2006－MPS－59 (21)
2006／5／26

近傍解コストの分布に基づく最適化手法の提案と評価

重弘裕二　増田達也

大阪工業大学 工学部 電気電子システム工学科

近傍探索法は，解に対して近傍操作を施すことにより，さらなる良い解を探す手法である．したがって，近傍操作を施す対象となる解をどのように選ぶかにより，探索の効率は変化する．従来の近傍探索法では，解自身の評価値に基づき近傍操作を施す対象となる解を選択するが，必ずしも，解自身の評価値が，その近傍に含まれる解の評価値を表す最適な指標であるとは限らない．そこで本稿では，各解に対して近傍解の評価値がどのように分布しているかをベイズ推定により予測し，その結果に基づき探索を行う多点探索法について考察する．まず評価値の分布を推定する方法について述べ，次に探索手法について説明する．また計算機実験の結果を示す．

Proposal of An Optimization Method Based on Cost Distribution of Neighbor Solutions

Yuji Shigehiro and Tatsuya Masuda

Department of Electrical and Electronic Systems Engineering,
Faculty of Engineering, Osaka Institute of Technology

In the neighbor search method for solving optimization problems, the neighbors of solutions of low cost are searched, assuming that the lower the cost of the solution is, the more good solutions its neighbor may include. However, the cost of the solution itself is not necessary the best indicator of the costs of solutions in its neighbor. In this paper, a new optimization method is proposed based on the cost distribution of neighbor solutions, which can be estimated from the costs of neighbor solutions obtained in the process of optimization, by using Bayes estimation.

1　はじめに

多くの組合せ最適化問題は NP 困難であり，実用上の問題において大域最適解を求めるのは難しいことが多い．そこで，より最適性の高い解を求めるために，メタヒューリスティックスと呼ばれるさまざまな手法が提案されている．

それらの多くは，「良い解同士は似通った構造を持っている」という概念に基づいて設計されている[1]．すなわち，それらの手法では，良い解に似通った構造を持つ解の中から別の良い解を探し出し，それらの中に，さらに良い解が見つかることを期待する．似通った構造を持つ解を求める手段が，局所探索法や Simulated annealing 法における近傍操作であり，遺伝的アルゴリズムにおける交差や突然変異の操作である．以下，本稿では近傍操作に基づいて探索を行う手法（近傍探索法）について考える．

近傍探索法は，「それまでに求まっている解の近傍の中から，それまでに求まっている最も良い解よりもさらに良い解（以下，改善解と呼ぶ）を見つけ出す」という処理を繰り返して，順次，より良い解を見つけ出していく手法である．したがって，効率良く解を探索するためには，改善解が含まれる可能性の高い近傍を，より優先して探索するべきであろう．既存の手法では解自身の評価値（コスト）に基づいて探索を行うので，「解自身が良ければ良いほど，その近傍に改善解が含まれている可能性が高い」という仮定に基づいて，探索を行っているものと考えることができる．しかし実際には，例えばその時点で最も良い解が，最も高い確率で改善解を含んでいるとは限らない．そこで本稿では，近

傍に含まれる解のコストの分布を予測し，より効率良く解探索を行う最適化手法について考察する．

2 近傍解のコストの分布

最適化問題とは一般に，$f(x)$ を $x \in F$ という条件の元で最小化するという問題である[1]．ここで x, f, F はそれぞれ解，目的関数，実行可能領域と呼ばれる．目的関数 f は，実行可能領域 F 内の各々の解 x に実数を対応付ける写像 $f: F \to R$ であり，実数値 $f(x)$ は x の評価値，またはコストとも呼ばれる．

近傍 N とは，解 x に F の部分集合を対応付ける写像 $N: F \to 2^F$ である．通常，近傍探索法では，x に似通った構造を持つ解の集合が $N(x)$ となるような近傍 N を考える．x の近傍に含まれる解 $x' \in N(x)$ を x の近傍解 と呼び，近傍解を 1 つ選び出す操作を近傍操作と呼ぶ．通常，近傍操作は x に少しの変形を加えたものを x' とする等により実現される．なお，本稿では対称な近傍 ($x' \in N(x) \Leftrightarrow x \in N(x')$) について考える．

近傍解を 1 つランダムに選び出す近傍操作が与えられているとする．このとき，ある解 x に対して，近傍操作により（ランダムに）得られる近傍解 $x' \in N(x)$ のコスト $f(x')$ が，ある実数値 v となる確率 $P(f(x') = v)$ を考えることができる．この確率を表す確率密度関数 $p(v)$ を本稿では近傍解コスト分布と呼び，特に解 x に依存していることを明示して $p(v|x)$ のように表す．

近傍探索法では，順次，より良い解を見つけ出すことにより探索を進めていく．本稿では，探索の過程のある時点において，それまでに見つかっている最も良い解を暫定最良解，それよりもさらに良い解を改善解と呼ぶ．暫定最良解を x_b とすると，ある解 x に 1 回近傍操作を施すことにより改善解が得られる確率は

$$\int_{-\infty}^{f(x_b)} p(v|x)dv$$

となる．これを本稿では x の改善解選択確率と呼ぶ．

例えば，暫定最良解 x_b と別の解 x_1 のコスト $f(x_b)$, $f(x_1)$，および近傍解コスト分布 $p(v|x_b)$, $p(v|x_1)$ が図 1 のような関係にあるとする．図中の斜線は改善解選択確率を表している（x_1 の改善解選択確率の方がより大きいとする）．従来の局所探索法では，最も良い解である x_b の近傍を探索するが，この場合は x_1 の近傍を探索した方が，より高い確率で改善解を見つけることができるはずである．

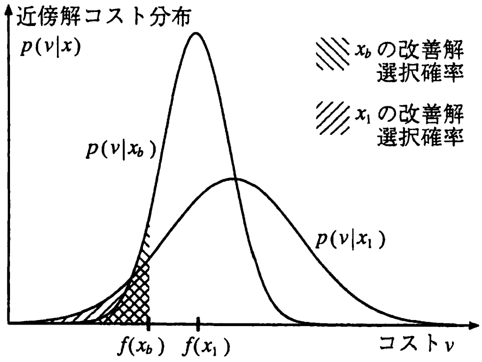

図 1: 近傍解コストの分布の例

3 コストの分布の推定

近傍解コスト分布は，問題や近傍により形状が異なり，多くの場合，正確に予測するのは難しい．本稿では，近傍解コスト分布が正規分布であると仮定し，探索中に行った近傍操作の結果から得られる情報を用いて，ベイズ推定により近傍解コスト分布を推定する．

3.1 近傍操作の結果に基づく推定

探索の過程において，解 x に対して施された i 回目の近傍操作の結果得られた近傍解を x_i とし，x_i と x のコストの差 $f(x_i) - f(x)$ を d_i で表す．ここで d_i が正規分布に従う確率変数であると考える．すると，文献 [2] 等に示されている通り，既に（n 回の近傍操作が x に対して施されて）$d_1, d_2, \ldots, d_n$ が得られているとき，次に得られるであろう d_{n+1} の予測分布は（一般化した） t 分布に従うと考えられる．具体的には

$$t = \frac{d_{n+1} - \bar{d}}{\sqrt{\dfrac{n+1}{n\nu}S^2}} \tag{1}$$

とおくとき，t は自由度 ν の t 分布に従う．ただし

$$\nu = n - 1$$

$$\bar{d} = n^{-1}\sum_{i=1}^{n} d_i$$

$$S^2 = \sum_{i=1}^{n}(d_i - \bar{d})^2$$

である．

次に得られるであろう近傍解のコスト $f(x_{n+1})$ は，d_{n+1} に解のコスト $f(x)$ を加えたものであるから，その予測分布は同様に，一般化した t 分布に従う．この予測分布を，以後，近傍解コスト分布と考える．

なお，近傍解のコストの差 d_i の総和と 2 乗和をそれぞれ

$$A = \sum_{i=1}^{n} d_i$$

$$B = \sum_{i=1}^{n} {d_i}^2$$

とすると $\bar{d} = n^{-1}A$，$S^2 = B - A^2/n$ であるので，n，A，B がわかれば，式 (1) を用いて近傍解コスト分布を推定することができる．

3.2 近傍の類似度に基づく推定

前節の方法は，注目している解に対して施された近傍操作の結果得られる情報に基づいているため，近傍操作がまだ施されていない，もしくは施された近傍操作の回数の少ない解に対しては，意味のある予測を行うことができない．そこで本節では，「構造の似通った解の近傍解コスト分布は似通っている」と仮定し，既に推定された他の解の近傍解コスト分布に基づいた推定を行う．具体的には，他の解の近傍解コスト分布に基づいて，注目している解の近傍解コスト分布の事前分布を想定し，ベイズ推定を行う．

文献 [2] 等には，事前分布を仮説的な標本という形で与える方法が紹介されている．それらを参考に，本稿では以下のように事前分布を設定する．

まず，解 x_a, x_b 間の近傍 N に基づく距離（以後，単に距離と呼ぶ）$D_N(x_a, x_b)$ を次のように定義する．

$$\begin{cases} D_N(x_a, x_b) = 1 \quad (x_b \in N(x_a)\text{ のとき}) \\ D_N(x_a, x_b) = D_N(x_a, x) + 1 \\ \qquad\qquad\qquad\qquad (x_b \in N(x)\text{ のとき}) \end{cases}$$

これは，x_a から始めて，近傍解へと移動する操作を $D_N(x_a, x_b)$ 回行うことにより，x_b に至ることができることを意味している．なお，このような値が定義できない場合は $D_N(x_a, x_b) = \infty$ とする．

次に，近傍解コスト分布の類似度を表すパラメータ r $(0 < r < 1)$ を導入する．ここでは，解 x とその近傍解 $x' \in N(x)$ の近傍解コスト分布 $p(v|x)$，$p(v|x')$ が似通っている度合が r であると考える．すると一般に，2 つの解 x_a と x_b の近傍解コスト分布 $p(v|x_a)$，$p(v|x_b)$ の類似度 $R_N(x_a, x_b)$ は，解 x_a と x_b の距離 $k = D_N(x_a, x_b)$ を用いて，$R_N(x_a, x_b) = r^k$ とするのが妥当である．

解 x' に対して既に n 回の近傍操作が施され，n 個の標本（前節の d_i）が得られているとき，本稿では，解 x に対して $n \cdot R_N(x, x')$ 個の仮説的な標本を事前分布として与えることとする．すなわち，前節のように x の近傍解コスト分布を推定する際，x 自身に施された近傍操作より得られた標本数 n，総和 A，2 乗和 B の値をそのまま用いるのではなく，x' に対して得られた標本数，総和，2 乗和に $R_N(x, x')$ を乗じた値を加えて，計算を行う．そのような解が複数ある場合は，各解に対応する仮説的標本を全て合わせたものを事前分布と考えることができる．

4 改善解選択確率に基づく探索

改善解を効率良く見つけるという観点から，「近傍探索の過程において（得られた全ての解を保持しておき），それまでに求まっている解の中から最も改善解選択確率が高い解を選んで，近傍操作を行う」という近傍探索法を考えることができる．実際には全ての解を保持しておくのは不可能であるため，最大保持数 s を設定し，それを越えたときには（越えた分だけ）解を破棄する．最大保持数の解を保持している状態では，「多数の解を保持しておき，1 つずつ近傍解を生成しては入れ換える」という，連続世代モデルの進化的計算法と同様な処理の流れになる．

提案する手法の概要を以下に示す．

Step 1 集合 S を初期化する．ランダムに解 x_0 を生成し，S に入れる．x_0 を暫定最良解とする．

Step 2 S に含まれる解の改善解選択確率を求め，最も高い解 x を選択する．

Step 3 x に近傍操作を施し，近傍解 x' を得る．x' がそれまでの暫定最良解より良い解であれば，新たな暫定最良解とする．

Step 4 x' を S に入れる．S の要素数が設定された最大保持数を越えた場合は，S から 1 つ解を選んで破棄する．

Step 5 終了条件を満たしていれば，暫定最良解を探索結果として終了．そうでない場合は Step 2 へ．

5 実験結果

評価のため，提案手法，多スタート局所探索法，および提案手法の Step 2 で常に暫定最良解を選択するという手法を巡回セールスマン問題へ適用した．その結果の一部を以下に示す．

50 都市の座標をランダムに定めて問題とし，巡回経路のユークリッド距離を目的関数とした．巡回順序を表す都市の順列を解の表現として交換近傍を用いた．終了条件として，解の評価を 25,000 回行ったところで処理を打ち切ることとした．

提案手法の解の最大保持数は $s = 500$，近傍解間の類似度は $r = 0.8$ とした．また，解の保持数が s を越えた場合は，ランダムに解を選んで破棄した．多スタート局所探索法では即時移動戦略を用いた．

各手法を 50 回ずつ実行し，解の評価回数に対する平均コストを求めた．その結果を図 2 に示す．提案手法では，序盤から局所探索法と変わらない速度で探索が進み，最終的にも局所探索法よりも良い解が得られている．これは，多点探索を行った効果も大きいと思われるが，改善解選択確率により解選択を行うことにより，単に暫定最良解を用いるよりもさらに良い結果が得られていることがわかる．

6 まとめ

近傍探索法において効率良く探索を進めるために改善解選択確率について考察し，それに基づく多点探索法を提案・評価した．紙面の都合で詳細には触れないが，現在までにわかっていることを以下に述べる．

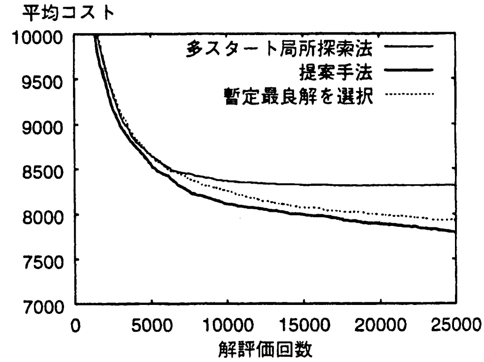

図 2: 各手法による解探索の様子

- 最大保持数を越えたときに破棄する解をどのように選ぶかは，探索の進行に大きく影響を及ぼす．解の選び方によっては，保持している解が均質化して探索が進まなくなる．
- ほとんどの場合，近傍解コスト分布は正規分布ではないと考えられる（巡回セールスマン問題の近傍解コスト分布を実際に調べたところ，探索の進行とともに形状が変化することがわかっている）．そのため，改善解選択確率には誤差が含まれる．

これらについては，今後の課題としたい．

参考文献

[1] 柳浦睦憲，茨木俊秀：組合せ最適化，朝倉書店 (2001).

[2] 繁桝算男：ベイズ統計入門，東京大学出版会 (1985).