Deep Learning

A course about theory & practice

Exponential Moving Average (EMA)

Marco Piastra

An aside: *moving averages*

Following non-stationary phenomena

- Average

Definition: $$\overline{v}_T := \frac{1}{T}\sum_{k=1}^{T} v_k$$

Running implementation:

$$\overline{v}_T = \frac{1}{T}(v_T + \sum_{k=1}^{T-1} v_k) = \frac{1}{T}(v_T + (T-1)\overline{v}_{T-1})$$

$$= \overline{v}_{T-1} + \frac{1}{T}(v_T - \overline{v}_{T-1}) = \frac{1}{T}\, v_T + (1-\frac{1}{T})\,\overline{v}_{T-1}$$

"the weight of newer observations diminishes with time"

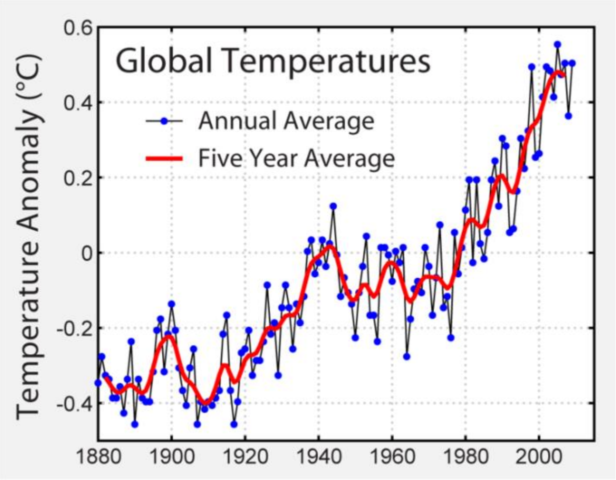
- Simple Moving Average (SMA)

$$\overline{v}_{T,n} := \frac{1}{n}\sum_{k=T-n}^{T} v_k$$

- Exponential Moving Average (EMA)

$$\overline{v}_{T,\alpha} := \alpha\, v_T + (1-\alpha)\,\overline{v}_{T-1,\alpha},\ \ \alpha \in [0,1]$$

"the weight of newer observations remains constant"



[image from wikipedia]

An aside: *moving averages*

- Exponential Moving Average (EMA)

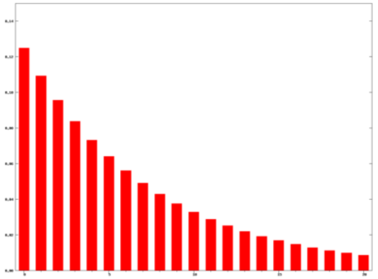
$$\overline{v}_{T,\alpha} := \alpha\, v_T + (1-\alpha)\, \overline{v}_{T-1,\alpha}, \;\; \alpha \in [0,1]$$

Expanding:

$$\begin{aligned}\overline{v}_{t,\alpha} &= \alpha\, v_t + (1-\alpha)\, \overline{v}_{t-1,\alpha} \\ &= \alpha\, v_t + (1-\alpha)(\alpha\, v_{t-1} + (1-\alpha)\overline{v}_{t-2,\alpha}) \\ &= \alpha\, v_t + (1-\alpha)(\alpha\, v_{t-1} + (1-\alpha)(\alpha\, v_{t-2} + (1-\alpha)\overline{v}_{t-3,\alpha})) \\ &= \alpha\,(v_t + (1-\alpha)\, v_{t-1} + (1-\alpha)^2\, v_{t-2}) + (1-\alpha)^3\, \overline{v}_{t-3,\alpha}\end{aligned}$$

$(1-\alpha)^{\Delta_t}$

"the weight of older observations diminishes with time"



[image from wikipedia]

The weight of past contributions decays as

$$(1-\alpha)^{\Delta_t}$$

A SMA with n previous values is approximately equal to an EMA with

$$\alpha = \frac{2}{n+1}$$