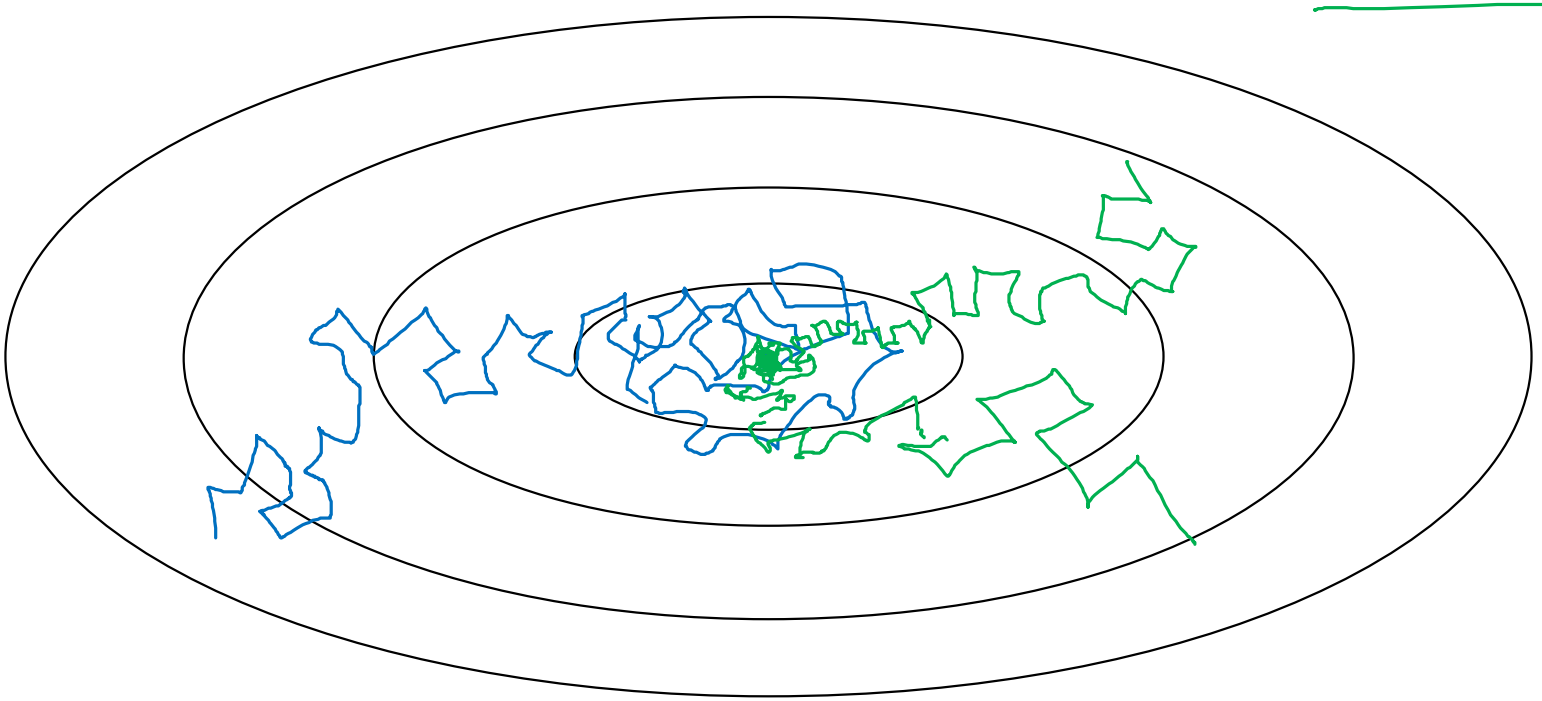
deeplearning.ai

Optimization Algorithms

Learning rate decay

Learning rate decay

Slowly reduce α

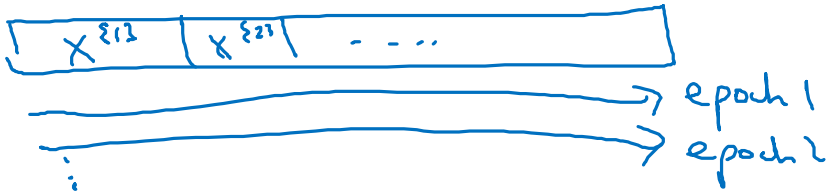


Andrew Ng

Learning rate decay

1 epoch = 1 pass through data.

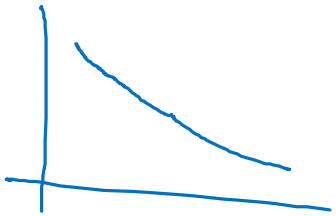
$$\alpha = \frac{1}{1 + \text{decay\_rate} * \text{epoch\_num}} \alpha_0$$



$\alpha_0 = 0.2$

decay_rate = 1

Epoch	α
1	0.1
2	0.67
3	0.5
4	0.4
⋮	⋮



Andrew Ng

Other learning rate decay methods

$\alpha = 0.95^{\text{epoch-num}} \cdot \alpha_0$ — exponentially decay.

$\alpha = \frac{k}{\sqrt{\text{epoch-num}}} \cdot \alpha_0$ or $\frac{k}{\sqrt{t}} \cdot \alpha_0$

formula

discrete staircase

Manual decay.