(a)

Let $L(\beta) = (y - X\beta)^T (y - X\beta) + \lambda \|\beta\|^2$

$$\frac{\partial L(\beta)}{\partial \beta} = 2X^T X\beta - 2X^T y + 2\lambda\beta = 0$$

$$(X^T X + \lambda I)\beta = X^T y$$

$$\hat{\beta}_\lambda = (X^T X + \lambda I)^{-1} X^T y$$

(b)

$$E\|\hat{\beta}_\lambda - \beta\|^2$$

$$= \|bias\|^2 + trace(var(\hat{\beta}_\lambda))$$

Note that

$$E\hat{\beta}_\lambda = (X^T X + \lambda I)^{-1} X^T \cdot X\beta$$

$$\text{var}\,\hat{\beta}_\lambda = (X^T X + \lambda I)^{-1} X^T X (X^T X + \lambda I)^{-1} \sigma^2$$

bias

$$= E\hat{\beta}_\lambda - \beta$$

$$= (X^T X + \lambda I)^{-1} X^T X \beta - \beta$$

$$= ((X^T X + \lambda I)^{-1} X^T X - I)\beta$$

(c)

$$E \| \hat{y}_\lambda - X\beta \|^2 = \| \text{bais} \|^2 + \text{trace}(\text{var}(\hat{y}_\lambda))$$

$$E \hat{y}_\lambda = X E \hat{\beta}_\lambda = X (X^T X + \lambda I)^{-1} X^T X \beta.$$

$$\text{bias} = E \hat{y}_\lambda - y_\lambda$$

$$= X (X^T X + \lambda I)^{-1} X^T X \beta - X\beta$$

$$= X[(X^T X + \lambda I)^{-1} X^T X - I]\beta$$

$$\text{var}(\hat{y}_\lambda) = \text{var}(X \hat{\beta}_\lambda)$$

$$= \text{var}(X (X^T X + \lambda I)^{-1} X^T y)$$

$$= X (X^T X + \lambda I)^{-1} X^T X (X^T X + \lambda I)^{-1} X^T \sigma^2.$$

(d)

$$X = \begin{bmatrix} -50 \\ -49 \\ \vdots \\ 50 \end{bmatrix}$$

$$X^T X = 2 \sum_{i=1}^{50} i^2$$

Denote $a = 2 \sum_{i=1}^{50} i^2 = 85850$

$$E\hat{\beta}_\lambda - \beta = \left((a+\lambda)^{-1} a - 1\right)\beta$$

$$= \frac{-\beta\lambda}{a+\lambda}$$

$$\operatorname{var} \hat{\beta}_\lambda = (a+\lambda)^{-1} a (a+\lambda)^{-1} = \frac{a}{(a+\lambda)^2}$$

So MSE is (b) vs

$$MSE(b) = \frac{\lambda^2}{(a+\lambda)^2}\beta^2 + \frac{a}{(a+\lambda)^2} = \frac{\lambda^2\beta^2 + a}{(a+\lambda)^2}$$

For c),

$$\text{bias} = X\left(\frac{a}{a+\lambda} - 1\right)\beta = X\frac{-\lambda\beta}{a+\lambda}$$

$$\|\text{bias}\|^2 = a \cdot \frac{\lambda^2\beta^2}{(a+\lambda)^2}$$

$$\text{var}(\hat{y}_\lambda) = XX^T(a+\lambda)^{-2}a$$

$$\text{trace}(\text{var}(\hat{y}_\lambda)) = a^2(a+\lambda)^{-2}.$$

$$MSE(c) = \frac{a^2}{(a+\lambda)^2} + \frac{a\lambda^2\beta^2}{(a+\lambda)^2}$$

$$= \frac{a(a+\lambda^2\beta^2)}{(a+\lambda)^2}$$

$$\mathrm{MSE}(b) = \frac{a+\lambda^2\beta^2}{(a+\lambda)^2} = \frac{4\lambda^2+a}{(\lambda+a)^2}$$

let $\dfrac{\partial \mathrm{MSE}}{\partial \lambda} = 0$

$$\frac{8\lambda(\lambda+a)^2 - 2(\lambda+a)(4\lambda^2+a)}{(\lambda+a)^4} = 0$$

$$8\lambda(\lambda+a) - 2(4\lambda^2+a) = 0$$

$$8\lambda a - 2a = 0$$

$$\lambda = \frac{1}{4}$$

$MSE(c) = a\,MSE(b)$.

So we get the same result.

$\lambda = \frac{1}{4}$.