2000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

準モンテカルロ法の誤差の解析手法の比較

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

s 次元の単位立方体 $I^s=[0,1]^s$ 上で定義された関数 $f(\mathbf{z})$ の積分

$$\theta_f=\int_{[0,1]^s} f(\mathbf{z})\mathrm{d}\mathbf{z} \tag{1}$$

を，I^s 内の点列 $\{\mathbf{z}_1,\ldots,\mathbf{z}_N\}$ による算術平均

$$S_N=\frac{1}{N}\sum_{i=1}^{N} f(\mathbf{z}_i) \tag{2}$$

によって推定することを考える．点列として準乱数と総称される決定論的な点列を用いる方法は，準モンテカルロ法と呼ばれ，多次元の数値積分法として有効であることが知られている．

準モンテカルロ法の誤差評価の方法としては，本来決定論的な準乱数列をランダム化して，モンテカルロ流の統計的な評価を行なうことが，現在のところ，ほぼ唯一の現実的な手法であると考えられている．

以前の報告[3]で，統計的な方法による誤差評価として，Owen[5]のscramble法と，より簡便なshift法を紹介し，それぞれの方法の有効性について数値実験を行なった結果を述べた．その後，Matoušek[2]によりOwenの方法を簡略化した方法が提案されている．今回の発表では，Matoušekの方法について検討するため，数値実験を行なった結果を述べたい．

2. 統計的な評価

準乱数としては，以下に定義する (t,m,s)-net[4] を用いる．

定義 1 整数 $b\geq 2$ が与えられたとき，I^s 中の区間で，

$$E=\prod_{j=1}^{s}\left[\frac{a_j}{b^{d_j}},\frac{a_j+1}{b^{d_j}}\right)$$

と表現できるものを 基数 b の基本区間という．ここで，d_j, a_j は非負整数で，$a_j<b^{d_j}$ を満たすものとする．m, t を非負整数とし，$t\leq m$ とする．I^s 内の総数 b^m 個の点列が，任意の体積 b^{t-m} の基本区間にちょうど b^t 個含まれるとき，基数 b の (t,m,s)-net であるという．

統計的な誤差評価法では，(t,m,s)-net $\{\mathbf{x}_i\}$ をもとにして，互いに確率的に独立な M 組の点列 $\{\mathbf{z}_i^{(r)}\}$ $(r=1,\ldots,M)$ を生成して，(2)によって $S_N^{(1)},\ldots,S_N^{(M)}$ を計算し，(1)の推定値を

$$\hat{\theta}_f=\frac{1}{M}\sum_{r=1}^{M} S_N^{(r)} \tag{3}$$

として，その誤差を分散

$$\hat{\sigma}_f^2=\frac{1}{M(M-1)}\sum_{r=1}^{M}(S_N^{(r)}-\hat{\theta}_f)^2$$

から，標準偏差 $\hat{\sigma}_f$ によって見積もる．

確率的に独立な点列を与える方法としては，次の4種類を考える．

1. **Owenの Scramble.** $\mathbf{x}_i=(x_i^1,\ldots,x_i^s)$ と座標成分で表示したとき，各成分を $x_i^j=\sum_{k=1}^{\infty}x_{ijk}b^{-k}$ と b 進展開する．$\{\mathbf{z}_i\}$, $\mathbf{z}_i=(z_i^1,\ldots,z_i^s)$, $z_i^j=\sum_{k=1}^{\infty}z_{ijk}b^{-k}$ を次のように決める．

$$\begin{aligned} z_{ij1} &= \pi_j(x_{ij1}),\\ z_{ij2} &= \pi_{jx_{ij1}}(x_{ij2}),\\ &\vdots\\ z_{ijk} &= \pi_{jx_{ij1}x_{ij2}\ldots x_{ij,k-1}}(x_{ijk}). \end{aligned}$$

各 π は $0,1,\ldots,b-1$ の置換で，全置換 $b!$ 個の上で一様に分布しているとする．π_j

は全ての i について各 x_i^j の第1桁を置換する．$\pi_{jx_{ij1}}$ は同様に第2桁を置換するが，第1桁の値に依存して決まる．以下同様に，第 k 桁の置換は，$k-1$ 桁までの値に依存して決まる．

2. **Shift.**
$\mathbf{u}$ は $[0,1)^s$ 上で一様分布するベクトルとする．$\mathbf{z}_i = \mathbf{x}_i + \mathbf{u} \pmod 1$ とする．

3. **Rrandom Linear Digit-scrambling.**
Owen の方法を簡略化したものを Matoušek[2] が提案している．ここでは，特に基数 b を素数とする．

$$z_{ijk} = h_{jk}x_{ijk} + g_{jk} (\mathrm{mod}\ b)$$

ここで，h_{jk} は $\{1, 2, \ldots, b-1\}$ から，g_{jk} は $\{0, 1, \ldots, b-1\}$ から，それぞれ互いに独立に選ぶ．

4. **Random Linear Scrambling.**
Matoušek[2] は上記の方法もう少し一般化して，

$$z_{ijk} = \sum_{l=1}^{k} h_{jkl}x_{ijl} + g_{jk} (\mathrm{mod}\ b)$$

とする方法も提案している．ここで，h_{jkk} は $\{1, 2, \ldots, b-1\}$ から，h_{jkl} $(l < k)$ と g_{jk} は $\{0, 1, \ldots, b-1\}$ から，それぞれ互いに独立に選ぶ．

注：shift を行なった (t, m, s)-net はもはや (t, m, s)-net ではない．しかし，領域内に一様に分布するという性質は保たれると考えてよいだろう．他の方法では (t, m, s)-net の性質は保存される．

注：Matoušek の主張は，$(0, m, s)$-net に random linear digit-scrambling や random linear scrambling を行なって得られる L^2-discrepancy の期待値が，既に Hickernell[1] によって計算された $(0, m, s)$-net に Owen の scramble を行なって得られる L^2-discrepancy の期待値と等しいというものである．

3． 数値実験

多次元数値積分法の評価によく利用される試験関数や，現実的な問題に現れる多次元数値積分について，前節で述べた4つの方法について，推定量の精度と誤差評価の有効性を検証した．点列は (t, m, s)-net を発生させ利用した．詳細については，当日述べる．

4． 結語

数値実験からは，推定値の精度で4つの方法に有意な差は見られなかった．一方計算の速度の点からは，方法2と3が速い．方法4はわずかに遅く，方法1は極めて遅い．実装面から考えると，方法2が最も簡単であるといえるだろう．以上より，今回の実験結果をみる限り，準モンテカルロ法の誤差見積もりの手段としては，もっとも単純な Shift 法が有効であると思われる．

参考文献

[1] Hickernell, F. J.: The Mean Square Discrepancy of Randomized Net, *ACM Trans. on Modeling Comput. Simul.*, Vol. 6(1996), No. 4, pp. 274–296.

[2] Matoušek, J.: On the L_2-Discrepancy for Anchored Boxes, *J. Complexity*, Vol. 14(1998), pp. 527–556.

[3] 諸星穂積, 伏見正則: 準モンテカルロ法による多次元数値積分の誤差評価, 日本オペレーションズ・リサーチ学会, 1996 年度春季研究発表会アブストラクト集, pp. 32–33,1996.

[4] Niederreiter, H.: *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.

[5] Owen, A. B.: Monte Carlo Variance of Scrambled Net Quadrature, *SIAM J. Numer. Anal.*, Vol. 34(1997), No. 5, pp. 1884–1910.