



deeplearning.ai

Optimization Algorithms

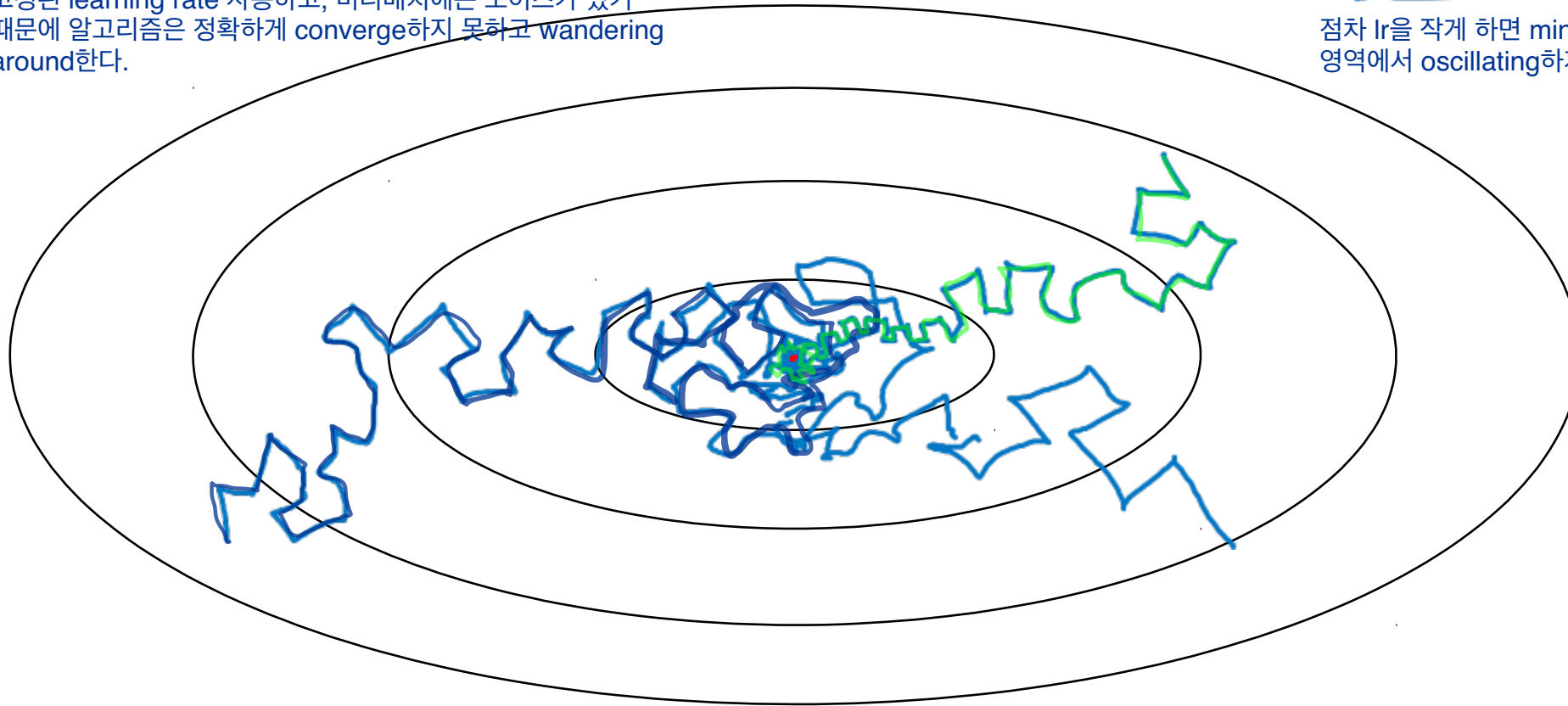
Learning rate
decay

Learning rate decay

고정된 learning rate 사용하고, 미니배치에는 노이즈가 있기 때문에 알고리즘은 정확하게 converge하지 못하고 wandering around한다.

Slowly reduce α

점차 lr을 작게 하면 minimum 주변 아주 타이트한 영역에서 oscillating하게 됨.



Learning rate decay

1 epoch = 1 pass through data.

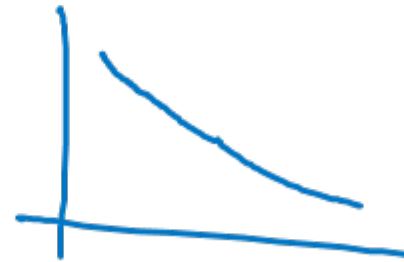
$$\alpha = \frac{1}{1 + \text{decay-rate} * \text{epoch-num}} \frac{\alpha_0}{\text{initial lr}}$$

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




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

decay rate



Other learning rate decay methods

formula {

- $\alpha = 0.95^{\text{epoch-num}} \cdot \alpha_0$ — exponentially decay.
exponentially decay
- $\alpha = \frac{k}{\sqrt{\text{epoch-num}}} \cdot \alpha_0$ or $\frac{k}{\sqrt{t}} \cdot \alpha_0$
-  discrete staircase
discrete staircase
- Manual decay.
manual decay