

9. Learning Rate Decay

Tags

[Learning rate decay](#)

[Other learning rate decay methods](#)

Clarification about Learning Rate Decay Video

Please note that in the next video, at time 3:35, the values for alpha should be:

Epoch 1: alpha 0.1

Epoch 2: alpha 0.067

Epoch 3: alpha 0.05

Epoch 4: alpha 0.04

The formula for learning rate decay is:

$$\alpha = \frac{1}{1 + \text{decayRate} \times \text{epochNumber}} \alpha_0$$

Learning rate decay
1 epoch = 1 pass through data.


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

Epoch

Epoch	α
1	0.1
2	0.067
3	0.05
4	0.04
...	...

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

Diagram showing data points $x^{(1)}, x^{(2)}, \dots$ being processed through epochs 1 and 2.



Video player controls: 3:35 / 6:44, Andrew Ng

Learning rate decay

Learning rate decay

1 epoch = 1 pass through data.

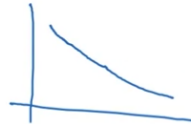
$$\alpha' = \frac{1}{1 + \text{decay-rate} * \text{epoch-num}} \alpha_0$$

Epoch	α
1	0.1
2	0.67
3	0.5
4	0.4
\vdots	\vdots



$$\alpha_0 = 0.2$$

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



Andrew Ng

Other learning rate decay methods

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

Andrew Ng