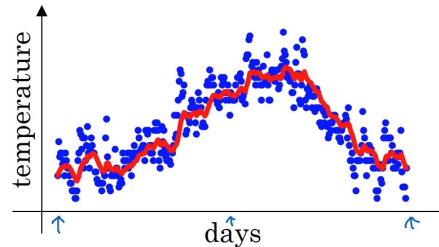


Exponentially weighted averages

Temperature in London

$\theta_1 = 40^\circ\text{F}$ 4°C \leftarrow
 $\theta_2 = 49^\circ\text{F}$ 9°C
 $\theta_3 = 45^\circ\text{F}$ \vdots
 \vdots
 $\theta_{180} = 60^\circ\text{F}$ 15°C
 $\theta_{181} = 56^\circ\text{F}$ \vdots
 \vdots



$$\begin{aligned}
 v_0 &= 0 \\
 v_1 &= 0.9 v_0 + 0.1 \theta_1 \\
 v_2 &= 0.9 v_1 + 0.1 \theta_2 \\
 v_3 &= 0.9 v_2 + 0.1 \theta_3 \\
 &\vdots \\
 v_t &= 0.9 v_{t-1} + 0.1 \theta_t
 \end{aligned}$$

Andrew Ng

The red line is the exponentially weighted average

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

Here you can think of v_t as averaging over the past $1/(1-\beta)$ days' temperature.

v_t is approximately
 average over
 $\approx \frac{1}{1-\beta}$ days'
 temperature.

High beta \rightarrow (eg: green line), the plot is smoother because it keeps track of a longer range of days' temperatures. Adapts slowly to newer temperatures

Very low beta \rightarrow (eg: red line), very noisy yellow line since you are very susceptible to new temperatures.

Exponentially weighted averages

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

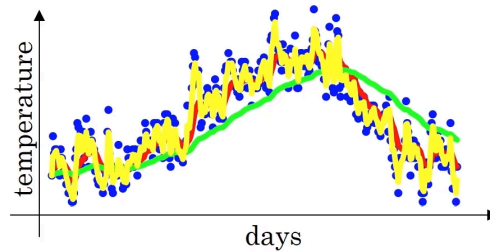
$\beta = 0.9$: ≈ 10 days' dependency

$\underline{\beta = 0.98}$: ≈ 50 days

$\beta = 0.5$: ≈ 2 days

V_e is approximately
 Over a other
 $\rightarrow \frac{1}{1-\beta}$ days'
 temperature.

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



Andrew Ng

Intuition for EWAs

Exponentially weighted averages

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

$$\begin{aligned}
 v_{100} &= 0.9v_{99} + 0.1\theta_{100} \\
 v_{99} &= 0.9v_{98} + 0.1\theta_{99} \\
 v_{98} &= 0.9v_{97} + 0.1\theta_{98} \\
 &\dots \\
 \rightarrow v_{100} &= 0.1\theta_{100} + 0.9 \cancel{v_{99}} (0.1\theta_{99} + 0.9 \cancel{v_{98}}) \\
 &= 0.1\theta_{100} + \underbrace{0.1 \times 0.9 \cdot \theta_{99}} + \underbrace{0.1 (0.9)^2 \theta_{98}} + \underbrace{0.1 (0.9)^3 \theta_{97}} + \underbrace{0.1 (0.9)^4 \theta_{96}} + \dots \\
 0.9^{10} &\approx 0.35 \approx \frac{1}{e} \\
 \frac{(1-\varepsilon)^{1/\varepsilon}}{\varepsilon} &\approx \frac{1}{e} \quad \varepsilon = 0.02 \rightarrow 0.98^{50} \approx \frac{1}{e}
 \end{aligned}$$

Andrew Ng

Explanation of the graphs on the top right

Top graph: just the temperature measurements over time.

Bottom one: exponentially decaying function of weightage factor for each of the elements on top.

So v_{100} = element wise product of those two graphs and sum them up