

การพยากรณ์อุปสงค์การส่งออกยางพาราของไทยไปมาเลเซีย
โดยวิธีการประมาณค่าพารามิเตอร์การถดถอยเชิงเส้น
พหุคูณแบบใหม่
Forecasting the Demand for Thai Rubber Export
to Malaysia Using New Parameter Estimation in
Multiple Linear Regression

กรองทิพย์ เนียมถนอม*

เสรี ชัดแฉ่ม**

* ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏพระนครศรีอยุธยา

** วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

*Corresponding author. Email : nong2513@gmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ และเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการวิเคราะห์การถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และ วิธีริดจ์ชนิดตัวประมาณจีเอ็ม เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมี ค่าผิดปกติ และการมีอัตราสหสัมพันธ์ของค่าคลาดเคลื่อน จำลองข้อมูลด้วยเทคนิคมอนติ คาร์โล ภายใต้สถานการณ์ 162 สถานการณ์ ใช้เกณฑ์ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของ การประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ และพยากรณ์ข้อมูลอุปสงค์การส่ง ออกยางพาราไทยไปประเทศมาเลเซีย สรุปผลการวิจัยได้ดังนี้ ตัวประมาณค่าสัมประสิทธิ์ การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่า ผิดปกติ และการมีอัตราสหสัมพันธ์ของค่าคลาดเคลื่อนคือ $\hat{b}_{NEW} = (Q'WQ + kI)^{-1}(Q'WR + kJ)$ ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นใช้ได้ดีกว่าวิธีอื่นภายใต้ สถานการณ์ 61 สถานการณ์ สมการพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราไทยไป ประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น คือ $z_y = 0.26453z_1 - 0.406258z_2 + 1.022741z_3$

คำสำคัญ: วิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป
วิธีการถดถอยรีดจ์ วิธีของนันทวัน จุลมุสิก วิธีรีดจ์ชนิดตัวประมาณจีเอ็ม
ยางพารา

Abstract

The purposes of this study were to developed and compare the estimator efficiency of linear multiple regressions coefficient among the methods of Ordinary Least Square, Generalized Least Square, Ridge Regressions, Julmusik's and Generalized M-estimator Ridge if the problems of multicollinearity, outliers and autocorrelation were occurred by using simulation of Monte Carlo technique under 162 situations using the mean square error of the coefficient estimates of the multiple linear regressions also to predict the Thai rubber exporting demand to Malaysia. The results of research were summarized as follows: The estimator of linear multiple regression in the case of problems of multicollinearity, outliers and autocorrelation were occurred as $\hat{b}_{NEW} = (Q'WQ + kI)^{-1}(Q'WR + kJ)$ The developed coefficient estimates of the multiple linear regressions were used better than those of other methods under 61 situations. The equation of Thai rubber exporting demand to Malaysia by using the developed estimator of linear multiple regressions coefficient was as follows: $z_y = 0.26453z_1 - 0.406258z_2 + 1.022741z_3$

Keywords: Ordinary Least Square Method, Generalized Least Square Method, Ridge Regression Method, Nantawan Julmusik Method, Generalized M-estimator Ridge Method, Rubber

บทนำ

ปัจจุบันวิธีการทางสถิติได้ถูกนำมาใช้ในการวิเคราะห์ข้อมูลกันอย่างแพร่หลาย การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการทางสถิติวิธีการหนึ่งที่ใช้ศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป เป้าหมายของการวิเคราะห์คือหาสมการทางคณิตศาสตร์ที่ใช้เป็นตัวแทนแสดงความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะประมาณค่าหรือพยากรณ์ค่าของตัวแปรหนึ่งจากค่าที่กำหนดให้ของตัวแปรอื่น ๆ ที่เกี่ยวข้องกัน ตัวแปรในตัวแบบการถดถอยจำแนกออกเป็น 2 ประเภทคือ ตัวแปรที่ต้องการพยากรณ์ เรียกว่าตัวแปรตาม (Dependent Variable) และตัวแปรที่มีอิทธิพลต่อตัวแปรตาม เรียกว่าตัวแปรอิสระ (Independent Variable) ถ้ามีตัวแปรอิสระเพียงตัวเดียว จะเรียกการวิเคราะห์การถดถอยเชิงเส้นเชิงเดียว (Simple Linear Regression Analysis) และถ้ามีตัวแปรอิสระหลายตัว จะเรียกว่าการวิเคราะห์

การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis) เขียนความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระได้ในรูปของเมทริกซ์ดังนี้ $Y = X\beta + \varepsilon$ เมื่อ Y แทนเวกเตอร์ของค่าสังเกตของตัวแปรตามที่มีมิติ $n \times 1$ X แทนเมทริกซ์ของค่าสังเกตของตัวแปรอิสระที่มีมิติ $n \times (p+1)$ β แทนเวกเตอร์ของสัมประสิทธิ์การถดถอยที่ไม่ทราบค่า มีมิติ $(p+1) \times 1$ ε แทนเวกเตอร์ของค่าคลาดเคลื่อนสุ่ม มีมิติ $n \times 1$ n แทนจำนวนค่าสังเกต และ p แทนจำนวนตัวแปรอิสระเนื่องจาก $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ เป็นพารามิเตอร์ของตัวแบบการถดถอยที่ไม่ทราบค่า หากสามารถประมาณค่าพารามิเตอร์เหล่านี้ได้ใกล้เคียงกับค่าจริงแล้ว จะทำให้การพยากรณ์ถูกต้องมากยิ่งขึ้น

ขั้นตอนในการวิเคราะห์การถดถอยมีหลายขั้นตอน ซึ่งขั้นตอนที่สำคัญคือการคัดเลือกตัวแปรอิสระที่มีอิทธิพลต่อการอธิบายความแปรผันของตัวแปรตาม และการประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอย การประมาณค่าพารามิเตอร์ของสัมประสิทธิ์การถดถอยมีด้วยกันหลายวิธี วิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่นิยมใช้กันมากที่สุดคือ วิธีการกำลังสองน้อยที่สุดสามัญ (Ordinary Least Square Method: OLS) ซึ่งจะมีตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณอยู่ในรูปของ $\hat{\beta} = (X'X)^{-1} X'Y$ โดยจะเป็นวิธีที่ทำให้ได้ตัวประมาณที่มีคุณสมบัติไม่เอนเอียง และมีความแปรปรวนต่ำที่สุด (Minimum Variance Unbiased Estimator: MVUE) และจะเรียกตัวประมาณที่มีคุณสมบัติดังกล่าวนี้ว่าตัวประมาณค่าไม่เอนเอียงเชิงเส้นที่ดีที่สุด (Best Linear Unbiased Estimator: BLUE) ภายใต้ข้อสมมุติคือและตัวแปรอิสระแต่ละตัวมีความเป็นอิสระกัน $E(\varepsilon) = 0, Cov(\varepsilon) = E(\varepsilon\varepsilon') = \sigma^2 I$

ในบางครั้งการวิเคราะห์การถดถอยเชิงเส้นพหุคูณตัวแปรอิสระที่นำมาศึกษาอาจมีความสัมพันธ์กันเกิดขึ้น ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรหรือตัวแปรอิสระมีความสัมพันธ์กัน เป็นปัญหาที่อาจพบได้โดยทั่วไป โดยเฉพาะข้อมูลทางด้านเศรษฐศาสตร์ ทำให้เมทริกซ์ความแปรปรวนและความแปรปรวนร่วมของตัวประมาณค่าสัมประสิทธิ์การถดถอย (Variance Covariance Matrix of the regression Coefficient) มีค่าสูงซึ่งจะส่งผลให้การทดสอบสมมุติฐานเกี่ยวกับค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ได้ขาดความแม่นยำ ผลกระทบที่เกิดขึ้นคือจะทำให้ค่าประมาณสัมประสิทธิ์การถดถอยมีค่าแตกต่างไปจากความเป็นจริงและยังทำให้มีเครื่องหมายไม่ตรงกับข้อเท็จจริง (ทรงศิริ, 2548: 182) มีวิธีการแก้ปัญหาโดยการตัดตัวแปรอิสระบางตัวที่มีความสัมพันธ์กันสูงกับตัวแปรอิสระตัวอื่นออกไปจากตัวแบบเชิงคณิตศาสตร์ โดยพิจารณาจากตัวแปรอิสระที่มีความสัมพันธ์กัน และตัดตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามน้อยกว่าออกไปกับวิธีการแก้ปัญหาที่ไม่ต้องการตัดตัวแปรอิสระตัวใดออกจากตัวแบบ เนื่องจากตัวแปรอิสระทุกตัวมีความสำคัญต่อตัวแปรตามเท่ากัน หรือความสัมพันธ์ระหว่างตัวแปรอิสระเกิดขึ้นไม่ชัดเจนพอ โดยได้มีผู้ศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณเมื่อเกิดปัญหาพหุสัมพันธ์ เช่น ในปี ค.ศ. 1970 Hoerl and Kennard ได้เสนอวิธีการแก้ปัญหาตัวแปรอิสระมีความสัมพันธ์กัน โดยไม่ต้องตัดตัวแปรอิสระทิ้ง เป็นการศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ให้ค่าเฉลี่ย

ของค่าคลาดเคลื่อนกำลังสองน้อยกว่าวิธีกำลังสองน้อยที่สุด เรียกว่าวิธีการถดถอยริดจ์ (Ridge Regression Method) โดยหลักการของวิธีนี้จะเป็นการนำค่าคงที่ k ที่มีค่าอยู่ระหว่าง 0 และ 1 มาบวกกับสมาชิกในแนวทแยงมุมของเมทริกซ์ $X'X$ เพื่อลดค่าคลาดเคลื่อนกำลังสองเฉลี่ยจากวิธีกำลังสองน้อยที่สุดให้มีความน้อยลง

การเก็บรวบรวมข้อมูลในบางครั้งอาจพบปัญหาข้อมูลที่เกิดขึ้นมีค่าผิดปกติเกิดขึ้นคือมีค่าสังเกตบางค่าสูงหรือต่ำกว่าค่าสังเกตส่วนใหญ่ หากพบว่าค่าผิดปกติเหล่านั้นเกิดจากธรรมชาติของข้อมูล จะต้องนำค่าผิดปกติเหล่านั้นมารวมในการวิเคราะห์การถดถอยด้วย (ทรงศิริ, 2548 หน้า 211) เนื่องจากข้อมูลที่มีค่าผิดปกติเหล่านั้นจะมีผลต่อการประมาณค่าพารามิเตอร์ของวิธีกำลังสองน้อยที่สุดสามัญ สำหรับการแก้ปัญหาที่มีนักสถิติหลายคนเสนอวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณที่สามารถช่วยลดอิทธิพลของข้อมูลที่มีค่าผิดปกติ ด้วยการให้น้ำหนักของค่าสังเกตที่ผิดปกติ้น้อยกว่าค่าสังเกตที่เป็นข้อมูลส่วนใหญ่ เรียกว่าการวิเคราะห์การถดถอยที่มีความแกร่ง (Robust Regression) โดยการวิเคราะห์การถดถอยที่มีความแกร่ง เป็นการวิเคราะห์การถดถอยที่ใช้หลักของการลดอิทธิพลของข้อมูลที่ผิดปกติลง โดยการสร้างสมการถดถอยสำหรับข้อมูลส่วนใหญ่ แล้วตรวจสอบข้อมูลที่ผิดปกติจากสมการถดถอยที่สร้างเป็นเกณฑ์ มีนักสถิติหลายคนได้ศึกษาและพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่มีความแกร่งหลายวิธี เช่น วิธีตัวประมาณค่าเอ็ม วิธีค่าสัมบูรณ์น้อยที่สุด วิธีตัวประมาณค่าเอ็มเอ็ม วิธีค่ามัธยฐานกำลังสองน้อยที่สุด วิธีตัวประมาณค่าแอลทีเอส และวิธีตัวประมาณค่าจีเอ็ม

การประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดสามัญ ยังมีข้อสมมุติที่สำคัญอีกประเด็นหนึ่งคือค่าคลาดเคลื่อนเป็นอิสระกัน ซึ่งในทางปฏิบัติข้อมูลทางด้านเศรษฐศาสตร์ที่เก็บรวบรวมมาจะเป็นข้อมูลอนุกรมเวลา ปัญหาที่พบบ่อยคือ ค่าคลาดเคลื่อนมีความสัมพันธ์กัน หรือเรียกว่ามีอัตสหสัมพันธ์ (Autocorrelation) จะส่งผลให้การประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดสามัญไม่มีประสิทธิภาพและส่งผลให้ตัวประมาณไม่มีคุณสมบัติตัวประมาณค่าไม่เอนเอียงเชิงเส้นที่ดีที่สุด วิธีการแก้ปัญหาเมื่อเกิดอัตสหสัมพันธ์ เช่น การเพิ่มตัวแปรอิสระ การเปลี่ยนรูปแบบความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระ หรือใช้การแปลงข้อมูล การแปลงแบบจำลองที่เกิดปัญหาไปเป็นแบบจำลองที่ไม่มีอัตสหสัมพันธ์ โดยการประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุดทั่วไป แทนวิธีกำลังสองน้อยที่สุดสามัญ ซึ่งมีข้อสมมุติดังนี้ $E(\varepsilon) = 0$ และ $Cov(\varepsilon) = E(\varepsilon\varepsilon') = \sigma^2 U$ โดยที่ U แทนเมทริกซ์อัตสหสัมพันธ์ของ ε มิติ $n \times n$ ที่มี $rank = n$ และเรียกข้อสมมุตินี้ว่าข้อสมมุติของวิธีกำลังสองน้อยที่สุดทั่วไป

จากปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และปัญหาอัตสหสัมพันธ์ล้วนเป็นปัญหาที่สำคัญทั้งสิ้น เมื่อต้องทำการพยากรณ์โดยใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ เนื่องจากปัญหาดังกล่าวจะส่งผลให้การประมาณค่าสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดไม่มีประสิทธิภาพ และส่งผลให้ตัวประมาณไม่มีคุณสมบัติตัวประมาณค่าไม่เอนเอียงเชิงเส้นที่ดีที่สุด โดยเฉพาะ

การพยากรณ์ข้อมูลทางเศรษฐศาสตร์นั้น ตัวแปรอิสระที่ใช้ในการทำนายตัวแปรตามจะมีหลายตัว จึงอาจจะมีความสัมพันธ์กันเกิดขึ้นได้ ประกอบกับข้อมูลทางเศรษฐศาสตร์เป็นข้อมูลที่จัดเก็บตามช่วงเวลาซึ่งเรียกว่าข้อมูลอนุกรมเวลา ทำให้ค่าสังเกตในอนุกรมเวลามีความสัมพันธ์กันได้ และค่าสังเกตบางค่าอาจมีค่าสูงหรือต่ำกว่าค่าสังเกตส่วนใหญ่ในการพยากรณ์ข้อมูลทางเศรษฐศาสตร์จึงจำเป็นต้องตระหนักถึงวิธีการประมาณค่าที่เหมาะสมกับข้อมูลเป็นอย่างมาก ซึ่งในงานวิจัยที่ผ่านมาได้มีการศึกษาเพื่อหาวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่สามารถแก้ปัญหาได้อย่างใดอย่างหนึ่งในกรณีเกิดอัตตสหสัมพันธ์ของความคลาดเคลื่อน ปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ ส่วนในกรณีมีปัญหาหว่ามกันมีผู้ศึกษากรณีเกิดปัญหาอัตตสหสัมพันธ์ของความคลาดเคลื่อนร่วมกับมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีพหุสัมพันธ์ระหว่างตัวแปรอิสระร่วมกับข้อมูลมีค่าผิดปกติเท่านั้น ยังไม่มีผู้ที่ศึกษากรณีเกิดปัญหาทั้งสามพร้อม ๆ กัน ในการศึกษาในผู้วิจัยต้องการจะพัฒนาวิธีการประมาณค่าสัมประสิทธิ์การถดถอย เพื่อแก้ปัญหาพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ และปัญหาอัตตสหสัมพันธ์ไปพร้อม ๆ กัน ซึ่งจะนำไปสู่วิธีการประมาณค่าพารามิเตอร์แบบใหม่ที่สามารถลดขั้นตอนในการประมาณค่าสัมประสิทธิ์การถดถอย และได้ค่าประมาณสัมประสิทธิ์การถดถอยที่มีประสิทธิภาพ

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ในกรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ข้อมูลมีค่าผิดปกติ และการมีอัตตสหสัมพันธ์ของความคลาดเคลื่อน
2. เพื่อเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ ที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุด วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการวิเคราะห์การถดถอยรีดจ์ วิธีของนันทวัน จุลมุลิก และวิธีรีดจ์ชนิดตัวประมาณจีเอ็ม ภายใต้สถานการณ์ 162 สถานการณ์ ได้แก่ 1) ขนาดตัวอย่าง 3 ขนาด คือ ขนาดตัวอย่างเท่ากับ 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 (Monte Carlo Technique)
3. เพื่อพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น

วิธีการดำเนินงาน

การวิจัยครั้งนี้ผู้วิจัยแบ่งขั้นตอนการดำเนินงานออกเป็น 3 ขั้นตอนดังนี้

ตอนที่ 1 การพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณ เมื่อเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตราสหสัมพันธ์ของค่าคลาดเคลื่อน ผู้วิจัยมีวิธีการดำเนินการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง พบว่ากรณีข้อมูลเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ปี ค.ศ. 1970 Hoerl and Kennard ได้เสนอวิธีการถดถอยริดจ์ในการแก้ปัญหาตัวแปรอิสระมีความสัมพันธ์เชิงเส้นต่อกัน วิธีนี้จะให้ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของการประมาณ β ลดลง เรียกว่าวิธีการถดถอยริดจ์ มีหลักการคือถ้าตัวแปรอิสระมีพหุสัมพันธ์เชิงเส้นกันแล้ว ค่าดีเทอร์มิแนนต์ (Determinant) ของ $X'X$ จะมีค่าเข้าใกล้ศูนย์จึงมีผลทำให้ความแปรปรวนของตัวประมาณของสัมประสิทธิ์การถดถอยด้วยวิธีกำลังสองน้อยที่สุดมีค่าสูงผิดปกติด้วย ดังนั้นจะทำให้ $|X'X|$ เพิ่มขึ้นได้โดยการบวกค่าคงที่ k ที่มากกว่า 0 เข้ากับสมาชิกทุกตัวบนเส้นทแยงมุมของเมทริกซ์ ($X'X$) ซึ่งจะทำให้ค่าเจาะจงมีค่ามากขึ้นและส่งผลให้ผลบวกกำลังสองของ β มีค่าลดลง ทำให้ตัวประมาณของสัมประสิทธิ์การถดถอยถูกต้องมากขึ้น ตัวประมาณค่า β ซึ่งเป็นตัวประมาณที่เอนเอียง คือ $\hat{\beta}_{RR} = (X'X + kI)^{-1} X'Y$; $k > 0$ ในปี ค.ศ. 1976 Swindel ได้เสนอเทคนิคการนำสารสนเทศเบื้องต้น (Prior Information) มาใช้ร่วมกับการถดถอยริดจ์ ทำให้ได้ตัวประมาณของสัมประสิทธิ์การถดถอย ดังนี้ $\hat{\beta}_{RJ} = (X'X + kI)^{-1} (X'Y + kJ)$ และเป็นตัวประมาณค่าสัมประสิทธิ์การถดถอยมีคุณสมบัติไม่เอนเอียง กรณีเกิดปัญหาข้อมูลมีค่าผิดปกติ ในปี ค.ศ. 1964 Huber ได้ศึกษาถึงตัวประมาณที่มีความแกร่ง (Robust Estimator) ที่เรียกว่าวิธีตัวประมาณเฮอร์เบิร์ตเป็นครั้งแรก ต่อมาในปี ค.ศ. 1996 Simpson and Montgomery ได้พัฒนาวิธีตัวประมาณเฮอร์เบิร์ตเป็นวิธีตัวประมาณจีเอ็ม (Generalized M-estimator Method) โดยตัวประมาณจีเอ็มมาจากหลักการพื้นฐานของการประมาณแบบความควรจะเป็นสูงสุด (Maximum Likelihood Estimation) นั่นคือหาตัวประมาณ $\hat{\beta}$ ที่ทำให้ฟังก์ชันความควรจะเป็นของค่าคลาดเคลื่อนมีค่าสูงสุด ได้สูตรการประมาณค่า $\hat{\beta}$ จากวิธี WLS อยู่ในรูปของ $\hat{\beta}_{GM} = (X'WX)^{-1} X'WY$ กรณีเกิดปัญหาอัตราสหสัมพันธ์ของค่าคลาดเคลื่อน แก้ปัญหาโดยการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยที่สุดทั่วไป ซึ่งอาศัยเทคนิคการแปลงข้อมูล (Transformation Technique) ภายใต้ข้อสมมุติของ GLS ตัวประมาณที่ไม่เอนเอียงของ $\hat{\beta}$ มีสูตรดังนี้

$\hat{\beta}_{GLS} = (X'U^{-1}X)^{-1} (X'U^{-1}Y)$ กรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระร่วมกับปัญหาอัตราสหสัมพันธ์ของค่าคลาดเคลื่อน ในปี พ.ศ. 2552 นันทวัน จุลมสิก ได้พัฒนาสูตรเพื่อประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณ เมื่อค่าคลาดเคลื่อนมีอัตราสหสัมพันธ์และมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ ซึ่งเป็นการแก้ปัญหาทั้งสองอย่างนี้ไปพร้อม ๆ กัน แล้วพัฒนาเป็นตัวประมาณตัวใหม่ได้ดังนี้

$\hat{\beta}_{NT} = (H'U^{-1}H + kI)^{-1} (H'U^{-1}N + kJ)$ กรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปร

อิสระร่วมกับปัญหาการมีข้อมูลที่มีค่าผิดปกติ ในปี ค.ศ. 1996 Simpson and Montgomery ได้พัฒนาวิธีตัวประมาณที่มีความแข็งแกร่งเรียกว่าวิธีตัวประมาณจีเอ็ม แล้วนำมารวมกับการประมาณโดยวิธีการถดถอยริดจ์ เรียกว่าวิธีดังกล่าวว่า วิธีริดจ์ชนิดตัวประมาณจีเอ็ม (Generalized M-estimator Ridge Method) ซึ่งรูปแบบการคำนวณค่าสัมประสิทธิ์การถดถอยโดยวิธีริดจ์ชนิดตัวประมาณจีเอ็ม คือ $\hat{\beta}_{GMR} = (X'W^{(r)}X + kI)^{-1} X'W^{(r)}Y; k > 0$

ตอนที่ 2 การเปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น และตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม โดยทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ภายใต้สถานการณ์ 162 สถานการณ์ ได้แก่ 1) ขนาดตัวอย่าง 3 ขนาด คือ ขนาดตัวอย่างเท่ากับ 20, 50 และ 100 2) ระดับการมีพหุสัมพันธ์ 3 ระดับ คือ ระดับต่ำ (0.3) ระดับกลาง (0.5) และระดับสูง (0.9) 3) ระดับการมี อัตราสหสัมพันธ์ 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% และใช้โปรแกรม MATLAB เวอร์ชัน 8.7 คำนวณหาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณของการถดถอยเชิงเส้นพหุคูณแบบใหม่ที่พัฒนาขึ้น และวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยริดจ์ วิธีของนันทวัน จุลมุสิก และวิธีริดจ์ชนิดตัวประมาณจีเอ็ม โดยใช้เกณฑ์ความแข็งแกร่งของ Ramsay ทำซ้ำในแต่ละสถานการณ์จำนวน 1,000 ครั้ง วิเคราะห์ค่าคลาดเคลื่อนกำลังสองเฉลี่ยของแต่ละวิธีและแต่ละสถานการณ์ เพื่อนำแต่ละวิธีและแต่ละสถานการณ์มาเปรียบเทียบกับ โดยวิธีที่ให้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ของการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่น้อยที่สุด จะเป็นวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่มีประสิทธิภาพมากที่สุด

ตอนที่ 3 การพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราไทยไปมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้น ในขั้นตอนนี้นำวิธีการประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุคูณแบบใหม่ มาประยุกต์ใช้กับข้อมูลทุติยภูมิเพื่อพยากรณ์อุปสงค์การส่งออกยางพาราไทยไปยังมาเลเซีย โดยข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลทุติยภูมิเกี่ยวกับยางพารา คือปริมาณการส่งออกยางพาราไทยไปมาเลเซีย ใช้ข้อมูลจากส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ปี พ.ศ. 2542 ถึง ปี พ.ศ. 2557 ผลิตภัณฑ์มวลรวมภายในประเทศของมาเลเซีย ราคาส่งออกยางพาราไทยใช้ข้อมูลจากสำนักงานคณะกรรมการเศรษฐกิจและสังคมภาคพื้นเอเชียและแปซิฟิก (Economic and Social Commission for Asia and Pacific) ปีพ.ศ. 2542 ถึง ปี พ.ศ. 2557

ผลการวิจัย

ตอนที่ 1 ผลการพัฒนาตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณกรณีมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และมีอัตราสหสัมพันธ์ของค่าคลาดเคลื่อนพัฒนาเป็นตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ได้ ดังนี้ $\hat{b}_{NEW} = (Q'WQ + kI)^{-1} (Q'WR + kJ)$ เมื่อ Q เป็นเมทริกซ์ของตัวแปรอิสระที่ใช้เทคนิคการแปลงรูป R เป็นเมทริกซ์ของตัวแปรตามที่ใช้เทคนิคการแปลงรูป W เป็นเมทริกซ์ทแยงมุมขนาด $n \times n$

ตอนที่ 2 ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าระหว่างตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นกับตัวประมาณค่าสัมประสิทธิ์การถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุดสามัญ วิธีกำลังสองน้อยที่สุดทั่วไป วิธีการถดถอยวิธีอื่น ๆ ของนันทวัน จุลมุลิก และวิธีวิธีจริงชนิดตัวประมาณจีเอ็ม ตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นใช้ได้ดีกว่าวิธีอื่นภายใต้สถานการณ์ 61 สถานการณ์ ใช้ได้ดังนี้

1. ข้อมูลมีค่าผิดปกติในตัวแปรตาม
 - 1.1 ขนาดตัวอย่างเท่ากับ 20 เมื่อข้อมูลมีค่าอัตราสหสัมพันธ์ 0.3, 0.5, 0.9 ก. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.3 ข้อมูลมีค่าผิดปกติ 1%, 5% ข. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.5 ข้อมูลมีค่าผิดปกติ 1%
 - 1.2 ขนาดตัวอย่างเท่ากับ 50, 100 เมื่อข้อมูลมีค่าอัตราสหสัมพันธ์ 0.3, 0.5, 0.9 ข้อมูลมีค่าผิดปกติ 1%, 5% ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.3, 0.5
 2. ในกรณีข้อมูลมีค่าผิดปกติในตัวแปรอิสระและตัวแปรตาม
 - 2.1 ขนาดตัวอย่างเท่ากับ 20 เมื่อข้อมูลมีค่าอัตราสหสัมพันธ์ 0.3, 0.5, 0.9 ข้อมูลมีค่าผิดปกติ 1%, 5% ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.3
 - 2.2 ขนาดตัวอย่างเท่ากับ 50 เมื่อข้อมูลมีค่าอัตราสหสัมพันธ์ 0.3, 0.5, 0.9 ก. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.3 ข้อมูลมีค่าผิดปกติ 1%, 5% ข. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.5 ข้อมูลมีค่าผิดปกติ 1%
 - 2.3 ขนาดตัวอย่างเท่ากับ 100 เมื่อข้อมูลมีค่าอัตราสหสัมพันธ์ 0.3, 0.5, 0.9 ก. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.3 ข้อมูลมีค่าผิดปกติ 1%, 5% ข. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.5 ข้อมูลมีค่าผิดปกติ 1% ค. ตัวแปรอิสระมีค่าความสัมพันธ์กับระดับ 0.9 ข้อมูลมีค่าผิดปกติ 5%
- และตัวประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นไม่เหมาะสมที่จะใช้ในกรณีข้อมูลมีค่าผิดปกติในตัวแปรอิสระ

ตอนที่ 3 ผลการพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราไทยไปมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นผลการพยากรณ์ข้อมูลอุปสงค์การส่งออกยางพาราไทยไปประเทศมาเลเซีย โดยวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณที่พัฒนาขึ้นคือ

สรุปผลการวิจัย

ผลการวิจัยที่ได้สอดคล้องกับ Huber (1964) ที่ได้เสนอวิธีตัวประมาณเอ็ม (M – estimator Method) เพื่อเป็นวิธีการประมาณค่าสัมประสิทธิ์การถดถอยสำหรับข้อมูลที่มีค่าผิดปกติในตัวแปรตาม หรือกรณีที่มีความคลาดเคลื่อนมีค่าผิดปกติ และ Simpson and Montgomery (1996) ได้พัฒนาตัวประมาณจีเอ็ม (Generalized M - Estimator Method: GM) เพื่อทำการสร้างตัวถ่วงน้ำหนักซึ่งจะพิจารณาทั้งค่าผิดปกติที่เกิดขึ้นในตัวแปรอิสระและตัวแปรตาม

ข้อเสนอแนะ

ข้อเสนอแนะในการนำไปใช้ กรณีเกิดปัญหาการมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ การมีข้อมูลที่มีค่าผิดปกติ และการมีอัตราสหสัมพันธ์ของค่าคลาดเคลื่อนไปพร้อมกัน ถ้าข้อมูลมีค่าผิดปกติที่ตัวแปรอิสระเพียงอย่างเดียวควรใช้วิธีของนันทวัน จุลมุลิกกรณีขนาดตัวอย่าง 20 และ 50 ถ้าขนาดตัวอย่าง 100 ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 และ 0.9 และวิธีการถดถอยริดจ์เมื่อขนาดตัวอย่าง 100 ควรใช้เมื่อตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 ข้อมูลมีค่าผิดปกติที่ตัวแปรตามเพียงอย่างเดียวควรใช้วิธีที่พัฒนาขึ้นกรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 และ 0.5 ใช้วิธีริดจ์ชนิดตัวประมาณจีเอ็มกรณี ตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.9 ข้อมูลมีค่าผิดปกติที่ตัวแปรอิสระและตัวแปรตามควรใช้วิธีที่พัฒนาขึ้นกรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.3 ใช้วิธีริดจ์ชนิดตัวประมาณจีเอ็มกรณีตัวแปรอิสระมีค่าความสัมพันธ์กันที่ 0.5 และ 0.9

ข้อเสนอแนะในการศึกษาต่อไป

การศึกษาครั้งนี้วิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณแบบใหม่ ที่พัฒนาขึ้น ผู้วิจัยได้ใช้เกณฑ์ความแกร่งของ Ramsay และประมาณ k ของ Hoerl, Kenard และ Baldwin เท่านั้น จึงควรมีการเปรียบเทียบกับเกณฑ์ความแกร่งและวิธีการประมาณค่า k อื่น ๆ และควรมีการเก็บรวบรวมข้อมูลยาวพาราได้มากกว่านี้ เนื่องจากในการวิจัยครั้งนี้เกิดข้อจำกัดจากหน่วยงานที่เก็บรวบรวมข้อมูลได้เก็บย้อนหลังแค่ปี พ.ศ. 2541 จึงไม่เกิดปัญหาการมีข้อมูลที่มีค่าผิดปกติ และการมีอัตราสหสัมพันธ์ของค่าคลาดเคลื่อนขึ้น

เอกสารอ้างอิง

- จิราวัลย์ จิตรถเวช. (2558). การวิเคราะห์การถดถอย. กรุงเทพฯ: โครงการส่งเสริมและพัฒนาเอกสารวิชาการสถาบันบัณฑิตพัฒนบริหารศาสตร์.
- ณัฐพร ภักดี. (2552). การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณ เมื่อตัวแปรอิสระมีพหุสัมพันธ์กันโดยวิธีการถดถอยริดจ์. วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- ทรงศิริ แด่สมบัติ. (2548). การวิเคราะห์การถดถอย (พิมพ์ครั้งที่ 3). กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.

- ัญญารัตน์ ไชยเนตรไกรสิน. (2551). ปัจจัยที่มีผลกระทบต่อและพยากรณ์ปริมาณการส่งออกยางพาราของประเทศไทยไปประเทศญี่ปุ่น. วิทยานิพนธ์เศรษฐศาสตรมหาบัณฑิต, เศรษฐศาสตร์, มหาวิทยาลัยรามคำแหง.
- นันทวัน จุลมุสิก. (2552). การประมาณค่าพารามิเตอร์ในการถดถอยเชิงเส้นพหุ เมื่อความคลาดเคลื่อนมีอัตตสหสัมพันธ์และมีพหุสัมพันธ์ระหว่างตัวแปรอิสระ. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะวิทยาศาสตร์, มหาวิทยาลัยบูรพา.
- เบญจวรรณ ระหงษ์. (2554). การเปรียบเทียบการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุกรณีเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์และเทคโนโลยี, มหาวิทยาลัยธรรมศาสตร์.
- มนตรี พิริยะกุล. (2545). เทคนิคการวิเคราะห์สมการถดถอย (ฉบับปรับปรุง). กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยรามคำแหง.
- ราชบัณฑิตยสถาน. (2553). พจนานุกรมศัพท์คณิตศาสตร์ ฉบับราชบัณฑิตยสถาน. 10. กรุงเทพฯ: นานามีบุ๊คส์พับลิเคชั่นส์.
- วรุณี รวยดี. (2552). การเปรียบเทียบวิธีการถ่วงน้ำหนักเฉพาะที่ในการวิเคราะห์การถดถอยสำหรับข้อมูลที่มีความสัมพันธ์กันและมีค่าผิดปกติ. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติ, คณะวิทยาศาสตร์, มหาวิทยาลัยเกษตรศาสตร์.
- วิธิตา พลาศรี. (2552). การประมาณค่าสัมประสิทธิ์ความถดถอยเชิงพหุสัมพันธ์ด้วยวิธีความถดถอยแบบตลบตรงแบบบริจ. วิทยานิพนธ์สถิติศาสตรมหาบัณฑิต, สาขาสถิติ, จุฬาลงกรณ์มหาวิทยาลัย.
- สถาบันวิจัยยาง. (2555). ข้อมูลวิชาการยางพารา 2555. กรุงเทพฯ: กรมวิชาการเกษตร กระทรวงเกษตรและสหกรณ์.
- สถาบันวิจัยยาง. (2555). สถิติยางประเทศไทย THAILAND RUBBER STATISTICS. กรุงเทพฯ: กรมวิชาการเกษตร กระทรวงเกษตรและสหกรณ์. ปีที่ 41 (2555) ฉบับที่ 4.
- อดุลย์เดช กรทอง. (2550). การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุด้วยวิธีดิฟเฟอเรนเชียลตัวประมาณจีเอ็ม. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต, สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, มหาวิทยาลัยนเรศวร.
- อรพรรณ ดันตระกูล. (2555). การเปรียบเทียบวิธีการประมาณค่าสัมประสิทธิ์การถดถอยที่แข็งแกร่งสำหรับการถดถอยเชิงเส้นพหุ เมื่อข้อมูลมีค่าผิดปกติ. 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.
- 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.
- 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.
- Huber, P. J. (1964). Robust Estimation of a Location Parameter. **Annals of Mathematical Statistic** 35 (1), 73-101.