A6: Linear Models – Question #1

1. $y_i = b_0 + b_1 z_{i1} + b_2 z_{i2} + e_i$, $z_{ij} = x_{ij} - m_j$, $\sum_i z_{ij} = 0$ this is the model & we just centered it

$SSE(b_0, b_1, b_2) = \sum_{i=1}^{N} (y_i - b_0 - b_1 z_{i1} - b_2 z_{i2})^2$

2. Let $e_i = y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}$

$$\frac{\partial SSE}{\partial b_0} = -2\sum_i e_i = 0, \quad \frac{\partial SSE}{\partial b_1} = -2\sum_i z_{i1} e_i = 0, \quad \frac{\partial SSE}{\partial b_2} = -2\sum_i z_{i2} e_i = 0$$

3. From $\sum_i e_i = 0 \Rightarrow \bar{e} = 0$

From the others: $\sum_i e_i z_{i2} = 0$

So, the residuals will be <u>orthogonal</u> to each centered predictor

4. Using $\sum_i e_i = 0$ and $\sum_i z_{ij} = 0$:

$$0 = \sum_i (y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}) = N\bar{y} - Nb_0 \Rightarrow b_0^* = \bar{y}$$

Let $\tilde{y}_i := y_i - \bar{y}$ and proceed with b_1, b_2

5. Normal Equations $Ab = c$ (for b_1, b_2)

$$\begin{bmatrix} \sum z_{i1}^2 & \sum z_{i1} z_{i2} \\ \sum z_{i1} z_{i2} & \sum z_{i2}^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum z_{i1} \tilde{y}_i \\ \sum z_{i2} \tilde{y}_i \end{bmatrix}$$

6. Divide by N and substitute $z_{ij} = x_{ij} - m_j$

$$\frac{1}{N} A = \begin{bmatrix} Var(x_1) & Cov(x_1, x_2) \\ Cov(x_1, x_2) & Var(x_2) \end{bmatrix}, \quad \frac{1}{N} c = \begin{bmatrix} Cov(x_1, y) \\ Cov(x_2, y) \end{bmatrix}$$

Intuition: OLS sets $\Sigma_{xx} b = \Sigma_{xy}$, so each slope reflects its contribution to explaining y after accounting for predictor inter-correlation; the intercept is $\bar{y}$ because predictors are centered.