Generalised linear hypothesis

The usual model $Y = XB + \mathcal{U}$, X $n \times p$, rank $X = p$

$\underset{\sim}{u}_i \sim N_d(\underset{\sim}{0}, \Sigma)$, $i = 1, \dots, n$, independent

$H_0: ABD = 0$, A $q \times p$ rank $A = q$, $q \le p$

$\Leftrightarrow A\Lambda = 0$, D $d \times v$ rank $D = v$, $v \le d$

$$Y = XB + \mathcal{U} \Rightarrow YD = XBD + \mathcal{U}D \Leftrightarrow Y_D = X\Lambda + \mathcal{U}_D$$

$$\underset{\sim}{u}_{Di}^T = \underset{\sim}{u}_i^T D \Rightarrow \underset{\sim}{u}_{Di} = D^T \underset{\sim}{u}_i \sim N_v(\underset{\sim}{0}, D^T\Sigma D)$$

$H = (A\hat{\Lambda})^T (A(X^TX)^{-1}A^T)^{-1} A\hat{\Lambda}$ cf. "corollary" 1 to theorem 8.5

$$= (A(X^TX)^{-1}X^TY_D)^T (A(X^TX)^{-1}A^T)^{-1} A(X^TX)^{-1}X^TY_D$$

$$= (A(X^TX)^{-1}X^TYD)^T (A(X^TX)^{-1}A^T)^{-1} A(X^TX)^{-1}X^TYD$$

$$= (A\hat{B}D)^T (A(X^TX)^{-1}A^T)^{-1} A\hat{B}D$$

$\sim W_v(q, D^T\Sigma D)$ when H_0 is true cf. theorem 8.5

$E = Y_D^T(I_n - X(X^TX)^{-1}X^T)Y_D$ cf. formula (8.12)

$$= D^TY^T(I_n - X(X^TX)^{-1}X^T)YD$$

$\sim W_v(n-p, D^T\Sigma D)$ cf. theorem 8.5

E and H independent

For rank $X = r < p$:

$H = (A\hat{B}D)^T(A(X^TX)^- A^T)^{-1} A\hat{B}D \sim W_v(q, D^T\Sigma D)$ when H_0 is true

$E = D^TY(I_n - X(X^TX)^- X^T)YD \sim W_v(n-r, D^T\Sigma D)$

$(\hat{B} = (X^TX)^- X^TY)$

Profile analysis / K populations

(K = 2 is earlier dealt with, cf. lecture note MA 11)

Consider samples from K d-dimensional normal distributed populations with equal variance matrix:

$$\underset{\sim}{x}_i^k \sim N_d(\underset{\sim}{\mu}_k, \Sigma), \quad i = 1, \dots, n_k, \quad \sum_{k=1}^{K} n_k = n$$
$$k = 1, \dots, K$$

$$\Leftrightarrow \quad X = \begin{bmatrix} \underset{\sim}{1}_{n_1} & \underset{\sim}{0} & \dots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{1}_{n_2} & \dots & \underset{\sim}{0} \\ \vdots & \vdots & & \vdots \\ \underset{\sim}{0} & \underset{\sim}{0} & \dots & \underset{\sim}{1}_{n_K} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\mu}_1^T \\ \underset{\sim}{\mu}_2^T \\ \vdots \\ \underset{\sim}{\mu}_K^T \end{bmatrix} + U, \quad \underset{\sim}{u}_i \sim N_d(\underset{\sim}{0}, \Sigma),$$

$i = 1, \dots, n$, independent

H_{01}: parallel profiles

$$\Leftrightarrow \quad \mu_{k-1,j-1} - \mu_{k-1,j} = \mu_{k,j-1} - \mu_{k,j}, \quad j = 2, \dots, d, \quad k = 2, \dots, K$$

$$\Leftrightarrow \quad C_1 \underset{\sim}{\mu}_{k-1} = C_1 \underset{\sim}{\mu}_k, \quad C_1 = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix}$$

$k = 2, \dots, K$ $\qquad ((d-1) \times d)$

$$\Leftrightarrow \quad (\underset{\sim}{\mu}_{k-1}^T - \underset{\sim}{\mu}_k^T) C_1^T = \underset{\sim}{0}^T, \quad k = 2, \dots, K$$

$$\Leftrightarrow \quad A_1 \begin{bmatrix} \underset{\sim}{\mu}_1^T \\ \vdots \\ \underset{\sim}{\mu}_K^T \end{bmatrix} C_1^T = O, \quad A_1 \ (K-1) \times K \text{ of same structure as } C_1$$

H_{02}: The profiles have same level

$$\Leftrightarrow \frac{1}{d}(\mu_{k-1,1} + \dots + \mu_{k-1,d}) = \frac{1}{d}(\mu_{k,1} + \dots + \mu_{k,d}) \quad k = 2, \dots, K$$

$$\Leftrightarrow \underset{\sim}{1}_d^T \underset{\sim}{\mu}_{k-1} = \underset{\sim}{1}_d^T \underset{\sim}{\mu}_k \quad , \quad k = 2, \dots, K$$

$$\Leftrightarrow (\underset{\sim}{\mu}_{k-1}^T - \underset{\sim}{\mu}_k^T) \underset{\sim}{1}_d = \underset{\sim}{0} \quad , \quad k = 2, \dots, K$$

$$\Leftrightarrow A_1 \begin{bmatrix} \underset{\sim}{\mu}_1^T \\ \vdots \\ \underset{\sim}{\mu}_K^T \end{bmatrix} \underset{\sim}{1}_d = \underset{\sim}{0}$$

H_{03}: no "main effect" from the d variables

$$\Leftrightarrow \frac{1}{K}(\mu_{1,j-1} + \dots + \mu_{K,j-1}) = \frac{1}{K}(\mu_{1,j} + \dots + \mu_{K,j}) \quad j = 2, \dots, d$$

$$\Leftrightarrow (\mu_{1,j-1} + \dots + \mu_{K,j-1}) - (\mu_{1,j} + \dots + \mu_{K,j}) = 0 \quad j = 2, \dots, d$$

$$\Leftrightarrow C_1(\underset{\sim}{\mu}_1 + \dots + \underset{\sim}{\mu}_K) = \underset{\sim}{0}$$

$$\Leftrightarrow C_1 [\underset{\sim}{\mu}_1 \dots \underset{\sim}{\mu}_K] \underset{\sim}{1}_K = \underset{\sim}{0}$$

$$\Leftrightarrow \underset{\sim}{1}_K^T \begin{bmatrix} \underset{\sim}{\mu}_1^T \\ \vdots \\ \underset{\sim}{\mu}_K^T \end{bmatrix} C_1^T = \underset{\sim}{0}^T$$

Both H_{01}, H_{02} and H_{03} are on the form

$ABD = 0$

$H_{01} \cap H_{02}$: coinciding profiles

$H_{01} \cap H_{03}$: parallel and horizontal profiles

$H_{01} \cap H_{02} \cap H_{03}$: coinciding and horizontal profiles

Test statistic for $H_{01} \cap H_{03}$:

$$T^2 = n (C_1 \bar{\underset{\sim}{x}})^T (C_1 S_p C_1^T)^{-1} C_1 \bar{\underset{\sim}{x}} \sim T^2(d-1, n-K)$$

cf. ex. 8.11

Note that $\bar{\underset{\sim}{x}} = \frac{1}{n} \sum_{k=1}^{K} n_k \bar{\underset{\sim}{x}}^k$, $\bar{\underset{\sim}{x}}^k = \frac{1}{n_k} \sum_{i=1}^{n_k} \underset{\sim}{x}_i^k$

$$S_p = \frac{\sum_{k=1}^{K} (n_k - 1) S_k}{\sum_{k=1}^{K} (n_k - 1)} = \frac{\sum_{k=1}^{K} (n_k - 1) S_k}{n - K}$$

$$E S_p = \frac{\sum_{k=1}^{K} (n_k - 1) E S_k}{n - K} = \frac{(n-K) \Sigma}{n - K} = \Sigma$$

$(n-K) S_p \sim W_d(n-K, \Sigma)$, as

$(n_k - 1) S_k \sim W_d(n_k - 1, \Sigma)$

$k = 1, \dots, K$, independent

Test for mean

(earlier developed, cf. lecture note MA 7)

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

$H_0: D^T \underset{\sim}{\mu} = \underset{\sim}{0}$, D^T $q \times d$, rank $D = q$

$Y = \underset{\sim}{1}_n \underset{\sim}{\mu}^T + U = XB + U$, $\underset{\sim}{u}_i \sim N_d(\underset{\sim}{0}, \Sigma)$, $i = 1, \ldots, n$ independent

$H_0: D^T \underset{\sim}{\mu} = \underset{\sim}{0} \Leftrightarrow \underset{\sim}{\mu}^T D = \underset{\sim}{0}^T \Leftrightarrow 1 \underset{\sim}{\mu}^T D = \underset{\sim}{0}^T \Leftrightarrow ABD = \underset{\sim}{0}^T$

$E = D^T Y^T (I_n - \underset{\sim}{1}_n (\underset{\sim}{1}_n^T \underset{\sim}{1}_n)^{-1} \underset{\sim}{1}_n^T) Y D$ cf. formula (8.72)

$= D^T Y^T (I_n - \frac{1}{n} \underset{\sim}{1}_n \underset{\sim}{1}_n^T) Y D$

$= D^T Q D$, cf. formula (1.13)

$H = D^T Y^T \underset{\sim}{1}_n (\underset{\sim}{1}_n^T \underset{\sim}{1}_n)^{-1} 1 \left(1 (\underset{\sim}{1}_n^T \underset{\sim}{1}_n)^{-1} 1\right)^{-1} 1 (\underset{\sim}{1}_n^T \underset{\sim}{1}_n)^{-1} \underset{\sim}{1}_n^T Y D$

cf. formula (8.71)

$= n D^T \bar{\underset{\sim}{y}} \bar{\underset{\sim}{y}}^T D$, rank $H = 1$

$T_q^2 = (n-1) \operatorname{tr}\left(n D^T \bar{\underset{\sim}{y}} \bar{\underset{\sim}{y}}^T D (D^T Q D)^{-1}\right)$ cf. formula (2.56)

$= n \bar{\underset{\sim}{y}}^T D (D^T S D)^{-1} D^T \bar{\underset{\sim}{y}} \sim T^2(q, n-1)$

hence the same test statistic as earlier found

cf. formula (3.17) with $A = D^T$ and $\underset{\sim}{b} = \underset{\sim}{0}$

Step down procedures

$Y = XB + U$, $\underset{\sim}{u}_i \sim N_d(\underset{\sim}{0}, \Sigma)$, $i = 1, \dots, n$, independent

$H_0 : AB = 0$

$Y_k = [\, y^{(1)} \cdots y^{(k)} \,]$, $B_k = [\, \underset{\sim}{\beta}^{(1)} \cdots \underset{\sim}{\beta}^{(k)} \,]$

Σ_k : k'th leading minor matrix in Σ

$$\Sigma_k = \begin{bmatrix} \Sigma_{k-1} & \underset{\sim}{\sigma}_{k-1,k} \\ \underset{\sim}{\sigma}_{k-1,k}^T & \sigma_{kk} \end{bmatrix} , \quad \underset{\sim}{y}_i^* = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{i,k-1} \end{bmatrix} , \quad k = 2, \dots, d$$

$$EY = XB \iff Ey_{ij} = \underset{\sim}{x}_i^T \underset{\sim}{\beta}^{(j)} , \quad \begin{matrix} i = 1, \dots, n \\ j = 1, \dots, d \end{matrix}$$

$$\sigma_k^2 = \frac{1}{(\Sigma^{-1})_{kk}} = \frac{\det \Sigma_k}{\det \Sigma_{k-1}}$$ cf. lemma 2.5

$$y_{ik} \,|\, \underset{\sim}{y}_i^* \sim N\left(\underset{\sim}{x}_i^T \underset{\sim}{\beta}^{(k)} + \underset{\sim}{\sigma}_{k-1,k}^T \Sigma_{k-1}^{-1} (\underset{\sim}{y}_i^* - B_{k-1}^T \underset{\sim}{x}_i) , \sigma_k^2 \right),$$

cf. theorem 2.1 (viii) and lemma 2.5

$\underset{\sim}{y}_1, \dots, \underset{\sim}{y}_n$ indep. $\Rightarrow$ $y_{1k} \,|\, \underset{\sim}{y}_1^*, \dots, y_{nk} \,|\, \underset{\sim}{y}_n^*$ indep.

Let $\underset{\sim}{\gamma}^{(k-1)} = \Sigma_{k-1}^{-1} \underset{\sim}{\sigma}_{k-1,k}$ and note that

$$\begin{aligned} \underset{\sim}{\sigma}_{k-1,k}^T \Sigma_{k-1}^{-1} (\underset{\sim}{y}_i^* - B_{k-1}^T \underset{\sim}{x}_i) &= (\underset{\sim}{y}_i^{*T} - \underset{\sim}{x}_i^T B_{k-1}) \Sigma_{k-1}^{-1} \underset{\sim}{\sigma}_{k-1,k} \\ &= (\underset{\sim}{y}_i^{*T} - \underset{\sim}{x}_i^T B_{k-1}) \underset{\sim}{\gamma}^{(k-1)} \end{aligned}$$

$$y^{(k)} \,|\, Y_{k-1} \sim N_n\left(X\underset{\sim}{\beta}^{(k)} + (Y_{k-1} - XB_{k-1}) \underset{\sim}{\gamma}^{(k-1)} , \sigma_k^2 I_n \right),$$

$k = 2, \dots, d$

Let $\eta^{(k)} = \beta^{(k)} - B_{k-1}\gamma^{(k-1)}$, $k = 2, \ldots, d$, $\eta^{(1)} = \beta^{(1)}$

$y^{(k)} \mid Y_{k-1} \sim N_n(X\eta^{(k)} + Y_{k-1}\gamma^{(k-1)}, \sigma_k^2 I_n)$, $k = 2, \ldots, d$

Also we have $y^{(1)} \sim N_n(X\beta^{(1)}, \sigma_{11} I_n)$

Assume $(a_i^T)^T \beta^{(j)}$, $i = 1, \ldots, q$, $j = 1, \ldots, d$, are estimable

$$AB = 0 \iff A\beta^{(k)} = 0 \ , \ k = 1, \ldots, d$$

$$\iff \begin{cases} A\beta^{(1)} = 0 \\ A(\beta^{(k)} - [\beta^{(1)} \ldots \beta^{(k-1)}]\gamma^{(k-1)}) = 0 \ , \ k = 2, \ldots, d \end{cases}$$

$$\iff A\eta^{(k)} = 0 \ , \ k = 1, \ldots, d$$

$H_{0k} : A\eta^{(k)} = 0$, $\bigcap_{k=1}^{d} H_{0k}$ equivalent to H_0

When testing H_{0k} in the model

$$y^{(k)} \mid Y_{k-1} = X\eta^{(k)} + Y_{k-1}\gamma^{(k-1)} + u_0^{(k)} \ , \quad u_0^{(k)} \sim N_n(0, \sigma_k^2 I_n)$$

we can use a test statistic F_k developed in the context of analysis of covariance, cf. p. 464 bottom (section 9.5.1)

$F_k \mid Y_{k-1} \sim F(q, n-r-k+1)$ when H_0 is true, cf. p. 465 top

$\Rightarrow$ $F_k \sim F(q, n-r-k+1)$ when H_0 is true, as the distribution does not depend on Y_{k-1}

$\Rightarrow$ F_k independent of Y_{k-1} when H_0 is true

$\Rightarrow$ F_k independent of $F_{k-1}, \ldots, F_1$ when H_0 is true

$\Rightarrow$ $F_1, \ldots, F_k$ independent when H_0 is true

$$1-\alpha = \prod_{k=1}^{d}(1-\alpha_k) \quad , \text{ where}$$

α is the significance level when testing H_0

α_k – – – – – H_{0k}

Multiple design model

Assume $\underset{\sim}{y}^{(j)} = \underset{\sim}{\theta}^{(j)} + \underset{\sim}{u}^{(j)}$, $j = 1, \dots, d$

$= X_j \underset{\sim}{\beta}^{(j)} + \underset{\sim}{u}^{(j)}$, X_j $n \times p_j$

and $H_{0j} : A_j \underset{\sim}{\beta}^{(j)} = \underset{\sim}{c}^{(j)}$, A_j $q_j \times p_j$

$\underset{\sim}{c}^{(j)}$ q_j-dim

$$H_0 = \bigcap_{j=1}^{d} H_{0j}$$

Ex.

EXAMPLE 8.1 Suppose we have two correlated regression models

$$E[y_{i1}] = \theta_{i1} = \beta_{10} + \beta_{11}x_i$$

and

$$E[y_{i2}] = \theta_{i2} = \beta_{20} + \beta_{21}x_i + \beta_{22}x_i^2 \qquad (i = 1, 2, \dots, n).$$

Here $\mathbf{Y} = [(y_{ij})] = (\mathbf{y}^{(1)}, \mathbf{y}^{(2)})$,

$$\boldsymbol{\theta}^{(1)} = \begin{pmatrix} \theta_{11} \\ \theta_{21} \\ \vdots \\ \theta_{n1} \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{11} \end{pmatrix} = \mathbf{X}_1 \boldsymbol{\beta}^{(1)}$$

and

$$\boldsymbol{\theta}^{(2)} = \begin{pmatrix} \theta_{12} \\ \theta_{22} \\ \vdots \\ \theta_{n2} \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \begin{pmatrix} \beta_{20} \\ \beta_{21} \\ \beta_{22} \end{pmatrix} = \mathbf{X}_2 \boldsymbol{\beta}^{(2)}.$$

The hypothesis of no regression on x then takes the form

$$H_0 : \underbrace{\beta_{11} = 0}_{H_{01}\uparrow}, \qquad \underbrace{\begin{pmatrix} \beta_{21} \\ \beta_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{H_{02}\uparrow},$$

so that

$$\mathbf{A}_1 = (0, 1) \quad \text{and} \quad \mathbf{A}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$X = [X_1 \; X_2 \; \dots \; X_d]$ $\quad n \times r$, $r = \sum_{j=1}^{d} r_j$, rank $X = r$

$\forall \underset{\sim}{\ell} \neq \underset{\sim}{0}$: $\underset{\sim}{y} = Y\underset{\sim}{\ell}$, $\Theta\underset{\sim}{\ell} = [\underset{\sim}{\theta}^{(1)} \dots \underset{\sim}{\theta}^{(d)}]\underset{\sim}{\ell}$

$$= [X_1 \underset{\sim}{\beta}^{(1)} \dots X_d \underset{\sim}{\beta}^{(d)}]\underset{\sim}{\ell}$$

$$= \sum_{j=1}^{d} X_j \underset{\sim}{\beta}^{(j)} \ell_j$$

$$= X\underset{\sim}{\beta} \text{ , where } \underset{\sim}{\beta} = \begin{bmatrix} \ell_1 \underset{\sim}{\beta}^{(1)} \\ \vdots \\ \ell_d \underset{\sim}{\beta}^{(d)} \end{bmatrix}$$

$H_{oj\ell}$: $\ell_j A_j \underset{\sim}{\beta}^{(j)} = \ell_j \underset{\sim}{c}^{(j)}$, $\bigcap_{\ell} \bigcap_{j=1}^{d} H_{oj\ell}$ equivalent to H_o

Let $q = \max(q_1, \dots, q_d)$

- $A_j^* = \begin{bmatrix} A_j \\ 0 \end{bmatrix} \begin{matrix} q_j \\ q - q_j \end{matrix}$, $A = [A_1^* \; A_2^* \dots A_d^*]$ $\quad q \times r$
- $\underset{\sim}{c}^{(j)*} = \begin{bmatrix} \underset{\sim}{c}^{(j)} \\ \underset{\sim}{0} \end{bmatrix} \begin{matrix} q_j \\ q - q_j \end{matrix}$, $C = [\underset{\sim}{c}^{(1)*} \; \underset{\sim}{c}^{(2)*} \dots \underset{\sim}{c}^{(d)*}]$ $\quad q \times d$
- $\underset{\sim}{c} = C\underset{\sim}{\ell}$

$H_{o\ell}$: $\sum_{j=1}^{d} A_j^* \ell_j \underset{\sim}{\beta}^{(j)} = \sum_{j=1}^{d} \ell_j \underset{\sim}{c}^{(j)*} \iff A\underset{\sim}{\beta} = \underset{\sim}{c}$

$\bigcap_{\ell} H_{o\ell}$ equivalent to H_o

Assume $A\underset{\sim}{\beta}$ is estimabel

Test statistic for $H_{o\ell}$:

$F_\ell = \frac{n-r}{q} \frac{\underset{\sim}{\ell}^T H \underset{\sim}{\ell}}{\underset{\sim}{\ell}^T E \underset{\sim}{\ell}} \sim F(q, n-r)$ when $H_{o\ell}$ is true ,

cf. lecture note 14 p. 3

As test statistic for H_0 we can use φ_{max}, the largest eigenvalue corresponding to HE^{-1}, cf. lecture note 14 p. 4

$$\left.\begin{array}{l} E \sim W_d(n-r, \Sigma) \\ H \sim W_d(q, \Sigma) \text{ when } H_0 \text{ is true} \end{array}\right\} \text{ independent}$$

cf. theorem 2.4 (ii)

Alternative test statistic for test of H_0:

$$\Lambda = \frac{\det E}{\det(E+H)} \sim U(d, q, n-r)$$

Note that $\underset{\sim}{\beta} = \begin{bmatrix} \underset{\sim}{\beta}^{(1)} & \underset{\sim}{0} & \cdots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\beta}^{(2)} & \cdots & \underset{\sim}{0} \\ \vdots & \vdots & & \\ \underset{\sim}{0} & \underset{\sim}{0} & \cdots & \underset{\sim}{\beta}^{(d)} \end{bmatrix} \underset{\sim}{l} = B_d \underset{\sim}{l}$, i.e.

$H_{0l}: AB_d \underset{\sim}{l} = C\underset{\sim}{l} \iff H_0: AB_d = C$

Also $\quad \underset{\sim}{y} = Y\underset{\sim}{l} = \Theta\underset{\sim}{l} + U\underset{\sim}{l} = XB_d\underset{\sim}{l} + U\underset{\sim}{l}$,

$\iff Y = XB_d + U$

$$E = Y^T(I_n - X(X^TX)^-X^T)Y$$

$$H = (A\hat{B}_d - C)^T(A(X^TX)^-A^T)^{-1}(A\hat{B}_d - C) \quad \text{cf. "corollary" 1 to theorem 8.5}$$

$$= (A(X^TX)^-X^TY - C)^T(A(X^TX)^-A^T)^{-1}(A(X^TX)^-X^TY - C)$$

$\hat{B}_d$ is not efficient, as it may contain estimates of some elements of B_d known to be zero.