$-2\sum_{i=1}^{N} z_{i1}(y_i-b_0-b_1z_{i1}-b_2z_{i2})$

$(y_i-b_0-b_1z_{i1}-b_2z_{i2})$

DS 3001 A6

Question 1

$\hat{y}_i = b_0 + b_1 z_{i1} + b_2 z_{i2}$ $\quad z_{ij} = x_{ij} - m_j$ $\quad \frac{1}{N}\sum_{i=1}^{N} z_{ij} = \emptyset$

1.1 $SSE = \sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2$

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

1.2

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

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

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

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

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

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

1.3 $e_i = y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}$

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

$$\boxed{\sum_{i=1}^{N} e_i = \emptyset} \quad \boxed{\frac{1}{N}\sum_{i=1}^{N} e_i = \emptyset}$$

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

$$\boxed{\sum_{i=1}^{N} z_{i1} e_i = \emptyset} \qquad \boxed{\sum_{i=1}^{N} z_{i2} e_i = \emptyset}$$

1.4 $\frac{\partial SSE}{\partial b_0} = -2\sum_{i=1}^{N}(y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}) = \emptyset$

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

*each z_{ij} is centered

$$\sum_{i=1}^{N} y_i - N b_0 = \emptyset \Longrightarrow b_0^* = \bar{y}$$

$$y_i - \bar{y} - b_1 z_{i1} - b_2 z_{i2}$$

$$\sum_{i=1}^{N} z_{i1}(\bar{y}_i - b_1 z_{i1} - b_2 z_{i2}) = 0$$

$$\sum_{i=1}^{N} z_{i2}(\bar{y}_i - b_1 z_{i1} - b_2 z_{i2}) = 0$$

1.5 $S_{yz_1} = \sum_i (y_i - \bar{y}) z_{i1}$, $S_{yz_2} = \sum_i (y_i - \bar{y}) z_{i2}$

$S_{z_1z_1} = \sum_i z_{i1}^2$, $S_{z_2z_2} = z_{i2}^2$, $S_{z_1z_2} = \sum_i z_{i1} z_{i2}$

$$\boxed{\underbrace{\begin{pmatrix} S_{z_1z_1} & S_{z_1z_2} \\ S_{z_1z_2} & S_{z_2z_2} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}}_{b} = \underbrace{\begin{pmatrix} S_{yz_1} \\ S_{yz_2} \end{pmatrix}}_{C}}$$

1.6

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

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

$z_{i1} = x_{i1} - m_1$, $z_{i2} = x_{i2} - m_2$

$$\frac{1}{N}\sum_i z_{i1}^2 = \frac{1}{N}\sum_i (x_{i1} - m_1)^2 = Var(x_1)$$

$$\frac{1}{N}\sum_i z_{i1} z_{i2} = \frac{1}{N}\sum_i (x_{i1} - m_1)(x_{i2} - m_2) = Cov(x_1, x_2)$$

$$\frac{1}{N}\sum_i (y_i - \bar{y}) z_{i1} = \frac{1}{N}\sum_i (y_i - \bar{y})(x_{i1} - m_1) = Cov(x_1, y)$$

$$A = \begin{pmatrix} Var(x_1) & Cov(x_1, x_2) \\ Cov(x_1, x_2) & Var(x_2) \end{pmatrix} \qquad C = \begin{pmatrix} Cov(x_1, y) \\ Cov(x_2, y) \end{pmatrix}$$

$$A\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = C$$

The slopes b1 and b2 measure how each predictor co-varies with the response while accounting for the other predictor. The matrix A captures the relationships among predictors, and the vector C captures the relationship between each predictor and y. We divide by N to take the average of the product. This gives us variance and covariance, which are the average of squared differences and the average of the product of the deviations of two variables from their respective means, respectively. it makes sense that dividing by N would give such values.