M 系列を用いる乱数発生

原 田 博 之*・柏 木 潤*

Random Number Generation by Use of M-Sequence

Hiroshi HARADA* and Hiroshi KASHIWAGI*

This paper proposes a new method of random number generation using a high-order M-sequence. This method uses Tausworthe's method for initializing seed matrix and uses Lewis & Payne's method for generation random numbers. Therefore, generated numbers are uniformly distributed in a high dimensional space and generation time becomes shorter.

When we use a high-order M-sequence for generating random numbers, selection of an initial n -tuple for generating M-sequence is important. If an unsuitable n -tuple is used, generated numbers are not random for a while. In order to solve this problem, in case of this method, initial n -tuples are chosen under the consideration of the partial properties of M-sequence.

Several statistical tests are applied to the random numbers generated by this method and it is shown that the numbers have good random properties.

Key Words: random number generation, M-sequence, partial properties of M-sequence

1. ま え が き

擬似乱数は、計算機シミュレーションにおける白色雑音として、あるいはモンテカルロ法などにさかんに用いられており、擬似乱数発生プログラムは計算機のライブラリプログラムとして最もよく利用されるものの一つである。現在、擬似乱数発生法として主に使用されている方法は Lehmer によって考案された合同法であり、(1)のような漸近式で乱数列 $\{X_i\}$ ($i=1, 2, \dots$)

$$X_{i+1} = \lambda X_i + \mu \pmod{p} \quad (1)$$

(λ, μ, p : 正の整数)

を発生する。しかし、この Lehmer 型合同法には

(1) λ や μ をどのように選んでも発生した擬似乱数の多次元分布が一様にならない¹⁾。

(2) 使用する計算機の語長を L とすると、擬似乱数の周期は最大でも 2^L にしかならないため、語長の短い計算機では周期が短すぎる。

などの欠点がある。このうち(2)については、倍精度整数を用いることにより解決できるが、(1)は合同法のもつ本質的欠点であるため、解決不可能である。

これに対して、高次M系列を用いる擬似乱数発生法は上記のような欠点をもたないため近年盛んに研究されている^{2)~4)}。しかし、これらの中にはM系列を発生するための初期値の選択が不適当なため、良好な擬似乱数が得られないものも見られる²⁾。本研究で述べる擬似乱数発生法はこのような事態を避けるため、M系列の部分的性質を考慮して、不規則性の良い初期値を選んだものである。

2. M 系列を用いる擬似乱数発生方法

M系列を用いる擬似乱数発生法として従来知られているものとしては、Tausworthe法⁵⁾、Lewis & Payne法⁶⁾の二つがある。本研究で述べる方法は、M系列のサンプリングに関する性質を利用し、両者の長所を取り入れたものである。

2.1 Tausworthe 法

この方法は(2)式のようにM系列 $\{a_i\}$ ($a_i=0$ または 1)の相続く L ビットを取り出し、2進の小数とした系列 W_i (Tausworthe 系列と呼ぶ)を擬似乱数とするものである。

$$W_i = 0.a_{iq+1}a_{iq+2}\dots a_{iq+L} \quad (\text{base } 2) \quad (2)$$

$$q \geq L \text{ かつ } (q, 2^n - 1) = 1$$

この Tausworthe 系列の周期は $2^n - 1$ (n : M系列の次数)となり、一周期についての統計的性質は理論的に明らかである。またM系列の次数 n を L に比べて十分大きく選べば、 $[n/L]$ 次元空間において均等分布す

† 第25回計測自動制御学会学術講演会で発表(昭61・7)

* 熊本大学工学部 熊本市黒髪 2-39-1

* Faculty of Engineering, Kumamoto University, Kumamoto

(Received January 24, 1987)

ることが知られている⁴⁾。しかし、この方法ではM系列をLビット発生して擬似乱数が1個得られるというように擬似乱数発生の速度が遅いのが欠点である。

2.2 Lewis & Payne 法

この方法では、M系列 $\{a_i\}$ を2進の小数列 W_i ($0 \leq i \leq n-1$) の第1ビットめに配置し、つぎに W_i の第2ビットめには $\{a_i\}$ より位相差 d だけ遅れたM系列 $\{a_{d+i}\}$ を並べていく。同様の操作をLビットまで繰返したとき数列 W_i (以下、初期行列と呼ぶ) は(3)式のようになる。

$$W_i = 0, a_i a_{d+i} \cdots a_{(L-1)d+i} \quad (\text{base } 2) \quad (3)$$

ここで、 W_i の各列間の位相差 d を列間遅れと称する。以上の方法で初期行列を設定した後、M系列の特性多項式に従って語同士の排他的論理和を取ることによって得られる系列を擬似乱数として用いるのが、Lewis & Payne 法である。この方法は擬似乱数発生に要する時間が短いという点で Tausworthe 法より優れているが、多次元分布の一様性は保証されていない。

2.3 本研究で用いた擬似乱数発生法

本研究で用いた擬似乱数発生法においても Lewis & Payne と同様に初期行列を設定するが、設定方法は以下のように異なる。この方法では、まず、M系列 $\{a_i\}$ をL個発生し、初期行列の第1語めの最上位ビットから最下位ビットへ順番に配置していく。この操作を第n語めまで行ったとき、初期行列 W_i ($0 \leq i \leq n-1$) は(4)式のように表わされる。

$$W_i = 0, a_{Li} a_{Li+1} \cdots a_{L(i+1)-1} \quad (\text{base } 2) \quad (4)$$

このとき、初期行列のビット長 L を $L=2^l$ にとると、初期行列 W_i の第jビットめを取り出した系列 $\{b_j\}$ は、

$$\{b_j\} = a_j, a_{2^l+j}, a_{2 \cdot 2^l+j}, \cdots, a_{(n-1) \cdot 2^l+j}$$

となり、もとのM系列 $\{a_i\}$ を 2^l 個ごとにサンプルした系列となる。したがって、 $\{b_j\}$ はもとのM系列と同じ特性多項式をもつM系列となる⁷⁾。またこのとき、初期行列の列間遅れ d はビット長が L であるため、

$$L \cdot d = 1 \pmod{2^n - 1} = 2^n$$

$$\therefore d = 2^n / L = 2^{n-l} \quad (5)$$

となる⁸⁾。

以上の結果より、上記の方法によって設定した初期行列は Lewis & Payne 法において列間遅れを(5)式のようにとったものと同じものとなる。さらに、この初期行列をもとに特性多項式に従って語同士の排他的論理和をとって得られる数列 W_k ($k \geq n$) は、

$$W_k = 0, a_{k \cdot 2^l} a_{k \cdot 2^l + 1} \cdots a_{(k+1) \cdot 2^l - 1} \quad (\text{base } 2)$$

となり、(2)式において、 $q=2^l$ とした場合と等しくなる。すなわち、この方法で発生した擬似乱数が

Tausworthe 系列となることがわかる。

このように、本研究で用いた擬似乱数発生法は、初期行列設定を Tausworthe 法で行い、Lewis & Payne 法によって擬似乱数列を発生するが、得られる数列は Tausworthe 系列と同じである。したがって、擬似乱数発生に要する時間は短く、得られる擬似乱数も多次元空間における一様分布が保証されたものとなる。

3. M系列発生に用いる初期nタップルの選定

前章で述べた方法で擬似乱数を発生する際に問題となるのは、M系列を発生するための初期値(以下、初期nタップルと呼ぶ)の選定である。これは、M系列の次数が高くなるほどその周期が長くなり、擬似乱数発生に用いられるのがM系列の1周期のうちのわずかな部分となるためである。このため、M系列の部分的性質(部分的自己相関関数など)が悪い位相点のnタップル、たとえば特性M系列の初期nタップルなどを用いた場合、得られる擬似乱数はきわめて不規則性の悪いものとなる⁹⁾。さらに、このような不規則性の悪いnタップルを位相原点としたとき、その $1/2^i$ 位相点 (i : 正の整数) のnタップルにおいても同様な性質がみられる⁹⁾。したがって、M系列発生のための初期nタップルの選定に際しては、不規則性の悪いnタップルだけでなく、それらの $1/2^i$ 位相点(円分位相点⁹⁾と呼ぶ)も避ける必要がある。

従来のM系列を用いる擬似乱数発生法では不規則性の悪いnタップルを避けるために、 $\log_2 n$ 次のM系列を用いたり¹⁰⁾、合同法乱数を用いたりして⁴⁾、初期nタップルを設定した。しかし、これらの方法で選んだ初期nタップルはM系列のどの位相点のnタップルに相当するのかわからない。このため、選んだ初期nタップルのすぐ後に不規則性の悪い位相点が存在する可能性がある。

そこで本研究では、あらかじめ不規則性のよい位相点のnタップルをいくつか選んでおき、擬似乱数発生プログラム(以下、乱数パッケージと呼ぶ)に入力した値に応じてそれらの中から一つを選び出し、M系列発生の初期nタップルとするという方法を用いた。この方法では、選んだ初期nタップルの位相が既知であるため、不規則性の悪いnタップルとの位相差も容易に求めることができる。

4. 試作した乱数パッケージ

本研究で試作した乱数パッケージは以下のようなものである。なお、ここで用いた計算機は1語16ビットのものである。

Table 1 Statistical tests

Test	Frequency	Serial (2nd)	Serial (3rd)	Combination	Run (above/below)	Run (up/down)	Poker
Degrees of Freedom	15	15	63	10	7	7	3

まず、M系列の特性多項式としては(6)式のような次数がメルセンヌ素数（その数 n および 2^n-1 がともに素数である整数）である89次の原始5項式を用いた。

$$f(x) = x^{89} + x^{32} + x^{12} + x + 1 \quad (6)$$

ここで特性多項式として原始5項式を使用したのは、原始3項式によって発生した擬似乱数の場合、それらを加算して得られる正規乱数が歪の大きい分布を示すからである⁷⁾。

つぎに、今回試作した乱数パッケージに用いた初期 n タプルは、以下の方法により特性M系列⁷⁾の j/k 位相点（ $k=3, 5, 7, 11, 13, 17, 19$; $1 \leq j \leq k-1$ ）68個の n タプルの中より選び出した。

- (1) 各位相点の n タプルを初期値としてM系列を発生し、2.3節で述べた方法により擬似乱数を発生する。
- (2) Table 1 に示す七つの検定項目¹¹⁾について、発生した擬似乱数が各検定項目の理論的分布（真に不規則な乱数の場合の分布）に従うかどうか χ^2 検定を行い χ^2 値を得る。この場合、 χ^2 値を求める際のサンプル数（乱数の数、乱数の組の数、連の数）は1,000とする。
- (3) 上記の χ^2 検定を連続して100回行い、得られた100個の χ^2 値が χ^2 分布からのランダムなサンプルとみなせるかどうかを判断するためKolmogorov-Smirnov検定（K-S検定）¹¹⁾を行う。

$$K_{100}^+ = 10 \max(F_{100}(x) - F(x)) \quad (7)$$

$$K_{100}^- = 10 \max(F(x) - F_{100}(x)) \quad (8)$$

を得る。ここで $F_{100}(x)$ は得られた χ^2 値の分布関数であり、 $F(x)$ は χ^2 分布の理論的分布関数である。

- (4) (1)~(3)を連続して20回行う。これにより、一つの検定項目につき2,000,000~10,000,000個の乱数を発生したことになる。
- (5) 初期 n タプルを位相差 q だけ遅らせて（以下この遅れ q をオフセットと呼ぶ）(1)~(4)を繰り返す。これを $0 \leq q \leq 15$ について行う。試作した乱数パッケージの場合、語長が 2^4 ビット、M系列の次数が89であるため、オフセットを q 加えることは(5)式より、初期行列 W_i

Table 2 Selected phases

2	1	1	2	3	4	1	3
3	5	7	7	7	7	11	11
9	10	1	4	5	12	3	7
11	11	13	13	13	13	17	17
8	9	10	11	13	1	3	4
17	17	17	17	17	19	19	19
5	12	13	14	15	16	17	18
19	19	19	19	19	19	19	19

(a) Division of initial value

(b) Algorithm of initializing seed matrix

Fig. 1 Initialization of seed matrix

の各列に $d=q \cdot 2^{85}$ の遅れを加えたことになる。

- (6) (1)~(5)を68個に位相点について行い、得られた K_{100}^+ , K_{100}^- の値（各位相点の一つの検定項目につき320個ずつ）のうち有意水準5%および1%で棄却されるものの数 r を求め、各検定項目における r の最大値が小さなものから順に32個の位相点を選ぶ。

以上の方法により選んだ位相点をTable 2に示す。

Fig. 2 Random number generator

```

C      MONTE CARLO SIMULATION
      DOUBLE INTEGER I
      DATA PT/3.141592/
      READ(5,*) IX
      CALL INTR(IX)
      PI=0.0
      DO 100 I=1,100000
        CALL UNIFR(X)
        CALL UNIFR(Y)
        R=X*X+Y*Y
        IF(R.LE.1.0) PI=PI+1.0
100    CONTINUE
      PI=PI/100000.0*4.0
      E=(PI-PT)/PT*100.0
      WRITE(6,1000) IX,PI,E
1000  FORMAT(' ',IX=' ',I7.3X,
*PI=' ',F7.4,3X,'E=' ',F6.2,'%')
      END
    
```

(a) Program

```

IX=      0      PI= 3.1404      E= -0.04%
IX=    767      PI= 3.1456      E=  0.13%
IX=   1985      PI= 3.1470      E=  0.17%
IX=     -1      PI= 3.1398      E= -0.06%
IX=  -12345      PI= 3.1380      E= -0.11%
    
```

(b) Results

Fig. 3 Monte Carlo simulation using proposed random package

試作した乱数パッケージは初期行列設定ルーチンと一様乱数発生ルーチンの二つの部分より構成される。このうち初期行列設定ルーチンでは、入力された初期値 IX を Fig. 1(a) のように分割し、Fig. 1(b) のようなアルゴリズムで位相点およびオフセットを選び、2.3節で述べた方法により初期行列を設定する。一方、一様乱数発生ルーチンでは Fig. 2 のように (6) の特性多項式に従って語同士の排他的論理和をとることにより、擬似乱数を発生する。

つぎに、試作した乱数パッケージの使用例を Fig. 3 に示す。Fig. 3(a) は、モンテカルロ法により円の面

積を計算し円周率 π を求めるプログラムである。プログラム中の $INTR(IX)$ は初期行列設定のサブルーチンであり、引数 IX を通じて与えられた初期値を用いて Fig. 1 のアルゴリズムにしたがって初期行列を設定する。一方、 $UNIFR(-)$ は一様乱数発生ルーチンであり、このサブルーチンと呼ぶたびに区間 $[0, 1)$ において一様分布する乱数一つ得られる。このプログラムを初期値 IX を 5 回変えて実行した結果を Fig. 3(b) に示すが、いずれの場合も誤差は $\pm 0.2\%$ 以下であった。

つぎに、この乱数パッケージの検定を以下のように行う。ここで、検定項目および自由度は Table 1 に示したものと同一とする。

- (1) 初期値 IX を与えて初期行列を設定し、擬似乱数を発生する。
- (2) 各検定項目についてサンプル数 1,000 個として χ^2 検定を行い χ^2 値を 1 個得る。
- (3) (2) を連続して 100 回行い、得られた 100 個の χ^2 値より (7), (8) 式に従って K_{100}^+ , K_{100}^- の値を求めそれらが有意水準 5% および 1% で棄却されるかどうかを判定する。

初期値 IX をランダムに 2,500 個選び (1) ~ (3) を行った場合、有意水準 5% および 1% で棄却される K_{100}^+ , K_{100}^- の割合を Table 3 に示す。これより、各検定項目について、良好な結果が得られていることがわかる。

5. 他の乱数発生法との比較

本章では、試作した乱数パッケージと他の乱数発生法との比較を行った結果を示す。比較の対象とした擬似乱数発生法の漸化式および周期は Table 4 のとおりである。

つぎに、これらの乱数発生法によって発生した擬似乱数の検定を行い、その結果を比較する。ここで、検定項目および自由度は Table 1 と同じものとし、検定方法としては試作した乱数パッケージの場合と同様に K-S 検定を 100 回行い得られた K_{100}^+ , K_{100}^- の値のうち有意水準 5% および 1% で棄却されるものの割合を求めるという方法を用いた。Table 5 に有意水準を 5% としたときの結果を示す。この表で試作した乱

Table 3 Results of K-S test

Test	Frequency	Serial (2nd)	Serial (3rd)	Combination	Run (above/below)	Run (up/down)	Poker
Level of significance							
5%	4.8	3.7	5.2	4.0	4.4	5.8	3.9
1%	0.9	0.6	1.0	0.6	1.0	1.6	0.8

Table 4 Methods of generating pseudo-random numbers

Method	Recurrence equation	Period of pseudo-random numbers
Proposed random package	$X_i = X_{i-88} \oplus X_{i-88} \oplus X_{i-77} \oplus X_{i-57}$	$2^{89} - 1$
Linear congruential (I) ¹²⁾	$X_i = 65539 \cdot X_{i-1} \pmod{2^{31}}$	$2^{31} - 1$
Linear congruential (II) ¹²⁾	$X_i = 16807 \cdot X_{i-1} \pmod{2^{31}-1}$	$2^{31} - 2$
Linear congruential (III) ¹²⁾	$X_i = 630360016 \cdot X_{i-1} \pmod{2^{31}-1}$	$2^{31} - 2$
Lagged-Fibonacci using + ¹³⁾	$X_i = X_{i-17} + X_{i-5} \pmod{2^{16}}$	$(2^{17}-1) \cdot 2^{14}$
Lagged-Fibonacci using - ¹³⁾	$X_i = X_{i-17} - X_{i-5} \pmod{2^{16}}$	$(2^{17}-1) \cdot 2^{14}$
Lagged-Fibonacci using $\cdot$ ¹³⁾	$X_i = X_{i-17} \cdot X_{i-5} \pmod{2^{16}}$	$(2^{17}-1) \cdot 2^{28}$
Middle-square ¹¹⁾	$X_i = \text{central 16 bits of } X_{i-1}^2$	unknown
Quadratic congruential ¹¹⁾	$X_i = X_{i-1} \cdot (X_{i-1} + 1) \pmod{2^{15}}$ $X_0 = 2 \pmod{4}$	unknown
Random number generator suggested for HP-25, HP-25C ¹⁴⁾	$X_i = \text{fractional part of } (\pi + X_{i-1})^5$	unknown
Uniform random numerator ¹⁵⁾	$X_i = \begin{cases} X_{i-1} + X_{i-2} + X_{i-3} + 1357 & (X_{i-2} < 5 \cdot 10^7) \\ X_{i-1} + X_{i-2} + X_{i-3} & (X_{i-2} \geq 5 \cdot 10^7) \end{cases}$	$> 10^{12}$

Table 5 Results of K-S test applied to random numbers generated by proposed random package and other methods. The numbers in the table show the rejected number among 100 tests under the significance level of 5% (#: all rejected)

Method \ Test	Frequency	Serial (2nd)	Serial (3rd)	Combination	Run (above/below)	Run (up/down)	Poker
Proposed random package	7	3	5	5	5	5	1
Linear congruential (I)	4	4	3	5	1	11	3
Linear congruential (II)	7	6	2	6	9	8	3
Linear congruential (III)	1	5	6	5	4	6	3
Lagged-Fibonacci using +	8	1	4	7	3	3	4
Lagged-Fibonacci using -	4	4	3	5	5	10	5
Lagged-Fibonacci using $\cdot$	3	4	6	10	1	2	3
Middle-square	#	#	#	#	#	#	#
Quadratic congruential	#	#	#	#	#	#	#
Random number generator suggested for HP-25, HP-25C	#	#	#	#	#	#	#
Uniform random numerator	6	36	8	4	#	#	#

数パッケージの検定結果は Table 3 に示した 2,500 回の検定結果のうちの最初の 100 回の部分である。この結果より、試作した乱数パッケージにより発生した擬似乱数はすべての検定項目について良好な結果が得られていることがわかる。一方、Lehmer 型合同法では乗数をうまく選んだ場合には特性のよい乱数が得られる ($\lambda=630,360,016$ の場合) が、乗数が不適当だと ($\lambda=65,539, \lambda=16,807$ の場合)、連の検定結果が悪い。これは、1 章で述べた合同法乱数の多次元空間における不均一な分布の影響と考えられる。また、Lagged-Fibonacci 法¹³⁾により発生した乱数では、算術演算として加算を用いた場合はほぼ良好な検定結果が得られているが、減算および乗算を用いた場合はそれぞれ、連 (up/down) の検定、組合せ検定の検定結果が悪い。さらに、その他の方法は乱数発生法としては不適当であると考えられる。

6. あとがき

M 系列の部分的性質を考慮した乱数パッケージを試作し、良好な検定結果が得られた。今回擬似乱数発生に用いた方法は初期行列設定に Tausworthe 法を用い、擬似乱数発生に Lewis & Payne 法を用いたものであるため、多次元分布における一様性が明らかであり、擬似乱数発生に要する時間も短いという利点がある。また、この乱数パッケージに使用した言語は FORTRAN であるため、他の計算機にも容易に移植することができる。

参考文献

- 1) G. Marsaglia: Random Number Fall Mainly in the Planes, Proc. Nat. Acad. Sci., **61**, 75/28 (1968)
- 2) 津田: レーマー型合同法によらない乱数について, bit, **12-9**, 64/75 (1980)
- 3) 柏木: M 系列の部分的性質を考慮した一様乱数の発生,

- 数理解析研究所講究録 498, 140/152, 京都大学数理解析研究所 (1983)
- 4) 伏見, 手塚: 多次元分布が一様な擬似乱数列の生成法, 応用統計学, **10**-3, 151/163 (1984)
 - 5) R. C. Tausworthe: Random Numbers Generated by Linear Recurrence Modulo Two, Math. Comp., **19**, 201/209 (1965)
 - 6) T. G. Lewis and W. H. Payne: Generalized Feedback Shift Register Pseudo-Random Number Algorithms, J. Assoc. Comp. Math., **20**, 456/468 (1973)
 - 7) 柏木: M系列再発見, 計測と制御, **20**-2, 236/245 (1981)
 - 8) 柏木: M系列による TLP 乱数の二, 三の性質, 計測自動制御学会論文集, **18**-8, 828/832 (1982)
 - 9) 柏木, 原田: M系列の円分位相, 計測自動制御学会論文集, **18**-10, 999/1004 (1982)
 - 10) 泉, 柏木: 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, **18**-9, 929/935 (1982)
 - 11) D. E. Knuth: The Art of Computer Programming Vol. 2 Seminumerical Algorithms, Addison-Wasley (1981)
 - 12) G. S. Fishman and L. R. Moore: A Statistical Evaluation of Multiplicative Congruential Random Number Generators with Modulus 2^n-1 , J. American Statistical Assoc., **77**-377, 129/136 (1982)
 - 13) G. Marsaglia: A Current View of Random Number Generators, Proc. of the 16th Symp. on the Interface Held in Amsterdam, 3/10 (1985)
 - 14) Hewlett-Packard Co., Cupertino, Calif.: HP-25 Applications Programs, rev. E., 116/117 (1976)
 - 15) C. G. Swain and M. S. Swain: A Uniform Random Number Generator that is Reproducible Hardware-Independent, and Fast, J. Chem. Inf. Sci., **20**-1, 56/58 (1980)