Minimization by use of differentiation

Show that $T^2 = \min_{\underset{\sim}{\mu} = K\underset{\sim}{\beta}} n(\underset{\sim}{\bar{x}} - \underset{\sim}{\mu})^T S^{-1} (\underset{\sim}{\bar{x}} - \underset{\sim}{\mu})$

$$
\begin{aligned}
R &= (\underset{\sim}{\bar{x}} - K\underset{\sim}{\beta})^T S^{-1} (\underset{\sim}{\bar{x}} - K\underset{\sim}{\beta}) \\
&= \underset{\sim}{\bar{x}}^T S^{-1} \underset{\sim}{\bar{x}} - \underset{\sim}{\bar{x}}^T S^{-1} K \underset{\sim}{\beta} - \underset{\sim}{\beta}^T K^T S^{-1} \underset{\sim}{\bar{x}} + \underset{\sim}{\beta}^T K^T S^{-1} K \underset{\sim}{\beta} \\
&= \underset{\sim}{\bar{x}}^T S^{-1} \underset{\sim}{\bar{x}} - 2 \underset{\sim}{\beta}^T K^T S^{-1} \underset{\sim}{\bar{x}} + \underset{\sim}{\beta}^T K^T S^{-1} K \underset{\sim}{\beta}
\end{aligned}
$$

Differentiation w r t $\underset{\sim}{\beta}$ leads to

$$
\begin{aligned}
\frac{dR}{d\underset{\sim}{\beta}} &= \underset{\sim}{0} - 2K^T S^{-1} \underset{\sim}{\bar{x}} + K^T S^{-1} K \underset{\sim}{\beta} + (\underset{\sim}{\beta}^T K^T S^{-1} K)^T \\
&= -2K^T S^{-1} \underset{\sim}{\bar{x}} + 2K^T S^{-1} K \underset{\sim}{\beta} = \underset{\sim}{0} \text{ for}
\end{aligned}
$$

$$\underset{\sim}{\beta} = \underset{\sim}{\beta}^* = (K^T S^{-1} K)^{-1} K^T S^{-1} \underset{\sim}{\bar{x}}$$

$$
\begin{aligned}
T^2 = n R_{\min} &= n\big(\underset{\sim}{\bar{x}}^T S^{-1} \underset{\sim}{\bar{x}} - 2 \underset{\sim}{\bar{x}}^T S^{-1} K (K^T S^{-1} K)^{-1} K^T S^{-1} \underset{\sim}{\bar{x}} \\
&\quad + \underset{\sim}{\bar{x}}^T S^{-1} K (K^T S^{-1} K)^{-1} K^T S^{-1} K (K^T S^{-1} K)^{-1} K^T S^{-1} \underset{\sim}{\bar{x}} \big) \\
&= n\big(\underset{\sim}{\bar{x}}^T S^{-1} \underset{\sim}{\bar{x}} - \underset{\sim}{\bar{x}}^T S^{-1} K (K^T S^{-1} K)^{-1} K^T S^{-1} \underset{\sim}{\bar{x}} \big) \\
&= n\big(\underset{\sim}{\bar{x}}^T S^{-1} \underset{\sim}{\bar{x}} - \underset{\sim}{\bar{x}}^T S^{-1} K \underset{\sim}{\beta}^* \big)
\end{aligned}
$$

$$T^2 \sim T^2(d-k, n-1) = \frac{(n-1)(d-k)}{n-d+k} F(d-k, n-d+k),$$

where k = rank K

Ex. **EXAMPLE 2.4** The height, in millimeters, of the ramus bone in the jaws of 20 boys was measured at ages 8, $8\frac{1}{2}$, 9, and $9\frac{1}{2}$, and the results from Elston and Grizzle [1962] are given in Table 3.4. A main objective of the study was to establish a standard growth curve for the use of orthodontists, and it is clear from the sample means that a straight line should provide a satisfactory fit for the age range considered. Following Grizzle and Allen [1969], we can use orthogonal polynomials and consider the model

$$\mu_j = \beta_0 + 2\beta_1(t_j - \bar{t}) \qquad (j = 1,2,3,4),$$

where age 8 is chosen as the time origin so that $t_j = j - 1$. Then $\boldsymbol{\mu} = \mathbf{K}\boldsymbol{\beta}$, where

$$\mathbf{K}\boldsymbol{\beta} = \begin{pmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},$$

MA9 9.3.2000

TABLE 3.4 Ramus Height of 20 Boys[a]

	Age in Years			
Individual	8	$8\frac{1}{2}$	9	$9\frac{1}{2}$
1	47.8	48.8	49.0	49.7
2	46.4	47.3	47.7	48.4
3	46.3	46.8	47.8	48.5
4	45.1	45.3	46.1	47.2
5	47.6	48.5	48.9	49.3
6	52.5	53.2	53.3	53.7
7	51.2	53.0	54.3	54.5
8	49.8	50.0	50.3	52.7
9	48.1	50.8	52.3	54.4
10	45.0	47.0	47.3	48.3
11	51.2	51.4	51.6	51.9
12	48.5	49.2	53.0	55.5
13	52.1	52.8	53.7	55.0
14	48.2	48.9	49.3	49.8
15	49.6	50.4	51.2	51.8
16	50.7	51.7	52.7	53.3
17	47.2	47.7	48.4	49.5
18	53.3	54.6	55.1	55.3
19	46.2	47.5	48.1	48.4
20	46.3	47.6	51.3	51.8
Mean	48.655	49.625	50.570	51.450
S. D.	2.52	2.54	2.63	2.73

[a]Reproduced from R. C. Elston and J. E. Grizzle [1962]. "Estimation of time-response curves and their confidence bands." *Biometrics*, **18**, 148–159, Table 2. With permission from The Biometric Society.

and **K** has orthogonal columns. Now

$$\bar{\mathbf{x}}' = (48.655,\ \ 49.625,\ \ 50.570,\ \ 51.450),$$

$$\mathbf{S} = \begin{pmatrix} 6.330 & 6.189 & 5.777 & 5.548 \\ - & 6.449 & 6.153 & 5.923 \\ - & - & 6.918 & 6.946 \\ - & - & - & 7.465 \end{pmatrix},$$

and

$$\begin{aligned} \boldsymbol{\beta}^* &= (\mathbf{K}'\mathbf{S}^{-1}\mathbf{K})^{-1}\mathbf{K}'\mathbf{S}^{-1}\bar{\mathbf{x}} \\ &= \begin{pmatrix} 50.050 \\ 0.465 \end{pmatrix}. \end{aligned}$$

Hence

$$\begin{aligned} T_0^2 &= n(\bar{\mathbf{x}} - \mathbf{K}\boldsymbol{\beta}^*)'\mathbf{S}^{-1}(\bar{\mathbf{x}} - \mathbf{K}\boldsymbol{\beta}^*) \\ &= n(\bar{\mathbf{x}}'\mathbf{S}^{-1}\bar{\mathbf{x}} - \bar{\mathbf{x}}'\mathbf{S}^{-1}\mathbf{K}\boldsymbol{\beta}^*) \\ &= 0.20113, \end{aligned}$$

and we compare

$$F_0 = T_0^2 \frac{20 - 4 + 2}{19(4 - 2)} = 0.095$$

with $F_{2,18}^{\alpha}$. Since F_0 is so low, we do not reject the straight line model.

Confidence intervals

$\underset{\sim}{x}_i \sim N_d(\underset{\sim}{\mu}, \Sigma)$, $i = 1, \dots, n$ independent

$\underset{\sim}{a}^T \bar{\underset{\sim}{x}} \sim N(\underset{\sim}{a}^T \underset{\sim}{\mu}, \frac{\underset{\sim}{a}^T \Sigma \underset{\sim}{a}}{n})$

Confidence intervals for pre specified linear combinations of $\mu_1, \dots, \mu_d$, $\underset{\sim}{a}^T \underset{\sim}{\mu}$, with confidence level $1-\alpha$:

$$\underset{\sim}{a}^T \underset{\sim}{\mu} = \underset{\sim}{a}^T \bar{\underset{\sim}{x}} \pm t_{1-\frac{\alpha}{2}}(n-1) \sqrt{\frac{\underset{\sim}{a}^T S \underset{\sim}{a}}{n}}$$

Confidence intervals for r pre specified linear combinations $\underset{\sim}{a}_i^T \underset{\sim}{\mu}$, $i = 1, \dots, r$, each with confidence level $1 - \frac{\alpha}{r}$:

$$\underset{\sim}{a}_i^T \underset{\sim}{\mu} = \underset{\sim}{a}_i^T \bar{\underset{\sim}{x}} \pm t_{1-\frac{\alpha}{2r}}(n-1) \sqrt{\frac{\underset{\sim}{a}_i^T S \underset{\sim}{a}_i}{n}}, \quad i = 1, \dots, r$$

(Bonferroni intervals)

Combined confidence level for the r intervals:

$$\begin{aligned} 1-\alpha_0 &= P\left(\bigcap_i \left(|t_i(n-1)| < t_{1-\frac{\alpha}{2r}}(n-1)\right)\right) \\ &= 1 - P\left(\bigcup_i \left(|t_i(n-1)| > t_{1-\frac{\alpha}{2r}}(n-1)\right)\right) \\ &\geq 1 - \sum_i P\left(|t_i(n-1)| > t_{1-\frac{\alpha}{2r}}(n-1)\right) \\ &= 1 - r\frac{\alpha}{r} = 1-\alpha, \end{aligned}$$

thus $1-\alpha_0 \geq 1-\alpha$

Assume that $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_r$ are ordered, so $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_q$ are linear independent and that $\underset{\sim}{a}_{q+1}, \dots, \underset{\sim}{a}_r$ are all linear dependent of $\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_q$

Let $A = [\underset{\sim}{a}_1 \dots \underset{\sim}{a}_q]^T$, A $q \times d$, rank $A = q$ $(\leq d)$,

and $\underset{\sim}{\varphi} = A \underset{\sim}{\mu}$

$$n(A\bar{\underset{\sim}{x}} - A\underset{\sim}{\mu})^T (ASA^T)^{-1} (A\bar{\underset{\sim}{x}} - A\underset{\sim}{\mu}) \sim T^2(q, n-1)$$

$$\Leftrightarrow \quad n(\hat{\underset{\sim}{\varphi}} - \underset{\sim}{\varphi})^T (ASA^T)^{-1} (\hat{\underset{\sim}{\varphi}} - \underset{\sim}{\varphi}) \sim T^2(q, n-1)$$

$$1 - \alpha = P\left((\hat{\underset{\sim}{\varphi}} - \underset{\sim}{\varphi})^T (ASA^T)^{-1} (\hat{\underset{\sim}{\varphi}} - \underset{\sim}{\varphi}) < T^2_{1-\alpha}(q, n-1) \frac{1}{n} \right)$$

$$= P\left(\sup_{\underset{\sim}{h} \neq \underset{\sim}{0}} \frac{(\underset{\sim}{h}^T(\hat{\underset{\sim}{\varphi}} - \underset{\sim}{\varphi}))^2}{\underset{\sim}{h}^T ASA^T \underset{\sim}{h}} < T^2_{1-\alpha}(q, n-1) \frac{1}{n} \right) \quad \text{cf. A7.6}$$

$$= P\left(\forall \underset{\sim}{h} \neq \underset{\sim}{0} : (\underset{\sim}{h}^T(\hat{\underset{\sim}{\varphi}} - \underset{\sim}{\varphi}))^2 < T^2_{1-\alpha}(q, n-1) \frac{\underset{\sim}{h}^T ASA^T \underset{\sim}{h}}{n} \right)$$

$$= P\left(\forall \underset{\sim}{h} \neq \underset{\sim}{0} : |\underset{\sim}{h}^T \hat{\underset{\sim}{\varphi}} - \underset{\sim}{h}^T \underset{\sim}{\varphi}| < \sqrt{T^2_{1-\alpha}(q, n-1) \frac{\underset{\sim}{h}^T ASA^T \underset{\sim}{h}}{n}} \right)$$

Hence confidence intervals with confidence level $1-\alpha$ for any $\underset{\sim}{h}^T \underset{\sim}{\varphi}$:

$$\underset{\sim}{h}^T \underset{\sim}{\varphi} = \underset{\sim}{h}^T \hat{\underset{\sim}{\varphi}} \pm \sqrt{T^2_{1-\alpha}(q, n-1) \frac{\underset{\sim}{h}^T ASA^T \underset{\sim}{h}}{n}} \quad \text{(Scheffé intervals)}$$

or – with $\underset{\sim}{h}^T \underset{\sim}{\varphi} = \underset{\sim}{h}^T A \underset{\sim}{\mu}$ –

$$\underset{\sim}{h}^T A \underset{\sim}{\mu} = \underset{\sim}{h}^T A \bar{\underset{\sim}{x}} \pm \sqrt{T^2_{1-\alpha}(q, n-1) \frac{\underset{\sim}{h}^T ASA^T \underset{\sim}{h}}{n}}$$

Any of the linear combinations $\underset{\sim}{a}_i^T \underset{\sim}{\mu}$, $i = 1, \dots, r$ arise from a suitable choice of $\underset{\sim}{h}$, as

$$\underset{\sim}{h}^T \underset{\sim}{\varphi} = (h_1 \underset{\sim}{a}_1^T + \dots + h_q \underset{\sim}{a}_q^T) \underset{\sim}{\mu}$$

For the combined confidence level $1-\alpha_0$ of the intervals

$$\underset{\sim}{a}_i^T \underset{\sim}{\mu} = \underset{\sim}{a}_i^T \bar{\underset{\sim}{x}} \pm \sqrt{T^2_{1-\alpha}(q, n-1) \frac{\underset{\sim}{a}_i^T S \underset{\sim}{a}_i}{n}}, \quad i = 1, \dots, r$$

we have $1 - \alpha_0 \geq 1 - \alpha$, as the r intervals constitute a subset of all the covered intervals

Bonferroni intervals are in general shorter than Scheffé intervals as

$$\left(t_{1-\frac{\alpha}{2q}}(n-1)\right)^2 = F_{1-\frac{\alpha}{q}}(1, n-1) < q\, F_{1-\alpha}(q, n-1) \quad \text{(empirical)}$$

$$\leq \frac{(n-1)q}{n-q} F_{1-\alpha}(q, n-1) = T^2_{1-\alpha}(q, n-1)$$

Confidence interval for any not pre specified $\underset{\sim}{h}^T \underset{\sim}{\mu}$ can be achieved by choosing $A = I_d$, hence $q = d$ and $\underset{\sim}{\varphi} = \underset{\sim}{\mu}$:

$$\underset{\sim}{h}^T \underset{\sim}{\mu} = \underset{\sim}{h}^T \bar{\underset{\sim}{x}} \pm \sqrt{T^2_{1-\alpha}(d, n-1) \frac{\underset{\sim}{h}^T S \underset{\sim}{h}}{n}} \quad , \quad \text{confidence level } 1-\alpha$$

Confidence intervals for contrasts can be achieved by choosing $A = FC_2$, where C_2 origins from section 3.4.2 and F is $(d-1) \times (d-1)$ non-singular.

Confidence interval for any contrast $\underset{\sim}{c}^T \underset{\sim}{\mu}$, $\underset{\sim}{c}^T = \underset{\sim}{h}^T A = \underset{\sim}{h}^T F C_2$, with confidence level $1-\alpha$:

$$\underset{\sim}{c}^T \underset{\sim}{\mu} = \underset{\sim}{c}^T \bar{\underset{\sim}{x}} \pm \sqrt{T^2_{1-\alpha}(d-1, n-1) \frac{\underset{\sim}{c}^T S \underset{\sim}{c}}{n}} \quad , \quad \underset{\sim}{c}^T \underset{\sim}{1}_d = 0$$

Bonferroni interval for one among r specified contrasts :

$$\underset{\sim}{c}^T \underset{\sim}{\mu} = \underset{\sim}{c}^T \bar{\underset{\sim}{x}} \pm t_{1-\frac{\alpha}{2r}}(n-1) \sqrt{\frac{\underset{\sim}{c}^T S \underset{\sim}{c}}{n}} \quad , \quad \underset{\sim}{c}^T \underset{\sim}{1}_d = 0 \ ,$$

confidence level $1 - \frac{\alpha}{r}$

EXAMPLE 3.5 Consider Example 3.2 (Section 3.4.2) and the cork data of Table 3.3. If we are interested in a confidence interval for each μ_j, then we can use (3.29) with $\mathbf{a}_1' = (1, 0, \ldots, 0)$, and so on, namely,

$$\bar{x}_j \pm k_\alpha \left(\frac{s_{jj}}{n} \right)^{1/2} \qquad (j = 1, 2, 3, 4).$$

With $n = 28$, $r = q = d = 4$, and $\alpha = 0.05$, the two methods give the following values of k_α: From Appendix D1, $t_{27}^{0.05/8} = 2.676$ [see (3.28)] for the Bonferroni method and $(T^2_{4,27,0.05})^{1/2} = \{[4(27)/24] F^{0.05}_{4,24}\}^{1/2} = 3.537$ [see (3.33) and (3.30)] for the Scheffé method. We require the elements of $\bar{\mathbf{x}}$ and the diagonal elements s_{jj} of $\mathbf{S}$, where

$$\bar{\mathbf{x}} = \begin{pmatrix} 50.536 \\ 46.179 \\ 49.679 \\ 45.179 \end{pmatrix} \begin{matrix} \mathrm{N} \\ \mathrm{E} \\ \mathrm{S} \\ \mathrm{W} \end{matrix} \quad \text{and} \quad \mathbf{S} = \begin{pmatrix} 290.406 & 223.753 & 288.438 & 226.271 \\ - & 219.930 & 229.060 & 171.374 \\ - & - & 350.004 & 259.541 \\ - & - & - & 226.004 \end{pmatrix}.$$

Using the Bonferroni intervals, we obtain the following:

$$\mu_1 : 50.54 \pm 8.62,$$

$$\mu_2 : 46.18 \pm 7.50,$$

$$\mu_3 : 49.68 \pm 9.46,$$

$$\mu_4 : 45.18 \pm 7.60.$$

Since the test for $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ was significant at the 1% level, we can see which contrasts $\mathbf{c}'\bar{\mathbf{x}}$ are responsible for the rejection of H_0. In Example 3.3 (Section 3.4.2) the contrast $\mu_1 - \mu_2 + \mu_3 - \mu_4$ (= N + S − E − W) with $\mathbf{c}' = (1, -1, 1, -1)$ was regarded as a likely candidate. Using (3.35), a 99% confidence interval for this contrast is

$$8.86 \pm \left\{ \frac{3(27)}{25} F^{0.01}_{3,25} \right\}^{1/2} \{128.718/28\}^{1/2},$$

that is, 8.86 ± 8.35 or $[0.51, 17.21]$, which does not contain zero. Corresponding intervals for pairwise contrasts are

$$\mu_1 - \mu_2 : 4.36 \pm 5.83,$$

$$\mu_1 - \mu_3 : 0.86 \pm 5.86,$$

$$\mu_1 - \mu_4 : 5.36 \pm 5.88,$$

$$\mu_3 - \mu_2 : 3.50 \pm 7.78,$$

$$\mu_2 - \mu_4 : 1.00 \pm 7.47,$$

$$\mu_3 - \mu_4 : 4.50 \pm 5.55,$$

and all of these contain zero. Since H_0 is rejected if and only if at least one interval for a contrast does not contain zero, we see that the contrast (N + S − E − W) is responsible for the rejection of H_0.

Test for blokwise independence / two blocks

$\underset{\sim}{x}_i \sim N_d(\underset{\sim}{\mu}, \Sigma)$, $i = 1, \dots, n$, independent, $n-1 \geq d$

Partition: $\underset{\sim}{x}_i = \begin{bmatrix} \underset{\sim}{y}_i \\ \underset{\sim}{z}_i \end{bmatrix}$ $\underset{\sim}{\mu} = \begin{bmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{bmatrix}$ $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

$\underset{\sim}{y}_i$ d_1-dim., $\underset{\sim}{z}_i$ d_2-dim., $d_1 + d_2 = d$, assume $d_1 \geq d_2$

H_0: $\underset{\sim}{y}_i$ and $\underset{\sim}{z}_i$ independent $\Leftrightarrow$ $\Sigma_{12} = 0$ cf. th. 2.1 (iv)

The likelihood ratio:

$$\ell = \frac{\sup_{\underset{\sim}{\mu}_1, \underset{\sim}{\mu}_2, \Sigma_{11}, \Sigma_{22}} L_1(\underset{\sim}{\mu}_1, \Sigma_{11}) L_2(\underset{\sim}{\mu}_2, \Sigma_{22})}{\sup_{\underset{\sim}{\mu}, \Sigma} L(\underset{\sim}{\mu}, \Sigma)}$$

$$= \frac{L_1(\hat{\mu}_1, \hat{\Sigma}_{11}) L_2(\hat{\mu}_2, \hat{\Sigma}_{22})}{L(\hat{\mu}, \hat{\Sigma})}$$

$$= \frac{(2\pi e)^{-\frac{nd_1}{2}} (\det \hat{\Sigma}_{11})^{-\frac{n}{2}} (2\pi e)^{-\frac{nd_2}{2}} (\det \hat{\Sigma}_{22})^{-\frac{n}{2}}}{(2\pi e)^{-\frac{nd}{2}} (\det \hat{\Sigma})^{-\frac{n}{2}}}$$

$$= \left(\frac{\det \hat{\Sigma}_{11} \det \hat{\Sigma}_{22}}{\det \hat{\Sigma}} \right)^{-\frac{n}{2}}$$

$$\ell^{\frac{2}{n}} = \frac{\det \hat{\Sigma}}{\det \hat{\Sigma}_{11} \det \hat{\Sigma}_{22}} = \frac{n^{d_1} n^{d_2} \det Q}{n^d \det Q_{11} \det Q_{22}} = \frac{\det Q}{\det Q_{11} \det Q_{22}}$$

$$= \frac{\det Q_{11} \det (Q_{22} - Q_{21} Q_{11}^{-1} Q_{12})}{\det Q_{11} \det Q_{22}} = \frac{\det Q_{22\cdot 1}}{\det Q_{22}}$$

$$= \frac{\det E}{\det (E + H)}$$, where $H = Q_{21} Q_{11}^{-1} Q_{12}$

$E = Q_{22\cdot 1} = Q_{22} - H$

$H \sim W_{d_2}(d_1, \Sigma_{22})$, H_0 true
$E \sim W_{d_2}(n-d_1-1, \Sigma_{22})$, H_0 true
} indep. cf. corollary to lemma 2.10

$$l^{\frac{2}{n}} = \Lambda = \frac{\det E}{\det(E+H)} \sim U(d_2, d_1, n-d_1-1), \; H_0 \text{ true},$$

cf. def. of U distribution, p. 40

$U(d_2, d_1, n-d_1-1) = U(d_1, d_2, n-d_2-1)$, thus interchange of y_i and $\underset{\sim}{z_i}$ leads to the same test statistic

H_0 is rejected for small values of l , equivalent are small values of Λ

$\Lambda = \prod_{j=1}^{d_2} (1-\theta_j)$ cf for instance formula (2.48)

The θ_j's are the ordered eigenvalues of

$$H(E+H)^{-1} = Q_{21} Q_{11}^{-1} Q_{12} Q_{22}^{-1} = S_{21} S_{11}^{-1} S_{12} S_{22}^{-1}$$

Use of the U-∩-principle:

H_{0ab} : $\underset{\sim}{a}^T \Sigma_{12} \underset{\sim}{b} = 0$ equiv. with $\dfrac{\underset{\sim}{a}^T \Sigma_{12} \underset{\sim}{b}}{\sqrt{\underset{\sim}{a}^T \Sigma_{11} \underset{\sim}{a}} \sqrt{\underset{\sim}{b}^T \Sigma_{22} \underset{\sim}{b}}} = 0$

H_{0ab} : $\sigma_{ab} = 0$ equivalent to $\rho_{ab} = 0$

ρ_{ab} is the correllation koefficient between $\underset{\sim}{a}^T \underset{\sim}{y}$ and $\underset{\sim}{b}^T \underset{\sim}{z}$

The empirical correllation coefficient r_{ab} :

$$r_{ab} = \frac{s_{ab}}{s_a s_b} = \frac{\underset{\sim}{a}^T S_{12} \underset{\sim}{b}}{\sqrt{\underset{\sim}{a}^T S_{11} \underset{\sim}{a}} \sqrt{\underset{\sim}{b}^T S_{22} \underset{\sim}{b}}} \quad , \quad \text{as}$$

$$s_a^2 = \frac{1}{n-1} \sum_i (\underset{\sim}{a}^T y_i - \underset{\sim}{a}^T \bar{y})^2$$

$$= \frac{1}{n-1} \sum_i (\underset{\sim}{a}^T y_i - \underset{\sim}{a}^T \bar{y})(\underset{\sim}{a}^T y_i - \underset{\sim}{a}^T \bar{y})^T$$

$$= \frac{1}{n-1} \underset{\sim}{a}^T \left(\sum_i (y_i - \bar{y})(y_i - \bar{y})^T \right) \underset{\sim}{a}$$

$$= \frac{1}{n-1} \underset{\sim}{a}^T Q_{11} \underset{\sim}{a} = \underset{\sim}{a}^T S_{11} \underset{\sim}{a}$$

analogous s_b^2 and s_{ab}

Acceptance area for H_{0ab} : $r_{ab}^2 \leq k \Leftrightarrow \frac{(\underset{\sim}{a}^T S_{12} \underset{\sim}{b})^2}{\underset{\sim}{a}^T S_{11} \underset{\sim}{a} \; \underset{\sim}{b}^T S_{22} \underset{\sim}{b}} \leq k$

$$\bigcap_a \bigcap_b \{(Y,Z) : r_{ab}^2 \leq k\} = \{(Y,Z) : \sup_{a,b} r_{ab}^2 \leq k\}$$

$$= \{(Y,Z) : \theta_{max} \leq k\},$$

where θ_{max} is the largest eigenvalue of $S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}$ or of $S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}$, cf A 7.7 (and A 1.4), or of $S_{21} S_{11}^{-1} S_{12} S_{22}^{-1}$ cf. A 1.4 $\Leftrightarrow$ θ_{max} is the largest root in

$$\det (S_{21} S_{11}^{-1} S_{12} S_{22}^{-1} - \theta I_{d_2}) = 0 \quad 1)$$

The test is called Roy's maximum root test

Pillai's test variable:

$$V^{(s)} = \sum_j \theta_j = tr\left(H(E+H)^{-1}\right) = tr\left(S_{21} S_{11}^{-1} S_{12} S_{22}^{-1}\right) \quad 2)$$

1) percentiles for θ_{max} in table D 14

2) — — $V^{(s)}$ — — D 16