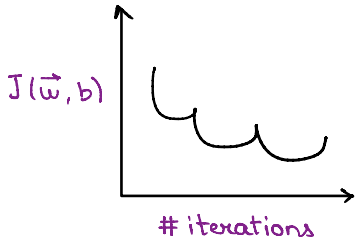


Identify problem with gradient descent



or learning rate too large

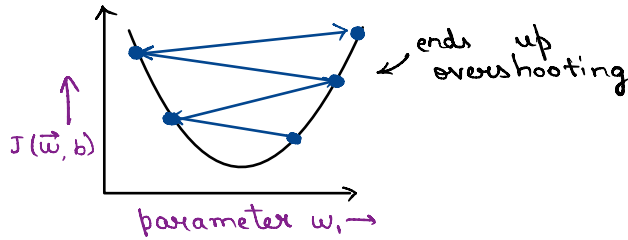
error

code :- ~~$w_1 = w_1 + \alpha d_1$~~
 ~~$w_2 = w_2 - \alpha d_2$~~

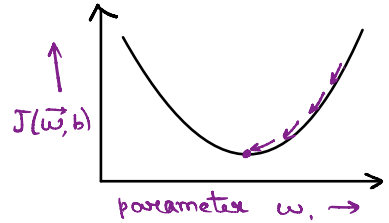
↓
correct

but something like this could be due to a bug

α is too big



Use smaller α



with a small enough α , $J(\vec{w}, b)$ should decrease with each iterations, but shouldn't be too small as it will take a lot of time to converge

Values of α to try :-

approx 0.009
 \downarrow
 ≈ 0.01

... 0.001 0.003 0.01 0.03 0.09 0.1 0.3 1
 $\swarrow *3 \nearrow \swarrow *3 \nearrow \swarrow *3 \nearrow \swarrow *3 \nearrow \swarrow *3 \nearrow \swarrow *3 \nearrow$

