5. a. $\tilde{X} = \begin{bmatrix} X \\ \sqrt{\lambda} I_p \end{bmatrix}$ $\quad \tilde{y} = \begin{bmatrix} y \\ 0_{p\times 1} \end{bmatrix}$

$$\hat{\tilde{\beta}}_{OLS} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y}$$

$$\tilde{X}^T \tilde{X} = [X^T \ \sqrt{\lambda} I_p] \begin{bmatrix} X \\ \sqrt{\lambda} I_p \end{bmatrix} = X^T X + \lambda I_p$$

$$\tilde{X}^T \tilde{y} = [X^T \ \sqrt{\lambda} I_p] \begin{bmatrix} Y \\ 0_{p\times 1} \end{bmatrix} = X^T y + 0 = X^T y$$

$$\boxed{\hat{\tilde{\beta}}_{OLS} = (X^T X + \lambda I_p)^{-1} X^T y = \hat{\beta}^R}$$

b. $\hat{\beta}^R = \tilde{Q}\left(\tilde{\Sigma}^T\tilde{\Sigma} + \lambda I_p\right)^{-1}\tilde{\Sigma}^T\tilde{P}^T y$

Is just a diagnal matrix so transpose is just itself

$\|\hat{\beta}^R\|_2 = \hat{\beta}^{R^T}\hat{\beta}^R$

$$= y^T\tilde{P}\tilde{\Sigma}\left(\tilde{\Sigma}^T\tilde{\Sigma}+\lambda I_p\right)^{-T}\overbrace{\tilde{Q}^T\tilde{Q}}^{I_p}\left(\tilde{\Sigma}^T\tilde{\Sigma}+\lambda I_p\right)^{-1}\tilde{\Sigma}^T\tilde{P}^T y$$

$$= y^T\tilde{P}\underbrace{\tilde{\Sigma}\left(\tilde{\Sigma}^T\tilde{\Sigma}+\lambda I_p\right)^{-2}\tilde{\Sigma}^T}_{r\times r \text{ diagnal matrix}}\tilde{P}^T y$$

$$\begin{array}{cc} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} & \begin{array}{l} r \\ N-r \end{array} \\ \begin{array}{cc} r & p-r \end{array} & \end{array} \begin{array}{cc} \begin{bmatrix} (\Sigma^2+\lambda I_r)^{-2} & 0 \\ 0 & \frac{1}{\lambda^2} I_{p-r} \end{bmatrix} & \begin{array}{l} r \\ N-r \end{array} \\ \begin{array}{cc} r & p-r \end{array} & \end{array} \begin{array}{cc} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} & \begin{array}{l} r \\ N-r \end{array} \\ \begin{array}{cc} r & p-r \end{array} & \end{array} = \Sigma\left(\Sigma^2+\lambda I_p\right)^{-2}\Sigma$$

diagnal matrix with (j,j) element $\frac{\sigma_j^2}{(\sigma_j^2+\lambda)^2}$

$$\|\hat{\beta}^R\|_2 = \sum_{j=1}^{r}\frac{(\tilde{P}^T y)_j^2\,\sigma_j^2}{(\sigma_j^2+\lambda)^2}$$

each term is proportional to $\frac{\sigma_j^2}{(\sigma_j^2+\lambda)^2}$ thus as λ increases we know each term is decreasing since a larger λ means a larger denominator so similarly as $\lambda\to\infty$ then each term goes to 0 thus $\|\hat{\beta}^R\|_2 \to 0$