On a modification of Lehmer algorithm

L.V. Nedorezov

Research Center for Interdisciplinary Environmental Cooperation RAS, Nab. Kutuzova 14, Saint-Petersburg 191187, Russia
e-mail: l.v.nedorezov@gmail.com

Keywords: Lehmer algorithm, pseudo-stochastic numbers

Abstract

In paper one of possible modifications of Lehmer algorithm of pseudo-stochastic number generation is under consideration. Modified map of interval [0,1] on [0,1] does not have stationary states. Some of properties of map were analyzed.

1. Introduction

D.H. Lehmer (1951) analyzed the following algorithm for obtaining of sequences of pseudo-stochastic numbers with uniform distribution on interval [0,1]:

$$x_{k+1} = \{Ax_k\}. \quad (1)$$

In formula (1) A is a positive (rather big)) number; $\{\cdot\}$ is a fractional part of number. x_0 is initial value for map (1), $x_0 \in [0,1]$. Formula (1) determines single-valued transformation of close interval [0,1] onto itself. If A is integer [0,1] is divided onto A sub-intervals of the length A^{-1}. Every sub-interval is «stretched» (transforms) on $[0,1]$. On every sub-interval transformation (1) has unstable stationary state.

It is obvious if initial value x_0 is equal to one of stationary states generated by (1) trajectory cannot be considered as a sequence of stochastic numbers (or pseudo- stochastic numbers). Moreover, it is easy to point out a set of initial values for trajectories which come into stationary state after 1, 2, 3… steps. Thus, it is not surprised that it is not easy to find initial value x_0 of variable which allows obtaining required pseudo-stochastic trajectory (Sobol, 1873; Ermakov, 1975; Tyurin, Figurnov, 1990; Mikhailov, Voitishek, 2006). When we are talking about “required pseudo-stochastic trajectory” it means that existing statistical tests don’t allow concluding that obtained values are not values of independent stochastic variables with uniform distribution on $[0,1]$.

On the other hand, map (1) can be easily modified in the following manner: every sub-interval of the length A^{-1} can be linear transformed on $[0,1]$, graphics of map (1) “covers” (when amount of A is rather big) square $[0,1]\times[0,1]$, and at the same time map hasn’t stationary states (Nedorezov, 1986):

$$x_{k+1} = \left\{ \frac{x_k}{h} + \varepsilon - (1-h)\left[\frac{x_k}{h}\right] \right\}. \tag{2}$$

Nonnegative parameters of map (2) are satisfied to inequalities $0<\varepsilon<h<1$. In (2) $[\cdot]$ is integer part of number, and $\{\cdot\}$ is fractional part of number. Parameters $h = A^{-1}$, $x_0 \in [0,1]$.

2. Some properties of map (2)

1. In figure 1 there is graphics of map (2) at $h=0.2$, $\varepsilon=0.1$. As we can see, function (2) is broken at every point mh, $m=0,1,2\ldots$ (when $x_k = 0.2$, 0.4, 0.6 and 0.8; graphics of map (2) has no intersections with bisectrix $x_{k+1}=x_k$; fig. 1). Really, if $x_k \to mh-0$ then

$$\left[\frac{x_k}{h}\right] = m-1,$$

$$\frac{x_k}{h} + \varepsilon - (1-h)\left[\frac{x_k}{h}\right] \to m + \varepsilon - (1-h)(m-1) = mh + 1 + (\varepsilon - h).$$

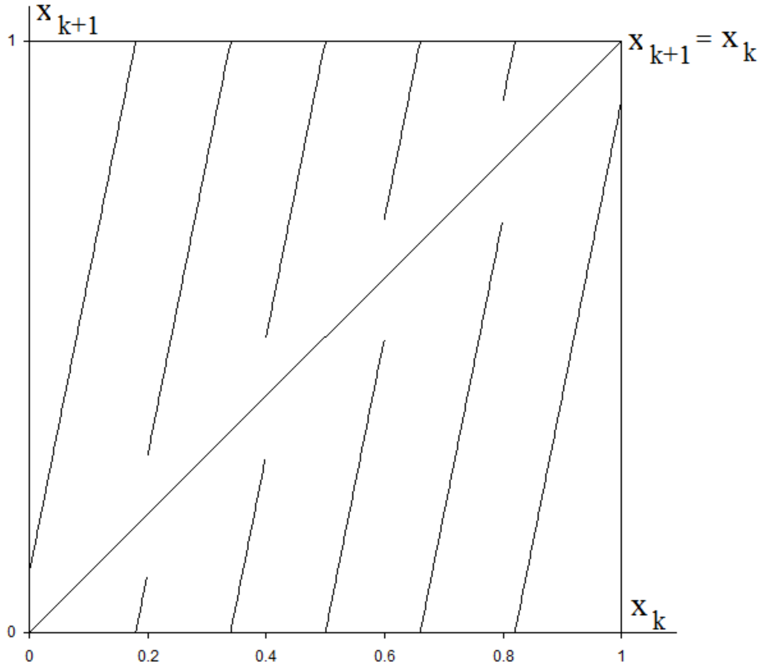

Fig. 1. Graphics of map (2) at $h=0.2$, $\varepsilon=0.1$.

Respectively, taking into account limits for map parameters we have

$$\left\{\frac{x_k}{h}+\varepsilon-(1-h)\left[\frac{x_k}{h}\right]\right\}\to mh+(\varepsilon-h)<mh\,.$$

If $x_k = mh$

$$\left\{\frac{x_k}{h}+\varepsilon-(1-h)\left[\frac{x_k}{h}\right]\right\}=m+\varepsilon-(1-h)m=mh+\varepsilon>mh\,.$$

If $x_k \to (m+1)h-0$ then

$$\left\{\frac{x_k}{h}+\varepsilon-(1-h)\left[\frac{x_k}{h}\right]\right\}\to mh+\varepsilon\,.$$

Thus, every interval $[mh,(m+1)h]$ transforms on $[0,1]$. Note that map (2) doesn't contain stationary states but it doesn't mean that it hasn't cycles of various lengths. It is easy to show that it has cycles of length 2.

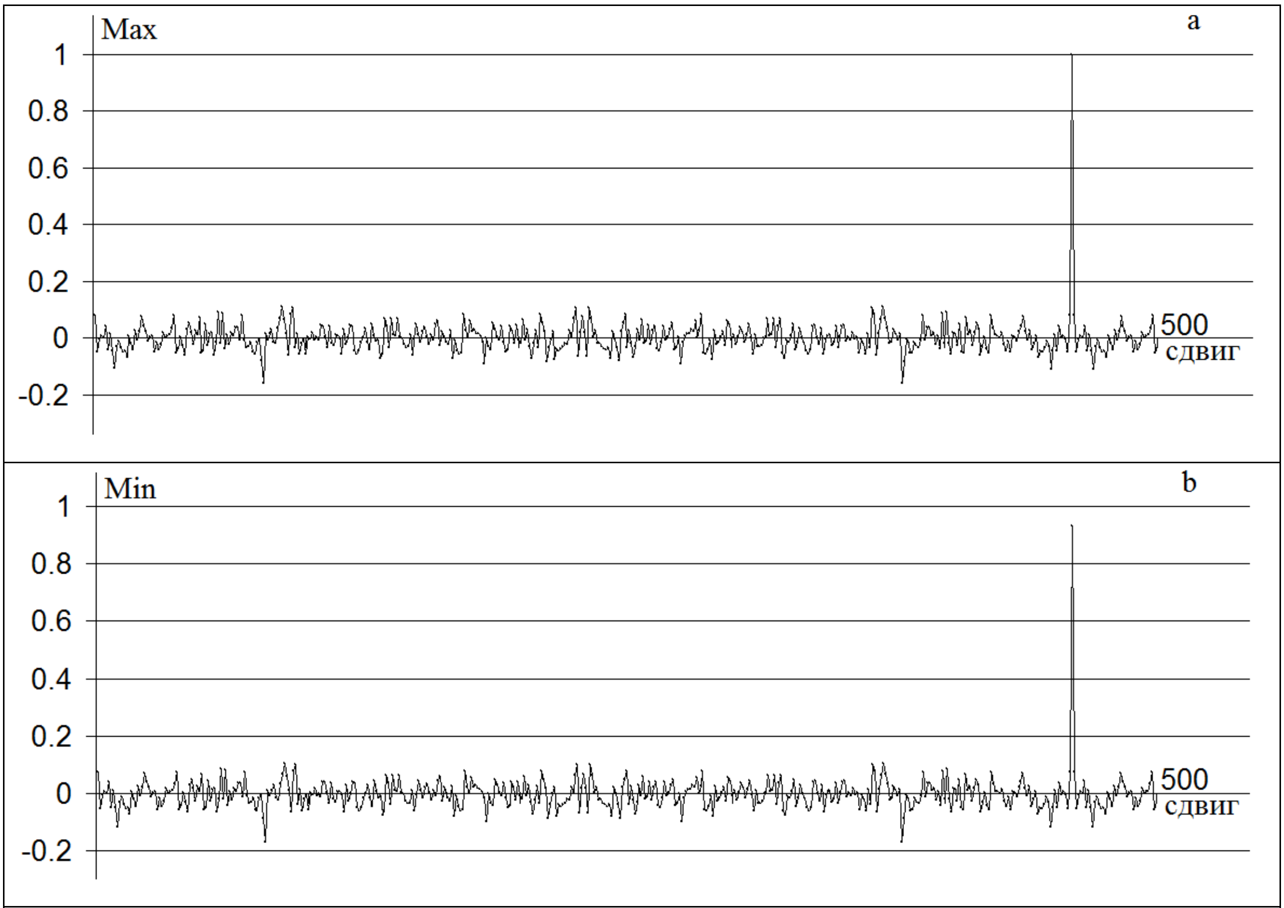

Fig. 2. Changing of maximum and minimum values of autocorrelation function at $h=2\cdot10^{-10}$, $\varepsilon=10^{-10}$ and various initial values of trajectories of the map (2). When lag is equal to 460 highest values 0.999983 and 0.93457 are observed for maximum and minimum respectively.

2. Analysis of behavior of autocorrelation function can be used as one of elements of testing of trajectories of map (2). In figure 2a and 2b 500 first values of maximums and minimums of autocorrelation functions are presented. Values of autocorrelation functions were calculated for various initial values of x_0 for 20 various samples (with sample sizes 60000; $h = 2 \cdot 10^{-10}, \varepsilon = 10^{-10}$). As we can see in this figure, absence of stationary states for map (2) doesn't lead to automatic appearance to cycles of rather big length only for every initial values (it is obvious that cyclic regimes will be observed for all values of map parameters because in computer we have a map of finite set onto itself). In considering case autocorrelation function has a jump when lag is equal to 460 (fig. 2 a, b). Such behavior of autocorrelation function is typical for a situation when asymptotically stable limit cycle of the respective length exists in phase space.

Numerical analysis of map (2) shows that for one and the same values of map parameters we may have cycles of various lengths for various initial values x_0. For example, for $h = 1.1973 \cdot 10^{-10}, \varepsilon = 1.0222 \cdot 10^{-10}$ we can observe cycle of the length 721 for $x_0 = 0.653954$ and cycle of the length 421 for $x_0 = 0.250949$; for $h = 1.0973 \cdot 10^{-10}, \varepsilon = 1.0222 \cdot 10^{-10}$ we can observe cycle of the length 572 for $x_0 = 0.7995417$ and cycle of the length 160 for $x_0 = 0.68037$. Provided numerical analysis of behavior of trajectories of the map (2) allows concluding that absence of stationary states doesn't lead to situation when search of suitable initial value isn't a problem. Below we consider modification of map (2) with periodic disturbance of trajectories: for fixed integer values m_0, n_0 and $k = 1,2,3\ldots$ initial values of trajectories were changed:

$$(x_{m_0 k})_0 = \{10^{n_0} \cdot x_{m_0 k}\}.$$

In this case final trajectory of modified map looks like a sequence of finite parts of trajectories of map (2) of the length m_0.

3. For $\varepsilon = 1.0822 \cdot 10^{-10}$, $h = 1.0979 \cdot 10^{-10}$, $x_0 = 0.238475$, $m_0 = 3700$, $n_0 = 4$ it was obtained (see fig. 3) that maximum value of autocorrelation function is equal to 0.068369 (sample size is equal to 60020, number of calculated values of autocorrelation function is equal to 15000). For these parameters length of stable limit cycle is equal to 3989. Comparing with standard algorithm Rnd of obtaining of pseudo-stochastic numbers we get (fig. 4) that maximum value is about 0.016 (and it is much less than 0.068369; sample size and calculated values of autocorrelation function were the same).

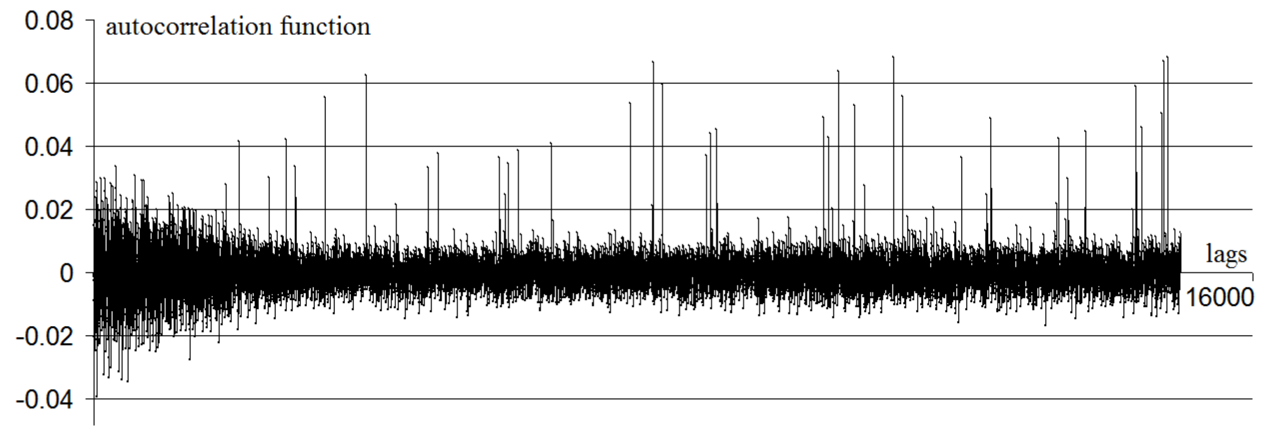

Fig. 3. Behavior of autocorrelation function for $\varepsilon = 1.0822 \cdot 10^{-10}$, $h = 1.0979 \cdot 10^{-10}$, $x_0 = 0.238475$, $m_0 = 3700$, and $n_0 = 4$.

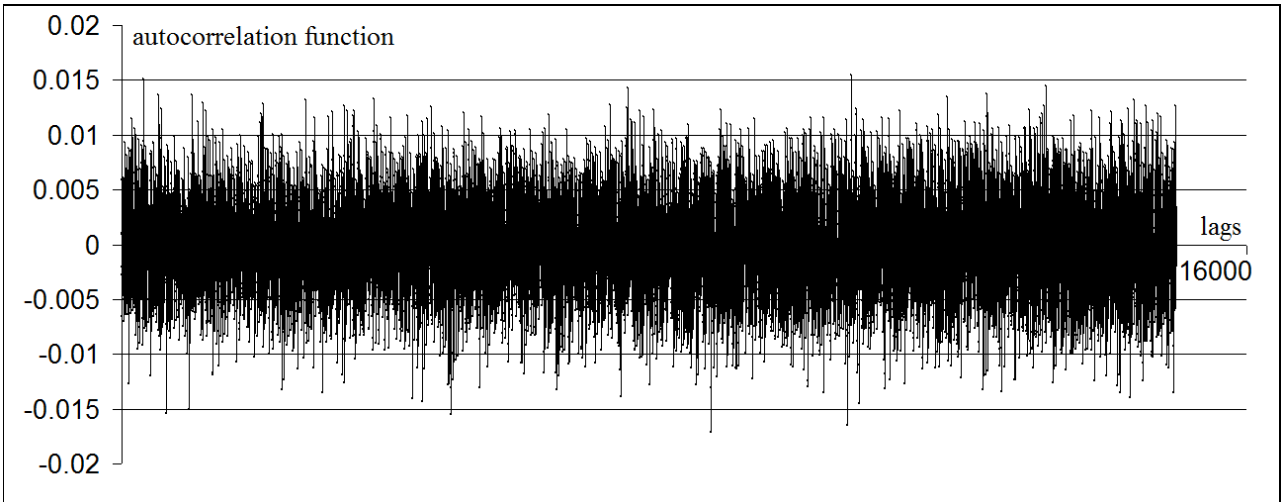

Fig. 4. Behavior of autocorrelation function for sequence of numbers obtained with standard function Rnd.

4. Analyses of distributions of pairs, triples etc. (which are constructed from elements of map (2) trajectories) are next important elements of testing of map properties. In this case we can choose one of two possible ways. First, we can divide N-dimensional cube onto subsets (for every subset we can easily calculate probability of event that point will appear in this subset; Sobol, 1873; Ermakov, 1975; Tyurin, Figurnov, 1990; Mikhailov, Voitishek, 2006). After a certain trials and calculation of numbers of points in all subsets, we can compare empirical and theoretical probabilities with the help of X^2-criterion. Second, we can choose one stochastic subset with easily determined theoretical probability, and after a certain number of trials we can compare empirical frequency with theoretical probability with t-test.

For parameters pointed out above we have that for big sample size (≥ 55000) we have to reject Null hypothesis (about equivalence of quota to probability) even with 0.1% significance level (Student t-criterion). Note that "theoretical probability" (subset of square $[0,1]\times[0,1]$ was found stochastically) was equal to 0.3867. It allows concluding that obtained trajectory of modified map (2) cannot be considered as a set of pseudo-stochastic numbers with uniform distribution.

For the same parameters and $x_0 = 0.367$ it was obtained (see fig. 3) that maximum value of autocorrelation function is equal to 0.0781. When sample size is big (≥ 60) we cannot reject Null hypothesis (about equivalence of quota to probability on the plane) even with 14.2% significance level (Student t-criterion). In all cases Null hypothesis cannot be rejected with 1% significance level. Subset of square $[0,1]\times[0,1]$ was found stochastically was equal to 0.500363.

Subset of cube $[0,1]\times[0,1]\times[0,1]$ was found stochastically was equal to 0.002062. In 44 cases we had a situation when Null hypothesis must be rejected with 5% significance level (for several samples with sizes from 31520 and up to 3223). In all other cases (5956) we cannot reject Null hypothesis with 5% significance level.

3. Conclusion

Modification of Lehmer algorithm didn't allow obtaining a situation when search of suitable initial value is easy than it is realized for original map. Analysis of modified map showed that for some initial values we may have short cycles of the length 2.

Periodic disturbances of cyclic trajectories (with rather big values of cycle length) can lead to appearance of combined trajectories (which are combinations of parts of trajectories of modified Lehmer algorithm) with rather small values of autocorrelation function and good distribution of points in square and three-dimensional cube. Thus, it can be considered as one of possible ways for further modifications of map (2) for obtaining of pseudo-stochastic values with uniform distribution on [0,1].

4. References

Ermakov S.M., 1975. Monte Carlo method and close problems. Moscow: Nauka, 472 p.

Lehmer D.H., 1951. Mathematical methods in large-scale computing units// Proceedings of a Second Symposium on Large-Scale Digital Calculating Machinery, 1949. Cambridge: Harvard University Press: 141—146.

Mikhailov G.A., Voitishek A.V., 2006. Numerical statistical modeling (Monte Carlo methods). Moscow: Academia, 368 p.

Nedorezov L.V., 1986. Methods of statistical trials. Methodological instructions for students. Krasnoyarsk: Krasnoyarsk State Polytechnic University, 33 p.

Sobol I.M., 1973. Numerical Monte Carlo methods. Moscow: Nauka, 312 p.

Tyurin Yu.N., Figurnov V.E., 1990. On checking of algorithms of obtaining of stochastic numbers// Theory of Probability and its Applications 35(1): 156-161.