Journal scientific and applied research, vol. 16, 2019
International Journal

Original Contribution ISSN 1314-6289

MONTE CARLO SAMPLING TECHNIQUES FOR COMPUTATION OF MULTIDIMENSIONAL INTEGRALS RELATED TO MIGRATION

Venelin Todorov[1,2], Valerij Dzhurov[3], Tosho Stanchev[3], Iliyan Tsvetkov[3], Yuri Dimitrov[4]

[1] INSTITUTE OF MATHEMATICS AND INFORMATICS, BULGARIAN ACADEMY OF SCIENCES, DEPARTMENT OF INFORMATION MODELING, ACAD. GEORGI BONCHEV STR., BLOCK 8, 1113 SOFIA, BULGARIA
[2] INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGIES, BULGARIAN ACADEMY OF SCIENCES, DEPARTMENT OF PARALLEL ALGORITHMS, ACAD. G. BONCHEV STR., BLOCK 25 A, 1113, SOFIA, BULGARIA
[3] ROUSSE UNIVERSITY "ANGEL KANCHEV"
[4] UNIVERSITY OF FORESTRY, SOFIA, DEPARTMENT OF MATHEMATICS AND PHYSICS

E-mail: vtodorov@math.bas.bg,venelin@parallel.bas.bg

ABSTRACT: *We study multidimensional integrals with applications to international migration forecasting. A comprehensive experimental study based on Latin Hypercube and Importance Sampling and Fibonacci based lattice rule has been done. This is the first time such a comparison has been made. The numerical tests show that the stochastic algorithms under consideration are efficient tool for computing multidimensional integrals. It is important in order to obtain a more accurate and reliable interpretation of the results which is a foundation in international migration forecasting.*

KEYWORDS: *Monte Carlo methods, multidimensional integrals, Latin hypercube sampling, Fibonacci lattice rule, Importance sampling, international migration forecasting.*

Introduction

Monte Carlo methods are among the most widely used stochastic methods for solving problems in different areas. Monte Carlo methods always produce an approximation to the solution of the problem or to some functional of the solution, but one can control the accuracy in terms of the probability error and risk evaluation for immigrant flows [4]. An important advantage of the Monte Carlo methods is that they are suitable for solving multi-dimensional problems [3], since the complexity increases polynomially, but not exponentially with the dimensionality [12,13,14]. Another advantage of the method is that it allows to compute directly functionals of the solution with the same complexity as to determine the solution. In such a way this class of methods can be considered as a good candidate for treating innovative problems related to modern areas in quantum physics and finance. The Monte Carlo methods for multidimensional integrals can be used in problems where the data is taken in randomized way such that locating targets or searching for radio signals. multi-dimensional problems [3], since the complexity increases polynomially, but not exponentially with the dimensionality.

Forecasting international migration is an important, yet difficult research task, characterized by the highest errors among the forecasts of all components of the demographic change [1]. Reasons for this include a lack of a comprehensive migration theory, difficulties in the theoretical framework of migration [5,6,24], uncertainty of potential explanatory variables, ignoring forced migration and policy elements in the forecasts, as well as poor data quality [3]. In order to improve accuracy of the international migration forecasts, attempts should be also made to improve the forecasting methodology [1,2,3].
The main drawback of a majority mathematical models of migration, apart from the event-history analysis, is that they themselves do not explicitly address the issue of uncertainty, important for preparing any forecast on their basis [4]. Although some of the models apply Markov chains [5,6,10,11], and can be therefore used to assess uncertainty using simulations, this possibility has not been explored up to date. However, the assessments of uncertainty may be also included in a majority of demographic models (cohort-component, multi-regional, or multi-state) by feeding them at input with stochastic forecasts of particular components of demographic change. The latter may involve econometric forecasts and time series models, both in the sample-theory and the Bayesian frameworks [7,23].

A fundamental problem in this methodology is the accurate evaluation of multidimensional integrals. High dimensional integrals are usually solved with Monte Carlo algorithms. Monte Carlo method is the only possible method for high-dimensional problems since its convergence is independent of the dimension. Monte Carlo methods give statistical estimates for the functional of the solution by performing random sampling of a certain random variable whose mathematical expectation is the desired functional. Monte Carlo methods are methods of approximation of the solution to problems of computational mathematics, by using random processes for each such problem, with the parameters of the process equal to the solution of the problem. The method can guarantee that the error of Monte Carlo approximation is smaller than a given value with a certain probability [8].

Description of the problem

Consider the problem of approximate integration of the multiple integral:

$$\int_{[0,1)^d} f(\mathrm{x})d\mathrm{x} = \int_0^1 d\mathrm{x}^{(1)} \int_0^1 d\mathrm{x}^{(2)} \ldots \int_0^1 d\mathrm{x}^{(\mathrm{d})} f(x^{(1)}, x^{(2)}, \ldots,, x^{(d)}),$$

where

$$x = (x^{(1)}, \ldots, x^{(d)}) \in [0,1)^d.$$

For small values of d, numerical integration methods such as Simpson's rule or the trapezoidal rule (see Davis and Rabinowitz [19]) can be used to approximate the integral (1). These methods, however, suffer from the so-called curse of dimensionality and be-come impractical as s increases beyond 3 or 4. The Crude Monte Carlo method has rate of convergence $O(N^{-1/2})$, where N is the number of samples, which is independent of the dimension of the integral, and that is why Monte Carlo integration is the only practical method for many high-dimensional problems.

Importance Sampling

Importance sampling is choosing a good distribution from which to simulate one's random variables. It involves multiplying the integrand by 1 (usually dressed up in a \tricky fashion") to yield an expectation of a quantity that varies less than the original integrand over the region of integration. For example, let h(x) be a density for the random variable X [21]. All we need to do to have a Monte Carlo estimator with zero variance is use and make sure that our density h is proportional to the function g. The ability to simulate independent random variables from h(x), or the ability to compute the density h(x), itself, implies that the normalizing constant of the distribution is computable, which in turn would imply that the original integral involving g(x) is computable. While h(x) might be roughly the same shape as g(x), serious difficulties arise if h(x) gets small much faster than g(x) out in the tails. In such a case, though it is improbable (by definition) that you will realize a value x_i from the far tails of h(x), if you do, then your Monte Carlo estimator will take $g(x_i)=h(x_i)$ for such an improbable x_i may be orders of magnitude larger than the typical values g(x)=h(x) that you see [25].

Latin Hypercube Sampling

The main problem of some of the widely used Monte Carlo methods such as Importance sampling is that a sample that is very close to another does not provide much new information about the function being integrated. One powerful variance-reduction technique that addresses this problem is called stratified sampling. Stratified sampling works by splitting up the original integral into a sum of integrals over sub-domains. In its simplest form, stratified sampling divides the domain G_d into N sub-domains (or stratas) and places a random sample within each of these intervals. A quantity of interest is the variance of the obtained approximation, considered as a random variable. It can be shown that stratified sampling can never result in higher variance than pure random sampling [8]. If N = 1, we have random sampling over the entire sample space (see [8]).

The Latin Hypercube Sampling (LHS) was described by McKay in 1979 [18]. If one wishes to ensure that each of the input variables x_i has all portions of its distribution represented by input values, we can divide the range of each x_i, in our case the interval [0,1], into M strata of equal marginal probability 1/M, and sample once from each stratum. In the case of integral approximation we must simply divide the interval [0,1] into M disjoint intervals, each of length 1/M and to sample one point from each of them. Let this sample be X_{kj}, for dimensions k = 1,..,d, j = 1,…, M. Those of them having first index k (k = 1,…, d) are the different components for the k-th dimension of the random points that are used for the integral's approximation. These components are matched at random. Thus the maximum number of combinations for a Latin Hypercube of M divisions and s variables (i.e., dimensions) can be computed by the formula $(M!)^{d-1}$. In the context of statistical sampling, a square grid containing sample positions is a Latin square if (and only if) there is only one sample in each row and each column. Thus following the described algorithm, we obtain a set of points with positions forming a Latin square [9]. Note that this sampling scheme does not require more samples for more dimensions (variables); this independence is one of the main advantages of LHS scheme. In LHS one must first decide how many sample points to use and for each sample point remember in which row and column the sample point was taken. Note that such configuration is similar to having N rooks on a chess board without threatening each other. A Latin Hypercube is the generalization of this concept to an arbitrary number of dimensions, whereby each sample is the only one in each axis-aligned hyperplane containing it. To prove that the variance of the LHS is smaller than the variance of IS we use a theorem proved in [18].

Generally, the time complexity of the algorithm depends on the integrand. However, it follows easily that the computational complexity of the LHS is linear, we will have only a constant number of additional operations compared to the regular Crude MC method and it is easy to show that the computational complexity of the crude Monte Carlo is linear.

Fibonacci based lattice rule

The monographs of Sloan and Kachoyan [20] and Wang and Hickernell [22] provide comprehensive expositions of the theory of integration lattices.

Let n be an integer, and $a =(a_1,...,a_s)$ be an integer vector modulo n. A set of the form [15]

$$P_n = \left\{ \left\{ \frac{ak}{n} \right\} = \left(\left\{ \frac{a_1 k}{n} \right\}, \dots, \left\{ \frac{a_s k}{n} \right\} \right) \mid k = 1, \dots, n \right\}$$

is called a lattice point set, where {x} denotes the fractional part of x. The vector a is called a lattice point or generator of the set. As one can see, the formula for the lattice point set is simple to program. The difficulty lies in finding a good value of a, such that the points in the set are evenly spread over the unit cube. The choice of good

generating vector, which leads to small errors, is not trivial. Complicated methods from theory of numbers have been used. We consider the following generating vector based on generalized Fibonacii numbers of corresponding dimensionality:

$$a = (1, F_{l+1}^{(s)}, ..., F_{l+s-1}^{(s)})), \quad n_l = F_l^{(s)},$$

where

$$F_{l+s}^{(s)} = F_l^{(s)} + F_{l+1}^{(s)} + ... + F_{l+s-1}^{(s)}, l = 0, 1, \dots$$

with initial conditions

$$F_0^{(s)} = F_1^{(s)} = \dots = F_{s-2}^{(s)} = 0, F_{s-1}^{(s)} = 1,$$

for l=0,1,…

The discrepancy of the set obtained by using the vector described above is asymptotically estimated in [21].

The number of calculation required to obtain the generating vector is O(ln n_l). The generation of a new point requires constant number of operations, thus to obtain a lattice set of the described kind consisting of n_l points, O(ln n_l) number of operations are necessary.

Numerical example

We want to compute the following 7 dimensional integral:

$$\int_{[0,1]^7} e^{1-\sum_{i=1}^{3} sin(\frac{\pi}{2}.x_i)} .arcsin\left(sin(1)+\frac{\sum_{j=1}^{7} x_j}{200}\right) \approx 0.7515.$$

We will use the IS algorithm with probability density function

$f(\overline{x}) = \frac{1}{\eta} e^{1-\sum_{i=1}^{3} sin(\frac{\pi}{2}.x_i)}$.

The value of η must be found separately. It is equal to the value of the integral

$$\int_{[0,1]^3} e^{1-\sum_{i=1}^{3} sin(\frac{\pi}{2}.x_i)}.$$

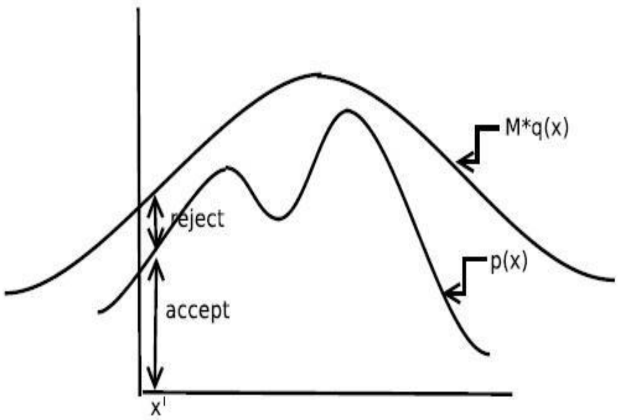

Figure 1. The acceptance-rejection method (the method of selection).

We evaluate the last integral with Crude Monte Carlo method for a number of samples N = 10^5. After that we use the method of selection (the acceptance-rejection method). The idea of the method is given by the Figure 1.

We make a comparison between the Latin Hypercube Sampling (LHS), the Importance sampling (IS) and the Fibonacci based lattice rule (FIBO) which is the goal of this paper. As we expected LHS outperforms IS by far. The results are given in the Tables 1 and 2 below. Each table contains information about the MC approach which is applied, the obtained relative error, the needed CPU time and the number of points. Table 1 shows the relative error for a given number of samples, while in Table 2 is presented the relative error for a fixed computational time, which is a measure of the computational complexity of the algorithms.

It can be seen than the CPU time for the Latin hypercube sampling and FIBO method is closer, while Importance sampling method needs much higher time. Also we see that for a large number of samples the LHS and FIBO outperforms the importance sampling. Therefore the LHS is definitely better from the Importance sampling method for computing multidimensional integrals.

Table 1. Relative error for 7 dimensional integral for a given number of samples.

N	FIBO	time	LHS	time	IS	time
1000	1.03e-3	0.15	2.46e-3	0.13	6.61e-3	3.1
5000	2.71e-2	0.75	9.03e-4	0.71	3.35e-3	15.8
10000	3.34e-4	1.52	4.43e-4	1.50	1.09e-3	30.2
25000	2.73e-4	3.51	1.10e-4	3.27	7.51e-4	74.5
10^5	1.62e-5	13.8	7.26e-5	14.2	5.41e-4	315
10^6	1.02e-6	104	6.26e-6	110	2.53e-4	3056

Table 2. Relative error for 7 dimensional integral and equal execution times.

time, sec.	FIBO	LHS	IS
0.1	1.38e-3	2.37e-3	5.51e-2
1	2.87e-4	3.37e-4	2.31e-2
5	1.16e-4	1.38e-4	8.05e-3
10	5.28e-5	8.78e-5	4.91e-3
20	2.26e-5	6.87e-5	2.58e-3
100	1.61e-6	7.01e-6	7.18e-4

In the Table 1 and 2 are presented the relative error for the 7 dimensional integrals with Fibonacci lattice sequence (FIBO), Latin hypercube sampling (LHS) and Importance sampling (IS) for a fixed number of points and for a fixed computational time. Obviously FIBO and LHS has the lowest computational complexity and are the fastest algorithm, while IS is slower. As can been seen from the results for 7 dimensional integral, the Importance sampling produces the worst results. It is interesting to see that FIBO and LHS gives errors of the same magnitude - see Table 1, but the Fibonacci lattice sequence has the advantage for a preliminary given time in seconds - see Table 2. So we can conclude that all stochastic algorithms under consideration are efficient tool for evaluation of multidimensional integrals related to models in migration forecasting.

Conclusion

In this paper we analyze the performance of different Monte Carlo methods for multidimensional integrals related to models in improving the international migration. Stochastic methods under consideration are an efficient way to solve problems in forecasting international migration. The problem of accurate evaluation of the presented by multidimensional integrals. This is the first time a particular 1-rank lattice rule based on Fibonacci generating vector is compared with the importance sampling technique. We make a comparison with the Latin Hypercube Sampling and it gives closer results to the Fibonacci method. It is a crucial element since this may be important for improving the international migration forecasting.

Acknowledgement

The author Venelin Todorov is supported by the Bulgarian National Fund of Science under Project DN 12/5-2017, "Efficient Stochastic Methods and Algorithms for Large-Scale Problems" and by the National Scientific Program "Information and Communication Technologies for a Single Digital Market in Science, Education and Security (ICT in SES)", financed by the Ministry of Education and Science in Bulgaria.
This material is financed by FNI-19, Faculty EEA 06/2019.

Reference

[1]. J. Alho, (1990). Stochastic methods in population forecasting. *International Journal of Forecasting*, 6 (4), 521–530.

[2]. M.J. Bayarri, and J. O. Berger (2004). The Interplay of Bayesian and Frequentist Analysis. *Statistical Science*, 19 (1), 58–80.

[3]. Jakub Bijak, Bayesian methods in international migration forecasting, 2008, In book International migration in Europe: Data, Models and Estimates, Chapter 12, https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470985557.ch12

[4]. Paul Bratley, Bennett Fox, Harald Niederreiter, Implementation and Tests of Low Discrepancy Sequences, ACM Transactions on Modeling and Computer Simulation, Volume 2, Number 3, July 1992, pages 195-213.

[5]. L.A. Brown, (1970). On the use of Markov chains in movement research. *Economic Geography*, 46 (Suppl.), 393–403.

[6]. H. Brücker and B. Siliverstovs (2005). *On the Estimation and Forecasting of International Migration: How Relevant is Heterogeneity Across Countries*. IZA Discussion Paper 170. Bonn: Institut zur Zukunft der Arbeit.

[7]. P. Congdon, 2003, *Applied Bayesian Modelling*. Chichester: John Wiley.

[8]. I. Dimov, Monte Carlo Methods for Applied Scientists, New Jersey, London, Singapore, World Scientific, 2008, 291p.

[9]. Dimov, I, Chernogorova, T., Vulkov, L.. Coordinate Transformation Approach

for Numerical Solution of Environmental Problems. Mathematical Problems in Meteorological Modelling, series Mathematics in Industry, 24, Springer International Publishing, 2016, ISBN:978-3-319-40155-3, ISSN:1612-3956, DOI:10.1007/978-3-319-40157-7_7, 117-127

[10]. Dimov, I., M. Nedjalkov**,** J. M. Sellier, S. Selberherr. Neumann Series Analysis of the Wigner Equation Solution. Progress in Industrial Mathematics at ECMI 2014, Mathematics in Industry, 22, Springer International Publishing, G. Russo, V. Capasso, G. Nicosia, V. Romano (ed), 2016, ISBN:978-3-319-23412-0, DOI:10.1007/978-3-319-23413-7_97, 701-707

[11]. Dimov, I., Zlatev, Z., Georgiev, K.**,** Margenov, S.**.** Numerical algorithms for scientific and engineering applications. Journal of Computational and Applied Mathematics, 310, Elsevier, 2017, ISSN:0377-0427, 1-4.

[12]. Henri Faure, Discrepance de suites associees a un systeme de numeration (en dimension s), Acta Arithmetica, Vol. 41, 1982, 337-351.

[13]. John Hammersley, Monte Carlo methods for solving multivariable problems, Proceedings of the New York Academy of Science, Volume 86, 1960, 844-874.

[14]. J.M. Hammersley, D.C. Handscomb, (1964). Monte Carlo Methods.

[15]. L.K. Hua, Y. Wang: Applications of Number Theory to Numerical analysis, 1981.

[16]. J.A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, 14 (4), 382–417.

[17]. Lin S., “Algebraic Methods for Evaluating Integrals in Bayesian Statistics,” UC Berkeley, May 2011.

[18]. McKay, M.D., Beckman, R.J., Conover, W.J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2), 23945 (1979)

[19]. Rabinowitz. P. and Davis P.: Methods of Numerical Integration. Academic Press, London, 2nd edition, (1984).

[20]. I.H. Sloan and P.J. Kachoyan, Lattice methods for multiple integration: Theory, error analysis and examples, (1987), J. Numer. Anal., 24, 116-128.

[21]. Van der Corput, J.G.: Verteilungsfunktionen I-VIII. Proc. Akad. Amsterdam, Vol. 38 (1935) 813-821, 1058-1066, Vol.39 (1936) 10-19, 19-26, 149-153, 339-344.

[22]. Y. Wang and F.J. Hickernell, An historical overview of lattice point sets, (2002).

[23]. W. Weidlich and G. Haag (eds.) (1988). *Interregional migration: dynamic theory and comparative analysis*. Berlin-Heidelberg: Springer

[24]. A. Zellner, (1971). *An Introduction to Bayesian Inference in Econometrics*. NY: John Wiley.

[25]. Lectures: Monte Carlo methods nd Importance sampling: http://ib.berkeley.edu/labs/slatkin/eriq/classes/guest_lect/mc_lecture_notes.pdf