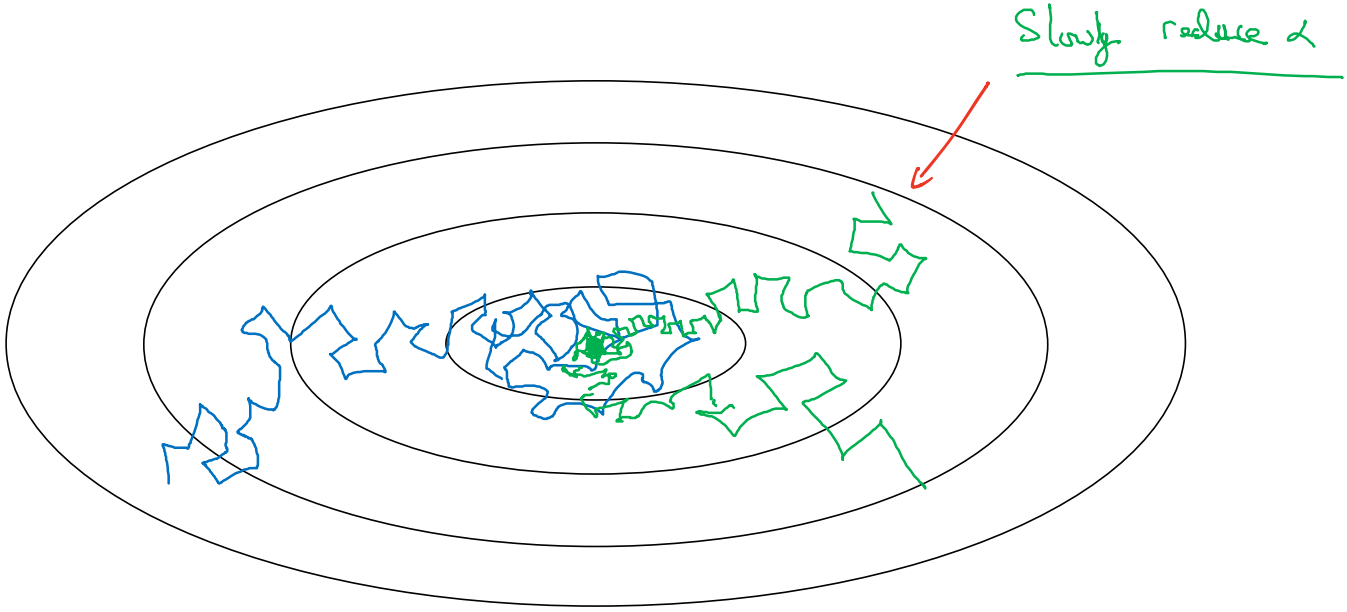
deeplearning.ai

Optimization Algorithms

Learning rate decay

Learning rate decay



Learning rate decay

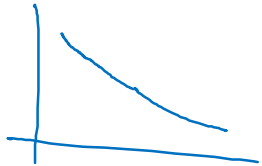
1 epoch = 1 pass through data.

$$\alpha = \frac{1}{1 + \text{decayRate} * \text{epochNum}} \alpha_0$$

$\alpha_0 = 0.2$

decay_rate = 1

Epoch	α
1	0.1
2	~~0.67~~ 0.067
3	~~0.5~~ 0.05
4	~~0.4~~ 0.04
⋮	⋮

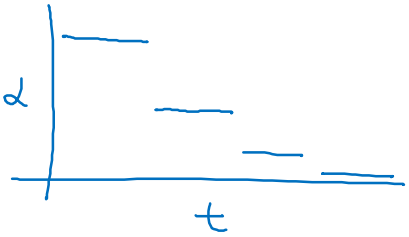


Other learning rate decay methods

formula

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

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



discrete

discrete staircase

Manual decay.