

組合せ最適化問題に対する近傍集合の解析

加 地 太 一

1 はじめに

プリント回路基盤において、ドリルで複数の穴をあける加工工程がある。このとき、基盤のドリルの穴あけに要する時間をできるだけ短くなるように機械を自動化したい。そのために、ドリル穴の位置を機械が次々と移動していくのに要する時間を最小化した工程システムを構築する。それによって、運用における効率的な生産計画が可能となる。実際、これによってあるメーカーでは年間一億円程度の費用削減に成功したとの話もある。さて、このような問題は本論文で対象とする組合せ最適化問題を用いて解くことができる。そして、この組合せ最適化問題は適切な計画、設計、方策など解くための応用範囲の広い問題でもある。たとえば、上記の回路基盤穿孔の問題では、組合せ最適化問題の一つである巡回セールスマン問題 (traveling salesman problem, TSP) を解く事によって解を求めることができる。

TSP とは “何ヶ所かの都市とその都市間の距離が与えられたとき、すべての都市を巡り元に戻る最短の道順を求める問題” である。この問題はすべての可能な道順を列挙してその中で最短な巡回路を見つければ解けるわけであるから、あまりにも単純でナンセンスな問題と思われるだろう。ところが、このナンセンスな問題が計算機でも解けない難物中の難物なのである。もちろん、5都市なら高々24通りの道順しかないので、その中で最短な巡回路が答えであり誰でも簡単に解ける問題である。ところが、その都市数が増えていくと道順の数がどうなるかがこの問題のネックとなるのである。都市数を n とすれば、

出発点から次の都市への行き方は $(n-1)$ 通りであり、さらに、その都市から次の都市への行き方は $(n-2)$ 通りと考えていけば、全巡回路の数は $(n-1)!$ 通りとなる。これはスターリングの公式 $(n! \approx \sqrt{2\pi n} (n/e)^n)$ によりほぼ n^n 通りと表すことができ、都市数の増加により指数関数的にすべての可能な道順の数が莫大に増加する。これを組合せ爆発と呼ぶ。たとえば、30都市の一つの巡回路のコストを一秒間に 10^{12} (1兆) 回計算できるスーパーコンピュータが仮にあった場合、30都市の全部の巡回路の数は $29! = \text{約} 8.8 \times 10^{30}$ 個であるので、30都市のすべての巡回路を計算するには 8.8×10^{18} 秒かかる。一年は 3.15×10^7 秒であるので、30都市の巡回セールスマン問題の答えを求めるには約2800億年 (2.8×10^{11}) 年もかかることになる。すなわち、たかだか、30都市を計算するのに数千億年、あるいは数兆億年という天文学的計算時間を要してしまう。また、これらの問題はいかに計算機のスピードが高速になろうが、その計算時間は雀の涙ほどしか改善されない。というか必要とする計算時間は残念ながら天文学的な値のままである。いわゆる、計算機が如何に速くてもその計算時間は問題のサイズ(都市数)に対して指数関数的に増加する傾向を示し、現実的な時間内で実行不可能な難しい問題の部類に入れられる。

このように、組合せ最適化問題の多くは本質的に複雑度が高く、厳密な最適解を効率よく求めるのは困難である。だが、最適値に近い値をもつ可能解でよければ効率よく求め得る可能性があり、実用上大きな意味がある。近似最適解をうまく求めるには、問題の解や構造に関する知識をいかに利用するかがポイントになるのでヒューリスティック解法と呼ばれる。さらに、近年ではそのヒューリスティックな知識を組み合わせるより高度なアルゴリズムを構成するためのメタヒューリスティックスの研究が盛んに行われている。その中には Local Search 法, Simulated Annealing 法, 遺伝アルゴリズム, Tabu Search 法など、最近話題のアプローチも含まれている。このメタヒューリスティックスにおいて、解の近傍がよりよいアルゴリズムを作るうえでの重要な概念の一つである。近傍とは任意の解に対して、何らかの操作を加えることにより移動した解の集合である。メタヒューリスティックスではこの近傍にもとづいて解

空間を探索し、より良い解へ到達させるものである。したがって、この近傍の良し悪しがアルゴリズムの優劣へとつながってくる。

そこで、本論分ではこの近傍の構造を確率的に解析して、近傍の特性を明らかにするものである。ここから得られる情報を近傍生成、近傍の移動をより効率的に実現する手立てへとつなげていくこと、また、メタヒューリスティックスの性能の解析へとつなげていきたい。本論分ではまず、解空間の特性をあらゆる基礎統計量を AR(1)モデルを構成することによって、汎用的に導出することを可能とした。さらに、情報量基準の考えをもとに、AR(1)モデルの正当性を補強しておく。この AR(1)モデルから得られた基礎統計量をもとに、近傍の構造をあらゆる確率モデルを構成する。

2 AR(1)プロセスによる解空間の解析

解の近傍の分布、あるいは、求まる局所解の分布などを推定する確率的解析を行うため、解空間の構造を統計的に明らかにしたい。そのために、解空間の特性を表す各種の統計量を考察していくことが必要となってくる。そこで、本論文では解空間が AR(1)プロセスと呼ばれる特徴的な性質を有するという仮定を検証し、そこから必要な統計量を導き出していく。そして、後述する節でその導出された統計量の有効性を検討する。確率的な解析では通常、モデルを構築しやすいように、特定の問題、および特定の近傍などを設定して、このような統計量を導出する。しかし、ここで提唱する AR(1)プロセスを用いることにより、多くの組合せ最適化問題、あるいは各種の近傍などに対応する汎用的な解析が可能となることが期待される。以下に、解空間の性質、AR(1)プロセスによるモデル化、そして、そこから導出される基礎統計量について考えていく。

組合せ最適化問題のすべての可能な解 x の集合を X としよう。そして、与えられた解からある基本操作で別の解を導出することを移動と呼び、その移動の集合によって、任意の解に対して近傍集合が定義できる。また、解 x に対しての評価値（コスト）を $f(x)$ で表すこととする。その写像 $f: X \rightarrow R$ を組合せ

最適化問題の評価値ランドスケープ (fitness landscape) と呼ぶこととする。評価値ランドスケープという言葉は理論生物学および理論物理学に端を発し、組合せの対象から実数値への写像に関する問題に広がりつつある。最近、Weiberger [5] は不偏的ランダムウォークの概念が評価値ランドスケープ構造を調査するための適切な手法であることを示した。ここで、“評価値ランドスケープは統計的に isotropic である” という重要な付加的条件を課することが必要である。

この評価値ランドスケープの特性を解析するために、まず、ランダムに選んだ点 (解) を出発点としてランダムに選ばれた近傍点 (近傍解) に移動し、この点 (解) から再びランダムに選ばれた近傍点 (近傍解) に移動することを繰り返すことによって得られた評価値 (fitness) の系列を考える。この評価値の系列の統計が、選ばれた出発点にかかわらず同じであるとき、考えている評価値ランドスケープは統計的に等方的なランドスケープ (statistically isotropic landscape) であるという。多くの組合せ最適化問題ではこの性質がみたまわっていると考えられるので、このモデルの特性を利用し、評価値ランドスケープの構造を分析することが可能である。

この評価値ランドスケープに着目し、組合せ最適化問題のランダムな探索過程のステップ k における評価値 $f(x_k)$ がつくる時系列は、あるランダムウォークの結果であると考えることができる。この系列が時系列モデルの一つのタイプである AR(1) プロセス (autoregressive process of order one) [5] としてモデル化することが可能であり、その考え方を利用して解析を試みるものである。

AR(1) プロセスは次の方程式をみたす過程である。

$$X_t = z_L X_{t-1} + W_t \quad (1)$$

ただし、 W_t は無相関確率変数の定常系列である。Weinberger は “AR(1) プロセスが、N-K 問題や組合せ最適化問題を含めて、ランドスケープの広いクラスの上でランダムウォークの統計をうまく捕らえるであらうと考えることには

十分納得のいく理由がある”と述べている。

組合せ最適化問題の評価値ランドスケープ上のランダムウォーク $\{x_1, x_2, \dots, x_N\}$ は、AR(1)プロセスとして、次の再帰方程式で支配されたモデル化が可能と考えられる。

$$\underline{f}(x_\kappa) = \mu + \rho(1) [\underline{f}(x_{\kappa-1}) - \mu] + \Delta \quad (2)$$

ここで、 $\underline{f}(x_\kappa)$ は解 x_κ のコストを確率変数とし考えたものである。また、 κ は時系列のステップを表す。この再帰的な性質が多く、組合せ最適化問題の解空間に見られ、統計的な性質を導き出すことを可能とする [2] [5]。ただし、 Δ は平均 0、分散 d^2 をもつ白色雑音であり、 $\rho(r)$ は r ステップの自己相関関数である。便宜上、 $\rho(1)$ を ρ で表すものとする。

AR(1)プロセスの特徴として

$$\rho(r) = \rho(1)^{|r|}, \quad r = 0, \pm 1, \pm 2, \dots \quad (3)$$

の式のように自己相関関数はステップ数 r の増加によって 0 への指数関数的な減衰性を示す。この性質が確認されたならば、解空間の評価値ランドスケープが AR(1)プロセスに従っていると判断できうる。また、時系列が AR(1)プロセスであり、かつ $\rho(1) < 1$ ならば、そのプロセスは定常過程となることが知られている [3]。定常過程の定義から、 $\underline{f}(x_\kappa)$ は定数の平均値 μ と分散 σ^2 を持つ。すなわち、

$$E[\underline{f}(x_\kappa)] = \mu \quad \text{for all } \kappa \quad (4)$$

$$E[(\underline{f}(x_\kappa) - \mu)^2] = \sigma^2 \quad \text{for all } \kappa \quad (5)$$

が成立する。また定常性により、 $Cov(\underline{f}(t_i), \underline{f}(t_j))$ はステップ差 $|i - j|$ にのみ依存する。共分散関数 (covariance function) を

$$R(h) = Cov(\underline{f}(x_h), \underline{f}(x_0)) = E[(\underline{f}(x_h) - \mu)(\underline{f}(x_0) - \mu)] \quad (6)$$

により定義すると、

$$\text{Cov}(\underline{f}(x_i), \underline{f}(x_j)) = R(|i - j|) \quad (7)$$

であり、また勿論

$$R(0) = \text{Var}(\underline{f}(x_k)) = \sigma^2 \quad (8)$$

である。ここで、定常過程において自己共分散は

$$R(r) = \frac{1}{N} \sum_{t=1}^{N-r} (\underline{f}(x_t) - E[\underline{f}(x_t)]) \cdot (\underline{f}(x_{t+r}) - E[\underline{f}(x_{t+r})]) \quad (9)$$

で見積もり可能となる [3]。したがって、1ステップの自己相関関数は

$$\rho = R(1)/R(0) \quad (10)$$

であり、この値は x_0 とその近傍との相関係数と同値である。また、 x_0 の任意の2つの近傍間の相関係数 ρ は、近傍解同士の共通する属性の比率が高いことから、近似的に ρ に近い定数と仮定する。このように、AR(1)プロセスの考えに基づき解空間の特徴的な統計量が推定可能となり、以後で述べる推定値の基礎統計量として活用できうるものとする。

評価値ランドスケープ上の時系列が AR(1)モデルであれば、いくつかの特性が得られるわけであるが、問題は実際に AR(1)プロセスであるかということに帰着される。AR(1)プロセスの特徴的な性質として指数的な減衰性を示す(3)式となる顕著な特性がある。この指数的減衰性を確認して、解空間が AR(1)プロセスに従っているかを確かめてみる。ただし、解空間を $\underline{f}(x)$ は平均 \bar{m} 、分散 σ^2 の正規分布 $\mathcal{N}(\bar{m}, \sigma^2)$ に従う確率空間と考える。この解空間モデルの設定は確率的解析における共通したアプローチでもあり、以後の分析においても同様な解空間を対象とする。特に、本研究における手始めの段階でもあるので、解析の検証実験の事例として、組合せ最適化問題の代表的な問題である TSP を使用し、扱いやすい問題の大きさから解析していくこととする。グラフの構造はランダムなコストを付与した完全グラフであり、頂点数100と500を対象とする。その解空間はコストの平均が0、分散は頂点数100の場合は10であり、頂

点数500の場合は50とする。ただし、この場合、ユークリッド平面的な性質は有さない。

図1は頂点数100の場合の結果であり、破線が理論的に導出した自己相関関数の振る舞いを示し、実線が対象の解空間上のサンプルから実際に自己共分散の見積もりを求める(9)式から自己相関関数の値を導出しその振る舞いを示した。また、図2は頂点数500の場合の同様な結果である。

結果として、解空間のサンプルから求めた自己相関関数値は理論的に導出した自己相関関数値と同様な振る舞いを示し、指数関数的な減衰性が認められる。したがって、その空間はAR(1)プロセスに従っているといつてよいであろう。

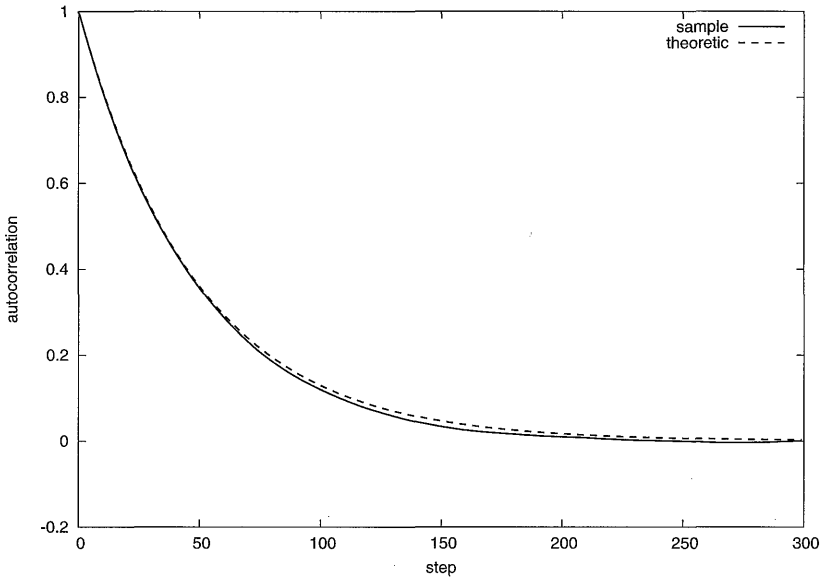


図1：頂点数100の自己相関関数の振る舞い

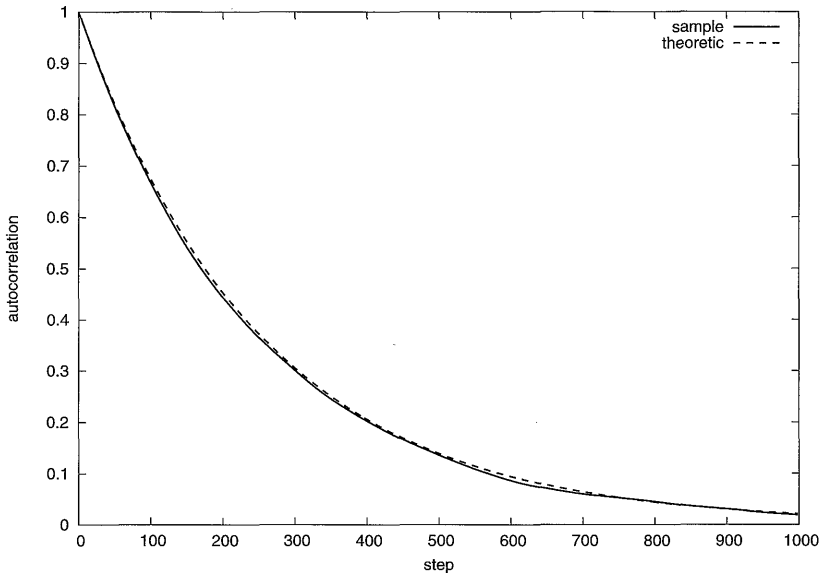


図 2 : 頂点数500の自己相関関数の振る舞い

3 情報量基準による AR(1)モデルの検証

前節では、自己相関関数が指数的減衰性を示す AR(1)プロセスの特徴的な性質が認められることを示した。ただし、AR(1)プロセスであるならば指数的減衰性となりうるが、指数的減衰性の性質があれば AR(1)であるとは必ずしもいえない。そこで、情報量基準などの考え方をを用いて、異なる観点より検討を加えその正当性を補強しておきたい。まず、確率過程における予測の理論から、時間 κ における時系列の値 \hat{F} はそれ以前の各時点の結合分布が多変量正規分布として表されるのならば、

$$\hat{F} = a_1 F_{\kappa-1} + a_2 F_{\kappa-2} + \cdots + a_p F_{\kappa-p} \quad (11)$$

が最適な予測子であると言われている [5]。ただし、各係数は Yule-Walker 方程式に従う。したがって、最良な自己回帰式モデルを推定することにより、

組合せ最適化問題の解のランダム移動のプロセスを決定することができる。具体的には、まず、自己回帰モデルとしていくつかのモデルの次数に対しての自己回帰係数、白色雑音の標準偏差などを最尤法を用いて推定する。次に、いくつかの次数に対して推定された自己回帰モデルに対してどのモデルが最善であるかを、客観的な基準として AIC (赤池の情報量基準) [3] を用いて決定し、自己相関関数の振る舞いのモデルとする。

ここで利用する最尤法は、得られたデータによって、任意の母数を取る尤もらしさを表す関数である尤度関数を用いて母数の値などを推定する方法である。そして、最適な予測子である自己回帰モデルの一般式は

$$X(t) = \alpha_1 X(t-1) + \alpha_2 X(t-2) + \dots + \alpha_M X(t-M) + W(t) \quad (12)$$

であるので、任意の次数 M に対しておのおのの自己回帰係数 α_i の値と $W(t)$ の分散 σ の値をこの最尤法で推定することとなる。ここで、時間とともに変化する時系列データを $\{x_1, x_2, \dots, x_N\}$ とすると、このモデルに対しての尤度関数の対数をとった対数尤度関数は

$$L(\alpha_1, \alpha_2, \dots, \alpha_M, \sigma) = -\frac{1}{2\sigma^2} \sum_{t=1}^N \left[x_t - \sum_{j=1}^M \alpha_j x^{t-j} \right]^2 - \frac{N}{2} \ln 2\pi\sigma^2 \quad (13)$$

となる。この対数尤度関数の値が最大であるところの $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_M, \hat{\sigma}$ が最も尤もらしいモデルを構成することとなる。その $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_M$ の値は、

$$\omega_{ij} = \sum_{t=1}^N x_{t-i} x_{t-j} \quad (14)$$

という量を導入することによって、

$$\begin{pmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1M} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{M1} & \omega_{M2} & \cdots & \omega_{MM} \end{pmatrix} \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \vdots \\ \hat{\alpha}_M \end{pmatrix} = \begin{pmatrix} \omega_{10} \\ \omega_{20} \\ \vdots \\ \omega_{M0} \end{pmatrix} \quad (15)$$

を解くことによって決定する。また、 $\hat{\sigma}$ は、

$$\hat{\sigma} = \frac{1}{N} \sum_{t=1}^N \left[x_t - \sum_{j=1}^M \hat{\alpha}_j x_{t-j} \right]^2 \quad (16)$$

から求まる。

さらに、実際のデータを分析する際の有効なモデル選択基準として情報量基準 AIC がある。その計算は

$$\text{AIC} = -2 (\text{最大対数尤度}) + 2 (\text{モデルのパラメータの数}) \quad (17)$$

で与えられる。この AIC の値が小さいほうが、より真のモデルに近いことを意味する値であり、この値の比較によってモデルの正当性を表す指標となるわけである。自己回帰モデルの場合、最大対数尤度は

$$L(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_M, \hat{\sigma}) = -\frac{N}{2} - \frac{N}{2} \ln 2\pi\hat{\sigma}^2 \quad (18)$$

となるので、これを(17)式に入れて本質的な項を取り出すと、

$$\text{AIC}(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_M, \hat{\sigma}) = N \ln \hat{\sigma}^2 + 2(M+1) \quad (19)$$

となる。

以上より、解空間から得られた時系列データに対して、(15)式から $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_M$ と(16)式から $\hat{\sigma}$ を求める。(19)式を計算し AIC が最小となるように、 $M, \hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_M, \hat{\sigma}$ を定めればよい。

そこで、表 1, 2 に頂点数が100と500に対する M が 0 から 4 の場合の自己回帰係数の推定値と情報量基準の値を記しておく。ただし、時系列上から取るデータによって多少の値の変化は見られるが、多くのケースの場合、 α_2 以降は 0 に近い値となり、実質上は 0 に近似できるものと考えても問題はないであろう。そこで、表 1, 2 では α_2 以降の値が他よりモデルに影響を与えるであろうケースを記しておく。この結果によると、頂点数が100、及び500の2つの場合において、共に $M=1$ 、すなわち AR(1)モデルのケースの情報量基準が最小となり、より、真のモデルに近いことを意味する。また、自己相関関数も(10)式と一致しておりモデルの正当性が示される。

表 1 : 自己回帰係数の推定と情報量基準 (頂点数100)

M	α_1	α_2	α_3	α_4	AIC
0					22582.0
1	0.979				-9236.9
2	0.986	-0.009			-8997.4
3	0.980	-0.02	0.018		-8760.6
4	0.968	0.012	0.01	-0.010	-8991.9

表 2 : 自己回帰係数の推定と情報量基準 (頂点数500)

M	α_1	α_2	α_3	α_4	AIC
0					39268.3
1	0.9959				-9443.7
2	1.0032	-0.0073			-9047.4
3	0.9962	0.0131	-0.0127		-9146.9
4	0.9834	0.0267	0.0034	-0.0173	-8866.7

4 近傍集合の分布に対する解析

評価値ランドスケープ上のランダムウォークから得られる統計的性質の知識として解空間の基本的諸量が導出された。ここで導かれた統計的性質の知識を用いて他の性質を推定できうる。そこからモデルの仮定より評価値ランドスケープ上の重要な局所的な特徴を推論していくこととなる。Weinberger は評価値ランドスケープ上で、状態 x の評価値が与えられたもとの次の状態 y の評価値の分布に関する期待値、分散の推定を示している。これは時系列上の 2 点間に着目し、次の移動する状態が振舞う予測のモデルにより導かれた値である。

しかし、組合せ最適化問題では多数の近傍点を有しており、近傍集合全体を捉えた分布のモデルとして表現はされていない。また、本節で導出される各種

の統計量を求めるには不十分なモデルでもある。そこで、複雑な近傍の構造を表すモデルで再検討し、単純な近傍点でなく近傍構造を確率的に表現していく。

特徴として、AR(1)から求められた統計的性質がこのモデルの上でもよい近似として働き、特定の問題、あるいは特定の近傍のみに限定することなく、多くの知識が推定できることとなる。また、AR(1)から求められた統計的性質と組み合わせて、ここで示した近傍構造のモデルにより、さらなる性質を導く可能性を広げていくものである。

モデルを考察するために、まず、任意の解を x_0 として、解 x_0 に対する近傍解の集合を $\{x_1, x_2, \dots, x_b\}$ と考える。 b は近傍解の個数である。 $f(x_0)$ を解 x_0 のコストの確率変数とすると、その近傍解の確率変数は $f(x_1), \dots, f(x_b)$ であり、それぞれがとる値は c_0, c_1, \dots, c_b で表す。ここで、任意の解 x_0 がコスト c_0 を持つときの解 x_0 の近傍構造を確率的に考察する。すなわち、 $f(x_0) = c_0$ が与えられた条件のもとで、 $(f(x_1), \dots, f(x_b))$ の条件付多変量分布の確率密度関数 $h(c_1, \dots, c_b | c_0)$ を導出することによりその近傍構造を明らかにしたい。このとき、 $\mathbf{X} = (f(x_1), \dots, f(x_b), f(x_0))$ の結合確率分布が多変量正規分布を示す ($\mathbf{X} \sim N_{b+1}(\mathbf{m}, \Sigma)$) と仮定したならば、 $h(c_1, \dots, c_b | c_0)$ は多変量正規分布 $N_b(\mathbf{m}', \Sigma')$ となることが言われている [4]。この平均ベクトル \mathbf{m}' と共分散行列 Σ' を導き、分布の特徴を特定する。そのために、結合確率分布 $\mathbf{X} \sim N_{b+1}(\mathbf{m}, \Sigma)$ のモデルを詳細にしておく必要がある。ここで対象となる \mathbf{X} , \mathbf{m} , Σ の要素は

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad (20)$$

の分割で表される。また、

$$\mathbf{X}_1 = (f(x_1), \dots, f(x_b))^t, \quad \mathbf{X}_2 = (f(x_0)) \quad (21)$$

であり、対応する平均ベクトルの各要素は

$$\mathbf{m}_1 = (m_1, \dots, m_b)^t, \quad \mathbf{m}_2 = (m_0) \quad (22)$$

となる。また、 Σ_{11} は $f(x_i), i = 1 \dots b$ と $f(x_j), j = 1 \dots b$ に対する共分散行列であ

り, その各要素を r_{ij} , $i, j = 1 \dots b$ と記す。 $\Sigma_{22} = (r_{00})$ は $\underline{f}(x_0)$ の分散となる。そして, Σ_{12} と $\Sigma_{21} = (r_{i0})$ は $f(x_0)$ と $f(x_i); i = 1, \dots, b$ との共分散行列を表すものである。

さて, ここで必要とされる各要素の値は, AR(1)の特性から導き出した基礎的な統計量を用いて決定することが可能である。これらの値を用いて有効な推定が可能であるかを検証したい。前述したように組合せ最適化問題の評価値ランドスケープが AR(1)であるところから, 導き出された(4)式から(10)式の性質, および, 近傍解との相関関数値などより,

$$m_i = \mu, \quad i = 0, 1, \dots, b \tag{23}$$

$$r_{ii} = \sigma^2, \quad i = 0, 1, \dots, b \tag{24}$$

$$r_{i0}, r_{0i} = \rho, \quad i = 1, \dots, b \tag{25}$$

$$r_{ij} = \nu, \quad i \neq j, \quad i, j = 1, \dots, b \tag{26}$$

として, $X \sim N_{b+1}(\mathbf{m}, \Sigma)$ のモデルを完成することができうる。

そこで, $h(c_1, \dots, c_b \mid c_0)$ の分布を導出し近傍の確率分布を特定したい。前述した $X \sim N_{b+1}(\mathbf{m}, \Sigma)$ の仮定のもと, X_2 の要素が $\underline{f}(x_0) = c_0$ として与えられた X_1 の条件付分布 $N_b(\mathbf{m}', \Sigma')$ の \mathbf{m}' と Σ' は

$$\mathbf{m}' = \mathbf{m}_1 + \Sigma_{12} \Sigma_{22}^{-1} (c_0 - m_2) = (\mathbf{m}'), \tag{27}$$

$$\Sigma' = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} = (r'_{ij}) \tag{28}$$

となることが知られている [4]。すでに AR(1)から得られた基礎統計量により, 上式の各要素の値は導くことが可能である。近傍の多変量正規分布の各々の確率変数の期待値は $m'_i = m' = \mu + \rho(c_0 - \mu); i = 1, \dots, b$ となり, 各々の分散は $r'_{ii} = \sigma^2(1 - \rho^2); i = 1, \dots, b$ として必要なモーメントが導出可能となる。

次に, 近傍の構成を示す値として, 解 x_0 のコストを c_0 とした場合, 解 x_0 近傍の最小値を AR(1)の情報に基づき推定したい。この値を求めるために重要な確率として以下のステップ確率がある [1]。

$$g(c_0, c) = \Pr \{ \forall i \in \{1, \dots, b\} \underline{f}(x_i) > c \mid \underline{f}(x_0) = c_0 \} \quad (29)$$

これは解 x_0 がコスト c_0 を持つとき、その近傍のすべてがコスト c より大である確率を示す。さらに、このステップ確率を用いて、コスト c_0 の解からコスト c の解へ移動する確率密度が

$$P(c_0, c) = \frac{-\frac{\partial}{\partial c} g(c_0, c)}{1 - g(c_0, c)} \quad (30)$$

の式で導出される [1]。

今回、この(30)式を用いて、

$$\int_{-\infty}^{\infty} cP(c_0, c) dc \quad (31)$$

を解 x_0 のコスト c_0 が与えられたものとしての解 x_0 の近傍の最小値の推定値とする。

問題はステップ確率(29)式を導出することに帰着される。この値を求めることにより、(30)、(31)式が求まることとなる。そのステップ確率は先に求めた AR(1) モデルからの基礎統計量を用いて導出した近傍の確率密度関数 $h(c_1, \dots, c_b \mid c_0)$ に対して

$$g(c_0, c) = \int_c^{\infty} \dots \int_c^{\infty} h(c_1, \dots, c_b \mid c_0) dc_1 \dots dc_b \quad (32)$$

のように求めることができる。これによって特定の問題、特定の近傍に限定することなくステップ確率が計算される。 $h(c_1, \dots, c_b \mid c_0)$ は多変量正規分布 $N_b(\mathbf{m}', \Sigma')$ に対応する。しかし、(32)式の多重積分の計算は数値的にも計算は容易でない。そこで、多次元変数を単一化して一重積分に簡単化したステップ確率の導出を示しておく。まず、Tong の定理 (Theorem 3.3.3) [4] を用いることにより、近傍集合を表す b 変量確率ベクトル $\mathbf{X}_1 = (\underline{f}(x_1), \dots, \underline{f}(x_b)) \sim N_b(\mathbf{m}', \Sigma')$ を独立した $b+1$ 変量標準正規分布の確率ベクトル $\mathbf{Y} = (y_0, y_1, \dots, y_b)$

$\sim \mathcal{N}_{b+1}(\mathbf{0}, I_{b+1})$ で構成できる。すなわち, $\mathbf{X}_1 = \mathbf{C}\mathbf{Y} + \mathbf{m}'$ で変換できる。ただし, \mathbf{C} は $b \times (b+1)$ の行列

$$\mathbf{C} = \alpha \begin{pmatrix} \beta & 1 & 0 \\ & \vdots & \ddots \\ \beta & 0 & 1 \end{pmatrix} \quad (33)$$

であり, ここで,

$$\alpha = \sigma \sqrt{1-\nu} \quad (34)$$

$$\beta = \sqrt{(\nu-\rho^2)/(1-\nu)} \quad (35)$$

であるものとする。この $\mathbf{X}_1 = \mathbf{C}\mathbf{Y} + \mathbf{m}'$ と \mathbf{Y} の独立した標準正規分布性からステップ確率(32)式を

$$\begin{aligned} & g(c_0, c) \\ &= Pr \{ \forall i \in [1, \dots, b] \underline{f}(x_i) > c \} = Pr \{ \forall i \in [1, \dots, b] y_i > -\beta y_0 + (c - \mathbf{m}') / \alpha \} \\ &= \int_{-\infty}^{\infty} Pr \{ \forall i \in [1, \dots, b] y_i > -\beta y_0 + (c - \mathbf{m}') / \alpha \mid y_0 = s \} Pr \{ y_0 = s \} ds \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{s^2}{2}\right) Pr \{ \forall i \in [1, \dots, b] y_i > -\beta s + (c - \mathbf{m}') / \alpha \} ds \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{s^2}{2}\right) \left(\frac{1}{\sqrt{2\pi}} \int_{-\beta s + (c - \mathbf{m}') / \alpha}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt \right)^b ds \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{s^2}{2}\right) \left(\frac{1}{2} \operatorname{erfc}\left(-\frac{\beta}{\sqrt{2}}s + \frac{1}{\alpha\sqrt{2}}(c - \mathbf{m}')\right) \right)^b ds \quad (36) \end{aligned}$$

と導出することができる。また, (30)式は

$$\begin{aligned} P(c_0, c) &= \frac{1}{\sqrt{2\pi}(1-g(c_0, c_0))} \int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} \\ &\quad \cdot \frac{b}{\sigma\sqrt{(2-2\nu)\pi}} \left(\frac{\operatorname{erfc}(u)}{2} \right)^{b-1} e^{-u^2} ds \quad (37) \end{aligned}$$

と計算される。ただし、

$$u = -s \sqrt{\frac{\nu - \rho^2}{2 - 2\nu}} + \frac{1}{\sigma\sqrt{2 - 2\nu}} (c - \rho c_0 + (\rho - 1)\mu) \quad (38)$$

である。以上より、(31)式が数値計算的に計算可能となる。

5 おわりに

計算困難な組合せ最適化問題の近似解を求めるための手法としてメタヒューリスティックスの研究、開発がなされている。この手法のベースとなる近傍構造を解析することは、近傍生成などにおける知識獲得を可能とし、メタヒューリスティックスの改良へとつなげる知見となりうる。また、アルゴリズムの良し悪しを解析するための確率的解析の基本的情報をも提供するものである。

そこで本論分では、近傍の分布の特性を確率的に解析するために、解空間の構造を統計的に明らかにした。解空間の特性を示す各種の統計量を導出するために、解空間が AR(1)プロセスと呼ばれる特徴的な性質を有するという仮定を検証し、そこから必要な統計量を導き出した。さらに、AR(1)モデルから導き出した統計量を用いて近傍構造を確率的にモデル化し、近傍の特性の解析に成功した。

確率的な解析では通常、モデルを構築しやすいように、特定の問題、および特定の近傍などを設定して、このような統計量を導出する。しかし、ここで提唱する AR(1)プロセスを用いることにより、多くの組合せ最適化問題、あるいは各種の近傍などに対応する汎用的な解析が可能となったところに特徴があり、汎用性に富んだ有効な解析方法が示されたものである。今後はさらにアルゴリズムの特性を示す解析モデルの構築、および、本論分で得られた近傍構造の情報をもとに有効な手法の開発を試みたい。

参考文献

- [1] Eikelder, H.M.M., Verhoeven, M.G.A., Vossen, T.W.M. and Aarts, E.H.L., "A Probabilistic Analysis of Local Search", in *Meta-Heuristics: Theory & Applications*, ed. Osman, I.O. and Kelly, J.P., Kluwer Academic Publishers, pp.605-618, 1996.
- [2] Kaji, T., "Probabilistic Analysis of Local Search using AR(1) Model in the Graph Partitioning Problem", *Proc. of the Fifth Metaheuristics International Conference*, Paper ID 38 (2003).
- [3] Priestley, M.B., "Spectral Analysis and Time Series", Academic Press, 1981.
- [4] Tong, Y.L., "The Multivariate Normal Distribution", Springer-Verlag, 1990.
- [5] Weinberger, E., "Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference", *Biological Cybernetics*, Vol.63, pp.325-336, 1990.