Programming Assignment #6. Problem #1

$$X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -3 & 2 \\ 1 & 2 & -1 \end{bmatrix}; \quad y = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

$$X^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -3 & 2 \\ 2 & 2 & -1 \end{bmatrix}; \quad X^T X = \begin{bmatrix} 3 & 0 & 3 \\ 0 & 14 & -6 \\ 3 & -6 & 9 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 5/8 & -1/8 & -7/24 \\ -1/8 & 1/8 & 1/8 \\ -7/24 & 1/8 & 7/24 \end{bmatrix}; \quad X^T y = \begin{bmatrix} 2 \\ -7 \\ 7 \end{bmatrix}$$

$$\hat{\beta} = (X^T X)^{-1}(X^T y) = \begin{bmatrix} 1/12 \\ -1/4 \\ 7/12 \end{bmatrix}$$

Problem 2 (a)

Model Matrix: $X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 1 & -3 & -6 \end{bmatrix}$ Observations $Y = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$

Suppose error $\epsilon = Y - X\beta$, where β is param vector.

Then Sum of Sq $L = \sum_{i=1}^{n} \epsilon_i^2 = (y - X\beta)^T(y - X\beta)$

The lease squares estimator $\hat{\beta}$ is the solution for β in $\frac{\partial L}{\partial \beta} = 0$. Minimizing $\frac{\partial L}{\partial \beta}$, ~~we get~~ we have normal equation.

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

$$X^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \\ 2 & 4 & -6 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 14 & 28 \\ 0 & 28 & 56 \end{bmatrix} \quad X^T Y = \begin{bmatrix} 0 \\ 14 \\ 28 \end{bmatrix}$$

$(X^T X)^{-1}$ is singular, $(X^T X)$ inverse doesn't exist!!

Problem 2 (b) Ridge Reg. Goal $\min_{\vec{\beta}} \|\vec{y} - A\vec{\beta}\|_2^2$ s.t. $\|\vec{\beta}\|_2 \le C^2$

$$\min_{\vec{\beta}} \sum_i (y_i - \vec{X}_i \vec{\beta})^2 \text{ s.t. } \underbrace{\beta_0^2 + \beta_1^2 + \beta_2^2}_{E} \le C^2$$

Lagrange Function $F(\beta_0, \beta_1, \lambda) = \sum_i (y_i - \vec{X}_i \vec{\beta})^2 + \lambda(\beta_0^2 + \beta_1^2 + \beta_2^2 - C^2)$

$\boxed{\min F \quad \beta_0, \beta_1, \beta_2, \lambda} \longrightarrow \boxed{\text{TRY Different } \lambda} \longrightarrow \boxed{\text{FIND BEST } \beta_0, \beta_1}$

$$\boxed{\min_{\vec{\beta}} \|\vec{y} - A\vec{\beta}\|_2^2 + \lambda \|\vec{\beta}\|_2^2}$$

Ridge Solution $\vec{\beta}^R = (X^T X + \lambda I)^{-1} X^T \vec{y}$

λ	$\vec{\beta}^R$	$\vec{\beta}$
10^{-5}	$\begin{bmatrix} 0 \\ 0.2 \\ 0.4 \end{bmatrix}$	
-10^{-5}	$\begin{bmatrix} 0 \\ 0.2 \\ 0.4 \end{bmatrix}$	

See second PDF