Technical Report

IRI-TR-11-02

Stochastic Approximations of Average Values using Proportions of Samples

IRI Technical Report

Alejandro Agostini
Enric Celaya

Institut de Robòtica i Informàtica Industrial

Abstract

In this work we explain how the stochastic approximation of the average of a random variable is carried out when the observations used in the updates consist in proportion of samples rather than complete samples.

Institut de Robòtica i Informàtica Industrial (IRI)
Consejo Superior de Investigaciones Científicas (CSIC)
Universitat Politècnica de Catalunya (UPC)
Llorens i Artigas 4-6, 08028, Barcelona, Spain

Tel (fax): +34 93 401 5750 (5751)
`http://www.iri.upc.edu`

Corresponding author:

Alejandro Agostini
tel: +34 93 401 5786
`agostini@iri.upc.edu`
`http://www.iri.upc.edu/people/agostini`

1 Basic Notation

To perform the explanation we will use the formula of the conventional average using the ratio of cumulative sums,

$$\bar{x}_N = \frac{\sum_{n=1}^{N} w_n x_n}{\sum_{n=1}^{N} w_n} \tag{1}$$

where $0 \le w_n \le 1$ indicates the proportion of the sample n.

For the cumulative sums we define the notation

$$[f] = \sum_{n=1}^{N} w_n f_n \tag{2}$$

which leads to a formulation of equation (1) as,

$$\bar{x}_N = \frac{[x]_N}{[1]_N}. \tag{3}$$

2 Stochastic Approximation of Average Values

The stochastic approximation formula for the average of a random variable using proportion of samples is derived as follows. First, using equations (1) and (3), we write the recursive formula to calculate the average $\hat{x}_N$,

$$\bar{x}_n = \bar{x}_{n-1}\left(1 - w_n \frac{1}{[1]_n}\right) + w_n \frac{1}{[1]_n} x_n \tag{4}$$

where $[1]_n$ is the accumulated proportions of samples at iteration n, $n = 1, \ldots, N$. For convenience, we use the notation $\upsilon_n = [1]_n$ to reference the total amount of proportions accumulated at iteration n.

Note that, υ plays the same role as the number of time steps t (or iterations n) normally used to indicate the amount of experiences collected in cases where the samples are provided in their full proportion $w = 1$, in which case υ and t (or n) would be equal.

Equation (4) has the form of the formula for stochastic approximations [2] with a learning coefficient given by $\eta = 1/\upsilon$,

$$\bar{x}_n = \bar{x}_{n-1}\left(1 - w_n \eta(\upsilon_n)\right) + w_n \eta(\upsilon_n) x_n \tag{5}$$

As usually occurs in applications with stochastic approximations, this learning coefficient η can be regulated to permit to reach the correct solution for the average. For instance, we have already seen the direct definition of the learning coefficient $\eta(\upsilon) = 1/\upsilon$, which is appropriate in cases where all the experiences are equally important for the average. However, there are other cases where new incomes should gain more importance over the old ones, like, for instance, in approximations of non-stationary averages. In those cases the function $\eta(\upsilon)$ should be defined so as to track the changes in the statistics in order to avoid a fast dropping to zero of the learning coefficient which may produce the average get stuck before reaching the correct solution. This may be accomplished by establishing the condition $\eta(\upsilon) > 1/\upsilon$ which makes new incomes more relevant than old ones, with a forgetting effect determined by the setting of η.

The definition of the learning coefficient $\eta(\upsilon)$ is not unrestricted if we want to guarantee convergence to a solution. To this, and as stated by the stochastic approximation theory [2], the function $\eta(\upsilon)$ should be defined in a way that it fulfils with the stochastic approximation requirements.

3 Stochastic Approximations using Cumulative Sums

To illustrate the forgetting effect produced with the regulation of the learning coefficient $\eta(\upsilon_n)$ we derive from equation (5) the following,

$$\frac{\bar{x}_n}{\eta(\upsilon_n)} = \bar{x}_{n-1}\left(\frac{1}{\eta(\upsilon_n)} - w_n\right) + w_n x_n. \tag{6}$$

Now, using the notation

$$\{x\}_n = \frac{\bar{x}_n}{\eta(\upsilon_n)}, \tag{7}$$

we get the following expression for equation (6),

$$\{x\}_n = \{x\}_{n-1}\left(\eta(\upsilon_{n-1})\left(1/\eta(\upsilon_n) - w_n\right)\right) + w_n x_n. \tag{8}$$

Equation (8) is the recursive form of a discounted sum of the form,

$$\{f\}_n = \sum_{i=1}^{n}\left(\prod_{j=i+1}^{n}\lambda(\upsilon_j)\right) w_i f_i, \tag{9}$$

with discount factor

$$\lambda(\upsilon_j) = \eta(\upsilon_{j-1})\left(1/\eta(\upsilon_j) - w_j\right), \tag{10}$$

and with generic expression

$$\{f\}_n = \{f\}_{n-1}\lambda(\upsilon_n) + w_n f_n. \tag{11}$$

From (11) we obtain, for the case of $f = 1$, the sequent formula

$$\{1\}_n = \{1\}_{n-1}\left(\eta(\upsilon_{n-1})\left(1/\eta(\upsilon_n) - w_n\right)\right) + w_n, \tag{12}$$

from which we derive the following expression:

$$\frac{1}{\{1\}_{n-1}}\left(\{1\}_n - w_n\right) = \eta(\upsilon_{n-1})\left(1/\eta(\upsilon_n) - w_n\right). \tag{13}$$

Using (13) we can obtain a relation between the discounted sum $\{1\}_i$ and the learning coefficient $\eta(\upsilon_i)$ from the solution to (13) given by

$$\{1\}_i = \frac{1}{\eta(\upsilon_i)}. \tag{14}$$

Finally, replacing (14) in (7) and rearranging we obtain the general expression for the mean calculated using the discounted sums (9),

$$\bar{x}_n = \frac{\{x\}_n}{\{1\}_n}. \tag{15}$$

3.1 Some Special Cases

Case $\eta(\upsilon_n) = 1/\upsilon$

For the case of a $\eta(\upsilon_n) = 1/\upsilon_n$, the discount factor (10) results to be,

$$\begin{aligned}\lambda(\upsilon_n) &= \frac{1}{\upsilon_{n-1}}(\upsilon_n - w_n),\\ &= \frac{1}{(\upsilon_n - w_n)}(\upsilon_n - w_n),\\ &= 1,\end{aligned}$$

in which case the updating is carried out without forgetting past values, and equation (15) has the same result as equation (3).

Case $w_n = 0$

When the proportion of the incoming sample is zero, we would expect that no forgetting occurs, and the cumulative sums used to calculate the average in (15) remain unchanged. This is what actually occurs with our definition of the discounted sum, which can be seen by instantiating the discount factor for the case of $w_n = 0$,

$$\begin{aligned}\lambda(\upsilon_n) &= \eta(\upsilon_{n-1})(1/\eta(\upsilon_n) - w_n),\\ &= \eta(\upsilon_n - w_n)(1/\eta(\upsilon_n) - w_n),\\ &= \eta(\upsilon_n)(1/\eta(\upsilon_n)),\\ &= 1,\end{aligned}$$

which produces equation (11) to be,

$$\{f\}_n = \{f\}_{n-1}. \tag{16}$$

Case of a constant η

In the case of a constant learning coefficient η, the discount factor (10) takes the form,

$$\begin{aligned}\lambda(\upsilon_n) &= \eta(1/\eta - w_n),\\ &= 1 - w_n\eta, \end{aligned} \tag{17}$$

and the discounted sum (9) turns to be,

$$\{f\}_n = \sum_{i=1}^{n}\left(\prod_{j=i+1}^{n}(1 - w_j\eta)\right)w_i f_i, \tag{18}$$

Comparison with the approach with forgetting factor $\lambda(n)^w$

In our former proposal for stochastic approximation using cumulative sums [1] we used a forgetting factor $\lambda(n)^w$ which depends on the number of iterations n (or time steps), rather than the accumulated proportions υ, and which component $\lambda(n)$ is related with an iteration dependent learning coefficient $\eta(n)$ in the following way,

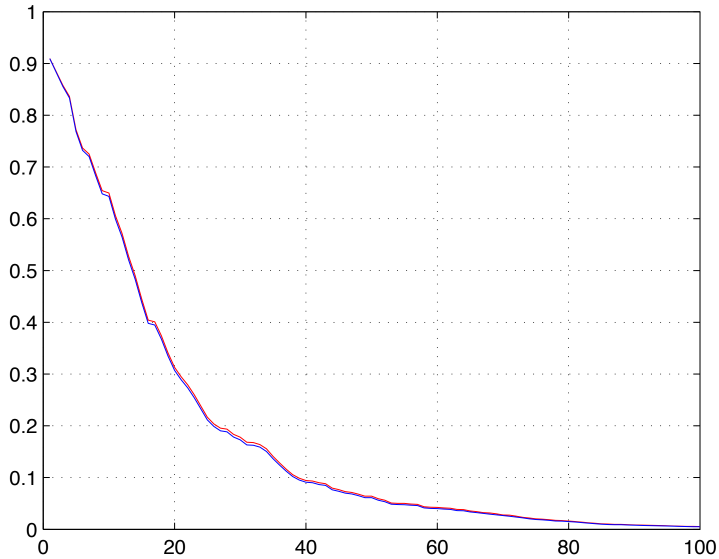

Figure 1: Comparison of the forgetting effects with a constant learning rate η produced using the old formula $(1-\eta)^w$ (blue curve) versus the new formula $(1-w\,\eta)$ (red curve).

$$\lambda(n) = \eta(n-1)\,(1/\eta(n)-1)\,. \tag{19}$$

In order to compare the new approach (10) with the one in [1] we set the learning coefficient η to be constant and equal in both cases, so as to eliminate the dependence on time. In this case, the discount factor $\lambda(n)^w$ results to be $(1-\eta)^w$, defining the following discounted sum,

$$\{f\}_n = \sum_{i=1}^{n} \left(\prod_{j=i+1}^{n} (1-\eta)^{w_j} \right) w_i f_i, \tag{20}$$

Then, we generate randomly samples f_i, $i = 1,\ldots,n = 100$, in the range $[0,10]$, and assign to each of them a proportion w_i also generated randomly, but in the range $[0,1]$. Comparing the forgetting effect produced with these samples in the discounted sum (18) with respect to the forgetting effect in the discounted sum (20) we get the behaviours plotted in figure 1.

As we can observe from the figure both approaches have nearly equal forgetting effects.

References

[1] A. Agostini and E. Celaya. Reinforcement learning with a Gaussian mixture model. In *Proc. International Joint Conference on Neural Networks (IJCNN'10). (Barcelona, Spain)*, pages 3485–3492, 2010.

[2] HJ Kushner and GG Yin. *Stochastic approximation algorithms and applications*. Springer-Verlag, 1997.

IRI reports

This report is in the series of IRI technical reports.
All IRI technical reports are available for download at the IRI website
`http://www.iri.upc.edu`.