



การคัดเลือกตัวแปรในแบบการถดถอยเชิงเส้นพหุ
โดยใช้วิธีการค้นหาแบบต้องห้าม

A Variable Selection in Multiple Linear Regression Models
Based on Tabu Search

กานต์ณัฐ ฦ บางช้าง^{1*} และ จิราวัลย์ จิตรถเวช¹

บทคัดย่อ

งานวิจัยนี้ได้เสนอวิธีการคัดเลือกตัวแปรในการวิเคราะห์การถดถอยตัวแบบเชิงเส้นพหุโดยใช้วิธีการค้นหาแบบต้องห้ามที่ใช้ฟังก์ชันเป้าหมายเป็นค่าความคลาดเคลื่อนกำลังสองเฉลี่ยและค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ แล้วนำผลมาเปรียบเทียบกับเทคนิคการคัดเลือกตัวแปรโดยวิธีการถดถอยแบบขั้นตอน เกณฑ์ที่ใช้ในการเปรียบเทียบ คือ ร้อยละของจำนวนครั้งที่คัดเลือกตัวแปรได้ถูกต้อง ในการศึกษาได้เปรียบเทียบการคัดเลือกตัวแปรอิสระเข้าสู่ตัวแบบในกรณีที่ไม่มีและมีปัญหาสหสัมพันธ์เชิงเส้นพหุ โดยใช้วิธีการจำลองข้อมูล จำนวนตัวแปรอิสระในตัวแบบเต็มรูปเท่ากับ 6 ตัวแปร และจำนวนตัวแปรอิสระที่ใช้ในการหาค่าตัวแปรตามเท่ากับ 4 ตัวแปร ขนาดตัวอย่างเท่ากับ 25 และ 100 กระทำซ้ำ 500 ครั้งในแต่ละสถานการณ์ กรณีที่ข้อมูลไม่มีปัญหาสหสัมพันธ์เชิงเส้นพหุ ความสามารถในการคัดเลือกตัวแปรของวิธีการถดถอยแบบขั้นตอนและการค้นหาแบบต้องห้ามที่ใช้ฟังก์ชันเป้าหมายทั้งสองมีร้อยละของความถูกต้องใกล้เคียงกัน แต่กรณีที่มีปัญหาสหสัมพันธ์เชิงเส้นพหุ วิธีการค้นหาแบบต้องห้ามที่มีฟังก์ชันเป้าหมายทั้งสองมีร้อยละของตัวแปรที่ถูกต้องสูงกว่าวิธีการถดถอยแบบขั้นตอน โดยวิธีการถดถอยแบบขั้นตอนจะให้ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระมีเครื่องหมายผิดไปจากที่กำหนดในการจำลอง

¹ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ คลองจั่น บางกะปิ กรุงเทพฯ 10240

*Corresponding Author; E-mail: kn_matu@msu.ac.th

ABSTRACT

This research has proposed a variable selection method based on the Tabu Search for multiple linear regression models. In this study two objective functions used in the Tabu Search are mean square error (MSE) and the mean absolute error (MAE). The results of Tabu Search are compared with the results obtained by the stepwise regression method based on the hit percentage criterion. The simulations cover both cases, without and with multicollinearity problems. There are six independent variables in full models with four independent variables that influence the dependent variable. The sample sizes are 25 and 100, repeated 500 times for each situation. Without the multicollinearity problem, the hit percentages of the stepwise regression method and Tabu Search using both objective functions are almost the same. With the multicollinearity problem, the hit percentages of Tabu Search using both objective functions are more than the hit percentage of the stepwise regression method. In particular, the regression coefficients have the wrong sign from simulation by the stepwise regression method.

บทนำ

สถิตินับว่าเป็นเครื่องมือสำคัญในการทำวิจัย สถิติจำแนกเป็น 2 ส่วน คือ สถิติเชิงพรรณนา (descriptive statistics) และสถิติเชิงอนุมาน (inferential statistics) โดยสถิติเชิงอนุมาน เป็นสถิติที่เกี่ยวข้องกับการวิเคราะห์ข้อมูลที่เก็บรวบรวมได้เพื่อสรุปผลเกี่ยวกับลักษณะเฉพาะ (characteristic) ของประชากร (ประชุม, 2553) ประกอบด้วย 2 ส่วนหลัก คือ การสรุปลักษณะประชากร ซึ่งอาศัยการทดสอบสมมติฐานทางสถิติ เพื่อนำไปสรุปถึงลักษณะของประชากร และการหาความสัมพันธ์ของตัวแปร สถิติที่ใช้ในการหาความสัมพันธ์มีอยู่หลายเทคนิค ซึ่งการเลือกใช้แต่ละเทคนิคทางสถิติขึ้นอยู่กับวัตถุประสงค์ของงานวิจัย การวิเคราะห์การถดถอย (regression analysis) เป็นวิธีการวิเคราะห์ทางสถิติที่เกี่ยวข้องกับการสร้างตัวแบบทางคณิตศาสตร์ เพื่อแสดงความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระจำนวน 1 ตัว หรือมากกว่า 1 ตัว ก็ได้ (วิชิต และจิราวัลย์, 2548) ซึ่งรูปแบบที่นิยมใช้กันอย่างแพร่หลาย คือ การวิเคราะห์การถดถอยเชิงเส้นพหุ (multiple linear regression analysis) โดยเป็นตัวแบบที่ความสัมพันธ์ของตัวแปรตามขึ้นอยู่กับตัวแปรอิสระมากกว่า 1 ตัวแปร ซึ่งทั้งตัวแปรอิสระและตัวแปรตามมีความสัมพันธ์ในรูปเชิงเส้นโดยมีวัตถุประสงค์ เพื่อการพยากรณ์ค่าของตัวแปรตามจากตัวแปรอิสระ ตลอดจนการอนุมานต่าง ๆ เกี่ยวกับตัวแปรตาม

การใช้ตัวแบบการถดถอยเชิงเส้นในการพยากรณ์ให้มีประสิทธิภาพขึ้นอยู่กับทางเลือกตัวแบบที่มีความเหมาะสม การได้ตัวแบบที่มีความเหมาะสมควรมาจากการคัดเลือกตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรตาม จำนวนที่ไม่มากหรือน้อยเกินไปให้อยู่ในสมการพยากรณ์ เนื่องจากตัวแบบการถดถอยที่มีตัวแปรอิสระมากเกินไปจนมีความจำเป็นทำให้ค่าพยากรณ์ที่ได้มีความคลาดเคลื่อนสูงและอาจเกิดปัญหาความสัมพันธ์เชิงเส้นพหุ (multicollinearity) ในทางตรงกันข้ามถ้าตัวแบบที่สร้างขึ้น ขาดตัวแปรอิสระที่สำคัญไป ทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูงได้เช่นกัน

ในปัจจุบันมีเทคนิคในการคัดเลือกตัวแปรอยู่หลายเทคนิคด้วยกัน ที่นิยมใช้กันอย่างแพร่หลาย คือ วิธีพิจารณาทุกตัวแบบที่เป็นไปได้ (all possible regression) วิธีเลือกตัวแปรอิสระแบบไปข้างหน้า (forward selection) วิธีตัดตัวแปรอิสระออกแบบถอยหลัง (backward elimination) และ วิธีการถดถอยแบบขั้นตอน (stepwise regression)

นอกจากเทคนิคการคัดเลือกตัวแปรที่กล่าวมาข้างต้นยังมีเทคนิคด้านการหาค่าเหมาะที่สุดโดยการจัดกลุ่ม (combinatorial) ที่สามารถใช้คัดเลือกตัวแปรอิสระที่เหมาะสมได้เช่นกัน เช่น วิธีการค้นหาแบบต้องห้าม เป็นกระบวนการสำหรับแก้ปัญหาเพื่อหาค่าเหมาะสมที่สุด โดยใช้ข้อจำกัดและเวลาในการประมวลผลให้น้อยที่สุด (Glover, 1990) ซึ่งเทคนิคนี้ใช้การกำหนดเกณฑ์ในการหยุดกระบวนการทำงาน (algorithm) คือ จำนวนรอบสูงสุดและการลดลงของค่าฟังก์ชันเป้าหมาย (objective function) ดังนั้นจึงสามารถนำแนวคิดของเทคนิคนี้มาช่วยในการคัดเลือกตัวแปรได้เช่นกัน โดยเป็นเทคนิคที่มีเป้าหมายในการคัดเลือกตัวแปรเข้าสู่ตัวแบบให้มีจำนวนน้อยที่สุดและใช้เวลาน้อยในการคัดเลือกตัวแปร และมีความคลาดเคลื่อนในการพยากรณ์ต่ำ หรือกล่าวได้ว่าเป็นเทคนิคที่ช่วยในการป้องกันการเกิดปัญหาของตัวแบบที่มีตัวแปรอิสระมากเกินไป (over specification) คือ ได้ตัวแบบที่ประกอบด้วยตัวแปรอิสระมากกว่าตัวแบบจริง นอกจากนี้ยังช่วยลดค่าใช้จ่ายและเวลาในการเก็บรวบรวมข้อมูลเกี่ยวกับตัวแปรอิสระที่จะนำมาใช้ในการพยากรณ์ตัวแปรตามในตัวแบบการถดถอยเชิงเส้นพหุ

ในกรณีที่ตัวแปรอิสระไม่มีสหสัมพันธ์เชิงเส้นพหุ การประมาณค่าโดยใช้วิธีกำลังสองน้อยที่สุด (ordinary least squares: OLS) สามารถประมาณค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระได้เป็นอย่างดี แต่ในกรณีที่ตัวแปรอิสระมีสหสัมพันธ์เชิงเส้นพหุวิธีการประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยที่สุด จะประมาณค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระผิดพลาดไปจากค่าจริง ซึ่งวิธีการแก้ไขที่นิยมใช้ คือ วิธีการถดถอยแบบบริดจ์ (Ridge regression) ที่ต้องใช้ดุลยพินิจในการพิจารณาเส้นทางของบริดจ์ (Ridge trace) ดังนั้นผู้วิจัยจึงนำแนวคิดของวิธีการค้นหาแบบต้องห้าม (tabu search) มาใช้ในการคัดเลือกตัวแปรอิสระในกรณีที่ตัวแปรอิสระมีสหสัมพันธ์เชิงเส้นพหุเช่นกัน

ทฤษฎีที่เกี่ยวข้อง

ตัวแบบการถดถอยเชิงเส้นพหุ (multiple linear regression model) มีตัวแบบทั่วไป ดังนี้

$$y = X\beta + \varepsilon$$

โดยที่ y คือ เวกเตอร์ของตัวแปรตาม y ขนาด $n \times 1$

X คือ เมทริกซ์ของตัวแปรอิสระขนาด $n \times (k+1)$

β คือ เวกเตอร์ของพารามิเตอร์ของตัวแบบขนาด $(k+1) \times 1$

ε คือ เวกเตอร์ของความคลาดเคลื่อนสุ่มที่เป็นอิสระต่อกันขนาด $n \times 1$ ภายใต้ข้อกำหนด

$$\varepsilon \sim N_n(0, \sigma^2 I_n)$$

เมื่อ \mathbf{I}_n คือ เมทริกซ์เอกลักษณ์ขนาด $n \times n$ โดย n คือ ขนาดตัวอย่าง และ k คือ จำนวนตัวแปรอิสระในตัวแบบ

ฐานคติของการวิเคราะห์การถดถอยเชิงเส้นพหุมีดังนี้

1. ความคลาดเคลื่อนเป็นตัวแปรสุ่มมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ σ^2 ความคลาดเคลื่อน ε_i และ ε_j เป็นอิสระต่อกัน เมื่อ $i \neq j$ นั่นคือ $\varepsilon_i \sim \text{i.i.d } N(0, \sigma^2)$

2. ตัวแบบการถดถอยเป็นตัวแบบเชิงเส้นของพารามิเตอร์

3. ตัวแปรอิสระแต่ละตัวเป็นตัวแปรเชิงคณิตศาสตร์

โดยในการวิจัยครั้งนี้ตัวประมาณที่ใช้ คือ ตัวประมาณภาวะน่าจะเป็นสูงสุด จากตัวแบบ คือ

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ จะได้ฟังก์ชันภาวะน่าจะเป็น คือ } L(\boldsymbol{\beta}, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\right)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\}\right)$$

$$\text{จะได้ } \ln L(\boldsymbol{\beta}, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

$$\text{พิจารณา } \frac{\partial}{\partial \boldsymbol{\beta}} \ln L(\boldsymbol{\beta}, \sigma^2) = 0 \text{ และ } \frac{\partial}{\partial \sigma^2} \ln L(\boldsymbol{\beta}, \sigma^2) = 0$$

$$\text{ดังนั้นจะได้ } \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \text{ และ } \hat{\sigma}^2 = \frac{1}{n} (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

หากตัวแปรอิสระเกิดปัญหาสหสัมพันธ์เชิงเส้นพหุอย่างสมบูรณ์จะส่งผลให้ $\mathbf{X}'\mathbf{X}$ เป็นเมทริกซ์เอกฐาน (singular matrix) กล่าวคือ $|\mathbf{X}'\mathbf{X}|$ มีค่าเท่ากับศูนย์ ทำให้ไม่สามารถหาค่า $(\mathbf{X}'\mathbf{X})^{-1}$ ได้ ในกรณีที่ศึกษาเมทริกซ์ $\mathbf{X}'\mathbf{X}$ มีค่าเข้าใกล้เมทริกซ์เอกฐาน หรืออาจมีลักษณะของ near singularity คือ $|\mathbf{X}'\mathbf{X}|$ มีค่าเข้าใกล้ศูนย์ แต่ยังสามารถหาค่า $(\mathbf{X}'\mathbf{X})^{-1}$ ได้แต่สมาชิกของ $(\mathbf{X}'\mathbf{X})^{-1}$ จะมีค่าสูง (มนตรี, 2544) จึงทำให้ $\hat{\boldsymbol{\beta}}$ ซึ่งคำนวณมาจาก $(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ เกิดความไม่เสถียร นอกจากนี้ยังทำให้ความคลาดเคลื่อนของ $\hat{\boldsymbol{\beta}}$ มีค่าสูง ส่งผลต่อการทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอย จึงอาจทำให้ผู้วิจัยตัดตัวแปรอิสระที่มีอิทธิพลต่อตัวแปรตามออกไปจากตัวแบบ

วิธีการถดถอยแบบขั้นตอน เป็นวิธีการเลือกตัวแปรอิสระเข้าในตัวแบบการถดถอยครั้งละหนึ่งตัวแปร ตัวแปรอิสระใดที่ถูกเลือกเข้ามาอยู่ในตัวแบบการถดถอยแล้วจะต้องมีการทดสอบว่าตัวแปรอิสระที่อยู่ในตัวแบบการถดถอยแล้วมีส่วนในการอธิบายความผันแปรของตัวแปรตาม ขณะที่ตัวแปรอิสระอื่นอยู่ในตัวแบบการถดถอย (Efroymsom, 1960) นั่นคือตัวแปรอิสระใดที่ถูกเลือกเข้ามาอยู่ในตัวแบบการถดถอยแล้วอาจจะถูกนำออกไปได้ภายหลังถ้าพบว่าตัวแปรอิสระนั้นไม่มีนัยสำคัญ หรือกล่าวได้ว่า วิธีการถดถอยแบบขั้นตอนเป็นวิธีที่รวมขั้นตอน

ของวิธีเลือกตัวแปรอิสระแบบไปข้างหน้าและวิธีตัดตัวแปรอิสระออกแบบถอยหลังเข้าไว้ด้วยกัน โดยขั้นตอนของวิธีการถอยแบบขั้นตอน แบ่งได้ ดังนี้

1) เหมือนกับวิธีเลือกตัวแปรอิสระแบบไปข้างหน้า คือ เริ่มต้นพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม โดยเลือกตัวแปรอิสระซึ่งมีค่าสัมประสิทธิ์สหสัมพันธ์กับตัวแปรตามสูงที่สุด เข้าสู่ตัวแบบเป็นตัวแรก

2) ภายหลังจากที่ตัวแปรอิสระถูกนำเข้าไปในตัวแบบแล้วจะมีการทดสอบนัยสำคัญของตัวแปรอิสระที่นำเข้าไปเป็นตัวแรก โดยใช้ค่าสถิติ t (t-value) หรือ ค่าสถิติเอฟ (F-value)

3) พิจารณาค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรอิสระกับตัวแปรตาม โดยจะเลือกตัวแปรอิสระที่ให้ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนกับตัวแปรตามสูงที่สุด เข้ามาอยู่ในตัวแบบเป็นตัวต่อมา

4) หลังจากนั้นจึงทำการทดสอบความมีนัยสำคัญของตัวแปรที่เข้ามาเป็นตัวสุดท้ายก่อน โดยพิจารณาจากค่าสถิติเอฟบางส่วน ถ้ามีค่าน้อยกว่าเกณฑ์ที่กำหนดไว้จะต้องนำตัวแปรนั้นออกจากตัวแบบ แต่ถ้าค่าสถิติเอฟบางส่วนมีค่ามากกว่าเกณฑ์ที่กำหนดไว้ ก็จะย้อนกลับไปทดสอบตัวแปรอิสระตัวก่อนหน้าที่เข้ามาอยู่ในตัวแบบ โดยพิจารณาจากค่าสถิติเอฟ บางส่วน เช่นเดียวกัน

ทำซ้ำในขั้นตอนที่สามและสี่ต่อไปเรื่อย ๆ จนกระทั่งไม่สามารถนำตัวแปรอิสระใดออกจากตัวแบบได้ หรือไม่สามารถนำตัวแปรอิสระใดเข้าสู่ตัวแบบได้อีก

การค้นหาแบบต้องห้าม (tabu search) เป็นวิธีการหาค่าตอบแบบมีเหตุผล (metaheuristic) วิธีหนึ่ง โดยวิธีดังกล่าวถูกดัดแปลงมาจากวิธีการทางปัญญาประดิษฐ์ (artificial intelligence) ให้มีความเหมาะสมในการใช้คำตอบเดิมในการตอบปัญหาเดิมที่มีวัตถุประสงค์เปลี่ยนแปลงไป และนำมาใช้ในการหาค่าตอบที่ใกล้เคียงกับค่าที่ดีที่สุด (Glover, 1990) ซึ่งนิยมใช้กับปัญหาในการตัดสินใจ เช่น ปัญหาการหาเส้นทางเดินของพนักงานขาย (travelling salesman problem) (Knox, 1989) ปัญหาการจัดช่องทางในการให้บริการในร้านสะดวกซื้อ ปัญหาการจัดตารางเวลาการทำงานในโรงงาน (job shop scheduling) เป็นต้น ขั้นตอนของวิธีการนี้ไม่ซับซ้อนและผลลัพธ์ที่ได้มีประสิทธิภาพสูง แนวคิดของวิธีการนี้ คือ ใช้การจดจำจากรอบการทำซ้ำที่ผ่านมา โดยใช้หน่วยความจำในเครื่องคอมพิวเตอร์มาช่วยในการประมวลผล นอกจากนี้ยังสามารถใช้ในการหาค่าตอบของปัญหาที่มีความซับซ้อนและมีตัวแปรที่ใช้พิจารณาเป็นจำนวนมาก วิธีการนี้ช่วยป้องกันการเกิดปัญหาค่าตอบที่ดีที่สุดเฉพาะที่ (local optimum) โดยแก้ไขการย้อนกลับไปหาค่าตอบเดิมและดำเนินการค้นหาคำตอบต่อไปจนกระทั่งได้คำตอบที่ดีที่สุดทั่วไป (global optimum) รูปแบบการค้นหาคำตอบประกอบด้วยการค้นหาที่สำคัญ 2 รูปแบบ คือ การค้นหาคำตอบโดยใช้หน่วยความจำระยะสั้น (short term memory) ซึ่งเป็นหน่วยความจำตามเวลา (recency base memory) หรือการค้นหาที่จดจำอดีตในการค้นหาที่ผ่านมาในระยะเวลาสั้น และการค้นหาคำตอบโดยการใช้หน่วยความจำระยะยาว (long term memory) ซึ่งถือเป็นหน่วยความจำตามความถี่ (frequency base memory) หมายถึง การค้นหาที่จดจำอดีตเพื่อช่วยให้การค้นหาคำตอบเป็นไปอย่างมีประสิทธิภาพมากขึ้น มีองค์ประกอบหลัก ดังนี้

1. ฟังก์ชันเป้าหมาย (objective function) ใช้ในการประกอบการตัดสินใจในการค้นหาคำตอบ โดยจะเป็นตัวประเมินการค้นหาคำตอบที่อยู่ในแต่ละขั้นตอน
2. ข้อจำกัดการเป็นต้องห้าม (tabu restriction) คือ การควบคุมการย้อนกลับ (reverse) หรือการวนรอบ (cyclic) อยู่ในคำตอบเดิม โดยการกำหนดข้อจำกัดการเป็นต้องห้ามซึ่งขึ้นอยู่กับแต่ละปัญหา เช่น กำหนดระยะเวลาหรือจำนวนรอบที่จะต้องอยู่ในรายการต้องห้ามเป็นอย่างน้อย ฯลฯ
3. รายการต้องห้าม (tabu list) เป็นรายการค่าของตัวแปรที่ถูกปรับเปลี่ยนในแต่ละรอบของการคัดเลือกค่าของตัวแปร เพื่อป้องกันไม่ให้เกิดค่าที่คำตอบเดิม
4. ข้อมูลความถี่ (frequency move) เป็นการใช้หน่วยความจำระยะยาวบันทึกการค้นหาค่าตลอดช่วงของการค้นหา โดยใช้หลักการสร้างความหลากหลาย (diversification) เพื่อค้นหาคำตอบในบริเวณที่แตกต่างจากที่ค้นพบมาแล้ว
5. เกณฑ์ความปรารถนา (aspiration criteria) เป็นรูปแบบการควบคุมที่ช่วยในการค้นหาคำตอบที่เหมาะสมเป็นไปอย่างรวดเร็ว โดยกำหนดเงื่อนไขสำหรับการค้นหาที่เป็นรายการต้องห้ามซึ่งให้คำตอบที่เหมาะสมที่สุดตั้งแต่การเริ่มค้นหาจนถึงปัจจุบัน จนสามารถยกเลิกการห้ามสำหรับการค้นหานั้นได้ อย่างไรก็ตามเกณฑ์ดังกล่าวนี้ขึ้นอยู่กับลักษณะของแต่ละปัญหา

ขั้นตอนทั่วไปในการทำงานสรุปได้ดังนี้

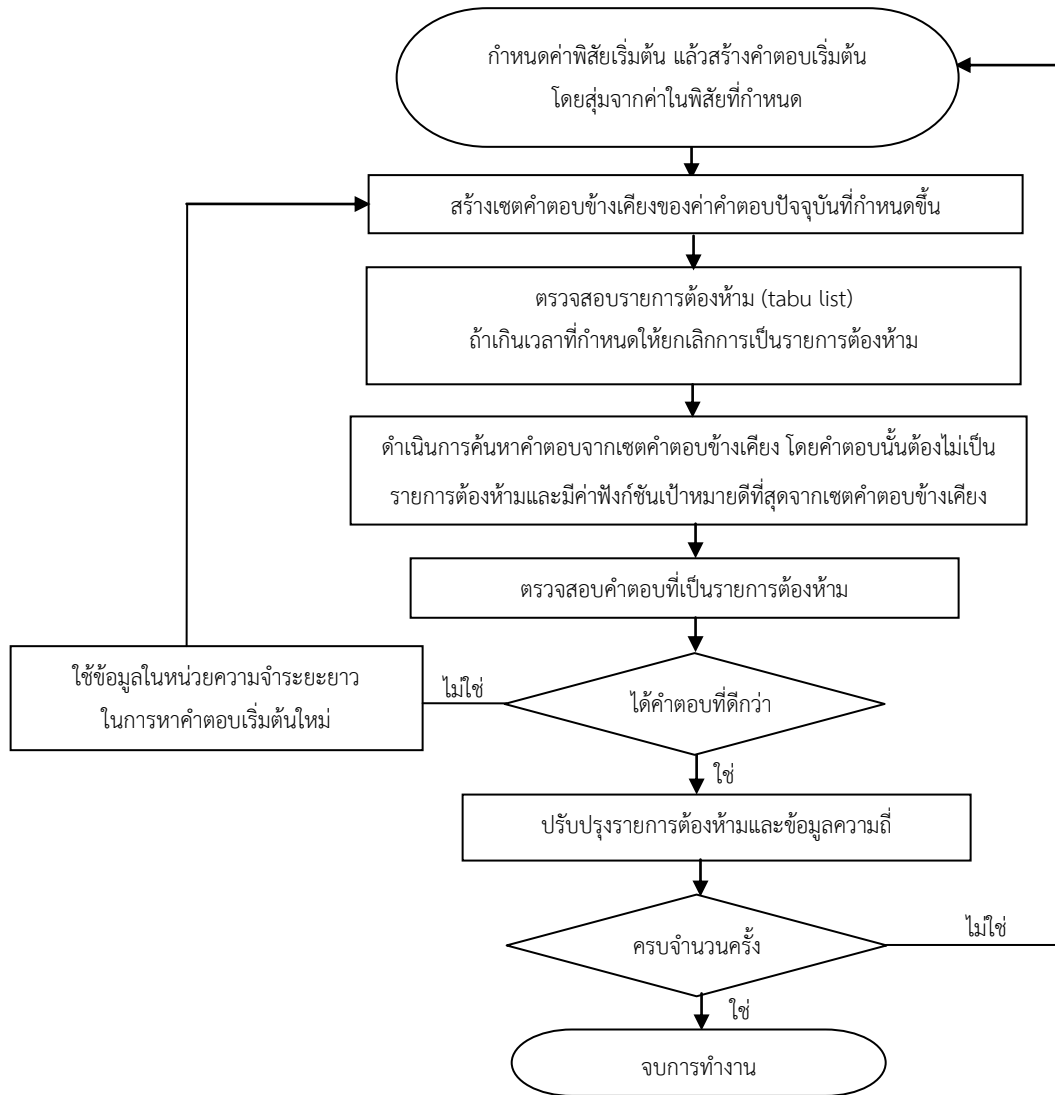
- 1) เริ่มจากการหาคำตอบเริ่มต้น โดยมีการจัดทำโครงสร้างหน่วยความจำมาใช้ในการเก็บค่า
- 2) ต่อมาสร้าง candidate และเลือกคำตอบจากใน candidate ที่มีความเหมาะสมที่สุด โดยจะหยุดการค้นหาคำตอบตามเงื่อนไขที่กำหนดไว้
- 3) ปรับปรุง (update) คำตอบโดยการแทนที่คำตอบที่ดีที่สุดในปัจจุบันด้วยคำตอบที่ดีกว่า จากนั้นจึงเก็บคำตอบบันทึกไว้ด้วย

ดังนั้นจะพบว่าวิธีการค้นหาแบบต้องห้าม (tabu search) เป็นวิธีการแก้ปัญหาที่ไม่ซับซ้อน แต่มีความสำคัญอยู่ที่การออกแบบโครงสร้างและการจัดการข้อมูลในหน่วยความจำ รวมถึงการกำหนดลักษณะ (attribute) ของ candidate ดังแสดงไว้ในภาพที่ 1

อนึ่ง การค้นหาแบบต้องห้ามมิได้ใช้วิธีการประมาณค่าสัมประสิทธิ์การถดถอย (β) โดยวิธีกำลังสองน้อยที่สุด หรือวิธีภาวะน่าจะเป็นสูงสุด ดังในวิธีที่ได้กล่าวมาทั้งหมด แต่การค้นหาแบบต้องห้ามจะกำหนดค่าของสัมประสิทธิ์การถดถอย (β) ที่ทำให้ฟังก์ชันเป้าหมายมีค่าต่ำสุด จึงไม่ต้องคำนวณค่า $(X'X)^{-1}$ ดังนั้นการค้นหาแบบต้องห้ามจึงสามารถใช้ได้กับข้อมูลทั้งที่ไม่มีและมีสหสัมพันธ์เชิงเส้นพหุ (multicollinearity)

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบการคัดเลือกตัวแปรอิสระ ในกรณีที่มีข้อมูลไม่มีและมีสหสัมพันธ์เชิงเส้นพหุระหว่างตัวแปรอิสระ โดยวิธีการค้นหาแบบต้องห้าม ที่ใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (mean squares error: MSE) และค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (mean absolute error: MAE) เป็นฟังก์ชันเป้าหมาย กับวิธีการถดถอยแบบขั้นตอน ในตัวแบบการถดถอยเชิงเส้นพหุ



รูปที่ 1 ขั้นตอนวิธีการค้นหาแบบต้องห้าม

วิธีการดำเนินงานวิจัย

งานวิจัยนี้ศึกษากรณีตัวแปรอิสระมีการแจกแจงเอกรูป (uniform distribution) และความคลาดเคลื่อนมีการแจกแจงปกติ (normal distribution) จำนวนตัวแปรอิสระในตัวแบบเต็มรูป (full model) เท่ากับ 6 ตัวแปร และจำนวนตัวแปรอิสระในตัวแบบแท้จริง (true model) ซึ่งใช้ในการหาค่าตัวแปรตาม เท่ากับ 4 ตัวแปร โดยมีขั้นตอนการดำเนินงาน ดังนี้

กรณีที่ข้อมูลไม่มีสหสัมพันธ์เชิงเส้นพหุ

- กำหนดขนาดประชากร $N=100,000$ และสร้างตัวแปรอิสระ สัมประสิทธิ์การถดถอย ความคลาดเคลื่อน ดังนี้ สร้างตัวแปรอิสระ 6 ตัวแปรที่เป็นอิสระต่อกัน ให้มีการแจกแจงเอกรูป โดย $X_1 \sim U(15,80)$

$X_2 \sim U(20,150)$ $X_3 \sim U(10,100)$ $X_4 \sim U(-50,50)$ $X_5 \sim U(-100,100)$ และ
 $X_6 \sim U(-80,400)$ รวมทั้งสร้างความคลาดเคลื่อนที่เป็นอิสระต่อกันให้มีการแจกแจงปกติ $\epsilon \sim N(0,2500)$
 และ $\epsilon \sim N(0,25)$

- กำหนดค่าสัมประสิทธิ์การถดถอย คือ $\beta' = (100, 24, 15, -8, 5, 0, 0)$ ในการสร้างตัวแปรตาม y หรืออีกนัยหนึ่ง ตัวแปรตาม y จะขึ้นอยู่กับตัวแปรอิสระ 4 ตัว คือ X_1, X_2, X_3 และ X_4 เท่านั้น

- สร้างตัวแปรตาม y จากตัวแบบการถดถอยต่อไปนี้ $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon$

- เมื่อได้ประชากรตามที่ต้องการ ดำเนินการสุ่มตัวอย่างขนาด 25 และ 100

- ทำการคัดเลือกตัวแบบที่เหมาะสมจากเทคนิคการคัดเลือกตัวแบบทั้ง 2 เทคนิค คือ วิธีการค้นหาแบบต้องห้าม และวิธีการถดถอยแบบขั้นตอน และคำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) จากการทำซ้ำจำนวน 500 ครั้งของตัวแบบที่ได้จากการคัดเลือกตัวแบบทั้ง 2 เทคนิค แล้วคำนวณร้อยละความถูกต้องของการคัดเลือกตัวแปรเข้าสู่ตัวแบบจากการทำซ้ำจำนวน 500 ครั้ง ในแต่ละเทคนิค

- เปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวแบบและร้อยละความถูกต้องของการคัดเลือกตัวแปรเข้าสู่ตัวแบบที่ได้จากเทคนิคการคัดเลือกตัวแปรทั้ง 2 เทคนิค

กรณีที่ข้อมูลมีสหสัมพันธ์เชิงเส้นพหุ

ขั้นตอนเช่นเดียวกับกรณีที่ข้อมูลไม่มีสหสัมพันธ์เชิงเส้นพหุ แต่มีการกำหนดค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 เท่ากับ 0.999 โดยใช้วิธีรากที่สอง (square root method)

- เมื่อได้ประชากรตามที่ต้องการ ดำเนินการสุ่มตัวอย่างขนาด 25 และ 100 แล้วคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 ในแต่ละตัวอย่างที่สุ่มมาได้

- เนื่องจากปัญหาสหสัมพันธ์เชิงเส้นพหุ เป็นปัญหาที่เกิดในตัวอย่าง (sample) ดังนั้นผู้วิจัยจึงพิจารณาแบ่งกลุ่มของตัวอย่างที่ได้จากการสุ่มตามค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง X_1 และ X_4 ดังนี้

$0.495 \leq r \leq 0.504$ จัดให้อยู่ในกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.50

$0.945 \leq r \leq 0.954$ จัดให้อยู่ในกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.95

$0.955 \leq r \leq 0.964$ จัดให้อยู่ในกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.96

$0.965 \leq r \leq 0.974$ จัดให้อยู่ในกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.97

$0.975 \leq r \leq 0.984$ จัดให้อยู่ในกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.98

ทำซ้ำเช่นนี้จนครบ 500 ครั้ง ในแต่ละกลุ่มค่าสัมประสิทธิ์สหสัมพันธ์และขนาดตัวอย่าง

- ทำการคัดเลือกตัวแบบที่เหมาะสมจากเทคนิคการคัดเลือกตัวแบบทั้ง 2 เทคนิค คือ วิธีการค้นหาแบบต้องห้าม และวิธีการถดถอยแบบขั้นตอน และคำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) จากการทำซ้ำจำนวน 500 ครั้งของตัวแบบที่ได้จากการคัดเลือกตัวแบบทั้ง 2 เทคนิค แล้วคำนวณร้อยละความถูกต้องของการคัดเลือกตัวแปรเข้าสู่ตัวแบบจากการทำซ้ำจำนวน 500 ครั้ง ในแต่ละเทคนิค

- เปรียบเทียบค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวแบบและร้อยละความถูกต้องของการคัดเลือกตัวแปรเข้าสู่ตัวแบบที่ได้จากเทคนิคการคัดเลือกตัวแปรทั้ง 2 เทคนิค

ผลการดำเนินงานวิจัย

ในบทความนี้ผู้วิจัยขอเสนอผลการวิจัยเฉพาะกรณี $n = 25$

ตารางที่ 1 ค่าเฉลี่ยของสัมประสิทธิ์การถดถอย ค่าความคลาดเคลื่อนมาตรฐาน ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยจำแนกตามค่าสัมประสิทธิ์สหสัมพันธ์และเกณฑ์ที่ใช้ในการคัดเลือกตัวแบบ เมื่อ $\varepsilon_i \sim N(0,2500)$, $n = 25$

สัมประสิทธิ์การถดถอย ค่าที่กำหนด	b_0	b_1	b_2	b_3	b_4	b_5	b_6	MSE	ร้อยละของ ตัวแบบที่ ถูกต้อง
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MSE , $r^* < 0.1$									
ค่าเฉลี่ย	99.7028	26.2503	15.3315	-10.3452	7.5213	0.0414	-0.0223	2509.13	88.60
ค่าความคลาดเคลื่อนมาตรฐาน	45.3017	0.5811	0.2933	0.4188	0.3717	0.1715	0.0760		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MAE , $r^* < 0.1$									
ค่าเฉลี่ย	98.9161	26.1159	15.3436	-10.2703	7.6571	-0.0817	-0.0248	6002.32	85.60
ค่าความคลาดเคลื่อนมาตรฐาน	73.1655	0.9835	0.4737	0.6763	0.6003	0.2770	0.1228		
วิธีการ : Stepwise , ฟังก์ชันเป้าหมาย : MSE , $r^* < 0.1$									
ค่าเฉลี่ย	100.9034	23.9646	15.0186	-8.0125	5.0032	0.0606	0.0799	2513.25	89.40
ค่าความคลาดเคลื่อนมาตรฐาน	45.3480	0.5817	0.2936	0.4592	0.3721	0.1717	0.0761		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MSE , $0.495 \leq r^* \leq 0.504$									
ค่าเฉลี่ย	104.3217	22.8864	15.4021	-7.0176	4.1275	-0.0898	0.0314	2511.32	79.40
ค่าความคลาดเคลื่อนมาตรฐาน	30.1552	1.1254	0.3010	0.4213	3.8976	0.2202	0.1621		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MAE , $0.495 \leq r^* \leq 0.504$									
ค่าเฉลี่ย	104.2415	22.9238	15.4273	-6.9981	4.0321	0.0731	0.0206	6179.11	79.20
ค่าความคลาดเคลื่อนมาตรฐาน	38.9916	2.1055	0.6541	0.6597	7.1125	0.4095	0.3088		
วิธีการ : Stepwise , ฟังก์ชันเป้าหมาย : MSE , $0.495 \leq r^* \leq 0.504$									
ค่าเฉลี่ย	59.1386	26.7712	15.0089	-8.0095	9.4512	-0.0054	0.0201	2517.24	45.20
ค่าความคลาดเคลื่อนมาตรฐาน	30.2548	1.1901	0.3129	0.4609	3.9905	0.2303	0.1639		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MSE , $0.945 \leq r^* \leq 0.954$									
ค่าเฉลี่ย	105.7548	20.9375	15.4692	-6.9312	3.8820	-0.1098	0.0654	2521.52	71.40
ค่าความคลาดเคลื่อนมาตรฐาน	20.0923	2.6750	0.2753	0.4431	2.9756	0.1864	0.0987		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MAE , $0.945 \leq r^* \leq 0.954$									
ค่าเฉลี่ย	105.7425	20.9416	15.4937	-6.7558	3.6995	0.1123	0.0456	6879.11	70.20
ค่าความคลาดเคลื่อนมาตรฐาน	33.1868	4.4184	0.4547	0.7318	4.9148	0.3079	0.1629		
วิธีการ : Stepwise , ฟังก์ชันเป้าหมาย : MSE , $0.945 \leq r^* \leq 0.954$									
ค่าเฉลี่ย	-34.1497	28.8845	15.0109	-8.0175	12.4643	-0.0181	0.0322	2523.11	15.60
ค่าความคลาดเคลื่อนมาตรฐาน	20.0987	2.6579	0.2754	0.4432	2.9765	0.1865	0.0987		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MSE , $0.955 \leq r^* \leq 0.964$									
ค่าเฉลี่ย	104.8070	20.9193	15.7283	-6.8772	3.6059	-0.2167	0.0568	2521.54	70.80
ค่าความคลาดเคลื่อนมาตรฐาน	15.9591	2.4090	0.2892	0.4151	2.5723	0.1588	0.0727		

ตารางที่ 1 (ต่อ)

สัมประสิทธิ์การถดถอย ค่าที่กำหนด	b_0	b_1	b_2	b_3	b_4	b_5	b_6	MSE	ร้อยละของ ตัวแบบที่ ถูกต้อง
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MAE , $0.955 \leq r^* \leq 0.964$									
ค่าเฉลี่ย	104.7409	20.9776	15.4020	-6.7680	3.4288	0.2389	0.0901	6880.34	70.00
ค่าความคลาดเคลื่อนมาตรฐาน	26.2321	3.9597	0.4754	0.6823	4.2282	0.2610	0.1194		
วิธีการ : Stepwise , ฟังก์ชันเป้าหมาย : MSE , $0.955 \leq r^* \leq 0.964$									
ค่าเฉลี่ย	-36.2087	29.8401	15.0093	-8.0199	15.2276	-0.1658	0.0327	2524.17	12.40
ค่าความคลาดเคลื่อนมาตรฐาน	15.9675	2.4103	0.2894	0.4153	2.5737	0.1589	0.0727		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MSE , $0.965 \leq r^* \leq 0.974$									
ค่าเฉลี่ย	103.7689	21.0725	15.2612	-6.7661	3.7082	-0.1176	0.0567	2521.24	71.00
ค่าความคลาดเคลื่อนมาตรฐาน	8.7749	2.2003	0.2880	0.4137	1.9198	0.1574	0.0736		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MAE , $0.965 \leq r^* \leq 0.974$									
ค่าเฉลี่ย	103.6011	21.2781	15.3523	-6.9222	3.6174	0.0875	0.0885	6787.11	70.40
ค่าความคลาดเคลื่อนมาตรฐาน	14.3964	3.6098	0.4726	0.6787	3.1497	0.2583	0.1207		
วิธีการ : Stepwise , ฟังก์ชันเป้าหมาย : MSE , $0.965 \leq r^* \leq 0.974$									
ค่าเฉลี่ย	-48.8832	30.6709	15.0092	-8.0156	16.9987	-0.1844	0.0717	2522.25	10.20
ค่าความคลาดเคลื่อนมาตรฐาน	8.7794	2.2014	0.2882	0.4139	1.9208	0.1575	0.0736		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MSE , $0.975 \leq r^* \leq 0.984$									
ค่าเฉลี่ย	104.4026	21.0057	15.7771	-6.7475	3.6870	-0.0812	0.0765	2520.52	71.20
ค่าความคลาดเคลื่อนมาตรฐาน	25.9240	2.0036	0.2861	0.4114	1.4284	0.1547	0.0722		
วิธีการ : Tabu , ฟังก์ชันเป้าหมาย : MAE , $0.975 \leq r^* \leq 0.984$									
ค่าเฉลี่ย	107.8179	21.4606	15.5949	-7.0149	3.7354	0.0487	0.0875	6809.35	69.40
ค่าความคลาดเคลื่อนมาตรฐาน	43.1861	3.3378	0.4766	0.6853	2.3795	0.2577	0.1203		
วิธีการ : Stepwise , ฟังก์ชันเป้าหมาย : MSE , $0.975 \leq r^* \leq 0.984$									
ค่าเฉลี่ย	-58.6637	30.6233	15.0132	-8.0128	20.1427	-0.0774	0.0621	2524.11	9.20
ค่าความคลาดเคลื่อนมาตรฐาน	25.9428	2.0051	0.2863	0.4117	1.4294	0.1548	0.0723		

หมายเหตุ: *ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4

สรุปผลการดำเนินงานวิจัย

กำหนดข้อความเพื่ออธิบายความหมาย ดังนี้

ตัวแปรอิสระ X_1 และ X_4 มีสหสัมพันธ์เชิงเส้นพหุในระดับต่ำ หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 มีค่าน้อยกว่า 0.1

ตัวแปรอิสระ X_1 และ X_4 มีสหสัมพันธ์เชิงเส้นพหุในระดับกลาง หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 มีค่าเท่ากับ 0.5

ตัวแปรอิสระ X_1 และ X_4 มีสหสัมพันธ์เชิงเส้นพหุในระดับสูง หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 มีค่าเท่ากับ 0.95 0.96 0.97 และ 0.98

กรณีตัวแปรอิสระ X_1 และ X_4 มีสหสัมพันธ์เชิงเส้นพหุในระดับต่ำ วิธีการคัดเลือกตัวแปรอิสระโดยวิธีการถดถอยแบบขั้นตอน จะได้ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระทุกตัวมีค่าใกล้เคียงกับค่าที่กำหนดในการจำลองและมีร้อยละของตัวแบบที่คัดเลือกได้ถูกต้องสูงที่สุด ในขณะที่วิธีการค้นหาแบบต้องห้ามเมื่อใช้ฟังก์ชันเป้าหมายทั้ง 2 ฟังก์ชัน จะให้ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระ $X_1(\beta_1)$ สูงกว่าค่าที่กำหนดในการจำลองประมาณร้อยละ 10 ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระ $X_3(\beta_3)$ ต่ำกว่าค่าที่กำหนดในการจำลองประมาณร้อยละ 20 และค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระ $X_4(\beta_4)$ สูงกว่าค่าที่กำหนดในการจำลองประมาณร้อยละ 50 แต่มีร้อยละของการคัดเลือกตัวแบบที่ถูกต้องต่ำกว่าวิธีการถดถอยแบบขั้นตอนประมาณร้อยละ 1 นอกจากนี้ยังพบว่า มีร้อยละของตัวแบบที่คัดเลือกได้ถูกต้องสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้นทั้งจากวิธีการค้นหาแบบต้องห้ามและวิธีการถดถอยแบบขั้นตอน

กรณีตัวแปรอิสระ X_1 และ X_4 มีสหสัมพันธ์เชิงเส้นพหุในระดับกลางและระดับสูง วิธีการคัดเลือกตัวแปรอิสระโดยใช้วิธีการค้นหาแบบต้องห้าม เมื่อใช้ฟังก์ชันเป้าหมายทั้ง 2 ฟังก์ชัน จะได้ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระเกือบทุกตัวมีค่าใกล้เคียงกับค่าที่กำหนดในการจำลองในทุกสถานการณ์และมีร้อยละของการคัดเลือกตัวแบบได้ถูกต้องมากกว่าวิธีการถดถอยแบบขั้นตอน นอกจากนี้พบว่าทั้งสองวิธีมีร้อยละของตัวแบบที่คัดเลือกได้ถูกต้องสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้นและใกล้เคียงกันเมื่อค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 เปลี่ยนไปจากระดับกลางไประดับสูง ในขณะที่การคัดเลือกตัวแปรอิสระโดยวิธีการถดถอยแบบขั้นตอน จะให้ค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระ (β_0) มีเครื่องหมายผิดไปจากที่กำหนดในการจำลอง เมื่อตัวแปรอิสระ X_1 และ X_4 มีสหสัมพันธ์เชิงเส้นพหุในระดับสูง นอกจากนี้ยังพบว่าค่าสัมประสิทธิ์การถดถอยของตัวแปรอิสระ $X_1(\beta_1)$ และ $X_4(\beta_4)$ มีค่าสูงกว่าค่าที่กำหนดในการจำลองทุกสถานการณ์ อีกทั้งพบว่าร้อยละของตัวแบบที่คัดเลือกได้ถูกต้องสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้นแต่จะมีค่าลดลงเมื่อค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X_1 และ X_4 เพิ่มขึ้น

ข้อเสนอแนะ

ในการวิจัยครั้งต่อไป ควรศึกษากรณีตัวแบบการถดถอยไม่เป็นไปตามข้อตกลงเบื้องต้นอื่นๆ เช่น ความแปรปรวนของความคลาดเคลื่อนมีค่าไม่คงที่ ความคลาดเคลื่อนไม่ได้มีการแจกแจงแบบปกติ เป็นต้น นอกจากนี้ อาจใช้วิธีการคัดเลือกตัวแบบด้วยวิธีการอื่น ๆ เช่น วิธีการค้นหาคำตอบแบบ genetic algorithm วิธีการค้นหาคำตอบแบบ simulated annealing

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณ รศ.ดร.วิจิต หล่อจิระชุมหัทกุล ที่ให้คำแนะนำทำให้งานวิจัยชิ้นนี้สำเร็จลุล่วงไปได้ด้วยดี

เอกสารอ้างอิง

- ประชุม สุวัฒน์. (2553). ทฤษฎีการอนุมานเชิงสถิติ. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: โครงการส่งเสริมและพัฒนาเอกสารวิชาการ สถาบันบัณฑิตพัฒนบริหารศาสตร์.
- มนตรี พิริยะกุล. (2545). เทคนิคการวิเคราะห์สมการถดถอย. กรุงเทพมหานคร: มหาวิทยาลัยรามคำแหง.
- วิจิต หล่อจีระชุนท์กุล และจิราวัลย์ จิตรถเวช. (2548). เทคนิคการพยากรณ์. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: โครงการส่งเสริมและพัฒนาเอกสารวิชาการ สถาบันบัณฑิตพัฒนบริหารศาสตร์.
- Efroymson, M. A. (1960). Multiple Regression Analysis: Mathematical Methods for Digital Computers. New York: Wiley.
- Glover, F. (1990). Tabu Search: a Tutorial. Interfaces. 20 (November): 74-94.
- Knox, J. (1989). The Application of Tabu Search to the Symmetric Travelling Salesman Problem. Boulder: Graduate School of Business, University of Colorado.

