ISYE 8803 HDDA

Homework 4

July 13th, 2021

1.

Based on the following properties we know:

$$\|y - XB\|_2^2 = (y - XB)^T (y - XB) = \left(y^T - B^T X^T\right)(y - XB) = y^T y - 2B^T X^T y$$

$X^T X = I$ where I is the identity matrix

$\hat{B}^{ols} = X^T y$

a)

Therefore to prove that:

$\widehat{\beta}^{ols} = X^T y$

We start with the following above property:

$= y^T y - 2\beta^T X^T y + \beta^T X^T X \beta$

$X^T X$ *simplies to the identity matrix and when multiplied by the* $\beta = 1$

Which simplifies to:

$= y^T y - 2\beta^T X^T y + \beta^T \beta$

Then we take the derivative w.r.t β *and set to* $= 0$ *to get the following:*

$$\frac{df}{d\beta} = 0$$

$-2X^T y + 2\beta = 0$

$\widehat{\beta}^{ols} = X^T y$

b)

$min(\beta) = ||y - x\beta||_2^2 + \lambda||\beta||_2^2$

$When\ building\ off\ of\ the\ previous\ answer\ but\ adding\ the\ lambda\ term\ for\ the$
$L2\ norm\ we\ get\ the\ following:$

$$-2X^Ty\ +2(\lambda+1)\widehat{\beta}^{\ ols}\ =\ 0$$

$Then\ now\ we\ solve\ for\ \widehat{\beta}^{\ ols}\ and\ get\ the\ following:$

$$\widehat{\beta}^{\ ols}\ =\ \frac{-2X^Ty}{2(\lambda+1)}$$

$$\widehat{\beta}^{\ ols}\ =\ \frac{-X^Ty}{(\lambda+1)}$$

$$\widehat{\beta}^{\ ols}\ =\ (\lambda+1)^{-1}X^Ty$$

$c)$

$$min(\beta) =\ ||y-x\beta||_2^2 + \lambda||\beta||_1$$

$When\ \ building\ off\ of\ the\ previous\ answer\ but\ adding\ the\ lambda\ term\ for\ the\ L1$
$norm\ we\ get\ the\ fol$

$$y^Ty-2\beta^TX^Ty\ +\beta^T\beta+\lambda||\beta||_1$$

$Now\ when\ taking\ the\ derivative\ and\ setting\ \ =\ 0\ we\ get\ the\ following:$

$$-2X^Ty\beta\ +\beta^T\beta+\lambda||\beta||_1 = 0$$

$We\ can\ see\ from\ 1a\ that\ \widehat{\beta}_{OLS}\ =\ \ X^Ty\ and\ \beta^T\beta\ \ = ||\beta||^2\ which\ simplies\ to\ the\ following:$

$$-2(\widehat{\beta}_{OLS})^T\beta\ + ||\beta||^2+\lambda||\beta||_1 = 0$$

$Which\ simplies\ to\ the\ equation\ derived\ in\ the\ hint\ of\ the\ question\ when\ factoring\ out$
$the\ summation\ for\ p:$

$$min(\beta_j)\ \beta_j^2-2(\widehat{\beta}_{jOLS})\beta_j\ \ +\lambda|\beta_j| = 0$$

$Now\ based\ on\ our\ formula\ we\ can\ derive\ the\ edge\ cases\ and\ show\ that\ Lasso\ regression\ is$

$a\ closed\ form\ solution:$

$$min(\beta_j)\ \beta_j^2 - 2(\widehat{\beta}_{jOLS})\beta_j\ + \lambda|\beta_j| + \lambda\ \cdot\ sign(\beta_j)\cdot\beta_j$$

When taking the derivate of the equation and solving for $\widehat{\beta}_{jOLS}$ we get the following:

$$2(\widehat{\beta}_{jLasso} - 2\widehat{\beta}_{jOLS}\ +)\lambda = 0$$

We can then use this equation for each inequality for $\frac{\lambda}{2}$ and solve for $\widehat{\beta}_{jLasso}$

$$2(\widehat{\beta}_{jLasso} - 2\widehat{\beta}_{jOLS}\ +)\lambda = 0$$

$$\widehat{\beta}_{jLasso} = \widehat{\beta}_{jOLS} - \frac{\lambda}{2}$$

We can reformulate the constraints to the following in order to simplify the calculation and formulate the table showing the closed form solution and possible outcomes:

<table>
<tr><td colspan="2">$0\ <\ \widehat{\beta}_{jOLS}\ \leq\ \frac{\lambda}{2}$</td><td>$\widehat{\beta}_{jOLS} > \frac{\lambda}{2}$</td></tr>
<tr><td>$\widehat{\beta}_{jOLS}\ > 0$</td><td>$\widehat{\beta}_{jOLS}\ <\ 0$</td><td rowspan="2">$\widehat{\beta}_{jLasso} = \widehat{\beta}_{jOLS} - \frac{\lambda}{2}$
Yields a positive amount</td></tr>
<tr><td rowspan="2">$\widehat{\beta}_{jOLS}\ > 0$
$= \widehat{\beta}_{jOLS} - \frac{\lambda}{2}$
Yield a negative amount but not less than $-\widehat{\beta}_{jOLS} - \frac{\lambda}{2}$</td><td rowspan="2">$\widehat{\beta}_{jLasso} = \widehat{\beta}_{jOLS} - \frac{\lambda}{2}$
$= -\widehat{\beta}_{jOLS} - \frac{\lambda}{2}$</td></tr>
<tr><td>$\widehat{\beta}_{jOLS}\ = 0$
$\widehat{\beta}_{jLasso} = -\frac{\lambda}{2}$</td></tr>
</table>