Optimization Algorithms

Learning rate decay
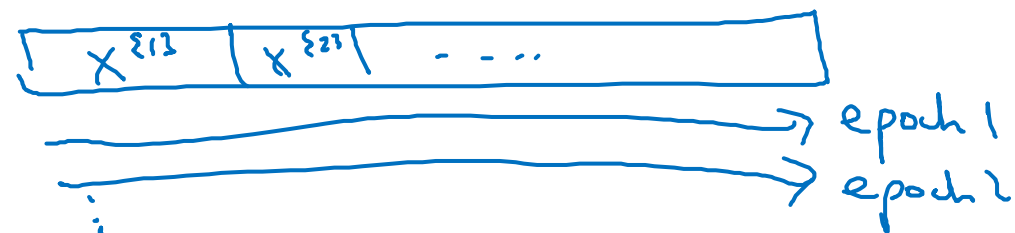
deeplearning.ai

Learning rate decay

Slowly reduce α

Andrew Ng
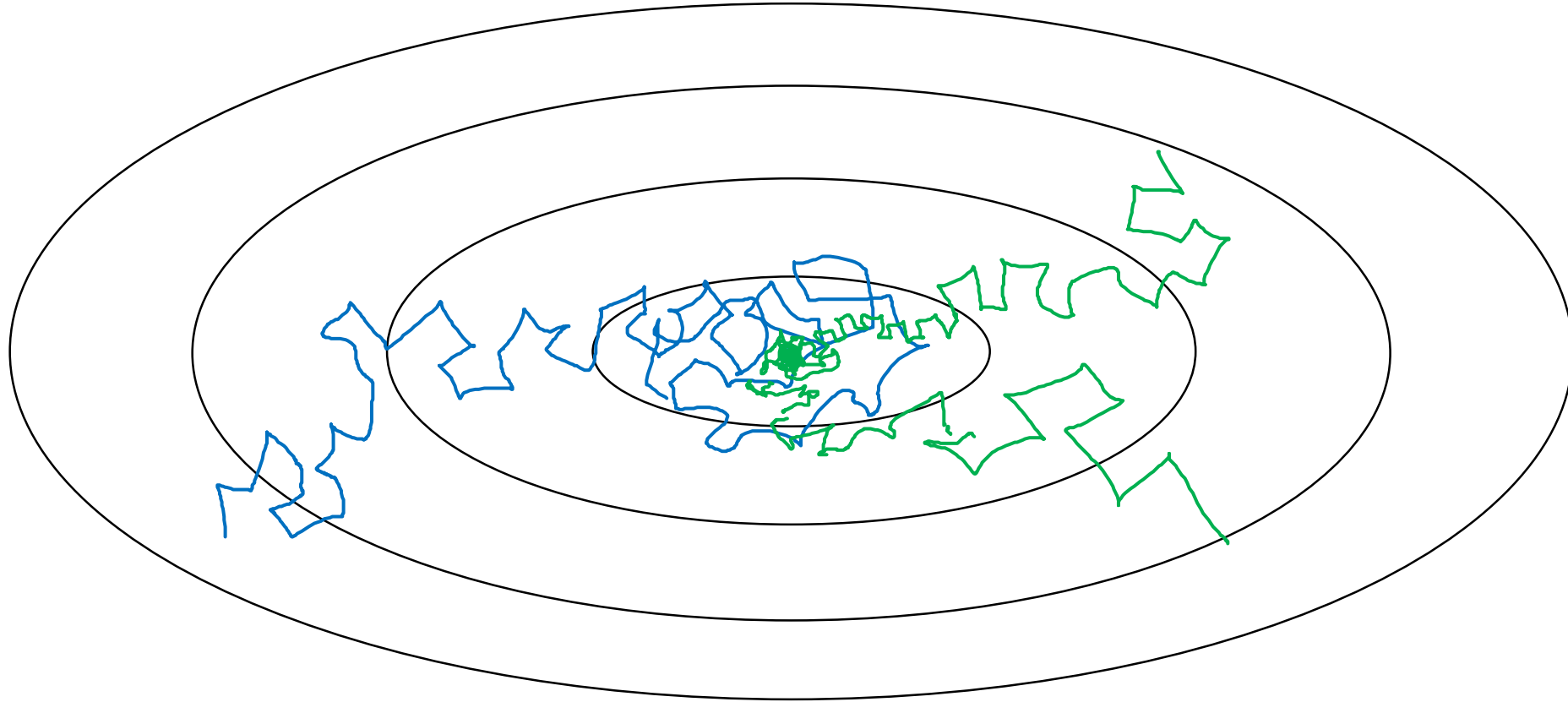
Learning rate decay

1 epoch = 1 pass through data.

$$\alpha = \frac{1}{1 + \text{decay-rate} * \text{epoch-num}} \alpha_0 \leftarrow$$

Epoch	α
1	0.1
2	0.67
3	0.5
4	0.4
⋮	⋮

$X^{\{1\}}$ $X^{\{2\}}$ $\cdots$

→ epoch 1
→ epoch 2
⋮

$\alpha_0 = 0.2$

decay-rate = 1

Andrew Ng
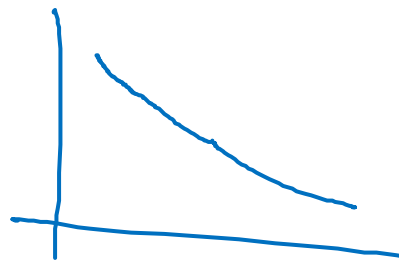
Other learning rate decay methods

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

formula

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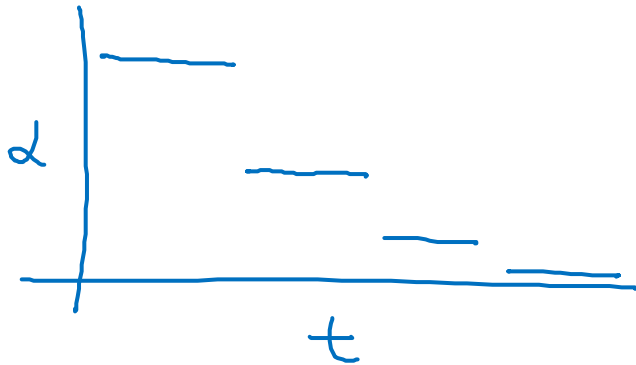
discrete staircase

Manual decay.