Journal scientific and applied research, vol. 13, 2018
Association Scientific and Applied Research
International Journal

Original Contribution ISSN 1314-6289

QUASI-MONTE CARLO METHODS FOR COMPUTATION OF MULTIDIMENSIONAL INTEGRALS RELATED TO BAYESIAN MODELS IN INTERNATIONAL MIGRATION FORECASTING

Venelin Todorov[1,2], Valentin Dimitrov[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: *The paper addresses selected methodological aspects of international migration forecasting. The new methods based on the Bayesian statistics have been recently developed. A fundamental problem in Bayesian statistics is the accurate evaluation of multidimensional integrals. A comprehensive experimental study based on Faure and Hammersley low discrepancy sequences and Fibonacci based lattice rule has been done. 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 in Bayesian statistics which is a foundation in international migration forecasting.*

KEYWORDS: *Quasi-Monte Carlo sequences, multidimensional integrals, Hammersley sequence, Fibonacci lattice rule, Faure sequence, international migration forecasting, Bayesian statistics.*

Introduction

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], 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].

Bayesian model for forecasting international migration

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 [3]. Although some of the models apply Markov chains [5,10,11], 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.

In the last few years an alternative approach based on the paradigm of Bayesian statistics has been developed in [3]. The methodology that would combine the advantages of the existing ones, including both the formality of the applied statistical tools, and including subjective expert judgment in the forecasting model is presented in [3]. This methodology allows for construction of forecasting models combining the formal methods with the subjective expertise [2,3].

Forecasting in the Bayesian approach is based on the construction of a probability distribution of the vector of future values of the variable under study, conditional on the vector of past (observed) values, and taking into account the posterior knowledge on the parameters of the forecasting model [7]. Bayesian methodology can reduce the estimation and prediction errors, in case the prior distribution is informative and consistent with the observations [3,4,9]. This is important in the small-sample studies (e.g., with population disaggregated by sex, age, regions, etc.), where the prior information has relatively more weight in the posterior result than the observations, unlike in large datasets [6,9]. The extreme estimates obtained from small-sample data are in this way corrected towards the prior expectations. The same applies to forecasting models based on short time series, where the Bayesian approach is a way to reduce uncertainty [23]. Additionally, the Bayesian methodology allows for a formal model selection in order to maximally utilize information from the sample, by comparing the posterior odds of different models given the data [22].

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].

Quasi-Monte Carlo algorithms for numerical integration

In the last few years new approaches have been developed that outperform standard Monte Carlo in terms of numerical efficiency. It has been found that there can be efficiency gains in using deterministic sequences rather than the random sequences which are a feature of standard Monte Carlo. These deterministic sequences are carefully selected so that they are well dispersed throughout the region of integration. Sequences with this property are known as low discrepancy sequences. These sequences are often more efficient than standard Monte Carlo in evaluating high dimensional integrals if the integrand is sufficiently regular.

They are usually superior to the Monte Carlo methods as they have a convergence rate of $((\log N)^s/N)$, where N is the number of samples and s is the dimensionality of the problem under consideration.

The standard M-dimensional Hammersley sequence [13] based on a number of samples N is simply composed of a first component of successive fractions 0/N, 1/N, ..., N/N, paired with M-1 1-dimensional van der Corput sequences [20], using as bases the first M-1 primes. The van der Corput sequence generates a sequence of points in [0,1] which never repeats. For positive index I, the elements of the van der Corput sequence are strictly between 0 and 1. In particular, the I-th element of the van der Corput sequence is computed by writing I in the base B (usually 2) and then reflecting its digits about the decimal point. Let $b_1,...,b_{s-1}$ be coprime positive integers greater than 1. For given s and N, the s-dimensional Hammersley set of size N is defined by [14]

$$x(n) = (g_{b_1}(n), \ldots, g_{b_{s-1}}(n), \frac{n}{N})$$

for $n = 1, ..., N$. Then the discrepancy of the set is obtained in [13]:

$$D_N^*(x(1), \ldots, x(N)) \leq C\frac{(\log N)^{s-1}}{N}$$

where C is a constant depending only on $b_1, ..., b_{s-1}$. The above estimation for the discrepancy of the Hammersley sequence means that this is a low

discrepancy sequence. The parameters of the algorithm are input: the integer I, the index of the element of the sequence 0 <= I; integer M, the spatial dimension, 1 <= M <= 100 and integer N, the "base" for the first component 1 <= N. Output is real R(M), the element of the sequence with index I.
The monographs of Sloan and Kachoyan [19] and Wang and Hickernell [21] 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]

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

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. We consider the following generating vector based on generalized Fibonacii numbers of corresponding dimensionality:

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

where

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

with initial conditions

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

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.

The Faure sequences [13] are a digital (0, s)-sequence over F_b with b denoting a prime (original case) or a prime power (general case) greater or equal to s. The s infinite generator matrices $\boldsymbol{C}^{(1)},\ldots,\boldsymbol{C}^{(s)}$ over $\mathbf{F}_b$ are defined by $\boldsymbol{C}^{(i)} = (c_{\mathrm{jr}}{}^{(i)})_{j,\,\mathrm{r}\geq 0}$

$$\text{with} \quad c_{\mathrm{jr}}{}^{(i)} = \binom{r}{j} a_i{}^{r-j},$$

where $\alpha_1,\ldots, \alpha_s$ denote s distinct elements from $\mathbf{F}_b$ and the conventions $\alpha^0 = 1$ for all $\alpha \in \mathbf{F}_b$ and $\binom{r}{j} = 0, \; j > r.$

For $\alpha = 1$, the resulting matrix is the infinite Pascal matrix modulo the characteristic of $\mathbf{F}_b$; for $\alpha = 0$, it is the infinite identity matrix. If $s = 1$ and $\alpha_1 = 0$, the resulting (0, 1)-sequence is identical to the van der Corput sequence in the same base [20]. The algorithm for the Faure sequence follows the method of Henri Faure in [12] for computing quasi-random numbers. It is a merging and adaptation of the routines INFAUR and GOFAUR from ACM TOMS 647. We use of persistent variables to improve the MATLAB implementation. The parameters of the Faure algorithm are described below. The input is an integer DIM_NUM, the spatial dimension, which should be at least 2. The other parameter is integer SEED, which is the seed, that indicates the index of the element of the sequence to be calculated. If SEED is negative, it is effectively replaced by a more suitable value. The output is a real QUASI(DIM_NUM), the next quasi-random vector. For the output the appropriate value of SEED have to be used on the next call, if the next element of the sequence is desired [4].

Numerical example and results

We will be interested in the following integrals that have widely used in Bayesian statistics:

$$\int_{\Omega} e^{-Nf(x)}\phi(x)dx,$$

where $f(x)$ and $\varphi(x)$ are s-dimensional polynomials and N is a natural number. This integrals are investigated by Shaowei Lin in [14]. We will test the performance of the Faure sequence (FAUR), the Hammersley sequence (HAM) and a particular lattice rule with generating vector, based on the generalized Fibonacci numbers of the corresponding dimensionality (FIBO). We will consider the following 7 and 15 dimensional integrals (Example 1 and Example 2 respectively):

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

$$\int_{[0,1]^{15}} \exp\left(-\prod_{i=1}^{15} x_i\right) \mathbf{dx} \approx 0.999.$$

Table 1: The relative error for 7 dimensional integral

N	FAUR	FIBO	HAM
100	9.81e-2	9.26e-2	1.64e-1
1000	1.03e-3	2.21e-2	6.59e-2
10000	3.34e-4	1.59e-3	6.45e-2
100000	3.40e-4	9.14e-5	8.17e-3

Table 2: The computational time for 7 dimensional integral

time in seconds	FAUR	FIBO	HAM
0.1	8.17e-2	2.17e-2	3.19e-1
1	2.26e-3	9.32e-3	8.12e-2
10	2.34e-4	2.47e-3	6.45e-2
60	1.10e-4	1.15e-4	8.69e-3

Table 3: The relative error for 15 dimensional integral

N	FAUR	FIBO	HAM
100	7.45e-3	3.04e-5	6.44e-4
1000	7.86e-4	3.04e-5	1.95e-4
10000	1.26e-4	1.23e-5	7.27e-5
100000	6.43e-5	7.99e-6	7.51e-5

Table 4: The computational time for 15 dimensional integral

time in seconds	FAUR	FIBO	HAM
0.1	5.56e-3	1.22e-4	5.94e-2
1	6.58e-4	7.80e-5	1.24e-3
10	1.06e-4	7.60e-5	9.46e-4
60	7.17e-5	1.38e-5	6.65e-5

In the Table 1 and 2 are presented the relative error for the 7 and 15 dimensional integrals with Fibonacci lattice sequence (FIBO), Faurre low discrepancy sequnce (FAUR) and Hammersley quasi-random sequence (HAM) for a fixed number of points. In Table 2 and 4 are presented the relative errors for 7 and 15 dimensional integrals with FAUR, FIBO and HAM for a fixed computational time which is a measure of the computational complexity. Obviously FIBO has the lowest computational complexity and is the fastest algorithm, while HAM and FAUR are slower, because they need an additional time for generating the corresponding low discrepancy sequences. As can been seen from the results for 7 dimensional integral, the low discrepancy sequence of Hammersley produces the worst results. It is interesting to see that Faure

sequence gives lowest relative errors for a given number of realizations of the random variable - see Table 1, but the Fibonacci lattice sequence has the advantage for a preliminary given time in seconds - see Table 2. For 15 dimensional integral Fibonacci gives the best results for a given number of realizations and fixed computational time - see Table 3, while Hammersley sequence and Faure sequence have very similar behavior - see Table 4. So we can conclude that all stochastic algorithms under consideration are efficient tool for evaluation of multidimensional integrals related to Bayesian models in migration forecasting. This is the first time a particular 1-rank lattice rule based on Fibonacci generating vector is compared with Hammersley and Faure quasi-random sequences.

Conclusion

In this paper we analyze the performance of different quasi-Monte Carlo methods for multidimensional integrals related to Bayesian based models in improving the international migration. Stochastic methods under consideration are an efficient way to solve problems in forecasting international migration based on Bayesian statistics. A fundamental problem in Bayesian statistics is the accurate evaluation of the presented multidimensional integrals. 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 Academy of Sciences through the „Program for Career Development of Young Scientists", Grant DFNP-17-88/28.07.2017, Project „Efficient Numerical Methods with an Improved Rate of Convergence for Applied Computational Problems", by the Bulgarian National Fund of Science under Project DN 12/5-2017, "Efficient Stochastic Methods and Algorithms for Large-Scale Problems", by the Bulgarian National Fund of Science under Project DN 12/4-2017, "Advanced Analytical and Numerical Methods for Nonlinear Differential Equations with Applications in Finance and Environmental Pollution".

References

[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]. Dzhurov V. Application of Summarized Functions for Information Source Protection, Journal Scientific and Applied Research, 2013, No 3, pp. 51-56.

[10]. Dzhurov V., Kostova M., Dzhurov K. Application of Probability Neural Networks for Classification of Explosives with Blasting Action. Mathematics in Industry, Cambridge Scholars Publishing, 2014, 254-285.

[11]. Dzhurov V., Slavova A., Kostova M., Tsakoumis A., Mladenov V.. Processing of Radioholographic Image with CNN Neural Network. WSEAS Transactions on Signal Processing, 2005, No 1, pp. 67-72.

[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]. Rogers, A. (1966). A Markovian Policy Model of Interregional Migration. *Papers of the Regional Science Assoc.*, 17, 205–224.

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

[20]. 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.

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

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

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