ADAGRAD Optimizer

SF.3.

Given Samples

x	y
0.2	3.4
0.4	3.8
0.6	4.2
0.8	4.6

Step 1:

$[x, y]$, $\eta = 0.1$, epochs = 1, $m = 1$, $c = 1$, $\varepsilon = 10^{-8}$, $g_m = 0$, $g_c = 0$

Step 2:

Iteration = 1

Step 3

Sample = 1

Step 4

$$g_m = -[y_i - mx_i - c]\, x_i$$

$$= -[3.4 - (1 \times 0.2) - (1)]0.2$$

$$= -[3.4 - (0.2) + 1]0.2$$

$$= [3.4 - 0.2 + 1] \times 0.2$$

$$= -[4.27] \times 0.2$$

$$= -0.854$$

$$g_c = -[y_i - mx_i - c]$$

$$= -[3.4 - (1 \times 0.2) - (1)]$$

$$= -[3.4 - 0.2 + 1]$$

$$= -[4.27]$$

Step 5:

$g_m = g_m + (g_m)^2$

$0 + (0.84)^2 = 0.7056$

$g_c = g_c + (g_c)^2$

$0 + (4.2)^2$

$= 17.64$

Step 6: $\Delta m = \frac{-0.1}{\sqrt{0.7056 + 10^{-8}}} \times (-0.84)$

$= -0.09999$

$\Delta c = \frac{-0.1}{\sqrt{17.64 + 10^{-8}}} \times (-4.2)$

$= -0.09999$

Step 7:

$m = m + \Delta m = 1 + 0.9999 = 1.9999$

$c = c + \Delta c = +1 + 0.9999 = -0.001$

Step 8:

sample = sample + 1

$= 1 + 1 = 2$

Step 9: 2 ? 2 = false

go to step 4

Step 10:

$$g_m = -[y_i - mx_i - c]x_i$$

$$-[3.8 - (0.4 \times 1.9999) + 0.001] \times 0.4$$

$$-[1.8011] \times 0.4$$

$$= -0.72044$$

$$g_c = -[y_i - mx_i - c]$$

$$-[3.8 - (0.4)(1.9999) + 0.001] \times 0.4$$

$$= -1.8011$$

Step 11:

$$g_m = g_m + (g_m)^2 = 0.7056 + 0.5190$$

$$= 1.2246$$

$$g_c = g_c + (g_c)^2 = 17.64 + 3.2437 = 20.8837$$

Step 12: $\delta m = \frac{-0.1}{\sqrt{1.2246 + 10^{-8}}} \times (-0.72044) = +0.065102$

$$\delta c = \frac{-0.1}{\sqrt{20.8837 + 10^{-8}}} \times (-1.8011) = 0.03941$$

Step 13: $m = 1.9999 + 0.065102 = 2.0650$

$$c = -0.001 + 0.3941 = 0.3931$$

Step 14: Sample = Sample + 1 = 2nd = 3rd time

Step 15: iter = iteration + 1

1 + 1 => 2

Step 16: iter > epochs => 2 > 2 => false

Goto step 17

Step 17: sample = 1

Step 18: $gm = -\left[3.4 - (2.0650 \times 0.2) - 0.3931\right] \times 0.2$

$gm = -\left[2.5939\right] \times 0.2 = -0.5187$

$gc = -2.5939$

Step 19: $Gm = Gm + (gm)^2 \Rightarrow 1.2246 + 0.2690 = 1.4936$

$Gc = Gc + (gc)^2 \Rightarrow 20.8839 + 6.7283 = 27.6122$

Step 20: $\Delta m = \frac{-0.1}{\sqrt{1.4936 + 10^{8}}} \times (-0.5187) = 0.01759$

$\Delta c = \frac{-0.1}{\sqrt{27.6122 + 10^{8}}} \times (-2.5939) = 0.04936$

Step 21:

$m = m + \Delta m \Rightarrow 2.0650 + 0.01759 = 2.08259$

$c = c + \Delta c \Rightarrow 0.3931 + 0.04936 = 0.44246$

Step 22: sample = sample + 1 => 1 + 1 => 2 > 2 false

Goto step 3

Step 23

$g_m = -[3.8 - (2.08289 \times 0.4) - 0.44246] \times 0.4$

$= -(2.5243) \times 0.4$

$= -1.00972$

$g_c = -2.5243$

Step 24: $G_m = G_m + (g_m)^2 = 1.4936 + (-1.00972)^2$

$= 2.5131$

$G_c = G_c + (g_c)^2 = 27.6122 + (-2.5243)^2$

$= 33.9842$

Step 25: $\Delta m = -\left(\frac{0.1}{\sqrt{2.5131 + 10^{-8}}} \times (-1.00972)\right) = 0.06369$

$\Delta c = -\left(\frac{0.1}{\sqrt{33.9842 + 10^{-8}}} \times (-2.5243)\right) = 0.0433$

Step 26: $m = m + \Delta m = 2.08289 + 0.06369$

$= 2.14658$

$c = c + \Delta c = 0.44246 + 0.0433 = 0.48576$

Step 27: samples = samples + 1 = 2 + 1 = 3

Goto next step

Step 28: iter = iter + 1 = 2 + 1 = 3 ? epochs

Step 29: print (m, c)

Step 30: calculate mean square error (mse)

$$\frac{1}{2\times2} \Sigma[y_i - \hat{y}_i]^2 = \frac{1}{4}\left[(3.4 - (2.14658 \times 0.2) - 0.48576)^2 + (3.8 - (2.41658 \times 0.4) - 0.48576)^2\right]$$

mse = 20631