Mettes notes on
Var(LS) > Var(Ridge)

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

and for calc "simplified" define: $W_\lambda = (X^TX + \lambda I)^{-1} X^TX$

because

$$\underset{\substack{\uparrow \\ LS}}{W_\lambda \hat{\beta}} = (X^TX + \lambda I)^{-1} \underbrace{X^TX \; (X^TX)^{-1}} X^TY = (X^TX + \lambda)^{-1} X^TY = \hat{\beta}_{ridge}$$

$$Var(\hat{\beta}_{ridge}) = Var(W_\lambda \hat{\beta}) = W_\lambda \underbrace{Var(\hat{\beta})}_{\sigma^2 (X^TX)^{-1}} W_\lambda^T$$

$$(X^TX + \lambda I)^{-1} \cancel{X^TX} \; \sigma^2 \cancel{(X^TX)^{-1}} \; (X^TX)(X^TX + \lambda X)^{-1}$$

$$= \sigma^2 \underbrace{(X^TX + \lambda I)^{-1} (X^TX)}_{W_\lambda} (X^TX + \lambda X)^{-1}$$

$$Var(\hat{\beta}) - Var(\hat{\beta}(\lambda)) = \sigma^2 \left[(X^TX)^{-1} - W_\lambda (X^TX)^{-1} W_\lambda^T \right]$$

Tricks:

$$W_\lambda (I + \lambda (X^TX)^{-1}) (X^TX)^{-1} (I + \lambda (X^TX)^{-1}) W_\lambda$$

$$\overbrace{(X^TX + \lambda I)^{-1} X^TX (I + \lambda (X^TX)^{-1})} (X^TX)^{-1} \overbrace{(I + \lambda (X^TX)^{-1}) X^TX (X^TX + \lambda I)^{-1}}$$

$$\underbrace{(X^TX + \lambda I)^{-1} (X^TX + \lambda I)}_{I} \; (X^TX)^{-1} \; \underbrace{(X^TX + \lambda I)(X^TX + \lambda I)^{-1}}_{I}$$

$$Var(\hat{\beta}) - Var(\hat{\beta}(\lambda)) =$$

$$\sigma^2 W_\lambda \left(\underbrace{(I + \lambda (X^TX)^{-1})(X^TX)^{-1}} (I + \lambda (X^TX)^{-1}) - (X^TX)^{-1} \right) W_\lambda^T$$

$$\sigma^2 W_\lambda \left(\underbrace{[(X^TX)^{-1} + \lambda (X^TX)^{-2}] (I + \lambda (X^TX)^{-1})} - (X^TX)^{-1} \right) W_\lambda^T$$

$$\sigma^2 W_\lambda \left(\left((X^TX)^{-1} + \lambda (X^TX)^{-2} + \lambda (X^TX)^{-2} + \lambda^2 (X^TX)^{-3} \right) - (X^TX)^{-1} \right) W_\lambda^T$$

$$= \underline{\sigma^2 W_\lambda \left(2\lambda (X^TX)^{-2} + \lambda^2 (X^TX)^{-3} \right) W_\lambda^T}$$ possible to stop here?

$$= \sigma^2 \left[(X^TX + \lambda I)^{-1} X^TX \left(2\lambda (X^TX)^{-2} + \lambda^2 (X^TX)^{-3} \right) X^TX (X^TX + \lambda I)^{-1} \right]$$

$$= \sigma^2 \left[(X^TX + \lambda I)^{-1} [2\lambda I + \lambda^2 (X^TX)^{-1}] (X^TX + \lambda I)^{-1} \right]$$

All matrices are symmetric and semi-positive definite

If (X^TX) is semi positive definite

if (X^TX) is spd then $(X^TX + \lambda I)$ is also spd $\quad \lambda \geq 0$

and $(X^TX)^{-1}$ is spd

and $(\lambda^2 (X^TX)^{-1} + 2\lambda I)$ is spd

$$\overset{P\Lambda P^T}{\downarrow}$$

$$a^T (X^TX) a \geq 0 \qquad a^T \overset{P\Lambda^{-1}P^T}{(X^TX)^{-1}} a = a^{*T} \overset{\text{diag}}{\Lambda^{-1}} a^* \geq 0$$

$$a^T (X^TX + \lambda I) a = a^T X^TX a + a^T \lambda a \geq 0$$

$$a^T (\lambda^2 (X^TX)^{-1} + 2\lambda I) a = \underset{\geq 0}{a \lambda^2 (X^TX)^{-1} a} + \underset{\geq 0}{a\, 2\lambda I a} \geq 0$$

This is the product of symmetric semi positive definite matrices also symmetric and semi positive definite?

$A \cdot B \cdot C$

$a^T A a \geqslant 0$

Yes, according to Wieringa p 12.