A Quickie on Exponential Smoothing

Javier R. Movellan

Let $X = X_1, \cdots,$ be independent identically distributed random variables with variance σ^2. We define the exponential running average of X with parameter $\alpha \in [0, 1]$ as follows

$$\hat{X}_t = \frac{1}{w_t}\alpha(X_t + (1-\alpha)X_{t-1} + \cdots + (1-\alpha)^{t-1}X_1) \tag{1}$$

where

$$w_t = \alpha(1 + (1-\alpha) + \cdots + (1-\alpha)^{t-1}) \tag{2}$$

1 Properties

Relation with running average

$$\lim_{\alpha \to 0} \hat{X}_t = \frac{\sum_{i=1}^{t} X_t}{t} \tag{3}$$

Recursive Update Equations

$$\hat{X}_t = \frac{1}{w_t}\bar{X}_t \tag{4}$$

where

$$\bar{X}_0 = 0 \tag{5}$$

$$w_0 = 0 \tag{6}$$

$$\bar{X}_t = \alpha X_t + (1-\alpha)\bar{X}_{t-1}, \quad \text{for } t > 0 \tag{7}$$

$$w_t = \alpha + (1-\alpha)w_{t-1}, \quad \text{for } t > 0 \tag{8}$$

Equivalent Variance Note

$$Var(\hat{X}_t) = \frac{\sigma^2\alpha^2}{w_t^2}\left(\sum_{s=0}^{t-1}(1-\alpha)^{2s}\right) \tag{9}$$

Using the following property about the sum of geometric series

$$a^0 + a^1 + a^2 + \cdots + a^{n-1} = \frac{1-a^n}{1-a} \tag{10}$$

we get that

$$w_t = 1 - (1-\alpha)^t \tag{11}$$

$$\sum_{s=0}^{t-1}(1-\alpha)^{2s} = \frac{1-(1-\alpha)^{2t}}{1-(1-\alpha)^2} \tag{12}$$

We can now compare this to the variance of a uniform averager of size n

$$Var(\frac{1}{n_t}\sum_{s=1}^{n_t} X_s) = \frac{\sigma^2}{n_t} \tag{13}$$

we want it to have to find the valu of n_t that would give the same variance as an exponential smoother with t observations. Thus

$$\frac{1}{n_t} = \frac{\alpha^2}{w_t^2}\left(\sum_{s=0}^{t-1}(1-\alpha)^{2s}\right) \tag{14}$$

and

$$n_t = \frac{(1-(1-\alpha)^t)^2(1-(1-\alpha)^2)}{\alpha^2(1-(1-\alpha)^{2t})} \tag{15}$$

Figure 1 shows the equivalent n_t as a function of time t for different values of α.

Asymptotics

- $\lim_{t\to\infty} w_t = 1$.
  This property tells us that for large t we don't really need to worry about the denominator w_t since its value will converge to 1. Thus, for large t we have $\alpha X_t + (1-\alpha)\hat{X}_{t-1}$
- $\lim_{t\to\infty} n_t = \frac{2-\alpha}{\alpha}$ This is useful to calculate the value of α that corresponds to a given uniform smoother

$$\alpha = \frac{2}{n+1} \tag{16}$$

- $\lim_{t\to\infty} \alpha \left(\sum_{s=0}^{t-1} (t-s)(1-\alpha)^s \right) = t - \frac{1-\alpha}{\alpha}$
  This property tells us that the expected lag for an exponential smoother is $(1-\alpha)/\alpha$

Appendix: Effective Number of Observation for Weighted Averages

Here we present some heuristics for computing the number of observations effectively used in a weighted average. Let

$$\bar{X}_t = \frac{\sum_{s=1}^{t} X_s w_s}{\sum_{s=1}^{t} w_s} \tag{17}$$

where $w_s \geq 0$ are fixed weights. Assuming independent identically distributed observations we have that

$$Var(\bar{X}_t) = \sigma^2 \frac{\sum_{s=0}^{t} w_s^2}{\left(\sum_{s=0}^{t} w_s \right)^2} \tag{18}$$

The variance of an average based on n observations with equal weights is σ^2/n thus, the effective number of observations in the weighted statistic is

$$n = \frac{\left(\sum_{s=0}^{t} w_s \right)^2}{\sum_{s=0}^{t} w_s^2} \tag{19}$$

For example, if all the weights are zero except for n weights that are 1, then the effective number of observations is n. If all the weights are zero, except for 10 weights that are $10, 9, \cdots 1$, then the effective number of observations is 7.857.

This heuristic does not capture the fact that the value of a weight, not its ratio, at times has meaning. For example, for a single observation with weight larger than zero the effective n is 1, regardless of how large that weight is.

2 History

- The first version of this document was written by Javier R. Movellan in August 2004, as part of the Kolmogorov project.

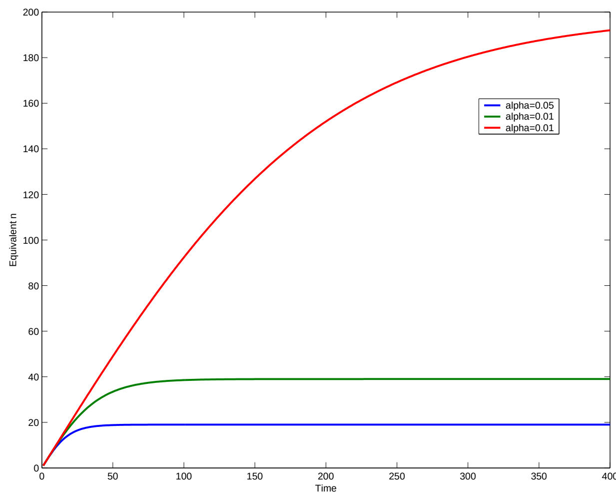

Figure 1: *Equivalent n_t for 3 different values of α.*