Problem 1

Since X is orthonormal, $XX^T = I$, $\hat{\beta^{ols}} = (X^TX)^{-1}X^Ty = X^Ty$.

$$\begin{aligned}(y - X\beta)^T(y - X\beta) &= (X^Ty)^TX^Ty + \beta^TX^TX\beta - 2\beta^TX^Ty \\ &= (\hat{\beta}^{ols})^T\hat{\beta}^{ols} + \beta^T\beta - 2\beta^T\hat{\beta}^{ols} \\ &= \sum_{j=1}^{p}(\beta_j - \hat{\beta}_j^{ols})^2 + (\beta_0 - \hat{\beta}_0^{ols})^2\end{aligned}$$

So we can optimize β componentwise.

Problem 2

See attachment.