Mert AGCAKOYUN

1) $MSE = \frac{1}{m} \sum_{i=1}^{m} (Q^T \cdot x^{(i)} - y^{(i)})^2$

Gradient of MSE functions:

$$\frac{\partial MSE}{\partial Q} = \frac{1}{m} \sum_{i=1}^{m} 2 \cdot (Q^T \cdot x^{(i)} - y^{(i)}) \cdot x^{(i)}$$

$$\frac{2}{m} \sum_{i=1}^{m} (Q^T \cdot x^{(i)} - y^{(i)}) \cdot x_0^{(i)}$$

Assume that $x_0^{(i)} = 1$;

$$\frac{\partial MSE}{\partial Q} = \frac{2}{m} \sum_{i=1}^{m} (Q^T \cdot (x^{(i)} - y^{(i)})$$

2) $\nabla MSE(Q) = \begin{bmatrix} \frac{\partial MSE(Q_0)}{\partial Q_0} \\ \vdots \\ \frac{\partial MSE(Q_n)}{\partial Q_n} \end{bmatrix} =$

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (Q^T \cdot x^{(i)} - y^{(i)})^2$$

$$(MSE)' = \frac{2}{m} \sum_{i=1}^{m} (Q^T \cdot x^{(i)} - y^{(i)}) \cdot d_{Q^T} \cdot x_0^{(i)}$$

$$(MSE)' = \frac{2}{m} \cdot d_{Q^T} \sum_{i=1}^{m} (Q^T \cdot x^{(i)} - y^{(i)}) \cdot x_0^{(i)}$$

from Q_0 to Q_n:

$$d(MSE) = \frac{d(MSE)}{dQ_0} + \cdots + \frac{d(MSE)}{dQ_n}$$

$$= [dQ_0 \ dQ_1 \ \ldots, dQ_n] \cdot \left[\frac{d(MSE)}{dQ_0} \ \ldots \ \frac{d(MSE)}{dQ_n}\right]$$

gradiant vector