Low discrepancy sequence

Arabin Kumar Dey

Assistant Professor

Department of Mathematics
Indian Institute of Technology Guwahati

14-th January, 2013

Low discrepancy sequence

- We wish approximate an integral $\int_0^1 f(u)du$ using a total of n evaluations of the function f.
- Suppose that the function is smooth and I use points $u_i = \frac{2i-1}{2n}$, $i = 1, 2, \cdots, n$.
- Note : $\int_0^{1/n} f(u)du = \frac{1}{n} f(u_1)$.
  For a point in the interval u_1

- How good is this approximation ? Using a Taylor series expansion around the point $f(u_1)$,

$$f(x) = f(u_1) + f^{'}(u_1)(x - u_1)$$

- and integrating both sides over the interval $[0, \frac{1}{n}]$ we obtain

$$\int_0^{1/n} f(u)du \approx \frac{1}{n} f(u_1) + \frac{f(u_1)}{4n^2}$$

- If the function f has a bounded first derivative this means that the integral over each subinterval

$$\int_{(j-1)/n}^{j/n} f(u)du = \frac{1}{n} f(u_1)$$

with an error that is less than

$$\text{constant} \times \frac{1}{n^2}$$

- Therefore the error in the sum over n such intervals is less than a constant$\times\frac{1}{n}$. How does this compare to a crude Monte Carlo integral ?
- We have seen that if we randomly select n uniform(0,1) poins u_i and use the crude estimator

$$\frac{1}{n}\sum_{i=1}^{n} f(u_i)$$

- then the estimator has variance

$$\frac{\sigma^2}{n}$$

where

$$\sigma^2 = var(f(U_i))$$

and standard error

$$\frac{\sigma}{\sqrt{n}}$$

2 dimension

- For large values of n notice that

$$constant \times \frac{1}{n} < \frac{\sigma}{\sqrt{n}}$$

. i.e. the error in the numerical integral is less than that in the monte carlo integral. **For large values of n, and for smooth functions f in one-dimension, numerical integration is better than Monte Carlo integration.**

- Suppose we now want to find an integral of the form

$$\int \int f(u_1, u_2) du_1 du_2$$

- where the integral is over the unit square. Again we wish to use n evaluation of the function. Suppose we use equally spaced points on a lattice in the two dimensional unit square and suppose $n = m^2$. If we define

$$u_i = \frac{2i-1}{2m}, \quad i = 1, 2, \cdots, m$$

- the error in this approximation is less than or equal to

$$constant \times \frac{1}{m} = constant \times \frac{1}{\sqrt{n}}$$

- Note that this is same order of magnitude as the standard error of a Monte Carlo integral.
- For dimensions higher than 2, for example for evaluating an integral like

$$\int\int\int f(u_1, u_2, u_3)du_1du_2du_3$$

the Monte Carlo integral as measured by the order of the error term than is a numerical integral based on placement of points on a lattice.

- Equally spaced points on the line, or in space have the advantage that they fill holes efficiently (they get reasonably close to all points in the space). **One disadvantage is that I need to know in advance how many points (n) are to be selected to that I can space them $1/n$ apart.**
  Is it possible to construct a sequence so that at least periodically the sequence consists of equally spaced points ?

Low Discrepancy Sequences

- low-discrepancy: successive numbers are added in a position as far as possible from the other numbers.
- i.e. avoiding clustering
- the numbers generated sequentially fill in the larger “gaps” between the previous numbers of the sequence.
- In dimension 1, the van der Corput (1935) sequence in base 2, starts from zero, and is confined in the interval $[0, 1)$.

Vander Corput Sequence, base b

- The n'th term of the van der Corput sequence, for base b, is generated as follows: - The decimal-base number n is expanded in the base b. For example, n = 4 in base 2 is 100 $(4 = 1 \times 4 + 0 \times 2 + 0 \times 1)$;
- The number in base b is reflected. In the example, 100 becomes 001;
- Map into interval [0,1). 001 becomes 0.001 (binary decimal) corresponds to the decimal number 1/8, that is 1/8 $(= 0 \times (1/2) + 0 \times (1/4) + 1 \times (1/8))$.

General Van der Corput

- In general, for base b if

$$n = \sum_{j=0}^{m} a_j(n) b^j$$

- Van der corput base $b(n) = \Phi_b(n) = \sum_{j=0}^{m} a_j(n) b^{-j-1}$

Halton sequence

Let $p_1, p_2, \cdots, p_m$ be a pairwise prime integers. The Halton sequence is defined as the sequence of vectors :
$x_i := (\phi_{p_1}(i), \phi_{p_2}(i), ..., \phi_{p_m}(i)), \quad i = 1, 2, \cdots$ Usually one takes $p_1, p_2, \cdots, p_m$ to the the first m prime numbers.

m	repres base 2	first comp	repres base 3	second comp	repres base 5	third comp
1	1	1/2	1	1/3	1	1/5
2	10	1/4	2	2/3	2	2/5
3	11	3/4	10	1/9	3	3/5
4	100	1/8	11	4/9	4	4/5
5	101	5/8	12	7/9	10	1/25
6	110	3/8	20	2/9	11	6/25
7	111	7/8	21	5/9	12	11/25
9	1000	1/16	22	8/9	13	16/25
10	1001	9/16	100	1/27	14	21/25