A-015

大規模なバイナリー2 次計画問題に対する Memetic Algorithm
An Memetic Algorithm for the Large Scale Binary Quadratic Programming Problem

脇坂 猛虎† 外山 史† 森 博志† 東海林 健二†
Taketora Wakisaka Fubito Toyama Hiroshi Mori Kenji Shoji

1. はじめに

バイナリー2 次計画問題(BQP)は，NP 困難な問題に含まれる組み合わせ最適化問題の 1 つであり，マシンスケジューリング問題や CAD 問題などに適用可能であるため，応用例が広い問題として知られている．BQP に対しては，多くの研究がなされており，特に Merz らの Memetic Algorithm(MA)[1]などによって非常に良い結果が報告されている．従来研究では BQP の良く知られたテスト問題として扱われている OR-Library[2]から入手可能な問題群が使用されているが，最大の問題サイズが 2500 変数となっている．これらの問題は近年の計算機の性能ならば数秒で実行可能なため，今後は OR-Library にはない大規模な問題に対して解の探索を行うことが重要であると考えられる．

本研究では，OR-Library にはない大規模な問題を対象とする．このような大規模な問題に対する従来手法として，IGKLS(Iterated Greedy algorithm with k-opt Local Search)[3]が提案されており，その有効性が示されている．そこで本研究では，従来研究として有効性のある MA と IGKLS を組み合わせることで，大規模な問題に対して更に効率の良い解の探索を可能とするアルゴリズムを提案する．

2. バイナリー2 次計画問題(BQP)

BQP とは$n \times n$の対称行列$Q = (q_{ij})$が与えられたとき，式(1)の目的関数を最大化する解xを求める問題である．ここで，解xは長さnのバイナリ列である．また，BQP の問題サイズはnの値によって表される．

$$f(x) = x^t Q x = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j \quad (1)$$

$$x_i \in \{0,1\} \quad \forall i = 1, \cdots, n$$

BQP において，解xにおけるk番目のビットx_kをフリップ(0 と 1 を入れ替える操作)して得られる解をx'とするとき，$\Delta = f(x') - f(x)$をゲインと呼び，ゲイン値$g_k = \Delta$は式(2)によって求められる．

$$g_k = q_{kk}(x_k'^2 - x_k^2) + 2(x_k' - x_k) \sum_{j=1, j \neq k}^{n} q_{jk} x_j \quad (2)$$

3. BQP に対する Memetic Algorithm

3.1 提案手法の概要

本研究では，大規模な BQP に対する新しい Memetic Algorithm を提案する．提案手法では，改良 IGKLS(H-IGKLS)[4]による単点探索を行う．H-IGKLS は従来の IGKLS[3]に変形 k-opt 局所探索法を組み込むことで，探索効率を向上させた手法である．次に，H-IGKLS で求めた最良解とランダムに生成した複数の初期解から親集団を構成し，多点探索の MA を適用する．これにより，大規模な解空間における更に効率の良い探索が可能となる．提案手法の流れを以下に示す．

1. 貪欲法による初期解の生成
2. H-IGKLS による解探索
3. 2.で求めた最良解とランダムな初期解で解集団を生成し，MA を適用
4. MA の処理を終了条件が満たされるまで実行

ここで，1.，2.の処理は H-IGKLS と同様のため，詳細は文献[4]を参照されたい．次節で，3.の MA の探索手法について述べる．3.，4.の処理において，最良解の更新状況に応じて解集団全体の収束を緩和するための再スタート戦略を適用する．

3.2 MA による探索処理

H-IGKLS による探索後の MA の処理の流れを以下に示す．

1. ランダムな初期解を複数個生成し，局所探索を行う
2. 局所探索を行った解と，H-IGKLS による最良解で親集団を構成
3. 親個体の選択，交叉処理
4. 生成した子個体で局所探索を行う
5. 親，子集団で，評価値の高い順に選択を行い，次世代集団を構成
6. 終了条件を満たすまで 3.~5.を繰り返す(3.~5.の処理の処理で一定回数評価値の更新が無い場合は 5.の直後に再スタート戦略)

提案手法で用いる局所探索は k-opt 局所探索法と変形 k-opt 局所探索法とした．k-opt 局所探索法は貪欲性の強い探索法で，従来の MA や IGKLS の局所探索法で用いられている．変形 k-opt 局所探索法はランダムな要素が含まれるため，探索パターンを増やすことができる特徴を持つ．3.の親個体の選択では，ランダムに 2 つの個体が選択される．ただし，解集団に含まれる親個体が重複なくすべての親が選択されるようにする．また 3.の交叉処理では，ハミング距離に応じて 2 種類の交叉法を用いた．具体的には，交叉をする親個体間のハミング距離が$n/10$を超えるときは Path-relinking[5]，$n/10$以下の時は Merz らによる Innovative Variation を適用する．これにより，交叉処理において解の探索パターンを広げることが可能になり，より解空間の探索効率が向上する．ここで，Path-relinking を行った際の局所探索は通常の k-opt 局所探索法とし，Innovative Variation[1]を行った際は変形 k-opt 局所探索法を使用する．これにより，ハミング距離が大きいときは収束を促し，小さいときは緩和を促す傾向を与えている．以下に，Path-relinking の処理の流れを示す．

†宇都宮大学大学院 工学研究科

1. 選択された親個体 A に対して，もう一方親個体 B とのハミング距離が小さくなるように 1 ビットずつ反転させる．このとき，最も評価値が良くなるビットを反転させ，1 ビットごとに解及びその評価値を保存しておく．
2. 親個体 B に対しても 1.と同様の処理を行う．ここで，ハミング距離を計算するもう一方の親個体は A となる．
3. 1.および 2.で求めた解の中での最良解が親個体の評価値より高ければその解を子個体とし，親個体より良い解が無い場合は親個体間のパスの中間地点の解を子個体とする．

なお，再スタート戦略については，一定の世代数最良評価値の更新が無かった場合，親集団に含まれる解すべてに対して *n*/3 ビットをランダムに反転することで，解の収束を緩和する．

4. 実験

大規模な問題に対する提案手法の有効性を示すために，従来手法の IGKLS[3] と，Merz らによる MA[1]，そして H-IGKLS[4]と比較実験を行った．本研究では OR-Library[2]にはない大規模な問題を対象としたアルゴリズムの探索効率の向上を目的としているため， 10,000 変数(n=10,000)の大規模な 10 個の問題を作成し，実験を行った．OR-Library にある問題の多くが行列 Q の密度(0 以外の値が含まれる割合)が 10%，値の取りうる範囲が[-100,100]であるため，この条件の下，ランダムに行列 Q を作成することにより，問題の生成を行った．実験はすべての手法を同じ計算機上(CPU:Intel(R) Xeon(R) X5650 2.67GHz，メモリ:64GB)で実行し，処理時間は 500 秒とし，試行回数は各問題に対して 10 回とした．実験で用いた MA の親集団の数は 50，子集団の数は 25 とし，30 世代連続して最良解の更新が無い場合は再スタート戦略を行う．

表 1 に，各問題名，従来手法と提案手法で得られたすべての解の中での最良評価値，各手法によって得られた 10 回の試行における最良評価値との差分を示す．表 1 より，最良の評価値は全ての問題に対して提案手法から得られた．

表 2 に，各問題名，従来手法と提案手法で得られたすべての解の中での最良評価値，各手法によって得られた 10 回の試行で得られた評価値の平均との差分を示す．表 2 より，10 種類中 7 種類の問題に対して提案手法が最も平均評価値が高い結果を得られた．

表 1: 10,000 変数の問題に対する最大評価値の比較

問題	最良評価値	評価値の差			
		MA	IGKLS	H-IGKLS	提案手法
1	11076178	125	3379	224	**0**
2	10846492	2522	3751	1085	**0**
3	11053493	1468	2772	350	**0**
4	11048669	484	2236	393	**0**
5	11041015	41	1556	618	**0**
6	11145449	28	51	142	**0**
7	10869303	3441	2030	63	**0**
8	11021727	1723	4065	64	**0**
9	10846303	2059	58	285	**0**
10	10904146	837	3615	560	**0**

表 2: 10,000 変数の問題に対する平均評価値の比較

問題	最良評価値	評価値の差			
		MA	IGKLS	H-IGKLS	提案手法
1	11076178	2931	4634	1206	**852**
2	10846492	5148	5597	**1560**	1873
3	11053493	4690	4451	1792	**1672**
4	11048669	2142	5208	959	**731**
5	11041015	1345	1961	995	**452**
6	11145449	2084	2822	1217	**833**
7	10869303	5819	3880	1955	**1250**
8	11021727	4610	5886	**1500**	2633
9	10846303	8100	3352	**1622**	1692
10	10904146	2606	6187	1007	**585**

5. 考察

3 つの従来手法と比べて提案手法は，最大評価値を全ての問題で算出することができたため，提案手法は大規模な問題に対して有効な結果となった．平均評価値に関しては，全種類の問題で従来手法を上回ることができなかった．この結果に関しては，MA を行うまでの探索の要となる H-IGKLS において十分な解の探索が行えていなかったことが大きな要因と考えられる．そのため，探索の前半部分となる H-IGKLS でより効率の良い探索を行える工夫が必要となる．また，今回の実験では Merz らによる Innovative Variation による交叉処理を用いたが，他の交叉処理の方が大規模な問題に関しては効果がある可能性もあるので，交叉処理を他の手法に差し替えて実験を行う必要があると考えられる．

実験では，解集団の収束を緩和するために再スタート戦略を行っているが，収束していく傾向自体は探索にとって必要な要素である．選択した親個体におけるハミング距離の親集団全体での平均値が *n* の 1%以下になった場合でも解の更新は行われることはあるため，再スタート戦略を行うタイミングは，ハミング距離の平均値に依存させた場合も検討を行う必要があると考えられる．

6. おわりに

本研究では、大規模なバイナリー2 次計画問題(BQP)に対するアルゴリズムとして，IGKLS と MA を組み合わせた手法を提案した．実験では，10,000 変数の大規模な BQP に対し従来手法との比較を行うことで，提案手法の有効性を確認した．

参考文献

[1] Peter Merz, Kengo Katayama, “Memetic Algorithm for the unconstrained binary quadratic programming”, problem”,Bio Systems, 78, pp99~118, 2004

[2] OR-Library
http://people.brunel.ac.uk/~mastjjb/jeb/orlib/bqpinfo.html

[3] 村上 剛基, 外山 史, 東海林 健二, 宮道 壽一, “バイナリー2 次計画問題に対する反復貪欲法”, 電気学会論文誌 C，Vol‘130, No. 6 , pp1089-1090, 2010

[4] 岡野 翔哉，外山 史, 森 博志，東海林 健二，“大規模なバイナリー2 次計画問題に対する反復貪欲法の改良”, 情報処理学会第 79 回全国大会, 3H-01, 2017

[5] Yang Wang, Zhipeng Lü, Fred Glover, Jin-Kao Hao, “Path relinking for unconstrained binary quadratic programming”,European Journal of Operational Research 223 , pp595~604, 2012