Problem 2:

a)

$$f(w)=\frac{1}{2n}\sum_{i=1}^{n}\left(y^{(i)}-\sum_{j=1}^{m+1} w_j x_j^{(i)}\right)^2$$

$$\frac{\partial f}{\partial w_j}=\frac{1}{\cancel{2}n}\sum_{i=1}^{n}\cancel{2}\left(y^{(i)}-\sum_{j=1}^{m+1} w_j x_j^{(i)}\right)\cdot\left(-x_j^{(i)}\right)$$

$$\frac{\partial f}{\partial w_j}=-\frac{1}{n}\sum_{i=1}^{n}\left(y^{(i)}-\sum_{j=1}^{m+1} w_j x_j^{(i)}\right)\cdot x_j^{(i)}$$

$$w_j := w_j - \alpha\cdot\frac{\partial f}{\partial w_j}$$ α = learning rate constant

$$w_j := w_j + \alpha\cdot\frac{1}{n}\sum_{i=1}^{n}\left(y^{(i)}-\sum_{j=1}^{m+1} w_j x_j^{(i)}\right)\cdot x_j^{(i)}$$

b)

$$f(w)=\frac{1}{2n}(y-Xw)^T(y-Xw)$$

$$(y-Xw)^T\cdot(y-Xw)=y^Ty-y^TXw-(Xw)^Ty+(Xw)^TXw$$

$$=y^Ty-2w^TX^Ty+w^TX^TXw$$

$$f(w)=\frac{1}{2n}(y^Ty-2w^TX^Ty+w^TX^TXw),$$ y^Ty does not depend on

$-\frac{1}{n}w^TX^Ty$ represents linear term w

$\frac{1}{2n}w^TX^TXw$ represents quadratic term w

$\therefore$ The two forms are represented the same c)

$$(y - Xw)^T(y - Xw) = y^Ty - y^TXw - (Xw)^Ty + (Xw)^TXw$$

$$(AB)^T = B^TA^T \rightarrow -y^TXw = -(Xw)^Ty = -X^Tw^Ty$$

$$= y^Ty - 2X^Tw^Ty + X^Tw^TXw$$

$$f(w) = \frac{1}{2n}y^Ty - \frac{\cancel{2}^{\,1}}{\cancel{2}n}X^Tw^Ty + \frac{1}{2n}X^Tw^TXw$$

$$\frac{\partial f}{\partial w_j} = \overset{0}{\cancel{\frac{\partial}{\partial w_j}(y^Ty)}} - \frac{\partial}{\partial w_j}\left(\frac{1}{n}X^Tw^Ty\right) + \frac{\partial}{\partial w_j}\left(\frac{1}{2n}X^Tw^TXw\right)$$

$$-\frac{\partial}{\partial w_j}\left(\frac{1}{n}X^Tw^Ty\right) = -\frac{1}{n}\cdot X^Ty \qquad \frac{\partial}{\partial w_j}\left(\frac{1}{2n}X^Tw^TXw\right) = \frac{\cancel{2}^{\,1}}{\cancel{2}n}\cdot X^T\cdot w\cdot X$$

$$\frac{\partial f}{\partial w_j} = \frac{1}{n}\left(X^TwX - X^Ty\right)$$

Problem 5:

a)

$$SSE = \sum_{i=1}^{n}\sum_{j=1}^{k} w^{(i,j)}\left\|x^{(i)} - \mu^{(j)}\right\|_2^2$$

$$\left\|x^{(i)} - \mu^{(j)}\right\|_2^2 = \left(x^{(i)} - \mu^{(j)}\right)^T\left(x^{(i)} - \mu^{(j)}\right) = x_i^Tx_i - 2\mu_j^Tx_i + \mu_j^T\mu_j$$

$$= k_{ii} - 2k_{ij} + k_{jj}$$

$$SSE = \sum_{i=1}^{n}\sum_{j=1}^{k} w^{(i,j)}\left(k_{ii} - 2k_{ij} + k_{jj}\right)$$

$$SSE = \sum_{i=1}^{n}\sum_{j=1}^{k} w^{(i,j)}\left(k_{ii} - \frac{2}{|j|}\sum_{l=1}^{n} w^{(l,j)}k_{il} + \frac{1}{|j|^2}\sum_{l=1}^{n}\sum_{q=1}^{n} w^{(l,j)}w^{(q,j)}k_{lq}\right)$$