$Y = B\beta + \varepsilon$ $\quad$ B has orthonormal cols.

"classic" objective : $\Psi(\beta) = \|y - B\beta\|_2^2 \quad \beta \in \mathbb{R}^p$.

"Ridge" obj. : $\bar{\Psi}(\beta) = \|y - B\beta\|_2^2 + \lambda\|\beta\|_2^2$

$$= \|\bar{y} - \bar{B}\beta\|_2^2.$$

And $\bar{B} = \begin{bmatrix} B \\ \sqrt{\lambda} I_p \end{bmatrix}$; $\bar{y} = \begin{bmatrix} y \\ 0_p \end{bmatrix}$.

Recall that LSE for β optimizes $\Psi(\beta)$ and the optimum $\quad \arg\max \Psi(\beta) = \underbrace{(B'B)}_{I_k}^{-1} B'y = B'y$.

We can also directly find optimum $\bar{\Psi}(\beta)$.

$$\arg\max \bar{\Psi}(\beta) = (\bar{B}'\bar{B})^{-1}\bar{B}'\bar{y}.$$

- $\bar{B}'\bar{B} = [B' \quad \sqrt{\lambda} I_p]\begin{bmatrix} B \\ \sqrt{\lambda} I_p \end{bmatrix} = B'B + \lambda I_p = I_p + \lambda I_p = \begin{pmatrix} 1+\lambda & & \\ & \ddots & \\ & & 1+\lambda \end{pmatrix}$
- $(\bar{B}'\bar{B})^{-1} = \frac{1}{1+\lambda} I_p$.
- $(\bar{B}'\bar{B})^{-1}\bar{B}' = \frac{1}{1+\lambda} I_p [B' + \sqrt{\lambda} I_p] = \frac{1}{1+\lambda}[B' + \sqrt{\lambda} I_p]$
- $(\bar{B}'\bar{B})^{-1}\bar{B}'\bar{y} = \frac{1}{1+\lambda}[B' + \sqrt{\lambda} I_p]\begin{bmatrix} y \\ 0_p \end{bmatrix} = \frac{1}{1+\lambda} B'y$.

hence $\arg\max \bar{\Psi}(\beta) = \frac{1}{1+\lambda} B'y = \frac{1}{1+\lambda} \arg\max \Psi(\beta)$

and $c(\lambda) = \frac{1}{1+\lambda}$