Optimization Algorithms

Learning rate decay

deeplearning.ai

Learning rate decay
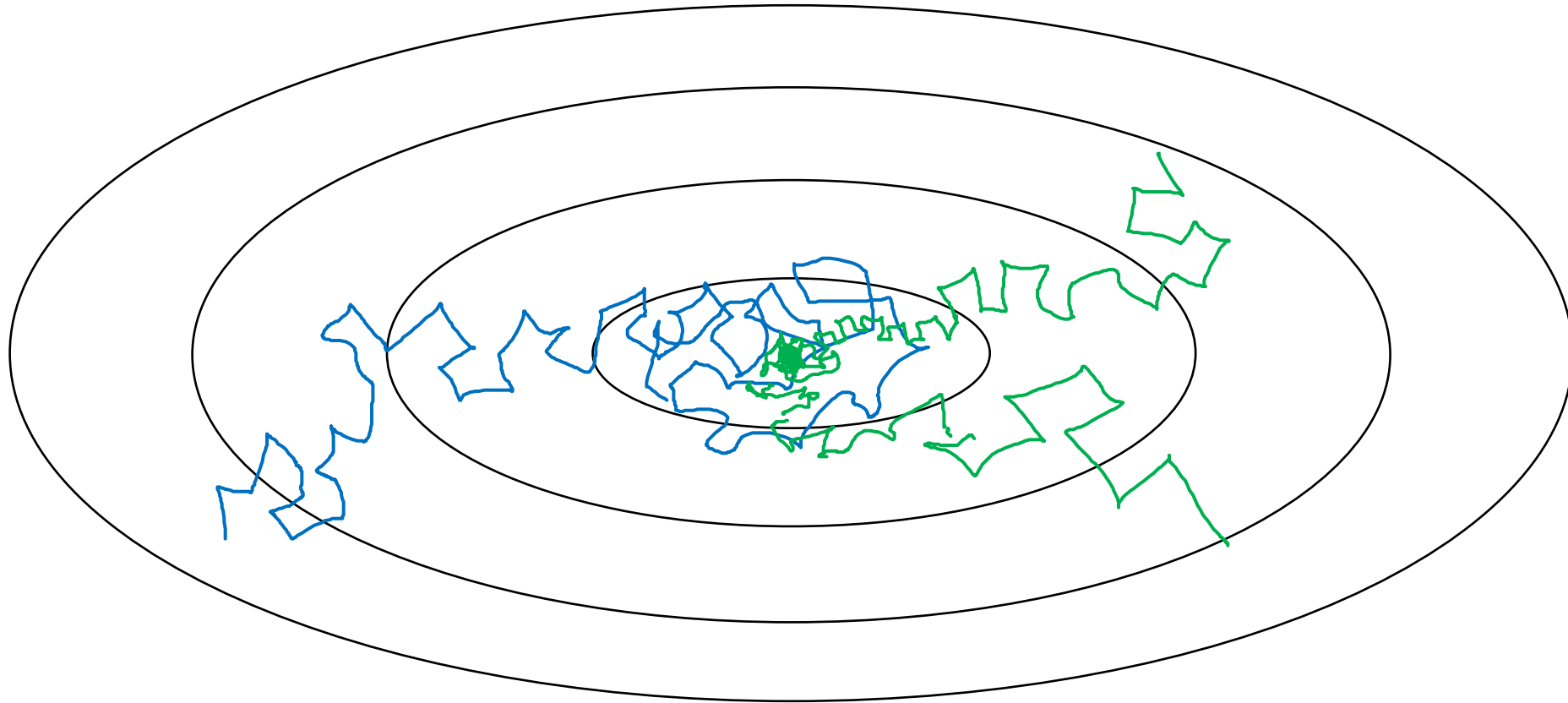
Learning rate decay

1 epoch = 1 pass through data.

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

Epoch	α
1	0.1
2	0.67
3	0.5
4	0.4
⋮	⋮

The values for alpha are not correctly calculated

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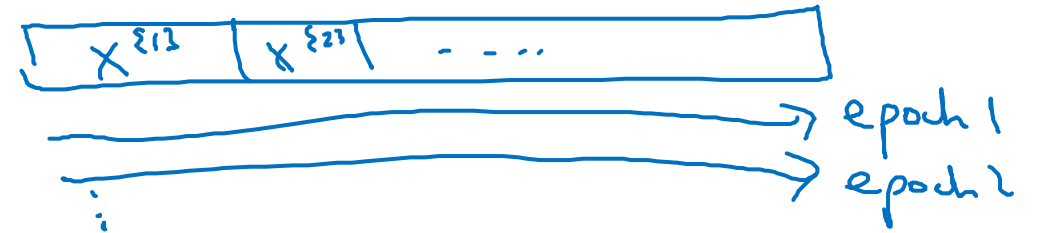
→ epoch 1
→ epoch 2

$\alpha_0 = 0.2$

decay.rate = 1
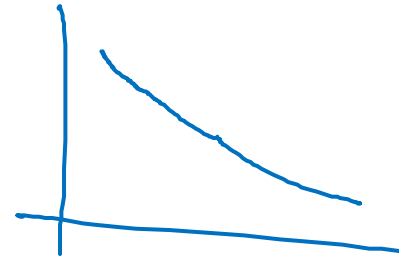
Other learning rate decay methods

formula $\begin{cases} & \\ & \\ & \\ & \\ & \\ & \\ & \end{cases}$

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

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

α |‾‾‾‾

discrete staircase

t

Manual decay.