A6

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

2. $\frac{\partial e_i}{\partial b_0} = -1, \quad \frac{\partial e_i}{\partial b_1} = -z_{i1}, \quad \frac{\partial e_i}{\partial b_2} = -z_{i2}$ $\qquad \frac{\partial SSE}{\partial b_0} = -2\sum_{i=1}^{N} e_i \overset{!}{=} 0 \Rightarrow \sum_{i=1}^{N} e_i = 0$

$\frac{\partial SSE}{\partial b_1} = -2\sum_{i=1}^{N} e_i z_{i1} \overset{!}{=} 0 \Rightarrow \sum_{i=1}^{N} e_i z_{i1} = 0$ $\qquad \frac{\partial SSE}{\partial b_2} = -2\sum_{i=1}^{N} e_i z_{i2} \overset{!}{=} 0 \Rightarrow \sum_{i=1}^{N} e_i z_{i2} = 0$

3. $\sum e_i = 0 \Rightarrow \frac{1}{N}\sum^{N} e_i = \bar{e} = 0$, avg error is 0

$\sum^{N} e_i z_{i1} = 0, \sum^{N} e_i z_{i2} = 0$, residuals are ortho. to each predictor

4. $0 = \sum_{i=1}^{N} (y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}) = \sum y_i - N b_0 - b_1 \sum z_{i1} - b_2 \sum z_{i2}$

$0 = \sum y_i - N b_0 \Rightarrow b_0^* = \frac{1}{N}\sum y_i = \bar{y}$

$\sum (y_i - \bar{y}) z_{i1} - b_1 \sum z_{i1}^2 - b_2 \sum z_{i1} z_{i2} = 0, \quad \sum (y_i - \bar{y}) z_{i2} - b_1 \sum z_{i1} z_{i2} - b_2 \sum z_{i2}^2 = 0$

5. $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad A = \begin{bmatrix} \sum z_{i1}^2 & \sum z_{i1} z_{i2} \\ \sum z_{i1} z_{i2} & \sum z_{i2}^2 \end{bmatrix},$

$$\begin{bmatrix} \sum z_{i1}^2 & \sum z_{i1} z_{i2} \\ \sum z_{i1} z_{i2} & \sum z_{i2}^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum (y_i - \bar{y}) z_{i1} \\ \sum (y_i - \bar{y}) z_{i2} \end{bmatrix} = \begin{bmatrix} \sum y_i z_{i1} \\ \sum y_i z_{i2} \end{bmatrix} = C$$

6. $$\begin{bmatrix} \frac{1}{N}\sum z_{i1}^2 & \frac{1}{N}\sum z_{i1} z_{i2} \\ \frac{1}{N}\sum z_{i1} z_{i2} & \frac{1}{N}\sum z_{i2}^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{N}\sum (y_i - \bar{y}) z_{i1} \\ \frac{1}{N}\sum (y_i - \bar{y}) z_{i2} \end{bmatrix}$$

$\frac{1}{N}\sum z_{i1}^2 = Var(X_1), \frac{1}{N}\sum z_{i2}^2 = Var(X_2), \frac{1}{N}\sum z_{i1} z_{i2} = Cov(X_1, X_2), \frac{1}{N}\sum (y_i - \bar{y}) z_{i1} = Cov(X_1, Y),$

$\frac{1}{N}\sum (y_i - \bar{y}) z_{i2} = Cov(X_2, Y)$, $\begin{bmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_1, X_2) & Var(X_2) \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} Cov(X_1, Y) \\ Cov(X_2, Y) \end{bmatrix}$ w/ $b_0^* = \bar{y}$

A is the covariance matrix of the predictors Σ_{XX}. C is the vector of covariances between each predictor & Y. So after scaling, $A = \Sigma_{XX}$ & $C = \Sigma_{XY}$ so $b = (\Sigma_{XX})^{-1} \Sigma_{XY}$ & $b_0 = \bar{y}$