



deeplearning.ai

# Optimization Algorithms

---

Understanding  
exponentially  
weighted averages

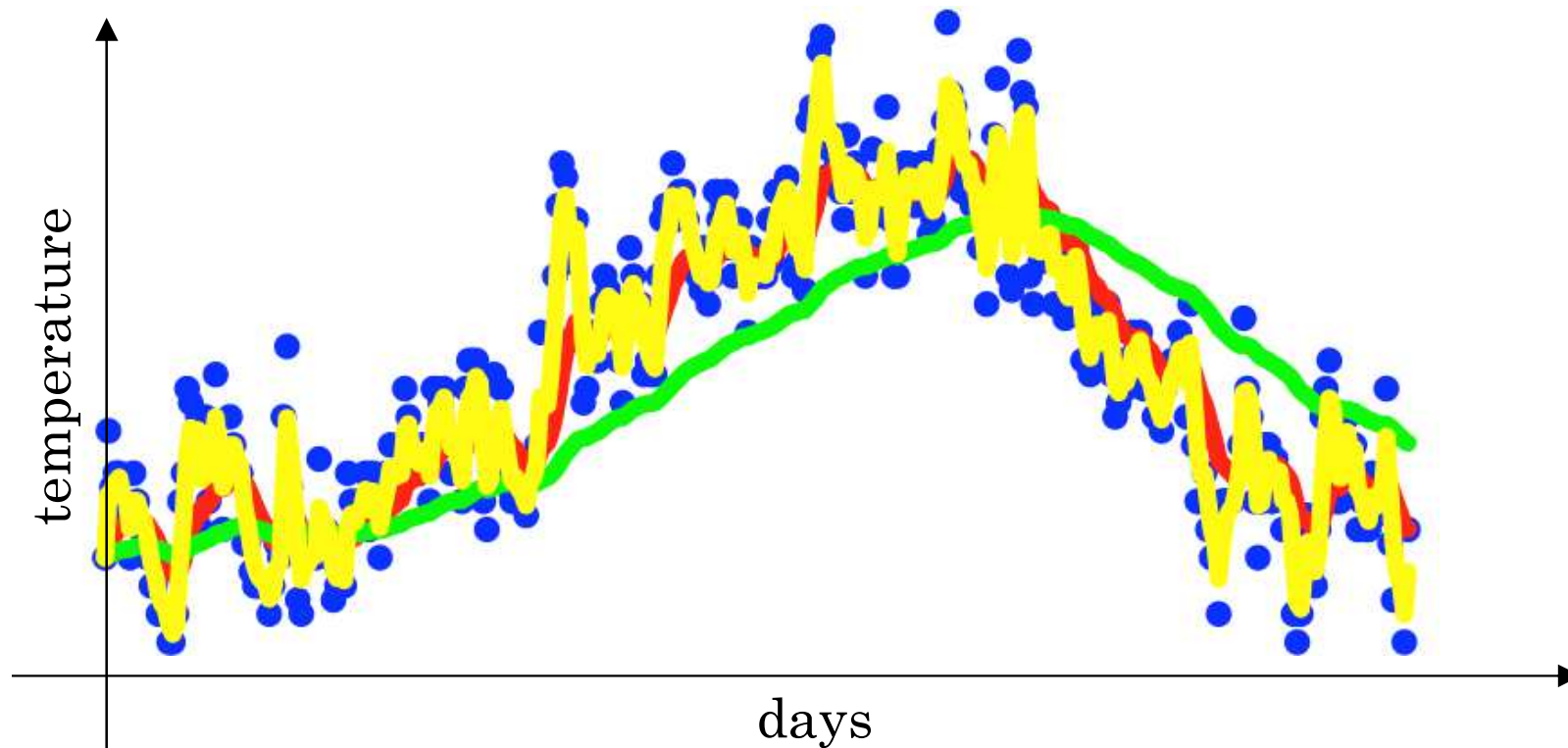
# Exponentially weighted averages

$$v_t = \beta v_{t-1} + (1 - \beta) \theta_t$$

$$\beta = 0.9$$

$$0.98$$

$$0.5$$



# Exponentially weighted averages

$$v_t = \beta v_{t-1} + (1 - \beta) \theta_t$$

$$v_{100} = 0.9v_{99} + 0.1\theta_{100}$$

$$v_{99} = 0.9v_{98} + 0.1\theta_{99}$$

$$v_{98} = 0.9v_{97} + 0.1\theta_{98}$$

...

$$\begin{aligned} \rightarrow v_{100} &= 0.1\theta_{100} + 0.9 \cancel{v_{99}} (0.1\theta_{99} + 0.9 \cancel{v_{98}}) \\ &= \underbrace{0.1\theta_{100}} + \underbrace{0.1 \times 0.9 \cdot \theta_{99}} + \underbrace{0.1 (0.9)^2 \theta_{98}} + \underbrace{0.1 (0.9)^3 \theta_{97}} + \underbrace{0.1 (0.9)^4 \theta_{96}} + \dots \end{aligned}$$

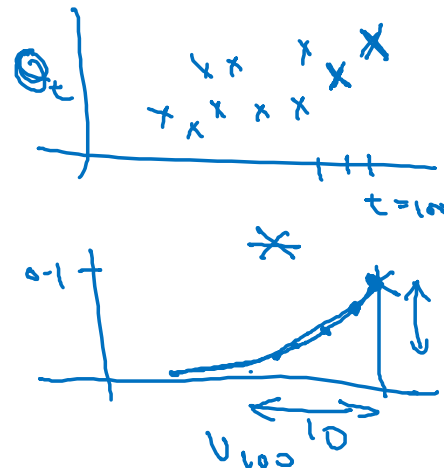
$$\underbrace{0.9^{10}}_{\approx 0.35} \approx \frac{1}{e}$$

$$\frac{(1-\epsilon)^{1/\epsilon}}{0.9} \approx \frac{1}{e}$$

$$0.98^?$$

$$\epsilon = 0.02 \rightarrow \underbrace{0.98^{50}}_{\approx \frac{1}{e}}$$

Andrew Ng



$$\approx \frac{1}{1-\beta}$$

$$\epsilon = 1-\beta$$

$$0.1\theta_{99} + 0.9v_{99}$$

# Implementing exponentially weighted averages

$$v_0 = 0$$

$$v_1 = \beta v_0 + (1 - \beta) \theta_1$$

$$v_2 = \beta v_1 + (1 - \beta) \theta_2$$

$$v_3 = \beta v_2 + (1 - \beta) \theta_3$$

...

$$V_0 := 0$$

$$V_0 := \beta v + (1 - \beta) \theta_1$$

$$V_0 := \beta v + (1 - \beta) \theta_2$$

⋮

---

$$\rightarrow V_0 = 0$$

Repeat {

Get next  $\theta_t$

$$V_0 := \beta V_0 + (1 - \beta) \theta_t \leftarrow$$

}