

令和元年6月25日現在

機関番号：32641

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K01204

研究課題名(和文) データ駆動型最適化モデルに対する統合的アプローチの探求

研究課題名(英文) Exploration of integrated approach to data-driven optimization models

研究代表者

後藤 順哉 (Gotoh, Jun-ya)

中央大学・理工学部・教授

研究者番号：40334031

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：確率計画やロバスト最適化といった不確実性下の意思決定のための最適化手法に、リスク管理、機械学習、信号処理といった文脈において発展してきた方法論を融合することで、いくつかの汎用的な方法論を提示した。得られた成果のうち中心的なものは2つある。1つ目は分布的ロバスト最適化モデルが、その内包するパラメータが十分小さいときには、いわゆる平均・分散モデルで近似できることから、パラメータのチューニングについての示唆を得たこと。2つ目は統計モデルの変数選択などの文脈で現れる基数制約について、連続最適化の表現とアルゴリズムを提示したことである。

研究成果の学術的意義や社会的意義

不確実性を考慮した意思決定の重要性は容易に認識されるが、その方法論は様々な形をとる。本研究課題の成果は最適化モデリングという、1つのレンズを通して、一見異なるものを眺めることで、既に多くの研究者コミュニティにおいてあまり疑問を持たれずに利用されている方法に対し、中立的立場から検討を行い、汎用性のある手法を提示したものとする。これにより、既存の方法論の限界や欠点、特徴が理解しやすくなると同時に、新たな方法論の創出にも寄与する可能性があると考えられる。

研究成果の概要(英文)：By combining optimization methods for decision making under uncertainty such as stochastic programming and robust optimization with methodologies developed in various fields and contexts such as risk management, machine learning, and signal processing, we provide general-purpose methods. Two outcomes form the major part of obtained results. For the first thing, based on the observation that a class of distributed robust optimization models can be approximated by the so-called mean-variance optimization model when a hyper-parameter is sufficiently small, we got suggestions for how to tune the parameter. The second is a method for presenting the so-called cardinality constraint that appears in contexts such as variable selection of statistical models by using a continuous optimization formulation. We also studied local search algorithms for those problems.

研究分野：数理最適化、オペレーションズリサーチ

キーワード：数理最適化 データ解析 ロバスト最適化 機械学習

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

実社会・工学に現れる最適化モデルの多くは確率変数を含む関数の期待値の最適化問題として定式化できる。しかし、通常、真の確率分布が未知なために、厳密な最適化は不可能である。代替手段として標本平均近似 (SAA) の最適化が行われるが、得られる最適解は真の最適解から離れ、事後パフォーマンスは劣化する。この問題は分野を超越して普遍的に現れるが、どのようにモデル化するかに絶対的な基準があるわけでもなく、解決策は各分野で個別に検討されてきた。SAA を補強するモデリングとして数理最適化分野で過去 20 年ほど人気があるのがロバスト最適化である。これは、最適化モデル内の確率変数や推定値を集合 (不確実性集合) として与え、その集合上で最も都合の悪い値に対する最適化を行う。中でも、想定している確率分布そのものの不確かさを扱った分布的ロバスト最適化が近年注目を集めている。しかしながら、不確実性集合はユーザ (意思決定者) が与える必要があり、どのように集合を与えたら得られる最適解の事後パフォーマンスが改善するのかといった理論は発展途上である。

2. 研究の目的

本研究では分布的ロバスト最適化を出発点とし、事後パフォーマンスの確率的保証を有したデータ駆動型最適化モデリングの枠組みを確立することを目標とした。具体的には、当時までに得られていた成果・準備を踏まえ、以下のような段階を計画していた。

- (1) 既存の分布的ロバスト最適化手法の事後パフォーマンスに対する理論の構築
- (2) (1)の一般的拡張。特に、CVaR に基づくロバスト化の理論の構築
- (3) 分布的ロバストな、新しい統計的モデル推定・機械学習手法の提案
- (4) 非凸正則化と非凸リスク尺度への拡張

(1)では、機械学習における汎化理論をベースにして、Kullback-Leibler ダイバージェンスを用いた基本的な分布的ロバスト最適化の事後パフォーマンスに対する理論的な裏づけを導くことを想定していた。(2)では、(1)の理論的枠組みを、凸解析や凸リスク尺度の理論 (Rockafellar & Uryasev, 2013 など) を用いて、より一般的な枠組みへと拡張、とりわけ CVaR の新たな利用法を提示することを想定していた。(3)では、(1)(2)の結果を踏まえ、在庫管理、ポートフォリオ選択などの経営科学、OR の問題の他、最小二乗法、最尤法、SVM などの統計的推定の既存手法に対する分布的ロバストな新しい方法の提案を目指していた。(4)では、(1)~(3)を踏まえ、信号処理やポートフォリオ選択などで利用される l_0 ノルムなどの非凸型正則化やバリューアットリスク (VaR) などの非凸リスク尺度へと議論を拡張することを想定していた。

3. 研究の方法

4. に示すように、大きく 2 つのトピックが主成果と言えるが、いずれも、(1)モデリング (定式化) 及び理論的分析と(2)数値計算に基づく実証分析の占める割合がほとんどである。具体的には以下に示すように、いずれも複数名による共同研究の形で実施し、流動的に役割を分担しながら進めた。

(1) 4. の(1)の課題について、特に、ロバスト最適化モデリングに対する理論的分析については、シンガポール国立大学の Andrew E.B. Lim 教授、ブリティッシュコロンビア大学の Michael J. Kim 助教と議論を行い、共著論文執筆を行うなど、国際的共同研究の形で行った。3名の居住地が離れていることもあり、不定期に skype や eメールによる議論、Dropbox による作業ファイルの共有など、インターネット技術も活用して進めた。4. の(2)の課題については、非凸最適化問題としての等価な定式化、アルゴリズムの構築について、東京大学の武田朗

子教授、東野克哉氏（当時、同大修士課程学生）と直接集まって議論を行い、eメールでの議論や作業ファイルのDropboxでの共有を行うなどしながら共著論文執筆を行った。

(2) 4.の(1)の数値計算例については、後藤が担当し、商用の数値最適化ソルバーRNUOPTを用い、ラップトップ上で実装・数値計算、および、そのとりまとめを行った。4の(2)の課題の計算・実証分析については、提案アルゴリズムを計算機上で実装し、数値計算ソフトMATLABを利用して求解を行った。計算機への実装・実験は東野氏が行った。

4. 研究成果

研究成果は大きく以下の2つのトピックにまとめられる。

(1) 分布的ロバスト最適化の定式化に関する理論的考察

(2) スパース最適化に対する非凸型正則化を用いた定式化とアルゴリズムの構築

これらについて、以下、やや詳しく記す。

(1) 1.に記したように、分布の不確かさを陽に考慮した方法論として、分布的ロバスト最適化(DRO)がある。これは想定する分布の範囲を予め集合(不確実性集合)や、ダイバージェンスによる罰則項により与えておいて、その想定範囲内での最悪な分布の下での標本平均に基づき目的関数を構成するアプローチであり、最も慎重・保守的な評価基準に基づく最適化モデルと言える。現実に応用するには、不確実性集合の大きさ、あるいは、ダイバージェンスの罰則を規定するあるハイパーパラメータによってそれを決める必要がある。そのパラメータの設定については、これといった基準がないまま、各ユーザに委ねられてきた。

本課題では滑らかなダイバージェンスによる罰則項を用いたDROを考え、その罰則パラメータが、得られるDRO解に与える感度について理論的な検討を加え、実際的な問題に対して数値計算を行い、その妥当性について検討を行った。当該DROの目的関数はパラメータ δ が0のとき、いわゆるSAAに等しく、 δ を大きくしていくと、より都合の悪い分布を想定することに対応する。そこで、 δ が十分に小さいとき、当該DROの目的関数が、経験分布に基づく期待利得(=SAAの目的関数)と利得の分散(に δ と定数を掛けたもの)の和でよく近似できること、 δ を0から大きくしていくと、標本の期待利得も分散も小さくなるのが期待できるが、特に、 δ が小さいときには分散の減少の方が、期待利得のそれよりも大きく減少することが期待できることを理論的に示した。また、実用的なパラメータ設定の方法として、リサンプリング法(具体的にはブートストラップ法)により期待利得と分散のフロンティアを推定し、そこから δ を選択するという方法の提案および数値例の提示を行った。

最近他の研究者から提示された方法は、不確実性集合が含まれる確率が信頼水準を満たすようにパラメータを設定するというものであった。しかしこの方法は信頼水準として95、99%といった、統計の検定や推定で慣習として用いられる数字を想定しているが、それらを用いると、過度に保守的な結果になってしまうこと、そもそも事後的な期待利得について何も言っていないことなどの欠点があると考えられる。それに比べ、本研究で提示した方法は、データ標本からリサンプリング法を用いるというデータ駆動的であり、(最終的な選択をユーザに委ねながらも)平均と分散のトレードオフという形に帰着し、直感的に決定できるという点で、実用的な示唆を与えると考えられる。この結果は数度の投稿、改訂を経て、Gotoh, Kim, Lim (2018)として、査読付き国際学術雑誌 Operations Research Letters に収録された。

(2) 機械学習や信号処理などの分野で、 l_1 ノルムの最小化に基づくスパース最適化モデルがここ10数年注目を集めている。たとえば統計モデルに採用するパラメータ数を一定値、たとえば K 以下に抑えるために、統計モデルのパラメータベクトルの l_1 ノルムを最小にすることで、

効率的に行える。一方、本来、選択する個数を K 以下に抑える（基数制約）を課す立場からすると、 l_1 制約は必ずしも K 以下を達成する保証はない。これは l_1 ノルムの持つ凸性が、基数制約を持つ非凸性を無視することに起因していると言える。

本課題では、 l_1 ノルムに代えて、 l_1 ノルムと最大 k ノルムと呼ばれる別のノルムとの差で表される非凸正則化項を考え、対応する正則化つき最適化が基数制約付き最適化と等価になる条件について明らかにした。得られた問題は、DC と呼ばれる構造を持つ、いわゆる非凸型の問題であり、一般に大域的最適解を（保証付きで）求めるのは困難である。そこで、本課題では、DC という特性を活かした効率的な局所探索アルゴリズム（PDCA）を提示し、数値実験によりその有効性を示した。特に、DCA と呼ばれる、DC 構造を持つ問題に対する一般的な局所探索アルゴリズムでは、各反復において凸型最適化問題を解くことが必要なため、機械学習などで現れる大規模問題に対しては効率的とは言えない。そこで各反復の最適化計算が解析解を持つように工夫することで、効率性を指向したのが PDCA の特徴である。その結果、ある線形モデルの回帰推定問題に対し、既存の類似のアプローチよりも効率よく、質の良い解が得られる例を示すことができた。これらの結果は Gotoh, Takeda, Tono (2018) として、査読付き国際学術雑誌 *Mathematical Programming* に収録された。

最後に「2. 研究の目的」との関連について触れておく。

研究を進めていく中で、当初の目論見とは異なる部分も出たものの、概ね順調に進めることができたと考えられる。

2 .の(1)は機械学習でよく見られる有限サンプルの下でのノンパラメトリックなパフォーマンス保証を記述した汎化理論をベースに、ロバスト最適化の事後パフォーマンスに対する理論的保証を目指そうと計画したものであった。しかしながら、機械学習においてもそうであるように、これらの汎化理論が提示する不等式は非常に緩く、それをハイパーパラメータの決定に用いるというアイデアは実用的でないということ、実際に数値的実験から判断し、別の理論的分析へと方向転換を行った。（この事後のパフォーマンスに関する理論的結果については、課題期間中に論文の採択には至っていないが、論文投稿し、現在も改訂を繰り返している。）

一方で、2 .の(2)および(3)については、平均・分散による近似のアイデアを、Kullback-Leibler ダイバージェンスを含む ϕ -ダイバージェンスにも広げたほか、CVaR を含む一般的な凸リスク尺度に対する DRO もやはり、経験的リスクとある量の分散に分解できることを示した（Gotoh, Kim, Lim, 2018）。

(4)は4 .の(2)に記したように、DC という構造を持った正則化に関するいくつかの結果を得た（Gotoh, Takeda, Tono, 2018）。非凸型のリスク尺度については引き続き検討を続けている。

5 . 主な発表論文等

〔雑誌論文〕(計5件)

1. Wei, X., Gotoh, J., Uryasev, S., "Peer-to-Peer Lending: Classification in Loan Application Process," *Risks*, 査読有, 6 巻, 2018 年, 129, DOI: 10.3390/risks6040129 (オープンアクセス)
2. Gotoh, J., Kim, M.J., Lim, A.E.B., "Robust Empirical Optimization is Almost the Same as Mean-Variance Optimization" *Operations Research Letters*, 査読有, 46, (4), 2018 年, pp.448-452. DOI: 10.1016/j.orl.2018.05.005 (オープンアクセス)
3. Gotoh, J., Takeda, A., Tono, K., "DC Formulations and Algorithms for Sparse

Optimization Problems" Mathematical Programming Series B, 査読有, 169, (1), 2018年, pp.141-176. DOI: 10.1007/s10107-017-1181-0

4. Gotoh, J., Uryasev, S., "Support Vector Machines Based on Convex Risk Functions and General Norms," Annals of Operations Research, 査読有, 249, (1), 2017年, pp.301-328. DOI: 10.1007/s10479-016-2326-x
5. Gotoh, J., Takeda, A., "CVaR Minimizations in Support Vector Machines," Akansu, A.N., Kulkarni, S.R., Malioutov, D.M. ed., Financial Signal Processing and Machine Learning, Wiley-IEEE Press, 2016年, pp.233-265. DOI: 10.1002/9781118745540.ch10

〔学会発表〕(計 18 件)

1. 後藤順哉, 福田琢巳, ノイズのない場合の k -疎復元に対する刈込 l_p 関数を用いた定式化と ADMM の適用, 日本 OR 学会 2019 年度春季研究発表会, 千葉工業大学, 2019 年 3 月
2. 後藤順哉, Kim, M.J., Lim, A.E.B., 分布的口バスト最適化の事後分析とパラメータ選択, 日本 OR 学会 2018 年度秋季研究発表会, 名古屋市立大学, 2018 年 9 月
3. Gotoh, J., Kim, M.J., Lim, A.E.B., "Out-of-sample analysis of distributionally robust optimization," 23rd ISMP (International Symposium on Mathematical Programming) 2018 Bordeaux, France, 2018 年 7 月
4. 後藤順哉, 尾崎裕介, 平均-分散規準に基づくサポートベクターマシン, 日本 OR 学会 2018 年度春季研究発表会, 東海大学高輪キャンパス, 2018 年 3 月
5. 後藤順哉, ノルムを用いた最適化モデリング- $\alpha+\beta$, 日本 OR 学会東北支部 2017 年度 SSOR (招待講演) 蔵王, 山形, 2017 年 11 月
6. Gotoh, J., Misawa, Y., "A DC Optimization Approach to Sparse Spline Regression," INFORMS Annual Meeting 2017, Houston, USA, 2017 年 10 月
7. 後藤順哉, ノルムを用いた最適化モデリング, 日本 OR 学会 2017 年度秋季研究発表会, 関西大学, 2017 年 9 月
8. Gotoh, J., Kim, M.J., Lim, A.E.B., "Robust Empirical Optimization is Almost the Same as Mean-Variance Optimization," SIAM Optimization Conference 2017, Vancouver, Canada, 2017 年 5 月
9. 後藤 順哉, 三澤 祐一, DC 表現によるスパース・スプライン回帰, 日本オペレーションズ・リサーチ学会春季研究発表会, 2017 年 3 月, 沖縄県市町村自治会館
10. Gotoh, J., Takeda, A., Tono, K., "DC Formulations and Algorithms for Sparse Optimization Problems," DS Seminar, NUS Business School Decision Sciences Department, Singapore, 2017 年 3 月
11. Gotoh, J., Takeda, A., Tono, K., "DC Formulations and Algorithms for Sparse Optimization Problems," Workshop on Risk Management Approaches in Engineering Applications, Gainesville, Florida, USA, 2016 年 11 月
12. Gotoh, J., Kim, M.J., Lim, A.E.B., "Robust Empirical Optimization is Almost the Same

as Mean-variance Optimization," INFORMS Annual Meeting 2016, Nashville, USA, 2016年11月

13. Gotoh, J., "Conditional Value-at-Risk and Its Applications in Optimization," Department of Industrial and Systems Engineering, KAIST, Daejeon, Korea, 2016年9月
14. 後藤順哉, Kim, M.J., Lim, A.E.B., ロバスト最適化~平均分散最適化, 日本オペレーションズ・リサーチ学会春季研究発表会, 慶應義塾大学矢上キャンパス(神奈川県横浜市), 2016年3月
15. Gotoh, J., Kim, M.J., Lim, A.E.B., "Robust Empirical Optimization is Almost the Same as Mean-variance Optimization," Workshop on Risk Management Approaches in Engineering Applications, Gainesville, Florida, USA, 2015年11月
16. Gotoh, J., Kim, M.J., Lim, A.E.B., "Robust Empirical Optimization is Almost the Same as Mean-variance Optimization," INFORMS Annual Meeting 2015, Pennsylvania Convention Center and Philadelphia Marriott Downtown Hotel, Philadelphia, USA, 2015年11月
17. 後藤順哉, 山田雄二, 単調な一般化加法モデルの二次錐制約を用いた定式化, 日本オペレーションズ・リサーチ学会 2015年秋季研究発表, 九州工業大学戸畑キャンパス, 2015年9月
18. Gotoh, J., Kim, M.J., Lim, A.E.B., "Two Perspectives on Robust Empirical Optimization," 22nd International Symposium on Mathematical Programming (ISMP 2015), Wyndham Grand Pittsburgh Downtown, Pittsburgh, USA, 2015年7月

〔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〕

ホームページ等: <http://www.indsys.chuo-u.ac.jp/~jgoto/>

6. 研究組織

後藤 順哉 (GOTOH, Jun-ya), 中央大学・理工学部・教授, 研究者番号: 40334031

(1) 研究分担者

なし

(2) 研究協力者

武田朗子 TAKEDA, Akiko / 高野祐一 TAKANO, Yuichi / 東野克哉 TONO, Katsuya / URYASEV, Stan / LIM, Andrew E.B. / KIM, Michael J.

科研費による研究は, 研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、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。