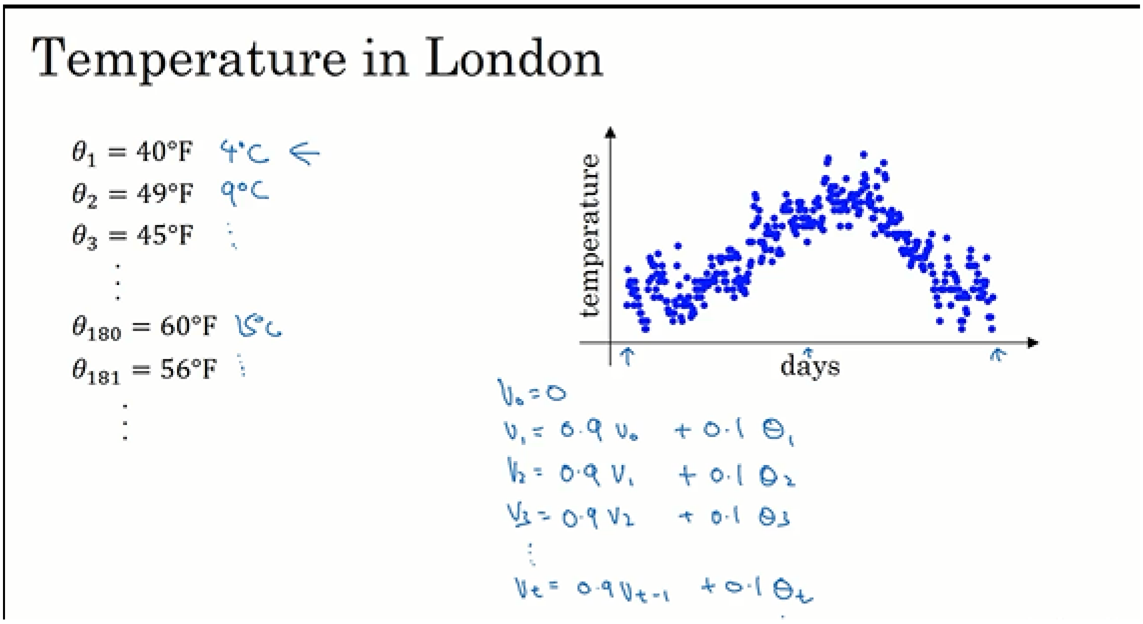
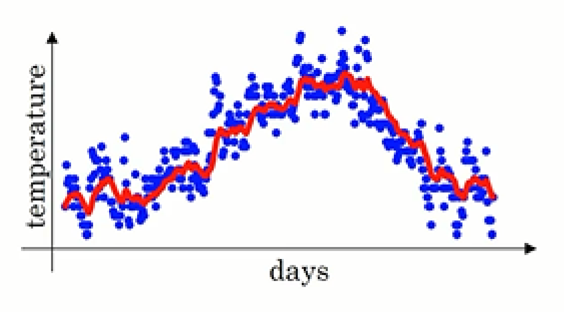
Exponentially weighted averages

Example



- $v_t = \beta v_{t-1} + (1-\beta)\theta_t$
  - $v_t\ as\ approx.\ average\ over\ \approx \frac{1}{1-\beta}\ days\ temperature$
  - $\beta = 0.9,\ approx.\ averg.\ over\ 10\ days$
  - β **describes how much to rely on the previous values, and how well to adapt to the present changes**
  - 

Understanding exponentially weighted averages

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

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

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

$\dots$

$v_{100} = 0.1\theta_{100} + 0.1 \times 0.9\theta_{99} + 0.1 \times 0.9^2\theta_{98} + \dots +$

each data values weighted by a corresponding exponential function value

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

$0.9^{10} \approx \frac{1}{e} \approx 0.35,\ it\ takes\ 10\ days\ to\ decay\ to\ its\ 0.35,\ when\ \beta = 0.9$

Implementing exponentially weighted

$v_\theta = 0$

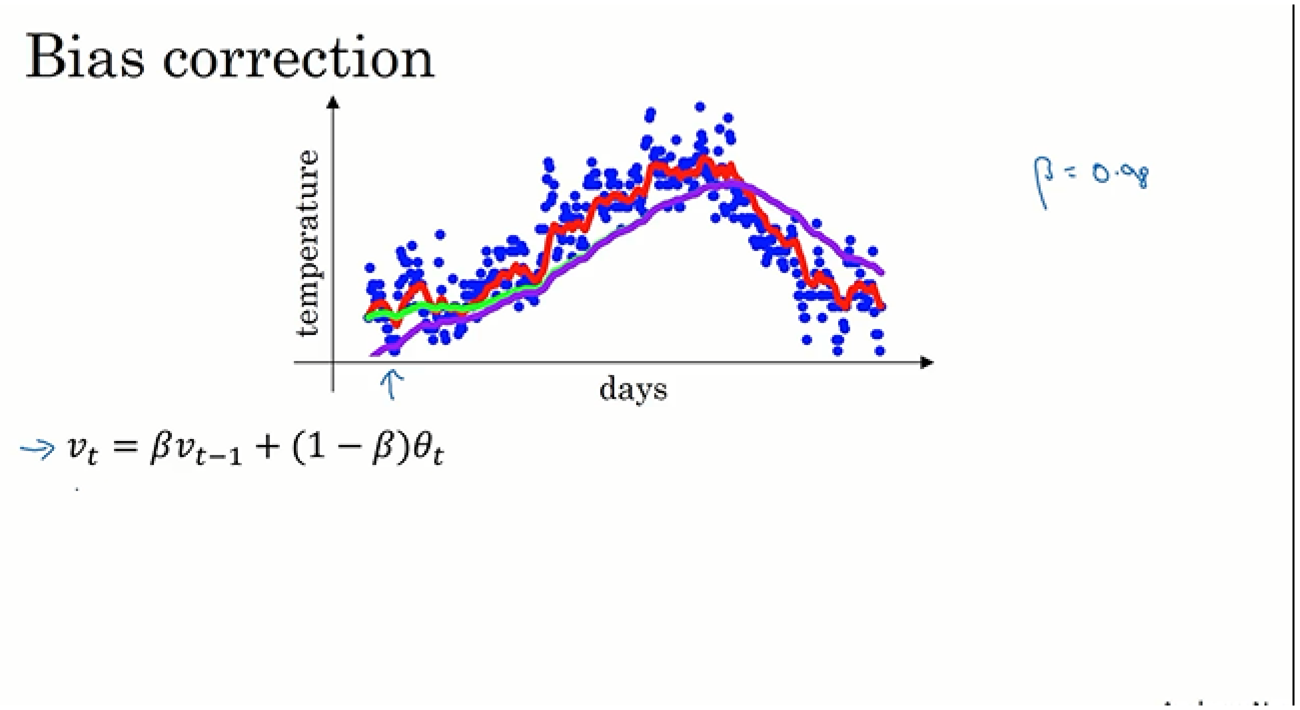
$repeat\ \{$

$get\ \theta_t$

$v_\theta := \beta v_\theta + (1-\beta)\theta_t$

$\}$

Bias correction in exponentially weighted averages



- the *red curve* plotted when $\beta = 0.9$
- the *green curve* plotted when $\beta = 0.98$
  - but actually, when $\beta = 0.98$, you will get the *purple curve*, which **starts a bit lower**, but when t is large, *purple* and *green* pretty much overlay
    - because
      - $v_0 = 0$
      - $v_1 = 0.98 * 0 + 0.02 * \theta_1 = 0.02 * \theta_1$
      - $v_2 = 0.98 \times 0.02 * \theta_1 + 0.02 * \theta_2$
      - $\ldots$
    - to fix this (remove the bias)

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