Assignment-11 : Nesterov Accelerated Gradient Descent

Manual Calculations

STEP-1: read [X,Y], m = 1, c = -1, η = 0.1, γ = 0.9, v_m = 0, v_c = 0, epochs = 2, no_of_samples = 2

X	Y
0.2	3.4
0.4	3.8

STEP-2: iter = 1

STEP-3: sample = 1

STEP-4: $g_m = -(y_i - (m + \gamma v_m)x_i - (c + \gamma v_c))x_i$

$= -(3.4 - (1 + (0.9)\times 0)\times 0.2 - ((-1) + 0))\times 0.2$

$= -(3.4 - 0.2 + 1)\times 0.2 = -(4.2 \times 0.2) = -0.84$

$g_c = -4.2$

STEP-5: $v_m = \gamma v_m - \eta g_m = (0.9)(0) - (0.1)(-0.84)$

$= 0.084$

$v_c = \gamma v_c - \eta g_c = (0) - (0.1)(-4.2) = 0.42$

STEP-6: $m = m + v_m = 1 + 0.084 = 1.084$

$c = c + v_c = -1 + 0.42 = -0.58$

STEP-7: sample = 1 + 1 = 2

STEP-8: if sample > no_of_samples ⇒ 2 > 2 ⇒ false

goto STEP 4

STEP-9: $g_m = -(3.8 - (1.084 + (0.9)\times(0.084))\times 0.4 - (-0.58 + (0.9)\times(0.42))\times 0.$

$\Rightarrow g_m = -(3.8 - (1.1596 \times 0.4) + 0.958) \times 0.4$

$= -(4.29416) \times 0.4 = -1.717664$

$g_c = -4.29416$

STEP-10: $V_m = \forall V_m - \eta g_m = (0.9)(0.084) - (0.1)(-1.717664)$

$= 0.2473664$

$V_c = \forall V_c - \eta g_c = (0.9)(0.42) - (0.1)(-4.29416) = 0.8074416$

STEP-11: $m = m + V_m = 1.084 + 0.24736 = 1.33136$

$C = C + V_c = -0.58 + 0.807416 = 0.227416$

STEP-12: Sample = 2+1 = 3

STEP-13: if sample > no-of_samples = 3 > 2 = true

goto next step

STEP-14: iter = 1+1 = 2

STEP-15: if iter > epochs $\Rightarrow$ 2 > 2 = false

go to step 3

STEP-16: Sample = 1

Step 17: $g_m = -(y_i - (m + \gamma v_m)x_i - (c + \gamma v_c))x_i$

$= -(3.4 - [[1.33136 + [(0.9) \times (0.24736)]] \times 0.2 -$

$(0.227416 + (0.9) \times 0.807416)$

$= -(3.4 - [1.553984] \times 0.2 - [0.95409])$

$= -(2.13511)$

$g_c = -(3.4 - 1.553984 - 0.95409) = -0.891926$

step 18: $v_m = \gamma v_m - \eta g_m = (0.9) \times 0.2473664 - (0.1) \times (-2.13511)$

$= 0.43614$

$v_c = \gamma v_c - \eta g_c = (0.9) \times 0.807416 - (0.1) \times (-0.891926)$

$= 0.815867$

STEP 19: $m = m + v_m = 1.3316 + 0.43614 = 1.76774$

$c = c + v_c = 0.227416 + 0.815867 = 1.043283$

STEP 20: sample = sample + 1 = 1 + 1 = 2

STEP 21: if sample > $n_s \Rightarrow 2 > 2 \rightarrow$ false

repeat step 4

STEP 22: $g_m = -(y_i - (m + \gamma v_m)x_i - (c + \gamma v_c))x_i$

$= -[3.8 - (1.76774 + (0.9) \times 0.43614)] \times 0.4 - (1.043283$

$+ (0.9) \times 0.815867)] \times 0.4$

$= -[3.8 - (2.160266) \times 0.4 - 1.7775633] \times 0.4$

$= -0.463332$

$g_c = -[3.8 - (2.160266 \times 0.4) - 1.7775633]$

$= -1.1583303$

STEP 23: $V_m = \gamma V_m - \eta \frac{\partial E}{\partial m}$

$= (0.9) \times 0.43614 - (0.1) \times (-0.463332)$

$= 0.4388592$

$V_c = \gamma V_c - \eta \frac{\partial E}{\partial c}$

$= (0.9) \times 0.815867 - (0.1) \times (-1.583303)$

$= 0.8501133$

STEP 24: $m = 1.76774 + 0.4388592 = 2.2065992$

$c = 1.043283 + 1.1583303 = 2.2016133$

STEP 25: sample = 2+1 = 3 > 2 ⇒ sample > epochs

goto step 26

STEP 26: iters = 2+1 = 3 ⇒ > epochs

Step 27

STEP 27: Print (m, c)

⇒ 2.2065992 , 2.2016133

STEP 28: Mean Squared Error

$= (3.4 - (2.2065992 \times 0.2) + -2.2016133)^2$

$+ (3.8 - (2.2065992 \times 0.4) - 2.2016133)^2$

$$= \frac{(0.57315) + (0.512293)}{2}$$

$$= \frac{1.085443}{2} = 0.542721$$