

2017年度修士論文

異なるダイバージェンスに基づく
分布的ロバスト最適化モデルの比較

A Comparison of Distributionally Robust
Optimization Models
Defined by Different Divergence Measures

中央大学大学院 理工学研究科
経営システム工学専攻 後藤研究室

16N7100026G

前田 彩

概要

ロバスト最適化は、最悪の確率分布を考慮した上で何らかの確率的な意思決定を行う最適化問題である。最悪の分布を定義する方法として、「その分布の台を名目分布の台（の部分集合）に限る場合」と「その分布の台を名目分布の台に限らない場合」の2つに分けることができる。最悪の分布を定義するために、従来使われていた ϕ -ダイバージェンスを用いる方法 [1] は前者であり、最近 [16, 20] によって提案されたワッサーズタイン距離を用いたロバスト最適化問題は、後者に対応する。一方で、ワッサーズタイン距離を名目分布の台に限定することで前者のロバスト最適化問題を定義することもできる。本研究では名目分布に経験分布を採用したうえで、ワッサーズタイン距離を用いた2つのロバスト最適化問題の比較を行う。具体的には、ロジスティック回帰、および、ポートフォリオ選択問題を取り上げ、それらのロバスト最適化問題を既述の2つの定式化に適用し、それらの事後パフォーマンスを比較した。標本の大きさに対して説明変数が少ないデータでは「その分布の台を名目分布の台に限らない場合」の方が、標本の大きさに対する説明変数が多いデータでは「その分布の台を名目分布の台に限らない場合」の方が良いパフォーマンスをしたことを示す。

目次

1	序論	3
2	分布的ロバスト最適化問題の定式化	5
2.1	期待報酬最大化問題	5
2.2	ワッサーズタイン距離 (Wasserstein Distance)	7
2.3	広義のワッサーズタイン距離を用いた場合の定式化	9
2.3.1	ロジスティック回帰	9
2.3.2	ポートフォリオ選択問題	10
2.4	狭義のワッサーズタイン距離を用いた場合の定式化	11
2.4.1	ロジスティック回帰	11
2.4.2	ポートフォリオ選択問題	12
3	平均分散フロンティア	13
4	数値実験	15
4.1	ロジスティック回帰	16
4.2	ポートフォリオ選択問題	25
5	結論	30
A	付録	33
A.1	広義のワッサーズタイン距離を用いた場合の定式化の変形	33
A.1.1	ロジスティック回帰の式変形	33
A.1.2	ポートフォリオ選択問題の式変形	36
A.2	狭義のワッサーズタイン距離を用いた場合の定式化の変形	37
A.2.1	ロジスティック回帰の変形	37
A.2.2	ポートフォリオ選択問題の変形	40
A.3	ポートフォリオ選択問題の平均分散フロンティア	41

1 序論

突然の自然災害、株価や為替の変動など過去のデータの予測からでは対応できないことが起こりうる。我々が完全なデータを収集することは現実的に不可能であり、不確実性を含む事象に対して悪い状況を考慮した意思決定することで、大きなリスクを避けることができる。

将来の予測をするなどの不確実なデータが含まれているとき、既知のデータから想定した範囲内における最悪な状況に対して最適解を導くロバスト最適化問題 (Distributionally Robust Optimization, DRO) の研究が注目を集めている [1, 3, 7, 16, 20]。最悪の状況を想定した意思決定をしているので、既知のデータと多少異なっていたとしても、不確実な事象の変化に対して強い手法である。金融工学のポートフォリオ選択では、株価や為替の大きな変動を想定したうえで投資比率を決めることができたり、製品の生産においても、使用環境の変化や部品の劣化などを考慮した外的作用に強い設計をすることができる [10, 14]。また、文字認識 [2] や画像処理 [17]、医療診断など様々な分野で使われはじめている。

一方で、DRO はいわゆる max-min 型の問題であり、よく使われる線形計画問題や凸計画問題などと比べると扱いにくい場合がある [1]。DRO を扱った研究の多くは、不確実なデータから起こりうる範囲をどう設定し、どう変換したら解きやすい問題になるのかが研究の主眼になっている。または、簡単な問題に変形できなくとも、最適解に近い解を求めるにはどうすればよいかという研究が進められている。解きやすいロバスト最適化問題を定式化するために、不確実なデータを含む 2 次錐計画問題や半正定値問題、線形計画問題などを扱いやすいロバスト最適化問題へ帰着させる研究などが進んでいる。また、不確実なデータに対して確率分布を仮説し、そこから最適化問題を解くという研究もある。不確実なデータから起こりうる範囲を定義する方法は、一般的に 2 つ存在する。1 つめは「最悪の分布の台を名目分布の台に限った部分集合」の場合であり、有名な方法として ϕ -ダイバージェンスを用いる方法 [7, 20, 1, 21] が知られている。2 つめは「最悪な分布の台を名目分布の台に限らない部分集合」の場合であり、ワッサーズタイン距離 [6, 8, 16] などを用いて定義できる。ここで言う名目分布とは最悪な分布を求めるときに基準とする分布のことであり、一般的に経験分布が使われることが多い。「最悪な分布の台を名目分布の台に限らない」ワッサーズタイン距離は、名目分布の台を限定することで「最悪の分布の台を名目分布の台に限定した」ロバスト最適化問題を定義することが可能である。

そこで本研究では、ワッサーズタイン距離を用いた分布的ロバスト最適化を用いて、この 2 つの DRO のパフォーマンスの比較を行う。ロバスト最適化を解きやすい問題に変形する研究やそれを利用して様々な分野で応用されてはいるが、不確実なデータから起こりうる範囲を定義する 2 つの方法に関して比較をした研究はまだ少

ない。ワッサーズタイン距離でこの2つを定義できれば、異なるダイバージェンスにおける比較をすることができる。それぞれの特徴を知ることが出来れば、時々にあわせてより良い解を得ることができるようになるのではないかと考えた。

また、ワッサーズタイン距離を用いた場合に限らず多くのロバスト最適化において、不確実なデータから起こりうる範囲は何かしらのパラメータに依存している。パラメータを大きくしたり小さくすることで、範囲の不確実なデータから起こりうる範囲を調節をすることができる。[11]にはこの範囲を調整するパラメータを動かすことで、パラメータ毎の最適値のパフォーマンスの振る舞いについて書かれている。[11]では、わずかなパラメータを与えることで、最適値を保障しつつ、最適値の変動（分散）を大幅に抑えられることが指摘されている。パラメータによるロバスト最適化問題の最適値と変動（分散）の関係をみることができ、それを平均分散フロンティアと呼んでいる。最適値をわずかに下げってしまうものの、分散をどれだけ小さくできたかという評価をすることができ、本研究の2つの最適化問題の比較に有効である。

本研究の構成は以下の通りである。2節でワッサーズタイン距離の定義を述べ、2つのワッサーズタイン距離を用いたロバスト最適化問題の定式化を示す。また、数値実験で取り扱うロジスティック回帰とポートフォリオ選択問題について具体的な定式化を示す。3節では、2つのDROを評価に用いる平均分散フロンティアを示し、4節で数値実験を行なった2つのDROについての比較をする。

2 分布的ロバスト最適化問題の定式化

この節では分布的ロバスト最適化問題についてまとめる。まず、2.1 節では期待報酬最大化問題について記し、一般的な分布的ロバスト最適化問題の定式化を示す。2.2 節でワッサーズタイン距離の定義を述べ、2.3 節で最悪な分布の台を経験分布に限らないワッサーズタイン距離を用いた分布的ロバスト最適化問題の定式化をし、2.4 節で最悪な分布の台を経験分布に限るワッサーズタイン距離を用いた最適化問題の定式化を示す。また、2.3 節と 2.4 節では具体的にロジスティック回帰とポートフォリオ選択問題の定式化を与える。

2.1 期待報酬最大化問題

本論文では以下のような最適化問題を考える。

$$\max_{\mathbf{x}} \mathbb{E}_{\mathbb{P}}[f(\mathbf{x}, \mathbf{Y})]. \quad (2.1)$$

ここで $f(\mathbf{x}, \mathbf{Y})$ は \mathbf{x} と \mathbf{Y} によって決まる報酬関数であり、 $\mathbf{x} \in \mathbb{R}^b$ は決定変数ベクトル、 \mathbf{Y} は \mathbf{x} とは独立な m 次元の確率変数ベクトルを表す。 \mathbb{P} は \mathbf{Y} が従う真の分布を表し、 $\mathbb{E}_{\mathbb{P}}$ は \mathbb{P} の下での期待値を表す。ポートフォリオ選択問題の効用関数の最大化問題など問題 (2.1) は多くの応用例を含む普遍的な構造を有している。現実における多くの場合、分布 \mathbb{P} は未知であるため、この問題の目的関数は未知となり、その意味で問題 (2.1) は実行ができない。

しかしながら、 \mathbb{P} が未知である代わりに \mathbb{P} に基づいて生成された \mathbf{Y} の標本 $\{\mathbf{Y}_1, \dots, \mathbf{Y}_n\}$ が与えて考えられることが多い。経験データ $\mathbf{Y}_1, \dots, \mathbf{Y}_n \in \mathbb{R}^m$ が与えられているとき、(2.1) の代わりに以下の問題を扱うことが多い。

$$\max_{\mathbf{x}} \left\{ \mathbb{E}_{\hat{\mathbb{P}}_n}[f(\mathbf{x}, \mathbf{Y})] = \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}, \mathbf{Y}_i) \right\}. \quad (2.2)$$

ただし、 $\hat{\mathbb{P}}_n$ は \mathbf{Y} の経験分布である。本論文ではこの問題を経験的最適化問題と呼ぶ。経験的最適化問題 (2.2) の目的関数およびその最適解は標本数 n が大きくなるにつれ問題 (2.1) の最適値および最適解に近づくことが知られているが、多くの場合、事後データのパフォーマンスを劣化させてしまう [11]。

これに対し分布的ロバスト最適化 (Distributionally Robust Optimization, DRO) は、名目分布から生じる範囲をあらかじめ決め、その中にある最悪な分布から最適な問題を解く方法である。ここでいう名目分布は不確実な範囲を決める基準となる分布であり、一般的に用いられる経験分布が使われている。本研究では、名目分布に経験分布を採用したうえで、以下のような形で定式化されるロバスト最適化問題について考える：

$$\max_{\mathbf{x}} \min_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \left\{ \mathbb{E}_{\mathbb{Q}}[f(\mathbf{x}, \mathbf{Y})] \right\} \quad (2.3)$$

このとき、 $\hat{\mathbb{P}}_n$ が \mathbf{Y} の経験分布、 \mathbb{Q} は代替分布である。 $\mathcal{M}(\hat{\mathbb{P}}_n)$ は経験分布から想定する代替分布の集合であり、不確実性集合と呼ぶ。経験分布 $\hat{\mathbb{P}}_n$ を中心とした不確実性集合 $\mathcal{M}(\hat{\mathbb{P}}_n)$ から目的関数を最大にする最悪な分布を導き、その最悪な分布をもとに最適化する問題である。最悪な場合を想定し決定しているので、経験的最適化問題 (2.2) と比べて不確実な変化に強い手法であると言える。

しかしながら、ロバスト最適化問題は一般的に扱いにくいとされている [1]。ロバスト最適化を扱った研究の多くは、不確実な分布から想定する範囲をどう設定し、どう変形したら解きやすい問題になるのかが研究の主な目的になっている。

不確実性集合 $\mathcal{M}(\hat{\mathbb{P}}_n)$ を定義する方法は一般的に 2 つ存在することが知られていて、「代替分布の台が経験分布の台に限る場合」と「代替分布の台が経験分布の台に限らない場合」である。名目分布とは不確実性集合の基準となる分布のことであり、一般的に経験分布を用いる。代替分布は不確実性集合に含まれる分布のことであり、その中から目的関数について最悪な分布を定義する。前者は、図 2.1(a) のような経験分布の台があるときに図 2.1(b) のように経験分布の台のみでしか定義されない。不確実性集合は経験分布についての確率変数の高さをどれだけ変えられるかによって集合の範囲が決まり、その中から最悪な分布を決定する。代替分布は名目分布である経験分布からの距離（ないしはそれに準ずる尺度）によって定義されるが、既存研究では ϕ ダイバージェンスを用いた方法が有名である [7, 20, 1, 21]。後者は図 2.1(c) のように代替分布が経験分布の台に限らず、決められた範囲の内であれば分布も高さも変えることができる。既存研究としては [6, 8, 16] によってワッサーズタイン距離を用いて不確実性集合を与えた分布的ロバスト最適化モデルが提案された。この 2 つを比べたとき、より自由度の高い後の方がより最悪な分布を想定できることが分かる。しかしながら、より最悪な分布を求めれば良いというわけではないので必ずしも後者が良いとは言えない。

そこで、本研究では「代替分布の台を経験分布の台に限る場合」と「代替分布の台を経験分布の台に限らない場合」について比較をする。ワッサーズタイン距離は、経験分布の台に限ることで、いわゆる Earth Mover's Distance (EMD) に帰着できる

が、これを用いた不確実性集合を採用することで、経験分布の台に限った代替分布に基づく分布的ロバスト最適化問題を定義することが可能である。

本研究では、代替分布が経験分布の台に限らないワッサーズタイン距離を用いた分布的ロバスト最適化問題と経験分布の台に限らないワッサーズタイン距離（EMD）を用いた分布的ロバスト最適化問題を比較する。次節でワッサーズタイン距離と EMD の定義についてまとめる。

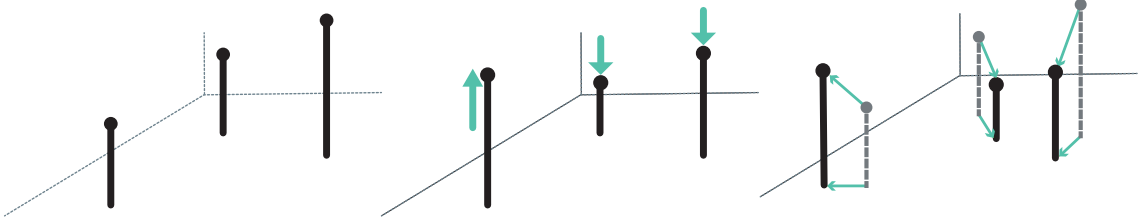


図 2.1: (a) 経験分布の台 (b) 代替分布の台を経験分布に限定する場合
(c) 代替分布の台を経験分布の台に限らない場合

2.2 ワッサーズタイン距離（Wasserstein Distance）

まず、経験分布の台に限らずに代替分布を定義する一般的なワッサーズタイン距離について説明し、経験分布の台に限った代替分布を定義するワッサーズタイン距離である Earth Mover's Distance（EMD）について定義する。

定義 1（ワッサーズタイン距離） 確率分布 \mathbb{P} に対する確率分布 \mathbb{Q} のワッサーズタイン距離 $W(\mathbb{Q}|\mathbb{P})$ を以下で定義する。

$$W(\mathbb{Q}|\mathbb{P}) := \left| \begin{array}{l} \inf_{\Pi \in M(\Xi^2)} \int_{\Xi^2} d(Y, Y') \Pi(dY, dY') \\ \text{s.t.} \quad \Pi(dY, \Xi) = \mathbb{Q}(dY) \\ \quad \quad \Pi(\Xi, dY') = \mathbb{P}(dY') \end{array} \right. \quad (2.4)$$

ここで、 Ξ は分布の台を含む空間とし、 $\Xi \times \Xi$ 上の確率分布の集合を $M(\Xi^2)$ とする。 Y は分布 \mathbb{Q} についての確率変数、 Y' は分布 \mathbb{P} についての確率変数であり、 Π は分布 \mathbb{P} と分布 \mathbb{Q} の確率変数の同時確率分布である。 任意のノルム $\|\cdot\|$ を用いて $d(Y, Y') = \|Y - Y'\|$ と与えるとする。

$d(Y, Y')$ は確率変数 Y' から確率変数 Y までの移動コストのことで、ワッサーズタイン距離 $W(\mathbb{Q}|\mathbb{P})$ は分布 \mathbb{P} から分布 \mathbb{Q} まで移動するときの最小コストを示しており、また、分布 \mathbb{Q} と分布 \mathbb{P} の 2 つの分布間の距離を表していると解釈できる。

これをふまえて経験分布 $\hat{\mathbb{P}}_n$ に対する代替分布 \mathbb{Q} のワッサーズタイン距離 $W(\mathbb{Q}|\hat{\mathbb{P}}_n)$ を用いた不確実性集合 $\mathcal{M}(\hat{\mathbb{P}}_n)$ を

$$\mathcal{M}(\hat{\mathbb{P}}_n) := \{\mathbb{Q} : W(\mathbb{Q}|\hat{\mathbb{P}}_n) \leq \varepsilon\} \quad (2.5)$$

と表す. ここで, $\varepsilon \geq 0$ は $\hat{\mathbb{P}}_n$ を中心としたワッサーズタイン球の半径である. 半径 ε が分布的ロバスト最適化問題におけるパラメータとなり, このパラメータを大きくしたり小さくすることで経験分布 $\hat{\mathbb{P}}_n$ を中心とした代替分布 \mathbb{Q} における不確実性集合 (2.5) の大きさを調節することができる.

次に, Earth Mover's Distance (EMD) について定義する. ワッサーズタイン距離と同様に, 確率分布 \mathbb{P} から確率分布 \mathbb{Q} まで移動するときの最小コスト問題であり, 確率分布 \mathbb{P} と確率分布 \mathbb{Q} の2つの分布間の距離を示す距離尺度である.

定義 2 (EMD) 確率分布 \mathbb{P} に対する確率分布 \mathbb{Q} の EMD である $W'(\mathbb{Q}|\mathbb{P})$ を以下で定義する.

$$W'(\mathbb{Q}|\mathbb{P}) := \left| \begin{array}{l} \min_{\pi_{ij}} \sum_{i=1}^n \sum_{j=1}^n d_{ij} \pi_{ij} \\ \text{s.t.} \quad \sum_{i=1}^n \pi_{ij} = p_j \quad (j = 1, \dots, n) \\ \sum_{j=1}^n \pi_{ij} = q_i \quad (i = 1, \dots, n) \\ \pi_{ij} \geq 0 \quad (i, j = 1, \dots, n) \end{array} \right. \quad (2.6)$$

ただし, 確率分布 $\mathbb{P} = (p_1, \dots, p_n)$, 確率分布 $\mathbb{Q} = (q_1, \dots, q_n)$ である. π_{ij} は p_j と q_i の同時確率分布, d_{ij} は任意のノルム $\|\cdot\|$ を用いて $d_{ij} = \|Y_i - Y_j\|$ と与えるとする.

すなわち, EMD は (1) 式において分布 \mathbb{Q} が分布 \mathbb{P} に対して絶対連続な分布であり, $p_i = 0$ ならば $q_i = 0$ であり, $q_i > 0$ ならば $p_i > 0$ でないといけない. これより, EMD(2.6) により定義したワッサーズタイン球 (2.5) を用いた分布的ロバスト最適化問題の「代替分布の台が経験分布の台に限る部分集合」を定義することができる. また, 確率分布 \mathbb{P} が経験分布 $\hat{\mathbb{P}}_n$, 確率分布 \mathbb{Q} が代替分布 \mathbb{Q} に対応しており, 以降からは $W'(\mathbb{Q}|\mathbb{P})$ を経験分布 $\hat{\mathbb{P}}_n$ に対する代替分布 \mathbb{Q} の $W'(\mathbb{Q}|\hat{\mathbb{P}}_n)$ と記す. (2.6) を用いた不確実性集合 $\mathcal{M}'(\hat{\mathbb{P}}_n)$ を示す.

$$\mathcal{M}'(\hat{\mathbb{P}}_n) := \{\mathbb{Q} : W'(\mathbb{Q}|\hat{\mathbb{P}}_n) \leq \varepsilon'\} \quad (2.7)$$

このとき, $\varepsilon' \geq 0$ は $\hat{\mathbb{P}}_n$ を中心としたワッサーズタイン球の半径である. 半径 ε' が分布的ロバスト最適化問題におけるパラメータとなり, このパラメータを動かすことで経験分布 $\hat{\mathbb{P}}_n$ を中心とした代替分布 \mathbb{Q} についての不確実性集合 (2.7) の大きさを調節することができる.

このように、「代替分布の台を経験分布の台に限らない」一般的なワッサーズタイン距離と「代替分布の台を経験分布の台に限る」ワッサーズタイン距離 (EMD) を区別することができた。以降からは、代替分布の台を経験分布の台に限らないワッサーズタイン距離を用いた分布的ロバスト最適化問題を「広義のワッサーズタイン距離を用いた分布的ロバスト最適化問題」とし、代替分布の台を経験分布の台に限る EMD を用いた分布的ロバスト最適化を「狭義のワッサーズタイン距離を用いた分布的ロバスト最適化問題」と呼ぶ。

次節から、2つのワッサーズタイン距離を用いた定式化についてまとめる。定式化を示した後に、具体的にロジスティック回帰と効用関数を用いたポートフォリオ選択問題を取り上げて定式化を表し、その定式化を変形した簡単な凸最適化問題を示す。

2.3 広義のワッサーズタイン距離を用いた場合の定式化

定義 1 と (2.3) から、代替分布の台が経験分布の台に限らない場合のワッサーズタイン距離を用いた分布的ロバスト最適化問題の定式化を以下のように表せる：

$$\sup_{\mathbf{x}} \inf_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [f(\mathbf{x}, \mathbf{Y})] | W(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon \}. \quad (2.8)$$

また、この分布的ロバスト最適化問題の解を $x(\varepsilon)$ とする。前述したように、ロバスト最適化問題を簡単に解くことはできない。ワッサーズタイン距離を用いたこの定式化は双対性を利用することで簡単な凸計画問題に帰着することができる [8, ?]。本研究では、具体的にロジスティック回帰とポートフォリオ問題を用いて、広義のワッサーズタイン距離を用いた場合の定式化とそれを変形させた等価な定式化を示す。

2.3.1 ロジスティック回帰

ロジスティック回帰は、データ分析の手法として最もよく用いられる回帰分析の 1 つである。回帰分析とは事前データから説明変数と被説明変数の関係を推定する手法であり、事後データの予測などに使われる。ロジスティック回帰は被説明変数が 2 値の場合に使われる手法であり、説明変数から被説明変数がどちらであるかを予測する。ロジスティック回帰を以下に示す。

$$\text{Prob}(y_i | \mathbf{z}_i) = \frac{1}{1 + \exp(-y_i(\mathbf{z}_i^T \mathbf{x} + x_0))}$$

ただし、 $y_i \in \{-1, 1\}$ が被説明変数、 $\mathbf{z}_i \in \mathbb{R}^m (i = 1, \dots, n)$ が説明変数であり、 x_0 と $\mathbf{x} \in \mathbb{R}^m$ が未知の回帰係数である。未知の回帰係数を求めるには、ロジスティックの対数尤度関数を最大化すればよく、対数尤度関数を次に示す。

$$l_{\mathbf{x}}(\mathbf{z}_i, y_i) = \log(1 + \exp(-y_i(\mathbf{z}_i^T \mathbf{x} + x_0))) \quad (2.9)$$

目的関数を $f(\mathbf{x}, (\mathbf{z}, \mathbf{y})) = l_{\mathbf{x}}(\mathbf{z}, \mathbf{y})$ として, 問題 (2.8) を表す:

$$\sup_{\mathbf{x}} \inf_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [l_{\mathbf{x}}(\mathbf{z}, \mathbf{y})] | W(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon \}. \quad (2.10)$$

双対性を利用して, (2.10) と等価な扱いやすい凸計画問題に変形することができる [5]. 詳細は付録に示す.

$$\left| \begin{array}{ll} \max_{\mathbf{x}, \lambda, s_i} & \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n s_i \\ \text{s.t.} & -l_{\mathbf{x}}(\hat{\mathbf{z}}_i, \hat{\mathbf{y}}_i) \geq s_i \\ & -l_{\mathbf{x}}(\hat{\mathbf{z}}_i, -\hat{\mathbf{y}}_i) - \lambda \kappa \geq s_i \\ & \|\mathbf{x}\| \leq \lambda \end{array} \right. \quad (2.11)$$

このとき, $\lambda, s_i (i = 1, \dots, n)$ は双対係数であり, 正の重み κ は 1 とした. $\|\cdot\|$ は任意のノルムである.

2.3.2 ポートフォリオ選択問題

ポートフォリオ選択問題とは投資対象が銘柄 m 個に対して投資の比率をどうするか決定する問題であり, 金融の分野でよく使われている. リスクを最小化する場合や, リターンを最大化する場合など様々な問題があるが, 今回は効用関数

$$U_{\mathbf{x}}(\mathbf{R}_i) = -\exp(-\alpha \mathbf{R}_i^T \mathbf{x})$$

を利用する. このとき, 投資比率 $\mathbf{x} \in \mathbb{R}^m$, n 期間の収益率 $\mathbf{R}_i \in \mathbb{R}^m (i = 1, \dots, n)$ である. α はリスク回避パラメータであり, $\alpha = 1$ とする. 収益 $\mathbf{R}_i^T \mathbf{x}$ について単調増加する関数であり, 収益があがるほど効用 (嬉しさ) も大きくなる最大化問題である. 目的関数を $f(\mathbf{x}, \mathbf{R}) = U_{\mathbf{x}}(\mathbf{R})$ として, 効用関数における広義のワッサーズタイン距離を用いた場合の定式化 (2.8) は以下のように表すことができる:

$$\sup_{\mathbf{x}} \inf_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [U_{\mathbf{x}}(\mathbf{R})] | W(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon, 1^T \mathbf{x} = 1, \mathbf{x} \geq 0 \}. \quad (2.12)$$

ポートフォリオ選択問題の制約である $1^T \mathbf{x} = 1, \mathbf{x} \geq 0$ も追加する. 双対性を利用して, (2.12) と等価な扱いやすい凸計画問題に帰着させる.

$$\left| \begin{array}{ll} \max_{\mathbf{x}, \lambda} & \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n U_{\mathbf{x}}(\mathbf{R}_i) \\ \text{s.t.} & \|\mathbf{x}\| \leq \lambda \\ & 1^T \mathbf{x} = 1 \\ & \mathbf{x} \geq 0 \end{array} \right.$$

このとき, $\lambda, s_i (i = 1, \dots, n)$ は双対係数であり, $\|\cdot\|$ は任意のノルムである.

2.4 狭義のワッサーズタイン距離を用いた場合の定式化

定義2のEMDを用いて, 代替分布の台が経験分布の台に限る場合の分布的ロバスト最適化問題の定式化を次に示す [11].

$$\max_{\mathbf{x}} \min_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [f(\mathbf{x}, \mathbf{Y})] | W'(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon' \} \quad (2.13)$$

これも双対性を利用して簡単な凸計画問題に帰着することができる. 広義のワッサーズタイン距離を用いた定式化と同様に, ロジスティック回帰とポートフォリオ問題を用いて, 広義のワッサーズタイン距離を用いた場合の定式化とそれを変形させた等価な定式化を示す.

2.4.1 ロジスティック回帰

前節のロジスティック損失関数 (2.9) を (2.13) で表す:

$$\max_{\mathbf{x}} \min_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [l_{\mathbf{x}}(\mathbf{z}, \mathbf{y})] | W'(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon' \} \quad (2.14)$$

双対性を利用して (2.14) と等価な扱いやすい凸計画問題に帰着させる [11]. 詳細は付録に示す.

$$\left| \begin{array}{ll} \max_{\mathbf{x}, \mu_j, \theta} & \varepsilon' \theta + \sum_{j=1}^n p_j \mu_j \\ \text{s.t.} & \theta \sum_{j=1}^n d_{ij} + \mu_j \leq l_{\mathbf{x}}(\mathbf{z}_i, y_i) \\ & \theta \geq 0 \end{array} \right. \quad (2.15)$$

このとき $\theta, \mu_j (j = 1, \dots, n)$ は双対係数であり, $\|\cdot\|$ は任意のノルムである.

2.4.2 ポートフォリオ選択問題

(2.12) と同様に, 負の効用関数 (2.12) を (2.13) で表す:

$$\max_{\mathbf{x}} \min_{\mathbb{Q} \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n}[U_{\mathbf{x}}(\mathbf{R})] | W'(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon', 1^T \mathbf{x} = 1, \mathbf{x} \geq 0 \}. \quad (2.16)$$

双対性を利用して (2.16) と等価な扱いやすい凸計画問題に変形することができる [11].

$$\left| \begin{array}{ll} \max_{\mathbf{x}, \boldsymbol{\mu}, \theta} & \varepsilon' \theta + \frac{1}{n} \sum_{j=1}^n \mu_j \\ \text{s.t.} & \theta \sum_{j=1}^n d_{ij} + \mu_j \leq U_i(\mathbf{x}) \\ & \theta \geq 0 \\ & 1^T \mathbf{x} = 1 \\ & \mathbf{x} \geq 0 \end{array} \right. \quad (2.17)$$

このとき $\theta, \boldsymbol{\mu} \in \mathbb{R}^n (j = 1, \dots, n)$ は双対係数であり, $\|\cdot\|$ は任意のノルムである.

3 平均分散フロンティア

2つのワッサースタイン距離を用いた分布的ロバスト最適化問題の比較を行うため、平均分散フロンティアを用いる。[11]によって、代替分布の範囲を調整するパラメータをわずかに与えるだけで、最適値を保障しつつ、最適値の変動（分散）を大幅に抑えられることが指摘されている。パラメータによるロバスト最適化問題の最適値と変動（分散）の関係を見ることができ、[11]では、それを平均分散フロンティアと呼んでいる。最適値をわずかに下げてしまったが、分散をどれだけ小さくできたかというように評価をすることができ、本研究の2つの最適化問題の比較に有効である。平均分散フロンティアは、異なるパラメータを与えたときの分布的ロバスト最適化問題の最適解 $x^*(\varepsilon)$ について平均と分散をプロットしたものである。平均と分散を以下のように示す。

$$\mu(\varepsilon) := \mathbb{E}_{\hat{\mathbb{P}}_n}[f(x^*(\varepsilon), \mathbf{Y})] = \frac{1}{n} \sum_{i=1}^n f(x^*(\varepsilon), \mathbf{Y}_i) \quad (3.1)$$

$$\sigma^2(\varepsilon) := \mathbb{V}_{\hat{\mathbb{P}}_n}[f(x^*(\varepsilon), \mathbf{Y})] = \frac{1}{n} \sum_{i=1}^n \{f(x^*(\varepsilon), \mathbf{Y}_i) - \mu(\varepsilon)\}^2 \quad (3.2)$$

図3.1にロジスティック回帰の事前データにおける平均分散フロンティアの例を示す。右上にあるパラメータ $\varepsilon = 0$ のときは一般的な経験的最適化問題である。 $\varepsilon = 0.01$ でわかるように、わずかに平均が減っているものの分散も小さくなっていることが分かる。分散を下げるのが分布的ロバスト最適化において大きな利点であり、最適値（平均）を保ちつつ、分散を小さくすることが分布的ロバスト最適化において良いパフォーマンスであるといえる。これから比較する「最悪な分布の台が経験分布の台に限った場合」と「最悪な分布の台が経験分布の台に限らない」最適化問題の比較するにあたって、パラメータに対する平均と分散に注目し、数値実験を進める。

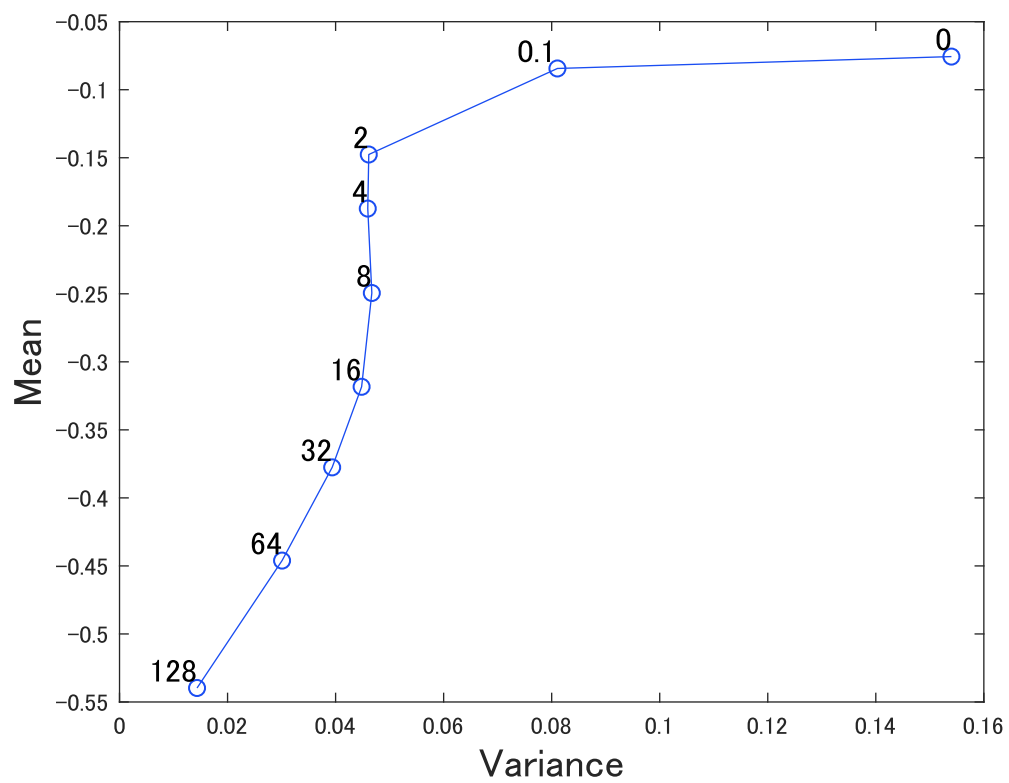


図 3.1: ロジスティック回帰における平均分散フロンティアの例

4 数値実験

実データを用いて, 2つのワッサーズタイン距離の分布的ロバスト最適化問題の数値実験を行い, 結果と考察を示す. 以降では, 広義なワッサーズタイン距離の分布的ロバスト最適化問題を B-WDR (Broad Sense Wasserstein Distributionally Robust Optimization) とし, 狭義なワッサーズタイン距離の分布的ロバスト最適化問題を N-WRO (Narrow Sense Wasserstein Distributionally Robust Optimization) と呼ぶ. 実装には CPU:Xeon(R) の PC で R x64 3.1.2, Rnuopt 19.0.0 を使用した.

まず, 検証に利用したブートストラップについてまとめ, ブートストラップを用いた平均分散フロンティアのアルゴリズムを示す. 4.1 節では実データの概要を示し, 実験結果と考察を示す. 4.2 節では, 実データの概要を示し, 実験結果と考察を示す.

ブートストラップ ブートストラップとは得られた標本から複数回の無作為抽出を行い, その無作為抽出した標本から母集団の性質を推定する統計的推定法である. 標本数 n , 説明変数 m 個のデータセット \mathbf{D} から重複を許して無作為抽出した n 個のブートストラップ標本を $J(j = 1, \dots, J)$ 個作成する. j 個目の標本を $\mathbf{D}^{(j)}$ とし, $\hat{\mathbb{P}}_n^{(j)}$ を $\mathbf{D}^{(j)}$ の経験分布とする. パラメータ $\varepsilon_l (l = 1, \dots, L)$ に対して分布的ロバスト最適化問題の解 $x^{(j)}(\varepsilon_l)$ を導き, その解から平均 $\mu^{(j)}(\varepsilon_l)$ と分散 $\sigma^{(j)}(\varepsilon_l)$ を求める. ε_l に対して, J 個の平均 $\mu^{(j)}(\varepsilon_l)$ と $\sigma^{(j)}(\varepsilon_l)$ 分散が出力されたので, ε_l ごとに平均と分散の平均をとり, ブートストラップによる解の平均と分散の関係を求めることができる. ブートストラップの平均分散フロンティアを Algorithm1 に示す. 本実験では, $J = 30$ で行った.

Algorithm 1 ブートストラップ法による平均分散フロンティア

Require: データセット $D = \{Y_1, \dots, Y_n\}$, パラメータ $\varepsilon \in \mathbb{R}^L (\geq 0)$

for $j = 1, \dots, J$ **do**

$D^{(j)}$: ブートストラップ標本を生成

for $l = 1, \dots, L$ **do**

$$x^{(1)}(\varepsilon_l) := \arg \min_{\mathbf{x}} \max_{\mathbb{Q}} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [f(\mathbf{x}, D^{(j)}) | W(\mathbb{Q} | \hat{\mathbb{P}}_n^{(j)}) \leq \varepsilon_l] \}$$

$$\mu_l^{(j)}(\varepsilon) = \mathbb{E}_{\hat{\mathbb{P}}_n} [f(x^{*(j)}(\varepsilon), D^{(j)})]$$

$$\sigma_l^{(j)}(\varepsilon) = \mathbb{V}_{\hat{\mathbb{P}}_n} [f(x^{*(j)}(\varepsilon), D^{(j)})]$$

end for

$$M_l(\varepsilon) = \frac{1}{J} \sum_{j=1}^J \mu_l^{(j)}(\varepsilon)$$

$$\Sigma_l(\varepsilon) = \frac{1}{J} \sum_{j=1}^J \sigma_l^{2(j)}(\varepsilon)$$

end for

return $(\Sigma_l(\varepsilon), M_l(\varepsilon)) : l = 1, \dots, L$

数値実験に利用するロジスティック回帰とポートフォリオ選択問題をまとめ、利用した実データを示した後に、実験結果と比較をまとめる。

4.1 ロジスティック回帰

ロジスティック回帰で使用する実データ WCBP, ILPD[15] の概要を表 4.1 に示す。標本数の半分を事前データとし、残りの半分を事後データとして利用した。

表 4.1: ロジスティックの数値実験に用いた実データ

データ名	標本数	説明変数	事前データの数	事後データの数
Breast Cancer Wisconsin	569	3	284	285
Indian Liver Patient Dataset	583	10	291	292

Breast Cancer Wisconsin は略称として WDBC, Indian Liver Patient Dataset の略称を ILPD と呼ぶ。また, WDBC は, 30 個ある説明変数から 3 つに絞ったデータを利用する。

事前データの比較 WDBC, ILPD の事前データの平均分散フロンティアを比較した。図 4.1, 図 4.2 は WDBC, ILPD の事前データの B-WRO と N-WRO の平均分散フロンティアである。B-WRO の $\varepsilon = 0$, N-WRO の $\varepsilon' = 0$ のときは、一般的な経験的最適化であり平均も分散も同じ結果が得られる。どちらもパラメータを与えることで、大幅に分散が減っていることがわかる。また、N-WRO の方が平均を保ちつつ分散を減らすことができている、良いパフォーマンスが出来ていると言える。

WDBC においてパラメータ $\varepsilon = 0.005, \varepsilon' = 2$ まで平均はほとんど変わらず分散を大きく下げているが、さらに大きいパラメータを与えると、B-WRO と N-WRO どちらも似たような振る舞いをしながら平均が小さくなっている。

ILPD は、N-WRO の方が緩やかに平均が落ちている。標本数に対する説明変数が多い ILPD にとって、最悪な分布を経験分布の台に限定していない B-WRO は不確実性集合が大きくなりより最悪な分布を定義したことが分かる。一方で、N-WRO は経験分布の台に限っているので、B-WRO より不確実性集合が小さく設定され、分散に対する平均のパフォーマンスが良くなったのではないかと考えられる。

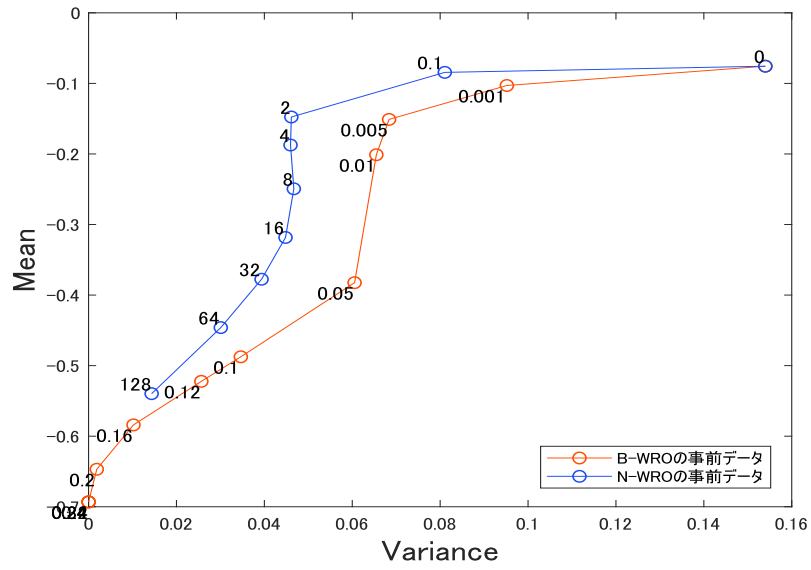


図 4.1: WDBC の事前データの平均分散フロンティアの比較

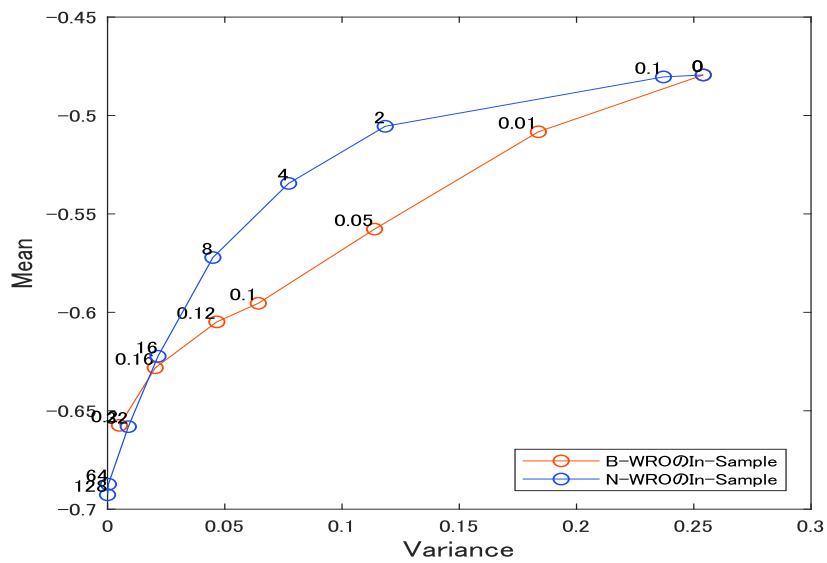


図 4.2: ILPD の事前データの平均分散フロンティアの比較

平均分散フロンティアの比較 WDBC, ILPD に対する B-WRO と N-WRO の事前とブートストラップ法, 事後から平均分散フロンティアを示し, 考察を述べる.

WDBC 図 4.3, 図 4.4 は WDBC の B-WRO, N-WRO における WDBC の事前データとブートストラップ, 事後データから平均分散フロンティアを示し, 図 1.5 では比較のため B-WRO と N-WRO の事後データの平均分散フロンティアを示す. 図 4.3, 図 4.4 より, どのフロンティアもパラメータを与えることで, 大幅に分散が小さくなった. ブートストラップの平均分散フロンティアは WDBC の場合, $\varepsilon = 0$ 以外はほとんど事前データと同じ結果になった. また, $\varepsilon = 0.05$ より大きい場合, 事後データの平均分散フロンティアが事前データを上回っている. 事前データの決定変数から, よりよい結果をだすことができた.

B-WRO のブートストラップの平均分散フロンティアは, $\varepsilon = 0$ の一般的な経験的最適化と $\varepsilon = 0.001$ では分散が約 $\frac{1}{2}$ になり, B-WRO でも一般的な経験的最適化と比べて $\varepsilon = 0.1$ を与えたほうが, 分散が半分になった. また, 図 1.5 から B-WRO と N-WRO を比較すると, 経験分布上だけでない B-WRO のほうがパフォーマンスが良くなった.

ILPD 図 4.6, 図 4.7 は WDBC の B-WRO, N-WRO における WDBC の事前データとブートストラップ, 事後データから平均分散フロンティアを示し, 図 1.8 では比較のため B-WRO と N-WRO の事後データの平均分散フロンティアを示す. WDBC と同様にどのフロンティアもパラメータを与えることで, 平均がわずかに減っているものの大幅に分散が小さくなった. B-WRO のブートストラップの平均分散フロンティアは, $\varepsilon = 0$ の一般的な経験的最適化と $\varepsilon = 0.001$ では分散が約 $\frac{1}{2}$ になり, B-WRO でも一般的な経験的最適化と比べて $\varepsilon = 0.1$ を与えたほうが, 分散が約 $\frac{1}{2}$ になった. 図 1.8 から B-WRO と N-WRO を比較すると, 経験分布上のみで最悪の分布を想定する N-WRO のほうがパフォーマンスが良くなった.

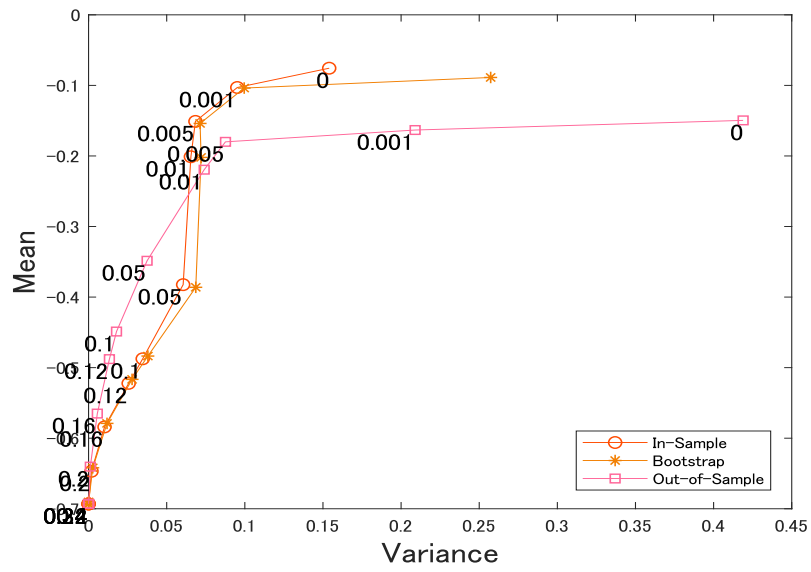


図 4.3: B-WRO における WDBC の事前データの平均分散フロンティア

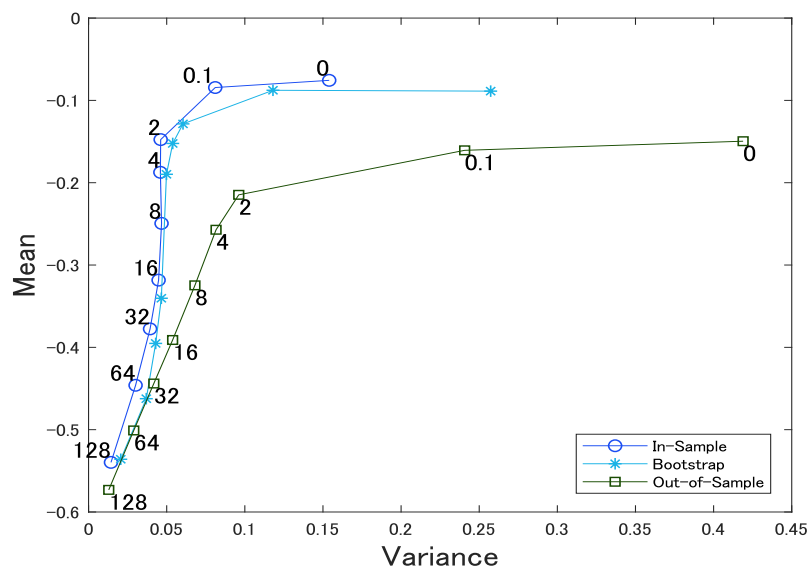
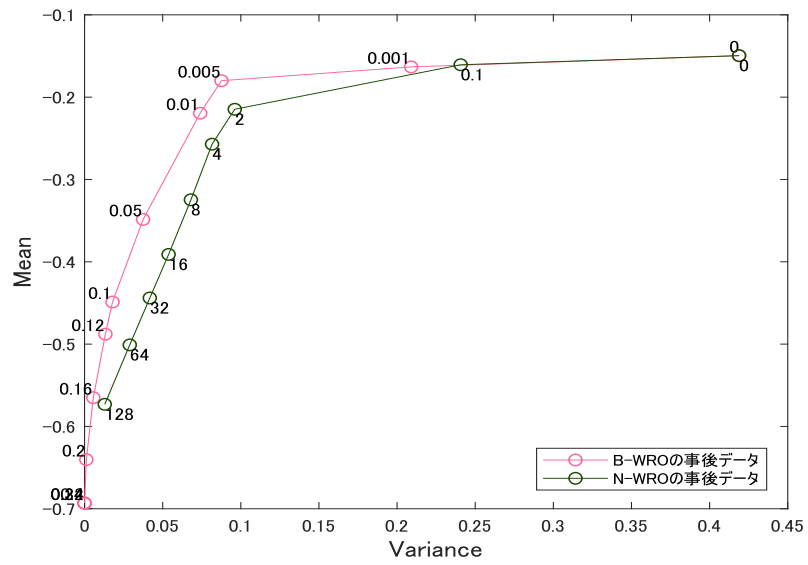


図 4.4: N-WRO における WDBC の事前データの平均分散フロンティア



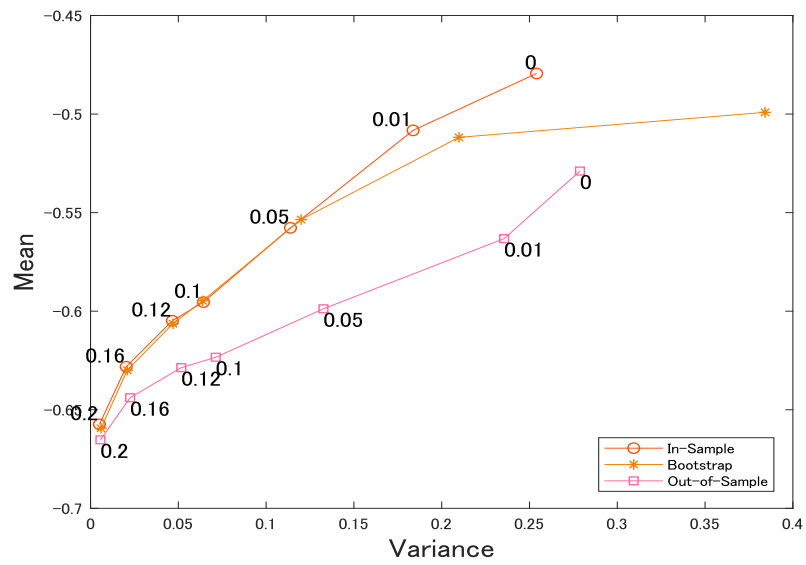


図 4.6: B-WRO における ILPD の平均分散フロンティア

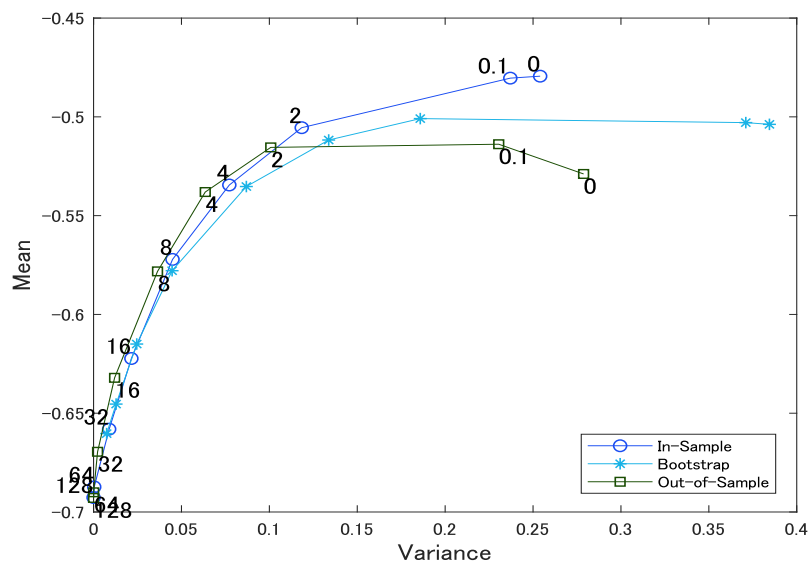


図 4.7: N-WRO における ILPD の平均分散フロンティア

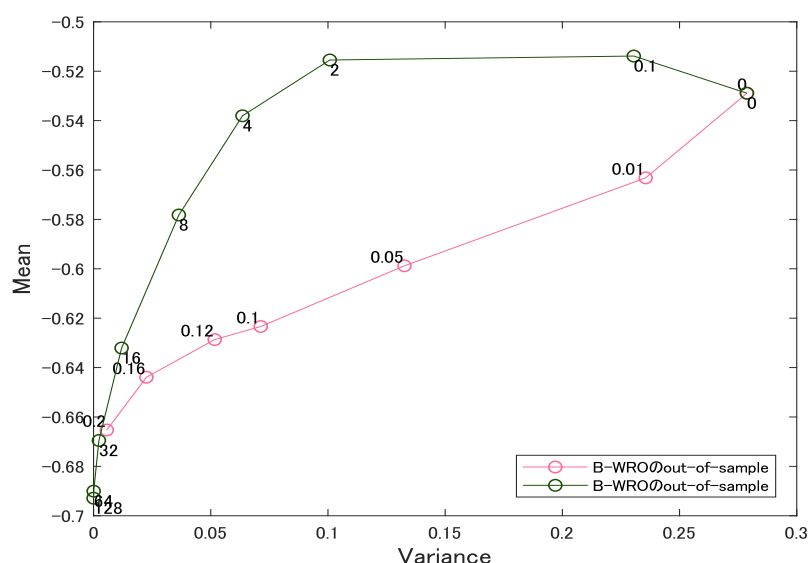


図 4.8: ILPD の事後データの平均分散フロンティア

ブートストラップにおける最適解の分布 ブートストラップ法で求めた解の分布を図 1.9, 図 1.10 に示す. パラメータを与えた分布ロバスト最適化の方が経験的最適化よりもそれぞれの解 x が収束していくことが見てとれる. 比較に利用したパラメータは B-WRO では $\varepsilon = 0, 0.01, 0.16$, N-WRO では $\varepsilon = 0, 0.1, 128$ であり, これは経験的最適解と, 平均をほとんど下げずに分散を大幅に下げることができたパラメータ, 与えた中で最も大きいパラメータをである. また, 表 4.2, 表 4.3 で概要を示す. どちらもパラメータが大きくなると解 x も分布の幅が狭くなっていく. B-WRO の $\varepsilon = 0.001$ と N-WRO の $\varepsilon = 0.001$ を比較すると, B-WRO の方が分布の幅が狭い, 図 1.5 において B-WRO のほうがパフォーマンスが良くなったことが図 1.9 でも言うことができる.

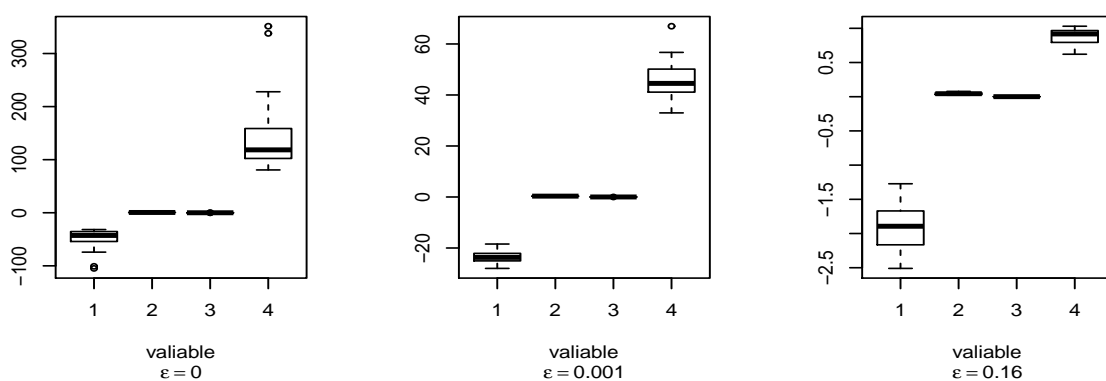


図 4.9: WDBC のブートストラップ法における B-WRO の解の分布

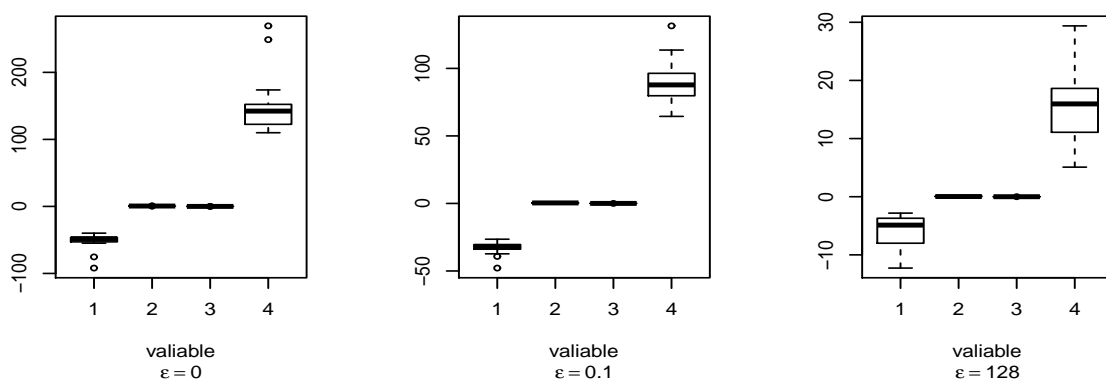


図 4.10: WDBC のブートストラップ法における N-WRO 解の分布

表 4.2: WDBC の B-WRO におけるブートストラップのロバスト解の分布の概要

	$\varepsilon = 0$				$\varepsilon = 0.001$				$\varepsilon = 0.16$			
	x_0	x_1	x_2	x_3	x_0	x_1	x_2	x_3	x_0	x_1	x_2	x_3
最小値	-91.99	0.3253	0.01554	104.6	-28.00	0.1462	0.01047	32.98	-2.510	0.02172	0.0005330	0.6210
第一四分位数	-51.36	0.4834	0.02095	115.4	-24.97	0.2784	0.01357	41.23	-2.164	0.03204	0.0007662	0.7960
中央値	-47.30	0.5565	0.02452	138.6	-23.68	0.3426	0.01404	44.54	-1.894	0.04418	0.0008594	0.9174
平均値	-49.82	0.5498	0.02567	145.0	-23.49	0.3180	0.01431	45.98	-1.926	0.04479	0.0008483	0.8733
第三四分位数	-41.86	0.6068	0.02787	150.0	-22.11	0.3709	0.01541	49.11	-1.674	0.05498	0.0009639	0.9659
最大値	-31.86	0.9059	0.04926	269.4	-18.45	0.4102	0.01754	66.96	-1.273	0.07796	0.0010493	1.0305

表 4.3: WDBC の N-WRO におけるブートストラップのロバスト解の分布の概要

	$\varepsilon' = 0$				$\varepsilon' = 0.1$				$\varepsilon' = 128$			
	x_0	x_1	x_2	x_3	x_0	x_1	x_2	x_3	x_0	x_1	x_2	x_3
最小値	-91.99	0.3253	0.01554	104.6	-47.85	0.2106	0.01229	64.46	-12.271	0.01487	0.002044	5.085
第一四分位数	-51.36	0.4834	0.02095	115.4	-33.63	0.3326	0.01501	79.81	-7.979	0.02767	0.002428	11.398
中央値	-47.30	0.5565	0.02452	138.6	-32.60	0.3751	0.01655	87.78	-4.884	0.03332	0.002754	15.962
平均値	-49.82	0.5498	0.02567	145.0	-33.01	0.3939	0.01734	90.53	-5.963	0.06532	0.003477	15.666
第三四分位数	-41.86	0.6068	0.02787	150.0	-30.54	0.4634	0.01935	95.21	-3.788	0.10703	0.004196	18.612
最大値	-31.86	0.9059	0.04926	269.4	-26.48	0.5703	0.02943	131.49	-2.823	0.15715	0.007255	29.379

ロジスティック回帰における数値実験をまとめる．分布的ロバスト最適化問題の利点である分散を減らすことは, WDBC と ILPD から十分言える．WDBC の図 1.5 と ILPD の図 1.8 より, 説明変数に対する変数のデータ数ではないかと考えた．ILPD は WDBC と標本数がほとんど変わらないが, WDBC の方が説明変数 3, ILPD は 11 である．説明変数に対する変数のデータ数が少ない場合は, 最悪な分布を経験分布上のみから想定する-WRO の方が良いパフォーマンスをしするのではないかと考えた．これは他のデータを利用して検証する必要がある．

4.2 ポートフォリオ選択問題

利用する実データは月次データ 5 Industry Portfolios[12], 10 Industry Portfolios[13] を利用する．数値実験に利用する実データは以下の表 4.4 に示す．

表 4.4: ポートフォリオ選択問題の数値実験に用いたデータ

	事前データ	事前データの時期	事後データ	事後データの期間
(1)	100	1926 年 7 月~1934 年 10 月	10	1934 年 11 月~1935 年 8 月
(2)	100	1968 年 1 月~1976 年 4 月	10	1976 年 5 月~1977 年 3 月
(3)	300	1926 年 7 月~1951 年 6 月	10	1951 年 7 月~1952 年 4 月
(4)	300	1968 年 1 月~1992 年月	10	1993 年 1 月~1993 年 10 月

5 Industry Portfolios の銘柄数は 5 個, 10 Industry Portfolios の銘柄数は 10 個である．銘柄が 5 個, 事前データの数 が 100, 時期が 1926 年 7 月~1934 年 10 月の場合であれば「5 - (1)」, 銘柄が 10 個, 事前データの数 が 100, 時期が 1926 年 7 月~1934 年 10

月の場合であれば「10 - (1)」とする。

事前とブートストラップにおける平均分散フロンティアの比較 事前データとブートストラップの平均分散フロンティアについて考察と比較を行う。今回のポートフォリオ選択問題では、5 銘柄と 10 銘柄の場合で数値実験を行った。事前データの期間が 1926 年 7 月~1934 年 10 月 ($n = 100$) である 5 - (1), 10 - (1) について比較し、事前データの期間が 1926 年 7 月~1951 年 6 月 ($n = 300$) である 5 - (3), 10 - (3) について比較する。

平均分散フロンティアは図 1.11, 図 1.12 に示した。図 1.11 の 5 - (1) は、パラメータを与えても平均はほとんど変わらずに分散が小さくなり、さらにパラメータを大きくすると平均が小さくなり分散も大きくなった。平均が小さくなり分散も大きくなるまで B-WRO と N-WRO がほとんど同じ動きをしていることから、分布的バスト最適化問題の最悪な分布がほとんど同じ形をしているのではないかと考えた。ブートストラップの平均分散フロンティアを見ると、B-WRO は事前データより分散が大きく、N-WRO はそうではなかった。図 1.12 の 10 - (1) は、今まで見せたような平均分散フロンティアを描いていない。銘柄数に対するデータ数が少ないことが原因と考えられる。

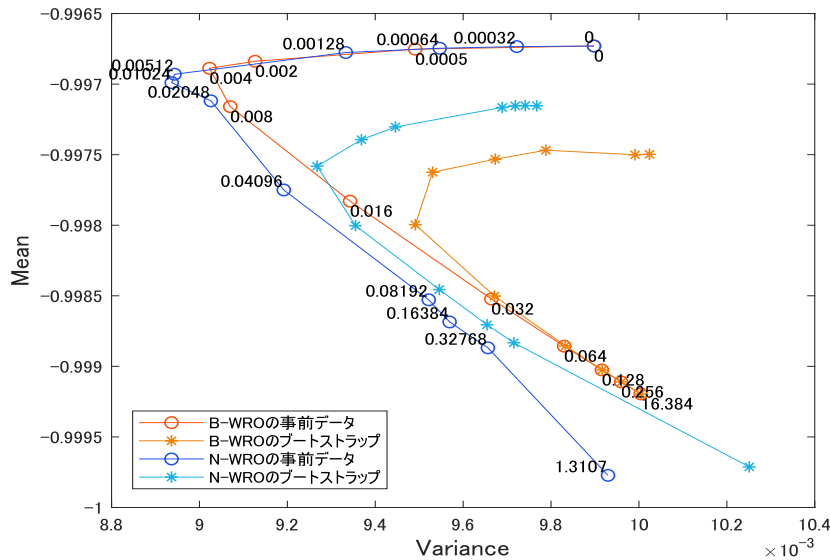


図 4.11: (1)5 銘柄 ($n = 100$) の平均分散フロンティア

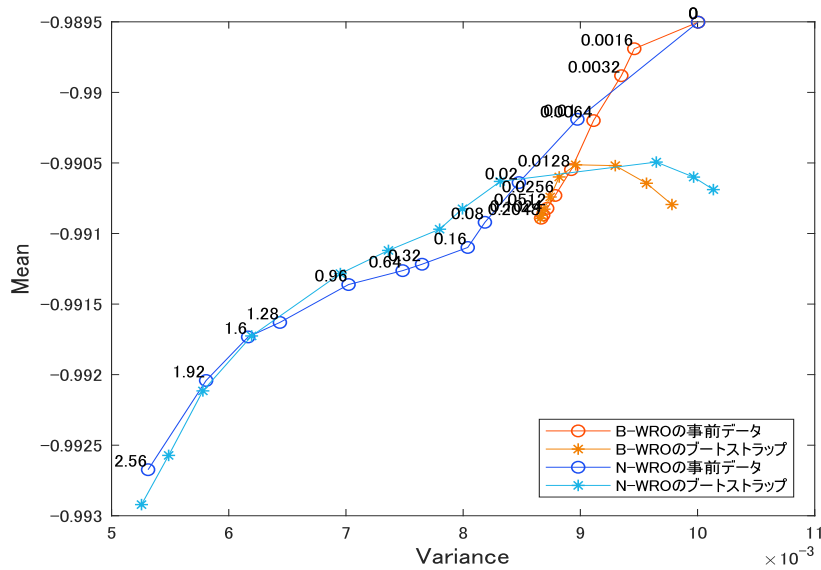


図 4.14: (3)10 銘柄 ($n = 300$) の平均分散フロンティア

図 1.13, 図 1.14 の 5-(1) 事前データの期間が 1926 年 7 月~1934 年 10 月, 事後データが 1934 年 11 月~1935 年 8 月である 5 銘柄のポートフォリオ選択問題と 10-(3) 事前データの期間が 1926 年 7 月~1951 年 6 月, 事後データ 1951 年 7 月~1952 年 4 月の 10 銘柄のポートフォリオ選択の事前データとブートストラップの平均分散フロンティアを示す。

B-WRO の $\varepsilon = 0$, N-WRO $\varepsilon' = 0$ のときは, 一般的な経験的最適化問題である。ロジスティック回帰と同様, パラメータを与えることで分散が小さくなっているが, 図 1.13 では, 平均が小さくなったまま分散が大きくなってしまったり, 図 1.14 では N-WRO が平均も分散も下がり続けるという結果になった。他の数値実験は, 付録に載せている。ポートフォリオ選択問題の実験結果から言える事前とブートストラップにおける平均分散フロンティアの振るまいについてまとめる。分布的ロバスト最適化問題の分散が小さくなるという特徴は, データ数が少なかった 10 - (1) を除いてどの結果でもいえた。異なるパラメータを与えても, 問題によっては平均と分散の値になることがある。その様子は N-WRO に多く見られた。異なるパラメータを与えても平均と分散が同じ値をとることで, 複数のプロットが重なることがある。さらに大きいパラメータを与えることで, 平均と分散の値が変わるが, 図 1.14 のように変化が緩やかではない。これは分布的ロバスト最適化における最悪な分布について, N-WRO は経験分布の台に限定した中で最悪の分布を想定するので, データ数が少ないために異なるパラメータを与えたときでも最悪の分布が同じ振る舞いになってしまうと考えられる。ロジスティックと異なり, 平均分散フロンティアな平均分散フ

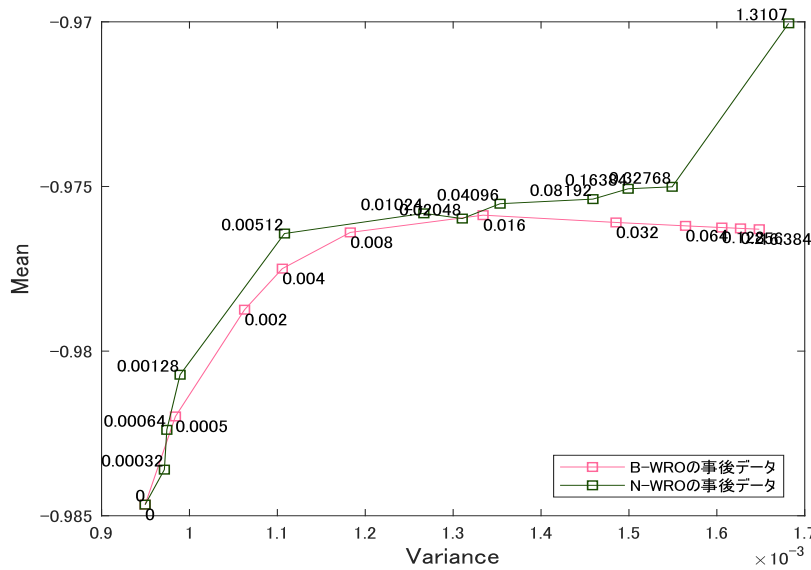


図 4.15: (1)5 銘柄 ($n = 100$) の事後データの平均分散フロンティア

フロンティアを描いた。

事後パフォーマンスの比較 事後データのパフォーマンスについて考察と比較をする。ポートフォリオ選択問題は事前データと事後データの時期によって、大きく変わってしまう。ロジスティック回帰の平均分散フロンティア事後データと違い、事後データが事前データと全く異なる振る舞いをすることが多い。図 4.15 は、事前データの期間が 1926 年 7 月~1934 年 10 月 ($n = 100$) である 5 - (1) から求めた最適解を直後の 10 ヶ月に当てはめて平均分散フロンティアを描いたものである。どちらの平均分散フロンティアもパラメータが 0 である経験的最適化問題が図の左下にあり、パラメータを大きくすると平均も分散も大きくなる。他の数値実験の事後パフォーマンスは、付録に載せている。

ポートフォリオ選択問題の数値実験についてまとめると、事前データの平均分散フロンティアについてはロジスティック回帰と同様に狭義のワッサーズタイン距離を用いた場合の方がパフォーマンスが良くなる傾向があった。しかし、事後データに当てはめた場合はどの期間やどの銘柄でも事前データとは異なる振る舞いをした。ロジスティック回帰とは異なり、政治や経済などに大きく影響を受けることが原因であり、数値実験では期間について工夫する必要がある。例えばリーマンショックなどの大きく株価が変動した時期や景気が横ばいである時期など、あらかじめ時期を決めて結果を見る必要があると考えた。

5 結論

本論文では、分布的ロバスト最適化問題において「代替分布の台が経験分布の台に限る場合」と「代替分布の台が経験分布の台に限らない場合」の2つ場合について比較を行った。「最悪な分布の台が経験分布の台に限らない場合」である広義のワッサーズタイン距離を用いた分布的ロバスト最適化問題と「最悪な分布の台が経験分布の台に限る場合」である狭義のワッサーズタイン距離を用いた分布的ロバスト最適化問題の定式化を示した。ロジスティック回帰とポートフォリオ選択問題について広義のワッサーズタイン距離を用いた分布的ロバスト最適化問題と狭義のワッサーズタイン距離を用いた分布的ロバスト最適化問題を具体的に示し、それを用いて数値実験を行った。

ワッサーズタイン距離を用いた場合に限らず、分布的ロバスト最適化問題はパラメータを与えることで最適値のパフォーマンスをわずかに下げるが、最適値の分散を大幅に下げるという利点がある。これを用いて評価を行った。

数値実験の結果として、ロジスティック回帰についてはデータ数に対して説明変数が少ない実データでは広義のワッサーズタイン距離を用いた分布的ロバスト最適化問題の方が良いパフォーマンスをし、データ数に対して説明変数が多い実データでは広義のワッサーズタイン距離を用いた分布的ロバスト最適化問題の方が良いパフォーマンスをした。

ポートフォリオ選択問題では、事前データについてはロジスティック回帰と似たような結果を得ることはできたが、事後データについては2つの最適化問題を比較した場合にこれといった結果を得ることが出来なかった。ポートフォリオ選択問題については、数値実験に利用する期間や標本数を工夫する必要がある。

今後の課題としては、狭義のワッサーズタイン距離を用いた分布的ロバスト最適化問題については計算時間がかかる傾向があった。より早い計算ができる方法を考えるべきである。

謝辞

本研究を進めるにあたりご指導頂いた後藤順哉先生に深く感謝いたします。研究室に入った当初は未熟で何も分からない状態でしたが、文献の読み方や研究の進め方などをご指導頂きました。後藤先生のご指導のもとで修士論文を形にすることが出来ました。また、論文をご精読いただきました副査の生田目先生、渡邊先生に感謝いたします。共に大学院生活を過ごした同期の皆様、および後藤研究室の皆様に感謝いたします。皆様のおかげで充実した研究室生活を送ることができました。

参考文献

- [1] A. Ben-Tal, N. Bayraksan, “Robust Solutions of Optimization Problems Affected by Uncertain Probabilities,” *Management Science* 59(2):341-357, 2013.
- [2] A. J. Newell, and L. D. Griffin, ”Multiscale histogram of oriented gradient descriptors for robust character recognition,” *Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE*, 2011.
- [3] A. N. Tikhonov, V. Y. Arsenin, *Solutions of ill-posed problems*, volume14 (Winston Washington, DC) ,1977.
- [4] A.Shapiro, “On duality theory of conic linear problems.” *Semi-infinite programming. Springer, Boston, MA,. 135-165*, 2001
- [5] D. Kuhn, S. S. Abadeh and P. M. Esfahani, “Distributionally Logistic Regression,” arXiv:1509.09259v3, 2015.
- [6] D. Wozabal, “Robustifying Convex Risk Measures for Linear Portfolios: A Nonparametric Approach,” *Operation Research*62(6):1302-1315, 2014.
- [7] G. Bayraksan, D. K. Love, “Data-Driven Stochastic Programming Using Phi-Divergences,” *The Optimaizations Research Revolution, 1-19(INFORMS)*, 2017.
- [8] J. Blanchet, K. Murthy, “Quantifying Distributional Model Risk Via Optimal Transport,” arXiv:1604. 01446, 2016.
- [9] J. Gotoh, M. J. Kim and A. E. B. Lim, “Calibration of Distributionally Robust Empirical Optimization Models,” arXiv:1711.06565v1, 2017.

- [10] H. Beyer, B. Sendhoff, "Robust optimization: a comprehensive survey." *Computer methods in applied mechanics and engineering* 196.33-34: 3190-3218, 2007.
- [11] J. Gotoh, "Robust Empirical Optimization with Wasserstein Uncertainty, *unpublished*, 2017.
- [12] K. R. French, 2017 *5 Industry Portfolios*. URL: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_5_ind_port.html.
- [13] K. R. French, 2017 *10 Industry Portfolios*. URL: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html.
- [14] L. E. Ghaoui, M. Oks, and F. Oustry. "Worst-case value-at-risk and robust portfolio optimization: A conic programming approach," *Operations research* 51.4: 543-556, 2003.
- [15] M. Lichman, UCI Machine Learning Repository, Irvine, CA: University of California, School of Information and Computer Science, 2013. URL: <http://archive.ics.uci.edu/ml>.
- [16] P. M. Esfahani, D. Kuhn, "Data-Driven Distributionally Robust Optimization Using the Wasserstein Metric: Performance Guarantees and Tractable Reformulations," *Mathematical Programming* ISSN 1436-4646, 2017.
- [17] R. Gao, A. J. Kleywegt, "Distributionally Robust Stochastic Optimization with Dependence Distance," arXiv preprint arXiv:1701.04200, 2017.
- [18] R. Gao, X. Chen and A. J. Kleywegt, "Distributional Robustness and Regularization in Statistical Learning," arXiv:1712.06050v2, 2017.
- [19] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)* 267-288, 1996.
- [20] R. Jiang, Y. Guan, "Data-driven chance constrained stochastic program," *Mathematical Programming* 1-37, 2015.
- [21] Z. Wang and P. W. Glynn, Y. Ye, "Likelihood Robust Optimization for Data-driven Problems," *Computational Management Science*, 351-359, 2016.

A 付録

A.1 広義のワッサースタイン距離を用いた場合の定式化の変形

A.1.1 ロジスティック回帰の式変形

ここでは、経験データ (z_i, y_i) と (z, y) を区別しやすいように、経験データを $(\hat{z}_i, \hat{y}_i) = \xi'_i, \Xi = \mathbb{R}^n \times \{-1, 1\}$ とし、 $(z, y) = \xi_i$ と表す。2つの距離は $((z, y), (\hat{z}_i, \hat{y}_i)) = \|z - \hat{z}_i\| + \frac{\kappa}{2}|y - \hat{y}_i|$ とし、 κ は正の重みである。

$$\begin{aligned}
& \inf_{Q \in \mathcal{M}(\hat{\mathbb{P}}_n)} \{ \mathbb{E}_{\hat{\mathbb{P}}_n}[l_{\mathbf{x}}(\xi)] | W(Q | \hat{\mathbb{P}}_n) \leq \varepsilon \} \\
&= \inf_{Q \in \mathcal{M}(\hat{\mathbb{P}}_n)} \int_{\Xi} -l_{\mathbf{x}}(\xi) Q(d\xi) \\
&= \left\{ \begin{array}{l} \inf_{\Pi \in M(\Xi^2)} \int_{\Xi} l_{\mathbf{x}}(\xi) \Pi(d\xi, \Xi) \\ \text{s.t.} \quad \int_{\Xi^2} d(\xi, \xi') \Pi(d\xi, d\xi') \leq \varepsilon \\ \Pi(\Xi, d\xi') = \hat{\mathbb{P}}_n(d\xi') \end{array} \right. \\
&= \left\{ \begin{array}{l} \inf_{Q^i} \frac{1}{n} \sum_{i=1}^n \int_{\Xi} l_{\mathbf{x}}(\xi) Q^i(d\xi) \\ \text{s.t.} \quad \frac{1}{n} \sum_{i=1}^n \int_{\Xi} d(\xi, \xi') Q^i(d\xi) \leq \varepsilon \\ \int_{\Xi} Q^i(d\xi) = 1 \end{array} \right. \\
&= \left\{ \begin{array}{l} \inf_{Q_{\pm 1}^i} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^n} l_{\mathbf{x}}(z, +1) Q_{+1}^i(dz) + l_{\mathbf{x}}(z, -1) Q_{-1}^i(dz) \\ \text{s.t.} \quad \frac{1}{n} \sum_{i=1}^n d((z, +1)) - (\hat{z}_i, \hat{y}_i) Q_{+1}^i(dz) + \sum_{i=1}^n d((z, -1)) - (\hat{z}_i, \hat{y}_i) Q_{-1}^i(dz) \leq \varepsilon \\ \int_{\mathbb{R}^n} Q_{+1}^i(dz) + Q_{-1}^i(dz) = 1 \end{array} \right.
\end{aligned}$$

1 つ目の制約式の左辺を変形させる.

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n d((z, +1)) - (\hat{z}_i, \hat{y}_i) \mathbb{Q}_{+1}^i(dz) + \sum_{i=1}^n d((z, -1)) - (\hat{z}_i, \hat{y}_i) \mathbb{Q}_{-1}^i(dz) \\
&= \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=+1} d((z, +1), (\hat{z}_i, +1)) \mathbb{Q}_{+1}^i(dz) + d((z, -1), (\hat{z}_i, +1)) \mathbb{Q}_{-1}^i(dz) \\
&\quad + \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=-1} d((z, +1), (\hat{z}_i, -1)) \mathbb{Q}_{+1}^i(dz) + d((z, -1), (\hat{z}_i, -1)) \mathbb{Q}_{-1}^i(dz) \\
&= \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=+1} \left\{ \|\hat{z}_i - z\| + \frac{\kappa}{2} |1 - 1| \right\} \mathbb{Q}_{+1}^i(dz) \\
&\quad + \left\{ \|\hat{z}_i - z\| + \frac{\kappa}{2} |1 - 1| \right\} \mathbb{Q}_{-1}^i(dz) \\
&\quad + \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=-1} \left\{ \|\hat{z}_i - z\| + \frac{\kappa}{2} |1 - (-1)| \right\} \mathbb{Q}_{+1}^i(dz) \\
&\quad + \left\{ \|\hat{z}_i - z\| + \frac{\kappa}{2} |1 - (-1)| \right\} \mathbb{Q}_{-1}^i(dz) \\
&= \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=+1} \|z - \hat{z}_i\| \mathbb{Q}_{+1}^i(dz) + \|z - \hat{z}_i\| \mathbb{Q}_{-1}^i(dz) + \kappa \mathbb{Q}_{-1}^i(dz) \\
&\quad + \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=-1} \|z - \hat{z}_i\| \mathbb{Q}_{-1}^i(dz) + \|z - \hat{z}_i\| \mathbb{Q}_{+1}^i(dz) + \kappa \mathbb{Q}_{+1}^i(dz)
\end{aligned}$$

変形後, 次のように示すことができた.

$$\left\{ \begin{array}{l} \inf_{\mathbb{Q}_{\pm 1}^i} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^n} l_x(z, +1) \mathbb{Q}_{+1}^i(dz) + l_x(z, -1) \mathbb{Q}_{-1}^i(dz) \\ \text{s.t.} \quad \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=+1} \left\{ \|\hat{z}_i - z\| \mathbb{Q}_{+1}^i(dz) + \|\hat{z}_i - z\| \mathbb{Q}_{-1}^i(dz) + \kappa \mathbb{Q}_{-1}^i(dz) \right\} + \\ \quad \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=-1} \left\{ \|\hat{z}_i - z\| \mathbb{Q}_{-1}^i(dz) + \|\hat{z}_i - z\| \mathbb{Q}_{+1}^i(dz) + \kappa \mathbb{Q}_{+1}^i(dz) \right\} \leq \varepsilon \\ \int_{\mathbb{R}^n} \mathbb{Q}_{+1}^i(dz) + \mathbb{Q}_{-1}^i(dz) = 1 \end{array} \right.$$

制約式の 1 つ目の左辺を変形させる.

$$\begin{aligned}
& \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=+1} \left\{ \|z - \hat{z}_i\| \mathbb{Q}_{+1}^i(dz) + \|z - \hat{z}_i\| \mathbb{Q}_{-1}^i(dz) + \kappa \mathbb{Q}_{-1}^i(dz) \right\} \\
&+ \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=-1} \left\{ \|z - \hat{z}_i\| \mathbb{Q}_{-1}^i(dz) + \|z - \hat{z}_i\| \mathbb{Q}_{+1}^i(dz) + \kappa \mathbb{Q}_{+1}^i(dz) \right\} \\
&= \frac{1}{n} \sum_{\hat{y}_i=1}^n \int_{\mathbb{R}^n} \|z - \hat{z}_i\| \{ \mathbb{Q}_{+1}^i(dz) + \mathbb{Q}_{-1}^i(dz) \} \\
&\quad + \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=+1} \kappa \mathbb{Q}_{-1}^i(dz) + \frac{1}{n} \int_{\mathbb{R}^n} \sum_{\hat{y}_i=-1} \kappa \mathbb{Q}_{+1}^i(dz) \\
&= \frac{1}{n} \int_{\mathbb{R}^n} \frac{\kappa}{n} \sum_{\hat{y}_i=+1} \mathbb{Q}_{-1}^i(dz) + \sum_{i=1}^n \|\hat{z}_i - z\| (\mathbb{Q}_{-1}^i + \mathbb{Q}_{+1}^i)
\end{aligned}$$

$$\left\{ \begin{array}{l} \inf_{\mathbb{Q}_{\pm 1}^i} \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}^n} l_{\mathbf{x}}(\mathbf{z}, +1) \mathbb{Q}_{+1}^i(d\mathbf{z}) + l_{\mathbf{x}}(\mathbf{z}, -1) \mathbb{Q}_{-1}^i(d\mathbf{z}) \\ \text{s.t.} \quad \frac{1}{n} \int_{\mathbb{R}^n} \frac{\kappa}{n} \sum_{\hat{y}_i=+1} \mathbb{Q}_{-1}^i(d\mathbf{z}) + \sum_{i=1}^n \|\hat{\mathbf{z}}_i - \mathbf{z}\| (\mathbb{Q}_{-1}^i + \mathbb{Q}_{+1}^i) \leq \varepsilon \\ \int_{\mathbb{R}^n} \mathbb{Q}_{+1}^i(d\mathbf{z}) + \mathbb{Q}_{-1}^i(d\mathbf{z}) = 1 \end{array} \right.$$

半無限双対を用いて, 変形させる [4].

$$\left\{ \begin{array}{l} \sup_{\lambda, s_i} \quad \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n s_i \\ \text{s.t.} \quad \inf_{\mathbf{x} \in \mathbb{R}^n} l_{\mathbf{x}}(\mathbf{z}, +1) - \lambda \|\hat{\mathbf{z}}_i - \mathbf{z}\| - \frac{1}{2} \lambda \kappa (1 - \hat{y}_i) \geq s_i \\ \inf_{\mathbf{x} \in \mathbb{R}^n} l_{\mathbf{x}}(\mathbf{z}, -1) - \lambda \|\hat{\mathbf{z}}_i - \mathbf{z}\| - \frac{1}{2} \lambda \kappa (1 + \hat{y}_i) \geq s_i \\ \lambda \geq 0 \end{array} \right.$$

[5] の定理 1 より, 次のように表すことができる.

$$\left\{ \begin{array}{l} \max_{\mathbf{x}, \lambda, s_i} \quad \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n s_i \\ \text{s.t.} \quad -l_{\mathbf{x}}(\hat{\mathbf{z}}_i, +1) - \frac{1}{2} \lambda \kappa (1 - \hat{y}_i) \geq s_i \\ -l_{\mathbf{x}}(\hat{\mathbf{z}}_i, -1) - \frac{1}{2} \lambda \kappa (1 + \hat{y}_i) \geq s_i \\ \|\mathbf{x}\| \leq \lambda \end{array} \right. \quad (\text{A.1})$$

$$= \left\{ \begin{array}{l} \max_{\mathbf{x}, \lambda, s_i} \quad \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n s_i \\ \text{s.t.} \quad -l_{\mathbf{x}}(\hat{\mathbf{z}}_i, \hat{y}_i) \geq s_i \\ -l_{\mathbf{x}}(\hat{\mathbf{z}}_i, -\hat{y}_i) - \lambda \kappa \geq s_i \\ \|\mathbf{x}\| \leq \lambda \end{array} \right. \quad (\text{A.2})$$

このとき, $\lambda, s_i (i = 1, \dots, n)$ は双対係数である. 広義のロバスト最適化問題の目的関数の中は次のように表すことが出来た. これより, 狭義のワッサーズタイン距離を用いたロバスト最適化問題 (2.8) は, 等価な凸計画問題に帰着された.

A.1.2 ポートフォリオ選択問題の式変形

ロジスティック回帰の定式化と同様に, ポートフォリオ選択問題も次のように表すことができる:

$$\begin{cases} \max_{\mathbf{x}, \lambda, s_i} & \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n s_i \\ \text{s.t.} & s_i \leq U_{\mathbf{x}}(\mathbf{R}_i) \\ & \|\mathbf{x}\| \leq \lambda \end{cases}$$

このとき, $\lambda, s_i (i = 1, \dots, n)$ は双対係数である. ポートフォリオ選択問題についても等価な凸計画問題に帰着することができた. さらに, これはより簡単な問題に表すことができる.

$$\begin{cases} \max_{\mathbf{x}, \lambda} & \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n U_{\mathbf{x}}(\mathbf{R}_i) \\ \text{s.t.} & \|\mathbf{x}\| \leq \lambda \end{cases}$$

これをもとに, ポートフォリオ選択問題の制約式を含んだ狭義のワッサーズタイン距離を用いたロバスト最適化問題 (2.8) と等価な凸計画問題に表すことができる:

$$\begin{cases} \max_{\mathbf{x}, \lambda} & \varepsilon \lambda + \frac{1}{n} \sum_{i=1}^n U_{\mathbf{x}}(\mathbf{R}_i) \\ \text{s.t.} & \|\mathbf{x}\| \leq \lambda \\ & \mathbf{1}^T \mathbf{x} = 1 \\ & \mathbf{x} \geq 0 \end{cases}$$

このとき, $\lambda, s_i (i = 1, \dots, n)$ は双対係数である.

A.2 狭義のワッサースタイン距離を用いた場合の定式化の変形

A.2.1 ロジスティック回帰の変形

まず, 式変形で使われる表記法を示す.

- 経験分布 $\hat{\mathbb{P}}_n := (p_1, \dots, p_n), (j = 1, \dots, n)$
- 代替分布 $\mathbb{Q} := (q_1, \dots, q_n), (i = 1, \dots, n)$
- 経験分布 $\hat{\mathbb{P}}_n$ と分布 \mathbb{Q} の同時確率分布 $\Pi := \pi_{ij}$
- p_j と q_i の距離 $d_{ij} := \|p_j - q_i\|$

また, 対数尤度関数について $l_{\mathbf{x}}(\mathbf{z}_i, \mathbf{y}_i) = l_i(\mathbf{x})$ と書く.

ロジスティック回帰のロバスト最適化問題 (2.14) の目的関数を取り出し, ワッサースタイン距離を制約式へ移す.

$$\begin{aligned} & \min_{\mathbb{Q}} \{ \mathbb{E}_{\hat{\mathbb{P}}_n} [l_{\mathbf{x}}(\mathbf{z}, \mathbf{y}) | W'(\mathbb{Q} | \hat{\mathbb{P}}_n) \leq \varepsilon'] \} \\ &= \left\{ \begin{array}{l} \min_{q_i, \pi_{ij}} \quad \sum_{i=1}^n l_i(\mathbf{x}) q_i \\ \text{s.t.} \quad \sum_{j=1}^n \sum_{i=1}^n d_{ij} \pi_{ij} \leq \varepsilon' \\ \sum_{i=1}^n \pi_{ij} = p_j \\ \sum_{j=1}^n \pi_{ij} = q_i \\ \pi_{ij} \geq 0 \end{array} \right. \end{aligned}$$

3 つめの制約式 $\sum_{j=1}^n \pi_{ij} = q_i$ を目的関数に代入し, q_i を消す.

$$\left\{ \begin{array}{l} \min_{q_i, \pi_{ij}} \quad \sum_{i=1}^n l_i(\mathbf{x}) \sum_{j=1}^n \pi_{ij} \\ \text{s.t.} \quad \sum_{j=1}^n \sum_{i=1}^n d_{ij} \pi_{ij} \leq \varepsilon' \\ \sum_{i=1}^n \pi_{ij} = p_j \\ \pi_{ij} \geq 0 \end{array} \right.$$

ここで、この問題の双対問題を導出する。双対係数は1つめの制約式に θ , 2つめに $\mu_j (j = 1, \dots, n)$ として、双対問題の目的関数は $\varepsilon'\theta + \sum_{j=1}^n p_j \mu_j$ である。制約式は

$$\theta \left(\sum_{j=1}^n \sum_{i=1}^n d_{ij} \pi_{ij} \right) + \sum_{j=1}^n \mu_j \left(\sum_{i=1}^n \pi_{ij} \right) \leq \sum_{i=1}^n l_i(\mathbf{x}) \left(\sum_{j=1}^n \pi_{ij} \right)$$

と示せるが、より簡単に示すために展開する。まず、左辺について展開する。

$$\begin{aligned} & \theta \sum_{j=1}^n \sum_{i=1}^n d_{ij} \pi_{ij} + \mu_1 \sum_{i=1}^n \pi_{i1} \cdots + \mu_n \sum_{i=1}^n \pi_{in} \\ = & \theta \sum_{j=1}^n d_{1j} (\pi_{11} + \pi_{12} \cdots + \pi_{1n}) + \mu_1 \pi_{11} + \mu_1 \pi_{21} + \cdots + \mu_1 \pi_{n1} \\ & + \theta \sum_{j=1}^n d_{2j} (\pi_{21} + \pi_{22} \cdots + \pi_{2n}) + \mu_2 \pi_{12} + \mu_2 \pi_{22} + \cdots + \mu_2 \pi_{n2} \\ & \vdots \\ & + \theta \sum_{j=1}^n d_{nj} (\pi_{n1} + \pi_{n2} \cdots + \pi_{nn}) + \mu_n \pi_{1n} + \mu_n \pi_{2n} + \cdots + \mu_n \pi_{nn} \\ = & \left(\theta \sum_{j=1}^n d_{1j} + \mu_1 \right) \pi_{11} + \left(\theta \sum_{j=1}^n d_{1j} + \mu_2 \right) \pi_{12} \cdots + \left(\theta \sum_{j=1}^n d_{1j} + \mu_n \right) \pi_{1n} \\ & \left(\theta \sum_{j=1}^n d_{2j} + \mu_1 \right) \pi_{21} + \left(\theta \sum_{j=1}^n d_{2j} + \mu_2 \right) \pi_{22} \cdots + \left(\theta \sum_{j=1}^n d_{2j} + \mu_n \right) \pi_{2n} \\ & \vdots \\ & \left(\theta \sum_{j=1}^n d_{nj} + \mu_1 \right) \pi_{n1} + \left(\theta \sum_{j=1}^n d_{nj} + \mu_2 \right) \pi_{n2} \cdots + \left(\theta \sum_{j=1}^n d_{nj} + \mu_n \right) \pi_{nn} \end{aligned}$$

次に右辺について展開する。

$$\begin{aligned} & \sum_{i=1}^n l_i(\mathbf{x}) \sum_{j=1}^n \pi_{ij} \\ = & +l_1(\mathbf{x}) \sum_{j=1}^n \pi_{1j} + l_2(\mathbf{x}) \sum_{j=1}^n \pi_{2j} \cdots + l_n(\mathbf{x}) \sum_{j=1}^n \pi_{nj} \\ = & +l_1(\mathbf{x}) \pi_{11} + l_1(\mathbf{x}) \pi_{12} \cdots + l_1(\mathbf{x}) \pi_{1n} \\ & +l_2(\mathbf{x}) \pi_{21} + l_2(\mathbf{x}) \pi_{22} \cdots + l_2(\mathbf{x}) \pi_{2n} \\ & \vdots \\ & +l_n(\mathbf{x}) \pi_{n1} + l_n(\mathbf{x}) \pi_{n2} \cdots + l_n(\mathbf{x}) \pi_{nn} \end{aligned}$$

展開した右辺と左辺を対応させ, π_{ij} を消す.

$$\begin{aligned}
& \left(\theta \sum_{j=1}^n d_{1j} + \mu_1 \right) \pi_{11} \leq l_1(\mathbf{x}) \pi_{11} \\
& \quad \vdots \\
& \left(\theta \sum_{j=1}^n d_{2j} + \mu_1 \right) \pi_{21} \leq l_2(\mathbf{x}) \pi_{21} \\
& \quad \vdots \\
& \left(\theta \sum_{j=1}^n d_{nj} + \mu_n \right) \pi_{nn} \leq l_n(\mathbf{x}) \pi_{nn}
\end{aligned}$$

これをまとめると $\theta \sum_{j=1}^n d_{ij} + \mu_j \leq l_i(\mathbf{x})$ と表すことができ, 狭義のロバスト最適化問題の目的関数の中は次のように表すことが出来る.

$$\left| \begin{array}{ll} \min_{\theta, \mu_j} & \varepsilon' \theta + \sum_{j=1}^n p_j \mu_j \\ \text{s.t.} & \theta \sum_{j=1}^n d_{ij} + \mu_j \leq l_i(\mathbf{x}) \\ & \theta \geq 0 \end{array} \right.$$

このとき, $\theta, \mu_j (1 = 1, \dots, n)$ は双対係数である.

これより, 狭義のワッサーズタイン距離を用いたロバスト最適化問題 (2.13) は, 等価な凸計画問題に帰着された.

$$\left| \begin{array}{ll} \min_{\mathbf{x}, \mu_j, \theta} & \varepsilon' \theta + \sum_{j=1}^n p_j \mu_j \\ \text{s.t.} & \theta \sum_{j=1}^n d_{ij} + \mu_j \leq l_i(\mathbf{x}) \\ & \theta \geq 0 \end{array} \right. \tag{A.3}$$

A.2.2 ポートフォリオ選択問題の変形

ポートフォリオ選択問題の式変形をする．まず式変形で使われる表記法を示す．

- 経験分布 $\hat{\mathbb{P}}_n := (p_1, \dots, p_n), (j = 1, \dots, n)$
- 代替分布 $\mathbb{Q} := (q_1, \dots, q_n), (i = 1, \dots, n)$
- 経験分布 $\hat{\mathbb{P}}_n$ と分布 \mathbb{Q} の同時確率分布 $\Pi := \pi_{ij}$
- p_j と q_i の距離 $d_{ij} := \|p_j - q_i\|$

ロジスティック回帰と同様に, EMD を制約式へ移動させ, $\sum_{j=1}^n \pi_{ij} = q_i$ を消す．

$$= \left\{ \begin{array}{l} \min_{q_i, \pi_{ij}} \sum_{i=1}^n U_{\mathbf{x}}(\mathbf{R}_i) q_i \\ \text{s.t.} \quad \sum_{j=1}^n \sum_{i=1}^n d_{ij} \pi_{ij} \leq \varepsilon' \\ \sum_{i=1}^n \pi_{ij} = p_j \\ \sum_{j=1}^n \pi_{ij} = q_i \\ \pi_{ij} \geq 0 \end{array} \right. = \left\{ \begin{array}{l} \min_{q_i, \pi_{ij}} \sum_{i=1}^n U_{\mathbf{x}}(\mathbf{R}_i) \sum_{j=1}^n \pi_{ij} \\ \text{s.t.} \quad \sum_{j=1}^n \sum_{i=1}^n d_{ij} \pi_{ij} \leq \varepsilon' \\ \sum_{i=1}^n \pi_{ij} = p_j \\ \pi_{ij} \geq 0 \end{array} \right.$$

ロジスティック回帰と同様に, 双対性を利用し変形させる．

$$\left\{ \begin{array}{l} \max_{\theta, \mu_j} \quad \varepsilon' \theta + \sum_{j=1}^n p_j \mu_j \\ \text{s.t.} \quad \theta \sum_{j=1}^n d_{ij} + \mu_j \leq U_{\mathbf{x}}(\mathbf{R}_i) \\ \theta \geq 0 \end{array} \right.$$

このとき, 1 つめの制約式に双対係数 $\theta \geq 0, \mu_j (j = 1, \dots, n)$ 制約式については, ロジスティック回帰の展開と変わらないので省略する．これを用いて, 効用関数に対す

る狭義のロバスト最適化問題 (2.16) を次に示す.

$$\begin{aligned}
 & \max_{\mathbf{x}, \boldsymbol{\mu}, \theta} \quad \varepsilon' \boldsymbol{\theta} + \frac{1}{n} \sum_{j=1}^n \mu_j \\
 & \text{s.t.} \quad \theta \sum_{j=1}^n d_{ij} + \mu_j \leq U_i(\mathbf{x}) \\
 & \quad \theta \geq 0 \\
 & \quad 1^T \mathbf{x} = 1 \\
 & \quad \mathbf{x} \geq 0
 \end{aligned}$$

A.3 ポートフォリオ選択問題の平均分散フロンティア

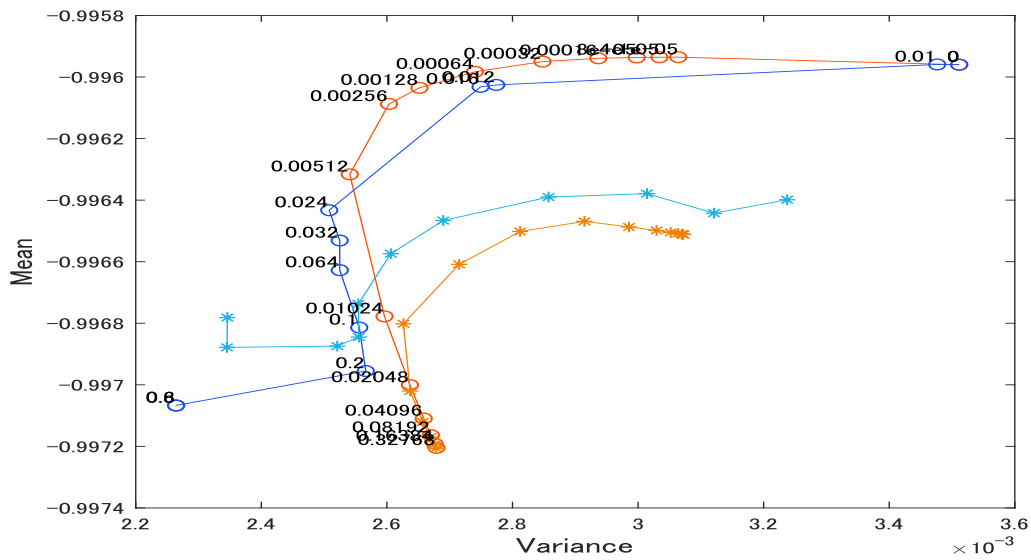


図 A.1: 5 Industry portfolio” (2) の平均分散フロンティア

”5 Industry portfolio” (2) の平均分散フロンティア

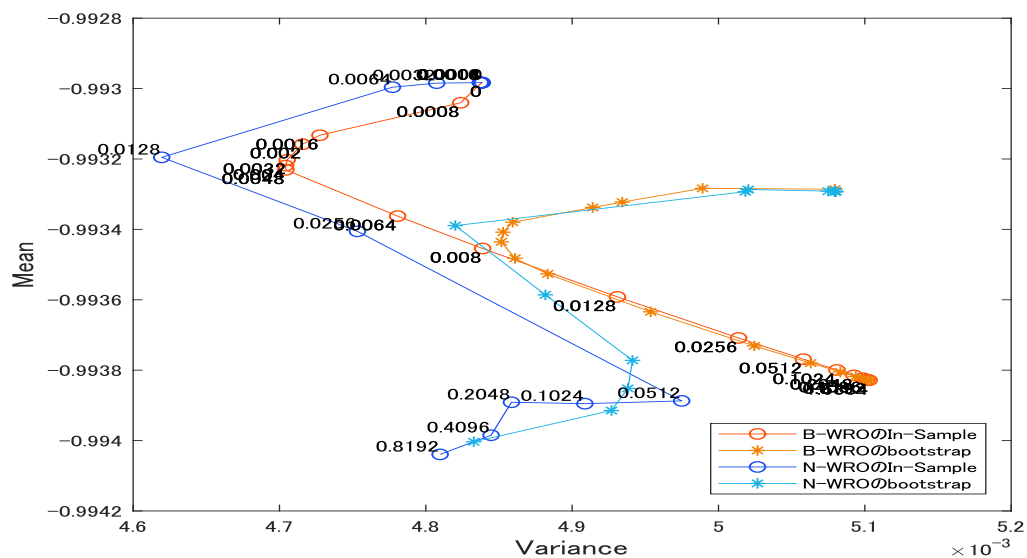


図 A.2: 5 Industry portfolio”(3) の平均分散フロンティア

”5 Industry portfolio”(3) の平均分散フロンティア

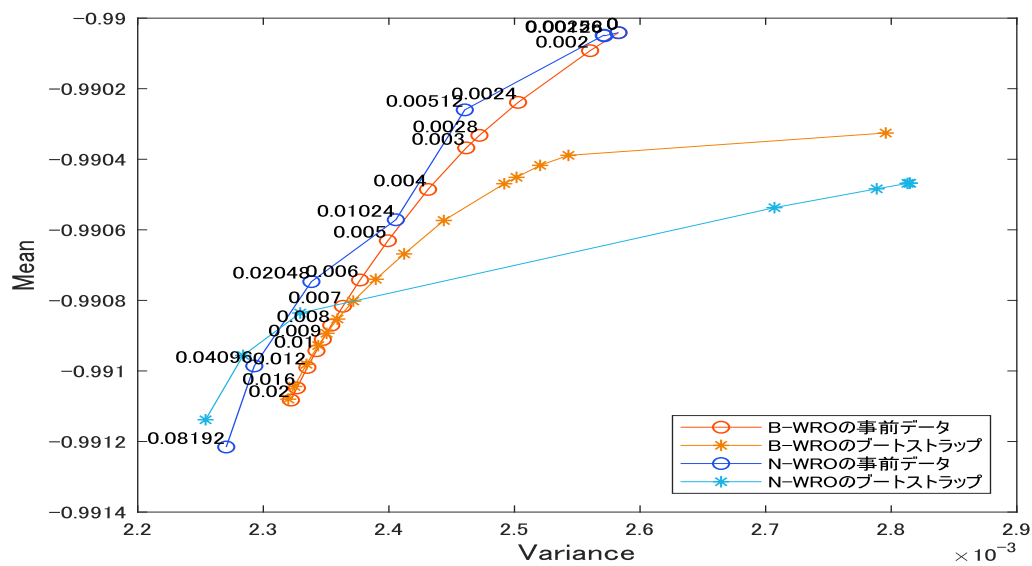


図 A.3: ”5 Industry portfolio”(4) の平均分散フロンティア

”5 Industry portfolio”(4) の平均分散フロンティア (2)=====

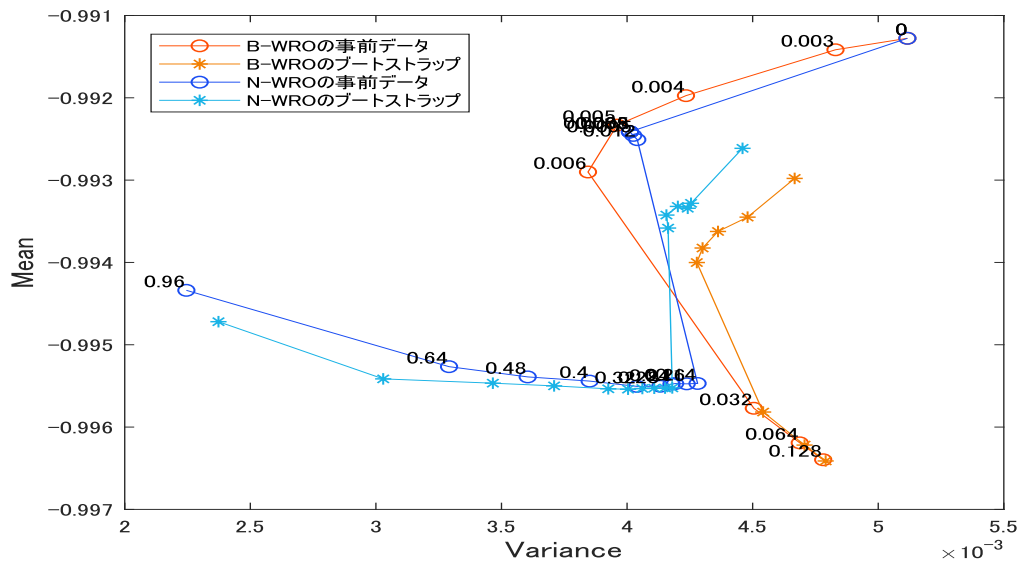


図 A.4: "10 Industry portfolio"(2) の平均分散フロンティア

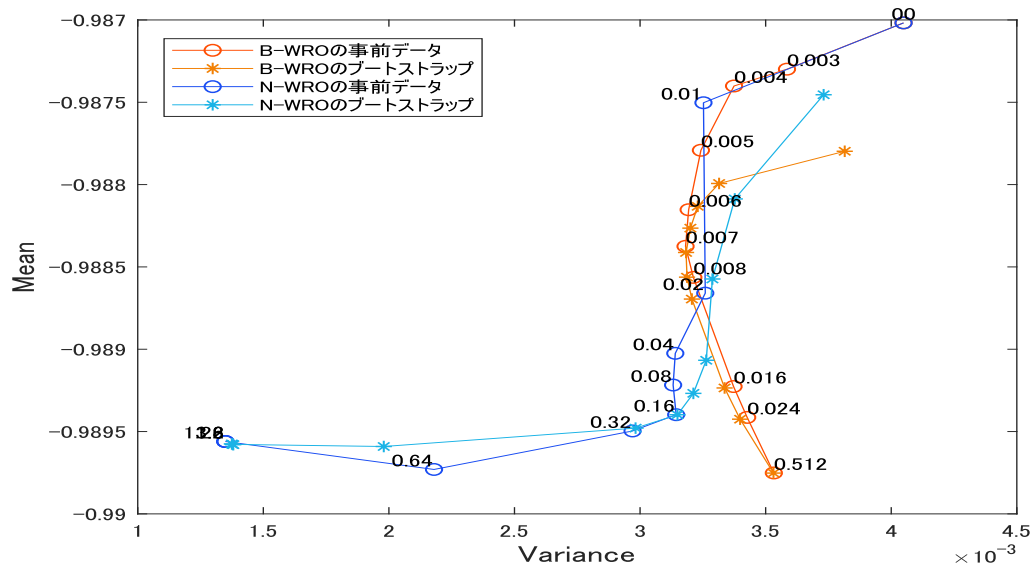


図 A.5: "10 Industry portfolio"(4) の平均分散フロンティア

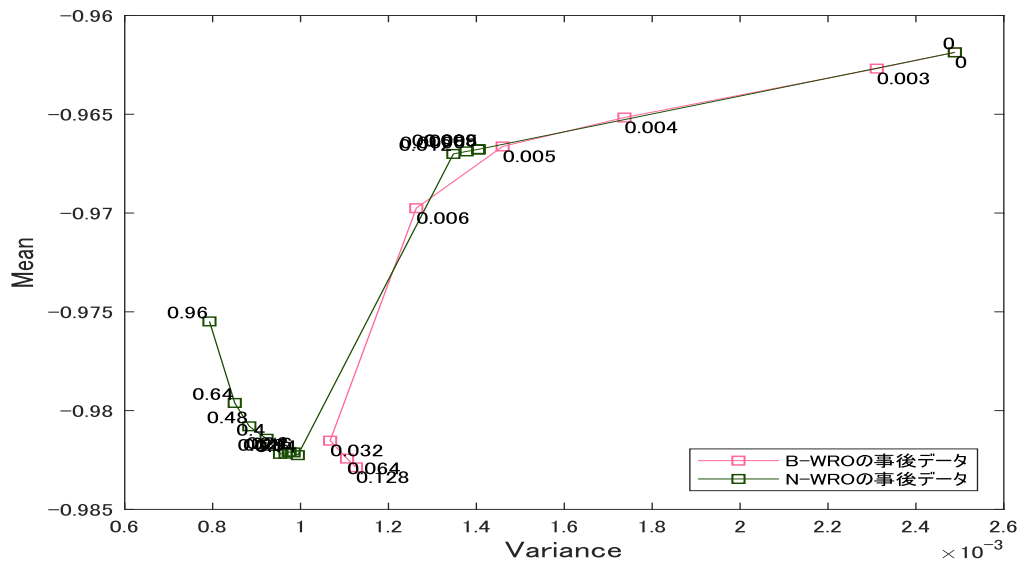


図 A.6: "5 Industry portfolio" (2) の事後データの平均分散フロンティア

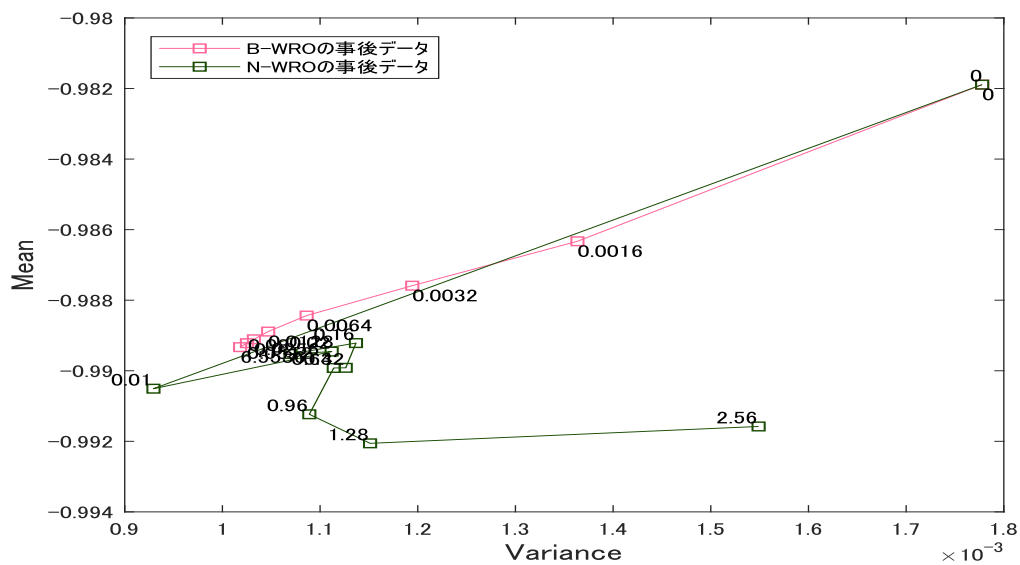


図 A.7: "5 Industry portfolio" (3) の事後データの平均分散フロンティア

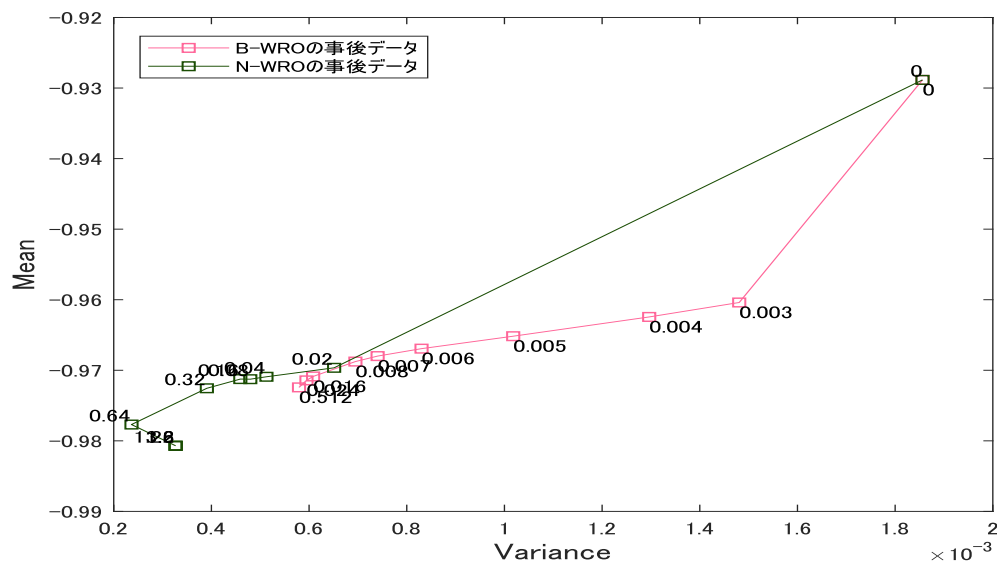


図 A.8: ”5 Industry portfolio”(4) の事後データの平均分散フロンティア

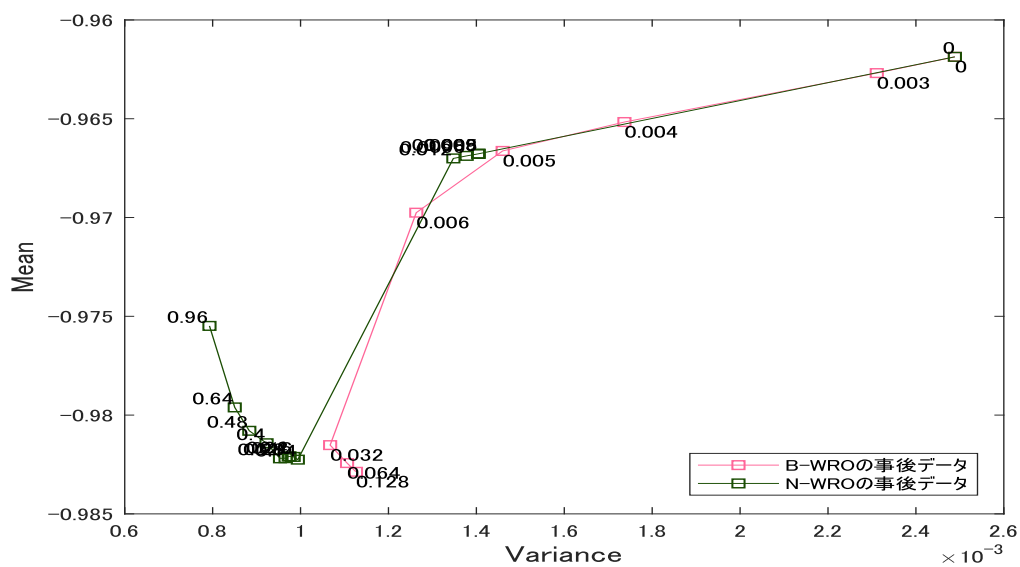


図 A.9: ”10 Industry portfolio”(2) の事後データの平均分散フロンティア

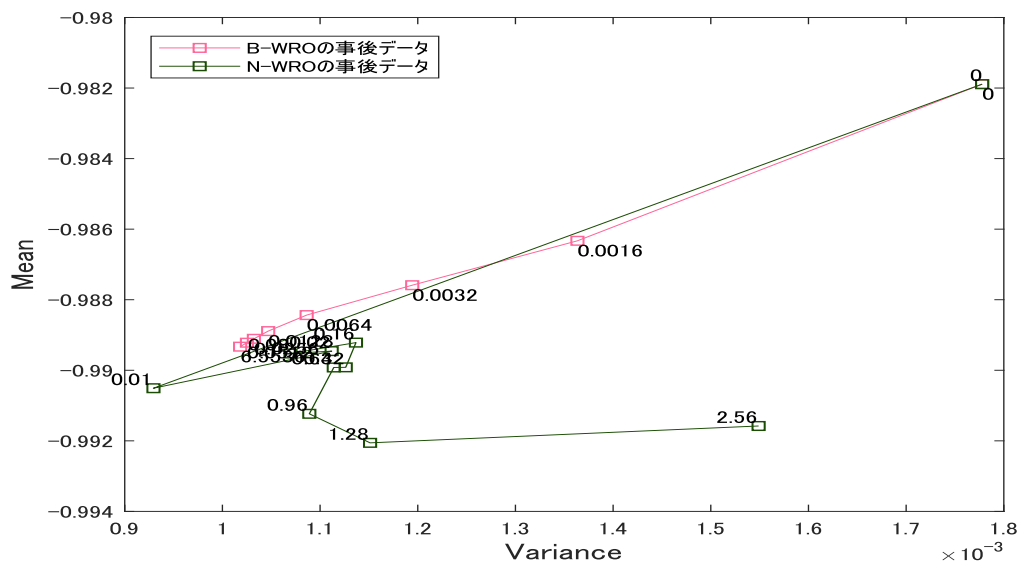


図 A.10: "10 Industry portfolio" (4) の事後データの平均分散フロンティア

