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

Examples $t^T y$	$\Rightarrow E(t^T y) = q^T b$
$y_1 = (1 \ 0 \ 0) y$	$\alpha+\beta-\gamma$
$y_2 = (0 \ 1 \ 0) y$	$\alpha+\beta-\gamma$
$y_3 = (0 \ 0 \ 1) y$	$\alpha+\beta+\gamma$
$y_1 + y_2 = (1 \ 1 \ 0) y$	$2\alpha + 2\beta - 2\gamma$
$\frac{1}{2}(y_1+y_2) = \frac{1}{2}(1 \ 1 \ 0) y$	$\alpha+\beta-\gamma$
$\frac{1}{2}(y_2+y_3) = \frac{1}{2}(0 \ 1 \ 1) y$	$\alpha+\beta$
$\frac{1}{2}(y_3-y_2) = \frac{1}{2}(0 \ -1 \ 1) y$	γ

No $t^T y$ s.t. $E(t^T y)$ = α } Not estimable

" " " " = β } functions

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

rank$(X) = 2$

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

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

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

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

$\alpha + \beta - \gamma = (1 \quad 1 \quad -1) b$

$q_1 = q_2$

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

$q_1 = q_2$

But $\alpha = (1 \quad 0 \quad 0) b$

$\Rightarrow q_1 \neq q_2$

Possible estimators

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

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

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

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

Not estimable

THE LINEAR MODEL

- Estimable Linear Functions and the Gauss-Markov Theorem

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

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

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

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

Proof:

$= t^T E(y)$

$$\boldsymbol{q}'\boldsymbol{b} \text{ estimable } \Leftrightarrow \exists \ \ \boldsymbol{t} \ni E(\boldsymbol{t}'\boldsymbol{y}) = \boldsymbol{t}'\boldsymbol{X}\boldsymbol{b} \equiv \boldsymbol{q}'\boldsymbol{b} \quad \forall\, \boldsymbol{b}$$
$$\Leftrightarrow \boldsymbol{q}' = \boldsymbol{t}'\boldsymbol{X} \text{ for some } \boldsymbol{t}.$$

If $rank\ (\boldsymbol{X}) = p$ 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 $\boldsymbol{X}'\boldsymbol{X}$ is non-singular and the solution of the normal equations is unique.

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

1. A **unique unbiased linear estimator**, $\boldsymbol{t}_0'\boldsymbol{y}$, of $\boldsymbol{q}'\boldsymbol{b}$ exists such that $\boldsymbol{t_0} \in V_r$. If $\boldsymbol{t}'\boldsymbol{y}$ is any unbiased estimator of $\boldsymbol{q}'\boldsymbol{b}$, then $\boldsymbol{t_0}$ is the projection of $\boldsymbol{t}$ on V_r.
2. $\mathrm{var}\ (\boldsymbol{t}_0'\boldsymbol{y}) \leq \mathrm{var}\ (\boldsymbol{t}'\boldsymbol{y})$ and the equality holds if and only if $\boldsymbol{t} = \boldsymbol{t_0}$.
3. $\boldsymbol{t}_0'\boldsymbol{y} = \boldsymbol{q}'\widehat{\boldsymbol{b}}$, where $\widehat{\boldsymbol{b}}$ is any set of least squares estimators of the parameters.

Proof:

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

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

Let $\boldsymbol{t} = \boldsymbol{t_0} + \boldsymbol{e}$, where $\boldsymbol{t_0} \in V_r$ and $\boldsymbol{e} \in V_e$. Then

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

$$E(y) = E\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \alpha + \beta - \gamma \\ \alpha + \beta - \gamma \\ \alpha + \beta + \gamma \end{pmatrix} \begin{array}{l} \longrightarrow \text{row } 1 \equiv R1 \\ \longrightarrow \text{row } 2 \equiv R2 \\ \longrightarrow \text{row } 3 \equiv R3 \end{array}$$

$q^T b$	$t^T X b$
$\alpha + \beta - \gamma$	$R1 + 0R2 + 0R3$ OR $0R1 + R2 + 0R3$ OR $\frac{1}{2}R1 + \frac{1}{2}R2 + 0R3$
$\alpha + \beta + \gamma$	$0R1 + 0R2 + R3 = R3$
$2\alpha + 2\beta - 2\gamma$	$R1 + R2 + 0R3$
γ	$0R1 - \frac{1}{2}R2 + \frac{1}{2}R3$

$$\left.\begin{array}{l} \alpha \\ \\ \beta \end{array}\right\} \text{ No linear combination of the rows of } Xb \text{ yields } \alpha \text{ or } \beta$$

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

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

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

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

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

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

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

$$= \mathrm{var}(\boldsymbol{t}_0'\boldsymbol{y}) + \mathrm{var}(\boldsymbol{e}'\boldsymbol{y})$$

$$\geq \mathrm{var}(\boldsymbol{t}_0'\boldsymbol{y})$$

3. $E(\boldsymbol{t}_0'\boldsymbol{y}) = \boldsymbol{t}_0'\boldsymbol{X}\boldsymbol{b} \equiv \boldsymbol{q}'\boldsymbol{b} \;\; \forall\, \boldsymbol{b}$. Therefore $\boldsymbol{q}' = \boldsymbol{t}_0'\boldsymbol{X}$. But $\boldsymbol{t}_0'(\boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{b}}) = \boldsymbol{0}$ since $\boldsymbol{t}_0 \perp V_e$. Therefore

$$\boldsymbol{t}_0'\boldsymbol{y} = \boldsymbol{t}_0'\boldsymbol{X}\widehat{\boldsymbol{b}} = \boldsymbol{q}'\widehat{\boldsymbol{b}}.$$

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

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

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

$$SSR = \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{b}} = \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{y} - SSE$$

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

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

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

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

Proof:

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

Therefore

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

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

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

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

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

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

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

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

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

$$Cov\ (\widehat{\boldsymbol{b}}) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{\sigma}^2.$$

Examples

1.

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

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

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

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

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

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

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 $\boldsymbol{q}'\boldsymbol{b}$ 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,

$$SSR = \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{y} = \frac{1}{8}Y_1^2 + \frac{1}{8}Y_2^2 + \frac{1}{8}Y_3^2$$
$$SSE = \boldsymbol{y}'\boldsymbol{y} - SSR$$

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 $\boldsymbol{e}$ is made at each observation, such that $\operatorname{var}(e) = \sigma^2$.

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

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

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

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

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

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

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 $\boldsymbol{q}'\boldsymbol{b}$ is

$$\frac{1}{2}\big(q_0 G + q_1(W_1 - G) + q_2(W_2 - G) + q_3(W_3 - G) + q_4(W_4 - G)\big)$$

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,

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

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

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