I －75

効率化モンテカルロ法における探索メカニズムの考察

(株)地崎工業 技術開発部　　正会員　須藤 敦史

1．はじめに

組み合わせ最適化問題は，多数の解候補の中から制約条件に適合する組み合わせを選択する決定論的な問題であり，連続変数の最適化手法を用いることができないため，最適解の探索は難しく，かつ常に局所解への滞留などの問題が生じる．このような組み合わせ最適化問題に対して遺伝的アルゴリズム:GA[1],Simulated Annealing:SA[2],Random Sampling[3]を基本とした手法などの発見的探索法が注目されている．これらの方法は複数の解候補の情報を利用した探索アルゴリズムを構成し，加えてアルゴリズムや局所解の回避に確率的な考え方を導入している．そこで本報告では，Random Sampling[4]を基本とした手法（修正インポータンスサンプリング法　以下 MIS 法と呼ぶ）の探索アルゴリズムに確率・統計的な考え方を導入する意義をマルコフ過程および情報理論により考察している．

2．探索問題の確率（マクロ）化

組み合わせ最適化問題は，すべての解候補を探索すれば最適解が求められる決定論的（ミクロ）問題である．しかし，実際には以下の問題が生じる．

(1)計算量が爆発的に増加するため，すべてを探索することは不可能　　（解候補と計算量の増加）

(2)直接情報（接線勾配）が得られないため，何らかの方法により探索情報を得なければならず，加えて探索情報が少ないため，これらを有効に利用した探索アルゴリズムを構成しなければならない．（探索情報・効率的なアルゴリズム構成問題）

(3)局所解への滞留問題が生じ，また変数が多い場合には問題を解くことは不可能（局所解への停留と解の唯一性の問題）

そこで発見的探索法は，探索アルゴリズムに以下の考え方を導入している．

(1)複数の解候補から探索情報を入手　　（多点探索法）

(2)集合特性（平均値，分散値）の着目による計算量の減少　　（集合の確率・統計的性質の利用）

(3)ダーウインの進化論(GA),物理現象(SA),離散マルコフ決定過程(MIS 法)などを最適化アルゴリズムに利用（自然・物理現象の最適化アルゴリズムの適用）

(4)仮説探索を繰り返すことで，より良い解（最良解）を探索する方法　　（繰り返し探索法）

したがって，発見的探索手法における最大の特徴はミクロ(決定論的)問題の解候補を集合的に扱うことで，問題のマクロ(確率)化を図り，近似的に解候補を得ようとしている点である．

ここでミクロ問題を確率表現で表すと，候補$x_1, x_2, \cdots, x_n$が生起する確率分布$p(x_1), p(x_2), \cdots, p(x_n)$の設定問題となり，このときの解候補の期待値は式(1)となる．

$$E[x]=\sum_{i=1}^{n} x_i p(x_i) \qquad 0 \le p_i \le 1, \sum_{i=1}^{n} p(x_i)=1 \tag{1}$$

よって，この確率分布を分散が小さくかつ平均値が最適解になるように設定すれば，生起する解候補の期待値が最適解となる確率は高くなる．したがって，探索アルゴリズムにおいて，いかに少ない手順や操作でこの確率分布の平均値を最適解の近傍に設定し，かつ小さい分散を求めていくかが重要となり，ほとんどの探索法で離散マルコフ決定過程によりこのアルゴリズムを組み立てている．

3．探索メカニズムの考察

MIS法や突然変異を用いないSimple GAは離散マルコフ決定過程によって定式化[5]できる．そこで，具体的

A Stochastic Basic Consideration of Conbinatorial Optimization Procedure using Monte Calro Method
By Atsushi SUTOH

にこれらの探索アルゴリズムを考察する.

まず，MIS 法や GA では初回の探索過程において解に関する事前情報が全くないため，解候補はそれぞれランダムに抽出される．次に MIS 法では目的とする解候補の存在範囲を初回の抽出領域から探して，次回の抽出範囲（候補領域）を各変数ごとに設定する．一方 SimpleGA では適応度の高い個体を取り出し，その遺伝子を交配させ，次回の個体集団を作成する.

したがって，両手法ともに k 回目の解候補集合は k-1 回目の解候補集合に依存する離散マルコフ過程を示しており，条件付き確率で表すと式(2)となる.

$$p(x^i_{(k)} \mid x^i_{(1)}, x^i_{(2)}, \cdots, x^i_{(k-1)}) = p(x^i_{(k)} \mid x^i_{(k-1)}) \tag{2}$$

ここで離散マルコフ過程を推移確率行列 P により確率分布 π_t から π_{t+1} への状態推移で示すと以下となる.

$$\pi_{t+1} = \pi_t \cdot P \tag{3}$$

π_t :状態確率分布$(1 \cdot M)$，M :候補の数

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix}$$: 推移確率行列

この推移確率行列の中で最適解となる要素の確率が 1，それ以外の要素が 0 になれば，常に最適解のみが生起する唯一の分布状態となる．しかし，どの要素が最適解であるかは不明なため，問題を解くためには何らかの方法で確率 1 になる要素を特定しなければならない.

そこで，問題を解くための唯一の情報は推移した解候補の個数 n_{ij} であるため，これにより確率 1 になる要素位置の特定する．いま，推移個数を n_{ij} とすると多項分布の確率分布は式(4)となる [6].

$$F(p_{11}, \cdots, p_{MM}) = \frac{n_i!}{\prod_{i=1}^{M}\prod_{j=1}^{M} n_{ij}!} \prod_{i=1}^{M}\prod_{j=1}^{M} p_{ij}^{n_{ij}} \qquad \because\ n_i = \sum_{j=1}^{M} n_{ij} \tag{4}$$

また，p_{ij} を未知パラメータとして対数尤度関数 L は式(5)となる.

$$\log L = \log\left(n_i! \Big/ \prod_{i=1}^{M}\prod_{j=1}^{M} n_{ij}!\right) + \sum_{i=1}^{M}\sum_{j=1}^{M} n_{ij} \log p_{ij} \tag{5}$$

式(5)の右辺第 1 項は定数となり，式(7)の制約条件で式(6)を最大にする推移確率 p_{ij} を求める問題となる.

$$\log L' = \sum_{i=1}^{M}\sum_{j=1}^{M} n_{ij} \log p_{ij} \tag{6}$$

$$p_{ij} \geq 0,\ i, j = 1, 2, \cdots M \quad , \quad \sum_{j=1}^{M} p_{ij} = 1,\ i = 1, 2, \cdots M \tag{7}$$

よって，推移確率行列の要素は式(8)のように推移した解候補の個数 n_{ij} で表される.

$$p_{ij} = n_{ij} / n_i \tag{8}$$

したがって，推移確率行列の要素推定値は，推移した候補数をその総数で割った値となり，同時にその分布も得られる．

一般に情報とは，ある事象の不確定性を減少させるものである．そこで情報理論では定量的な評価指標として情報量という考え方を導入しており，ある生起確率を有する事象の平均的な情報量として，以下に示す情報エントロピーを定義している[7]．

$$H(X)=\sum_{i=1}^{N}-p_i(a)\log p_i(a) \qquad (9)$$

ここで式(9)のグラフを Fig. 1 に示す（$\log p$ の底はeとし，$0\log 0=0$）．この関数はグラフより$p=0$と$p=1.0$で$y=0$（最小値），$p=1/e$で$y=1/e$（最大値）をとることがわかかる．

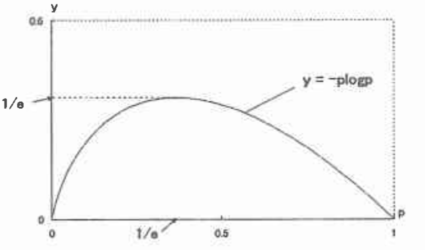

Fig. 1　$y=-p(x)\log p(x)$ のグラフ

グラフより一様分布からの選抜はその情報エントロピーが最大になる．よって，候補からのランダムな抽出は，解候補に関する事前情報が全くない場合において，偏りのない探索の実行を意味している．（解候補の抽出は情報エントロピー最大）

ここで，サイコロを振る確率モデルを考えると情報エントロピーは以下となる．

$$H(X)=\sum_{x}^{6}-\frac{1}{6}\log\frac{1}{6}=1.792 \qquad (10)$$

つまり，一つの目が出現する不確定性の評価値を$-\log p(a_i)=-\log 1/6$とすれば，全体事象における不確定さの平均量が情報エントロピーである．次に，1と6の目が候補から削除された場合の情報エントロピーは式(11)となり，さらに2と5の目が候補から削除されたすると，情報エントロピーは式(12)となる．

$$H(X)=\sum_{x}^{4}-\frac{1}{4}\log\frac{1}{4}=1.386 \qquad (11) \qquad\qquad H(X)=\sum_{x}^{2}-\frac{1}{2}\log\frac{1}{2}=0.693 \qquad (12)$$

ここで，候補のが少なくなるにつれて情報エントロピーは小さくなっており，同時に候補数も少なくなっていることから，解候補の不確定性は減少している．

したがって，情報エントロピー最小となるような解候補は，ある特定な組み合わせ（最適解）になることを意味しており，前記の離散マルコフ過程で示した推移確率行列における最適解となる要素確率を1，それ以外を0にする唯一の分布を形成する操作と同じである．（最適解は情報エントロピー最小）

一方，この探索過程における解候補は一つ前の解候補状態に依存するマルコフ連鎖を示しており，条件付き確率で表される．ここで，以下に示す確率モデルX，Yとその結合確率モデル$Z=X\cdot Y$が定義される事象の条件付き確率の情報量を考える．

$$X=\begin{pmatrix} a_1 & a_2\cdots & a_M \\ p_1 & p_2\cdots & p_M \end{pmatrix} \qquad (13.\mathrm{a}) \qquad\qquad Y=\begin{pmatrix} b_1 & b_2\cdots & b_M \\ q_1 & q_2\cdots & q_M \end{pmatrix} \qquad (13.\mathrm{b})$$

$$Z=X\cdot Y=\begin{pmatrix} a_i\cap b_j & i=1,2,\cdots M \\ r_{ij} & j=1,2,\cdots,M \end{pmatrix} \qquad (13.\mathrm{c})$$

事象a_iが得られたときのb_jに対する条件付き確率の情報エントロピーは次式となる．

$$H(Y|X)=\sum_{j=1}^{M}\sum_{i=1}^{M}-P(a_i)P(b_j|a_i)\log P(b_j|a_i) \tag{14}$$

また，上式は式(15)より式(16)となる．

$$P(b_j|a_j)=\frac{P(a_i\cap b_j)}{P(a_i)} \tag{15}$$

$$H(Y|X)=\sum_{j=1}^{M}\sum_{i=1}^{M}-P(a_i\cap b_j)\log P(b_j|a_i) \tag{16}$$

ここで，条件付き確率の情報エントロピーにおいて事象a_iとb_jが生起した確率$P(a_i\cap b_j)$を実際に推移した解候補の数n_{ij}で置き換えると対数尤度関数L'となる．

加えて，次回の抽出確率は離散マルコフ過程における推移確率行列の要素p_{ij}を求める問題と同様に，条件付き確率$P(b_j|a_i)$における情報エントロピーを最大（偏りのない生起確率）にするように式(16)から求められ，その値は式(6)と同様に推移した解候補数をその総和で割った値である．

4．まとめ

本報告では，発見的探索法の離散マルコフ過程および情報理論による考察を行った結果，以下に示す結論が得られた．

(1) 発見的探索手法は，組み合わせ最適化（ミクロ）問題の確率（マクロ）化を図ることで，効率的に近似解を得ようとする解法である．つまり，解候補が生起する確率分布を，制約条件に適合した解候補（サンプル）の情報より，平均値が最適解近傍にかつ分散をできるだけ小さく設定してゆく問題に置き換えている．

(2) 離散マルコフ過程では，推移確率行列の唯一の分布（最適解となる要素確率を１，それ以外を０）が最適解を生起する状態となり，具体的には推移（制約条件に適合した）した候補の数から，その要素位置は推定される．

(3) 情報理論では，解候補（サンプル）は情報エントロピーが最大（偏りのない一様分布）になるように抽出し，情報エントロピーが最小となる特定な組み合わせが最適解となる．

(4) 離散マルコフ過程における推移確率行列の解候補（サンプル）による推定式と，情報エントロピーによる推定式は同じになり，その値は推移した解候補数およびその総和より求められる．

参考文献

1)D.E.Goldberg: Genetic Algorithm in Search, Optimization and Machine Learning,Addison-Wesly, 1983.

2) E.Aarts and J.Korst: Simulated Annealing and Boltzmann Machines-A Stochastic Approach to Combinatorial Optimization and Neural Computing, Jhon Wiley,1989.

3) Rubinstein,R.Y.: Simulation and Monte Carlo Method,John Wiley & Sons,New York,1981.

4) 須藤敦史,星谷勝,宮沢和樹:遺伝的要素を考慮したｲﾝﾎﾟｰﾀﾝｽｻﾝﾌﾟﾘﾝｸﾞによる離散型変数を有するｼｽﾃﾑの最適化, 土木学会論文集,第519号, I-32, pp.223-232,1995.

5) 山村雅幸, 織田悦子, 小林重信：ﾏﾙｺﾌ過程によるSimple GAの解析,日本機械学会,第2回FANｼﾝﾎﾟｼﾞｳﾑ講演論文集,pp.383-388,1992.

6)西尾勝：自然科学の統計学, 東京大学出版会, 1996.　7)有本卓：確率•情報エントロピー, 森北出版, 1994.