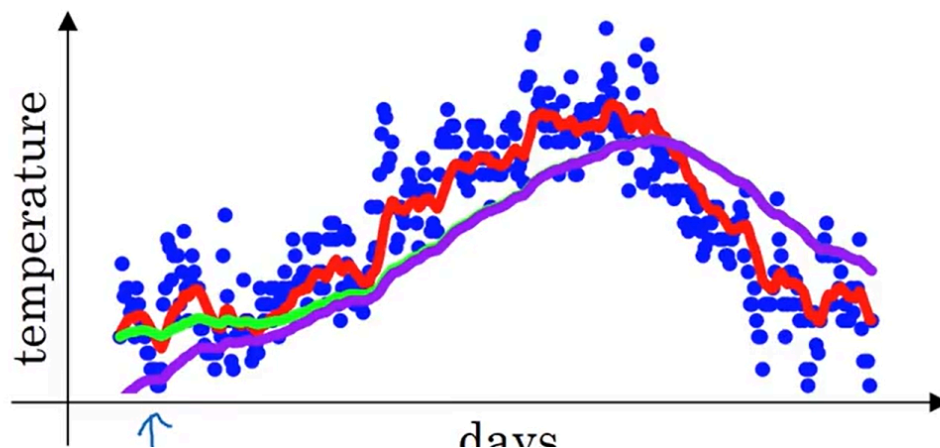


Bias correction in exponentially weighted averages

beta = 0.9 → red line

beta = 0.98 → green line (actually you don't get the green line, you get the purple line)

The purple line starts off low.

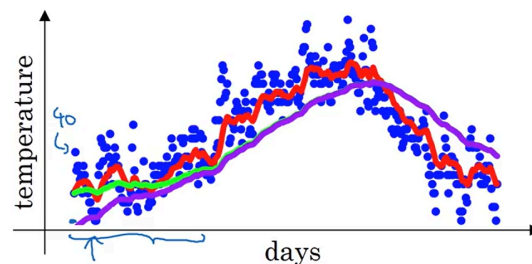


Initial EWA is bad, therefore use the following formula:

(essentially do the same EWA as before, but just divide each of those v_t 's by $(1 - \beta^t)$)

And also note that $1 - \beta^t$ is the sum of the coefficients of the summands, so it is a weighted average—it's at the scale of new temperatures.

Bias correction



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

$$v_0 = 0$$

$$v_1 = \cancel{0.98 v_0} + 0.02 \theta_1$$

$$\begin{aligned} v_2 &= 0.98 v_1 + 0.02 \theta_2 \\ &= 0.98 \times 0.02 \times \theta_1 + 0.02 \theta_2 \\ &= 0.0196 \theta_1 + 0.02 \theta_2 \end{aligned}$$

$$\frac{v_t}{1 - \beta^t}$$

$$t=2: 1 - \beta^t = 1 - (0.98)^2 = 0.0396$$

$$\frac{v_2}{0.0396} = \frac{0.0196 \theta_1 + 0.02 \theta_2}{0.0396}$$

Andrew Ng

Now, as t becomes large $1 - \beta^t$ will be 1 meaning bias correction is switched off.