# **Gradient Descent**

Krishna Kumar Veeraputhiran

Grand Canyon University

DSC 540 – O500 – Machine Learning for Data Science

Dr. Aiman Darwiche

August 25,2021

# **Gradient Descent**

In this assignment we will discuss on the questions regarding gradient descent and finding the minima of the given function after two iterations.

Part 2

Find the minima of E(w) = 2w12+2w1w2+5w22, with starting point as w=[2,-2]T. Let us assume that the learning rate to be 0.1. The initial starting point of the function can be found out by substituting the value of w in E(w). So it can be written as below,

E(w) = 2(2)2 + 2(2)(-2) + 5(-2)2

* 8 -8 +20 => 20

The initial value of E(w) is 20. Now to find out the value of gradient descent at each iteration, we need to identify the differential values of E(w) with respect to both w1 and w2 respectively.

d E(w)/dw1 ( keeping w2 as constant) = 2\* 2 (w1) + 2 w2 -> **4w1 + 2 w2** -> *Eqn (1)* Note that the differential of a variable is 1 and differential of a constant is 0 and the differential of xn = n\*xn-1 . So in the above equation w2 is treated as constant. Applying the same rules to identify the differential of E(w) with respect to w2 keeping w1 as constant.

d E(w)/dw2 (keeping w1 as constant) = 2w1 + 2\* 5w2 -> **2w1 + 10 w2**-> *Eqn (2)*

Let us calculate the values at first two iterations for E(w) now. We know that the initial value of w1 = 2 and w2 = -2. According to Gopal (2019), the gradient descent equation can be given as,

w k+1 = wk – η\* dE/dw

Iteration 1 for w1, w11 = w1 – η d E(w)/ dw1 with initial value of w1 = 2 , w2 = -2 and η = 0.1 and using Eqn (1)

w11 = 2 – 0.1 \* (4 \* 2 + 2 (-2))

w11 = 2 – 0.1 \* (8-4)

w11 = 2- 0.4

**w11 = 1.6**

Similarly we can derive w2 using Eqn (2),

w21 = -2 – 0.1 \* (2\*2 + 10 (-2))

w21 = -2 – 0.1 \* (4-20)

w21 = -2 + 1.6

**w21 = -0.4**

Substituting the above values in E(w) we would get the first iteration value

E(w) = 2w112+2w11w21+5w212

= 2(1.6)2 + 2(1.6)(-0.4) + 5(-0.4)2

= 5.12 – 1.28 + 0.8

**E(w) = 4.64**

Now using the value of w11 and w21 we can calculate the value of E(w) for iteration 2.

w12 = w11 – η (4w11 + 2w21)

= 1.6 – 0.1 (4\*1.6 + 2\*(-0.4))

= 1.6 – 0.1 (6.4-0.8)

= 1.6-0.56

**w12 = 1.04**

w22 = w21 – η (2w11 + 10 w21)

= -0.4 -0.1 (2\*1.6 + 10\*(-0.4))

= -0.4 – 0.1 (-0.8)

= -0.4 + 0.08

**w22= -0.32**

Substituting w12 and w22 in E(w), we get the minima for the function post second iteration.

E(w) = 2w122+2w12w22+5w222

= 2(1.04)2 + 2 (1.04\* (-0.32)) + 5 (-0.32)2

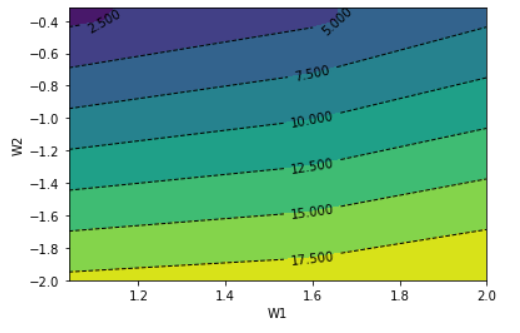
= 2(1.08) + 2 (-0.333) + 5 (0.102)

= 2.16 – 0.66 + 0.51

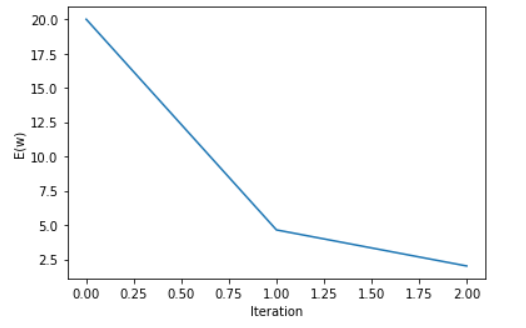
**E(w) = 2.01**

Hence, we can say that the minima after two iteration for E(w) is 2.01 and w = [1.04, -0.32]

The Contour plot can be shown in the below figure. The python code is presented as a separate file.



The learning curve for the function based on the iteration can be given as below,



# **References**

Gopal, M. (2019). *Applied machine learning*. McGraw-Hill Education.