

Week 2-3 Learning Rate Decay

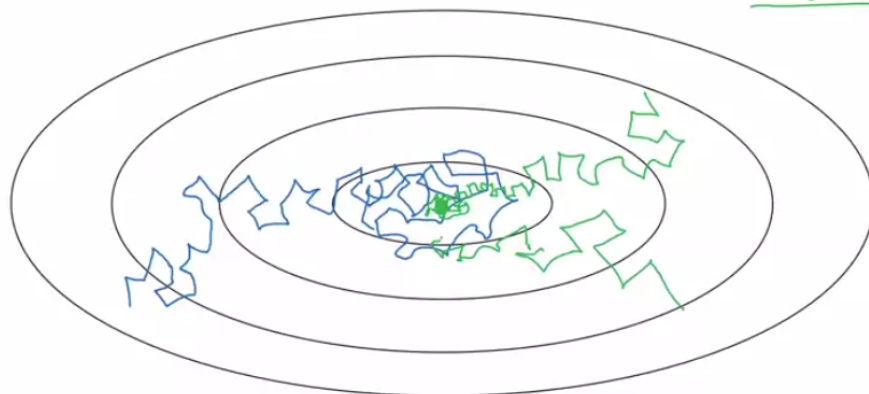
笔记本: DL 2 - Deep NN Hyperparameter Tuning, Regularization & Optimization

创建时间: 2021/1/9 11:28

更新时间: 2021/1/9 11:38

Learning rate decay

Slowly reduce α



Learning rate decay

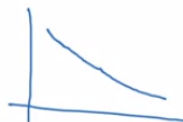
1 epoch = 1 pass through data.

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

Epoch	α
1	0.1
2	0.067
3	0.05
4	0.04
\vdots	\vdots



$\alpha_0 = 0.2$
 $\text{decay-rate} = 1$




0.1 -> 0.067 (typo)

other decay formulas

Other learning rate decay methods

formula {

$$\alpha = 0.95^{\text{epoch-num}} \cdot \alpha_0 \quad - \text{exponentially decay.}$$
$$\alpha = \frac{k}{\sqrt{\text{epoch-num}}} \cdot \alpha_0 \quad \text{or} \quad \frac{k}{\sqrt{t}} \cdot \alpha_0$$


discrete staircase

Manual decay.

For me, I would say that learning rate decay is usually lower down on the list of things I try.