# บทที่ 2

## ทฤษฎีที่เกี่ยวข้องกับการวิจัย

การวิเคราะห์ความถดถอยพหุคือ การวิเคราะห์รูปแบบหรือสมการของความสัมพันธ์ของตัวแปรอิสระกับตัวแปรตาม โดยแสดงให้เห็นว่าตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กับตัวแปรตามมากหรือน้อยเพียงใด และความสัมพันธ์เป็นไปในเชิงบวกหรือเชิงลบ การวิเคราะห์โดยใช้เทคนิคดังกล่าวจะต้องมีข้อสมมติที่ว่า ตัวแปรอิสระแต่ละตัวต้องเป็นอิสระซึ่งกันและกัน แต่จากที่ได้กล่าวมาแล้วในตอนที่ 1.1 กรณีที่มีการพิจารณาตัวแปรอิสระหลาย ๆ ตัว เป็นไปได้ยากที่ตัวแปรอิสระจะเป็นอิสระซึ่งกันและกัน ดังนั้น การหลีกเลี่ยงปัญหาดังกล่าวโดยการตัดตัวแปรใดตัวแปรหนึ่งออกไปอาจทำให้ค่าประมาณสัมประสิทธิ์ความถดถอยพหุที่ได้คลาดเคลื่อนไป และความคลาดเคลื่อนนี้จะมากหากตัวแปรที่ถูกตัดทิ้งไปมีความสัมพันธ์สูงกับตัวแปรอื่น ๆ ที่มีอยู่ในสมการ

ต่อไปจะขอกล่าวถึงการประมาณค่าสัมประสิทธิ์การถดถอยพหุของตัวแบบการถดถอยเชิงเส้นด้วยวิธีริดจ์รีเกรสชั่นและวิธีลาเท้นรูทธ์รีเกรสชั่น ซึ่งมีรายละเอียดต่าง ๆ ดังนี้

### 2.1 ตัวแบบทั่วไป

ตัวแบบทั่วไปของการวิเคราะห์การถดถอยพหุเชิงเส้นสามารถเขียนได้ ดังนี้

$$\tilde{y} = X\tilde{\beta} + \tilde{\varepsilon} \qquad (2.1.1)$$

เมื่อ $\tilde{y}$ คือ เมตริกซ์ของตัวแปรตามขนาด $n \times 1$

$X$ คือ เมตริกซ์ของตัวแปรอิสระขนาด $n \times p$

$\tilde{\beta}$ คือ เวกเตอร์ของพารามิเตอร์ที่ไม่ทราบค่าขนาด $p \times 1$

$\underset{\sim}{\varepsilon}$ คือ เวกเตอร์ของความคลาดเคลื่อนที่เกิดขึ้น ซึ่งมีขนาด $n \times 1$

$n$ คือ ขนาดตัวอย่าง

$p$ คือ จำนวนตัวแปรอิสระ

จากตัวแบบ (2.1.1) ประมาณค่าพารามิเตอร์ $\underset{\sim}{\beta}$ ด้วยวิธีกำลังสองน้อยที่สุดจะได้ค่าประมาณ

$$\underset{\sim}{\hat{\beta}} = (x'x)^{-1}x'y \qquad (2.1.2)$$

ซึ่งการประมาณ $\beta$ โดยวิธีนี้จะได้ค่าตัวประมาณ $\beta$ ที่มีคุณสมบัติไม่เอนเอียง กล่าวคือ

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

ตัวประมาณ $\beta$ จะให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองน้อยที่สุดในบรรดาตัวประมาณที่ไม่เอนเอียง ดังนั้นค่าความผิดพลาดเขียนได้ในรูปของ

$$\underset{\sim}{e} = \underset{\sim}{y} - x\underset{\sim}{\beta}$$

ซึ่งค่าผลบวกของความคลาดเคลื่อนกำลังสองมีค่าเท่ากับ $\underset{\sim}{e}'\underset{\sim}{e}$ ถ้าให้ $\phi(\underset{\sim}{\hat{\beta}})$ เป็นผลบวกความคลาดเคลื่อนกำลังสองเมื่อใช้ตัวประมาณ $\underset{\sim}{\beta}$ จะได้

$$\phi(\underset{\sim}{\hat{\beta}}) = (\underset{\sim}{y} - x\underset{\sim}{\hat{\beta}})'(\underset{\sim}{y} - x\underset{\sim}{\hat{\beta}}) \qquad (2.1.3)$$

เนื่องจากข้อมูลตัวอย่างที่นำมาศึกษาการถดถอยพหุเชิงเส้นส่วนมากมีตัวแปรอิสระที่มีความสัมพันธ์ภายในต่อกันในระดับต่างๆ ซึ่งจะมีผลทำให้ดีเทอร์มิแนนท์ของเมตริกซ์ $x'x$ มีค่าเข้าใกล้ 0 เพราะฉะนั้นถ้าประมาณ $\underset{\sim}{\beta}$ ด้วย $\underset{\sim}{\hat{\beta}}$ จากวิธีกำลังสองน้อยที่สุด อาจมีผลทำให้เมตริกซ์ของค่าความแปรปรวนและค่าความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\beta}$ ไม่มีขนาดเล็กที่สุด เราจึงควรพิจารณา

คุณสมบัติของตัวประมาณค่า $\underset{\sim}{\beta}$ 2 กรณี คือเมตริกซ์ค่าความแปรปรวนและค่าความแปรปรวนร่วมของตัวประมาณ $\underset{\sim}{\beta}$ และค่าเฉลี่ยของกำลังสองระยะทางจาก $\underset{\sim}{\beta}$ ไปยัง $\beta$ ซึ่งสามารถเขียนอยู่ในรูปฟังก์ชันของ $X'X$ และ $\sigma^2$ ได้ดังต่อไปนี้

กำหนดให้ความแปรปรวนของเวกเตอร์ตัวประมาณ $\underset{\sim}{\hat{\beta}}$ มีค่าเป็น $VAR(\underset{\sim}{\hat{\beta}})$ ดังนั้น

$$Var(\underset{\sim}{\hat{\beta}}) = \sigma^2 (X'X)^{-1} \qquad (2.1.4)$$

และให้ระยะทางจาก $\underset{\sim}{\hat{\beta}}$ ถึง $\underset{\sim}{\beta}$ มีค่าเป็น $L_1$ ดังนั้น

$$L_1^2 = (\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})'(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta}) \qquad (2.1.5)$$

ค่าเฉลี่ยกำลังสองของระยะทางจาก $\underset{\sim}{\hat{\beta}}$ ไปยัง $\underset{\sim}{\beta}$ เป็นดังนี้

$$E(L_1^2) = \sigma^2 \, Trace(X'X)^{-1} \qquad (2.1.6)$$

$E(L_1^2)$ จะมีค่าสมมูล (Equivalent) กับ $E(\beta\beta)$

$$E(\underset{\sim}{\hat{\beta}}'\underset{\sim}{\hat{\beta}}) = \underset{\sim}{\hat{\beta}}'\underset{\sim}{\hat{\beta}} + \sigma^2 \, Trace(X'X)^{-1} \qquad (2.1.7)$$

เมื่อ $\epsilon$ มีการแจกแจงแบบปกติจะได้

$$Var(L_1^2) = 2\sigma^4 \, Trace(X'X)^{-2} \qquad (2.1.8)$$



จะเห็นได้ว่าจากสมการ (2.1.4),(2.1.5)และ(2.1.6) ค่าความแปรปรวนของเวกเตอร์ตัวประมาณ $\underset{\sim}{\hat{\beta}}$ ($Var(\underset{\sim}{\hat{\beta}})$) ค่าเฉลี่ยของกำลังสองของระยะทางจาก $\underset{\sim}{\beta}$ ไปยัง $\underset{\sim}{\hat{\beta}}$ ($E(L^2_1)$) และค่าความแปรปรวนของกำลังสองของระยะทางจาก $\underset{\sim}{\beta}$ ไปยัง $\underset{\sim}{\hat{\beta}}$ ($V(L^2_1)$) เหล่านี้เป็นฟังก์ชันของ $\underset{\sim}{\beta}$ เพื่อความสะดวกในการเปรียบเทียบการประมาณค่า $\underset{\sim}{\beta}$ ควรแปลงเมตริกซ์ $(X'X)^{-1}$ ให้อยู่ในรูปฟังก์ชันของค่าไอเก็นแวลลิว (eigenvalue) ของเมตริกซ์ $X'X$ โดยใช้คุณสมบัติที่สำคัญข้อหนึ่งของค่าไอเก็นแวลลิวของ $X'X$ กล่าวคือ ถ้า $\lambda_i$ เป็นค่าไอเก็นแวลลิว ของ $X'X$ เมื่อ $i$ = 1,2,...,p เมื่อ p คือ จำนวนตัวแปรอิสระ แล้ว $\Sigma \lambda_i$ = trace(XX)

$$(\lambda_{max} = \lambda_1) > \lambda_2 > \lambda_3 > \ldots > (\lambda_p = \lambda_{min}) > 0$$

จากสมการ (2.1.6) ค่าเฉลี่ยความผิดพลาดยกกำลังสองระหว่าง $\underset{\sim}{\beta}$ กับ $\hat{\underset{\sim}{\beta}}$ (risk function) อาจเขียนอยู่ในรูปฟังก์ชันของค่าไอเก็นแวลลิวของ $X'X$ ดังนี้

$$E[L_1^2] = \sigma^2 \sum_{i=1}^{p} (1/\lambda_i) \qquad (2.1.9)$$

และจากสมการ (2.1.7) ค่า $Var[L_1^2]$ อาจเขียนอยู่ในรูปฟังก์ชันของค่าไอเก็นแวลลิวของ $X'X$ ดังนี้

$$Var[L_1^2] = 2\sigma^4 \sum_{i=1}^{p} (1/\lambda_i)^2 \qquad (2.1.10)$$

จากสมการ (2.1.9) และสมการ (2.1.10) จะได้ค่าเฉลี่ยกำลังสองของระยะทางจาก $\underset{\sim}{\beta}$ ไปยัง $\hat{\underset{\sim}{\beta}}$ หรือ $E(L^2_1)$ และค่าความแปรปรวนกำลังสองของระยะทางจาก $\underset{\sim}{\beta}$ ไปยัง $\hat{\underset{\sim}{\beta}}$ หรือ $Var[L^2_1]$ มีค่าเป็น $\sigma^2/\lambda_{min}$ และ $2\sigma^4/\lambda^2_{min}$ ตามลำดับ

แต่ในกรณีที่ตัวแปรอิสระเกิดมีสภาพที่ไม่เหมาะสม (เกิดความสัมพันธ์ระหว่างตัวแปรอิสระในอัตราที่สูง) ค่าไอเก็นแวลลิวของเมตริกซ์ $X'X$ จะมีค่าต่ำมาก (เนื่องจาก $X'X$ มีค่าเท่ากับผลคูณของค่าไอเก็นแวลลิว และมีค่าต่ำเมื่อระดับความสัมพันธ์ของตัวแปรอิสระอยู่ในอัตราที่สูง) ซึ่งมีผลทำให้ระยะทางจาก $\underset{\sim}{\beta}$ ไปยัง $\hat{\underset{\sim}{\beta}}$ มีค่ามากพิจารณาค่า $E(L^2_1)$ และ $Var[L^2_1]$ จะเห็นว่าค่าทั้งสองสูงขึ้น

## 2.2 การประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยวิธีริดจ์รีเกรสชั่น

Hoerl และ Kennard (1970) ได้เสนอวิธีริดจ์รีเกรสชั่นเพื่อแก้ปัญหาพหุสัมพันธ์ ซึ่งวิธีนี้จะให้ค่าสัมประสิทธิ์ความถดถอยพหุที่ให้ค่าเฉลี่ยความคลาดเคลื่อนต่ำลง โดยอาศัยหลักการที่ว่าเมื่อ $|X'X|$ มีค่าเข้าใกล้ศูนย์ จะมีผลทำให้ไม่สามารถหาค่า $(X'X)^{-1}$ และผลบวกกำลังสองของ $\underset{\sim}{\beta}$ มีค่ามาก

ผิดจากความเป็นจริง ดังนั้นจึงพยายามทำให้ $|x'x|$ เพิ่มขึ้น โดยการบวกค่าคงที่ k ที่มากกว่าศูนย์ เข้ากับสมาชิกทุกตัวบนเส้นทแยงมุม ซึ่งจะช่วยทำให้คาแรคเตอร์ริสติกรูทซ์ มีค่ามากขึ้น และผลบวกกำลังสองของ $\underset{\sim}{\hat{\beta}}$ ลดลง

ค่าประมาณสัมประสิทธิ์ความถดถอยพหุด้วยวิธีริดจ์รีเกรสชั่น คือ

$$\underset{\sim}{\hat{\beta}}_R = (x'x + kI)^{-1}x'y \quad ; \quad k > 0 \qquad (2.2.1)$$

และให้ $[x'x + kI]^{-1} = W$ ดังนั้นจากสมการ(2.2.1) จะได้ว่า

$$\underset{\sim}{\hat{\beta}}_R = Wx'y \qquad (2.2.2)$$

เราสามารถจัดตัวประมาณค่า $\underset{\sim}{\hat{\beta}}_R$ ให้อยู่ในรูปของ $\underset{\sim}{\hat{\beta}}$ ได้ดังนี้

$$\underset{\sim}{\hat{\beta}}_R = [I + k(x'x)^{-1}]^{-1}\underset{\sim}{\hat{\beta}} \qquad (2.2.3)$$

$$= Z\underset{\sim}{\hat{\beta}} \qquad (2.2.4)$$

โดยที่ $Z = [I + k(x'x)^{-1}]^{-1}$

ให้ $E_i(W)$ และ $E_i(Z)$ เป็นค่าไอเก้นแวลลิวของ W และ Z ตามลำดับ ซึ่งได้จากการแก้สมการคาแรคเตอร์ริสติก (Characteristic equations)

$$|W - E_iI| = 0$$

และ $$|Z - E_iI| = 0$$

$$E_i(W) = 1/(\lambda_i + k) \qquad (2.2.5)$$

$$E_i(Z) = \lambda_i/(\lambda_i + k) \qquad (2.2.6)$$

เราอาจเขียน Z ให้อยู่ในรูปฟังก์ชันของ W ได้ดังนี้

$$Z = I - k(X'X + kI)^{-1}$$
$$= I - kW \qquad (2.2.7)$$

จากสมการ (2.2.3) ผลบวกกำลังสองของ $\hat{\beta}_R$ จะมีค่าน้อยกว่าผลบวกกำลังสองของ $\hat{\beta}$ เมื่อ $k > 0$ นั่นคือ

$$\hat{\beta}_R'\hat{\beta}_R < \hat{\beta}\hat{\beta} \qquad (2.2.8)$$

ซึ่งสามารถพิสูจน์ข้อความข้างต้นได้ดังนี้

จากนิยาม $\hat{\beta}_R = Z\hat{\beta}$ โดยที่ $X'X$ และ $Z$ มีคุณสมบัติเป็นเมตริกซ์สมมาตรที่เป็นบวกแน่นอน (symmetric positive definite) จะได้ว่า

$$\hat{\beta}_R'\hat{\beta}_R = (Z\hat{\beta})'(Z\hat{\beta})$$
$$= \sum_{i=1}^{p} E_i^2(Z)\hat{\beta}\hat{\beta}$$
$$< E^2_{i\,(max)}(Z)\hat{\beta}'\hat{\beta} \qquad (2.2.9)$$

เมื่อ $E_{max}(Z) = \lambda_{max} / (\lambda_{max} + k)$ ; $\lambda_{max}$ เป็น eigenvalue ที่มีค่ามากที่สุดของเมตริกซ์ $X'X$ ดังนั้น

$$\hat{\beta}_R'\hat{\beta}_R < \hat{\beta}'\hat{\beta}$$



เนื่องจากผลบวกของความคลาดเคลื่อนกำลังสองของค่าตัวประมาณ $\beta$ เป็นดังนี้

$$\phi^*(\hat{\beta}_R) = (y - X\hat{\beta}_R)(y - X\hat{\beta}_R)$$

จากการประมาณค่าพารามิเตอร์ $\beta$ ด้วยวิธีกำลังสองน้อยที่สุด เราสามารถเขียน $\phi(\hat{\beta})$ ซึ่งเป็นผลบวกของความคลาดเคลื่อนกำลังสองของตัวประมาณ $\hat{\beta}$ เป็น $y'y - \hat{\beta}'X'y$ ดังนั้น เพื่อให้เห็นความแตกต่างของ $\phi^*(\hat{\beta}_R)$ และ $\phi(\hat{\beta})$ เราสามารถเขียน $\phi^*(\hat{\beta}_R)$ ในเทอม

ของ $y'y$ และ $x'y$ ได้ดังนี้

$$\phi^*(\hat{\beta}_R) = y'y - (\hat{\beta}_R)x'y - k(\hat{\beta}_R)'(\hat{\beta}_R)$$

จาก $\phi(\hat{\beta})$ และ $\phi^*(\hat{\beta}_R)$ สรุปได้ว่าผลบวกของความคลาดเคลื่อนกำลังสองของวิธีริดจ์-รีเกรสชั่น จะให้ค่าน้อยกว่าผลบวกของความคลาดเคลื่อนกำลังสองที่คำนวณโดยวิธีกำลังสองน้อยที่สุด

$\hat{\beta}_R$ เป็นการแปลงเชิงเส้น (linear transformation) ของตัวประมาณกำลังสองน้อยที่สุด กล่าวคือ

$$\begin{aligned}
\hat{\beta}_R &= (x'x + kI)^{-1}x'y \\
&= (x'x + kI)^{-1}(x'x)(x'x)^{-1}x'y \\
&= (x'x + kI)^{-1}x'x[(x'x)^{-1}x'y] \\
&= ([I + k(x'x)^{-1}]^{-1})\beta
\end{aligned}$$

โดยที่ $\hat{\beta}_R$ สัมพันธ์กับ $\hat{\beta}$ ในลักษณะที่ว่า $\hat{\beta}_R = [I + k(x'x)^{-1}]^{-1}\hat{\beta}$ และ เมื่อ $k = 0$ จะได้ว่า $\hat{\beta}_R = \hat{\beta}$ กล่าวคือ ตัวประมาณกำลังสองจะเป็นกรณีเฉพาะของตัวประมาณริดจ์รีเกรสชั่น เมื่อ $k = 0$ แต่ $\hat{\beta}_R$ จะเป็นตัวประมาณที่เอนเอียงของ $\beta$ โดยที่

$$E(\hat{\beta}_R) = E(Z\hat{\beta}) = Z\beta$$



เนื่องจาก $E(\hat{\beta}) = \beta$ ดังนั้นการประมาณค่าพารามิเตอร์ $\beta$ ด้วยวิธีกำลังสองน้อยที่สุดจะได้ตัวประมาณค่า $\beta$ ที่ไม่เอนเอียง แต่ในการประมาณค่าพารามิเตอร์ $\hat{\beta}_R$ ด้วยวิธีริดจ์รีเกรสชั่น จะเป็นตัวประมาณที่เอนเอียง ดังสมการข้างต้น



ค่าความคลาดเคลื่อนกำลังสองของตัวประมาณสัมประสิทธิ์ความถดถอยพหุโดยวิธีริดจ์รีเกรส-ชั่นเป็นดังนี้

$$\begin{aligned}
\mathrm{Cov}(\hat{\beta}_R) &= E[\hat{\beta}_R - E(\hat{\beta}_R)][\hat{\beta}_R - E(\hat{\beta}_R)]' \\
&= E[\hat{\beta}_R - Z\hat{\beta}][\hat{\beta}_R - Z\hat{\beta}]'
\end{aligned}$$

$$= E[Z(\hat{\beta} - \beta)][Z(\hat{\beta} - \beta)]'$$

$$= E[Z(\hat{\beta} - \beta)][(\hat{\beta} - \beta)'(Z)']$$

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

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

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

$$Cov(\hat{\beta}_R) = E[Z(X'X)^{-1}X'\varepsilon(\varepsilon'X(X'X)^{-1}Z]'$$

$$= Z(X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1}Z'$$

$$= \sigma^2 Z(X'X)^{-1}Z'$$

จากสมการที่ (2.2.3) และ (2.2.4) จะได้ว่า

$$Z = [I+k(X'X)^{-1}]^{-1}$$

$$= [I+k(X'X)^{-1}]^{-1}(X'X)^{-1}(X'X)$$

$$= \{X'X[I+k(X'X)^{-1}]\}^{-1}(X'X)$$

$$= (X'X + kI)^{-1}(X'X)$$

ดังนั้น

$$Cov(\hat{\beta}_R) = \sigma^2 Z(X'X)^{-1}Z'$$

$$= \sigma^2(X'X + kI)^{-1}X'X(X'X)^{-1}(X'X)(X'X + kI)^{-1}$$

$$= \sigma^2(X'X + kI)^{-1}(X'X)(X'X + kI)^{-1} \quad (2.2.10)$$

$$MSE(\hat{\beta}_R) = trace\ Cov(\hat{\beta}_R) + \beta'(Z - I)'(Z - I)\beta$$

$$= Variance + (bias)^2$$

$$= \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2\beta'[X'X + kI]^{-2}\beta \quad (2.2.11)$$

สามารถพิสูจน์ (2.2.11) ได้ดังนี้

$$MSE(\hat{\beta}_R) = E[(\hat{\beta}_R-\beta)'(\hat{\beta}_R-\beta)]$$

$$\hat{\beta}_R = [I+k(x'x)^{-1}]^{-1}\hat{\beta} = Z\beta$$
$$MSE(\hat{\beta}_R) = E[Z\hat{\beta} - \beta]'[Z\hat{\beta} - \beta]$$

พิจารณาจาก $Z\hat{\beta} - \beta$ จะได้ว่า $Z\hat{\beta} - \beta = Z\hat{\beta}-Z\beta+Z\beta-\beta$
$$= Z(\hat{\beta}-\beta)+(Z-I)\beta$$

นั่นคือ $(Z\hat{\beta} - \beta)' = (\hat{\beta} - \beta)'Z' + \beta'(Z - I)'$

ดังนั้น
$$MSE(\hat{\beta}_R) = E((\hat{\beta}-\beta)'Z'+\beta'(Z-I)')[Z(\hat{\beta}-\beta)+(Z-I)\beta]$$
$$= E[(\hat{\beta}-\beta)'Z'Z(\hat{\beta}-\beta)+(\hat{\beta}-\beta)'Z'(Z-I)\beta$$
$$+\beta'(Z-I)'Z(\hat{\beta}-\beta)+\beta'(Z-I)'(Z-I)\beta]$$

เนื่องจาก $E(\hat{\beta}-\beta) = 0$

ดังนั้น
$$MSE(\hat{\beta}_R) = E[(\hat{\beta}-\beta)'Z'Z(\hat{\beta}-\beta)]+\beta'(Z-I)'(Z-I)\beta$$
$$= a + b$$

โดยที่ $a = E(\hat{\beta}-\beta)'Z'Z(\hat{\beta}-\beta)$

จะได้ว่า $\hat{\beta}-\beta = \beta+(x'x)^{-1}x'\varepsilon -\beta$

เมื่อ
$$a = E(\hat{\beta}-\beta)'Z'Z(\hat{\beta}-\beta)$$
$$= E[\varepsilon'x(x'x)^{-1}Z'Z(x'x)^{-1}x'\varepsilon]$$
$$= \sigma^2 trace[x(x'x)^{-1}Z'Z(x'x)^{-1}x']$$
$$= \sigma^2 trace[(x'x)^{-1}Z'Z] \qquad (2.2.12)$$

โดยที่ $Z = [I+k(x'x)^{-1}]^{-1}$ เป็นเมตริกซ์สมมาตร

$$(x'x)^{-1}Z'Z = (x'x)^{-1}ZZ = (x'x)^{-1}[I+k(x'x)^{-1}]^{-1}[I+k(x'x)^{-1}]^{-1}$$
$$= [\{I+k(x'x)^{-1}\}(x'x)]^{-1}[I+k(x'x)^{-1}]^{-1}$$
$$= (x'x+kI)^{-1}(x'x)(x'x)^{-1}\{I+k(x'x)^{-1}\}^{-1}$$
$$= (x'x+kI)^{-1}(x'x)[\{I+k(x'x)^{-1}\}(x'x)]^{-1}$$

$$= (X'X+kI)^{-1}(X'X)(X'X+kI)^{-1}$$

กล่าวคือ จาก(2.2.12) และ (2.2.13) จะได้

$$\begin{aligned} MSE(\hat{\underset{\sim}{\beta}}_R) &= \sigma^2 trace[(X'X+kI)^{-1}(X'X)(X'X+kI)^{-1}] \\ &\quad + \underset{\sim}{\beta}'(Z-I)'(Z-I)\underset{\sim}{\beta} \\ &= trace[Cov(\hat{\underset{\sim}{\beta}}_R)]+\beta'(Z-I)'(Z-I)\underset{\sim}{\beta} \\ &= trace[Cov(\hat{\underset{\sim}{\beta}}_R)]+(bias)^2 \end{aligned}$$

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง คือ

$$MSE(\hat{\underset{\sim}{\beta}}_R) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k)^2} + k^2 \underset{\sim}{\beta}_R'(X'X+kI)^{-2}\underset{\sim}{\beta}_R \quad (2.2.14)$$

เมื่อ $\lambda_i$ คือ ไอเก็นแวลลิวของ $X'X$

$\sigma^2$ คือ ค่าความแปรปรวนที่ได้จากวิธีกำลังสองน้อยที่สุด ซึ่งเท่ากับ

$$\frac{1}{n-p}[\,y'y - \hat{\underset{\sim}{\beta}}'X'y\,]$$

$\hat{\underset{\sim}{\beta}}$ คือ ค่าสัมประสิทธิ์ความถดถอยที่ได้จากวิธีกำลังสองน้อยที่สุด ซึ่งเท่ากับ $(X'X)^{-1}X'y$



จากการประมาณค่าสัมประสิทธิ์ความถดถอย $\hat{\underset{\sim}{\beta}}$ ด้วยวิธีกำลังสองน้อยที่สุดจะได้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ดังนี้

$$\begin{aligned} MSE(\hat{\underset{\sim}{\beta}}) &= Var(\hat{\underset{\sim}{\beta}}) + (bias)^2 \\ &= \sigma^2(X'X)^{-1} + 0 \end{aligned}$$

สำหรับค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณสัมประสิทธิ์การถดถอยพหุด้วยวิธีริดจ์รีเกรสชั่น ซึ่งมีค่าเป็น $E[L^2_1(k)]$ อาจพิจารณาได้ดังนี้

$$
\begin{aligned}
E[L^2_1(k)] &= E[(\hat{\underset{\sim}{\beta}}_R - \underset{\sim}{\beta})'(\hat{\underset{\sim}{\beta}}_R - \underset{\sim}{\beta})] \\
&= E[(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})'Z'Z(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta})] + (Z\underset{\sim}{\beta} - \underset{\sim}{\beta})'(Z\underset{\sim}{\beta} - \underset{\sim}{\beta}) && (2.2.11) \\
&= \sigma^2 \operatorname{trace}(x'x)Z'Z + \underset{\sim}{\beta}'(Z - I)'(Z - I)\underset{\sim}{\beta} && (2.2.12) \\
&= \sigma^2[\operatorname{trace}(x'x + kI)^{-1} - k \operatorname{trace}(x'x + kI)^{-2}] && (2.2.13) \\
&\quad + k^2\underset{\sim}{\beta}'(x'x + kI)^{-2}\underset{\sim}{\beta} \\
&= \sigma^2 \Sigma\, \lambda_i/(\lambda_i + k)^{-2} + k^2\underset{\sim}{\beta}'(x'x + kI)^{-2}\underset{\sim}{\beta} && (2.2.14) \\
&= r_1(k) + r_2(k) && (2.2.15)
\end{aligned}
$$

จะเห็นว่า $E[L_1^2(k)]$ เป็นค่าเฉลี่ยความคลาดเคลื่อนกำลังสองซึ่งอยู่ในรูปผลบวกของฟังก์ชัน $r_1(k)$ และ $r_2(k)$

โดยที่ $r_1(k)$ เป็นค่าความแปรปรวนของตัวประมาณ $\hat{\underset{\sim}{\beta}}_R$ $(Var(\hat{\underset{\sim}{\beta}}_R'))$ เมื่อ พิจารณาเทอม $r_2(k)$ ซึ่งเป็นระยะทางจาก $Z\hat{\underset{\sim}{\beta}}$ ไปยัง $\hat{\underset{\sim}{\beta}}$ จะเห็นว่าเป็นศูนย์เมื่อ $k = 0$ โดยที่ Z เป็นเมตริกซ์เอกลักษณ์การคูณ ซึ่งสอดคล้องกับค่าความเอนเอียงกำลังสองของ $\hat{\underset{\sim}{\beta}}$ ที่มีค่าเป็นศูนย์ สำหรับกรณีที่ $k > 0$ อาจพิจารณา $r_2(k)$ ในเทอมของความเอนเอียงของ $\hat{\underset{\sim}{\beta}}$ กำลังสอง กล่าวคือ $r_2(k) = (bias^2(\hat{\underset{\sim}{\beta}}_R))$

ส่วนเทอม $r_1(k)$ นั้นเป็นผลบวกของความแปรปรวนของตัวประมาณค่า $\hat{\underset{\sim}{\beta}}_R$ ซึ่งเขียนในเทอมของ y ได้ดังนี้



$$
\begin{aligned}
\hat{\underset{\sim}{\beta}}_R &= Z\hat{\underset{\sim}{\beta}} \\
&= Z(x'x)^{-1}x'y && (2.2.16) \\
V(\hat{\underset{\sim}{\beta}}_R) &= Z(x'x)^{-1}x'\,\sigma^2\,x(x'x)^{-1}Z' \\
&= \sigma^2 \operatorname{trace} Z(x'x)^{-1}Z' && (2.2.17)
\end{aligned}
$$

เนื่องจากความแปรปรวนของตัวประมาณ $\hat{\underset{\sim}{\beta}}_R$ ทั้งหมด $V(\hat{\underset{\sim}{\beta}}_R)$ คำนวณได้จากผลบวกของสมาชิกทุกตัวบนเส้นทแยงมุมของเมตริกซ์ $\sigma^2 Z(x'x)^{-1}Z$ เพื่อให้สอดคล้องกับ $r_1(k)$ ของสมการ (2.2.15) ซึ่งเป็นฟังก์ชันที่อยู่ในเทอมของไอเก็นแวลลิวของ $x'x$ และ k ดังนั้นเมตริกซ์ $x'x$ และ

Z สามารถปรับให้เป็นเมตริกซ์อยู่ในรูปของ $\lambda_i$ ซึ่งเป็นค่าไอเก็นแวลลิว ของ $X'X$ เมื่อ $i = 1,2,\ldots,p$ และ p หมายถึง จำนวนตัวแปรอิสระ x

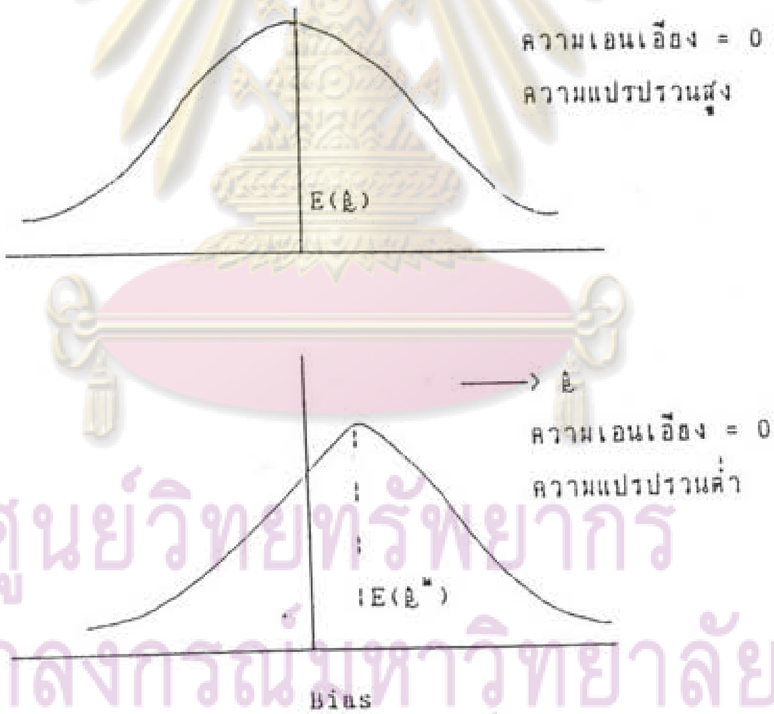
$$\mathrm{trace}(X'X)^{-1} = \Sigma\,(1/\lambda_i) \tag{2.2.18}$$

$$\mathrm{trace}(Z) = \Sigma\,\lambda_i/(\lambda_i+k) \tag{2.2.19}$$

จากสมการ (2.2.17) จะได้ว่า

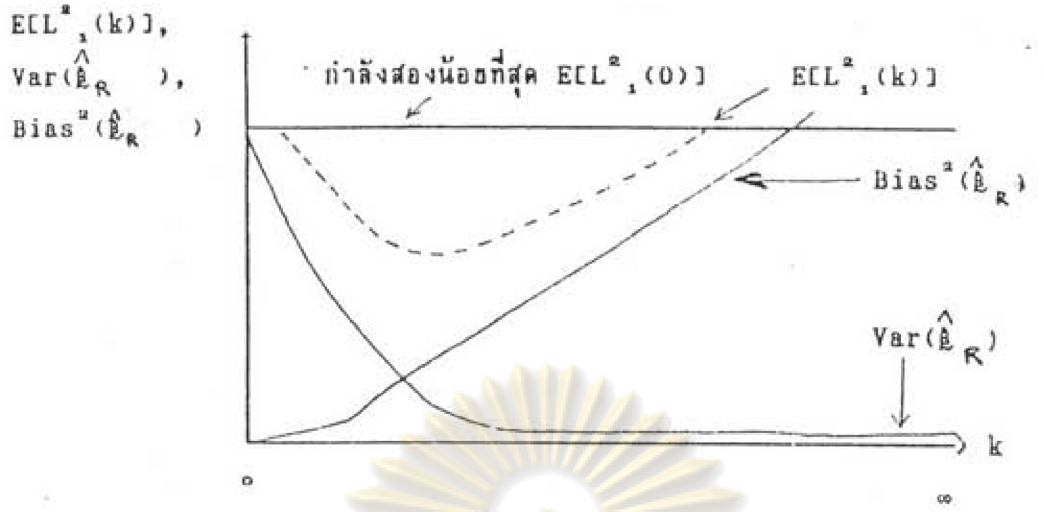
$$\sigma^2\,\mathrm{trace}\;Z(X'X)^{-1}Z' = \sigma^2\,\Sigma\,\lambda_i/(\lambda_i + k)^2$$

ดังนั้น $E[L^2_1(k)] = \mathrm{Var}(\hat{\beta}_R) + \mathrm{Bias}^2(\hat{\beta}_R)$




รูปที่ 2.2.1

แสดงการกระจายของตัวประมาณพารามิเตอร์ $\beta$ ของกำลังสองน้อยที่สุดและริดจ์รีเกรสชั่น




รูปที่ 2.2.2

แสดงกราฟของ $E[L^2_1(k)]$ , $[MSE(\hat{\underset{\sim}{\beta}}_R)]$ , $Var(\hat{\underset{\sim}{\beta}}_R)$ , $Bias^2(\hat{\underset{\sim}{\beta}}_R)$

จากรูปที่ 2.2.2 เป็นกราฟแสดงความสัมพันธ์ระหว่างความแปรปรวนของตัวประมาณ $\hat{\underset{\sim}{\beta}}_R$ ความเอนเอียงกำลังสองของตัวประมาณ เมื่อกำหนดให้ k (0,∝] กล่าวคือความแปรปรวนจะมีค่าลดลงเรื่อยๆ เมื่อ k มีค่าเพิ่มขึ้นและความเอนเอียงกำลังสองจะมีค่าแปรผันตาม k รูปเส้นไข่ปลาเป็นกราฟที่แสดง ค่า $E[L^2_1(k)]$ ซึ่งเป็นค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่เกิดจากตัวประมาณค่า $\hat{\underset{\sim}{\beta}}_R$ เมื่อ k > 0

กราฟของฟังก์ชัน $r_1(k)$ จะมีลักษณะเป็นฟังก์ชันที่ลดลง (monotonically decreasing function) ในรูปฟังก์ชันของ k และกราฟของฟังก์ชัน $r_2(k)$ จะมีลักษณะเป็นฟังก์ชันที่เพิ่มขึ้น (monotonically increasing function) ในรูปฟังก์ชันของ k จากรูปที่ 2.2.2 จะเห็นว่า $E[L^2_1(k)]$ มีค่าน้อยกว่า $E[L^2_1(0)]$ โดยที่ $E[L^2_1(0)]$ เป็นค่าเฉลี่ยความคลาดเคลื่อนกำลังสองที่เกิดจากการประมาณค่าพารามิเตอร์ $\underset{\sim}{\beta}$ โดยกำลังสองน้อยที่สุด

จากรูปที่ 2.2.2 แสดงให้เห็นการเปลี่ยนแปลงของค่าต่างๆที่กล่าวมาแล้วข้างต้น ซึ่งสามารถคำนวณหาขีดจำกัด (Limit) ของความแปรปรวนของตัวประมาณ $\hat{\underset{\sim}{\beta}}_R$ และความเอนเอียงกำลังสองของตัวประมาณ $\hat{\underset{\sim}{\beta}}_R$ ได้จากอนุพันธ์ของแต่ละฟังก์ชันดังนี้

$$\lim_{k \to 0} (\partial r_1 / \partial k) = -2 \sigma^2 \Sigma (1/\lambda^2_i) \qquad (2.2.20)$$

$$\lim_{k \to 0} (\partial r_2 / \partial k) = 0 \qquad (2.2.21)$$

ดังนั้นอนุพันธ์ของความแปรปรวนของตัวประมาณ $\hat{\underset{\sim}{\beta}}_R$ จะมีค่าเป็นลบและจะลู่เข้าสู่ $-2\sigma^2 \Sigma(1/\lambda_i^2)$ เมื่อ $k \to 0$ โดยที่ผลคูณของเมตริกซ์ของตัวแปรอิสระ $(X'X)$ เป็นออร์ธอโกนอลและ $r_1(k)$ มีค่าเข้าสู่ $-\infty$ ในกรณีที่เมตริกซ์ของตัวแปรอิสระมีสภาพไม่เหมาะสมนั้น ค่าไอเก็นแวลลิว $\lambda_p \to 0$

ส่วนอนุพันธ์ของความเอนเอียงกำลังสองน้อยที่สุดมีค่าค่อย ๆ ลดลงและเป็นศูนย์ที่จุดกำเนิด เมื่อ $k \to 0$ คุณสมบัติของสมการ (2.2.20) และ (2.2.21) จะเป็นจริงเมื่อ $k > 0$ ซึ่งสรุปได้ว่าค่าความเอนเอียงเพิ่มขึ้นเรื่อยๆ ขณะที่ $Var(\hat{\underset{\sim}{\beta}}_R)$ ลดลง ซึ่งจะทำให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองลดลงด้วย

## 2.3 สูตรและทฤษฎีที่เกี่ยวข้องในการประมาณค่า k ในวิธีริดจ์รีเกรสชั่น

เนื่องจากการกำหนดค่า k สามารถทำได้หลายวิธีด้วยกัน แต่สำหรับในการวิจัยครั้งนี้ จะศึกษาในกรณีที่ค่า k ประมาณได้จากวิธี Hoerl, Kennard and Baldwin (HKB) และวิธี Lawless and Wang (LW) ซึ่งมีรายละเอียดดังนี้

### 2.3.1 <u>วิธี Hoerl, Kennard and Baldwin</u> (HKB)

วิธีนี้เสนอโดย Hoerl, Kennard และ Baldwin(1975) เนื่องจากว่าวิธีริดจ์รีเกรสชั่น จะให้ค่า MSE ต่ำ ถ้า $k_i = \sigma^2/\underset{\sim}{\beta}'\underset{\sim}{\beta}$ เมื่อต้องการค่า k เพียงค่าเดียว จึงหาค่า k มาจากส่วนเฉลี่ยฮาร์โมนิค (Harmonic Mean) ของ k โดยที่

$$\begin{aligned} 1/k &= \Sigma(1/k_i)/p \\ &= \Sigma(\hat{\underset{\sim}{\beta}}_i^2/\sigma^2)/p \\ &= \Sigma\hat{\underset{\sim}{\beta}}_i^2/(p\,\sigma^2) \end{aligned}$$

จะได้ว่า

$$k = p\,\hat{\sigma}^2/\hat{\underset{\sim}{\beta}}'\hat{\underset{\sim}{\beta}}$$

### 2.3.2 <u>วิธี Lawless and Wang</u> (L.W.)

วิธีนี้เสนอโดย Lawless and Wang (1979) โดย

กำหนดให้
$$k = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p} \theta \hat{\alpha}_j^2}$$

เมื่อ $\hat{\alpha}_j^2 = V'\hat{\beta}$ โดยที่ $V = (v_1, v_2, \ldots, v_p)$ หมายถึงค่าไอเก็นเวกเตอร์ ที่ขึ้นอยู่กับค่าไอเก็นแวลลิวเท่ากับ $\theta_j$

## 2.4 การประมาณค่าสัมประสิทธิ์ความถดถอยพหุโดยวิธีการลาเท้นรูทธ์รีเกรสชั่น

วิธีการนี้เป็นเทคนิคที่ใช้การวิเคราะห์จากลาเท้นรูทธ์ และลาเท้นเวกเตอร์ ของเมตริกซ์สหสัมพันธ์ (Correlation matrix) ที่ได้รวมทั้งคอลัมน์ ของตัวแปรอิสระและตัวแปรตามอยู่ในเมตริกซ์เดียวกัน ดังนั้นจึงกำหนดให้

$Y^*_i$ เป็นตัวแปรตามที่มีการปรับค่า (standardize) แล้ว

และ $X_i$ เป็นตัวแปรอิสระที่มีการปรับค่า (standardize) แล้วเช่นกันเพราะฉะนั้น $A = [\, Y^* : X \,]$ เป็นเมตริกซ์ที่มีขนาด $n \times (k+1)$

$A'A$ เป็นเมตริกซ์สหสัมพันธ์ของตัวแปรตามและตัวแปรอิสระ นอกจากนี้ค่าลาเท้นรูทธ์และลาเท้นเวกเตอร์ มาจาก

$$|A'A - \lambda I| = 0 \text{ และ } (A'A - \lambda_j I)\gamma_j = 0 \text{ ตามลำดับ}$$

โดยที่ $j = 0, 1, \ldots, k$

$\gamma_j' = (\gamma_{0j} \; \gamma_{1j} \; \ldots \; \gamma_{kj})$ เป็นลาเท้นเวกเตอร์ที่ j

และ $\gamma_j^{o\prime} = (\gamma_{1j} \; \gamma_{2j} \; \ldots \; \gamma_{kj})$ ซึ่งจะมีทุกคอลัมน์ของ $\gamma_j$ ยกเว้น คอลัมน์แรก

ให้ $\Gamma = (\gamma_0 \ \gamma_1 \ \dots \ \gamma_k)$ และ $\Lambda = \mathrm{diag}(\lambda_0, \lambda_1, \dots, \lambda_k)$

ดังนั้น $\Gamma'(A'A)\Gamma = \Lambda$ และ $A'A = \Gamma\Lambda\Gamma'$

สังเกตว่า

$$A\gamma_j = \begin{bmatrix} Y^*_1\gamma_{0j} + \sum_{r=1}^{k} X_{1r}\gamma_{rj} \\ Y^*_2\gamma_{0j} + \sum_{r=1}^{k} X_{2r}\gamma_{rj} \\ \vdots \\ Y^*_n\gamma_{0j} + \sum_{r=1}^{k} X_{nr}\gamma_{rj} \end{bmatrix} \qquad (2.4.1)$$

และ

$$\lambda_j = (A\gamma_j)'(A\gamma_j) = \gamma_j'(A'A)\gamma_j$$

$$= \sum_{i=1}^{n} \left(Y^*_i\gamma_{0j} + \sum_{r=1}^{k} X_{ir}\gamma_{rj}\right)^2 \qquad (2.4.2)$$

ถ้า $\lambda_j$ ใด ๆ มีค่าเท่ากับ 0 และ $\gamma_{0j} \neq 0$ แสดงว่ามีความเป็นอิสระเชิงเส้น (Linear dependence) ระหว่างบางคอลัมน์หรือทุกคอลัมน์ของ A แต่ถ้า $\lambda_j$ ใด ๆ มีค่าเท่ากับ 0 และ $\gamma_{0j} = 0$ แสดงว่า มีความเป็นอิสระเชิงเส้น ระหว่างคอลัมน์ของX ซึ่งจะมีความสัมพันธ์เป็นดังนี้

$$\sum_{r=1}^{k} X_{ir}\gamma_{rj} = 0$$

ในทางปฏิบัติจริง จะไม่มีค่าลาเท้นรูทซ์เป็นศูนย์ แต่บางค่าอาจจะมีค่าน้อยมาก ค่าที่น้อยแต่ไม่เป็นศูนย์ของลาเท้นรูทซ์ หมายถึงว่า ดีเทอร์มิแนนท์ของเมตริกซ์นั้นเข้าใกล้ศูนย์

จากสมการ (2.4.1) จะมีสมการ k+1 สมการ และให้ทุกค่าของ $\text{ช}_{oj} \neq 0$ พิจารณาเมื่อ j = 0

$$A\text{ช}_o = \begin{bmatrix} Y^*_1\text{ช}_{oo} + \sum_{r=1}^{k} X_{1r}\text{ช}_{ro} \\ Y^*_2\text{ช}_{oo} + \sum_{r=1}^{k} X_{2r}\text{ช}_{ro} \\ \vdots \\ Y^*_n\text{ช}_{oo} + \sum_{r=1}^{k} X_{nr}\text{ช}_{ro} \end{bmatrix} \qquad (2.4.3)$$

จากสมการ (2.4.2) และข้อกำหนดที่ว่า $\lambda_j = 0$ จะได้ว่า

$$A\text{ช}_j = 0$$

ดังนั้น (2.4.3) จะได้ว่า

$$Y^*_1 \gamma_{oo} + \sum_{r=1}^{k} X_{1r}\gamma_{ro} = 0 \qquad (2.4.4)$$

$$Y^*_2 \gamma_{oo} + \sum_{r=1}^{k} X_{2r}\gamma_{ro} = 0 \qquad (2.4.5)$$

$$\vdots$$

$$Y^*_n \gamma_{oo} + \sum_{r=1}^{k} X_{nr}\gamma_{ro} = 0 \qquad (2.4.6)$$

พิจารณา (2.4.4)

$$Y^*_1 \gamma_{oo} = -\sum_{r=1}^{k} X_{1r}\gamma_{ro}$$

$$Y^*_1 = -\gamma^{-1}_{oo} \sum_{r=1}^{k} X_{1r}\gamma_{ro}$$

$$\frac{(Y_1 - \overline{Y})}{\eta} = -\gamma^{-1}_{oo} \sum_{r=1}^{k} X_{1r}\gamma_{ro}$$

$$Y_1 = \overline{Y} - \eta\gamma^{-1}_{oo} \sum_{r=1}^{k} X_{1r}\gamma_{ro}$$

ตัวประมาณของ $Y_1$ คือ $\overline{Y}_1$

$$\hat{Y}_1 = \overline{Y} - \eta\gamma^{-1}_{oo} \sum_{r=1}^{k} X_{1r}\gamma_{ro}$$

ในกรณีที่ทำการพิจารณาข้อมูลที่ i ใดๆ จะได้ว่า

$$\hat{Y}_i = \bar{Y} - \delta^{-1}_{oo} \sum_{r=1}^{k} X_{ir}\delta_{ro} \qquad (2.4.7)$$

เช่นเดียวกัน เมื่อพิจารณาสมการในการพยากรณ์ตัวแปรที่ j จะได้สมการดังนี้

$$\hat{Y}_{ij} = \bar{Y} - \delta^{-1}_{oj} X \delta^{o}_{j} \quad ; j = 0,1,2,\ldots,k \quad (2.4.8)$$

ซึ่งจากรูปแบบของสมการทั่วไป คือ $\hat{Y} = b_o 1 + Xb$

จาก (2.4.8) จะได้ว่า $b_o = \bar{Y}$ , $b = -\eta \delta^{-1}_{oj} \delta^{o}_{j}$

สำหรับค่าความผันแปรที่เกิดขึ้นจากความไม่แน่นอนของข้อมูลรอบเส้นถดถอย (residual sum of squares) สำหรับวิธีการนี้ คือ

$$\sum_{i=1}^{n} (Y_i - \hat{Y}_{ij})^2 = (Y - \hat{Y}_j)'(Y - \hat{Y}_j) = \eta^2 \lambda^*_j$$

เมื่อ $\lambda^*_j = \lambda_j / \delta^2_{oj}$

<u>ตัวประมาณผลรวมของ $\hat{Y}_i$</u>

$$\hat{Y} = \sum_{j=1}^{k} a_j\delta_{oj}\hat{Y}_j \qquad (2.4.9)$$

โดยที่ $\sum_{j=0}^{k} a_j\delta_{oj} = 1$

**ข้อจำกัด**

$$\hat{Y} = \sum_{j=0}^{k} a_j\delta_{oj}(\bar{Y} - \eta\, \delta^{-1}_{oj}X\delta^{o}_{j})$$

$$\hat{Y} = \bar{Y} - \eta X \sum_{j=0}^{k} a_j \gamma^{o}_{j} \qquad (2.4.10)$$

และจากสมการ (2.4.2) จะได้ว่า

$$b = -\eta \sum_{j=0}^{k} a_j \gamma^{o}_{j}$$

ค่าความผันแปรที่เกิดขึ้นจากความไม่แน่นอนของข้อมูลรอบเส้นถดถอย ของ Y คือ

$$(Y - \hat{Y})'(Y - \hat{Y}) = \eta^2 a' \Lambda a = \eta^2 \sum_{j=0}^{k} a_j^2 \lambda_j$$

เมื่อ $a = (a_0 \; a_1 \; \ldots \; a_k)$

ถ้าทำให้ฟังก์ชันมีค่าต่ำสุด สูตรค่าความผันแปรที่เกิดขึ้นจากความไม่แน่นอนของข้อมูลรอบเส้นถดถอยของตัวประมาณ Y จะทำให้ได้ตัวประมาณกำลังสองน้อยที่สุด

วิธีในการหาตัวประมาณนี้ ใช้วิธีการลากรานเจียน (Lagrangian Method)

$$f(a) = \eta^2 \sum_{j=0}^{k} a_j^2 \lambda_j - 2\mu_o \left(\sum_{j=0}^{k} a_j \gamma_{oj} - 1\right) \qquad (2.4.11)$$



เมื่อ $\mu_o$ คือ ตัวคูณลากรานเจียน (Lagrangian Multiplier)

$$\frac{\partial f(a_j)}{\partial a_j} = 0$$

จะได้ $$a_j = \eta^{-2} \gamma_{oj}^{-1} \mu_o \qquad ; \; j=0,1,2,\ldots,k \qquad (2.4.12)$$

เมื่อ $\sum_{j=0}^{k} a_j \sigma_{oj} = 1$

ทำการย้ายข้างของสมการ (2.4.12) จะได้

$$\mu_o = \eta^2 \left( \sum_{j=0}^{k} \lambda_j^{*-1} \right)^{-1} \quad ; \text{ เมื่อ } \lambda_j^* = \lambda_j / \sigma_{oj}^2$$

แทนค่า $\mu_o$ ลงใน (2.4.12)

$$a_j = r_j \lambda_j^{-1} \left( \sum_{l=0}^{k} \lambda_l^{*-1} \right)^{-1} \quad ; j = 0,1,2,\ldots,k \qquad (2.4.13)$$

จากสมการ (2.4.10) แทนค่า $a_j$ ในสมการ $b = -\eta \sum_{j=0}^{k} a_j \mathring{r}_j$

$$\text{เพราะฉะนั้น } b = -\eta \sum_{j=0}^{k} a_j \mathring{r}_j = -\eta \sum_{j=0}^{k} (\sigma_{oj} \lambda_j^{-1} \mathring{r}_j) \left( \sum_{l=0}^{k} \lambda_l^{*-1} \right)^{-1} \quad (2.4.14)$$

จาก (2.4.10) จะได้

$$SSE(x_1, x_2, \ldots, x_k) = \eta^2 \sum_{j=0}^{k} ((\sigma_{oj} \lambda_j^{-1})^2 \lambda_j) \left[ \left( \sum_{l=0}^{k} \lambda_l^{*-1} \right)^{-1} \right]^2$$

$$= \sum_{j=0}^{k} (\sigma_{oj}^2) \lambda_j^{-1} \left[ \left( \sum_{l=0}^{k} \lambda_l^{*-1} \right)^{-1} \right]^2$$

$$= \eta^2 \sum_{j=0}^{k} (\sigma_{oj}^2 / \lambda_j) \left( \sum_{l=0}^{k} \lambda_l^{*-1} )^{-1} \right)^2 \quad ; \lambda_j^* = \lambda_j / \sigma_{oj}^2$$

$$= \eta^2 \sum_{j=0}^{k} (\lambda^{*}_{j}{}^{-1})^{-1} (\sum_{l=0}^{k} \lambda^{*}_{l}{}^{-1})^{-1})^2$$

$$= \eta^2 \sum_{j=0}^{k} (\lambda^{*}_{j}{}^{-1})^{-1}) = \mu_o$$

<u>ตัวประมาณ b และค่าความผันแปรที่เกิดขึ้นจากความไม่แน่นอนของข้อมูลรอบเส้นถดถอยที่เกิดจากวิธีลาเท้นรูทธ์รีเกรสชั่น</u>

ให้ $\gamma_o, \gamma_1, \ldots, \gamma_{p-1}$ ขึ้นอยู่กับค่าดีเทอร์มิแนนท์ของ $A'A$ ที่ไม่สามารถนำมาพยากรณ์ได้ (non-predictive near singularities) จากวิธีการข้างต้น จะทำการปรับโดยกำหนดให้

$$a_o = a_1 = a_2 = \ldots = a_{p-1} = 0$$

จากสมการ (2.4.11) จะได้

$$a_j = \gamma_{oj} \lambda_j^{-1} (\sum_{l=p}^{k} \lambda_l^{*-1})^{-1} \quad ; j = p, p+1, \ldots, k \qquad (2.4.15)$$

และค่า b จะเป็นดังนี้

$$b^* = -\eta (\sum_{l=p}^{k} \lambda_l^{*-1})^{-1} (\sum_{j=p}^{k} \gamma_{oj} \lambda_j^{-1} \gamma^o_{j}) \qquad (2.4.16)$$

และค่าความผันแปรที่เกิดขึ้นจากความไม่แน่นอนของข้อมูลรอบเส้นถดถอย เป็นดังนี้

$$SSE^*(x_1, x_2, \ldots, x_k) = \eta^2 (\sum_{l=p}^{k} \lambda^{*}_{l}{}^{-1})^{-1} \qquad (2.4.17)$$

การที่ให้ $a_j$ เป็นศูนย์ เนื่องมาจากว่าค่าดีเทอร์มิแนนท์ของเมตริกซ์ $X'X$ ที่เข้าใกล้ศูนย์ มีบางค่าของ $\lambda_j$ และ $\gamma_{oj}$ ของ $A'A$ มีค่าเข้าใกล้ศูนย์ หรือน้อยมากและค่า $a_j$ จะขึ้นอยู่กับลาเท้นเวกเตอร์ ที่มีค่ามากเมื่อเปรียบเทียบกับ $a_j$ ที่เหลือ ดังนั้นเมื่อเกิดขึ้นอย่างนี้ จะทำให้เทอมของ $a_j\gamma^{o}_{j}$ เมื่อ$j=0,1,\ldots,p-1$ มีผลกระทบทำให้ค่า b มีค่ามากผิดปกติ แต่เมื่อทำการขจัดเทอมที่ผิดปกติเหล่านี้ออกไป จะมีผลทำให้ค่า b ที่ได้มีความแม่นยำใกล้เคียงกับค่า $\beta$ มากขึ้น ดังนั้นตัวประมาณลาเท้นรูทช์ จึงน่าจะเป็นตัวประมาณที่ให้ผลได้ดีกว่าวิธีกำลังสองน้อยที่สุด และวิธีริดจ์รีเกรสชั่น

## 2.5 การแปลงสมการความถดถอยพหุเป็นสมการมาตรฐาน

$$w_{ij} = \frac{x_{ij} - \bar{x}_j}{s_{jj}^{1/2}} \quad ; i = 1, 2, \ldots, n \quad j = 1, 2, \ldots, p$$

$$y_i = \frac{y_i - \bar{y}}{s_{yy}^{1/2}} \quad ; i = 1, 2, \ldots, n$$

โดยที่ $x_{ij}$ คือ ค่าของตัวแปรอิสระตัวที่ j จากตัวอย่างชุดที่ i

$y_i$ คือ ค่าของตัวแปรตามจากตัวอย่างชุดที่ i

$$s_{jj} = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2$$

ในการแปลงค่าตัวแปรแบบนี้จะได้ค่าตัวแปรอิสระใหม่ $w_j$ มีค่าเฉลี่ยเท่ากับ 0 และ $\sqrt{\Sigma (w_{ij} - w_i)^2} = 1$ เมื่อเขียนสมการความถดถอยพหุ ในเทอมของตัวแปรใหม่เหล่านี้ จะได้

$$y_i = b_1 w_{i1} + b_2 w_{i2} + \ldots + b_p w_{ip} + \underset{\sim}{\varepsilon}_i \quad ; \quad i = 1, 2, \ldots, n$$

หากประมาณค่าสัมประสิทธิ์ความถดถอยพหุโดยวิธีกำลังสองน้อยที่สุดจะได้ว่า $\hat{b} = (W'W)^{-1}W'y$ ในการใช้วิธีการแปลงข้อมูลนี้ จะได้เมตริกซ์ของความสัมพันธ์ $W'W$ ดังนี้

$$
W'W = \begin{bmatrix}
1 & r_{12} & r_{13} & r_{14} & \cdots\cdots & r_{1p} \\
r_{12} & 1 & r_{23} & r_{24} & \cdots\cdots & r_{2p} \\
r_{13} & r_{23} & 1 & r_{34} & \cdots\cdots & r_{3p} \\
. & & & & & . \\
. & & & & & . \\
r_{1p} & r_{2p} & & & & 1
\end{bmatrix}
$$

โดยที่ $r_{ij} = \sum_{k=1}^{n} \dfrac{(X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j)}{(S_{ii}\ S_{jj})^{1/2}}$

$$= \frac{S_{ij}}{(S_{ii}\ S_{jj})^{1/2}}$$

$r_{ij}$ เป็นสหสัมพันธ์ระหว่างตัวแปรอิสระ $X_i$ และ $X_j$ ในทำนองเดียวกัน

$$
W'y = \begin{bmatrix}
r_{1y} \\
r_{2y} \\
. \\
. \\
r_{ky}
\end{bmatrix}
$$

โดยที่ $r_{jy} = \sum_{u=1}^{n} \dfrac{(X_{uj} - \bar{X}_j)(y_u - \bar{y})}{(S_{jj}\ S_{yy})^{1/2}}$

$$= \frac{S_{jy}}{(S_{jj}\ S_{yy})^{1/2}}$$

$r_{jy}$ เป็นสหสัมพันธ์ระหว่างตัวแปรอิสระ $X_j$ และตัวแปรตาม $y$

ถ้าใช้วิธีการแปลงข้อมูลนี้ จะได้ว่า $z'z = (n-1)w'w$ ค่าสัมประสิทธิ์ความถดถอยพหุที่ได้จากตัวแปรอิสระนี้จะเรียกว่าสัมประสิทธิ์ความถดถอยมาตรฐาน