様　式　C－19、F－19－1、Z－19　(共通)

科学研究費助成事業　　研究成果報告書

令和　元 年　 5 月 29 日現在

機関番号：12608
研究種目：基盤研究(C)（一般）
研究期間：2015～2018
課題番号：15K00032

研究課題名（和文）錐最適化理論を用いた種別構成問題に対する効率的な計算手法の構築

研究課題名（英文）An Efficient Numerical Method for Optimal Contribution Problem based on Conic Optimization

研究代表者
山下　真（Yamashita, Makoto）

東京工業大学・情報理工学院・准教授

研究者番号：20386824

交付決定額（研究期間全体）：（直接経費）　2,700,000円

研究成果の概要（和文）：種別構成問題は混合整数二次錐最適化問題として定式化できる。本研究では、錐最適化アプローチに基づいて効率的なアルゴリズムを構築し、高速な計算手法を開発した。二次錐分割による線形近似では、最適解が得られる理論的枠組みの中で、従来提案されてきた計算手法の10分の1程度に計算時間を短縮した。また、steep ascent method は、最適解の保証はないものの良質な解を数秒程度という短時間で得ることを可能とした。

研究成果の学術的意義や社会的意義
種別構成問題は、育種学などから生じる数理最適化問題である。例えば「数千種の遺伝子から割合を決めて樹木園を作るときに、遺伝子の多様性を維持できる範囲で利益を最大にする割合の算出」の計算などを数理モデル化したものである。この計算時間が短縮されることで、遺伝子のパラメータなどを変えての複数回のシミュレーションの実施が容易になる。また、混合整数二次最適化問題は種別構成問題以外にも多くの応用を持っていることが知られており、本研究で開発した高速な計算解法は、それらの応用にも適用可能な範囲があると考えられる。

研究成果の概要（英文）：Optimal contribution problems arising from tree breeding or other optimization problems can be formulated as a mixed-integer second-order cone programming problem. We propose fast numerical methods that exploits conic optimization approaches. The proposed linear approximation reduces the computation time to about 1/10 of an existing method for obtaining exact solutions. Another new method, a steep ascent method, cannot theoretically guarantee the optimality, but it successfully outputs a favorable solution in several seconds.

研究分野：数理最適化

キーワード：応用数学　数理最適化　錐最適化　育種学　種別構成問題

１．研究開始当初の背景

（１）　種別構成問題は、育種学や畜産学、金融工学などに現れる最適化計算をモデル化した数理最適化問題である。特に、多様性を維持するための制約が二次錐により表現されることが従前の研究により指摘されており、決定変数が連続変数の場合には錐最適化理論を用いると短時間で求解可能である。

（２）　育種学などで実用的に利用するためには決定変数が整数制約であるモデルに対する計算手法への拡張も必要である。しかし、種別構成問題を混合整数二次錐計画問題として捉え汎用的ソフトウェアを適用してしまうと、小規模な問題にも長時間の計算時間を要することがあり、パラメータを変更しながら複数回のシミュレーションを行うなどが困難であった。このことから、種別構成問題への効率的な計算手法の構築が重要な課題として認識されている。

２．研究の目的

（１）本研究では、種別構成問題が二次錐で表現されることに着目し、二次錐計画問題や半正定値計画問題などを対象とする錐最適化理論を用いることで、種別構成問題に対する効率的な計算方法の構築を目的とした。

（２）特に、育種学における二次錐では、データ行列が疎行列であるという特徴を持っており、この特徴を利用してアプローチできれば、他の混合整数二次錐計画問題への適用可能性も高まる。疎行列を活用した計算方法の構築も本研究では重視した。

３．研究の方法

（１）育種学における既存手法では、Outer Approximation に基づいて多数の線形制約により二次錐近似を行っている。本研究では、n 次元の二次錐が n 個の 3 次元に分割できる、という数学的特徴を利用することによって、線形制約の生成方法を新しく提案する。また、線形制約についても従来は劣勾配を用いて計算を行っていたが、本研究では計算効率を向上させるために直交射影を用いて線形制約を生成することとする。

（２）混合整数計画問題の近似解を得る方法として半正定値計画緩和手法や二次錐緩和手法など錐最適化理論に基づく緩和手法が知られており、本研究では、これらの緩和手法により得られる近似解の評価を行った。また、これらの近似解を用いることで、種別構成問題に対する良好な解を得る高速な計算方法の構築も行った。

４．研究成果

（１）二次錐分割による線形近似

育種学における種別構成問題では$x^T A x \leq 2\ \theta$という凸２次制約により多様性維持の制約が課されており、行列Aの正定値性を利用して二次錐として表現することが可能である。従来の Outer Approximation では、切除平面法を利用した反復法であり、ある反復の点x^kが$x^T A x \leq 2\ \theta$を満たしていない場合に、劣勾配に基づいて$2(Ax^k)^T x \leq (x^k)^T A x^k + 2\theta$という新しい線形制約を追加して再度混合整数計画問題を解く、という反復法であった。この方法は有限回数の反復で厳密解が得られるという理論的保証があるという点でそれまでの既存手法に対する優位性があったが、計算時間が長時間になるという課題があった。特に、各反復で追加される線形制約が一つのみであることから、厳密解を得るまでに多くの反復が必要である。

本研究では、n 次元二次錐が n 個の 3 次元二次錐によって表現可能なことに着目した。一般に n 次元二次錐は$x_1 \geq \sqrt{\sum_{i=2}^n x_i^2}$という非線形制約式で表現されるが、数学的には以下のような制約と等価である。

$$x_1 \geq \sum_{i=2}^{n} w_i\,,\ w_i x_1 \geq x_i^2 \,(i\,=\,2, \dots, n),\ w_i \geq 0\ (i\,=\,1, \dots,\,n)$$

特に、$w_i x_i \geq x_i^2$は回転しているが二次錐の一種である。したがって、ある反復の点で多様性の制約$x^T A x \leq 2\ \theta$を満たされないときに、複数の回転二次錐を満たしていないと一般には考えられる。つまり、満たされていない回転二次錐の数だけ、新しく線形制約が生成可能である。各反復で追加される線形制約が増える分だけ、全体の反復回数を抑える効果が期待できる。

また、回転二次錐に関する線形制約についても劣勾配から直交射影に本研究では変更することとした。多様性の制約である$x^T A x \leq 2\ \theta$に対しては、線形制約を直交射影により生成しようとしても解析的には求まらず、別途何らかの反復計算を導入する必要がある。しかし、本研究で

は二次錐を分解したことにより、直交射影の対象となる二次錐は低次元である。直交射影の計算は 3 次多項式に帰着され、Cardano の公式の適用により、解析的に線形制約を得ることが可能である。

以上の二次錐分割と直交射影により、本研究では Cone Decomposition Method (以下、CDM) を新規に開発した。育種学の種別構成問題に対して計算時間をまとめたのが以下の表 1 である。

表１：CPLEX，OPSEL，CDM の計算時間（秒）

m	CPLEX	OPSEL	CDM
200	8735.24	606.94	9.89
1050	6.64	11654.61	12.83
2045	3.83	14.42	2.61
5050	10810.49	10873.73	179.05
10100	11657.20	8787.37	1431.12
15222	10979.49	OOM	1111.49

ここで、mは種数である。例えば$m = 200$であれば、200 種の中から$N = 50$種を選抜する問題である。CPLEX は混合整数二次錐計画問題に対する汎用ソルバーであり、OPSEL は Outer Approximation に基づくソフトウェアである。また、OOM は 8GB 以上が必要となりメモリ不足 (out of memory) で異常終了したことを示す。各ソフトウェアの計算時間の上限は 10800 秒(3 時間)に設定してある。

表 1 を見ると分かるように CDM は 3 つの方法の中で高速な求解に成功している。例えば、$m = 5050$では、CPLEX や OPSEL と比較しても 60 分の 1 程度の計算時間での求解を達成しており、十分な高速化を得ている。

また、本研究で提案した Cone Decomposition Method も有限の反復回数で厳密解が得らえるという理論的な性質を保持している。

（２）錐最適化による緩和手法と Steep ascent method の構築

混合整数計画問題の近似値を得るうえで、錐最適化を用いた緩和手法が有効であることが知られている。特に、半正定値計画緩和は二次錐緩和よりも良好な近似値が理論的には得られる。本研究においても半正定値計画緩和の近似値の精度を評価した。その結果をまとめたのが以下の表 2 である。

表 2：種別構成問題に対する半正定値計画緩和の数値評価

m	下限	平均値	上限
200	16.161	25.812	30.340
1050	5.075	32.305	112.600
2045	279.259	446.089	2007.212
5050	5.775	284.965	806.205

厳密な最適値は下限と上限の間に存在する可能性が高いとされるが、下限と上限の差が大きいことから、半正定値計画緩和を用いても十分な近似値を得られないことを本研究では数値的に示した。

次に、本研究では錐最適化緩和手法によって得られた解を利用する Steep ascent method を開発した。均等展開の場合の種別構成問題においては、各変数は$x_i \in \{0,1\}$とバイナリー変数である。半正定値計画緩和あるいは二次錐緩和では、バイナリー変数を$0 \leq x_i \leq 1$のように上下限制約の変数に置換するなどして近似値を得る。Steep ascent method では、上下限制約で求めた解を参考にして、まずは、各変数を$x_i \in \{0,1\}$に割り当てる。以降は、$x_i = 0$となっている変数x_iと$x_j = 1$となっている変数x_jを入れ替えて目的関数値が向上すれば、その入れ替えを維持して、また別の変数の交換を行う。ここで用いる目的関数には、多様性維持のための二次錐制約からの違反量をペナルティとして元々の目的関数に加えたものを採用している。この Steep ascent method は必ずしも厳密な最適解が得られる保証はないが、局所的には最適解となっており、良好な解を得られる可能性が高い。

錐最適化緩和手法と Steep ascent method の計算時間を示すのが、次の表 3 である。

表 3：錐最適化緩和手法と Steep ascent method の計算時間

	目的関数値		計算時間（秒）	
m	半正定値計画緩和	二次錐計画緩和	半正定値計画緩和	二次錐計画緩和
200	25.207	25.090	1.30	0.06
1050	24.846	24.831	27.96	0.09
2045	438.457	438.386	145.59	0.09
5050	42.431	42.691	2221.40	0.37
10100	44.662	46.568	5582.46	0.87
15222	460.409	460.769	17441.93	2.56

表 3 より、半正定値計画緩和を用いても二次錐計画緩和を用いても最終的に得られる目的関数値に極端な差がないことが見て取れる。緩和の時点では半正定値計画緩和は二次錐計画緩和よりも優れた近似値を算出するが、Steep ascent method の初期点という観点からは 2 つの緩和は同等程度であると考えられる。

また、計算時間に着目すると、二次錐計画緩和のほうが極めて短時間で計算可能である。二次錐計画緩和では、多様性維持に制約に表れる行列Aの疎性を活用できることが計算時間短縮に直結している。さらに半正定値計画緩和では大規模な$m = 10100, 15222$では計算途中で数値的不安定性が表面化したが、二次錐計画緩和ではそのような振る舞いは観察されなかった。疎性の活用により求解すべき問題のデータ量が削減されたことが数値的安定性の向上に寄与したと推察され、種別構成問題に対する二次錐計画緩和の持つ優位性の一つと捉えられる。

CPLEX，OPSEL と Steep ascent method（以下、SA）を比較した結果が表 4 である。SA については上記の結果から二次錐計画緩和を利用したものを採用している。なお、表 1 とは異なるパラメータを利用しているため CPLEX，OPSEL の結果は表 1 と同じではない。また、CPLEX，OPSEL は誤差が 1%未満となるか 10800 秒（3 時間）を経過した時点で計算を終了し、その時点での目的関数値を記している。（CPLEX の一部の結果では 3 時間経過までに目的関数値が得られなかったものがある。）

表 4：CPLEX，OPSEL，SA の比較結果

	目的関数値			計算時間（秒）		
m	CPLEX	OPSEL	SA	CPLEX	OPSEL	SA
200	25.190	25.191	25.090	4270.77	1779.13	0.06
1050	--	24.858	24.831	> 10800	> 10800	0.09
2045	436.213	435.826	438.386	0.37	16.08	0.09
5050	42.456	42.702	42.691	2.02	> 10800	0.37
10100	--	46.252	46.568	> 10800	> 10800	0.87
15222	459.135	459.040	460.769	39.20	> 10800	2.56

目的関数値を比較すると CPLEX，OPSEL，SA は、ほぼ同等の目的関数値を得られている。CPLEX，OPSEL が厳密解を求める手法であり、SA は厳密解を得られる保証がないため、目的関数値のみで比較することは困難であるが、SA によって得られた解が実用的には十分と考えられる良好な解であることが示されている。また、SA の計算時間は、CPLEX や既存手法の OPSEL と比較しても短時間である。最大の問題であっても 3 秒程度で解が得られている。

育種学における種別構成問題では、パラメータなどを変更しながら多数回にわたり求解を行う必要があり、本研究で構築した Steep ascent method は計算効率の観点から効果が高いと考えられる。

（３）上記２点に加えて、錐最適化問題を解く主要なアルゴリズムである内点法についても理論的解析や数値実験による評価を行った。

５．主な発表論文等

〔雑誌論文〕（計４件）

① Sena Safarina, Satoko Moriguchi, Tim J. Mulllin, and Makoto Yamashita, “Conic relaxation approaches for equal deployment problems,” To appear in Discrete Applied Mathematics, 査読有.

② Makoto Yamashita, Tim J. Mullin, and Sena Safarina, “An efficient second-order

cone programming approach for optimal selection in tree breeding," Optimization Letters, 査読有, 2018, Vol. 12, No.7, pp1683-1697, DOI: 10.1007/s11590-018-1229-y.
③ Yaguang Yang and Makoto Yamashita, "An arc-search O(nL) Infeasible-Interior-Point Algorithm for Linear Programming," Optimization Letters, 査読有, 2018, Vol. 12, No. 4, pp 781-798, DOI:10.1007/s11590-017-1142-9.
④ Sena Safarina and Makoto Yamashita, "An Application of Polyhedral Relaxations to Optimal Contribution Selection of Tree Breeding Problem," RIMS Koukyuroku, 査読無, 2018, No. 2069, pp 62-73.

〔学会発表〕(計８件)

① Makoto Yamashita and Kei Takemura, "A path-following method for semidefinite programming without Slater condition," International Symposium on Mathematical Programming 2018, University of Bordeaux (Bordeaux, France), 2018年7月5日.
② Sena Safarina and Makoto Yamashita, "Cone Decomposition Method for Mixed-Integer SOCP arising from tree breeding," International Symposium on Mathematical Programming 2018, University of Bordeaux (Bordeaux, France), 2018年7月4日.
③ Sena Safarina, 山下真,「最適構成問題に対する LPP 緩和に基づいた整数計画問題による定式化」, 京都大学数理解析研究所研究集会 「数理最適化の発展：モデル化とアルゴリズム」, 京都大学数理解析研究所 (京都, 日本), 2017年8月24日.
④ Makoto Yamashita, Tim J Mullin, and Sena Safarina, "Steep-Ascent Method for MI-SOCP arising from Tree Breeding," SIAM Conference on Optimization 2017, Sheraton Vancouver Wall Centre (Vancouver, Canada), 2017年5月24日.
⑤ Sena Safarina, Makoto Yamashita, "Lifted-Polyhedral-Programming Approach for Optimal Contribution Problems," SIAM Conference on Optimization 2017, Sheraton Vancouver Wall Centre, (Vancouver, Canada), 2017年5月24日.
⑥ Makoto Yamashita, Tim J Mullin, and Sena Safarina, "A Fast SOCP-based Method for Optimal Selection Problem in Tree Breeding," INFORMS Annual Meeting 2016,
Music City Center & Omni Nashville Hotel (Nashville, USA), 2016年11月16日.
⑦ Makoto Yamashita, Sena Safarina, and Tim J. Mullin, "Mixed-integer SOCP in optimal contribution selection of tree breeding," Workshop on Advances in Optimization, TKP Shinagawa カンファレンスセンター (品川, 東京), 2016年8月12日.
⑧ Sena Safarina, Tim J. Mullin, and Makoto Yamashita, "An Efficient Second-Order Cone Programming Approach for Optimal Selection in Tree Breeding," ICCOPT (International Conference on Continuous Optimization) 2016 Tokyo, 政策研究大学院大学 (六本木), 2016年8月8日.

〔図書〕(計０件)

〔産業財産権〕
○出願状況（計０件）

○取得状況（計０件）

〔その他〕

データおよび計算ファイルの一部を以下のホームページにおいて公開
Data and source code for optimal contribution problem from tree breeding
http://www.opt.c.titech.ac.jp/yamashita/software/OCP/OCP01.html

６．研究組織

(1)研究分担者
なし

(2)研究協力者
研究協力者氏名：福田　光浩, 小島　政和、中田　和秀, Kim Sunyoung, Mullin Tim J., Safarina Sena
ローマ字氏名：(FUKUDA, mituhiro), (KOJIMA, masakazu), (NAKATA, kazuhide), (KIM sunyoung), (MULLIN, tim j.), (SAFARINA, sena)

ついては、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。