数理解析研究所講究録
537巻 1984年 100-111

有限体での演算を基礎とした多次元擬似乱数の発生法

(Generation of n-space Uniform Pseudorandom Numbers Based on Computations in Finite Fields)

統計数理研究所(The Institute of Statistical Mathematics) 仁木 直人(Naoto Niki)

One of the most important randomness criteria for a sequence of pseudo-random numbers seems to depend upon properties of the joint distribution of n consecutive numbers. The idea of this article is to generate a sequence of n-dimensional vectors uniformly distributed in n-space unit cube for a suitably high value of n. If we have such a sequence of vectors

$$U = \{\underline{u}_0, \underline{u}_1, \underline{u}_2, \cdots\}$$
$$= \{(u_{01}, u_{02}, \cdots, u_{0n}), (u_{11}, u_{12}, \cdots, u_{1n}), \cdots\}$$

of period T, for any divisor k of n ($1 \leq k \leq n$) the sequence comprising the components of vectors

$$u = \{u_{01}, u_{02}, \cdots, u_{0n}, u_{11}, u_{12}, \cdots, u_{1n}, \cdots\}$$

is k-space equidistributed, that is, the distribution of all (Tn/k) k-tuples of the sequence is uniform in k-space unit cube. If k does not divide n and divides T, it is required to test seperately the k-space uniformity of the sequence u. However, if n has many divisors, we may expect that it should be also k-space equidistributed even for such k.

The generation method is as follows:
Given a large prime p and an integer n, we can find by probabilistic algorithms an irreducible polynomial of degree n,

$$g(x) = a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_0,$$

over Z_p and a primitive element, $f(x)$, in the finite field $F=Z_p[x]/[g(x)]$. If $f(x)$ is linear, say,

$$f(x) = cx + d$$

each element except 0 in F

$$h_i(x) = b_{i,n-1} x^{n-1} + b_{i,n-2} x^{n-2} + \cdots + b_{i,0} \quad (i=0,1,2,\ldots)$$

is generated by the following recurrence formulae modulo p

$$b_{i+1,j} = c\, a_j\, b_{i,n-1} + c\, b_{i,j-1} + d\, b_{i,j} \quad (1 \leq j \leq n-1)$$

$$b_{i+1,0} = c\, a_j\, b_{i,n-1} + d\, b_{i,0}$$

with an arbitrary non-zero polynomial $h_0(x)$ of degree n-1 (or less). The sequence of coefficients of $(1/p)h_i(x)$ would behave as a sequence of pseudo-random vectors equidistributed in n-space unit cube with a period of length $T=p^n-1$. We may use a primitive irreducible polynomial as $g(x)$, but the author feels that the corresponding recurrence formulae are too simple to generate sufficiently random sequences.

The computer algebra systems (REDUCE, MACSYMA etc) may be much useful for us to find an irreducible polynomial and a "linear" primitive element.

1. はじめに

モンテカルロ法やシミュレーションでは、しばしば大量の多次元一様乱数が必要とされる。多次元超立方体での多重モンテカルロ積分や、固定複数個の一様乱数からある特定の分布に従う乱数に変換して使用する場合（精度の高いWalker(1974)の方法，Niki(1979)など）は、その典型例である。また、常に固定次元のベクトルとして使われるのではなくとも、多次元一様性が保証されている乱数列については、より低い次元での独立性がある程度保証されていることになる。

擬似乱数の発生法として最もよく使われている乗算合同法は、連続した数個の乱数列を1個の多次元乱数とみなしたとき、乗数をうまく選べば、ある程度高い次元までの良い一様性を実現できる（一様性の評価についてはNiederreiter(1978)を参照）。しかし、乗算合同法は周期が比較的短く（法として用いた数の大きさ以下）、また下位の桁について存在する擬周期の問題を考慮すると、大量の多次元乱数として用いることには問題に応じた工夫が必要であろう。

乗算合同法の欠点を克服するために考案されたM系列に基づく発生法（例えば、伏見・手塚(1981)参照）は、周期全体を見るときには、ほとんど理想的な乱数列を発生するが、その部分列の統計的性質は必ずしも良くないとの報告が多い（Lindholm(1968)，仁木(1983)，栗田(1983)など）。その改良のため、M系列の特性多項式を3項式から5項式以上とする提案もあるが、発生速度の大幅な犠牲を伴う上、Arvillias and Maritsas(1978)型の周期分割を行う際に必要な初期値設定の問題が残ろう。

ここで議論するのは、多次元一様擬似乱数を直接発生する原理とその実現法である。発

生法自体は、乗算合同法とM系列発生法をその特殊ケースとして含む、一般的な方法といえよう。同様な一般的発生法として、線型結合による方法（Knuth(1981)，3.2.2）があるが、有限体の理論に基づく点は同じでも、基本的な発想と具体的な発生アルゴリズムが異なる。

2．多次元擬似乱数の発生原理

いまｎ次元単位超立方体（ｎ≧１）上の擬似的な一様乱数列

$$U = \{ \underline{u}_0, \underline{u}_1, \underline{u}_2, \cdots \}$$

を発生させる場合を考える。すなわち、各$\underline{u}$はｎ次元ベクトルで、

$$\underline{u} \in [0, 1]$$

とする。このためには、必要な有効桁数に応じて充分大きい素数ｐを選び、ｐを法とする剰余類体を

$$Zp = \{ 0, 1, \cdots, p-1 \} \sim GF(p)$$

とするとき、

$$Zp^n \sim GF(p^n)$$

で一様分布する（あるいは一様分布にごく近い分布をする）ベクトル列

$$V = \{ \underline{v}_0, \underline{v}_1, \underline{v}_2, \cdots \}$$

を生成し、

$$\underline{x} = (1/p)\,\underline{v}$$

とすればよいであろう。

多項式環Ｚｐ〔ｘ〕の任意のｎ次既約多項式

$$g(x) = x^n - a_{n-1}x^{n-1} - \cdots - a_0$$

$$= x^n + (p-a_{n-1})x + \cdots + (p-a_0) \in Zp[x]$$

に対して、Ｚｐ〔ｘ〕のｇ（ｘ）を法とする剰余類体　Ｆ＝Ｚｐ〔ｘ〕／〔ｇ（ｘ）〕を考えると、Ｚｐ　は　Ｆ　に同形であり、Ｆに属す多項式の係数からなるｎ次元ベクトル（各係数はＺｐの要素）の集合として表現できる。いま、

$$T = p^n - 1$$

とし、Ｆの任意の原始元をｆ（ｘ）とすると

$$F^* = F - \{0\}$$

$$= \{ f(x), (f(x))^2, \cdots, (f(x))^T = 1 \}$$

であるから、多項式ξのべき乗列によってＺｐ　から零ベクトルを除いた集合上で一様分布するｎ次元ベクトル列が生成され、その周期は　Ｔ　である。

よって、もし　Ｚｐ〔ｘ〕のあるｎ次既約多項式　ｇ（ｘ）　を見つけ、かつ

$$F = Zp[x] / [g(x)]$$

のある１個の原始元

$$\xi = f(x)$$

が見出されれば、Ｆ上のべき計算、あるいは、適当な０でない初期値から

$$h_{i+1}(x) = f(x)\, h_i(x) \quad [\mathrm{mod}\ p\ \mathrm{and}\ g(x)]$$

なる漸化式を用いて　$h_i(x)$　の係数として目的とする　ｖ　を生成することができる。

素数ｐを法とする乗算合同法は

$$n = 1, \quad g(x) = x, \quad \xi = c \in Zp$$

の場合と考えればよい。ただし、cはZpの原始根を用いる。

また、Z_2 上の原始既約3項式

$$x^n + x^m + 1 \quad (n>m)$$

に従うM系発生法が

$$p=2, \quad g(x) = x^n + x^{n-m} + 1, \quad \xi = x$$

に対応していることは、ただちに示すことができる。ただし、pが小さいので必要な有効桁にするには別な工夫（Tausworthe(1965), Lewis and Payne(1973)）が必要である。

この代表的なふたつの方法とも、上で述べた発生法の最も簡単な形態といえる。その簡単さゆえに、非常に効率よく発生できるようになっていることは疑いないが、反面、『真の乱数列』の代替列としての性質が犠牲にされていることも事実であろう。『発生効率』と『乱数としての性質』とのバランスを考慮した、『適度な複雑さ』を求めていく必要が痛感される。

3. 多次元擬似乱数発生法の実現

前節で述べた発生法を実現するためには、Fの構成に必要な既約多項式g（x）およびFの原始元ξ＝f（x）を見つけなければならない。

まず、Zp〔x〕内のn次既約多項式を見出す問題については、Rabin(1980) が実際的な方法を提示している。簡単に述べれば、『勝手なZp〔x〕内のn次多項式をとってくれば、それが既約である確率はほぼ　1／n』ということである。すなわち、ランダムに

n個の整数（∈Zp）を選んで、それらをxの0次からn－1次までの項の係数とするn次のモニック多項式をつくり、既約性の判定を行えばよい。既約であればそれをg（x）として採用し、可約であれば新たにn個の整数（∈Zp）を選び、既約なものが見つかるまで繰返す。

Zp〔x〕内のn次多項式g（x）が既約であるための必要十分条件は、

（1）　$g(x) \mid (x^{p^n} - x)$,

（2）$\ell_1, \cdots, \ell_k$ をnの全ての素数の因数とするとき

$$GCD\{g(x), (x^{p^{\ell_i}} - x)\} = 1, \quad 1 \leq i \leq k$$

で与えられる。または、法pのもとで多項式の因数分解を行うBerlekamp のアルゴリズム（Knuth(1981) 4.6.2 参照）の前半を利用して、因数多項式の数が1であるかどうかを確かめてもよい。電子計算機による数式処理システム（例えば、Hearn(1983)）を使用すれば、その判定はさほど困難ではない。

次に、原始元f（x）を見出す方法であるが、これも既約多項式をみつける場合と同じ戦略をとる。F* の要素（Tだけある）のうち原始元は

$$r(T) = \varphi(T)/T = (1 - 1/p_1)(1 - 1/p_2) \cdots\cdots$$

の割合で含まれている。ここにφ（m）はオイラーの関数であり、p_1, p_2, …はTを割り切る相異なる素数を表わす。r（T）の値は比較的大きく、1／6を下回ることは稀である。そこで同様に、ランダムに選んだn個の整数（∈Zp）によりξをひとつ定め、それがFの原始元であるかどうか判定する。原始元でなければ、原始元が現れるまで新たなξを作って判定することを繰返せばよい。

原始元であるための必要十分条件は、（n－1）次以下の多項式f（x）について有限

体Fにおける演算を行い、全てのp_1，p_2，…について、

$$\{f(x)\}^{T/p_i} \neq 1 \qquad (i=1, 2, \cdots)$$

が成立することである（原始元の定義から明らか）。この判定も、数式処理システムを用いて（Fにおける演算を行う環境設定で）、２乗計算の繰返しにより

より

$$\{f(x)\}^{2^0}, \{f(x)\}^{2^1}, \cdots, \{f(x)\}^{2^i}, \cdots$$

をまず求め、次いでf（x）のべきのビット・パターンに従って乗算を行っていけば、比較的容易に実行できる。

但し、実用的な擬似乱数の発生法という観点からは、発生に要する計算が比較的簡単である必要があろう。そのためには、原始元f（x）として低次の多項式が見つかれば好都合である。実際、ξ が１次式

$$f(x) = cx + d$$

であれば、既約多項式g（x）を法とする演算を行うことは、

$$x^n \rightarrow a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_0$$

なる置換則が設定されていることと同等であるから、

$$h_i(x) = b_{i\,n-1}x^{n-1} + b_{i\,n-2}x^{n-2} + \cdots + b_{i\,0}$$

とすれば、

$$h_{i+1}(x) = f(x)\,h_i(x)$$

は、

$$b_{i+1,j} = c\,a_j\,b_{i,n-1} + c\,b_{i,j-1} + d\,b_{i,j} \qquad (1 \leq j \leq n-1)$$

$$b_{i+1,0} = c\,a_0\,b_{i,n-1} + d\,b_{i,0}$$

を法pのもとで計算することにより求めることができる。あらかじめ定数$c\,a_j$ （$0 \leq j \leq n-1$）を求めておいて、jの大きい方から$b_{i+1,j}$を計算するような手続きとすれば、1個の乱数（n次元乱数をn個からなる乱数列と見なすと）は3回の乗算と2回の加算および1回（場合によってはそれ以上）の剰余計算により発生できることになる。

但し、1次式という形の元の集合に限定したとき、その中に原始元がかなり高い比率で含まれているかどうかの保証を著者は知らない。しかし、上記の確率的な原始元探索を実行してみると、かなりの高い頻度で1次式の原始元が見出されることが、経験の上から言うことはできる。もし、あまりに見つからなければ、別の既約多項式を用いるのも一策であろう。

発生に要する演算を減らすという観点からは、f（x）が単に1次式というばかりではなく、$c=1$ あるいは $d=0$ または 1 であれば、乗算・加算の回数が減ることになる。しかし、問題がふたつある。ひとつは、このように強い制限を満たす原始元が存在するかどうか不明であり、存在しているとしても数が少ない（従って見つからない）可能性もある、という点である。もうひとつは、例えば

$$\xi = x$$

が原始元（すなわち g（x）は原始既約多項式）であったとすると、発生に要する計算が、

$$b_{i+1,j} = a_j\, b_{i,n-1} + b_{i,j-1} \qquad (1 \leq j \leq n-1)$$

$$b_{i+1,0} = a_0\, b_{i,n-1}$$

と余りに簡単化しすぎて、種々の使い方に対する耐性（真の乱数の代替物と見なし得るかどうかの）が、乗算合同法やM系列法と同程度にまで、小さくなる危険を感じる。

4. おわりに

多次元一様擬似乱数の発生法に関して、その原理と実現法を述べてきた。ただし、ここで取扱ったのは乱数列の1周期に亘る性質のみで、実際の使用形態に応じた理論的・実験的評価は全く残されている。周期に亘る性質だけを考える上では全て同等であった既約多項式および原始元の組合わせにも、部分列の統計的性質の良し悪しによって、優劣がついてくることになる。『この既約多項式とこの原始元の組合わせを用いれば、性質の良い乱数列が得られる』と自信をもって言うことができるようになるまでには、さらに深い検討を要す。

既約多項式および原始元の判定の実際的な方法は、現在のところ、計算機との頻繁な会話を行いながら実行しているが、数式処理にあまり馴染みのない研究者でも容易に判定ができるように、ＲＥＤＵＣＥなどの数式処理システムに必要な手続きを組込むことも考えている。とくに数式処理システムを用いなくとも判定計算が可能ではあるが、関連する計算の多くを全く別に行う必要があることを考慮すると、ＦＯＲＴＲＡＮなどの数値計算用の汎用言語で手続きを記述するのは得策ではないように思う。

参考文献

Arvillias, A. C., and Maritsas, D. G. (1978). Partitioning the period of m-sequences and application to pseudorandom number generation, J. ACM, 25, 675-686.

伏見正則・手塚　集(1981). 多次元分布が一様な擬似乱数列の生成法, 応用統計学, 10, 151-163.

Hearn, A. C. (ed.) (1983). REDUCE user's manual, Rand Corp., Santa Monica.

Knuth, D. E. (1981). The art of computer programming, Vol. 2 (2nd ed.), Reading, Massachusettes, Addison-Wesley.

栗田良春(1983). M系列のＬ－ｔｕｐｌｅのｗｅｉｇｈｔ　ｄｉｓｔｒｉｂｕｔｉｏｎの偏りについて, 数理解析研講究録４９８, 153-171.

Lewis, T. G. and Payne, W. H. (1973). Generalized feedback shift register pseudorandom number algorithm, J. ACM, 20, 456-468.

Lindholm, J. H. (1968). An analysis of the pseudo-randomness properties of subsequences of long m-sequences, IEEE trans. Informa. Th., IT-14, 569-576.

Niederreiter, H. (1978). The serial test for linear congruential pseudo-random numbers, Bull. Amer. Math. Soc., 84, 273-274.

Niki, N. (1979). Multi-folding the normal distribution and mutual transformation between uniform and normal random variables., Ann. Inst. Statist. Math., 31, 125-140.

仁木直人(1983). パーソナル・コンピュータのための物理乱数発生器, 統数研彙報, 31, 33-49.

Rabin, M. O. (1980). Probabilistic algorithms in finite fields, SIAM J. Comput., 9, 273-280.

Tausworthe, R. C. (1965). Random numbers generated by linear recurrence modulo

two, Math. Comput., 19, 201-209.

Walker, A. J. (1974). New fast method for generating discrete random numbers with arbitrary frequency distributions, Electr. Letters, 10, 127-128.