

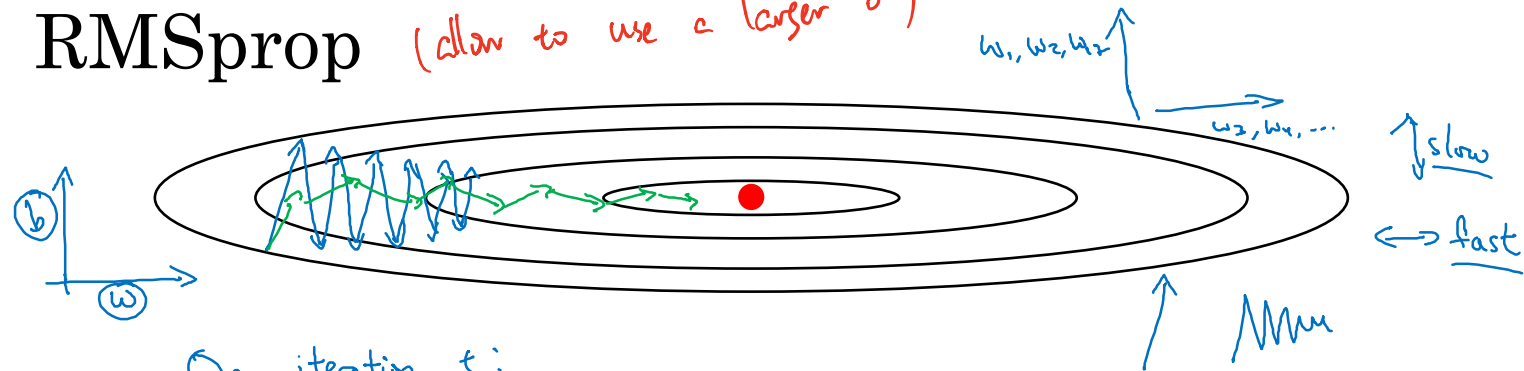


deeplearning.ai

Optimization Algorithms

RMSprop

RMSprop *(allow to use a larger α)*



On iteration t :

Compute dw, db on current mini-batch

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

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

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

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

$\epsilon = 10^{-8}$: prevent divided a very small num