DS Linear Models Assignment

$$\hat{y}_i = b_0 + b_1 z_{i1} + b_2 z_{i2} \qquad z_{ij} = \underbrace{x_{ij}}_{\text{og. value of } j \text{ for observation } i} - \underbrace{\mu_j}_{\mu \text{ of } j}$$

$$\frac{1}{N}\sum_{i=1}^{N} z_{ij} = 0 \qquad e_i = y_i - \hat{y}_i = y_i - x_i \cdot b$$

1. SSE of model

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

2. Partial derivatives w/ respect to b_0, b_1 & b_2

$$\frac{\partial SSE}{\partial b_0} = -2\sum e_i = 0$$

$$\frac{\partial SSE}{\partial b_1} = -2\sum e_i z_{i1} = 0$$

$$\frac{\partial SSE}{\partial b_2} = -2\sum e_i z_{i2} = 0$$

3. verify avg. error is 0 & $e \cdot z = 0$ @ optimum (just like linear regression case)

$\sum_i e_i = 0$ so avg. error rate is 0

$\sum_i e_i z_{i1} = 0$, $\sum_i e_i z_{i2} = 0$ so residuals are orthogonal to each regressor (like OLS)

4. Show optimal intercept $b_0^* = \bar{y}$, eliminate b_0^* from remaining equations (focus on b_1 & b_2)

$\sum_i e_i = 0 \qquad \sum(y_i - b_0 - b_1 z_{i1} - b_2 z_{i2}) = \sum y_i - Nb_0 - b_1 \underbrace{\sum_i z_{i1}}_{=0} - b_2 \underbrace{\sum_i z_{i2}}_{=0} = 0$

$Nb_0 = \sum_i y_i \qquad b_0^* = \bar{y}$

$\tilde{y}_i = y_i - \bar{y}$ remaining FOCs:

$$\boxed{\sum_i (\tilde{y}_i - b_1 z_{i1} - b_2 z_{i2}) z_{i1} = 0, \quad \sum_i (\tilde{y}_i - b_1 z_{i1} - b_2 z_{i2}) z_{i2} = 0}$$

5. Matrix equation $Ab = C$ (normal equations)

$$\boxed{A = \begin{pmatrix} \sum z_{i1}^2 & \sum z_{i1} z_{i2} \\ \sum z_{i1} z_{i2} & \sum z_{i2}^2 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad C = \begin{pmatrix} \sum y_i z_{i1} \\ \sum y_i z_{i2} \end{pmatrix}}$$

6. Matrix A, vector C

$z_{ij} = x_{ij} - m_j$, w/ $m_j = \bar{x}_{\cdot j}$

$$\frac{1}{N} \sum z_{i1}^2 = Var(x_1), \quad \frac{1}{N} \sum z_{i2}^2 = Var(x_2), \quad \frac{1}{N} \sum z_{i1} z_{i2} = Cov(x_1, x_2)$$

$$\frac{1}{N} \sum y_i z_{i1} = Cov(x_1, y), \quad \frac{1}{N} \sum y_i z_{i2} = Cov(x_2, y)$$

$$\boxed{\begin{pmatrix} Var(x_1) & Cov(x_1, x_2) \\ Cov(x_1, x_2) & Var(x_2) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} Cov(x_1, y) \\ Cov(x_2, y) \end{pmatrix}}$$

OLS slopes take covariance between residuals & each residual and set them to equal 0. The slopes solve system of combining 2 regressors to the covariance of x_1, x_2 w/ y & [illegible]. If x_1 & x_2 are correlated, each slope cancels out from off-diagonal covariance terms.