Assignment 4

Question 1: a) $Y = \dfrac{\alpha x}{\gamma + x}$ → reciprocate each side.

$$\frac{1}{Y} = \frac{1}{\frac{\alpha x}{\gamma + x}}$$

$$\frac{1}{Y} = 1 \times \frac{\gamma + x}{\alpha x}$$

$$\frac{1}{Y} = \frac{\gamma + x}{\alpha x}$$

$$\frac{1}{Y} = \frac{\gamma}{\alpha x} + \frac{x}{\alpha x}$$

$$\frac{1}{Y} = \frac{\gamma}{\alpha x} + \frac{1}{\alpha}$$

$$\frac{1}{Y} = \frac{\gamma}{\alpha} \times \frac{1}{x} + \frac{1}{\alpha}$$

$$Y^* = \frac{1}{Y}, \quad \beta_0 = \frac{1}{\alpha}, \quad \beta_1 = \frac{\gamma}{\alpha}, \quad x^* = \frac{1}{x}$$

$$Y^* = \beta_0 + \beta_1 x^* \ \#$$

b) Given that $\beta_0 = \frac{1}{\alpha}$ and $\beta_1 = \frac{\gamma}{\alpha}$

Then, $\hat{\beta}_0 = \frac{1}{\hat{\alpha}}$ and $\hat{\beta}_1 = \frac{\hat{\gamma}}{\hat{\alpha}}$

$$\hat{\alpha} = \frac{1}{\hat{\beta}_0} \ \#$$

$$\hat{\beta}_1 \hat{\alpha} = \hat{\gamma}$$

$$\frac{\hat{\beta}_1}{\hat{\beta}_0} = \hat{\gamma}$$

$$\hat{\gamma} = \frac{\hat{\beta}_1}{\hat{\beta}_0}$$

c) Linearised model

$$Q(\beta_0, \beta_1) = \sum_{i=1}^{n} \left(y_i^* - (\beta_0 + \beta_1 x_i^*) \right)^2$$

minimising the Linearised model would give us a pretty nice solution for estimates $\hat{\beta}_0$ and $\hat{\beta}_1$.
And given that $\hat{\alpha} = \frac{1}{\hat{\beta}_1}$ and $\hat{\gamma} = \frac{\hat{\beta}_1}{\hat{\beta}_0}$, we can back transform to get $\hat{\alpha}$ and $\hat{\gamma}$.

Saturated Model

$$Q(\alpha, \gamma) = \sum_{i=1}^{n} \left(y_i - \left(\frac{\alpha x}{\gamma + x} \right) \right)^2$$

→ minimising the saturated Model would not give us any nice solution.

Also, $\boxed{Y_i = \dfrac{\alpha x_i}{\gamma + x_i} + \epsilon_i}$, $\epsilon_i \sim N(0, \sigma^2)$ and $\boxed{Y_i^* = \beta_0 + \beta_1 x_i^* + \epsilon_i^*}$, $\epsilon_i^* \sim N(0, \sigma^2)$

Back transforming $\epsilon_i^* = \frac{1}{\epsilon_i}$ from the linearised model would yield a different error term than what was originally in the saturated model.

$\epsilon_i = \frac{1}{\epsilon_i^*}$

$\epsilon_i^* \neq \epsilon_i$

d) i) we know that degrees of freedom is $n-k$ where k is the number of parameters being estimated.

In our case $k=2$, β_0 and β_1

$n-2 = 13$

$n = 13+2$

$n = 15$ #

ii)

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{15} \end{bmatrix} \quad x = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_{15} \end{bmatrix}$$

intercept

$$\underbrace{y = \begin{bmatrix} 489 \\ 476 \\ 513 \\ 382 \end{bmatrix}}_{\text{length}} \quad x = \underbrace{\begin{bmatrix} 1 & 4 \\ 1 & 3 \\ 1 & 7 \\ 1 & 1 \end{bmatrix}}_{\text{Age}}$$

iii) $\hat{\alpha} = \dfrac{1}{\hat{\beta}_0}$, $\hat{\gamma} = \dfrac{\hat{\beta}_1}{\hat{\beta}_0}$ from part b)

$y = \hat{\beta}_0 + \hat{\beta}_1 x \rightarrow$ given $\hat{\beta}_0 = 6.075 \times 10^{-3}$ ☆

To find $\hat{\beta}_1$ we use the test statistic of $\hat{\beta}_1$:

$$t_{\hat{\beta}_1} = \frac{\hat{\beta}_1 - \beta_1}{\frac{S_E}{\sqrt{S_{xx}}}}$$

$$8.423 = \frac{\hat{\beta}_1 - 0}{2.897 \times 10^{-4}}$$

$$(8.423)(2.897 \times 10^{-4}) = \hat{\beta}_1$$

$$\hat{\beta}_1 = 0.0024401 \text{ ☆}$$

$$\hat{\alpha} = \frac{1}{6.075 \times 10^{-3}} \qquad \hat{\gamma} = \frac{0.0024401}{6.075 \times 10^{-3}}$$

$$\hat{\alpha} = 164.6091 \qquad \hat{\gamma} = 0.40166$$

iv) $CI = \hat{\beta}_0 \pm t_{n-2,\frac{\alpha}{2}} \left(Se\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{Sxx}} \right)$

= estimate ± (critical val)(standard error)

$\hat{\beta}_0 = 6.075 \times 10^{-3}$

$t_{n-2,\frac{\alpha}{2}} = t_{15-2,\frac{0.10}{2}}$

$= t_{13,0.05}$

= (R) qt(0.05, 13, lower.tail = FALSE)

$= 1.7709$

$Se\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{Sxx}} = 9.825 \times 10^{-5}$

$CI = 6.075 \times 10^{-3} \pm (1.7709)(9.825 \times 10^{-5})$

Lower limit = 0.00590

Upper limit = 0.00624

CI = (0.00590, 0.00624) #

v) Prediction Interval $= \hat{y} \pm \left(t_{n-2,\frac{\alpha}{2}}\right)\left(Se\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{Sxx}}\right)$

$= \hat{y} \pm (t^*)(SE)$

Given $\hat{y}$ or the predicted value "fit" = 0.006481724

lower bound, L = 0.006023172

$L = \hat{y} - (t^*)(SE)$

$L + (t^*)(SE) = \hat{y}$

$t^*SE = \hat{y} - L$

$= 0.006481724 - 0.006023172$

$= 0.000458552$

sub t^*SE into U

Upper bound, $U = \hat{y} + (t^*)(SE)$

$= 0.006481724 + 0.000458552$

$= 0.006940276$

Linearised Model 90% Prediction Interval = (0.006023172, 0.006940276)

back transform to find Y from the linearised Model Y^*.

Find $Y^* = \frac{1}{Y}$

$Y = \frac{1}{Y^*}$

Lower $= \frac{1}{0.006940276}$ $= 144.0865$

Upper $= \frac{1}{0.006023172}$ $= 166.02548$

The model $Y = \frac{\alpha x}{\gamma + x}$ has a 90% Prediction Interval between 144.09 cm and 166.03 cm.

Question 2:

a) Assuming that the columns of x^* are dependent.

Then there exist a constant $\underline{\alpha} \in \mathbb{R}^n$ where $\underline{\alpha} \neq 0$.

$$x = \underbrace{\begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{np} \end{bmatrix}}_{n \times p} \quad A = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}}_{p \times p}$$

$$\underline{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_j \end{bmatrix}$$

$$x^* = \underbrace{\begin{bmatrix} x_{11}a_{11} + x_{12}a_{21} + \cdots + x_{1p}a_{p1}, & \cdots, & x_{11}a_{1p} + x_{12}a_{2p} + \cdots + x_{1p}a_{pp} \\ \vdots & \vdots & \vdots \\ x_{n1}a_{11} + x_{n2}a_{21} + \cdots + x_{1p}a_{p1}, & \cdots, & x_{n1}a_{1p} + x_{n2}a_{2p} + \cdots + x_{np}a_{pp} \end{bmatrix}}_{n \times p}$$

$$\underline{\alpha}\, x^* = \begin{bmatrix} \alpha_1 x_{11}a_{11} + \alpha_1 x_{12}a_{21} + \cdots + \alpha_1 x_{1p}a_{p1}, & \cdots, & \alpha_1 x_{11}a_{1p} + \alpha_1 x_{12}a_{2p} + \cdots + \alpha_1 x_{1p}a_{pp} \\ \vdots & \vdots & \vdots \\ \alpha_j x_{n1}a_{11} + \alpha_j x_{n2}a_{21} + \cdots + \alpha_j x_{1p}a_{p1}, & \cdots, & \alpha_j x_{11}a_{1p} + \alpha_j x_{12}a_{2p} + \cdots + \alpha_j x_{1p}a_{pp} \end{bmatrix}$$

If we take $\underline{\alpha}\, x A = x(\underline{\alpha} A) \rightarrow (\underline{\alpha} A) = \underline{k}$

$= x\, \underline{k}$ → given that the columns of x are linearly independent then $\underline{k}$ must be 0. But this cannot happen due the fact that matrix A is invertible meaning $\det(A) \neq 0$ and given that $\underline{\alpha}$ cannot be 0 as well, this means that $\underline{k}$ cannot be 0 and therefore our assumption that the columns of x^* are dependent is contradicted.

By contradiction, the columns of x^* are linearly independent.

b) $$\begin{aligned} X^*(X^{*T}X^*)^{-1}X^{*T} &= XA((XA)^T XA)^{-1}(XA)^T \\ &= XA\left((XA)^{-1}((XA)^T)^{-1}\right)A^T X^T \\ &= XA\left(A^{-1}X^{-1}(A^T X^T)^{-1}\right)A^T X^T \\ &= XA\left(A^{-1}X^{-1}(X^T)^{-1}(A^T)^{-1}A^T X^T\right. \\ &= X(AA^{-1})X^{-1}(X^T)^{-1}(A^T)^{-1}A^T X^T \\ &= X I X^{-1}(X^T)^{-1}(A^{-1})^T A^T X^T \\ &= X I X^{-1}(X^T)^{-1}(AA^{-1})^T X^T \\ &= X I (X^T X)^{-1} I X^T \\ &= X(X^T X)^{-1}X^T \;\# \end{aligned}$$

c) $\hat{n} = \hat{n}^*$

Given that $\hat{\beta} = (X^T X)^{-1} X^T Y$

$$\begin{aligned} X\hat{\beta} &= X^*\hat{\beta}^* \\ &= \left[X^*(X^{*T}X^*)^{-1}X^{*T}\right]Y \rightarrow \text{using the property from b)} \\ &= X(X^T X)^{-1}X^T Y \\ &= X\hat{\beta} \;\# \end{aligned}$$

Therefore, $\hat{n} = \hat{n}^*$ because $X^*\hat{\beta}^* = X\hat{\beta}$ #