บทที่ 2

ทฤษฎีและตัวสถิติที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณทั้ง 4 วิธี ได้แก่ วิธีกำลังสองน้อยสุด (Ordinary Least Squares (OLS) method) วิธีกำลังน้อยสุดที่ถูกจำกัด (Restricted Least Squares (RLS) method) วิธีริดจ์รีเกรสชันที่ถูกจำกัด (Restricted Ridge Regression (RRR) method) วิธีลิวที่ถูกจำกัด (Restricted Liu (RL) method) ซึ่งรายละเอียดต่าง ๆ มีดังนี้

1. การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุด (OLS)

"สมมติว่าเราต้องการประมาณ $\underset{\sim}{y} \in \Re^n$ โดยเวกเตอร์ของการประมาณซึ่งอยู่ในรูปของ $\underset{\sim}{\hat{y}}_1 = X\underset{\sim}{\beta}$ เมื่อ $X_{n\times(p+1)}$ เป็นเมทริกซ์คงที่และ $\underset{\sim}{\beta}$ เป็นเวกเตอร์ของพารามิเตอร์ซึ่งจะทำให้ค่าของระยะทางกำลังสองมีค่าน้อยที่สุด กล่าวคือ เลือกค่า $\underset{\sim}{\beta}$ ซึ่งทำให้

$$
\begin{aligned}
(1.1)\qquad \{d(\underset{\sim}{y},\underset{\sim}{\hat{y}}_1)\}^2 &= (\underset{\sim}{y}-\underset{\sim}{\hat{y}}_1)'(\underset{\sim}{y}-\underset{\sim}{\hat{y}}_1)\\
&= (\underset{\sim}{y}-X\underset{\sim}{\beta})'(\underset{\sim}{y}-X\underset{\sim}{\beta})\\
&= f(\underset{\sim}{\beta})
\end{aligned}
$$

มีค่าน้อยที่สุด ซึ่งค่าของ $\underset{\sim}{\beta}$ ดังกล่าวจะได้จากการหาอนุพันธ์รวม (total differential) แล้วกำหนดให้เท่ากับ 0

$$
\begin{aligned}
df &= (-X\,d\underset{\sim}{\beta})'(\underset{\sim}{y}-X\underset{\sim}{\beta}) + (\underset{\sim}{y}-X\underset{\sim}{\beta})'(-X\,d\underset{\sim}{\beta})\\
&= -2(\underset{\sim}{y}-X\underset{\sim}{\beta})'X\,d\underset{\sim}{\beta}
\end{aligned}
$$

ค่า $df = 0$ ณ ค่า $\underset{\sim}{\hat{\beta}}$ ของ $\underset{\sim}{\beta}$ จะสอดคล้องกับสมการ

$$(1.2)\qquad (\underset{\sim}{y}-X\underset{\sim}{\beta})'X = \underset{\sim}{0}'$$

กล่าวคือ $$X'X\,\hat{\underset{\sim}{\beta}} = X'\,\underset{\sim}{y}$$

ซึ่งเราเรียกสมการนี้ว่าสมการปกติ (normal equation) สำหรับค่า $\hat{\underset{\sim}{\beta}}$ "[1] และจะหาตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณได้ดังนี้

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

ตัวประมาณในสมการ (1.3) มีคุณสมบัติเป็นตัวประมาณไม่เอนเอียงของ $\underset{\sim}{\beta}$ กล่าวคือ

$$\begin{aligned}
E(\hat{\underset{\sim}{\beta}}) &= E[(X'X)^{-1}X'\,\underset{\sim}{y}] \\
&= E[(X'X)^{-1}X'(X\,\underset{\sim}{\beta}+\underset{\sim}{\varepsilon})] \\
&= E[(X'X)^{-1}X'X\,\underset{\sim}{\beta}+(X'X)^{-1}X'\,\underset{\sim}{\varepsilon})] \\
&= E(\underset{\sim}{\beta})+(X'X)^{-1}X'\,E(\underset{\sim}{\varepsilon}) \\
&= \underset{\sim}{\beta}
\end{aligned}$$

ตัวประมาณ $\hat{\underset{\sim}{\beta}}$ ในสมการ (1.3) ที่ได้จากการประมาณค่าด้วยวิธีกำลังสองน้อยสุดเป็นตัวประมาณไม่เอนเอียงซึ่งมีความแปรปรวนต่ำสุดในบรรดาตัวประมาณไม่เอนเอียง แต่ในการประมาณค่า $\hat{\underset{\sim}{\beta}}$ ด้วยวิธีกำลังสองน้อยสุดมีข้อสมมติที่จำเป็นข้อหนึ่งคือ ตัวแปรอิสระแต่ละตัวต้องไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระตัวอื่น ซึ่งในทางปฏิบัติเป็นไปได้น้อยมากเพราะตัวแปรอิสระต่าง ๆ ที่นำมาศึกษาส่วนใหญ่อาจมีความสัมพันธ์กัน คือ มีสภาพไม่เหมาะสม (ill – conditioned) ซึ่งอาจมีผลทำให้การประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยสุดไม่ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำที่สุด เราสามารถพิจารณาผลของตัวแปรอิสระที่มีพหุสัมพันธ์ (multicollinearity) กันโดยพิจารณาจากคุณสมบัติ 2 ประการคือ เมทริกซ์ความแปรปรวนร่วมของตัวประมาณ $\hat{\underset{\sim}{\beta}}$ และค่าเฉลี่ยกำลังสองของความแตกต่างระหว่าง $\hat{\underset{\sim}{\beta}}$ และ $\underset{\sim}{\beta}$ ดังต่อไปนี้

---

[1] ธีระพร วีระถาวร, ตัวแบบเชิงเส้น ทฤษฎีและการประยุกต์ (กรุงเทพฯ: วิทยพัฒน์, 2541), หน้า 56-57.

(1.4) $$\operatorname{cov}(\hat{\underset{\sim}{\beta}}) = \sigma^2 (X'X)^{-1}$$

ให้ L_1 คือความแตกต่างระหว่าง $\hat{\underset{\sim}{\beta}}$ และ $\underset{\sim}{\beta}$ ดังนั้น

(1.5) $$L_1^2 = (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})$$

จะได้ค่าเฉลี่ยกำลังสองของความแตกต่างระหว่าง $\hat{\underset{\sim}{\beta}}$ และ $\underset{\sim}{\beta}$ ดังนี้

(1.6) $$E(L_1^2) = \sigma^2 tr(X'X)^{-1}$$

โดยที่ $E(L_1^2)$ จะมีค่าสมมูล (equivalent) กับ $E[(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})]$ ซึ่งทำให้สามารถเขียนได้ว่า

$$\begin{aligned} E(L_1^2) &= E[(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})] \\ &= E[\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}} - 2\hat{\underset{\sim}{\beta}}'\underset{\sim}{\beta} + \underset{\sim}{\beta}'\underset{\sim}{\beta}] \\ &= E(\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}) - \underset{\sim}{\beta}'\underset{\sim}{\beta} \end{aligned}$$

จากนั้นแทนค่า $E(L_1^2)$ ในสมการ (1.6) จะได้

$$E(\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}) = \underset{\sim}{\beta}'\underset{\sim}{\beta} + \sigma^2 tr(X'X)^{-1}$$

ถ้า $\underset{\sim}{\varepsilon}$ มีการแจกแจงปกติจะได้ว่า

(1.7) $$Var(L_1^2) = 2\sigma^4 tr(X'X)^{-2}$$

จากสมการ (1.4) , (1.6) และ (1.7) เราจะพบว่าค่า $\operatorname{cov}(\hat{\underset{\sim}{\beta}})$, $E(L_1^2)$ และ $Var(L_1^2)$ ต่างอยู่ในรูปฟังก์ชันผลบวกของสมาชิกในแนวทแยงมุมของเมทริกซ์ $X'X$ ดังนั้นเพื่อความสะดวกในการเปรียบเทียบค่า $\underset{\sim}{\hat{\beta}}$ จึงควรแปลงให้อยู่ในรูปฟังก์ชันของค่าลักษณะเฉพาะ (eigenvalue) ของ

เมทริกซ์ $X'X$ โดยใช้คุณสมบัติข้อหนึ่งของค่าลักษณะเฉพาะคือ ถ้า λ_i เป็นค่าลักษณะเฉพาะของเมทริกซ์ $X'X$ แล้ว $\sum_{i=1}^{p+1}\lambda_i = tr(X'X)$

กำหนดให้ค่า λ_i เป็นค่าลักษณะเฉพาะของเมทริกซ์ $X'X$ ซึ่ง $\lambda_{\max}$ เป็นค่าลักษณะเฉพาะที่มีค่ามากที่สุดและ $\lambda_{\min}$ เป็นค่าลักษณะเฉพาะที่มีค่าน้อยที่สุด ดังนั้น $(\lambda_{\max} = \lambda_1) \geq \lambda_2 \geq \ldots \geq (\lambda_{p+1} = \lambda_{\min}) > 0$

จากสมการ (1.6) และ (1.7) เราสามารถเขียนค่า $E(L_1^2)$ และ $Var(L_1^2)$ ให้อยู่ในรูปฟังก์ชันค่าลักษณะเฉพาะของเมทริกซ์ $X'X$ ได้ดังนี้

$$E(L_1^2) = \sigma^2 \sum_{i=1}^{p+1} \frac{1}{\lambda_i}$$

และ

$$Var(L_1^2) = 2\sigma^4 \sum_{i=1}^{p+1} \left(\frac{1}{\lambda_i}\right)^2$$

เนื่องจากในกรณีที่ตัวแปรอิสระมีสภาพไม่เหมาะสม กล่าวคือ ค่าลักษณะเฉพาะบางค่าของเมทริกซ์ $X'X$ จะมีค่าน้อยมาก ๆ ซึ่งมีผลทำให้ความแตกต่างระหว่าง $\underset{\sim}{\hat{\beta}}$ และ $\underset{\sim}{\beta}$ มีค่ามาก ทำให้การประมาณค่า $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยสุดจะให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่มีค่าสูงขึ้น ดังนั้นในกรณีที่เกิดปัญหาพหุสัมพันธ์กันตัวประมาณ $\underset{\sim}{\hat{\beta}}$ จึงเป็นตัวประมาณที่ไม่เหมาะสม

2. การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีกำลังสองน้อยสุดที่ถูกจำกัด (RLS)

เนื่องจากวิธีกำลังสองน้อยสุดเป็นวิธีที่นิยมใช้กันมาก แต่มีปัญหาในกรณีที่ตัวแปรอิสระมีพหุสัมพันธ์กันทำให้ความแปรปรวนมีค่าสูง ค่าสัมประสิทธิ์การถดถอยพหุคูณที่ได้ขาดความเที่ยงตรงจึงได้มีวิธีพื้นฐานที่สร้างขึ้นเพื่อที่จะปรับปรุงความเที่ยงตรงของตัวประมาณสัมประสิทธิ์กำลังสองน้อยสุดที่เรียกว่า**วิธีกำลังสองน้อยสุดที่ถูกจำกัด** หลักการของวิธีนี้ยังคงใช้หลักการของวิธีกำลังสองน้อยสุดเหมือนเดิมเพียงแต่เพิ่มข้อมูลข้อจำกัดเกี่ยวกับการรวมเชิงเส้นของสัมประสิทธิ์การถดถอยพหุคูณที่ไม่ทราบค่าดังนี้

$$R\underset{\sim}{\beta} = \underset{\sim}{r} \qquad (2.1)$$

เมื่อ R เป็นเมทริกซ์แสดงรูปแบบความสัมพันธ์ของพารามิเตอร์ $\underset{\sim}{\beta}$ ขนาด $q\times(p+1)$

โดยที่ $q < (p+1)$

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

และ $\underset{\sim}{r}$ เป็นเวกเตอร์ที่ได้จากการรวมเชิงเส้นของสัมประสิทธิ์การถดถอยพหุคูณขนาด $q\times 1$

จากตัวแบบในสมการ (2.1) ได้นำมาเป็นข้อจำกัดในการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณวิธีกำลังสองน้อยสุดที่ให้ผลรวมความคลาดเคลื่อนกำลังสองมีค่าต่ำสุด โดยมีรูปแบบดังนี้

$$\underset{\underset{\sim}{\beta}}{Minimize}\,(\underset{\sim}{y}-X\underset{\sim}{\beta})'(\underset{\sim}{y}-X\underset{\sim}{\beta})$$

ภายใต้ข้อจำกัด

$$R\underset{\sim}{\beta} = \underset{\sim}{r}$$

เมื่อมีข้อจำกัดเกิดขึ้นเราสามารถใช้เซตของตัวคูณลากรองจ์[2] (Lagrange multiplier) $\underset{\sim}{\lambda}(q\times 1)$ มาช่วยหาค่า $\underset{\sim}{\beta}$ ซึ่งทำให้ $g(\underset{\sim}{\beta})$ มีค่าน้อยที่สุดเมื่อ

$$\begin{aligned}
g(\underset{\sim}{\beta}) &= (\underset{\sim}{y}-X\underset{\sim}{\beta})'(\underset{\sim}{y}-X\underset{\sim}{\beta}) + 2\underset{\sim}{\lambda}'(R\underset{\sim}{\beta}-\underset{\sim}{r})\\
dg &= (-X\,d\underset{\sim}{\beta})'(\underset{\sim}{y}-X\underset{\sim}{\beta}) + (\underset{\sim}{y}-X\underset{\sim}{\beta})'(-X\,d\underset{\sim}{\beta})\\
&\quad + 2\underset{\sim}{\lambda}'R\,d\underset{\sim}{\beta} + 2\,d\underset{\sim}{\lambda}'(R\underset{\sim}{\beta}-\underset{\sim}{r})\\
&= -2\,d\underset{\sim}{\beta}'(X'\underset{\sim}{y}-X'X\underset{\sim}{\beta}-R'\underset{\sim}{\lambda}) + 2\,d\underset{\sim}{\lambda}'(R\underset{\sim}{\beta}-\underset{\sim}{r}) \qquad (2.2)
\end{aligned}$$

เราให้สมการ (2.2) = 0 และให้การเทียบสัมประสิทธิ์ จะได้ว่า

$$X'\underset{\sim}{y}-X'X\underset{\sim}{\beta}-R'\underset{\sim}{\lambda} = \underset{\sim}{0}$$

และ

$$R\underset{\sim}{\beta}-\underset{\sim}{r} = \underset{\sim}{0}$$

---

[2] ธีระพร วีระถาวร, ตัวแบบเชิงเส้น ทฤษฎีและการประยุกต์ (กรุงเทพฯ: วิทยพัฒน์, 2541), หน้า 210.

ค่าประมาณ $\underset{\sim}{\beta}^{*}$ ของ $\underset{\sim}{\beta}$ จะต้องสอดคล้องกับ

(2.3) $$X'X\,\underset{\sim}{\beta}^{*} + R'\,\underset{\sim}{\lambda} = X'\,\underset{\sim}{y}$$

(2.4) และ $$R\,\underset{\sim}{\beta}^{*} = \underset{\sim}{r}$$

ผู้อ่านจะเห็นได้ว่าสมการ (2.3) ซึ่งทำให้ได้ค่าประมาณ $\underset{\sim}{\beta}^{*}$ จะไม่เหมือนกับสมการปกติที่ไม่มีข้อจำกัด จากสมการ (2.3) จะได้ว่า

$$
\begin{aligned}
\underset{\sim}{\beta}^{*} &= (X'X)^{-1}(X'\,\underset{\sim}{y} - R'\,\underset{\sim}{\lambda}) \\
&= \underset{\sim}{\hat{\beta}} - (X'X)^{-1}R'\,\underset{\sim}{\lambda} \qquad (2.5)
\end{aligned}
$$

แทนค่า $\underset{\sim}{\beta}^{*}$ ในสมการ (2.4) จะได้ว่า

$$R\,\underset{\sim}{\hat{\beta}} - R(X'X)^{-1}R'\,\underset{\sim}{\lambda} = \underset{\sim}{r}$$

ดังนั้น $$\underset{\sim}{\lambda} = -[R(X'X)^{-1}R']^{-1}(\underset{\sim}{r} - R\,\underset{\sim}{\hat{\beta}})$$

แทนค่า $\underset{\sim}{\lambda}$ ในสมการ (2.5) จะได้ตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณวิธีกำลังสองน้อยสุดที่ถูกจำกัดอยู่ในรูปของ

$$
\begin{aligned}
\underset{\sim}{\beta}^{*} &= \underset{\sim}{\hat{\beta}} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(\underset{\sim}{r} - R\,\underset{\sim}{\hat{\beta}}) \\
&= \underset{\sim}{\hat{\beta}} + S^{-1}R'(RS^{-1}R')^{-1}(\underset{\sim}{r} - R\,\underset{\sim}{\hat{\beta}}) \qquad (2.6)
\end{aligned}
$$

เมื่อ $S = X'X$

พิจารณาค่าคาดหวังของตัวประมาณ $\underset{\sim}{\beta}^{*}$ จะได้ว่า

$$
\begin{aligned}
E(\underset{\sim}{\beta}^{*}) &= E[\underset{\sim}{\hat{\beta}} + S^{-1}R'(RS^{-1}R')^{-1}(\underset{\sim}{r} - R\,\underset{\sim}{\hat{\beta}})] \\
&= E(\underset{\sim}{\hat{\beta}}) + E[S^{-1}R'(RS^{-1}R')^{-1}(\underset{\sim}{r} - R\,\underset{\sim}{\hat{\beta}})]
\end{aligned}
$$

$$= \underset{\sim}{\beta} + S^{-1}R'(RS^{-1}R')^{-1}(\underset{\sim}{r} - R\underset{\sim}{\beta})$$
$$= \underset{\sim}{\beta} + S^{-1}R'[RS^{-1}R']^{-1}\underset{\sim}{\delta}$$

เมื่อ $\underset{\sim}{r} - R\underset{\sim}{\beta} = \underset{\sim}{\delta} = \underset{\sim}{0}$ เป็นกรณีที่ข้อจำกัดนั้นเป็นจริง ตัวประมาณ $\underset{\sim}{\beta}^*$ จะเป็นตัวประมาณไม่เอนเอียง (unbiased estimator) ส่วนในกรณีที่ข้อจำกัดนั้นไม่เป็นจริงคือ $\underset{\sim}{r} - R\underset{\sim}{\beta} = \underset{\sim}{\delta} \neq \underset{\sim}{0}$ ตัวประมาณ $\underset{\sim}{\beta}^*$ จะเป็นตัวประมาณเอนเอียง (biased estimator)

เมทริกซ์ความแปรปรวนร่วมของ $\underset{\sim}{\beta}^*$ สามารถเขียนได้เป็น

$$\begin{aligned}
\text{cov}(\underset{\sim}{\beta}^*) &= E[(\underset{\sim}{\beta}^* - E(\underset{\sim}{\beta}^*))(\underset{\sim}{\beta}^* - E(\underset{\sim}{\beta}^*))'] \\
&= E[(I - S^{-1}R'(RS^{-1}R')^{-1}R)(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'(I - R'(RS^{-1}R')^{-1}RS^{-1})] \\
&= \sigma^2(I - S^{-1}R'(RS^{-1}R')^{-1}R)S^{-1}(I - R'(RS^{-1}R')^{-1}RS^{-1}) \\
&= \sigma^2(S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\
&\quad + S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}R'(RS^{-1}R')^{-1}RS^{-1}) \\
&= \sigma^2(S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1})
\end{aligned}$$

ถ้าพิจารณาความแตกต่างของความแปรปรวนร่วมระหว่าง $\underset{\sim}{\beta}^*$ กับ $\hat{\underset{\sim}{\beta}}$ จะได้ว่า

$$\text{cov}(\hat{\underset{\sim}{\beta}}) - \text{cov}(\underset{\sim}{\beta}^*) = \sigma^2[S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}]$$

เมื่อ $S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$ เป็นเมทริกซ์กึ่งบวกแน่นอน (positive semidefinite matrix) ทำให้ตัวประมาณ $\underset{\sim}{\beta}^*$ ที่ได้จากวิธีกำลังสองน้อยสุดที่ถูกจำกัดมีความแปรปรวนน้อยกว่าตัวประมาณ $\hat{\underset{\sim}{\beta}}$ ที่ได้จากวิธีกำลังสองน้อยสุด

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\underset{\sim}{\beta}^*$ จะอยู่ในรูปของ

$$\begin{aligned}
MSE(\underset{\sim}{\beta}^*) &= tr(\text{cov}(\underset{\sim}{\beta}^*)) + [(E(\underset{\sim}{\beta}^*) - \underset{\sim}{\beta})'(E(\underset{\sim}{\beta}^*) - \underset{\sim}{\beta})] \\
&= \sigma^2 tr(S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}) \\
&\quad + [(S^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta})'(S^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta})]
\end{aligned}$$

$$= \sigma^2 tr(S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1})$$
$$+ [\underset{\sim}{\delta}'(RS^{-1}R')^{-1}RS^{-1}S^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}]$$
$$= \sigma^2 tr(A) + \underset{\sim}{\delta}'(RS^{-1}R')^{-1}RS^{-2}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}$$

(2.7) $$= \sigma^2 tr(A) + \underset{\sim}{\delta}^{*'} S^{-2} \underset{\sim}{\delta}^{*}$$

เมื่อ $A = S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$, $\underset{\sim}{\delta} = r - R\underset{\sim}{\beta}$

และ $\underset{\sim}{\delta}^{*} = R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}$

เมื่อพิจารณา $MSE(\underset{\sim}{\beta}^{*})$ จะพบว่าในกรณีที่ข้อจำกัดนั้นเป็นจริง $\underset{\sim}{\delta} = \underset{\sim}{0}$ จะได้ว่า

(2.8) $$MSE(\underset{\sim}{\beta}^{*}) = \sigma^2 tr(A)$$

3. **การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีริดจ์รีเกรสชันที่ถูกจำกัด (RRR)**

ในปี ค.ศ. 1992 ซาคาร์ (Sarkar) ได้เสนอวิธีริดจ์รีเกรสชันที่ถูกจำกัด โดยได้นำข้อดีของวิธีริดจ์รีเกรสชันและวิธีกำลังสองน้อยสุดที่ถูกจำกัดมาผสมผสานกันเพื่อแก้ปัญหาการเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ **หลักการของวิธีนี้คือ พยายามที่จะลดค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณ $\underset{\sim}{\beta}^{*}$ ให้ต่ำลงโดยใช้หลักการของวิธีริดจ์รีเกรสชัน** เมื่อพิจารณาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณ $\underset{\sim}{\beta}^{*}$ การที่จะลดค่าเฉลี่ยความคลาดเคลื่อนกำลังสองให้ต่ำลงจะต้องลดค่า $(X'X)^{-1}$ ให้ต่ำลง ซึ่งก็คือทำให้ $|X'X|$ มีค่าเพิ่มมากขึ้น โดยการบวกค่าคงที่ที่มากกว่าศูนย์กับสมาชิกทุกตัวบนเส้นทแยงมุมของเมทริกซ์ $X'X$ โดยที่เราสามารถเขียนสมการปกติของตัวประมาณด้วยวิธีริดจ์รีเกรสชันที่ถูกจำกัดได้เป็น

$$(X'X + kI)\underset{\sim}{\beta}_R^{*}(k) = X'\underset{\sim}{y} + Sg$$

เมื่อ $g = S^{-1}R'(RS^{-1}R')^{-1}(\underset{\sim}{r} - R\hat{\underset{\sim}{\beta}})$

ดังนั้นตัวประมาณ $\underset{\sim}{\beta}_R^{*}(k)$ จะอยู่ในรูปของ

(3.1) $$\underset{\sim}{\beta}_R^{*}(k) = (X'X + kI)^{-1}(X'\underset{\sim}{y} + Sg) \qquad , \quad k > 0$$

เราสามารถเขียน $\underset{\sim}{\beta}_R^*(k)$ ให้อยู่ในรูปฟังก์ชัน $\underset{\sim}{\beta}^*$ โดยแทนค่า $X'\underset{\sim}{y}+Sg = X'X\underset{\sim}{\beta}^*$

ดังนั้น
$$\begin{aligned}\underset{\sim}{\beta}_R^*(k) &= (X'X+kI)^{-1}X'X\underset{\sim}{\beta}^* \\ &= [I+k(X'X)^{-1}]^{-1}\underset{\sim}{\beta}^* \\ &= [I+kS^{-1}]^{-1}\underset{\sim}{\beta}^* \\ &= W\underset{\sim}{\beta}^*\end{aligned}$$

เมื่อ $W=[I+kS^{-1}]^{-1}$ และตัวประมาณ $\underset{\sim}{\beta}_R^*(k)=\underset{\sim}{\beta}^*$ ถ้า $k=0$

เมื่อพิจารณาค่าคาดหวังของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ จะได้ว่า

$$\begin{aligned}E(\underset{\sim}{\beta}_R^*(k)) &= E(W\underset{\sim}{\beta}^*) \\ &= WE(\underset{\sim}{\beta}^*) \\ &= W(\underset{\sim}{\beta}+S^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}) \\ &= W\underset{\sim}{\beta}+WS^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}\end{aligned}$$

จากค่าคาดหวังของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ จะพบว่าตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ เป็นตัวประมาณเอนเอียงสำหรับ $\underset{\sim}{\beta}$ ในกรณีที่ $k=0$ และ $\underset{\sim}{\delta}=\underset{\sim}{0}$ ตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ จะเป็นตัวประมาณไม่เอนเอียง และค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\underset{\sim}{\beta}_R^*(k)$ จะอยู่ในรูปของ

$$\begin{aligned}MSE(\underset{\sim}{\beta}_R^*(k)) &= E[(\underset{\sim}{\beta}_R^*-\underset{\sim}{\beta})'(\underset{\sim}{\beta}_R^*-\underset{\sim}{\beta})] \\ &= E[(\underset{\sim}{\hat\beta}-\underset{\sim}{\beta})'(I-S^{-1}R'(RS^{-1}R')^{-1}R)'W'W(I-S^{-1}R'(RS^{-1}R')^{-1}R)(\underset{\sim}{\hat\beta}-\underset{\sim}{\beta})] \\ &\quad +[(W\underset{\sim}{\beta}+WS^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}-\underset{\sim}{\beta})'(W\underset{\sim}{\beta}+WS^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}-\underset{\sim}{\beta})] \\ &= \sigma^2 tr(S^{-1})(I-S^{-1}R'(RS^{-1}R')^{-1}R)'W'W(I-S^{-1}R'(RS^{-1}R')^{-1}R) \\ &\quad +[(WS^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}+(W-I)\underset{\sim}{\beta})'(WS^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}+(W-I)\underset{\sim}{\beta})]\end{aligned}$$

(3.2) $$= \sigma^2 tr(WAW')+[WS^{-1}\underset{\sim}{\delta}^*-kS(k)^{-1}\underset{\sim}{\beta}]'[WS^{-1}\underset{\sim}{\delta}^*-kS(k)^{-1}\underset{\sim}{\beta}]$$

เมื่อ $\underset{\sim}{\delta}^*=R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}$

และ $S(k)^{-1}$ เป็นเมทริกซ์ผกผันของ $S(k)=(S+kI)$

ในกรณีที่ $\underset{\sim}{\delta} = \underset{\sim}{0}$ คือข้อจำกัดเป็นจริง $MSE(\underset{\sim}{\beta}^*_R(k))$ จะมีรูปแบบ

(3.3) $$MSE(\underset{\sim}{\beta}^*_R(k)) = \sigma^2 tr(WAW') + k^2\, \underset{\sim}{\beta}'\, S(k)^{-2}\, \underset{\sim}{\beta}$$

ตัวแบบทั่วไป $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ สามารถเขียนให้อยู่ในรูปแบบใหม่โดยใช้เมทริกซ์เชิงตั้งฉาก (orthogonal matrix) ได้ดังนี้

(3.4) $$\underset{\sim}{y} = Z\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$$

เมื่อ $Z = XQ$, $\underset{\sim}{\alpha} = Q'\underset{\sim}{\beta}$, $Z'Z = Q'X'XQ = \Lambda$, $QQ' = Q'Q = I$

$\underset{\sim}{y}$ เป็นเวกเตอร์ของตัวแปรตามขนาด n

X เป็นเมทริกซ์ของตัวแปรอิสระขนาด n×(p+1) และ X มีค่าลำดับชั้นเต็มเท่ากับ p+1

$\underset{\sim}{\beta}$ เป็นเวกเตอร์ของสัมประสิทธิ์การถดถอยขนาด p+1

$\underset{\sim}{\varepsilon}$ เป็นเวกเตอร์ของความคลาดเคลื่อนขนาด n โดยที่ $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$, $\text{cov}(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

Q เป็นเมทริกซ์เชิงตั้งฉากขนาด (p+1)×(p+1) โดยแต่ละแนวตั้งของ p+1 คือเวกเตอร์ลักษณะเฉพาะ (eigenvector) ที่สอดคล้องกับค่าลักษณะเฉพาะของ $X'X$ กล่าวคือ $Q = (q_1, q_2, ..., q_{p+1})$ เมื่อ $q_1, q_2, ..., q_{p+1}$ เป็นเวกเตอร์ลักษณะเฉพาะที่สอดคล้องกับค่าลักษณะเฉพาะ $\lambda_1, \lambda_2, ..., \lambda_{p+1}$ ตามลำดับ

Z เป็นเมทริกซ์ที่แต่ละคอลัมภ์ของ Z เป็นตัวแปรอิสระชุดใหม่ซึ่งเป็นผลรวมเชิงเส้น (linear combination) ของตัวแปรอิสระ

และ Λ เป็นเมทริกซ์ขนาด (p+1)×(p+1) ซึ่งสมาชิกในแนวทแยงมุมเป็นค่าลักษณะเฉพาะของเมทริกซ์ $X'X$ และสมาชิกนอกแนวทแยงมุมเป็น 0 กล่าวคือ $\Lambda = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_{p+1})$

เมื่อพิจารณาคุณสมบัติ $\underset{\sim}{\alpha} = Q'\underset{\sim}{\beta}$ เราสามารถเขียนสมการได้เป็น $\underset{\sim}{\beta} = Q\underset{\sim}{\alpha}$ เมื่อเราศึกษาคุณสมบัติของ $\underset{\sim}{\alpha}$ ก็จะทราบถึงคุณสมบัติของ $\underset{\sim}{\beta}$ ได้เช่นเดียวกัน ดังนั้นเพื่อความสะดวกในการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ $\underset{\sim}{\beta}^*_R(k)$ เราจึงศึกษาโดยอาศัยคุณสมบัติของ $\underset{\sim}{\alpha}$ ดังนี้คือ

1. **ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ ในกรณีที่ข้อจำกัดเป็นจริง $\underset{\sim}{\delta} = \underset{\sim}{0}$**

จากสมการ (3.4) จะได้ว่า

$$WAW' = Q(\Lambda + kI)^{-1}\Lambda B\Lambda(\Lambda + kI)^{-1}Q'$$

เมื่อ $B = Q'AQ$ เป็นเมทริกซ์กึ่งบวกแน่นอน (positive semidefinite matrix)

กำหนดให้ค่า $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{p+1} > 0$ จะได้ว่า

$$tr(WAW') = \sum_{i=1}^{p+1} \lambda_i^2 b_{ii} /(\lambda_i + k)^2$$

โดยที่ $b_{ii} \geq 0$ เป็นสมาชิกในแนวทแยงมุมของ B

และ $$\underset{\sim}{\beta}' S(k)^{-2} \underset{\sim}{\beta} = \underset{\sim}{\alpha}'(\Lambda + kI)^{-2} \underset{\sim}{\alpha} = \sum_{i=1}^{p+1} \alpha_i^2 /(\lambda_i + k)^2$$

ทำให้ $MSE(\underset{\sim}{\beta}_R^*(k))$ กรณีที่ข้อจำกัดเป็นจริงในสมการ (3.4) สามารถเขียนให้อยู่ในรูปแบบใหม่ได้ดังนี้

$$MSE(\underset{\sim}{\beta}_R^*(k)) = \sigma^2 \sum_{i=1}^{p+1} \lambda_i^2 b_{ii} /(\lambda_i + k)^2 + k^2 \sum_{i=1}^{p+1} \alpha_i^2 /(\lambda_i + k)^2$$

(3.5) $$= \gamma_1(k) + \gamma_2(k)$$

เนื่องจาก $\gamma_1(k)$ เป็นค่าความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ ดังนั้นฟังก์ชัน $\gamma_1(k)$ จะมีลักษณะเป็นฟังก์ชันลดลงทางเดียว (monotonically decreasing function) ในรูปฟังก์ชันของ k และ $\gamma_2(k)$ เป็นค่าความเอนเอียงกำลังสองของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ จะได้ว่าฟังก์ชัน $\gamma_2(k)$ มีลักษณะเป็นฟังก์ชันที่เพิ่มขึ้นทางเดียว (monotonically increasing function) ในรูปฟังก์ชันของ k

เราสามารถหาค่าจำกัดของความแปรปรวนร่วมและความเอนเอียงกำลังสองของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ โดยใช้ฟังก์ชันของ $\gamma_1(k)$ และ $\gamma_2(k)$ ซึ่งค่าจำกัดของความแปรปรวนร่วม

และค่าความเอนเอียงกำลังสองหาได้จากอนุพันธ์ของ $\gamma_1(k)$ เทียบกับ k และ $\gamma_2(k)$ เทียบกับ k ตามลำดับ สำหรับค่าจำกัดของแต่ละฟังก์ชันเป็นดังนี้

$$\lim_{k \to 0^+} \left(\frac{\partial y_1}{\partial k} \right) = -2\sigma^2 \sum_{i=1}^{p+1} \frac{b_{ii}}{\lambda_i} \tag{3.6}$$

$$\lim_{k \to 0^+} \left(\frac{\partial y_2}{\partial k} \right) = 0 \tag{3.7}$$

ดังนั้นอนุพันธ์ของความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ จะมีค่าเป็นลบและลู่เข้าสู่ $-2\sigma^2 \sum_{i=1}^{p+1} \frac{b_{ii}}{\lambda_i}$ เมื่อ $k \to 0^+$ โดยที่ผลคูณของเมทริกซ์ของตัวแปรอิสระ $X'X$ มีลักษณะเชิงตั้งฉาก กล่าวคือ ตัวแปรอิสระ X จะต้องไม่มีความสัมพันธ์กัน และอนุพันธ์ของความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ จะมีค่าเข้าสู่ $-\infty$ เมื่อ $k \to 0^+$ และ $\lambda_i \to 0$ กล่าวคือ เมทริกซ์ของตัวแปรอิสระ $X'X$ มีสภาวะไม่เหมาะสม ส่วนอนุพันธ์ของความเอนเอียงกำลังสองมีค่าลดลงและจะเป็นศูนย์ ณ จุดกำเนิด เมื่อ $k \to 0^+$

คุณสมบัติของสมการ (3.6) และ (3.7) จะเป็นจริงเมื่อ $k > 0$ ซึ่งสรุปได้ว่าจะมีความเอนเอียงเล็กน้อยและค่าความแปรปรวนร่วมจะมีค่าลดลงมาก จึงส่งผลให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองมีค่าลดลงและมีค่าน้อยกว่าค่าเฉลี่ยความคลาดเคลื่อนกำลังสองจากวิธีกำลังสองน้อยสุดที่ถูกจำกัดในกรณีที่ข้อจำกัดเป็นจริง

ช่วงของค่า k ที่ทำให้ $MSE(\underset{\sim}{\beta}_R^*(k)) < MSE(\underset{\sim}{\beta}^*)$ สามารถหาได้จากการนำ $MSE(\underset{\sim}{\beta}_R^*(k))$ ในสมการ (3.5) หาอนุพันธ์อันดับที่ 1 เทียบกับ k ซึ่งผลดังกล่าวอยู่ในรูปแบบของ

$$\frac{\partial}{\partial k} MSE(\underset{\sim}{\beta}_R^*(k)) = 2\sum_{i=1}^{p+1} \frac{\lambda_i}{(\lambda_i + k)^3} (k\alpha_i^2 - \sigma^2 \lambda_i b_{ii}) \tag{3.8}$$

จากสมการ (3.8) จะเห็นได้ว่าเมื่อ $0 < k < \frac{\sigma^2}{\tilde{\alpha}_{\max}^2}$ เมื่อ $\tilde{\alpha}_{\max}^2$ เป็นค่าที่มากที่สุดของ $\tilde{\alpha}_i^2$ ซึ่ง $\tilde{\alpha}_i^2 = \frac{\alpha_i^2}{(\lambda_i b_{ii})}$ จะทำให้ตัวประมาณริดจ์รีเกรสชันที่ถูกจำกัดมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุดที่ถูกจำกัดในกรณีที่ข้อจำกัดนั้นเป็นจริง

2. ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ $\underset{\sim}{\beta}_R^*(k)$ ในกรณีที่ข้อจำกัดไม่เป็นจริง $\underset{\sim}{\delta} \neq \underset{\sim}{0}$

เนื่องจาก $WS^{-1} = S(k)^{-1}$ ทำให้ $MSE(\underset{\sim}{\beta}_R^*(k))$ ในสมการ (3.2) มีรูปแบบใหม่ดังนี้

$$MSE(\underset{\sim}{\beta}_R^*(k)) = \sigma^2 tr(WAW') + [S(k)^{-1}\underset{\sim}{\delta}^* - kS(k)^{-1}\underset{\sim}{\beta}]'[S(k)^{-1}\underset{\sim}{\delta}^* - kS(k)^{-1}\underset{\sim}{\beta}]$$

$$= \sum_{i=1}^{p+1} \frac{1}{(\lambda_i + k)^2}[\sigma^2\lambda_i^2 b_{ii} + k^2\alpha_i^2 + \delta_i^{*^2} - 2k\alpha_i\delta_i^*] \tag{3.9}$$

จากสมการ (3.9) หาอนุพันธ์อันดับที่ 1 เทียบกับ k จะได้ว่า

$$\frac{\partial}{\partial k} MSE(\underset{\sim}{\beta}_R^*(k)) = 2\sum_{i=1}^{p+1} \frac{1}{(\lambda_i + k)^3}[k(\lambda_i\alpha_i^2 + \alpha_i\delta_i^*) - (\sigma^2\lambda_i^2 b_{ii} + \delta_i^{*^2} + \alpha_i\lambda_i\delta_i^*)] \tag{3.10}$$

จากสมการ (3.10) จะเห็นได้ว่าเมื่อ $0 < k < k^* < \infty$ โดยที่ $k^* = \dfrac{\min\limits_i(\sigma^2\lambda_i^2 b_{ii} + \delta_i^* + \alpha_i\lambda_i\delta_i^*)}{\max\limits_i(\lambda_i\alpha_i^2 + \alpha_i\delta_i^*)}$ จะทำให้ตัวประมาณริดจ์รีเกรสชันที่ถูกจำกัดมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุดที่ถูกจำกัดในกรณีที่ข้อจำกัดนั้นไม่เป็นจริง

ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีริดจ์รีเกรสชันที่ถูกจำกัดจะต้องเลือกค่า k ให้มีค่าเหมาะสม เพื่อที่จะทำให้ค่าความแปรปรวนร่วมของตัวประมาณสัมประสิทธิ์การถดถอยพหุคูณลดลงและค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณริดจ์รีเกรสชันที่ถูกจำกัดมีค่าต่ำสุด

การประมาณค่าพารามิเตอร์ k [3]

วิธีในการประมาณค่าพารามิเตอร์ k มีหลายวิธีการเช่น วิธีของ Hoerl Kennard and Baldwin (1975) วิธีของ TZE-SAN-LEE และวิธีของ McDONALD Galarneau (1975) ซึ่งจากผลงานวิจัยของเจษฎาพร ยุทธนพิบูลย์ชัย (พ.ศ. 2533) ไม่สามารถสรุปได้แน่นอนว่าวิธีการประมาณค่า k วิธีการใดที่ให้ผลสรุปชัดเจน แต่จากผลงานวิจัยของจิรายุส พู่มนตรี (พ.ศ. 2534) ได้มีการเสนอวิธีการประมาณค่า k โดยใช้วิธีการ Binary Search ซึ่งเป็นวิธีการหนึ่งในทางคอมพิวเตอร์ที่ใช้ในการค้นหาข้อมูลและผลสรุปที่ได้คือ วิธีการ Binary Search จะให้ผลในการประมาณค่า k ดีที่สุด แต่วิธีการนี้มีข้อจำกัดคือ ใช้กับข้อมูลที่มีลักษณะแบบไม่ต่อเนื่องและใช้ในการค้นหาข้อมูลที่มีลักษณะแบบเรียงลำดับ ดังนั้นจึงต้องมีการเรียงข้อมูลจากน้อยไปหามากหรือจากมากไปหาน้อย จากข้อจำกัดดังกล่าวจึงทำให้วิธีการแบบ Binary Search ไม่เหมาะสมที่จะนำมาใช้ ผู้วิจัยจึงพยายามหาวิธีการค้นหาข้อมูลอื่น ๆ ซึ่งพบว่าวิธีการค้นหาข้อมูลทางคอมพิวเตอร์ที่สามารถนำมาประยุกต์ใช้คือ วิธีการค้นหาข้อมูลแบบลำดับ (Sequential Search) เพราะวิธีการนี้สามารถใช้ได้กับข้อมูลที่มีลักษณะทั้งแบบต่อเนื่องและแบบไม่ต่อเนื่องและข้อมูลที่จะค้นหาจะเรียงลำดับหรือไม่เรียงลำดับก็ได้ แต่ข้อเสียของวิธีการนี้คือ เราอาจใช้เวลาในการค้นหานานเมื่อข้อมูลที่ต้องการค้นหาอยู่ ณ ลำดับไกล ๆ จากคุณสมบัติของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของสัมประสิทธิ์การถดถอยพหุคูณลักษณะกราฟที่ได้จะมีจุดต่ำสุดเพียงจุดเดียวและไม่มีจุดวกกลับ ดังนั้นการพิจารณาที่จะใช้วิธีการค้นหาแบบลำดับจึงมีความเหมาะสมมากกว่า ผู้วิจัยจึงได้ใช้วิธีการในการประมาณค่า k โดยใช้วิธีการค้นหาแบบลำดับ ซึ่งมีหลักการดังนี้

กำหนดให้ k_{opt} คือ ค่า k ที่ทำให้ $MSE(\underset{\sim}{\beta}_R^*(k))$ มีค่าต่ำสุด

k_{int} คือ ค่า k ณ จุดเริ่มต้นในการค้นหาข้อมูล

k_{01} คือ ค่า k ณ จุดที่มีการเว้นช่วงห่างของค่า k เท่ากับ 0.01

$k_{BEF\,01}$ คือ ค่า k ณ จุดที่อยู่ก่อนค่า k ที่เหมาะสมที่มีการเว้นช่วงห่างของค่า k เท่ากับ 0.01

---

[3] ธันยากร ต้นชลขันธ์,"การเปรียบเทียบการประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยวิธีกำลังสองน้อยที่สุด วิธีริดจ์รีเกรสชัน และวิธีที่ใช้หลักการของริดจ์และสไตน์ ในกรณีที่เกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ," (วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย,2538), หน้า 19.

k_{AFT01} คือ ค่า k ณ จุดที่อยู่หลังค่า k ที่เหมาะสมที่มีการเว้นช่วงห่างของค่า k เท่ากับ 0.01

k_{001} คือ ค่า k ณ จุดที่มีการเว้นช่วงห่างของค่า k เท่ากับ 0.001

k_{STOP} คือ ค่า k ที่ใช้ตรวจสอบหยุดการค้นหาค่า k ที่เหมาะสมเมื่อทำการค้นหาข้อมูลทางด้านขวาของค่า k ที่เหมาะสม

d คือ ค่าคงที่ที่กำหนดขึ้นเพื่อเป็นช่วงห่างของค่า k

<u>ขั้นที่ 1</u> กำหนดให้ค่า $k_{\text{int}} = 0.0$ และ ค่า $d = 0.01$

<u>ขั้นที่ 2</u> คำนวณค่า $k_{01} = k_{\text{int}} + d$

<u>ขั้นที่ 3</u> คำนวณหา $MSE(\underset{\sim}{\beta}_R^*(k_{\text{int}}))$ และหา $MSE(\underset{\sim}{\beta}_R^*(k_{01}))$ แล้วทำการเปรียบเทียบ

ก) ถ้า $MSE(\underset{\sim}{\beta}_R^*(k_{01})) > MSE(\underset{\sim}{\beta}_R^*(k_{\text{int}}))$

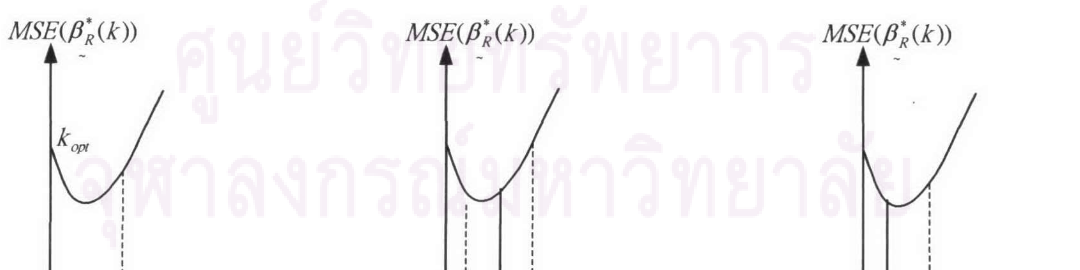
จะได้ว่า $k_{opt} = k_{\text{int}}$ ยุติการประมวลผล

ข) ถ้า $MSE(\underset{\sim}{\beta}_R^*(k_{01})) \le MSE(\underset{\sim}{\beta}_R^*(k_{\text{int}}))$ เราจะกำหนดให้

$$k_{\text{int}} = k_{01}$$

แล้วทำการวิเคราะห์ขั้นที่ 2 ใหม่จนกว่าจะเข้ากรณี ก)

เนื่องจากลักษณะของ $MSE(\underset{\sim}{\beta}_R^*(k))$ เป็นรูปโค้งหงาย การหาจุดที่ต่ำสุดจึงขึ้นอยู่กับการกำหนดช่วงห่างของค่า k ซึ่งจากการวิจัยพบว่าลักษณะความเป็นไปได้ของการจะพบจุดที่ต่ำสุดเมื่อช่วงห่างไม่ละเอียดพออาจแสดงได้ดังกรณีที่ 1-3



<u>กรณีที่ 1</u> <u>กรณีที่ 2</u> <u>กรณีที่ 3</u>

<u>กรณีที่ 1</u> ถ้าเรากำหนดช่วง d ห่างเกินไป จุดที่จะได้จุดต่ำสุดอาจจะเป็นจุดเริ่มแรกที่กำหนดคือ $k_{opt} = 0.0$ ซึ่งกรณีนี้ผู้วิจัยจะทำการแก้ปัญหาโดยให้มีการกำหนดช่วงห่างของการ

ค้นหาให้ละเอียดขึ้นแล้วทำการค้นหาข้อมูลแบบลำดับจากจุดเริ่มต้น $k_{opt} = 0.0$ จนถึงจุดที่ $k_{opt} = 0.0+0.01$ ใหม่อีกครั้งโดยใช้ช่วงห่างของการค้นหาเป็น $d = 0.001$

<u>กรณีที่ 2</u> ถ้าจุดของค่า k ที่ค้นหาได้ตกอยู่หลังค่า k ที่เหมาะสมจริง ๆ จึงจะเกิดกรณีนี้ขึ้น ผู้วิจัยได้แก้ปัญหาโดยการเพิ่มช่วงห่างของการค้นหาให้ละเอียดขึ้นและจะค้นหาข้อมูลทางด้านซ้ายของค่า k_{opt} อีกครั้งโดยจะค้นหาข้อมูลแบบลำดับจากจุด $k_{opt} - 0.01$ ไปจนถึงจุด k_{opt} โดยใช้ช่วงห่างของการค้นหาเป็น $d = 0.001$

<u>กรณีที่ 3</u> ถ้าจุดของค่า k ที่ค้นหาได้ตกอยู่ก่อนหน้าค่า k ที่เหมาะสมจริง ๆ จึงจะเกิดกรณีนี้ขึ้น ผู้วิจัยได้แก้ปัญหาโดยการเพิ่มช่วงห่างของการค้นหาให้ละเอียดขึ้นและจะค้นหาข้อมูลทางด้านขวาของค่า k_{opt} อีกครั้งโดยจะค้นหาข้อมูลแบบลำดับจากจุด k_{opt} ไปจนถึงจุด $k_{opt} + 0.01$ โดยใช้ช่วงห่างของการค้นหาเป็น $d = 0.001$

ดังนั้นผู้วิจัยจึงทำการค้นหาค่า k ที่เหมาะสมอีกครั้งโดยกำหนดช่วงความห่างของการค้นหาใหม่ คือ ให้ $d = 0.001$ โดยจะทำการตรวจสอบก่อนว่าค่า k_{opt} ที่ได้มีค่าเท่ากับศูนย์หรือไม่ ถ้ามีค่าเท่ากับศูนย์ก็คือกรณีที่ 1 ผู้วิจัยจะทำการค้นหาข้อมูลแบบลำดับเฉพาะทางด้านขวาของค่า k_{opt} แต่ถ้าค่า k_{opt} มีค่าไม่เท่ากับศูนย์ ผู้วิจัยจะทำการค้นหาข้อมูลแบบลำดับทางด้านซ้ายของค่า k_{opt} ก่อนถ้าไม่พบก็จะทำการค้นหาทางด้านขวาของค่า k_{opt} อีกครั้งซึ่งขั้นตอนต่าง ๆ ในการค้นหามีดังนี้

ให้ $d = 0.001$

<u>ขั้นที่ 1</u> ตรวจสอบว่าค่า k_{opt} เท่ากับศูนย์หรือไม่

ก) ถ้า $k_{opt} = 0$ ให้ทำการวิเคราะห์ขั้นที่ 3 (ค้นหาทางด้านขวาของค่า k_{opt})

ข) ถ้า $k_{opt} \neq 0$ (มากกว่าศูนย์) ให้ทำการวิเคราะห์ขั้นที่ 2 (ค้นหาทางด้านซ้ายของค่า k_{opt}

<u>ขั้นที่ 2</u> ให้ $k_{BEF01} = k_{opt} - 0.01$

ก) ให้ $k_{001} = k_{BEF01} + d$

เปรียบเทียบค่า k_{001} กับ k_{opt}

ถ้า $k_{001} \geq k_{opt}$ ให้ทำการวิเคราะห์ขั้นที่ 3

ถ้า $k_{001} < k_{opt}$

คำนวณหา $MSE(\underset{\sim}{\beta}_R^*(k_{BEF01}))$ และหา $MSE(\underset{\sim}{\beta}_R^*(k_{001}))$ แล้วทำการเปรียบเทียบ

- ถ้า $MSE(\underset{\sim}{\beta}_R^*(k_{001})) > MSE(\underset{\sim}{\beta}_R^*(k_{BEF01}))$

จะได้ว่า $k_{opt} = k_{BEF01}$

ยุติการประมวลผล

- ถ้า $MSE(\underset{\sim}{\beta}_R^*(k_{001})) \le MSE(\underset{\sim}{\beta}_R^*(k_{BEF01}))$

เราจะกำหนด $k_{BEF01} = k_{001}$

แล้วทำการวิเคราะห์ข้อ ก) ในขั้นที่ 2 จนกว่าจะเข้าใกล้กรณีที่

$k_{001} \ge k_{opt}$ หรือ $MSE(\underset{\sim}{\beta}_R^*(k_{001})) > MSE(\underset{\sim}{\beta}_R^*(k_{BEF01}))$

<u>ขั้นที่ 3</u> ให้ $k_{STOP} = k_{opt} + 0.01$

ก) ให้ $k_{001} = k_{opt}$

ข) ให้ $k_{AFT01} = k_{001} + d$

เปรียบเทียบค่า k_{AFT01} กับ k_{STOP}

ถ้า $k_{AFT01} \ge k_{STOP}$ ให้ยุติการประมวลผล

ถ้า $k_{AFT01} < k_{STOP}$

เราจะคำนวณหา $MSE(\underset{\sim}{\beta}_R^*(k_{AFT01}))$ และคำนวณหา

$MSE(\underset{\sim}{\beta}_R^*(k_{001}))$ แล้วทำการเปรียบเทียบ

- ถ้า $MSE(\underset{\sim}{\beta}_R^*(k_{AFT01})) > MSE(\underset{\sim}{\beta}_R^*(k_{001}))$

จะได้ว่า $k_{opt} = k_{001}$ และยุติการประมวลผล

- $MSE(\underset{\sim}{\beta}_R^*(k_{AFT01})) \le MSE(\underset{\sim}{\beta}_R^*(k_{001}))$

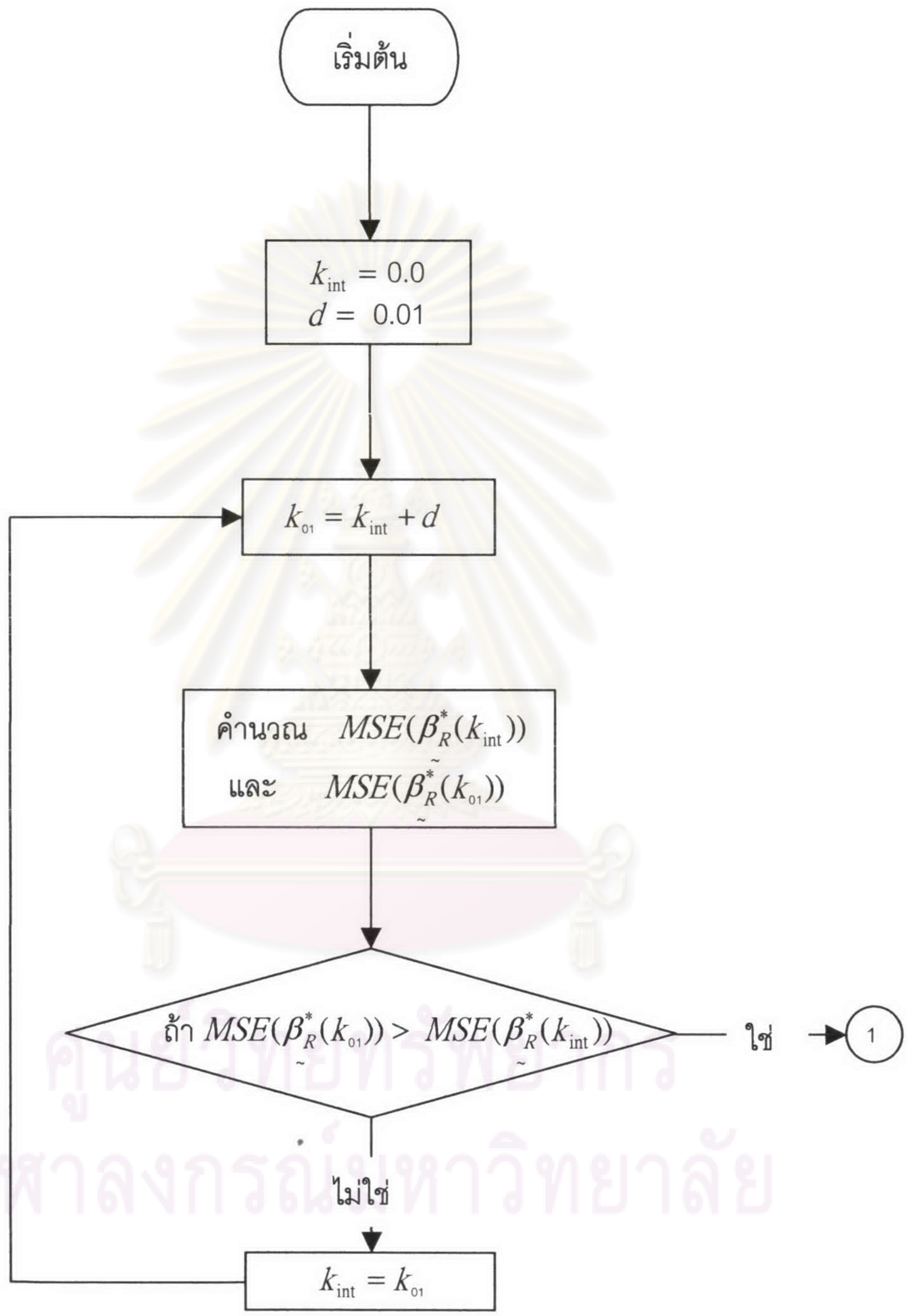
เราจะกำหนด $k_{001} = k_{AFT01}$

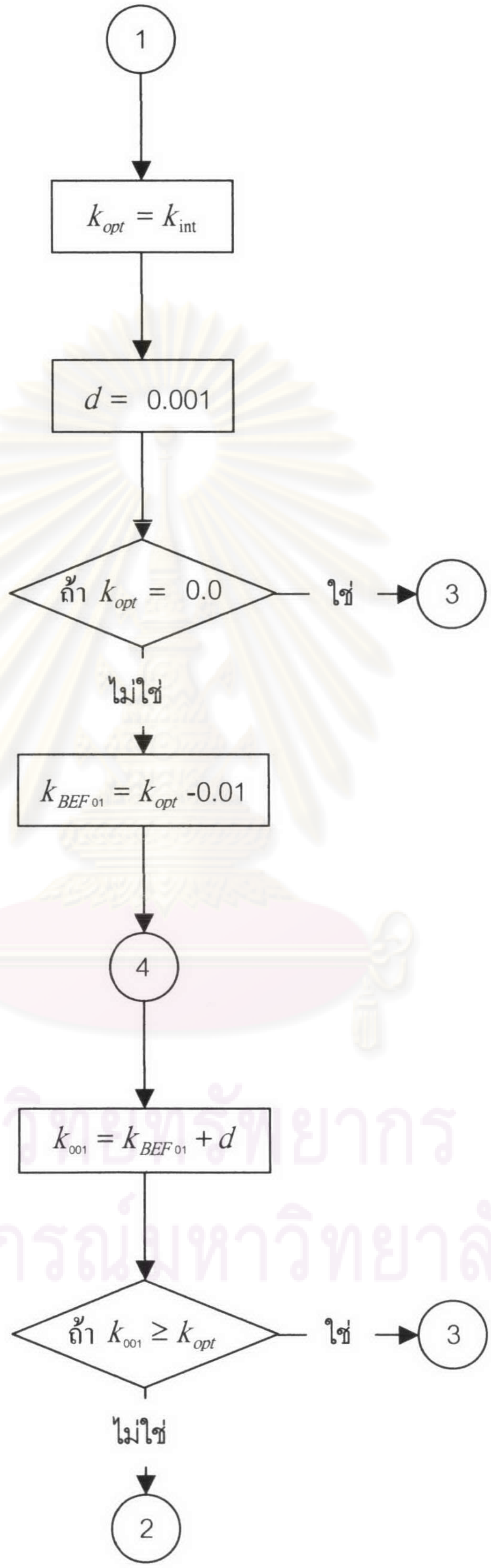
และทำการวิเคราะห์ข้อ ข) ในขั้นที่ 3 จนกว่าจะเข้ากรณีที่

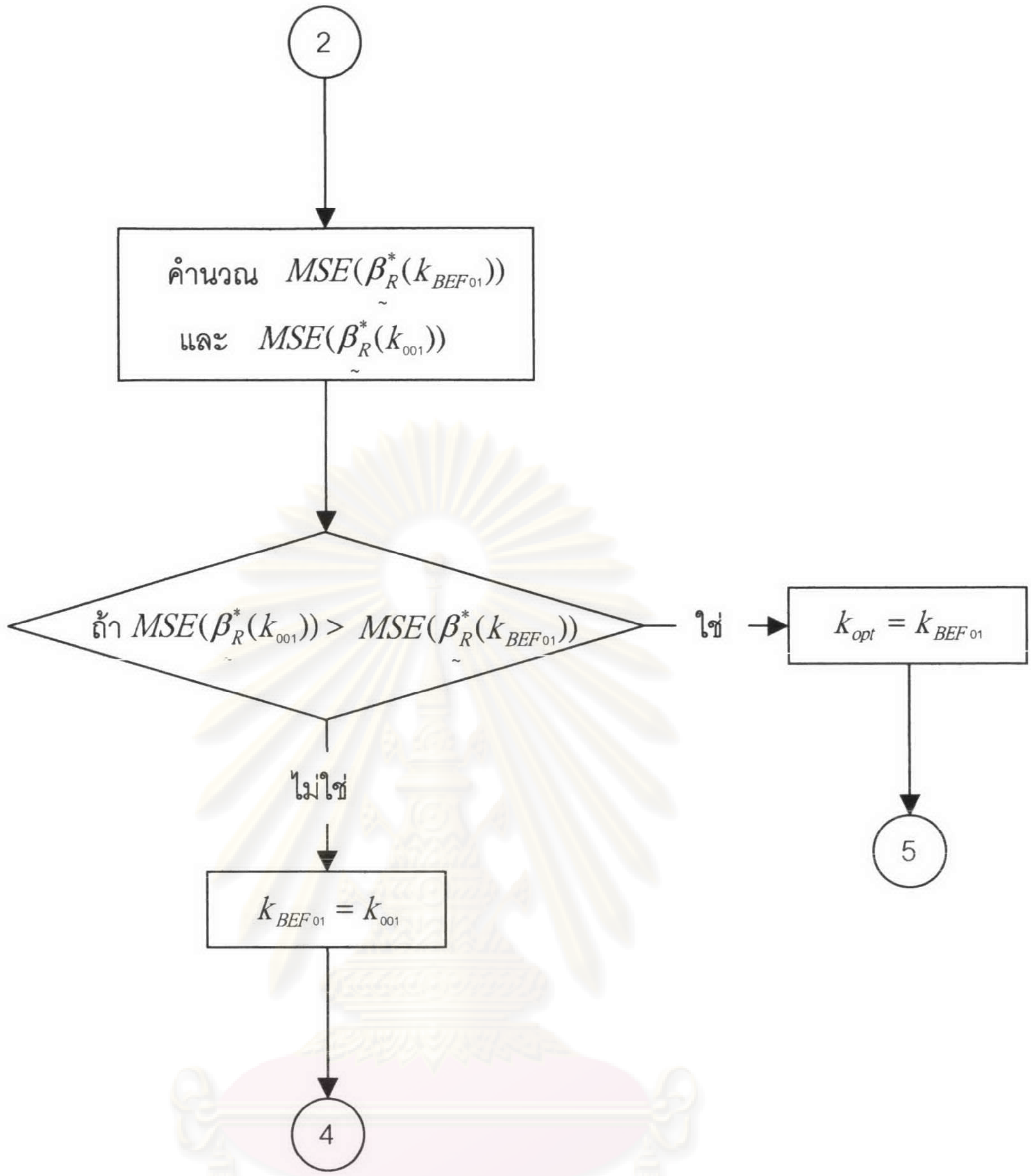
$k_{AFT01} \ge k_{STOP}$ หรือ $MSE(\underset{\sim}{\beta}_R^*(k_{AFT01})) > MSE(\underset{\sim}{\beta}_R^*(k_{001}))$

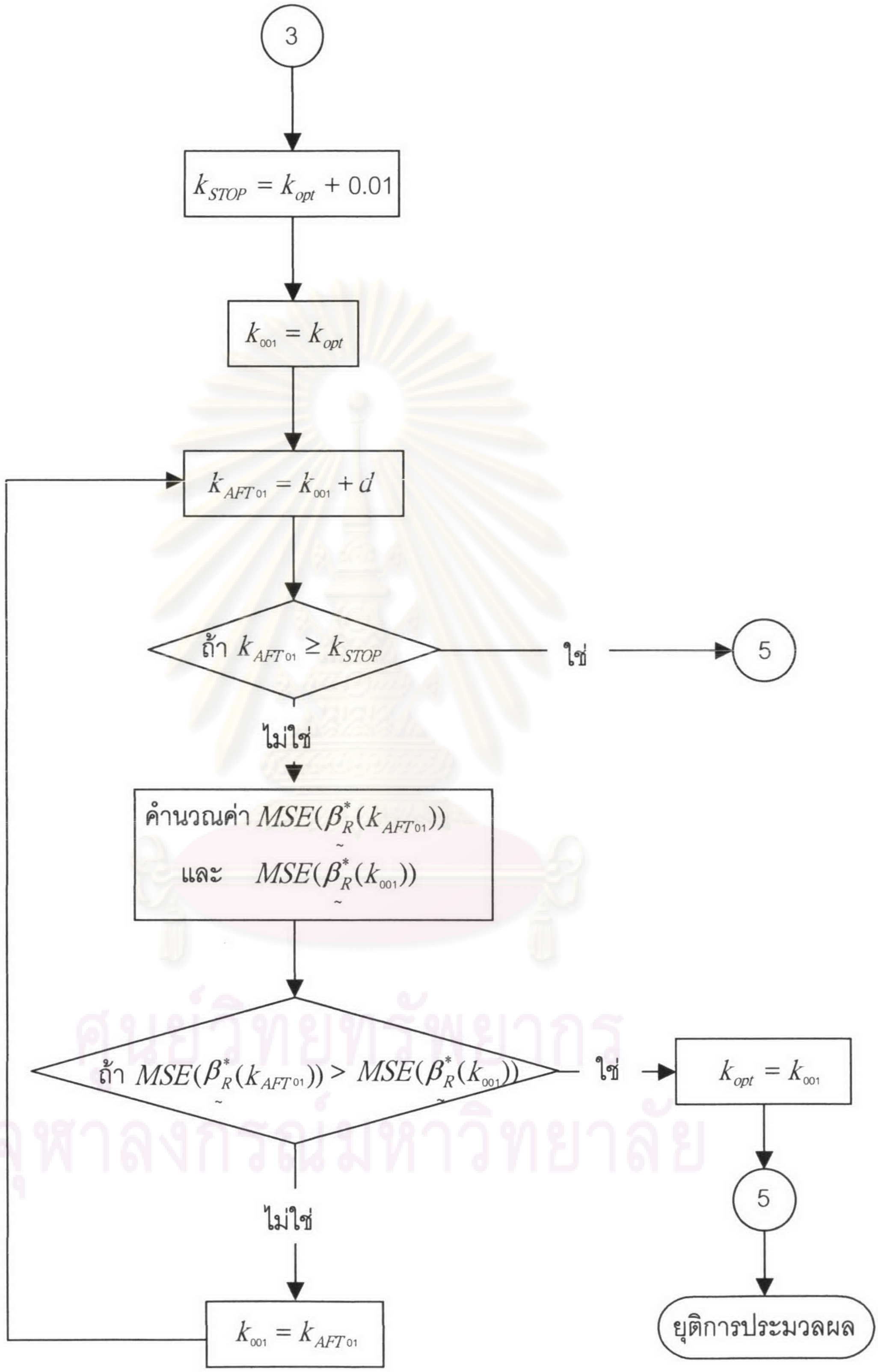
เนื่องจาก ณ จุดที่ค่า $d = 0.001$ และ $d = 0.0001$ จะให้ค่า $MSE(\underset{\sim}{\beta}_R^*(k))$ ที่มีค่าใกล้เคียงกัน ดังนั้นผู้วิจัยจึงไม่ทำการขยายช่วงค่า d จาก $d = 0.001$ เป็น $d = 0.0001$ สำหรับขั้นตอนต่าง ๆ ที่อธิบายมาแล้วสามารถเขียนเป็นผังงานเพื่อให้เห็นลักษณะการทำงานที่ชัดเจนได้ดังผังงานที่ 1

ผังงานที่ 1 แสดงการค้นหาค่า k โดยใช้วิธีการค้นหาแบบลำดับ (Sequential Search)




1
$k_{opt} = k_{\text{int}}$
$d =$ 0.001
ถ้า k_{opt} = 0.0
ใช่
3
ไม่ใช่
$k_{BEF01} = k_{opt}$ -0.01
4
$k_{001} = k_{BEF01} + d$
ถ้า $k_{001} \geq k_{opt}$
ใช่
3
ไม่ใช่
2


2
คำนวณ $MSE(\underset{\sim}{\beta}^*_R(k_{BEF01}))$
และ $MSE(\underset{\sim}{\beta}^*_R(k_{001}))$
ถ้า $MSE(\underset{\sim}{\beta}^*_R(k_{001})) > MSE(\underset{\sim}{\beta}^*_R(k_{BEF01}))$
ใช่
$k_{opt} = k_{BEF01}$
5
ไม่ใช่
$k_{BEF01} = k_{001}$
4


3
$k_{STOP} = k_{opt} + 0.01$
$k_{001} = k_{opt}$
$k_{AFT01} = k_{001} + d$
ถ้า $k_{AFT01} \geq k_{STOP}$
ใช่
5
ไม่ใช่
คำนวณค่า $MSE(\underset{\sim}{\beta}^*_R(k_{AFT01}))$
และ $MSE(\underset{\sim}{\beta}^*_R(k_{001}))$
ถ้า $MSE(\underset{\sim}{\beta}^*_R(k_{AFT01})) > MSE(\underset{\sim}{\beta}^*_R(k_{001}))$
ใช่
$k_{opt} = k_{001}$
5
ไม่ใช่
$k_{001} = k_{AFT01}$
ยุติการประมวลผล

4. การประมาณค่าสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีลิวที่ถูกจำกัด (RL)

ในปี ค.ศ. 1999 เซลาห์ฮัททิน คาซิแรนลาร์ (Selahattin Kaciranlar) , ซาดุลเลาะห์ ซาคัลลิโอกลู (Sadullah Sakallioglu) และ ฟิคลิ แอคเดนนิส (Fikri Akdeniz) ได้เสนอวิธีลิวที่ถูกจำกัดในการแก้ปัญหาการเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ โดยได้นำข้อดีของวิธีลิวและวิธีกำลังสองน้อยสุดที่ถูกจำกัดมาผสมผสานกัน หลักการของวิธีนี้จะใช้หลักการของตัวประมาณลิวที่จะลดค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของการประมาณ $\underset{\sim}{\beta}^*$ ให้ต่ำลง กล่าวคือการบวกค่าคงที่ค่าหนึ่งให้กับสมาชิกในแนวทแยงมุมของเมทริกซ์ $X'X$ เนื่องจากตัวประมาณริดจ์มีปัญหาในการประมาณค่าพารามิเตอร์ k ดังนั้นตัวประมาณที่ใช้หลักการของลิวจึงแก้ปัญหาดังกล่าวโดยการบวกค่าคงที่ k ที่มีค่าเท่ากับหนึ่งให้กับสมาชิกทุกตัวในแนวทแยงมุมของเมทริกซ์ $X'X$ และเทอมของ $X'\underset{\sim}{y}+Sg = X'X\underset{\sim}{\beta}^*$ จึงนำ $d\underset{\sim}{\beta}^*$ มาบวกเพราะพารามิเตอร์ d ที่ใช้หลักการของตัวประมาณสไตน์หาได้ง่าย ดังนั้นการประมาณสัมประสิทธิ์การถดถอยพหุคูณด้วยวิธีลิวที่ถูกจำกัดอยู่ในรูปของ

(4.1) $$\underset{\sim}{\beta}_d^*(d) = (X'X+I)^{-1}(X'\underset{\sim}{y}+Sg+d\underset{\sim}{\beta}^*)$$

เราสามารถเขียน $\underset{\sim}{\beta}_d^*(d)$ ให้อยู่ในรูปของฟังก์ชัน $\underset{\sim}{\beta}^*$ โดยแทนค่า $X'\underset{\sim}{y}+Sg = X'X\underset{\sim}{\beta}^*$ ในสมการ (4.1) จะได้ว่าตัวประมาณดังกล่าวอยู่ในรูปของ

$$\begin{aligned}\underset{\sim}{\beta}_d^*(d) &= (X'X+I)^{-1}(X'X\underset{\sim}{\beta}^*+d\underset{\sim}{\beta}^*)\\ &= (X'X+I)^{-1}(X'X+dI)\underset{\sim}{\beta}^*\\ &= (S+I)^{-1}(S+dI)\underset{\sim}{\beta}^*\\ &= F_d\underset{\sim}{\beta}^*\end{aligned}$$

เมื่อ $F_d = (S+I)^{-1}(S+dI)$ ถ้า $d=1$ จะทำให้ $\underset{\sim}{\beta}_d^*(d) = \underset{\sim}{\beta}^*$

เมื่อพิจารณาค่าคาดหวังของ $\underset{\sim}{\beta}_d^*(d)$ จะได้ว่า

(4.2) $$E(\underset{\sim}{\beta}_d^*(d)) = F_d\underset{\sim}{\beta} + F_dS^{-1}R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}$$

จากค่าคาดหวังของตัวประมาณ $\underset{\sim}{\beta}_d^*(d)$ จะทำให้ทราบว่าตัวประมาณ $\underset{\sim}{\beta}_d^*(d)$ เป็นตัวประมาณเอนเอียง เนื่องจาก $E(\underset{\sim}{\beta}_d^*(d)) \neq \underset{\sim}{\beta}$ ในกรณีที่ $\underset{\sim}{\delta} = \underset{\sim}{0}$ คือข้อจำกัดเป็นจริง และ $d = 1$ ตัวประมาณ $\underset{\sim}{\beta}_d^*(d)$ จะเป็นตัวประมาณไม่เอนเอียง

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\underset{\sim}{\beta}_d^*(d)$ จะอยู่ในรูปของ

$$
\begin{aligned}
MSE(\underset{\sim}{\beta}_d^*(d)) &= E[(\underset{\sim}{\beta}_d^* - \underset{\sim}{\beta})'(\underset{\sim}{\beta}_d^* - \underset{\sim}{\beta})] \\
&= tr(Cov(\underset{\sim}{\beta}_d^*)) + [E(\underset{\sim}{\beta}_d^*) - \underset{\sim}{\beta}]^2 \\
&= \sigma^2 tr(F_d(S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1})(I - R'(RS^{-1}R')^{-1}RS^{-1})F_d^{/}) \\
&\quad + \left\| F_d\,\underset{\sim}{\beta} + F_d S^{-1}\,\underset{\sim}{\delta}^* - \underset{\sim}{\beta} \right\|^2 \\
&= \sigma^2 tr(F_d(S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1})F_d^{/}) + \left\| F_d\,\underset{\sim}{\beta} + F_d S^{-1}\,\underset{\sim}{\delta}^* - \underset{\sim}{\beta} \right\|^2
\end{aligned}
$$

(4.3) $$= \sigma^2 tr(F_d A F_d') + \left\| F_d\,\underset{\sim}{\beta} + F_d S^{-1}\,\underset{\sim}{\delta}^* - \underset{\sim}{\beta} \right\|^2$$

เมื่อ $\underset{\sim}{\delta}^* = R'(RS^{-1}R')^{-1}\underset{\sim}{\delta}$

และ $A = S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$

จากสมการ (4.3) ในกรณีที่ข้อจำกัดนั้นเป็นจริงคือ $\underset{\sim}{\delta} = \underset{\sim}{0}$ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\underset{\sim}{\beta}_d^*(d)$ จะอยู่ในรูปของ

(4.4) $$MSE(\underset{\sim}{\beta}_d^*(d)) = \sigma^2 tr(F_d A F_d') + (d-1)^2\,\underset{\sim}{\beta}'(S+I)^{-2}\,\underset{\sim}{\beta}$$

คุณสมบัติของตัวประมาณลิวที่ถูกจำกัดโดยใช้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองเป็นเกณฑ์

1. **ในกรณีที่ข้อจำกัดเป็นจริง $\underset{\sim}{\delta} = \underset{\sim}{0}$**

เนื่องจาก S เป็นเมทริกซ์บวกแน่นอน (positive definite matrix) ทำให้สามารถหาเมทริกซ์เชิงตั้งฉาก (orthogonal matrix) ที่ทำให้ $Q'SQ = \Lambda$ เมื่อ $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_{p+1})$ โดยที่ $QQ' = Q'Q = I$ และ $B = Q'AQ$ เมื่อ $A = S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$ เป็นเมทริกซ์กึ่งบวกแน่นอน (positive semidefinite matrix) ทำให้สมาชิกในแนวทแยงมุม b_{ii} ของ

สมาชิก B มีค่ามากกว่าหรือเท่ากับศูนย์ทั้งหมด ใช้ประโยชน์จากคุณสมบัติของฟังก์ชันรอยเมทริกซ์ (trace function) ทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของ $\underset{\sim}{\beta}_d^*(d)$ ในสมการ (4.4) มีรูปแบบดังนี้

$$MSE(\underset{\sim}{\beta}_d^*(d)) = \sigma^2 \sum_{i=1}^{p+1} \frac{(\lambda_i + d)^2}{(\lambda_i + 1)^2} b_{ii} + (d-1)^2 \sum_{i=1}^{p+1} \frac{\alpha_i^2}{(\lambda_i + 1)^2} = f(d) \tag{4.5}$$

เมื่อ $\underset{\sim}{\alpha} = Q' \underset{\sim}{\beta}$

สังเกตว่าเมื่อกำหนด $\underset{\sim}{\beta}$ และ σ^2 ฟังก์ชัน $f(d) = MSE(\underset{\sim}{\beta}_d^*(d))$ ในสมการ (4.5) เป็นฟังก์ชันนูนในรูปของ d และอนุพันธ์อันดับที่ 1 เทียบกับ d มีรูปแบบดังนี้

$$f'(d) = 2\sigma^2 \sum_{i=1}^{p+1} \frac{(\lambda_i + d)^2 b_{ii}}{(\lambda_i + 1)^2} + 2\sum_{i=1}^{p+1} \frac{(d-1)\alpha_i^2}{(\lambda_i + 1)^2} = 2\sum_{i=1}^{p+1} \frac{(\lambda_i + d) b_{ii} \sigma^2 + (d-1)\alpha_i^2}{(\lambda_i + 1)^2} \tag{4.6}$$

เนื่องจาก $B = Q'AQ$ และ $QQ' = Q'Q = I$ ทำให้ $A = QBQ'$ จากสมการ (2.8) และใช้คุณสมบัติของฟังก์ชันรอยเมทริกซ์จะได้ว่า $MSE(\underset{\sim}{\beta}^*)$ มีรูปแบบดังนี้

$$MSE(\underset{\sim}{\beta}^*) = \sigma^2 tr(A) = \sigma^2 tr(QBQ') = \sigma^2 tr(B) = \sigma^2 \sum_{i=1}^{p+1} b_{ii} = f(1) \tag{4.7}$$

พิจารณาที่ $d = 1$ อนุพันธ์อันดับที่ 1 ของ f ในสมการ (4.5) จะกลายเป็น

$$f'(1) = 2\sum_{i=1}^{p+1} \frac{\sigma^2}{\lambda_i + 1} b_{ii} \tag{4.8}$$

เนื่องจาก $\lambda_i > 0$ (i = 1, 2, ..., p+1) และ $\sigma^2 > 0$ ทำให้ $f'(1)$ จะมากกว่าศูนย์ได้เมื่อ $B \neq 0$ เท่านั้น แต่ $B \neq 0$ ถ้า q < (p+1) เท่านั้น จากข้อสมมติ q < (p+1) ในตัวแบบข้อจำกัดจึงทำให้ $f'(1) > 0$ ซึ่งหมายความว่าเมื่อ $d < 1$ อัตราการเพิ่มขึ้นของ $MSE(\underset{\sim}{\beta}_d^*(d))$ เมื่อเทียบกับ d จะมีค่าน้อยกว่าอัตราการเพิ่มขึ้นของ $MSE(\underset{\sim}{\beta}^*)$ เมื่อเทียบกับ d แสดงว่าถ้า $d < 1$ จะทำให้ $f(d) < f(1)$ นั่นก็คือทำให้ $MSE(\underset{\sim}{\beta}_d^*(d)) < MSE(\underset{\sim}{\beta}^*)$ เมื่อ

พิจารณา $L_\tau(f) = \{d : f(d) \le \tau\}$ สำหรับแต่ละ τ เราสามารถหาความแม่นยำของช่วงค่าพารามิเตอร์ d ได้ เนื่องจาก f เป็นฟังก์ชันนูนโดยแท้ (strictly convex function) ทำให้ $L_\tau(f)$ เป็นฟังก์ชันนูน (convex function) สำหรับแต่ละ τ ด้วย จาก $f(d) = f(1)$ เราสามารถหาค่าพารามิเตอร์ d ได้ 2 ค่าที่ทำให้ $f(d) = f(1)$ เพราะเป็นฟังก์ชันพอลิโนเมียลดีกรี 2 ค่าพารามิเตอร์ d ทั้ง 2 ค่านี้คือ

(4.9) $$d_l = 1 - 2\sum_{i=1}^{p+1} \frac{\sigma^2 b_{ii}}{\lambda_i + 1} \Bigg/ \sum_{i=1}^{p+1} \frac{\sigma^2 b_{ii} + \alpha_i^2}{(\lambda_i + 1)^2} \quad \text{และ} \quad d_u = 1$$

เมื่อ $d_l < d_u = 1$

จากที่กล่าวมาข้างต้น เราจึงสรุปได้ว่าในกรณีที่ข้อจำกัดเป็นจริง ถ้า $\sigma^2 > 0$, $q < p+1$ และ d_l และ d_u เป็นไปตามสมการ (4.9) ตัวประมาณลิวที่ถูกจำกัดจะมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุดที่ถูกจำกัด เมื่อ $d_l < d < 1$

เนื่องจาก $f(d) = MSE(\underset{\sim}{\beta}_d^*(d))$ เป็นฟังก์ชันนูน เราสามารถแก้ปัญหาเพื่อหาค่าพารามิเตอร์ d ที่ทำให้ $f(d)$ มีค่าต่ำสุดได้ โดยการหาอนุพันธ์อันดับที่ 1 ของ $f(d)$ เทียบกับ d แล้วให้เท่ากับศูนย์ ซึ่งค่า d_{opt} ที่ได้จะมีเพียงคำตอบเดียว คือ

(4.10) $$d_{opt} = 1 - \sigma^2 \sum_{i=1}^{p+1} \frac{b_{ii}}{\lambda_i + 1} \Bigg/ \sum_{i=1}^{p+1} \frac{b_{ii}\sigma^2 + \alpha_i^2}{(\lambda_i + 1)^2}$$

จากสมการ (4.10) เมื่อกำหนด $\underset{\sim}{\beta}$ และ σ^2 การใช้ $d = d_{opt}$ จะทำให้ $MSE(\underset{\sim}{\beta}_d^*(d))$ มีค่าต่ำสุด แต่ว่า d_{opt} ขึ้นอยู่กับค่าพารามิเตอร์ที่ไม่ทราบค่าดังกล่าว จึงได้มีการประมาณค่าพารามิเตอร์ที่ไม่ทราบค่า α_i และ σ^2 ขึ้นโดยใช้การประมาณวิธี OLS เราเรียกการประมาณ d_{opt} นี้ว่า ตัวประมาณที่ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำสุดที่มาจาก OLS (OLS-based minimum MSE estimator)

(4.11) $$\hat{d}_{OLS} = 1 - \hat{\sigma}^2 \sum_{i=1}^{p+1} \frac{b_{ii}}{\lambda_i + 1} \Bigg/ \sum_{i=1}^{p+1} \frac{b_{ii}\hat{\sigma}^2 + \hat{\alpha}_i^2}{(\lambda_i + 1)^2}$$

เนื่องจากตอนนี้เราพิจารณาที่ข้อจำกัดเป็นจริง ตัวประมาณ d_{opt} ที่ดีที่สุดควรได้มาจากการแทนที่ α_i และ σ^2 ใน d_{opt} ที่ไม่ใช่การประมาณด้วยวิธี OLS แต่แทนที่ด้วย $\tilde{\alpha}_i$ และ

$\tilde{\sigma}^2$ ตามลำดับ ที่มาจากการประมาณด้วยวิธี RLS เราเรียกการประมาณ d_{opt} ที่ได้นี้ว่า ตัวประมาณที่ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำสุดที่มาจาก RLS (RLS-based minimum MSE estimator) ตัวประมาณที่ได้มีรูปแบบดังนี้

(4.12) $$\hat{d}_{RLS} = 1 - \tilde{\sigma}^2 \sum_{i=1}^{p+1} \frac{b_{ii}}{\lambda_i + 1} \Big/ \sum_{i=1}^{p+1} \frac{b_{ii}\tilde{\sigma}^2 + \tilde{\alpha}_i^2}{(\lambda_i + 1)^2}$$

2. ในกรณีที่ข้อจำกัดไม่เป็นจริง $\underset{\sim}{\delta} \neq \underset{\sim}{0}$

ในกรณีที่ข้อจำกัดไม่เป็นจริงตัวประมาณ $\underset{\sim}{\beta}_d^*(d)$ และ $\underset{\sim}{\beta}^*$ จะเป็นตัวประมาณค่าเอนเอียงทั้งคู่ เราสามารถแทน $F_d - I = (d-1)(S+I)^{-1}$ ในสมการ (4.3) ซึ่งผลดังกล่าวจะมีรูปแบบดังนี้

$$MSE(\underset{\sim}{\beta}_d^*(d)) = \sigma^2 tr(F_d A F_d') + \left\| (d-1)(S+I)^{-1} \underset{\sim}{\beta} + F_d S^{-1} \underset{\sim}{\delta}^* \right\|^2$$

(4.13) $$= \sum_{i=1}^{p+1} \frac{1}{(\lambda_i + 1)^2} \left[\left(\sigma^2 b_{ii} + \frac{\tilde{\delta}_i^2}{\lambda_i^2} \right) (\lambda_i + d)^2 + (d-1)^2 \alpha_i^2 + \frac{2(d-1)\alpha_i \tilde{\delta}_i (\lambda_i + d)}{\lambda_i} \right]$$

โดยที่ $\underset{\sim}{\tilde{\delta}} = Q' \underset{\sim}{\delta}^*$ จากสมการ (4.13) เราจะให้ $MSE(\underset{\sim}{\beta}_d^*(d)) = h(d)$ ซึ่ง $h(d)$ เป็นฟังก์ชันนูนโดยแท้ใน d

จากสมการ (2.7) ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ $\underset{\sim}{\beta}^*$ ในกรณีที่ข้อจำกัดไม่เป็นจริงสามารถเขียนได้อีกรูปแบบหนึ่งดังนี้

$$MSE(\underset{\sim}{\beta}^*) = \sum_{i=1}^{p+1} \left(\sigma^2 b_{ii} + \frac{\tilde{\delta}_i^2}{\lambda_i^2} \right)$$

จากสมการ (4.13) หาอนุพันธ์อันดับที่ 1 เทียบกับ d และแทนค่า $d = 1$ จะได้ว่า

$$h'(d) = \sum_{i=1}^{p+1} \frac{2}{(\lambda_i + 1)^2} \left[\sigma^2 b_{ii}(\lambda_i + d) + (d-1)\alpha_i^2 + \frac{\alpha_i \tilde{\delta}_i (\lambda_i + d) + (d-1)\alpha_i \tilde{\delta}_i}{\lambda_i} + \frac{\tilde{\delta}_i^2 (\lambda_i + d)}{\lambda_i^2} \right]$$

และ
$$h'(1)=\sum_{i=1}^{p+1}\frac{2}{(\lambda_i+1)}\left[\sigma^2 b_{ii}+\frac{\alpha_i\tilde{\delta}_i}{\lambda_i}+\frac{\tilde{\delta}_i^2}{\lambda_i^2}\right]$$

นอกจากนี้
$$h(1)=\sum_{i=1}^{p+1}\left(\sigma^2 b_{ii}+\frac{\tilde{\delta}_i^2}{\lambda_i^2}\right)=MSE(\underset{\sim}{\beta}^*)$$

เมื่อพิจารณาจะเห็นได้ว่าเงื่อนไขที่จะทำให้ $h'(1)>0$ ได้นั้น $1+\alpha_i\lambda_i\tilde{\delta}_i^+$ จะต้องมากกว่าหรือเท่ากับศูนย์เมื่อ $i=1,2,\ldots,p+1$ โดยที่ $(.)^+$ คือ อินเวอร์ส Moore – Penrose แต่ถ้า $h'(1)>0$ หมายความว่าเมื่อ $d<1$ จะทำให้ $h(d)<h(1)$ ซึ่งก็คือ ค่า $MSE(\underset{\sim}{\beta}_d^*(d))<MSE(\underset{\sim}{\beta}^*)$ ซึ่งการเปรียบเทียบดังกล่าวจะนำไปสู่ทฤษฎีดังต่อไปนี้

1. ถ้า $h'(1)>0$ หรือ ถ้า

$$tr\left\{(S+I)^{-1}S^{-1}\underset{\sim}{\delta}^*(\underset{\sim}{\beta}+S^{-1}\underset{\sim}{\delta}^*)'\right\}>-\sigma^2 tr\left\{(S+I)^{-1}A\right\}$$

ดังนั้น พารามิเตอร์ d ที่ $d<1$ ตัวประมาณลิวที่ถูกจำกัดมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุดที่ถูกจำกัด เงื่อนไขที่ทำให้ $h'(1)>0$ ได้นั้น $1+\alpha_i\lambda_i\tilde{\delta}_i^+$ จะต้องมากกว่าหรือเท่ากับศูนย์เมื่อ $i=1,2,\ldots,p+1$

2. ถ้า $h'(1)<0$ หรือ ถ้า

$$tr\left\{(S+I)^{-1}S^{-1}\underset{\sim}{\delta}^*(\underset{\sim}{\beta}+S^{-1}\underset{\sim}{\delta}^*)'\right\}<-\sigma^2 tr\left\{(S+I)^{-1}A\right\}$$

ดังนั้น พารามิเตอร์ d ที่ $d>1$ ตัวประมาณลิวที่ถูกจำกัดมีค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุดที่ถูกจำกัด

3. ถ้า $h'(1)=0$ หรือ ถ้า

$$tr\left\{(S+I)^{-1}S^{-1}\underset{\sim}{\delta}^*(\underset{\sim}{\beta}+S^{-1}\underset{\sim}{\delta}^*)'\right\}=-\sigma^2 tr\left\{(S+I)^{-1}A\right\}$$

ดังนั้น ตัวประมาณลิวที่ถูกจำกัดไม่สามารถให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองต่ำกว่าตัวประมาณกำลังสองน้อยสุดที่ถูกจำกัด

จากที่กล่าวมาข้างต้น $MSE(\underset{\sim}{\beta}_d^*(d))<MSE(\underset{\sim}{\beta}^*)$ ถ้าพารามิเตอร์ d มีค่าอยู่ในช่วงเปิดที่มีจุดปิดดังนี้

$$d_1=1 \quad \text{และ} \quad d_2=1-\frac{2tr\left\{(S+I)^{-1}(\sigma^2 A+S^{-1}\underset{\sim}{\delta}^*\underset{\sim}{\beta}'+S^{-2}\underset{\sim}{\delta}^*\underset{\sim}{\delta}^{*\prime})\right\}}{tr\left\{(S+I)^{-2}(\sigma^2 A+\underset{\sim}{\beta}\underset{\sim}{\beta}'+2S^{-1}\underset{\sim}{\delta}^*\underset{\sim}{\beta}'+S^{-2}\underset{\sim}{\delta}^*\underset{\sim}{\delta}^{*\prime})\right\}}$$

เมื่อพิจารณาจุดปิด d_2 จุดปิด d_2 สามารถมากกว่า หรือ เท่ากับ หรือ น้อยกว่า $d_1 = 1$