deeplearning.ai
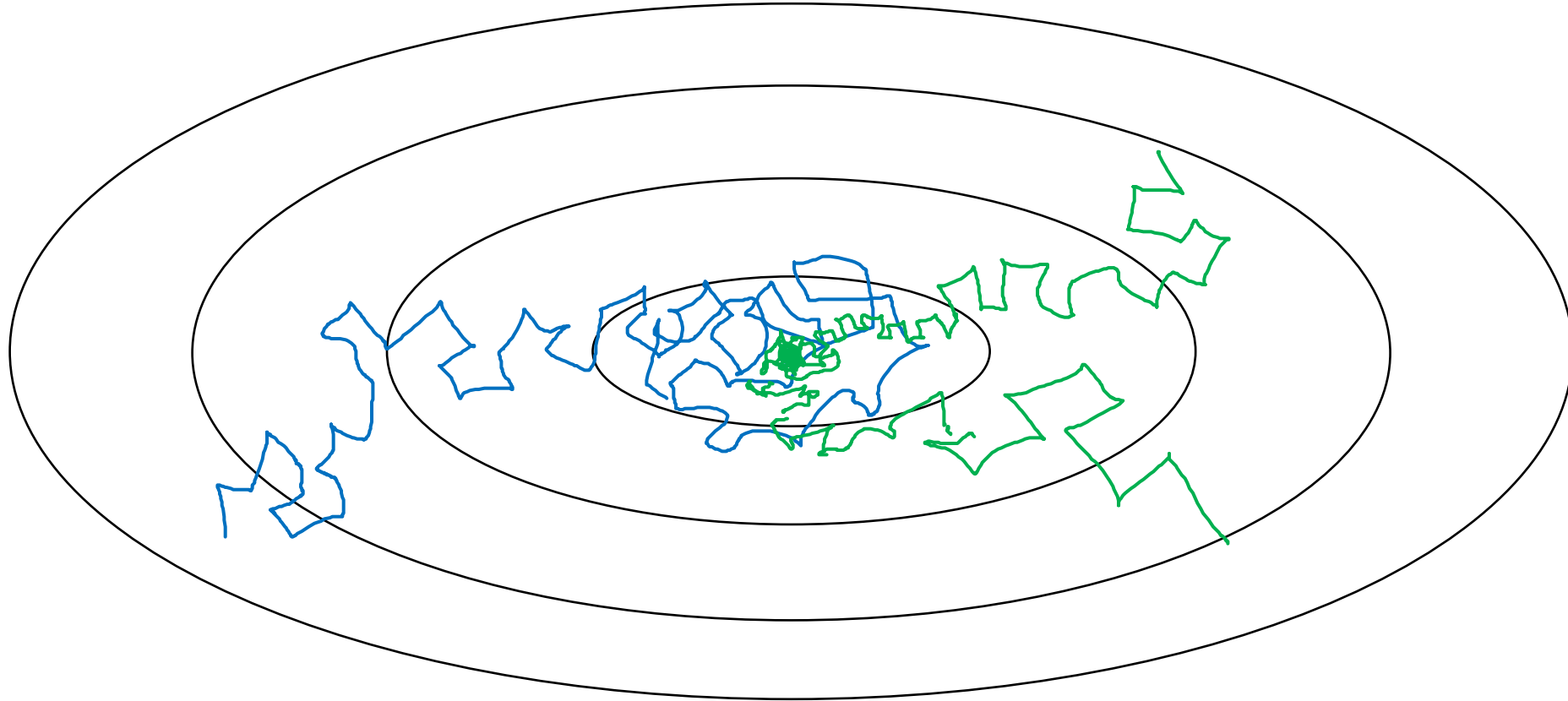
Optimization Algorithms

Learning rate decay

Learning rate decay

Slowly reduce α

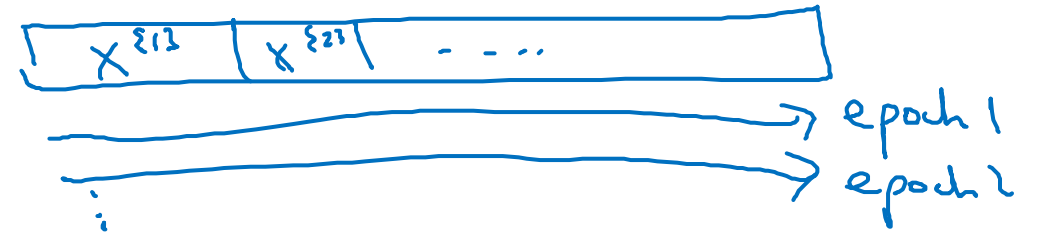# Learning rate decay

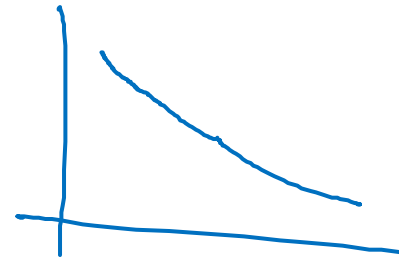1 epoch = 1 pass through data.

$$\alpha = \frac{1}{1 + decay\text{-}rate * epoch\text{-}num} \, \underline{\alpha_o} \; \longleftarrow$$

| $x^{\{1\}}$ | $x^{\{2\}}$ | - - - - |

→ epoch 1
→ epoch 2
⋮

$\alpha_o = 0.2$
decay-rate = 1

Epoch	α
1	0.1
2	0.67
3	0.5
4	0.4
⋮	⋮

Other learning rate decay methods

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

formula

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

discrete staircase

Manual decay.