

Applying Stochastic Approximation Method with Delayed Observations in Exponential Distribution Case

R. A. Atwa
Department of Mathematics
Faculty of science
Zagazig University
Zagazig, Egypt

ABSTRACT

The main purpose of this work is investigated a loss system, which can serve as a model of modified Robbins-Monro stochastic approximation in the presence of delayed observations. Here we confine ourselves to the case of exponential distribution. The results achieved for the loss system enable to conclude about the efficiency of the procedure and to give a hint for the choice of the number of servers in the modified loss system.

Keywords

Stochastic approximation, efficiency of the procedure

1. INTRODUCTION

Stochastic approximation is a sequential method, one will be interested to know an optimum stopping rule for a given situation. A modeling and analysis of stochastic approximation procedure has become an important field of performance analysis –(cf. [1], [3], [5], [7], [8], [10], [11], [12], [14], [18], [21], [22]). Typically, in stochastic procedures, the observations of an experiment follow each other after fixed time-intervals; the point of the next observation is corrected according to the result of the preceding one. However, in some situations, as in biological or lifetime experiments, the result of an observation becomes known only after a random time delay.

The Robbins-Monro stochastic approximation procedure is an iterative algorithm to find zeros of functions, which cannot be computed directly, but only estimated with estimation subject to random error. This method is significant in many fields such as biological, medical and lifetime experiments.

Kiefer and Wolfowitz [12], considered the Robbins-Monro method as their point of departure, they gave a method of determining the location of maximum of regression function (if it exist) and established asymptotic properties of the procedure.

There is another sequential method the so called up-and-down method. This method is used in many fields where it is desirable to estimate a critical measurements for some response. It has merits and demerits which indicate situations where stochastic approximation can be exploited with advantage are appended [20].

Newton-Raphson method is an iterative technique of numerical analysis, where this technique is frequently employed in problem such that locating a point, where a given function behaves in finding the roots of equation, or locating a point of maximum of a function [20].

Loss system is a service system which no queuing is allowed and so arriving customers will be lost when all servers are busy.

Erlang loss system is queuing model which customers arrive according to a poisson process and are served individually by one of n servers, each of which has an identical general distribution of service time.

Many papers dealt with loss system to investigate its properties such as [17], which deals with discrete-time analysis of the general three-server loss system, and it focuses on determining the proportion of the customers that will be lost in the long run.

In addition the Equilibrium results for the M/G/K group-arrival loss system are studied by [4], they considered this loss system, under statistical equilibrium and two cases of acceptance policy. In the first case, the system works under the partial acceptance policy. In the second case, the system works under the all-or-nothing acceptance policy. In both cases customer depart individually, while the joint service time distribution of the accepted members of a group may depend both on its initial and accepted size plus an additional condition.

Jonckheere and Leskelä [9], developed stochastic comparison and coupling techniques to study how multiclass loss system with two layers of servers, where each server at the first layer dedicate to a certain customer class, and the servers at the second layer can handle all customer classes, is affected by packing of customers. Arriving customers are preferentially directed to the first layer. Altered service rate, and altered server configurations. This analysis leads to easily computable upper and lower bounds for the performance of the system.

Recently, the Robbins-Monro stochastic approximation was employed in clinical applications to find the optimal dose [2]. It is meaningful to ask, whether and how stochastic approximation can be applied. Mahmoud and Rasha [5], applied Robbins and Monoro stochastic approximation procedure [19] in the presence of compound delayed observations. They partitioned experiments (or observations) into parts, where each part is treated as a sub-experiment. Some of those parts or sub-experiments are lost due to the delay of the preceding part during the time interval between any two consecutive parts or compound observations. Thus, Bernolli and Binomial delay distributions of parts and compound observations were assumed.

At the beginning, all series are open, all k 's equal to 1, all $x_1^{(k)}$'s equal to the same constant. At time t , an experiment is

made at point $x_t^{(i)}$. The i th series is then closed at the same point $x_t^{(i)}$, till time $(t + d(t) + 1) - 0$, when it opens at point

$$x_{t+1}^{(i)} = x_t^{(i)} - a_t (m(x_t^{(i)}) + e((t + d(t) + 1), x_t^{(i)})).$$

Here $m(x)$ is a function whose zero point θ is to be found, $e(v + 1, x)$ is the observational error corresponding to an observation of m made at point x and becoming known during the interval $[v, v + 1)$; $x_t^{(i)}$ is the current approximation to θ at time $t - 0$; $a_t, t \in \mathbb{N}$, is a zero sequence of positive constants, typically $a_t = \frac{a}{t+t_0}$, t_0 is non-negative; $d(t)$ is [the integer part of] the delay of the result of an experiment made at time t .

If there is no series open at time $t - 0$, no experiment is made at time t and a time-loss is thus incurred. If l is the steady state probability of such time-loss, then its complement $e = 1 - l$ is called the efficiency of the procedure.

For finding θ , the average of current approximations over all series, $\theta_t = \frac{1}{K} \sum_{k=1}^K x_t^{(k)}$, has been chosen as a global approximation to θ at time $t - 0$. Under usual assumptions on function m and errors $e(t, x)$, not repeated here, and under independence of delays $d(t)$, the normed

$$p_{ij} = \begin{cases} \binom{K}{j} e^{-\theta a j} (1 - e^{-\theta a})^{K-j} & i = 0, 1; j = 0, \dots, K, \\ 0 & i = 2, 3, \dots, K; j = 0, \dots, j - 2, \\ \binom{K-i+1}{j-i+1} e^{-\theta a(j-i+1)} (1 - e^{-\theta a})^{K-j} & i = 2, 3, \dots, K; j = i - 1, \dots, K. \end{cases} \quad (2)$$

From the formed system, we can eliminate from the system as successive transitions from state to state and occur deterministically, with probability 1. The resulting matrix of transition probabilities is called the reduced transition matrix.

From the reduced transition matrix we find the following:

- i All states are irreducible closed sets; therefore they contain persistent non-null states.
- ii All states have period 1 because

$$p_{jj}(1) > 0 \quad \forall j. \quad (3)$$

From the previous conditions (see [8]) all states are ergodic and there is a unique stationary distribution π that can be calculated by solving the system of equations:

$$P^T \pi = \pi, \quad (4)$$

Together with the added requirement:

$$1^T \pi = 1,$$

where T denotes the transpose of the matrix, π is the stationary distribution matrix of the compound Markov chain; and P is the matrix of transition probabilities.

2.1 Methodology

To solve the system (4), we apply the following steps:

- i Assume the value K to form the transition matrix P and substitute in (4).
- ii One of the remaining equations in (4) can always be deleted, another one is to be added, namely the requirement

approximation $t^{\frac{1}{2}}(\theta_t - \theta)$ are asymptotically normally distributed, with parameters 0 and σ^2/e , where σ^2 is the asymptotic variance of the same normed approximation in a procedure with no delays. Hence e is also the relative asymptotic efficiency of θ_t as a statistical estimator.

2. DESCRIPTIVE OF THE SYSTEM

At the beginning, assume that an experiment starts its service at time 0. If a customer is served without delay, then it will leave before time a , the unit time, where the next one arrives at time a and starts its service without delay. If a customer is served with delay, then it will leave after time a , where the service is delayed such that the time delay starts from the arrival of the next part.

Here we applied stochastic approximation with delayed observations in exponential distribution case, i.e.

$$f(t) = \theta e^{-\theta t}, \quad t > 0, \theta > 0 \quad (1)$$

It is shown that the stationary transition probabilities p_{ij} of the compound Markov chain with states $\{0, 1, 2, \dots, K\}$ are given by:

$$\sum_{\alpha} \pi_{\alpha} = 1$$

- iii The reduced system is solved for π_{α} , using the Matlab program.
- iv The solution π_{α} is used to compute the efficiency e of the procedure where $e = \pi_0$.

2.2 Results and discussion

Tables from 1 to 4 are given four different cases for (a), to calculate the asymptotic efficiencies for parameter $\theta = 0.1, 0.2, 0.4, 1, 2, 3, 4, 5$, that is for $K = 2(1)8$. The consequent recommendation for the choice of K, taking the value of a , in consideration seems to be reasonable even for more general delay distribution not drastically different from the Exponential one. The headings $E(d)$, in the tables are then made instead of θ .

3. REFERENCES

- [1] Albert, A.E. and Gamder, L.A. 1967 Stochastic Approximation and Nonlinear Regression. MIT press, Cambridge.
- [2] Cheung, Y. and Elkind, M.S. V. 2010. Stochastic approximation with virtual observations for dose finding on discrete levels. Biometrika, 97, 109-121.
- [3] Chung, K. L. 1954. On A Stochastic Approximation Method. Ann. Math. Statist. 25, 463-483.
- [4] Dimous, S. and Fakinos, D. 2010. Equilibrium results for the M/G/K group-arrival loss system. Societed de Estadistica e Investigacion Operativa.
- [5] Doob, J. L. 1953. Stochastic Processes. Wiley New York; Chapman and Hall, London.

- [6] Dupač, V. and Herkenrath, U. 1985. Stochastic Approximation with delayed observations. *Biometrika*, 72, 683-685.
- [7] Fabian, V. 1971. Stochastic Approximation. In *Optimizing Methods in Statistics*. (J. S. Rustagi, ed.), Academic press, New York. 439-470.
- [8] Harold, J. Kushner, G. and George Yin. 1997. *Stochastic Approximation and Recursive Algorithm and applications*. Applications of Mathematics, 35.
- [9] Joseph, VR. Tian, Y. and Wu, CFJ. 2007. Adaptive Designs for Stochastic Root-Finding. *Statistica Sinica*, 17, 1549-1565.
- [10] Kesten, H. 1958. Accelerated Stochastic Approximation. *Ann. Math. Statist.*, 29, 41-59.
- [11] Kiefer, J. and Wolfowitz, J. 1952. Stochastic Approximation of the Maximum of a regression function, *Ann. Math. Statist.*
- [12] Konev, V. and Pergamenshchikov, S. 2003. Sequential estimation in Stochastic Approximation with Autoregressive errors in observations. *Sequential Analysis*, 22, 1-29.
- [13] Lai, T. 2003. Stochastic approximation. *The Annals of Statistics*, 31, 391-406.
- [14] M. A. Mahmoud and A.A. Rasha, Stochastic Approximation with Compound Delayed Observations. *Mathematical & Computational Applications*, 10, 283-289.
- [15] Mahmoud, M. A. and Rasha A. A. Stochastic Approximation and Compound Delayed Observations with Independent Random Time Delay Distribution. *Arabian Journal of Science and Engineering*, 36, 1549-1558, 2011.
- [16] Mokaddis, G. S. Metwally S. A. and Ahmed, A. M. 2000. The General Three-Server Queuing Loss System: Discrete-Time Analysis. *Information and Management Sciences*, 11, 81-95.
- [17] Mukerjee, H.G. 1981. A Stochastic Approximation by Observation on Discrete Lattice Using Isotonic Regression. *Ann. Math. Statist.* 9, 1020- 1025.
- [18] Robbins, H. and Monro, S. 1951. A stochastic approximation method. *Ann. Math. Statist.*, 22, 400-407.
- [19] Wasan, M.T. 1969. *Stochastic Approximation*. Cambridge Univ. Press, New York.
- [20] Xu, Z. and Dai, Y. 2008. A Stochastic Approximation Frame Algorithm with Adaptive Directions. *Numerical Mathematics: Theory, Methods and Applications*, 1, 460-474.
- [21] Zhu, Y. and Yin, G. 1994. Stochastic approximation in real time: a pipe line approach. *J. Comput. Math.*, 12, 21-30.

Table 1. Percentage asymptotic efficiencies of the proposed procedure with K series and Exponentially-distributed delay d , with parameter θ , at $a = 1$ case

θ	$E(d)$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
0.1	10	0.48	0	0	0	0	0	0
0.2	5	1.81	0	0	0	0	0	0
0.4	5/2	6.51	0.2	0	0	0	0	0
1	1	29.24	12.59	3.08	0.44	0.03	0	0
2	1/2	65.89	60.04	48.1	36.59	26.1	10.16	17.21
3	1/3	86.03	85.13	80.21	75.3	70.42	60.79	65.57
4	1/4	94.65	94.52	92.7	90.87	89.05	85.4	87.23
5	1/5	98	97.99	97.31	96.64	95.97	94.63	95.3

Table 2. Percentage asymptotic efficiencies of the proposed procedure with K series and Exponentially-distributed delay d , with parameter θ , at $a = 2$ case

θ	$E(d)$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
0.1	10	1.81	0	0	0	0	0	0
0.2	5	6.51	0.2	0	0	0	0	0
0.4	5/2	20.94	5.44	0.69	0.04	0	0	0
1	1	65.89	60.4	48.1	36.59	26.1	17.21	10.16
2	1/2	94.65	94.52	92.7	90.87	89.05	87.23	85.4
3	1/3	99.26	99.26	99.01	98.77	98.52	98.27	98.02
4	1/4	99.9	99.9	99.87	99.83	99.8	99.77	99.73
5	1/5	99.99	99.99	99.98	99.98	99.97	99.97	99.96

Table 3. Percentage asymptotic efficiencies of the proposed procedure with K series and Exponentially-distributed delay d , with parameter θ , at $a = 3$ case

θ	$E(d)$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
0.1	10	3.86	0.04	0	0	0	0	0
0.2	5	13.16	1.52	0.08	0	0	0	0
0.4	5/2	37.56	22.16	8.57	2.27	0.37	0.03	0
1	1	86.03	85.13	80.21	75.3	70.42	65.57	60.75
2	1/2	99.26	99.26	99.01	98.77	98.52	98.27	98.02
3	1/3	99.96	99.96	99.95	99.94	99.93	99.91	99.9
4	1/4							
5	1/5							

Table 4. Percentage asymptotic efficiencies of the proposed procedure with K series and Exponentially-distributed delay d , with parameter θ , at $a = 4$ case

θ	$E(d)$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
0.1	10	6.51	0.2	0	0	0	0	0
0.2	5	20.94	5.44	0.69	0.04	0	0	0
0.4	5/2	53.03	42.95	27.54	15.32	6.92	2.46	0.64
1	1	94.65	94.52	92.7	90.87	89.05	87.23	85.4
2	1/2	99.9	99.99	99.87	99.83	99.8	99.77	99.73
3	1/3							
4	1/4							
5	1/5							

Table 5. Percentage asymptotic efficiencies of the proposed procedure with K series and Exponentially-distributed delay d , with parameter θ , at $a = 5$ case

θ	$E(d)$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$
0.1	10	9.65	0.62	0.02	0	0	0	0
0.2	5	29.24	12.59	3.08	0.44	0.03	0	0
0.4	5/2	65.89	60.4	48.1	36.59	26.1	17.21	10.16
1	1	98	97.99	97.31	96.64	95.97	95.3	94.63
2	1/2	99.99	99.99	99.98	99.98	99.97	99.97	99.96
3	1/3							
4	1/4							
5	1/5							

100.0 in all empty cells.