5. 一様乱数の発生法†

伏　見　正　則††

はじめに

本稿では，ソフトウェアによる一様乱数の発生法について述べる．アルゴリズムとしては，合同法とM系列によるものとが有名で，重要であると思われるが，前者については Knuth の本[11]に大変によい解説があるので，ここではごく簡単に述べるに止める．後者については，以前に解説を試みたことがある[5]が，その後の研究によって得られた結果に重点を置いて述べることにする．

1. 合同法など

計算機を使って，四則演算によって“でたらめのように見える数列”（擬似乱数列）を作り出そうという試みは，von Neumann によって始められたといわれている．1946年頃に彼が提案した**平方採中法**(middle-square method) は，“前回使った乱数を平方して中央の部分のけたを取り出したものを今回使う乱数とする”というものであったが，その後，この方法によって作り出される数列は短いサイクルに陥りやすいことが判明し，ほとんど使われなくなってしまった．平方採中法にとって代わったのが，1948年頃に **Lehmer** が提案した**線形合同法** (linear congruential method) であり，現在までにきわめて多数の人々によって研究され，利用されてきている．この方法では，法 $M=0$, 乗数 $\lambda>0$, 増分 $\mu\geq 0$ (いずれも整数) を適切に選び，適当な整数 x_0 から出発して，漸化式

$$x_t=\lambda x_{t-1}+\mu \pmod M,\ t\geq 1$$

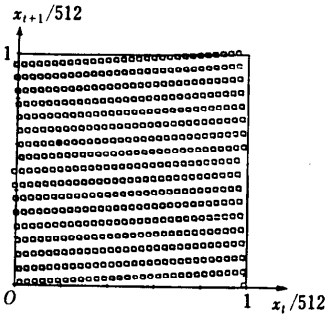
により数列 $\{x_t\}$ を生成する．$\mu=0$ の場合を**乗算合同法** (multiplicative congruential method), $\mu\neq 0$ の場合を**混合合同法** (mixed congruential method) と呼ぶならわしである．パラメータ M, λ, μ の選び方については，きわめて多くの研究がされてきているが，その結果は Knuth の本に詳しいので，本稿では割愛する．

合同法の長所は，アルゴリズムが簡単であること，そしてその割には多くの問題に対して実用上ほぼ満足できる乱数列が作れることが経験的に知られていることであろう．（もちろん，“パラメータを適切に選べば”という制限つきであるが．）短所としては，使用する計算機のけた数によって周期が制限されること，多次元分布が一様でないことなどが挙げられよう．後者の性質は Marsaglia[13),14)] によって指摘されたもので，**多次元の疎結晶構造**ともいわれ，次のように述べられる．(**図-1** 参照.)合同法によって生成される数列の相続く k 個の要素を座標成分とする点列，$\{(x_t, x_{t+1}, \ldots, x_{t+k-1})\}$ は，(λ と μ をどう選んでも) $(k!M)^{1/k}$ 枚以下の等間隔に並んだ平行な超平面の上に載ってしまう．この枚数は k が増すとともに減少し，たとえば $M=2^{32}$ のときには，$k=3$ で 2953, $k=6$ で 120, $k=10$ で 41 となる．これらの枚数は上界であり，パラメータの選び方によってはずっと少なくなってしまうことがあることに注意する必要がある．例えば $M=2^{32}$ または 2^{31} で，$\lambda=65539$, $\mu=0$ とすると，$k\geq 3$ ではわずか 15 枚の (超) 平面上にすべての点が載ってしまう．このパラメタが現在でもときどき使われているのは問題であろう．

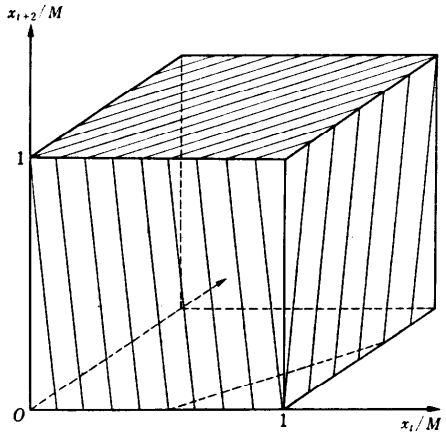
多次元空間内の点列が平行な超平面上に並んでしまうことは避けられないにしても，これらの超平面の間隔がある程度小さければ，実用上はさほど問題にならないであろうと考えられる．このような観点から，与えられた乗数 λ の良さを判定するためのアルゴリズムが初め Coveyou & MacPherson[3)] によって提案され，**スペクトル検定**と名づけられた．(増分 μ は，超平面の間隔に影響を及ぼさない.) このアルゴリズムは，後にいろいろな人々によって改良された．改良版については Knuth[11)] を見よ．スペクトル検定を使って，良い乗数 λ を各次元について試行錯誤によって探

† Algorithms for Generating Uniform Random Numbers by Masanori FUSHIMI (Department of Mathematical Engineering and Instrumentation Physics, Faculty of Engineering, University of Tokyo).

†† 東京大学工学部計数工学科



(a)　$x_t=17x_{t-1}+511 \pmod{512}$ によって生成される2次元の点列.



(b)　$x_t=65539\,x_{t-1} \pmod{M}$, $M=2^{31}$ または 2^{32}, とすると, $(x_t/M,\ x_{t+1}/M,\ x_{t+2}/M)$ を座標とする単位立方体内の点は, すべて15枚の平行な平面のいずれかに載ってしまう.

図-1　多次元疎結晶構造の例

すことができる. そのような探索の結果が, $2\leq k\leq 5$ について文献8) に, $2\leq k\leq 6$ について Knuth の本に載っている.

合同法乱数の多次元疎結晶構造が指摘されてから, 他の形の乱数発生法に対する関心が高まってきた. これらの方法の詳細については Knuth の本[11]を参照するとよいが, 次節で述べるM系列による方法以外については, 理論的な解明は十分にはなされていないことに注意しておこう.

2. M系列に基づく方法

ガロア体 GF(2) 上の p 次の原始多項式

$$f(z)=1+c_1z+c_2z^2+\ldots+z^p \qquad (2.1)$$

を特性多項式とする漸化式

$$a_t=c_1a_{t-1}+c_2a_{t-2}+\ldots+a_{t-p} \pmod 2 \qquad (2.2)$$

を任意の初期条件 $(a_0, a_1, \ldots, a_{p-1})\neq(0, 0, \ldots, 0)$ の下に解いて得られる数列 $\{a_t\}$ のことを**M系列**という. これは周期列であり, その周期は $T=2^p-1$ である*.

例　$f(z)=1+z^3+z^5$ は6次の原始多項式のひとつであり, これを特性多項式とする GF(2) 上の差分方程式は

$$a_t+a_{t-3}+a_{t-5}=0 \pmod 2$$

で, これは漸化式

$$a_t=a_{t-3}+a_{t-5} \pmod 2$$

と等価である. $a_0=a_1=a_2=a_3=a_4=0$ 以外の任意の初期条件を与えたとき, この漸化式によって生成される数列の1周期分は

{...1111100011011010100001001011 00...}

となる.

2.1 横型系列

Tausworthe[16] は, M系列の連続する $l(\leq p)$ 個の要素を並べて l ビットの2進小数

$$x_t=0.a_{\sigma t}a_{\sigma t+1}\ldots a_{\sigma t+l-1} \qquad (2.3)$$

を作り, 系列 $\{x_t\}$ を区間 $[0,1)$ 上の一様乱数列として使うことを提案した. 本稿では, この型の系列を仮に**横型系列**と呼ぶことにする. この系列は, $\sigma\geq l$ で σ と T が互いに素ならば, $1\leq k\leq\lfloor p/\sigma\rfloor$ の範囲の任意の k について k 次均等分布**をする (したがって多次元疎結晶構造を持たない) こと, 系列の自己相関関数も, 遅れが $(T-l)/\sigma$ 以下ならば理想的な一様乱数列のものとほとんど一致することが Tausworthe によって示された. また連の性質等が文献18), 19) で研究された.

2.2 縦型系列

横型系列を定義式 (2.3) に忠実に従って生成すると, かなり時間がかかる. そこで, 発生を高速化するという観点から, 同一のM系列の位相をずらしたものを各ビットに配する次の形の系列が提案され, 研究された. (文献12), 9), 1) など.)

$$y_t=0.a_{t+\tau_1}a_{t+\tau_2}\cdots a_{t+\tau_l} \qquad (2.4)$$

この形の系列を以後**縦型系列**と呼ぶことにする. なお $\tau_1=0$ とおいても一般性を失わないので, そうすることにする. 系列 $\{y_t\}$ が計算機で高速に発生できるの

---

* GF(2), 特性多項式, M系列の詳しい性質等については, 例えば次の文献を参照するとよい. S.W. Golomb: Shift Register Sequences, revised ed., Aegean Park Press, Laguna Hills, Cal. (1982). G. Hoffmann de Visme 著, 伊理正夫・伊理由美訳: 2値系列, 共立出版 (1977).

** $\{x_t\}$ の相続く k 個の要素を座標成分とする点を $\boldsymbol{x}_t$ と書くことにすると, k 次元単位立方体内の (座標成分が l ビットの2進小数で表わせる) 任意の点 $\boldsymbol{x}$ に対して $P(\boldsymbol{x}_t=\boldsymbol{x})\fallingdotseq 2^{-kl}$ が成り立つとき, 系列 $\{x_t\}$ は k 次均等分布をするという. ここで確率 P は1周期全体にわたる相対頻度を意味する.

は，その各ビットが同一の漸化式 (2.2) を満たし，したがって2進整数の系列 $\{Y_i\}=\{2^l y_i\}$ が漸化式

$$Y_i=c_1Y_{i-1}\oplus c_2Y_{i-2}\oplus\cdots\oplus Y_{i-p} \qquad (2.5)$$

によって生成できるからである．ここに⊕は排他的論理和 (exclusive or) を表わし，多くの計算機できわめて高速に処理できる演算である．縦型系列の中では，特性多項式 $f(z)$ の項数が少ないものほど発生速度が速く，3項式

$$f(z)=1+z^q+z^p \quad (q<p) \qquad (2.6)$$

を選ぶともっとも速くなるので，そのように選んだ発生法が **Lewis & Payne**[12] によって提案され，また Arvillias & Maritsas[1], Payne & McMillen[15] 等によって研究された．

縦型系列を生成するアルゴリズムを設計する際に注意すべきことは，①位相 $\tau_i(2\leq i\leq l)$ をいかにして適切に選ぶか，②M系列の位相が隔たった部分の要素（遠隔項）をいかにして速く発生するかである．$\{y_i\}$ のビット間の位相差の最小値を

$$s_0=\min_{1\leq i\neq j\leq l}|\tau_i-\tau_j| \qquad (2.7)$$

とするとき，$\{y_i\}$ の理論的な自己相関関数

$$R(s)=\frac{1}{T}\sum_{i=1}^{T}(y_i-\bar{y})(y_{i+s}-\bar{y})$$

の値は，$1\leq s<s_0$ ならばほぼ0で，$s=s_0$ では大きくなる．したがって $\tau_i(2\leq i\leq l)$ は，s_0 が実際に使用する乱数の個数より大きくなるように選ぶことが望ましい．ところが Lewis & Payne のアルゴリズムでは，遠隔項は基本的には漸化式 (2.2) を繰り返し使用して求めるという方法を用いているので，所要時間の制約を考えると s_0 はあまり大きくできない．そこで彼等のプログラムでは $\tau_i=100(i-1)p$ と設定していて，したがって $s_0=100p$ となるが，これでは必ずしも十分な大きさとは言えない*．そこでその後，漸化式 (2.2) を多数回使用することなしに十分離れた遠隔項を求める方法が研究された．なかでも **Arvillias & Maritsas**[1] は，特性多項式が3項式 (2.6) で，q が2のべき乗の場合には，$l\leq q$ の範囲内ならば，$s_0\fallingdotseq T/q$ となるような $\{y_i\}$ の初期値が高速に設定できることを示した．

はなはだ奇妙なことであるが，Lewis & Payne 以後の縦型系列に関する研究では，**多次元分布の均等性**に関する条件が考慮されなかった．実際，Lewis & Payne のプログラムによって発生される系列の多次元分布は必ずしも均等でないし，Arvillias & Maritsas が例示した“realistic generator”によって生成される系列の2次元以上の分布は均等分布とは著しくかけ離れている[17]．そこで Fushimi & Tezuka[4] は，縦型系列が多次元の均等分布をするための条件を調べ，次の必要十分条件を求めた．

定理 1　縦型系列 $\{y_i\}$ が k 次均等分布をするための必要十分条件は，この系列の相続く k 個の要素を構成するM系列の要素 kl 個が1次独立であることである．

この定理に基づいて，われわれは，①多次元分布の均等性が保証された系列を発生する方法，および②任意に与えられた縦型系列発生アルゴリズムによって生成される系列の多次元分布が均等であるかどうかを理論的に検定する方法を提案した[4]．これらはきわめて一般的な理論であるが，実用上は $\{y_i\}$ の相隣るビット間の位相差がすべて一定，すなわち

$$\tau_i-\tau_{i-1}=\tau,\ 2\leq i\leq l \qquad (2.8)$$

という制約を課しても特に問題はないであろうと考えられる．そこで，以下ではこの制約を課した縦型系列のみを考察の対象とする．

2.3 縦型系列と横型系列の同値関係

横型系列および縦型系列を構成するM系列の特性多項式およびパラメータを明示するために，以後必要に応じて $\{x_i(f;\sigma)\}$, $\{y_i(f;\tau)\}$ と書くことにする．M系列 $\{a_i\}$ の要素を τ 番目ごとに系統的にサンプリングして得られる系列 $\{a_{\tau t};\ t=0,1,2,\ldots\}$ は，τ が周期 T と互いに素ならばやはり周期 T のM系列であることが知られている．この系列の特性多項式を f_τ と書くことにする．また，二つの系列が位相をずらせば一致するという意味で同値であることを記号≃を使って表わすことにする．次の定理が成り立つ[6]．

定理 2　σ,τ が T と互いに素ならば

$$\{x_i(f;\sigma)\}\simeq\{y_i(f_\sigma;\sigma^{-1})\},$$
$$\{y_i(f;\tau)\}\simeq\{x_i(f_\tau;\tau^{-1})\}.$$

である．ここに，σ^{-1},τ^{-1} は，T を法とする乗算に関する，σ,τ の逆元を表わす．

この定理により，$\tau(=s_0)$ がきわめて大きく，かつ多次元の均等分布が保証された縦型系列の初期値を高速に設定する方法が得られる．まず，乱数のビット数 l が2のべき乗である場合について述べる．$\tau=2^r/l$ とおく．そうすると τ も2のベキ乗であり，M系列の基

* 津田[19]は，$p=89$, $q=51$, $l=15$ として Lewis & Payne の方法によって発生した10万個の乱数について統計的検定を行ったところ，ビットの独立性に関する検定の成績が大変に悪かったと報告しているが，これは使用した乱数の個数10万が $s_0=8900$ をはるかに超えているところに原因があるものと考えられる．

本的性質により $f_t=f$ となる．また，$\tau^{-1}=l$ である．したがって定理2により

$$\{y_t(f;\tau)\}\simeq\{x_t(f;l)\}$$

が成り立ち，M系列の最初の lp 個の要素を（$\sigma=l$ として）横型に並べれば，縦型系列の初期値の設定が完了することになる．この方法では，Lewis & Payne の方法と異なり，"無駄な" M系列の要素は一切計算しない．

l が2のべき乗でない場合には，いくつかの方法が考えられるが，次のようにするのが手軽でよいであろう．l より大きい最小の2のべき乗数を 2^d とする．$\sigma=2^d$ として，M系列の要素を横型に並べる．（これは，2^d ビットの乱数列の初期値を前記の方法で設定し，その先頭の l ビットだけを使用することに相当する．）

2.4 原始3項式を用いる場合の例

以上の議論は，どのような原始多項式を用いる場合にもあてはまるが，実際には乱数の発生速度を速くするために3項式 (2.6) が使われることが多い．そこで以下この場合について述べる．

1000次以下のすべての原始3項式が文献21), 22)に載っている．この中から，使用する計算機のビット数，発生させたい点列の次元，等を考慮してその都度適切なものを選んでもよいが，"大は小を兼ねる" という意味で一般的に使えそうなものをあらかじめ用意しておくとすれば，①次数 p が十分に大きく，②周期 T が素数である，ものが無難である．このように選んだ場合の縦型系列は，$\lfloor p/l \rfloor$ 以下の任意の次元で均等分布をする．上記の基準にかなうものとしては，

$$z^{521}+z^{32}+1,\quad z^{607}+z^{273}+1$$

などがある．ここでは前者を取り上げ，$l=32$ ビットの縦型系列の初期値設定手順の具体例を示そう．

1° M系列の初期値 $a_0, a_1, \cdots, a_{520}$ を任意に与える．（ただし，すべてが0ではないように．）

2° 漸 化 式

$$a_t=a_{t-521}+a_{t-32}\ (\mathrm{mod}\ 2)$$

を用いて $a_t(521\leq t\leq 16671)$ を求める．

3° $0\leq t\leq 520$ について

$$Y_t=a_{32t}a_{32t+1}\cdots a_{32t+31}\ （2進整数）$$

とする．

この方針に基づく FORTRAN のプログラムが文献7) に載っている．（ただし，1語＝32ビットで，負の整数は補数表示をするタイプの計算機の使用を前提にしているので，実際には上記の Y_t の最下位ビットを切り落とした系列を生成している．）

上記の手順は1ビットずつ処理してもよいが，論理演算を用いて 32 ビットずつ並列に処理して高速化することもできる．この場合，手順は次のようにかわる．

1° 32ビットの2進整数 $Y_t(0\leq t\leq 16)$ を任意に与える．（ただし，すべてが0ではないように．）

2° Y_{16} を次式により更新する．

$$Y_{16}=M^{32}((L^{23}Y_{16}+R^{9}Y_0)\oplus Y_{15})$$

3° 漸 化 式

$$Y_t=M^{32}((L^{23}Y_{t-17}+R^{9}Y_{t-16})\oplus Y_{t-1})$$

を用いて $Y_t(17\leq t\leq 520)$ を求める．

この手順中の＋は論理和，⊕は排他的論理和，L^{23} は23ビット左論理シフト，R^9 は9ビット右論理シフト，M^{32} は下位32ビットのみを取り出すマスク演算を表わす．（1語＝32ビットの計算機を用いるなら，マスク演算はもちろん不要である．）

上記の2種類のアルゴリズムのいずれを用いるにしても，手順 1° で初期値を与える部分は，例えば合同法乱数などを使ってランダムに設定するのがよい．0あるいは1が極端に多いビット列を与えると，かなり長い間性質のよくない数が生成されることが多いことが経験的に知られている．

3. おわりに

乱数に関する最近の研究としては，本稿で述べたものの他に，さまざまな分布に従う乱数の効率的な発生法，カオスと乱数の関係，差異 (discrepancy) の小さい乱数（準乱数），等々があり，当初の予定ではこれらの話題についても述べるつもりであったが，すでに予定の紙数に達してしまったので，割愛せざるをえない．これらについては，それぞれ研究を進めておられる方々が，別の機会に解説して下さることを期待したい．

参 考 文 献

1) Arvillias, A.C. and Maritsas, D.G.: Partitioning the period of *m*-sequences and application to pseudorandom number generation, *J. ACM*, Vol. 25, pp. 675-686 (1978).
2) Bright, H. and Enison, R.: Quasi-random number sequences from a long TLP generator with remarks on application to cryptography, *Computing Surveys*, Vol. 11, pp. 357-370 (1979).
3) Coveyou, R.R. and Macpherson, R.D.: Fourier

analysis of uniform random number generators, *J. ACM*, Vol. 14, pp. 100-119 (1967).
4) Fushimi, M. and Tezuka, S.: The *k*-distribution of the generalized feedback shift register pseudorandom numbers, to appear in *Comm. ACM*.
5) 伏見正則：擬似乱数の発生法について，情報処理，Vol. 21, No. 9, pp. 968-974 (1980).
6) 伏見正則：M系列に基づく乱数発生法に関する相反定理とその応用，本学会論文誌に掲載予定.
7) 伏見正則，手塚 集：多次元分布が一様な擬似乱数列の生成法，応用統計学，Vol. 10, pp. 151-163 (1981).
8) 原田紀夫：スペクトル検定に対する乱数列発生法の最適係数，情報処理，Vol. 15, No. 3, pp. 180-188 (1974).
9) Hurd, W. J.: Efficient generation of statistically good pseudonoise by linearly interconnected shift registers, *IEEE Trans. Computers*, Vol. C-23, pp. 146-152 (1974).
10) 泉 照之，柏木 潤：2値乱数源用高次M系列の初期値，計測自動制御学会論文集，Vol. 18, pp. 929-935 (1982).
11) Knuth, D. E.: *The Art of Computer Programming*, Vol. 2: *Seminumerical Algorithms* 2nd ed., Addison-Wesley, Reading, Mass. (1981)；渋谷政昭（訳）：準数値算法／乱数，サイエンス社 (1981).
12) Lewis, T. G. and Payne, W. H.: Generalized feedback shift register pseudorandom number algorithms, *J. ACM*, Vol. 21, pp. 456-468 (1973).
13) Marsaglia, G.: Random numbers fall mainly on the planes, *Proc. Nat. Acad. Sci.*, Vol. 61, pp. 25-28 (1968).
14) Marsaglia, G.: The structure of linear congruential sequences, in Zaremba, S. K. (Ed.): *Applications of Number Theory to Numerical Analysis*, Academic Press, New York (1972).
15) Payne, W. H. and McMillen, K. L.: Orderly enumeration of nonsingular binary matrices applied to text encryption, *Comm. ACM*, Vol. 21, pp. 259-263 (1978).
16) Tausworthe, R. C.: Random numbers generated by linear recurrence modulo two, *Mathematics of Computation*, Vol. 19, pp. 201-209 (1965).
17) 手塚 集，伏見正則：最大周期列を用いたある種の一様乱数発生法の欠陥とその改善，応用統計シンポジウム（第2回）講演予稿集 (1980).
18) Tootill, J. P. R., Robinson, W. D. and Adams, A. G.: The runs up-and-down performance of Tausworthe pseudo-random number generators, *J. ACM*, Vol. 18, pp. 381-399 (1971).
19) Tootill, J. P. R., Robinson, W. D. and Eagle, D. J.: An asymptotically random Tausworthe sequence, *J. ACM*, Vol. 20, pp. 469-481 (1973).
20) 津田孝夫：レーマー型合同法によらない乱数について，**bit**, Vol. 12, pp. 1180-1191 (1980).
21) Zierler, N. and Brillhart, J.: On primitive trinomials (mod 2), *Inf. Control*, Vol. 13, pp. 541-554 (1968).
22) Zierler, N. and Brillhart, J.: On primitive trinomials (mod 2) II, *Inf. Control*, Vol. 14, pp. 556-569 (1969).

（昭和57年12月3日受付）