

$$E(y) = E \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \alpha + \beta - \gamma \\ \alpha + \beta - \gamma \\ \alpha + \beta + \gamma \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}}_{\text{rank}(X)=2} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

$$\text{Examples } t^T y \quad \Rightarrow \quad E(t^T y) = q^T b$$

$$y_1 = (1 \ 0 \ 0) y$$

$$\alpha + \beta - \gamma$$

$$y_2 = (0 \ 1 \ 0) y$$

$$\alpha + \beta - \gamma$$

$$y_3 = (0 \ 0 \ 1) y$$

$$\alpha + \beta + \gamma$$

$$y_1 + y_2 = (1 \ 1 \ 0) y$$

$$2\alpha + 2\beta - 2\gamma$$

$$\frac{1}{2}(y_1 + y_2) = \frac{1}{2}(1 \ 1 \ 0) y$$

$$\alpha + \beta - \gamma$$

$$\frac{1}{2}(y_2 + y_3) = \frac{1}{2}(0 \ 1 \ 1) y$$

$$\alpha + \beta$$

$$\frac{1}{2}(y_3 - y_2) = \frac{1}{2}(0 \ -1 \ 1) y \quad \gamma$$

No $t^T y$ s.t. $E(t^T y)$

α } Not estimable
 β } functions

$$\text{rank}(X) = 2$$

$$\text{Augmented matrix } \begin{pmatrix} q' \\ X \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & q_3 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\text{rank} \begin{pmatrix} q' \\ X \end{pmatrix} = 2 \Rightarrow q_1 = q_2$$

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

Example $E(t^T y) = q^T b$

$$\alpha + \beta - \gamma = (\textcircled{1} \textcircled{0} \textcircled{1}) b \quad \rightarrow b$$

$\underline{\underline{q_1 = q_2}}$

$$\alpha + \beta + \gamma = (\textcircled{1} \textcircled{1} \textcircled{1}) b$$

$\underline{\underline{q_1 = q_2}}$

But $\alpha = (1 \ 0 \ 0) b$
 $\Rightarrow q_1 \neq q_2$

$$b = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

Possible estimators

$$y_1 = (1 \ 0 \ 0) y$$

$$y_2 = (0 \ 1 \ 0) y$$

$$\frac{1}{2}(y_1 + y_2) = \frac{1}{2}(1 \ 1 \ 0) y$$

$$y_3 = (0 \ 0 \ 1) y$$

Not estimable

THE LINEAR MODEL

- Estimable Linear Functions and the Gauss-Markov Theorem

Definition: The linear function of the parameters, $\mathbf{q}'\mathbf{b}$, is estimable if it has an unbiased linear estimator, that is, if a $\mathbf{t}'\mathbf{y}$ exists such that

$$E(\mathbf{t}'\mathbf{y}) = \mathbf{q}'\mathbf{b}.$$

Theorem: $\mathbf{q}'\mathbf{b}$ estimable $\Leftrightarrow \text{rank } \mathbf{X} = \text{rank} \begin{pmatrix} \mathbf{q}' \\ \mathbf{X} \end{pmatrix}$.

The implication is that $\mathbf{q}'\mathbf{b}$ is estimable if and only if a \mathbf{t} exists such that $\mathbf{q}' = \mathbf{t}'\mathbf{X}$.

Proof:

$$\begin{aligned} \mathbf{q}'\mathbf{b} \text{ estimable} &\Leftrightarrow \exists \mathbf{t} \ni E(\mathbf{t}'\mathbf{y}) = \mathbf{t}'\mathbf{X}\mathbf{b} \equiv \mathbf{q}'\mathbf{b} \quad \forall \mathbf{b} \\ &\Leftrightarrow \mathbf{q}' = \mathbf{t}'\mathbf{X} \text{ for some } \mathbf{t}. \end{aligned}$$

If $\text{rank}(\mathbf{X}) = p$ all linear functions of the parameters are estimable, because the condition of the Theorem is always met in this case. All the parameters are then also estimable. In this case $\mathbf{X}'\mathbf{X}$ is non-singular and the solution of the normal equations is unique.

Theorem (Gauss-Markov): Consider the estimable function $\mathbf{q}'\mathbf{b}$.

1. A unique unbiased linear estimator, $\mathbf{t}_0'\mathbf{y}$, of $\mathbf{q}'\mathbf{b}$ exists such that $\mathbf{t}_0 \in V_r$. If $\mathbf{t}'\mathbf{y}$ is any unbiased estimator of $\mathbf{q}'\mathbf{b}$, then \mathbf{t}_0 is the projection of \mathbf{t} on V_r .
2. $\text{var}(\mathbf{t}_0'\mathbf{y}) \leq \text{var}(\mathbf{t}'\mathbf{y})$ and the equality holds if and only if $\mathbf{t} = \mathbf{t}_0$.
3. $\mathbf{t}_0'\mathbf{y} = \mathbf{q}'\hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is any set of least squares estimators of the parameters.

Proof:

1. Let $\mathbf{t}'\mathbf{y}$ be any unbiased linear estimator:

$$E(\mathbf{t}'\mathbf{y}) = \mathbf{q}'\mathbf{b}.$$

Let $\mathbf{t} = \mathbf{t}_0 + \mathbf{e}$, where $\mathbf{t}_0 \in V_r$ and $\mathbf{e} \in V_e$. Then

$q^T b$ is a linear combination of the rows of $x b$

$$E(y) = E \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \alpha + \beta - \gamma \\ \alpha - \beta - \gamma \\ \alpha + \beta + \gamma \end{pmatrix} \begin{matrix} \rightarrow \text{row 1} \equiv R_1 \\ \rightarrow \text{row 2} \equiv R_2 \\ \rightarrow \text{row 3} \equiv R_3 \end{matrix}$$

$q^T b$	$t^T x b$
$\alpha + \beta - \gamma$	$R_1 + 0R_2 + 0R_3$ OR $0R_1 + R_2 + 0R_3$ OR $\frac{1}{2}R_1 + \frac{1}{2}R_2 + 0R_3$

$\alpha + \beta + \gamma$	$0R_1 + 0R_2 + R_3 \equiv R_3$
---------------------------	--------------------------------

$2\alpha + 2\beta - 2\gamma$	$R_1 + R_2 + 0R_3$
------------------------------	--------------------

γ	$0R_1 - \frac{1}{2}R_2 + \frac{1}{2}R_3$
----------	--

α	}	No linear combination of the rows of $x b$ yields α or β
β		

$$E(\mathbf{t}'_0 \mathbf{y}) = E(\mathbf{t}' \mathbf{y}) = \mathbf{q}' \mathbf{b}$$

since $E(\mathbf{e}' \mathbf{y}) = \mathbf{0}$. Therefore $\mathbf{t}'_0 \mathbf{y}$ is an unbiased estimator of $\mathbf{q}' \mathbf{b}$ and $\mathbf{t}_0 \in V_r$. Suppose that the same holds for $\mathbf{t}'_1 \mathbf{y}$. Then

$$E(\mathbf{t}'_0 \mathbf{y}) - E(\mathbf{t}'_1 \mathbf{y}) = (\mathbf{t}_0 - \mathbf{t}_1)' \mathbf{X} \mathbf{b} \equiv \mathbf{0} \quad \forall \mathbf{b}.$$

Therefore $\mathbf{t}_0 - \mathbf{t}_1 \in V_e$. But $\mathbf{t}_0 \in V_r, \mathbf{t}_1 \in V_r$ and $(\mathbf{t}_0 - \mathbf{t}_1) \in V_r$. Therefore $\mathbf{t}_0 - \mathbf{t}_1 = \mathbf{0}$ and $\mathbf{t}_1 = \mathbf{t}_0$.

$$2. \text{var}(\mathbf{t}' \mathbf{y}) = \mathbf{t}' \mathbf{t} \sigma^2$$

$$= (\mathbf{t}_0 + \mathbf{e})'(\mathbf{t}_0 + \mathbf{e}) \sigma^2$$

$$= \mathbf{t}'_0 \mathbf{t}_0 \sigma^2 + \mathbf{e}' \mathbf{e} \sigma^2 \text{ because } \mathbf{t}_0 \perp \mathbf{e}$$

$$= \text{var}(\mathbf{t}'_0 \mathbf{y}) + \text{var}(\mathbf{e}' \mathbf{y})$$

$$\geq \text{var}(\mathbf{t}'_0 \mathbf{y})$$

3. $E(\mathbf{t}'_0 \mathbf{y}) = \mathbf{t}'_0 \mathbf{X} \mathbf{b} \equiv \mathbf{q}' \mathbf{b} \quad \forall \mathbf{b}$. Therefore $\mathbf{q}' = \mathbf{t}'_0 \mathbf{X}$. But $\mathbf{t}'_0 (\mathbf{y} - \mathbf{X} \hat{\mathbf{b}}) = \mathbf{0}$ since $\mathbf{t}_0 \perp V_e$. Therefore

$$\mathbf{t}'_0 \mathbf{y} = \mathbf{t}'_0 \mathbf{X} \hat{\mathbf{b}} = \mathbf{q}' \hat{\mathbf{b}}.$$

The estimator $\mathbf{q}' \hat{\mathbf{b}} = \mathbf{t}'_0 \mathbf{y}$ is called the best estimator of $\mathbf{q}' \mathbf{b}$. Conversely, for any $\mathbf{d} \in V_r, \mathbf{d}' \mathbf{y}$ is the best estimator of the estimable linear function $E(\mathbf{d}' \mathbf{y}) = \mathbf{d}' \mathbf{X} \mathbf{b}$. Therefore V_r is called the estimation space and the linear set

$$L_r = \{\mathbf{t}' \mathbf{y} : \mathbf{t} \in V_r\}$$

is called the estimation set. The sum of squares of the estimation set is:

$$SSR = \hat{\mathbf{b}}' \mathbf{X}' \mathbf{X} \hat{\mathbf{b}} = \hat{\mathbf{b}}' \mathbf{X}' \mathbf{y} = \mathbf{y}' \mathbf{y} - SSE$$

Suppose that the columns of $\mathbf{E}: n \times (n - r)$ generate V_e . Then

$$SSE = \mathbf{y}' \mathbf{E} (\mathbf{E}' \mathbf{E})^{-1} \mathbf{E}' \mathbf{y}.$$

It also follows that

$$E(SSE) = (n - r) \sigma^2$$

and

$$MSE = SSE/(n - r)$$

is an unbiased estimator of σ^2 , based on all the degrees of freedom for error.

Theorem: Suppose that the elements of $\mathbf{Q}'\mathbf{b}$: $s \times 1$ are estimable with best estimators the elements of $\mathbf{T}'\mathbf{y}$. Then

- $\mathbf{h}'\mathbf{Q}'\mathbf{b}$ is estimable with best estimator $\mathbf{h}'\mathbf{T}'\mathbf{y}$ and
- $\text{rank}(\mathbf{Q}) = \text{rank}(\mathbf{T})$.

Proof:

- $E(\mathbf{h}'\mathbf{T}'\mathbf{y}) = \mathbf{h}'\mathbf{Q}'\mathbf{b}$ and $\mathbf{T}\mathbf{h} \in V_r$.
- $E(\mathbf{T}'\mathbf{y}) = \mathbf{T}'\mathbf{X}\mathbf{b} \equiv \mathbf{Q}'\mathbf{b} \quad \forall \mathbf{b}$.

Therefore

$$\mathbf{Q}' = \mathbf{T}'\mathbf{X}$$

and $\text{rank}(\mathbf{Q}) \leq \text{rank}(\mathbf{T})$.

Also, $\mathbf{T}'\mathbf{y} = \mathbf{Q}'\hat{\mathbf{b}} = \mathbf{Q}'(\mathbf{X}'\mathbf{X})^*\mathbf{X}'\mathbf{y}$ for all values of \mathbf{y} ,

$$\mathbf{T}' = \mathbf{Q}'(\mathbf{X}'\mathbf{X})^*\mathbf{X}'$$

and $\text{rank}(\mathbf{T}) \leq \text{rank}(\mathbf{Q})$.

It follows that $\text{rank}(\mathbf{Q}) = \text{rank}(\mathbf{T})$.

For the covariance matrix of $\mathbf{T}'\mathbf{y}$, the best estimator of $\mathbf{Q}'\mathbf{b}$, it follows that:

$$\begin{aligned} \text{Cov}(\mathbf{T}'\mathbf{y}) &= \mathbf{T}'\mathbf{T}\sigma^2 \\ &= \mathbf{Q}'(\mathbf{X}'\mathbf{X})^*\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^*\mathbf{Q}\sigma^2 \\ &= \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^*\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^*\mathbf{X}'\mathbf{T}\sigma^2 \\ &= \mathbf{T}'\mathbf{X}(\mathbf{X}'\mathbf{X})^*\mathbf{X}'\mathbf{T}\sigma^2 \text{ since } \mathbf{X}(\mathbf{X}'\mathbf{X})^*\mathbf{X}' \text{ is idempotent} \\ &= \mathbf{Q}'(\mathbf{X}'\mathbf{X})^*\mathbf{Q}\sigma^2. \end{aligned}$$

If $\text{rank}(\mathbf{X}) = p$ then \mathbf{b} is estimable and

$$\text{Cov}(\hat{\mathbf{b}}) = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2.$$

Examples

1.

$$E(\mathbf{y}) = \begin{pmatrix} \mu + \alpha + \beta + \gamma \\ \mu + \alpha + \beta + \gamma \\ \mu - \alpha - \beta + \gamma \\ \mu - \alpha - \beta + \gamma \\ \mu + \alpha + \beta - \gamma \\ \mu + \alpha + \beta - \gamma \\ \mu - \alpha - \beta - \gamma \\ \mu - \alpha - \beta - \gamma \end{pmatrix} = \begin{pmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix}$$

and $cov(\mathbf{y}) = \sigma^2 \mathbf{I}$. Consider the linear function $\mathbf{q}'\mathbf{b} = q_1\mu + q_2\alpha + q_3\beta + q_4\gamma$.

$$\text{rank} \begin{pmatrix} q_1 & q_2 & q_3 & q_4 \\ 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} = \text{rank} \begin{pmatrix} q_1 & q_2 & q_2 - q_3 & q_4 \\ 1 & +1 & 0 & +1 \\ 1 & +1 & 0 & +1 \\ 1 & -1 & 0 & +1 \\ 1 & -1 & 0 & +1 \\ 1 & +1 & 0 & -1 \\ 1 & +1 & 0 & -1 \\ 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & -1 \end{pmatrix} = \text{rank}(\mathbf{X}) = 3$$

if and only if $q_2 = q_3$. $\mathbf{q}'\mathbf{b}$ is therefore estimable if and only if $q_2 = q_3$.

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

$$\text{and } \mathbf{X}'\mathbf{y} = \begin{pmatrix} y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 \\ y_1 + y_2 - y_3 - y_4 + y_5 + y_6 - y_7 - y_8 \\ y_1 + y_2 - y_3 - y_4 + y_5 + y_6 - y_7 - y_8 \\ y_1 + y_2 + y_3 + y_4 - y_5 - y_6 - y_7 - y_8 \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_2 \\ Y_3 \end{pmatrix}$$

A solution of the normal equations is:

$$\hat{\mu} = Y_1/8, \hat{\alpha} = Y_2/8, \hat{\beta} = 0, \hat{\gamma} = Y_3/8.$$

The best estimator of $\mathbf{q}'\mathbf{b}$ is

$$q_1\hat{\mu} + q_2\hat{\alpha} + q_3\hat{\beta} + q_4\hat{\gamma} = q_1Y_1/8 + q_2Y_2/8 + q_4Y_3/8.$$

Furthermore,

$$SSR = \hat{\mathbf{b}}' \mathbf{X}' \mathbf{y} = \frac{1}{8} Y_1^2 + \frac{1}{8} Y_2^2 + \frac{1}{8} Y_3^2$$

$$SSE = \mathbf{y}' \mathbf{y} - SSR$$

and

$$MSE = \frac{1}{5} SSE$$

2. Suppose that each combination of three out of four objects with weights w_1, w_2, w_3 and w_4 is weighed twice. A constant error, g , is attributed to the scale used and a random error e is made at each observation, such that $\text{var}(e) = \sigma^2$.

$$E(\mathbf{y}) = \begin{pmatrix} g + w_1 + w_2 + w_3 \\ g + w_1 + w_2 + w_3 \\ g + w_1 + w_2 + w_4 \\ g + w_1 + w_2 + w_4 \\ g + w_1 + w_3 + w_4 \\ g + w_1 + w_3 + w_4 \\ g + w_2 + w_3 + w_4 \\ g + w_2 + w_3 + w_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} g \\ w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$$

and $\text{cov}(\mathbf{y}) = \sigma^2 \mathbf{I}$. Consider the linear function $\mathbf{q}'\mathbf{b} = q_0 g + q_1 w_1 + q_2 w_2 + q_3 w_3 + q_4 w_4$.

$$\text{rank} \begin{pmatrix} q_0 & q_1 & q_2 & q_3 & q_4 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} = \text{rank} \begin{pmatrix} 3q_0 - q_1 - q_2 - q_3 - q_4 & q_1 & q_2 & q_3 & q_4 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} = \text{rank}(\mathbf{X}) = 4$$

if and only if $3q_0 = q_1 + q_2 + q_3 + q_4$. Therefore $\mathbf{q}'\mathbf{b}$ is estimable if and only if $3q_0 = q_1 + q_2 + q_3 + q_4$.

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 8 & 6 & 6 & 6 & 6 \\ 6 & 6 & 4 & 4 & 4 \\ 6 & 4 & 6 & 4 & 4 \\ 6 & 4 & 4 & 6 & 4 \\ 6 & 4 & 4 & 4 & 6 \end{pmatrix}$$

$$\text{and } \mathbf{X}'\mathbf{y} = \begin{pmatrix} y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 \\ y_1 + y_2 + y_3 + y_4 + y_5 + y_6 \\ y_1 + y_2 + y_3 + y_4 + y_7 + y_8 \\ y_1 + y_2 + y_5 + y_6 + y_7 + y_8 \\ y_3 + y_4 + y_5 + y_6 + y_7 + y_8 \end{pmatrix} = \begin{pmatrix} G \\ W_1 \\ W_2 \\ W_3 \\ W_4 \end{pmatrix}$$

where W_i is the sum of the weights when the i -th subject is weighed. The normal equations are:

$$\begin{aligned} 8g + 6w_1 + 6w_2 + 6w_3 + 6w_4 &= G \\ 6g + 6w_1 + 4w_2 + 4w_3 + 4w_4 &= W_1 \\ 6g + 4w_1 + 6w_2 + 4w_3 + 4w_4 &= W_2 \\ 6g + 4w_1 + 4w_2 + 6w_3 + 4w_4 &= W_3 \\ 6g + 4w_1 + 4w_2 + 4w_3 + 6w_4 &= W_4 \end{aligned}$$

These normal equations can be reduced to 4 equations in 5 unknowns. To find a solution, set

$$g + w_1 + w_2 + w_3 + w_4 = 0.$$

(It should not be an estimable function). A solution of the normal equations is:

$$\hat{g} = G/2, \hat{w}_i = W_i/2 - G/2$$

The best estimator of $\mathbf{q}'\mathbf{b}$ is

$$\frac{1}{2}(q_0G + q_1(W_1 - G) + q_2(W_2 - G) + q_3(W_3 - G) + q_4(W_4 - G))$$

It follows, for instance, that the best estimator of $w_1 - w_2$ is:

$$\hat{w}_1 - \hat{w}_2 = \frac{1}{2}(W_1 - W_2) = \frac{1}{2}(y_5 + y_6 - y_7 - y_8).$$

Furthermore,

$$\begin{aligned} SSR &= \hat{\mathbf{b}}'\mathbf{X}'\mathbf{y} \\ &= \frac{1}{2}G^2 + W_1(W_1/2 - G/2) + W_2(W_2/2 - G/2) + W_3(W_3/2 - G/2) + W_4(W_4/2 - G/2) \\ &= \frac{1}{2}\sum_{i=1}^4 W_i^2 - G^2 \end{aligned}$$

$$SSE = \mathbf{y}'\mathbf{y} - SSR$$

$$\text{and } MSE = \frac{1}{4}SSE.$$