Low Discrepancy Sequences

Anne GILLE-GENEST

March 1, 2012

Premia 14

Contents

Implementation of the Low Discrepancy Sequences
QMC simulation

Quasi Monte Carlo simulation consists in approximating the integral $\int_{[0,1]^d} f(u)du$ by $\frac{1}{N}\sum_{i=1}^{N} f(u_i)$ where $\{\xi_i\}$ are quasi-random numbers, that means they are generated from low-discrepancy sequences. As we already have explained it, such sequences neither are random nor pseudo-random but deterministic and successive values are not independent. However they satisfy good properties of equidistribution on $[0,1]^d$ and we have that $\frac{1}{N}\sum_{i=1}^{N} f(\xi_i) \to \int_{[0,1]^d} f(u)du$.

In the following sections we describe some low discrepancy sequences. We explain their construction and discuss some of their properties, especially on their discrepancy.
General references about the Quasi-Monte Carlo simulation are [2], [7], [8], [6] or [4].
The implementation of the sequences are described in the implemented part.

1 Tore-SQRT sequences

They are d-dimensional sequences, obtained by considering the multiples of suitable irrational numbers modulo 1.
• **Tore sequence**
It is defined by :

$$\xi_n = (\{n.\alpha\}) = (\{n.\alpha_1\}, \ldots, \{n.\alpha_d\})$$

where $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ such that $(1, \alpha_1, \ldots, \alpha_d)$ are linearly independent on $\mathbb{Q}$.
$\{x\} = x - [x]$ denotes the fractional part of x.

• **SQRT sequence**
It is a particular case of the Tore sequence with

$$\alpha = (\sqrt{p_1}, \ldots, \sqrt{p_d})$$

where $(p_1, \ldots, p_d)$ are the first d prime numbers.

If $\alpha_1, \ldots, \alpha_d$ are algebric, then the discrepancy satisfies:

$$D_n^*(\xi) = O\left(\frac{1}{n^{1-\varepsilon}}\right), \forall \varepsilon > 0$$

Click there to reach the implemented part: implementation.

2 Van der Corput and Halton sequences

2.1 Van der Corput sequence

This is a one-dimensional sequence defined by the *radical-inverse function* φ_p in base p:

$$\varphi_p(n) = \sum_{i=0}^{R(n)} \frac{a_i}{p^{i+1}}$$

where the coefficients a_i are given by the digit expansion in base p of n :

$$n = \sum_{i=0}^{R(n)} a_i p^i$$

$R(n)$ denotes the maximum index for which $a_{R(n)}$ is not equals to 0. Its value depends on n and p by the relation $p^{R(n)} \leq n < p^{R(n)+1}$, that is:

$$R(n) = \left[\frac{\log n}{\log p}\right]$$

Discrepancy of the Van der Corput sequence satisfies the following majoration:

$$D_n^*(\xi) \leq \frac{1}{n}\frac{p\log(pn)}{\log(p)} = O\left(\frac{\log n}{n}\right)$$

Remarks: (ref Article of Alan Jung and Silvio Galanti)
- For $n < p$, there is only one positive coefficient a in the decomposition in base p, that is $a_0 = n$. Thus $\varphi_p(n) = \frac{n}{p}$ and the sequence is increasing.
- There are cycles of length p in this sequence. Each subsequence of length p (indices kp to $(k+1)p-1$) is increasing in magnitude proportionnaly to power of $1/p$, and covers uniformly the interval $[0, 1)$. Consequences of this property will be studied for multidimensional sequences (especially Halton sequence).

2.2 Halton sequence

The Halton sequence is a d-dimensional generalization of the Van Der Corput sequence. Let $(p_1, \ldots, p_d)$ be the d first prime numbers, then ξ_n is defined by:

$$\xi_n = (\varphi_{p_1}(n), \ldots, \varphi_{p_d}(n))$$

where $\varphi_{p_i}(n)$ is the Van der Corput sequence in base p_i.
The Halton sequence satisfies :

$$D_n^*(\xi) \leq \frac{1}{n}\prod_{i=1}^{d}\frac{p_i \log(p_i n)}{\log(p_i)} = O\left(\frac{\log^d(n)}{n}\right)$$

with a constant $C^d = \prod_1^d \frac{p_k - 1}{2\log p_k}$.
This constant grows to infinity super-exponentially with dimension.
Click there to reach the implemented part: implementation.

2.3 Permuted (Generalized) Halton sequence

Orthogonal projections of points from the Halton sequence show non uniform distribution for some dimensions (see Morokoff and Caflish [6], Jung ??? or Bratley and Fox ????). This non-uniformity is due to cycles of length p_i for each one-dimensional sequence.
To break correlations between the inverse radical functions of different dimensions, we realize permutations of coefficients a_i.
We consider $(\Pi_{p_i}, 1 \leq i \leq d)$ d permutations over $\{0, \ldots, p_{i-1}\}$ such that $\Pi_{p_i}(0) = 0$.
Each term of the *permuted Halton sequence* is defined by:

$$S_{p_i}(n) = \frac{\Pi_{p_i}(a_0)}{p_i} + \cdots + \frac{\Pi_{p_i}(a_{R(n)})}{p_i^{R(n)+1}}$$

with $n = \sum_{i=0}^{R(n)} a_i p^i$.
The global sequence is given by :

$$\xi_n = (S_{p_1}(n), \ldots, S_{p_d}(n))$$

There is no optimal choice for the permutations. We present 3 approaches to modify the Halton sequence
• An algorithm was suggested by Braaten and Weller [1] for $d \leq 16$ with a possible extension to a larger d (however with significant computation).

• Reverse-Radix Algorithm
An other algorithm (see Kocis and Whiten [5]) consists in reversing the binary digits of integers, expressed using a fixed number of base 2 digits and removing any values that are too large.
This algorithm can be applied for very large values of dimensions.

• Halton Sequence Leaped:
This other variant for the Halton sequence consists in using only every Lth Halton number subject to the condition that L is a prime different from all bases $p_1, \ldots, p_d$ (see Kocis and Whiten [5]).

3 Faure sequence

This is a d-dimensional sequence.
The Faure sequence is a permutation of the Halton sequence, but it uses the same base r for each dimension. We choose r as the smallest odd prime integer such that $r \geq d$.
Note that the k-th dimension of a d-dimensional Faure sequence is different from the k-th dimension of a d'-dimensional Faure sequence as soon as the base r is different.
With usual notations, a_i are the coefficients of the r-adic decomposition of n

$$n = \sum_{i=0}^{R(n)} a_i r^i$$

We consider the following transformation T :

$$T : x = \sum_{k=0}^{R(n)} \frac{a_k}{r^{k+1}} \mapsto T(x) = \sum_{k=0}^{R(n)} \frac{b_k}{r^{k+1}}$$

with $b_k = \sum_{i=k}^{R(n)} C_i^k a_i \mod r$ and C_i^k denote binomial coefficients.
The coefficients b_k are a permutation of the a_k.
Precision ... and reference ?????

The *Faure sequence* is defined by using successive transformations T^k:

$$\xi_n = \Big(\varphi_r(n), T(\varphi_r(n)), \ldots, T^{d-1}(\varphi_r(n))\Big)$$

where φ_r is the Van der Corput sequence in base r.

The discrepancy of the sequence satisfies :

$$D_n^*(\xi) \leq C^d \frac{\log^d(n)}{n}$$

where C^d is a constant dependent on d and r : $C = \frac{1}{d!}(\frac{r-1}{2\log r})^d$.
The constant C^d tends to 0 with dimension.

The Faure sequence exhibits cycles of length r but cycles are not composed of increasing terms, except for the first dimension. For the same dimension, the Faure sequence has generally a smaller base than the Halton one, thus cycles are smaller too. Because we use the smallest prime number greater than the dimension d and not the d-th prime number.

Click there to reach the implemented part: implementation.

4 Generalized Faure sequence

This is a d-dimensional sequence. Let r be the smallest odd prime integer, such that $r \geq d$.
The digit expansion of n in base r is given by $n = \sum_{i=0}^{R(n)} a_i(n) r^i$.
The *Generalized Faure sequence* is defined by :

$$\xi_n = \left(\sum_{k=0}^{R(n)} \frac{\xi_{n,k}^{(1)}}{r^{k+1}}, \dots, \sum_{k=0}^{R(n)} \frac{\xi_{n,k}^{(d)}}{r^{k+1}} \right)$$

with

$$\xi_{n,k}^{(j)} = \sum_{s=0}^{R(n)} c_{k,s}^{(j)} a_s(n), \quad j \leq d, k \leq R(n)$$

$c^{(j)} = (c_{k,s}^{(j)})_{0 \leq k \leq R(n), 0 \leq s \leq R(n)}$ and $c^{(j)} = A^{(j)} P^{j-1}$ where $A^{(j)}$ is a lower triangular inversible matrix such that $(a_{i,l}) \in \mathbb{F}_r$ and $P = (C_s^k)$ for $k \leq R(n), s \leq R(n)$ is built with the binomial coefficients.

The discrepancy of the sequence satisfies:

$$D_n^*(\xi) \leq C(d,r) \frac{\log^d(n)}{n}$$

where $C(d,r) \approx \frac{1}{d!} (\frac{r}{2 \log r})^d$.

Click there to reach the implemented part: implementation.

5 Nets and (t,s)-sequences

(t,s)-sequences are a group of sequences with a very regular distribution behaviour. Their points are placed into certain equally sized volumes of the unit cube for sequences of a fixed length. Chapter 4 of Niederreiter [7] well describes theoretical aspects for such sequences. We just summarize in this section some definitions and properties of those sequences.

Definitions

- An *elementary interval* $E \in I^d$ is defined as $E = \prod_{i=1}^{d}[a_i b^{-d_i}, (a_i + 1)b^{-d_i}]$ where $a_i, d_i > 0$ are integers satisfying $0 \leq a_i \leq b^{d_i}$ for $1 \leq i \leq d$.
- Let $0 \leq t \leq m$ be integers. A (t, m, s)-*net* in base b is a point set P of b^m points in I^s such that the number of points in E is equal to b^t for every elementary interval E in base b with $\Pi(E) = b^{t-m}$.
- Let $t \geq 0$ be an integer. A sequence $x_0, x_1, \ldots$ of points in I^s is a (t, s)-*sequence* in base b if, for all integers $k \geq 0$ and $m > t$, the point set constituting of the x_n with $kb^m \leq n \leq (k+1)b^m$ is a (t, m, s)-net in base b.

Properties:

- Any (t, m, s)-net in base b is also a (u, m, s)-net in base b for integers $t \leq u \leq m$.
  The same property holds for (t, s)-sequences.
  Then smaller values of t mean stronger regularity properties.
- The discrepancy of a (t, m, s)-net P in base b with $m > 0$ satisfies:

  $$ND_N(P) \leq B(s, b)b^t(\log N)^{s-1} + O(b^t(\log N)^{s-2})$$

  where

  $$B(s,b) = \begin{cases} (\frac{b-1}{2\log b})^{s-1} \text{ if either } s = 2 \text{ or } b = 2, s = 3, 4 \\ \frac{1}{(s-1)!}(\frac{\lfloor b/2 \rfloor}{\log b})^{s-1} \text{ otherwise} \end{cases}$$

- The discrepancy of the first N terms of a (t, s)-sequence P in base b satisfies:

  $$ND_N(P) \leq C(s, b)b^t(\log N)^s + O(b^t(\log N)^{s-1})$$

  where

  $$C(s,b) = \begin{cases} \frac{1}{s}(\frac{b-1}{2\log b})^{s} \text{ if either } s = 2 \text{ or } b = 2, s = 3, 4 \\ \frac{1}{s!}\frac{b-1}{2\lfloor b/2 \rfloor}(\frac{\lfloor b/2 \rfloor}{\log b})^{s} \text{ otherwise} \end{cases}$$

- For $m \geq 2$, a $(0, m, s)$-net in base b can only exist if $s \leq b + 1$.
  A $(0, s)$-sequence in base b can only exists if $s \leq b$.

Examples:

- The Van der Corput sequence is a $(0, 1)$ sequence in base b. In fact, if we consider the b^m points x_n with $kb^m \leq n < (k+1)b^m$ $(k \geq 0, m \geq 1)$, every b-adic interval $[ab^{-m}, (a+1)b^{-m}]$ contains exactly one point x_n.
- The s-dimensional Sobol sequence is a (τ, s)-sequence in base 2, where $\tau = \sum_{i=1}^{s} deg(P_i) - s$. It is called a LP_τ-sequence. Sobol sequence is described in the next point.
- The s-dimensional Faure sequence in base r is a $(0, s)$-sequence where r is the smallest prime integer greater or equal than s.

6 Sobol sequence

The Sobol sequence is a d-dimensional sequence in base 2 and it is a (τ, d)-sequence. It is one of the most used sequences for Quasi-Monte Carlo simulation. It was first developped by Sobol [3] and it has been proved to have some additional uniformity property under some initialization conditions (see [9]). Its construction is based on primitive polynomials in the field $\mathbb{Z}_2$ and XOR operations.

Each dimension is a permutation of the Halton sequence with base 2 whenever $N = 2^d$. These permutations are generated from irreductible polynomials in $\mathbb{Z}_2$. But they allow for certain correlations to develop, then they can produce regions where no points fall until N becomes very large.

The *Sobol sequence* is defined by:

$$\xi_n = \left(a_0 V_0^{(1)} \oplus \cdots \oplus a_{R(n)} V_{R(n)}^{(1)}; \ldots ; a_0 V_0^{(d)} \oplus \cdots \oplus a_{R(n)} V_{R(n)}^{(d)} \right)$$

where the $V_i^{(j)}$ are direction numbers (expressed as binary fraction) obtained from d different primitive polynomials and a_i denote the coefficients of the digit expansion of n in base $b = 2$, given by: $n = \sum_{i=0}^{R(n)} a_i 2^i$.
$\oplus$ represents the bitwise exclusive OR operator (XOR). For explanation about XOR operation or primitive polynomials, we refer the reader to the Numerical Recipes in C ?????.

To implement this sequence, we use an other expression for ξ_n depending only on the previous point and one direction number. This principle is detailed in the implemented part and is due to Antonov and Saleev ?????.

The discrepancy of the sequence satisfies:

$$D_n^*(\xi) \leq C_d \frac{(\log n)^d}{n} + O\left(\frac{(\log n)^{d+1}}{n}\right)$$

where $C_d = \frac{2^{t(d)}}{d!(\log 2)^d}$ grows superexponentially with dimension,
and for $K > 0, K\frac{d \log d}{\log \log d} \leq t(d) \leq \frac{d \log d}{\log 2} + O(d \log \log d)$. $t(d)$ grows superlinearly with dimension.

• **Definition of the constants** V**:**
- For each $j \leq d$ we first choose a primitive polynomial $P(j)$ with degree $s(j)$:

$$P(j) = x^{s(j)} + b_1 x^{s(j)-1} + \cdots + b_{s(j)-1} x + 1$$

and we select $s(j)$ odd integers $c_i^{(j)}$ such that

$$c_i^{(j)} < 2^{i+1}, \quad 0 \leq i < s(j)$$

The choice for constants $c_i^{(j)}$ is not a easy step. Sobol' article ????? gives some explanations about this problem.
- Once we have chosen $P(j)$ and the $c_i^{(j)}$ for $i < s(j)$, we use the coefficients b_i through the recurrence relation :

$$c_i^{(j)} = 2b_1 c_{i-1}^{(j)} \oplus 2^2 b_2 c_{i-2}^{(j)} \oplus 2^{s(j)-1} b_{s(j)-1} c_{i-s(j)}^{(j)} \oplus 2^{s(j)} c_{i-s(j)}^{(j)} \oplus c_{i-s(j)}^{(j)}$$

to determine the $c_i^{(j)}$ for $i \geq s(j)$.
- Finally we calculate V by:

$$V_i^{(j)} = \frac{c_i^{(j)}}{2^{i+1}}$$

• **Uniformity property**: An additional uniformity property of the sequence is called by Sobol the *property A*.

- We define a *binary segment* of length 2^s as a set of points P_i whose subscripts satisfy the inequality $l2^s \leq i < (l+1)2^s$ where $l = 0, 1, \ldots$.
We divide up the s-dimensional unit cube I^s by the planes $x_k = \frac{1}{2}$ into 2^s multidimensional small cubes, which represent binary parallelepipeds.

- **Property A**: If in any binary segment of length 2^s of the sequence $P_0, \ldots, P_i, \ldots$, all the points belong to different small cubes, then we say that the sequence satisfies property A.

Sobol [9] proved a sufficient and necessary condition on the direction numbers so that the property A is verified. A table of good numerical values for V is given for a dimension $s \leq 16$.

- **Property A'**: The property A can be extended to the property A' defined as follows.
We divide up the s-dimensional unit cube I^s by the planes $x_k = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ into 2^{2s} multidimensional small cubes. If in any binary segment of the sequence $P_0, \ldots, P_i, \ldots$ of length 2^{2s}, all the points belong to different small cubes, then we say that the sequence possesses the property A'.

- Remark about the link between property A or A' and the dimension s: Note that the property A (resp. A') holds for subsequences of length 2^s (resp. 2^{2s}). In practice if s increases, it becomes difficult to verify the condition because we need to simulate at least 2^s (2^{2s}) points.

Click there to reach the implemented part: implementation.

7 Niederreiter sequence

The *Niederreiter sequence* is a s-dimensional (t, s)-sequence in base b whose theoretical aspects are described in Niederreiter [7]. It is defined as:

$$\xi_n = \left(\sum_{j=0}^{R(n)} \frac{y_{n,j}^{(1)}}{b^{j+1}}, \ldots, \sum_{j=0}^{R(n)} \frac{y_{n,j}^{(s)}}{b^{j+1}} \right)$$

with $n = \sum_{r=0}^{R(n)} a_r(n) b^r$ and

$$y_{n,j}^{(i)} = \sum_{r=0}^{R(n)} c_{j,r}^{(i)} a_r(n) \ \in \mathbb{F}_b$$

$C^{(i)} = (c_{j,r}^{(i)})$ is called the generator matrix of the i-th coordinate. An algorithm to compute the values is given in Niederreiter [7]. Inititialization of the $(c_{j,r}^{(i)})$ is done at the beginning of the simulation.

The discrepancy of the sequence satisfies:

$$D_n^*(\xi) = O\left(\frac{(\log n)^s}{n}\right)$$

Construction of the $c_{jr}^{(i)}$: (in the next version)
The method is based on the formal Laurent series.

Remark: If b is a prime power and s an arbitrary dimension such that $s \leq b$, we can choose $P_1, \ldots, P_s$ as the linear polynomials $P_i(x) = x - b_i$ where $b_q, \ldots, b_s$ are distinct elements of F_b. Then the Niederreiter sequence is a $(0, s)$-sequence in base b and we have for $1 \geq i \geq s$ and $j \geq 1$:
$c_{jr}^{(i)} = 0$ if $0 \leq r < j-1$
$c_{jr}^{(i)} = (r/j-1)b_i^{r-j+1}$ if $r \geq j-1$.

Click there to reach the implemented part: implementation.

8 General remarks on low discrepancy sequences

- Quasi-random numbers combine the advantage of a random sequence that points can be added incrementally, with the advantage of a lattice that there is no clumping of points.

- For large dimension s, the theoretical bound $(\log N)^s/N$ may only be meaningful for extremly large values of N. The bound in Koksma-Hlawka inequality gives no relevant information until a very large number of points is used.
  Low discrepancy sequences are very useful for low dimension. In high dimension s, a lattice can only be refined by increasing the number of points by a factor 2^s.

- Orthogonal projections: if a d-dimensional sequence is uniformly distributed in I^d, then two-dimensional sequences formed by pairing coordinates should also be uniformly distributed. The appearance of non-uniformity in these projections is an indication of potential problems in using a quasi-random sequence for integration. This problem is developped in Morokoff and Caflish [6]. We will see that procedures like scrambling permutation can be suggested to improve the uniformity property while preserving the discrepancy.

References

[1] E. BRAATEN and G. WELLER. An improved low-discrepancy sequence for multidimensional quasi-monte carlo integration. *Journal of Comput. Phys.*, (33):249–258, 1979. 4

[2] H.NIEDERREITER. Points sets ans sequences with small discrepancy. *Monatsh.Math*, 104:273–337, 1987. 2

[3] I.M.SOBOL. The distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Computational Math.and Math.Phys.*, 7(4):86–112, 1967. 8

[4] INRIA. *Probabilites numeriques. Chap 1: suites a discrepance faible et integration numerique.* 2

[5] L. KOCIS and W.J. WHITEN. Computational investigations of low discrpeancy sequences. *ACM Transactions on Mathematical Software*, 23(2):266–294, June 1997. 4, 5

[6] W.J. MOROKOFF and R.E. CAFLISH. Quasi-random sequences and their discrepancies. *SIAM, Journal of Scientific Computing*, 15(6):1251–1279, nov 1994. 2, 4, 11

[7] H. NIEDERREITER. Random number generation and quasi-monte carlo methods. *SIAM*, 1992. 2, 6, 10

[8] H. NIEDERREITER P.J.S. SHIUE. Monte carlo and quasi-monte carlo methods in scientific computing. *Lecture Notes in Statistics,Ed Springer*, 106, 1995. 2

[9] I.M. SOBOL. Uniformly distributed sequencs with an additional uniformity property. *USSR Comput. Maths. Math. Phys*, 16:236–242, 1976. 8, 10