ปัญหาการลู่เข้าในการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้น

อัญชลี ทองกำเหนิด [1] และ ลี่ลี อิงศรีสว่าง [2]

มหาวิทยาลัยเกษตรศาสตร์ จตุจักร กรุงเทพฯ 10903

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาสถานการณ์ที่มีผลต่อการลู่เข้าสู่ (converge) ค่าพารามิเตอร์ สำหรับการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้น โดยมี 3 ปัจจัยที่นำมาพิจารณาได้แก่ 1) การเลือกค่าเริ่มต้น ได้แก่การเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการเลือกค่าเริ่มต้นโดยใช้ grid search 2) ลักษณะการแจกแจงของข้อมูลในที่นี้พิจารณา 2 ตัวแบบคือ Negative Exponential และ Monomolecular และ 3) ขนาดของตัวอย่าง ที่ใช้ในการศึกษานี้ได้แก่ตัวอย่างขนาด n=20 n=30 n=50 และ n=100 ตามลำดับ ผลจากการจำลองสถานการณ์ต่างๆ ทั้งหมด และทำซ้ำ 500 รอบ พบว่า เมื่อขนาดของตัวอย่างเพิ่มขึ้นจาก n=20 เป็น n=100 ทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์ และจำนวนรอบเฉลี่ยที่ converge เข้าสู่ค่าพารามิเตอร์ของวิธี Gauss Newton, วิธี Marquardt และวิธี Newton ให้ผลที่ไม่แตกต่างกันคือ เมื่อขนาดของตัวอย่างเพิ่มมากขึ้นจะทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มมากขึ้นจนเกือบเท่ากับ 100 เปอร์เซ็นต์ และส่งผลให้จำนวนรอบเฉลี่ยในการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้นด้วย นอกจากนั้นเมื่อขนาดของตัวอย่างเพิ่มขึ้นวิธีการเลือกค่าเริ่มต้นที่แตกต่างกันก็แทบจะไม่มีผลต่อการ converge เข้าสู่ค่าพารามิเตอร์ ซึ่งจากผลการทดลองเมื่อ n=100 พบว่าวิธีการที่ให้ผลดีที่สุดคือ วิธี Marquardt ที่เลือกค่าเริ่มต้นโดยใช้ grid search โดยทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์ และจำนวนรอบเฉลี่ยในการ converge เข้าสู่ค่าพารามิเตอร์ เท่ากับ 100 เปอร์เซ็นต์ และ 39.58 รอบ ในตัวแบบ Negative Exponential และเท่ากับ 98 เปอร์เซ็นต์ และ 47.98 รอบ ในตัวแบบ Monomolecular ตามลำดับ

คำสำคัญ : การถดถอยไม่เชิงเส้น / ปัญหาการลู่เข้า / การประมาณค่าพารามิเตอร์

---

[1] *นิสิตปริญญาเอก ภาควิชาสถิติ คณะวิทยาศาสตร์*

[2] *ผู้ช่วยศาสตราจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์*

Convergence Problem in Nonlinear Regression for Parameter Estimation

Unchalee Tonggumnead [1] and Lily Insrisawang [2]

Kasetsart University, Jatujuk, Bangkok 10903

Abstract

The purpose of this research is to study three effective factors for convergence problem in nonlinear regression parameter estimation : 1) choice of initial value, considering from : Fekedulegn's method and Grid Search, 2) distribution of data, considering from : Negative Exponential and Monomolecular and 3) sample size : n =20 , n=30 , n=50 and n =100 respectively. Simulations was performed and repeated 500 times for each scenario. The result showed that in most situations, when the sample size increase from n=20 to n=100 Gauss Newton method, Marquardt and Newton method will make no difference in percent of convergence to parameter and the number of average iteration, namely, when the sample size increase, it will make percent of convergence to parameter increase almost 100 percent. In addition, when the sample size increase from n=20 to n=100, choice of different initial value does not effect for convergence to parameter. From this research the best method is Marquardt by using Grid Search. It make percent of convergence to parameter and the number of average iteration equal to 100 percent and 39.58 iterations in Negative Exponential model and equal to 98 percent and 47.98 iterations in Monomolecular model respectively.

Keywords : Nonlinear Regression / Convergence Problem / Parameter Estimation

[1] *Ph.D. (Statistics) Student, Department of Statistic, Faculty of Science.*

[2] *Assistant Professor, Department of Statistic, Faculty of Science.*

1. บทนำ

การศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ (Regressor variable / predictor variable) X และตัวแปรตาม (response variable) Y ที่สำคัญวิธีหนึ่งคือการวิเคราะห์การถดถอย (regression analysis) ซึ่งตัวแบบการถดถอยสามารถแบ่งได้เป็น 2 ประเภทใหญ่ๆ คือ ตัวแบบการถดถอยเชิงเส้น (linear regression model) และตัวแบบการถดถอยแบบไม่เชิงเส้น (nonlinear regression model) โดยที่ตัวแบบเชิงเส้นหมายถึง ตัวแบบการถดถอยที่สามารถเขียนได้ในลักษณะเชิงเส้นของพารามิเตอร์ (linear in parameter) ในขณะที่ตัวแบบการถดถอยไม่เชิงเส้นหมายถึง ตัวแบบการถดถอยที่ไม่สามารถเขียนได้ในแบบเส้นของพารามิเตอร์ (nonlinear in parameter) [1]

การหาค่าสัมประสิทธิ์การถดถอยหรือการประมาณค่าพารามิเตอร์ในรูปแบบการถดถอยสามารถแบ่งเป็น 2 ประเภทใหญ่ ๆ คือ

1) วิธีกำลังสองน้อยที่สุด โดยสร้างสมการปกติจากรูปแบบการถดถอยไม่เชิงเส้น

$$Y_i = f(\underline{X}_i,\underline{\beta}) + \varepsilon_i$$

และหาตัวประมาณ b_j หรือเวกเตอร์ของตัวประมาณ $\underline{b}$ ที่ทำให้ผลรวมกำลังสองของความคลาดเคลื่อน (SSE) มีค่าน้อยที่สุด

2) ใช้วิธีการทางคณิตศาสตร์โดยการคำนวณแบบวนซ้ำ (iterative methods) ได้แก่ วิธี Gauss-Newton, Marquardt, Steepest Descent, Newton เป็นต้น

ซึ่งข้อเสียของวิธีกำลังสองน้อยที่สุดคือสมการปกติเขียนเป็นฟังก์ชั่นเชิงเส้นของตัวประมาณ b_j ไม่ได้ ทำให้แก้สมการหรือเขียนตัวประมาณ b_j ในเทอมของตัวแปร X และ Y ไม่ได้จึงมักไม่ใช้วิธีกำลังสองน้อยที่สุดในการประมาณค่าพารามิเตอร์ ในขณะที่การใช้วิธีการทางคณิตศาสตร์โดยการคำนวณแบบการวนซ้ำ อาจพบปัญหาในเรื่องของการลู่เข้าสู่ค่าพารามิเตอร์ ซึ่งในบางกรณีอาจใช้จำนวนรอบในการลู่เข้าสู่ค่าพารามิเตอร์สูง Jukic and Scitovski [2] และ Fekedulegn [3] ได้ศึกษาเกี่ยวกับปัจจัยที่มีผลต่อการลู่เข้าสู่ พบว่าค่าเริ่มต้นมีผลต่อการลู่เข้าสู่ค่าพารามิเตอร์ โดยถ้าเลือกค่าเริ่มต้นเหมาะสมจะทำให้จำนวนรอบในการลู่เข้าสู่ค่าพารามิเตอร์ลดลง

เนื่องจากการศึกษาในครั้งนี้ต้องการศึกษาสถานการณ์การลู่เข้าสู่ค่าพารามิเตอร์ โดยจากงานวิจัยที่กล่าวมาแล้วดังข้างต้นพบว่า ปัญหาในการลู่เข้าสู่ค่าพารามิเตอร์ขึ้นอยู่กับการเลือกค่าเริ่มต้น ดังนั้นในงานวิจัยนี้จึงต้องการศึกษาปัจจัยที่มีผลต่อการลู่เข้าสู่ค่าพารามิเตอร์ของรูปแบบการถดถอยไม่เชิงเส้น ซึ่งปัจจัยที่นำมาพิจารณาในงานวิจัยนี้ประกอบไปด้วย 3 ปัจจัย ได้แก่ 1) การเลือกค่าเริ่มต้น 2) ลักษณะการแจกแจงของข้อมูล และ 3) ขนาดของตัวอย่าง โดยการศึกษาครั้งนี้ได้พิจารณาการเลือกค่าเริ่มต้นโดยใช้วิธีการของ Fekedulegn [3] และการเลือกค่าเริ่มต้นโดยใช้ grid search (ซึ่งเป็น option หนึ่งใน SAS program) ลักษณะการแจกแจงของข้อมูลในการศึกษาครั้งนี้คือ Negative Exponential และ Monomolecular และขนาดของตัวอย่างที่ใช้ในงานวิจัยนี้ได้แก่ ตัวอย่างขนาด n=20 n=30 n=50 และ n=100 ตามลำดับ

2. วัตถุประสงค์ของการวิจัย

เพื่อศึกษาสถานการณ์การลู่เข้าสู่ค่าพารามิเตอร์ ในการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้น เมื่อมีการกำหนดค่าเริ่มต้นของพารามิเตอร์ ลักษณะการแจกแจงของข้อมูล และขนาดตัวอย่างที่แตกต่างกัน

3. วิธีดำเนินการวิจัย

การวิจัยในครั้งนี้ใช้ข้อมูลที่ได้จากการจำลองข้อมูลโดยทำซ้ำ 500 รอบ ในแต่ละสถานการณ์ ใช้โปรแกรม SAS เวอร์ชัน 9.1 โดยกำหนดให้

1. รูปแบบการถดถอยไม่เชิงเส้น งานวิจัยนี้อยู่ในกลุ่มของตัวแบบการเจริญเติบโต (growth model) 2 ตัวแบบ ได้แก่ Negative Exponential และ Monomolecular

a. Negative Exponential

$$Y=\beta_0(1-\exp(-\beta_2 X)) + \varepsilon \quad , \quad \varepsilon \sim N(0,1)$$

b. Monomolecular

$$Y=\beta_0(1-\beta_1\exp(-\beta_2 X)) + \varepsilon \quad , \quad \varepsilon \sim N(0,1)$$

ซึ่งทั้ง 2 ตัวแบบเป็นตัวแบบการถดถอยไม่เชิงเส้นในตระกูลของตัวแบบการเจริญเติบโตที่ไม่สามารถแปลงให้อยู่ในรูปเชิงเส้นตรงได้ (non – intrinsically linear)

2. Generate ข้อมูลที่มีรูปแบบ Negative Exponential และ Monomolecular โดย Y_i เป็นอิสระต่อกัน (response independence) X เป็นตัวแปรอิสระไม่ต่อเนื่องที่มีค่าห่างเท่าๆ กัน (equal space) ε_i มีการแจกแจงปกติมาตรฐาน generate ข้อมูล 4 ขนาดโดย 1) ตัวอย่างขนาด n=20 2) ตัวอย่างขนาด n=30 3) ตัวอย่างขนาด n=50 และ 4) ตัวอย่างขนาด n=100 ดังนั้นสามารถสรุปรูปแบบในการ generate ข้อมูลทุกรูปแบบได้ดังต่อไปนี้

- ข้อมูลที่อยู่ในรูปตัวแบบ Negative Exponential ตัวอย่างขนาด 20
- ข้อมูลที่อยู่ในรูปตัวแบบ Negative Exponential ตัวอย่างขนาด 30
- ข้อมูลที่อยู่ในรูปตัวแบบ Negative Exponential ตัวอย่างขนาด 50
- ข้อมูลที่อยู่ในรูปตัวแบบ Negative Exponential ตัวอย่างขนาด 100
- ข้อมูลที่อยู่ในรูปตัวแบบ Monomolecular ตัวอย่างขนาด 20
- ข้อมูลที่อยู่ในรูปตัวแบบ Monomolecular ตัวอย่างขนาด 30
- ข้อมูลที่อยู่ในรูปตัวแบบ Monomolecular ตัวอย่างขนาด 50
- ข้อมูลที่อยู่ในรูปตัวแบบ Monomolecular ตัวอย่างขนาด 100

3. กำหนดค่าเริ่มต้นของพารามิเตอร์ตามวิธีของ Fekedulegn [3] และการใช้ grid search

4. ประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้นด้วยวิธี Gauss-Newton Marquardt และ Newton โดยใช้ option PROC NLIN ใน SAS program คำนวณจำนวนครั้งในการวนซ้ำที่ลู่เข้าสู่ค่าพารามิเตอร์ ค่าพารามิเตอร์เฉลี่ย (mean of β) S.E.(β) เฉลี่ย และจำนวนรอบเฉลี่ย (mean of iteration) ที่ลู่เข้าสู่ค่าพารามิเตอร์

3.1 วิธี Gauss Newton

Bates and Watts [4] ได้อธิบายวิธีของ Gauss Newton หรือวิธี linearization เป็นวิธีการหาตัวประมาณของ b_j ของพารามิเตอร์ β_j ซึ่งมีขั้นตอนดังนี้

กำหนด $b_j^{(0)}$ เป็นค่าเริ่มต้นหรือค่าประมาณในรอบ (iteration) ที่ 0 ของพารามิเตอร์ β_j กระจายฟังก์ชั่น $f(\underline{X}_i,\underline{\beta})$ ในรูปของอนุกรมเทเลอร์ (Taylor series) รอบจุด $b_j^{(0)}$ และประมาณฟังก์ชั่น $f(\underline{X}_i,\underline{\beta})$ ในเทอมเส้นตรงได้เป็น

$$f(\underline{X}_i,\underline{\beta}) \approx f(\underline{X}_i,\underline{b}^{(0)}) + \sum_{j=1}^{p} \frac{\partial f(\underline{X}_i,\underline{\beta})}{\partial \beta_j}\bigg|_{\underline{\beta}=\underline{b}^{(0)}} (\beta_j - b_j^{(0)}) \quad \textbf{(1)}$$

สามารถเขียนได้เป็นฟังก์ชั่นใหม่คือ

$$f(\underline{X}_i,\underline{\beta}) \approx f_i^{(0)} + \sum_{j=1}^{p} D_{ij}^{(0)} \gamma_j^{(0)} \quad \textbf{(2)}$$

ที่มี $f_i^{(0)} = f(\underline{X}_i,\underline{b}^{(0)})$, $\gamma_j^{(0)} = \beta_j - b_j^{(0)}$

และ $$D_{ij}^{(0)} = \frac{\partial f(\underline{X}_i,\underline{\beta})}{\partial \beta_j}\bigg|_{\underline{\beta}=\underline{b}^{(0)}}$$

จาก (2) เขียนเป็นตัวแบบการถดถอยเชิงเส้นตรงได้โดย

$$Y_i \approx f_i^{(0)} + \sum_{j=1}^{p} D_{ij}^{(0)} \gamma_j^{(0)} + \varepsilon_i \quad \textbf{(3)}$$

ลบ $f_i^{(0)}$ ออกจากทั้งสองข้างของ (3) จะได้

$$Y_i^{(0)} \approx \sum_{j=1}^{p} D_{ij}^{(0)} \gamma_j^{(0)} + \varepsilon_i \quad \textbf{(4)}$$

ซึ่ง $Y_i^{(0)} = Y_i - f_i^{(0)}$ โดย $Y_i^{(0)}$ เป็นค่าความคลาดเคลื่อนในรอบที่ 0 เขียนตัวแบบประมาณในเทอมของเมตริกซ์ได้เป็น

$$\underline{Y}^{(0)} = D^{(0)} \underline{\gamma}^{(0)} + \underline{\varepsilon} \quad \textbf{(5)}$$

เมื่อ

$$\underline{Y}^{(0)} = \begin{bmatrix} Y_1 - f_1^{(0)} \\ \vdots \\ Y_n - f_n^{(0)} \end{bmatrix}, \quad \underline{\gamma}^{(0)} = \begin{bmatrix} \gamma_1^{(0)} \\ \vdots \\ \gamma_p^{(0)} \end{bmatrix} \text{ และ } \quad D_{nxp}^{(0)} = \begin{bmatrix} D_{11}^{(0)} \cdots D_{1p}^{(0)} \\ \vdots \\ D_{n1}^{(0)} \cdots D_{np}^{(0)} \end{bmatrix}$$

เรียกตัวแบบที่ (5) ว่าตัวแบบการประมาณ (approximate model) ด้วยวิธีกำลังสองน้อยที่สุด หา $\underline{a}^{(0)}$ ซึ่งเป็นเวกเตอร์ของตัวประมาณของเวกเตอร์ของพารามิเตอร์ $\underline{\gamma}^{(0)}$ ได้โดย $\underline{a}^{(0)} = (D'^{(0)} D^{(0)})^{-1} D'^{(0)} \underline{Y}^{(0)}$ เนื่องจาก $\underline{\gamma}^{(0)} = \underline{\beta} - \underline{b}^{(0)}$ ในรอบที่ 1 ซึ่งเป็นเวกเตอร์ของตัวประมาณ b_j ในรอบที่ 1 จาก $\underline{a}^{(0)} = \underline{b}^{(1)} - \underline{b}^{(0)}$ ได้ $\underline{b}^{(1)} = \underline{b}^{(0)} + \underline{a}^{(0)}$ เขียนผลรวมกำลังสองของความคลาดเคลื่อนในรอบที่ 0 และรอบที่ 1 ในเทอมของ $\underline{b}^{(0)}$ และ $\underline{b}^{(1)}$ เป็น $SSE^{(0)}$ และ $SSE^{(1)}$ ตามลำดับ โดย $SSE^{(0)} = \sum (Y_i - f_i^{(0)})^2$ และ $SSE^{(1)} = \sum (Y_i - f_i^{(1)})^2$ ในรอบที่ 2 หา $\underline{b}^{(2)}$ ซึ่งเป็นเวกเตอร์ของตัวประมาณ b_j ในรอบที่ 2 ด้วยวิธีการทำนองเดียวกันได้ $\underline{b}^{(2)} = \underline{b}^{(1)} + \underline{a}^{(1)}$ การประมาณจะทำต่อไปจนถึงรอบที่ S+1 ซึ่งเป็นรอบที่ให้ผลต่างของค่าประมาณของพารามิเตอร์ β_j ได้แก่ $(b_j^{(s+1)} - b_j^{(s)})$ หรือผลต่างของค่า SSE ได้แก่ $(SSE^{(s+1)} - SSE^{(s)})$ ในรอบที่ s+1 และรอบที่ s ลู่เข้าสู่ค่าคงที่น้อยที่สุด ดังนั้น $b_j^{(s+1)}$ เป็นค่าประมาณของพารามิเตอร์ β_j

Raymond [5] ได้กล่าวว่ากระบวนการวนซ้ำจะหยุดการทำงานเมื่อ

$$\left|\frac{b_j^{s+1} - b_j^s}{b_j^s}\right| < 10^{-6},\ j = 1, 2,...,p$$

3.2 วิธี Marquardt

พิจารณาตัวแบบการถดถอย $Y_i = f(\underline{X}_i,\underline{\beta}) + \varepsilon_i$ $i = 1, 2,...,n$ เมื่อ Y_i เป็นค่าสังเกตที่ i ของตัวแปรตาม Y $\underline{X}'_i$ เป็นเวกเตอร์ของค่าสังเกตของตัวแปรอิสระ q ตัวแปร ขนาด q จากหน่วยตัวอย่างที่ i ที่มี $\underline{X}'_i = (x_{1i}.x_{2i}...x_{qi})$ และ $\underline{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์ขนาด p ที่มี $\underline{\beta} = (\beta_1.\beta_2...\beta_p)$ $\varepsilon_i \sim$ i.i.d. $N(0,\sigma^2)$ และ $f(\underline{X}_i,\underline{\beta})$ เป็นค่าความคาดหวังของ Y_i ค่าผลต่างกำลังสองของความคลาดเคลื่อน (residual sum of squares) คือ

$$S(\underline{\beta}) = \sum_{i=1}^{n} \left(Y_i - f(\underline{X}_i,\underline{\beta})\right)^2$$

ให้ $\underline{b}_0 = (b_{10}, b_{20},...b_{k0})'$ เป็นเวกเตอร์ของค่าเริ่มต้น ซึ่งกระบวนการวนซ้ำในการประมาณค่าพารามิเตอร์สามารถกำหนดได้โดย

$$(\underline{D}'\underline{D} + \lambda I_p)\hat{\underline{\theta}} = \underline{D}'(\underline{y} - \underline{f}) \qquad \textbf{(6)}$$

เมื่อ $\underline{D}' = \dfrac{\partial f(\underline{x}_i,\underline{\beta})}{\partial \underline{\beta}}$

การคำนวณ vector of increments ที่ k iteration อยู่ในรูปของ

$$(\underline{D}'_k\underline{D}_k + \lambda I_p)\hat{\underline{\theta}}_k = \underline{D}'_k(\underline{y} - \underline{f}_k) \qquad \textbf{(7)}$$

เมื่อ $\hat{\underline{\theta}}_k = \underline{b}_{k+1} - \underline{b}_k$

$\underline{b}_{k+1} = \underline{b}_k + (\underline{D}'_k\underline{D}_k)^{-1}\underline{D}'_k(\underline{y} - \underline{f}_k)$

จากสมการ (7) เป็นการหาค่า λ ที่ทำให้ residual sum of square ลดลงกล่าวคือ $S(\underline{b}_{k+1}) < S(\underline{b}_k)$ λ เป็นค่าคงที่ ซึ่งมีค่าไม่น้อยกว่า 0 และ I คือ เมตริกซ์เอกลักษณ์ (identity) [5]

3.3 วิธี Newton

จากตัวแบบการถดถอย $Y_i = f(\underline{X}_i,\underline{\beta}) + \varepsilon_i$ $i = 1, 2,...,n$ Y_i เป็นค่าสังเกตที่ i ของตัวแปรตาม Y $\underline{X}'_i$ เป็นเวกเตอร์ของค่าสังเกตของตัวแปรอิสระ q ตัวแปร ขนาด q จากหน่วยตัวอย่างที่ i ที่มี $\underline{X}'_i = (x_{1i}.x_{2i}...x_{qi})$ $\underline{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์หรือค่าสัมประสิทธิ์การถดถอยขนาด p ที่มี $\underline{\beta}=(\beta_1.\beta_2...\beta_p)$ $\varepsilon_i \sim$ i.i.d. $N(0,\sigma^2)$

$$S(\underline{\beta}) = \sum_{i=1}^{n} \left(Y_i - f(\underline{X}_i,\underline{\beta})\right)^2$$

ค่าประมาณพารามิเตอร์ $\hat{\beta}$ โดยกระบวนการวนซ้ำสามารถคำนวณได้ดังนี้

$$\hat{\underline{\beta}}_{k+1} = \hat{\underline{\beta}}_k - H^{-1}(r^2, \hat{\underline{\beta}}_k)\, J(r^2, \hat{\underline{\beta}}_k) \qquad \textbf{(8)}$$

เมื่อ $r(\underline{\beta}) = \left(y_i - f(x_i,\underline{\beta})\right)$ แทน เวกเตอร์ของความคลาดเคลื่อน

$J(f,\underline{\beta}) = \frac{\partial f(\underline{\beta})}{\partial \underline{\beta}}$ คือ Jacobean matrix ของ function $f(\beta)$

$H(f,\underline{\beta}) = \frac{\partial f^2(\underline{\beta})}{\partial \underline{\beta}\ \partial \underline{\beta}'}$ คือ Hessian matrix

สมการที่ (8) จะทำการวนซ้ำต่อไปถ้า

$$r(\hat{\underline{\beta}}_{k+1})' r(\hat{\underline{\beta}}_{k+1}) < r(\hat{\underline{\beta}}_k)' r(\hat{\underline{\beta}}_k) \quad \text{หรือ} \quad S(\hat{\underline{\beta}}_{k+1}) < S(\hat{\underline{\beta}}_k)$$

นั่นเอง [6]

3.4 การกำหนดค่าเริ่มต้นด้วยวิธีของ Fekedulegn

Fekedulegn [3] ได้เสนอวิธีเลือกค่าเริ่มต้นที่ตัวแบบอยู่ในกลุ่มของรูปแบบการเจริญเติบโต โดยพิจารณาจากคุณสมบัติหรือลักษณะทั่วไปของตัวแบบ ซึ่งการกำหนดค่าเริ่มต้นจะเริ่มจากการกำหนดค่าเริ่มต้นของ β_0, β_2, β_3 และ β_1 ตามลำดับ

ค่าเริ่มต้นของ β_0 จะพิจารณาจากค่าสูงสุดของตัวแปรตาม Y

ค่าเริ่มต้นของ β_2 จะพิจารณาจาก

$$\beta_{2s} = \frac{(Y_2 - Y_1)/(X_2 - X_1)}{\beta_{0s}}$$

เมื่อ Y_1 และ Y_2 เป็นค่าของตัวแปรตามที่สอดคล้องกับ X_1 และ X_2 (ซึ่งจะพิจารณาที่ช่วงใดของข้อมูลก็ได้) และ β_{0s} แทนค่าเริ่มต้นของ β_0

ค่าเริ่มต้นของ β_3 มีค่าอยู่ระหว่าง 0 ถึง 1 ($0 < \beta_3 < 1$) สำหรับตัวแบบที่อยู่ในรูปของ Chapman-Richards growth model และ β_3 มีค่าเป็นบวก ($\beta_3 > 0$) สำหรับ Von Bertalanffy growth model และ Weibull growth model

ค่าเริ่มต้นของ β_1 เป็นค่าเริ่มต้นของตัวแบบที่เรียกว่า start of growth ซึ่งหมายความว่าเป็นค่าเริ่มต้นที่ตัวแปรอิสระ X มีค่าเท่ากับ 0 หรืออาจกล่าวได้ว่า Y(0) คือค่าของตัวแปรตาม Y ที่จุด start of growth ซึ่งในความเป็นจริงแล้วคือ 0 แต่ในทางปฏิบัติจะเลือกเป็นค่าบวกที่มีขนาดเล็ก

3.5 การกำหนดค่าเริ่มต้นโดยใช้ grid search

การกำหนดค่าเริ่มต้นในการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้นในกรณีที่ไม่ทราบว่าจะกำหนดค่าเริ่มต้นเป็นอะไรสามารถทำได้โดยใช้ grid search ซึ่งจะเป็น option หนึ่งใน SAS program ตัวอย่างการใช้ grid search เช่น กำหนดตัวแบบการถดถอยไม่เชิงเส้น

อยู่ในรูป $Y = e^{-\theta x} + \varepsilon$ สามารถประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้นโดยให้ SAS program ค้นหาค่าเริ่มต้นที่เหมาะสมได้โดยใช้คำสั่ง

```
PROC NLIN;
PARAMETERS G = -100 TO 100 BY .1;
MODEL Y = EXP(-(G*x));
RUN;
```

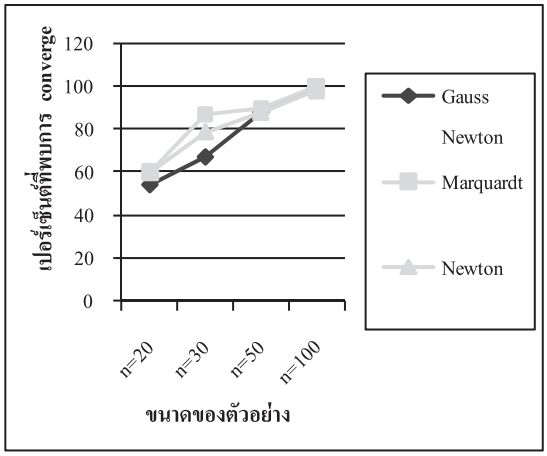
รูปที่ 1 แสดงการใช้คำสั่ง grid search ใน SAS program

SAS program จะค้นหาค่าเริ่มต้นที่เหมาะสมของพารามิเตอร์ให้ โดยจะค้นหาค่าเริ่มต้นของพารามิเตอร์ที่เหมาะสมตั้งแต่ -100 โดยเพิ่มทีละ 0.1 จนกระทั่งถึง 100 โดย G ใน SAS program แทนค่าพารามิเตอร์ θ x แทนตัวแปรอิสระซึ่งในตระกูลของตัวแบบการเจริญเติบโตจะแทนด้วยเวลา

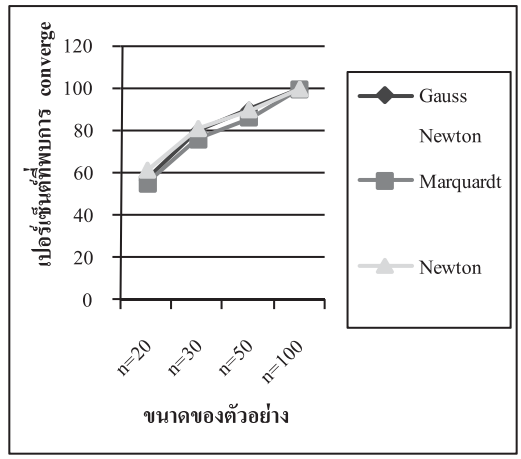
4. ผลการวิจัย

ส่วนที่ 1 พิจารณาเปอร์เซ็นต์ที่ลู่เข้าสู่ค่าพารามิเตอร์จากกระบวนการทำซ้ำ 500 รอบ กรณีเลือกค่าเริ่มต้นด้วยวิธีที่แตกกัน และขนาดตัวอย่างแตกต่างกัน

1.1 เปอร์เซ็นต์ที่พบการ converge ของตัวแบบ Negative Exponential เมื่อเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการใช้ Grid search แสดงดังรูปที่ 2 และ 3



รูปที่ 2 เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn



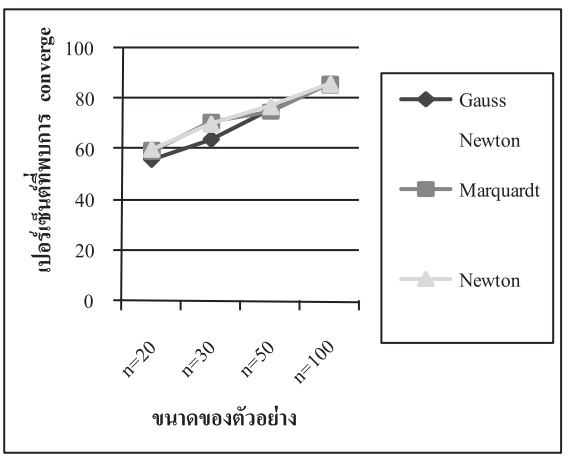
รูปที่ 3 เลือกค่าเริ่มต้นโดยใช้ Grid search

จากรูปที่ 2 และ 3 แสดงให้เห็นว่าใน รูปแบบ Negative Exponential ทั้งวิธีการเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการเลือกค่าเริ่มต้นโดยใช้ grid search จะให้ผลในลักษณะเช่นเดียวกัน คือ เมื่อขนาดของตัวอย่างเพิ่มขึ้นจะทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้นโดย

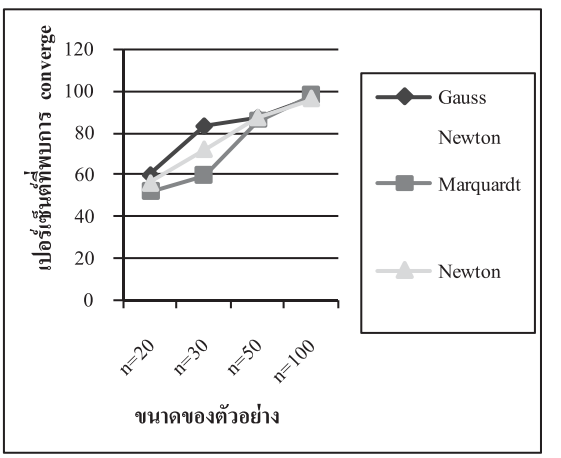
1.1.1 กรณี รูปแบบ Negative Exponential ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn เมื่อ n = 100 พบว่าวิธีที่ดีที่สุดคือวิธี Marquardt และวิธี Newton โดยมีเปอร์เซ็นต์ที่พบการ converge 100 เปอร์เซ็นต์ ทั้ง 2 วิธี

1.1.2 กรณี รูปแบบ Negative Exponential ที่เลือกค่าเริ่มต้นโดยใช้ Grid search เมื่อ n = 100 พบว่าวิธี Gauss Newton, วิธี Marquardt และวิธี Newton มีเปอร์เซ็นต์ที่พบการ converge เท่ากับ 100 เปอร์เซ็นต์ทั้ง 3 วิธี

1.2 เปอร์เซ็นต์ที่พบการ converge ของตัวแบบ Monomolecular เมื่อเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการใช้ Grid search แสดงดังรูปที่ 4 และ 5



รูปที่ 4 เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn



รูปที่ 5 เลือกค่าเริ่มต้นโดยใช้ Grid search

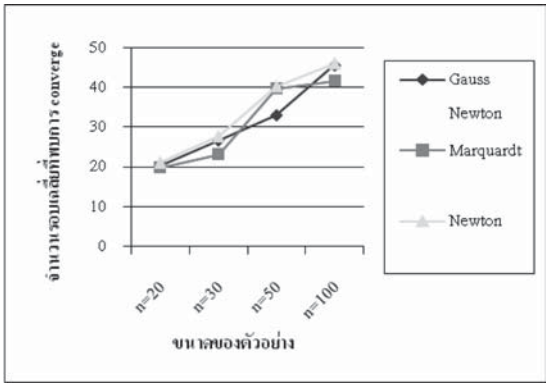
จากรูปที่ 4 และ 5 แสดงให้เห็นว่าในรูปแบบ Monomolecular ทั้งวิธีการเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการเลือกค่าเริ่มต้นโดยใช้ grid search ให้ผลในลักษณะเช่นเดียวกัน ทั้งวิธี Gauss Newton, วิธี Marquardt และวิธี Newton คือ เมื่อขนาดของตัวอย่างเพิ่มขึ้นจะทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้น โดย

1.2.1 กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn พบว่า เมื่อ n = 100 ทั้งวิธี Gauss Newton, วิธี Marquardt และวิธี Newton มีเปอร์เซ็นต์ที่พบการ converge ไม่แตกต่างกันคือเท่ากับ 86 ทั้ง 3 วิธี

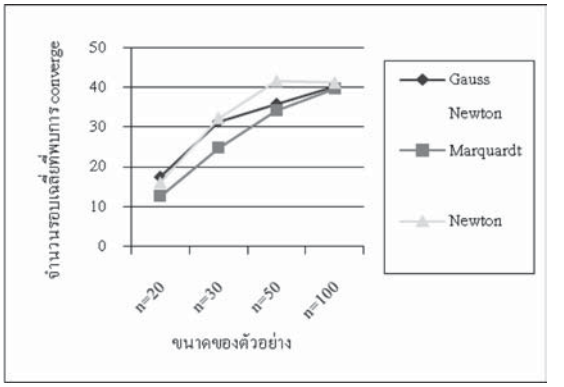
1.2.2 กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยใช้ Grid search พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือวิธี Marquardt มีเปอร์เซ็นต์ที่พบการ 98 เปอร์เซ็นต์

ส่วนที่ 2 พิจารณาจำนวนรอบเฉลี่ยที่ลู่เข้าสู่ค่าพารามิเตอร์ จากกระบวนการทำซ้ำ 500 รอบ กรณีเลือกค่าเริ่มต้นด้วยวิธีที่แตกกัน และขนาดตัวอย่างแตกต่างกัน

2.1 จำนวนรอบเฉลี่ยที่พบการ converge ของตัวแบบ Negative Exponential เมื่อเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการใช้ Grid search แสดงดังรูปที่ 6 และ 7



รูปที่ 6 เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn



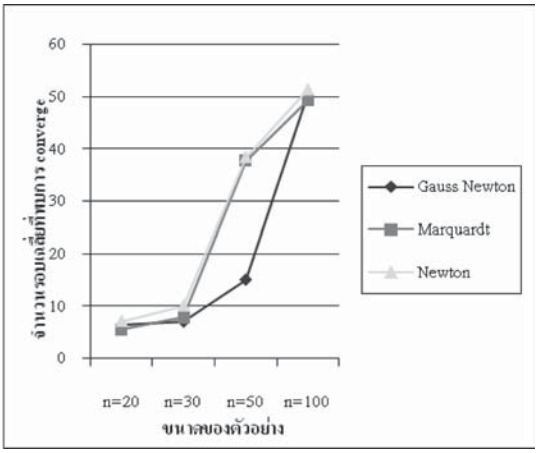
รูปที่ 7 เลือกค่าเริ่มต้นโดยใช้ Grid search

จากรูปที่ 6 และ 7 แสดงให้เห็นว่าในรูปแบบ Negative Exponential ทั้งวิธีการเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการเลือกค่าเริ่มต้นโดยใช้ grid search ให้ผลในลักษณะเช่นเดียวกัน ทั้งวิธี Gauss Newton, วิธี Marquardt และวิธี Newton คือ เมื่อขนาดของตัวอย่างเพิ่มขึ้นจะทำให้จำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้น โดย

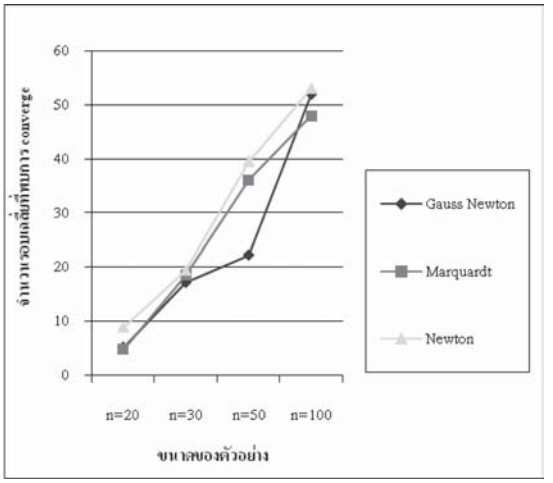
2.1.1 กรณี รูปแบบ Negative Exponential ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือวิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดเท่ากับ 41.49 รอบ

2.1.2 กรณี รูปแบบ Negative Exponential ที่เลือกค่าเริ่มต้นโดยใช้ Grid search พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือวิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดท่ากับ 39.58 รอบ

2.2 จำนวนรอบเฉลี่ยที่พบการ converge ของตัวแบบ Monomolecular เมื่อเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการใช้ Grid search แสดงดังรูปที่ 8 และ 9



รูปที่ 8 เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn



รูปที่ 9 เลือกค่าเริ่มต้นโดยใช้ Grid search

จากรูปที่ 8 และ 9 แสดงให้เห็นว่าในรูปแบบ Monomolecular ทั้งวิธีการเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการเลือกค่าเริ่มต้นโดยใช้ grid search ให้ผลในลักษณะเช่นเดียวกัน ทั้งวิธี Gauss Newton, วิธี Marquardt และวิธี Newton คือ เมื่อขนาดของตัวอย่างเพิ่มขึ้นจะทำให้จำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้น โดย

2.2.1 กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือ วิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดเท่ากับ 48.34 รอบ

2.2.2 กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยใช้ Grid search พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือ วิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดเท่ากับ 47.98 รอบ

5. สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

1. รูปแบบ Negative Exponential

พิจารณาจากเปอร์เซ็นต์ที่พบการ converge

รูปแบบ Negative Exponential ทั้งวิธีการเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn และการเลือกค่าเริ่มต้นโดยใช้ grid search จะให้ผลในลักษณะเช่นเดียวกัน คือ เมื่อขนาดของตัวอย่างเพิ่มขึ้นจะทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้นโดยกรณีที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn เมื่อ n = 100 พบว่าวิธีที่ดีที่สุดคือวิธี Marquardt และวิธี Newton โดยมีเปอร์เซ็นต์ที่พบการ converge 100 เปอร์เซ็นต์ ทั้ง 2 วิธี กรณีที่เลือกค่าเริ่มต้นโดยใช้ Grid search เมื่อ n = 100 พบว่าวิธี Gauss Newton, วิธี Marquardt และวิธี Newton มีเปอร์เซ็นต์ที่พบการ converge เท่ากับ 100 เปอร์เซ็นต์ทั้ง 3 วิธี

พิจารณาจากจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์

กรณี รูปแบบ Negative Exponential ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือวิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดเท่ากับ 41.49 รอบ กรณี รูปแบบ Negative Exponential ที่เลือกค่าเริ่มต้นโดยใช้ Grid search พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือวิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดท่ากับ 39.58 รอบ

2. รูปแบบ Monomolecular

พิจารณาจากเปอร์เซ็นต์ที่พบการ converge

กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn พบว่า เมื่อ n = 100 ทั้งวิธี Gauss Newton, วิธี Marquardt และวิธี Newton มีเปอร์เซ็นต์ที่พบการ converge ไม่แตกต่างกันคือเท่ากับ

86 ทั้ง 3 วิธี กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยใช้ Grid search พบว่าเมื่อ n = 100 วิธีที่ดีที่สุดคือวิธี Marquardt มีเปอร์เซ็นต์ที่พบการ 98 เปอร์เซ็นต์

พิจารณาจากจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์

กรณีรูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn พบว่า เมื่อ n = 100 วิธีที่ดีที่สุดคือ วิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดเท่ากับ 48.34 รอบ

กรณี รูปแบบ Monomolecular ที่เลือกค่าเริ่มต้นโดยใช้ Grid search พบว่า เมื่อ n = 100 วิธีที่ดีที่สุด วิธี Marquardt มีจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์น้อยที่สุดเท่ากับ 47.98 รอบ

5.2 อภิปรายผลการวิจัย

จากการศึกษาสถานการณ์การลู่เข้าสู่ค่าพารามิเตอร์ในการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้นด้วยวิธี Gauss-Newton, Marquardt และ Newton ในสถานการณ์ต่างๆ โดยปัจจัยที่นำมาพิจารณาในงานวิจัยนี้ประกอบไปด้วย 3 ปัจจัย ได้แก่ 1) การเลือกค่าเริ่มต้น 2) ลักษณะการแจกแจงของข้อมูล และ 3) ขนาดของตัวอย่าง พบว่าปัจจัยด้านการเลือกค่าเริ่มต้นนั้น มีผลต่อเปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์และจำนวนรอบเฉลี่ยที่พบการ converge เข้าสู่ค่าพารามิเตอร์ ซึ่งสอดคล้องกับการศึกษาของ Jusic [7] และ Fekedulegn [3] ที่พบว่าค่าเริ่มต้นที่ดีมีผลต่อเปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์ โดยได้มีการศึกษาเกี่ยวกับการเลือกค่าเริ่มต้นที่เหมาะสมด้วยวิธีการต่างๆ โดย Jusic [7] ได้ศึกษาเกี่ยวกับการเลือกค่าเริ่มต้นที่เหมาะสมของตัวแบบ Exponential ระหว่างการแปลงให้อยู่ในรูปแบบเชิงเส้นและไม่ได้แปลงให้อยู่ในรูปแบบเชิงเส้น ในขณะที่ Fekedulegn [3] ได้ศึกษาเกี่ยวกับการเลือกค่าเริ่มต้นที่เหมาะสมของตัวแบบการเจริญเติบโตซึ่งใช้วิธีการเลือกค่าเริ่มต้นโดยพิจารณาจากลักษณะทั่วไปหรือธรรมชาติของตัวแบบกับการเลือกค่าเริ่มต้นโดยการคาดคะเนหรือการเดา (guess) โดยที่งานวิจัยของ Jusic [7] และ Fekedulegn [3] ไม่ได้พิจารณาที่ขนาดของตัวอย่าง ซึ่งในงานวิจัยในครั้งนี้ได้นำปัจจัยด้านขนาดของตัวอย่างเข้ามาพิจารณาร่วมด้วยพบว่า การเลือกค่าเริ่มต้นด้วยวิธีการที่แตกต่างกันคือเลือกค่าเริ่มต้นโดยวิธีของ Fekedulegn [3] และการใช้ Grid search จะทำให้การประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้นด้วยวิธี Gauss-Newton, Marquardt และ Newton มีประสิทธิภาพที่ไม่แตกต่างกันเมื่อขนาดของตัวอย่างเพิ่มขึ้นจนถึงระดับหนึ่ง (ซึ่งในที่นี้คือ n=100) ซึ่งจะทำให้เปอร์เซ็นต์ที่พบการ converge เข้าสู่ค่าพารามิเตอร์เกือบเท่ากับ 100 เปอร์เซ็นต์ แต่ในขณะเดียวกันจะส่งผลให้จำนวนรอบเฉลี่ยในการ converge เข้าสู่ค่าพารามิเตอร์เพิ่มขึ้นด้วย ซึ่งจะเพิ่มขึ้นมากน้อยเท่าไรขึ้นอยู่กับลักษณะการแจกแจงของข้อมูลหรือตัวแบบที่นำมาพิจารณา

5.3 ข้อเสนอแนะ

เพื่อเป็นแนวทางในการศึกษาถึงสถานการณ์การลู่เข้าสู่ค่าพารามิเตอร์ ในการประมาณค่าพารามิเตอร์ของตัวแบบการถดถอยไม่เชิงเส้น ผู้วิจัยจึงขอเสนอแนะแนวทางในการศึกษาต่อไปดังนี้

- ควรพิจารณาการเลือกค่าเริ่มต้นที่เหมาะสมในลักษณะอื่นๆ เช่น วิธี Linearization เป็นการกำหนดค่าเริ่มต้นในกรณีที่ตัวแบบสามารถแปลงให้อยู่ในรูปแบบเชิงเส้นได้ การแก้ระบบสมการ (solving a system of equations) วิธี graphical เป็นต้น

- เนื่องจากทราบว่าเมื่อค่าเริ่มต้นที่เลือกใกล้เคียงกับ final solution จะทำให้จำนวนรอบที่ใช้ในการ converge ลดลง ดังนั้นจึงควรหาวิธีในการกำหนดค่าเริ่มต้นให้ใกล้เคียงกับ final solution

6. เอกสารอ้างอิง

1. Seber, G.A.F., and Wild, C.J., 1989, *Nonlinear Regression*, John Wiley & Sons, Inc., New York.

2. Jukic, D., and Scitovski, R., 1997, “Existence of Optimal Solution for Exponential Model by Least Square”, *Computational and Applied Mathematics*, Vol. 78, pp. 317 – 328.

3. Fekedulegn, D., Mac Siurtain, M.P., and Colbert, J.J., 1999, “Parameter Estimation of Nonlinear Growth Models in Forestry”, *Silva Fennica*, Vol. 33, No. 4,

pp. 327 – 336.

4. Bates, D.M. , and Watts, D.G., 1988, *Nonlinear Regression Analysis and It Applications*, John Wiley & Sons, Inc., New York.

5. Raymond, H. Myers, Douglas C. Montgomery and G. Geoffrey Vining., 2002, *Generalized Linear Model*, John Wiley & Sons, Inc., New York, pp.

6. SAS Institute Inc, 2003., Cary, NC, USA.

7. Jukic D., 1995, “The Problem of Initial Approximation for a Special Nonlinear Least Squares Problem”, *Appl. Math. Comput*, Vol. 84, pp. 25 - 32.