

準モンテカルロ法

伏見 正則

1. はじめに

1945年頃, J. von Neumann と S. M. Ulam は乱数を使って決定論的な種々の数学的問題を解く方法を考え, このような方法の総称としてモンテカルロ法 (以下 MC 法と略す) という名称を用いた. 現在では, 対象が本来確率的な要素を含む場合も含めて, 乱数を用いて実験する方法のことを MC 法と呼ぶことが多い.

MC 法では, 何回か実験を繰り返し, 各回に得られた測定値の算術平均をもって, 知りたい特性値の推定値とするのが普通である. 1 回の実験では区間 $[0, 1]$ 上の s 個の一様乱数を用いて測定値を得るものとすれば, 知りたい特性値は形式的に s 重積分

$$I = \int_0^1 \cdots \int_0^1 f(x_1, \dots, x_s) dx_1 \cdots dx_s \quad (1)$$

の形に書き表わされる. われわれは, これを

$$A_N = \frac{1}{N} \sum_{n=1}^N f_n \quad (2)$$

によって推定する. ただし, f_n は n 回目の実験で使用した乱数の組に対応する f の値である. 確率変数としての A_N の期待値 $E(A_N)$ は I に等しく (すなわち A_N は I の不偏推定量であり), 各回の実験で使用する乱数の組が互いに独立であれば, A_N の分散 $\text{Var}(A_N)$ は

$$\text{Var}(A_N) = \frac{\sigma^2}{N},$$

$$\sigma^2 = \int_0^1 \cdots \int_0^1 [f(x_1, \dots, x_s) - I]^2 dx_1 \cdots dx_s$$

となることはよく知られている.

MC 法の欠点は, 誤差 $(A_N - I)$ の目安としての A_N の標準偏差 $\sqrt{\text{Var}(A_N)}$ が, N を増していったときに $N^{-1/2}$ という遅い速度でしか小さくならず, たとえば有効数字を十進で 1 桁増やすためには N を 100 倍に増やさなければならないというような効率の悪さである. 一方, MC 法の長所は, 関数 f が微分できないとか不連続であるなど, 解析的には性質が悪くても適用でき, また上

記の性質が次元 s によらないところにある.

ところで, f が解析的性質が良い関数である場合には台形公式などの数値積分法を多次元に拡張したものを使って I の近似値を求めることもできる. しかしこの場合には, 1 つの次元方向を m 等分するとすれば, 全部で $(m+1)^s$ 回 f の値を計算しなければならず, 計算量が次元 s とともに指数関数的に増大するという欠点がある.

そこで, 虫のよい要望が出てくる: 上の 2 つの方法の“中間的な”方法で, 有効数字の桁数を増すために関数計算の回数 N を急激に大きくしなくてもよく, かつ, 計算量が次元とともにあまり増えない方法はないものだろうか? —このような要望にある意味で応えるのが準モンテカルロ法 (Quasi-Monte Carlo method, 以下 QMC 法と略す) と呼ばれる耳よりな (耳なれない?) 方法である. これにはいくつかの方法があり, 関数 f の解析性がきわめて良い場合に有効な優良格子点 (good lattice points) 法もそのひとつであるが, 本稿では解析性が良くなくても使える“差異の小さい点列 (low-discrepancy sequence) を用いる方法”について述べる.

2. 点列の差異と積分の誤差

(1) 式の積分領域, すなわち s 次元の単位超立方体 C_s 内の N 個の点の集合 $S_N = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ の (s 次元) 差異 (discrepancy) というのは,

$$D_N^{(s)} = \sup_J \left| \frac{\nu(J; N)}{N} - \text{vol}(J) \right|$$

で定義される量である. ここで上限 \sup は

$$J = [0, t_1] \times \cdots \times [0, t_s] \subseteq C_s$$

という形のすべての (超) 直方体にわたってとるものとする. また, $\nu(J; N)$ は J に含まれる S_N の点の個数を表わし, $\text{vol}(J)$ は J の体積である. 1 次元 ($s=1$) の場合には, $D_N^{(s)}$ は S_N が一様分布からのランダムサンプルと見なせるかどうかの検定に使われるコルモゴロフ・スミルノフ統計量に一致する. $D_N^{(s)}$ は, その多次

元への一種の拡張であり、 S_N の C_s 内における一様性を表わす1つの尺度である。

差異の小さい点列が定積分 (1) の評価に有効であるとされる根拠は、次の Koksma-Hlawka の不等式である。

$$|I - A_N| \leq V(f) D_N^{(s)} \quad (3)$$

ここに $V(f)$ は関数 f の Hardy-Krause の意味での変動と呼ばれ、 C_s における f の変化の激しさを表わす量である。その厳密な定義は省略するが、連続性や微分可能性といった条件が成り立たなくても有界になりうる量である。

3. 差異の小さい点列の例

差異の小さい点列を構成する方法はいろいろ提案されているが、そのうちのいくつかを紹介しよう。

1) Halton の方法

$$x_n = (\phi_{b_1}(n-1), \dots, \phi_{b_s}(n-1)), \quad n=1, 2, \dots \quad (4)$$

ここに b_1, \dots, b_s は、どの2つをとっても互いに素な自然数の組であり、 $\phi_b(n)$ は radical-inverse function と呼ばれる関数であり、 n の b 進展開を $n = \sum_{j=0}^{\infty} a_j b^j$ とすると、 $\phi_b(n) = \sum_{j=0}^{\infty} a_j b^{-j-1}$ である。(4)は無限点列であり、最初の N 点からなる集合の差異は次の不等式を満たす。

$$D_N^{(s)} \leq B_s N^{-1} (\log N)^s + O(N^{-1} (\log N)^{s-1}) \quad (5)$$

$$B_s = \prod_{k=1}^s \frac{b_k - 1}{2 \log b_k} \quad (6)$$

2) Hammersley の方法

$$x_n = \left(\phi_{b_1}(n-1), \dots, \phi_{b_{s-1}}(n-1), \frac{n-1}{N} \right), \\ n=1, \dots, N \quad (7)$$

b_1, \dots, b_{s-1} に対する条件は1)と同様であり、差異は

$$D_N^{(s)} \leq B_{s-1} N^{-1} (\log N)^{s-1} + O(N^{-1} (\log N)^{s-2}) \quad (8)$$

を満たす。

これらの点列を使うと、誤差の上界評価 (3) が $O(N^{-1} (\log N)^s)$ あるいは $O(N^{-1} (\log N)^{s-1})$ となり MC 法における誤差の確率的評価 (標準偏差) が $O(N^{-1/2})$ であるのといちじるしい対照をなす。また、1)と2)を比べると、1)はオーダの意味で劣っており、使い途がないと思われるかもしれないが、 N を順次増し、任意のところで止められるというメリットがある。なお、一般に、ある無限点列の差異に対して (5) の形の上界評価 (係数 B_s は上記のものと違ってよい)

が成り立つならば、その点列の最後の座標成分を (上記の (4) から (7) を作ったと同様に) $(n-1)/N$ で置き換えて得られる N 個の点の集合の差異に対しては (8) 式の形の上界評価ができることが知られている。

係数 B_s (6) を最小にするためには、 $b_1=2, b_2=3, \dots$ というふうに、 b_k として小さい方から k 番目の素数をとるのがよい。しかし、そうしても $\log B_s = O(s \log s)$ であり、 B_s は s とともに急激に大きくなってしまふ。

3) Sobol' の方法

無限列を生成する方法で、差異に関する評価式は (5) と同じ形であるが、係数 B_s が $\log B_s = O(s \log \log s)$ となり、 $s \geq 5$ では (6) 式の B_s より小さくなる。系列の構成法は次のとおりである。

$x_n = (x_n^{(1)}, \dots, x_n^{(s)})$ の各座標成分 $x_n^{(k)}$ をそれぞれ独立に次の漸化式を用いて計算する。

$$x_0^{(k)} = 0 \\ x_{n+1}^{(k)} = x_n^{(k)} \oplus v_m^{(k)}, \quad 1 \leq k \leq s, \quad n \geq 1 \quad (9)$$

ここに、 \oplus は 2 進法での繰り上りなしの足し算を表わす。また $m(n)$ は、 n の 2 進展開を $\dots d_3 d_2 d_1$ とするとき、 $d_m = 0$ となる最小の m を表わす。 $v_m^{(k)}$ は次の漸化式によって定まる 2 進小数である。(記号を簡単にするため、上つき添字 k は省略する)

$$v_m = c_1 v_{m-1} \oplus c_2 v_{m-2} \oplus \dots \oplus c_{p-1} v_{m-p+1} \oplus v_{m-p} \oplus \\ [v_{m-p}/2^p], \quad m \geq p+1 \quad (10)$$

ただし係数 c_j は、 $x^p + c_1 x^{p-1} + \dots + c_{p-1} x + 1$ がガロア体 $GF(2)$ 上の原始多項式となるように定め、また各次元 k ごとに異なる原始多項式を使用するものとする。漸化式 (10) を使うための初期値 $v_m (1 \leq m \leq p)$ は、 $v_m < 1$ で、かつ小数点以下第 m ビットが 1、それより下位のビットはすべて 0 となるように定める。

4) Faure の方法

以上の3つの方法で構成される点列の差異の上界評価式の主要項の係数 B_s は、いずれも $s \rightarrow \infty$ のとき無限大になる。これに対して、本方法によって得られる系列に対する B_s は

$$B_2 = \frac{3}{16(\log 2)^2}, \quad B_s = \frac{1}{s!} \left(\frac{r_s - 1}{2 \log r_s} \right)^s \quad (s \geq 3)$$

r_s : s 以上で最小の素数

であり、すべての $s \geq 2$ について他より小さく、 $s \rightarrow \infty$ のとき $B_s \rightarrow 0$ となる。

与えられた次元 s に対して、 s 以上で最小の素数 (上記の r_s) を r と書くことにする。 $x_n (n \geq 1)$ の各座標成分 $x_n^{(k)}$ は、 r 進小数として次のように定める。

$$x_n^{(k)} = \sum_{j=0}^{\infty} y_j^{(k)}(n) r^{-j-1}$$

ここに $y_j^{(k)}(n)$ は、 $n-1$ の r 進展開を

$$n-1 = \sum_{j=0}^{\infty} a_j(n) r^j$$

とすると、

$$y_j^{(k)}(n) = \sum_{l=j}^{\infty} \binom{l}{j} (k-1)^{l-j} a_l(n) \pmod{r}$$

で与えられる。(便宜上、 $k=1$ の場合は $0^0=1$ と解釈するものとする)

4. 高速に発生できる系列

前節で述べた方法には、いずれも点列の生成に時間がかかるという欠点がある。差異の小さい点列の生成が遅いのであれば、QMC法を用いるよりも通常の乱数によるMC法を使った方が同一時間内にもっと精度の良い解が得られるという場合もありうる。そこで本節では、最近筆者ら [9] が考案した高速に生成できる差異の小さい点列を紹介する。生成に要する時間は、一番速い通常の乱数の生成とほとんど同じである。

4.1 生成法

2進小数の列 $\{u_n\}$ を漸化式

$$u_n = u_{n-p} \oplus u_{n-p+q}, \quad n \geq p+1 \quad (11)$$

によって作る。ただしパラメタの組 (p, q) は表1から選ぶ。これを用いて点列 $\{x_n\}$ を次のように定める。

$$x_n = (u_n, u_{n+1}, \dots, u_{n+s-1}), \quad n=1, 2, \dots, 2^p-1 \quad (12)$$

これに $x_0 = \mathbf{0}$ を加えた $N=2^p$ 個の点を使用する。

漸化式 (11) を使用するための初期値 u_1, \dots, u_p の設定法の原理は次のとおりである。まず GF(2) 上の p 次未満の多項式 $f_0(x)$ を任意に選ぶ (たとえば $f_0(x)=1$)。次に表1の多項式の組 $M(x), g(x)$ を使って、多項式の列 $f_1(x), \dots, f_p(x)$ を次の漸化式により生成する。

$$f_n(x) = g(x)f_{n-1}(x) \pmod{M(x)} \quad (13)$$

そして、各 n について $f_n(x)/M(x)$ を形式的に $L_n(x) = \sum_{j=1}^{\infty} l_j^{(n)} x^{-j}$ という形のローラン級数に展開し、これから得られる2進小数 $L_n(2)$ を適当なビット数 (16ビット, 32ビットなど) で打ち切ったものを u_n とする。このように初期値の設定にはやや時間がかかるが、これが完了すると、あとは1つの x_n を生成するのに1回の \oplus 演算をすればよいので、きわめて高速に点列が生成できる。しかも、 N 個の点の集合 $S_N = \{x_0, x_1, \dots,$

$x_{N-1}\}$ は $f_0(x)$ の選び方によらないので、いちど u_1, \dots, u_p を計算して記録しておけば、以後はこれを使って即座に点列の生成を開始できるのである。

4.2 理論的背景

表1に示す p, q および $M(x), g(x)$ は次のようにして選ばれたものである。まず、GF(2) 上の p 次の任意の原始多項式 $M(x)$ と、 $p-1$ 次以下の任意の多項式 $g(x)$ を用いて上記のようにして生成される点列 S_N の s 次元“メリット数”を次式により定義する。

$$\rho^{(s)} = \min \sum_{k=1}^s \{\deg(h_k(x)) + 1\} \quad (14)$$

ここで \min は方程式

$$\sum_{k=1}^s g^{k-1}(x) h_k(x) = 0 \pmod{M(x)} \quad (15)$$

のすべての非零多項式解 $(h_1(x), \dots, h_s(x))$ についてとるものとし、また $\deg(0) = -1$ と定義する。そうすると、メリット数と差異の間には次の関係がある。

$$D_N^{(s)} = O((\log N)^{s-1/2^{\rho^{(s)}}}), \quad N=2^p$$

したがって差異を小さくするためにはメリット数を大きくすればよいことになる。

2次元の場合、メリット数は $g(x)/M(x)$ の連分展開

$$g(x)/M(x) = 1/(A_1(x) + 1/(A_2(x) + \dots + 1/A_r(x)))$$

の部分商 $A_1(x), \dots, A_r(x)$ と次の関係があることが知られている。

$$\rho^{(2)} = p+2 - \max_{1 \leq i \leq r} \deg(A_i(x)) \quad (16)$$

したがって、 p を固定した場合、部分商がすべて1次式 (x または $x+1$) となる多項式の対 (g, M) がメリット数を最小にする。このような多項式対をフィボナッチ多項式対と呼ぶ。(この命名は、フィボナッチ数列 $(1, 1, 2, 3, 5, 8, \dots)$ の連続する2項の比の正則連分展開の部分商がすべて1に等しいという事実との類似による)。フィボナッチ多項式対は次の漸化式により生成できる。

$$\begin{aligned} F_0(x) &= 1, \quad F_1(x) = A_1(x) \\ F_i(x) &= A_i(x)F_{i-1}(x) + F_{i-2}(x), \quad i \geq 2 \end{aligned} \quad (17)$$

ここに、 $A_i(x) (i \geq 1)$ は x または $x+1$ のいずれでもよい。フィボナッチ数列の生成に使う漸化式 ($F_i = F_{i-1} + F_{i-2}$) と違い、 $A_i(x)$ として2つのものを使えるので、(17)式によって生成されるのは多項式の二分木であり、フィボナッチ多項式対 $(F_{i-1}(x), F_i(x))$ は同じ次数のものが一般に多数あることに注意されたい。

一方、 p 次の原始多項式 $M(x)$ と、 $p-1$ 次以下の多

表 1 (18) 式を満たすフィボナッチ多項式対 $G(p, q) : (g, M)$

(数字は非零項のべき指数を表わす)

$G(15, 1) : M$	0 1 5 7 9 11 12 14 15
g	0 3 5 10 11 12 13 14
$G(17, 5) : M$	0 4 5 6 11 14 15 16 17
g	7 9 12 15 16
$G(18, 7) : M$	0 1 2 3 4 5 8 10 13 14 18
g	0 1 3 4 6 8 12 14 15 17
$G(20, 3) : M$	0 2 4 6 10 12 13 14 15 16 20
g	1 3 4 5 6 7 9 10 16 17 19
$G(22, 1) : M$	0 1 5 6 7 9 10 12 13 14 15 16 18 19 22
g	0 3 6 8 14 16 19 21
$G(23, 5) : M$	0 1 4 5 7 8 9 11 13 14 16 17 18 19 20 21 23
g	1 3 6 7 8 9 11 18 22
$G(25, 3) : M$	0 1 6 9 11 14 16 18 19 23 25
g	0 3 7 9 11 12 13 14 21 24
$G(28, 13) : M$	0 1 3 4 5 8 9 10 11 12 15 20 21 22 23 24 26 27 28
g	0 1 2 3 4 5 6 7 9 10 11 12 13 14 15 16 17 18 19 21 22 24 26 27
$G(31, 6) : M$	0 3 5 6 7 8 12 13 15 16 18 19 20 23 24 25 26 27 29 30 31
g	1 3 8 9 10 11 12 13 14 22 23 30

項式 $g(x)$ を用いて、漸化式 (13) によって生成される多項式の系列 $\{f_n(x)\}$ から、前記のように $f_n(x)/M(x)$ のローラン級数展開を基にして 2 進小数の系列 $\{u_n\}$ を作ったとき、これが漸化式 (11) を満たすための条件は

$$g^p(x) + g^q(x) + 1 = 0 \pmod{M(x)} \quad (18)$$

が成り立つことである。ただし $x^p + x^q + 1$ は GF(2) 上の原始 3 項式である。したがってフィボナッチ多項式対 $(F_{p-1}(x), F_p(x))$ の中で、 $g(x) = F_{p-1}(x), M(x)$

$= F_p(x)$ が (18) 式を満たすものを探せば、2 次元メリット数が最小の数列を高速に生成できることになる。表 1 は、 $10 \leq p \leq 31$ の範囲でこのような全数探索を行なった結果得られたものである。ただし、表に示してある各 p について複数の対が見つかったので、ある基準により選んだ 1 対ずつを挙げてある。

3 次元以上のメリット数については (16) 式のような簡単な表現は知られていないので、(14) および (15) 式にもとづいて計算する必要がある。表 2 は、このようにして得られたメリット数を示したものである。

表 2 $G(p, q)$ により生成される点列のメリット数

	$\rho^{(2)}$	$\rho^{(3)}$	$\rho^{(4)}$	$\rho^{(5)}$	$\rho^{(6)}$
$G(15, 1)$	16	12	11	7	7
$G(17, 5)$	18	14	12	11	7
$G(18, 7)$	19	14	13	12	11
$G(20, 3)$	21	14	14	12	12
$G(22, 1)$	23	17	17	15	13
$G(23, 5)$	24	16	15	15	15
$G(25, 3)$	26	20	19	17	15
$G(28, 13)$	29	24	23	18	18
$G(31, 6)$	32	24	24	22	20

表 3 QMC 法と MC 法の誤差の比較

(数値は相対誤差 $\times 10^6$ を示す)

方法 \ 関数	関数				
	①	②	③	④	⑤
$G(17, 5)$	21	40	29	3	7
線形合同法	10373	19816	16503	1432	4084
	2612	34696	3698	583	3101
	2763	34342	3610	500	2393
	853	8887	2510	115	369
	365	275	1577	201	463

4.3 数値計算例

本節で述べた方法の効果を確認するために、表1のG(17,5)を使って次の5つの5変数関数について計算を行なってみた。

- ① $\exp(-\sum_{k=1}^5 x_k)$
- ② $x_1 x_2 x_3 x_4 x_5 \exp(\sum_{k=1}^5 x_k^2)$
- ③ $\exp(-\sum_{k=1}^5 x_k) \sin(\sum_{k=1}^5 x_k)$
- ④ $\sqrt{1 + \sum_{k=1}^5 x_k}$
- ⑤ $1 + x_1 + 2x_1 x_2^2 + 3x_1 x_2^2 x_3^3 + 4x_1 x_2^2 x_3^3 x_4^4$
 $+ 5x_1 x_2^2 x_3^3 x_4^4 x_5^5$

表3は、相対誤差 $|(A_N - I)/I|$ の 10^6 倍を示したものである。比較のために、使用した計算機 (Sun-4) に内蔵されている線形合同法

$$X_n = 2736731631558 X_{n-1} + 138 \pmod{2^{48}}$$

による乱数生成関数 drand48() を使って得られた結果も示してある。後者については、乱数の初期値 X_0 を5通り変えて実験を行なっている。

表3に見られるとおり、本節で述べた方法による誤差は、通常のMC法による誤差よりはるかに小さい。MC法では、誤差を1桁小さくするのに100倍のサンプル数を必要とすることを思い起こせば、われわれの方法の有効性がさらにはっきりする。

5. おわりに

差異の小さい点列を用いる準モンテカルロ法を紹介し、特にわれわれが最近考察した点列の高速発生法とその効果について述べた。これは、別稿の著者手塚集氏との共同研究によるものであり、同氏の貢献に感謝する。

参考文献

[1] Andre, D. A., Mullen, G. L., and Niederreiter, H.: Figures of merit for digital multistep

pseudorandom numbers. *Mathematics of Computation*, Vol. 54 (1990), 737-748.

- [2] Bratley, P., and Fox, B. L.: Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Trans. Mathematical Software*, Vol. 14 (1988), 88-100.
- [3] Faure, H.: Discr pance de suites associ es   un syst me de num ration (en dimension s). *Acta Arithmetica*, Vol. 41 (1982), 337-351.
- [4] Mullen, G. L., and Niederreiter, H.: Optimal characteristic polynomials for digital multistep pseudorandom numbers. *Computing*, Vol. 39 (1987), 155-163.
- [5] Niederreiter, H.: Quasi-Monte Carlo methods and pseudo-random numbers. *Bulletin of the American Mathematical Society*, Vol. 84 (1978), 957-1041.
- [6] Niederreiter, H.: Point sets and sequences with small discrepancy. *Monatshefte f r Mathematik*, Vol. 104 (1987), 273-337.
- [7] Sobol', I. M.: The distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, Vol. 7 (1967), 86-112.
- [8] Tezuka, S.: On the discrepancy of GFSR pseudorandom numbers. *Journal of ACM*, Vol. 34 (1987), 939-949.
- [9] Tezuka, S., and Fushimi, M.: Calculation of Fibonacci polynomials for GFSR sequences with low discrepancies. Submitted.
- [10] 水田成仁: 差異の小さい点列の高速発生法に関する研究. 東京大学工学部計数工学科卒業論文, 1991.

