1. Responda las siguientes preguntas.

a) Suponga que $\underline{\mathbf{x}}_0 = [1, x_{01}, \cdots, x_{0k}]$ es un punto en el que no se comete extrapolación, luego $\underline{\mathbf{x}}_0(\mathbf{X}^T\mathbf{X})^{-1}\underline{\mathbf{x}}_0^T < 1$.

b) Considere a la entrada h_{ii} de la matriz $n \times n$ definida como: $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, se tiene que $\sum_{i=1}^{n} h_{ii}$ es igual al número de covariables en el modelo.

c) En un modelo de regresión suponga que $2p > n$ y que $h_{33} > \frac{2p}{n}$, luego la observación 3 es un punto de balanceo.

d) Una observación es influencial si $|\text{DFFITS}_i| > 2\sqrt{\frac{k}{n}}$.

a) Verdadero

$h_{00} = \underline{X}_0(\mathbb{X}^T\mathbb{X})^{-1}\underline{X}_0^T < \max\{h_{ii}\}$, si eso se cumple no hay extrapolación.

Como no se comete extrapolación, se tiene que

$$h_{00} < \max\{h_{ii}\} \leq 1 \Rightarrow h_{00} < 1$$

¿Por qué el $\max\{h_{ii}\} \leq 1$? Para todo i se cumple que $h_{ii} \leq 1$, en particular se cumple para el máximo

¿Por qué $h_{ii} \leq 1$? H es idempotente, vamos a calcular la entrada ii de H^2, esto es, el producto punto ente la i-ésima fila de H y la i-ésima columna de H.

$$\begin{bmatrix} h_{i1} & h_{i2} & \cdots & h_{in} \end{bmatrix} \begin{bmatrix} h_{1i} \\ h_{2i} \\ \vdots \\ h_{ni} \end{bmatrix} = h_{i1} \cdot h_{1i} + h_{i2} \cdot h_{2i} + \cdots + h_{in} \cdot h_{ni}$$

$$= h_{i1}^2 + h_{i2}^2 + \cdots + h_{in}^2 = h_{ii}$$

$$h_{ii} = h_{i1}^2 + h_{i2}^2 + \cdots + h_{in}^2$$

$$= \underset{\geq 0}{h_{i1}^2} + \underset{\geq 0}{h_{i2}^2} + \cdots + \underset{\geq 0}{h_{i,i-1}^2} + h_{ii}^2 + \underset{\geq 0}{h_{i,i+1}^2} + \cdots + \underset{\geq 0}{h_{in}^2} \geq h_{ii}^2$$

$\Rightarrow h_{ii} \geq h_{ii}^2 \Rightarrow 1 \geq h_{ii}$ ó $h_{ii} \leq 1$

\# β's en el modelo

\# parámetros

b) $\bar{h} = \dfrac{\sum_{1\leq i\leq n} h_{ii}}{n} = \dfrac{p}{n} \Rightarrow \sum_{1\leq i\leq n} h_{ii} = p$

$p = k+1 \qquad \Rightarrow \sum h_{ii} = p \neq K$

↳ \# Covariables

Falso

c) $2p > n \Rightarrow \dfrac{2p}{n} > 1 \Rightarrow h_{33} > \dfrac{2p}{n} > 1 \Rightarrow h_{33} > 1$ ¿?

d)

Según el criterio es p y no K

$$|\mathrm{DFFITS}_i| > 2\sqrt{\frac{k}{n}}.$$

Un punto es influencial si $|\mathrm{DFFITS}_i| > 2\sqrt{\dfrac{p}{n}}$

$\beta_0 = \beta_2,$ $\qquad\qquad\qquad \cancel{\beta_1}^{\to 0} + \beta_2 = \beta_0 + \cancel{\beta_3}^{\to 0}$

$\beta_3 = 0$

$\beta_1 = 0$

$$\begin{cases} H_0: \beta_0 = \beta_2,\ \beta_3 = 0,\ \beta_1 = \beta_3,\ \beta_1 + \cancel{\beta_2} = \cancel{\beta_0} + \cancel{\beta_3} \\ H_1: \text{Al menos una de las igualdades no se cumple} \end{cases}$$

Expresemos H_0 de forma matricial

$$H_0: \underbrace{\begin{array}{cccc} \beta_0 & \beta_1 & \beta_2 & \beta_3 \\ \end{array}}_{}$$

$$H_0: \underbrace{\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}}_{\underline{\beta}} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\underline{0}}$$

$\beta_0 - \beta_2 = 0 \Rightarrow \beta_0 = \beta_2$

Vamos a escribir el modelo reducido

Se escalona $L \Rightarrow$ $\begin{array}{cccc} \beta_0 & \beta_1 & \beta_2 & \beta_3 \end{array}$ $\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ y con L escalonada

Se plantea el modelo reducido

$$Y_i = \underbrace{\beta_0}_{\beta_2} + \cancel{\beta_1}^{\to 0} X_{1i} + \beta_2 X_{2i} + \cancel{\beta_3}^{\to 0} X_{3i} + \varepsilon_i, \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$Y_i = \beta_2 \underbrace{(1 + X_{2i})}_{X_{2i}^*} + \varepsilon_i, \quad \varepsilon_i \sim^{iid} (N 0, \sigma^2)$$

$$F_0 = \frac{[SSE(MR) - SSE(MF)]/3}{\underbrace{MSE(MF)}_{\frac{SSE(MF)}{n-p}}} \sim F_{3,n-p}$$