5.a) show that $J(\theta) = \frac{1}{2} \sum_{i=1}^{n} . w^{(i)} (\theta^T . x^{(i)} - y^{(i)})^2$ can be written as $J(\theta) = (X.\theta - \vec{y})^T . W (X.\theta - \vec{y})$

$$J(\theta) = 1/2 . \sum_{i=1}^{n} w^{(i)} . (\theta^T . x^{(i)} - y^{(i)})^2$$

$$= \sum_{i=1}^{n} (\theta^T . x^{(i)} - y^{(i)}) (\tfrac{1}{2} w^{(i)}) . (\theta^T . x^{(i)} - y^{(i)})$$

$$= \sum_{i=1}^{n} \underbrace{(X.\theta - \vec{y})}_{\text{it's a n dim. vector}}{}_i^T . (1/2 . w^{(i)}) . (X.\theta - \vec{y})_i$$

let $W_{ij} = \begin{cases} 1/2\, w^{(i)} , & \text{if } i = j \\ 0 , & \text{otherwise} \end{cases}$

then:

$$J(\theta) = \sum_{i,j=1}^{n} (X.\theta - \vec{y})_i^T . W_{ij} . (X\theta - \vec{y})_i$$

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

ii) when all $w^{(i)}$'s equal to 1, we have the normal eq: $X^T X . \theta = X^T$ and $\theta = (X^T . X)^{-1} . X^T . \vec{y}$ minimizes $J(\theta)$.

Generalize the normal eq. to weighted setting and give θ that minimizes $J(\theta)$.

$$\frac{\partial J(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta} (\theta^T . X^T . W . X . \theta - 2 . \vec{y}^T . W . X . \theta + \vec{y}^T . W . \vec{y})$$

$$= 2 X^T . W . X . \theta - 2 . X^T . W^T . \vec{y} + 0$$

taking $\frac{\partial J(\theta)}{\partial \theta} = 0$ we find $\theta = (X^T . W . X)^{-1} . X^T . W . \vec{y}$