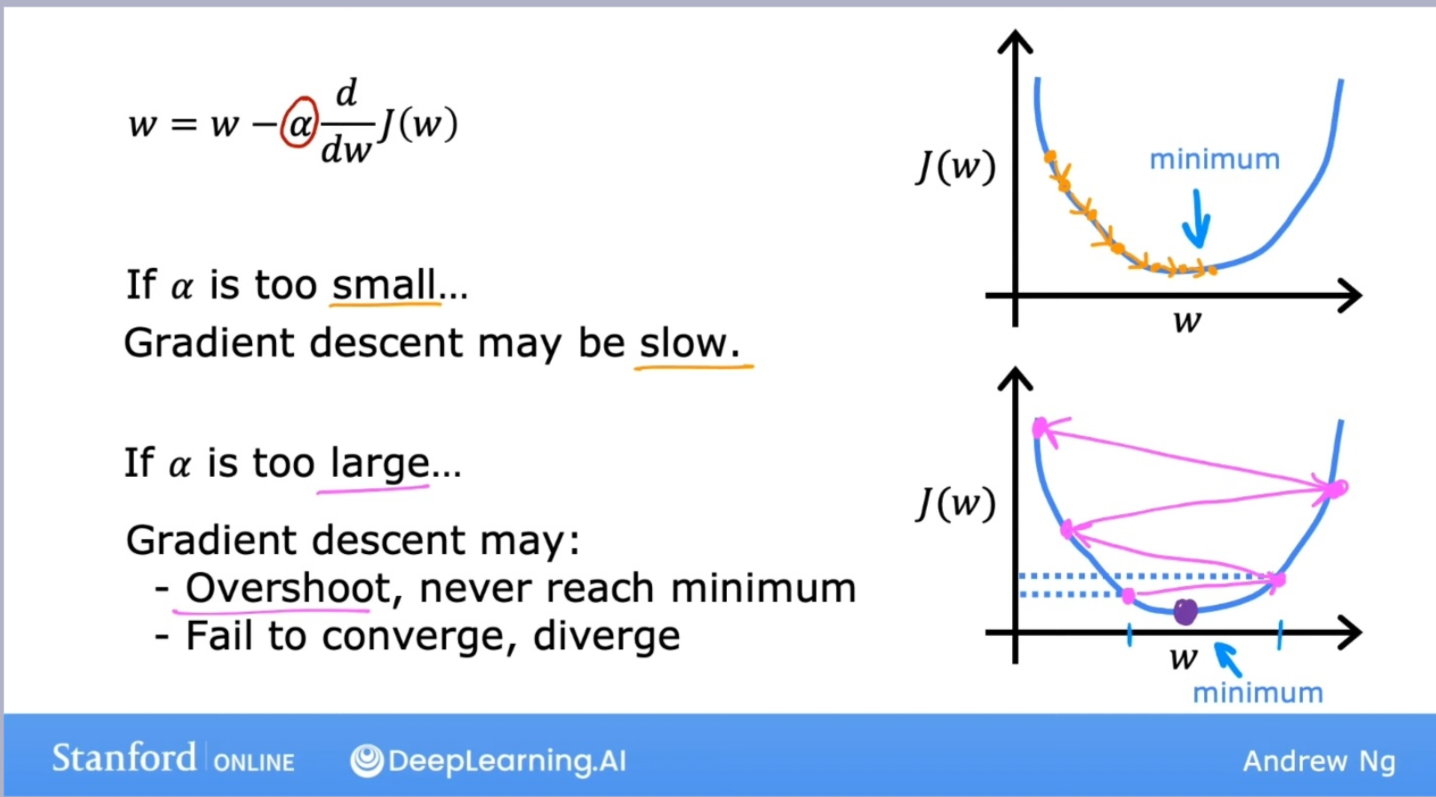
→ Here we will see how different learning rate effects the cost func graph.



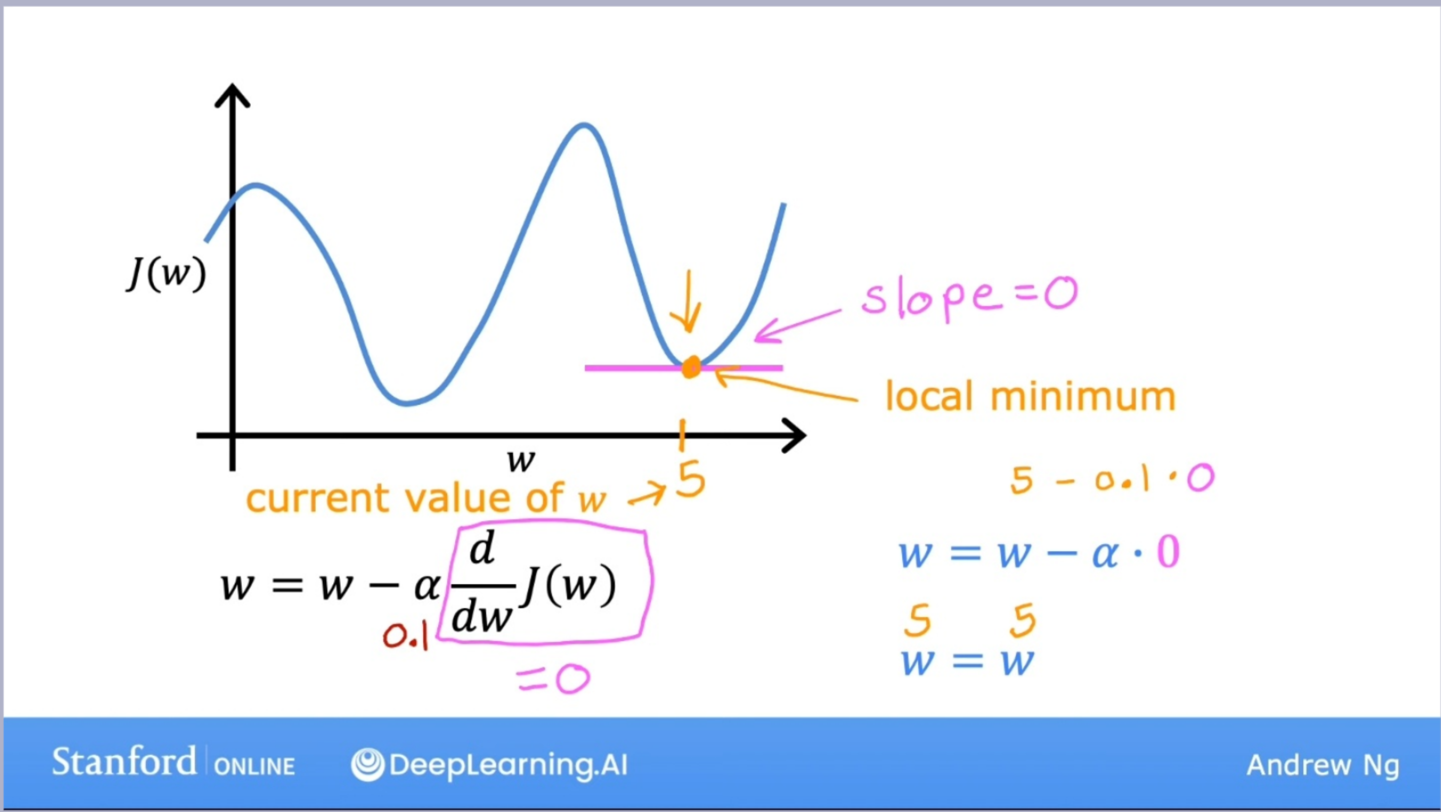
→ If we choose α as a small value, it will take tiny steps toward local minima

→ so it will be slow.

→ If α is too high, it may work opposite.

→ Cause it will take huge steps and it may increase the $J(w)$. since the derivatives term is changing.

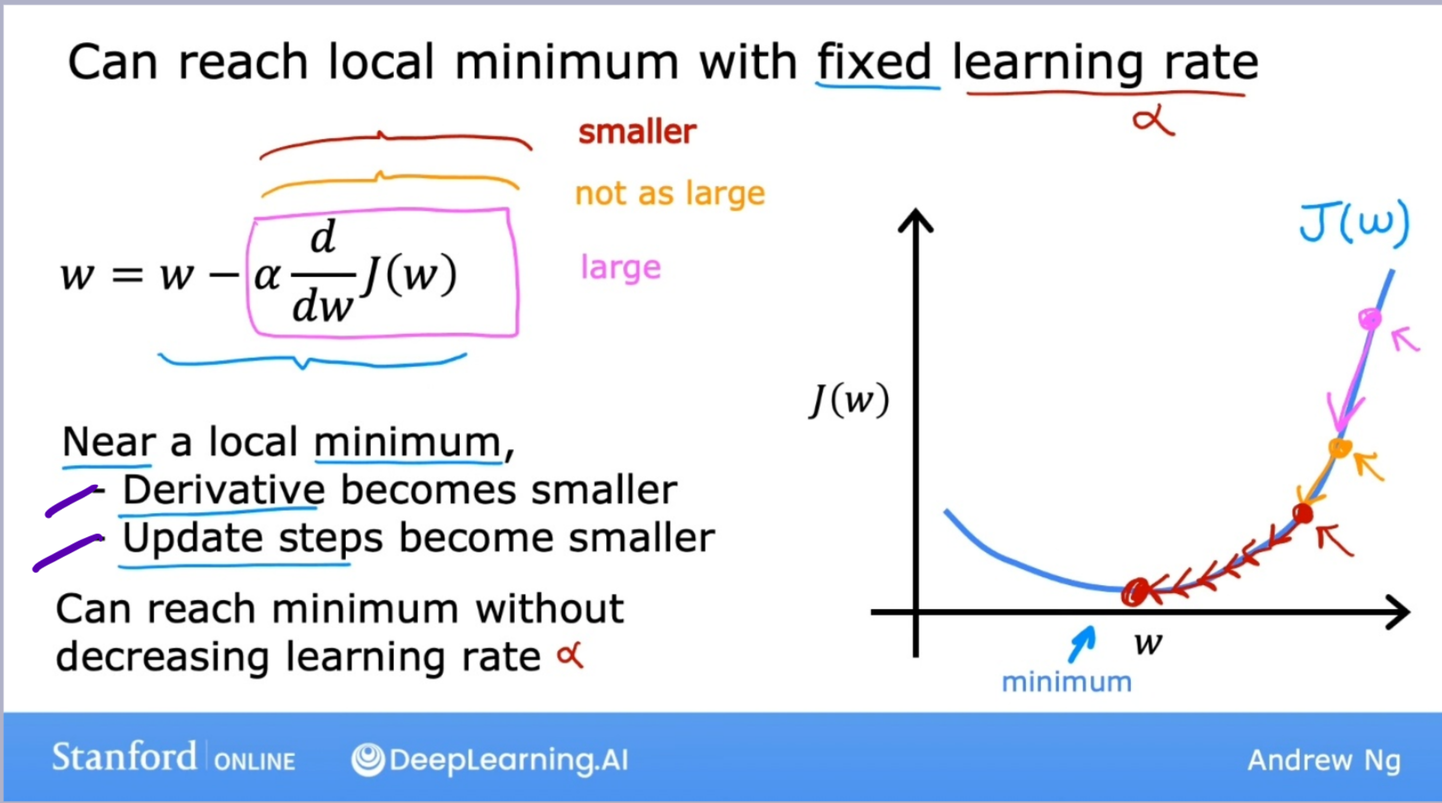
→ The value of w and b will increase.



→ The main drawback of this formula is, it always stucks on local minima.

→ cause, at the bottom of the curve the slope is 0. So, it becomes $W=W-0$

→ the value of w won't change.



→ We do not need to change the learning rate. It will work with fixed α

→ At first step it will take large step. after that it will take small steps as reaching toward local minima.

→ cause the slope value is decreasing as we reaching toward local minima.