RIESZ-RAIKOV SUMS AND IRRATIONAL ROTATIONS

Katusi FUKUYAMA (Kobe)

Abstract. In this note, we will consider on an effectiveness of quasi Monte-Carlo method. We will prove some central limit theorem for Riesz-Raikov sums, and by using that we will explain why the quasi Monte-Carlo method is not no more effective in case when the integrand is complicated functions. This partly solves the Sugita's conjecture. We will also give a new proof of a previous result.

§1. Monte-Carlo and quasi Monte-Carlo method

In this note, we study some topics related to random numbers and numerical integrations. Among various methods of numerical integration, we here treat Monte-Carlo method and quasi Monte-Carlo method. Let us first explain these two methods, especially in one-dimensional case.

Suppose that a Riemann integrable function f on unit interval $[0,1]$ is given. The purpose is to evaluate the integral $\int_0^1 f(x)\,dx$ numerically. If we can prepare $[0,1]$-valued uniform i.i.d. $\{\xi_n\}$, we have the following law of the large numbers:

$$\frac{1}{N}\sum_{n=1}^{N} f(\xi_n) \longrightarrow \int_0^1 f(x)\,dx \quad \text{a.s.,} \tag{1.1}$$

The Monte-Carlo method evaluates the integral by using this theorem. More precisely, by using $[0,1]$-valued uniform random numbers $\{\tilde{\xi}_n\}$, we have an approximate value of the integral by calculating $(1/N)\sum_{n=1}^{N} f(\tilde{\xi}_n)$ for large enough N.

It is important for practical purpose to ask how to prepare random numbers and how large we should take N. As to the first question, it seems that we still don't have complete answer. The question seems to be too difficult and we do not study further here. We only mention that the statement (1.1) contains the phrase 'almost surely'and there definitely exists a non-trivial exceptional set, so that even in the case when we have i.i.d., it is not easy to pick up correct ω which ensures the convergence.

On the other hand, we have a good answer to the second question. By precise version of the law of large numbers or the law of iterated logarithm, we have the following estimate of errors:

$$\frac{1}{N}\sum_{n=1}^{N} f(\xi_n) - \int_0^1 f(x)\,dx = O\big(N^{\varepsilon-1/2}\big) \quad \text{a.s.} \quad \text{for all} \quad \varepsilon > 0, \tag{1.2}$$

in which ε can not be 0.

By the way, to escape from the above ambiguity of exceptional ω, the following quasi Monte-Carlo method seems to be effective. Let α be an irrational number and ω_0 be an arbitrary real number. Then, by Weyl's theorem, we have

$$\frac{1}{N}\sum_{n=1}^{N} f(\omega_0 + n\alpha) \longrightarrow \int_0^1 f(x)\,dx, \tag{1.3}$$

where f is extended over $\mathbf{R}$ periodically. The first advantage of the quasi Monte-Carlo method is that there is no ambiguity in the statement, and the second is the following estimate of errors: If α is an algebraic irrational number, we have

$$\frac{1}{N}\sum_{n=1}^{N} f(\omega_0 + n\alpha) - \int_0^1 f(x)\,dx = O\big(N^{\varepsilon-1}\big) \quad \text{for all} \quad \varepsilon > 0. \tag{1.4}$$

By these results, the quasi Monte-Carlo method seems to be much effective than the Monte-Carlo method, but it is not generally correct.

For example, to integrate a function on high dimensional space, the quasi Monte-Carlo method is not effective. An extensive treatment on this phenomenon is recently given by Sugita & Takanobu [16].

And when the integrand f is irregular in some sense, the quasi Monte-Carlo method is not effective, either. In this paper, we give some treatment on this phenomenon. The story goes back to the results given by Sugita [14].

2. Sugita's results and its refinements

Let S_m be a one-dimensional simple random walk. Sugita [14] considers the following scheme to evaluate the probability $P(S_m = a)$ numerically.

Prepare the Lebesgue probability space $\big(\Omega := [\,0,1), P(d\omega) := d\omega\big)$ and Rademacher functions r_i. Recall that r_i is a function on $\mathbf{R}$ with period 1 defined by $r_i(\omega) := r_1(2^{i-1}\omega)$ and $r_1(\omega) := \mathbf{1}_{[\,0,1/2)}(\omega) - \mathbf{1}_{[\,1/2,1)}(\omega)$ $(\omega \in \Omega)$.

By putting $S_m := \sum_{i=1}^{m} r_i$, a simple random walk $\{S_m\}$ is defined on this space. Since $P(S_m = a)$ is given by the integral $\int_0^1 \mathbf{1}_{\{S_m=a\}}(\omega)d\omega$ on this space, by virtue of the quasi-Monte Carlo method, it holds that

$$P(S_m = a) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{\{S_m=a\}}(\omega_0 + n\alpha). \tag{2.1}$$

for all $\omega_0 \in \Omega$ and $\alpha \notin \mathbf{Q}$. When α is an algebraic irrational, the theoretical rate of convergence in (2.1) is of $O(N^{\varepsilon-1})$, but Sugita [14] reports that this rate cannot be observed in numerical experiments and only the rate $O(N^{\varepsilon-1/2})$ is observed when m is large. He claims that the quasi Monte-Carlo method is not effective in this scheme. He explains this phenomenon by his conjecture that the sequence $S_m(\omega)$, $S_m(\omega+\alpha)$, $S_m(\omega+2\alpha)$, ... is nearly independent when m is large. His result is as follows:

Theorem A. *The correlation $R_\alpha^{(m)}(n)$ of stationary sequence $\boldsymbol{X}_\alpha^{(m)} = \{S_m(\omega)/\sqrt{m}, S_m(\omega+\alpha)/\sqrt{m}, S_m(\omega+2\alpha)/\sqrt{m}, \dots\}$ decays when $m \to 0$ and its order is of $O(m^{-1/2+\varepsilon})$ for every n, $\varepsilon > 0$, and for almost every α.*

As a refinement of this theorem, we have the following results (Cf. [5].):

Theorem 1. *(1) For almost every α, the stationary sequence $\boldsymbol{X}_\alpha^{(m)}$ converges to Gaussian i.i.d. as $m \to \infty$, in sense of convergence of every finite dimensional distributions.*
(2) For almost every α and every n, $\limsup_{m\to\infty} \sqrt{m/\log\log m}\, R_\alpha^{(m)}(n) = \sqrt{2/3}$.
(3) Hausdorff dimension of the set of α for which (1) does not hold is 1.

By (1) and (2), we can say that Sugita's conjecture is correct for almost every α, and that Theorem A gives almost best possible result, but by (3) we conclude that the exceptional set for this phenomenon is also large.

There is, however, other interpretation of this result apart from the conjecture. There is some possibility that random numbers can be generated by $\boldsymbol{X}_\alpha^{(m)}$ for large m. For the purpose of practical use, we must be serious about the exceptional set of α. We need some test to distinguish good α and bad one. In the next section, we give results for every α and try to give some answer to this demand.

By the way, we would like to mention that, by using similar but more simple scheme, Sugita [15] gives effective random number generator.

3. Riesz-Raikov sums and irrational rotation

To study limit behavior of $\boldsymbol{X}_\alpha^{(m)}$ for every α, the following treatment as Riesz-Raikov sums is appropriate. Let us first begin with the introduction of this notion.

Let $\theta > 1$, $\alpha \in \mathbf{R}$, and g be a function on $\mathbf{R}$ with period 1 satisfying

$$\int_0^1 g(t)\,dt = 0 \quad \text{and} \quad 0 < \int_0^1 \left|g(t)\right|^2 dt < \infty. \tag{3.1}$$

Let us put

$$X^{(m)}(\omega) := \frac{1}{\sqrt{m}} \sum_{k=1}^{m} g(\theta^{k-1}\omega).$$

The sum in the right hand side is called Riesz-Raikov sum. If we put $\theta := 2$ and $g := r_1$, we have $g(\theta^{k-1}\omega) = r_k(\omega)$ and $X^{(m)} = S_m/\sqrt{m}$ and thereby we can regard the Riesz-Raikov sum as an extension of simple random walk $\{S_m\}$.

As before, we have the next central limit theorem for Riesz-Raikov sums. (see Berkes [1], [2] and Fukuyama [6].)

Theorem B. *Let $\theta > 1$ and g satisfy* (3.1) *and*

$$\sum_{k=0}^{\infty} \left\| s_{2^{k+1}} - s_{2^k} \right\|_2 < \infty, \tag{3.2}$$

where $\|\cdot\|_2$ denotes $L^2[0,1]$-norm and s_i denotes i-th subsum of the Fourier series of g. Then for any measurable $\Omega \subset [0,1)$ with $|\Omega| > 0$,

$$X^{(m)} \xrightarrow{\mathscr{D}} N(0,v)$$

holds on $\big(\Omega, d\omega/|\Omega|\big)$, where the limiting variance $v = v_{g,\theta}$ is given as follows: In case

$$\theta^r \notin \mathbf{Q} \qquad (r \in \mathbf{N}) \tag{3.3}$$

is satisfied, $v = \|g\|_2^2$; Otherwise, by using p and q given by

$$\theta^r = \frac{p}{q} \ \text{where}\ r = \min\{\, n \in \mathbf{N} \mid \theta^n \in \mathbf{Q}\,\},\ p, q \in \mathbf{N} \ \text{and}\ \gcd(p,q) = 1, \tag{3.4}$$

we have

$$v = \|g\|_2^2 + 2\sum_{k=1}^{\infty} \int_0^1 g(p^k t) g(q^k t)\, dt < \infty.$$

As before, we operate irrational rotation to $X^{(m)}$ and define $\boldsymbol{X}_\alpha^{(m)}$, i.e.,

$$\boldsymbol{X}_\alpha^{(m)} := \big\{X_{\alpha;n}^{(m)}\big\}_{n\in\mathbf{Z}} \quad \text{where} \quad X_{\alpha;n}^{(m)}(\omega) := X^{(m)}(\omega + n\alpha)\ .$$

Next theorem characterizes the possible limit distribution of $\boldsymbol{X}_\alpha^{(m)}$ as $m \to \infty$. There is no restriction or ambiguity on α.

Theorem 2. *Suppose that* (3.1) *and* (3.2) *are satisfied. Then for all* $\Omega \subset [0,1)$ *with* $|\Omega| > 0$, $\theta > 1$ *and* $\alpha \in \mathbf{R}$, *on probability space* $\big(\Omega, d\omega/|\Omega|\big)$, *the sequence* $\{\boldsymbol{X}^{(m)}_\alpha\}_{m\in\mathbf{N}}$ *is relatively compact in the sense of convergence in distribution on* $\mathbf{R}^{\mathbf{Z}}$, *and all the possible limit is stationary Gaussian sequence, i.e., for any subsequence of* $\{\boldsymbol{X}^{(m)}_\alpha\}_{m\in\mathbf{N}}$ *there exists a subsequence which converges to a stationary gaussian sequence. This subsequence can be taken independently of* Ω *and* g, *and the limit distribution does not depend on* Ω.

By this theorem, we see that only stationary gaussian distribution can appear as the limit distribution.

The next theorem gives a relation between convergence of the correlation and convergence in law. It means that, by calculating the correlation numerically, we can test whether α is good one or not. It has a practical significance.

Theorem 3. *Let us assume* (3.1) *and* (3.2). *Let* $\{m_j\}_{j\in\mathbf{N}}$ *be a subsequence of* $\mathbf{N}$ *and* $\Omega \subset [0,1)$ *satisfy* $|\Omega| > 0$. *For any* n *and* $l \in \mathbf{Z}$, *the convergence of correlation*

$$\lim_{j\to\infty} E\Big(X^{(m_j)}_{\alpha;n} X^{(m_j)}_{\alpha;l}\Big) = r_{n\,l} \tag{3.5}$$

holds if and only if $\boldsymbol{X}^{(m_j)}$ *converges as* $j \to \infty$ *to the stationary gaussian sequence with mean vector* $\mathbf{0}$ *and correlation matrix* $\{r_{n\,l}\}_{n,l\in\mathbf{Z}}$.

As before, the limit of $\boldsymbol{X}^{(m)}$ is gaussian i.i.d. for almost all α. The next theorem state this result.

Theorem 4. *Let us assume* (3.1) *and* (3.2). *For almost all* α *with respect to the Lebesgue measure, the sequence* $\boldsymbol{X}^{(m)}_\alpha$ *converges in distribution to gaussian i.i.d. The exceptional set of* α *can be taken to be independent of* Ω *and* g.

Let us next consider about the fluctuation of the correlation. From now on we only consider the case when $\theta > 1$ is an integer.

Theorem 5. *Let* $\Omega = [0,1)$, $\theta > 1$ *be an integer and* g *satisfies* (3.1), $v = v_{g,\theta} > 0$ *and*

$$\sum_{k=0}^{\infty} k \big\| s_{2^{k+1}} - s_{2^k} \big\|_2 < \infty. \tag{3.6}$$

Then the correlation on $\big(\Omega, d\omega\big)$ *obeys*

$$\limsup_{m\to\infty} \sqrt{\frac{m}{\log\log m}}\, E\Big(X^{(m_j)}_{\alpha;n} X^{(m_j)}_{\alpha;l}\Big) = \beta \quad a.e. \quad \alpha$$

for some positive constant β.

Unlike as it appears, proof of $\beta > 0$ is more difficult rather than the law of the iterated logarithm itself.

The next result claim that the exceptional set of α is not small with respect to the Hausdorff dimension.

Theorem 6. *Let $\theta > 1$ be an integer and g satisfies* (3.1)*, $v = v_{g,\theta} > 0$ and*

$$\sum_{k=0}^{\infty} 2^{k\gamma} \left\| s_{2^{k+1}} - s_{2^k} \right\|_2 < \infty \quad \textit{for some} \quad \gamma > 0 \tag{3.7}$$

The set of α for which the limit distribution of $\boldsymbol{X}_\alpha^{(m)}$ is dependent has Hausdorff dimension 1*.*

By the way, the regularity conditions (3.2), (3.6) and (3.7) are derived from some smoothness conditions on g. We now note these implications. Let us define the L^2-modulus of continuity $\omega_2(\delta)$ of function g by

$$\omega_2(\delta) := \sup_{|h| \le \delta} \left(\int_0^1 \left| g(t+h) - g(t) \right|^2 dt \right)^{1/2}.$$

The condition (3.2), (3.6) or (3.7) are derived from

$$\int_0^1 \frac{\omega_2(y)}{y}\, dy < \infty, \qquad \int_0^1 \frac{\omega_2(y) \log 1/y}{y}\, dy < \infty \quad \text{or} \quad \int_0^1 \frac{\omega_2(y)}{y^{1-\gamma}}\, dy < \infty, \tag{3.8}$$

respectively. (Cf. Zygmund [18], (3.3) of pp. 241).

Obviously Hölder continuous functions and functions of bounded variation satisfy these conditions.

Proofs of all results stated in this section can be found in [7].

4. Rademacher functions

In this section, we restrict ourself to the case of Rademacher functions, and give a sketch of proofs of our result putting $g = r_1$.

The next lemma plays an key.

Lemma 1. *For any sequence $\{\alpha_i\}$ of real numbers, the sequence $\{r_i'(\omega) := r_i(\omega + \alpha_i)\}$ is an i.i.d. on (Ω, P).*

Here we will give new proof. It is enough to prove

$$P\left(r_1' = \varepsilon_1, \ldots, r_n' = \varepsilon_n\right) = \frac{1}{2^n}, \quad \text{where} \quad \varepsilon_1, \ldots, \varepsilon_n \in \{-1, +1\}. \tag{4.1}$$

Let us put $\Psi(\omega) := \bigl(r_1'(\omega), \dots, r_n'(\omega)\bigr)$. (4.1) reduces to $\Psi^{-1}(\varepsilon_1, \dots, \varepsilon_n) = 1/2^n$. It is clear that $r_i'(\omega + 1/2^i) = -r_i'(\omega)$ holds, and thereby $r_i'(\omega + j/2^i) = (-1)^j r_i'(\omega)$ holds. If we take $j = 1, \dots, 2^n - 1$ arbitrary, we can write $j = (2j' + 1)2^m$. Hence we have

$$r_{n-m}'(\omega + j/2^n) = r_{n-m}'\bigl(\omega + (2j' + 1)/2^{n-m}\bigr) = -r_{n-m}'(\omega).$$

Thus we have $\Psi(\omega + j/2^n) \neq \Psi(\omega)$, and we can collude that $\Psi(\omega + j/2^n)$ $(j = 0, \dots, 2^n - 1)$ are different from each other, and thereby $\Psi(\omega + j/2^n)$ $(j = 0, \dots, 2^n - 1)$ takes all values in $\{-1, +1\}^n$.

It implies that $\Psi^{-1}(\varepsilon_1, \dots, \varepsilon_n) + j/2^n$ $(j = 0, \dots, 2^n - 1)$ do not intersect each other and the union of these coincides with Ω. Thus we can conclude that $\Psi^{-1}(\varepsilon_1, \dots, \varepsilon_n)$ has measure $1/2^n$. $\square$

To prove (1) of Theorem 1, it is enough to prove that the law of $a_1 X_{1;\alpha}^{(m)} + \cdots + a_k X_{k;\alpha}^{(m)}$ converges to the standard normal law for any $a_1, \dots, a_k$ with $a_1^2 + \cdots + a_k^2 = 1$.

To prove this, we use the next theorem due to McLeish [11].

Theorem B. *Let* $\{\, \zeta_{m,j}\,;\, 1 \le j \le k_m \,\}$ *be a triangular array of random variables and put* $L_m := \prod_{j \le k_m} \bigl(1 + \sqrt{-1} t \zeta_{m,j}\bigr)$. *The law of* $\sum_{j \le k_m} \zeta_{m,j}$ *converges to the standard normal distribution as* $m \to \infty$, *provided that the following four conditions are satisfied for all* $t \in \mathbf{R}$:

(1) $EL_m \to 1$ *as* $m \to \infty$;

(2) The sequence $\{L_m\}_{m \in \mathbf{N}}$ *is uniformly integrable;*

(3) $\sum_{j \le k_m} \zeta_{m,j}^2 \to 1$ *in probability as* $m \to \infty$;

(4) $\max_{j \le k_m} |\zeta_{m,j}| \to 0$ *in probability as* $m \to \infty$.

Putting $\zeta_{m,i}(\omega) := \bigl(1/\sqrt{m}\,\bigr) \sum_{j=1}^{k} a_j r_i(\omega + j\alpha)$, we apply this theorem. As to condition (1), by Lemma 1, we have $EL_m = 1$. (2) and (4) are easily verified. (3) is derived from the following:

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} r_i(\omega + p\alpha) r_i(\omega + q\alpha) = 0, \quad \text{a.s.}$$

It is easily proved that the above formula holds for almost all (ω, α). By applying Fubini's theorem, we complete the proof.

For the proof of the rest of the result, see [5].

References

[1] Berkes, I., On the asymptotic behavior of $\sum f(n_k x)$, Main theorems, Z. Wahr. verw. Geb., **34** (1976) 319–345

[2] Berkes, I., On the asymptotic behavior of $\sum f(n_k x)$, Applications, Z. Wahr. verw. Geb., **34** (1976) 347–365

[3] Billingsley, P., Ergodic theory and information. Wiley, New York, 1967

[4] Erdős, P. & Taylor, S. J., On the set of points of convergence of a lacunary trigonometric series and the equidistribution properties of related sequences, Proc. London Math. Soc. (3), **7** (1957) 598–615

[5] Fukuyama, K., The central limit theorem for Rademacher system, Proc. Japan Acad. Ser. A, **70** (1994) 253–256

[6] Fukuyama, K., The central limit theorem for Riesz-Raikov sums, Prob. Theory Related Fields, **100** (1994) 57–75

[7] Fukuyama, K., Riesz-Raikov sums and Weyl transform, Monte Carlo Methods and Appl., **2** (1996) 271–293

[8] Hawkes, J., Probabilistic behaviour of some lacunary series, Z. Wahr. verw. Geb. **53** (1980) 21–33

[9] Heyde, C. & Scott, J., Invariance principles for the law of the iterated logarithm for martingales and process with stationary increments, Ann. Prob., **1** (1973) 428–436

[10] Kupier, L. & Niederreiter, H., Uniform distribution of sequences, Wiley, New York, 1974.

[11] McLeish, D. L., Dependent central limit theorem and invariance principles, Annals of Prob. **2** (1974) 620–628

[12] Salem, R. & Zygmund, A., On lacunary trigonometric series I, Proc. Nat. Acad. Sci. U. S. A., **33** (1947) 333–338

[13] Salem, R. & Zygmund, A., On lacunary trigonometric series II, Proc. Nat. Acad. Sci. U. S. A., **34** (1948) 54–62

[14] Sugita, H., A private communication, (1993)

[15] Sugita, H., Pseudo-random number generator by means of irrational rotation, Monte Carlo methods and Appl., **1** (1995) 35-57

[16] Sugita, H. & Takanobu, S., Limit theorem for symmetric statistics with respect to Weyl transformation: Disappearance of dependency, preprint

[17] Takahashi, S., On the distribution of values of the type $\sum f(q^k t)$, Tôhoku Math. J., **14** (1962) 233–243

[18] Zygmund, A., Trigonometric series I, Cambridge university press, Cambridge, 1959

Katusi FUKUYAMA *Department of Mathematics, Kobe University, Rokko, Kobe, 657 Japan*

E-mail: `fukuyama@math.s.kobe-u.ac.jp`