

I-34 積率適合度による擬似乱数シードの選別に関する基礎的研究

Basic Study on Selection of Optimum Seed of Pseudorandom Numbers Based on Moment Matching

松保重之¹

Shigeyuki MATSUHO

抄録: 限られた個数の擬似乱数を使ったモンテカルロ法によって解を推定する場合、効率の良い計算ができる擬似乱数シードの選別方法を、紹介する。所用個数の擬似乱数を発生させる場合、それらの積率が理論値に最も近くなるような最適な乱数シードを、計算機の空き時間を利用して、予め選別しておいて、必要時(モンテカルロ計算時)に用いることを考える。数値計算例では、いくつかの問題に対して、紹介手法の有効性を例証する。領域積分型モンテカルロ法の例として、円周率を求める問題、指数関数の4重積分を求める問題、および、2次元正規確率分布の上側超過確率を求める問題を考える。領域積分型モンテカルロ法として、遺伝的アルゴリズムへの応用例を示す。

Abstract: In case that there are not enough pseudorandom numbers, precision of computational results based on Monte-Carlo method varies widely depending on a seed of the pseudorandom numbers. This paper shows a method to select an optimum seed of pseudorandom numbers. The selection of the optimum seed is carried out based on Moment Matching method. The optimum seed realizes efficient computation of Monte-Carlo method. Effectiveness of the introduced method is illustrated through several numerical examples. The numerical examples include some integral calculus problems and Genetic Algorithm computational problem.

キーワード: モンテカルロ法, 擬似乱数, シード, 効率化

Keywords: Monte-Carlo Method, Pseudo-random Numbers, Seed, Efficiency

1. まえがき

たとえば、人命に関わる問題、膨大な予算を必要とする問題、長時間の現象を対象とした問題を扱う場合、便利な問題解決法はシミュレーションである。ただし、実際の現象はバラツキを有するので、シミュレーションに關与するパラメータもバラツキを持たせて、所要量も平均値(最確値)で報告されるべきである。すなわち、シミュレーションのパラメータのバラツキは実測データに基づく乱数を用いて生成し、そのバラツキを考慮したシミュレーション(実験)を多数回行う。そうすれば、各試行での所要量はバラツキを有するので、所要量を平均値として求める。この手法を事象再現型モンテカルロ法¹⁾⁻⁵⁾と言う。なお、期待値演算の本質は積分なので、対象問題を領域積分の形に定式化できる場合には、その積分に対してモンテカルロ法を応用して問題解決を図る²⁾場合もある。これを領域積分型モンテカルロ法¹⁾⁻⁵⁾と言う。いずれにしても、これらは乱数を利用しているため、得られる解の精度、計算効率は、用いる擬似乱数の精度に依存している⁵⁾。特に、擬似乱数が所要の確率分布に従っているかどうかは、重要であると考えられる。このような観点から著者は、発生擬似乱数の精度を改善する方法として、一種の積率適合(モーメント・マッチング)

法を紹介した⁶⁾。積率適合法とは、発生させようとする擬似乱数の積率(確率モーメント)が、その理論値に近い値になるように乱数を発生させる方法である。従来の積率適合法では、発生させた擬似乱数に対称変量法を適用したり、2次再抽出(quadratic resampling)法を適用したりして擬似乱数の精度向上を図っている⁷⁾⁻⁸⁾。しかし、対称変量法は対称確率分布に従う擬似乱数を対象として、奇数次積率のみの改善を目的としたもので、2次再抽出法は2次積率のみの改善を考えたものである。これらの欠点を解消しようとしたのが、文献⁶⁾の提案手法である。しかし、文献⁶⁾では、発生擬似乱数の全てを計算機メモリ(配列)に一時保存していたため、計算に使用できる擬似乱数の数が限られ、所要精度の解を得るのに問題があった。また、従来の積率適合法も、文献⁶⁾の提案法も、擬似乱数の精度向上は、ある程度期待できるが、擬似乱数の所要発生時間は長くなる。擬似乱数の高精度化のためにロスする計算時間以上に、モンテカルロ計算による解の推定精度の向上が実現できれば計算効率は向上するが、解析参照時間が短い問題、対象問題の基本変数が少ない問題等では、モンテカルロ法の計算効率向上の程度は小さくなる。

以上のことを考慮して、本研究では、文献⁶⁾で紹介した手法を、モンテカルロ計算の実行前に、擬似乱数の頻度

1: 正会員 博(工) 阿南工業高等専門学校 准教授 建設システム工学科
(〒774-0017 徳島県阿南市見能林町青木265, Tel: 0884-23-7203, E-mail: matsuh@anan-nct.ac.jp)

検定として用いる。ただし、従来の擬似乱数の検定のように、単に擬似乱数発生システムの良否を検定するのではなく、モンテカルロ計算の収束性が良い擬似乱数列を、多くの乱数列の中から選別する。そして、乱数列は乱数シードによって識別して用いる。無限個の乱数を使って所用の確率分布を正確に再現できる理想的な(乱数の)母集団を想定すれば、どのシードの乱数を使っても計算値(推定値)は、正解値に収束するものと考えられる。しかしながら、実際には、我々は、限られた時間内に解を推定しなければならず、また、擬似乱数の周期性の問題、計算機資源(特にメモリ容量)には限りがあること等も視野に入ると、計算の効率化が必要不可欠である。実際に、限られた個数の擬似乱数でモンテカルロ計算を行った場合、シードが異なれば、それらの計算結果も、各シードによって大きなバラツキが出ることは、良く経験するところである。そこで、本研究では、限られた個数の擬似乱数を使って解を推定する場合、効率の良い計算ができる擬似乱数シードを選別する方法を、紹介する。すなわち、所用個数の擬似乱数を発生させた場合、それらの積率が理論値に最も近くなるような最適な乱数シードを予め選別しておいて、必要時に用いることを考える。モンテカルロ計算を実際に行う前に、計算機の空き時間を利用して、種々の条件で、そのような最適な乱数シードをデータベース化しておけば、必要時に最適な乱数シードを参照して普通のモンテカルロ計算を行うだけで効率計算が期待できる。効率化のために浪費する計算時間等は全く無くなる。

数値計算例では、いくつかの典型的な問題に対して、本研究で紹介した手法を適用して、擬似乱数の確率特性の改善効果に関する検討を行う。モンテカルロ法は、冒頭で言及したように、領域積分型と事象再現型とに大別されるので、それぞれについて例題を考える。すなわち、領域積分型モンテカルロ法の例として、円周率を求める問題、指数関数の4重積分を求める問題、および、2次元正規確率分布の上側超過確率を求める問題を考える。所要量を期待値演算としてモンテカルロ計算で求める場合、考慮する(バラツキを有する)パラメータ(これは基本変数とも言う)の数だけの多重積分が必要となるため、そのベンチマーク的な例題として円周率の問題と4重積分の問題を考える。建設分野への応用問題と言うわけではないが、提案手法を建設分野に有効利用するための基礎的データを与える例題である。もちろん、建設分野には多重積分を必要とする問題も多数あり、例題の考え方は容易に応用可能である。また、2次元正規確率分布の上側超過確率の問題は、建設分野の確率統計問題として良く出てくる典型的な計算である。一方、事象再現型モンテカルロ計算の例としては、ごく簡単な遺伝的アルゴリズム(単純GAを対象とする)への応用例を示す。遺伝的アルゴリズムは、生物の進化の過程をシミュレーションしたものであるが、交叉確率、突然変異確率等に乱数を用いるモンテカルロ計算である。

ただし、領域積分の形に定式化できないので、事象再現型モンテカルロ計算である。この応用例は、建設分野では、最適化問題等に応用することができる。

2. 頻度検定による擬似乱数シードの選別

擬似乱数の具備すべき条件^{9),10)}は、a)速やかに多数個発生できること、b)周期があるとすれば十分長い周期を有すること、c)再現性があること、d)良好な統計的性質を持つこと、等が挙げられる。このうち、a)高速生成の条件は、計算機で発生させる擬似乱数であれば満足している。次のb)の条件は、メルセンヌ・ツイスター等の乱数を使えば相当長い周期は確保できるものの大規模問題等では極めて多くの乱数が必要とされる場合があるので注意が必要である。また、c)の条件は、同一の擬似乱数列を何度でも発生できるか否かを問うている。擬似乱数は、数式を用いて発生させているので、シードを指定すれば再現性があり、物理乱数を用いない限り自動的に満足している。通常、このような条件を利用することは無いが、本研究では、頻度検定の結果を計算機メモリに保存して計算機資源を消耗することのないように、積極的に利用する。最後のd)の条件は、擬似乱数の利用目的にあった検定を、利用前に実施すべきであるが、現状はほとんど実施されていない。統計的検定としては、頻度検定、系列相関検定、組合せ検定、ギャップ検定、連の検定等⁹⁾⁻¹¹⁾が知られている。これらのうち、頻度検定は重要な検定となることが多い。特に、確率統計の処理問題では、必要不可欠な検定である。以下、文献⁶⁾で紹介した積率適合法を再度簡単に紹介し、その方法を対象問題へ応用する方法を紹介する。

(1) 擬似乱数の一様性の評価

擬似乱数が所要の確率分布への程度適合しているかを計算プログラムで検定する方法としては、a)乱数の定義区間を適当に(等)分割して、各区間の頻度を観測する方法⁹⁾⁻¹²⁾、b)カイ2乗適合度検定でのカイ2乗値を利用する方法^{13),14)}、c)コルモゴロフ検定での最大絶対偏差値を利用する方法^{13),14)}等が考えられる。しかし、方法a)とb)では区間分割をどのように分割するかと言う問題、方法c)では発生擬似乱数の並べ替えをする問題が生ずる。そこで、文献⁶⁾では、発生させた擬似乱数列の積率を観測し、それらの理論値との一致度によってその擬似乱数列の精度を評価した。完全な確率情報は、確率密度関数で与えられるが、頻度分布で比較することは、プログラム作成上、上述のような欠点がある。そこで、無限大次数までの全ての確率モーメントは、完全な確率情報と等価であることを利用して、対象擬似乱数の良否を、擬似乱数列の積率で評価するのが提案手法の目的である。従来の積率適合の手法⁷⁾⁻⁸⁾も、同様な目的で開発されたものである。具体的には、以下のようにして、高精度の一様擬似乱数系列を選別した。

区間[0, 1]で一様分布する確率変数Xの原点回りのn次原点積率(理論値)は, 一般に,

$$E[X^n] = \frac{1}{n+1} \quad (1)$$

のように簡単に与えられる. そこで, 当該モンテカルロ計算に必要な数のN個の擬似乱数から構成される擬似乱数列をM系列(たとえば1000系列)発生させ, L次までの積率適合を考える. この時, j番目の擬似乱数系列中のi番目の擬似乱数 X_{ij} ($i = 1, \dots, N; j = 1, \dots, M$)に対して, k次の標本原点積率は,

$$\hat{E}[(X_j)^k] = \frac{1}{N} \sum_{i=1}^N (X_{ij})^k \quad (2)$$

で計算でき, j番目の擬似乱数系列についての原点積率の2乗誤差は,

$$S_j = \sum_{k=1}^L \left(\hat{E}[(X_j)^k] - \frac{1}{k+1} \right)^2 \quad (3)$$

で与えられる. 故に, 文献⁶⁾では, 上記の2乗誤差 S_j が最小となる擬似乱数列を一番良好な擬似乱数列とし, その擬似乱数列を実際のモンテカルロ計算に用いた. なお, $\hat{\bullet}$ は, \bullet の推定値であることを示す記号である.

(2) 乱数シードの選別

文献⁶⁾では, 最小2乗誤差 S_j の擬似乱数列を探索するために, 最適な擬似乱数列の候補を計算機メモリに一時保存した. しかし, 擬似乱数シードが決まれば, 同一の擬似乱数列が完全に再現されるので, 上記2.(1)での記述中, 「系列」を「シード」と読み替えれば, 擬似乱数をメモリに保存する必要がなくなる. さらに, 文献⁶⁾では, 最適な擬似乱数列を試行ごとに毎回計算していたので, 効率を計算する場合に, 式(2), (3)の所要計算時間もモンテカルロ計算の所要時間に算入した. しかし, 最小の2乗誤差 S_j となるような擬似乱数シードを予め調査してデータベース化しておけば, 2度とこのような最小2乗誤差 S_j を計算する必要が無い. したがって, 本研究では, モンテカルロ計算を実行する前の頻度検定の代わりに, 式(3)の2乗誤差 S_j が最小となるような乱数シードを選別しておくことにする. そして, 実際にモンテカルロ計算する場合には, 選別した乱数シードを用いることにする. ただし, 最適な乱数シードを選別する場合には, より効果のあるシードを選別するために, 計算で必要とする擬似乱数の個数ごとに, 乱数シ

ードを選別すべきことに注意を要する. その結果は, 擬似乱数の所要個数ごとに作表しておけば, 同じ計算環境であれば, 種々の問題に利用可能である.

なお, 本研究では, 計算機の空き時間を利用して, 予め作成した高精度擬似乱数のシード表に基づいて計算するので, 計算効率^{3),15)-18)}は, 同一のモンテカルロ法に基づく限り, 単純に, 解の推定分散の比として表わされる(推定に伴う分散は, 後述の式(8)を参照).

3. 数値計算例

領域積分型モンテカルロ計算の計算例として, 円周率を求める問題と指数関数の4重積分を求める問題と2次元正規確率分布の上側超過確率を求める問題を考える. 事象再現型モンテカルロ計算の計算例としては, 遺伝的アルゴリズムの最大値探索問題を考える. 計算には, 富士通 Fortran & C Ver.4.0 (使用CPUとOSはWindows XP on Pentium IV 2.8GHz)上で提供されるC言語を用いた.

(1) 計算条件

乱数シードの選別で用いる乱数精度の指標 S_j の計算では, なるべく少ない計算で最小の S_j (式(3)参照)が得られる積率次数を試算し, 5次迄の積率(式(2)と(3)で $L=5$)を考える. ちなみに, 信頼性工学では2次までの積率で信頼性指標を定義¹⁹⁾⁻²⁰⁾し, 不規則関数論の相関理論²¹⁾でも2次までの積率で確率過程をモデル化している. このような状況に鑑みて, $L=3$ 以上として, 計算時間との兼ね合いで $L=5$ と決めた. また, 乱数シードの選別も, なるべく多くのシードから求める方が良いが, これも計算時間との関係から1000個のシードから選別することにする.

なお, 乱数シードの選別を行わない場合のモンテカルロ計算では, 乱数生成システムがデフォルトで有する乱数シードを用いることにする. これは, 普通の工学者は, 特別な理由がない限り, 擬似乱数の検定を行うことはなく, デフォルトの乱数シードを用いてモンテカルロ計算を行うのが普通である状況を反映させたものである.

さらに, モンテカルロ計算では, 擬似乱数の精度のみの影響を調べるために, 分散還元法^{9),15)-18),22)}等を援用しない入門的(Crude)モンテカルロ法^{9),15)-18),22)}を用いる. 入門的モンテカルロ法では,

$$I = \int_a^b f(x) dx \quad (4)$$

と言う定積分は, N個の一様擬似乱数 $\xi_1, \xi_2, \dots, \xi_N$ を用いて, 次式で近似評価される:

$$\hat{I} \cong \frac{b-a}{N} \sum_{i=1}^N f(\xi_i) \quad (5)$$

表 - 1 各種乱数を用いて入門的 MC 法によって求めた円周率

| No. | 試行回数 | デフォルトのシード | | | | 選別したシード | | | |
|-----|-------|-----------|--------|--------|--------|---------|--------|--------|--------|
| | | 普通乱数 | | MT乱数 | | 普通乱数 | | MT乱数 | |
| | | 計算値 | 誤差% | 計算値 | 誤差% | 計算値 | 誤差% | 計算値 | 誤差% |
| 1 | 10 | 3.4374 | 9.4155 | 3.0419 | 3.1728 | 3.0831 | 1.8628 | 3.1635 | 0.6977 |
| 2 | 100 | 3.2065 | 2.0655 | 3.2963 | 4.9254 | 3.1446 | 0.0955 | 3.1473 | 0.1831 |
| 3 | 1000 | 3.1188 | 0.7242 | 3.1082 | 1.0615 | 3.1405 | 0.0361 | 3.1425 | 0.0300 |
| 4 | 10000 | 3.1330 | 0.2742 | 3.1441 | 0.0798 | 3.1418 | 0.0073 | 3.1407 | 0.0269 |

表 - 2 各種乱数を用いて入門的 MC 法によって求めた 4 重積分 (式(6)参照)

| No. | 試行回数 | デフォルトのシード | | | | 選別したシード | | | |
|-----|---------|-----------|--------|----------|--------|----------|--------|----------|--------|
| | | 普通乱数 | | MT乱数 | | 普通乱数 | | MT乱数 | |
| | | 計算値I | 誤差% | 計算値I | 誤差% | 計算値I | 誤差% | 計算値I | 誤差% |
| 1 | 1000 | 0.071059 | 2.3946 | 0.069366 | 0.0456 | 0.070623 | 1.7655 | 0.070943 | 2.2271 |
| 2 | 10000 | 0.070461 | 1.5317 | 0.069554 | 0.2253 | 0.067763 | 2.3551 | 0.069295 | 0.1480 |
| 3 | 100000 | 0.069744 | 0.4986 | 0.069720 | 0.4648 | 0.069138 | 0.3745 | 0.069718 | 0.4620 |
| 4 | 1000000 | 0.069427 | 0.0431 | 0.069502 | 0.1500 | 0.069443 | 0.0656 | 0.069342 | 0.0800 |

表 - 3 デフォルトのシードの各種乱数を用いて入門的 MC 法によって求めた上側超過確率 (式(7)参照)

| No. | 試行回数 | 普通の乱数 | | | MT乱数 | | |
|-----|---------|-----------|---------|--------|-----------|---------|--------|
| | | 計算値P | 誤差% | 変動係数 | 計算値P | 誤差% | 変動係数 |
| 1 | 1000 | 0.0005676 | 13.2042 | 1.3276 | 0.0007399 | 13.1272 | 1.1627 |
| 2 | 10000 | 0.0006406 | 2.0446 | 0.3950 | 0.0006709 | 2.5883 | 0.3860 |
| 3 | 100000 | 0.0006604 | 0.9730 | 0.1230 | 0.0006527 | 0.2038 | 0.1237 |
| 4 | 1000000 | 0.0006554 | 0.2089 | 0.0390 | 0.0006519 | 0.3177 | 0.0392 |

表 - 4 シード選別した各種乱数を用いて入門的 MC 法によって求めた上側超過確率 (式(7)参照)

| No. | 試行回数 | 普通の乱数 | | | MT乱数 | | |
|-----|---------|-----------|---------|--------|-----------|--------|--------|
| | | 計算値P | 誤差% | 変動係数 | 計算値P | 誤差% | 変動係数 |
| 1 | 1000 | 0.0005740 | 12.2332 | 1.3202 | 0.0006810 | 4.1347 | 1.2119 |
| 2 | 10000 | 0.0006536 | 0.0541 | 0.3910 | 0.0006547 | 0.1017 | 0.3907 |
| 3 | 100000 | 0.0006550 | 0.1483 | 0.1235 | 0.0006581 | 0.6200 | 0.1232 |
| 4 | 1000000 | 0.0006506 | 0.5252 | 0.0392 | 0.0006578 | 0.5745 | 0.0390 |

(2) 円周率の計算

提案手法で選別した乱数シードを用いて、検証が簡単で、この種の計算のベンチマーク的な計算例として、円周率を入門的モンテカルロ法によって計算した。

式(5)による入門的モンテカルロ計算を行った結果を表 - 1 に示す。表 - 1 の左半分には、処理系の組み込み関数として一般に提供されている普通の擬似乱数をシード選別しないまま用いた場合の結果と、MT (メルセンヌ・ツイスター) 乱数²³⁾⁻²⁴⁾と呼ばれる擬似乱数をシード選別しな

いま用いた場合の結果を示す。一方、表 - 1 の右側には、シード選別を行った場合の普通擬似乱数を用いた場合とMT乱数を用いた場合とを示している。いずれも正解値との誤差を百分率で示している。

表 - 1 より、シード選別を行うことによって、モンテカルロ計算による解の推定精度が向上することが確認できる。また、普通の擬似乱数の場合とMT乱数の場合との間では、明らかな結果の違いは見られないようである。

表 - 5 多次元分布状況を考慮した積率による選別を行った各種乱数を用いて求めた上側超過確率(式(7)参照)

| No. | 試行回数 | 普通の乱数 | | | MT乱数 | | |
|-----|---------|-----------|--------|--------|-----------|--------|--------|
| | | 計算値P | 誤差% | 変動係数 | 計算値P | 誤差% | 変動係数 |
| 1 | 1000 | 0.0006805 | 4.0531 | 1.2124 | 0.0006880 | 5.2064 | 1.2058 |
| 2 | 10000 | 0.0006956 | 6.3592 | 0.3790 | 0.0006477 | 0.9578 | 0.3928 |
| 3 | 100000 | 0.0006554 | 0.2189 | 0.1235 | 0.0006531 | 0.1328 | 0.1237 |
| 4 | 1000000 | 0.0006539 | 0.0160 | 0.0391 | 0.0006550 | 0.1486 | 0.0391 |

(3) 4重積分の計算

モンテカルロ法のメリットの1つとして、モンテカルロ法は、次元数に関係なく同じ計算公式を使って多次元積分を求めることができる¹⁵⁾⁻¹⁸⁾。しかしながら、多次元積分の計算は、多くの時間を必要とし、効率化が必要不可欠である。ここでは、解の精度の検証が可能な多次元積分として、以下のような4重積分¹⁶⁾

$$I = \int_0^1 \int_0^1 \int_0^1 \int_0^1 (e^{x_1 x_2 x_3 x_4} - 1) dx_1 dx_2 dx_3 dx_4 = 0.0693976 \quad (6)$$

を考えた。

入門的モンテカルロ法によって式(6)を計算した結果を表 - 2 に示す。表記法は表 - 1 と同様である。表 - 2 より、多次元の場合は、シード選別を行った擬似乱数を使っても、モンテカルロ法による解の推定精度に余り影響しないことが分かる。また、円周率の計算時と同様に、普通の擬似乱数の場合とMT乱数の場合との間では、明らかな結果の違いは見られないようである。

(4) 2次元正規確率分布の上側超過確率の計算

たとえば、信頼性評価問題では、当該基本変数の同時確率密度関数を破壊領域にわたって積分し破壊確率を求める¹⁹⁾⁻²⁰⁾が、通常、小さな確率となるため、多くの擬似乱数を必要とする。したがって、このような積分においても、擬似乱数の精度向上が計算結果に与える影響を知ることには、モンテカルロ法に基づく信頼性評価を行う際には重要である。そこで、多次元確率変数の簡単な確率計算問題として、平均値0、分散1の正規分布に従う2変数 x_1 と x_2 を考え、それらの相関係数が $r = 0.6$ の場合の上側確率

$$P = \int_{3.0}^{\infty} \int_{2.0}^{\infty} \frac{1}{2\pi\sqrt{1-r^2}} \exp\left\{-\frac{x_1^2 - 2rx_1x_2 + x_2^2}{2(1-r^2)}\right\} dx_2 dx_1 \quad (7)$$

を求める。正解値は $P = 0.000654$ である²⁵⁾。

式(5)によって式(7)を計算した結果を表 - 3 と表 - 4 に示す。表 - 3 は、普通の擬似乱数とMT乱数に対してシード選別しないままモンテカルロ計算に用いた場合の計算結果である。表 - 4 は、これらの擬似乱数に対してシード選別してモンテカルロ計算に用いた場合の計算結果で

ある。なお、表中の変動係数とは、式(5)による推定誤差に基づく変動係数であり、次式によって評価することができる¹⁸⁾。

$$\rho = \sqrt{\text{Var}[\hat{I}]/\hat{I}} \quad \text{Var}[\hat{I}] = \frac{1}{N} \left\{ (b-a) \int_a^b [f(x)]^2 dx - I^2 \right\} \quad (8)$$

表 - 3 と表 - 4 とを比較することにより、この場合も、シード選別を行った擬似乱数が、モンテカルロ計算による解の推定精度に余り影響しないことが分かる。また、普通の擬似乱数の場合とMT乱数の場合との間でも、明らかな結果の違いは見られないようである。

今まで示した計算では、1次元確率特性が良好な擬似乱数のシードを、多くの乱数シードの中から式(2)と式(3)の評価式を使って選別したが、多次元積分の問題では、余り良い結果を得ることができなかった。そこで、同様に、対象とする多次元確率分布への分布状況が良好な擬似乱数シードを選別すれば、多次元積分を求める問題でも、効率の良いモンテカルロ計算を実施することができるのではないかと考えた。

そこで、多次元確率分布への分布状況を考慮して選別した擬似乱数シードを作表した(この表は、各読者の計算環境によって異なる)。その結果を使って、式(7)の上側超過確率の計算を式(5)の入門的モンテカルロ法で行った。

数値計算の結果を表 - 5 に示す。表 - 3 と表 - 5 との比較により、選別シードの擬似乱数を用いれば、効率化を実現できることが分かる。また、表 - 3 と表 - 5 とを比較すると、誤差が大きいかもかわらず推定誤差に基づく変動係数が小さい場合等が存在する。これは、たとえば、誤差が大きくても、大きな数値を推定するほど推定が容易なることを反映したものである。事象再現型計算の場合に対応させて言えば、頻繁に生起する事象を推定することは容易であるが、稀にしか生起しない事象を推定するのは困難であることに対応する。したがって、推定誤差に基づく変動係数のみを推定精度の評価指標として使う場合は、注意が必要である。

(5) 単純GAによる最大値探索問題への適用

領域積分型モンテカルロ法では、積分領域を陰の形で

表 - 6 単純GAによる最大値探索状況(各世代における適応関数の値の変化)

| 世代 | デフォルトの乱数シードを用いる場合 | | | 選別された乱数シードを用いる場合 | | |
|----|-------------------|-------|-------|------------------|-------|-------|
| | 平均 | 最大値 | 歴代最大値 | 平均 | 最大値 | 歴代最大値 |
| 1 | 1.608 | 3.740 | 3.740 | 2.145 | 3.847 | 3.847 |
| 2 | 2.270 | 3.914 | 3.914 | 2.176 | 3.724 | 3.847 |
| 3 | 2.778 | 3.914 | 3.914 | 2.569 | 3.724 | 3.847 |
| 4 | 2.546 | 3.865 | 3.914 | 2.827 | 3.84 | 3.847 |
| 5 | 2.427 | 3.338 | 3.914 | 2.967 | 3.995 | 3.995 |
| 6 | 2.708 | 3.338 | 3.914 | 3.106 | 3.995 | 3.995 |
| 7 | 2.531 | 3.914 | 3.914 | 3.23 | 3.995 | 3.995 |
| 8 | 2.543 | 3.914 | 3.914 | 3.573 | 3.995 | 3.995 |
| 9 | 2.650 | 3.914 | 3.914 | 3.182 | 3.995 | 3.995 |
| 10 | 2.464 | 3.914 | 3.914 | 3.221 | 3.995 | 3.995 |
| 11 | 2.533 | 3.670 | 3.914 | 3.388 | 3.999 | 3.999 |
| 12 | 2.396 | 3.670 | 3.914 | 3.38 | 3.999 | 3.999 |
| 13 | 2.518 | 3.670 | 3.914 | 3.458 | 3.999 | 3.999 |
| 14 | 2.896 | 3.670 | 3.914 | 3.66 | 3.999 | 3.999 |
| 15 | 2.584 | 3.447 | 3.914 | 3.716 | 4 | 4 |
| 16 | 2.545 | 3.447 | 3.914 | 3.702 | 4 | 4 |
| 17 | 2.652 | 3.363 | 3.914 | 3.861 | 4 | 4 |
| 18 | 2.559 | 3.829 | 3.914 | 3.728 | 4 | 4 |
| 19 | 2.600 | 3.829 | 3.914 | 3.856 | 3.995 | 4 |
| 20 | 2.664 | 3.377 | 3.914 | 3.632 | 3.998 | 4 |
| 21 | 2.473 | 3.150 | 3.914 | 3.454 | 3.999 | 4 |
| 22 | 2.819 | 3.970 | 3.970 | 3.778 | 3.999 | 4 |
| 23 | 2.706 | 3.304 | 3.970 | 3.651 | 3.998 | 4 |
| 24 | 2.761 | 3.413 | 3.970 | 3.754 | 3.998 | 4 |
| 25 | 2.833 | 3.332 | 3.970 | 3.599 | 3.995 | 4 |
| 26 | 2.699 | 3.332 | 3.970 | 3.616 | 3.995 | 4 |
| 27 | 2.662 | 3.150 | 3.970 | 3.748 | 3.995 | 4 |
| 28 | 2.575 | 3.332 | 3.970 | 3.683 | 3.999 | 4 |
| 29 | 2.510 | 3.740 | 3.970 | 3.462 | 3.995 | 4 |
| 30 | 2.450 | 3.970 | 3.970 | 3.547 | 3.995 | 4 |
| 31 | 2.736 | 3.829 | 3.970 | 3.567 | 3.995 | 4 |
| 32 | 2.693 | 3.970 | 3.970 | 3.384 | 3.995 | 4 |
| 33 | 3.017 | 3.970 | 3.970 | 3.630 | 3.995 | 4 |
| 34 | 3.217 | 3.846 | 3.970 | 3.551 | 3.995 | 4 |
| 35 | 3.225 | 3.846 | 3.970 | 3.254 | 3.998 | 4 |
| 36 | 2.815 | 3.970 | 3.970 | 3.384 | 3.998 | 4 |
| 37 | 3.087 | 3.970 | 3.970 | 3.575 | 3.998 | 4 |
| 38 | 2.983 | 3.846 | 3.970 | 3.827 | 3.998 | 4 |
| 39 | 3.265 | 3.848 | 3.970 | 3.766 | 3.998 | 4 |
| 40 | 3.582 | 3.848 | 3.970 | 3.647 | 3.998 | 4 |
| 41 | 3.317 | 3.848 | 3.970 | 3.379 | 3.999 | 4 |

あれ陽の形であれ, 定式化する必要があり, その記述には, 十分な検討を要する場合がある. そのような場合, 注

目事象を計算機内で正確に再現する方法(事象再現型モンテカルロ法)が便利である. 事象再現型モンテカルロ法

は、きわめて多くの問題に適用可能で、安定的に解を得ることができるが、効率が悪く、その効率化に関する研究は重要課題である。ここでは、事象再現型モンテカルロ法の一つである遺伝的アルゴリズム²⁶⁾を対象に、本研究で紹介した方法で選別した最適乱数シードの適用を考える。遺伝的アルゴリズムとしては、単純GAを考える。そして、単純GAに基づき、次式(9)の多峰関数の最大値を求める問題を考える。

$$f(x) = 2 \cos(16x) \sin(x) + 2 \quad (9)$$

これは $x = \pi/2$ の時に最大値 4.0 をとる関数であり、単純GAの計算では、これを適応関数とする。変数 x のコーディングは 0 と 1 の 2 進表現を行い、解空間 $(0 \leq x < 1)$ を 2^{12} 個の離散値に等分する。単純GAは、個体数 20、交叉確率 0.4、突然変異確率 0.1、計算世代数 41 として計算を行う。

デフォルトの擬似乱数シードを用いて計算を行った結果を表 - 6 の左半分に、最適乱数シードの適用を行った単純GAによる計算結果を表 - 6 の右半分に示す。世代ごとの線列の評価値(適合度)を示している。表中、「平均」とは 20 個の個体で得られる評価関数値の平均値である。また、「最大値」とは、20 個の個体で得られる評価関数値の最大値である。「歴代最大値」とは、過去の世代を通じての歴代の最大値を示している。表 - 6 より、右半分の結果は、表 - 6 の左半分の結果に比べ、各値が全体的に大きくなり、正解値 4.0 が出やすくなっているのが分かる。その結果、最適乱数シードの適用によって、容易に正解値が出現している状況を確認することができる。

4. あとがき

本研究では、効率の良い計算ができる擬似乱数シードの選別を、計算機の空き時間(モンテカルロ計算の前)を利用して実施する方法、および、そのような結果を擬似乱数の必要個数ごとに最適乱数シードとして作表しておく方法を提案した。そして、いくつかの典型的な例題に提案手法を適用して、擬似乱数の確率特性の改善効果に関する検討を行った。得られた主な結果は以下のとおりである。

- (1) 本研究で紹介した手法によって作表した最適乱数シードを用いてモンテカルロ計算を実施すれば、効率の良いモンテカルロ計算を行うことができる。
- (2) 多次元の積分問題に対しては、1次元確率特性に着目して選別した乱数シードを用いても、計算の効率化が余り期待できない。多次元確率分布への分布状況の考慮が必要である。
- (3) 事象再現型モンテカルロ計算の一例として遺伝的アルゴリズムの最大値探索問題を考え、本研究で紹介

した手法を適用した結果、効率的な計算が可能であることが確認できた。

数値計算は、限定された条件下での結果ではあるが、紹介した乱数シードの選別法は汎用性のある方法であり、今後、種々の分野へ応用されることが期待される。

参考文献

- 1) 松保重之: 限界状態再現のシミュレーションに基づく構造信頼性評価に関する基礎的研究, 第 18 回材料・構造信頼性シンポジウム講演論文集, pp.5-9, 2001 年 11 月.
- 2) 松保重之: 限界状態再現型モンテカルロ法に基づく構造信頼性評価に関する検討, 第 19 回材料・構造信頼性シンポジウム講演論文集, pp.11-15, 2002 年 11 月.
- 3) 松保重之: 構造信頼性評価のための事象再現型モンテカルロ法の効率化, 材料, Vol.54, No.3, pp.314-319, 2005 年 3 月.
- 4) 松保重之: 事象再現型モンテカルロ法による橋梁信頼性評価に関する基礎研究, 鋼構造年次論文報告集, Vol.16, pp.147-152, 2008 年 11 月.
- 5) 松保重之: 事象再現 MC 法による橋梁信頼性評価で用いる乱数生成法について, 鋼構造年次論文報告集, Vol.17, pp.227-232, 2009 年 11 月.
- 6) 松保重之: 構造信頼性評価に用いる乱数に関する基礎的考察, 第 20 回材料・構造信頼性シンポジウム講演論文集, pp.122-126, 2004 年 12 月.
- 7) 石川達也, 内田善彦: MC 法によるプライシングとリスク量の算出について, 日本銀行金融研究所 IMES Discussion Paper, 2002-J-4, pp.12-25, 2002 年.
- 8) 松保重之: モンテカルロ法の積率調整法に基づく効率化に関する考察, 第 23 回材料・構造信頼性シンポジウム講演論文集, pp.67-71, 2008 年 12 月.
- 9) 三根 久編著: モンテカルロ法・シミュレーション, pp.25-48, コロナ社, 1994 年 10 月.
- 10) 津田孝夫: モンテカルロ法とシミュレーション(改訂版), pp.7-63, 培風館, 1977 年 11 月.
- 11) 宮武修, 脇本和昌: 乱数とモンテカルロ法, pp.36-50, 森北出版, 1978 年 8 月.
- 12) Alfredo H-S. Ang and Wilson H. Tang 著, 伊藤 学, 亀田弘行訳: 土木・建築のための確率・統計の基礎, pp.215-257, 丸善, 1977 年 6 月.
- 13) Alfredo H-S. Ang and Wilson H. Tang 著, 伊藤 学, 亀田弘行訳: 土木・建築のための確率・統計の基礎, pp.258-281, 丸善, 1977 年 6 月.
- 14) 室津義定, 米澤政昭, 邵 暁文: システム信頼性工学, pp.31-47, 共立出版, 1996 年 4 月.
- 15) R. Y. Rubinstein: Simulation and the Monte Carlo Method, pp.114-157, John Wiley & Sons, New York, 1981.
- 16) 津田孝夫: モンテカルロ法とシミュレーション(改訂版), pp.84-106, 培風館, 1977 年 11 月.

- 17) 三根 久編著: モンテカルロ法・シミュレーション, pp.30-40, コロナ社, 1994年10月.
- 18) 室津義定, 米澤政昭, 邵 暁文: システム信頼性工学, pp.105-120, 共立出版, 1996年4月.
- 19) 室津義定, 米澤政昭, 邵 暁文: システム信頼性工学, pp.73-103, 共立出版, 1996年4月.
- 20) 長 尚: 基礎知識としての構造信頼性設計(改訂新版), pp.71-119, 山海堂, 1995年4月.
- 21) 高岡宣善: 工学のための応用不規則関数論, pp.42-66, 共立出版, 1975年2月.
- 22) 湯前祥二, 鈴木輝好: モンテカルロ法の金融工学への応用, pp.59-98, 朝倉書店, 2000年6月.
- 23) M. Matsumoto and T. Nishimura: Mersenne Twister, A 623-dimensionally Equidistributed Uniform Pseudorandom Number Generator, ACM Trans. on Model Comput. Simul., Vol.8, pp.3-30, 1998.
- 24) 湯前祥二, 鈴木輝好: モンテカルロ法の金融工学への応用, pp.30-40, 朝倉書店, 2000年6月.
- 25) 山内二郎編: 統計数値表 JSA-1972, 日本規格協会, pp.124-125, 1972年.
- 26) Goldberg, D.E.: Genetic Algorithms in Search, Addison-Wesley Publishing Company, Inc, 1989.

(2010.5.28 受付)