deeplearning.ai

Optimization Algorithms

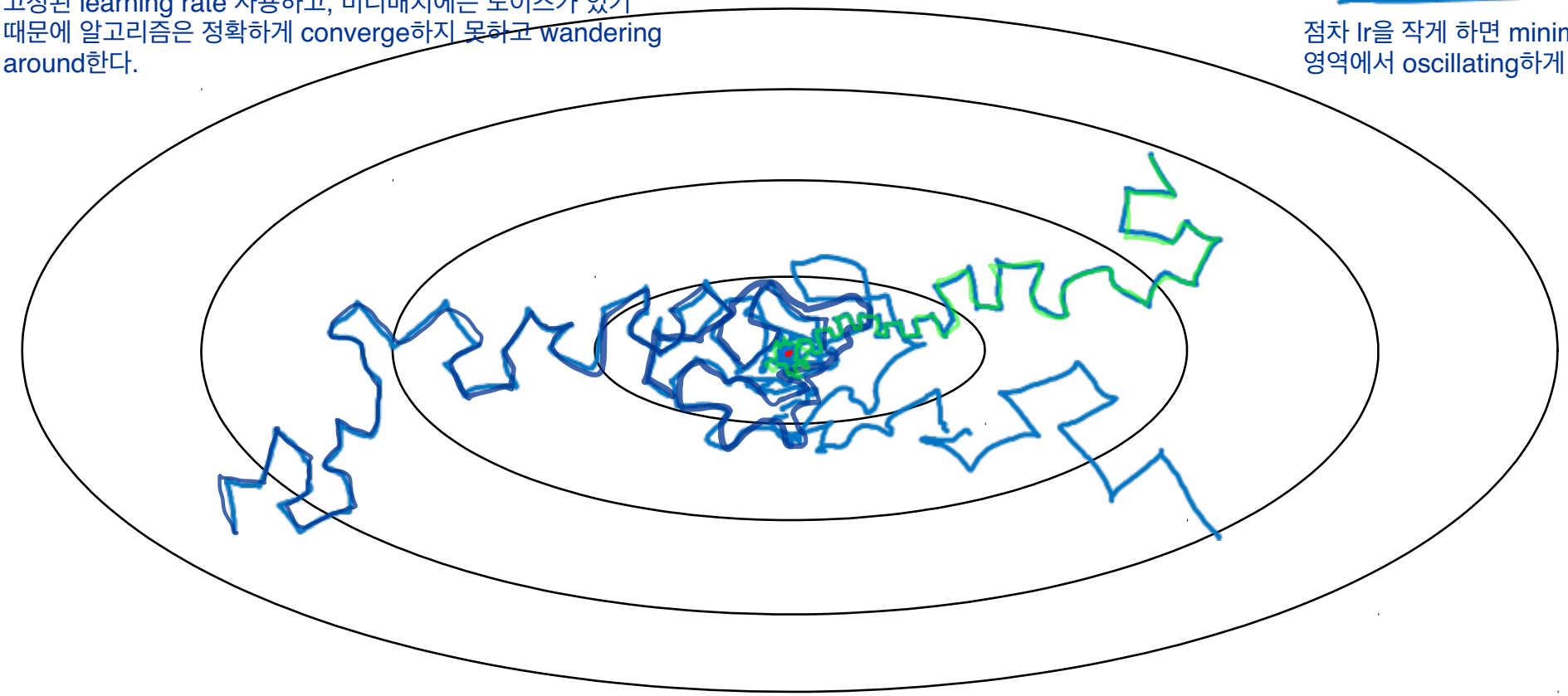
Learning rate decay

Learning rate decay

Slowly reduce α

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

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



Andrew Ng

Learning rate decay

1 epoch = 1 pass through data.

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

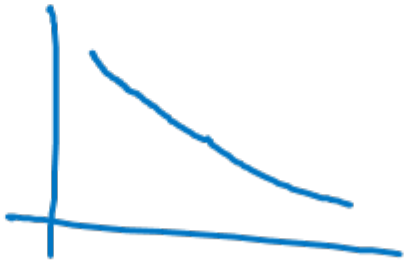
α_0: initial lr

Epoch	α
1	0.1
2	0.67
3	0.5
4	0.4
⋮	⋮

$\alpha_0 = 0.2$

decay_rate = 1

decay rate



Other learning rate decay methods