An adaptable, convergent moving average for online machine-learning

David Owen*

February 26, 2021

Abstract

A weight scheme is presented for a moving average that can adapt to early changes (like an EMA), but converges (like a simple average). Online forms for both the mean and variance are given. The proposed average is well-suited for machine-learning applications, such as Q-learning.

1 Introduction

In problems where an expected value is calculated, it is sometimes useful to use an average that can "forget" older data. The weight scheme for an Exponential Moving Average (EMA) [1] can be considered, but its additional parameter and lack of convergence can cause other difficulties.

I present a weight scheme based on a window of a fixed proportion of the data seen so far, with the following properties:

- It "forgets" older data, like an EMA;
- It converges, like a simple average; and,
- It has a convenient online form, including for calculating the weighted sample variance.

2 The problem

There are some problems where we must take some kind of average, but we'd like the average to be able to forget earlier values that may be biased or inaccurate in some way.

For example, in the Q-learning algorithm by Watkins [3], early explorations from a particular state may lead through many low-scoring paths before considering any high-scoring ones. Even after the high-value paths are discovered,

*david@devquant.com

the values of the many low-value paths may dominate the state's expected value for some length of time, causing other less-optimal sibling states to be selected instead.

In cases like these, we might pick an EMA for the property that it conceptually has a window that gradually forgets older data, allowing the state's expected value to more quickly adjust to the newer information.

However, picking the window size can be hard. If it's too small, the algorithm will fail to account for important examples. If it's too large, it will adapt slowly. Some problems are hard to estimate (how many relevant training examples would the game of Go have, at any given time during training?), and some problems may actually change in size as training proceeds.

Even if the window size is well-tuned, the EMA's lack of convergence can cause a reinforcement-learning algorithm to constantly bounce around between many sub-optimal alternatives instead of settling on the best options.

3 Initial formulation

In the reinforcement-learning situation, as well as many others, averages are calculated using online-update formulations. For samples $x_1 \ldots x_n$ and a conceptual window size W, the average y_n can be calculated recursively as:

$$y_n = \left(1 - \frac{1}{W}\right) y_{n-1} + \frac{1}{W} x_n \tag{1}$$

For a simple average, $W = n$; for an EMA, $W = k$ where k is some constant.

Let's now assume instead that we want the window size to grow in proportion to the number of samples we've seen. As a concrete example, let's say that $W = \frac{n}{2}$, meaning the window is half our sample (the most recent half), giving:

$$y_n = \left(1 - \frac{2}{n}\right) y_{n-1} + \frac{2}{n} x_n \tag{2}$$

However, it would be convenient to have y_1 expressible in terms of only x_1 and not on any y_i. Shifting W by 1 in n such that

$$W = \frac{n+1}{2} \tag{3}$$

gives the desired result:

$$y_n = \left(1 - \frac{2}{n+1}\right) y_{n-1} + \frac{2}{n+1} x_n \tag{4}$$

4 Closed form and frequency weights

The recurrence relation (4) can be solved into a closed form. Maxima [2] gives the following:

$$y_n = \frac{2\left(\left(\sum_{i=2}^{n} i x_i\right) + x_1\right)}{n\,(n+1)} \tag{5}$$

The x_1 term is simply the summand at $i = 1$, and can be subsumed there:

$$y_n = \frac{2\sum_{i=1}^{n} i x_i}{n(n+1)} \tag{6}$$

To bring this into the standard form of a weighted average, let's define weights w_i. Here, i is obviously a factor of w_i, but it would be nice to normalize $w_n = 1$ so that the weights have a frequency interpretation. Therefore, we let:

$$\frac{i}{n} = w_i \tag{7}$$

Substitution yields:

$$y_n = \frac{2\sum_{i=1}^{n} w_i x_i}{n+1} \tag{8}$$

Summing the weights gives:

$$\frac{n+1}{2} = \sum_{i=1}^{n} w_i \tag{9}$$

Finally, substituting that into (8) demonstrates that the definition for w_i results in a properly-scaled weighted average:

$$y_n = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \tag{10}$$

5 Convergence and comparison to other averages

Without loss of generality, we'll assume that random variables are drawn from a distribution having an expectation of 0 and variance of 1.

The weighted sample mean converges to the true mean when the expected variance of the sample mean tends toward 0.

$$\mathrm{Var}(y_n) = \mathrm{Var}\left(\frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}\right) \tag{11}$$

For constant variance $\mathrm{Var}(x_i) = 1$, this simplifies to:

$$\mathrm{Var}(y_n) = \frac{\sum_{i=1}^{n} w_i^2}{\left(\sum_{i=1}^{n} w_i\right)^2} \tag{12}$$

The simple average uses the weight scheme:

$$w_i = 1 \tag{13}$$

Giving variance:

$$\mathrm{Var}(y_n) = \frac{1}{n} \tag{14}$$

Which converges:

$$\lim_{n\to\infty} \operatorname{Var}(y_n) = 0 \tag{15}$$

The EMA, for a parameter $0 < a < 1$, uses the weight scheme:

$$w_i = a^{n-i} \tag{16}$$

This gives variance:

$$\operatorname{Var}(y_n) = \frac{(a-1)(a^n+1)}{(a+1)(a^n-1)} \tag{17}$$

Which does not converge:

$$\lim_{n\to\infty} \operatorname{Var}(y_n) = -\frac{a-1}{a+1} \tag{18}$$

The proposed average uses the weight scheme:

$$w_i = \frac{i}{n} \tag{19}$$

Giving variance:

$$\operatorname{Var}(y_n) = \frac{2(2n+1)}{3n(n+1)} \tag{20}$$

Which, like the simple average's, converges:

$$\lim_{n\to\infty} \operatorname{Var}(y_n) = 0 \tag{21}$$

Notably, in the limit, the proposed average will have only 1/3 part more variance than the simple average.

The convergence behavior of the three averages may be compared in figure 1. The EMA's parameter is set to $a = 0.8$ for illustration.

6 Variance, online form

Welford provided a stable online form for calculating variance [4]. By reweighting the running sum, the variance of the proposed average can be calculated in a straight-forward way.

Let $S_{n-1,p}$ be the sum-of-squares calculated through x_{n-1} but weighted as if there were p samples. Assume frequency weights, with the sample x_n having a weight of 1; the previous sum-of-squares will be reweighted to accomodate it.

$$S_{n-1,p} = \sum_{i=1}^{n-1} w_{i,p}(x_i - y_{n-1})^2 \tag{22}$$

The weights are normalized according to p:

$$w_{i,p} = \frac{i}{p} \tag{23}$$

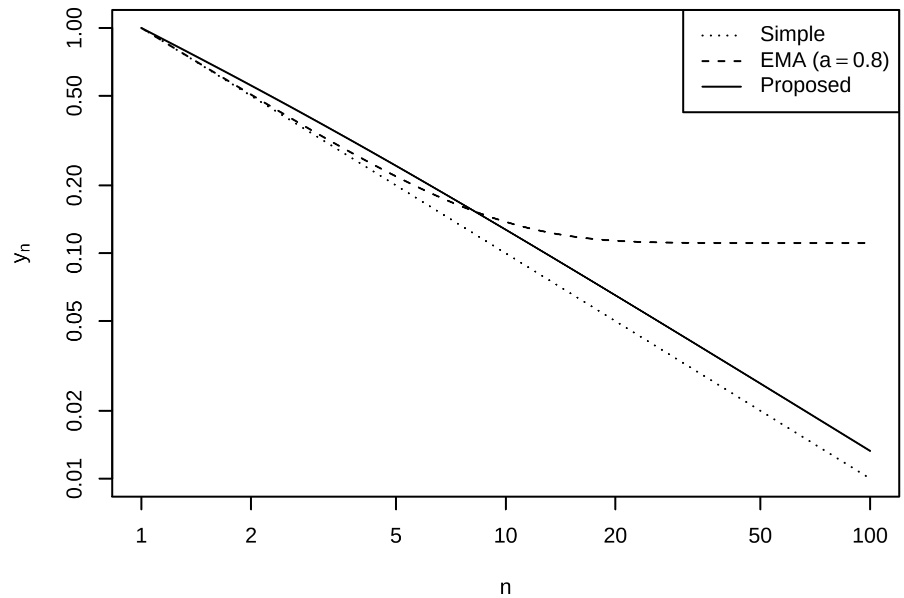

Figure 1: Variance of y_n as n increases

Giving:

$$S_{n-1,p} = \frac{\sum_{i=1}^{n-1} i(x_i - y_{n-1})^2}{p} \tag{24}$$

At $p = n - 1$, this gives the formula:

$$S_{n-1,n-1} = \frac{\sum_{i=1}^{n-1} i(x_i - y_{n-1})^2}{n-1} \tag{25}$$

And at $p = n$ it gives:

$$S_{n-1,n} = \frac{\sum_{i=1}^{n-1} i(x_i - y_{n-1})^2}{n} \tag{26}$$

The two being related by a simple re-weighting:

$$S_{n-1,n} = \frac{S_{n-1,n-1}\,(n-1)}{n} \tag{27}$$

Therefore, Welford's update can be applied to the proposed average as:

$$S_n = \frac{n-1}{n} S_{n-1} + (x_n - y_{n-1})\,(x_n - y_n) \tag{28}$$

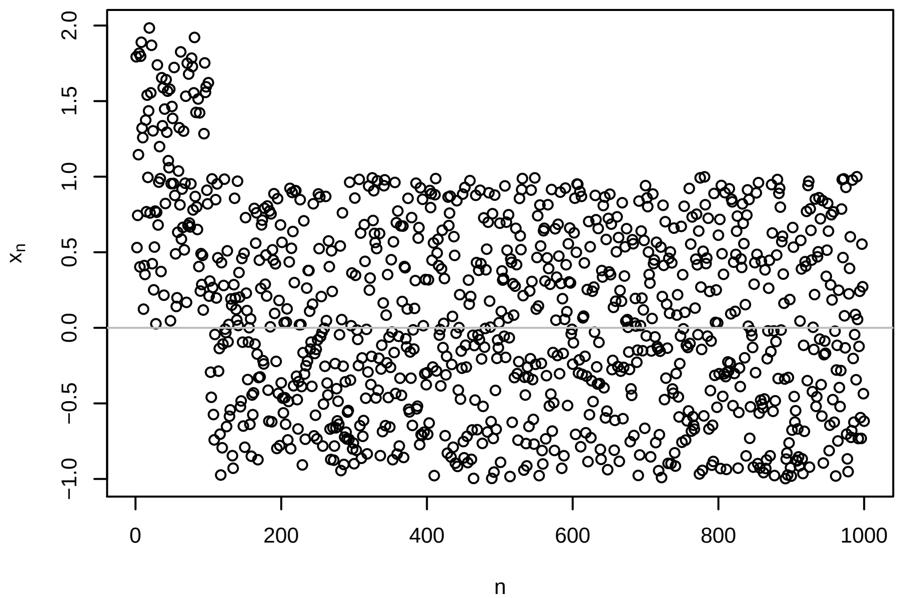

Figure 2: Random samples

7 Example behavior

Figure 2 shows 100 samples drawn from the uniform distribution from 0 to 2, followed by 900 samples from -1 to 1.

Averages are calculated from the data and displayed in figure 3. It demonstrates the adaptability of the simple average, EMA with $W = 100$, and the proposed average to the level-change in the data. The proposed average is clearly more adaptable than the simple average, and nearly as much as the EMA.

8 Conclusions and further work

The proposed average seems well-suited to the needs it was developed for, namely: an online average and variance that is adaptable to early changes but nevertheless converges.

It may be noted that (3) may be generalized to $W = \frac{n+k}{1+k}$, which leads to closed-form weights $w_i(k)$ with order $O\left(w_i(k)\right) = k$. This might form a class of useful weighted averages that could be explored in the future.

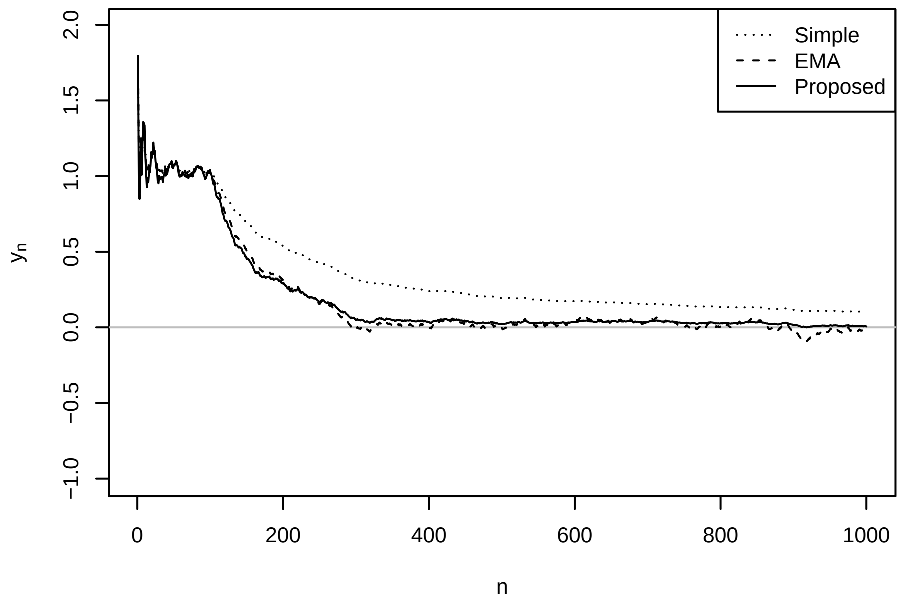

Figure 3: Progression of y_n

References

[1] Charles Holt. Forecasting seasonals and trends by exponentially weighted moving averages. O.N.R. Research Memorandum 52, Carnegie Institute of Technology, 1957.

[2] Macsyma group at Project MAC and volunteer contributors. Maxima. `http://maxima.sourceforge.net/`.

[3] Christopher John Cornish Hellaby Watkins. *Learning from Delayed Rewards*. PhD thesis, University of Cambridge, May 1989.

[4] B. P. Welford. Note on a method for calculating corrected sums of squares and products. *Technometrics*, 4(3):419–420, Aug 1962.