Roman Muñoz
CS - 4033

Problem 1.

1)

$$J(\beta) = \sum_{i=1}^{n} w_i \cdot (x_i^T \beta - y_i)^2$$

Let X be a $n \times p$ matrix with x_i^T on the ith row and Y be an n-dimensional vector with y_i in its ith element.
Then Let W be a diagonal matrix of w_ith terms.

$$\Rightarrow \| X^T \beta - Y \|^2 = (X^T \beta - Y)^T (X^T \beta - Y)$$

$$J(\beta) = W \cdot \| X^T \beta - Y \|^2 = W \cdot (X^T \beta - Y)^T (X^T \beta - Y)$$

2)

$$\frac{\partial J(\beta)}{\partial \beta} = (X^T \beta - Y)^T \cdot (W + W^T) \cdot \frac{\partial}{\partial \beta} (X^T \beta - Y)$$

$$0 = (X^T \beta - Y) \, W \cdot X^T$$

$$0 = X^T W^T X \beta - X^T W Y$$

$$X^T W^T X \beta = X^T W Y$$

$$\beta = (X^T W X)^{-1} (X^T W Y)$$

3)

Choice of Weights	$w_l = 1, w_h = 1$	$w_l = 1, w_h = 10$	$w_l = 1$ $w_h = 50$	$w_l = 1, w_h = 0.1$
Error on all testing instances	0.139 ± 0.007	0.167 ± 0.004	0.224 ± 0.007	0.142 ± 0.008
Error on high crime rate group	0.309 ± 0.039	0.259 ± 0.048	0.250 ± 0.054	0.378 ± 0.036
Error on low crime rate group	0.122 ± 0.005	0.161 ± 0.005	0.222 ± 0.009	0.119 ± 0.004
Avg. low crime test instances	1428	1428.9	1429.7	1430
Avg. high crime test instances	66	65.1	64.3	64

\# Instances in the High Crime Group = $64.85 \approx 65$
\# Instances in the Low Crime Group = $1429.15 \approx 1429$