deeplearning.ai

Optimization Algorithms

Exponentially weighted averages

(Better Algos than Gradient Desc)

Temperature in London

(Pre-Req to better Algos than Gradient descent)

$\theta_1 = 40°\text{F}$ (Jan)

$\theta_2 = 49°\text{F}$

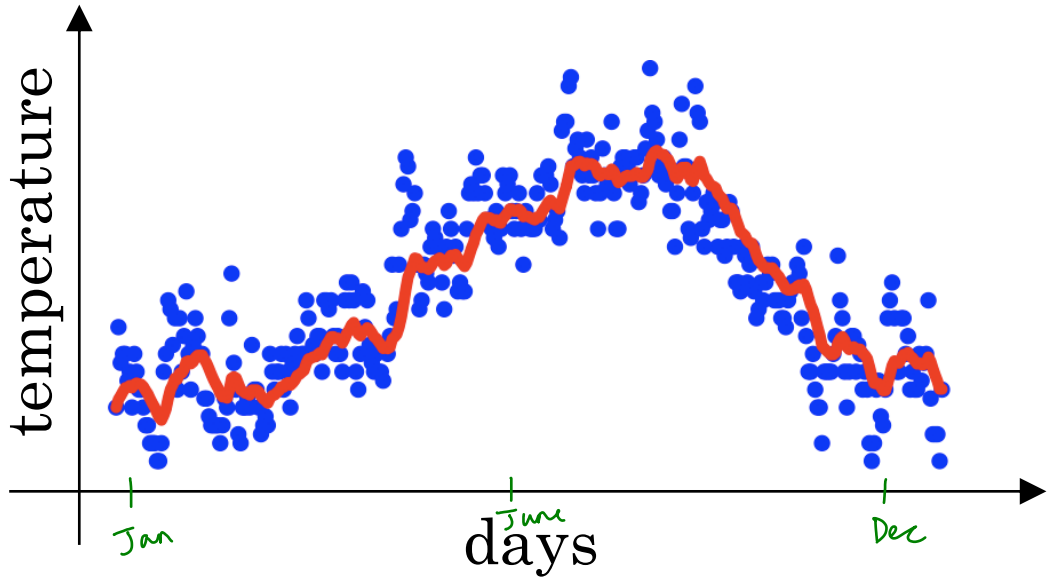
$\theta_3 = 45°\text{F}$

$\vdots$

$\theta_{180} = 60°\text{F}$

$\theta_{181} = 56°\text{F}$ (June)

$\vdots$



$V_0 = 0$

$V_1 = 0.9V_0 + 0.1\theta_1$

$V_2 = 0.9V_1 + 0.1\theta_2$

$\vdots$

$V_t = 0.9V_{t-1} + 0.1\theta_t$

Exponentially weighted avg.

Daily Avg or Moving Avg

Exponentially weighted averages

$V_t = \beta V_{t-1} + (1-\beta)\theta_t$

$\beta = .9$

V_t = approx avg over $\frac{1}{1-\beta}$ days temp

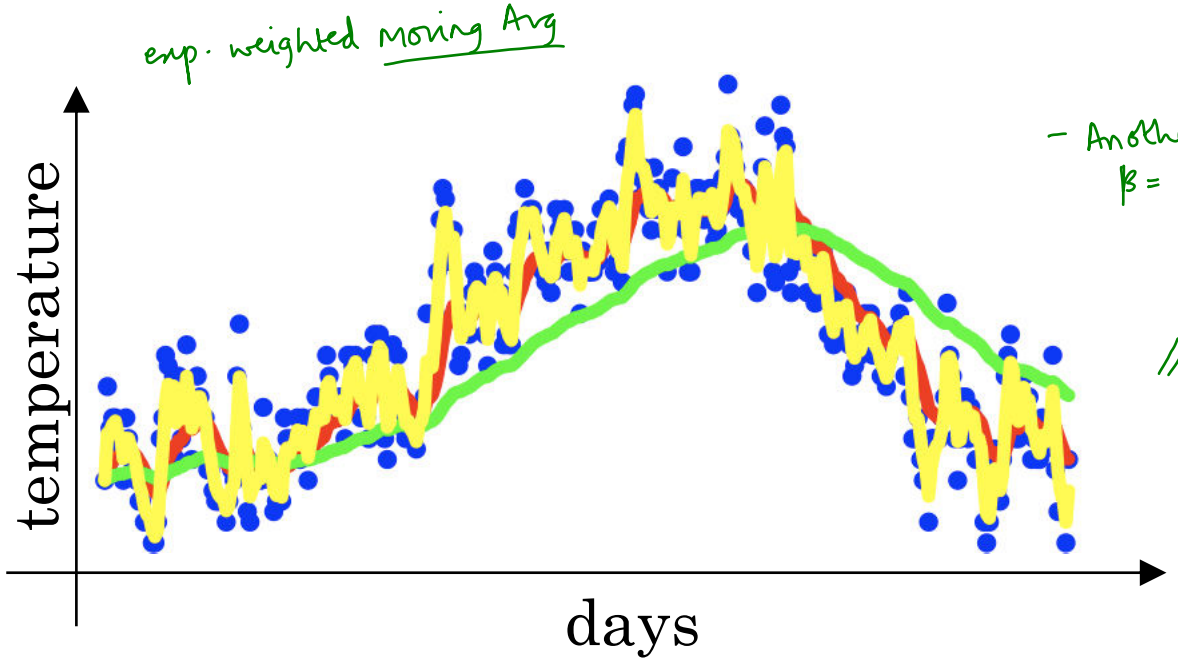
$\Rightarrow$ approx avg over $\frac{1}{1-0.9} = 10$ days temp (Red plot)

If $\beta = 0.98$

$\frac{1}{1-\beta} = 50$, then we would get avg of the last 50 days temp. (Green plot)

$\Rightarrow$ when $\beta \uparrow$, curve becomes smoother

Why? Averages over a larger window, less susceptible to 1 day's erratic weather change compared to 10 day Avg



- Another way to see this

$\beta = .98 \Rightarrow .98$ weight is assigned to the previous values

.02 weight is assigned to today's value

//ly, If $\beta = .5$, Avg over 2 days $\Rightarrow$ much more noisy (yellow) can Adapt more quickly to past/present temp change