



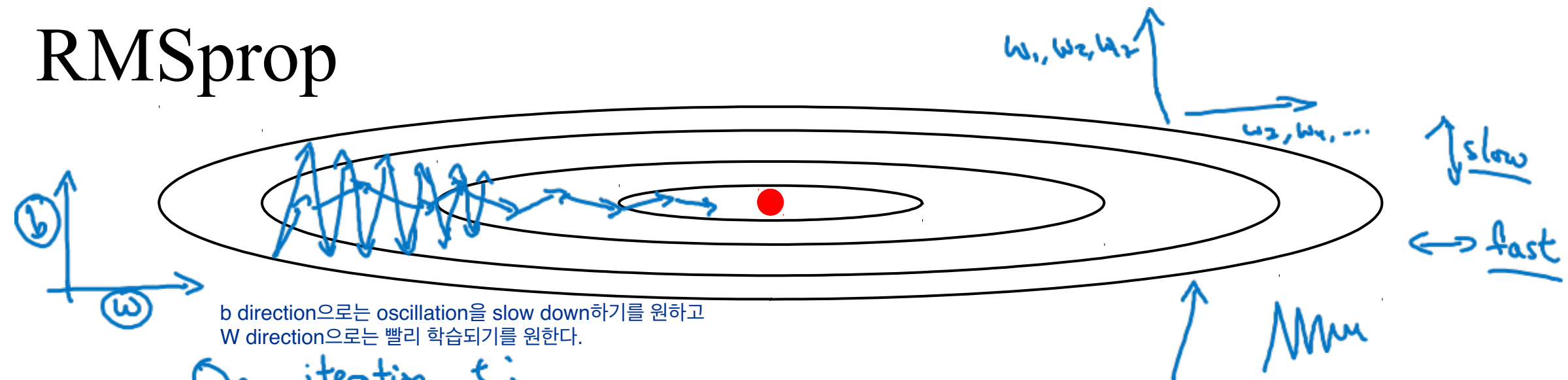
deeplearning.ai

Optimization Algorithms

RMSprop

root mean square prop

RMSprop



b direction으로는 oscillation을 slow down하기를 원하고
W direction으로는 빨리 학습되기를 원한다.

On iteration t :

Compute dw, db on current mini-batch

$$S_{dw} = \beta_2 S_{dw} + (1 - \beta_2) \underbrace{dw^2}_{\text{element-wise}} \leftarrow \text{small}$$

$$\rightarrow S_{db} = \beta_2 S_{db} + (1 - \beta_2) \underline{db^2} \leftarrow \text{large}$$

$$w := w - \frac{\alpha dw}{\sqrt{S_{dw} + \epsilon}} \leftarrow$$

$$b := b - \frac{\alpha db}{\sqrt{S_{db} + \epsilon}} \leftarrow$$

$$\epsilon = 10^{-8}$$

0으로 나누는걸 방지해서 계산의 안정성을
더하기 위해 epsilon term을 추가

고로 S_{dw} 는 작게, S_{db} 는 크게 되기를 원한다. 왜냐하면
아래 업데이트 식에서 나눠주는 sqrt term이 작으면
파라미터가 크게 변하고, 크면 작게 변하기 때문이다.