Journal of Basic and Applied Engineering Research
p-ISSN: 2350-0077; e-ISSN: 2350-0255; Volume 5, Issue 5; July-September, 2018, pp. 391-394

http://www.krishisanskriti.org/Publication.html

Quasi Monte Carlo with Digital Nets and Sequences

Mahesh S. Naik

Research Scholar, Shri Jagdishprasad Jhabarmal Tibrewala University, Chudela, Rajasthan, India
E-mail: nike.mahesh@gmail.com

Abstract—*Quasi-Monte Carlo methods of numerical integration use low discrepancy sequence to approximate the integral and achieves the result faster as compared to Monte Carlo methods which uses pseudo random numbers. The paper describes digital nets and sequences, one of the major low discrepancy sequence construction given by Niederreiter. The Paper reviews various aspects of this technique and focuses randomizing the sequences produced by this techniques.*

1. INTRODUCTION

The Monte Carlo method is a statistical method of evaluation of mathematical functions using random samples. Monte Carlo methods need random numbers for computation. There is always some error possible, but larger the number of random samples taken, more accurate is the result.

In mathematical form, the Monte Carlo method is finding the definite integral of a function by choosing a large size of independent variable samples at random from a region, taking average of the resulting dependent-variable values, and then dividing by the size of the region over which the random samples were chosen. This is different from the traditional method of approximating a definite integral, in which independent-variable samples are selected at equally-spaced points within an interval or region.

2. QUASI-MONTE CARLO METHODS

Quasi-Monte Carlo simulation is the traditional Monte Carlo simulation but using quasi-random sequences instead of random numbers.

The quasi-random sequences, also known as low-discrepancy sequences, improve the performance of Monte Carlo simulations, by decreasing the computational times and increasing the accuracy.

Quasi-random sequences are more evenly scattered throughout the region over which the Monte Carlo integral is calculated, uniformity of quasi random sequences improves the accuracy of the integral evaluation. Consider the integral of some function f(x) being evaluated using simulation. The idea is to use random points for the numerical evaluation of an integral, using random points to determine the area under the function.

The integral of the function f(x) is area under the curve, which is approximately the total area times the fraction of points that fall under the curve of f(x). This method for evaluation of an integral is useful only for the multi-dimensional case and complicated functions. The integral evaluation is better if the points are more uniformly scattered in the entire area.

The Monte Carlo simulation can be viewed as a problem of integral evaluation. To calculate an expected value we have to evaluate an integral (or a summation for discrete probability density case).

The Monte Carlo method solves multidimensional integrals, the expression for the Monte Carlo approximation for the multidimensional integral over the unit hypercube is given by:

$$\int_0^1 \cdots \int_0^1 g(x_1, \ldots, x_n)\, dx_1 \ldots, dx_n \approx \frac{1}{N} \sum_{j=1}^{N} g(\Gamma_j)$$

This estimate is an approximation of an integral over the n-dimensional unit cube, obtained by averaging the values of the integrand at independent random points uniformly distributed in the cube.

The pseudo-random sequence of numbers looks like random numbers. However, pseudo random numbers are generated with deterministic algorithm like the linear congruential random generator. the implementation of these pseudo random sequences are, in general, of volatile type such that the seed (initial value of a sequence) depends of an external feeder like the computer clock. For example, in Excel the function Rand() returns a different number from U[0, 1] .

By running the simulations, it can be verified that in most cases the simulation error with traditional Monte Carlo (pseudo random numbers) is higher or much higher than the error with quasi-Monte Carlo simulation.

Even pseudo-random numbers from a reliable random generator, which have "the same relevant statistical properties as a sequence of random numbers" (Ripley, 1987), in many situations exhibit a very slow rate of convergence which is the main problem of a Monte Carlo simulation. The main problem in finance and in many other fields is the evaluation of an integral of a function, e.g., the options payoff function with

variable(s) assuming values that obeys some probability density function(s). In such cases it is possible to skip some statistical properties (like independence), if not required in the practical problem. In this types of problems the idea of low discrepancy sequences is applicable.

The quasi-random sequences have low discrepancy property that is a measure of uniformity for the distribution of the points. This is very useful to represent a uniform distribution .for the multi-dimensional case, this is a measure of no large gaps and no clustering of points in the d-dimensional hypercube. Following is the expression for discrepancy of a sequence,

$$D_N^*(x^1, x^2, \ldots, x^N) = \sup_{x \in I^S} |S_N(G,) - N x_1 x_2, \cdots, x_S|$$

The following table illustrates the statistical properties of quasi-random sequence (used a van der Corput sequence in base 2) compared with the (theoretical) Uniform [0, 1] distribution, and with two typical pseudo-random sequences (generated with Excel).

Statistical Properties of Quasi-Random x Uniform x Pseudo Random
N = 1,000 simulations

	Quasi-Random (b2)	Uniform [0, 1]	Pseudo-Random (2 series)	
Minimum	9.76563e-4	0	1.81917e-5	3.27922e-4
Maximum	0.998047	1	0.999004	0.999982
Mean	0.498866	0.5	0.492246	0.512323
Median	0.498047	0.5	0.484037	0.521946
Standard Deviation	0.288661	0.288675	0.2851	0.291302
Variance	0.083325	0.083333	0.081282	0.084857
Skewness	2.042816e-3	0	0.039117	-0.077605
Kurtosis	1.794803	1.8	1.833614	1.773279

Hence, quasi-random sequence presents a better performance than typical pseudo-random sequences for all four probabilistic moments, indicating that quasi-random sequence is more representative of U[0, 1] than pseudo-random numbers. The quasi-random numbers were developed by number-theoreticians.

Quasi-Monte Carlo simulation originates from the discipline of Number Theory. Concepts for the measure of dispersion like discrepancy replacing the concept of variance, the use of tools like change of number base , the use of properties of prime numbers , the use of the coefficients of primitive polynomials, and so on, comprise the tools of number theory.

There are some problems when using Quasi Monte Carlo:

- First, quasi-Monte Carlo methods are valid for integration problems, but may not be directly applicable to simulations.
- Also the improved accuracy of quasi-Monte Carlo methods is generally lost for problems of high dimension or problems in which the integrand is not smooth.

2.1 Quasi-random sequences (Low Discrepancy Sequences)

Quasi-random numbers are also called low discrepancy sequences. The discrepancy of a sequence is a measure of its uniformity and is defined as follows:

Given a set of points $x^1, x^2, \ldots, x^N \in I^S$ and a subset $G \subset I^S$, define the counting function $S_N(G)$ as the number of points $x_i \in G$. For each $x=(x_1, x_2, \ldots, x_s) \in I^S$, let G_x be the rectangular s -dimensional region $G_x = [0, x_1) \times [0, x_2) \times L \times [0, x_S)$ with volume $x_1 x_2 \ldots x_N$. Then the discrepancy of the points x^1, x^2 ,....., x^N is given by:

The discrepancy is therefore computed by comparing the actual number of sample points in a given volume of multidimensional space with the number of sample points that should be there assuming a uniform distribution.

It can be shown that the discrepancy of the first N terms of quasi-random sequence has the form:

$$D_N^*(x^1, x^2, \ldots, x^N) \le C_S (\log N)^S + O((\log N)^{S-1}), \text{ for all } N \ge 2$$

The principal aim in the construction of low-discrepancy sequences is thus to find sequences in which the constant C_S is as small as possible. The concept of low-discrepancy is associated with the property that the successive numbers are added in a position as away as possible from the others numbers that is, avoiding clustering. The sequence is constructed based in the schema that each point is repelled from the others. So, if the idea for the points is maximally avoiding of each other, the job for the numbers generated sequentially is to fill in the larger "gaps" between the previous numbers of the sequence.

Various sequences have been constructed to achieve this goal. Here we explain two of them,

2.2. Vander Corput sequence

This is the basic low discrepancy sequence

Steps to obtain van der corput sequence

1. The nth point of the sequence is expressed in base p, wher p is a prime number.
   $N=a_0p_0+a_1p_1+\ldots\ldots a_Np_N$, whre N is the smallest integer for which $a_i=0$.
2. Thus the number n in base p is $a_0a_1a_2\ldots a_N$.
3. This number is reflected around the decimal point which gives $0.a_N\ldots..a_2a_1$
4. Now $x_n=\frac{a_N}{p}+\frac{a_{N-1}}{p^2}+\ldots\ldots\ldots\ldots+\frac{a_1}{p^{N-1}}+\frac{a_0}{p^N}$

For Example:

For base 3 the 23rd point is given by

$23=2x3^2+1x3^1+2x3^0$

Thus the number equivalent in base 3 is 212

Its reflection around the decimal point 0.212

Thus $x_{23}=2 \text{ x } \frac{1}{3^1} +1 \text{ x}\frac{1}{3^2}+2 \text{ x}\frac{1}{3^3}=\frac{23}{27}$

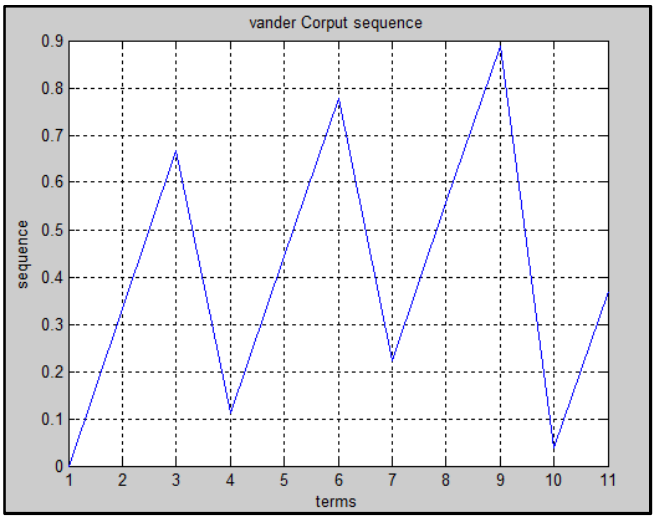
The sequence spreads out after 2n-1 points

Cycle length differs with change in base.

For base 3

No.	Expressed in terms of base	Sequence
0	0	0
1	1	1/3
2	2	2/3
3	10	1/9
4	11	4/9
5	12	7/9
6	20	2/9

First ten terms of Vander Corput Sequence with base 3



2.3 Faure Sequences

The Faure sequence uses only one base for all dimensions & it uses a permutation of the vector elements for each dimension.

The base of a Faure sequence is the smallest prime number that is larger than or equal to the number of dimensions in the problem, or equal 2 for one dimensional problem.

We can generate a number in the interval[0,1) by reflecting the expansion in base r about the decimal point. Let us see through an illustration.

We can write 7 in base 3,

$7=2(3^1)+1(3^0)=21$

Thus $_3(7)=1/3+2/9=5/9$

Similarly $_3(8)=2/3+2/9=8/9$

The first 9 numbers in this sequence are

9/27,18/27,3/27,12/27,21/27,6/27,15/27,24/27,1/27

The general expression for n in terms of base r is

n=

only a finite no. Of $a_j(n)$ will be non zero.

Thus the corresponding random no as per this procedure is

$_r(n)=$

Let us see the procedure to obtain s dimensional fauresequence .

Let r be the smallest prime no. That is ≥2& ≥s

$a^k_j(n)={}^iC_j a_i^{k-1}(n) \bmod r,$

the next level of coefficients is obtained by multiplying an upper triangular matrix with elements

$$\begin{matrix} {}^0C_0 & {}^1C_0 & {}^2C_0 & {}^3C_0 \\ 0 & {}^1C_1 & {}^2C_1 & {}^3C_1 \\ 0 & 0 & {}^2C_2 & {}^3C_2 \\ 0 & 0 & 0 & {}^3C_3 \end{matrix}$$

The successive points in the Faure sequence is given by ,

$^k_r(n)=a^k_j(n)^{-j-1}$, 2≤k≤s

3. (T,M,S) NETS AND (T,S) SEQUENCE

(t,m, s)-nets and *(t, s)*-sequences were developed by Niedreitter with an idea of providing excellent deterministic sample points for quasi-Monte Carlo methods.The motivation behind development of (t,m,s) nets and (t,s) sequences was to all intervals should have fairly equal proportion of points of a sequence.

Consider $[0,1)^s$ a half open s dimensional cube, an elementary interval in $[0,1)^s$ in base b is of the form

$$\prod_{i=1}^{s}[\frac{a_i}{b^{d_i}}, \frac{a_i+1}{b^{d_i}})$$

Where each $d_i \geq 0$ also for each ai, $0 \leq a_i \leq b^{di}$.

Volume of this elementary sub interval is $b^{\Sigma d_i}$.

Let $s \geq 1, b \geq 2 \text{ and } 0 \leq t \leq m$

then (t, m, s)*net in base b is a set of* b^m*points in the unit half open cube* defined above.

A (t,s) sequence in base b is an infinite sequence x0,x1,x2,……. of points in $[0,1)^s$ with the property that for every m>t and k>=0 the set { $[x_i]$ / kb^m<i< $(k+1)b^m$} is (t,m,s) net in base b.

For applications of these , it is important to have a nested sequence of (t_i,m_i, s)-nets(with mi tending to infinity) in the same dimension s and with all t_i bounded above by a constant t.

This allows the user to improve estimates on an integral without losing the benefit of preliminary computations. Such a sequence of nets can be developed via (t, s)-sequences.

3.1 Scrambling

Quasi-Monte Carlo methods are deterministic in nature and therefore error estimationas in statistical Monte Carlo methods is not possible. Randomization can be introduced in a quasi-Monte Carlo method by scramblingi.e. by randomizing thedeterministic sample points used in the method. A simple scrambling scheme like *random shifts can be applied to any intervak* I^s.

Let $\mathbf{x}_1,\ldots,\mathbf{x}_N \in I^s$ be arbitrary then

$\mathbf{y}_n = \{\mathbf{x}_n + \mathbf{r}\}$

ris a random vector uniformly distributed over I^s and {·} denotes reductionmodulo 1 in each coordinate of a point in R^s.

Random permutation of digits and Linear scrambling are other methods of Scrambling to introduce randomization.

4. CONCLUSION

Monte Carlo method play an important role in obtaining approximation to a problem that can't be solved analytically. The convergence is improved further of these methods by introducing Quasi Monte Carlo Methods which use deterministic points for the evaluation of the integral. Several constructed sequences like Vander Corput, halton, Sobol give improved results while running simulation problems in diverse fields compared to Monte Carlo Methods. One of the disadvantages occurs with increase in dimension. The (t,m,s) nets and (t,s) help in achieving the goal of uniformity in much better way compared to other low discrepancy sequences. They are constructed keeping in mind the fair filling of space which is the requirement of uniformity. The only problem of error estimation in this case can be achieved by efficient ways of introducing randomization.

REFERENCES

Niederreiter, H. Random Number Generation and Quasi-Monte Carlo Methods, SIAM, 1992.

[1]. Faure, H. Discrepance de suites associees a un systeme de numeration (en dimension s). Acta Arith. 41, 337-351 1982.

[2]. "Point Sets and Sequences with Small Discrepancy:' Monatsh. Math., 104 (1987), 273-337.

[3]. Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia, PA, 1992.

[4]. P. P. Boyle.Options: A Monte Carlo Approach.Journal of Financial Economics,4:323-338,1977.

[5]. C.Joy,P.P.Boyle and K.Seng Tan.Quasi-monte carlo methods in numericalfinance.Management Science,42:926-938,1996.

[6]. L.C.G. Rogers,and Z. Shi. The value of an Asian option. Journal of AppliedProbability, 32:1077–88,1995.

[7]. R.Y.Rubinstein, and D.P.Kroese.Simulation and the Monte Carlo Method.Wiley-Interscience,2007.

[8]. P.Wilmott,S.Howison, and J. Dewynne. The Mathematics of Financial

[9]. Derivatives.Cambridge University Press,1996.