Tres vehículos se encuentran situados en los puntos $0 < \beta_1 < \beta_2 < \beta_3$ de una carretera recta. Para estimar la posición de los vehículos se toman las siguientes medidas (todas ellas sujetas a errores aleatorios de medición independientes con distribución normal de media 0 y varianza σ^2):

- Desde el punto 0 medimos las distancias a los tres vehículos, denotándolas Y_1, Y_2 e Y_3.
- Nos trasladamos al primer vehículo y medimos las distancias a los otros dos, dando dos nuevas medidas Y_4 e Y_5.
- Nos trasladamos al segundo vehículo y medimos la distancia al tercero, dando una medida adicional, Y_6.

a) Expresa el problema de estimación como un modelo de regresión múltiple indicando claramente cuál es la matriz de diseño.
b) Calcula la distribución del estimador de mínimos cuadrados del vector de posiciones $(\beta_1, \beta_2, \beta_3)'$.
c) Se desea calcular un intervalo de confianza de nivel 95 % para la posición del primer vehículo β_1 a partir de 6 medidas (obtenidas de acuerdo con el método descrito anteriormente) para las que la varianza residual resultó ser $s_R^2 = 2$. ¿Cuál es el margen de error del intervalo?

Solución:

a)

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_6 \end{pmatrix} \sim N_6\left(\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_2-\beta_1 \\ \beta_3-\beta_1 \\ \beta_3-\beta_2 \end{pmatrix}, \sigma^2\mathbf{I}_6\right) = N_6\left(\mathbb{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_6\right) = \mathbb{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

donde $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)'$ es el vector de parámetros,

$$\mathbb{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

es la matriz del diseño y $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_6)' \sim N_6(\mathbf{0}_6, \mathbf{I}_6)$.
b)

$$\mathbb{X}'\mathbb{X} = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \Rightarrow (\mathbb{X}'\mathbb{X})^{-1} = \frac{1}{4}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

$$\hat{\boldsymbol{\beta}} = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbf{Y} = \frac{1}{4}\begin{pmatrix} 2Y_1 + Y_2 + Y_3 - Y_4 - Y_5 \\ Y_1 + 2Y_2 + Y_3 + Y_4 - Y_6 \\ Y_1 + Y_2 + 2Y_3 + Y_5 + Y_6 \end{pmatrix} \sim N_3(\boldsymbol{\beta}, \sigma^2(\mathbb{X}'\mathbb{X})^{-1})$$

c) $\hat{\beta}_1 \sim N(\beta_1, \sigma^2 q_{11})$ con $(\mathbb{X}'\mathbb{X})^{-1} = (q_{ij})_{i,j=1}^3$. El error típico de $\hat{\beta}_1$ es s.e.$(\hat{\beta}_1) = s_R^2\, q_{11} = 1$. Se cumple que $t(\beta_1) = \hat{\beta}_1/\text{s.e.}(\hat{\beta}_1) \sim t_3$, pues los g.l. de los residuos son $n-$número de parámetros a estimar $= 6 - 3 = 3$. Por tanto,

$$\text{IC}_{0.95}(\beta_1) = (\hat{\beta}_1 \mp t_{3;0.025}) = (\hat{\beta}_1 \mp 3.182).$$