不確実を伴うチョイス・ブラインドネスにおけるポートフォリオ最適化モデルの構築

A Portfolio Optimization Model for Choice Blindness with Uncertainty

清水　豪士

Department of Information Systems Engineering,
Graduate School of Engineering
u155016@st.pu-toyama.ac.jp

12:10-12:35 Friday, December 10, 2021,
Toyama Prefectural University.

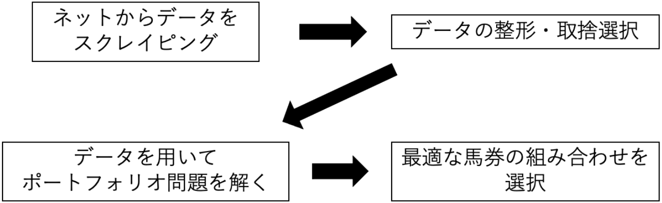
本研究の背景

- 投資・資産運用においてなぜ○○株式を選択したかといった理由が曖昧な場合，リスクがあってもそのリスクに気づかず選択してしまう場合がある．
- その結果，資産がマイナスになってもなぜマイナスに陥ったのかの部分が曖昧なままになってしまう．
- その際に，ポートフォリオ最適化モデルを用いることで，リスクを抑えてリターンを最大化する最適な資産の組み合わせを見つけることができる．
- そこで，本研究では競馬を具体的な例として取り上げ，競馬の馬券購入の組み合わせをポートフォリオ最適化モデルに当てはめることによって，最適な馬券の組み合わせを選択できるモデルを構築する．

1.2 本研究の目的

本研究の目的

1. 競走馬の走破タイムの予測をし，それをもとにレースに勝利する馬を予測する．
2. スクレイピングしてきたデータと算出したデータをもとに競馬におけるポートフォリオ最適化問題を解き，リスクを抑えてリターンを最大化できる馬券の組み合わせの選択を行う．



2.1 チョイス・ブラインドネスについて

チョイス・ブラインドネス：選択盲

- 「意思決定において，意図と結果の違いを検出できない」現象．
- 選択する理由が，実際の選択に反映されていないことに気づいていないこと．
- 自分で選択したと思いこんだモノに対して，たとえそれが間違いであったとしても理由をこじつけて正しいと判断しようとする心理的錯覚のこと．
- チョイス・ブラインドネス・パラダイムによってこの現象が存在することが発覚した．

チョイス・ブラインドネス・パラダイム

- 意思と選択の関係を操作するために，ミスディレクションや手品を用いて，意思決定者に実際には選んでいない結果を提示し，反応を引き出すという実験手法のこと．

2.2 チョイス・ブラインドネス実験

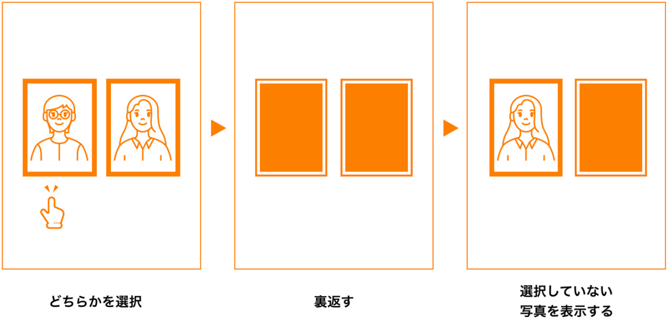


図 1: チョイス・ブラインドネス・パラダイム実験

- 自分が選択した人物写真ではない方を表示されても，多くの人がすり替えに気づかないうえに，参加者の多くは自分が選択していない写真を前に，なぜこっちを選んだかと語り，理由を後付けしていることがわかった [1].

[1] Johansson, P., Hall, L., Sikstrom, S., Olsson, A., "Failure to detect mismatches between intention and outcome in a simple decision task.", Science, 310, pp. 116-119, 205.

2.3 ヒューリスティック

ヒューリスティック

- 問題を解決する時間が極力最低限で済むように，経験則や先入観によって効率良く答えを導き出す思考法.
- 論理的に分析をして答えを出すのではなく，直感的に答えを出しているため，導き出された答えが必ず正しいとは限らない.

ヒューリスティックとチョイス・ブラインドネス

- ヒューリスティックに基づいて答えを選択する場合，直感的に答えを出しているため，なぜ選んだのといった理由の部分が明確でない場合が多い.
- そのため，チョイス・ブラインドネスの現象に陥りやすく，自分の選択に至ったプロセスを意識できない場合が考えられる.

3.1 ポートフォリオについて

ポートフォリオ

- ポートフォリオとは安全資産と危険資産の最適保有率，または自分の持っている金融資産（株，保険，不動産など）の按分率のこと．
- 似た単語でアセットアロケーションというのもあるが，どちらも資産配分の考え方という点では同じだが，配分する対象をより細分化したものをポートフォリオという．

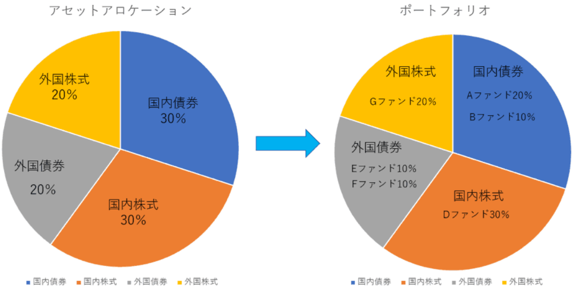


図 2: アセットアロケーションとポートフォリオ

3.2 ポートフォリオ最適化問題

ポートフォリオ最適化モデル

- リスクとリターンを考慮して，最適な資産の組み合わせを見つける問題を解くための数理モデル．
- 平均・分散モデルは「ポートフォリオの期待収益率が投資家の要求する期待収益率以上であるという制約のもとで，ポートフォリオの収益率の分散を最小化する投資比率を求める」2 次計画問題として定式化されている．

ロバストポートフォリオ最適化

- 不確実性が高まっている市場環境では，資産運用を行うことが難しいとされており，ロバストポートフォリオ最適化を利用することで，損失を抑えながら安定的な資産運用が可能になる [2]．
- 通常の平均・分散モデルと比べた場合，よりロバストなデータに対応できるため，ロバストポートフォリオ最適化は,「どの程度の範囲で起こりうる不確実性を想定するか」という情報を与えることでそれに応じて投資家の望む保守的なポートフォリオが構築できる．

[2] 山本 零， 石橋 拓弥 “不確実性下での資産運用-ロバストポートフォリオ最適化の活用-”，ファイナンシャル・プランニング研究, No. 10, pp. 71-79, 2010.

3.3 ポートフォリオ

- 市場に N 個の資産 S_i があるものとし R_i を資産 S_i の収益率を表す確率変数とする.
- 投資家は投資比率 x_i を決定するものとする.
- このときポートフォリオの収益率 $R(x)$ は次のように表すことができる.

$$R(x) = \sum_{i=1}^{N} R_i x_i$$

- ここで, r_i を資産 i の期待収益率, σ_{ij} を資産 i と資産 j の共分散としたときポートフォリオ x のリターンとリスクの指標である期待収益率と分散は

$$E[R(x)] = \sum_{i=1}^{N} r_i x_i$$

$$V[R(x)] = \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij} x_i x_j$$

3.3 ポートフォリオ

通常

$$\text{最大化} \sum_{i=1}^{N} r_i x_i$$

$$\text{条件} \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij} x_i x_j \leq \sigma_T^2$$

$$\sum_{i=1}^{N} x_i = 1$$

$$0 \leq x_i \leq u_i, i = 1, 2, ..., N$$

ロバスト

$$\underset{x}{\text{最大化}}\{ \underset{r \in U_r}{\text{最小化}} \sum_{i=1}^{N} r_i x_i \}$$

$$\text{条件} \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij} x_i x_j \leq \sigma_T^2$$

$$\sum_{i=1}^{N} x_i = 1$$

$$0 \leq x_i \leq u_i, i = 1, 2, ..., N$$

- ここで，σ_T はターゲットとするリスク上限であり，u_i は資産 i の最大投資比率を表す定数.
- ロバストの方の r は $r = (r_1, r_2, ..., r_N)^T$ を表す期待収益率ベクトルであり，U_r は期待収益率ベクトルの不確実性集合である.
- ロバストの方は通常の平均・分散モデルと違い，期待収益率を 1 つの値に特定せず，想定しうる範囲 U_r の中で最悪（最小）のケースがおこる場合を想定して，ポートフォリオの期待収益率を最大化する問題として定式化されている.

4.1 競馬におけるチョイス・ブラインドネス

オッズ

- 競馬において払い戻しの金額を知るためにはオッズを知ることが重要である．
- オッズとは，馬券が的中した場合，どのくらいの配当がつくのかを倍率で表したもの．
- 人気になればなるほどオッズは低くなっていき，払戻金が少なくなる．

競馬におけるチョイス・ブライドネス

- オッズが低いから人気な馬だといったヒューリスティック的な考えや，自分がこの馬が好きがからといった好みだけでの理由で馬券を選択した場合，もし外れた際に選択した理由が曖昧，弱いため，なぜ外れたかを曖昧な状態のまましてしまったり，人がこう言っていたからこうなんだといったように理由を後付けしてしまう．

4.2 競馬におけるポートフォリオ問題

競馬におけるポートフォリオ問題

- 競馬において理想的な立ち回りは損失を失くし，利益を最大化できる馬券を購入すること．
- 複数の馬券を購入した中で，どれか 1 つの馬券が当たった際に払い戻し金額が支払った額より小さくなる場合があり，勝つためにはこの現象を避けないといけない．
- そこで，競馬にポートフォリオ問題を適用することでプラスに持っていくための最適な馬券の組み合わせを選択できる．

従来研究

- ポートフォリオ問題における期待値を馬券が当たった時の払い戻し，分散を競走馬のオッズとして捉え，ポートフォリオ最適化モデルを構築し，期待値と分散を求めることで最適な馬券の組み合わせを選択することが可能になった[3]．

[3]秋田 貴成，“競走馬の強さの分析とポートフォリオ問題”，中央大学 卒業研究論文 2016 年, pp. 1-29, 2016

5.1 実験の概要

実験の流れ

1. 競馬サイト（netkeiba）からデータをスクレイピング.
2. スクレイピングしてきたデータを整形する.
3. 整形したデータを用いてデータ分析にかけ，競走馬の走破タイムを予測する.
4. 予測した走破タイムを馬券を買う馬を選択する.
5. 馬券を買う際に利益を最大化できる選択を行う.

競馬における特徴量

- 血統：競走馬の父馬，母父馬
- 馬の能力：体重，年齢，適正距離，今までの成績 など
- 騎手の能力：体重，今までの成績，技術 など
- 馬場状態，天気，枠順，レースグレード など

5.2 実験の結果

netkeiba のスクレイピング

netkeiba.com のスクレイピングを行う.

- natkeiba の数年分のレース結果を取得した.
- 得られる情報は着順，騎手，人気順，オッズなどの情報

レースに関してのスクレイピング

着順	枠番	馬番	馬名	性齢
1	4	7	オイデヤスダ	牝3
2	1	1	グルアーブ	牝3
3	2	3	トーレスデル	牝3
4	3	6	デュープス	牝3
5	6	12	ディヴィニテ	牝3
6	7	13	ノヘア	牝3
7	7	14	エクメディノ	牝3
8	3	5	シンキングア	牝3

競走馬に関してのスクレイピング

日付	開催	天気	R	レース名1
2021/11/7	5東京2	晴	11	アルゼンチン
2021/10/10	4阪神2	晴	11	京都大賞典(G
2021/5/15	2東京7	晴	10	緑風S(3勝ク
2021/4/25	2阪神10	晴	9	白鷺特別(2勝
2021/2/13	1東京5	晴	9	箱根特別(2勝
2020/12/12	6阪神3	曇	10	境港特別(2勝
2020/11/23	5阪神7	晴	10	猪名川特別(2
2020/10/25	4京都6	晴	9	鳴滝特別(2勝
2020/9/27	2中京7	晴	11	神戸新聞杯(G

図 3: スクレイピング結果

5.2 実験の結果

走破タイムの予測

今年のステイヤーズ S の競走馬の走破タイムの予測を行う．
競走馬についての以下の変数を用いて重回帰分析にかけた．

目的変数：タイム
説明変数：レース間隔，頭数，枠番，馬番，ターフ，距離，馬場状態，体重

結果と考察

- 予測タイムと実際のタイムの間で差が大きい．
- タイムを予測するにあたって「距離」が強く影響していることがわかった．
- 今回使用した 8 種類の説明変数では少ない．
- また，サンプル数，説明変数の種類の両方を増やす必要がある．

6. おわりに

まとめ

- 修士論文の概要の説明をした.
- 不確実性を伴うことの例として競馬を取り上げ，競馬におけるポートフォリオ問題に取り組む.
- また，現段階でできているところまでの紹介を行った.

今後の課題

- 必要な特徴量の取得と選別
- 今まで取り組んだ内容を基に今後はロバストなポートフォリオ問題最適化モデルを作成する.

ロバストポートフォリオ

ロバストポートフォリオ

- 期待収益率や共分散を各資産の過去の収益率を用いて推定する場合，市場が安定していれば妥当性があるが，近年のような不確実性の高い環境では，その妥当性が小さくなると考えられる．
- 平均・分散モデルは入力する期待収益率の値が異なると，その解も大きく変化することが知られている．
- つまり期待収益率の情報が間違っている場合には，真の最適なポートフォリオと大きくかけ離れたポートフォリオを構築していしまう可能性が高い．

$$\underset{x}{\text{最大化}}\{\underset{r\in U_r}{\text{最小化}}\sum_{i=1}^{N} r_i x_i\}$$

$$\text{条件}\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}x_i x_j \le \sigma_T^2$$

$$\sum_{i=1}^{N} x_i = 1$$

$$0 \le x_i \le u_i, i = 1, 2, ..., N$$

続き

矩形の不確実性集合

期待収益率の推定値 $\hat{r}_i$ とその許容乖離 δ_i を用いて以下で表すことができる

$$U_r = \{r \mid |r_i - \hat{r}_i| \le \delta_i,\ i = 1,2,\ldots,N\}$$

空売りを許容しない場合，期待収益率の最悪ケースは $(\hat{r}_i - \delta_i)$ となるため目的関数は以下のように変形される

$$\underset{x}{\text{最大化}} \sum_{i=1}^{N} (\hat{r}_i - \delta_i) x_i$$

矩形のロバスト最適化モデルでは，各資産の期待収益率が独立に考慮されているため，全資産の期待収益率がそれぞれ最悪になる，いわば究極の保守的ポートフォリオを作り上げてしまう．

これは通常の平均・分散モデルと比べると，モデルに入力する期待収益率を $\hat{r}_i$ から $(\hat{r}_i - \delta_i)$ に変更したものと解釈でき，各資産のリスクと投資家が想定する不確実性に基づいて，期待収益率を保守的に修正して最適化を行っている，と解釈することができる．

続き

楕円形の不確実性集合

楕円形の期待収益率は以下で表すことができる

$$U_r = \{ r \mid (r - \hat{r})^T \Sigma_r^{-1} (r - \hat{r}) \leq \delta^2 \}$$

$\sum_r$ は各資産の期待収益率の共分散行列，δ は不確実性の大きさを表す定数である．
矩形では各資産間の相関を考慮していなかったが，楕円形ではそのような極端なケースが避けられている．
楕円型の不確実性集合を採用した場合，目的関数は以下のようになる．

$$\underset{x}{\text{最大化}} \sum_{i=1}^{N} \hat{r}_i x_i - \delta \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij}^r x_i x_j}$$

σ_{ij}^T は $\sum_r$ の第 ij 要素を表す．
つまり，目的関数は期待収益率の項に不確実性の関数が追加された形になり，δ は投資家のリスク回避度を表現していると見ることができる．

続き

楕円形の不確実性集合

ここで与えるパラメータとして期待収益率の共分散 σ_{ij}^T は収益率分布に定常性を仮定した場合，収益率の共分散 σ_{ij} を用いて以下のように導出される．

$$\sigma_{ij}^r = \frac{1}{T}\sigma_{ij},\ i=1,2,\ldots,N;\ j=1,2,\ldots,N$$

よって，楕円形の目的関数は以下のように書き換えることができる．

$$\underset{x}{\text{最大化}} \sum_{i=1}^{N} \hat{r}_i x_i - \frac{\delta}{\sqrt{T}}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\sigma_{ij}x_i x_j}$$

この式の第 2 項はポートフォリオの標準偏差を表しており，$\delta/\sqrt{T}$ をリスク回避度とする効用関数型の平均・分散モデルと解釈できる．
つまりリスク回避度を投資家の許容する不確実性水準に基づいて保守的に決定して平均・分散モデルを解くことと等しいと考えられる．