Week 6 From lec 9:

$\hat{\beta}_{ridge} = W\hat{\beta}_{ols}, \quad W = \left(I + \lambda(X^TX)^{-1}\right)^{-1}$

$Var[\hat{\beta}_{ridge}] = W\,Var[\hat{\beta}_{ols}]\,W^T = \sigma^2 W (X^TX)^{-1} W$

$Var[\hat{\beta}_{ols}] - Var[\hat{\beta}_{ridge}] = \sigma^2\left((X^TX)^{-1} - W(X^TX)^{-1}W^T\right)$

$= \sigma^2 W\left[W^{-1}(X^TX)^{-1}(W^T)^{-1} - (X^TX)^{-1}\right]W^T$

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

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

$= \sigma^2 W\left[2\lambda(X^TX)^{-2} + \lambda^2(X^TX)^{-3}\right]W^T \geq 0$

SPD: $z^TAz \geq 0$

$A^T = A$

$\sum_i \sum_j x_i^T x_i x_j^T x_j z_i z_j \geq 0$

feb. 25-08:18

ITLS

6.1)

a) iv

b) ii

c) $E[(\hat{y}_0 - y)^2] = (E[\hat{y}_0] - y)^2 + Var[\hat{y}_0] + Var[\varepsilon]$

iii

d) iv

e) v

feb. 25-08:34