

## A STOCHASTIC SEARCH ALGORITHM WITH AN APPLICATION TO MULTIDIMENSIONAL INTEGRATION

K.A. Borovkov

Steklov Mathematical Institute, Vavilov Street 42, 117966 Moscow GSP-1, Russia

1. Recently Lalley and Robbins [2] suggested an asymptotically minimax strategy of stochastic search for the so-called “princess and monster” game. According to this strategy, the movement of the monster in a convex domain  $\Omega$  on the plane, while it is looking for the princess, can be described by a piecewise linear ergodic semi-Markov process  $X(t)$ , whose stationary distribution proves to be uniform on  $\Omega$ . Thus, for any integrable function  $f$  on  $\Omega$  one has

$$\lim_{u \rightarrow \infty} \frac{1}{u} \int_0^u f(X(t)) dt = \frac{I}{\mu(\Omega)} \quad \text{a.s.}, \quad I = \int_{\Omega} f(x) \mu(dx),$$

where  $\mu$  is the Lebesgue measure on the plane.

In the present paper we describe a generalization of this type of a process to the multidimensional case. We suggest a variant of the Monte Carlo method for multidimensional integration which might turn out to be rather effective in a number of situations. Moreover, the stationary distribution of the embedded sequence of hitting points of our process on the boundary  $\partial\Omega$  of the domain  $\Omega$  appears to be uniform on  $\Omega$ . Thus, the suggested procedure allows us to evaluate surface integrals, too.

Recall that, in the multidimensional case, the crude approach enclosing a given domain in a cube and then taking a random sample from the uniform distribution on this cube turns out to be useless. In fact, the efficiency of such a sampling procedure is very low. (For example, the ratio of the volume of a given ball in an  $m$ -dimensional Euclidean space to the volume of the smallest possible cube enclosing it is  $\pi^{m/2} / (2^m \Gamma(\frac{1}{2}m + 1))$ , tending to zero very fast as  $m \rightarrow \infty$ . Thus, almost all the simulated random samples will be lost.) Our approach, estimating the integral to be evaluated by the average value of integrals over the segments of straight lines along which the monster crosses the domain  $\Omega$  (in this respect it is similar to stereology), does not imply such losses. Also many other practical methods have this property, too (see e.g. [1, 4]). But, as already mentioned above, our method possesses an essential advantage: its unique feature is that it allows us to evaluate simultaneously also the surface integral over  $\partial\Omega$ . It should be noted that a direct application of our method leads to an estimate, not of

$I$ , but the ratio  $I/\mu(\Omega)$ . In order to obtain an estimate of  $I$  itself one needs an estimate of the volume of the domain  $\Omega$  or of the area of  $\partial\Omega$  (cf. (8)). Of course, we can estimate  $\mu(\Omega)$  by using our method another time (either starting from some "good" domain  $\Omega'$  with known parameters or by evaluating a known volume of some domain  $\Omega'' \subset \Omega$  using our procedure in  $\Omega$ ). Moreover, it is worth pointing out that knowing the volume of  $\Omega$  we can, according to the properties of our method, evaluate the area of the surface  $\partial\Omega$ , and vice versa, see formula (3) below.

Let  $\Omega$  be a convex domain in  $R^m$  having the property that its boundary  $\partial\Omega$  belongs to the class  $C^1$  (in fact it could be piecewise smooth as well). Let  $\nu_0$  be the area measure on  $\partial\Omega$ , i.e.  $\nu(\cdot) = \nu_0(\cdot)/\nu_0(\partial\Omega)$  is the uniform distribution on  $\partial\Omega$ . The direction of the movement of the random walk process will be given by the points on the unit sphere  $S = S^{m-1}$  (with the center at the origin). For this we assume that for each  $x = (x_1, \dots, x_m) \in \partial\Omega$  an orthogonal transformation  $F(x): R^m \rightarrow R^m$  is given such that the image of  $e_m$  ( $e_k$ ,  $k = 1, \dots, m$ , are the elements of the canonical orthonormal basis in  $R^m$ ) under  $F(x)$  is  $n_x$ , the inner normal vector to  $\partial\Omega$  of unit length at the point  $x \in \partial\Omega$ .

A random walk process will be defined in the following way. Let  $\{M_i^0\}_{i \geq 1}$  be a sequence of i.i.d. random vectors on  $S_+ = S \cap \{x_m > 0\}$  such that for any  $x \in S_+$

$$(1) \quad P(M_i^0 \in ds_x) = C_m x_m \lambda(ds_x),$$

where  $ds_x$  is a surface element on  $S$  at the point  $x = (x_1, \dots, x_m)$ ,

$$C_m = \pi^{(1-m)/2} \Gamma(\frac{1}{2}(m+1))$$

and  $\lambda$  is the area measure on  $S$ . The random walk  $X(t)$ ,  $t \geq 0$ , starts at an arbitrary point  $P_0 \in \partial\Omega$  proceeding first with the unit speed in the direction of the vector  $M_1 = F(P_0)M_1^0$  until it hits  $\partial\Omega$  (at the point  $P_1$ , say). Then, at this point the random walk process again chooses a new direction given by  $M_2 = F(P_1)M_2^0$  and proceeds in this direction, still with the unit velocity, until hitting  $\partial\Omega$  for another time (at the point  $P_2$ ), and so on. Denote by  $T_n$  the time when the process arrives at the point  $P_n$ , moving along the  $n$ th segment of its trajectory.

The main result of the paper is the following

**Theorem.** *The Markov chain  $\{P_i\}_{i > 0}$  on  $\partial\Omega$  and the semi-Markov process  $\{X(t), t \geq 0\}$  in  $\Omega$ , defined above, are ergodic and their stationary distributions are uniform on  $\partial\Omega$  and  $\Omega$ , respectively. Furthermore, for any integrable function  $f: \Omega \rightarrow R$*

$$(2) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \int_0^{T_n} f(X(t)) dt = \frac{B_m}{\nu_0(\partial\Omega)} \int_{\Omega} f(x) \mu(dx) \quad \text{a.s.}$$

where  $B_m = \pi^{1/2} m \Gamma(\frac{1}{2}(m+1)) / \Gamma(\frac{1}{2}m)$  and  $\mu = \mu^{(m)}$  is the Lebesgue measure on  $R^m$ .

Therefore, by this Theorem for any  $\nu_0$ -integrable function  $g$  on  $\partial\Omega$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n g(P_i) = \frac{1}{\nu_0(\partial\Omega)} \int_{\partial\Omega} g(x) \nu_0(dp_x) \quad \text{a.s.}$$

where  $dp_x$  is the surface element on  $\partial\Omega$  at a point  $x$ . Moreover, by choosing  $f \equiv 1$  in (2), we get

$$(3) \quad \lim_{n \rightarrow \infty} \frac{1}{n} T_n = \frac{B_m \mu(\Omega)}{\nu_0(\partial\Omega)} \quad \text{a.s.},$$

and therefore the limiting relation

$$(2') \quad \lim_{u \rightarrow \infty} \frac{1}{u} \int_0^u f(X(t)) dt = \frac{1}{\mu(\Omega)} \int_{\Omega} f(x) \mu(dx) \quad \text{a.s.}$$

is valid in the multidimensional case as well.

**Remark.** It is worth pointing out that at first sight the suggested process could seem to be similar to the process considered by Turchin [4]. In his paper a Markov chain  $\{X_n\}$  taking values in  $\Omega$  with uniform limiting distribution was described. The mechanism of this chain is in short the following one. Given  $X_n$  we draw a straight line passing through this point in one of directions taken from a given set of them (for example, we could take at random one of the elements of a basis in  $R^m$ ), and then  $X_{n+1}$  is selected according to the uniform probability distribution on the section of  $\Omega$  on this line. Thus, the constructions of these two processes are quite different. The essence of this difference is that in [4] there is a Markov random walk in  $\Omega$ , while we have a random walk on  $\partial\Omega$ . In fact, the “generating mechanism” in [4] consists of a random sampling procedure of points from the “punctures” of  $\Omega$ , and in our case the crucial point is the “reflecting law”. (The possibility to evaluate the surface integrals follows from the special choice of this law.) Furthermore, the definition of the estimator as an arithmetic mean of linear integrals over the punctures looks preferable even for the process introduced in [4]. Clearly, such an integral over the  $n$ th puncture will be equal to the conditional expectation of  $f(X_n)$  given  $X_{n-1}$  and the direction of the  $n$ th puncture. Thus its variance in non-trivial cases will be less than the variance of  $f(X_n)$ . In fact, it is a priori obvious that the integral over the puncture gives more information about the integral than simply the value of the function at one point. Moreover, from the computational point of view an evaluation of such a linear integral is usually not difficult and does not require much time. This was confirmed in our simulations.

2. We present here some notes on a practical realization of the method. Formulae (2) and (2') suggest the use of the statistics

$$(4) \quad \frac{\mu(\Omega)}{u} \int_0^u f(X(t)) dt \quad \text{and} \quad \frac{\nu_0(\partial\Omega)}{n} \int_0^{T_n} f(X(t)) dt$$

for large  $u$  and  $n$  as estimators of  $I$ . (The mixing rate of  $X(t)$ ,  $t \geq 0$ , is exponential, so that the error for a wide class of functions  $f$  will be of the order  $u^{-1/2}$  or  $n^{-1/2}$ ). Since the process  $X(t)$ ,  $t \geq 0$ , is ergodic, the random walk can start from any point on  $\partial\Omega$ . The main problem is then to define, at the points  $P_i$ , the "correct sampling method" of the direction in which the random walk process, will next cross the domain  $\Omega$ .

The simplest way to generate the random vectors  $M_i^0$  is the following one. It is easy to see that for the vector  $M_i^0 = (Z_1, \dots, Z_m)$  the density function of the angle  $\Theta = \arccos(Z_m)$  of its deviation from  $e_m$  is proportional to  $\cos(\theta) \sin^{m-1}(\theta)$ . Thus,

$$\mathbf{P}(\Theta < \theta) = \sin^{m-1}(\theta), \quad 0 \leq \theta \leq \frac{1}{2}\pi.$$

Therefore  $\sin(\Theta) = (Z_1^2 + \dots + Z_{m-1}^2)^{1/2}$  has the distribution function  $t^{m-1}$ ,  $0 \leq t \leq 1$ , while the vector  $(\sin(\Theta))^{-1/2}(Z_1, \dots, Z_{m-1})$  is uniformly distributed over  $S^{m-2}$ . Hence the random vectors

$$M_1^0 \quad \text{and} \quad (Y_1 V/R, \dots, Y_{m-1} V/R, (1 - V^2)^{1/2})$$

have the same distribution; here  $Y_1, \dots, Y_{m-1}$  are independent standard normal variables and  $R = (Y_1^2 + \dots + Y_{m-1}^2)^{1/2}$ . Moreover,  $V = U^{1/(m-1)}$ , where the random variable  $U$  is uniformly distributed on  $[0, 1]$ , and is independent of  $Y_1, \dots, Y_{m-1}$ .

The next step is to define the local rule  $F(x)$  of "rotation" of the vector  $M_i^0$  for the actual choice of direction  $M_i$  of the movement from a point  $x \in \partial\Omega$ . For this it will suffice, for example, to define locally the curvilinear systems of coordinates on  $\partial\Omega$ . If such a system is defined through a parametrization  $x_i = x_i(\xi_1, \dots, \xi_{m-1})$ ,  $i = 1, \dots, m$ , on a given part  $\Gamma \subset \partial\Omega$  of the surface then the vectors

$$j_k = j_k(x) = j_k^0 / |j_k^0|, \quad k = 1, \dots, m-1 \quad (j_{k,i}^0 = \partial x_i / \partial \xi_k)$$

will form an orthonormal basis of the system; here  $|\cdot|^2 = (\cdot, \cdot)$ . The matrix of the transformation  $F(x)$  in this part  $\Gamma$  of the boundary will have  $m$  columns  $j_1, \dots, j_{m-1}, n_x$  (all these vectors are considered to be columns). This scheme is easy to implement when  $\Gamma$  admits a simple parametrization, say when  $\Gamma$  is a part of a sphere, cylinder, plane etc. In more complicated situations, after

having determined the normal vector  $n_x$  one could then apply an orthogonalization process to the set of vectors  $e_k - (n_x, e_k)n_x$ ,  $k = 1, \dots, m$ .

The integral

$$\int_{T_i}^{T_{i+1}} f(X(t)) dt$$

is just the integral of  $f$  over the line segment  $P_i P_{i+1}$ . To evaluate it one could make use of any suitable numerical integration formula.

Finally, one can obtain estimates of the variances of our estimators, too. For this we first compute the sample covariances of the integrals of the random walk over the segments  $P_i P_{i+1}$  and  $P_{k+1} P_{k+i+1}$   $k$  steps apart. Then the well-known formula for the limiting variance of the sum of values of a functional on a Markov chain can be applied.

To test the efficiency of the method we have carried out several evaluations of integrals of the type

$$I = \int_{V_m} f(x) \mu(dx), \quad V_m = \{x : |x| \leq 1\},$$

for smooth as well as discontinuous functions  $f$ . (Of course, in this special case one could readily generate random vectors that are uniformly distributed in  $\Omega = V_m$ . It is considered only as an illustrative example.) The integrals over the segments were calculated by the trapezoidal quadrature formula by using the step size  $|P_i P_{i+1}|/40$ .

For  $m = 20$  the advantage of our method, when compared to the crude one, becomes overwhelming. To obtain an estimate with the same order of the error it requires thousands times less computer time. Typical results of the simulation are given in the following table.

$f(x)$	$10^2 \cdot I$	$10^2 \cdot \bar{I}_{10}$	$10^2 \cdot \sigma_{10}$
$ x ^5$	2.0646	2.0665	.1147
$2 \text{ sign}(x_m)$	2.5807	2.6046	.1373
$\mathbf{1}_A$	1.2903	1.2932	.1719

Here for each of the functions there were 10 series of the random walk process consisting of 1000 steps for  $f = |x|^5$  and, respectively, 100 steps in other cases. Furthermore,

$$\bar{I}_{10} = \frac{1}{10} \sum_{i=1}^{10} I_i,$$

where  $I_i$  is the estimate of  $I$  resulting from the  $i$ th series,

$$\sigma_{10} = \left( \frac{1}{9} \left( \sum_{i=1}^{10} I_i^2 - 10 \cdot \bar{I}_{10}^2 \right) \right)^{1/2},$$

$\mathbf{1}_A$  is the indicator function of the set  $A = .5^{1/20}V_{20}$  (a ball with the volume  $\frac{1}{2}\mu(V_{20})$ ).

#### References

- [1] DEVROYE, L.: Non-uniform random variate generation. - Springer-Verlag, New York, 1986.
- [2] LALLEY, S., and H. ROBBINS: Stochastic search in a convex region. - Probab. Theory Related Fields 77, 1988, 99–116.
- [3] REVUZ, D.: Markov chains. - North-Holland, Amsterdam, 1975.
- [4] TURCHIN, V.F.: On the computation of multidimensional integrals by the Monte-Carlo method. - Theory Probab. Appl. 16, 1971, 720–724 (English translation of Teor. Veroyatnost. i Primenen. 16, 1971, 738–743).