ADAGRAD OPTIMIZER

AFEEFA (18K41A0503)

Manual Calculations for 2 iterations with 1st 2 samples.

Sample(i)	x_i^a	y_i^a
1	0.2	3.4
2	0.4	3.8
3	0.6	4.2
4	0.8	4.6

step1 epochs = 2, m = 1, c = -1, $G_m = G_c = 0$ (sum of squared of gradient)

$\eta = 0.1$, $\varepsilon = 10^{-8}$, ns = 2

step2 iter = 1

step3 sample = 1

step4 $g_m = -(y_i - mx_i - c)x_i$

$g_c = -(y_i - mx_i - c)$

$g_m = -(3.4 - 0.2 + 1)(0.2)$

$g_m = -0.84$

$g_c = -(3.4 - 0.2 + 1) = -4.2$

step5 $G_m = G_m + (g_m)^2$

$G_m = 0 + (-0.84)^2 = 0.7056$

$G_c = G_c + (g_c)^2$

$G_c = 0 + (-4.2)^2 = 17.64$

step6 $\Delta m = \frac{-\eta}{\sqrt{G_m + \varepsilon}} g_m$

$= \frac{-0.1}{\sqrt{0.7056 + 10^{-8}}} (-0.84)$

$= \frac{0.084}{0.84}$

$\Delta m = 0.10$

$\Delta c = \frac{-\eta}{\sqrt{G_c + \varepsilon}} g_c$

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

$\Delta c = \frac{0.42}{4.9} = 0.90$

step7 $m = m + \Delta m = 1 + 0.10 = 1.10$

$c = c + \Delta c = -1 + 0.10 = -0.90$

step8 sample = sample + 1 = 2

step9 if (sample > ns)

2 > 2 ✗

goto step4

step4 $g_m = -(y_i - m x_i - c) x_i$

$g_m = -(3.8 - (1.10)(0.4) + 0.90)(0.4) = -1.704$

$g_c = -(3.8 - (1.10)(0.4) + 0.90)$

$g_c = -4.26$

$g_m = -1.704$

$g_c = -4.26$

step 5: $G_m = G_m + (g_m)^2$

$G_m = 0.7056 + (-1.704)^2$

$G_m = 3.6092$

$G_c = G_c + (g_c)^2$

$= 17.64 + (-4.26)^2$

$G_c = 35.78$

step 6: $\Delta m = \frac{-\eta}{\sqrt{G_m + \varepsilon}} g_m$

$= \frac{-0.1}{\sqrt{3.609 + 10^{-8}}} (-1.704) = \frac{0.1704}{\sqrt{3.609 + 10^{-8}}}$

$\Delta m = 0.0896$

$\Delta c = \frac{-\eta}{\sqrt{G_c + \varepsilon}} g_c$

$= \frac{-0.1}{\sqrt{35.78 + 10^{-8}}} (-4.26)$

$= \frac{0.426}{\sqrt{35.78 + 10^{-8}}}$

$\Delta c = 0.0712$

step 7 $m = m + \Delta m = 1.10 + 0.0896 = 1.1896$

$C = c + \Delta c = -0.90 + 0.0712 = -0.828$

step8: sample = sample+1 = 3

step9: if(sample<=ns)
3>2 ✓

step10: iter = iter+1
= 2

step11: if(iter>epochs)
2>2 ✗

step3: sample = 1

step4: $g_m = -(y_i - mx_i - c)x_i$

$= -(3.4 - (1.18)(0.2) + 0.82)(0.2)$

$g_m = -0.796$

$g_c = -(y_i - mx_i - c)$

$g_c = -(3.4 - (1.18)(0.2) + 0.82) = -3.98$

step5: $G_m = G_m + (g_m)^2$

$= 3.6092 + (-0.796)^2$

$G_m = 4.23$

$G_c = G_c + (g_c)^2$

$= 35.78 + (-3.98)^2$

$G_c = 51.62$

step6: $\Delta m = \frac{-\eta}{\sqrt{G_m + \varepsilon}} g_m$

$= \frac{-0.1\ (-0.796)}{\sqrt{4.23 + 10^{-8}}} = 0.0384$

$$\Delta c = \frac{\Delta \eta}{\sqrt{G_c + \varepsilon}} g_c$$

$$= \frac{-0.1(-3.98)}{\sqrt{51.62 + 10^{-8}}} \varepsilon$$

$$\Delta c = 0.055$$

step7 $m = m + \Delta m = 1.1896 + 0.0384 = 1.228$

$c = c + \Delta c = -0.828 + 0.055 = -0.773$

step8 sample = 1 = 2

step9 if (sample > ns)

2 > 2 X

step4: $g_m = -(y_i - mx_i - c)x_i$

$= -(3.8 - (1.22)(0.4) + 0.77)(0.4)$

$g_m = -1.63$

$g_c = -(3.8 - (1.22)(0.4) + 0.77)$

$g_c = -4.082$

step5 $G_m = G_m + (g_m)^2$

$= 4.23 + (-1.63)^2 = 1.573$

$G_c = G_c + (g_c)^2$

$G_c = 51.62 + (-4.08)^2 = 68.26$

step6: $\Delta m = \frac{-\eta}{\sqrt{G_m + \varepsilon}} g_m$

$= \frac{-0.1}{\sqrt{1.573 + 10^{-8}}} (-1.63)$

$\Delta M = 0.129$

$\Delta c = \frac{-\eta}{\sqrt{G_c + \varepsilon}} g_c$

$= \frac{-0.1}{\sqrt{68.26 + 10^{-8}}} (-4.082)$

$\Delta c = 0.049$

step7: $m = m + \Delta m = 1.228 + 0.129 = 1.357$

$c = c + \Delta c = -0.773 + 0.049 = -0.724$

step 8: sample+1 = 2+1 = 3

step 9: if(sample>ns)

3 > 2 ✓

next step

step 10: iter+1 = 2+1 = 3

step 11: if(iter>epochs)

3 > 2 ✓

next step

step 12: print(m, c)

m = 1.357

c = −0.724

} 2 iterations on 2 samples