deeplearning.ai

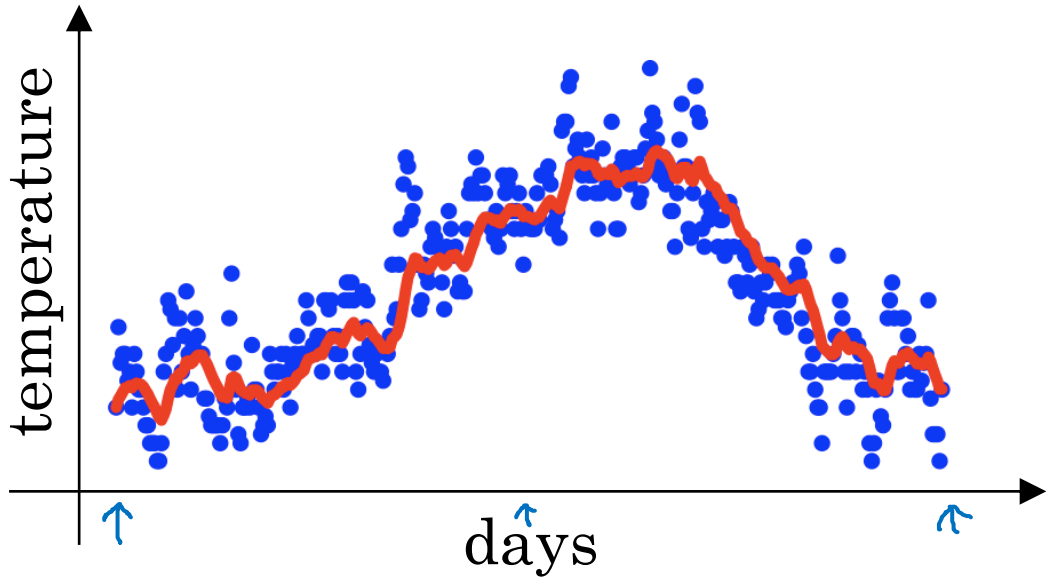
Optimization Algorithms

Exponentially weighted averages

Also called exponentially weighted moving averages in statistics.

Temperature in London

$\theta_1 = 40°\text{F}$ 4°C ←
$\theta_2 = 49°\text{F}$ 9°C
$\theta_3 = 45°\text{F}$ ⋮
⋮
$\theta_{180} = 60°\text{F}$ 15°C
$\theta_{181} = 56°\text{F}$ ⋮
⋮



$v_0 = 0$
$v_1 = 0.9\, v_0 + 0.1\, \theta_1$
$v_2 = 0.9\, v_1 + 0.1\, \theta_2$
$v_3 = 0.9\, v_2 + 0.1\, \theta_3$
⋮
$v_t = 0.9\, v_{t-1} + 0.1\, \theta_t$

Andrew Ng

Exponentially weighted averages

Moving

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

$\beta = 0.9$: $\approx$ 10 days' temperature.

$\beta = 0.98$: $\approx$ 50 days

$\beta = 0.5$: $\approx$ 2 days

EWA jus adapts slowly when beta is so large.

V_t as approximately average over

$\rightarrow \approx \frac{1}{1-\beta}$ days' temperature.

$\frac{1}{1-0.98} = 50$