

Q2.2 Part B

Generalizing to weighted setting $\sum_{i=1}^n \frac{1}{2} (y_i - x_i^T \theta)^2$

$$J(\theta) = (X\theta - y)^T W (X\theta - y)$$

$$\frac{\partial J}{\partial \theta} = \frac{\partial (X\theta - y)^T}{\partial \theta} \cdot \frac{\partial (X\theta - y)}{\partial \theta} = 2X^T W (X\theta - y) = 0$$

$$\Rightarrow X^T W X \theta - X^T W y = 0$$

$$\theta = (X^T W X)^{-1} (X^T W y)$$

#

$$\sum_{i=1}^n x_i x_i^T w_i = X^T W X$$

$$\sum_{i=1}^n (y_i - x_i^T \theta) x_i w_i = (y - X\theta)^T W (y - X\theta)$$

$$\sum_{i=1}^n (y_i - x_i^T \theta) w_i = 0 \Rightarrow \bar{y} = \bar{y}$$

$$\sum_{i=1}^n (y_i - x_i^T \theta) w_i = 0$$

By setting $w_i = \frac{1}{2}$, then we arrive at (a)
 (ii) $w_i = \frac{1}{2}$ is a valid diagonal matrix, $w_i = \frac{1}{2}$

Part 2

Let us consider the following

$$(y - X\theta)^T (y - X\theta) = (y - X\theta)^T (y - X\theta)$$

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

$$0 = y^T y - \theta^T X^T X \theta - 2y^T X \theta$$

$$y^T X \theta = \theta^T X^T y$$

$$y^T X \theta = \theta^T X^T y$$