

国内株式投資信託における資金フローの非線形モデリング

Nonlinear Modeling of Money Flows in Japanese Mutual Funds

吉田 遼平^{1*} 中道 拓馬¹ 田中 陸² 鈴木 智也^{1,2}
Ryohei Yoshida¹ Takuma Nakamichi¹ Riku Tanaka² Tomoya Suzuki^{1,2}

¹ 茨城大学大学院理工学研究科

¹ Graduate School of Ibaraki University

² 大和アセットマネジメント株式会社

² Daiwa Asset Management Co.Ltd.

Abstract: The purpose of this study is to visualize nonlinear relationships which are quite ambiguous in the correlation diagram by the approach of explainable artificial intelligence (XAI). By using this approach for a practical marketing problem, we could visualize the nonlinear relationship between fund performances and money flows, which is asymmetric in inflows and outflows and is consistent with investors psychology based on the behavioral economics.

1 はじめに

本稿の目的は、従来の相関分析では不明瞭な関係性を Explainable AI (以下, XAI) 技術を活用することで可視化し、マーケティング等の実務に生かすことである。その事例として、国内投資信託における資金フローの要因分析を行う。当然ながら運用会社や販売会社にとって、資金フローの流入を増やし、流出を抑える方策を見つきたい。

国内外の先行研究によれば、論点は大きく2つに大別される。まず第一に、投資信託の特徴（過去の成績、販売手数料、分配金の頻度、設定年数、etc.）と資金フローの関係を重回帰分析によって検証する方法 [1] が一般的である。近年においては、目論見書に含まれる単語のユニークさとの関係に着眼した事例 [2] もある。しかし国内においては、その解釈に論点が集中している。運用成績が良いファンドほど売られてしまうなど、直感に反する不思議な現象が幾つか観測されており、その妥当性や解釈を巡る議論が多い [3-6]。文献 [3,4] は行動経済学（プロスペクト理論）の観点から解釈しており、投資家心理による売り急ぎや損失の先送りを指摘している。その結果、利小損大の運用スタイルになり、さらに回転売買にともなう税金や再投資コストも加わり、投資家らは本来得られる投資信託のリターンを享受できない問題が起こる。一方、文献 [5] は販売会社の乗換営業による回転売買に着眼し、特に若いファンドかつ同一運用会社のファンド内で資金が出入りしやす

い傾向を指摘している。さらに資金を流入と流出に分けて重回帰分析を行い、詳細に関係性を分析した点が新しい。しかし分析期間が2010年1月～2014年12月迄と古いため、本研究では2019年12月迄に延長して分析を行う。

その動機として、2014年7月に金融庁の年次報告書 [7] により実際に販売会社の乗換営業が多い事実が示され、文献 [6] が指摘するように投資家（顧客）らの利益が最優先されないプリンシパル・エイジェント問題が発生していた。これを受けて金融業界は、販売額重視から預かり資産残高重視への転換が求められており [8]、本稿で提案する可視化手法はこの転換を常時モニタリングするツールとして活用できる可能性がある。

その手法として、文献 [5] と同様に資金を流入と流出に分けつつも、機械学習（主に LightGBM [9]）を導入することにより、従来の重回帰分析では対応できない非線形な関係性も検出可能にする。しかし一方でモデルの解釈性が低下するが、SHAP (Sharpley Additive exPlanation) [10] を導入することで、むしろ効果的な可視化表現ができる可能性がある。

2 分析データ

投資信託には様々なタイプがあるが、本研究では最も標準的な追加型国内株式投信（インデックス型やブルベア型などを除く）670ファンドを分析対象にする。分析期間は2010年1月～2019年12月とし、分析に用いる月足データは株式会社 NTT データエービック社の FundMonitor データベースより取得した。

*連絡先： 茨城大学大学院理工学研究科機械システム工学専攻
〒316-8511 茨城県日立市中成沢町 4-12-1
E-mail: 20nm505n@vc.ibaraki.ac.jp

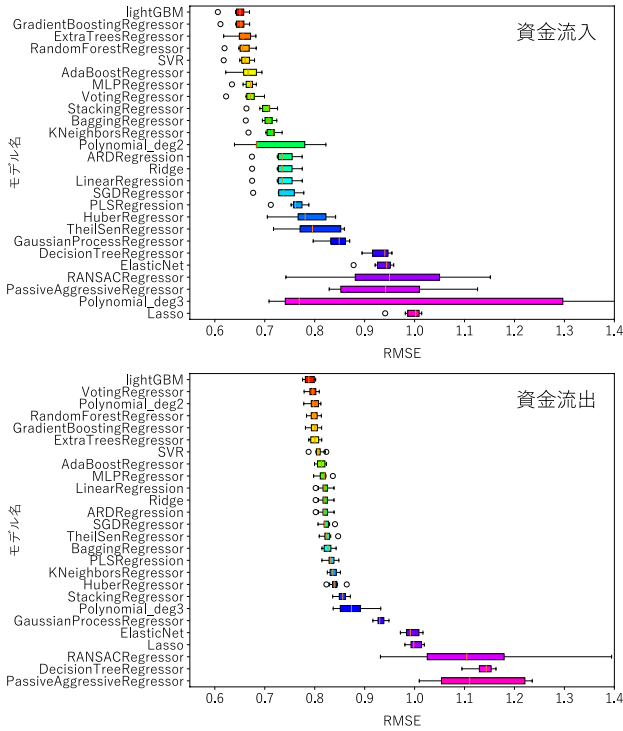


図 3: 代表的な機械学習モデル (24 種) による RMSE

標準化を施し、各モデルのハイパーパラメータはデフォルト値を用いる。学習用データは全期間からランダムに抽出した 70% を使用し、残りの 30% を評価用データとする。学習に用いる乱数のシードを変えながら学習と評価を 10 回繰り返し、それぞれの RMSE (Root Mean Square Error) を図 3 に示す。

結果として、資金流入および資金流出のどちらにおいても LightGBM が最良であり、高い汎用性を有している。従来より先行研究で用いられてきた重回帰モデルは LinearRegression に相当するが、LightGBM 等の決定木ベースの機械学習による誤差が小さい点を鑑みると、式 (1)(2) の関数 F は非線形的なダイナミクスを有するものと示唆される。そこで次節では SHAP [10] を用いて、図 2 では確認できない非線形ダイナミクスの可視化を試みる。

3.2 SHAP による可視化

SHAP は協力ゲーム理論のシャープレイ値 (Shapley Value) の考え方により、目的変数に対する各説明変数の貢献度を求める手法である。式 (1)(2) のパネルデータに関するデータ番号を k とすると、説明変数 $x_j(k)$ の SHAP 値 $\phi_j(k)$ は以下のように算出される。

$$\phi_j(k) = \sum_{S \subseteq \Omega_j} \frac{|S|!(|\Omega| - |S| - 1)!}{|\Omega|!} [\hat{F}(S \cup x_j(k)) - \hat{F}(S)] \quad (3)$$

ここで Ω はデータ番号 k の全説明変数の集合であり、 Ω_j は Ω から $x_j(k)$ を除去した集合である。さらに S は Ω_j の部分集合であり、 \hat{F} は機械学習によって推定された関数 F である。つまり $\phi_j(k)$ は、 \hat{F} に説明変数 $x_j(k)$ を追加した時の推定値の差分 (貢献度) を評価しており、全ての組合せについて平均化したものである。

さらに、説明変数 x_j の重要度 I_j およびポジネガ極性 p_j について、以下のように算出する。

$$I_j = \frac{1}{K} \sum_{k=1}^K |\phi_j(k)| \quad (4)$$

$$p_j = \frac{\sigma_{\phi_j x_j}}{\sigma_{\phi_j} \sigma_{x_j}} \quad (5)$$

ここで K はパネルデータの総数であり、 p_j は $\{\phi_j(k) | k = 1 \sim K\}$ と $\{x_j(k) | k = 1 \sim K\}$ の相関係数である。なお相関係数では線形の関係性しか評価できないが、 ϕ_j と x_j の相関図をプロットすることで、目的変数との非線形関係を可視化できる可能性がある。

3.3 実験

上記の有用性を検証すべく、2 章で示した実データに対して適用する。なお SHAP 分析においては解釈の都合上、各変数に標準化を施さない。機械学習モデルとして LightGBM を用いるが、本実験では Optuna [11] を用いた交差検証法によりハイパーパラメータを最適化した。なお評価指標は前述と同様に RMSE とした。

図 4 に、各説明変数 x_j の重要度 I_j とポジネガ極性 p_j を示す。なおコントロール変数として用いた $y^{\text{in}}(t-1)$ と $y^{\text{out}}(t-1)$ は非表示とする¹。重要度の結果として、資金流入および資金流出において設定年数やファンド収益に関する説明変数が上位を占めている。さらにポジネガ極性を参照すると、先行研究 [3-6] で議論されてきたように、ファンド収益の向上は資金流入を促進するものの、資金流出にも顕著なポジティブ極性を持っている。また設定年数は資金流入および資金流出に対してネガティブ極性を持っており、やはり古いファンドほど資金フローが減退している。

これらの傾向をより詳細に分析すべく、 ϕ_j と x_j の相関図により資金フローに与える影響を可視化する。なお議論に重要な変数として、対数設定年数、年間収益率、 α 値に着眼する。結果を図 5 に示す。

可視化を通じて初めて認識できる新事実を列挙する。

- 設定年数 4 ~ 5 年 (対数目盛で 1.5) で不連続な相転移が起っている。資金流入の方が謙虚であるが、資金流出でも確認できる。アクティブファン

¹ 説明変数と正の自己相関を有するため、高い重要度およびポジティブ極性を持つことは自明である。

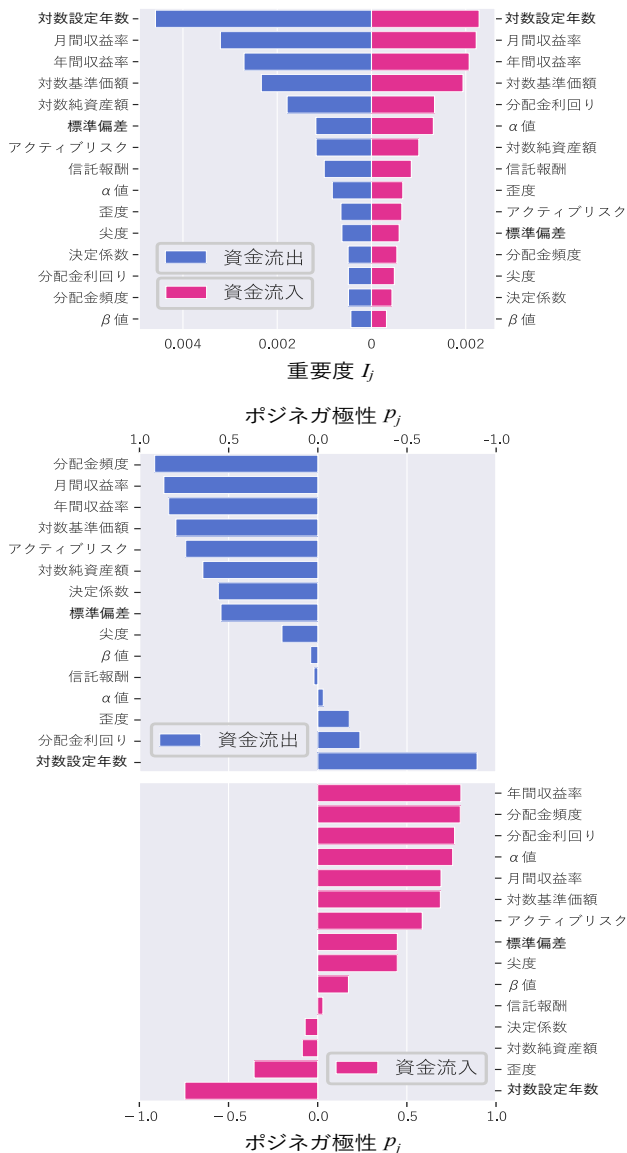


図 4: 説明変数の重要度 I_j とポジネガ極性 p_j

ドにおいては償還期限を 4,5 年に設定するケースが多く、この期限を境に相が転移すると考えられる。たとえ償還期限を延長しても、前述と同様に古いファンドの相では資金流入出は少ない。

- 設定年数 2 年 (対数目盛で 0.8) 付近にも不連続な変化点がある。設定間も無いファンドほど資金流入も流出も突出しており、投資家らはファンドの実力を見極めるには短い期間で売買行動をしている。これは投資行動としては非合理的であり、行動経済学的な短期的な利益確定心理 [3,4] や販売会社による乗換営業 [5] の可能性を支持している。
- 年間収益率は資金流入と流出でほぼ同じ関係性を示している。つまりファンドの成績が良いほど資

金流入もするが流出も促進され、先行研究が指摘している通りである。なお 0.2 付近で不連続な変化点を確認でき、売買コスト (信託報酬や購入手数料) による損益分岐点だと考えられる。この損益分岐点を境に $\phi_j(k)$ の符号が反転しており、損益分岐点以下では損失回避心理により売却 (資金流出) が抑制される。

- 年間収益率が 1.0 (年間で基準価額が倍) 以上は特殊なためか、有意義な結論を見出せない。
- 本研究は主にアクティブファンドを分析対象としているため、 α 値の方が運用能力の評価指標として合理性が高い。そのためか、年間収益率とは全く異なる非線形ダイナミクスを確認できる。しかも資金流入と流出で非対称的である。
- α 値が高いほど資金が流入しており、この傾向は年間収益率より明確である。アクティブファンドの運用努力は高い α の実現であり、資金流入の増加によって報われている。
- $\alpha > 0$ でも不合理な資金流出は発生するが、年間収益率よりも顕著ではない。
- これまでの議論と明確な違いとして、 $\alpha < 0$ では資金流出が促進されている。つまり損失回避心理は働かず、投資家らは α 値に敏感に反応している。やはりアクティブファンドの商品価値は $\alpha > 0$ の実現にあり、市場リスクに対する要求リターン (株主資本コスト) に劣後することは許されない。
- 特に V 字の関係は相関係数が 0 になるため、可視化による分析は有意義である。

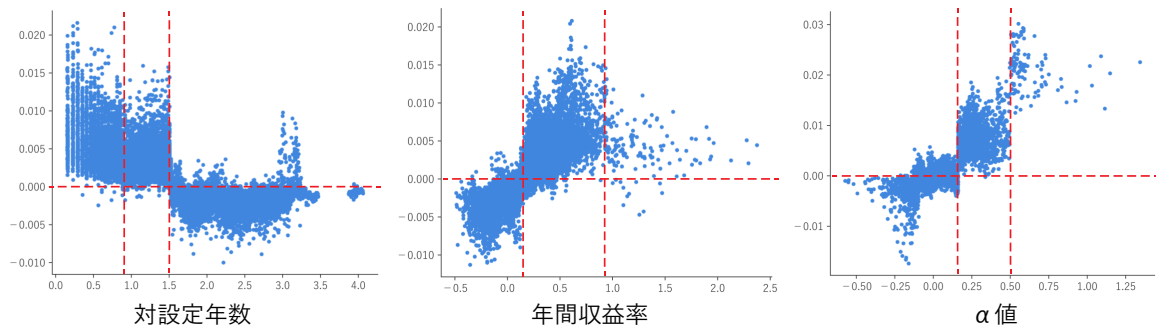
4 資金フロー構造の時代変化

1 章で述べたように、2014 年 7 月の金融庁報告書を契機に金融業界はビジネスモデルの転換が求められている。また投資家心理も時代に応じて変化している可能性がある。そこで本手法を用いて、2014 年前後のメカニズムの変化について分析する。なお機械学習にとって十分な学習サンプル数 K を確保すべく、5 年間で 1 期間として 1 年毎にシフトしながら変化を分析する。具体的な期間を表 1 に示す。

重要度 I_j を図 6 に、ポジネガ極性 p_j を図 7 に示す。主に時代変化が大きい変数を太線で示す。得られる知見は以下の通りである。

- 資金流入においては、年間収益率と分配金利回りの重要性が上昇している。それぞれポジティブ極

(a) 資金流入 $y^{\text{in}}(k)$ に対する SHAP 値 $\phi_j(k)$



(b) 資金流出 $y^{\text{out}}(k)$ に対する SHAP 値 $\phi_j(k)$

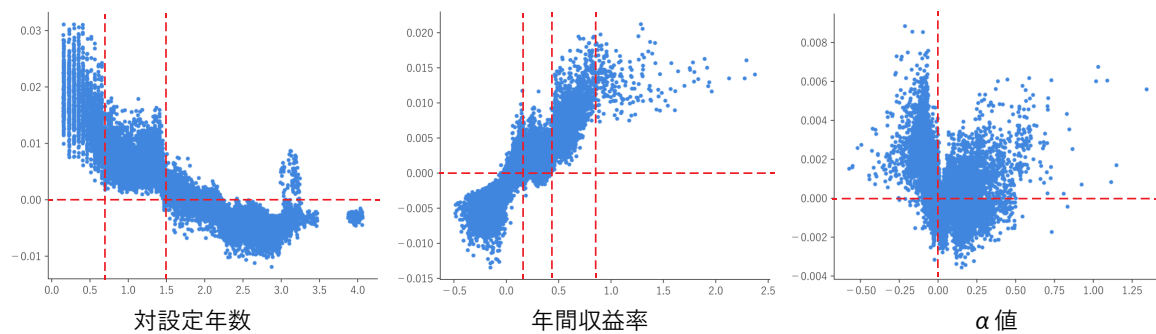


図 5: SHAP による非線形性の可視化

表 1: 5 年毎の分析期間

No.	期間
term 1	2010 年 1 月～2014 年 12 月
term 2	2011 年 1 月～2015 年 12 月
term 3	2012 年 1 月～2016 年 12 月
term 4	2013 年 1 月～2017 年 12 月
term 5	2014 年 1 月～2018 年 12 月
term 6	2015 年 1 月～2019 年 12 月

性を持つため、運用成績が良く分配金を多く支給できるファンドが評価されつつある。同様に、資金流出における分配金利回りがネガティブ極性になっており、資金流出を抑制している。

- しかし資金流出においては、年間収益率がポジティブ極性のまま重要性が上昇している。つまり短期的な利益確定の売却が顕著になりつつある。
- 対数設定年数については、顕著な変化は見受けられない。つまり古いファンドの資金フローは乏しく、若いファンドほど資金流入出が活発になる性質は不変である。

- 重要度は小さいが、信託報酬のポジネガ極性は弱くなり、資金流出において決定係数がポジティブ極性になっている。つまりアクティブファンドにも関わらず市場インデックスと相関が高いファンドは敬遠される傾向にある。

次に、対数設定年数における不連続変化を図 8 に、 α 値における非線形性を図 9 に示す。認識できる時代変化は以下の通りである。

- 若いファンドに資金流入する傾向は、近年において更に顕著になっている。2015 年以降は AI 技術の台頭により技術革新関連ファンドが急増しており、運用成績を見極める前に投資テーマで購入される傾向が高まっている可能性がある。
- 逆に、若いファンドから資金流出する傾向は、近年において弱まっている。設定間も無く解約することは購入手数料を無駄にするため、投資家にとって望ましくない。
- しかし若いファンドほど資金が回転しやすい傾向は不変である。
- 一方 $\alpha > 0$ において、近年では不合理な資金流出は徐々に縮小しており、さらに資金流入は徐々に

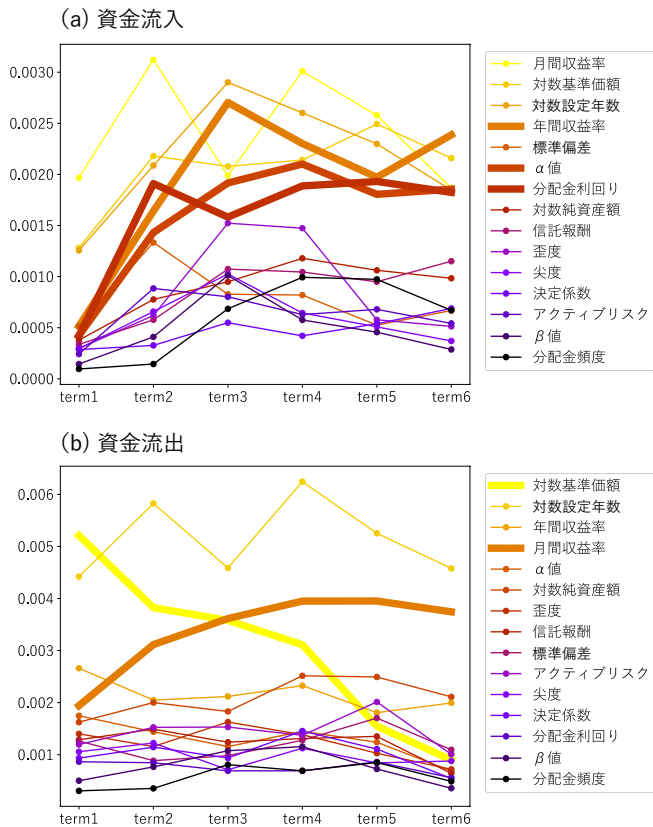


図 6: 重要度 I_j の時代変化

に拡大している。ファンドの運用努力が報われる健全な状態になりつつある。

- しかし $\alpha < 0$ における敏感な資金流出は不変であり、アクティブファンドにとって α 値は重要な評価指標である。

以上の考察は、投資家らの短期的な利益確定心理もしくは販売会社による乗換営業によるものか判別は難しいが、最終的には投資家自身が売買判断をしている。短期的な回転売買は中長期的な経済成長および時間分散効果を享受しやすく売買コストも増える。資産残高重視へのビジネス転換にとって、投資家らの金融リテラシーを高める取り組みも必要だと思われる。

5 まとめ

非線形モデルとして LightGBM を導入することで当てはめ精度を向上できることを確認し、さらに SHAP を導入することで従来の相関分析では不明瞭だった変数間の関係性を目視できるように表現力を高めた。その結果、投資信託への資金流入と資金流出には異なる

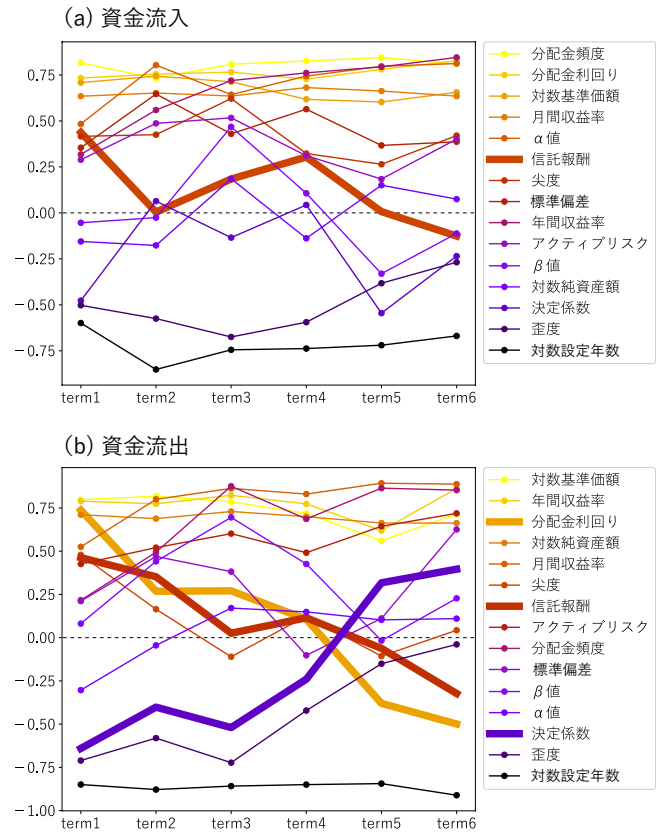


図 7: ポジネガ極性 p_j の時代変化

非線形ダイナミクスが内在し、時代と共に少しずつ変化している様子を確認した。

今後の課題として、SHAP はモデル出力を構成する各説明変数の貢献度を定量化したものに過ぎず、変数間の因果関係を決定づけるものではない。そこで本稿による可視化は仮説構築など前処理的な定性分析として活用し、より厳密には因果推論等の定量分析を併用する必要がある。その際には従来の線形重回帰分析による t 検定も強力な分析ツールとなる。

本稿の内容は筆者個人の見解であり、所属組織の公式見解ではありません。本研究は JSPS 科研費 (20K11969) の助成により行われました。

参考文献

- [1] L. Davidson and T. Strauts: “What factors drive investment flows?” *Morningstar Quantitative Research*, 33 pages, 2015.
- [2] L. Kostovetsky and J. B. Waner: “Measuring innovation and product differentiation: Evidence from mutual funds,” *The Journal of Finance*, 75(2), 779–823, 2020.

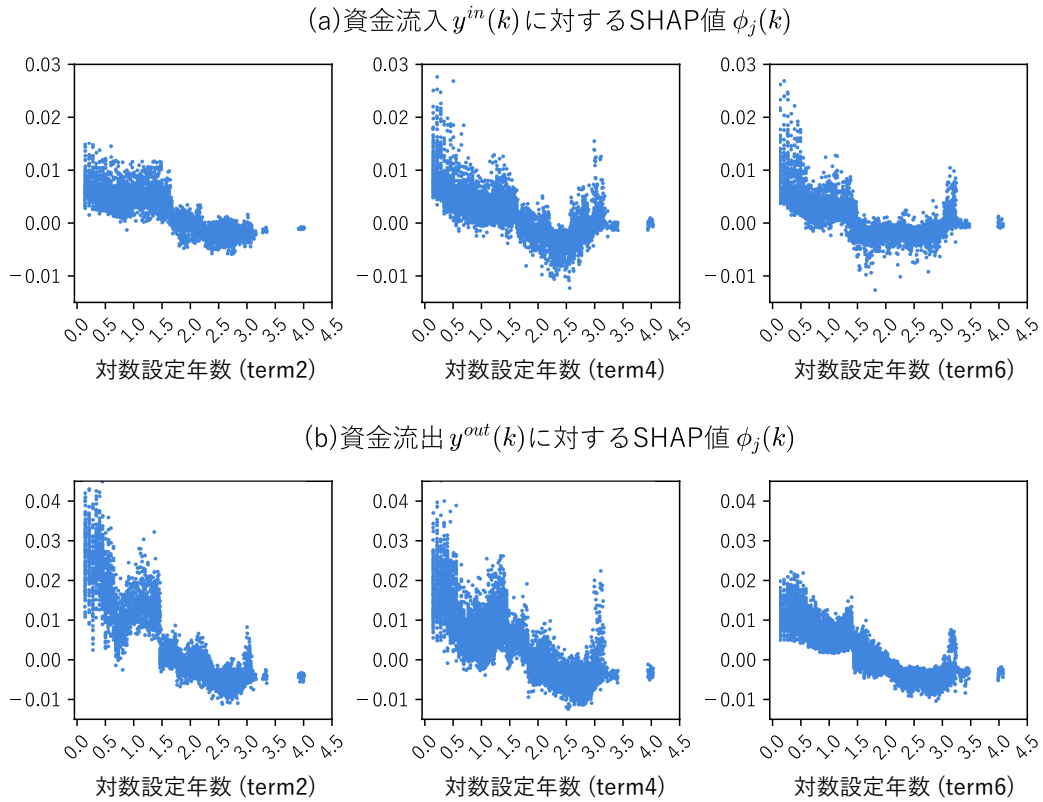


図 8: SHAP による不連続変化 (対数設定年数) の可視化

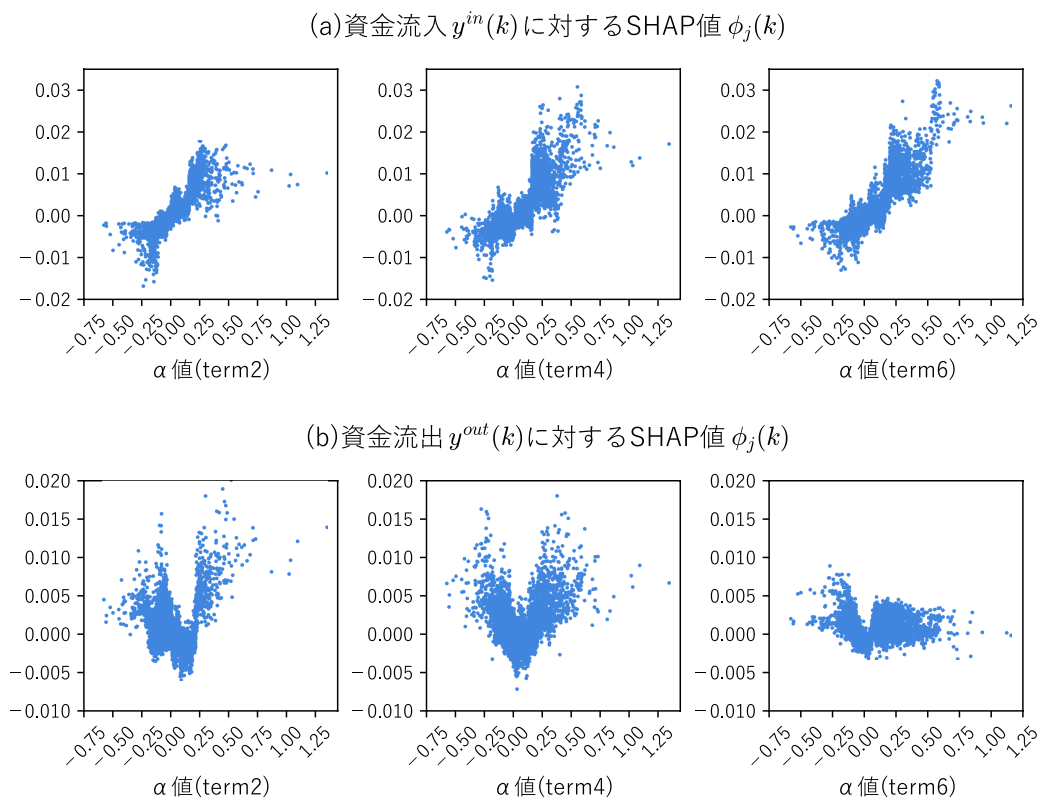


図 9: SHAP による非線形性 (α 値) の可視化

- [3] 金子久, “個人投資家の投資行動と普及への展望,” 証券アナリストジャーナル, vol.41, no.7, pp.18-31, 2003.
- [4] 半田拓, “個人投資家から見た投資信託の投資家リターン分析—投資家に真のリターンをもたらす投資行動を求めて—,” 証券アナリストジャーナル, vol.44, no.8, pp.71-81, 2006.
- [5] 阿萬弘行, 高橋陽二, “日本における投資信託の資金フローと市場構造,” 証券アナリストジャーナル, vol.58, no.5, pp.76-87, 2020.
- [6] 山本健, 米澤泰博, 花村泰廣, “わが国の株式投資信託の Flow-Performance 関係: 序説,” 証券経済研究, vol.46, pp.1-15, 2004.
- [7] 金融庁, “金融モニタリングレポート,” <https://www.fsa.go.jp/news/26/20140704-5/01.pdf> 2014.
- [8] 日本経済新聞, “投信乗り換え販売増. 金融庁報告書見直し促す,” https://www.nikkei.com/article/DGKDASDF04H0G_U4A700C1EE8000/ 2014.
- [9] G. Ke, et al., “LightGBM: A highly efficient gradient boosting decision tree,” *NIPS2017*, pp.3146-3154, 2017.
- [10] Scott M. Lundberg and Su-In Lee, “A unified approach to interpreting model predictions,” *NIPS2017*, pp.4768-4777, 2017.
- [11] T. Akiba, et al., “Optuna: A next-generation hyperparameter optimization framework,” *KDD 2019*, pp.2623-2631, 2019.
- ・年間収益率
直近 12ヶ月の各月間収益率に 1 を加えた数値を掛け合わせ、その後に 1 を引いて算出する.
 - ・標準偏差, 歪度, 尖度
それぞれ直近 12ヶ月の月間収益率より算出する.
 - ・ β 値, α 値
各ファンドの月間収益率を, 市場インデックスの月間収益率で単回帰して推計する. その単回帰係数が β 値, 切片が α 値である.
 - ・決定係数
各ファンドの月間収益率を, 市場インデックスの月間収益率で単回帰して推計する. 結果として, 両者の共分散の 2 乗をそれぞれの分散で除したものに等しい. この値が高いほどファンドの収益率を市場インデックスの収益率で説明できる割合が高く, パッシブ運用に近いことを示す.
 - ・アクティブリスク
直近 12ヶ月の月間アクティブリターン (各ファンドの月間収益率 - 市場インデックスの月間収益率) の標準偏差を算出し, それに $\sqrt{12}$ を掛けて年率に換算する.
 - ・分配金利回り
直近 12ヶ月の分配金の合計金額を, 当月基準価額で除して算出する.

A Appendix

本稿に用いる各変数の算出方法を示す. それぞれ株式会社 NTT データエービック社 FundMonitor に準拠する. 同情報端末では下記以外の指標も提供しているが, 互いに相関が高いものを含むため, 本研究では比較的相関が小さく, ファンド評価において代表的な指標に厳選した.

・月間収益率

月中に分配金の支払いがあった場合には, その時点で再投資したものとして収益率を算出する.

月間収益率

$$= \frac{\text{当月末基準価額} + \text{分配金} \cdot \frac{\text{当月末基準価額}}{\text{月中分配基準価額}} - \text{前月末基準価額}}{\text{前月末基準価額}}$$