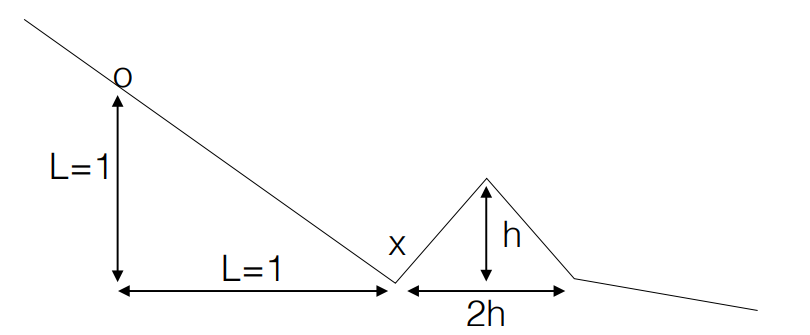
# Assignment 2

1.

2.



1. SGD

I set the x axis for O is 0.

For the first optimize, the gradient = -1 and the a = 0.3. so the O will move from 0 to 0.3.

For the second optimize, the gradient = -1 and the a = 0.3. so the O will move from 0 to 0.6.

For the third optimize, the gradient = -1 and the a = 0.3. so the O will move from 0 to 0.9.

For the fourth optimize, the gradient = -1 and the a = 0.3. so the O will move from 0 to 1.2.

For the fifth optimize, the gradient = 1 and the a = 0.3. so the O will move from 0 to 0.9.

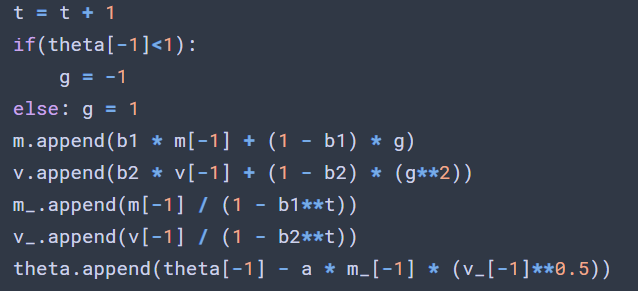
For the sixth optimize, the gradient = -1 and the a = 0.3. so the O will move from 0 to 1.2.

And converge.

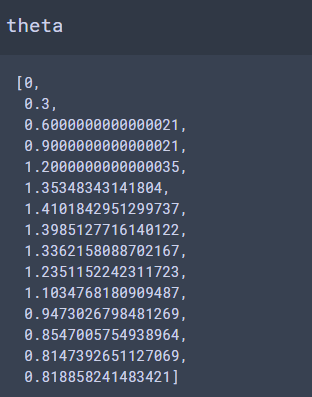
1. Adam

If we want to escape from the local min x(x axis=1), we need the point to move to the position that larger than the peek(x axis > 1 + h). since h is known, I assume that h is positive infinite, and to see with given parameter and adam, what is the max length that the point can move. And the h should be smaller than that value.

The following is the code for adam.



And result is as follow



We can see that at the beginning and the gradient = -1, the Theta is optimized just like the SGD, but when theta = 1.2, the gradient change to 1. But unlike the SGD, the next theta is not 0.9 but with the momentum, the next theta = 1.35. With this momentum, the max value that O can reach is 1.41. And after that the O start to go back to the local min x. So if we want to escape form the local min, the highest point(1.41) should have the gradient < 0.

So the max value of h is 1.41 – 1 = 0.41.

3.

