

Let us consider a sample dataset have one input ( $x_i$ ) and one output ( $y_i$ ) and number of samples 4. Develop a LR model using nestrov accelerated gradient (NAG) optimises.

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

• Do manual calculations for 2 iterations with 1<sup>st</sup> 2 samples.

Step-1:  $[X, Y], m=1, c=-1, \eta=0.1, \text{epochs}=2, \gamma=0.9, \bar{v}_m=\bar{v}_c=0, n_s=2$

Step-2:  $\text{itr}=1$

Step-2: Sample=1

$$\begin{aligned} \text{Step-4: } g_m &= \frac{\partial E}{\partial m} = -(y_i - (m + \gamma \bar{v}_m)x_i - (c + \gamma \bar{v}_c))x_i \\ &= -(3.4 - (1 + (0.9)0)0.2 - (-1 + (0.9)0)0.2) \\ &= -0.84 \end{aligned}$$

$$\begin{aligned} g_c &= \frac{\partial E}{\partial c} = -(y_i - (m + \gamma \bar{v}_m)x_i - (c + \gamma \bar{v}_c)) \\ &= -(3.4 - (1 + 0.9) \times 0)0.2 \\ &= -(-1 + (0.9)0) \\ &= -4.2 \end{aligned}$$

$$\begin{aligned} \text{Step-5: } \bar{v}_m &= \gamma \bar{v}_m - \eta g_m \\ &= (0.9)0 - (-0.1) \times (-0.84) \\ &= -0.084 \end{aligned}$$

$$\begin{aligned} \bar{v}_c &= \gamma \bar{v}_c - \eta g_c \\ &= (0.9)0 - (-0.1) \times (-4.2) \\ &= -0.42 \end{aligned}$$

$$\begin{aligned} \text{Step-6: } m &= m + \bar{v}_m = 1 - 0.084 = 0.916 \\ c &= c + \bar{v}_c = -1 - 0.42 = -1.42 \end{aligned}$$

step-7: sample + = 1  
1 + 1 = 2

②

step-8: if (sample > ns)  
goto step-9  
else  
goto step-4

$$\text{step-4: } g_m = \frac{\partial E}{\partial m} = -(3.8 - (0.916 + (0.9 \times -0.089)) \times 0.4 - (-1.42) + (0.98 - 0.034) \times 0.4) \\ = -1.983$$

$$g_c = \frac{\partial E}{\partial c} = -4.959$$

$$\text{step-5: } \Delta m = \eta g_m \\ = (0.9 \times -0.089) - (-0.1 \times -1.983) \\ = -0.2739 \\ \Delta c = (0.9 \times -0.42) - (-1 \times -4.959) \\ = 0.8739$$

$$\text{step-6: } m + = \Delta m \\ = 0.916 - 0.2739 \\ = 0.6421 \\ c + = \Delta c \\ = -1.42 - 0.8739 \\ = -2.2939$$

Step-7: sample + = 1  
1 + 1 = 2

step-8: if (sample > ns)  
goto step-11  
2 > 2  
else  
goto step-3

step-3: sample = 1

$$\text{step-4: } \frac{\partial E}{\partial m} = -(3.4 - (0.642 + (0.9 \times 0.273)) \times 0.2 - (-2.293) + (0.9 \times -0.273) \times 0.2) \\ g_m = -1.171 \\ g_c = \frac{\partial E}{\partial c} = -5.859$$

(3)

$$\begin{aligned}\text{Step-5: } v_m &= \delta v_m - \eta j_m \\ &= [(0.9) \times (-0.273)] - (-0.1 \times -1.81) \\ &= -0.3627\end{aligned}$$

$$\begin{aligned}v_c &= \delta v_c - \eta j_c \\ &= (0.9)(-0.873) - (-0.1)(-5.859) \\ &= -1.3707\end{aligned}$$

$$\begin{aligned}\text{Step-6: } m+ &= v_m \\ &= 0.6421 + (-0.3627) \\ &= 0.2794 \\ c+ &= v_c \\ &= -2.2939 - 1.3707 \\ &= -3.6646\end{aligned}$$

$$\begin{aligned}\text{Step-7: } \text{sample}+ &= 1 \\ 1+1 &= 2\end{aligned}$$

$$\begin{aligned}\text{Step-8: } \text{if}(\text{sample} > \text{ns}) \\ &\text{goto step-9} \\ \text{else} \\ &\text{goto step-4}\end{aligned}$$

$$\begin{aligned}\text{Step-4: } j_m &= \frac{\partial E}{\partial m} = -13.8 - (0.2794 + (0.9 \times -0.3627)) \times 0.4 - (-3.6646 + (0.9)) \\ &= -2.985\end{aligned}$$

$$j_c = \frac{\partial E}{\partial c} = -7.4645$$

$$\begin{aligned}\text{Step-5: } v_m &= [0.9 \times -0.3627] - [-0.1 \times -2.985] \\ &= -0.6249 \\ v_c &= [0.9 \times -1.3707] - [-0.1 \times 7.4645] \\ &= -1.9200\end{aligned}$$

$$\begin{aligned}\text{Step-6: } m+ &= v_m \\ &= 0.2974 + (-0.6249) \\ &= -0.3275 \\ c+ &= v_c = -3.6646 - 1.9200 \\ &= -4.6446\end{aligned}$$

step-7: sample += 1  
2 + 1 = 3

step-8: if (sample > n\_s)  
    goto step-9  
else  
    goto step-4

step-9: itr += 1  
2 + 1 = 3

step-10: if (itr > epochs)  
    goto step-4  
else  
    goto step-3

step-11: print m, c  
m = 0.3275  
c = -4.6446