Ex 2.6

Given $RSS(\Theta) = \sum_{i=1}^{n} (y_i - f_\theta(x_i))^2$

We will assume there are m repeated values of x. Then, without loss of generality we may reorder the RSS and seperate into two parts:

$$(*) \quad RSS(\Theta) = \underbrace{\sum_{i=1}^{m} (y_i - f_\theta(x_m))^2}_{\text{Repeated } x \text{ labels}} + \underbrace{\sum_{i=m+1}^{N} (y_i - f_\theta(x_i))^2}_{\text{Remaining } x\text{'s}}$$

We fill focus on LHS:

$$\sum_{i=1}^{m} (y_i^2 - 2f_\theta(x_m)y_i + f_\theta(x_m)^2)$$

$$= -2 f_\theta(x_m) \sum_i y_i + m f_\theta(x_m)^2 + \sum_i y_i^2$$

$$= m(-2f_\theta(x_m) \tfrac{1}{m}\sum y_i + f_\theta(x_m)^2) + \sum_i y_i^2$$

Note: $\frac{1}{m}\sum_{1}^{n} y_i = \bar{y}_m$

$$= m(-2f_\theta(x_m)\bar{y}_m + f_\theta(x_m)^2) + \sum y_i^2$$

$$= m(\bar{y}_m - f_\theta(x_m))^2 \underbrace{- m\bar{y}_m + \sum y_i^2}_{\text{Notice this is constant wrt } \theta}$$

$\Rightarrow$ Minimising $(*)$ wrt. θ is equivalent to minimising.

$$m(\bar{y}_m - f_\theta(x_m))^2 + \sum_{i=m+1}^{N} (y_i - f_\theta(x_i))^2$$

$$\Rightarrow RSS(\Theta) = \sum_{i=1}^{N-m+1} w_i(\bar{y}_i - f_\theta(x_i))^2$$

Where $W_i =$ the number of times observation i is repeated in the data.

More generally, for multiple repeated x values by induction we obtain:

$$RSS(\theta) = \sum_{i=1}^{N_u} W_i \left(\bar{y}_i - f_\theta(x_i) \right)^2$$

Where N_u is the number of uniques x's in the data and W_i represents the number of times observation x_i was repeated.