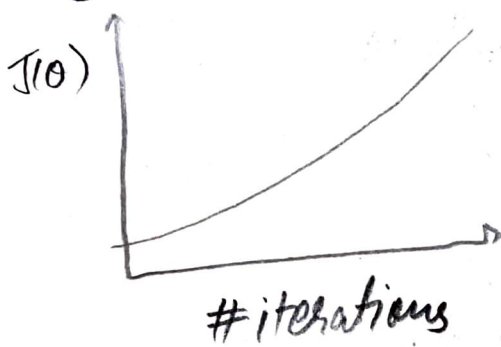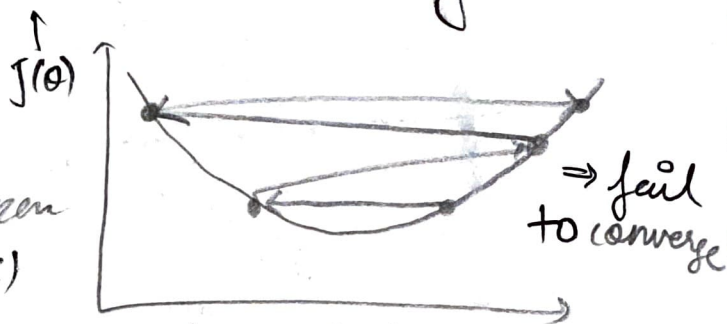If $J(\theta)$ v/s # iteration looks like
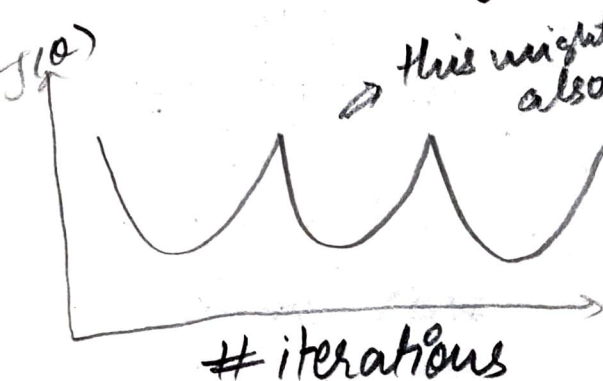

$\Rightarrow$ this means α (i.e learning rate is too large) and dec^ the value of α

$J(\theta)$

iterations

$\Rightarrow$ fail to converge

this might also happen the code have bug(s)

$J(\theta)$

$\rightarrow$ this might also happen when α is large if α is large

$\theta \rightarrow$

iterations

$\Rightarrow$ for sufficiently small α, $J(\theta)$ should decrease on every iteration (theorem)

$\Rightarrow$ but if α is too small then, can be slow to converge

Summary,

$\Rightarrow$ α is too small $\Rightarrow$ slow convergence

$\Rightarrow$ α is too large $\Rightarrow$ $J(\theta)$ may not converge with every iteration as it does decrease or convergence may happen but slowly.

$(\alpha = \cdots, 0.001, 0.01, 0.1, \cdots, 1 \cdots$ so on$)$

x3 $\rightarrow$ 0.003 x3 $\rightarrow$ 0.03 $\rightarrow$ try to choose like this in practice