Sobol sequence implementation

CD

08/05/2022

Contents

Source : p311 (pdf) of Monte Carlo Methods in Financial Engineering, Glasserman (2003)

In contrast to Halton or Faure sequence, Sobol are (t, d)-sequences in base 2 for all d with t values depending on d: they are permutation of Van den Corput sequence in base 2 only.

1 Principle

V a generator matrix of direction (binary) constructed as binary expansion of numbers $v_1, \ldots, v_r$. y is the state variable and x its output in $[0, 1]$.

Prop : V is upper triangular.

For a number k, the binary coefficients are denoted by $\boldsymbol{a}(k) = (a_0(k), \ldots, a_{r-1}(k))$ such that

$$k = a_0(k)2^0 + \cdots + a_{r-1}(k)2^{k-1}.$$

Recurrence is

$$\begin{pmatrix} y_1(k) \\ \vdots \\ y_r(k) \end{pmatrix} = V \begin{pmatrix} a_0(k) \\ \vdots \\ a_{r-1}(k) \end{pmatrix} \bmod 2, \tag{1}$$

$$x_k = y_1(k)/2 + \cdots + y_r(k)/(2^r).$$

Special case : V being the identity matrix corresponds to Van der Corput sequence. (1) rewrites as

$$\boldsymbol{y}(k) = a_0(k)v_1 \oplus a_1(k)v_2 \oplus \cdots \oplus a_{r-1}(k)v_r \oplus, \tag{2}$$

where $\oplus$ denotes the XOR operation.

For a d-dimensional sequence, we need d direction numbers $(v_j)_j$. Sobol's method relies on the use primitive polynomial of binary coefficients

$$P(x) = x^q + c_1x^{q-1} + \cdots + c_{q-1}x + 1 \tag{3}$$

which are irreducible and the smallest power dividing the polynomial P is $p = 2^q - 1$. the coefficients c_q, c_0 always equals 1.

1.1 Examples of direction number and primitive polynomials

Polynomials up to degree 5 are given in the following table table 5.2 of Glasserman (2003).

Degree p	Primitive polynomial $P(x)$	Binary coefficients $(c_q, \dots, c_0) = \boldsymbol{c}(k)$	Associated number $c_q 2^q + \cdots + c_0 2^0 = k$
0	1	1	1
1	$x+1$	$(1,1)$	$3 = 2^1 + 2^0$
2	x^2+x+1	$(1,1,1)$	$7 = 2^2 + 2^1 + 2^0$
3	x^3+x+1	$(1,0,1,1)$	$11 = 2^3 + 2^1 + 2^0$
3	x^3+x^2+x+1	$(1,1,0,1)$	$13 = 2^3 + 2^2 + 2^0$
4	x^4+x+1	$(1,0,0,1,1)$	$19 = 2^4 + 2^1 + 2^0$
4	$x^4+x^3+x^2+x+1$	$(1,1,1,1,1)$	$25 = 2^4 + 2^3 + 2^2 + 2^1 + 2^0$
5	x^5+x+1	$(1,0,0,0,1,1)$	$35 = 2^5 + 2^1 + 2^0$
5	$x^5+x^4+x^2+x+1$	$(1,1,0,1,1,1)$	$59 = 2^5 + 2^4 + 2^2 + 2^1 + 2^0$
5	$x^5+x^3+x^2+x+1$	$(1,0,1,1,1,1)$	$47 = 2^5 + 2^3 + 2^2 + 2^1 + 2^0$
5	$x^5+x^4+x^3+x^2+1$	$(1,1,1,1,0,1)$	$61 = 2^5 + 2^4 + 2^3 + 2^2 + 2^0$
5	$x^5+x^4+x^2+x+1$	$(1,1,0,1,1,1)$	$55 = 2^5 + 2^4 + 2^2 + 2 + 2^0$
5	x^5+x^3+x+1	$(1,0,1,0,0,1)$	$41 = 2^5 + 2^3 + 2^0$

Table 1: Primitive polynomial

Polynomial (3) defines a recurrence relation between m_j as

$$m_j = 2c_1 m_{j-1} \oplus 2^2 c_2 m_{j-2} \oplus \cdots \oplus 2^{q-1} c_{q-1} m_{j-q+1} \oplus 2^q m_{j-q} \oplus m_{j-q}. \tag{4}$$

The direction numbers are $v_j = m_j / 2^j$ using (4) and a set of initial values $m_1, \dots, m_q$.

1.2 Examples

Consider $k = 13 = 2^3 + 2^2 + 2^0$ so that the primitive polynomial is $x^3 + x^2 + 1$. (4) writes as

$$m_j = 2m_{j-1} \oplus 8m_{j-3} \oplus m_{j-3}.$$

Initializing with $m_1 = 1$, $m_2 = m_3 = 3$ leads to

$$m_4 = (2 \times 3) \oplus (8 \times 1) \oplus 1 = (1111)_2 = 15.$$

$$m_5 = (2 \times 15) \oplus (8 \times 3) \oplus 3 = (00101)_2 = 5.$$

R functions `sobol.directions.mj()` and `sobol.directions.vj()` compute integers m_j and direction numbers v_j.

```r
p13 <- int2bit(13)
m1<-1; m2<-m3<-3
#mj
sobol.directions.mj(c(m1,m2,m3), p13, 2, echo=FALSE, input="real", output="real")
```

```
## [1]  1  3  3 15  5
```

```r
#V matrix
head(sobol.directions.vj(c(m1,m2,m3), p13, 2, echo=FALSE, output="binary"))
```

```
##      v1 v2 v3 v4 v5
## [1,]  1  1  0  1  0
## [2,]  0  1  1  1  0
## [3,]  0  0  1  1  1
## [4,]  0  0  0  1  0
## [5,]  0  0  0  0  1
## [6,]  0  0  0  0  0
```

In this example, the generator is

$$\boldsymbol{V} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

This gives the following sequence for integers $k = 1, 2, 3, 31$. V produces a permutation of Van der Corput as it uses a primitive polynomial.

k	$\boldsymbol{a}(k)$	$\boldsymbol{V}\boldsymbol{a}(k) = \boldsymbol{y}(k)$	x_k
1	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	1/2
2	$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	3/4
3	$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	1/4
31	$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$	31/32

2 A faster implementation using Gray code

Antanov and Saleev point out Sobol's method simplifies we use the Gray code representation of k rather than the binary representation. The Gray code is such that exactly one bit changes from k to $k+1$ and is defined as

$$\boldsymbol{g}(k) = \boldsymbol{a}(k) \oplus \boldsymbol{a}(\lfloor k/2 \rfloor).$$

This is a shift in the string representation. For instance, the Gray code of 3 and 4 are

$$\boldsymbol{g}(3) = \boldsymbol{a}(3) \oplus \boldsymbol{a}(1) = (011)_2 \oplus (001)_2 = (010)_2,$$

$$\boldsymbol{g}(4) = \boldsymbol{a}(4) \oplus \boldsymbol{a}(2) = (100)_2 \oplus (010)_2 = (110)_2.$$

The Gray code representations of integers $0, \ldots, 2^r - 1$ are a permutation of the sequence of strings formed by the usual binary representations. So the asymptotic property of the Gray code $\boldsymbol{g}(k)$ remains the same as the one of the binary representation $\boldsymbol{a}(k)$.

	Binary/Gray representation						
integer k	1	2	3	4	5	6	7
binary $\boldsymbol{a}(k)$	001	010	011	100	101	110	111
Gray $\boldsymbol{g}(k)$	001	011	010	110	111	101	100

Recurrence (2)

$$\boldsymbol{x}(k) = g_0(k)v_1 \oplus g_1(k)v_2 \oplus \cdots \oplus g_{r-1}(k)v_r \tag{5}$$

Consider two consecutive integers k and $k+1$ differing in the lth bit. So

$$\begin{aligned} \boldsymbol{x}(k+1) &= g_0(k+1)v_1 \oplus g_1(k+1)v_2 \oplus \cdots \oplus g_{r-1}(k+1)v_r & (6)\\ &= g_0(k)v_1 \oplus g_1(k)v_2 \oplus \dots (g_l(k) \oplus 1)v_l \cdots \oplus g_{r-1}(k+1)v_r = \boldsymbol{x}(k+1) \oplus v_l & (7) \end{aligned}$$

Starting from 0, we never to calculate a Gray code, only l is needed to use (7). Otherwise we need the Gray code of the starting point.

3 Multivariate sequence.

Sobol initiates multivariate sequence based on hypercube distribution. For each coordinate $i \in 1, \dots, d$, we need a generator

$$\boldsymbol{V}^{(i)} = (v_1^{(i)}, \dots, v_r^{(i)}).$$

The determinant should non zero modulo 2.

3.1 Example in dimension 3

Consider direction numbers

$$(m_1, m_2, m_3) = (1, 1, 1), (m_1, m_2, m_3) = (1, 3, 5), (m_1, m_2, m_3) = (1, 1, 7).$$

The associated generator matrices using the first three primitive polynomials of Table 1 are

$$\boldsymbol{V}^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \boldsymbol{V}^{(2)} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \boldsymbol{V}^{(3)} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Table 2 gives the initial values to consider up to dimension 10.

i	m_1	m_2	m_3	m_4	m_5
1	1				
2	1				
3	1	1			
4	1	3	7		
5	1	1	5		
6	1	3	1	1	
7	1	1	3	7	
8	1	3	3	9	9
9	1	3	7	13	3
10	1	1	5	11	27

Table 2: Initial values

Using primitive polynomial of Table 1, the 5 direction numbers are computed. Below, we reproduce the first ten rows of Table 5.3 of Glasserman (2003).

```
first_prim_poly <- c(1, 3, 7, 11, 13, 19, 25, 35, 59, 47)
initmj <- list(
 1,
 1,
 c(1 , 1 ),
 c(1 , 3 , 7 ),
 c(1 , 1 , 5 ),
 c(1 , 3 , 1 , 1 ),
 c(1 , 1 , 3 , 7 ),
```

```r
 c(1 , 3 , 3 , 9 , 9 ),
 c(1 , 3 , 7 , 13 , 3 ),
 c(1 , 1 , 5 , 11 , 27  ))
firstmj <- function(i)
  sobol.directions.mj(initmj[[i]], first_prim_poly[i],
                      8-length(initmj[[i]]), echo=FALSE,
                      input="real", output="real")

t(sapply(1:length(first_prim_poly), firstmj))
```

```
##       [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
##  [1,]    1    1    1    1    1    1    1    1
##  [2,]    1    3    5   15   17   51   85  255
##  [3,]    1    1    7   11   13   61   67   79
##  [4,]    1    3    7    5    7   43   49  147
##  [5,]    1    1    5    3   15   51  125  141
##  [6,]    1    3    1    1    9   59   25   89
##  [7,]    1    1    3    7   31   47  109  173
##  [8,]    1    3    3    9    9   17   83  243
##  [9,]    1    3    7   13    3   35   89    9
## [10,]    1    1    5   11   27   53   69   25
```

Page 319