(assume ✱ in LNp 52)

Note. Suppose $W_i = \text{span}\{X_i\}$, GLS = OLS iff $\text{span}\{X_i\} = \text{span}\{\Sigma^{-1}X_i\}$ (LM, LNp 6-2)

⊙ Consider the sequential ANOVA: for $i = 1, \ldots, k$,

Under the linear model in LNp 49 → $H_0^{(i)}: \boldsymbol{\beta}_i = \mathbf{0}$ $(\omega_i = V_{i-1})$ vs. $H_A^{(i)}: \boldsymbol{\beta}_i \neq \mathbf{0}$ $(\Omega_i = V_i)$

and suppose $W_i = V_i \ominus V_{i-1} \subset S_j$. (e.g., $W_3 \subset S_2 \Rightarrow i=3 \rightarrow j=2$)

check (P1) in LNp 2-35

(L1) $SS_{W_i} \equiv RSS_{\omega_i} - RSS_{\Omega_i} = \|\boldsymbol{P}_{W_i}\boldsymbol{Y}\|^2$ and $df_{\omega_i} - df_{\Omega_i} = \dim(W_i) \equiv r_i$

OLS

(L2) $\boldsymbol{P}_{W_i}\boldsymbol{Y} = \boldsymbol{P}_{W_i}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{P}_{W_i}\boldsymbol{\varepsilon}$, where

$Y = X\beta + \varepsilon$

check (P2) in LNp 2-35

Recall. For $v \in S_j$, $v \rightarrow \Sigma v = e_j v$

– $\boldsymbol{P}_{W_i}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{P}_{W_i}(\boldsymbol{X}_i\boldsymbol{\beta}_i + \cdots + \boldsymbol{X}_k\boldsymbol{\beta}_k) = \boldsymbol{P}_{W_i}(\boldsymbol{u}_i + \cdots + \boldsymbol{u}_k)$

By (N1) in LNp 2-30 and $\varepsilon \sim N(0, \Sigma)$

– $\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon} \sim \underline{N}(\mathbf{0}, \boldsymbol{P}_{W_i}\boldsymbol{\Sigma}\boldsymbol{P}_{W_i}^T)$, where

Suppose $W_i = \text{span}\{B_i\} \Rightarrow$ columns of $B_i \in W_i \subset S_j$. Then, $P_{W_i} = B_i(B_i^TB_i)^{-1}B_i^T$, and

$$\boldsymbol{P}_{W_i}\boldsymbol{\Sigma}\boldsymbol{P}_{W_i}^T = \boldsymbol{P}_{W_i}e_j\boldsymbol{P}_{W_i}^T = e_j\boldsymbol{P}_{W_i}$$

$\Sigma P_{W_i}^T = \underbrace{\Sigma B_i}_{=e_jB_i}(B_i^TB_i)^{-1}B_i^T = e_jB_i(B_i^TB_i)^{-1}B_i^T = e_jP_{W_i}$

the reason why we split $\mathbb{R}^N$ into 3 strata S_1, S_2, S_3

$P_{W_i}^3 = P_{W_i}$

and $\boldsymbol{P}_{W_i} = \boldsymbol{P}_{W_i}^{=}$.

– Thus, $\boldsymbol{P}_{W_i}\boldsymbol{Y} \sim \underline{N}(\boldsymbol{P}_{W_i}(\boldsymbol{u}_i + \cdots + \boldsymbol{u}_k), e_j\boldsymbol{P}_{W_i})$

$(e_jP_{W_i})(\frac{1}{e_j}P_{W_i})(e_jP_{W_i}) = e_jP_{W_i}$

$(e_jP_{W_i})^-$

check (P3) in LNp 2-35

(L3) $\|\boldsymbol{P}_{W_i}\boldsymbol{Y}\|^2 = \boldsymbol{Y}^T\boldsymbol{P}_{W_i}\boldsymbol{Y} = (\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})^T\boldsymbol{P}_{W_i}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$

$$= \|\boldsymbol{P}_{W_i}\boldsymbol{X}\boldsymbol{\beta}\|^2 + 2(\boldsymbol{X}\boldsymbol{\beta})^T\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^T\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon},$$

$= 0$ under null

By (N6) in LNp 2-30

where $\boldsymbol{\varepsilon}^T\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^T\boldsymbol{P}_{W_i}^T\boldsymbol{P}_{W_i}\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon} = e_j(\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon})^T(\boldsymbol{P}_{W_i}/e_j)(\boldsymbol{P}_{W_i}\boldsymbol{\varepsilon}) \sim e_j\chi^2_{r_i}$

FYI. $\|\boldsymbol{P}_{W_i}\boldsymbol{Y}\|^2/e_j \sim$ noncentral $\chi^2_{r_i}(\|\boldsymbol{P}_{W_i}\boldsymbol{X}\boldsymbol{\beta}\|^2/e_j)$ ← noncentral parameter

(L4) $E(SS_{W_i}) = E(\|\boldsymbol{P}_{W_i}\boldsymbol{Y}\|^2) = \|\boldsymbol{P}_{W_i}(\boldsymbol{u}_i + \cdots + \boldsymbol{u}_k)\|^2 + r_ie_j$

cf.

(L5) For the jth error space T_j $(\subset S_j)$, $j = 1, 2, 3$,

Intuition? check LNp 48

cf.

– $SS_{T_j} \equiv \|\boldsymbol{P}_{T_j}\boldsymbol{Y}\|^2 = \boldsymbol{Y}^T\boldsymbol{P}_{T_j}\boldsymbol{Y} = \boldsymbol{Y}^T\left(\boldsymbol{P}_{S_j} - \sum_{W_i \subset S_j}\boldsymbol{P}_{W_i}\right)\boldsymbol{Y}$

$$= (\boldsymbol{Y}^T\boldsymbol{P}_{S_j}\boldsymbol{Y}) - \textstyle\sum_{W_i \subset S_j}(\boldsymbol{Y}^T\boldsymbol{P}_{W_i}\boldsymbol{Y}) = \|\boldsymbol{P}_{S_j}\boldsymbol{Y}\|^2 - \sum_{W_i \subset S_j}\|\boldsymbol{P}_{W_i}\boldsymbol{Y}\|^2$$

Can use $\|P_{T_j}Y\|^2$, $j = 1, 2, 3$, to estimate e_1, e_2, e_3 or $\sigma_R^2, \sigma_W^2, \sigma_S^2$

$= 0$, ∵ $T_j \subset V_k^\perp$

– $\boldsymbol{P}_{T_j}\boldsymbol{Y} = \boldsymbol{P}_{T_j}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{P}_{T_j}\boldsymbol{\varepsilon} = \boldsymbol{P}_{T_j}\boldsymbol{\varepsilon} \sim \underline{N}(\mathbf{0}, e_j\boldsymbol{P}_{T_j})$

$P_{T_1}\varepsilon = P_{T_1}(A_R\varepsilon^R + A_W\varepsilon^W + A_S\varepsilon^S)$

$P_{T_2}\varepsilon = P_{T_2}(A_W\varepsilon^W + A_S\varepsilon^S)$

$P_{T_3}\varepsilon = P_{T_3}(A_S\varepsilon^S)$

– $\|\boldsymbol{P}_{T_j}\boldsymbol{Y}\|^2 = \boldsymbol{\varepsilon}^T\boldsymbol{P}_{T_j}\boldsymbol{\varepsilon} \sim e_j\chi^2_{\dim(T_j)}$

$\frac{SS_{T_j}}{\dim(T_j)} \xrightarrow{\text{unbiased estimator}} e_j$

– $E(SS_{T_j}) = E(\|\boldsymbol{P}_{T_j}\boldsymbol{Y}\|^2) = \dim(T_j)e_j = \left(\dim(S_j) - \sum_{W_i \subset S_j}\dim(W_i)\right)e_j$

(L6) Since all W_i's and T_j's are orthogonal subspaces of the eigenspaces of $\boldsymbol{\Sigma}$,

e.g. $\text{cov}(P_{W_i}Y, P_{T_j}Y) = P_{W_i}\Sigma P_{T_j}^T = P_{W_i}e_jP_{T_j}^T = e_jP_{W_i}P_{T_j} = 0$

$\Rightarrow$ all $\boldsymbol{P}_{W_i}\boldsymbol{Y}$'s and $\boldsymbol{P}_{T_j}\boldsymbol{Y}$'s are *independent* random vectors

similar for $P_{W_i}\Sigma P_{W_\ell}$ and $P_{T_j}\Sigma P_{T_\ell}$

$\Rightarrow$ all $\|\boldsymbol{P}_{W_i}\boldsymbol{Y}\|^2$'s and $\|\boldsymbol{P}_{T_j}\boldsymbol{Y}\|^2$'s are *independent* random variables

• A geometric view:

same structure as an LM with $\Sigma = \sigma^2 I$