

## 科学研究費助成事業 研究成果報告書

平成 27 年 5 月 4 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2011～2014

課題番号：23710174

研究課題名(和文)統計的学習理論のポートフォリオ最適化問題への適用

研究課題名(英文)Applying statistical learning theory to portfolio optimization problems

## 研究代表者

武田 朗子(Takeda, Akiko)

東京大学・情報理工学(系)研究科・准教授

研究者番号：80361799

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：数理最適化・金融工学・統計的学習の三つの分野の融合領域を開拓することを研究目的とした。本研究課題以前は、金融工学でよく知られたリスク尺度(VaRやCVaR)を統計的学習の研究分野に取り入れる”というスタンスで研究を進めてきたが、本課題では、逆に、統計的学習で研究が進められている汎化誤差理論を金融工学に持ち込み、予測精度の高いポートフォリオ最適化モデルを構築する”ことを目的に研究を進めた。その結果、高い汎化能力(新たなデータに対する予測能力)の理論的保証のついたポートフォリオ最適化モデルを構築し、そのモデルを効率的に解くための手法を考案することができた。

研究成果の概要(英文)：In this study, we aimed to explore different but related research areas (mathematical optimization, financial engineering and machine learning). Before this study, we have incorporated risk measures, which are investigated in financial engineering (mainly, VaR and CVaR), into machine learning problems. On the other hand, this study brought generalization theories, which are developed in machine learning, into financial engineering based on an expectation that we might obtain a new portfolio optimization model having high prediction accuracy. Indeed, we could construct such a new portfolio optimization model which has a theoretical guarantee on high prediction accuracy and devise an efficient algorithm to solve the optimization model. Our numerical experiments supported the theoretical guarantee of the proposed model.

研究分野：数理最適化

キーワード：ポートフォリオ最適化問題 金融リスク尺度 統計的学習 ロバスト最適化 サポートベクターマシン

1. 研究開始当初の背景

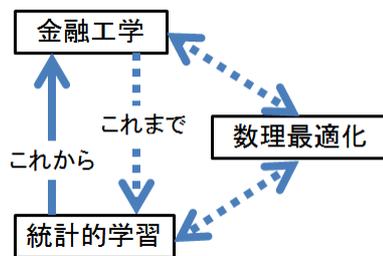
(1)金融資産への最適な投資配分を決定するための数理計画問題はポートフォリオ最適化問題と呼ばれている。マーコヴィッツ(H.Markowitz)は1952年に平均・分散モデルとよばれるポートフォリオ最適化問題を構築し、ポートフォリオ理論の基礎を築いた。それ以降、ポートフォリオ最適化は数理最適化や金融工学等の分野で研究が進められている。

(2)観測されたデータから統計的手法を用いて新たな知識を導出するための手法は統計的学習手法と呼ばれ、音声認識やスパム検知、画像認証などに用いられている。研究代表者はこれまでに、金融工学でよく知られたリスク尺度 (value-at-risk: VaR や conditional VaR: CVaR) を統計的学習の世界に持ち込み、統計的学習分野において代表的な分類モデルの1つである nu-SVM (Scholkopf 等)がノルム制約付き CVaR 最小化モデルとして見なせることを示した。

2. 研究の目的

(1)数理最適化・金融工学・統計的学習の三つの分野の融合領域を開拓することを研究目的とした。

(2)本研究課題以前は“金融工学でよく知られたリスク尺度 (VaR や CVaR) を統計的学習の研究分野に取り入れる”というスタンスで研究を進めてきたが、本課題では、逆に“統計的学習で研究が進められている汎化誤差理論を金融工学に持ち込み、予測精度の高いポートフォリオ最適化モデルを構築する”ことを目的に研究を進めた (次図を参照)。



(3)統計的学習の知見を取り入れて汎化能力 (新たなデータに対する予測能力) の高いポートフォリオ最適化モデルを構築すること、その最適化モデルを解くための最適化解法を考案することを目的として研究を行った。具体的には、

- ノイズの少ないデータセットに対するポートフォリオ最適化法
- データノイズにロバストな意思決定を行うためのポートフォリオ最適化法
- 大量のデータセットに対して予測に有用な特徴を選択して最適な意思決定をおこなうポートフォリオ最適化法
- より一般化したリスク尺度に基づくポートフォリオ最適化法

について研究を行った。以降、「3. 研究の方

法」, 「4. 研究成果」を、この順に報告する。

3. 研究の方法

(1)金融分野の理論・実務両面から脚光を浴びている CVaR と呼ばれるリスク尺度を最小にするポートフォリオ最適化問題 (金融資産への投資配分決定問題) を取りあげた。ノルム制約付き CVaR 最小化に基づくポートフォリオ最適化モデルについて、統計的学習分野からの知見を取り入れて汎化誤差上界値を導出して理論分析を行うとともに、数値実験による実証分析を行なった。

(2)ノルム制約付き VaR 最小化が汎化誤差の最もタイトな上界値を与えること、またノイズの多いデータセットに対してロバストな推定結果を与えることを理論的に示した。しかし、ノルム制約付き VaR 最小化が「ノルム制約の付いた整数計画問題」という求解困難な問題に帰着されるため、ノルム制約付き VaR 最小化モデルを解くためのアルゴリズムを考案する必要があった。Larsen 等[2002]によって提案された VaR 最小化解法に対してノルム制約も扱えるように修正を加え、ノルム制約付き VaR 最小化の近似解法を構築した。

(3)ポートフォリオの組換えには手数料がかかるため、組換え量が少なくすむようにポートフォリオモデルの改良を検討した。“ベンチマークとポートフォリオのパフォーマンスの乖離”を、L2 ノルムを用いて表現すれば、「ポートフォリオモデル固有の制約のついた回帰モデル」としてトラッキング・ポートフォリオモデルを見なすことができる。統計的学習分野では、正則化項 (ノルムを用いて表わされる項) を回帰モデルに加えることにより、モデルの汎化能力が向上することが知られている。その知見を取り入れて、既存のトラッキング・ポートフォリオモデルに正則化項を加え、汎化能力の高いモデルを構築した。正則化項のノルムを L1, L2, L0, L0 と L2 を組み合わせたものに変えることで、予測に有用な特徴 (金融資産) を選択して精度よい推定が出来るかを調べた。

(4) CVaR を含む広いクラスのリスク尺度であるコヒーレントリスクを用いた最適ポートフォリオモデルを検討した。このモデルは min-max 問題で記述され、確率分布が不確実なロバスト最適化モデルとして見なされる。一方、統計的学習分野で提案されている様々な判別モデルに対して、min-max 問題を用いた統一的な定式化で表現できることを示した。既存の判別手法の違いは、2段目の max 問題の制約領域 U (不確実性集合) の定義に現れることを明らかにし、今まで異なる仮定や想定のもとで提案されてきた判別手法の本質的な違い、そして、コヒーレントリスクとの関係を明確に示した。

4. 研究成果

(1)ノルム制約付き CVaR 最小化モデルの理

### 論的・実証的検証：

各資産の平均収益率に関する過去データを訓練データと見なし、ポートフォリオ最適化問題を「ある一定以上の収益を上げられるか否か」という分類問題と見立てることにより、ノルム制約付き CVaR 最小化モデルをポートフォリオ最適化のために用いることを提案した。分類モデルの汎化誤差の評価方法を真似ることにより、ポートフォリオ最適化モデルの汎化誤差に関して分類モデルと同様の結果が得られた。また、ノルム制約付き CVaR 最小化モデルの汎化誤差を評価することで、モデルの妥当性を理論的に示すことができた。また、日経 225 等の金融データを用いた数値実験を通して既存のポートフォリオ最適化モデルによる汎化誤差と提案手法による汎化誤差を比較することにより、提案手法が高い予測能力を持つことを確認した。

### (2)-①ノルム制約付き VaR 最小化ポートフォリオモデルの提案：

ノルム制約付き VaR 最小化が汎化誤差の最もタイトな上界値を与えることを理論的に示したものの、ノルム制約付き VaR 最小化が「ノルム制約の付いた整数計画問題」という求解困難な問題に帰着されるため、実データを用いてノルム制約付き VaR 最小化モデルを厳密に解くことは難しい。しかしながら、リスク尺度 VaR は観測データ(各資産収益率)のはずれ値の影響を受けにくいいため、本モデルは現実の金融データに対してあてはまりがよいと予想される。そこで、ノルム制約付き VaR 最小化ポートフォリオモデルを提案し、統計的な観点から理論的な特徴づけを試みた。ロバスト統計で使われる破局点を用いてモデルのロバスト性を評価する方法を確立した。その結果、高いロバスト性を達成するために選択すべきハイパーパラメータの範囲を導出することができた。

### (2)-②ノルム制約付き VaR 最小化ポートフォリオモデルの解法の考案：

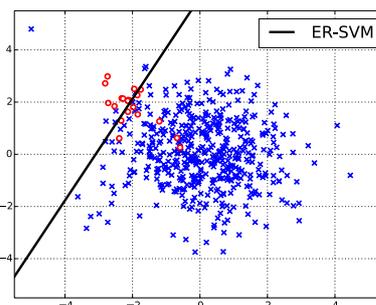
4-(2)-①においてノルム制約付き VaR 最小化ポートフォリオモデルを理論的に分析した後、実際にモデルの解を求めてモデルのあてはまりのよさを評価した。ノルム制約付き VaR 最小化問題を厳密に解こうとすると多くの計算時間が必要となり現実的なサイズのポートフォリオを扱うことは困難となるため、厳密解法ではなく近似解法を考案した。具体的には、Larsen 等[2002]によって提案された VaR 最小化解法をノルム制約も扱えるように改良し、ノルム制約付き VaR 最小化の近似解法を提案することを検討した。数値実験を通して関連手法との比較を行ない、データに外れ値が含まれていても、提案モデルがあまり外れ値の影響を受けずに済むこと、また通常の判別手法に劣らない計算時間でよい近似解を求められることを確認した。

### (2)-③機械学習・バイオインフォマティクス

### への VaR 最小化法の適用

4-(2)-②の研究成果：VaR 最小化法をポートフォリオ最適化問題に限らず、ノイズを含んだ実データに対する判別問題やバイオインフォマティクスへ適用し、ノイズの多い実データやクラス間のサンプル数に偏りが大きなデータセットに対してロバストな推定が行えることを確認した。

実際、クラス間でデータ数に偏りが大きなデータセットに対して、既存の判別手法は実行不可能となり、判別平面を求めることが出来なかった。一方で提案手法では、次図が示すように、妥当な判別平面を求めることができた。



### (3)スパースポートフォリオモデルの考案：

ポートフォリオモデルの正則化項のノルムを L1, L2, L0, L0 と L2 を組み合わせたものに変えることで、予測精度や解の疎性がどのように変化するか、数値実験を通して調べた。疎なポートフォリオ(最適解のベクトル要素にゼロが多く含まれている状態)は保有資産銘柄数が少ないことを意味する。投資家にとって、取引手数料は無視できないコストであり、疎なポートフォリオは取引手数料を小さく押さえるという意味では好ましいといえる。一方、あまりに疎なポートフォリオでは事後パフォーマンスが低下してしまう。そこで、L0 と L2 を組み合わせたノルムを正則化項として用いることにより、疎でありながら事後パフォーマンスのよりポートフォリオが達成できることを数値実験により確かめた。

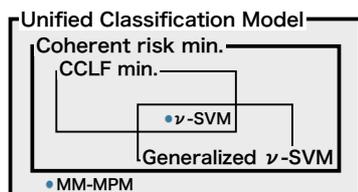
### (4)-①コヒーレントリスク最小化モデル：

min-max 問題で記述されるコヒーレントリスク最小化モデルは、確率分布が不確実なロバスト最適化モデルとして見なされる。確率分布集合を変えることで、コヒーレントリスクは、CVaR のみならずより広いクラスのリスク尺度を表現できる。一方で、min-max 問題の定式化は扱いにくく、この定式化のまま最適解を求めることはあまり試みられていなかった。しかし、様々なリスク尺度を試し判別モデルを構築するためには、min-max 問題のまま最適解を求める必要がある。そこで、first-order method と呼ばれる最適化手法を用いて、アルゴリズムを考案した。数値実験を通して、CVaR 最小化に特化した解法と比較した結果、コヒーレントリ

スク最小化による汎用解法（提案法）の方が、高速に CVaR 最小化問題を解くことができることを確認した。

#### (4)-②統一的な機械判別学習モデル：

統計的学習分野で提案されている様々な判別モデルに対して、数理最適化手法の一つであるロバスト最適化モデルを用いて、統一的に min-max 問題として表現できることを示した（提案モデルは下図の Unified Classification Model, その他の CCLF min, Generalized nu-SVM, nu-SVM, MM-MPM は既存の機械学習法）。



既存の判別手法の違いは、ロバスト最適化モデルで用いられる入力（不確実性集合）の定義に現れることを明らかにし、今まで異なる仮定や想定のもとで提案されてきた判別手法の本質的な違いを明確に示した。

さらに、min-max 問題で記述されるコヒーレントリスク最小化モデル（4-(4)-①を参照）との定式化を比較することにより、

- ✓ 統一的な機械判別学習モデル (Unified Classification Model) はコヒーレントリスク最小化モデルの一般化として見なせること、
- ✓ また、既存学習モデル (CCLF min, Generalized nu-SVM, nu-SVM) はコヒーレントリスク最小化モデルとして解釈できること、

を明らかにし、機械学習法と金融リスク尺度最小化モデルとの関係を明確に示した。

#### 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 20 件）

①Akiko Takeda and Takafumi Kanamori, Using Financial Risk Measures for Analyzing Generalization Performance of Machine Learning Models, Neural Networks, 査読あり, 57, 2014, pp. 29-38

DOI:10.1016/j.neunet.2014.05.006

②Jun-ya Gotoh, Akiko Takeda and Rei Yamamoto, Interactions between Financial Risk Measures and Machine Learning Methods, Computational Management Science, 査読あり, 11 (4), 2014, pp. 365-402

DOI:10.1007/s10287-013-0175-5

③Akiko Takeda, Hiroyuki Mitsugi, and Takafumi Kanamori, A Unified Classification Model Based on Robust Optimization, Neural Computation, 査読あり,

25 (3), 2013, pp. 759-804

DOI:10.1162/NECO\_a\_00412

④Akiko Takeda, Mahesan Niranjan, Jun-ya Gotoh and Yoshinobu Kawahara, Simultaneous Pursuit of Out-of-Sample Performance and Sparsity in Index Tracking Portfolios, Computational Management Science, 査読あり, 10 (1), 2013, pp. 21-49

DOI:10.1007/s10287-012-0158-y

⑤Jun-ya Gotoh and Akiko Takeda, Minimizing Loss Probability Bounds for Portfolio Selection, European Journal of Operational Research, 査読あり, 217 (2), 2012, pp. 371-380

DOI:10.1016/j.ejor.2011.09.012

〔学会発表〕（計 16 件）

①武田朗子, 数理最適化の視点から機械学習の分野へ～ロバスト最適化法と金融リスク尺度最小化の応用～, OR学会「最適化の理論と応用」研究部会 (SOTA), 2013年2月2日, 東京大学 (東京都文京区)。

②武田朗子, 機械学習方法を用いたポートフォリオ選択モデルの提案, JAFEE: ファイナンスの意思決定解析研究部会, 2012年4月7日, 芝浦工業大学 (東京都中央区)。

③武田朗子, ロバスト最適化法の機械学習への適用, 第23回 RAMP シンポジウム, 2011年10月25日, 関西大学千里山キャンパス百周年記念会館ホール, (大阪府吹田市)。

④武田朗子, 不確実性を考慮した最適化手法: ロバスト最適化法, システム制御情報チュートリアル講座 2011, 2011年6月23日, 常翔学園大阪センター (大阪府大阪市)。

〔図書〕（計 0 件）

該当なし

〔産業財産権〕

○出願状況（計 0 件）

該当なし

○取得状況（計 0 件）

該当なし

〔その他〕

ホームページ等

<http://www.opt.mist.i.u-tokyo.ac.jp/~takeda/index-j.html>

#### 6. 研究組織

(1) 研究代表者

武田朗子 (TAKEDA, Akiko)

東京大学・大学院情報理工学系研究科

・准教授

研究者番号: 80361799

(2) 研究分担者

該当なし

(3) 連携研究者  
該当なし