Quiz#1 of ESE 417 (09/12)

Consider an objective function $g(\boldsymbol{w}, x) = w_0 + w_1 x + w_2 x^2$, where $\boldsymbol{w} = [w_0 \quad w_1 \quad w_2]^T$ is the weight vector and x is the input.

(1) Find the gradient of this objective function with respect to the weight vector $\boldsymbol{w}$**.** Then evaluate the gradient at $x = 1.0$.

(2) The initial value of the weight vector is $\boldsymbol{w^0} = [0.5 \quad 0.5 \quad 0.5]^T$. Update the weight vector using the ***gradient descent method*** to find $\boldsymbol{w^1}$. Let the learning rate to be $\eta = 0.1$

(3) Evaluate the objective function with this new weight vector $\boldsymbol{w^1}$ at $x = 1.0$.

(1) $\nabla g(\vec{w}, x) = \begin{pmatrix} \frac{\partial g}{\partial w_0} \\ \frac{\partial g}{\partial w_1} \\ \frac{\partial g}{\partial w_2} \end{pmatrix} = \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix}$ $\quad \nabla g(\vec{w}, 1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

(2) $\vec{w}^0 = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$ Taylor Series of the objective function

$$g(\vec{w} + \Delta\vec{w}) \approx g(\vec{w}) + \nabla g(\vec{w}) \Delta\vec{w}$$

substitute $\Delta\vec{w}$ with $-\eta \nabla g(\vec{w})$, $\eta > 0$, constant

$$g(\vec{w} - \eta \nabla g(\vec{w})) \approx g(\vec{w}) - \eta \|\nabla g(\vec{w})\|_2^2$$

$\|\nabla g(\vec{w})\|_2^2 > 0$

$$g(\vec{w} - \eta \nabla g(\vec{w})) < g(\vec{w})$$

$$\vec{w}^1 = \vec{w}^0 - \eta \nabla g(\vec{w}^0)$$

$$\Rightarrow \vec{w}^1 = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix} - 0.1 \begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} = \begin{pmatrix} 0.5 - 0.1 \\ 0.5 - 0.1x \\ 0.5 - 0.1x^2 \end{pmatrix}$$

(3) $x = 1$

$$\vec{w}^1 = \begin{pmatrix} 0.5 - 0.1 \\ 0.5 - 0.1 \\ 0.5 - 0.1 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.4 \\ 0.4 \end{pmatrix}$$

$$g(\vec{w}^1, x) = w_0^1 + w_1^1 x + w_2^1 x^2 = 0.4 + 0.4 \times 1 + 0.4 \times 1 = 1.2$$