1998年度日本オペレーションズ・リサーチ学会
春季研究発表会

擬似乱数発生法の設計における数理計画法の応用について

01604870　東京大学　諸星　穂積　MOROHOSI Hozumi
01501020　東京大学　伏見　正則　FUSHIMI Masanori

1　はじめに

擬似乱数は確率的シミュレーションの基本的道具であり，多数の発生法が提案されている．一方，乱数の定義については，Kolmogorov や Chaitin 等によって精力的に多くの研究が行われた．Kolmogorov[4] は，von Mises の Collective の概念に基づき有限乱数列の存在について一つの解答を与えているが，具体的にそれをどう構成するか全く別の問題である．

Fushimi[2] は Kolmogorov が例示した部分列抽出規則の中で最も単純な等間隔抽出（詳細は次節で述べる）を考えたとき，k 次均等分布の意味で良い乱数列を，M 系列を基に生成する算法の一般的設計法を提案した．また，M 系列乱数について，[5] では，漸近的ランダム性（asymptotic randomness, a.r. 性）という概念を提案している．この性質は，乱数の上位ビットほど，均等分布の次数が大きくなる，というもので乱数列として望ましい性質であると考えられる．

本論では，等間隔の部分列抽出規則を考慮した上で，a.r. 性を近似的に満たす乱数列を生成する算法の設計法を提案する．その中で，乱数の設計という問題が，組合せ理論における binary matroid の circuit の数え上げ問題と，0-1 整数計画問題とに深く関わっていることを述べ（それらはいずれも厳密な解を求めることが難しい問題である），近似解法によってこれらの問題を解いて乱数の設計を行なうアルゴリズムを提案する．

2　基本的事項

等間隔部分列抽出規則 S：数列 $\{x_t;\ t = 1,2,3\ldots\}$ から等間隔に抽出する規則のこと．間隔を自然数 n とすると抽出された数列は $\{x_{nt};\ t = 1,2,3,\ldots\}$ となる．抽出間隔 n の集合を N とする．

k 次均等分布：周期が T の l ビットの 2 進数の系列 $\{x_t;\ t = 1,2,\ldots,T\}$ が与えられたとき，k 次元のベクトル $(x_t, x_{t+1},\ldots,x_{t+k-1})$ の 1 周期にわたる頻度分布が，このベクトルがとり得る 2^{kl} 個のすべての値の上で一様分布となるならばもとの数列 $\{x_t\}$ は k 次均等分布をするという．

漸近的ランダム性：乱数列が周期 T，長さ l ビットの数列で表されるとすると，達成可能な均等分布の最大次数 k は

$$2^{kl} \leq T \qquad (1)$$

を満たす．(1) において，乱数列の上位 $l' < l$ ビットについて考えれば，より高次の均等分布が実現可能な点に着目する．ある数列が漸近的ランダム性を有するとは，その数列において任意の $0 < l' \leq l$ に対して，(1) で決まる $k' = \lfloor \log_2 T/l' \rfloor$ の次元の均等分布が実現されていることをいう．

M 系列：GF(2) 上の原始多項式

$$f(z) = 1 + c_1 z + \cdots + c_p z^p, \quad c_p = 1 \qquad (2)$$

を特性多項式とする漸化式

$$a_t = c_1 a_{t-1} + \cdots + c_p a_{t-p} \pmod 2 \qquad (3)$$

と，すべてが 0 ではない初期値 $(a_1,\ldots,a_p)$ から生成される 0-1 系列のこと．これは，周期 $2^p - 1$ の系列である．

Tausworthe 列：M 系列 $\{a_t\}$ から次のようにして構成される l ビットの 2 進数の系列 $\{x_t\}$ のこと．

$$x_t = 0.a_{\sigma t+1}\cdots a_{\sigma t+l} \qquad (4)$$

ただし，σ は $2^p - 1$ と互いに素となる自然数である．

Tausworthe 列の k 次均等分布：Tausworthe 列が k 次均等分布するための必要十分条件は，$\{x_t;\ t = 1,\ldots,k\}$ の各ビットを構成する M 系列の kl 個の要素が線形独立になることである．ここで，線形独立というのは，以下の意味で使う．M 系列の任意の要素は，漸化式 (3) を繰り返し使うことで $e_1 a_1 + \cdots + e_p a_p$ と書くことができる．すなわち，M 系列の各要素にに唯一の重みベクトル $(e_1,\ldots,e_p)$ が対応する．対応する重みベクトルが線形独立であることを，M 系列の要素が線形独立であるということにする．

一般化 Tausworthe 列：Tausworthe 列 $\{x_t\}$ のビットを入れ換えて

$$x'_t = 0.a_{\sigma t+j(1)}\cdots a_{\sigma t+j(l)} \qquad (5)$$

（ここで，$(j(1),\ldots,j(l))$ は $(1,\ldots,l)$ の置換である）と表現される数列のこと．Tausworthe 列の均等分布の最大次数は，$m = \lfloor p/l \rfloor$ である．しかし，もとの Tausworthe 列が m 次均等分布をしていても，抽出規則 S を適用して得られる数列 $\{x_{nt}\}, n \in \mathrm{N}$ が，m 次均等分布をする保証はない．そこで，一般化 Tausworthe 列を構成して，この数列の上位 l' ビットに注目したとき，抽出規則 S を適用しても m 次均等分布が保証されるようにすることを考える．

3 設計のアルゴリズム

一般化 Tausworthe 列を構成することにより，抽出規則 S による部分列の上位ビットにおいて m 次均等分布を実現させる Fushimi[2] のアルゴリズムは次のようなものである．

1. 原始多項式 $f(z)$，パラメータ σ，乱数列のビット長 l，部分列抽出の間隔 n の集合 N を決める．
2. 均等分布の最大次数 $m = \lfloor p/l \rfloor$ を計算する．
3. 各 $n \in N$ に対して，$x_{ni} = 0.a_{\sigma ni+1} \cdots a_{\sigma ni+l}$ $(i = 0, \ldots, m-1)$ の各ビット $a_{\sigma ni+j}$ $(i = 0, \ldots, m-1;\ j = 1, \ldots, l)$ を $a_1, \ldots, a_p$ で表したときの重みベクトル $e_j^{(i)}$ を求め，それらを列ベクトルとして並べて $p \times lm$ の行列 E_n を作成する．
4. 各 $n \in N$ に対し E_n の列ベクトルの中で極小な線形従属関係をすべて求め，$\mathcal{C}_n$ とする．$\mathcal{C}_n$ の要素は，極小従属関係がある列ベクトルの集合である．
5. $\mathcal{C}_n$ より，ビット間の従属関係を表す行列 G を次のように構成する：$C \in \mathcal{C}_n$ が $C = \{e_{j_1}^{(i_1)}, \ldots, e_{j_\nu}^{(i_\nu)}\}$ であったとする．これに対応して l 次元の 0-1 ベクトル $\boldsymbol{g} = (g_1, \ldots, g_l)$ を $g_{j_1} = \cdots = g_{j_\nu} = 1$，その他の成分を 0 として定義する．すべての $C \in \mathcal{C}_n$ $(n \in N)$ に対して，$\boldsymbol{g}$ をつくり，それらを行ベクトルとして並べて G とする．
6. 以下の 0-1 整数計画問題を解く．

$$\begin{aligned} \text{ILP:}\quad \min.\quad & z_0 = \sum_{j=1}^{l} z_j \\ \text{s. t.}\quad & G\boldsymbol{z} \geq \boldsymbol{1}, \\ & z_j = 0, 1 \quad (j = 1, \ldots, l). \end{aligned}$$

ただし，$\boldsymbol{z} = (z_1, \ldots, z_l)^{\mathrm{T}}$, $\boldsymbol{1} = (1, \ldots, 1)^{\mathrm{T}}$ である．

7. ILP の解として $z_{j_1}, \ldots, z_{j_{l'}}$ が 0，$z_{j'_1}, \ldots, z_{j'_{l-l'}}$ が 1 となるものが得られたとする．
$x'_t = 0.a_{\sigma t+j_1} \cdots a_{\sigma t+j_{l'}} a_{\sigma t+j'_1} \cdots a_{\sigma t+j'_{l-l'}}$ とする．$\{x'_t\}$ の上位 l' ビットについて m 次均等分布が実現している．

上記のアルゴリズムに基づいて，近似的に a. r. 性を有する乱数列の設計を以下のように行う：上記のアルゴリズムで得られた系列 $\{x'_t\}$ の上位 l' ビットからなる系列について，再びアルゴリズムを適用すれば，上位 l'' ビットが $m'(= \lfloor p/l' \rfloor)$ 次均等分布する系列 $\{x''_t\}$ が得られる．以下同様に最上位の 1 ビットに到達するまで繰返す．このようにして得られた一般化 Tausworthe 列は，上位ビットに着目するほど均等分布の次数が大きくなっており，近似的には，a. r. 性を満たしているということが出来るだろう．

アルゴリズム中の 4. で $\mathcal{C}_n$ を厳密に求めることは，所要計算時間の観点から難しい．そこで次のような heuristics をもちいる：行列 E_n に Gauss-Jordan の消去法を適用することで，列ベクトルを基底ベクトルと非基底ベクトルに分ける．各非基底ベクトルが，どの基底ベクトルと従属関係にあるかは，ただちに分かり，この従属関係は極小である．binary matroid（例えば [1] 参照）の理論によれば，これら非基底ベクトルが示す極小従属関係からすべての極小従属関係を求めることができるのだが，その操作が所要計算時間の点で難しいのである．そこで，すべての極小従属関係を求めることはあきらめて，非基底ベクトルが示す極小従属関係だけから，ILP の制約条件の行列 G を構成することにする．この結果，ILP の制約条件がビット間の従属性の記述に関して不十分になったとしても，その解に従って上位 l' を選択したものが均等分布を実現している可能性はある．ビット置換した結果の乱数列の上位 l' ビットが均等分布を実現しているかどうかは，容易に検証できる（選択した l' ビットに対応する重みベクトルの線形独立性を調べればよい）ので，上位 l' ビットの独立性が保証される限りは，アルゴリズムを続ける．

なお ILP（これは集合被覆問題として知られる問題である）についても，やはり，厳密解を求めることは計算量の観点から難しいので，適当な近似解法を用いることにする．

4 数値例

いくつかの GF(2) 上の原始多項式を選んで，上記アルゴリズムを適用してみた．近似的に a.r. 性を満たす乱数発生法をいくつか得ることができた．詳細は，当日発表する．

参考文献

[1] Fournier, J. C., Binary Matroids, in *Combinatorial Geometries* (N. White, ed.) (*Encyclopedia of Mathematics and Its Applications, Vol. 29*), Cambridge University Press, Cambridge, 1987.

[2] Fushimi, M., Designing a Uniform Random Number Generator Whose Subsequences Are k-Distributed, *SIAM J. Comput.*, Vol. 17, pp. 89–99, 1988.

[3] Knuth, D. E., *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms*, 2nd ed., Addison-Wesley, Reading, MA, 1981.

[4] Kolmogorov, A. N., On Tables of Random Numbers, *Sankhya*, Vol. 25A, pp. 369–376, 1963.

[5] Tootill, J. P. R., W. D. Robinson, and D. J. Eagle, An Asymptotically Random Tausworthe Sequence, *J. ACM*, Vol. 20, pp. 469–481, 1973.