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ พบว่าไม่ควรใช้ตัวประมาณค่าสัมประสิทธ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดสามัญ ซึ่งสอดคล้องกับข้อสมมติของตัวแบบสมการถดถอย ในกรณีข้อมูลมีค่าผิดปกติที่ตัวแปรอิสระควรใช้ตัวประมาณค่าสัมประสิทธ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีของนันทวัน จุลมุสิก หรือวิธีการถดถอยริดจ์ ส่วนในกรณีข้อมูลมีค่าผิดปกติที่ตัวแปรตาม และข้อมูลมีค่าผิดปกติที่ตัวแปรอิสระและตัวแปรตามควรใช้ตัวประมาณค่าสัมประสิทธ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น หรือวิธีริดจ์ชนิดตัวประมาณจีเอ็ม ซึ่งสอดคล้องกับหลักการของ Huber (1964) ที่เสนอวิธีตัวประมาณเอ็มในการประมาณค่าสัมประสิทธิ์การถดถอยสำหรับข้อมูลที่มีค่าผิดปกติในตัวแปรตาม หรือกรณีที่ค่าคลาดเคลื่อนมีค่าผิดปกติ และหลักการของ Simpson and Montgomery (1996) ที่พัฒนาตัวประมาณจีเอ็มขึ้นมาเพื่อประมาณค่าสัมประสิทธิ์การถดถอยสำหรับข้อมูลที่มีค่าผิดปกติเกิดขึ้นทั้งในตัวแปรอิสระและตัวแปรตาม เนื่องจากวิธีริดจ์ชนิดตัวประมาณจีเอ็มเป็นการนำวิธีตัวประมาณค่าจีเอ็มมารวมกับวิธีการถดถอยริดจ์เพื่อประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณในกรณีเกิดปัญหาการมีความสัมพันธ์ระหว่างตัวแปรอิสระและการมีค่าผิดปกติเกิดขึ้นพร้อมกัน ส่วนวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้นมานั้นเป็นตัวประมาณค่าที่ผู้วิจัยพัฒนามาจากวิธีริดจ์ชนิดตัวประมาณจีเอ็มร่วมกับวิธีกำลังสองน้อยที่สุดทั่วไป เพื่อประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณในกรณีเกิดปัญหาการมีความสัมพันธ์ระหว่างตัวแปรอิสระ การมีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนเกิดขึ้นพร้อมกัน

ปัจจัยที่มีผลต่อวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม คือขนาดตัวอย่าง ค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์

ของค่าคลาดเคลื่อน พบว่าขนาดตัวอย่าง ค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน มีผลต่อค่าคลาดคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี โดยค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณจะแปรผกผันกับขนาดตัวอย่าง และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณจะแปรผันตามค่าพหุสัมพันธ์ระหว่างตัวแปรอิสระ ระดับเปอร์เซนต์ค่าผิดปกติ และค่าอัตตสหสัมพันธ์ของค่าคลาดเคลื่อน สอดคล้องกับผลการศึกษาของอดุลย์เดช กรงทอง (2550) ที่พบว่าเมื่อขนาดตัวอย่างเพิ่มขึ้นค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็มมีแนวโน้มลดลง และเมื่อเปอร์เซนต์ค่าผิดปกติในตัวแปรตามและ/หรือตัวแปรอิสระเพิ่มจาก 1% เป็น 5% ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็มมีแนวโน้มเพิ่มขึ้นเป็นส่วนใหญ่ และสอดคล้องกับผลการศึกษาของวารุณี รวยดี (2551) ที่พบว่าประสิทธิ์ภาพของการประมาณค่าพารามิเตอร์จะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นและจะด้อยประสิทธิภาพลงเมื่อระดับอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนเพิ่มขึ้น

2. การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้นนั้น สมการที่ได้อาจจะไม่เหมาะสมเท่าที่ควร ด้วยเหตุเพราะจากการการทดสอบข้อสมมติของตัวแบบการถดถอยของข้อมูลที่นำมาวิเคราะห์พบว่าข้อมูลชุดนี้เกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระเพียงปัญหาเดียว (ภาคผนวก ข) จึงควรใช้วิธีการถดถอยริดจ์ในการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณในการพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย

ข้อเสนอแนะ

การวิจัยเพื่อศึกษาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนไปพร้อมกัน เพื่อเป็นแนวทางสำหรับการวิจัย ผู้วิจัยขอเสนอแนะ ดังนี้

ข้อเสนอแนะในการนำไปใช้

1. กรณีข้อมูลที่นำมาวิเคราะห์เกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนไปพร้อมกัน โดยข้อมูลมีค่าผิดปกติที่ตัวแปรอิสระเพียงอย่างเดียว ในกรณีขนาดตัวอย่าง 20 และ 50 ควรใช้วิธีของนันทวัน จุลมุสิก ถ้าขนาดตัวอย่าง 100 ควรใช้วิธีของ นันทวัน จุลมุสิก เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันที่ระดับปานกลาง (0.5) และระดับสูง (0.9) และควรใช้วิธีการถดถอยริดจ์เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันระดับต่ำ (0.3)

2. กรณีข้อมูลที่นำมาวิเคราะห์เกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนไปพร้อมกัน โดยข้อมูลมีค่าผิดปกติที่ตัวแปรตามเพียงอย่างเดียว ควรใช้วิธีที่พัฒนาขึ้นกรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ระดับต่ำ

(0.3) และระดับปานกลาง (0.5) และระดับปานกลาง (0.5) และควรใช้วิธีริดจ์ชนิดตัวประมาณจีเอ็ม กรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ระดับสูง (0.9)

3. กรณีข้อมูลที่นำมาวิเคราะห์เกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนไปพร้อมกัน โดยข้อมูลมีค่าผิดปกติที่ตัวแปรอิสระและตัวแปรตาม ควรใช้วิธีที่พัฒนาขึ้นกรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ระดับต่ำ (0.3) ใช้วิธีริดจ์ชนิดตัวประมาณจีเอ็มกรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ระดับปานกลาง (0.5) และระดับสูง (0.9)

ข้อเสนอแนะในการศึกษาต่อไป

1. การศึกษาครั้งนี้วิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น ผู้วิจัยได้กำหนดข้อมูลมีค่าผิดปกติเกิดขึ้นในระดับรุนแรง ใช้เกณฑ์ความแกร่งของ Ramsay และประมาณ k ของ Hoerl, Kenard, and Baldwin เท่านั้น จึงควรมีการเปรียบเทียบกับระดับความรุนแรงของค่าผิดปกติ เกณฑ์ความแกร่ง และวิธีการประมาณค่า k อื่น ๆ

2. ควรมีการเก็บรวบรวมข้อมูลยางพาราได้มากกว่านี้ เนื่องจากในการวิจัยครั้งนี้เกิดข้อจำกัดจากหน่วยงานที่เก็บรวบรวมข้อมูลได้เก็บย้อนหลังแค่ปี พ.ศ. 2541 ข้อมูลที่นำมาวิเคราะห์จึงไม่เกิดปัญหาการมีข้อมูลที่มีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของค่าคลาดเคลื่อนขึ้น

บรรณานุกรม

กัลยา วานิชบัญชา. (2546). *การวิเคราะห์สถิติ: สถิติสำหรับการบริหารและวิจัย* (พิมพ์ครั้งที่ 7). กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

กิตติ วารีรัตน์. (2555). การแก้ไขปัญหาการประมาณค่าพารามิเตอร์ของตัวแบบความถดถอยพหุเชิงเส้นที่เกิดปัญหาพหุสัมพันธ์ด้วยวิธีการวิเคราะห์ตัวประกอบหลัก. วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย.

กิตติพงษ์ ไตรทิพย์พานิชย์. (2550). *การเปรียบเทียบสถิติเดอร์บินวัตสัน ตัวสถิติวอลลิสกับตัวสถิติบูทสแตรปสำหรับการตรวจสอบอัตตสหสัมพันธ์ของความคลาดเคลื่อนในตัวแบบถดถอยเชิงเส้น.* วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย.

จินดา สุวินัยตระกูล. (2545). *การประมาณค่าสัมประสิทธิ์การถดถอย กรณีปัญหามีความสัมพันธ์ระหว่างตัวแปรอิสระและมีค่าผิดปกติ.*วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, มหาวิทยาลัยเชียงใหม่.

จิราวัลย์ จิตรถเวช. (2558). การวิเคราะห์การถดถอย. กรุงเทพฯ: โครงการส่งเสริมและพัฒนาเอกสารวิชาการสถาบันบัณฑิตพัฒนบริหารศาสตร์.

จุฑารัตน์ พรหมทัต. (2553). *การวิเคราะห์ขีดความสามารถในการแข่งขันเพื่อการส่งออกยางพาราไปสาธารณรัฐประชาชนจีน.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์เกษตร, คณะวิทยาศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.

ณรงค์ คงสังข์. (2549). *ปัจจัยที่กำหนดอุปสงค์การนำเข้ายางพาราไทยของจีน ญี่ปุ่น และมาเลเซีย.* วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

ณัฐพร ภักดี. (2552). *การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณเมื่อตัวแปรอิสระมีพหุสัมพันธ์กันโดยวิธีการถดถอยริดจ์.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

ณิชกุล จันทร์มณี. (2551). *ปัจจัยที่ทำให้เกิดความผันผวนของราคายางพาราภายในประเทศ.*วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

เด่นนภา จุลเพชร. (2555). *การเปรียบเทียบการประมาณค่าพารามิเตอร์ระหว่างวิธีกำลังสองน้อยสุดเชิงส่วนและวิธีกำลังสองน้อยสุดด้วยวิธีการแปลงข้อมูลแบบกรามชมิดต์ สำหรับตัวแบบความถดถอยพหุเชิงเส้นที่เกิดพหุสัมพันธ์.* วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย.

ทรงศิริ แต้สมบัติ. (2548). *การวิเคราะห์การถดถอย* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.

ทัศนีย์ ลักษณะ. (2549). การ*วิเคราะห์การส่งออกยางธรรมชาติของประเทศไทย.*วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

ธัญญารัตน์ ไชยเนตรไกรสิน. (2551). *ปัจจัยที่มีผลกระทบและพยากรณ์ปริมาณการส่งออกยางพาราของประเทศไทยไปประเทศญี่ปุ่น.* เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

นันทวัน จุลมุสิก และกิดาการ สายธนู. (2553). การประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณเมื่อค่าคลาดเคลื่อนมีอัตตสหสัมพันธ์และมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ. *วารสารวิทยาศาสตร์ มข.*, *38*(2), 289-302.

เบญจวรรณ ระหงษ์. (2557). การเปรียบเทียบการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุโดยวิธีริดจ์รีเกรสชันที่มีค่าเบื้องต้น กับวิธีมูนิซไคเบรียริดจ์รีเกรสชัน. *วารสารวิชาการมหาวิทยาลัยธนบุรี*, *8*(17), 10-14.

ประชุม สุวัตถี. (2553). *ทฤษฎีการอนุมานเชิงสถิติ* (ปรับปรุงครั้งที่ 3). โครงการส่งเสริมและพัฒนาเอกสารวิชาการ: สถาบันบัณฑิตพัฒนบริหารศาสตร์.

ปราโมทย์ เดชะอำไพ. (2555). *พื้นฐานแมทแลบ* (พิมพ์ครั้งที่ 2). กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ปริญญา สงวนสัตย์. (2556). *คู่มือการใช้งาน MATLAB ฉบับสมบูรณ์.* นนทบุรี: ไอดีซี พรีเมียร์.

ปิยะนาฎ หาดารา. (2549). *การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุเมื่อข้อมูลมีค่าผิดปกติ.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, มหาวิทยาลัยนเรศวร.

พรรณรวีย์ จันทรมาศ. (2556). *ความสัมพันธ์ระหว่างมูลค่าการส่งออกยางพารากับรายได้ในภาคเกษตรของประเทศไทย.* วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

พลพจน์ กิจชัยสวัสดิ์. (2549). อุปสงค์การส่งออกยางธรรมชาติของไทยไปสาธารณรัฐประชาชนจีน. วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

มนตรี พิริยะกุล. (2545). *เทคนิคการวิเคราะห์สมการถดถอย* (ฉบับปรับปรุง). กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยรามคำแหง.

มัลลิกา บุนนาค. (2551). *สถิติเพื่อการวิจัยและตัดสินใจ* (พิมพ์ครั้งที่ 7). กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

มานพ วราภักดิ์. (2548). *ทฤษฎีความน่าจะเป็น.* กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

มานพ วราภักดิ์. (2550). *การจำลอง (Simulation).* กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ยุภาวดี สำราญฤทธิ์. (2552). *การเปรียบเทียบวิธีการแก้ไขปัญหาอัตตสหสัมพันธ์อันดับหนึ่งของความคลาดเคลื่อนในสมการถดถอยเชิงเส้นอย่างง่าย.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

ราชบัณฑิตยสถาน. (2553). *พจนานุกรมศัพท์คณิตศาสตร์ ฉบับราชบัณฑิตยสถาน* 10. กรุงเทพฯ: นานมีบุ๊คส์พับลิเคชั่นส์.

รุ่งโรจน์ เอียดเกิด. (2546). *วิธีประมาณที่แกร่งสำหรับการถดถอยเชิงเส้นพหุ เมื่อข้อมูลมีค่าผิดปกติ.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ลัญฉกร วุฒิสิทธิกุลกิจ และคณะ. (2551). *MATLAB: การประยุกต์ใช้งานทางวิศวกรรมไฟฟ้า* (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

วราฤทธิ์ พานิชกิจโกศลกุล และสุภารัตน์ นิวิศพงศ์. (2552). การเปรียบเทียบวิธีการพยากรณ์ข้อมูลล่วงหน้าหนึ่งคาบเวลาสำหรับตัวแบบอัตตสหสัมพันธ์อันดับที่หนึ่ง เมื่อข้อมูลมีค่าผิดปกติ. *วารสารวิชาการ มหาวิทยาลัยหอการค้าไทย, 29*(4), 54-46.

วริดา พลาศรี. (2552). *การประมาณค่าสัมประสิทธิ์ความถดถอยเชิงพหุสัมพันธ์ด้วยวิธีความถดถอยบูตสแตรปแบบริดจ์.* วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะพาณิชยศาสตร์และการบัญชี, จุฬาลงกรณ์มหาวิทยาลัย.

วารุณี รวยดี. (2551). การเปรียบเทียบวิธีการถ่วงน้ำหนักเฉพาะที่ในการวิเคราะห์การถดถอยสำหรับข้อมูลที่มีความสัมพันธ์กันและมีค่าผิดปกติ. *วารสารวิทยาศาสตร์ประยุกต์, 7*(2), 75-83.

วิทยากร อัศดรวิเศษและคณะ. (2555). *การประยุกต์ใช้ MATLAB.*กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

วิรัชช พานิชวงศ์. (2549). *การวิเคราะห์ถดถอย* (พิมพ์ครั้งที่ 4). กรุงเทพฯ: ศูนย์ผลิตตำราเรียน สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ศราวุธ อินแป้น. (2551). *ความได้เปรียบโดยการเปรียบเทียบในการส่งออกยางพาราไปประเทศสหรัฐอเมริกา.* วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

ศศิธร สว่างวงศ์. (2558). ปัจจัยที่มีอิทธิพลต่อการส่งออกยางพาราของประเทศไทยไปสู่ประเทศจีน. วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

ศิรประภา สุขสำโรง. (2557). *ความสัมพันธ์ของปัจจัยภายในและปัจจัยภายนอกประเทศที่มีต่อการส่งออกยางพาราชนิดยางแผ่นรมควันชั้น 3 ไปประเทศจีน.* วิทยานิพนธ์บริหารธุรกิจมหาบัณฑิต, สาขาธุรกิจระหว่างประเทศ, คณะบริหารธุรกิจ, มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี.

ศิริรัตน์ วงศ์ประกรณ์กุล. (2539). การจำลองโดยใช้เทคนิคมอนติคาร์โล (Monte-Carlo Simulation Technique). *วารสารวิทยาศาสตร์ มข., 24*(4), 240-246.

โศรฎา แข็งการ และกนต์ธร ชำนิประศาสน์. (2556). *การใช้ MATLAB สำหรับงานทางวิศวกรรม* (ปรับปรุงครั้งที่ 2). นครราชสีมา: มหาวิทยาลัยเทคโนโลยีสุรนารี.

สถาบันวิจัยยาง. (2555). ข้อมูลวิชาการยางพารา 2555. กรุงเทพฯ: *กรมวิชาการเกษตร กระทรวงเกษตรและสหกรณ์.*

สถาบันวิจัยยาง. (2555). สถิติยางประเทศไทย THAILAND RUBBER STATISTICS. กรุงเทพฯ: *กรมวิชาการเกษตร กระทรวงเกษตรและสหกรณ์. 41*(4).

สมลักษณ์ หอมสิน. (2551). *การผลิตและการส่งออกยางพาราในประเทศไทย.*วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

สัญชัย บริบูรณ์. (2550). ปัจจัยที่กำหนดการส่งออกยางพาราของประเทศไทยไปตลาดจีน. วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์ธุรกิจ, คณะเศรษฐศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.

สุภัทรา ไชยศร. (2549). *การส่งออกยางแผ่นรมควันชั้น 3 ของไทยไปสหรัฐอเมริกา.* วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, คณะเศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.

สุรางค์ ผลาวัลย์. (2548). ปัจจัยที่กำหนดอุปสงค์และอุปทานของการส่งออกยางพาราของไทย. วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์ธุรกิจ, คณะเศรษฐศาสตร์ มหาวิทยาลัยเกษตรศาสตร์.

เสาวณีย์ ก่อวุฒิกุลรังสี. (2547). *การผลิตยางธรรมชาติ*. ปัตตานี: สำนักพิมพ์มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี.

อดุลย์เดช กรงทอง. (2550). *การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุด้วยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, มหาวิทยาลัยนเรศวร.

อรพรรณ ตันตระกูล. (2555). *การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่แกร่งสำหรับการถดถอยเชิงเส้นพหุ เมื่อข้อมูลมีค่าผิดปกติ.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะวิทยาศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.

อิศรา ชูบำรุง. (2554). การวิเคราะห์ความสามารถในการแข่งขันการส่งออกยางพาราธรรมชาติของประเทศไทย. วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์, มหาวิทยาลัยสุโขทัยธรรมธิราช.

อุบลรัตน์ ภู่ระหงษ์. (2549). *การศึกษาเปรียบเทียบวิธีการถดถอยแบบแกร่งในการประมาณค่าพารามิเตอร์ในการวิเคราะห์การถดถอยเชิงเส้นพหุ.* วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะวิทยาศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.

Alma, O. G. (2011). Comparison of Robust Regression Methods in Linear Regression. *Int. J. Contemp. Math. Sciences, 6(9),* 409-421.

Al-Hassan Y. M. (2010). Performance of new ridge regression estimator. *Journal of the Association of Prob Universities for Basic and Applied Science, 2010(9),* 23-26.

Bagheri, A., H. Midi, M. Ganjali & S. Eftekhari. (2010). A Comparison of Various Influential Points Diagnostic Methods and Robust Regression Approaches: Reanalysis of Interstitial Lang Disease Data. *Applied Mathematical Scienc, 4 (28),* 1367-1386.

Bayhan, G. M., & Bayhan, M. (1998). Forecasting using autocorrelated errors and multicollinear predictor variable. *Computer Industrial Engineering, 34(2)*, 413-421.

Cochrane, D., & Orcutt, G. N. (1949). Application of Least Squares to Relationships Containing Autocorrelated Error Terms. *Journal of the American Statistical Associations, 1949(44),* 32-61.

Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Crouse, R. H., Jin, C., & Hanumara, R. C. (1995). Unbiased ridge estimation with prior information and ridge trace. *Communication in Statistics Theory and Method, 24(9)*, 2341-2354.

Dorugade A. V. (2013). New ridge parameter for ridge regression. *Journal of the Association of Arob Universities for Basic and Applied Science, 2013.*

Draper, N. R., & Smith. H. (1998). *Applied Regression Analysis* (3rd ed.). New York: John Wiley & Sons.

Durbin, J., & Watson, G. S. (1951). Testing for Serial Correlation in Least Squares Regression II. *Biometrika, 1951*(38), 159-178.

Economic and Social Commission for Asia and Pacific. (2010). *Key Indicatorsfor Asia and the Pacific 2010*. New York: The United Nations.

Economic and Social Commission for Asia and Pacific. (2015). *Key Indicatorsfor Asia and the Pacific 2015*. New York: The United Nations.

El-Dereny, M., & Rashwan, N. I. (2001). Solving Multicollinearity Problem Using Ridge Regression Model. *Math Science, 6(12),* 585-600.

Heikkila, J. (2001). *Robust Regression.*Department of Electrical Engineering, University of Oulu.

Hoerl, A. E., & Kennard, R. W. (1970a). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics, 12,* 55-67.

Hoerl, A. E., & Kennard, R. W. (1970b). Ridge regression: application to nonorthogonal problems. *Technometrics, 12,* 69-83.

Hoerl, A. E., Kennard, R. W., & Baldwin, K. F. (1975). Ridge regression: Some simulation. *Communications in Statistics A, 4*, 105-123.

Huber, P. J. (1964). Robust Estimation of a Location Parameter. *Annals of Mathematical Statistic, 35(1)*, 73-101.

International Rubber Study Group. (2007), *Rubber statistical bulletin 2007*, London: Wembley.

International Rubber Study Group. (2011), *Rubber statistical bulletin 2011*, London: Wembley.

Lee, J., & Lund, R. (2004). Revising simple linear regression with autocorrelated errors. *Biometrika, 91*, 240-245.

Montgomery, D. C., Peck, E. A, & Vining G. G. (2001*). Introduction to linear regression analysis* (3rd ed.). New York: John Wiley and Sons Inc.

Mosteller, F., & J. W. Tukey. (1997). *Data Analysis and Regression.*Wiley, United State of America.

Muniz Gisela, & Kibria B. M. Golam. (2008). *"On Some Ridge Regression Estimator: An Empirical Comparisons".*

Muniz Gisela, Kibria B. M. Golam & Shukur Ghazi. (2009). *"On Developing Ridge Regression parameters: AnGraphical investigation".*

Neter, J., M. Kutner, W. Wasserman & C. Nachtsheim. (2005). *Applied Linear Statistical Model* (4th ed.). Singapore: McGraw-Hill.

Nicholson, W. (1994). *Microeconomic theory: Basic principles and extensions* (6th ed.). New York: Thomson Learning Academic Resource Center.

Nicholson, W. (2000). *Microeconomic theory: Basic principles and extensions* (8th ed.). New York: Thomson Learning Academic Resource Center.

Ozkale, M. R. (2008). A jackknifed ridge estimator in the linear regression model with heteroscedastic or correlated errors. *Statistics and Probability Letters, (78),* 3159-3169.

Ozkale, M. R. (2009). A Stochastic restricted. *Journal of Multivariate Analysis, (100)*, 1706-1716.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association, 79(388),* 871-880.

Rousseeuw, P. J., & A. M. Leroy. (2003). *Robust Regression and Outlier Detection.* New York: Wiley.

Savin, N. E., & K. J. White. The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes of many Regressors, *Econometrica 45, 8*(November, 1977): 1992-1995.

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika, (52),* 591-611.

Simpson, D. G., & Chang, Y. I. (1997). Reweighting approximate GM estimators: asymptotics and residual-based graphics. *Journal of Statistical Planning and Inference, (57),* 273-293.

Simpson, J. R., & D. C. Montgomery. (1996). A biased-robust regression technique for the combined outlier-multicollinearity problem. *Journal of Statistical Computation and Simulation, 56(1),* 1-22.

Swindel, B. F. (1976). Good estimators based on prior information. *Communications in Statistics, 5,* 1065-1075.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Melbourne: Addison-Wesley.

Weisberg, S. (1985). *Applied Linear Regression* (2nd ed.). New York: John Wiley & Sons.

ภาคผนวก

ภาคผนวก ก

การเขียนโปรแกรมชุดคำสั่งเพื่อประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี

```
%*******Simulation Main Program in Matlab *******************************
clear all        %Reset all varibles in memory
totcase=1000;    %no. cases
n_sample_size=[20 50 100];  %n sample size n=[20 50 100];
for superloop=1:3
n=n_sample_size(superloop);
filename=['outputN' num2str(n) '.txt']; %create three outputfiles by sample size
fid=fopen(filename,'w');  %Open output file
%   MSE_b(n_set,rhoauto,rhomulicol,errorind,caseind,method)
MSE_b=zeros(totcase,3,3,2,3,6);  %Keep Mean Sqaure Error of Beta
%*******Start case*******************************************************
for n_set=1:totcase          %n_set 1..totcase;
clc;
disp(n_set);
%******Monte Carlo Simulation for generate data set
%******Box Muller u is uniform z is normal dist **************************
u = rand(6,n);                              %Uniform Random 6 variables
r = sqrt(-2*log(u(1,:)));theta = 2*pi*u(2,:);%Polar form P(r,theta)=f(u1,u2)
z(:,1) = r.*cos(theta);z(:,2) = r.*sin(theta);   %  x1<-z(:,1) & x2<-z(:,2)
r = sqrt(-2*log(u(3,:)));theta = 2*pi*u(4,:);%Polar form P(r,theta)=f(u3,u4)
z(:,3) = r.*cos(theta);z(:,4) = r.*sin(theta);   %  x3<-z(:,3) & x4<-z(:,4)
r = sqrt(-2*log(u(5,:)));theta = 2*pi*u(6,:);%Polar form P(r,theta)=f(u5,u6)
z(:,5) = r.*cos(theta);z(:,6) = r.*sin(theta);
%********z1..z6 is Nomal Distibution
%********z1 for x1,z2 for x2,z3 for x3
u=z(:,1);m=mean(u);s=std(u);zu=(u-m)./s;z(:,1)=zu;
u=z(:,2);m=mean(u);s=std(u);zu=(u-m)./s;z(:,2)=zu;
u=z(:,3);m=mean(u);s=std(u);zu=(u-m)./s;z(:,3)=zu;
x=[z(:,1) z(:,2) z(:,3)];   %Matrix x dim nx3
%********z5 is u , z6 is er
%********error(i)=rho*error(i-1)+u(i)
u=z(:,5);m=mean(u);s=std(u);           % u  is N(mu,s)
zu=(u-m)/s;m=mean(zu);s=std(zu);       % zu is N(0,1)
er=z(:,6);m=mean(er);s=std(er);        % er is N(u,s)
zer=(er-m)/s;m=mean(zer);s=std(zer);   % zer is N(0,1)
zer=zer*var(u);                        % Zer is N(0,sigmau^2)
```

```
%*******Generate error with AR(1) autocorrelation rho=0.3,0.5,0.9
error_auto(1,1)=0;error_auto(2,1)=0;error_auto(3,1)=0;
mrhoauto=[0.3 0.5 0.9];
for i=1:3     %0.3 0.5 0.9
  for j=2:n
    error_auto(i,j)=mrhoauto(i)*zer(j-1)+zu(j);
  end
end
%*******Generate Multicollinearity between x1 and x2
%*******Wichern and Churchill (1978)
mrhomulicol=[0.3 0.5 0.9];
for i=1:3                    %rho multicol=0.30, 0.50, 0.90
  for j=1:2
     xmm(:,j)=(1-mrhomulicol(i)^2)*x(:,j)+mrhomulicol(i)*x(:,3);
  end
  xm(i,:,:)=[xmm(:,:) x(:,3)];
end
%****** Outliers x3
    x3=z(:,3);xsort=sort(x3);
    p1=(n+1)/4;pl1=floor(p1);pu1=ceil(p1);
    q1 = xsort(pl1)+(p1-pl1)*(xsort(pu1)-xsort(pl1));
    p3=(n+1)*3/4;pl3=floor(p3);pu3=ceil(p3);
    q3 = xsort(pl3)+(p3-pl3)*(xsort(pu3)-xsort(pl3));
    iqr=q3-q1;
    R1 =q1-4*iqr;           % error 1
    R2 =q1-0.95*4*iqr;      % error 2
    Rn2=q3+0.90*4*iqr;      % error n-2
    Rn1=q3+0.95*4*iqr;      % error n-1
    Rn =q3+4*iqr;           % error n
    for i=1:n
      if x3(i)==xsort(1)     %Minimun or first data
         r1=i;
      end
      if x3(i)==xsort(2)     %Minimun or first data
         r2=i;
      end
```

```
      if x3(i)==xsort(n-1)   %Rn-2
         r3=i;
      end
      if x3(i)==xsort(n-2)   %Rn-1
         r4=i;
      end
      if x3(i)==xsort(n)     %Rn
         r5=i;
      end
    end
    % Rn error 1% of n=20 50 100
    x3noerr=x3;
    x3err1=x3noerr;x3err1(r5)=Rn;          % error Rn
    x3err2=x3err1;x3err2(r1)=R1;      % error Rn R1
    x3err3=x3err2;x3err3(r4)=Rn1;     % error Rn R1 Rn-1
    x3err4=x3err3;x3err4(r2)=R2;      % error Rn R1 Rn-1 R2
    x3err5=x3err4;x3err5(r3)=Rn2;     % error Rn R1 Rn-1 R2 Rn-2
%********************Start Calculate
const=ones(n,1);                   % [1 1 ...1]'
  for rhoauto=1:3                  % [0.3 0.5 0.9]
    for rhomulicol=1:3              %rho multicol=0.30, 0.50, 0.90
      for errorind=1:2            %  error 1% 5%
%%%%%%%
  if errorind==1  %   1% n=20,50,100
      x3_error=x3err1;
  else
      if n==20     % 5% n=20 error 2 value
       x3_error=x3err2
      end
      if n==50     % 5% n=50 error 3 value
       x3_error=x3err3
      end
      if n==100    % 5% n=100 error 5 value
       x3_error=x3err5
      end
  end
```

```
%********calculate y=1+x1+x2+x3+error
 for i=1:n

y_noerr(i)=const(i)+xm(rhomulicol,i,1)+xm(rhomulicol,i,2)+xm(rhomulicol,i,3)+error_au
to(rhoauto,i);

y_real(1,:)=const(i)+xm(rhomulicol,i,1)+xm(rhomulicol,i,2)+x3_error(i)+error_auto(rhoau
to,i);
  x_case(1,i,:)=[const(i) xm(rhomulicol,i,1) xm(rhomulicol,i,2) x3_error(i)];
  x_case(2,i,:)=[const(i) xm(rhomulicol,i,1) xm(rhomulicol,i,2) xm(rhomulicol,i,3)];
  x_case(3,i,:)=[const(i) xm(rhomulicol,i,1) xm(rhomulicol,i,2) x3_error(i)];
 end
%****** Outliers in Case 2&3 set y  error for
    y_out=y_noerr;ysort=sort(y_out);
    p1=(n+1)/4;pl1=floor(p1);pu1=ceil(p1);
    q1 = ysort(pl1)+(p1-pl1)*(ysort(pu1)-ysort(pl1));
    p3=(n+1)*3/4;pl3=floor(p3);pu3=ceil(p3);
    q3 = ysort(pl3)+(p3-pl3)*(ysort(pu3)-ysort(pl3));
    iqr=q3-q1;
    R1 =q1-4*iqr;           % error 1
    R2 =q1-0.95*4*iqr;      % error 2
    Rn2=q3+0.90*4*iqr;      % error n-2
    Rn1=q3+0.95*4*iqr;      % error n-1
    Rn =q3+4*iqr;           % error n
    for i=1:n
      if y_out(i)==ysort(1)     %Minimun or first data
         r1=i;
      end
      if y_out(i)==ysort(2)     %Minimun or first data
         r2=i;
      end
      if y_out(i)==ysort(n-1)   %Rn-2
         r3=i;
      end
      if y_out(i)==ysort(n-2)   %Rn-1
         r4=i;
```

```
      end
      if y_out(i)==ysort(n)      %Rn
         r5=i;
      end
    end
    % Rn error 1% of n=20 50 100
    ynoerr=y_out;
    yerr1=y_noerr;yerr1(r5)=Rn;    % error Rn
    yerr2=yerr1;yerr2(r1)=R1;      % error Rn R1
    yerr3=yerr2;yerr3(r4)=Rn1;     % error Rn R1 Rn-1
    yerr4=yerr3;yerr4(r2)=R2;      % error Rn R1 Rn-1 R2
    yerr5=yerr4;yerr5(r3)=Rn2;     % error Rn R1 Rn-1 R2 Rn-2
%%%%%%%
  if errorind==1  %   1% n=20,50,100
      y_error=yerr1;
  else
      if n==20     % 5% n=20 error 2 value
       y_error=yerr2
      end
      if n==50     % 5% n=50 error 3 value
       y_error=yerr3
      end
      if n==100    % 5% n=100 error 5 value
       y_error=yerr5
      end
  end
%**************************
for i=1:n
  y_real(2,i)=y_error(i);
  y_real(3,i)=y_error(i);
end
for caseind=1:3                    %Run 1,2,3 case
% Y=BX+e
  y=[y_real(caseind,:)]';            % y nx1
  x=[x_case(caseind,:,1)' x_case(caseind,:,2)' x_case(caseind,:,3)' x_case(caseind,:,4)']; %
x nx(p+1)
```

```
  beta0=[1 1 1 1]';                  % Beta initilite (p+1)x1
%%%%%%% Computing Ordinary Least Squares (OLS) method
  [B_OLS,stderr_OLS_b,mse_OLS] = lscov(x,y);
  SSE=(y'*y)-(B_OLS'*x'*y);
  Covb=(SSE/(n-2))*inv(x'*x);
  MSEb(1)=(beta0-B_OLS)'*(beta0-B_OLS)/4;
%%%%%%%  Computing General Least Squares (GLS) method
errorG=y-(B_OLS'*x')';
sumr1=0;sumr21=0;sumr22=0;
sumr2=sum((errorG).^2);
if sumr2==0
  sumr2=1;
end
for i=2:n
    sumr1=sumr1+errorG(i)*errorG(i-1);
end
r_phi=sumr1/sumr2;
invM=zeros(n,n);
invM(1,1)=1;invM(1,2)=-r_phi;
for i=2:n-1
    invM(i,i-1)=-1*r_phi;
    invM(i,i)=1+(r_phi).^2;
    invM(i,i+1)=-1*r_phi;
end
invM(n,n-1)=-r_phi;invM(n,n)=1;
%  M = .2*ones(length(x(:,1))) + .8*diag(ones(size(x(:,1)))); % M nxn
b_GLS=inv(x'*invM*x)*(x'*invM*y);
ssq_b_GLS=(beta0-b_GLS)'*(beta0-b_GLS)/4;
MSEb(2)=ssq_b_GLS;
%%%%%%%  Computing  Ridge regression (RR)
%  Computing k by Hoerl and Kennard (1970) k=p*mse_OLS/(b_OLS'*b_OLS)
   k=3*mse_OLS/(B_OLS'*B_OLS);  %p=3 (x1,x2,x3)
   Im=eye(3);
   H=[x(:,2) x(:,3) x(:,4)];        % No constant term
   xx=(H'*H);   %x'x
   xxkI=xx+k*Im;   %x'x
```

```
   xy=(H'*y);    %x'y
   xxklinv=inv(xx);
   b_RR=xxklinv*xy;
   y_RR=(b_RR'*H')';
   err=y-y_RR;
   SSErr=y'*y-b_RR'*xy;
   MSErr=SSErr/(n-2);
   Covb_rr=MSErr*xxklinv*xx*xxklinv;
   ssq_b_RR=(beta0(1:3)-b_RR)'*(beta0(1:3)-b_RR)/3;
   MSEb(3)=ssq_b_RR;
% R%%%%%%%  Computing Nuthawan (NT)
   k=3*mse_OLS/(B_OLS'*B_OLS);  %p=3 (x1,x2,x3)
   Im=eye(3);
   Jm=(B_OLS(2)+B_OLS(3)+B_OLS(4))/3;
   N=[y_real(caseind,:)]'; % N nx1
   H=[x(:,2) x(:,3) x(:,4)];            % No constant term nxp
   M = .2*ones(length(H(:,1))) + .8*diag(ones(size(H(:,1)))); % M nxn
   hh=H'*inv(M)*H;   %H'*invM*H  pxn nxn nxp
   hhklinv=inv(hh+k*Im);
   hn=(H'*inv(M)*N);   %x'y       pxn nxn nx1
   hnkJ=hn+k*Jm;
   b_NT=hhklinv*hnkJ;
   ssq_b_NT=(beta0(1:3)-b_NT)'*(beta0(1:3)-b_NT)/3;
   MSEb(4)=ssq_b_NT;
% *******  Computing Iteratively Reweighted Least Square: (IRLS) by Beaton and
Huber(1964)
% *******  Computing Generalized M-estimator Ridge (GMR) Ramsay
%****** GMR *******************************************
% init W_H of Ramsay
kR=1;repR=0.001;XR=x;WR=zeros(n,n);
[B_OLS,se_bOLS,mse_OLS]=lscov(XR,y);
errorR=y-(B_OLS'*XR')';
Rb0(kR)=B_OLS(1);Rb1(kR)=B_OLS(2);Rb2(kR)=B_OLS(3);Rb3(kR)=B_OLS(4);
S1=(abs((y'*y)-(B_OLS'*XR'*y)))/(n-4);             %Find MSE of OLS
k1=(4*S1)/abs(B_OLS'*B_OLS);
if sum(errorR) <=0.0001
```

```
   BgmrR=B_OLS;
end
while (kR>=1)
 kR=kR+1;
 DR=zeros(1,n);
 sR=1.4826*median(abs(errorR-median(errorR)));zR=zeros(n,1);
 zzR=errorR/sR;
 aR=0.30;      % a is Tunning constant of Ramsay
 for i=1:n
  piR(i)=(median(abs(zzR)))/zzR(i);
  zR(i)=errorR(i)/(piR(i)*sR);
  DR(i)=exp(-0.30*abs(zR(i)));
 end
WR=diag(DR);
B_R=inv((XR'*WR)*XR)*((XR'*WR)*y);
Rb0(kR)=B_R(1);Rb1(kR)=B_R(2);Rb2(kR)=B_R(3);Rb3(kR)=B_R(4);
diffR=[abs(Rb0(kR)-Rb0(kR-1)),abs(Rb1(kR)-Rb1(kR-1)),abs(Rb2(kR)-Rb2(kR-
1)),abs(Rb3(kR)-Rb3(kR-1))];
if diffR <= [repR,repR,repR,repR]
    Rb0(kR)=B_R(1);Rb1(kR)=B_R(2);Rb2(kR)=B_R(3);Rb3(kR)=B_R(4);
    break;
else
    if kR>=100
     break;
    end
end
b0H=Rb0(kR);b1H=Rb1(kR);b2H=Rb2(kR);b3H=Rb3(kR);
errorR=y-(B_R'*XR')';
%Y_hatH=b0H+b1H*x(:,1)+b2H*x(:,2)+b3H*x(:,3);
%errorH=y-Y_hatH;
BgmrR=inv((XR'*WR*XR)+k1*eye(4))*((XR'*WR)*y);
end
MSEb(5)=(beta0-BgmrR)'*(beta0-BgmrR)/4;
%************** Krongthip *************
XK=x;         %[ones(n,1) x1 x2 x3];
[B_OLS,se_bOLS,mse_OLS]=lscov(XK,y);
```

```
errorK=y-(B_OLS'*XK')';
sumr1=0;sumr2=sum((errorK).^2);
for i=2:n
    sumr1=sumr1+errorK(i)*errorK(i-1);
end
r_kt=sumr1/sumr2;
P=zeros(n,n);
P(1,1)=sqrt(1-r_kt^2);
for i=2:n
    P(i,i-1)=-1*r_kt;P(i,i)=1;
end
R=P*y;
Q=P*XK;
W=WR;               %IRLS method
[B_OLS,se_bOLS,mse_OLS]=lscov(XK,y);
k=(4*mse_OLS)/abs(B_OLS'*B_OLS);
Jm=(B_OLS(1)+B_OLS(2)+B_OLS(3))/3;
B_KT=inv((Q'*WR*Q)+k*eye(4))*((Q'*WR*R)+k*Jm);
MSEb(6)=(beta0-B_KT)'*(beta0-B_KT)/4;
%**************
   for method=1:6
    MSE_b(n_set,rhoauto,rhomulicol,errorind,caseind,method)=MSEb(method);
   end
end
end
end
end
end  % 1000 cases
for rhoauto=1:3
    for rhomulicol=1:3
      for errorind=1:2
         for caseind=1:3
  fprintf(fid,'Y=BX rho auto=%4.2f rho multi=%4.2f n=%i Running=%i
\n',mrhoauto(rhoauto),mrhomulicol(rhomulicol),n,totcase);
  if errorind==1
    fprintf(fid,'Case %i error 1 percent \n',caseind);
```

```
  else
    fprintf(fid,'Case %i error 5 percents \n',caseind);
  end
   fprintf(fid,'            OLS MSEb=%f
\n',mean(MSE_b(:,rhoauto,rhomulicol,errorind,caseind,1)));
   fprintf(fid,'            GLS MSEb=%f
\n',mean(MSE_b(:,rhoauto,rhomulicol,errorind,caseind,2)));
   fprintf(fid,'            Ridge RR MSEb=%f
\n',mean(MSE_b(:,rhoauto,rhomulicol,errorind,caseind,3)));
   fprintf(fid,'            Nuthawan NT MSEb=%f
\n',mean(MSE_b(:,rhoauto,rhomulicol,errorind,caseind,4)));
   mn1=0;sn1=0;
   mn2=0;sn2=0;
   for i=1:totcase
       if isnan(MSE_b(i,rhoauto,rhomulicol,errorind,caseind,5))
       else
           mn1=mn1+1;
           sn1=sn1+MSE_b(i,rhoauto,rhomulicol,errorind,caseind,5);
       end
       if isnan(MSE_b(i,rhoauto,rhomulicol,errorind,caseind,6))
       else
           mn2=mn2+1;
           sn2=sn2+MSE_b(i,rhoauto,rhomulicol,errorind,caseind,6);
       end
   end
   fprintf(fid,'            GMR MSEb=%f \n',sn1/mn1);
   fprintf(fid,'            Krongthip MSEb=%f \n',sn2/mn2);
         end
      end
    end
end
fclose(fid);
disp('********* end of sample size');
end
disp('********* end of running');
```

ภาคผนวก ข

ข้อมูลยางพารา การทดสอบข้อสมมติของตัวแบบการถดถอยค่าประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี

ข้อมูลเรื่อง การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย ตัวแปรที่ใช้ได้แก่

ตัวแปรตาม

Y คือ ปริมาณการนำเข้ายางพาราไทยของประเทศมาเลเซีย (ตัน)

ตัวแปรอิสระ

X_1 คือ ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย (ล้านริงกิต)

X_2 คือ ราคาส่งออกยางพาราไทย (บาทต่อตัน)

X_3 คือ ปริมาณการนำเข้ายางพาราไทยของประเทศมาเลเซียในปีที่ผ่านมา (ตัน)

ตารางภาคผนวกที่ 1 ข้อมูลยางพารา

ปี พ.ศ.	Y	X_1	X_2	X_3
2542	119,516	300,764	17,482	144,997
2543	195,044	356,401	18,688	119,516
2544	289,735	352,579	18,093	195,044
2545	362,108	383,213	19,793	289,735
2546	434,957	418,769	27,526	362,108
2547	483,083	474,048	33,821	434,957
2548	488,333	543,578	36,814	483,083
2549	509,479	596,784	50,194	488,333
2550	578,983	665,340	47,618	509,479
2551	548,359	769,949	54,093	578,983
2552	630,033	712,857	38,498	548,359
2553	584,034	821,434	63,807	630,033
2554	490,918	911,733	86,250	584,034
2555	512,022	971,252	64,484	490,918
2556	578,640	1,018,821	49,234	512,022
2557	537,408	1,106,580	42,565	578,640

ที่มา : จาก 1. ส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์

2.*Key Indicatorsfor Asia and the Pacific 2010* (p. 335), by Economic and Social Commission for Asia and Pacific, 2010, New York: The United Nations.

3.*Key Indicators for Asia and the Pacific 2015* (p. 338), by Economic and Social Commission for Asia and Pacific, 2015, New York: The United Nations.

การทดสอบข้อสมมติของตัวแบบการถดถอย

1. ทดสอบความสัมพันธ์ของตัวแปร จากเมทริกซ์ความสัมพันธ์ (Correlation matrix)

ในการวิจัยครั้งนี้ใช้การตรวจสอบค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายจากเมตริกซ์สหสัมพันธ์ โดยพิจารณาดูว่าตัวแปรอิสระคู่ใดบ้างที่มีค่าสัมประสิทธิ์สหสัมพันธ์สูงเกิน 0.8 ไม่ว่าจะเป็นบวกหรือลบ ให้สงสัยว่าตัวแปรอิสระคู่นั้นอาจเกิดปัญหาพหุสัมพันธ์กัน (วิรัชช พานิชวงศ์, 2549, หน้า 165-167)

ตารางภาคผนวกที่ 2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ

	Y	X_1	X_2
X_1	.729**		
Sig.	.001		
X_2	.664**	.780**	
Sig.	.005	.000	
X_3	.938**	.813**	.814**
Sig.	.000	.000	.000

** มีความสัมพันธ์กันอย่างมีนัยสำคัญ .01

สมมุติฐานของการทดสอบความเป็นอิสระของตัวแปร

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

จากเมทริกซ์สหสัมพันธ์ จะเห็นได้ว่าตัวแปรอิสระ X_3 และ X_1 มีความสัมพันธ์กันที่ระดับ.813 ค่า Sig. เท่ากับ .000 จึงทำให้ปฏิเสธ H_0 ณ ระดับนัยสำคัญ .01 นั่นคือ X_3 กับ X_1 มีความสัมพันธ์กัน และตัวแปรอิสระ X_3 และ X_2 มีความสัมพันธ์กันที่ระดับ .814 ค่า Sig. เท่ากับ .000 จึงทำให้ปฏิเสธ H_0 ณ ระดับนัยสำคัญ .01 นั่นคือ X_3 กับ X_2 มีความสัมพันธ์กัน ส่วนตัวแปรอิสระ X_1 และ X_2 มีความสัมพันธ์กันที่ระดับ .780 ค่า Sig. เท่ากับ .000 จึงทำให้ปฏิเสธ H_0 ณ ระดับนัยสำคัญ .01 นั่นคือ X_1 กับ X_2 มีความสัมพันธ์กัน ดังนั้นสรุปได้ว่า X_1 กับ X_2 มีความสัมพันธ์กันแต่ไม่ได้มากพอจนเกิดปัญหา multicollinearity ในขณะที่ตัวแปร X_3 กับ X_1 และ X_3 กับ X_2 มีความสัมพันธ์กัน จนเกิดปัญหา multicollinearity

2. ทดสอบค่าความคลาดเคลื่อนมีอัตตสหสัมพันธ์

ในการวิจัยครั้งนี้ใช้วิธีการทดสอบของ Durbin-Watson Test (DW)

วิธีทดสอบของเดอร์บินและวัตสันเป็นเครื่องมือทดสอบสหสัมพันธ์ในตัวเองใน AR(1) ที่มีผู้นิยมใช้กันค่อนข้างกว้างขวาง เดอร์บินและวัตสัน เสนอตัวทดสอบสำหรับสมมติฐานต่อไปนี้ ดังนี้

$$\mathrm{H}_0 : \phi = 0$$
$$\mathrm{H}_1 : \phi > 0$$

ปฏิเสธสมมติฐานว่าง ณ ระดับนัยสำคัญ α ถ้า $d < d_L$

ยอมรับสมมติฐานว่าง ถ้า $d > d_U$ และไม่อาจตัดสินใจได้ ถ้า $d_L < d < d_U$

$H_0: \phi = 0$

$H_1: \phi < 0$

ปฏิเสธสมมติฐานว่าง ณ ระดับนัยสำคัญ α ถ้า $d > 4 - d_L$

ยอมรับสมมติฐานว่าง ถ้า $d < 4 - d_U$ และไม่อาจตัดสินใจได้ ถ้า $4 - d_U < d < 4 - d_L$

เปิดตารางค่าสถิติเดอร์บิน-วัตสันที่ระดับนัยสำคัญ α เท่ากับ .05 และตัวอย่างขนาด16 จากภาคผนวก ง จะได้ค่า $d_L = 1.106,\ d_U = 1.371$

จะได้ $d < 4 - d_U = 4 - 1.371 = 2.629,\ d > 4 - d_L = 4 - 1.106 = 2.894$

ตารางภาคผนวกที่ 3 ค่าเดอร์บินและวัตสัน

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.954[a]	.910	.888	48861.6842	2.040

a. Predictors: (Constant), x3, x1, x2

b. Dependent Variable: y

ผลการคำนวณได้ค่า d = 2.0402 หรือค่า d ใกล้ 2 จึงสรุปได้ว่าไม่เกิดปัญหาอัตตสหสัมพันธ์

3. ทดสอบข้อมูลมีค่าผิดปกติ

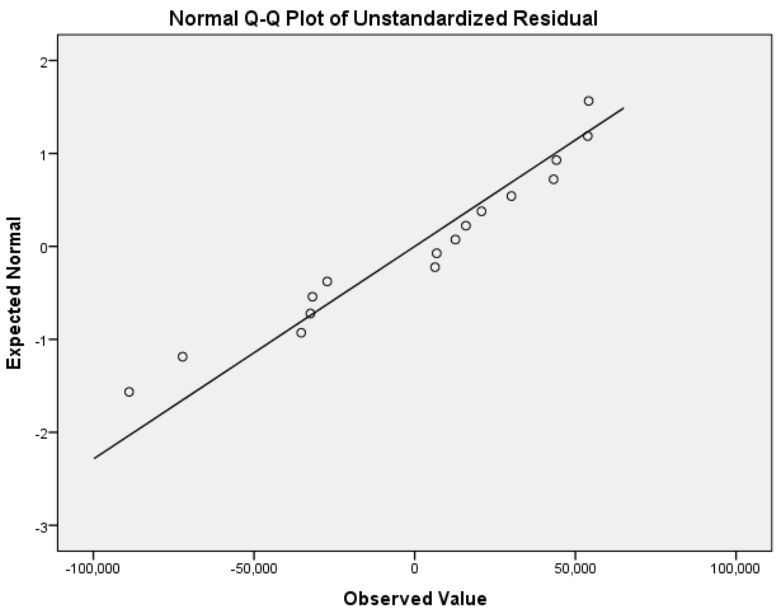
ในการวิจัยครั้งนี้ใช้วิธีการตรวจสอบว่าข้อมูลมีค่าผิดปกติหรือไม่ โดยพิจารณาจากข้อมูลที่ค่าของตัวแปรตามและตัวแปรอิสระอยู่ในช่วง $[Q_1 - 1.5(IQR), Q_3 + 1.5(IQR)]$ เป็นข้อมูลที่ไม่มีค่าผิดปกติคำนวณค่าต่าง ๆ ได้ดังนี้

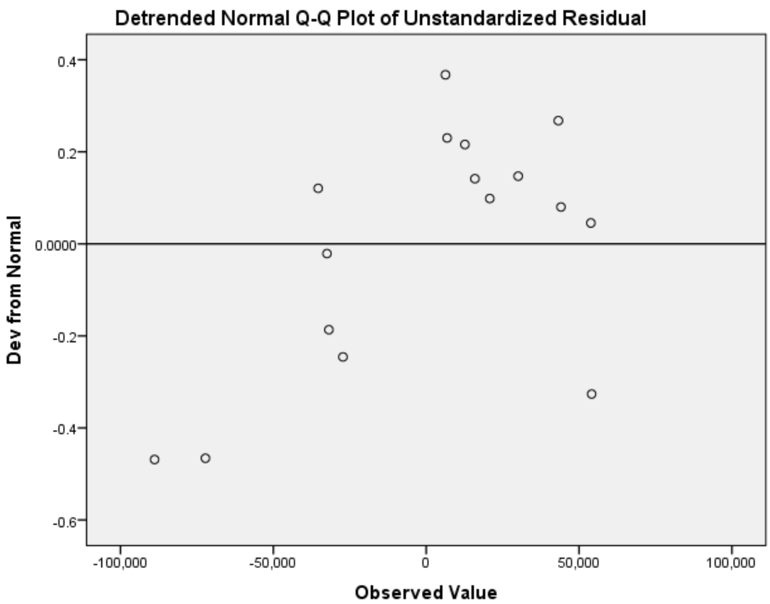
ตารางภาคผนวกที่ 4 ค่า Q_1, Q_3และ IQR

	Y	X_1	X_2	X_3
Q_1	380,320.250	392,102.000	21,726.250	307,828.250
Q_3	571,069.750	889,158.250	53,118.250	571,069.750
IQR	190,749.50	497,056.25	31,392	263,241.50
Q_1-1.5(IQR)	94,196	-353,482.375	-25,361.75	-87,034
Q_3+1.5(IQR)	857,194	1,634,742.625	100,206.25	965.932

นำข้อมูล $Y,\ X_1,\ X_2$ และ X_3 จากตารางภาคผนวกที่ 4 มาพิจารณาพบว่าไม่มีข้อมูลของตัวแปรใดที่อยู่นอกช่วง $[Q_1 - 1.5(IQR), Q_3 + 1.5(IQR)]$ เลย จึงสรุปได้ว่าไม่เกิดปัญหาข้อมูลมีค่าผิดปกติ

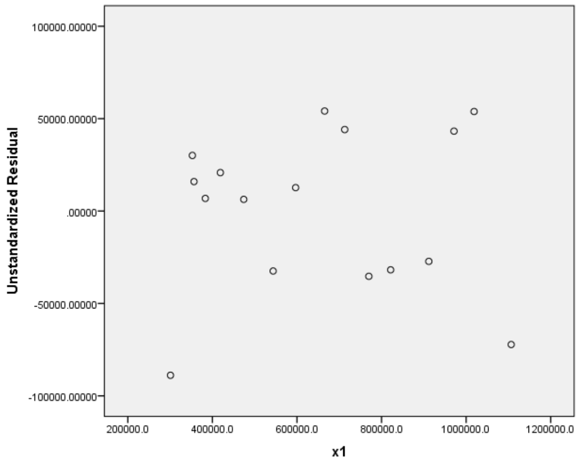
4. ทดสอบค่าคลาดเคลื่อนมีการแจกแจงปกติ

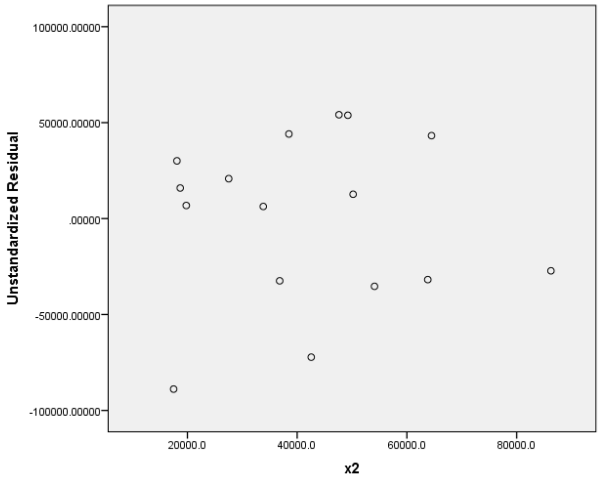


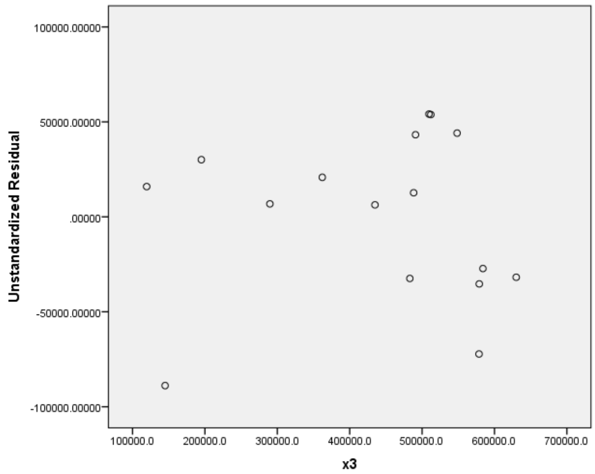


การพิจาณาการลงจุดความน่าจะเป็นแบบ Normal Q-Q Plot พบว่าจุดบนกราฟเรียงตัวกันในแนวเส้นตรง และการลงจุดความน่าจะเป็นแบบ Detrended Normal Q-Q Plot พบว่าจุดบนกราฟกระจายแบบสุ่ม สรุปได้ว่าค่าความคลาดเคลื่อนมีการแจกแจงปกติ

5. ทดสอบค่าคลาดเคลื่อนมีความแปรปรวนคงที่







การพิจาณาการลงจุดพบว่าจุดบนกราฟกระจายอย่างสุ่มและไม่มีรูปแบบ สรุปได้ว่าค่าคลาดเคลื่อนมีความแปรปรวนคงที่

ค่าประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี

ตารางภาคผนวกที่ 5 ค่าประมาณค่าสัมประสิทธิ์การถดถอย เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของสัมประสิทธิ์การถดถอย และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณทั้ง 6 วิธี

วิธีการประมาณค่าพารามิเตอร์	$\hat{\alpha}$	$Cov(\hat{\alpha})$	MSE
OLS	$\begin{bmatrix} 0.290831 \\ -0.568511 \\ 1.177736 \end{bmatrix}$	$\begin{bmatrix} 0.030796 & -0.019259 & -0.006753 \\ -0.019259 & 0.034162 & -0.010731 \\ -0.006753 & -0.010731 & 0.021467 \end{bmatrix}$	0.193757
GLS	$\begin{bmatrix} 0.279170 \\ -0.568230 \\ 1.184574 \end{bmatrix}$	$\begin{bmatrix} 0.020501 & -0.013106 & -0.005805 \\ -0.013106 & 0.022534 & -0.008112 \\ -0.005805 & -0.008112 & 0.015799 \end{bmatrix}$	0.059933
RR	$\begin{bmatrix} 0.290831 \\ -0.568511 \\ 1.177736 \end{bmatrix}$	$\begin{bmatrix} 0.019696 & -0.012591 & -0.005576 \\ -0.012591 & 0.021649 & -0.007793 \\ -0.005576 & -0.007793 & 0.015178 \end{bmatrix}$	0.057577
NT	$\begin{bmatrix} 0.304202 \\ -0.526872 \\ 1.123495 \end{bmatrix}$	$\begin{bmatrix} 0.122743 & 0.062715 & 0.047140 \\ 0.062715 & 0.121406 & 0.048501 \\ 0.047140 & 0.048501 & 0.111128 \end{bmatrix}$	0.345132
GMR	$\begin{bmatrix} 0.289175 \\ -0.512846 \\ 1.034279 \end{bmatrix}$	$\begin{bmatrix} 0.107227 & -0.021805 & -0.002150 \\ -0.021805 & 0.117061 & -0.010653 \\ -0.002150 & -0.010653 & 0.094190 \end{bmatrix}$	0.164075
NEW	$\begin{bmatrix} 0.274685 \\ -0.495796 \\ 1.057818 \end{bmatrix}$	$\begin{bmatrix} 0.038748 & -0.028177 & -0.007791 \\ -0.028177 & 0.045690 & -0.016016 \\ -0.007791 & -0.016016 & 0.028441 \end{bmatrix}$	0.139717

ภาคผนวก ค

การเขียนโปรแกรมชุดคำสั่งเพื่อพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซีย
โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่

```
%*******Simulation Main Program in Mathlab ********************************
clear all        %Reset all varibles in memory
n=15;
p=3;
filename=['outputR' num2str(n) 'ext.txt'];
fid=fopen(filename,'w');  %Open output file
% Prepare data
year_rubber=[2541:1:2556]';
y_quan1=[144997 119516 195044 289735 362108 434957 483083 488333];
y_quan2=[509479 578983 548359 630033 584034 490918 512022 578640 537408];
y_quan=[y_quan1(2:8) y_quan2]';
x_gdp1=[283243 300764 356401 352579 383213 418769 474048 543578];
x_gdp2=[596784 665340 769949 712857 821434 911733 971252 1018821 1106580];
x_gdp=[x_gdp1(2:8) x_gdp2]';
x_price1=[20605 17482 18688 18093 19793 27526 33821 36814];
x_price2=[50194 47618 54093 38498 63807 86250 64484 49234 42565];
x_price=[x_price1(2:8) x_price2]';
x_quan=[y_quan1(1:8) y_quan2(1:8)]';
z=[x_gdpx_pricex_quan];
fprintf(fid,'I  Y  X1 X2 X3 \n');
for i=1:n
fprintf(fid,'%3i   %.0f  %.0f  %.0f %.0f \n',i,y_quan(i),x_gdp(i),x_price(i),x_quan(i));
end
fprintf('\n\n');
u=z(:,1);m=mean(u);s=std(u);zu=(u-m)./s;z(:,1)=zu;
u=z(:,2);m=mean(u);s=std(u);zu=(u-m)./s;z(:,2)=zu;
u=z(:,3);m=mean(u);s=std(u);zu=(u-m)./s;z(:,3)=zu;
const=ones(n,1);                   % [1 1 ...1]'
x=[z(:,1) z(:,2) z(:,3)];   %Matrix x dim nx3
u=y_quan;m=mean(u);s=std(u);           % u  is N(mu,s)
y=(u-m)/s;
fprintf(fid,'I  Y quan  Z1 GDP Z2 price Z3quan \n');
for i=1:n
fprintf(fid,'%3i  %.4f  %.4f %.4f %.4f\n',i,y(i),x(i,1),x(i,2),x(i,3));
end
fprintf('\n\');
```

```
%%%%%%% Computing Ordinary Least Squares (OLS) method
[B_OLS,stderr_OLS_b,mse_OLS] = lscov(x,y);
SSE_OLS=(y'*y)-(B_OLS'*x'*y);
Covb=(SSE_OLS/(n-2))*inv(x'*x);
fprintf(fid,'\n\nOLS Y=BX \n');
fprintf(fid,'B1=%f B2=%f B3=%f \n',B_OLS);
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(1,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(2,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(3,1:3));
fprintf(fid,'MSE=%f \n\n',SSE_OLS/(n-2));
%%%%%%%  Computing General Least Squares (GLS) method
errorG=y-(B_OLS'*x')';
sumr1=0;sumr21=0;sumr22=0;
sumr2=sum((errorG).^2);
ifsumr2==0
sumr2=1;
end
for i=2:n
sumr1=sumr1+errorG(i)*errorG(i-1);
end
r_phi=sumr1/sumr2;
fprintf(fid,'rho phi =%.4f \n',r_phi);
invM=zeros(n,n);
invM(1,1)=1;invM(1,2)=-r_phi;
for i=2:n-1
invM(i,i-1)=-1*r_phi;
invM(i,i)=1+(r_phi).^2;
invM(i,i+1)=-1*r_phi;
end
invM(n,n-1)=-r_phi;invM(n,n)=1;
fprintf(fid,'M^-1 or inv M \n\n');
for i=1:n
for j=1:n
fprintf(fid,'%.4f ',invM(i,j));
end
fprintf(fid,'\n');
```

```
end
%***************************************************************
B_GLS=inv(x'*invM*x)*(x'*invM*y);
SSE_GLS=(y'*y)-(B_GLS'*x'*y);
Covb=(SSE_GLS/(n-2))*inv(x'*x);
fprintf(fid,'GLS Y=BX \n');
fprintf(fid,'B1=%f B2=%f B3=%f \n',B_GLS);
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(1,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(2,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(3,1:3));
fprintf(fid,'MSE=%f \n\n',SSE_GLS/(n-2));
%%%%%%%  Computing  Ridge regression (RR)
%  Computing k by Hoerl and Kennard (1970) k=p*mse_OLS/(b_OLS'*b_OLS)
   k=3*mse_OLS/(B_OLS'*B_OLS);  %p=3 (x1,x2,x3)
Im=eye(3);
   H=x    % No constant term
xx=(H'*H);   %x'x
xxkI=xx+k*Im;   %x'x
xy=(H'*y);   %x'y
xxkIinv=inv(xx);
b_RR=xxkIinv*xy;
y_RR=(b_RR'*H)';
errNT=y-y_RR;
SSErr=y'*y-b_RR'*xy;
MSErr=SSErr/(n-2);
Covb=MSErr*xxkIinv*xx*xxkIinv;
fprintf(fid,'Ridge Regression (RR) Y=BX no constant \n');
fprintf(fid,'Im(n,n) is idenmaxtrix   k=%.4f \n',k);
fprintf(fid,'B1=%f B2=%f B3=%f \n',b_RR);
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(1,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(2,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(3,1:3));
fprintf(fid,'MSE=%f \n\n',SSErr/(n-2));
% R%%%%%%%  Computing Nuthawan (NT)
   k=3*mse_OLS/(B_OLS'*B_OLS);  %p=3 (x1,x2,x3)
Im=eye(3);
```

```
Jm=(B_OLS(1)+B_OLS(2)+B_OLS(3))/3;
   N=y;                 % N nx1
   H=x;            % No constant term nxp
   M = .2*ones(length(H(:,1))) + .8*diag(ones(size(H(:,1)))); % M nxn
hh=H'*inv(M)*H;   %H'*invM*H  pxnnxnnxp
hhkIinv=inv(hh+k*Im);
hn=(H'*inv(M)*N);   %x'ypxnnxnnx1
hnkJ=hn+k*Jm;
b_NT=hhkIinv*hnkJ;
y_NT=(b_NT'*H')';
errNT=N-y_NT;
SSEnt=N'*N-b_NT'*hn;
MSEnt=SSEnt/(n-2);
Covb=MSEnt*(hhkIinv*(H'*inv(M)*H)*hhkIinv);
fprintf(fid,'Nuthawan (NT) Y=BX no constant \n');
fprintf(fid,'k=%.4f \n',k);
fprintf(fid,'B1=%f B2=%f B3=%f \n',b_NT);
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(1,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(2,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(3,1:3));
fprintf(fid,'MSE=%f \n\n',SSEnt/(n-2));
% *******  Computing Iteratively Reweighted Least Square: (IRLS) by Beaton and
          Huber(1964)
% *******  Computing Generalized M-estimator Ridge (GMR) Ramsay
%****** GMR *****************************************
% init W_R of Ramsay
kR=1;repR=0.001;XR=x;WR=zeros(n,n);
[B_OLS,se_bOLS,mse_OLS]=lscov(XR,y);
errorR=y-(B_OLS'*XR')';
Rb0(kR)=B_OLS(1);Rb1(kR)=B_OLS(2);Rb2(kR)=B_OLS(3);
S1=(abs((y'*y)-(B_OLS'*XR'*y)))/(n-3);          %Find MSE of OLS
k1=(3*S1)/abs(B_OLS'*B_OLS);
if sum(errorR) <=0.0001
BgmrR=B_OLS;
end
while (kR>=1)     %  & (kR<=10)
```

```
kR=kR+1;
 DR=zeros(1,n);
sR=1.4826*median(abs(errorR-median(errorR)));zR=zeros(n,1);
zzR=errorR/sR;
aR=0.30;      % a is Tunning constant of Ramsay
for i=1:n
piR(i)=(median(abs(zzR)))/zzR(i);
zR(i)=errorR(i)/(piR(i)*sR);
DR(i)=exp(-0.30*abs(zR(i)));
end
WR=diag(DR);
B_R=inv((XR'*WR)*XR)*((XR'*WR)*y);
Rb0(kR)=B_R(1);Rb1(kR)=B_R(2);Rb2(kR)=B_R(3);
diffR=[abs(Rb0(kR)-Rb0(kR-1)),abs(Rb1(kR)-Rb1(kR-1)),abs(Rb2(kR)-Rb2(kR-1))];
ifdiffR<= [repR,repR,repR]
Rb0(kR)=B_R(1);Rb1(kR)=B_R(2);Rb2(kR)=B_R(3);
break;
else
ifkR>=100
break;
end
end
b0H=Rb0(kR);b1H=Rb1(kR);b2H=Rb2(kR);
errorR=y-(B_R'*XR')';
%Y_hatH=b0H+b1H*x(:,1)+b2H*x(:,2)+b3H*x(:,3);
%errorH=y-Y_hatH;
BgmrR=inv((XR'*WR*XR)+k1*eye(3))*((XR'*WR)*y);
end
SSE_GRM=(y'*y)-(B_R'*x'*y);
Covb=(SSE_GRM/(n-
        2))*(inv(XR'*WR*XR)+k1*eye(3))*(XR'*WR*WR'*XR)*(inv(XR'*WR*XR)+k1*eye(3));
fprintf(fid,'GMR with Ramsay Y=BX \n');
fprintf(fid,'   k=%.4f \n\n',k1);
fprintf(fid,'W \n');
for i=1:n
for j=1:n
```

```
fprintf(fid,'%.4f ',WR(i,j));
end
fprintf(fid,'\n');
end
fprintf(fid,'B1=%f B2=%f B3=%f \n',B_R);
fprintf(fid,'Cov(b)=%f    %f    %f\n',Covb(1,1:3));
fprintf(fid,'Cov(b)=%f    %f    %f\n',Covb(2,1:3));
fprintf(fid,'Cov(b)=%f    %f    %f\n',Covb(3,1:3));
fprintf(fid,'MSE=%f \n\n',SSE_GRM/(n-2));
%************** Krongthip *************
XK=x;         %[ones(n,1) x1 x2 x3];
[B_OLS,se_bOLS,mse_OLS]=lscov(XK,y);
errorK=y-(B_OLS'*XK')';
sumr1=0;sumr2=sum((errorK).^2);
for i=2:n
sumr1=sumr1+errorK(i)*errorK(i-1);
end
r_kt=sumr1/sumr2;
P=zeros(n,n);
P(1,1)=sqrt(1-r_kt^2);
for i=2:n
P(i,i-1)=-1*r_kt;P(i,i)=1;
end
R=P*y;
Q=P*XK;
W=WR;                 %IRLS method
[B_OLS,se_bOLS,mse_OLS]=lscov(XK,y);
k=(3*mse_OLS)/abs(B_OLS'*B_OLS);
Jm=(B_OLS(1)+B_OLS(2)+B_OLS(3))/3;
B_KT=inv((Q'*W*Q)+k*eye(3))*((Q'*W*R)+k*Jm);
SSE_KT=(y'*y)-(B_KT'*x'*y);
Covb=(SSE_KT/(n-2))*inv((Q'*W*Q)+k*eye(3))*(Q'*W*W'*Q)*inv((Q'*W*Q)+k*eye(3));
fprintf(fid,'Krongthip Y=BX \n');
fprintf(fid,'  k=%.4f Jm=%.4f  \n\n',k,Jm);
fprintf(fid,'P(nxn) \n');
for i=1:n
```

```
for j=1:n
fprintf(fid,'%.4f ',P(i,j));
end
fprintf(fid,'\n');
end
fprintf(fid,'I  R=P*y_quan  Q=P*X \n');
for i=1:n
fprintf(fid,'%3i  %.4f  %.4f %.4f %.4f\n',i,R(i),Q(i,1),Q(i,2),Q(i,3));
end
fprintf('\n');
fprintf(fid,'B1=%f B2=%f B3=%f \n',B_KT);
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(1,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(2,1:3));
fprintf(fid,'Cov(b)=%f   %f   %f\n',Covb(3,1:3));
fprintf(fid,'MSE=%f \n\n',SSE_KT/(n-2));
fclose(fid);
disp('********* end of running');
```

ภาคผนวก ง

ตารางสถิติเดอร์บิน-วัตสัน

ค่าสถิติเดอร์บิน-วัตสันที่ระดับนัยสำคัญ α เท่ากับ .01 และ .05 และตัวอย่างขนาด n

Signific.	1%		5%		Signific.	1%		5%	
n	d_L	d_U	d_L	d_U	n	d_L	d_U	d_L	d_U
6	0.390	1.142	0.610	1.400	31	1.147	1.273	1.363	1.496
7	0.435	1.036	0.700	1.365	32	1.160	1.282	1.373	1.502
8	0.497	1.003	0.763	1.332	33	1.172	1.291	1.383	1.508
9	0.554	0.998	0.824	1.320	34	1.184	1.299	1.393	1.514
10	0.604	1.001	0.879	1.320	35	1.195	1.307	1.402	1.519
11	0.653	1.010	0.927	1.324	36	1.206	1.315	1.411	1.525
12	0.697	1.023	0.971	1.331	37	1.217	1.323	1.419	1.530
13	0.738	1.038	1.010	1.340	38	1.227	1.330	1.427	1.535
14	0.776	1.054	1.045	1.350	39	1.237	1.337	1.435	1.540
15	0.811	1.070	1.077	1.361	40	1.246	1.344	1.442	1.544
16	0.844	1.086	1.106	1.371	45	1.288	1.376	1.475	1.566
17	0.874	1.102	1.133	1.381	50	1.324	1.403	1.503	1.585
18	0.902	1.118	1.158	1.391	55	1.356	1.427	1.528	1.601
19	0.928	1.132	1.180	1.401	60	1.383	1.449	1.549	1.616
20	0.952	1.147	1.201	1.411	65	1.407	1.468	1.567	1.629
21	0.975	1.161	1.221	1.420	70	1.429	1.485	1.583	1.641
22	0.997	1.174	1.239	1.429	75	1.448	1.501	1.598	1.652
23	1.018	1.187	1.257	1.437	80	1.466	1.515	1.611	1.662
24	1.037	1.199	1.273	1.446	85	1.482	1.528	1.624	1.671
25	1.055	1.211	1.288	1.454	90	1.496	1.540	1.635	1.679
26	1.072	1.222	1.302	1.461	95	1.510	1.552	1.645	1.687
27	1.089	1.233	1.316	1.469	100	1.522	1.562	1.645	1.694
28	1.104	1.244	1.328	1.476	150	1.611	1.637	1.720	1.746
29	1.119	1.254	1.341	1.483	200	1.664	1.684	1.758	1.778
30	1.133	1.263	1.352	1.489					

ที่มา: Savin, N.E., and K.J. White. The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes of many Regressors, *Econometrica 45,* 8 (November 1977): 1992-1995

ภาคผนวก จ

การพิสูจน์ค่าคาดหวัง ค่าความแปรปรวนร่วม และค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่า
วิธีการถดถอยริดจ์

วิธีการถดถอยริดจ์ (RidgeRegression Method)

1. ค่าคาดหวังของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์

1.1 ค่าคาดหวังของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ของ Hoerl and Kennard (1970)

$$\hat{\boldsymbol{\beta}}_{RR} = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{Y}$$

$$= (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$= \left[(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1} + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y})$$

$$= \left[I + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}\hat{\boldsymbol{\beta}}$$

$$= \mathbf{Z}\hat{\boldsymbol{\beta}}$$

$$\mathrm{E}(\hat{\boldsymbol{\beta}}_{RR}) = \mathbf{Z}\mathrm{E}(\hat{\boldsymbol{\beta}}) = \mathbf{Z}\boldsymbol{\beta}$$

เมื่อ $\mathbf{Z} = \left[\mathbf{I} + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}$ และในกรณีที่ k=0 จะพบว่า $\hat{\boldsymbol{\beta}}_{RR} = \hat{\boldsymbol{\beta}}$

1.2 ค่าคาดหวังของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์ที่มีค่าเบื้องต้นของ Swindel (1976)

$$\hat{\boldsymbol{\beta}}_{RJ} = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{Y} + K\mathbf{J})$$

$$= (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y} + k\mathbf{J})$$

$$= \left[\mathbf{I} + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}\left[\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J} + \mathbf{J} + k(\mathbf{X}'\mathbf{X})^{-1}\mathbf{J}\right]$$

$$= \left[\mathbf{I} + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}\left[(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) + (\mathbf{J} + k(\mathbf{X}'\mathbf{X})^{-1}\mathbf{J})\right]$$

$$= \left[\mathbf{I} + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) + \mathbf{J}$$

$$\mathrm{E}(\hat{\boldsymbol{\beta}}_{RJ}) = \left[\mathbf{I} + k(\mathbf{X}'\mathbf{X})^{-1}\right]^{-1}\mathrm{E}(\hat{\boldsymbol{\beta}}_{OLS} - \mathbf{J}) + \mathrm{E}(\mathbf{J})$$

เนื่องจาก $\bar{\mathbf{J}} = \sum_{j=1}^{p}\frac{\hat{\beta}_j}{p}$ จะได้ว่า $E(\mathbf{J}) = \boldsymbol{\beta}$; $\mathbf{J} \sim (\boldsymbol{\beta}, \frac{\sigma^2\mathbf{I}}{k})$

ดังนั้น $E(\hat{\boldsymbol{\beta}}_{RJ}) = \boldsymbol{\beta}$

2. ความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์

พิสูจน์

$$Cov(\hat{\boldsymbol{\beta}}_R) = \mathrm{E}\left[\hat{\boldsymbol{\beta}}_R - \mathrm{E}(\hat{\boldsymbol{\beta}}_R)\right]\left[\hat{\boldsymbol{\beta}}_R - \mathrm{E}(\hat{\boldsymbol{\beta}}_R)\right]'$$

$$= \mathrm{E}\left[\mathbf{Z}\hat{\boldsymbol{\beta}} - \mathbf{Z}\boldsymbol{\beta}\right]\left[\mathbf{Z}\hat{\boldsymbol{\beta}} - \mathbf{Z}\boldsymbol{\beta}\right]'$$

$$= \mathrm{E}\left[\mathbf{Z}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right]\left[\mathbf{Z}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right]'$$

$$= \mathrm{E}\left[\mathbf{Z}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right]\left[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{Z}'\right]$$

..............................(1)

พิจารณา

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) - \boldsymbol{\beta}\right]$$

$$= \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} - \boldsymbol{\beta}$$

$$= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$$

............................(2)

แทนค่า $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$ ในสมการ (2) จะได้ว่า

$$Cov(\hat{\boldsymbol{\beta}}_R) = \mathrm{E}\left[\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}(\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'\right]$$

$$= \mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\left[\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\right]\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}' \quad ; \mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \sigma^2$$

$$= \sigma^2\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'$$

ดังนั้นค่า $Cov(\hat{\boldsymbol{\beta}}_R) = \sigma^2\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'$

3. ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าสัมประสิทธิ์การถดถอยริดจ์

พิสูจน์

$$\mathrm{MSE}(\hat{\boldsymbol{\beta}}_R) = \mathrm{E}\left[\hat{\boldsymbol{\beta}}_R - \boldsymbol{\beta}\right]'\left[\hat{\boldsymbol{\beta}}_R - \boldsymbol{\beta}\right]$$

จาก $\hat{\boldsymbol{\beta}}_R = \mathbf{Z}\hat{\boldsymbol{\beta}}$

ดังนั้น $\mathrm{MSE}(\hat{\boldsymbol{\beta}}_R) = \mathrm{E}\left[\mathbf{Z}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right]'\left[\mathbf{Z}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right]$

พิจารณา $\mathbf{Z}\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\beta}} = \mathbf{Z}\hat{\boldsymbol{\beta}}-\mathbf{Z}\boldsymbol{\beta}+\mathbf{Z}\boldsymbol{\beta}-\boldsymbol{\beta} = \mathbf{Z}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+(\mathbf{Z}-\mathbf{I})\boldsymbol{\beta}$

ดังนั้น $\mathrm{MSE}(\hat{\boldsymbol{\beta}}_R) = \mathrm{E}\{\left[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{Z}'+\boldsymbol{\beta}'(\mathbf{Z}-\mathbf{I})'\right]\left[\mathbf{Z}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+(\mathbf{Z}-\mathbf{I})\boldsymbol{\beta}\right]\}$

$$= \mathrm{E}[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{Z}'\mathbf{Z}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{Z}'(\mathbf{Z}-\mathbf{I})'\boldsymbol{\beta}$$

$$+\boldsymbol{\beta}'(\mathbf{Z}-\mathbf{I})'\mathbf{Z}\,(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+\boldsymbol{\beta}'(\mathbf{Z}-\mathbf{I})'(\mathbf{Z}-\mathbf{I})\boldsymbol{\beta}]$$

เนื่องจาก $\mathrm{E}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) = 0$

ดังนั้น $\mathrm{MSE}(\hat{\boldsymbol{\beta}}_\mathrm{R}) = \mathrm{E}[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{Z}'\mathbf{Z}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] + (\mathbf{Z}\boldsymbol{\beta}-\boldsymbol{\beta})'(\mathbf{Z}\boldsymbol{\beta}-\boldsymbol{\beta})$

$$= \mathrm{a} + \mathrm{b}$$

พิจารณา $\mathrm{a} = \mathrm{E}[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\mathbf{Z}'\mathbf{Z}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]$

จะได้ว่า $\hat{\boldsymbol{\beta}}-\boldsymbol{\beta} = \boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}-\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}$

ดังนั้น $a = \mathrm{E}[\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}]$

$$= \mathrm{E}[\boldsymbol{\varepsilon}'\{\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}\boldsymbol{\varepsilon}]$$

$$= \sigma^2\mathrm{trace}\{\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\}$$

$$= \sigma^2\mathrm{trace}\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Z}\}$$

แต่ $Z = [\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}$ เป็น Symmetric Matrix

$$= \sigma^2\mathrm{trace}\{(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}[\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}\}$$

$$= \sigma^2\mathrm{trace}\{[\{\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}\}(\mathbf{X}'\mathbf{X})]^{-1}[\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}\}$$

$$= \sigma^2\mathrm{trace}\{(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}]^{-1}\}$$

$$= \sigma^2\mathrm{trace}\{(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X})[\{\mathbf{I}+k(\mathbf{X}'\mathbf{X})^{-1}\}(\mathbf{X}'\mathbf{X})]^{-1}\}$$

$$= \sigma^2\mathrm{trace}\{(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\}$$

$$= \sigma^2\sum_{i=1}^{p}\frac{\lambda_i}{(\lambda_i+k)^2}$$

พิจารณา $b = (\mathbf{Z\beta} - \boldsymbol{\beta})'(\mathbf{Z\beta} - \boldsymbol{\beta})$

$$= [(\mathbf{X'X} + k\mathbf{I})^{-1}(\mathbf{X'X})\boldsymbol{\beta} - \boldsymbol{\beta}]'[(\mathbf{X'X} + k\mathbf{I})^{-1}(\mathbf{X'X})\boldsymbol{\beta} - \boldsymbol{\beta}]$$

$$= \{[(\mathbf{X'X} + k\mathbf{I})^{-1}[(\mathbf{X'X} + k\mathbf{I}) - k\mathbf{I}]\boldsymbol{\beta} - \boldsymbol{\beta}\}'$$

$$\{[(\mathbf{X'X} + k\mathbf{I})^{-1}[(\mathbf{X'X} + k\mathbf{I}) - k\mathbf{I}]\boldsymbol{\beta} - \boldsymbol{\beta}\}$$

$$= \{[\mathrm{I} - k(\mathbf{X'X} + k\mathbf{I})^{-1}]\boldsymbol{\beta} - \boldsymbol{\beta}\}'\{[\mathrm{I} - k(\mathbf{X'X} + k\mathbf{I})^{-1}]\boldsymbol{\beta} - \boldsymbol{\beta}\}$$

$$= \{[\boldsymbol{\beta} - k(\mathbf{X'X} + k\mathbf{I})^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}\}'\{[\boldsymbol{\beta} - k(\mathbf{X'X} + k\mathbf{I})^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}\}$$

$$= \{[-k(\mathbf{X'X} + k\mathbf{I})^{-1}\boldsymbol{\beta}\}'\{[-k(\mathbf{X'X} + k\mathbf{I})^{-1}\boldsymbol{\beta}\}$$

$$= k^2\boldsymbol{\beta}'[(\mathbf{X'X} + k\mathbf{I})^{-1}]^2\boldsymbol{\beta}$$

ดังนั้น $\mathrm{MSE}(\hat{\boldsymbol{\beta}}_{\mathrm{R}}) = \mathrm{E}[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{Z'Z}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})] + (\mathbf{Z\beta} - \boldsymbol{\beta})'(\mathbf{Z\beta} - \boldsymbol{\beta})$

$$= \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2\boldsymbol{\beta}'[(\mathbf{X'X} + k\mathbf{I})^{-1}]^2\boldsymbol{\beta}$$

ภาคผนวก ฉ

การปรับข้อมูลให้อยู่ในรูปมาตรฐาน

การปรับข้อมูลให้อยู่ในรูปมาตรฐาน (Standardized form)

การวิเคราะห์การถดถอยโดยใช้ข้อมูลเดิม (Initial form หรือ Deviation from Zero) มีปัญหาสำคัญอยู่ประการหนึ่งคือข้อมูลเดิมมักจะมีค่าที่เป็นปริมาณค่อนข้างสูง ซึ่งมีผลทำให้ผลรวมกำลังสอง (Sum Square) และผลรวมเทอมไขว้ (Sum of Cross Product Term) จะมีค่าสูงมาก การแก้ปัญหานี้คือการวิเคราะห์ข้อมูลโดยอาศัยการแปลงรูปด้วยการปรับข้อมูลให้อยู่ในรูปมาตรฐาน

จากสมการ

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon_i \quad ; \; i = 1, 2, 3, \ldots, n \qquad (1)$$

เมื่อ Y_i แทน ค่าสังเกตที่ i ของตัวแปรตามที่ทราบค่า

$X_{i1}, X_{i2}, \ldots, X_{ip}$ แทน ค่าสังเกตที่ i ของตัวแปรอิสระ p ตัวที่ทราบค่า

$\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ แทน สัมประสิทธิ์การถดถอยที่ไม่ทราบค่า

ε_i แทน ค่าคลาดเคลื่อนสุ่มของค่าสังเกตที่ i

n แทน จำนวนค่าสังเกต

และ p แทน จำนวนตัวแปรอิสระ

หรือเขียนในรูปของเมทริกซ์ดังนี้

$$\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon} \qquad (2)$$

เมื่อ $\mathbf{Y}$ แทน เวกเตอร์ของค่าสังเกตของตัวแปรตามที่ทราบค่า มีมิติ n x 1

$\mathbf{X}$ แทน เมทริกซ์ของค่าสังเกตของตัวแปรอิสระที่ทราบค่า มีมิติ n x (p+1)

$\boldsymbol{\beta}$ แทน เวกเตอร์ของสัมประสิทธิ์การถดถอยที่ไม่ทราบค่า มีมิติ (p+1) x 1

$\boldsymbol{\varepsilon}$ แทน เวกเตอร์ของค่าคลาดเคลื่อนสุ่ม มีมิติ n x 1

n แทน จำนวนค่าสังเกต

และ p แทน จำนวนตัวแปรอิสระ

การปรับข้อมูลให้อยู่ในรูปมาตรฐานเป็นการลดขนาดของข้อมูล Y และ X's ผลลัพธ์คือ

$$z_{yi} = \frac{Y_i - \bar{Y}}{s_y}$$

$$z_{ij} = \frac{X_{ij} - \bar{X}_j}{s_j}$$

สมการ (1) จึงเปลี่ยนรูปเป็น

$$z_{yi} = \alpha_1 z_{i1} + \alpha_2 z_{i2} + ... + \alpha_p z_{ip} + z_{\varepsilon i} \quad ; i = 1, 2, 3, ..., n \qquad \text{.............................. (3)}$$

เมื่อ z_{yi} แทน ค่าสังเกตที่ i ของตัวแปรตามที่ทราบค่าที่อยู่ในรูปมาตรฐาน
$z_{i1}, z_{i2}, ..., z_{ip}$ แทน ค่าสังเกตที่ i ของตัวแปรอิสระ p ตัวที่ทราบค่าที่อยู่ในรูปมาตรฐาน
$\alpha_1, \alpha_2, ..., \alpha_p$ แทน สัมประสิทธิ์การถดถอยที่ไม่ทราบค่า
$z_{\varepsilon i}$ แทน ค่าคลาดเคลื่อนสุ่มมาตรฐานของค่าสังเกตที่ i
n แทน จำนวนค่าสังเกต
และ p แทน จำนวนตัวแปรอิสระ

หรือเขียนในรูปของเมทริกซ์ ดังนี้

$$\mathbf{N} = \mathbf{H}\boldsymbol{\alpha} + \mathbf{e} \qquad \text{.............................. (4)}$$

เมื่อ $\mathbf{N}$ แทน เวกเตอร์ของตัวแปรตามที่อยู่ในรูปมาตรฐานมีมิติ n x 1
$\mathbf{H}$ แทน เมทริกซ์ของตัวแปรอิสระที่อยู่ในรูปมาตรฐาน มีมิติ n x p
$\boldsymbol{\alpha}$ แทน เวกเตอร์ของสัมประสิทธิ์การถดถอยที่ไม่ทราบค่า มีมิติ p x 1
$\mathbf{e}$ แทน เวกเตอร์ของค่าคลาดเคลื่อนสุ่มมาตรฐานมีมิติ n x 1
n แทน จำนวนค่าสังเกต
และ p แทน จำนวนตัวแปรอิสระ

จากสมการที่ (3) และสมการที่ (4) ตัวแปรเดิมทุกตัวถูกแปลงเป็นรูปมาตรฐาน ซึ่งกรณีนี้ค่าของตัวแปรจะวิ่งอยู่ในช่วง -3 ถึง +3 ด้วยความน่าจะเป็นไม่น้อยกว่า .99และสัมประสิทธิ์ β_j เปลี่ยนเป็น α_j โดยค่า α_j มีค่าอยู่ระหว่าง -1 ถึง +1 ด้วยประเด็นนี้ค่าประมาณสัมประสิทธิ์ของสมการที่ (3) และสมการที่ (4) จึงเป็นค่าที่เปรียบเทียบกันได้ ทำให้เราอ่านได้จากสมการถดถอยทันทีว่า ตัวแปรใดมีอิทธิพลต่อ Y มากกว่ากัน

ผลงานวิจัย

สุภาพร ศรีหามี, วสันต์ เดือนแจ้ง และกรองทิพย์ เนียมถนอม. (2553). การศึกษาระดับพฤติกรรมทางคุณธรรมจริยธรรมของนักเรียนโรงเรียนมัธยมสาธิตมหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา. *ครุศาสตร์สาร*, *4*(1), 148-165.

ชัยยา น้อยนารถ, กิตติศักดิ์ จังพานิช, อัญชลี ชุ่มบัวทอง, พัชรา เกรียงไกร, ณัฏฐ์พัชร์ วณิชย์กุล, นุชนาฏ เลี้ยงอำนวย, วสันต์ เดือนแจ้ง, กรองทิพย์ เนียมถนอม และพูลพงศ์ สุขสว่าง. โมเดลความสัมพันธ์เชิงสาเหตุการรับรู้โครงสร้างมหาวิทยาลัยและความผูกพันต่อองค์กรที่ส่งผลต่อความพึงพอใจในการปฏิบัติงานของอาจารย์ใน 3 จังหวัดชายแดนภาคใต้. *วารสารนานาชาติมหาวิทยาลัยขอนแก่น สาขามนุษยศาสตร์และสังคมศาสตร์*, *6*(3). (in press).

กรองทิพย์ เนียมถนอม, เสรี ชัดแช้ม และกิดาการ สายธนู. การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปประเทศมาเลเซียโดยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้นพหุคูณแบบใหม่. *ครุศาสตร์สาร*, *11*(1). (in press).