STAT380: Assignment 2

Vivienne Crowe ID:40071153

4/2/21

1

1.1

A.5
First note that for ordinary least squares:

$$\hat{\beta} = (\mathbf{X^T X})^{\mathbf{-1}}\mathbf{X y}$$

Since $\mathbf{X}$ is square and full rank, it is therefore inveritiable, and so $(\mathbf{X^T X})^{\mathbf{-1}} = \mathbf{X^{-1}}(\mathbf{X^T})^{\mathbf{-1}}$. So it follows that:

$$\hat{\beta} = \mathbf{X^{-1}}$$

Now I'll show that $\mathbf{X^{-1}} = \mathbf{V\Sigma^{-1}U^T}$.

$$\begin{aligned}
\mathbf{X}(\mathbf{V\Sigma^{-1}U^T}) &= (\mathbf{V\Sigma^{-1}U^T})(\mathbf{U\Sigma V^T}) \\
&= (\mathbf{V\Sigma^{-1}\Sigma V^T}) \\
&= (\mathbf{V V^T}) \\
&= \mathbf{I}
\end{aligned}$$

Now onto the solution for ridge regression, the problem statement, in vector notation is:

$$\begin{aligned}
&\min_{\beta} \quad (\mathbf{y} - \mathbf{X}\beta)^{\mathbf{T}}(\mathbf{y} - \mathbf{X}\beta) - \lambda\beta^{\mathbf{T}}\beta \\
&\min_{\beta} \quad \mathbf{y^T y} - \mathbf{2y^T X}\beta + \beta^{\mathbf{T}}\mathbf{X^T X}\beta) - \lambda\beta^{\mathbf{T}}\beta
\end{aligned}$$

Now to solve for where the derivative is equal to zero.

$$\begin{aligned}
0 &= \frac{\partial}{\partial\beta}\mathbf{y^T y} - \mathbf{2y^T X}\beta + \beta^{\mathbf{T}}\mathbf{X^T X}\beta - \lambda\beta^{\mathbf{T}}\beta \\
0 &= -2\mathbf{y^T X} + (\mathbf{X^T X} + \mathbf{X^T X})\beta - \mathbf{2}\lambda\beta \\
\mathbf{y^T X} &= (\mathbf{X^T X} + \lambda\mathbf{I})\beta \\
\implies \hat{\beta} &= (\mathbf{X^T X} + \lambda\mathbf{I})^{\mathbf{-1}}\mathbf{y^T X}
\end{aligned}$$