1. SSE

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

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

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

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

3. $$\frac{\partial SSE}{\partial b_0} = -2\sum_{i=1}^{N} (y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}) = 0$$

$$-2\sum_{i=1}^{N} e_i = 0 \Rightarrow \text{Avg err} = 0$$

$$\sum z_i e_i = 0 \qquad \Rightarrow e \cdot z = 0 \text{ at optimum}$$

4. $b_0^* = \bar{y}$

$$\sum_{i=1}^{N} (y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}) = 0 \quad \left(\sum z_{ij} = 0\right)$$

$$\Downarrow$$

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

$$b_0 = \bar{y}$$

5. $\hat{y}_i = \bar{y} + b_1 z_{i1} + b_2 z_{i2}$

Assuming $\frac{1}{N} \sum_{i=1}^{N} z_{ij} = 0$ for 1 and 2

$$Z = \begin{bmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \\ \vdots & \vdots \\ z_{1N} & z_{2N} \end{bmatrix} \qquad A$$

$$b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad b$$

$$y_i - \bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \qquad C = \|y - Zb\|^2$$

$$Z^T Z b = \|y - Zb\|^2$$

6. $\frac{1}{N} Z^T Z b = \frac{1}{N} \|y - Zb\|^2$

Where $A = \frac{1}{N} Z^T Z$ and $C = \frac{1}{N} \|y - Zb\|^2$

so $Ab = C$

$z_{ij} = x_{ij} - m_j$

A is a matrix of Var and Cov between preds
C is a vector of cov between pred. + outcome

So A gives how predictors relate to eachother
and C gives how each pred. relates to y