数理解析研究所講究録
第 1620 巻 2009 年 204-210

Low discrepancy sequences

Makoto Mori
Dept. Math.College of Humanities and Sciences
Nihon University

July 8, 2008

1 introduction

To calculate the numerical integration $\int_0^1 f\,dx$ of a function f on the unit interval $[0,1]$, the most natural way is to use Riemaninan sum

$$\frac{1}{n}\sum_{i=0}^{n-1} f\left(\frac{i}{n}\right).$$

But in general, the speed of convergence is very slow. Another way to calculate numerical integration is to use random variables. By the law of large numbers, for independent and identically uniformly distributed random variable $X_1, X_2, \ldots$ on $[0,1]$,

$$\frac{f(X_1)+\cdots+f(X_n)}{n}$$

converges to its expectation $\int_0^1 f\,dx$, and by the central limit theorem, its speed of convergence is of order $\frac{1}{\sqrt{n}}$. This method is called Monte–Carlo method.

Instead of using random variables, quasi Monte–Carlo method uses a sequence, which we call quasi random sequence $x_1, x_2, \ldots$, and approximate the integral $\int_0^1 f\,dx$ by its average

$$\frac{f(x_1)+\cdots+f(x_n)}{n}.$$

In this article, we will construct one of the best quasi random numbers called low discrepancy sequences from the view point of ergodic theory.

2 low discrepancy sequence

Let us consider a sequence $x_1, x_2, \dots \in [0,1]$. We define its discrepancy by

$$D(N) = \sup_J \left| \frac{\#\{x_i \in J : i \leq n\}}{n} - |J| \right|,$$

where sup is taken over all the intevals contained in $[0,1]$, and $|J|$ is the Lebesgue measure of an interval J.

We know that $D(N)$ is greater than or equal to the order $\frac{\log N}{N}$. A sequence which satisfies

$$D(N) = O\left(\frac{\log N}{N}\right)$$

is called of low discrepancy. This sequence converges much faster than the sequence obtained by independent and identically distributed random variables. For a function f with bounded variation, it is also well known that there exists a constant C such that

$$\left| \frac{f(x_1) + \cdots + f(x_N)}{N} - \int_0^1 f\,dx \right| \leq C\,V(f)D(N),$$

where $V(f)$ is the total variation of f.

3 Perron–Frobenius operator

We will consider a piecewise linear transformation $F\colon [0,1] \to [0,1]$. We will denote the Perron–Frobenius operator by

$$Pf(x) = \sum_{y:\ F(y)=x} f(y)|F'(y)|^{-1}.$$

Then this operator determines the ergodic property of the dynamical system. Let us assume that the transformation is expanding, that is, the lower Lyapunov number $\xi > 0$, where

$$\xi = \liminf_{n\to\infty} \frac{1}{n} \operatorname*{ess\,inf}_{x\in[0,1]} \log |F^{n\prime}(x)|.$$

Then

1. P has eigenvalue 1, and there exists a non negative eigenfunction ρ such that $\int \rho\,dx = 1$. Then ρ is the density of an invariant probability measure μ.

2. If the eigenspace of the eigenvalue 1 is simple, then the dynamical system $([0,1], \mu, F)$ is ergodic.

3. Moreover, if there exists no eigenvalue except 1, then the dynamical system is mixing.

Hereafter, we consider piecewise linear transformations, and assume that the dynamical system is mixing. Moreover to construct the uniformly distributed quasi random sequence, we assume that $|F'| \equiv \beta > 1$. Then there exists a fintite number of intervals on which F is monotone. We denote the indices of intervals $\mathcal{A} = \{a, b, \ldots\}$, and denote subintervals by $\langle a \rangle$ and so on. We call $\mathcal{A}$ an alphabet.

A finite sequence $a_1 \cdots a_n$ $(a_1, \ldots, a_n \in \mathcal{A})$ is called a word, and define

1. $|w| = n$,
2. $\langle w \rangle = \bigcap_{i=1}^{n} F^{-i+1}(\langle a_i \rangle)$,
3. wx is a point whcih satisfies $wx \in \langle w \rangle$ and $F^{|w|}(wx) = x$, if it exists,
4. an end point of $\langle a \rangle$ is called Markov if ifor a sequence $a_n \in \langle a \rangle$ and $a_n \to a$ the image $F(a_n)$ converges to one of the end point of $\langle b \rangle$ $(b \in \mathcal{A})$. The endpoints $\sup \langle a \rangle$ and $\inf \langle a \rangle$ which are not Markov is called non–Markov endpoints.

Now we will consider the convergence for $f \in L^1$ and $g \in L^\infty$

$$\int f(x)\, g(F^n(x))\, dx = \int P^n f(x)\, g(x)\, dx \to \int f\, dx \int g\, d\mu.$$

This speed of convergence is called the decay rate of correlation. As an original definition P is the operator from L^1 to itself. In this case, all the points inside the unit circle are eigenvalues, that is, the decay rate of correlation has no meaning. Fortunately, the eigenfunction of the eigenvalue 1 belongs to BV, the set of functions with bounded variation. Thus we restrict the domain of P to BV. Then the essential spectral radius equals $\beta^{-1} = e^{-\xi}$, and the decay rate of correlation equals the second eigenvalue of P. If there exists no eigenvalue in the annuls $\beta^{-1} < |z| \leq 1$ except 1, then the decay rate of correlation

$$\int f(x) g(F^n(x))\, dx - \int f\, dx \int g\, d\mu = O(\beta^{-n}).$$

This is the fastest speed of convergence, and we will show that from the dynamical system which satisfies the above condition, we can construct low discrepancy sequences.

4 van der Corput sequence

The van der Corput sequence is the one of the most famous low discrepancy sequence. This is made as follows:

1. Let us consider a sequence of natural number $1, 2, 3, \ldots$.
2. Express them by binary digits
$$1, 10, 11, 100, 101, 110, 111, 1000, \ldots.$$
3. Reverse the order
$$1, 01, 11, 001, 101, 011, 111, 0001, \ldots.$$
4. Add 0. in front of the sequence,
$$0.1, 0.01, 0.11, 0.001, 0.101, 0.011, 0.111, 0.0001, \ldots.$$

We will see this sequence from the view point of dynamical system. For a piecewise linear transformation F with alphabet $\mathcal{A}$, let us consider a natural order on $\mathcal{A}$. Choose any $x \in [0,1]$, and arrange the inverse images of x in this order, say $ax, bx, \ldots$. Moreover, we define order $wx < w'x$ either one of the following holds:

1. $|w| < |w'|$,
2. $w = a_1 \cdots a_n$ and $w' = b_1 \cdots b_n$, $a_{k+1} \cdots a_n = b_{k+1} \cdots b_n$, and $a_n < b_n$.

We call the sequence $\{wx\colon wx \text{ exists}\}$ arranged in the above order a van der Corput sequence associated with F. The original van der Corput sequence is the case that $F(x) = 2x \pmod 1$ and the initial point $x = \frac{1}{2}$.

Theorem 1 *Assume that the dynamical system is mixing, and moreover, assume that there exists no eigenvalues of P in the annulus $\beta^{-1} < |z| \le 1$ except 1. Then the discrepancy of van der Corput sequence is*
$$D(N) = O\left(\frac{(\log N)^{k+1}}{N}\right),$$
where k is the number of non Markov endpoints $\sup\langle a\rangle$ *and* $\inf\langle a\rangle$ $(a \in \mathcal{A})$.

5 Higher dimensional cases

It is proved that for $d = 2$ that

$$D(N) \geq C\frac{(\log N)^d}{N},$$

and conjectured it will hold even for $d \geq 3$. Thus we also call a sequence of low discrepancy if

$$D(N) = O\left(\frac{(\log N)^d}{N}\right).$$

For 1 dimensional cases, as we said before, the essential spectral radius satisfies

$$\frac{1}{\beta} = e^{-\xi}.$$

For higher dimensional cases, we define ξ using Jacobian instead of $F^{n\prime}$. Moreover, instead of the set of functions BV of bounded variation, let $\mathcal{B}$ be the set of functions for which there exists a partition $f = \sum_{w \in \mathcal{W}} C_w 1_{\langle w \rangle}$ such that for any $0 < r < 1$

$$||f||_r = \inf \sum_{m=0}^{\infty} r^m \sum_{|w|=m} |C_w| < \infty,$$

where inf is taken over all decompositions of f

$\mathcal{B}$ becomes a locally convex space with semi norms $|| \cdot ||_r$ $(0 < r < 1)$. Because we can choose a decomposition such that $\sum_{|w|=m} |C_w|$ is less than or equal to the total variation of f, thus all the functions with bounded variation belongs to $\mathcal{B}$. Then even for higher dimensional cases, we can prove the similar results as in 1 dimensional cases, if the essential spectrum radius equals $e^{-\xi}$ and there exists no eigenvalues except 1 in $e^{-\xi} < |z| < 1$, then we can get the low discrepancy sequence in a same manner. However, it is not easy to construct transformations for which the essential spectral radius equals $e^{-\xi}$.

The famous low discrepancy sequence is a Halton sequence, which uses several integers. Then, for large d, the constant C may become large. So, we want to construct low discrepancy sequences using binary expansions.

We find two examples for two dimensional cases.

An example Define a sequence $s_0 = 00\cdots$, and $s_1 = 101000100000001\cdots$. Then define

$$F\begin{pmatrix} x_1x_2\cdots \\ y_1y_2\cdots \end{pmatrix} = \begin{pmatrix} x_2x_3\cdots \\ y_2y_3\cdots \end{pmatrix} + \begin{pmatrix} s_{y_1} \\ s_{x_1} \end{pmatrix}.$$

Then this F generates low discrepancy sequence. This sequence s_1 is constructed as follows:

1. As the first letter, choose 1. Then the first letter 0 of s_0 and the first letter of s_1 generate all the words 0 and 1 with length 1.

2. As the second letter, we may choose any letter. So we choose 0. We define 101 as the first three letters of s_1. Then the first two letters of $s_0 = 00$ and the first two letter s $s_1 = 10$, and the first two letters of $\theta s_1 = 01$, where θ is the shift operator to left. Then these three words $00, 10, 01$ generate all the words with length two.

3. As the fourth letter of s_1, we may choose any letter, so we choose 0 again. Then we define $s_1 = 10100\ldots$, then the first three letters of $s_0, s_1, \theta s_1, \theta^2 s_1$ generate all the words with length three. Here addition is done digits by digits.

We continue this procedure, and get s_1. From the construction, it is not difficult to see that any rectangle of the form $J = \langle w_1 \rangle \times \langle w_2 \rangle$ with $|w_1| + |w_2| = 2n$ for some n, $F^n(J) = [0,1]^2$. Using this fact, we can show that the essential spectra ot this transformation equals $\frac{1}{4}$, and there exists no eigenvalue which satisfies $\frac{1}{4} < |z| < 1$.

Unfortunately, we can not construct similar transformation with $d \geq 3$. Thus, now we are studying another examples.

2nd example Let $\mathcal{A} = \{(i,j) \colon i,j \in \{0,1\}\}$, and

$$\begin{aligned} \langle 0,0 \rangle &= [0, \tfrac{1}{2}] \times [0, \tfrac{1}{2}), \quad \langle 0,1 \rangle = [\tfrac{1}{2}, 1] \times [0, \tfrac{1}{2}), \\ \langle 1,0 \rangle &= [0, \tfrac{1}{2}] \times [\tfrac{1}{2}, 1], \quad \langle 1,1 \rangle = [\tfrac{1}{2}, 1] \times [\tfrac{1}{2}, 1]. \end{aligned}$$

We denote a point $(x,y) \in [0,1]^2$ by

$$(i_1, j_1)(i_2, j_2) \cdots \qquad (i_k . j_k \in \{0,1\}),$$

where

$$x = 0.i_1 i_2 \cdots, \qquad y = 0, j_1 j_2 \cdots .$$

Now we denote $a_k = (i_k, j_k)$. We will consider permutaions σ_w on $\mathcal{A}$ with word w as index. Let

$$F(a_1 a_2 \cdots) = \sigma_{a_1}(a_2) \sigma_{a_1 a_2}(a_3) \sigma_{a_2 a_3}(a_4) \cdots .$$

By computer experiments, we found several candidates. Moreover, we believe that we can construct higher dimensional low discrepancy sequences using permutaions on the symbols.

参考文献

[1] Low discrepancy sequences generated by piecewise linear Maps, Monte Carlo methods and Applications, vol. 4 No.2, 141-162(1998).

[2] Construction of two dimensional low discrepancy sequences, Monte Carlo methods and Applications vol. 8 No.2, 159-170(2002)

[3] Spectra of Perron–Frobenius operator and new construction of two dimensional low discrepancy sequences, to appear in Monte Carlo Methods and Applications.