3

โมเดลเชิงเส้นและการประมาณค่า

การประมาณค่าอิทธิพลต่างๆจากโมเดลเชิงเส้น เป็นขั้นตอนหนึ่งที่มีความเกี่ยวข้องกับการประเมินพันธุกรรมสัตว์ ปัจจุบันมีวิธีการทางสถิติที่ใช้ในการประมาณอิทธิพลเหล่านี้อยู่หลายแบบ และแม้ว่าจะมีโปรแกรมคอมพิวเตอร์ที่ช่วยในการวิเคราะห์ค่าเหล่านี้อยู่บ้างแล้วก็ตาม แต่นักปรับปรุงพันธุ์ควรทำความเข้าใจถึงหลักการ คุณสมบัติ และข้อจำกัดของตัวประมาณที่ได้มาด้วยวิธีการต่างๆ ซึ่งจะช่วยให้เข้าใจและสามารถนำไปใช้ได้อย่างถูกต้อง เนื่องจากในทางปรับปรุงพันธุ์สัตว์นิยมใช้โมเดลผสม (mixed model) จึงมีการประมาณค่าอิทธิพลหลัก 2 แบบ ได้แก่การประมาณอิทธิพลคงที่ (fixed effects) และอิทธิพลสุ่ม (random effects) ซึ่งจะมีความเกี่ยวข้องกับการประมาณค่าการผสมพันธุ์ (breeding value) และในการประมาณชั้นสูงต่อๆไป

เนื้อหาสังเขป

โมเดลเชิงเส้น โมเดลแบบบวกสะสมและสมการถดถอย : : การสร้างโมเดลเชิงเส้นทั่วไป : : การหาคำตอบของระบบสมการ : : การประมาณค่าด้วยวิธี Ordinary Least Squares (OLS), Generalized Least Squares (GLS), Maximum Likelihood (ML) และ Best Linear Unbiased Estimation (BLUE)

I. โมเดลเชิงเส้น โมเดลบวกสะสม และสมการถดถอย

- *Linear model*
- *Additive model*
- *Regression*

ในทางปรับปรุงพันธุ์สัตว์นั้น เรานิยมแสดงความสัมพันธ์ระหว่างลักษณะการผลิต (production traits) หรือค่าสังเกต (observations) กับอิทธิพลที่มีต่อค่าสังเกตนั้นในรูปของผลบวกของอิทธิพลต่างๆ เช่น หากกำหนดให้การให้นมของโคนมขึ้นกับอิทธิพลเนื่องจากพ่อพันธุ์ (s) และฝูงการจัดการ (h) เราจะได้โมเดลเป็น $y_{ijk} = \mu + s_i + h_j + \varepsilon_{ijk}$ เมื่อ μ เป็น overall mean, ε_{ijk} เป็นความคลาดเคลื่อน, y_{ijk} เป็นปริมาณน้ำนมที่วัดได้จากสัตว์ตัวที่ k ซึ่งเกิดจากพ่อพันธุ์ i และอยู่ในฝูงที่ j ในทางสถิติสามารถเรียกรูปสมการที่มีค่าอิทธิพลต่างๆ อยู่ในรูปของผลบวกลักษณะนี้ว่าโมเดลแบบบวกสะสม (additive model) ซึ่งมีความหมายเช่นเดียวกับความสัมพันธ์ทางคณิตศาสตร์ในรูปของโมเดลเชิงเส้น (linear model) ซึ่งมีรูปทั่วไปเป็น $y = a + bx$ ดังนั้นจึงเห็นได้ว่าโมเดลแบบบวกสะสมและโมเดลเชิงเส้นมีรูปแบบเช่นเดียวกัน และเนื่องจากอิทธิพลที่แท้จริงของพ่อพันธุ์แต่ละตัวและอิทธิพลจากแต่ละฝูงเป็นพารามิเตอร์ที่ไม่ทราบค่า (unknown parameter) จึงต้องมีการประมาณขึ้นด้วยวิธีการทางเมทริกซ์และสถิติ ซึ่งจะได้กล่าวต่อไป

สมการถดถอย (regression) เป็นรูปหนึ่งของโมเดลเชิงเส้น โดยพบว่ารูปทั่วไปของสมการถดถอยจะมีลักษณะ เป็น $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_m X_m + \varepsilon$ เมื่อ y เป็นค่าสังเกตหรือตัวแปรตาม (dependent variable), $\beta_0, \beta_1, \beta_2, ..., \beta_m$ เป็นค่าสัมประสิทธิ์ความถดถอย (regression coefficients) ซึ่งเป็นพารามิเตอร์ที่ไม่ทราบค่า (unknown parameter), $X_1, X_2, ..., X_m$ เป็นตัวแปรอิสระ (independent variables) ที่มีลักษณะเป็นตัวแปรต่อเนื่อง (continuous variable) ซึ่งบางครั้งอาจเรียกว่า regressor และ ε เป็นความคลาดเคลื่อน

ในกรณีที่บางตัวแปรของอิทธิพลในโมเดลเชิงเส้นมีลักษณะเป็นตัวแปรไม่ต่อเนื่อง (discrete variable) หรือสามารถจำแนกเป็นกลุ่มได้ชัดเจน (classification variable) ซึ่งในการวิเคราะห์จะมีการแปลงตัวแปรนั้นเป็นตัวแปรรหัส (dummy variable) โมเดลแบบนี้จะถูกเรียกว่า general linear model ซึ่งจากโมเดลการให้น้ำนมของโคนมข้างต้น เราจะได้ว่า β_0 = overall mean, $\beta_1, \beta_2, ...$ เป็นอิทธิพลเนื่องจากพ่อพันธุ์ตัวที่ 1, 2, ... ตามลำดับ และหากพ่อพันธุ์มีทั้งหมด 3 ตัว ค่าสัมประสิทธิ์ต่อไปซึ่งได้แก่ $\beta_4, \beta_5, ... \beta_m$ จะเป็นอิทธิพลเนื่องจากฝูงที่ 1, 2, ... ตามลำดับ ในขณะที่ $X_1, X_2, ..., X_m$ ของโมเดลเชิงเส้นจะมีค่าเป็น 0 หรือ 1 ที่แสดงความสัมพันธ์ระหว่างค่าสังเกตกับการปรากฏของอิทธิเนื่องจากพ่อพันธุ์และฝูง ($\beta_0, \beta_1, \beta_2, ..., \beta_m$) นั้นๆ

โดยนิยาม โมเดลเชิงเส้น (linear model) หมายถึงฟังก์ชั่นเชิงเส้น (linear function) ของพารามิเตอร์ $\beta_0, \beta_1, \beta_2, ..., \beta_m$ แม้ว่าตัวแปร $X_1, X_2, ..., X_m$ จะอยู่ในรูปที่ไม่ใช่เชิงเส้นก็ตาม เช่น $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2^2 + \varepsilon$, $y = \beta_0 + \beta_1 e^X + \varepsilon$, และ

$y = \beta_0 + \beta_1 ln(X) + \varepsilon$ เป็นต้น หากพารามิเตอร์ไม่อยู่ในรูปฟังก์ชั่นเชิงเส้นจะไม่ถูกเรียกว่าโมเดลเชิงเส้น เช่น $y = \beta_0 e^{\beta_1 X} + \varepsilon,\ \ y = 1/(1+e^{\beta_0+\beta_1 X+\varepsilon})$ เป็นต้น

จะเห็นได้ว่าความสัมพันธ์เชิงเส้นข้างต้นสามารถเขียนให้อยู่ในรูปของเมทริกซ์ได้เป็น $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์สุ่ม (random vector) ของค่าสังเกต, $\boldsymbol{X}$ เป็น เมทริกซ์ของตัวแปรอิสระหรือ incidence matrix หรือ design matrix หรือ dummy-coding matrix, $\boldsymbol{\beta}$ เป็นเวคเตอร์ของพารามิเตอร์ และ $\boldsymbol{\varepsilon}$ เป็นเวคเตอร์สุ่มของความคลาดเคลื่อน

II. การสร้างโมเดลเชิงเส้น (Setting Up Linear Model)

Example 3.1

การสร้างความสัมพันธ์ระหว่างอิทธิพลต่างๆ ที่มีต่อค่าสังเกตในรูปของโมเดลเชิงเส้นเป็นขั้นตอนแรกที่นำไปสู่วิธีการประมาณทางสถิติอื่นๆต่อไป หากกำหนดให้โมเดลเชิงเส้นของการให้ผลผลิตนมในฟาร์มมีรูปสมการเป็น

$$y_{ij} = \mu + h_i + \varepsilon_{ij}$$

เมื่อ y_{ij} เป็นค่าปริมาณน้ำนม, μ เป็นค่า overall mean, h_i เป็นอิทธิพลเนื่องจากฝูงสัตว์ที่ i , และ ε_{ij} เป็นค่า error ของสัตว์จากฝูงที่ i ตัวที่ j โดยมีข้อมูลดังนี้

h1	h2	h3
10	10	10
11	11	15
12		20

จากโมเดลการให้นมของสัตว์ตามสมการข้างต้น เราสามารถดัดแปลงให้อยู่ในรูปของสมการถดถอยหรือโมเดลเชิงเส้นได้เป็น

$$y_{ij} = \mu + \underbrace{h_1 + h_2 + h_3}_{h_i} + \varepsilon_{ij}$$

ซึ่งจากข้อมูลในตารางข้างต้น สามารถแสดงความสัมพันธ์ของอิทธิพลต่างๆ ต่อการให้นมได้เป็น

$$\begin{aligned}
10 &= \mu + h_1 + 0 + 0 + \varepsilon_{11}\\
11 &= \mu + h_1 + 0 + 0 + \varepsilon_{12}\\
12 &= \mu + h_1 + 0 + 0 + \varepsilon_{13}\\
10 &= \mu + 0 + h_2 + 0 + \varepsilon_{21}\\
11 &= \mu + 0 + h_2 + 0 + \varepsilon_{22}\\
10 &= \mu + 0 + 0 + h_3 + \varepsilon_{31}\\
15 &= \mu + 0 + 0 + h_3 + \varepsilon_{32}\\
20 &= \mu + 0 + 0 + h_3 + \varepsilon_{33}
\end{aligned}$$

จากข้อมูลข้างต้น สังเกตว่าค่าสังเกตที่ 1, 2 และ 3 นั้น ถึงแม้จะได้รับอิทธิจากการจัดการหรือการเลี้ยงดูในฝูงเดียวกันแต่การตอบสนองการผลิตจะไม่เท่ากันทั้งนี้เนื่องจากความคลาดเคลื่อนของสัตว์แต่ละตัว (ε_{ij}) ที่แตกต่างกัน

- *Dummy coding*

ซึ่งสามารถสร้างรหัส (dummy coding) การปรากฏของอิทธิพลต่างๆ เมื่อกำหนดให้รหัส 1 = ปรากฏ และ 0 = ไม่ปรากฏ ได้เป็น

$$
\begin{array}{rcccccccccc}
y_{ij} & & \mu & & h_1 & & h_2 & & h_3 & & \varepsilon_{ij} \\
10 & = & 1 & + & 1 & + & 0 & + & 0 & + & \varepsilon_{11} \\
11 & = & 1 & + & 1 & + & 0 & + & 0 & + & \varepsilon_{12} \\
12 & = & 1 & + & 1 & + & 0 & + & 0 & + & \varepsilon_{13} \\
10 & = & 1 & + & 0 & + & 1 & + & 0 & + & \varepsilon_{21} \\
11 & = & 1 & + & 0 & + & 1 & + & 0 & + & \varepsilon_{22} \\
10 & = & 1 & + & 0 & + & 0 & + & 1 & + & \varepsilon_{31} \\
15 & = & 1 & + & 0 & + & 0 & + & 1 & + & \varepsilon_{32} \\
20 & = & 1 & + & 0 & + & 0 & + & 1 & + & \varepsilon_{33}
\end{array}
$$

- *Linear model in matrix form*

ระบบสมการเชิงเส้นข้างต้นสามารถเขียนในรูปของเมทริกซ์ได้เป็น

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

ซึ่งมีรายละเอียดดังนี้

$$
\underbrace{\begin{bmatrix} 10 \\ 11 \\ 12 \\ 10 \\ 11 \\ 10 \\ 15 \\ 20 \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{X}} \underbrace{\begin{bmatrix} \mu \\ h_1 \\ h_2 \\ h_3 \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon 3 \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{31} \\ \varepsilon_{32} \\ \varepsilon_{33} \end{bmatrix}}_{\boldsymbol{\varepsilon}}
$$

- *Design matrix*
- *Incidence matrix*

เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์ของค่าสังเกต (response variable), $\boldsymbol{\beta}$ เป็นเวคเตอร์ของอิทธิพลคงที่ (parameter), $\boldsymbol{X}$ ถูกเรียกว่า design matrix หรือ incidence matrix ที่สัมพันธ์กับการปรากฏของอิทธิพลคงที่ $\boldsymbol{\beta}$ และ $\boldsymbol{\varepsilon}$ เป็นเวคเตอร์ของอิทธิพลสุ่มของความคลาดเคลื่อน (error)

2.1 ข้อกำหนดของโมเดล

- *Model assumption*

โดยทั่วไปในการเขียนโมเดลเชิงเส้นใดๆ ควรระบุข้อกำหนดต่างๆ ของอิทธิพลต่างๆ ที่จำเพาะกับโมเดลนั้นไว้ด้วย เช่นการแจกแจง, ค่าคาดคะเน และความแปรปรวนของอิทธิพลเหล่านั้นเป็นต้น เช่น หากกำหนดให้เขียนโมเดลข้างต้นเป็น $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ เมื่อ $\boldsymbol{\varepsilon} \sim (\boldsymbol{0}, \boldsymbol{I}\sigma_e^2)$ ซึ่งแสดงถึง $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ และ $V(\boldsymbol{\varepsilon}) = \boldsymbol{I}\sigma_e^2$ และมีหมายความว่าค่าเฉลี่ยของความคลาดเคลื่อนของประชากรมีค่าเป็น $\boldsymbol{0}$ และความแปรปรวนของความคลาดเคลื่อนของประชากรมีค่าเป็น $\boldsymbol{I}\sigma_e^2$ โดยมีโครงสร้าง ดังนี้

$$\boldsymbol{I}\sigma_e^2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \sigma_e^2$$

$$= \begin{bmatrix} \sigma_e^2 & 0 & \cdots & 0 \\ 0 & \sigma_e^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_e^2 \end{bmatrix}$$

- *Independent error*
- *Uncorrelated error*

ซึ่งหมายความว่าความคลาดเคลื่อนจากสัตว์แต่ละตัวมีความแปรปรวนค่าเท่ากับ σ_e^2 โดยความคลาดเคลื่อนจากสัตว์แต่ละตัวจะเป็นอิสระต่อกัน เรียก independent error หรือ uncorrelated error (สังเกตว่าค่า covariance เท่ากับ 0)

2.2 การหาค่าคาดคะเนและความแปรปรวนของโมเดล

- *Model expectation and variance*

นอกจากการกำหนดค่าคาดคะเน (expectation) และความแปรปรวน (variance) ของอิทธิพลต่างๆ ในโมเดลแล้ว การระบุค่าคาดคะเนและความแปรปรวนของโมเดลเป็นสิ่งจำเป็นที่ต้องระบุ เช่นกัน ทั้งนี้เนื่องจากในแต่ละโมเดลประกอบด้วยอิทธิพลคงที่ (fixed effects) และอิทธิพลสุ่ม (random effects) ที่แตกต่างกัน ซึ่งในการหาคาดคะเนและความแปรปรวนของตัวแปร (variable) อื่นๆ ใช้หลักการการหาคาดคะเนของเวคเตอร์และเมทริกซ์ที่กล่าวไว้แล้วในบทที่ 2 ตามปกติ หากกำหนดให้

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad \text{โดย } E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \text{ และ } V(\boldsymbol{\varepsilon}) = \boldsymbol{I}\sigma_e^2$$

เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์ของค่าสังเกต, $\boldsymbol{\beta}$ เป็นเวคเตอร์ของอิทธิพลคงที่, $\boldsymbol{X}$ เป็น incidence matrix ที่แสดงการปรากฏของอิทธิพลคงที่ $\boldsymbol{\beta}$ ในค่าสังเกตและ $\boldsymbol{\varepsilon}$ เป็นเวคเตอร์ของอิทธิพลสุ่มของความคลาดเคลื่อน

$$
\begin{aligned}
V(\boldsymbol{y}) &= V(\boldsymbol{X\beta}+\boldsymbol{\varepsilon}) \\
&= V(\boldsymbol{X\beta})+V(\boldsymbol{\varepsilon})+2Cov(\boldsymbol{X\beta},\boldsymbol{\varepsilon}) \\
&= \boldsymbol{I}\sigma_e^2 \qquad ;V(\boldsymbol{X\beta})=0, Cov(\boldsymbol{X\beta},\boldsymbol{\varepsilon})=0 \\
Cov(\boldsymbol{y},\boldsymbol{\varepsilon}) &= Cov(\boldsymbol{X\beta}+\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}) \\
&= Cov(\boldsymbol{X\beta},\boldsymbol{\varepsilon})+Cov(\boldsymbol{\varepsilon},\boldsymbol{\varepsilon}) \\
&= \boldsymbol{I}\sigma_e^2 \qquad ;Cov(\boldsymbol{X\beta},\boldsymbol{\varepsilon})=0, Cov(\boldsymbol{\varepsilon},\boldsymbol{\varepsilon})=V(\boldsymbol{\varepsilon})=\boldsymbol{I}\sigma_e^2
\end{aligned}
$$

สังเกตว่า variance ของ fixed effects ($\boldsymbol{\beta}$) มีค่าเป็น 0 และกำหนดให้ covariance ของ fixed effects กับ random มีค่าเป็น 0 เช่นกัน

III. การหาคำตอบของระบบสมการ (Solving System of Linear Equation)

เนื่องจากค่าพารามิเตอร์ ($\boldsymbol{\beta}$) เป็นค่าของประชากรซึ่งเราไม่ทราบค่า ดังนั้นเราจึงต้องมีการประมาณค่าของอิทธิพลเหล่านั้นขึ้นมา ตัวประมาณของอิทธิพลเนื่องจากปัจจัยคงที่ (fixed effect) เรียกว่า estimator ส่วนตัวประมาณของอิทธิพลเนื่องจากปัจจัยสุ่มเรียกว่า predictor ส่วนค่าอิทธิพลที่ประมาณได้อาจเรียกรวมๆว่าค่าประมาณ (estimates) เราสามารถสร้างตัวประมาณจากระบบสมการเชิงเส้นข้างต้นได้หลายวิธี ซึ่งในบทนี้จะได้กล่าวถึงวิธีการประมาณค่าของอิทธิพลเนื่องจากตัวแปรคงที่ที่นิยม 4 วิธี ได้แก่i) Ordinary Least Squares (OLS), ii) Generalized Least Squares (GLS), iii) Maximum Likelihood (ML), และ iv) Best linear Unbiased Estimator (BLUE)

3.1 การสร้างตัวประมาณด้วยวิธี Ordinary Least Squares (OLS)

กำหนดให้โมเดลเชิงเส้นมีลักษณะดังนี้

$$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}, \qquad \text{โดย } E(\boldsymbol{\varepsilon})=\boldsymbol{0} \text{ และ } V(\boldsymbol{\varepsilon})=\boldsymbol{I}\sigma_e^2$$

เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์ขนาด $n\times 1$ ของค่าสังเกต, $\boldsymbol{\beta}$ เป็นเวคเตอร์ขนาด $m\times 1$ ของอิทธิพลคงที่, $\boldsymbol{X}$ เป็น incidence matrix ขนาด $n\times m$ ที่แสดงการปรากฏของอิทธิพลคงที่ $\boldsymbol{\beta}$ ในแต่ละค่าสังเกต, $\boldsymbol{\varepsilon}$ และเวคเตอร์ขนาด $n\times 1$ ของความคลาดเคลื่อน โดยมีค่าเฉลี่ยเป็นศูนย์และมีความแปรปรวนเป็น $\boldsymbol{I}\sigma_e^2$ (uncorrelated error variance)

- *Uncorrelated error variance*

ตัวประมาณแบบ ordinary least squares (OLS) หมายถึงตัวประมาณค่าของ $\boldsymbol{\beta}$ ที่ให้ค่ากำลังสองของความคลาดเคลื่อน ($\boldsymbol{\varepsilon}$) ของโมเดลต่ำสุด เมื่อ $\boldsymbol{\varepsilon}=\boldsymbol{y}-\boldsymbol{X\beta}$ ดังนั้นตัวประมาณจึงได้จากการ minimize ค่า sum of squares of error หรือ $Min\{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\}$ ซึ่งพบว่าค่าประมาณแบบ OLS จะประเมินได้จาก

- *OLS estimator*

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$$

▪ *Example 3.2*

▪ จากตัวอย่างที่ 3.1 ค่าประมาณของ $\boldsymbol{\beta}$ ด้วยวิธี OLS สามารถคำนวณได้ดังนี้

$$\boldsymbol{X'X} = \begin{bmatrix} 8 & 3 & 2 & 3 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 3 & 0 & 0 & 3 \end{bmatrix}, \quad \boldsymbol{X'y} = \begin{bmatrix} 99 \\ 33 \\ 21 \\ 45 \end{bmatrix}$$

เนื่องจาก $\boldsymbol{X'X}$ เป็น not of full rank ดังนั้น

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (\boldsymbol{X'X})^{-}\boldsymbol{X'y} \\ &= \begin{bmatrix} 8 & 3 & 2 & 3 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 3 & 0 & 0 & 3 \end{bmatrix}^{-} \begin{bmatrix} 99 \\ 33 \\ 21 \\ 45 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/3 \end{bmatrix}^{-} \begin{bmatrix} 99 \\ 33 \\ 21 \\ 45 \end{bmatrix} \\ &= \begin{bmatrix} 0.0 \\ 11.0 \\ 10.5 \\ 14.0 \end{bmatrix} = \begin{bmatrix} \hat{\mu} \\ \hat{h}_1 \\ \hat{h}_2 \\ \hat{h}_3 \end{bmatrix} \end{aligned}$$

ข้อสังเกต

▪ *OLS estimator*
▪ *Estimated effects*

$\hat{\boldsymbol{\beta}}$ เรียกว่าเป็นตัวประมาณแบบลีสท์สแควร์ (OLS estimator) ซึ่งในที่นี้เป็นเวคเตอร์ของค่าประมาณของอิทธิพล (estimated effects) เนื่องจากฝูง (herd) ซึ่งได้แก่ $\hat{h}_1, \hat{h}_2, \hat{h}_3$ นอกจากนี้ค่า $\hat{\mu}+\hat{h}_1, \hat{\mu}+\hat{h}_2, \hat{\mu}+\hat{h}_3$ ถูกเรียกว่าค่าเฉลี่ยลีสท์สแควร์ (least squares means, LSM) เนื่องจากเป็นค่าเฉลี่ยที่คำนวณจากค่าของอิทธิพลต่างๆที่ประมาณขึ้นด้วยวิธี least squares ดังนั้น

$$\begin{aligned} LSM(herd\,1) &= \hat{\mu}+\hat{h}_1 = 0+11.0 = 11.0 \\ LSM(herd\,2) &= \hat{\mu}+\hat{h}_2 = 0+10.5 = 10.5 \\ LSM(herd\,3) &= \hat{\mu}+\hat{h}_3 = 0+14.0 = 14.0 \end{aligned}$$

เนื่องจาก $\boldsymbol{X'X}$ เป็น not of full rank matrix จึงทำให้ $\hat{\boldsymbol{\beta}}$ มีคำตอบได้หลายรูปแบบขึ้นอยู่กับ generalized inverse ที่ใช้

Solution 1 (set $\hat{\mu}=0$)	Solution 2 (set $\hat{h}_1=0$)	Solution 3 (set $\hat{h}_2=0$)	Solution 4 (set $\hat{h}_2=0$)
$\hat{\mu}=0.0$	$\hat{\mu}=11.0$	$\hat{\mu}=10.5$	$\hat{\mu}=15.0$
$\hat{h}_1=11.0$	$\hat{h}_1=0.0$	$\hat{h}_1=0.5$	$\hat{h}_1=-4.0$
$\hat{h}_2=10.5$	$\hat{h}_2=-0.5$	$\hat{h}_2=0.0$	$\hat{h}_2=-4.5$
$\hat{h}_3=15.0$	$\hat{h}_3=4.0$	$\hat{h}_3=4.5$	$\hat{h}_3=0.0$

อย่างไรก็ตาม แม้ว่าค่าประมาณของอิทธิพล $\hat{\mu}, \hat{h}_1, \hat{h}_2, \hat{h}_3$ จะแตกต่างกันไป แต่ linear combination ($\boldsymbol{k}'\hat{\boldsymbol{\beta}}$) ของค่าประมาณของอิทธิพลเหล่านี้ เช่น ค่าเฉลี่ยลีสท์สแควร์ ($\hat{\mu}+\hat{h}_1, \hat{\mu}+\hat{h}_2, \hat{\mu}+\hat{h}_3$) จะไม่เปลี่ยนแปลง

3.2 การสร้างตัวประมาณด้วยวิธี Generalized Least Squares (GLS)

กำหนดให้โมเดลเชิงเส้นมีลักษณะดังนี้

$$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}, \text{ โดย } E(\boldsymbol{\varepsilon})=\boldsymbol{0} \text{ และ } V(\boldsymbol{\varepsilon})=\boldsymbol{V}$$

เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์ขนาด $n\times 1$ ของค่าสังเกต, $\boldsymbol{\beta}$ เป็นเวคเตอร์ขนาด $m\times 1$ ของอิทธิพลคงที่, $\boldsymbol{X}$ เป็น incidence matrix ขนาด $n\times m$ ที่แสดงการปรากฏของอิทธิพลคงที่ $\boldsymbol{\beta}$ ในแต่ละค่าสังเกต, $\boldsymbol{\varepsilon}$ และเวคเตอร์ขนาด $n\times 1$ ของความคลาดเคลื่อน โดยมีค่าเฉลี่ยเป็นศูนย์และมีโครงสร้างของความแปรปรวนและความแปรปรวนร่วมระหว่างความคลาดเคลื่อนเป็น $\boldsymbol{V}$ (correlated error variance)

- *Correlated error variance*
- *Weighted least square*

ตัวประมาณแบบ generalized least squares (GLS) หมายถึงตัวประมาณค่าของ $\boldsymbol{\beta}$ ที่ให้ค่ากำลังสองของความคลาดเคลื่อนของโมเดลต่ำสุดเช่นกัน แต่สามารถใช้กับโมเดลที่มีความคลาดเคลื่อนที่มีความสัมพันธ์กัน โดยวิธีการ weighted least squares จะถูกนำมาใช้ในการจัดการ correlated error ดังนั้นตัวประมาณจึงได้จากการ minimize ค่า standardized residual sum of squares หรือ $Min\{\boldsymbol{\varepsilon}'\boldsymbol{V}^{-1}\boldsymbol{\varepsilon}\}$ ซึ่งพบว่าค่าประมาณแบบ GLS จะประเมินได้จาก

- *GLS estimator*

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

ข้อสังเกต

- *Comparisons of OLS and GLS estimator*

ข้อเปรียบเทียบบางประการจากตัวประมาณแบบ Ordinary Least Squares (OLS) และ Generalized Least Squares (GLS) สามารถสรุปได้ดังนี้

	OLS	GLS
ข้อกำหนดของ error term	$\boldsymbol{\varepsilon} \sim (\boldsymbol{0}, \boldsymbol{I}\sigma_e^2)$	$\boldsymbol{\varepsilon} \sim (\boldsymbol{0}, \boldsymbol{V})$
สูตรตัวประมาณของ $\boldsymbol{\beta}$	$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$	$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$
Var($\hat{\boldsymbol{\beta}}$)	$(\boldsymbol{X}'\boldsymbol{X})^{-}\sigma_e^2$	$(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}$
Var($\boldsymbol{y}$)	$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\sigma_e^2$	$\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'$

3.3 การสร้างตัวประมาณด้วยวิธี Maximum Likelihood (ML)

- *Maximum likelihood estimator*

กำหนดให้โมเดลเชิงเส้นมีลักษณะเช่นเดียวกับโมเดลเชิงเส้นในการสร้างตัวประมาณแบบ GLS ในหัวข้อ 3.2 ตัวประมาณแบบ Maximum Likelihood (ML) หมายถึงตัวประมาณค่าของ $\boldsymbol{\beta}$ ที่ให้ภาวะน่าจะเป็น (likelihood) ที่ค่าประมาณที่ได้จะเป็นค่าของพารามิเตอร์สูงสุด ดังนั้นตัวประมาณจึงได้จากการ maximize likelihood function ($\mathscr{L}$) หรือ $Max\{\mathscr{L}\}$ เมื่อ $\mathscr{L}$ เป็น joint distribution function ของ $\boldsymbol{y}$ ซึ่งพบว่าภายใต้การแจกแจงแบบปกติ (normality) ค่าประมาณแบบ ML จะประเมินได้จาก

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

3.4 การสร้างตัวประมาณด้วยวิธี Best Linear Unbiased Estimator (BLUE)

- *Best linear unbiased estimator*

กำหนดให้โมเดลเชิงเส้นมีลักษณะเช่นเดียวกับโมเดลเชิงเส้นในการสร้างตัวประมาณแบบ GLS ในหัวข้อ 3.2 ตัวประมาณแบบ BLUE หมายถึงตัวประมาณค่าที่อยู่ในรูป linear combination ของค่าสังเกต ($\hat{\boldsymbol{\beta}} = \boldsymbol{L}'\boldsymbol{y}$) โดยมีคุณสมบัติไม่เอนเอียง (unbiasedness) กล่าวคือค่าคาดคะเนของของค่าประมาณมีค่าเท่ากับค่าพารามิเตอร์หรือ $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$ และเป็นตัวประมาณที่ดีที่สุด (best estimator) ซึ่งหมายถึงเป็นตัวประมาณที่มีค่าความแปรปรวนของความคลาดเคลื่อนของตัวประมาณต่ำสุด (estimation error variance ต่ำสุด) หรือมีค่า $Min\{E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^2\}$ ซึ่งพบว่าค่าประมาณแบบ BLUE จะประเมินได้จาก

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

ข้อสังเกต

ทั้งตัวประมาณแบบ GLS, ML, และ BLUE มีรูปเดียวกัน ดังนั้นอาจกล่าวได้ว่าการประมาณค่าอิทธิพลต่างๆ ด้วยสมการ $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$ จะทำให้ค่าประมาณที่ได้มีคุณสมบัติเป็นตัวประมาณที่ 1) มีค่ากำลังสองของ sum of squares of error (SSE) ของโมเดลต่ำที่สุด 2) ให้ภาวะน่าจะเป็นที่ค่าประมาณที่ได้จะเป็นค่าพารามิเตอร์ของประชากรสูงสุดหากข้อมูลมีการแจกแจงปกติ และ 3) มีคุณสมบัติไม่เอนเอียง (unbiasedness) และมีค่า estimation error variance ของตัวประมาณต่ำสุด

▪ *Example 3.3*

▪ จากตัวอย่างที่ 3.1 หากกำหนดให้ $\boldsymbol{V}$ มีค่าดังนี้

$$\boldsymbol{V} = \begin{bmatrix} 20 & 5 & 5 & 0 & 0 & 0 & 0 & 10 \\ 5 & 20 & 5 & 0 & 0 & 0 & 0 & 0 \\ 5 & 5 & 20 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 30 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 20 & 1 \\ 10 & 0 & 0 & 0 & 0 & 1 & 1 & 20 \end{bmatrix}$$

ค่าประมาณค่าของ $\boldsymbol{\beta}$ ด้วยวิธี GLS, ML และ BLUE สามารถคำนวณได้ดังนี้

1) คำนวณ $\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}$ และ $(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}$

$$\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 20 & 5 & 5 & 0 & 0 & 0 & 0 & 10 \\ 5 & 20 & 5 & 0 & 0 & 0 & 0 & 0 \\ 5 & 5 & 20 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 30 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 20 & 1 \\ 10 & 0 & 0 & 0 & 0 & 1 & 1 & 20 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.28 & 0.09 & 0.06 & 0.13 \\ 0.09 & 0.11 & 0 & -0.02 \\ 0.06 & 0 & 0.06 & 0 \\ 0.13 & -0.02 & 0 & 0.15 \end{bmatrix}$$

$$(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 9.54 & 0 & 1.32 \\ 0 & 0 & 17.50 & 0 \\ 0 & 1.32 & 0 & 6.75 \end{bmatrix}$$

2) คำนวณ $\boldsymbol{X'V^{-1}y}$

$$\boldsymbol{X'V^{-1}y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 20 & 5 & 5 & 0 & 0 & 0 & 0 & 10 \\ 5 & 20 & 5 & 0 & 0 & 0 & 0 & 0 \\ 5 & 5 & 20 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 30 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 30 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 20 & 1 \\ 10 & 0 & 0 & 0 & 0 & 1 & 1 & 20 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3.52 \\ 0.73 \\ 0.60 \\ 2.19 \end{bmatrix}$$

3) เนื่องจาก $\boldsymbol{X'V^{-1}X}$ เป็นเมทริกซ์ not of full rank ดังนั้น

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X'V^{-1}X})^{-}\boldsymbol{X'V^{-1}y}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 9.54 & 0 & 1.32 \\ 0 & 0 & 17.50 & 0 \\ 0 & 1.32 & 0 & 6.75 \end{bmatrix} \begin{bmatrix} 3.75 \\ 1.10 \\ 0.60 \\ 2.05 \end{bmatrix}$$

$$= \begin{bmatrix} 0.00 \\ 9.87 \\ 10.50 \\ 15.73 \end{bmatrix} = \begin{bmatrix} \hat{\mu} \\ \hat{h}_1 \\ \hat{h}_2 \\ \hat{h}_3 \end{bmatrix}$$

- *Some properties of estimators*

3.5 สรุปคุณสมบัติบางประการและที่มาของตัวประมาณแบบต่างๆ

ข้อเปรียบเทียบบางประการจากตัวประมาณแบบ OLS, GLS, ML และ BLUE สามารถสรุปได้ดังตาราง

	OLS	GLS	ML*	BLUE
Derive by minimizing sum squares of residual	x	x		
Derive by minimizing mean squares of estimation error				x
Derive by maximizing likelihood function			x	
Unbiased estimator	x	x	x	x
Best estimator (minimize error variance)		x	x	x

Note: * ML ภายใต้ normality assumption

IV. ที่มาทางทฤษฎีของตัวประมาณ (Theoretical Background of Estimator)*

ในส่วนนี้จะเน้นการพิสูจน์ที่มาของตัวประมาณแบบต่างๆที่กล่าวไว้ในส่วนที่ III โดยตัวประมาณแบบลีสท์สแควร์ชนิด OLS ถูกสร้างจากพื้นฐานของโมเดล $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ เมื่อ $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ และ $V(\boldsymbol{\varepsilon}) = \boldsymbol{I}\sigma_e^2$ ในขณะที่ตัวประมาณแบบลีสท์สแควร์ชนิด GLS, Maximum likelihood (ML) และ Best Linear Unbiased estimator (BLUE) ถูกสร้างจากพื้นฐานของโมเดล $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ เมื่อ $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ และ $V(\boldsymbol{\varepsilon}) = \boldsymbol{V}$

4.1 ที่มาของตัวประมาณแบบ Ordinary Least Squares (OLS)

กำหนดให้โมเดลเชิงเส้นมีลักษณะดังนี้

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad \text{โดย } E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \text{ และ } V(\boldsymbol{\varepsilon}) = \boldsymbol{I}\sigma_e^2$$

ตัวประมาณแบบ OLS ถูกสร้างจากการ minimize ค่า sum of squares of error ของโมเดลหรือ $Min\{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}\}$ ซึ่งมีขั้นตอนดังนี้

1) หาค่าความคลาดเคลื่อน ($\boldsymbol{\varepsilon}$) จากโมเดลพบว่า

$$\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{X\beta}$$

▪ *Construct residual SS*

2) สร้าง residual sum of squares หรือ sum of squares of error (SSE) ซึ่งจากสมการใน 1) ทำให้สามารถสร้าง SSE ในรูปของ quadratic form ของพารามิเตอร์ ($\boldsymbol{\beta}$) ที่ต้องการประมาณค่าได้ ดังนี้

$$\begin{aligned} Q(\boldsymbol{\beta}) &= \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} \\ &= (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) \\ &= (\boldsymbol{y}' - \boldsymbol{\beta}'\boldsymbol{X}')(\boldsymbol{y} - \boldsymbol{X\beta}) \\ &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta} \end{aligned}$$

▪ *Minimize residual SS*

3) จาก quadratic form หรือ SSE ที่ได้ ทำให้มีค่าต่ำสุดได้โดยหาอนุพันธ์ตาม $\boldsymbol{\beta}$ แล้วเทียบให้เท่ากับศูนย์ซึ่งจะได้ว่า

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\beta}} Q(\boldsymbol{\beta}) &= \frac{\partial}{\partial \boldsymbol{\beta}} \{\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{X\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta}\} \\ &= 0 - \boldsymbol{X}'\boldsymbol{y} - \boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X\beta} \\ &= -2\boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X\beta} \end{aligned}$$

เมื่อเทียบอนุพันธ์ให้เท่ากับ 0 จะได้ว่า

$$\boldsymbol{X}'\boldsymbol{X\beta} = \boldsymbol{X}'\boldsymbol{y}$$

---

* เนื้อหาขั้นสูง สำหรับผู้ที่สนใจในเชิงทฤษฎี

▪ *Normal equation*

4) สมการที่เกิดจากการเทียบอนุพันธ์อันดับแรกให้เท่ากับศูนย์ข้างต้นเรียกว่า normal equation ซึ่งจากสมการนี้ทำให้ประมาณ $\boldsymbol{\beta}$ ได้จาก

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{y}$$

4.2 ที่มาของตัวประมาณแบบ Generalized Least Squares (GLS)

กำหนดให้โมเดลเชิงเส้นมีลักษณะดังนี้

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \text{ โดย } E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \text{ และ } V(\boldsymbol{\varepsilon}) = \boldsymbol{V}$$

ตัวประมาณแบบ GLS ถูกสร้างจากการ minimize ค่า standardized residual sum of squares ของโมเดลหรือ $Min\{\boldsymbol{\varepsilon}'\boldsymbol{V}^{-1}\boldsymbol{\varepsilon}\}$ ซึ่งมีขั้นตอนดังนี้

1) Standardize โมเดลด้วย error term โดยใช้ $\boldsymbol{V}^{-1/2}$ คูณเข้าทั้งสองข้างของสมการ ดังนั้น

$$\begin{aligned}\boldsymbol{V}^{-1/2}\boldsymbol{y} &= \boldsymbol{V}^{-1/2}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= \boldsymbol{V}^{-1/2}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{V}^{-1/2}\boldsymbol{\varepsilon}\end{aligned}$$

2) หาค่าความคลาดเคลื่อน (ในรูป standardized) จากโมเดล

$$\boldsymbol{V}^{-1/2}\boldsymbol{\varepsilon} = \boldsymbol{V}^{-1/2}\boldsymbol{y} - \boldsymbol{V}^{-1/2}\boldsymbol{X}\boldsymbol{\beta}$$

▪ *Construct standardized residual SS*

3) สร้าง sum of squares of error (SSE) ซึ่งจากสมการใน 1) ทำให้สามารถสร้าง Standardized residual SS ในรูปของ quadratic form ของพารามิเตอร์ ($\boldsymbol{\beta}$) ที่ต้องการประมาณค่าได้ ดังนี้

$$\begin{aligned}Q(\boldsymbol{\beta}) &= (\boldsymbol{V}^{-1/2}\boldsymbol{\varepsilon})'(\boldsymbol{V}^{-1/2}\boldsymbol{\varepsilon}) \\ &= (\boldsymbol{V}^{-1/2}\boldsymbol{y} - \boldsymbol{V}^{-1/2}\boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{V}^{-1/2}\boldsymbol{y} - \boldsymbol{V}^{-1/2}\boldsymbol{X}\boldsymbol{\beta}) \\ &= (\boldsymbol{y}'\boldsymbol{V}^{-1/2} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1/2})(\boldsymbol{V}^{-1/2}\boldsymbol{y} - \boldsymbol{V}^{-1/2}\boldsymbol{X}\boldsymbol{\beta}) \\ &= \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta}\end{aligned}$$

▪ *Minimize standardized residual SS*

4) จาก quadratic form หรือ SSE ที่ได้ ทำให้มีค่าต่ำสุดได้โดยหาอนุพันธ์ตาม $\boldsymbol{\beta}$ แล้วเทียบให้เท่ากับศูนย์ซึ่งจะได้ว่า

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\beta}}Q(\boldsymbol{\beta}) &= \frac{\partial}{\partial\boldsymbol{\beta}}\{\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta}\} \\ &= \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} - \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} \\ &= -2\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta}\end{aligned}$$

เมื่อเทียบอนุพันธ์ให้เท่ากับ 0 จะได้ว่า

$$\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

5) สมการที่เกิดจากการเทียบอนุพันธ์อันดับแรกให้เท่ากับศูนย์ข้างต้นเรียกว่า normal equation ซึ่งจากสมการนี้ทำให้ประมาณ $\boldsymbol{\beta}$ ได้จาก

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

4.3 ที่มาของตัวประมาณแบบ Maximum Likelihood (ML)

กำหนดให้โมเดลเชิงเส้นมีลักษณะดังนี้

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \text{ โดย } E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \text{ และ } V(\boldsymbol{\varepsilon}) = \boldsymbol{V}$$

ตัวประมาณแบบ Maximum Likelihood (ML) หมายถึงตัวประมาณค่าของ $\boldsymbol{\beta}$ ที่ให้ภาวะน่าจะเป็นที่ค่าประมาณจะเป็นค่าของพารามิเตอร์สูงสุด โดยการ maximize likelihood function ($\mathscr{L}$) หรือ $Max\{\mathscr{L}\}$ ซึ่งพบว่าค่าประมาณแบบ ML จะประเมินได้มีขั้นตอนดังนี้

1) สร้างฟังก์ชันภาวะน่าจะเป็นของพารามิเตอร์

▪ *Likelihood function*

จากทฤษฎีทางสถิติ ฟังก์ชันภาวะน่าจะเป็น (likelihood function) ของพารามิเตอร์จากค่าสังเกต ($\boldsymbol{y}$) สามารถอธิบายได้จากฟังก์ชันของการแจกแจงร่วม (joint density หรือ products of density) ของค่าสังเกตเมื่อกำหนดพารามิเตอร์ (likelihood function of parameters given data is a product of density function of data given parameters)

เนื่องจากในการสร้าง likelihood function ต้องทราบการแจกแจงร่วม ดังนั้นในการสร้างตัวประมาณด้วยวิธีนี้จึงต้องทราบชนิดของการแจกแจง ซึ่งโดยทั่วไปนิยมสร้างข้อกำหนดให้มีการแจกแจงปกติ (normality assumption) ดังนั้นโมเดลเชิงเส้นจึงมีรูปแบบเป็น

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \text{ โดย } \boldsymbol{\varepsilon} \sim MVN(\boldsymbol{0}, \boldsymbol{V})$$

▪ *Construct likelihood function*

เมื่อ MVN หมายถึง multivariate normal distribution ดังนั้น likelihood function ($\mathscr{L}$) จะมีรูปดังนี้

$$\begin{aligned}\mathscr{L}\ (\boldsymbol{\beta}, \boldsymbol{V} \mid \boldsymbol{y}) &= \prod f(\boldsymbol{y} \mid \boldsymbol{\beta}, \boldsymbol{V}) \\ &= (2\pi)^{-n/2} \mid \boldsymbol{V} \mid^{-1/2} exp\{-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta})\}\end{aligned}$$

▪ *Construct log likelihood*

2) ทำให้ค่า likelihood function สูงสุด ซึ่งในการหาค่าสูงสุดของฟังก์ชันที่อยู่ในรูปผลคูณหรือ exponential family นิยมแปลงฟังก์ชันนั้นให้อยู่ในรูปลอการิทึมก่อนแล้วจึงหาอนุพันธ์ ดังนั้น log likelihood function (L) จะมีค่าดังนี้

$$
\begin{aligned}
\mathscr{L} &= (2\pi)^{-n/2}\,|\boldsymbol{V}|^{-1/2}\,exp\{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{X\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X\beta})\} \\
\ln \mathscr{L} &= -\frac{n}{2}\ln(2\pi)-\frac{1}{2}\ln|\boldsymbol{V}|-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{X\beta})'\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X\beta}) \\
L &= -\frac{n}{2}\ln(2\pi)-\frac{1}{2}\ln|\boldsymbol{V}|-\frac{1}{2}(\boldsymbol{y}'-\boldsymbol{\beta}'\boldsymbol{X}')\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X\beta}) \\
&= -\frac{n}{2}\ln(2\pi)-\frac{1}{2}\ln|\boldsymbol{V}|-\frac{1}{2}(\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y}-\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}-\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{X\beta} \\
&\quad +\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta})
\end{aligned}
$$

- *Maximize log likelihood*

3) ทำให้มีค่าสูงสุดได้โดยหาอนุพันธ์ตาม $\boldsymbol{\beta}$ แล้วเทียบให้เท่ากับศูนย์ซึ่งจะได้ว่า

$$
\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\beta}}L &= \frac{\partial}{\partial\boldsymbol{\beta}}\{-\frac{n}{2}\ln(2\pi)-\frac{1}{2}\ln|\boldsymbol{V}|^{-1/2}-\frac{1}{2}(\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{y}-\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y} \\
&\quad -\boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}+\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta})\} \\
&= -\frac{1}{2}(-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}+2\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}) \\
&= -\frac{1}{2}(-2\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}+2\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}) \\
&= \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}-\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}
\end{aligned}
$$

เมื่อเทียบอนุพันธ์ให้เท่ากับศูนย์จะได้ว่า

$$\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}=\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

6) จาก normal equation ที่ได้ ซึ่งจากสมการนี้ทำให้ประมาณ $\boldsymbol{\beta}$ ได้จาก

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

4.4 ที่มาของ Best Linear Unbiased Estimator (BLUE)

กำหนดให้โมเดลเชิงเส้นมีลักษณะดังนี้

$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}$, โดย $E(\boldsymbol{\varepsilon})=\boldsymbol{0}$ และ $V(\boldsymbol{\varepsilon})=\boldsymbol{V}$

ตัวประมาณแบบ BLUE หมายถึงตัวประมาณที่อยู่ในรูป linear combination ของค่าสังเกต ($\hat{\boldsymbol{\beta}}=\boldsymbol{L}'\boldsymbol{y}$) โดยมีคุณสมบัติไม่เอนเอียง (unbiasedness) และมีเป็นตัวประมาณที่ดีที่สุด (best estimator) ซึ่งหมายถึงเป็นตัวประมาณที่มีค่า estimation error variance ต่ำสุด

ดังนั้น การประมาณค่าด้วยวิธีนี้ใช้หลักการสร้างตัวประมาณจากฟังก์ชันเชิงเส้นของค่าสังเกต $\boldsymbol{y}$ กล่าวคือสร้างตัวประมาณให้อยู่ในรูป $\boldsymbol{L'y}$ เมื่อ $\boldsymbol{L}$ เป็น linear combination matrix ($\hat{\boldsymbol{\beta}} = \boldsymbol{L'y}$) จากนั้นต้องสร้างตัวประมาณให้มีคุณสมบัติไม่เอนเอียง (unbiasedness) และเป็น best estimator หรือสร้างตัวประมาณให้มีความแปรปรวนของความคลาดเคลื่อนของการประมาณต่ำสุด (estimation error variance ต่ำสุด) ซึ่งค่า estimation error variance จะคำนวณจาก mean squares of estimation error ดังนั้นหากกำหนดให้ $\boldsymbol{\delta}$ เป็น estimation error แล้ว วิธีการนี้จึงต้องการตัวประมาณที่สร้างจาก $Min\{E(\boldsymbol{\delta\delta'})\}$ ซึ่งในการหาตัวประมาณจากสมการข้างต้นต้องใช้หลายขั้นตอน ดังนี้

- *Estimation error*

1) หาค่าความคลาดเคลื่อนของตัวประมาณ (estimation error, $\boldsymbol{\delta}$) หากกำหนดให้ $\boldsymbol{\beta}$ เป็นพารามิเตอร์ที่ไม่ทราบค่า และ $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณที่สร้างขึ้น ค่า estimation error หาได้จาก

$$\boldsymbol{\delta} = \hat{\boldsymbol{\beta}} - \boldsymbol{\beta}$$

- *Construct mean squares of estimation error*

2) สร้าง mean squares of estimation error ในรูปของ matrix form ของพารามิเตอร์ ($\boldsymbol{\beta}$) ที่ต้องการประมาณค่าได้ ดังนี้

$$\begin{aligned} M(\boldsymbol{\beta}) &= E(\boldsymbol{\delta\delta'}) \\ &= E\{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\} \\ &= E\{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}' - \boldsymbol{\beta}')\} \\ &= E\{(\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}' - \boldsymbol{\beta}\hat{\boldsymbol{\beta}}' - \hat{\boldsymbol{\beta}}\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}')\} \end{aligned}$$

กำหนดให้ $\hat{\boldsymbol{\beta}} = \boldsymbol{L'y}$ ดังนั้น

$$\begin{aligned} M(\boldsymbol{\beta}) &= E\{(\boldsymbol{L'y})(\boldsymbol{L'y})' - \boldsymbol{\beta}(\boldsymbol{L'y})' - (\boldsymbol{L'y})\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}')\} \\ &= E\{(\boldsymbol{L'y})(\boldsymbol{y'L}) - \boldsymbol{\beta}(\boldsymbol{y'L}) - (\boldsymbol{L'y})\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}')\} \\ &= E\{\boldsymbol{L'yy'L} - \boldsymbol{\beta y'L} - \boldsymbol{L'y\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}'\} \\ &= \boldsymbol{L'}E(\boldsymbol{yy'})\boldsymbol{L} - \boldsymbol{\beta}E(\boldsymbol{y'})\boldsymbol{L} - \boldsymbol{L'}E(\boldsymbol{y})\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}' \end{aligned}$$

คำนวณค่า Expectation เพื่อแทนค่าในสมการ $M(\boldsymbol{\beta})$

จาก $E(\boldsymbol{y}) = E(\boldsymbol{X\beta} + \boldsymbol{\varepsilon}) = E(\boldsymbol{X\beta}) + E(\boldsymbol{\varepsilon})$, เมื่อ $E(\boldsymbol{\varepsilon}) = 0$ ดังนั้น $E(\boldsymbol{y}) = \boldsymbol{X\beta}$

$$\begin{aligned} E(\boldsymbol{yy'}) &= E\{(\boldsymbol{X\beta} + \boldsymbol{\varepsilon})(\boldsymbol{X\beta} + \boldsymbol{\varepsilon})' \\ &= E\{(\boldsymbol{X\beta} + \boldsymbol{\varepsilon})(\boldsymbol{\beta'X'} + \boldsymbol{\varepsilon'}) \\ &= E\{\boldsymbol{X\beta\beta'X'} + \boldsymbol{\varepsilon\beta'X'} + \boldsymbol{X\beta\varepsilon'} + \boldsymbol{\varepsilon\varepsilon'}\} \\ &= \boldsymbol{X\beta\beta'X'} + E(\boldsymbol{\varepsilon})\boldsymbol{\beta'X'} + \boldsymbol{X\beta}E(\boldsymbol{\varepsilon'}) + E(\boldsymbol{\varepsilon\varepsilon'}) \\ &= \boldsymbol{X\beta\beta'X'} + \boldsymbol{V} \qquad ; E(\boldsymbol{\varepsilon}) = 0, E(\boldsymbol{\varepsilon\varepsilon'}) = V(\boldsymbol{\varepsilon}) = \boldsymbol{V} \end{aligned}$$

แทนค่า $E(\boldsymbol{y})$, $E(\boldsymbol{y'})$ และ $E(\boldsymbol{yy'})$ ลงในสมการ $M(\boldsymbol{\beta})$ จะได้ว่า

$$\begin{aligned} M(\boldsymbol{\beta}) &= \boldsymbol{L'}(\boldsymbol{X\beta\beta'X'} + \boldsymbol{V})\boldsymbol{L} - \boldsymbol{\beta}(\boldsymbol{\beta'X'})\boldsymbol{L} - \boldsymbol{L'}(\boldsymbol{X\beta})\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}' \\ &= \boldsymbol{L'X\beta\beta'X'L} + \boldsymbol{L'VL} - \boldsymbol{\beta\beta'X'L} - \boldsymbol{L'X\beta\beta'} + \boldsymbol{\beta}\boldsymbol{\beta}' \end{aligned}$$

- *Find unbiased condition*

3) สร้างเงื่อนไขที่จะทำให้ตัวประมาณ $\hat{\boldsymbol{\beta}} = \boldsymbol{L}'\boldsymbol{y}$ มีคุณสมบัติไม่เอนเอียง (unbiasedness) ซึ่งพบว่า $\hat{\boldsymbol{\beta}}$ จะมีคุณสมบัติดังกล่าวเมื่อ $E(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \boldsymbol{0}$

จาก

$$\begin{aligned} E(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) &= E(\boldsymbol{\beta} - \boldsymbol{L}'\boldsymbol{y}) \\ &= E(\boldsymbol{\beta}) - \boldsymbol{L}'E(\boldsymbol{y}) \\ &= \boldsymbol{\beta} - \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta} \qquad ; E(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta} \\ &= \boldsymbol{\beta}(\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X}) \end{aligned}$$

ดังนั้น $\hat{\boldsymbol{\beta}}$ จะมีคุณสมบัติไม่เอนเอียงเมื่อ $(\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X})\boldsymbol{\beta} = \boldsymbol{0}$ และเนื่องจาก $\boldsymbol{\beta}$ เป็นค่าจำนวนจริงใดๆ ทำให้สมการข้างต้นจะเป็นจริงได้ต่อเมื่อ $(\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X}) = \boldsymbol{0}$

- *Construct restricted mean squares of estimation error with Lagrange multiplier.*

4) สร้างเงื่อนไข unbiased estimator ที่กำหนดไว้ใน 3) ให้กับ $M(\boldsymbol{\beta})$ โดยการใช้เทคนิคของการเพิ่ม Lagrange multiplier ให้กับฟังก์ชัน ดังนั้นหากกำหนดให้ $\boldsymbol{\theta}$ เป็น Lagrange multiplier vector จะได้ว่าฟังก์ชันที่เราต้องการทำให้มีค่าต่ำสุดจะเพิ่มส่วนของผลคูณของ $\boldsymbol{\theta}$ กับ condition ที่ต้องการ (condition ที่ต้องการในที่ได้จากข้อ 3) ซึ่งได้แก่ $(\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X}) = \boldsymbol{0}$) ดังนั้น

$$\tilde{M}(\boldsymbol{\beta}) = M(\boldsymbol{\beta}) + (\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X})\boldsymbol{\theta}$$

เพื่อความสะดวกในขั้นต่อไป เราสามารถกำหนด Lagrange muliplier ให้อยู่ในรูปสองเท่าของ condition matrix เนื่องจากหาก $(\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X}) = \boldsymbol{0}$ แล้ว $2(\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X}) = \boldsymbol{0}$ ดังนั้นตัวประมาณจะถูกสร้างจาก

$$\begin{aligned} \tilde{M}(\boldsymbol{\beta}) &= M(\boldsymbol{\beta}) + (\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X})2\boldsymbol{\theta} \\ &= \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} + \boldsymbol{L}'\boldsymbol{V}\boldsymbol{L} - \boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} - \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}' + (\boldsymbol{I} - \boldsymbol{L}'\boldsymbol{X})2\boldsymbol{\theta} \\ &= \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} + \boldsymbol{L}'\boldsymbol{V}\boldsymbol{L} - \boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} - \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{\beta}\boldsymbol{\beta}' + 2\boldsymbol{\theta} - 2\boldsymbol{L}'\boldsymbol{X}\boldsymbol{\theta} \\ &= \boldsymbol{L}'\boldsymbol{V}\boldsymbol{L} + \boldsymbol{\beta}\boldsymbol{\beta}' + 2\boldsymbol{\theta} - 2\boldsymbol{L}'\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} - \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} \end{aligned}$$

ดังนั้นในขั้นต่อไป ตัวประมาณจะถูกสร้างจากการ $Min\{\tilde{M}(\boldsymbol{\beta})\}$

- *Minimize estimation error variance*

5) ทำการ $Min\{\tilde{M}(\boldsymbol{\beta})\}$ เพื่อหา $\boldsymbol{L}$ และ $\boldsymbol{\theta}$ ที่ทำให้ estimator error variance มีค่าต่ำสุด โดยหาอนุพันธ์ของ $\tilde{M}(\boldsymbol{\beta})$ ตาม $\boldsymbol{L}$ และ $\boldsymbol{\theta}$ แล้วเทียบให้เท่ากับศูนย์ซึ่งจะได้ว่า

a) หาค่าอนุพันธ์ตาม $\boldsymbol{L}$

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{L}}\tilde{M}(\boldsymbol{\beta}) &= \frac{\partial}{\partial \boldsymbol{L}}\{\boldsymbol{L}'\boldsymbol{V}\boldsymbol{L} + \boldsymbol{\beta}\boldsymbol{\beta}' + 2\boldsymbol{\theta} - 2\boldsymbol{L}'\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} - \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L}\} \\ &= 2\boldsymbol{V}\boldsymbol{L} - 2\boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' - \boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' + 2\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} \\ &= 2\boldsymbol{V}\boldsymbol{L} - 2\boldsymbol{X}\boldsymbol{\theta} - 2\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}' + 2\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L} \\ &= 2\{\boldsymbol{V}\boldsymbol{L} - \boldsymbol{X}\boldsymbol{\theta} - \boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'(\boldsymbol{I} + \boldsymbol{X}'\boldsymbol{L})\} \end{aligned}$$

เนื่องจากข้อกำหนดในขั้นตอนที่ 3) ต้องการให้ตัวประมาณที่จะสร้างขึ้นมี $(\boldsymbol{I}-\boldsymbol{L}'\boldsymbol{X})=(\boldsymbol{I}-\boldsymbol{X}'\boldsymbol{L})=\boldsymbol{\theta}$ ซึ่งทำให้ได้ว่าค่า $\boldsymbol{L}$ สมการอนุพันธ์นี้จะเป็นจริงเมื่อ $\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'(\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{L})=\boldsymbol{\theta}$ ดังนั้นเมื่อเทียบอนุพันธ์ให้เท่ากับศูนย์จะได้ว่า

$$\begin{aligned}\boldsymbol{V}\boldsymbol{L}-\boldsymbol{X}\boldsymbol{\theta}&=\boldsymbol{\theta}\\ \boldsymbol{V}\boldsymbol{L}&=\boldsymbol{X}\boldsymbol{\theta}\\ \boldsymbol{L}&=\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\theta}\qquad\cdots(i)\end{aligned}$$

b) หาค่าอนุพันธ์ตาม $\boldsymbol{\theta}$

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\theta}}\tilde{M}(\boldsymbol{\beta})&=\frac{\partial}{\partial\boldsymbol{\theta}}\{\boldsymbol{L}'\boldsymbol{V}\boldsymbol{L}+\boldsymbol{\beta}\boldsymbol{\beta}'+2\boldsymbol{\theta}-2\boldsymbol{L}'\boldsymbol{X}\boldsymbol{\theta}-\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L}-\boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'+\boldsymbol{L}'\boldsymbol{X}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{L}\}\\ &=2\boldsymbol{I}-2\boldsymbol{X}'\boldsymbol{L}\\ &=2(\boldsymbol{I}-\boldsymbol{X}'\boldsymbol{L})\end{aligned}$$

ดังนั้นเมื่อเทียบอนุพันธ์ให้เท่ากับศูนย์จะได้ว่า

$$\begin{aligned}\boldsymbol{I}-\boldsymbol{X}'\boldsymbol{L}&=0\\ \boldsymbol{X}'\boldsymbol{L}&=\boldsymbol{I}\qquad\cdots(ii)\end{aligned}$$

เมื่อแทนค่า $\boldsymbol{L}$ จาก (i) ลงใน (ii) จะได้ว่า

$$\begin{aligned}\boldsymbol{X}'(\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\theta})&=\boldsymbol{I}\\ \boldsymbol{\theta}&=(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\qquad\cdots(iii)\end{aligned}$$

เมื่อแทนค่า $\boldsymbol{\theta}$ จาก (iii) ลงใน (i) จะได้ว่า

$$\boldsymbol{L}=\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}$$

6) เมื่อแทนค่า $\boldsymbol{L}'=(\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}$ ลงในสมการ $\hat{\boldsymbol{\beta}}=\boldsymbol{L}'\boldsymbol{y}$ ทำให้ได้ว่าตัวประมาณของ $\boldsymbol{\beta}$ ด้วยวิธี BLUE มีรูปแบบเป็น

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

V. สรุป

- โมเดลเชิงเส้น (linear model) มีรูปทั่วไปเป็น $y=a+bx$ ในทางสถิติสามารถเรียกรูปสมการที่มีค่าอิทธิพลต่างๆ อยู่ในรูปของผลบวกลักษณะนี้ว่าโมเดลแบบบวกสะสม (additive model) หากพิจารณาโมเดลเชิงเส้นในรูปของการประเมินอิทธิพลเนื่องจากตัวแปรอิสระ (independent variables) ที่มีต่อตัวแปรตาม (dependent variable) โดยมีวัตถุประสงค์หลักเพื่อสร้างสมการทำนาย (prediction equation) จะนิยมเรียกว่าโมเดลเชิงเส้นนั้นว่าสมการถดถอย (regression)

- โมเดลเชิงเส้นสามารถเขียนให้อยู่ในรูปของเมทริกซ์ได้เป็น $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์สุ่ม (random vector) ของค่าสังเกต, $\boldsymbol{X}$ เป็นเมทริกซ์ของตัวแปรอิสระหรือ incidence matrix หรือ design matrix หรือ dummy-coding matrix, $\boldsymbol{\beta}$ เป็นเวคเตอร์ของพารามิเตอร์ และ $\boldsymbol{\varepsilon}$ เป็นเวคเตอร์สุ่มของความคลาดเคลื่อน
- ในแต่ละโมเดลควรระบุข้อกำหนดให้ชัดเจน เช่น $\boldsymbol{\varepsilon} \sim (\boldsymbol{0}, \boldsymbol{I}\sigma_e^2)$ แสดงถึง $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ และ $V(\boldsymbol{\varepsilon}) = \boldsymbol{I}\sigma_e^2$ หมายถึงค่าเฉลี่ยของความคลาดเคลื่อนของประชากรมีค่าเป็น $\boldsymbol{0}$ โดยมีความแปรปรวนของความคลาดเคลื่อนเป็น $\boldsymbol{I}\sigma_e^2$ เมื่อทราบข้อกำหนดของโมเดลทำให้สามารถหาค่าคาดคะเน (expectation) และความแปรปรวน (variance) อื่นๆ ของโมเดลได้ เช่น เมื่อ $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ โดยมี $\boldsymbol{\varepsilon} \sim (\boldsymbol{0}, \boldsymbol{I}\sigma_e^2)$ ดังนั้น $E(\boldsymbol{y}) = \boldsymbol{X\beta}$ และ $V(y) = \boldsymbol{I}\sigma_e^2$ เป็นต้น
- การประมาณค่าพารามิเตอร์ของโมเดลเชิงเส้นเมื่อมีแต่อิทธิพลคงที่ (fixed effects) ที่กล่าวถึงในบทนี้มี 4 วิธีได้แก่วิธี OLS (Ordinary Least Squares), GLS (Generalized Least squares), ML (Maximum Likelihood) และ BLUE (Best Linear Unbiased Estimator) ตัวประมาณที่ได้จาก OLS จะมีค่าเป็น $\hat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-}\boldsymbol{X'y}$ ในขณะที่ตัวประมาณที่ได้จาก GLS, ML และ BLUE จะเหมือนกัน ได้แก่ $\hat{\boldsymbol{\beta}} = (\boldsymbol{X'V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X'V}^{-1}\boldsymbol{y}$ อย่างไรก็ตามในการใช้การประมาณค่าแบบ OLS ต้องการข้อกำหนดว่าความแปรปรวนของความคลาดเคลื่อนมีค่าเท่ากับ $\boldsymbol{I}\sigma_e^2$ หรือไม่มี covariance ระหว่างความคลาดเคลื่อน หากข้อมูลที่ได้อาจไม่เป็นไปตามข้อกำหนดดังกล่าวควรใช้ตัวประมาณแบบ GLS, ML หรือ BLUE เนื่องจากมีการ standardize โมเดลด้วย $\boldsymbol{V}$ ซึ่งเป็นเมทริกซ์ความแปรปรวนของค่าสังเกต อย่างไรก็ตามตัวประมาณแบบ ML ต้องการข้อกำหนดที่ว่าการแจกแจงของข้อมูลต้องเป็น multivariate normal distribution ในขณะที่ตัวประมาณแบบอื่นไม่จำเป็นต้องมีข้อกำหนดนี้

บรรณานุกรม

Henderson, C.R. 1973. Sire evaluation and genetic trends. pp 10-41. In Proceeding of The Animal Breeding and Genetics Symposium in Honor of Dr. Jay L. Lush. American Society of Animal Science, IL.

Lynch, M., and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc., NY.

Raymond, H.M., and J.S. Milton. 1991. A First Course in the Theory of Linear Statistical Models. PWS-KENT Publishing Company, Boston.

Schaeffer, L.R. 1994. Course Note in Linear model. Guelph University, Guelph.

Searle, S.R. 1982. Matrix Algebra Useful for Statistics. John Wiley and Sons, Toronto.

คำถามท้ายบท III

1. กำหนดให้

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad , \boldsymbol{\varepsilon} \sim \boldsymbol{(0,V)}$$

เมื่อ $\boldsymbol{V} = \boldsymbol{R}\sigma_e^2$ โดย $\boldsymbol{R}$ เป็นเมทริกซ์ค่าสหสัมพันธ์ของความคลาดเคลื่อน จงพิสูจน์ว่าหากแปลงตัวแบบให้อยู่ในรูป $\boldsymbol{R}^{-1/2}\boldsymbol{y}$ แล้วจะได้ว่า $\boldsymbol{\varepsilon} \sim \boldsymbol{(0, I\sigma_e^2)}$

2. จงพิสูจน์ว่า $E(\boldsymbol{K'\hat{\beta}}) = \boldsymbol{K'\beta}$ และ $Var(\hat{\boldsymbol{\beta}}) = \boldsymbol{(X'X)}^{-}\sigma_e^2$ เมื่อ $\hat{\boldsymbol{\beta}}$ เป็นตัวประมาณแบบ OLS

3. กำหนดให้น้ำหนักแรกเกิดในโคจากงานวิจัยครั้งหนึ่ง มีข้อมูลดังตาราง

ID	Sex	Breed	BW
1	F	B	34
2	F	B	36
3	M	B	28
4	M	A	32
5	F	A	41
9	F	C	38

และกำหนดให้ตัวแบบของน้ำหนักแรกเกิดเป็น

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad , \boldsymbol{\varepsilon} \sim \boldsymbol{(0,V)}$$

เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์ของค่าสังเกตน้ำหนักแรกเกิด, $\boldsymbol{X}$ เป็น incidence matrix, $\boldsymbol{\beta}$ เป็นเวคเตอร์เนื่องจากอิทธิพลคงที่เนื่องจากเพศและพันธุ์, และ $\boldsymbol{\varepsilon}$ เป็นความคลาดเคลื่อน โดยมีความแปรปรวนเป็น $\boldsymbol{V}$

โดยกำหนดให้

$$\boldsymbol{V} = \begin{bmatrix} 10 & 1 & 1 & 1 & 1 & 1 \\ 1 & 10 & 1 & 1 & 1 & 1 \\ 1 & 1 & 10 & 5 & 5 & 5 \\ 1 & 1 & 5 & 10 & 5 & 5 \\ 1 & 1 & 5 & 5 & 10 & 5 \\ 1 & 1 & 5 & 5 & 5 & 10 \end{bmatrix}$$

i) จงประมาณค่าของอิทธิพลเนื่องจากพันธุ์และเพศที่มีต่อน้ำหนักลูกโค ด้วยวิธี OLS

ii) จงประมาณค่าของอิทธิพลในข้อ i) อีกครั้งด้วยวิธี GLS

iii) กำหนดให้ $\boldsymbol{V} = \boldsymbol{I}\sigma_e^2$ จากนั้นทำซ้ำในข้อ ii) อีกครั้ง

iv) เปรียบเทียบค่าประมาณที่ได้จากข้อ i, ii, iii พร้อมทั้งวิจารณ์

v) คำนวณค่า LS-MEAN หรือ Least squares mean ของแต่ละเพศและแต่ละพันธุ์

vi) ตรวจสอบ estimability ต่อไปนี้

a) $\beta_F - \beta_M$

b) $\beta_A - \beta_B$

c) $LSM(breedA) - LSM(breedB)$

d) $LSM(sexF) - LSM(breedA)$

เมื่อ β_F, β_M เป็นอิทธิพลเนื่องจากเพศเมียและเพศผู้ และ เป็นอิทธิพลเนื่องจากพันธุ์ A และ B ตามลำดับ LSM = Least squares mean

4. กำหนดให้

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad , \boldsymbol{\varepsilon} \sim \boldsymbol{(0, R)}$$

เมื่อ $\boldsymbol{y}$ เป็นเวคเตอร์ของค่าสังเกต, $\boldsymbol{X}$ เป็น incident matrix, $\boldsymbol{\beta}$ เป็นเวคเตอร์เนื่องจากอิทธิพลคงที่, และ $\boldsymbol{\varepsilon}$ เป็นความคลาดเคลื่อน โดยมีความแปรปรวนเป็น $\boldsymbol{R}$ จงสร้างตัวประมาณค่าของ $\boldsymbol{\beta}$ ด้วยวิธี Maximum likelihood