CS2180: Aritificial Intelligence Lab 6

Name & Roll. No:

February 15, 2019

Q1) Gradient Descent in 1-D: Consider the function to be $f(w) = \frac{1}{2}w^2$
a) Perform gradient descent to find the minimum of f_1. For $\alpha = 0.1$, plot the output of the algorithm at each step. [25 Marks]
b) Plot the output of the algorithm for $\alpha = -0.1$, $\alpha = 1$, $\alpha = 1.5$, $\alpha = 2$, $\alpha = 2.5$. [15 Marks]
c) Implement gradient descent with line search. [10 Marks]

Q2) Repeat previous question for a) $f(x) = \frac{1}{2}w^2 - 5w + 3$. [20 Marks]
b) $f(x) = \frac{1}{1+e^{-w}}$. [10 Marks]

Q3) Gradient Descent in 2D: Let $x \in \mathbb{R}^2$. Consider the functions $f_1(w) = w(1)^2 + w(2)^2 + 5w(1) - 3w(2) - 2$ and $f_2(w) = 10w(1)^2 + w(2)^2$
a) Show the gradient and contour plots for f_1 and f_2 [10 Marks]
b) Perform gradient descent to find the minimum of f_1 and f_2. [10 Marks]

The gradient descent procedure in 1-dimnension is given by

$$w_{t+1} = w_t - \alpha \frac{dL}{dw}|_{w=w_t} \tag{1}$$

The gradient in d-dimension is denoted by ∇L, and it is a function from $\mathbb{R}^d \rightarrow \mathbb{R}^d$, i.e., at any input point in $\mathbb{R}^d$, the gradient function output the direction of maximum change (the direction is a vector in $\mathbb{R}^d$). Thus at input $w_0 \in \mathbb{R}^d$, the gradient outputs
$\nabla L(w_0) = (\frac{\partial L}{\partial w(1)}|_{w(1)=w_0(1)}, \frac{\partial L}{\partial w(2)}|_{w(2)=w_0(2)}, \ldots, \frac{\partial L}{\partial w(d)}|_{w(d)=w_0(d)})$. The gradient descent procedure in d-dimension is given by

$$w_{t+1} = w_t - \alpha \nabla L(w_t), \tag{2}$$

which is same as

$$w_{t+1}(i) = w_t(i) - \alpha \frac{\partial L}{\partial w(i)}|_{w(i)=w_t(i)} \tag{3}$$