THE LINEAR MODEL

• GENERALIZED t- AND F-TESTS

It is now assumed that the random vector $\boldsymbol{y}$ in the linear model has the $N(\boldsymbol{\eta}, \boldsymbol{\sigma^2 I})$ distribution.

Theorem: Let $\boldsymbol{T}'\boldsymbol{y}$ be the best estimator of $\boldsymbol{Q}'\boldsymbol{b}$. Then, the error sum of squares, SSE, has the $\sigma^2\chi^2(n-r)$ distribution, independent of $\boldsymbol{T}'\boldsymbol{y}$.

Proof:

Suppose that $\boldsymbol{V_e}$ is generated by the columns of $\boldsymbol{E}: n \times (n-r)$. Then $\boldsymbol{E}'\boldsymbol{y}$ has the $N(\boldsymbol{0}, \sigma^2\boldsymbol{E}'\boldsymbol{E})$ distribution. It follows that $SSE = \boldsymbol{y}'\boldsymbol{E}(\boldsymbol{E}'\boldsymbol{E})^{-\boldsymbol{1}}\boldsymbol{E}'\boldsymbol{y}$ has the $\sigma^2\chi^2(n-r)$ distribution. Since $\boldsymbol{T}'\boldsymbol{y}$ and $\boldsymbol{E}'\boldsymbol{y}$ are independent $(\boldsymbol{Cov}\ (\boldsymbol{T}'\boldsymbol{y}, \boldsymbol{E}'\boldsymbol{y}) = \boldsymbol{T}'\boldsymbol{E}\boldsymbol{\sigma^2} = \boldsymbol{0})$, $\boldsymbol{T}'\boldsymbol{y}$ and SSE are independent.

The Generalized $\boldsymbol{t}$-Test

Suppose that $\boldsymbol{q}'\boldsymbol{b}$ is an estimable function with best estimator $\boldsymbol{t}'\boldsymbol{y}$. Therefore

$\boldsymbol{t}'\boldsymbol{y} \sim N(\boldsymbol{q}'\boldsymbol{b}, \sigma^2\boldsymbol{t}'\boldsymbol{t})$, independent of SSE/σ^2, which has the $\chi^2(n-r)$ distribution. Under the null hypothesis:

$$H_0: \boldsymbol{q}'\boldsymbol{b} = m$$

for a given value of m, it follows that

$$\begin{aligned} t &= \frac{\boldsymbol{t}'\boldsymbol{y} - m}{\sqrt{\boldsymbol{t}'\boldsymbol{t}\sigma^2}} / \sqrt{MSE/\sigma^2} \\ &= \frac{\boldsymbol{t}'\boldsymbol{y} - m}{\sqrt{\boldsymbol{t}'\boldsymbol{t}MSE}} \\ &= \frac{\boldsymbol{t}'\boldsymbol{y} - m}{\sqrt{\text{estimated var } (\boldsymbol{t}'\boldsymbol{y})}} \end{aligned}$$

has a $\boldsymbol{t}$ distribution with $n-r$ degrees of freedom under H_0.

A $(1-\epsilon)$ confidence interval for $\boldsymbol{q}'\boldsymbol{b}$ is:

$$\boldsymbol{t}'\boldsymbol{y} \pm t_{\frac{\epsilon}{2}}(n-r)\sqrt{\text{estimated variance of } \boldsymbol{t}'\boldsymbol{y}}$$

where (estimated variance of $\boldsymbol{t}'\boldsymbol{y}$) $= \boldsymbol{t}'\boldsymbol{t}MSE = \boldsymbol{q}'(\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{q}MSE$.

Examples

1. Suppose

$$\begin{pmatrix} \mu+\alpha+\beta+\gamma \\ \mu+\alpha+\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu-\alpha-\beta-\gamma \\ \mu-\alpha-\beta-\gamma \end{pmatrix} = \begin{pmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix}$$

and $cov\,(\boldsymbol{y}) = \sigma^2 \boldsymbol{I}$. Consider the linear function $\boldsymbol{q}'\boldsymbol{b} = q_1\mu + q_2\alpha + q_3\beta + q_4\gamma$. Further, assume $\boldsymbol{y}$ has a normal distribution. Then, variance of best estimator of $\boldsymbol{q}'\boldsymbol{b}$ is given by (independent?)

$$\mathrm{var}\left(q_1\hat{\mu} + q_2\hat{\alpha} + q_3\hat{\beta} + q_4\hat{\gamma}\right) = q_1^2\mathrm{var}(\hat{\mu}) + q_2^2\mathrm{var}(\hat{\alpha}) + q_3^2\mathrm{var}\left(\hat{\beta}\right) + q_4^2\mathrm{var}(\hat{\gamma}).$$

Thus, we need to compute $\mathrm{Cov}\left(\widehat{\boldsymbol{b}}\right)$, which is $\mathrm{Cov}\left(\widehat{\boldsymbol{b}}\right) = (\boldsymbol{X}'\boldsymbol{X})^*\sigma^2$. How can we calculate $(\boldsymbol{X}'\boldsymbol{X})^*$?

First, $\boldsymbol{q}'\boldsymbol{b}$ is estimable iff $q_2 = q_3$. Using the fact that $\widehat{\boldsymbol{b}} = (\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{X}'\boldsymbol{y}$, where

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

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

we have $\begin{pmatrix} Y_1/8 \\ Y_2/8 \\ 0 \\ Y_3/8 \end{pmatrix} = \begin{pmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_2 \\ Y_3 \end{pmatrix}$

that gives

$$(\boldsymbol{X}'\boldsymbol{X})^* = \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & 1/8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 0 & 0 & 1/8 \end{pmatrix}$$

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

Therefore, $\mathrm{Cov}(\widehat{\boldsymbol{b}}) = \begin{pmatrix} 1/8 & 0 & 0 & 0 \\ 0 & 1/8 & 8 & 0 \\ 0 & 8 & 8 & 0 \\ 0 & 0 & 0 & 1/8 \end{pmatrix} \sigma^2$ and under normality assumption, the estimators are independent. Furthermore,

$$\mathrm{var}\big(q_1\hat{\mu} + q_2\hat{\alpha} + q_3\hat{\beta} + q_4\hat{\gamma}\big) = \mathrm{var}(q_1 Y_1/8 + q_2 Y_2/8 + q_4 Y_3/8) = \frac{1}{8}(q_1^2 + q_2^2 + q_4^2)\sigma^2.$$

For the null hypothesis $H_0{:}\,\gamma = 0$, it follows that $\boldsymbol{q} = (0,0,0,1)'$ and

$$\mathrm{var}\ (\hat{\gamma}) = \frac{1}{8}\sigma^2$$

and (Recall $t = \frac{\boldsymbol{t}'\boldsymbol{y} - \boldsymbol{q}'\boldsymbol{b}}{\sqrt{\boldsymbol{t}'\boldsymbol{t}\sigma^2}} / \sqrt{MSE/\sigma^2}$)

$$t_5 = \hat{\gamma} \Big/ \sqrt{\frac{1}{8} MSE}.$$

Also, to construct A $(1-\epsilon)$ confidence interval for $\alpha + \beta - \gamma$, we have

$$\hat{\alpha} + \hat{\beta} - \hat{\gamma} = \frac{1}{4}(y_5 + y_6 - y_3 - y_4)$$

and

$$\mathrm{var}\ (\hat{\alpha} + \hat{\beta} - \hat{\gamma}) = \frac{1}{4}\sigma^2$$

Noting that $q = (0,1,1,-1)'$, a $(1-\epsilon)$ confidence interval for $\alpha + \beta - \gamma$ is

$$\hat{\alpha} + \hat{\beta} - \hat{\gamma} \pm \frac{1}{2} t_{\epsilon/2}(5)\sqrt{MSE}.$$

2. Suppose that $y_{11}, y_{12}, \cdots, y_{1n_1}$ is a random sample from $N(\mu_1, \sigma^2)$ distribution and that $y_{21}, y_{22}, \cdots, y_{2n_2}$ is an independent sample from a $N(\mu_2, \sigma^2)$ distribution. The linear model can be expressed as:

$$E(\boldsymbol{y}) = E\begin{pmatrix} y_{11} \\ y_{12} \\ \cdots \\ y_{1n_1} \\ y_{21} \\ y_{22} \\ \cdots \\ y_{2n_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ \cdots & \cdots \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ \cdots & \cdots \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

It follows that

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix}$$

$$\hat{\mu}_i = \bar{y}_i$$

$$SSE = \sum_{i=1}^{2} \sum_{j=1}^{n_i} y_{ij}^2 - n_1 \bar{y}_1^2 - n_2 \bar{y}_2^2 = \sum_{i=1}^{2} \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_i\right)^2$$

and

$$MSE = SSE/(n_1 + n_2 - 2).$$

Since

$$\text{var}\,(\bar{y}_1 - \bar{y}_2) = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\sigma^2$$

It follows that the t-statistic for the hypothesis:

$$H_0: \mu_1 = \mu_2$$

is given by

$$t_{n_1+n_2-2} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) MSE}}.$$

The Generalized $\boldsymbol{F}$-Test

Suppose that the elements of $\boldsymbol{Q}'\boldsymbol{b}$, are s linearly independent estimable functions with best estimators the elements of $\boldsymbol{T}'\boldsymbol{y}$.

Then $\boldsymbol{T}'\boldsymbol{y} \sim N(\boldsymbol{Q}'\boldsymbol{b}, \sigma^2\boldsymbol{T}'\boldsymbol{T})$, independent of $\frac{SSE}{\sigma^2}$, which has the $\chi^2(n-r)$ distribution. Therefore

$$\boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}/\sigma^2$$

has the $\chi^2(s)$ distribution under the hypothesis $\boldsymbol{Q}'\boldsymbol{b} = \boldsymbol{0}$. Suppose that V_s and L_s are the vector space and linear set generated by the columns of $\boldsymbol{T}$ and the elements of $\boldsymbol{T}'\boldsymbol{y}$. Then

$$SSH = \boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}$$

is the sum of squares for L_s.

Under the hypothesis:

$$H_0: \boldsymbol{Q}'\boldsymbol{b} = \boldsymbol{0}$$

it follows that

$$F = \frac{SSH}{SSE}\frac{n-r}{s}$$

has an F distribution with s and $n-r$ degrees of freedom.

$$F \sim F(s, n-r).$$

Note that $V_s \subset V_r, L_s \subset L_r$ and that SSH is the square of the length of the projection of $\boldsymbol{y}$ on V_s.

The Principle of Conditional Error

Instead of calculating the hypothesis sum of squares

$$SSH = \boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}$$

directly, it can be determined more conveniently by using the conditional error sum of squares. The model under the hypothesis is

$$E(\boldsymbol{y}) = \boldsymbol{X_0 b_0}$$

with the new estimation space V_{r-s} such that

$$V_r = V_{r-s} \oplus V_s$$

with V_{r-s} and V_s mutually orthogonal. A necessary and sufficient set of conditions are:

1. $E(\boldsymbol{T}'\boldsymbol{y}) = 0$ under this model,
2. $V_{r-s} \subset V_r$ and
3. $\text{rank}\,(\boldsymbol{X_0}) = r - s$.

This implies that the new error space (the conditional error space)

$$V_c = V_s \oplus V_e$$

and the new error set (the conditional error set)

$$L_c = L_s \oplus L_e.$$

The error sum of squares under the new model (the conditional error sum of squares) is

$$SSC = SSE + SSH$$

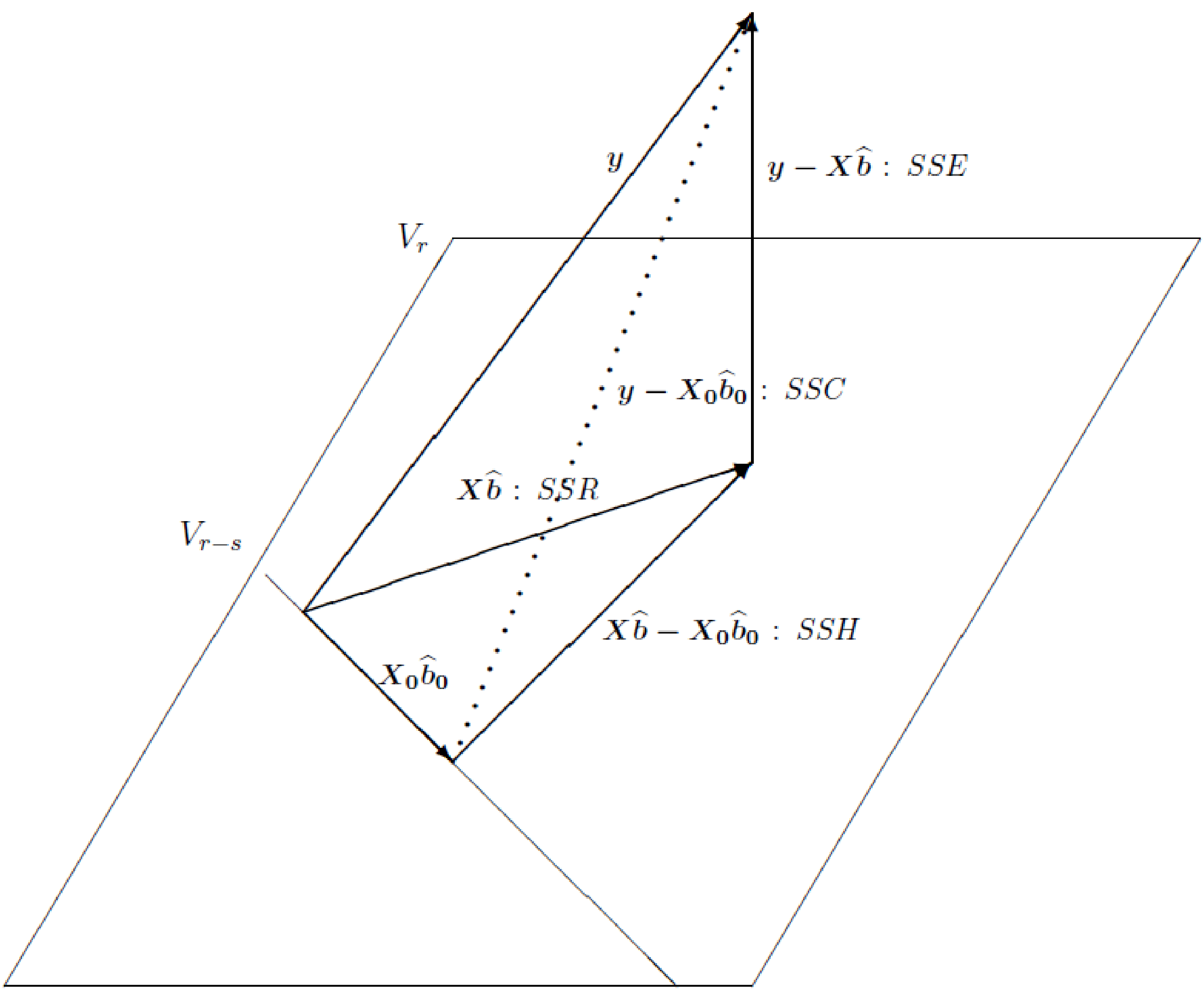
and the hypothesis sum of squares:

$$SSH = SSC - SSE.$$

Also:

$$L_r = L_s \oplus L_{r-s}.$$
$$SSR = SSH + SS_{r-s}.$$

Schematic Representation



Examples

1. The random vector $\boldsymbol{y}$ has the $N(\boldsymbol{\eta}, \sigma^2\boldsymbol{I})$ distribution with

$$\boldsymbol{\eta} = \begin{pmatrix} \alpha+\beta \\ \alpha-\beta \\ \alpha+\gamma \\ \alpha-\gamma \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

The parameters are estimable with best estimators:

$$\hat{\alpha} = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$
$$\hat{\beta} = \frac{1}{2}(y_1 - y_2)$$
$$\hat{\gamma} = \frac{1}{2}(y_3 - y_4).$$

The error sum of squares is $SSE = \boldsymbol{y}'\boldsymbol{y} - 4\hat{\alpha}^2 - 2\hat{\beta}^2 - 2\hat{\gamma}^2$.

Consider the hypothesis $H_0\colon \beta = \gamma = 0$. The hypothesis can be expressed as $H_0\colon \boldsymbol{Q}'\boldsymbol{b} = \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = 0$.

With $\boldsymbol{T}'\boldsymbol{y} = \begin{pmatrix} (y_1-y_2)/2 \\ (y_3-y_4)/2 \end{pmatrix}$. Since the elements of $\boldsymbol{T}'\boldsymbol{y}$ are mutually orthogonal, it follows that

$$SSH = (y_1 - y_2)^2/2 + (y_3 - y_4)^2/2$$

and the F-test for the hypothesis is

$$F = \frac{SSH}{SSE}\frac{1}{2}.$$

By using the principle of conditional error, it follows that under the hypothesis

$$E(\boldsymbol{y}) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\alpha$$

and

$$SSC = \boldsymbol{y}'\boldsymbol{y} - (y_1 + y_2 + y_3 + y_4)^2/4.$$

The hypothesis sum of squares is therefore

$$SSH = SSC - SSE = 2\hat{\beta}^2 + 2\hat{\gamma}^2.$$

2. Suppose

$$\begin{pmatrix} \mu+\alpha+\beta+\gamma \\ \mu+\alpha+\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu-\alpha-\beta+\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu+\alpha+\beta-\gamma \\ \mu-\alpha-\beta-\gamma \\ \mu-\alpha-\beta-\gamma \end{pmatrix} = \begin{pmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & +1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & -1 & -1 & +1 \\ 1 & +1 & +1 & -1 \\ 1 & +1 & +1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha \\ \beta \\ \gamma \end{pmatrix}$$

and $cov\ (\boldsymbol{y}) = \sigma^2 \boldsymbol{I}$. Further, suppose that the random vector $\boldsymbol{y}$ has a normal distribution and consider the hypothesis

$$H_0: \alpha+\beta = 0, \quad \gamma = 0.$$

In this case

$$\boldsymbol{Q}'\boldsymbol{b} = \begin{pmatrix} \alpha+\beta \\ \gamma \end{pmatrix}$$

and

$$\boldsymbol{T}'\boldsymbol{y} = \begin{pmatrix} Y_2/8 \\ Y_3/8 \end{pmatrix} = \begin{pmatrix} (y_1+y_2-y_3-y_4+y_5+y_6-y_7-y_8)/8 \\ (y_1+y_2+y_3+y_4-y_5-y_6-y_7-y_8)/8 \end{pmatrix}.$$

Since Y_2 and Y_3 are mutually orthogonal, it follows that SSH is the total of the individual sums of squares of Y_2 and Y_3, namely

$$SSH = Y_2^2/8 + Y_3^2/8.$$

If the principle of conditional error is used, it follows that under the null hypothesis

$$E(\boldsymbol{y}) = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \mu.$$

Therefore

$$SSC = \boldsymbol{y}'\boldsymbol{y} - \frac{1}{8} Y_1^2$$

and

$$SSE = \boldsymbol{y}'\boldsymbol{y} - \frac{1}{8}Y_1^2 - \frac{1}{8}Y_2^2 - \frac{1}{8}Y_3^2$$

such that

$$SSH = SSC - SSE = \frac{1}{8}Y_2^2 + \frac{1}{8}Y_3^2.$$

The F-test for the hypothesis is

$$F = \frac{SSH}{SSE}\frac{5}{2}.$$

In the case of the hypothesis

$$H_0: \mu = \alpha + \beta = \gamma.$$

it follows, however, that the hypothesis can be given as

$$H_0: \boldsymbol{Q}'\boldsymbol{b} = \begin{pmatrix} \mu - \alpha - \beta \\ \mu - \gamma \end{pmatrix} = 0.$$

In this case

$$\boldsymbol{T}'\boldsymbol{y} = \begin{pmatrix} Y_1/8 - Y_2/8 \\ Y_1/8 - Y_3/8 \end{pmatrix} = \begin{pmatrix} (y_3 + y_4 + y_7 + y_8)/4 \\ (y_5 + y_6 + y_7 + y_8)/4 \end{pmatrix}.$$

Again SSH can be directly determined from the expression

$$SSH = \boldsymbol{y}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{y}$$

It, however, follows easier that the elements of

$$\begin{pmatrix} (y_3 + y_4 + y_5 + y_6 + 2y_7 + 2y_8)/4 \\ (y_3 + y_4 - y_5 - y_6)/4 \end{pmatrix}$$

form a mutually orthogonal basis for L_s. The hypothesis sum of squares is:

$$SSH = (y_3 + y_4 + y_5 + y_6 + 2y_7 + 2y_8)^2/12 + (y_3 + y_4 - y_5 - y_6)^2/4.$$

By using the principle of conditional error, the model under the hypothesis is

$$E(\boldsymbol{y}) = \begin{pmatrix} 3 \\ 3 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \mu.$$

Under this model

$$\hat{\mu} = (3y_1 + 3y_2 + y_3 + y_4 + y_5 + y_6 - y_7 - y_8)/24 = Y_0/24, \text{ say}.$$

Therefore

$$SSC = \boldsymbol{y}'\boldsymbol{y} - Y_0^2/24$$

and the hypothesis sum of squares is

$$SSH = SSC - SSE = Y_1^2/8 + Y_2^2/8 + Y_3^2/8 - Y_0^2/24.$$

It must be shown algebraically that the two expressions for SSH are the same.

The F-test for the hypothesis is

$$F = \frac{SSH}{SSE}\frac{5}{2}.$$

- ## Multiple Correlation

Consider, under the assumption: $\mathbf{1} \in V_r$, with $\mathbf{1}' = (1,1,\cdots,1)$, the null hypothesis

$$H_0: E(\boldsymbol{y}) = \mathbf{1}\mu\,.$$

It implies that the expected values of all the elements of $\boldsymbol{y}$ are the same. The conditional error sum of squares is

$$SSC = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2.$$

A measure of how well the model fits, is the ratio of the hypothesis sum of squares and the conditional error sum of squares:

$$R^2 = \frac{SSC - SSE}{SSC}\,,$$

which is the coefficient of determination. The positive square root, R, is the multiple correlation coefficient. The F-test for the hypothesis is:

$$F_{r-1,n-r} = \frac{R^2}{1-R^2}\frac{n-r}{r-1}\,.$$

Note that $0 \le R \le 1$. Furthermore

$$1 - R^2 = \frac{SSE}{SSC}$$

the ratio of the total variation of the observations which is not explained by the model, while R^2 is that ratio of the total variation of the observations which is explained by the model.

Since $\mathbf{1} \in V_r$, is $\mathbf{1}'\boldsymbol{P} = \mathbf{1}'$, with $\boldsymbol{P} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{X}'$. Then it follows for the estimated y-values, $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{b}}$, that

$$\mathbf{1}'\boldsymbol{X}\hat{\boldsymbol{b}} = \mathbf{1}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^*\boldsymbol{X}'\boldsymbol{y} = \mathbf{1}'\boldsymbol{y}$$

and

$$\sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i \text{ or } \bar{\hat{y}} = \bar{y}\,.$$

Consequently, the mean of the estimated y-values is the same as the mean y-value (provided that $\mathbf{1} \in V_r$).

Then it follows that:

$$
\begin{aligned}
SSH = \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{b}} - n\bar{\hat{y}}^2 &= \widehat{\boldsymbol{b}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{y}^2 \\
&= \sum_{i=1}^{n} \hat{y}_i^2 - n\bar{\hat{y}}^2 = \sum_{i=1}^{n} \hat{y}_i y_i - n\bar{y}\bar{\hat{y}} \\
&= \sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)^2 = \sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)(y_i - \bar{y}).
\end{aligned}
$$

Consequently, the coefficient of determination is

$$
\begin{aligned}
R^2 &= \frac{SSH}{SSC} \\
&= \frac{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)(y_i - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \\
&= \frac{\left\{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)(y_i - \bar{y})\right\}^2}{\left\{\sum_{i=1}^{n} (y_i - \bar{y})^2\right\}\left\{\sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}}\right)^2\right\}}
\end{aligned}
$$

i.e. the square of the correlation coefficient between the y-values and the estimated y-values.