## Exercise F

Consider the pathwise algorithm for the elastic-net. At each step we compute $\hat{\beta}_k$ as

$$\frac{1}{1+(1-\alpha)\lambda} S_{\alpha\lambda}\left(\frac{1}{m}\sum_{i=1}^{m} z_{ik}\, r_i^{(j)}\right), \text{ with } r_i^{(j)} = y_i - \sum_{k\neq j} z_{ik}\beta_k$$

However, under orthogonal predictors, we get

$$\sum_{i=1}^{m} z_{ij}\, r_i^{(j)} = \sum_{i=1}^{m} z_{ij} y_i - \underbrace{\sum_{k\neq j}\sum_{i=1}^{m} z_{ij} z_{ik}\, \beta_k}_{= 0\,!} = \sum_{i=1}^{m} z_{ij} y_i ,$$

which does not depend on the other values of $\beta$. Hence we get convergence in one step and

$$\hat{\beta}_j = \frac{1}{1+(1-\alpha)\lambda} S_{\alpha\lambda}\left(\frac{1}{m}\sum_{i=1}^{m} z_{ij} y_i\right)$$

$$= \frac{1}{1+(1-\alpha)\lambda} S_{\alpha\lambda}\left(\mathrm{cov}(z,y)\right).$$

This covers both lasso and ridge as special cases.