Assignment3_Solution

Consider the function $J(w) = w_1^2 + w_2^2 + 4w_1 - 6w_2 - 7$. Answer Questions (1-6):

1). What is the theoretical value of $min(J(w))$?

ans). $\frac{dJ}{dw_1} = 2w_1 + 4$

$\frac{dJ}{dw_2} = 2w_2 - 6$

$argmin_w(J(w_1)) \ : \ \frac{dJ}{dw_1} = 0 \Rightarrow w_1 =- 2$

$argmin_w(J(w_2)) \ : \ \frac{dJ}{dw_2} = 0 \Rightarrow w_2 = 3$

$min(J(w_1 =- 2, w_2 = 3) \ = \ (-2)^2 + 3^2 + 4(-2) - 6(3) - 7$

$= \ -20$

2). What is the theoretical value of the first component of $rgmin_w(J(w)$?

ans). $argmin_w(J(w_1)) \ : \ \frac{dJ}{dw_1} = 0$

$\Rightarrow w_1 =- 2$

3). What is the theoretical value of the first component of $rgmin_w(J(w)$?

ans). $argmin_w(J(w_2)) \ : \ \frac{dJ}{dw_2} = 0$

$\Rightarrow w_2 = 3$

4). Starting from an initial guess of $[w_1, w_2] = [5, 5]$, for a value of learning rate = 0.3, the value of w_1 at the 5^{th} iteration will be =_______ (Assume that the initial guess is the 1^{st} iteration)

ans). $w_1 =- 1.8208$

5). Starting from an initial guess of $[w_1, w_2] = [5, 5]$, for a value of learning rate = 0.3, the value of w_2 at the 5^{th} iteration will be =_______ (Assume that the initial guess is the 1^{st} iteration)

ans). $w_2 = 3.0512$

6). The absolute value of the difference between the values of J at the 5^{th} iteration and the 4^{th} iteration is ___________

ans). 0.1824

Matlab code for questions 4 to 6:

```
clc
clear
syms w1 w2             % defining weights w1 and w2 as system variables
J = @(w1,w2) w1.^2+w2.^2+4*w1-6*w2-7;   % defining cost function J
dJ_w1 =@(w1) 2*w1+4;        % defining grading of cost function with respect to w1
dJ_w2 =@(w2) 2*w2-6;        % defining grading of cost function with respect to w2
w1 = 5; w2 = 5;        % defining initial weights
alpha = 0.3;           % defining learning rate
W(1,:) = [w1,w2];      % assigning weights variables to a weight matrix
cost(1) = J(w1,w2);    % assigning cost with respect to weights w1 and w2 to a cost matrix
for i=2:5      % initial weights are taken as 1st iteration. Thus, running for 4 more iterations
```

```
        w1 = w1 - alpha*dJ_w1(w1);  % updating weight 1 using gradient descent method
        w2 = w2 - alpha*dJ_w2(w2);  % updating weight 2 using gradient descent method
        W(i,:) = [w1,w2];           % assigning these weights to weight matrix
        cost(i) = J(w1,w2);         % assigning cost to cost matrix
end
fprintf('At 5th iteration, w1 = %0.4f\n',W(5,1))
fprintf('At 5th iteration, w2 = %0.4f\n',W(5,2))
fprintf('Absolute difference b/w 5th and 4th cost is %0.4f\n', abs(cost(5)-cost(4)))
```

Based on the data provided below, answer questions from 7 to 10.

We consider a function we wish to minimize. $J(\mathbf{w}) = \frac{1}{10}\sum_i^5\left(y^{(i)} - w_1x^{(i)} - w_0\right)^2$ where the constants $x^{(i)}, y^{(i)}$ are provided in the table below

i	$x^{(i)}$	$y^{(i)}$
1	0	0.8822
2	0.25	1.2165
3	0.50	1.3171
4	0.75	1.7930
5	1.00	1.9826

7). The dimension of w is ________

ans). 2

8). If our initial guess of $w = [0,0]$, then for a learning rate of 1, the value of w_1 at the 5^{th} iteration is ________ (assume that the initial guess is the 1^{st} iteration.)

ans). $w_1 = 1.0506$

9). If our initial guess of $w = [0,0]$, then for a learning rate of 1, the value of w_2 at the 5^{th} iteration is ________ (assume that the initial guess is the 1^{st} iteration.)

ans). $w_2 = 0.7922$

10). If our initial guess of $w = [0,0]$, then for a learning rate of 1, the value of J at the 5^{th} iteration is ________ (assume that the initial guess is the 1^{st} iteration.)

ans). $w_1 = 0.0088$

Matlab code for questions 7 to 10:

```
clc
clear
x = [0;0.25;0.5;0.75;1];            % defining x vector
y = [0.8822;1.2165;1.3171;1.7930;1.9826];   % defining y vector
m = size(x,1);                      % m is the number of examples
X = [ones(m,1),x];                  % appending ones (bias) to input vector
w = zeros(size(X,2),1);             % defining weight matrix
fprintf('Dimension of w is \t %d\n', length(w))
alpha = 1;                  % defining learning rate
```

Assignment3_Solution

```
for i=2:5          % initial weights are taken as 1st iteration. Thus, running for 4 more iterations
        dJ_w = [1/m*sum((X*w-y).*X)]';          % gradient of cost with respect to weights
        w = w - alpha*dJ_w;                     % updating weights using gradient descent method
        J = 1/(2*m)*sum((y-X*w).^2); % cost with respect to updated weights
end
fprintf('At 5th iteration, w1 = %0.4f\n',w(1,1))
fprintf('At 5th iteration, w2 = %0.4f\n',w(2,1))
fprintf('cost at 5th iteration are %0.4f\t \n',J)
```