擬似乱数生成器 Mersenne Twister について
On Mersenne Twister, a pseudorandom number generator

数学専攻　山田 雅哉
Masaya Yamada

序

本論文は，擬似乱数生成のアルゴリズム Mersenne Twister を提唱した論文 M. Matsumoto,T. Nishimura, MersenneTwister: a 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Transaction on Modeling and Computer Simulation 8 (1998) に基づく総合報告である．乱数は従来，数値解析におけるモンテカルロ法や統計学におけるシミュレーションで重要な役割を果たしていたが，近年では情報伝達の安全性を保証する暗号技術の基盤としても重要性を増して来ている．Mersenne Twister は長周期と高次元均等分布で優れた擬似乱数生成のアルゴリズムである．Mesenne Twister は代数学の，特に有限体の上の線型代数の非常に優れた応用であることをはっきりさせた．また，方々で流布している Mersenne Twister のプログラムソースを Delphi で書き直して実際に動かすところまで持ち込んだ．Mesenne Twister が高速である一つの根拠はすべてをビット演算で処理していることであるが，C や Pascal にはビット演算が組み込まれている，Java にはビット演算が組み込まれていないことなど，実地に体験することができた．

1　乱数の定義

1.1. 一つの確率空間の独立な確率分布に従う確率変数の実現値を乱数とよぶ．乱数に要請される主な性質として
(1) 一様性．すべての数が等確率，つまり等頻度で一様に出現する．
(2) 独立性．規則性がなく，各項は他の項と独立で相関性がない．
(3) 予測不可能性．過去の数列から次の数を予測できない．
(4) 再現不可能性．同じ数列を再現できない．再現するには乱数列そのものを保存しておくしかない．
が挙げられる．

1.2. 疑似乱数の用途は大きく分けて二つある．

(1) モンテカルロ法

(2) 暗号技術

2　線形漸化式の解数列の周期

定義 2.1. S を集合，$\{s_k\}_{k\geq 0}$ を S の元の列とする．整数 $N \geq 0$ と $p \geq 1$ が存在して $i \geq N$ なら $s_{i+p} = s_i$ となるとき，S の元の列 $\{s_k\}_{k\geq 0}$ は準周期的であるという．

定義 2.2. S を集合，$\{s_k\}_{k\geq 0}$ を S の元の準周期列とし，

$$\mathcal{P} = \{(N,p) \in \mathbb{N}\times\mathbb{N}\,;\, p \geq 1,\, i \geq N \text{ なら } s_{i+p} = s_i \text{ が成立する }\}$$

とおく．さらに，(N,p) を辞書式順序を持つ $\mathbb{N}\times\mathbb{N}$ の部分集合 $\mathcal{P}$ の最小元とするとき p を準周期列 $\{s_k\}_{k\geq 0}$ の周期，$\{s_N, s_{N+1}, \ldots, s_{N+p-1}\}$ を準周期列 $\{s_k\}_{k\geq 0}$ の循環節という．

定理 2.3. S を有限集合 $\neq\emptyset$，$s \in S$ とし，$f : S \to S$ を写像とする．$s_0 = s$, $s_k = f(s_{k-1})$ $(k \geq 1)$ によって S の元の列 $\{s_k\}_{k\geq 0}$ を定義する．このとき，$\{s_k\}_{k\geq 0}$ は準周期的．

補注 2.4. 準周期的 $\{s_k\}_{k\geq 0}$ の周期は $N = \#S$ を超えない.

これ以降，線形漸化式が最大周期を実現する条件について考察する.

定理 2.5. q を素数巾，n を整数 ≥ 2 とし，$A \in M(n, \mathbb{F}_q)$ とする．このとき，次の条件は同値.
(a) $\Psi_A(t)$ が次数 n の原始多項式.
(b) $\Phi_A(t)$ が (次数 n の) 原始多項式.
(c) $A \in GL(n, \mathbb{F}_q)$ で，A の位数が $q^n - 1$.
(d) $\boldsymbol{x} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$ が存在して，$\{\boldsymbol{x}, A\boldsymbol{x}, A^2\boldsymbol{x}, \ldots, A^{q^n-2}\boldsymbol{x}\}$ は $\mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$ と一致する.
(e) 任意の $\boldsymbol{x} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$ に対して，$\{\boldsymbol{x}, A\boldsymbol{x}, A^2\boldsymbol{x}, \ldots, A^{q^n-2}\boldsymbol{x}\}$ は $\mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$ と一致する.

命題 2.6. n を整数 ≥ 3 とし，$2^n - 1$ が素数であると仮定する．$f(t) \in \mathbb{F}_2[t]$ を n 次多項式とする．このとき，次の条件は同値.
(a) $f(t)$ は既約多項式.
(b) $f(t)$ は原始多項式.
(c) $t^{2^n} \equiv t \mod f(t)$.

3 Mersenne Twister の定義

定義 3.1. w, n, r, m $(w > 0,\ 0 \leq r < w,\ 0 < m \leq n)$ を整数とする.

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ a_{w-1} & a_{w-2} & a_{w-3} & \ldots & a_0 \end{pmatrix} \in GL(w, \mathbb{F}_2)$$

とする．w 次元行ベクトル $\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{n-1}$ を初期条件にもつ n 階線漸化式

$$(\#) \quad \boldsymbol{X}_{k+n} = \boldsymbol{X}_{k+m} + (\boldsymbol{X}_k \text{ の上位 } w-r \text{ ビット} \quad \boldsymbol{X}_{k+1} \text{ の下位 } r \text{ ビット})A$$

によって $\mathbb{F}_2$ に成分をもつ w 次元行ベクトルの列 $\{\boldsymbol{X}_k\}_{k\geq 0}$ を定義する.

3.2. 線形漸化式 $(\#)$ を 1 階化する.

$$B = \left(\begin{array}{cccc|cc|c} 0 & I_w & 0 & 0 & & & \\ 0 & 0 & I_w & 0 & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ 0 & & & & & & \\ I_w & & & & & & \\ 0 & & & & & & \\ \vdots & & & & \ddots & & \\ 0 & & & & 0 & I_w & 0 \\ \hline 0 & & & & 0 & 0 & I_{w-r} \\ \hline S & & & & 0 & 0 & 0 \end{array}\right), \quad S = \begin{pmatrix} 0 & I_r \\ I_{w-r} & 0 \end{pmatrix} A$$

とおく．このとき,

$$(\tilde{\#}) \quad (\boldsymbol{X}_{k+n}\ \boldsymbol{X}_{k+n-1}\ \ldots\ \boldsymbol{X}_{k+1} \text{ の上位 } w-r \text{ ビット}) = (\boldsymbol{X}_{k+n-1}\ \boldsymbol{X}_{k+n-2}\ \ldots; \boldsymbol{X}_k \text{ の上位 } w-r \text{ ビット})B$$

が成立する.

補注 3.3. $2^{nw-r}-1$ が素数であると仮定する．このとき，命題から，線型漸化式 $(\tilde{\#})$ が最大周期性 $(2^{nw-r}-1)$ を持つ $\Leftrightarrow$ B の固有多項式 $\Phi_B(t)$ が $\mathbb{F}_2[t]$ において既約.

補注 3.4. $a \geq 2$, $n \geq 2$ を整数とする．a^n-1 が素数なら，$a=2$ で，n は素数.

2^p-1 の形をした素数をメルセンヌ素数と呼び，p をメルセンヌ指数と呼ぶ．$2^{19937}-1$ はメルセンヌ素数である．本論文では $mt19937$ について扱っているため，メルセンヌツイスターの最大周期は，$2^{19937}-1$ である.

3.5. MT19937 のパラメーターを列挙しておく.

$$\begin{aligned}(w,n,m,r) &= (32,624,397,31),\\ \boldsymbol{a} &= \text{行列 } A \text{ の最下位行ベクトル} = \text{9908B0DF (16 進法表記)}\\ &= 10011001000010001011000011011111\end{aligned}$$

4 高次元均等分布性

定義 4.1. $\{s_k\}_{k\geq 0}$ を $\mathbb{F}_q^\nu$ の元の純巡回列とし，$S=\{s_k\,;\,k\geq 0\}$ とおく．$\xi_j : \mathbb{F}_q^\nu \to \mathbb{F}_q$ を線型写像とする.

$$w_j(s_k) = \sum_{i=0}^{\infty} \xi_j(s_{k+i})t^i \in \mathbb{F}_q[[t]]$$

とおく.

定義 4.2. $\xi_j : \mathbb{F}_q^\nu \to \mathbb{F}_q$ $(1 \leq j \leq v)$ を線型写像とする.

$$w(s_k) = (w_1(s_k), w_2(s_k), \ldots, w_v(s_k)) = \Big(\sum_{i=0}^{\infty} \xi_1(s_{k+i})t^i, \sum_{i=0}^{\infty} \xi_2(s_{k+i})t^i, \ldots, \sum_{i=0}^{\infty} \xi_v(s_{k+i})t^i\Big)$$

とおく.

定義 4.3. $(t^{-1},0,\ldots,0),(0,t^{-1},\ldots,0),\ldots,(0,0,\ldots,t^{-1}),w(s_0)$ によって生成される $\mathbb{F}_q((t))^v$ の部分 $\mathbb{F}_q[t^{-1}]$ 加群を L で表わす.

記号 4.4. 写像 $\Xi_k : S \subset \mathbb{F}_q^\nu \to (\mathbb{F}_q^v)^k$ を

$$\begin{aligned}&\Xi_k(s_i) =\\ &(\xi_1(s_i),\xi_2(s_i),\ldots,\xi_v(s_i),\xi_1(s_{i+1}),\xi_2(s_{i+1}),\ldots,\xi_v(s_{i+1}),\ldots,\xi_1(s_{i+k-1}),\xi_2(s_{i+k-1}),\ldots,\xi_v(s_{k-1+1}))\end{aligned}$$

によって定義する.

定理 4.5. $S=\mathbb{F}_q^\nu \setminus \{\mathbf{0}\}$ と仮定する．次の条件は同値
(a) $\Xi_k : S \subset \mathbb{F}_q^\nu \to (\mathbb{F}_q^v)^k$ は全射.
(b) $L+(t^k\mathbb{F}_q[[t]])^v = \mathbb{F}_q((t))^v$.
(c) $w(S)+(t^k\mathbb{F}_q[[t]])^v = \mathbb{F}_q[[t]]^v$.

定義 4.6. 定理の同値な条件をみたすとき，線型周期列 S は k 次元均等分布であるという.

定義 4.7. v ビット精度での均等分布の次元とは，上の定理をみたす最大の k と定義する．これを $k(v)$ と表す.

4.8. 自明な上限として，

$$k(v)v \leq \mathrm{d}imS$$

を得る．各 $v = 1, 2, \ldots, w$ について

$$k(v) = [\mathrm{d}imS/v]$$

のとき，各ビットでの高次元均等分布性が最良であるという．

MT19937 において各パラメータは

$$\begin{aligned}
&u = 11,\\
&s = 7,\\
&b = \mathrm{9D2C5680} = 10011101001011000101011010000000,\\
&t = 15,\\
&c = \mathrm{EFC60000} = 11101111110001100000000000000000,\\
&l = 16
\end{aligned}$$

と指定されている．

$1 \leq v \leq 32$ において $k(v)$ を列挙する．

$k(1) = 19937$	$k(2) = 9968$	$k(3) = 6240$	$k(4) = 4984$
$k(5) = 3738$	$k(6) = 3115$	$k(7) = 2493$	$k(8) = 2492$
$k(9) = 1869$	$k(10) = 1869$	$k(11) = 1248$	$k(12) = 1246$
$k(13) = 1246$	$k(14) = 1246$	$k(15) = 1246$	$k(16) = 1246$
$k(17) = 623$	$k(18) = 623$	$k(19) = 623$	$k(20) = 623$
$k(21) = 623$	$k(22) = 623$	$k(23) = 623$	$k(24) = 623$
$k(25) = 623$	$k(26) = 623$	$k(27) = 623$	$k(28) = 623$
$k(29) = 623$	$k(30) = 623$	$k(31) = 623$	$k(32) = 623$

参考文献

[1] M. Matsumoto and T. Nishimura, Mersenne Twister: a 623-dimensionally equidistributed uniform pseudo-random number generator, ACM Trans.on Modeling and Computer Simulation 8 (1998) 3-30

[2] 中村憲，数論アルゴリズム，朝倉書店

[3] 岡本栄司，暗号理論入門，共立出版

[4] O. ゴールドライヒ (著), 岡本龍明・藤崎英一郎 (訳), 現代暗号・確率的証明・擬似乱数, シュプリンガー・フェアラーク東京