1999年度日本オペレーションズ・リサーチ学会
秋季研究発表会

集合被覆問題に対する局所探索法について

02103204 京都大学 *岸田 正博 KISHIDA Masahiro
01704164 京都大学 柳浦 睦憲 YAGIURA Mutsunori
01001374 京都大学 茨木 俊秀 IBARAKI Toshihide

1 はじめに

集合被覆問題 (SCP) は代表的な組合せ最適化問題の一つであり, NP 困難であることが知られている. この問題に対しては様々な近似解法が提案されており, 代表的なものとしては, Beasley と Chu による遺伝アルゴリズム [2], Caprara, Fischetti, Toth や Ceria, Nobil, Sassano によるラグランジュ乗数に基づく重みを用いた欲張り法 [3][4], Jacobs と Brusco によるアニーリング法 [5] などがある. 本研究では, 同時に 3 つまでの集合を出し入れすることによって得られる解集合を近傍とする局所探索法を提案する. これは, 従来のアルゴリズムに比べ大きな近傍となっているが, 単純に近傍を広げただけでは効率が悪くなるので, 大きな近傍を扱う際には, 解の質を悪くすることなく効率的に近傍を探索する工夫を行っている. また, 実行不可能解も許すことによって, 探索の柔軟性を高めている.

2 集合被覆問題

集合被覆問題とは, 要素集合 $M=\{1,...,m\}$ と, Mの部分集合族 S_j, $j \in N=\{1,...,n\}$ が与えられたとき, Mの全ての要素をカバーするようにいくつかの集合を選び, 選んだ集合に付けられた重みの総和を最小にする問題である. 0-1 変数 $\boldsymbol{x} \in \{0,1\}^n$を用いると, この問題は以下のように定式化される.

$$\text{minimize} \quad cost(\boldsymbol{x}) = \sum_{j\in N} c_j x_j \tag{1}$$

$$\text{subject to} \quad \sum_{j\in N} a_{ij} x_j \geq 1,\ i \in M \tag{2}$$

$$x_j \in \{0,1\},\ j \in N \tag{3}$$

a_{ij}: 集合 S_jが要素 i をカバーするなら 1, そうでなければ 0

c_j: 集合 S_jの重み (c_jは正整数と仮定する)

x_j: 集合 S_jが解に含まれるなら 1, そうでなければ 0

3 ラグランジュ緩和問題の利用

集合被覆問題に対しては, ラグランジュ緩和問題を解くことによって得られる情報が非常に有用であることが知られており, 多くのアルゴリズムにおいて利用されている. 本研究では, ラグランジュ緩和問題から得られた情報を用いて, 最適解に含まれる可能性が高いと思われる集合だけを取り出すことにより問題サイズを縮小している. これは, 集合数 n が非常に大きい場合, 解の質の向上とアルゴリズムの高速化の両面において有効である.

制約条件 (2) に関するラグランジュ乗数ベクトル $\boldsymbol{v}\in R^m_+$ に対し,

$$c_j(\boldsymbol{v}) = c_j - \sum_{i\in S_j} v_i$$

は, 集合 S_j に関する相対コストと呼ばれる. これらを用いて, ラグランジュ緩和問題は

$$L(\boldsymbol{v}) = \min_{\boldsymbol{x}} \sum_{j\in N} c_j(\boldsymbol{v}) x_j + \sum_{i\in M} v_i \tag{4}$$

$$\text{subject to} \quad x_j \in \{0,1\}, \quad j \in N \tag{5}$$

と書かれる. ラグランジュ緩和問題の最適値 $L(\boldsymbol{v})$ は, 集合被覆問題の最適解 $\boldsymbol{x}^*$に対し, 必ず $L(\boldsymbol{v}) \leq cost(\boldsymbol{x}^*)$ を満たし, 集合被覆問題の下界値を与える. ラグランジュ双対問題は, 下界値 $L(\boldsymbol{v})$ を最大化するラグランジュ乗数ベクトル $\boldsymbol{v}^*$を見つける問題である. しかし, 最適なラグランジュ乗数ベクトル $\boldsymbol{v}^*$を計算するのは, 大規模な問題に対し非常に計算時間がかかる. そのため, 通常, 短い計算時間で近似的にラグランジュ乗数ベクトルを求めるために劣勾配法が利用される.

問題サイズの縮小は, n 個の集合 S_jの中から, 有効と考えられる一部の集合を残すことによって構成される. 本研究では以下の方法により, 探索の対象となる集合を選択した.

1. 得られたラグランジュ乗数ベクトルに対する相対コストが $c_j(\boldsymbol{v}) \leq \gamma$ (γは実数値を取るパラメータ) であるような集合 S_jを全て選ぶ.
2. 1 で選んだ集合のいずれにも含まれないような要素 i があれば, 要素 i を含む集合の中で相対コスト $c_j(\boldsymbol{v})$ が最小の集合を 1 つ選ぶ.

4 局所探索法

局所探索法は, 現在の解 $\boldsymbol{x}$ の近傍 $N(\boldsymbol{x})$ 内に $\boldsymbol{x}$ より良い解があればそれに置き換える操作を, 近傍内に改善解が

なくなるまで反復する方法である．近傍 $N(\boldsymbol{x})$ は解 $\boldsymbol{x}$ に多少の変更を加えて得られる解集合である．近傍内により良い解が存在しない解を局所最適解と呼ぶ．局所探索法の動作は，探索空間，解の評価関数，および近傍により決定される．以下では，これらの詳細について述べる．

4.1 探索空間と解の評価関数

本研究では，探索空間として $\{0,1\}^n$ を考える．すなわち，集合被覆問題の制約条件 (2) を満たさない実行不可能解も探索空間に含める．そこで，制約条件 (2) を考慮したペナルティ関数を以下のように定義し，評価関数に用いる．各要素 i に対する制約条件 (2) の制約違反を表す関数 $\theta_i(\boldsymbol{x})$ を，

$$\theta_i(\boldsymbol{x}) \;=\; \max\left\{\, 1-\sum_{j\in N} a_{ij}x_j,\ 0 \,\right\},$$

ペナルティ係数を $p_i(>0)$ とし，解の評価関数を，

$$pcost(\boldsymbol{x}) \;=\; \sum_{j\in N} c_j x_j + \sum_{i\in M} p_i\theta_i(\boldsymbol{x}) \qquad (6)$$

と定義する．ペナルティ係数 p_i は，探索の状況に応じて動的に変化させる．p_i の制御法については 4.3 節で述べる．

4.2 近傍とその探索

近傍 $N_h(\boldsymbol{x})$ を，$\boldsymbol{x}$ からのハミング距離が h 以下の解集合と定義し，N_3 を用いる．探索を効率化するため，$h\geq 2$ に対しては，$N_{h-1}(\boldsymbol{x})$ に改善解がない時に限り $N_h(\boldsymbol{x})$ を調べる．また，単純に $|N_h|=\binom{n}{h}$ 通りの可能性を全て調べるのではなく，改善の可能性のない解の探索を省くような工夫を加えている．

4.3 ペナルティ係数の更新

局所探索法を 1 度行っただけでは未探索の領域にさらによい解が隠れているという危惧が残る．また，本研究で提案している局所探索法は，探索空間に実行不可能解も含めているため，ペナルティ係数 p_i の値を十分大きくしない限り，1 度の探索で必ず実行可能解が得られるという保証はない．そこで，局所探索法が局所最適解を求めて停止した時点で各要素のペナルティ係数 p_i を更新し，前回の局所最適解を初期解としてさらに局所探索法を続けるという方法をとる．まず，p_i の初期値は

$$p_i := \min\left\{ \frac{c_j}{|S_j|} \;\middle|\; i\in S_j \right\}$$

と定める．次に，現在の局所最適解 $\boldsymbol{x}$ と暫定値 UB に対して，$cost(\boldsymbol{x}) < UB-1$ のときは，

$$p_i := p_i\left(1+\theta_i(\boldsymbol{x})\max\left\{\frac{UB-1-cost(\boldsymbol{x})}{UB},\ \Delta^+\right\}\right),$$

そうでないときは，

$$p_i := p_i\left(1+\min\left\{\frac{UB-1-cost(\boldsymbol{x})}{UB},\ \Delta^-\right\}\right)$$

と更新する．ここで，Δ^+ と Δ^- は，$\Delta^+>0,\ -1<\Delta^-<0$ をみたすパラメータである．

すなわち，$cost(\boldsymbol{x}) < UB-1$ ならば，この $\boldsymbol{x}$ は実行不可能解であるが (実行可能解であれば暫定値 UB が更新されるので不等式をみたさない)，$\boldsymbol{x}$ に集合を追加して実行可能解を得ることで，暫定解を更新する可能性があるので，カバーされてない要素のペナルティ係数を増加させる．逆に，$cost(\boldsymbol{x}) \geq UB-1$ ならば，$\boldsymbol{x}$ に集合を追加しても，暫定解は更新できないので，各要素のペナルティ係数を減少させる．

4.4 欲張り法の併用

前節で述べたペナルティ係数の更新ルールにより良質な実行可能解が得られる可能性は高くなるが，必ずしも十分な頻度で実行可能解を得ることはできない．しかし，局所探索法で得られた局所最適解は，少しの変形で良質な実行可能解にできると考えられる．そこで，本アルゴリズムでは，得られた局所最適解に対して以下の欲張り法を適用し，実行可能解を得る．すなわち，x_j の値を 0 から 1 に変化させたとき $pcost(\boldsymbol{x})$ の増加量を最小化するような集合 S_j を選び $x_j:=1$ とする，という操作を実行可能解が得られるまで繰り返す．さらに，得られた解に対して，実行可能解のみを探索空間とした，近傍 $N_h(\boldsymbol{x})$, $(h\leq 3)$ に基づく局所探索法を適用する．

5 まとめ

集合被覆問題に対し，より大きな近傍を探索する局所探索法に基づく近似解法を提案した．詳しい計算結果は当日発表する．

参考文献

[1] J.E. Beasley, "A Lagrangian Heuristic for Set Covering Problems," *Naval Research Logististics*, 37 (1990) 151-164.

[2] J.E. Beasley and P.C. Chu, "A Genetic Algorithm for Set Covering Problem," *European Journal of Operational Research*, 94 (1996) 392-404.

[3] A. Caprara, M. Fischetti and P. Toth, "A Heuristic Method for the Set Covering Problem," *Proceedings of the Fifth IPCO Conference*, Springer-Verlag, (1996) 72-81.

[4] S. Ceria, P. Nobili and A. Sassano, "A Lagrangian-Based Heuristic for Large-Scale Set Covering Problems," *Mathematical Programing*, 81 (1998) 215-228.

[5] L.W. Jacobs and M.J. Brusco, "A Local-Search Heuristic for Large Set-Covering Problems," *Naval Research Logististics*, 42 (1995) 1129-1140.