deeplearning.ai

Optimization Algorithms

RMSprop

RMSprop

w_1, w_2, w_2

$w_3, w_4, \ldots$

slow

fast
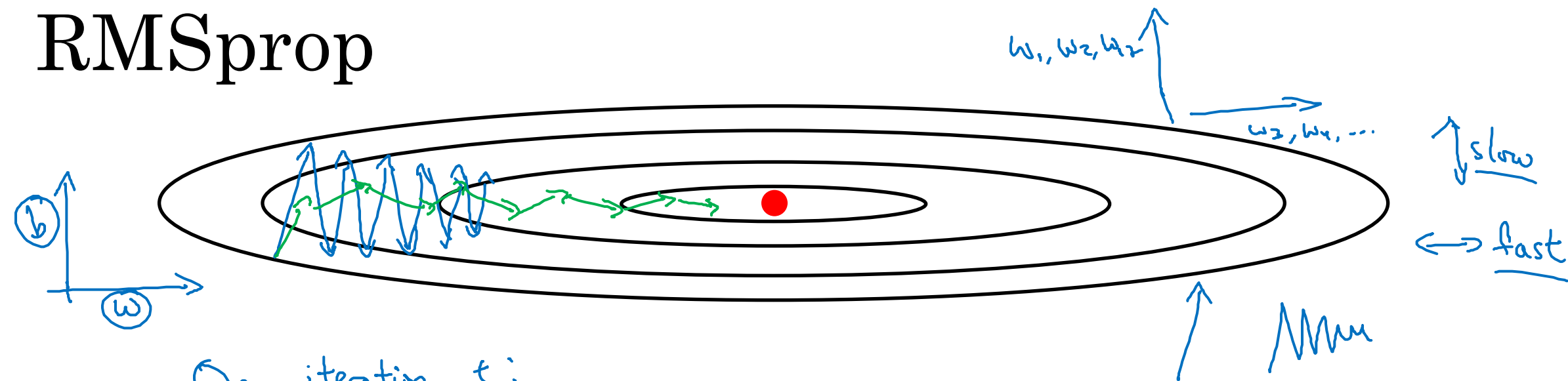
On iteration t:

Compute dW, db on current mini-batch

$$\underline{S_{dW}} = \beta_2 S_{dW} + (1-\beta_2) \underline{dW^2} \longleftarrow \text{small} \quad \text{element-wise}$$

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

$$W := W - \alpha \frac{dW}{\sqrt{S_{dW}+\varepsilon}} \leftarrow$$

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

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

Andrew Ng