1. Consider the following full rank linear model

$y = Xb + e \quad e \sim N(0, \sigma^2 I), \quad X \in \mathbb{R}^{n\times p}$

$\hat{y} = X\hat{b}$, $\hat{b}$ is the OLS estimate of b

(a) Prove that $\sum_{i=1}^{n} \hat{y}_i (y_i - \hat{y}_i) = 0$

Denote $H := X(X^TX)^{-1}X^T$, where X is data matrix. we know H is idempotent

then $\hat{Y} = X\hat{b} = X(X^TX)^{-1}X^TY = HY$

$$\sum_{i=1}^{n} \hat{y}_i (y_i - \hat{y}_i) = \hat{Y}^T(Y - \hat{Y})$$

$$= Y^TH(I-H)Y$$

$$= Y^T(H-H)Y = 0$$

(b) Prove that $\sum_{i=1}^{n} Var(\hat{y}_i) = p\sigma^2$

Since $Var\,\hat{Y} = Var(HY)$

$$= H[Var\,Y]H$$

$$= H\sigma^2 I H$$

$$= \sigma^2 H$$

then $\sum_{i=1}^{n} Var\,\hat{y}_i = tr(Var\,\hat{Y}) = tr(\sigma^2 H) = p\sigma^2$

2. Mean Dispersion Error (MDE) of an estimator $\hat{\beta}$: $M(\hat{\beta}, \beta) = \mathbb{E}(\hat{\beta}-\beta)(\hat{\beta}-\beta)^T$

MDE I criterion : $\hat{\beta}_2$ is called MDE I superior to $\hat{\beta}_1$

if $\Delta(\hat{\beta}_1, \hat{\beta}_2) \triangleq M(\hat{\beta}_1, \beta) - M(\hat{\beta}_2, \beta) \geq 0$ (Semi-positive definite)

Given a theorem $I - aa^T \geq 0$ iff $a^Ta \leq 1$

Q: $Y = X\beta + e$, $\mathbb{E}e = 0$ $Var\,e = \sigma^2 I$, $X = [1, \tilde{X}]$, X is full column rank

$\hat{\beta}_1 = (1+\rho)^{-1}\hat{\beta}$, $\rho > 0$ and known

Prove $\beta^TX^TX\beta \leq \sigma^2 \Rightarrow \Delta(\hat{\beta}, \hat{\beta}_1) \geq 0$ (See next page)

$$\mathbb{E}\hat{\beta} = (X^TX)^{-1}X^T\mathbb{E}Y = (X^TX)^{-1}X^TX\beta = \beta \quad \Rightarrow \mathbb{E}\hat{\beta}_1 = (1+\rho)^{-1}\beta$$

$$\begin{aligned} MDE(\hat{\beta},\beta) &= \mathbb{E}(\hat{\beta}-\beta)(\hat{\beta}-\beta)^T = Var\,\hat{\beta} \\ &= (X^TX)^{-1}X^T[Var\,Y]\,X(X^TX)^{-1} \\ &= \sigma^2(X^TX)^{-1} \end{aligned}$$

$$\begin{aligned} MDE(\hat{\beta}_1,\beta) &= \mathbb{E}(\hat{\beta}_1-\beta)(\hat{\beta}_1-\beta)^T \\ &= \mathbb{E}(\hat{\beta}_1-(1+\rho)^{-1}\beta-\tfrac{\rho}{1+\rho}\beta)(\hat{\beta}_1-(1+\rho)^{-1}\beta-\tfrac{\rho}{1+\rho}\beta)^T \\ &= Var\,\hat{\beta}_1 + \frac{\rho^2}{(1+\rho)^2}\beta\beta^T \\ &= \frac{1}{(1+\rho)^2}Var\,\hat{\beta} + \frac{\rho^2}{(1+\rho)^2}\beta\beta^T \end{aligned}$$

then
$$\begin{aligned} \Delta(\hat{\beta},\hat{\beta}_1) &= \left(1-\frac{1}{(1+\rho)^2}\right)Var\,\hat{\beta} - \frac{\rho^2}{(1+\rho)^2}\beta\beta^T \\ &= \frac{2\rho+\rho^2}{(1+\rho)^2}\sigma^2(X^TX)^{-1} - \frac{\rho^2}{(1+\rho)^2}\beta\beta^T \\ &=: M \end{aligned}$$

To show M is semi-positive definite, need to show

for $\forall\, v$ compatiable vector, M satisfies $v^TMv \geq 0$

i.e. $\frac{2\rho+\rho^2}{(1+\rho)^2}\sigma^2 v^T(X^TX)^{-1}v - \frac{\rho^2}{(1+\rho)^2}v^T\beta\beta^Tv \geq 0.$

Since X is full column rank,

then for $\forall\, v$, we can find a u such that $v = (X^TX)^{\frac{1}{2}}u$ $\left(u = (X^TX)^{-\frac{1}{2}}v\right)$

Since $\beta^TX^TX\beta \leq \sigma^2$, i.e. $\left[\frac{(X^TX)^{\frac{1}{2}}\beta}{\sigma}\right]^T\left[\frac{(X^TX)^{\frac{1}{2}}\beta}{\sigma}\right] \leq 1$

then by the given thm, $I - \sigma^{-2}(X^TX)^{\frac{1}{2}}\beta\beta^T(X^TX)^{\frac{1}{2}} \geq 0 \Rightarrow (X^TX)^{\frac{1}{2}}\beta\beta^T(X^TX)^{\frac{1}{2}} \leq \sigma^2 I$

then
$$\begin{aligned} v^TMv &= \frac{2\rho+\rho^2}{(1+\rho^2)}\sigma^2u^T(X^TX)^{\frac{1}{2}}(X^TX)^{-1}(X^TX)^{\frac{1}{2}}u - \frac{\rho^2}{(1+\rho)^2}u^T(X^TX)^{\frac{1}{2}}\beta\beta^T(X^TX)^{\frac{1}{2}}u \\ &\geq \frac{2\rho+\rho^2}{(1+\rho^2)}\sigma^2u^Tu - \frac{\rho^2}{(1+\rho)^2}\sigma^2u^Tu \\ &= \frac{2\rho}{1+\rho^2}\sigma^2u^Tu \\ &\geq 0 \quad // \end{aligned}$$

3. $\lambda^T\beta$ estimatable function of β in $Y = X\beta + \varepsilon$

$\mathbb{E}Y = X\beta$ and $X \in \mathbb{R}^{n\times p}$, $\operatorname{rank}(X) = k < p \le n$

$\hat\beta$ be any solution to $X^TX\beta = X^TY$, r be any solution to $X^TXr = \lambda$

For two estimators $\lambda^T\hat\beta$ and r^TX^TY, prove

(a) $\mathbb{E}(\lambda^T\hat\beta) = \mathbb{E}(r^TX^TY) = \lambda^T\beta$

$$\mathbb{E}(\lambda^T\hat\beta) = \mathbb{E}(r^T\underline{X^TX\hat\beta}) = \mathbb{E}(r^TX^TY) = r^TX^T\mathbb{E}Y = \underline{r^TX^TX}\beta = \lambda^T\beta$$

(b) r^TX^TY is invariant to the choice of r.

Denote r_1 and r_2 are two solutions to $X^TXr = \lambda$, then

$$r_1^TX^TY - r_2^TX^TY = r_1^TX^TX\hat\beta - r_2^TX^TX\hat\beta = \lambda\hat\beta - \lambda\hat\beta = 0$$

then r^TX^TY is invariant to the choice of r.

4. Consider model

$$Y = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}}_{=:X} \underbrace{\begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}}_{b} + \varepsilon \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2 I_9)$$

(a) State the conditions when λ^Tb is estimable. $\lambda^T = (\lambda_0\ \lambda_1\ \lambda_2\ \lambda_3\ \lambda_4\ \lambda_5\ \lambda_6)$

By the Theorem 2 in Lecture Notes.

λ^Tb is estimable iff $\lambda^TH = \lambda^T$, $H = GX^TX$ G is a general inverse of (X^TX)

Now try to find a G

$$\text{Since } X^TX = \begin{bmatrix} 9 & 3 & 3 & 3 & 3 & 3 & 3 \\ 3 & 3 & 0 & 0 & 1 & 1 & 1 \\ 3 & 0 & 3 & 0 & 1 & 1 & 1 \\ 3 & 0 & 0 & 3 & 1 & 1 & 1 \\ 3 & 1 & 1 & 1 & 3 & 0 & 0 \\ 3 & 1 & 1 & 1 & 0 & 3 & 0 \\ 3 & 1 & 1 & 1 & 0 & 0 & 3 \end{bmatrix}, \quad \operatorname{rank}(X^TX) = 5 \quad (r_1 = \tfrac{1}{2}(r_2 + \cdots + r_7))$$

$(r_2 + r_3 + r_4 = r_5 + r_6 + r_7)$

Partition $X^TX = \begin{bmatrix} P & Q \\ S & R \end{bmatrix}$ $\quad R = \begin{bmatrix} 3 & 0 & 1 & 1 & 1 \\ 0 & 3 & 1 & 1 & 1 \\ 1 & 1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 3 & 0 \\ 1 & 1 & 0 & 0 & 3 \end{bmatrix}$ $\quad \text{rank}(R) = 5.$

By block matrix inversion theorem, if $R = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, then

$$R^{-1} = \begin{bmatrix} (A-BD^{-1}C)^{-1} & -(A-BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A-BD^{-1}C)^{-1} & D^{-1}+D^{-1}C(A-BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$$

Denote $A = 3I_2$, $B = C^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, $D = 3I_3$, then $D^{-1} = \frac{1}{3}I_3$

$$\begin{cases} (A-BD^{-1}C)^{-1} = (3I_2 - J_2)^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \frac{1}{3}(I_2 + J_2) \\ -(A-BD^{-1}C)^{-1}BD^{-1} = -\frac{1}{3}(I_2+J_2)B\frac{1}{3}I_3 = -\frac{1}{3}B = -\frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\ -D^{-1}C(A-BD^{-1}C)^{-1} = -\frac{1}{3}I_3C\frac{1}{3}(I_2+J_2) = -\frac{1}{3}C = -\frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \\ D^{-1}+D^{-1}C(A-BD^{-1}C)^{-1}BD^{-1} = \frac{1}{3}I_3 + \frac{1}{3}I_3C\frac{1}{3}(I_2+J_2)B\frac{1}{3}I_3 = \frac{1}{3}I_3 + \frac{2}{9}J_3 \end{cases}$$

then $G = \begin{bmatrix} 0 & 0 \\ 0 & R^{-T} \end{bmatrix}$ $\quad \Rightarrow \quad H = G(X^TX)$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{5}{9} & \frac{2}{9} & \frac{2}{9} \\ 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{9} & \frac{5}{9} & \frac{2}{9} \\ 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{9} & \frac{2}{9} & \frac{5}{9} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) Is $\alpha_1 + \alpha_2$ estimable?

$\lambda^T = [0\ 1\ 1\ 0\ 0\ 0\ 0]$, $\lambda^T H = [0\ {-1}\ 1\ 0\ 0\ 0\ 0] \neq \lambda^T$, so NO.

(c) Is $\beta_1 + \beta_2 + \beta_3$ estimable?

$\lambda^T = [0\ 0\ 0\ 0\ 1\ 1\ 1]$, $\lambda^T H = [3\ 3\ 0\ 0\ 1\ 1\ 1] \neq \lambda^T$, so NO

(d) Is $\mu + \alpha_2$ estimable?

$\lambda^T = [1\ 0\ 1\ 0\ 0\ 0\ 0]$, $\lambda^T H = [0\ {-1}\ 1\ 0\ 0\ 0\ 0] \neq \lambda^T$, so NO

(e) Is $6\mu + 2\alpha_1 + 2\alpha_2 + 2\alpha_3 + 3\beta_1 + 3\beta_3$ estimable?

$\lambda^T = [6\ 2\ 2\ 2\ 3\ 0\ 3]$, $\lambda^T H = [6\ 2\ 2\ 2\ 3\ 0\ 3] = \lambda^T$, so YES

(f) Is $\alpha_1 - 2\alpha_2 + \alpha_3$ estimable?

$$\lambda^T = [0\ 1\ -2\ 1\ 0\ 0\ 0],\quad \lambda^T H = [0\ 1\ -2\ 1\ 0\ 0\ 0] = \lambda^T, \text{ so YES}$$

5. Consider model $Y = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \beta + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I_3)$

Show $\lambda_1\beta_1 + \lambda_2\beta_2 + \lambda_3\beta_3$ is estimable iff $\lambda_1 = \lambda_2 + \lambda_3$

$$X := \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \text{ then } X^TX = \left[\begin{array}{c|cc} 3 & 2 & 1 \\ \hline 2 & 2 & 0 \\ 1 & 0 & 1 \end{array}\right] =: \begin{bmatrix} P & Q \\ S & R \end{bmatrix} \quad \text{rank}(X) = \text{rank}(R) = 2.$$

$$\text{then } R^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \text{ then general inverse of } X^TX = G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow H := GX^TX = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

By the Thm 2 in Lecture Notes $\lambda^T\beta$ is estimable iff $\lambda^T H = \lambda^T$ $\lambda = [\lambda_1\ \lambda_2\ \lambda_3]$

$$\Leftrightarrow [\lambda_2 + \lambda_3\ \ \lambda_2\ \ \lambda_3] = [\lambda_1\ \lambda_2\ \lambda_3]$$

$$\Leftrightarrow \lambda_1 = \lambda_2 + \lambda_3 \ //$$

6. $t_{ij} = \frac{2\pi}{\sqrt{g}}\sqrt{l_i} + \varepsilon_{ij} \qquad i = 1, \dots, k \quad j = 1, \dots, n_i \qquad \varepsilon_{ij} \sim N(0, \sigma^2)$

Obtain the best unbiased estimate of $\frac{2\pi}{\sqrt{g}}$ and an estimate of its variance.

the model can be written as matrix form.

$$T = Lb + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I_{\sum_{i=1}^k n_i}),\ T = \begin{bmatrix} t_{11} \\ \vdots \\ t_{1n_1} \\ \vdots \\ t_{k1} \\ \vdots \\ t_{kn_k} \end{bmatrix},\quad L = \begin{bmatrix} \sqrt{l_1}\mathbf{1}_{n_1} \\ \vdots \\ \sqrt{l_i}\mathbf{1}_{n_i} \\ \vdots \\ \sqrt{l_k}\mathbf{1}_{n_k} \end{bmatrix},\quad b = \frac{2\pi}{\sqrt{g}}$$

then by Gaussian-Markov Thm. the B.L.U.E of b is

$$\hat{b} = (L^TL)^{-1}L^TT = \frac{\sum_{i=1}^k \left[\sqrt{l_i}\sum_{j=1}^{n_i} t_{ij}\right]}{\sum_{i=1}^k n_i l_i}$$

an estimate of variance is $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^k\sum_{j=1}^{n_i}(t_{ij} - \sqrt{l_i}\hat{b})^2$

7. $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, $i=1,\dots,k$ $j=1,\dots,n_i$ $\varepsilon_{ij} \sim N(0,\sigma^2)$

(a) Is the null hypothesis $H_0 : \frac{\mu+\alpha_1}{a_1} = \frac{\mu+\alpha_2}{a_2} = \cdots = \frac{\mu+\alpha_k}{a_k}$ testable

Rewritten model as matrix form $Y = X\beta + \varepsilon$,

$$X = \begin{bmatrix} & 1_{n_1} & 0 & \cdots & 0 \\ & \vdots & \ddots & & \\ 1_N & 0 & \cdots & 1_{n_i} \cdots & 0 \\ & \vdots & & \ddots & \\ & 0 & \cdots & 0 & 1_{n_k} \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_k \end{bmatrix}, \quad \varepsilon \sim N(0, \sigma^2 I_N), \; N = \sum_{i=1}^{k} n_i$$

Since $H_0 \Leftrightarrow \frac{\mu+\alpha_1}{a_1} - \frac{\mu+\alpha_j}{a_j} = 0$ for $\forall\ 2 \le j \le k$

$\Leftrightarrow (\frac{1}{a_1} - \frac{1}{a_j})\mu + \frac{1}{a_1}\alpha_1 - \frac{1}{a_j}\alpha_j = 0.$ for $\forall\ 2 \le j \le k$

then H_0 is equivalent to $K^T\beta = 0$

$$K^T = \begin{bmatrix} \frac{1}{a_1} - \frac{1}{a_2} & \frac{1}{a_1} & -\frac{1}{a_2} & 0 & \cdots & 0 \\ \vdots & & & & & \\ \frac{1}{a_1} - \frac{1}{a_j} & \frac{1}{a_1} & \cdots & -\frac{1}{a_j} & \cdots & 0 \\ \vdots & & & & & \\ \frac{1}{a_1} - \frac{1}{a_k} & \frac{1}{a_1} & \cdots & 0 & \cdots 0 & \frac{1}{a_k} \end{bmatrix} \in \mathbb{R}^{(k-1)\times(k+1)}$$

Since

$$X^TX = \begin{bmatrix} N & n_1 & n_2 & \cdots & n_k \\ n_1 & n_2 & 0 & \cdots & 0 \\ n_2 & 0 & n_2 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ n_k & 0 & 0 & \cdots & n_k \end{bmatrix} \text{ whose general inverse } G = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{n_1} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{n_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{n_k} \end{bmatrix}$$

then

$$H := GX^TX = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

then we have $K^TH = K^T$,

thus $H_0 : K^T\beta = 0$ is testable

(b) Derive the testing procedure.

① Construct the test statistic

$$F(Y) = \frac{Q/(k-1)}{SSE/n-(k+1)} \sim F_{k-1,\, n-(k+1)}$$

(continue on the next page)

$$Q = (K^T\hat{\beta})^T (K^T G K)^{-1} (K^T\hat{\beta})$$

$$SSE = (Y - X\hat{\beta})^T (Y - X\hat{\beta})$$

$$\hat{\beta} = G X^T Y$$

② Given the observations y, calculate the value of test statistic $F(y)$

③ If $F(y) \geq F_{k-1, n-(k+1), \alpha}$ the upper α-quantile of $F_{k-1, n-(k+1)}$

then reject the H_0, otherwise accept it.

8. $\begin{cases} y_i = \mu + \alpha_1 + \varepsilon_i & i = 1, \cdots, n \\ y_i = \mu + \alpha_2 + \varepsilon_i & i = n+1, \cdots, n+m \end{cases}$

(a) Is $\beta = (\mu \ \alpha_1 \ \alpha_2)^T$ identifiable?

Write this model in matrix form

$$Y = X\beta + \varepsilon, \text{ where } X = \left[\begin{array}{c|cc} & 1_n & 0 \\ 1_{n+m} & & \\ & 0 & 1_m \end{array}\right], \ \beta = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \end{bmatrix}$$

For $\forall k$, the model can be rewritten as

$$Y = X\beta + \varepsilon = X\tilde{\beta} + \varepsilon, \text{ where } \tilde{\beta} = \begin{bmatrix} \mu + k \\ \alpha_1 - k \\ \alpha_2 - k \end{bmatrix} \neq \beta$$

then β is not identifiable.

(b) When a constraint $\alpha_1 + \alpha_2 = 0$ is imposed on β, is β identifiable?

If there is another $\tilde{\beta} = (\tilde{\mu} \ \tilde{\alpha}_1 \ \tilde{\alpha}_2)^T$ such that $\begin{cases} Y = X\beta + \varepsilon = X\tilde{\beta} + \varepsilon \\ \tilde{\alpha}_1 + \tilde{\alpha}_2 = \alpha_1 + \alpha_2 = 0 \end{cases}$

then $\begin{cases} \mu + \alpha_1 = \tilde{\mu} + \tilde{\alpha}_1 \\ \mu + \alpha_2 = \tilde{\mu} + \tilde{\alpha}_2 \\ \alpha_1 + \alpha_2 = \tilde{\alpha}_1 + \tilde{\alpha}_2 = 0 \end{cases} \Rightarrow \begin{cases} \tilde{\mu} = \mu \\ \tilde{\alpha}_1 = \alpha_1 = -\tilde{\alpha}_2 = -\tilde{\alpha}_2 \end{cases}$

thus under this constraint, β is identifiable.

(c) When a constraint $\mu = 0$ is imposed on β, is β identifiable?

Similarly. $\begin{cases} \mu + \alpha_1 = \tilde{\mu} + \tilde{\alpha}_1 \\ \mu + \alpha_2 = \tilde{\mu} + \tilde{\alpha}_2 \\ \mu = \tilde{\mu} = 0 \end{cases} \Rightarrow \begin{cases} \tilde{\mu} = \mu = 0 \\ \tilde{\alpha}_1 = \alpha_1 \\ \tilde{\alpha}_2 = \alpha_2 \end{cases}$

thus under this constraint, β is identifiable.

9. Quantile regression model $Y_i = X_i \beta_\tau + \varepsilon_{i\tau}$ X_i and $\varepsilon_{i\tau}$ are correlated.

$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}\, G(\beta)$, $G(\beta) = \rho_\tau(Y - X\beta)$ $\rho_\tau(u) = \begin{cases} \tau u & u > 0 \\ (\tau - 1) u & u \le 0 \end{cases}$

$\sqrt{n}(\hat{\beta} - \beta) \sim N(0, A^{-1} B A^{-T})$ can be proved. Find such A and B

Since $G(\beta) = (\tau - \frac{1}{2})(Y - X\beta) + \frac{1}{2}|Y - X\beta|$

then $g(\beta) = \frac{\partial}{\partial \beta} G(\beta) = -(\tau - \frac{1}{2}) X^T - \frac{1}{2} X^T \operatorname{sgn}(Y - X\beta)$

$\dot{g}(\beta) = \frac{\partial^2}{\partial \beta^2} G(\beta) = \frac{\partial}{\partial \beta} g(\beta) = \frac{1}{2} X^T X \cdot 2\, \delta(Y - X\beta) = X^T X\, \delta(Y - X\beta)$

If $X \perp\!\!\!\perp \varepsilon$ ...