Exercise C

Suppose the predictors Z are orthogonal, that is

$$Z^T Z = \operatorname{diag}(\tilde{z}_1^T \tilde{z}_1, \dots, \tilde{z}_p^T \tilde{z}_p) = \begin{bmatrix} \tilde{z}_1^T \tilde{z}_1 & & & \emptyset \\ & \ddots & \tilde{z}_j^T \tilde{z} & \vdots \\ \emptyset & & \ddots & \\ & \dots & & \tilde{z}_p^T \tilde{z}_p \end{bmatrix}$$

$$\Rightarrow (Z^T Z)^{-1} = \operatorname{diag}(1 / \tilde{z}_1^T \tilde{z}_1, \dots, 1 / \tilde{z}_p^T \tilde{z}_p)$$

$$\Rightarrow \hat{\beta} = (Z^T Z)^{-1} \underbrace{Z^T y} = \begin{bmatrix} (\tilde{z}_1^T \tilde{z}_1)^{-1} & & \emptyset \\ & \ddots & \\ \emptyset & & (\tilde{z}_p^T \tilde{z}_p)^{-1} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{m} \tilde{z}_{i1} y_i \\ \vdots \\ \sum_{i=1}^{m} \tilde{z}_{ip} y_i \end{bmatrix}$$

Hence, $\hat{\beta}_j = \dfrac{\tilde{Z}_j^T y}{\tilde{z}_j^T \tilde{z}_j}$.

Now, suppose $y \sim N(Z\beta, \sigma^2 I_m)$. Then, $\hat{\beta} \sim N_p(\beta, \sigma^2 (Z^T Z)^{-1})$, because $\hat{\beta} = (Z^T Z)^{-1} Z^T y$ is a linear transformation of a Gaussian random variable, then

$$\mathbb{E}[\hat{\beta}] = (Z^T Z)^{-1} Z^T \mathbb{E}[y] = \cancel{(Z^T Z)^{-1} Z^T Z}\, \beta = \beta$$

$$\mathrm{var}(\hat{\beta}) = (Z^T Z)^{-1} Z^T \underbrace{\mathrm{var}(y)}_{\sigma^2 I_p} \left[(Z^T Z)^{-1} Z^T \right]^T =$$

$$= \sigma^2 \cancel{(Z^T Z)}^{-1} \cancel{Z^T Z} (Z^T Z)^{-1}$$

$$= \sigma^2 \underbrace{(Z^T Z)}_{}{}^{-1} .$$

However, $(Z^T Z)^{-1}$ is diagonal and hence the covariances are zero. However, in the Gaussian case uncorrelation $\iff$ independence.