**2.1**

According to the gradient descend formula which is Wt=Wt-1-a\*dWt-1

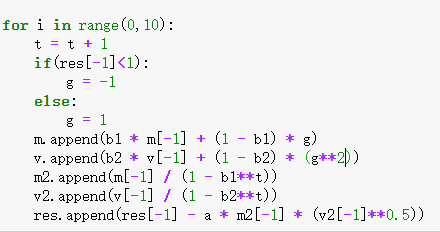
And in the left side of the graph, we can calculate its gradient which is -1

We define the initial point which is w0, then the next point which is w1, its value can be calculated by applying the formula which is w0-(learning rate) \*(-1)=w0-0.3\*(-1)=w0+0.3

Hence, we get the value of w1 which equals to w0+0.3, we can repeat the same process to calculate w2=w1-0.3\*(-1)=w1+0.3=w0+0.6; w3=w2-0.3\*(-1)=w2+0.3=w1+0.9; w4=w3-0.3\*(-1)=w0+1.2; (Here, since 1.2>1, the gradient change to 1) w5=w4-0.3\*(1)=w4-0.3=w0+0.9, hence we find that w5=w3, which means that the algorithm will stuck at that point

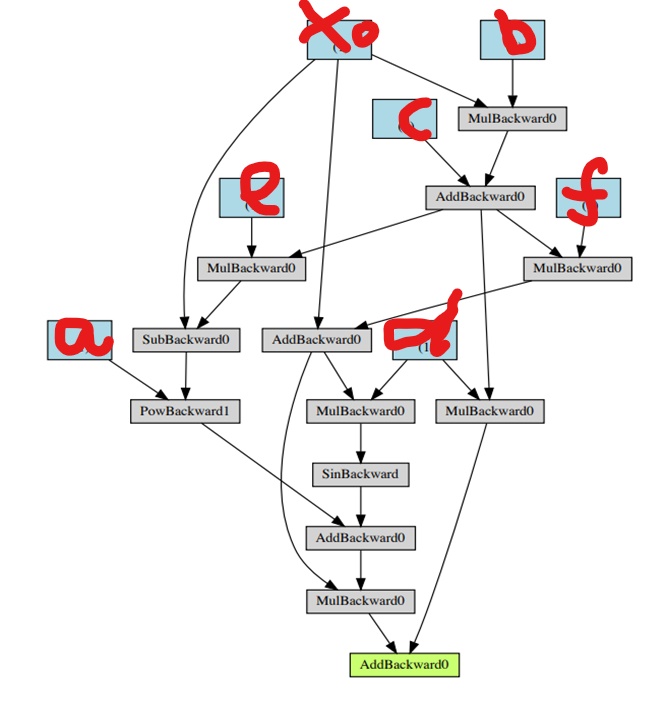
**2.2**

By defining the adam optimization function and run it 10 times



I can get the maximum height which is 0.410

**3**

****