金融情報学研究会(第27回)

日時　2021 年 10 月 9 日(土)

会場　オンライン

SIG-FiN
JSAI Special Interest Group on Financial Informatics

人工知能学会
金融情報学研究会

第27回 人工知能学会 金融情報学研究会(SIG-FIN)

2021 年 10 月 9 日(土) オンライン

01. 文章極性を考慮したニュースステキスト分析による経済動向予測 …… 1
川崎拓海，穴田一(東京都市大学)

02. 金融文書を用いた事前学習言語モデルの構築と検証 …… 5
鈴木雅弘，坂地泰紀，平野正徳，和泉潔(東京大学)

03. 機械学習を用いた統合報告書のESG関連ページの推定 …… 11
河村康平(成蹊大学)，高野海斗(野村アセットマネジメント)，酒井浩之，永並健吾(成蹊大学)，
中川慧(野村アセットマネジメント)

04. PDF 形式の金融文書における項目-数値間関係を考慮したテーブル情報抽出 …… 19
青野有華，市川幸史，近藤浩史，加藤淳也(日本総合研究所)

05. 有価証券報告書におけるリスク階層構造分析 …… 26
藤井元雅(東京理科大学)，坂地泰紀，佐々木一(東京大学)，増山繁(東京理科大学)

06. 特許ドメイン特化型 BERT による脱炭素関連特許技術の「見える化」 …… 32
前原義明，久々宇篤志，長部喜幸(日本特許情報機構)

07. アナリストレポートから生成した極性指標を用いた市場動向分析 …… 36
田口怜，渡邊光，平野正徳，鈴木雅弘，坂地泰紀，和泉潔(東京大学)，
平松賢士(アイフィスジャパン)

08. 因果情報を用いた経済数値予測 …… 44
和泉潔，坂地泰紀，佐野仁美(東京大学)

09. 確率的にホールドの長さが決まる人工市場モデル …… 50
片平啓(筑波大学)，陳昱(東京大学)，秋山英三(筑波大学)

10. 人工市場を用いたショートサイドの市場非効率性に関する分析 …… 56
則武誉人，逸見龍太(三井住友 DS アセットマネジメント)，
水田孝信(スパークスアセットマネジメント)

11. 財市場を内包した Agent-based Computational Economics の構築に関する一提案 …… 64
高島幸成(長岡大学)，八木勲(工学院大学)

12. 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 …… 71
吉田遼平，中道拓馬(茨城大学)，田中陸(大和アセットマネジメント)，
鈴木智也(茨城大学，大和アセットマネジメント)

13. 外国為替証拠金取引における解釈可能な売買戦略の構築 …… 79
内田純平，穴田一(東京都市大学)

14. 機械学習による為替フォワード取引期間の判別モデル …… 87
雉子波晶(茨城大学)，杉本誠忠(ワイジェイ FX)，酒本隆太(岡山大学)，
鈴木智也(茨城大学，大和アセットマネジメント)

15. Thompson Samplingを用いた複数ポートフォリオの合成戦略 …… 94
藤島圭吾(東京大学)，中川慧(野村アセットマネジメント)

16. 国債市場情報を用いた機械学習に基づく経済予測モデルの構築 …… 101
水門善之(東京大学，野村証券)，和泉潔，坂地泰紀(東京大学)

文章極性を考慮したニュースステキスト分析による経済動向予測

Economic Trend Prediction by News Analysis Considering Text Polarity

川崎　拓海 [1*]　　穴田　一 [1]
Takumi Kawasaki　　Hajime Anada

[1] 東京都市大学大学院
[1] Tokyo City University Graduate School

Abstract: In recent years, various researches in the field of economic prediction have been carried out using fundamental analysis and technical analysis with numerical information. Considering news articles containing not only numerical information but also textual information means that we can pay attention to public opinion, and thus we can make more accurate economic trend prediction which are difficult to predict only by numerical information. In this study, we propose a news text analysis for economic trend prediction using polarity dictionaries.

はじめに

近年，金融予測の分野ではローソク足の画像を用いた分析やファンダメンタル分析，数値情報を用いたテクニカル分析などによる様々な研究が行われている．その中でも数値情報だけでなくテキスト情報も含まれているニュース記事を考慮することは，世論に目を向けることを意味し，数値情報だけでは説明が難しい市場の予測を精度高く行える可能性があると考えられる．そこで本研究では，テキストマイニング手法を用いてニュース記事から株価の上昇・下落の予測を行った．テキストマイニング手法を用いた金融予測についても様々な研究が行われているが，本研究では，新聞記事の予測前営業日と予測当日のテキストを用いて株価の上昇下落を予測した和泉らの研究[1]を基に，日本語評価極性辞書[2][3]と金融に関係する単語を分析する金融専門極性辞書[4]を用いたニュースステキスト分析による東証株価指数(TOPIX)の株価予測を提案し，その有効性を確認した．本研究では予測前営業日の見出しのテキストデータをまとめ，文中の否定文を考慮した極性単語を抽出し，一定割合以上出現した単語の中で株価上昇確率を共に満たすものを特徴語とする．そしてテキストにその特徴語が出現した際，TOPIXの株価は上昇するか否かを勾配ブースティング決定木(GBDT)に学習させて，その有効性を検証した．

*連絡先：東京都市大学大学院　総合理工学研究科
〒158-8557 東京都世田谷区玉堤１丁目２８番１号
E-mail: g2181429@tcu.ac.jp

既存研究

テキストの時系列出現パターン

従来のテキスト分類を用いた市場予測では，テキストの時系列性に着目し，直近m個のテキストの特徴ベクトルから二値分類した値 y_{t+1}を求める．

$$y_{t+1} = f(x_{t,\dots,}x_{t-m-1})$$

ここで，fは手法を表し，x_tは時刻 tにおけるテキストの特徴語ベクトルを表す．

テキストの時系列出現パターン

和泉ら[1]は新聞記事の予測前営業日x_{t-1}と予測当日x_tのテキストで，単語の出現パターンを作成した．予測前営業日のテキストx_{t-1}では出現していないが予測当日x_tでは出現している場合 ”新出”．予測前営業日のテキストx_{t-1}に出現している，かつ予測当日のテキストx_tにも出現している場合 ”続出”．予測前営業日のテキストx_{t-1}には出現してい

るが予測当日のテキストx_tには出現していない場合 ”消滅”と定義した.

特徴語の抽出

和泉らは，日本経済新聞の予測前営業日と予測当日の記事のリード(第一段落)と見出しを結合して Mecab を用いて形態素解析を行い，TeamExtract で専門用語を抽出し，特徴語とした. TeamExtract は形態素解析で分割された専門用語を再度組み合わせ，専門用語として抽出するものである．これを訓練期間内に出現した記事のテキストデータに用いた．出現パターンを考慮した専門用語の出現数を調べ，k回以上出現したものの中から，テキストに出現パターンを考慮した単語が出てきた時，株価が上昇した確率が θ以上のものと$1-\theta$以下のもの $(\theta > 0.55)$を取り出した.

SVM を用いた株価予測

和泉らの研究では抽出した特徴語で株価の上昇・下落を予測するために SVM を用いた. SVM とは互いに一番近いベクトルの距離を最大化することで未知データを 2 クラスのどちらかに分類する手法である．既存研究では単語の特徴量が多いので，カーネルトリック法という非線形分離型の分類器を用いて実験を行っている.

抽出した l個の特徴語の出現パターンを $p_1, \dots, p_l$とし，訓練期間内のテキストに出現パターン p_iの単語 iが生じている場合，i次元の特徴量を1そうでない時は0とした. 出力を当日の株価の利益率が 0または正のとき1，負の場合は−1とし， 作られた l次元の専門用語に関する特徴ベクトルと株価の出力の関係を SVM に学習させた.

提案手法

和泉らの研究での全体の平均正解率は 71.4%であるが，悪い年は 56.3%と不安定である．これは 単語の出現数や出現パターンのみ考慮していて，単語の印象を考慮していないことが要因であると考えられる．なぜなら人に良い印象を与える単語が出現すると株価が上昇し，人に悪い印象を与える単語が出現すると株価が下落すると考えたからである．そこで提案手法では肯定文中極性単語と否定文中極性単語を考慮した特徴量抽出を行った.

今回極性単語を扱うにあたり，日本語評価極性辞書と金融専門極性辞書を利用した．日本語評価極性辞書とは様々な用法や名詞に対し, Negative・Positive と極性が振られている辞書である．また，金融極性辞書とは金融専門単語についてネガティブ・ポジティブ度を極性値として表した辞書であり，−1以上 1 以下の数値データで表されている.

この２つの極性辞書から感情分析ツールの１つである Oseti を用いて否定文を考慮した極性単語の抽出を行った. Oseti とは形態素解析ツール Mecab を用いて文章極性スコアを算出するものである．単語に”せず”や”ない”等の否定が掛かっている場合，その単語の極性を反転させスコアを求める．よって Positive(Negative)の極性単語に否定が掛かっている場合 Negative(Positive)とし，肯定文中極性単語と否定文中極性単語をそれぞれ抽出した。

本研究では IT・経済ニュースの記事に対して２つの辞書から得られる極性単語を用いたネガティブ・ポジティブ分析(以下ネガポジ分析とする)による経済動向予測を提案する．まず訓練データ内において 1 日に数件ずつ掲載されている IT・経済ニュースの見出しから, Oseti を用いて否定文中か肯定文中かを考慮した極性単語を抽出した. Oseti に用いる金融専門極性辞書の極性値ηは$\eta > \eta_{th}$ の場合 Positive, $\eta < -\eta_{th}$ の場合 Negative と分類された極性単語を抽出した．得られた極性辞書の単語がk回以上出現した中から株価上昇割合 θ_1以上， 株価下落割合 θ_2以上の単語を取り出し特徴語とした.

取り出されたl個の特徴語に対し，訓練期間内のテキストに特徴語が生じている場合，特徴量を 1，存在しない特徴語に関しては 0 とし，勾配ブースティング決定木に学習させた.

結果

提案手法の有効性を確認するためロイターニュース IT・経済ニュースの見出しを用いて，予測対象を半年ごとに分けた 2018 年 7 月~2020 年 12 月までの TOPIX-連動型上昇投資信託(ETF)とし,予測前営業日のニュースの見出しで上昇下落の予測を行った．訓練データの期間は直近の過去 3 年間を用いた．また，予測前日の終値と予測対象日の終値の差分を TOPIX -ETF の上昇・下落の基準とした.

極性値の閾値をη_{th} = 0.03 とし， 得られた極性単語で 10 回以上出現した単語の中から予測当日の株価の上昇割合 η が 0. 75 以上と株価の下落割合 η が 0. 65 以上のパターンを抽出し，特徴語として用いた．モデルのパラメータはグリットサーチを行い,

最適なパラメータを選択した.

予測結果は表 1 の混同行列を用いて評価する.

表 1　混同行列の例

実際のクラス	Negative	TN(True Negative)	FN(False Positive)
	Positive	FP(False Negative)	TP(True Positive)
		Negative	Positive
		機械学習モデルの予測	

True は予測が正しく False は予測が正解のクラスと異なったことを表す. 表 1 を元に Accuracy(正解率) や Precision (適合率) , Recall (再現率) を求め, グラフ化した結果を図 1 に示し, F 値も求め, それぞれの結果を表 2 に示した.

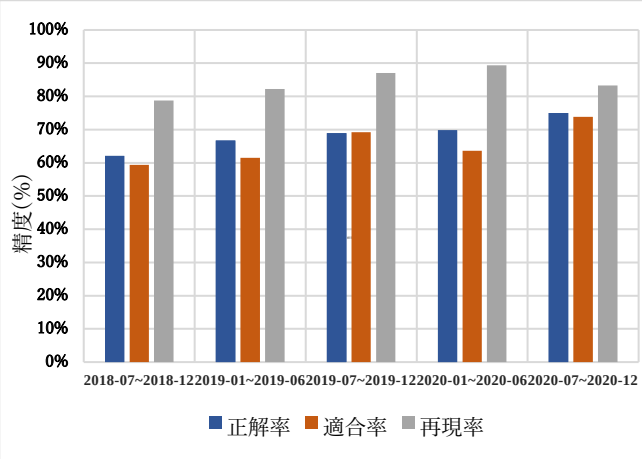


図 1　混同行列を用いた結果

表 2　混同行列を用いた結果

テスト期間	F 値	正解率	適合率	再現率
2018-07~2018-12	**0.68**	**62.1%**	**59.4%**	**78.8%**
2019-01~2019-06	**0.67**	**66.7%**	**61.5%**	**82.2%**
2019-07~2019-12	**0.77**	**68.9%**	**69.2%**	**87.1%**
2020-01~2020-06	**0.74**	**69.8%**	**63.6%**	**89.4%**
2020-07~2020-12	**0.78**	**75.0%**	**73.8%**	**83.3%**
全体の平均	**0.73**	**67.5%**	**64.4%**	**84.2%**

表 2 より, 先行研究の全体の正解率が約 70%に対し, 提案手法の精度は低く, 全体の正解率が 67.5%という結果となった. しかし, 悪い年の正解率でも約 62%と既存手法より安定した結果を得ることができた. 実際に上昇すると予測するデータのうち, 上昇すると予測できた割合を示す再現率の平均が約 84%と高い精度で予測できていた.

また, 株価の急落が起こった 2020 年 1 月~6 月の予測と 2020 年 7 月以降の期間においての予測が精度高く出来ていた. 金融業界では正解率が常に 55%以上あれば有用であると評価されていて, 今回それを越えることができていた.

特徴語として得られた極性単語は発表で述べる.

考察

2018 年 7 月~2019 年 6 月の精度が低かったが, これは米国の業績悪化の影響で起こった大幅な急落が主な原因で, 訓練期間のデータではうまく学習できなかったと考えられる. この期間の精度は, 異なるニュース記事で予測を行った場合にも同じように悪くなった.逆に, 急落・急騰の激しい時期であった 2020 年 1 月から 2020 年 12 月にかけての正解率が, ほかの年と比べ上がっていた. これは直近 3 年間を訓練データとしているので急落が激しかった 2017 年 1 月~2019 年 12 月にかけての暴落をうまく学習したからだと考えられる.

今後の課題

前日と予測当日の終値の差が 1 円以上あれば上昇・下落と株価の変化がほとんど無い横ばいの変動時でも二値で分類してしまうため, うまく分析できていない. 今後は一定の閾値を設けて上昇・横ばい・下落の多値分類を行うことで精度を高めていきたい.

また, 予測前営業日のみで当日の株価予測を行っているため時系列性を考慮できていない. よって予測前営業日だけでなく, 2 日間や 3 日間の出現パターンを考慮して予測行っていきたい.

参考文献

[1] 和泉潔, 松井藤五郎 : 新聞記事の時系列テキスト分析による株式市場の動向予測, 第 30 回人工知能学会, 3L3-OS-16a-6 (2016).

[2] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一, 意見抽出のための評価表現の収集, 自然言語処理, Vol.12, No.3, pp.203-222, 2005. / Nozomi Kobayashi, Kentaro Inui, Yuji Matsumoto, Kenji Tateishi. Collecting Evaluative Expressions for Opinion Extraction, Journal of Natural Language Processing 12(3), 203-222 (2005)

[3] 東山昌彦, 乾健太郎, 松本裕治, 述語の選択選好性に着目した名刺評価極性の獲得, 言語処理学会第 14 回年次大会論文集, pp.584-587, 2008. / Masahiko Higashiyama, Kentaro Inui, Yuji Matsumoto. Learning Sentiment of Nouns from Selectional Preferences of Verbs and Adjectives, Proceedings of the 14th Annual Meeting of

the Association for Natural Language Processing, pp.584-587 (2008)

[4] Ito T., Sakaji H., Tsubouchi K., Izumi K., Yamashita T. Text-Visualizing Neural Network Model: Understanding Online Financial Textual Data. In: Phung D., Tseng V., Webb G., Ho B., Ganji M., Rashidi L. (eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2018. Lecture Notes in Computer Science, Springer, vol 10939, pp 247-259 (2018).

[5] 中川裕志，森辰則，湯本紘彰:出現頻度と連接頻度に基づく専門用語抽出，自然言語処理, Vol. 10, No. 1, pp. 27-45 (2003).

[6] 東山昌彦，乾健太郎，松本裕治:述語の選択選好性に着目した名刺評価極性の獲得， 言語処理学会第 14 回年次大会論文集, pp.584-587 (2008).

金融文書を用いた事前学習言語モデルの構築と検証

Construction and Validation of a Pre-Trained Language Model Using Financial Documents

鈴木 雅弘 [1*]　坂地 泰紀 [1]　平野 正徳 [1]　和泉 潔 [1]
Masahiro Suzuki[1]　Hiroki Sakaji[1]　Masanori Hirano[1]　Kiyoshi Izumi[1]

[1] 東京大学大学院工学系研究科
[1] School of Engineering, The University of Tokyo

Abstract: BERT を始めとする事前学習言語モデルは，様々な自然言語処理のタスクにおいて成果を上げている．これらのモデルの多くは Wikipedia やニュース記事などの一般的なコーパスを用いているため，専門的な単語が使用される金融分野においては十分な効果が得られない．本研究では決算短信や有価証券報告書から事前学習言語モデルを構築する．また金融ドメインのタスクによって汎用モデルとの性能を比較する．

1　はじめに

近年，決算短信や有価証券報告書，ニュース記事や証券レポートなど，インターネットで閲覧可能な金融文書が豊富に存在する．金融関連のテキストの分析は投資やマーケット分析に役立つ一方で，毎日大量に作成されるテキストを人手によって全て分析することは難しい．そのため，近年盛んにおこなわれているのが，金融文書に自然言語処理 (NLP) を適用する金融テキストマイニングである．機械学習を用いた金融関連のツイートのセンチメント分析 [1][2] をはじめとして，金融分野における自然言語処理に，機械学習を適用する研究が多く存在する [3][4].

本研究では，日本語金融コーパスによって事前学習を行った BERT モデルと ELECTRA モデルを提案する．Word2vec[5] や GloVe[6] などによる分散表現は，教師なしのデータから知識を抽出し，テキストマイニングにおける重要な手法となっている．しかし金融ドメインにおいては特殊な単語が使用されるため，これらの単純な分散表現によるアプローチでは十分な効果が得られない．BERT[7] は事前学習によって各言語タスクの精度を大幅に改善した．BERT は Attention 機構をベースとした Transformer[8] によって主に構成される．まず大規模言語コーパスから事前学習し，その後出力に近いレイヤーのみを学習させるファインチューニングを組み合わせる．また ELECTRA[9] は BERT に GAN[10] のアイデアを加え，さらに Generator において最尤法 [11] を適用した．その結果 GLUE において，BERT より少ない計算量で高い精度を示した．日本語においても Wikipedia の記事から事前学習された BERT モデルが提案されている [12][1]．しかし金融コーパスと一般的なコーパスとの間で語彙や表現の違いが大きいため，一般的なコーパスのみで学習したモデルは金融テキストマイニングのタスクに最適とは言えない．また英語では FinBERT[13] として，Wikipedia や金融に関するニュース記事などを組み合わせたコーパスから構築した事前学習モデルが提案されている．しかしこれらはニュース記事にとどまり，金融の専門用語が多く用いられているわけではない．またファインチューニングや再事前学習 [14] を行うことも考えられるが，これらはネットワークの重みを変更するのみで，入力文をトークン化する時に必要なトークンの語彙を変更することはできない．本研究では，金融コーパスと Wikipedia を組み合わせたコーパスから Small サイズの事前学習 BERT モデルと ELECTRA モデルを構築する．また比較対象として Wikipedia のみからも Small サイズの事前学習モデルをそれぞれ構築する．これらの事前学習モデルを用い，金融ドメインのテキストを対象とした 2 つの評価実験を行い性能を評価する．

本研究の貢献は以下の通りである．

- 金融ドメインの文書と Wikipedia からなるコーパスを組み合わせ，金融特有の専門単語を反映した金融 BERT モデル・金融 ELECTRA モデルを構築した．
- 構築した金融モデルを，金融ドメインで研究されているタスクを対象に実験を行い，一般的なコー

*連絡先：東京大学大学院工学系研究科
〒 113-8686 東京都文京区本郷 7-3-1
E-mail: b2019msuzuki@socsim.org

[1] https://github.com/cl-tohoku/bert-japanese

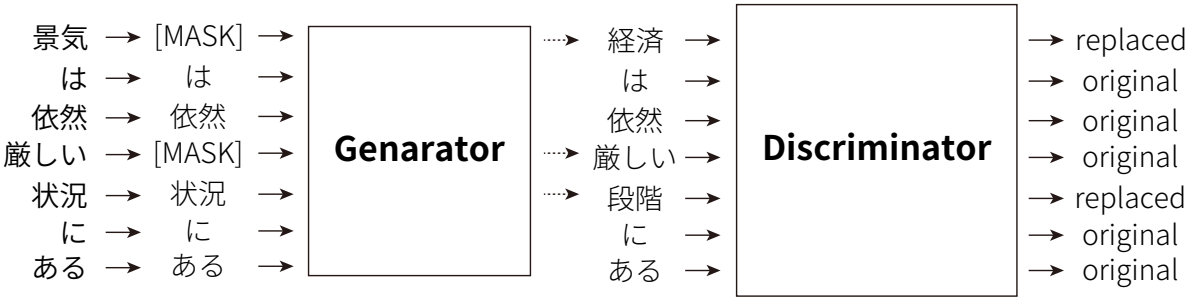


図 1: ELECTRA の概要図．Generator では Masking の対象となったトークンの予測を行う．Discriminator では Generator から出力されたトークンが元のトークンと置き換えられているかを判別する．例では「景気」「厳しい」「状況」の 3 単語が Masking の対象となり，そのうち「景気」「厳しい」の 2 単語が [MASK] に置換され Generator に入力される.「状況」は [MASK] に置換されずにそのまま入力される．Generator によって 3 単語がそれぞれ「経済」「厳しい」「段階」と予測される.「景気」「状況」の 2 単語がそれぞれ Generator によって「経済」「段階」に置換されたため，Discriminator では”replaced”のラベルが付与される．それ以外の単語は元のトークンから置換されなかったため”original”のラベルが付与される．

パスである Wikipedia のみから作成したモデルよりも高い性能を持つことを示した．

2 モデルの構築

BERT や ELECTRA といった言語モデルでは，大規模なコーパスに対してタスクを与えて事前学習 (pre-training) し，各タスクでファインチューニングするという 2 段階から構成される．

2.1 BERT の事前学習タスク

BERT の事前学習は，単語の穴埋め (Masked LM) と 2 文の接続性の判定 (Next sentence prediction) の 2 つのタスクの学習によって行われる．Masked LM では，各入力のトークンのうち 15%が Masking の対象となり，事前学習ではこの対象となったトークンを予測する．Masking の対象となったこれらのトークンのうちさらにそのうち 80%のトークンが [MASK] トークンに置換され，10%がランダムに別のトークンに置換される．残った 10%のトークンは元のトークンのまま入力される．Next sentence prediction では，入力のうち 50%は実際に存在する連続した 2 文を [SEP] トークンでつなぐ．残りの 50%はランダムにサンプリングしたドキュメントから抽出し，実際には連続しない 2 文を [SEP] トークンでつなぐ．各入力について [SEP] トークンの前後の 2 文が実際に連続しているかを学習する．BERT の論文 [7] では 2 文を [SEP] トークンで接続して入力したとのみ記載があるが，本研究では計算の効率化のため，入力上限である 128 トークンになるまでサンプルの先頭と後尾にトークンを追加する．

2.2 ELECTRA の事前学習タスク

ELECTRA の事前学習は，入力文の一部のトークンを置き換え，置き換えたトークンを検知するタスク (Replaced Token Detection) によって行われる．ELECTRA は図 1 のように Generator と Discriminator の 2 つのアーキテクチャによって構成され，それぞれに与えられたタスクを同時に学習するマルチタスクによって事前学習を行う．Generator と Discriminator のどちらも，BERT と同様に Transformer の Encoder を重ねたものである．ファインチューニングで使用されるのは Discriminator 部分のみである．Generator は Discriminator が学習しやすいように，Discriminator の 1/4 から 1/3 のサイズ (ELECTRA+は Discriminator と同じサイズ) に設定する．入力トークンの 15%を Masking の対象とする．そのうち 85%のトークンを [MASK] トークンに置換し，残りの 15%は元のトークンのまま Generator に入力する．Generator は Masking の対象となったトークンが元々どのトークンであったかを予測する．このタスクは BERT における Masked LM と似たタスクである．Discriminator には，Generator が予測したトークンを入力する．Discriminator では，入力されたトークンが Generator によって置き換えられたかを判別する 2 値分類タスクを行う．その際，Generator によって正しく予測されたトークンは Generator によって置換されていないものとラベリングする．ELECTRA

表 1: 各コーパスによって構築された語彙から,「デリバティブ取引には，先物取引やスワップ取引がある」という文をトークン化する例．[CLS] は文頭を，[SEP] は文末などを表す．”##”はサブワードに分割された語のうち，先頭でないものに付与される．

コーパス	トークン
金融	[CLS] / デリバティブ / 取引 / に / は / ， / 先物 / 取引 / や / スワップ / 取引 / 等 / が / ある / ． / [SEP]
Wikipedia	[CLS] / デリ / ##バ / ##ティブ / 取引 / に / は / ， / 先 / ##物 / 取引 / や / スワ / ##ップ / 取引 / 等 / が / ある / ． / [SEP]

における損失関数は式 (1) によって計算される．

$$\mathcal{L}_{\mathrm{ELECTRA}} = \mathcal{L}_{\mathrm{MLM}} + 50\mathcal{L}_{\mathrm{Disc}} \tag{1}$$

ここで $\mathcal{L}_{\mathrm{MLM}}, \mathcal{L}_{\mathrm{Disc}}$ はそれぞれ Generator，Discriminator におけるタスクの Loss である．BERT と同様，計算の効率化のため入力上限である 128 トークンになるまでサンプルの先頭と後尾にトークンを追加する．

2.3 日本語における事前学習

BERT や ELECTRA では英語のコーパスを用いており，入力文をトークン化する際に半角スペースで分割し，その後 WordPiece[15] によるサブワード分割を行う．しかし，日本語の文章は半角スペースで分割することができない．そのため，本研究ではまず MeCab[16] によって形態素解析を行い，その後 WordPiece によるサブワード分割を行う．金融コーパスと Wikipedia のそれぞれによって構築された語彙によって，表 1 のように文をトークンに分割することが可能になる．表 1 の場合，金融コーパスによる語彙では「デリバティブ」や「先物」，「スワップ」を 1 語とし扱うのに対し，Wikipedia による語彙ではサブワードを用いて「デリ/##バ/ ##ティブ」「先/##物」「スワ/##ップ」のように分割して扱う．このように，Wikipedia による汎用的語彙には含まれないものの，金融文書においては登場する単語を，金融コーパスからモデルを作成することで扱うことができる．

2.4 使用データ

事前学習に用いる金融コーパスのテキストデータとして，3 種類のデータを用いる．1 つ目は 2012 年 10 月 9 日から 2020 年 12 月 31 日にかけて開示された決算短信等のデータである．2 つ目は EDINET[2]にて，2018 年 2 月 8 日から 2020 年 12 月 31 日にかけて開示された有価証券報告書等の 2 種類データを用いる．3 つ目は Wikipedia の日本語記事によるコーパスである．これら 3 つのデータセットから，金融コーパス (約 4,700 万文) を作成した．金融コーパスのデータサイズは約 8GB となった．また，金融コーパスとの比較のために，Wikipedia のテキストデータ (約 2,000 万文) のみから作成したコーパスも用いる．

[2]https://disclosure.edinet-fsa.go.jp/

表 2: 事前学習のハイパーパラメータ．Generator Size は，Transformer エンコーダーの層の数は 12，隠れ層の数は 256，Transformer エンコーダーの FFN の層数は 1024，Transformer エンコーダーの Attention Head の数は 4，Embedding の次元数は 128，学習率は 5e-4 で共通である．

パラメータ	BERT	ELECTRA	ELECTRA+
Generator Size	-	1/4	1/1
Train Steps	1.45M	1M	1M

2.5 実験設定

サブワード分割のためのトークナイザーの学習についての実験設定は東北大学[3]によって作成されたモデルを参考にした．MeCab の辞書は IPAdic を用い，語彙数は 32,768 とした．このうち 5 語を未知語を表す [UNK]，文頭に挿入される [CLS]，2 文の間や入力の最後に挿入される [SEP]，入力長を揃えるために入力される [PAD]，Masked LM タスクの際に用いられる [MASK] に割り当てた．また，新たにファインチューニングなどの際に単語を追加するために 10 語を，1 文字の単語のために 6,129 語を割り当てた．各モデルのパラメータは [9] において用いられている Small モデルを参考に，表 2 のように設定した．ELECTRA+モデルは，Google が公開している ELECTRA-Small モデル[4]の Generator のサイズが Discriminator のサイズと同じであることを

[3]https://github.com/cl-tohoku/bert-japanese
[4]https://github.com/google-research/electra

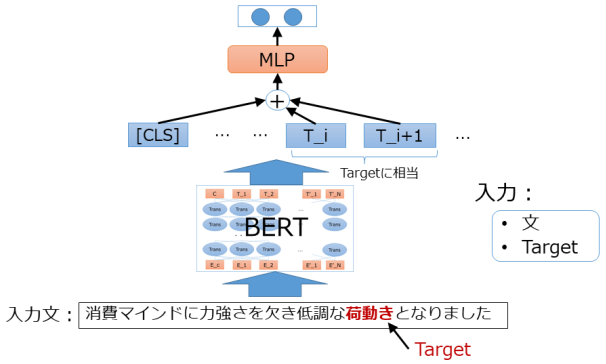


図 2: chABSA-detaset を用いた Aspect-Based Sentiment Analysis に関する実験のネットワーク概要図.

参考に作成した．Learning Rate は 10,000 ステップまで Warmup を行い，そこから線形に減衰させた．表 2 における Learning Rate は，10,000 ステップにおける値である．実装は PyTorch ベースの実装[5]を用いた．

3 評価実験

金融コーパスから構築したモデル (金融*) と Wikipedia から構築したモデル (Wiki*) に対し，ファインチューニングによる評価実験を行い性能を評価する．構築した金融モデルを評価するために，因果関係を含んでいる文を抽出するタスクを行う．因果関係を含んでいる文の抽出については，坂地ら [17] が取り組んでおり，彼らは経済・金融の記事を分析するために因果関係抽出を行っている．因果関係の抽出は，2 つのタスクで構成されており，一つ目が因果関係を含んでいる文を抽出するタスク，もう一つが因果関係を示す原因と結果表現を抽出するタスクである．ここでは，坂地らが機械学習を用いて取り組んでいた因果関係を含む文の抽出を対象に金融 BERT の性能検証を行う．この実験においては，日経新聞記事から作成されたデータを以下のように分割して用いる．1,305 文 (うち 571 文が因果関係文) を学習データに，327 文 (うち 138 文が因果関係文) を検証データに，413 文 (うち 189 文が因果関係文) をテストデータに分割し，実験に使用する．

加えて，坂地らは決算短信を対象に，同様の実験を行っており [18]，こちらのデータを対象にも実験を行う．この実験においては，決算短信から作成されたデータを以下のように分割して用いる．1,850 文 (うち 243 文が因果関係文) を学習データに，463 文 (うち 60 文が因果関係文) を検証データに，578 文 (うち 91 文が因果関係文) をテストデータに分割し，実験に使用する．

[5]https://github.com/huggingface/transformers

表 3: 日経新聞記事データを対象に因果関係を含む文の抽出における評価実験結果．Prec, Recall, F1 はマクロ平均である．

	Acc	Prec	Recall	F1
金融 BERT	**0.891**	**0.890**	**0.890**	**0.890**
金融 ELECTRA	0.872	0.876	0.867	0.869
金融 ELECTRA+	0.879	0.878	0.878	0.878
WikiBERT	0.877	0.887	0.870	0.873
WikiELECTRA	0.845	0.851	0.839	0.842
WikiELECTRA+	0.828	0.827	0.828	0.827

表 4: 決算短信データを対象に因果関係を含む文の抽出における評価実験結果．Prec, Recall, F1 はマクロ平均である．

	Acc	Prec	Recall	F1
金融 BERT	**0.929**	0.860	**0.882**	**0.870**
金融 ELECTRA	0.917	0.837	0.866	0.850
金融 ELECTRA+	0.903	0.815	0.826	0.821
WikiBERT	0.926	0.852	0.880	0.865
WikiELECTRA	0.843	0.421	0.500	0.457
WikiELECTRA+	0.929	**0.870**	0.860	0.864

表 5: chABSA-dataset を対象とした評価実験結果．Prec, Recall, F1 はマクロ平均である．

	Acc	Prec	Recall	F1
金融 BERT	**0.884**	**0.881**	0.880	**0.881**
金融 ELECTRA	0.881	0.877	**0.882**	0.879
金融 ELECTRA+	0.847	0.847	0.838	0.842
WikiBERT	0.862	0.864	0.853	0.857
WikiELECTRA	0.845	0.851	0.839	0.842
WikiELECTRA+	0.599	0.580	0.565	0.558

さらに，本論文では，TIS 株式会社が公開している chABSA-dataset[6]を用いて，Aspect-Based Sentiment Analysis に関する実験を行う．この実験では，図 2 のように入力に文と表現を入力し，その表現に関する文内でのセンチメントを出力するという問題設定にする．データセットには，Positive，Negative，Neutral のタグが付与されていたが，Neutral が他のタグに対して大幅に少なかったことから，Neutral を除外して実験を行う．ここでは，4,479 件 (うち 2,776 件が Positive) を学習データに，1,194 件 (うち 690 件が Positive) を検証データに，1,492 件 (うち 868 件が Positive) をテストデータに分割し，実験に使用する．

[6]https://github.com/chakki-works/chABSA-dataset

4 結果と考察

表 3 に日経新聞記事を対象に，また表 4 に決算短信を対象に，因果関係を含む文の抽出を行った結果をそれぞれ示す．また表 5 に chABSA-dataset を用いてセンチメント出力を行った結果を示す．

表 3，4，5 より，各実験において金融モデルを用いた方が F1 値が高くなった．このことより，金融におけるタスクでは Wikipedia をコーパスとしたモデルよりも，金融テキストをコーパスとしたモデルの方が性能が良いことを示した．モデル間の比較では，BERT モデルが最も F1 値が高かった．金融の ELECTRA モデルと ELECTRA+モデルは，日経新聞記事データを対象とした実験ではモデルサイズの大きい ELECTRA+モデルの方が F1 値が高かったものの，決算短信と chABSA-dataset を対象とした実験ではモデルサイズの小さい ELECTRA モデルの方が F1 値が高かった．これは ELECTRA の Generator が Discriminator の 1/4 から 1/3 のサイズでより精度が高かったこと [9] と同様の現象が発生していると考えられる．実際 ELECTRA の論文では ELECTRA+のような Generator と Discriminator のサイズが同じモデルについての言及はない．また決算短信データを対象とした実験での WikiELECTRA と chABSA-dataset を対象とした実験での WikiELECTRA+は，他のモデルより F1 値が大きく低くなった．これは ELECTRA の学習が不安定なことによると考えられる．今回は金融ドメインから決算短信と有価証券報告書の 2 種類のデータのみを使用したが，今後は，新聞記事の金融面など金融に関連した別のテキストデータも併せて学習することで，より高い性能を示す金融 BERT が構築できると考えられる．

5 まとめ

本論文では，決算短信等のデータと有価証券報告書等のデータを Wikipedia の日本語記事と組み合わせて金融ドメインの事前学習モデルを構築し，その性能を確認した．金融ドメインで研究されている因果関係を含む文の抽出タスクを対象に，日経新聞記事から作成された評価データと決算短信から作成された評価データを用いて実験を行い，Wikipedia から作成したモデルよりも高い性能を示した．また，chABSA-dataset を用いたセンチメント分析においても，金融 BERT や金融 ELECTRA が，Wikipedia から作成したモデルよりも高い性能を示した．

今後の課題として，因果関係を含む文の抽出やセンチメントの分析だけではなく，株価予想や要約などの他のタスクにおいて高い性能を示すことができるかを検証していく．また，Optimzier やモデルの改良により ELECTRA の学習を安定することを目指す．さらに作成した各モデルの公開に向けた準備を進めていく．

謝辞

本研究は，JSPS 科研費（JP21K12010）の助成を受けました．

参考文献

[1] Nuno Oliveira, Paulo Cortez, and Nelson Areal. The impact of microblogging data for stock market prediction: Using twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications*, Vol. 73, pp. 125 – 144, 2017.

[2] Gabriele Ranco, Darko Aleksovski, Guido Caldarelli, Miha Grčar, and Igor Mozetič. The effects of twitter sentiment on stock price returns. PLoS ONE, 2015.

[3] B. Shravan Kumar and Vadlamani Ravi. A survey of the applications of text mining in financial domain. *Knowledge-Based Systems*, Vol. 114, pp. 128–147, 2016.

[4] Li Guo, Feng Shi, and Jun Tu. Textual analysis and machine leaning: Crack unstructured data in finance and accounting. *The Journal of Finance and Data Science*, Vol. 2, No. 3, pp. 153–170, 2016.

[5] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. CoRR, 2013.

[6] Pennington Jeffrey, Socher Richard, and Manning Christopher. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*, pp. 1532–1543, 2014.

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186, 2019.

[8] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30 (NeurIPS)*, pp. 5998–6008. 2017.

[9] Kevin Clark, Minh-Thang Luong, Quoc V. Le, and Christopher D. Manning. Electra: Pre-training text encoders as discriminators rather than generators, 2020.

[10] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

[11] Massimo Caccia, Lucas Caccia, William Fedus, Hugo Larochelle, Joelle Pineau, and Laurent Charlin. Language gans falling short, 2020.

[12] 柴田知秀, 河原大輔, 黒橋禎夫. Bert による日本語構文解析の精度向上. 言語処理学会 第 25 回年次大会, 2019.

[13] Zhuang Liu, Degen Huang, Kaiyu Huang, Zhuang Li, and Jun Zhao. Finbert: A pre-trained financial language representation model for financial text mining. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pp. 4513–4519, 2020. Special Track on AI in FinTech.

[14] 仁木裕太, 坂地泰紀, 和泉潔, 松島裕康. 再事前学習した bert を用いた金融文書中の因果関係知識有無の判別. 人工知能学会全国大会論文集, 2020.

[15] Mike Schuster and Kaisuke Nakajima. Japanese and korean voice search. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5149–5152, 2012.

[16] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to japanese morphological analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP 2004)*, pp. 230—237, 2004.

[17] 坂地泰紀, 増山繁. 新聞記事からの因果関係を含む文の抽出手法. 電子情報通信学会論文誌 D, Vol. J94-D, No. 8, pp. 1496–1506, 2011.

[18] 坂地泰紀, 酒井浩之, 増山繁. 決算短信 pdf からの原因・結果表現の抽出. 電子情報通信学会論文誌 D, Vol. J98-D, No. 5, pp. 811–822, 2015.

機械学習を用いた統合報告書の ESG 関連ページの推定

Estimation of ESG-related pages in integrated reports using machine learning

河村 康平 [1]　高野 海斗 [2]　酒井 浩之 [1]
永並 健吾 [1]　中川 慧 [2]
Kohei Kawamura[1], Kaito Takano[2], Hiroyuki Sakai[1]
Kengo Enami[1], Kei Nakagawa[2]
[1] 成蹊大学
[1] SEIKEI University
[2] 野村アセットマネジメント株式会社
[2] Nomura Asset Management Co., Ltd.

Abstract: 近年の資産運用分野では，財務情報である売上や利益だけでなく，非財務情報である環境(Environment)，社会(Social)，企業統治(Governance)の 3 つの観点を考慮して投資を行う「ESG 投資」が世界的に広まりつつある．そのため ESG 投資において，企業による ESG 情報を判断材料として獲得することは重要である．日本においては，企業の自社の ESG 情報を開示する手段として統合報告書がある．決算短信や有価証券報告書のような他の金融テキストとは異なり，統合報告書には企業の財務情報に加え非財務情報が開示されており，ESG 投資を行う上で重要な情報源であると言える．しかし，統合報告書の中には 100 ページを超えるものもあり，人手で ESG 情報を探すには多くの時間と労力が必要となる．統合報告書において ESG 情報が存在する位置を自動で推定することが可能になれば，例えば，企業の成長・持続可能性を判断し，投資リスクを抑えることが可能となる，そこで本研究では，ESG に関連する内容についての記述が含まれるページを ESG 関連ページと定義し，機械学習手法を用いて統合報告書から ESG 関連ページを推定する手法を提案する．

1. はじめに

近年資産運用の分野では，財務情報である売上や利益だけでなく，非財務情報である環境(Environment)，社会(Social)，企業統治(Governance)に対する企業の取り組みを考慮して投資を行う「ESG 投資」が世界的に広まりつつある[1]．これは，短期的な利益追求による金融危機の反動から企業の長期的な価値が重要視されるようになり，目の前の売上や利益にとらわれない長期的・持続的な展望への関心が高まっていることが背景にある．

資産運用分野での世界的な ESG 投資に対する関心の高まりを受け，投資家など様々なステークホルダーは ESG 情報などの非財務情報の開示を企業に求めるようになり，企業の ESG 情報(環境，社会，企業統治の問題に対する取り組みについての情報)の開示の重要性は高まり続けている．日本においても，ESG 情報の開示企業は，近年増加傾向にある[2]．

一方で，機関投資家や個人投資家にとって企業が開示している ESG 情報の獲得・利用には 2 つの問題が存在する．1 つは企業の ESG 情報の開示媒体が多様なことにより ESG 情報の満足な活用が出来ていないことが挙げられる．企業は ESG 情報の開示手段として統合報告書，CSR 報告書，有価証券報告書，自社ホームページなど多くの媒体を使用しており，全ての内容を確認するには多くの労力と時間が必要となる．2 つ目は，比較可能性の欠如である．統合報告書，CSR 報告書，サステナビリティレポートなどの任意開示は ESG 情報のような非財務情報を多く含む一方で，財務情報のように開示に関する統一的な基準が存在しないために，情報開示の量と質共に各企業の裁量に委ねられることが多い，そのため．同業他社，あるいは同企業の年ごとの比較や分析などが容易でなく，投資家の投資判断材料とするには困難である．

上記の 2 つの問題を解決するためには，企業が公

開している媒体から ESG 情報を自動で抽出し，ESG 情報の比較や定量的分析が可能な形式に変換できる仕組みが必要である．我々は，このような仕組みの実現を目指し，第一段階として，企業の ESG 情報が掲載されている媒体として統合報告書に注目し，統合報告書の ESG 情報が含まれる可能性の高いページを，機械学習モデルにより自動で推定することを試みる．

統合報告書とは，売上や利益などの財務情報に加え，ESG 情報などの非財務情報が開示された報告書である．[3]の調査では，企業への「ESG への取り組みについて情報を開示している媒体をお答えください（複数選択可）」という問に対して，解答数 500 のち 50.2%の企業が「統合報告書」と解答している．これは，ホームページと IR 説明会資料を除いた中では最も高い結果であり，統合報告書は企業が発行する報告書の中では ESG 情報を獲得しやすい媒体であると言える．しかし統合報告書に対して機械学習や統計的分析を適用することは今まで困難とされてきた．その理由は，統合報告書は任意開示であり，開示に関する統一的な基準が存在せず，開示内容が企業毎に大きく異なるからである．また，開示形式についても，決算短信や有価証券報告書のような他の金融テキストのように共通した形式が存在せず，企業毎に独自のデザインや図表を用いており，PDF ファイルをテキストデータに変換する際にレイアウトが崩れてしまうなどの問題もあった．

本研究で扱うデータは，統合報告書の PDF ファイルをテキスト情報に変換したデータである．そのため，先述したように文の前後が変わる，文が分割される，文字化けが起こるなど元のレイアウトの情報を損なうこともある．提案手法では，このような文やレイアウトが崩れたテキストデータに対して，自動生成した学習データを用いて機械学習モデルを学習し，統合報告書のページ単位でマルチラベル分類を行うことで，E,S,G それぞれについて関連性の推定を試みる．さらに，本研究で扱うような形式が崩れたテキストデータに対して，機械学習モデルの適用がどの程度可能なのかを検討する．

2. 関連研究

本研究の特徴は以下の通りである．

（1）対象は統合報告書
（2）学習データの自動生成
（3）ページ単位でのマルチラベル分類

これらの点を踏まえて関連研究を挙げていく．

ESG 情報の抽出に関連する研究として[4]がある．[4]は有価証券報告書から ESG 関連文（ESG に関連する文）を抽出する手法を提案した．[4]と本研究の違いは対象が統合報告書であることと，ESG 情報の抽出をページ単位で行うことである．統合報告書から ESG 情報を文単位で取り出すことは可読性や分析のしやすさから需要があるが，統合報告書 PDF のテキスト化の際に文の形が崩れてしまうことが多いことから，本研究ではページ単位で抽出することにした．また，統合報告書には有価証券報告書などの金融テキストに比べ ESG 情報が豊富に含まれるため文として抽出するメリットは少なく，統合報告書はページごとに内容がある程度まとまっている点からも，統合報告書においてはページ単位での ESG 情報の抽出が有効であると考える．

統合報告書を用いたテキストマイニング分析に関連する研究として[5]がある．[5]では，環境報告書やサステナビリティ報告書，統合報告書の経営トップメッセージのテキストに対して，深層学習により自動的に社会，及び環境ラベルを付与することで，3 つの媒体の違いによる環境および社会情報開示の記述量の傾向を分析した．[5]は，テキストマイニングの分析対象を統合報告書の中の経営トップメッセージのみに絞り込んでいるが，統合報告書には，ESG 情報が経営トップメッセージ以外の場所にも豊富に掲載されているため，より広範囲の情報を利用できた方が投資において良い判断材料になるはずである．我々は，統合報告書の一部ではなく，全てのページを機械学習の適用範囲とするため，統合報告書の広範囲の内容から ESG 情報を獲得することができる．また，本手法により抽出した ESG 関連ページを用いることで，統合報告書の広範囲の内容を対象としたテキストマイニング分析が可能になると考える．

金融テキストに対してページ単位で機械学習モデルを適用した研究として，[6]がある．[6]では株主招集通知という金融テキストから，ルールベースで自動生成した学習データを用いて深層学習モデルを学習し，株主招集通知から自動的に重要ページを抽出する手法を提案した．[6]では本研究のように形式が崩れたテキストデータを扱っており，このようなデータに対して様々なモデルの提案や考察を行った．[6]と本研究は，単語の出現情報を利用したルールベースにより学習データを自動生成する点で共通しているが，本研究の処理対象は統合報告書であるため，学習データの生成に使用する単語が異なる．ESG 関連ページには必ず出現するような単語は存在しないが，高い確率で出現する単語は存在するため，我々はこのような単語を利用し学習データを自動生成する．また，ページに対して単一ラベルではなくマルチラベルを付与する点でも異なる．

3. 提案手法

統合報告書は，1 つのページに ESG の複数のトピックの内容が開示されることもある．本研究では，統合報告書のあるページについて，そのページが E,S,G のそれぞれのトピックに関連するかどうかをマルチラベル分類により推定する手法を提案する．

ただし ESG については，一般にどのような話題が該当するのか明確に定められておらず，投資家によって解釈が異なる．本研究では，表 1 に示す話題を ESG と定義する．また，図 1 に ESG 関連ページの例を示す．

表 1. 本研究における ESG の定義

トピック	話題
E	気候変動，生物多様性，汚染，資源利用，有害物質の排出と廃棄物
S	ダイバーシティと機会均等，地域社会，人材育成，顧客の健康と安全に対する責任，非差別，労働基準・雇用
G	企業倫理，反競争的行為，汚職・腐敗の防止，リスクマネジメント，コーポレート・ガバナンス，ステークホルダー・エンゲージメント

図 1. ESG 関連ページの例

3. 1. 提案手法の概要

提案手法の概要を以下に示す．

Step1: 統合報告書 PDF ファイルをテキストデータに変換し，さらにルールベースによりテキストデータを整形する．

Step2: Step1 で取得した統合報告書のテキストデータを用いて E,S,G それぞれに対する特徴語を抽出する．

Step3: Step1 で収集したページに対して，特徴語を用いたルールベースによりラベルを付与し，学習データを自動生成する．

Step4: 学習したモデルを用いて，統合報告書から ESG 関連ページを推定する．

3. 2. 統合報告書 PDF のテキスト抽出

統合報告書の各ページを機械学習モデルに入力するために，まず PDF ファイルをテキストデータに変換する必要がある．PDF ファイルをテキストデータに変換するために，MacOS に標準インストールされている「Automator」を使用した．「Automator」を用いることで，PDF ファイルからテキストデータをページごとに取得することが可能である．また．他のソフトに比べて比較的段組みを考慮して PDF ファイルをテキストデータに変換出来る．しかし「Automator」による変換により，不要な空欄や改行が挿入されるため，処理がしやすいように可能な限りきれいな形に整えるために，ルールベースによるテキストデータの整形を行う．具体的には以下のようなルールを設ける．

・数値を全て 0 に変換する．
・空欄を空文字に変換する．
・句点“。”を改行に変換する．
・品詞が助詞・助動詞で終わった場合，後ろの文と連結する．など

4 つめに示した文の連結を行うルールは，本来繋がりのない文が後ろにあった場合は適切に文を連結出来ないが，このような場合よりも，連結ルールにより正しく文を修正できた場合の方が多かったため，ある程度の誤りは許容することにした．

3. 3. 特徴語の抽出

特徴語は以下の手法で獲得される．

Step1: 統合報告書から単語“ESG”を含むページを取得し，取得したページの文集合から Word2Vec により 100 次元の分散表現を学習する．

Step2: 統合報告書に出現する名詞nと，$w \in$｛“環境”，“社会”，“ガバナンス”｝に対して正規化自己相互情報量$npmi(w,n)$を計算する．

Step3: ある企業tの統合報告書に名詞nが出現する確率$P(w,n)$に基づくエントロピー$e(n)$を計算する．

Step4: Step2，Step3 で計算した$npmi(w,n)$と$e(n)$が共に上位の名詞nを特徴語候補とする．

Step5: Step1 で学習した 100 次元ベクトル空間上で，特徴語候補をk-means 法($k = 15$)でクラスタリングする．

Step6: Step6 でクラスタリングされたクラスタから E, S，G ごとに対応するクラスタを選択し，それらの中の名詞を人手で選別することで，各クラスタ内の名詞を特徴語として獲得する．

Step2 において，環境，社会，ガバナンスのそれぞれと多く共起する単語は環境，社会，ガバナンスのそれぞれと関連性が高いという考えに基づき，$w \in$｛“環境”，“社会”，“ガバナンス”｝と，統合報告書に出現する名詞nについて式 1 により自己相互情報量を求める．なお，本研究ではある 2 つの単語が同じページに出現することを共起とする．

$$npmi(w,n) = \frac{pmi(w,n)}{-\log_2 P(w,n)} \quad (1)$$

$$pmi(w,n) = \log_2 \frac{P(w,n)}{P(w)P(n)} \quad (2)$$

$P(x)$: 統合報告書のページ集合において単語xが出現する確率

$P(x,y)$: 統合報告書のページ集合において単語x,yが同時に出現する確率

式(1)は式(2)の値を正規化したものであり，x,yの希少度による影響を軽減し，ノイズとなる単語（人名など）を除くことができる．

Step3 における，ある企業tの統合報告書に名詞nが出現する確率$P(t,n)$に基づくエントロピー$e(n)$は式(3)で求める．

$$e(n) = -\sum_t P(t,n)\log_2 P(t,n) \quad (3)$$

$e(n)$は任意の企業の統合報告書集合を 1 つの文書とみなし，それらの文書集合において名詞nが満遍なく出現している場合に高い値をとる尺度である．つまり$e(n)$に閾値を設けることにより特定の企業の統合報告書に出現する単語や，統合報告書のテキストデータ化の際に発生したノイズを除去することができる．

Step2 で得た$npmi$，及び Step3 で得たeが共に上位である名詞集合は E,S,G のいずれかに高い関連性があり，かつ会社名や人名などの ESG とは関係ない名詞を除いたものとなっている．

最終的に，表 2 に示すような特徴語を E,S,G それぞれに対して獲得できる．

表 2. 獲得する特徴語の例

トピック	特徴語
E	環境負荷，カーボン，温室効果ガス，生物多様性
S	障がい者，育児休業，地方創生，教育支援，男女
G	リスク管理，コンプライアンス，透明性，法令遵守，是正措置

3.4. 学習データの自動生成

ページ単位で ESG の関連性の推定を行うためには，学習データをページ単位で用意しなければならないが，1 つのページにはタイトルや本文のような文字列が多く含まれるため，学習に必要な十分な量の学習データを手作業で作成するには多くの時間と労力が必要となる．そこで本研究では，統合報告書において ESG の特徴語が多く出現するページは，その特徴語のトピックに関連しているという考えに基づき，単語の出現情報を用いたルールを設定し，学習データを自動生成する．具体的には，統合報告書の学習用のページ集合において，E,S,G の各トピックについて，特徴語が重複無しで 5 個以上出現するページに対してそのトピックにラベル 1 を，特徴語が 1 つも出現しなかったページに対してそのトピックについてラベル 0 を付与する．ただし，学習データの精度を高めるために，上記のルールに少しでも当てはらないページは学習データから除く．例えば，あるページに G の特徴語が 1～4 個出現した場合，そのページが G に関連するかの判定が困難であるため，G のラベルを付与せず，このページは学習データには追加しない．

ラベル 1 を付与するための条件である特徴語の閾値は，閾値を小さくすると学習データとして抽出できるページ数が増える一方でラベル付与の精度が下がり，閾値を大きくすると学習データとして抽出できるページ数が減る一方でラベル付与の精度が上がるというトレードオフの関係を考慮した結果，5 という値を選択した．

ルールベースにより抽出した学習データは，最終的に 3 次元の multi-hot 形式のラベルが付与される．ここで，ラベルの 1 次元目は E に関連するかどうかの 2 値(関連があれば 1，非関連であれば 0)を取る．同様に 2 次元目は S，3 次元目は G に対応する．

3.5. 分類モデル

本研究で扱うタスクにおいて，どのような分類器が適切であるかを比較するために，3 つのモデルを提案する．

3.5.1 モデル 1

入力を統合報告書のあるページとし，そのページが E,S,G のそれぞれについて関連するかを判別する分類器として，サポートベクトルマシン(SVM)を選択した．すなわち，統合報告書のページが E,S,G のそれぞれに関連するかそうでないかを判別する分類器を，E,S,G の 3 種類生成し，テストデータとなる統合報告書のページが E,S,G のぞれぞれに関連するか判定する(One vs Rest)．したがって，例えばページが E に関連するかを判別するための学習データは, E に関連するページが正例，E に非関連なページが負例（具体的には multi-hot ベクトルの 1 次元目が 1 のページが正例，0 のページが負例）となる．

各分類器の入力は[7]の素性選択手法を参考に，学習データから取得した素性および素性値を使用したベクトル(Bag-of-Words)とする．

3.5.2 モデル 2

入力を統合報告書のあるページとし，そのページが E,S,G のそれぞれについて関連するかを判別する深層学習によるマルチラベル分類器として多層パーセプトロン(MLP)を選択した．入力はモデル 1 で使用した 3 種類の分類器についての素性をまとめたものを素性としたベクトル(Bag-of-Words)とする．

モデルの入力層のノード数を入力ベクトルの次元数と同じとし，隠れ層は 1,000 次元が 3 層，300 次元が 3 層，25 次元が 3 層の計 9 層とする．出力層は E,S,G に対応する 3 次元とする．活性化関数はランプ関数(ReLU)を使用し，出力層はシグモイド関数を使用する．損失関数には BCE(Binary Cross Entropy)を適用する．

3.5.3 モデル 3

モデル 1,2 のようなベクトルを入力で受け取る分類器は, 入力であるページに素性が含まれない場合，分類器の誤判定につながる可能性がある．そこで，入力を統合報告書のあるページとし，そのページに含まれる複数の文の文脈の情報を考慮して，かつ分類の手がかりとなる文を取捨選択し，そのページが E,S,G のそれぞれについて関連するかを判別するマルチラベル分類器として，双方向 LSTM(BiLSTM)と Attention を組み合わせた分類器を選択した．分類器の処理の流れは，統合報告書のページに含まれる各文を入力として受け取り，各文を単語埋め込み層により系列に変換し，双方向 LSTM に通すことで文の文脈を考慮した表現に変換，さらに Attention による重み付き和を計算することにより，ページの各文の情報を圧縮した表現を得る．最終的に圧縮した表現を線形変換することで E,S,G それぞれの関連性を推定する．単語埋め込み層の出力は 300 次元とし，双方向 LSTM の出力は 600 次元とする. 出力層は E,S,G に対応する 3 次元とし，活性化関数はシグモイド関数を使用する．損失関数には BCE(Binary Cross Entropy)を適用する．

本研究では，あるページの入力に対して，ページに含まれる各文に対応する双方向 LSTM の出力を $H = (h_1, h_2, \dots, h_t)$としたとき, 式(6)によりページに含まれる各文の情報をまとめた表現xを得る．

$$u_i = \tanh(W h_i) \tag{4}$$

$$\alpha_i = \frac{\exp(u_i^T v)}{\sum_i^t \exp(u_i^T v)} \tag{5}$$

$$x = \sum_i \alpha_i h_i \tag{6}$$

ただし, Wとvは学習により求められるパラメータであり，αは入力された文の注目度の重みを表す．

4. 実装と評価

本手法の評価を行うため，本手法を実装した．実装にあたり 2015 年から 2019 年までに発行された上場企業 412 社の 1,251 個の統合報告書 PDF ファイルの約 80,000 ページを学習データの自動生成および分類器の学習に使用した．

特徴語の抽出手法を実行することにより，最終的に E の特徴語を 187 個，S の特徴語を 176 個，E の特徴語を 152 個それぞれ抽出し，これらの特徴語を使用し, 学習データを自動生成した, ただし E,S,G のどのトピックにも関連しないラベル[0,0,0]のページが他のラベルに比べて多く抽出されたため，分類モデルの学習の際にはデータの偏りを軽減するために, ラベル[0,0,0]のデータからランダムにサンプリングした 1,000 件をラベル[0,0,0]の学習データとする．

また，本手法で自動生成した学習データを用いて学習された分類器の性能を評価するための評価データとして，学習データの自動生成に使用していない 2020 年もしくは 2021 年に発行された 4 企業の統合報告書 PDF ファイルの計 372 ページに対して，各企業の web ページにて公開されている GRI スタンダード対照表を参考に，人手にて正解ラベルを付与し作成した．本手法で自動生成した学習データと，人手にて作成した評価データの，ラベルごとのページ

表 4. 各モデルの評価結果

	E(P)	E(R)	E(F)	S(P)	S(R)	S(F)	G(P)	G(R)	G(F)
モデル 1	87.6	77.3	82.1	71.5	72.7	72.1	68.3	71.7	69.9
モデル 2	78.7	80.7	79.7	62.6	75.8	68.6	69.9	77.4	73.4
モデル 3	78.4	88.2	83.0	64.0	82.0	71.9	65.8	82.4	73.2

表 5. 学習データの評価結果

	E(P)	E(R)	E(F)	S(P)	S(R)	S(F)	G(P)	G(R)	G(F)
ルール	92.5	97.4	94.9	92.5	92.5	92.5	92.5	100	96.1

数を表 3 に示す．ただし，○は該当するトピックに関連することを示す．

表 3. 自動生成した学習データと，人手にて作成した評価データのページ数

ラベル	E	S	G	学習データ	評価データ
[0, 0, 0]				13870	89
[1, 0, 0]	○			375	62
[0, 1, 0]		○		797	57
[0, 0, 1]			○	923	74
[1, 1, 0]	○	○		97	5
[1, 0, 1]	○		○	203	22
[0, 1, 1]		○	○	1228	36
[1, 1, 1]	○	○	○	628	27

提案モデル 3 は，入力としてあるページに含まれる複数の文を受け取るが，ページごとに含まれる文の数は異なるため，モデルに入力する文の数を統一する必要がある．本手法では入力文の数の上限を 100 とした．ただしページに含まれる文については，分類の判断に悪影響を与える可能性のある明らかなノイズを予め除去するため，以下の条件を全て満たす文字列とする．

- 文字数が 4 以上かつ 220 以下
- “http”を含まない
- 同じ文字の連結ではない（¥¥¥¥¥など）
- 同じページに重複して出現していない

本手法の評価を行うために，評価データを使用し，3 つの提案モデルについて，モデルの出力である 3 次元ベクトルの次元ごとに 2 値分類で評価を行った．評価指標として適合率(P)，再現率(R)，F1-score(F)を求めた．表 4 に各提案モデルに対する評価値を示す．表 4 の見方については，先頭に“E”のついた評価指標は出力である 3 次元ベクトルの，1 次元目における 2 値分類の評価を表しており，同様に“S”は 2 次元目，“G”は 3 次元目における 2 値分類の評価を表す．

5. 考察

以下の 3 つの内容に対して考察を行う．

(1)学習データを自動生成したことによる分類性能への影響

(2)形式崩れの多いテキストデータに対して本研究のようなタスクを行う場合，どのような機械学習モデルが適切なのか

(3)崩れたテキストデータに対しての機械学習モデルによる文単位での処理の有効性

まず(1)について述べる．本研究では，ルールベースにより学習データを自動生成しているが，モデルの適切な学習のためには，学習データのラベル付与の精度が重要であり，モデルの分類性能に影響する．そこで，本手法で自動生成された学習データの評価を行った．自動生成された学習データは，18,121 ページと数が膨大であるため，ルールにより付与されたラベルごとにランダムに 10 件選んだ，合計 80 ページに対し，人手で確認し正解ラベルを付与することで，学習データの定量的な評価を行う．ルールにより自動生成したラベルと人手により付与した正解ラベルを用いて，ESG の各トピックごとの 2 値分類の適合率(P)，再現率(R)，F1-score(F)を表 5 に示す．

表 5 より，各トピックごとに，ルールによる学習データのラベル付与の精度を確認すると，高い精度でラベルを付与できているが，表 4 の評価データを用いた提案モデルによる評価値は，学習データの評価値と大きな差がある．これは，ルールに当てはまるページに対しては正しくラベル付与し学習データとして用意できたが，ルールに当てはまらなかったページに対してモデルは学習できず，適切な分類が出来なかったためと考える．

また，本研究の学習データの自動生成手法におい

て学習データのラベル付与の誤りや見逃しが起こる原因としては，ラベル付与に単語の出現情報を利用していることが起因していると考える．まず，ルールにより誤って関連ありと判断される場合(0→1)について考察する．本手法では学習データを自動生成する際にページの内容については考慮せずに，あるページに特徴語が一定数出現したかどうかでそのトピックに関連があるか判断するため，ESG に非関連なページに対しても，特徴語を含んでいると ESG に関連があると誤認識してしまう場合がある．この問題が頻繁に発生するのが目次ページである．目次ページは，ESG 情報や他の内容が PDF のどの位置に掲載されているか示すページであり，目次ページ自体に ESG 情報は含まれないため ESG に非関連と認識することが望ましいが，本手法のルールでは ESG に関連ありと誤認識されてしまう．このようなページが学習データに追加されると，分類モデルの適合率の低下に繋がる可能性があるため，学習データから極力取り除くべきである．

次にあるルールによるラベル付与の見逃し(1→0)について考察する．我々は特徴語を一定数含むページを学習データとしたが，本手法で全ての特徴語は獲得できていないため，獲得できなかった特徴語に関する話題のページについては，そのトピックのラベルを付与できない．例えば，環境に配慮した取り組みを表す“グリーン調達”という単語は E の特徴語として適切だが，本手法では特徴語として抽出できなかったため，あるページに“グリーン調達”が出現しても E に非関連と認識されてしまう場合がある．このようなラベル付与を見逃したページは分類モデルの再現率の低下を招くと考えられるため，学習データに追加させないほうがよい．以上より，分類モデルの性能を向上させるには，本手法でのルールベースによる学習データの自動生成手法の改善が必要であることが分かった．

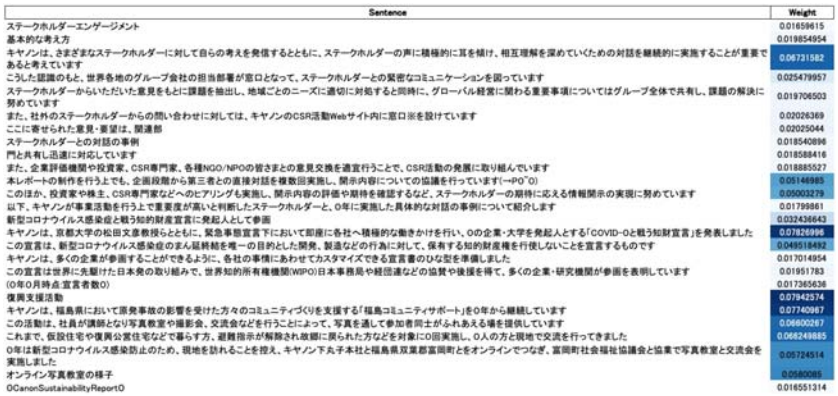
次に(2)について述べる．従来，本タスクで扱うような形式が崩れたテキストデータに対して機械学習を適用する場合は，単語単位に分割して行われてきた．これは，文が崩れても単語の出現情報は失われにくいという点で有効だが，単語の順序情報を失うという欠点も存在する．本研究では，入力に単語の出現情報を利用した分類器として提案モデル 1,2 を，文の文脈情報を利用した分類器として提案モデル 3 を選択し，形式崩れの多いテキストデータに対してどのような機械学習モデルが適切かの考察を行う．

まず，各提案モデルによる統合報告書のページへのラベル付与の正確性を評価するため，表 5 の適合率に注目する．表 5 より適合率は，全ての提案モデルで E が 8 割，S と G は 6,7 割となっており．モデルによる適合率の大きな違いは見られなかったが，ESG の全てのトピックで最も高い適合率を得たのは提案モデル 1 であった．一方で，提案モデル 1,2 に比べて，提案モデル 3 は適合率が若干低い傾向にある．これは，モデルの特徴に起因すると考える．提案モデル 3 は入力であるページの各文を受け取るため，提案モデル 1, 2 よりも受け取る情報が多くなる．統合報告書の ESG 情報を含むページには，ESG と関連がない話題が含まれていることもあるため，広範囲の情報を受け取る提案モデル 3 は，学習時にそのような話題についても ESG の特徴と学習したため，他のモデルに比べ適合率が下がったと考える．

次に各提案モデルが統合報告書のページに対して，ラベル付与をどの程度取りこぼしていないかを考察するため，表 5 の再現率に注目する．表 5 より，全ての提案モデルが ESG の全てのトピックにおいて 7 割を超えており，ESG 情報を比較的見落とすこと無く認識できている．特に，提案モデル 3 は ESG の各トピックについて再現率が 8 割を超えており，他の提案モデルよりも多くの ESG 情報を認識している．提案モデル 1,2 は入力としてベクトルを受け取るため，入力ページをベクトルに変換する必要があるが，変換する際にページに含まれる文の文脈情報は無視され，かつ学習に使用した素性以外の単語の出現情報は無視される．そのため細かな ESG 情報を認識しにくい一方で，提案モデル 3 は入力として，入力ページに含まれる各文を受け取るため，各文の文脈情報を捉えることができ，ESG 情報の取りこぼしが減り再現率が高い結果になったと考える．ただし，単語の出現情報を利用するモデルは情報の見落としにより再現率が低くなると予想していたが，そのような提案モデル 1，2 において，全てのトピックで再現率が 7,8 割を達成しているため，使用した素性が有効であったと考える．

以上の考察より，形式崩れの多いテキストデータを用いる本タスクにおいて，ESG の関連性推定の正確性を優先する場合は，提案モデル 1, 2 のような単語の出現情報を用いたモデルが適切であり，ESG 情報の取りこぼしを減らすことを優先する場合は，提案モデル 3 のような文の文脈情報を用いたモデルが適切な選択であると考える．

最後に(3)について述べる．本研究では崩れたテキストデータから ESG 情報をページ単位で抽出する際に，文単位での処理が有効であるかを調査するために提案モデル 3 を選択した．提案モデル 3 は，入力情報に対して分類の手がかりとしての重みを計算することで，情報の取捨選択が可能な Attention を持つ．ノイズが多いデータを文単位で扱うためには，不要な文をフィルタリングし，重要な文を強調して



Sentence	Weight
ステークホルダーエンゲージメント	0.01659615
基本的な考え方	0.019854954
キヤノンは、さまざまなステークホルダーに対して自らの考えを発信するとともに、ステークホルダーの声に積極的に耳を傾け、相互理解を深めていくための対話を継続的に実施することが重要であると考えています	0.067315182
こうした認識のもと、世界各地のグループ会社の担当部署が窓口となって、ステークホルダーとの緊密なコミュニケーションを図っています	0.025479957
ステークホルダーからいただいた意見をもとに課題を抽出し、地域ごとのニーズに適切に対処すると同時に、グローバル経営に関わる重要事項についてはグループ全体で共有し、課題の解決に努めています	0.019706503
また、社外のステークホルダーからの問い合わせに対しては、キヤノンのCSR活動Webサイト内に窓口※を設けています	0.02026369
ここに寄せられた意見・要望は、関連部	0.02025044
ステークホルダーとの対話の事例	0.018540896
門と共有し迅速に対応しています	0.018588416
また、企業評価機関や投資家、CSR専門家、各種NGO/NPOの皆さまとの意見交換を適宜行うことで、CSR活動の発展に取り組んでいます	0.018885527
本レポートの制作を行うにでも、企画段階から第三者との直接対話を複数回実施し、開示内容についての協議を行っています(→P07〇)	0.05146965
このほか、投資家や株主、CSR専門家などへのヒアリングも実施し、開示内容の評価や期待を確認するなど、ステークホルダーの期待に応える情報開示の実現に努めています	0.05003279
以下、キヤノンが事業活動を行うとで重要度が高いと判断したステークホルダーと、〇年に実施した具体的な対話の事例について紹介します	0.01799861
新型コロナウイルス感染症と戦う知的財産宣言に発起人として参画	0.032436643
キヤノンは、京都大学の松田文彦教授らとともに、緊急事態宣言下において即座に各社へ積極的な働きかけを行い、〇の企業・大学を発起人とする「COVID-〇と戦う知財宣言」を発表しました	0.07826996
この宣言は、新型コロナウイルス感染症のまん延終結を唯一の目的とした開発、製造などの行為に対して、保有する知的財産権を行使しないことを宣言するものです	0.049518492
キヤノンは、多くの企業が参画することができるように、各社の事情にあわせてカスタマイズできる宣言書のひな型を準備しました	0.017014954
この宣言は世界に先駆けた日本発の取り組みで、世界知的所有権機関(WIPO)日本事務局や経団連などの協賛や後援を得て、多くの企業・研究機関が参画を表明しています	0.01951783
(〇年〇月時点:宣言者数〇)	0.017365638
復興支援活動	0.07842574
キヤノンは、福島県において原発事故の影響を受けた方々のコミュニティづくりを支援する「福島コミュニティサポート」を〇年から継続しています	0.07740967
この活動は、社員が講師となり写真教室や撮影会、交流会などを行うことによって、写真を通して参加者同士がふれあえる場を提供しています	0.06600267
これまで、仮設住宅や復興公営住宅などで暮らす方、避難指示が解除され故郷に戻られた方などを対象に〇回実施し、〇人の方と現地で交流を行ってきました	0.068249885
〇年は新型コロナウイルス感染防止のため、現地を訪れることを控え、キヤノン下丸子本社と福島県双葉郡富岡町とをオンラインでつなぎ、富岡町社会福祉協議会と協業で写真教室と交流会を実施しました	0.05724514
オンライン写真教室の様子	0.0580085
〇CanonSustainabilityReport〇	0.016551314

図 2. 統合報告書のページを提案モデル 3 に入力したときの Attention による文の注目度の例.
正解ラベル[0,1,1], 予測ラベル[0,1,1]

処理を行う必要があると考えAttentionを採用したが，本タスクにおいてAttentionが有効なのかを調査した．図 2 に，統合報告書のページが入力されたときの，Attention による文の重みを示す．図 2 は，復興支援活動(S)とステークホルダーエンゲージメント(G)に関連するページを入力した際の文の注目度である．「復興支援活動」や「キヤノンは、さまざまなステークホルダーに対して...」のような重要な文と,「門と協力し迅速に対応します」や「基本的な考え方」のような分類の手がかりにならないような文を正確に判別できており，このページに対して G,S に関連があると正しく推定できている．この例は，Attention を用いることで，途切れた文が含まれるようなノイズの多いテキストデータでもある程度正確に文の重要性を判別することができることを示しており，Attention により提案モデル 3 は高い再現率を得たと考える．また，提案モデル 3 のようなアーキテクチャは統合報告書だけでなく，他の崩れたテキストデータを対象にした文書やページ単位の分類タスクにおいても有効と考える．

6. まとめ

本研究では，企業の ESG 情報の開示媒体の 1 つである統合報告書から，ESG 関連ページを自動で推定する手法を提案した．本タスクでは扱うデータは形式の崩れたテキストデータであることを考慮し，複数の機械学習モデルを提案し，適切なモデルについて考察を行った．その結果，ESG の関連性推定の正確性を優先する場合は単語の出現情報を利用したモデル，ESG 情報の網羅性を優先する場合は，文の文脈情報を利用したモデルが適切であることを示した．

今後は，学習データの自動生成手法を見直し，精度を改善することで分類モデルの性能向上を目指す．

参考文献

[1] Global Sustainable Investment Alliance: GLOBAL SUSTAINABLE INVESTMENT REVIEW 2021. http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf

[2] 年金積立金管理運用独立行政法人（GPIF）: 2020 年度 ESG 活動報告. https://www.gpif.go.jp/investment/GPIF_ESGReport_FY2020_J.pdf

[3] 一般社団法人生命保険協会: 企業価値向上に向けた取り組みに関するアンケート集計結果一覧（2020 年度版）企業様向けアンケート. https://www.seiho.or.jp/info/news/2021/pdf/20210416_4-3.pdf

[4] 土橋諒太，中田和秀: BERT を用いた有価証券報告書からの ESG 関連文抽出，第 26 回 金融情報学研究会，2021.

[5] 中尾悠利子，石野亜耶，岡田斎: ニューラルネットワークによるサステナビリティ情報のテキスト分析――経営トップメッセージの環境・社会記述分析への適用，企業と社会フォーラム学会誌，第 8 号，pp.57-72，2019.

[6] 高野海斗，酒井浩之，中川慧: 学習データの自動生成による深層学習を用いた株主招集通知の重要ページ抽出，人工知能学会論文誌 2021 年 36 巻 1 号 p.WI2-G_1-19.

[7] 酒井浩之，松下和暉，北島良三: 学習データの自動生成による決算短信からの業績要因文の抽出，日本知能情報ファジィ学会誌，Vol.31，No.2，pp.653‐661 (2019).

[8] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, Eduard Hovy: Hierarchical Attention Networks for Document Classification, Proceedings of NAACL-HLT 2016, pages 1480–1489.

PDF 形式の金融文書における 項目-数値間関係を考慮したテーブル情報抽出

Table Information Extraction from Financial PDF Documents Considering Item-Value Relation

青野有華 [1] * 市川幸史 [1] 近藤浩史 [1] 加藤淳也 [1]
Yuka Aono[1] Koji Ichikawa[1] Hirofumi Kondo[1] Junya Kato[1]

[1] 株式会社日本総合研究所
[1] The Japan Research Institute, Limited

Abstract: 有価証券報告書などの金融文書において，重要な情報はテーブル形式で記載されることもあり，テーブル内の情報抽出は金融データの更なる利活用に向けて重要な役割を果たすと期待される．しかし，企業が共通して開示する文書であっても，企業によってテーブル形式が異なることや，情報抽出の難しい PDF 形式で開示される文書も存在することから，現状テーブル情報が抽出され，十分に活用されているとは言い難い．そこで本研究では，PDF 形式で開示されている日本語金融文書内のテーブルからの情報抽出を試みた．我々の手法では，PDF 内の罫線情報を利用しテーブル領域およびテーブル内セルを抽出した．その上でセル内での改行とセルの区切りを区別するために，セル内項目情報および数値情報に着目した BERT ベースの分割判定モデルを構築した．実験では，2 種類の PDF 形式の金融文書に含まれるテーブルを対象とした性能評価実験を行い，我々の提案手法が優れた性能を発揮することを確認した．

1 はじめに

上場企業は有価証券報告書や四半期報告書など様々な文書を開示している．これらに加え近年では，SDGs に対する社会的な意識の高まりなどもあり，従来とは異なる情報開示の流れもある．金融庁は「ディスクロージャーワーキング・グループ」[1]という審議会を開き，企業情報の開示や提供のあり方を述べている．また，新しい開示項目として，新型コロナウイルス感染症や ESG に関する開示例を挙げた「記述情報の開示の好事例集 2020」[2]を発行している．このような官庁の動きもあり，今後は企業による情報開示は更なる広がりをみせると考えられる．

開示情報は，企業の評価や株式投資の意思決定において重要な情報源である．一方，情報開示の広がりに伴い，多くの情報から必要な情報のみを読み取り，評価することは時間やコストを多く必要とする．

この課題に対応するため，近年では機械学習や深層学習を活用した金融文書のテキストマイニングの研究がなされている (例えば [1])．また，企業によって文書の形式は異なり，それらを一覧化，データベース化することは，金融文書に対するテキストマイニング研究分野の最終的な目標の一つともいえる．近年の研究傾向として，金融文書内のテキスト情報に対しての研究は多くなされている（例えば [2]）．一方で，金融文書にはテーブル形式で記載された情報もあるが，金融文書を対象としてテーブルからの情報抽出に重点をおいた研究は少ない．

そこで本研究では金融文書内のテーブル情報を正確に抽出することを目標とする．一般に有価証券報告書などは XBRL 形式での開示がなされているが，PDF 形式でのみ開示されている文書もある．そのため本研究では特に，金融文書の中でも，有価証券報告書に添付される「株主総会招集通知」などの，PDF 形式の文書を対象とする．

PDF からのテーブル情報の抽出には主に 2 点の課題がある．1 つめの課題は，罫線やテキストの位置関係のみから，適切にテーブルを検出し，構造を把握することである．この課題に対して本研究では，PDF 内の線分情報と文字情報を用いて解決を試みた．

2 つめの課題は，テーブルのセル内に複数の情報が含まれている場合の分割である．この場合，罫線を区切りとした単純なセル抽出では，これらが 1 つの情報として抽出される (図 1)．この状態では項目 (例えば科目) と数値 (例えば期末残高) の関連が不明確であるため，分析を行う際に特定の科目に関する合計を算出することや，項目のフィルタ検索をすることが難しい．そこで本研究では，BERT [3] を用いて，1 セル内に複数の情報を含むセルを，意味のある項目単位に分割することを試みた．その際に，金融文書のテーブルでは「科目」といった列 (項目列) に対して，「残高」列 (数値列) などの数値情報が対応していることが多いことから，数値情報の数を利用してラベル付きデータの構築や分割数の制限を行った．

提案手法の評価には株主総会招集通知にある「関連当事者との取引に関する注記」にあるテーブルと有価証券報告書にある「大株主の状況」にあるテーブルを対象とした[3]．その結果，我々の手法はそれぞれ F 値 0.995，0.981 を達成し，既存のオープンソースのソフ

*連絡先: aono.yuka@jri.co.jp
[1] https://www.fsa.go.jp/singi/singi_kinyu/tosin/20180628.html
[2] https://www.fsa.go.jp/news/r2/singi/20201106-3.html
[3]「大株主の状況」は XBRL 形式のデータが存在するが，実験のため PDF 形式に変換し評価を行った．

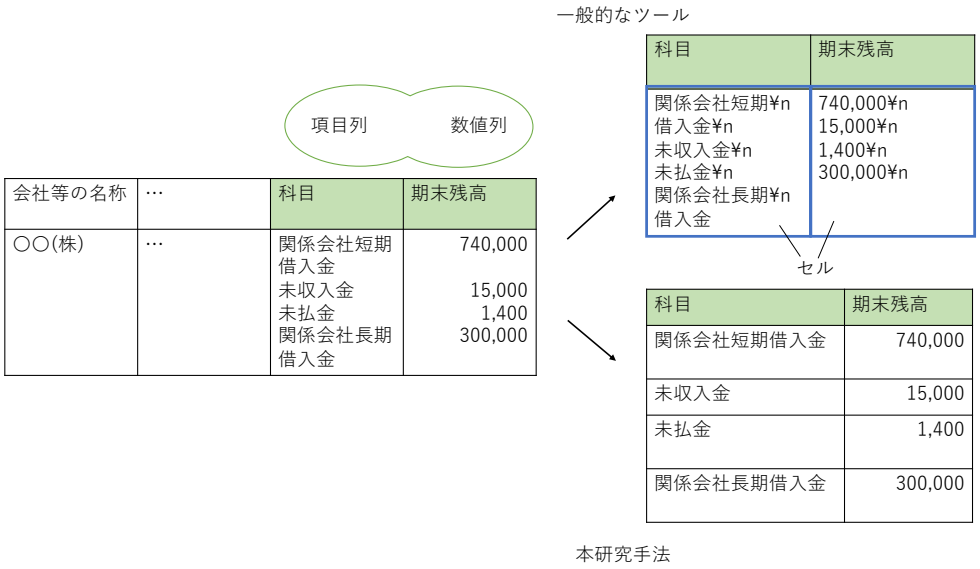


図 1: 本研究で対象とする抽出例

トウェアを利用した場合に比べてそれぞれ 0.35，0.38 ポイントの性能改善を確認した．

以下，2 章では問題設定を述べ， 3 章では提案手法について，4 章では実験，5 章では結論，6 章では今後の課題を述べる．

2 問題設定

本研究では PDF 形式で開示された文書内のテーブルからの情報抽出を行う．具体的には，金融文書内の，あるセクションに企業間で共通して記載されるテーブルのうち，「項目」列とそれに紐づく同一テーブル内の「数値」列を抽出し，項目-数値間が 1 対 1 対応した情報を取得する．したがって，前提として以下の情報が与えられることを想定する．

1. 抽出対象のテーブルのおおよその位置（セクション名など）
2. 「項目」列の列名
3. 「数値」列の列名

分かり易さのため, 図 1 を使って具体例を挙げると,「項目」列の列名とは「科目」を指し,「数値」列の列名とは「期末残高」を指す．なお，本問題設定は企業評価等の実業務を想定して定めた．実業務では意味のある数値を企業間で比較する場面が多くあり，テーブル全体の情報よりも，一部の項目と対応する数値を抽出したいというケースが多いと想定する．この想定において，本問題設定は有効と考える．

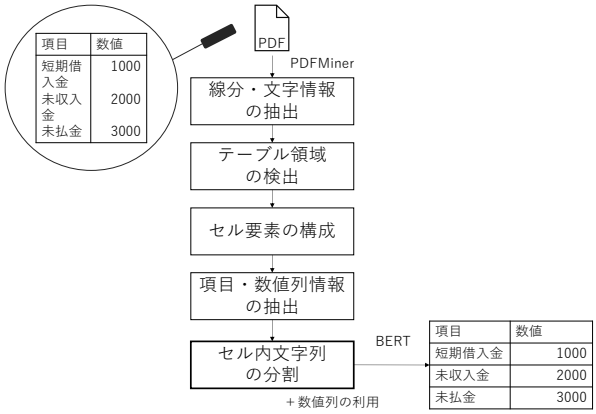


図 2: 提案手法の流れ

3 提案手法

提案手法では，図 2 に示す手順に従いテーブル内情報の抽出を行う．まず，PDF 文書を入力として，PDF 文書の線分情報と文字情報からテーブル領域を検出する．その後，テーブル領域内のデータを抽出するために，セル・列の構成や，機械学習モデルを用いたセル内文字列の分割を行う．最終的に，図 1 で示したような，入力として与えた「項目」列の列名および「数値」列の列名をもつ，2 列のテーブルデータを出力する．以下で，処理の詳細を紹介する．

3.1 線分・文字情報の抽出

第一ステップとして, PDF 文書から, 文書に含まれる線分情報と文字情報の抽出を行う．本研究では PDF 構成要素を PDFMiner[4]を用いて抽出し，抽出された線分要素と文字要素をテーブル情報の抽出のために用いる．この時, セクション名によるキーワードマッチングなどにより, 抽出対象のテーブルが含まれるおおよそのページ範囲は特定できていることを想定とする．以下, 線分要素を，その線分がもつ 2 端点の座標 $(x_1, y_1), (x_2, y_2)$ により

$$l = (x_1, y_1, x_2, y_2) \tag{1}$$

で表す．また文字要素を，その文字要素が存在する領域および文字により

$$c = (b, t) \tag{2}$$

で表す．ただしここで $b = (x_1, y_1, x_2, y_2)$ は文字が存在する左下の点座標 (x_1, y_1) および右上の点座標 (x_2, y_2) により形成される矩形領域を，また t は領域 b 内に含まれる文字を表す．抽出ツールにより，PDF の各ページごとに以下のような線分要素の集合 L，文字要素の集合 C が得られるとする:

$$L = \{l_1, l_2, \dots\}, C = \{c_1, c_2, \dots\}. \tag{3}$$

ただし，本研究では線分要素 l は水平あるいは垂直な線分のみを対象とし，それ以外のものは無視する．また要素に付随する文字フォントや線分の太さといったスタイル情報は使用しない.

3.2 テーブル領域の検出

次に，上記の線分要素集合 L を用いてテーブル領域の抽出を行う．一般にテーブルの構造は縦横の線分で構成されており，各線分は外枠または内罫線のいずれかを表している．そこで，接点あるいは交点のある線分同士を同一集合としたクラスターを作成し，これに基づきテーブル領域の検出を行う．この時，集合内の少なくとも 1 線分と接点あるいは交点が存在する線分もまた同一集合とする．例えば，線分 l_1 と線分 l_2，線分 l_2 と線分 l_3 で接点あるいは交点がある場合，線分 l_1, l_2, l_3 は同一クラスターとなる．この時，各クラスターごとにクラスター内の線分要素の座標の最小・最大値を用いて $(x_{\min}, y_{\min}), (x_{\max}, y_{\max})$ をテーブルが存在する矩形領域 b^{table} の左下および右上の座標として検出する[5]．さらに，クラスターに含まれる線分の数 n とクラスターが作る矩形領域 b^{table} の縦横の長さについてそれぞれに下限

$$n_{\text{lower}} \leq n\,, \tag{4}$$
$$w_{\text{lower}} \leq x_{\max} - x_{\min}\,, \tag{5}$$
$$h_{\text{lower}} \leq y_{\max} - y_{\min} \tag{6}$$

を設定し，よりテーブルらしい矩形領域のみを抽出する．ここで $n_{\text{lower}}, w_{\text{lower}}, h_{\text{lower}}$ は事前に定めた閾値である．この操作により，我々は各テーブル領域 b^{table} 内の線分要素の集合

$$L_{\text{table}} = \{l | l \subset b^{\text{table}}, l \in L\} \tag{7}$$

および，文字要素の集合

$$C_{\text{table}} = \{c | c \subset b^{\text{table}}, c \in C\} \tag{8}$$

を得ることができる．ただしここで $a \subset b$ は要素 a が領域 b 内に存在することを表している[6].

3.3 セル要素の構成

次に, 得られたテーブル領域内の線分要素集合 L_{table} および文字要素集合 C_{table} から，テーブルの各セル領域の検出およびセルに内包される文字列を構成し，セル要素を取得する.

まず, テーブル領域内の線分情報 L_{table} により, テーブル中の線分によって囲われた矩形領域（セル領域）b^{cell} を検出する．これは文字集合 C_{table} の各要素 c の矩形領域 b の中心から最も距離の近い上下左右の線分を特定することで行われる．すなわち，特定された上下左右の線分の交点により形成される矩形領域をセル領域 b^{cell} として検出する．この操作により，テーブル領域 b^{table} 内に存在するセル領域の集合 $\{b^{\text{cell}} | b^{\text{cell}} \subset b^{\text{table}}\}$ および，各セル領域内の文字要素の集合 $C_{\text{cell}} = \{c | c \subset b^{\text{cell}}, c \in C_{\text{table}}\}$ を得ることができる.

その後，セル領域内の文字列 t^{cell} を C_{cell} を用いて以下の手順により構成する: まず文字要素 $c \in C_{\text{cell}}$ の矩形領域中心の y 座標 y_c から，セル内に存在する行ごとの文字列を構成する．具体的には，2 つの文字要素 $c, c' \in C_{\text{cell}}$ について，$|y_c - y_{c'}| < \delta$ ならば, c, c' を同一行に存在する文字とみなす．ここで δ はある閾値である．そして，得られた同一行とみなされた文字要素の集合に対しそれぞれ x 座標の小さい順にソートし，1 行分の文字列とみなすことで，文字列の集合 $\{s_1, s_2, \dots\}$ を構成する.

最後に，各文字列をそれぞれの y 座標の大きい順にソートし，文字列の間に [SEP] トークンを挿入した上で結合し，これをセル要素内の文字列とする．この操作により我々はセル要素

$$e^{\text{cell}} = (b^{\text{cell}}, t^{\text{cell}}) \tag{9}$$

を得る．ただしここでの文字列 t^{cell} は

$$t^{\text{cell}} = s_1[\text{SEP}]s_2[\text{SEP}]s_3 \dots \tag{10}$$

のように表される文字列である.

3.4 項目・数値列情報の抽出

次に, テーブル領域内のセル要素の集合 $\{e^{\text{cell}} | e^{\text{cell}} \subset b^{\text{table}}\}$ および線分要素の集合 L_{table} を用いて，列成分を構成する．これは集合 L_{table} から，垂直な線分要素

[4] https://github.com/pdfminer/pdfminer.six

[5] PDF ではページの左下を起点とし，右上ほど x, y 座標の値が大きくなる.

[6] より厳密には，この包含関係 $\subset$ はある小さい範囲内でのはみ出しを許容することにする．また，PDF によってはページ全体を囲う不可視の線分が抽出されることがあり，上記のテーブル検出方法においてしばしばページ全体がテーブル領域として取得される場合がある．これを取り除くため，実験では検出された 2 つのテーブル領域が $b \subset b'$ となっている場合，b' を除く処理を行っている.

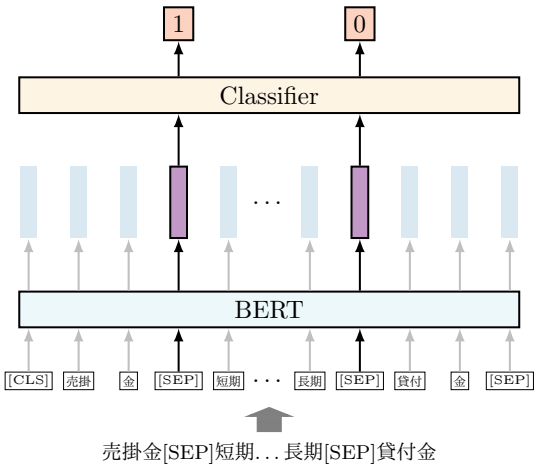


図 3: セル内文字列分割判定モデル概略図

の x 座標を抽出することによって行われる．すなわち，全ての垂直な線分要素 $l_{\text{vertical}} \in L_{\text{table}}$ の x 座標情報から集合 $X_{\text{table}} = \{x_0, x_1, \ldots | x_0 < x_1 < \ldots\}$ を取得した後，$x_{e^{\text{cell}}} \subset (x_i, x_{i+1})$ を満たす e^{cell} を i カラム目の要素とする．ただしここで $x_{e^{\text{cell}}}$ はセル要素 e^{cell} が x について占める領域を表す．この操作により，テーブル中の列要素の集合

$$\{\text{COL}_1, \text{COL}_2, \ldots\} \tag{11}$$

を得る．ただしここで COL_i は

$$\text{COL}_i = \{e^{\text{cell}} | x_{e^{\text{cell}}} \subset (x_i, x_{i+1}), e^{\text{cell}} \subset b^{\text{table}}\} \tag{12}$$

のように表されるセル要素の集合である．

「項目」列および「数値」列情報の抽出は以下のように行う: 上記の手法で得られたテーブル中の各列のうち，列中のセル要素内文字列 t^{cell} に ([SEP] トークンを除いた後) 設定したキーワードがマッチするものを抽出する．本研究においては，我々は項目列名のキーワードおよび数値列名のキーワードにより，2 つのカラム $\text{COL}_{\text{項目}}$ と $\text{COL}_{\text{数値}}$ を取得する．また，ここで我々は項目列と数値列が紐づいている状況を想定する．すなわち，$\text{COL}_{\text{項目}}$ と $\text{COL}_{\text{数値}}$ に属するそれぞれのセル要素について，その中心位置が同一の y 座標をもつものが 1 対 1 対応している状況を想定する．この状況においては，項目列中のあるセル要素内の文字列 $t^{\text{cell}}_{\text{項目}}$ に対して，数値列中の 1 つのセル要素内の文字列 $t^{\text{cell}}_{\text{数値}}$ を 1 対 1 で紐付けすることができる．これにより，各行 $(t^{\text{cell}}_{\text{項目}}, t^{\text{cell}}_{\text{数値}})$ を値にもつ 2 列のテーブルを構成することができる．

3.5 セル内文字列の分割

最後に，得られたテーブルの各行 $(t^{\text{cell}}_{\text{項目}}, t^{\text{cell}}_{\text{数値}})$ に対して分割処理を行う．3.4 節までの手法で抽出された文字列のペア $(t^{\text{cell}}_{\text{項目}}, t^{\text{cell}}_{\text{数値}})$ には，罫線の省略 (すなわち，結合セルの存在) により，本来複数の行に属すべき文字列が結合した状態で抽出されている場合がある (図 1 右上参照)．正確な情報抽出のためにはこれを適切に分割する必要があるが，特に $t^{\text{cell}}_{\text{項目}}$ においては，2 つの理由により分割を機械的に行うことが難しい．第一に，図 1 に示すように，抽出された文字列が折り返しを含んでいる場合があることである．このような状況では，単純な改行記号による分割は項目を過剰に出力してしまうことになる．第二に，行間のマージンが，PDF によって異なりうる点である．例えばマージンサイズを基準とした分割判定では，マージンの閾値を決める必要があるが，この数値を適切に決定することは難しい．また特に，図 1 に示すように折り返しによって生じるマージンと行間のマージンが等しい場合も存在し，マージンのみによってはこれを適切に分割することができない．上記の問題に対処するため，我々は項目の意味内容，および項目に紐づく数値情報を使用した機械学習ベースの分割手法を採用する．

項目セル内文字列の分割に取り組むにあたって，我々は問題を以下のような [SEP] トークンの分類タスクとみなす．すなわち，$t^{\text{cell}}_{\text{項目}}$ を [SEP] トークン部分で区切った系列を

$$X = (s_1, [\text{SEP}], s_2, [\text{SEP}], \ldots, s_n) \tag{13}$$

と書いたとき，X 中の各 [SEP] トークンに対し分割・結合のラベルを適切に割り当てることを目的とする．

上記の問題に対し，我々は機械学習ベースの分類モデルを構築することで解決を目指す．すなわち，各 s_i について tokenizer を通して得られるトークン $t_{i1}, t_{i2}, \ldots$ と [SEP] トークンによる系列

$$S = (t_{11}, t_{12}, \ldots, [\text{SEP}], t_{21}, t_{22}, \ldots, [\text{SEP}], \ldots) \tag{14}$$

を入力に，各 [SEP] トークンに対し結合・分割の 2 値分類を行うモデルを，教師付き学習により構築する[7]．

本研究ではこのセル内文字列の分割判定モデルに図 3 に示すような BERT をベースとしたアーキテクチャを使用する．すなわち，tokenizer により得られた系列 S を BERT に入力し，各トークンの埋め込み表現を得た後，この埋め込み表現を分類器を通し [SEP] トークンのラベル確率を出力する．このとき，モデルは人手または後述する手法により自動でアノテーションされたラベル付きデータを使用し fine-tuning を行う．

3.6 数値列の利用

多くの場合，項目列中の文字列 $t^{\text{cell}}_{\text{項目}}$ の分割数は，それに紐づいている数値列中の文字列 $t^{\text{cell}}_{\text{数値}}$ から推察することができる．なぜなら，この $t^{\text{cell}}_{\text{数値}}$ 内の数値情報は通常 $t^{\text{cell}}_{\text{項目}}$ 内の各項目に 1 対 1 対応しており，また数値は折り返しを含まないゆえに単純な不要語の除去と [SEP] トークンのカウントによりその分割数 (および $t^{\text{cell}}_{\text{数値}}$ の正確な分割) を得ることができるためである．すなわち，我々は，期待される X の分割数 $n(X)$ を，X に紐づく数値列中の文字列 $t^{\text{cell}}_{\text{数値}}$ により

$$n(X) = \#_{[\text{SEP}]}(\text{PROCESS}(t^{\text{cell}}_{\text{数値}})) \tag{15}$$

[7] 文字列 s_i を tokenize する際，[SEP] トークンの位置によっては不適切な分割やモデルの語彙外のトークンが生じうるが，本研究の実験においてこの影響は軽微であった．より厳密な取り扱いは将来の研究課題とする．

のように取得する．ただしここで $\#_{[\mathrm{SEP}]}(t)$ は文字列 t 中の [SEP] トークンの数であり，PROCESS(t) は文字列 t に対する不要語の除去などの簡単な処理を表す．

我々は，この分割数の情報を弱いラベル付きデータの構築および上記の分割判定モデルの分割数への制約に利用し，アノテーションコストの削減および性能の向上を目指す．

弱いラベル付きデータの自動構築

3.5 節で導入したモデルではラベル付きデータにより fine-tuning を行う必要があるが，このようなラベル付きデータを十分な量用意することが難しい場合もある．そこで我々は数値情報 $t^{\mathrm{cell}}_{\text{数値}}$ より得られる分割数から，自明にラベルが定まるサンプルを教師データとして利用する手法を提案する．すなわち，X を抽出された文字列，$n(X)$ を $t^{\mathrm{cell}}_{\text{数値}}$ より示唆される分割数としたとき，自明にラベルが定まるサンプルとは X が以下の 2 通りのどちらかを満たす場合である: (1) $n(X) = 0$, (2) $n(X) = \#_{[\mathrm{SEP}]}(X)$．ここで (1) の場合は，$X$ 中の全ての [SEP] トークンに結合ラベルを，また，(2) の場合は，X 中の全ての [SEP] トークンに分割ラベルを付与することができる．

分割数への制約

上記の分割判定モデルでは，意味内容によるセルの分割を行うことができるが，現れる文字列が出現頻度の少ない場合や人間でも判断の難しいケースの場合は誤った分割となりうる．また上記の弱いラベル付きデータにより fine-tuning を行った場合，モデルは与えられた文字列に対し文字列内の全ての [SEP] トークンに同一のラベルを付与する傾向が生じ，正確な分割が得られにくい．さらに，モデルが予測した分割が $t^{\mathrm{cell}}_{\text{数値}}$ の分割数と不一致である場合，出力された項目と数値情報を 1 対 1 に対応づけることができない．

そこで，提案手法では数値列から得られる分割数 $n(X)$ を X 内の合計分割数とみなし，この制限の元で適切な分割ラベルを付与する．具体的には，モデルによる各 [SEP] $\in X$ の分割ラベルへのスコアのうち上位 $n(X)$ については分割ラベルを，その他は結合ラベルを付与することでこれを実現する．これにより，より正確なラベル付けや上述した同一ラベル出力の抑制に加え，分割された項目文字列と数値文字列を 1 対 1 で対応づけできるようになり，より正確なテーブル情報抽出ができると期待される．

4 実験

4.1 データセット

本研究では，表 1 に記載するように，開示されている 2 つの金融系 PDF 文書内の項目を対象とし，データセットを作成した (関係当事者データセット，大株主データセット)．その際，2020 年 1 月 1 日から 2020 年 12 月 31 日までの期間に開示された文書をモデル学習用データセット，2021 年 1 月 1 日から 2021 年 6 月 30 日までの期間を評価用データセットの作成に使用した．評価用データセットでは上記期間に開示された PDF 文書の一部をサンプルし，文書内で，表 1 記載のセクション中に存在する項目列および数値列中の値を人手で抽出する形で Ground-truth を作成した．このアノテーションは著名 4 名でそれぞれ別の PDF 文書群に対して行った．また，簡単のため，ページを跨ぐようなテーブルについては，項目名・数値名が存在するページ部分のみを評価対象とした．

モデル学習用データセットは 3.6 節 で述べた自動構築手法に加え，人手で [SEP] トークンにラベルをアノテーションしたデータセットも用意した．人手アノテーションは上記期間内の PDF で関係当事者データセットとして 2000 サンプル ([SEP] トークン数 4829)，大株主データセットとして 800 サンプル ([SEP] トークン数 3180) について行った．また，同期間内の PDF で自動構築したデータセットのサンプル数はそれぞれ 1483 サンプル ([SEP] トークン数 2772)，3225 サンプル ([SEP] トークン数 8606) であった．実験では，自動構築したデータセットにより fine-tuning したモデルと，人手アノテーションしたデータセットにより fine-tuning したモデルの性能をそれぞれ評価した．

4.2 評価指標

実験の評価指標として，(マイクロ)Precision，Recall, F1 スコアを採用した．すなわち，ある PDF i に対し，それがもつ真の項目-数値のペアの集合を T_i，また手法により抽出された項目-数値のペアの集合を P_i としたとき，Precision，Recall を以下のように定義した．

$$\mathrm{Precision} = \frac{|\sum_i T_i \cap P_i|}{|\sum_i P_i|}, \tag{16}$$

$$\mathrm{Recall} = \frac{|\sum_i T_i \cap P_i|}{|\sum_i T_i|}. \tag{17}$$

ここで $|A|$ は集合 A の要素数を，$A \cap B$ は集合 A および B の共通要素の集合を表す． 共通要素数では，T, P の各要素に文字列正規化[8]を行った後，完全にペアが等しくなったもののみを共通要素として数えた．

4.3 ベースライン

我々の手法との比較対象として，PDF テーブル抽出を対象とした標準的なオープンソースである Tabula[9] と Camelot[10]を採用した．実験では，それぞれが出力するテーブルデータのうち，表 1 で指定した項目列・数値列名を列に含むテーブルを抽出し，そこに含まれる項目列・数値列をもって手法の予測とした．本実験では両者ともデフォルトのパラメータ設定を採用した．

4.4 実験設定の詳細

分割判定モデルで使用する BERT としては，東北大学が公開している日本語 Wikipedia による事前学習済みモデル (bert-base-japanese-whole-word-masking)[11] を使用した．BERT により得た各トークンごとの埋め込みを，Dropout 層および 1 層の全結合層を通した後，

[8] Unicode 正規化 (NFKC)，大文字小文字の統一を行った．
[9] https://github.com/chezou/tabula-py
[10] https://github.com/camelot-dev/camelot
[11] https://github.com/cl-tohoku/bert-japanese

表 1: データセット

データセット	文書	セクション	項目列名	数値列名	評価用項目数	学習用データ数 (人手 / 自動)
関連当事者	株主総会招集通知	関連当事者との取引に関する注記	科目	期末残高	1485	2000 / 1483
大株主	有価証券報告書	大株主の状況	氏名又は名称	所有株式数	1131	800 / 3225

表 2: 実験結果

手法	関連当事者			大株主		
	Precision	Recall	F1	Precision	Recall	F1
Tabula	0.540	0.115	0.190	0.410	0.682	0.512
Camelot	0.720	0.580	0.642	0.722	0.520	0.604
自動アノテーション	0.997	0.992	0.995	0.935	0.908	0.921
人手アノテーション	0.997	0.992	0.995	0.995	0.966	0.981

[SEP] トークンに対して softmax 関数を適用し，[SEP] トークンに対するラベル確率とした．学習はクロスエントロピー損失の最小化によって行い，最適化アルゴリズムは AdamW[4] を採用した．実験では学習率を 2×10^{-5} (学習率スケジューリングなし)，ドロップアウト率を 0.1，バッチサイズを 32，トークン最大長は 256 にそれぞれ設定した．学習用データセットは訓練用データセット, 検証用データセットとして 8:2 に分割し，訓練用データセットを使用してモデル訓練を行った．また学習のイテレーションは検証用データセットの分割・結合ラベルに対する正解率に基づき，early-stopping により打ち切った．

4.5 結果

表 2 に実験の結果を示す．関連当事者データセットでは, 人手アノテーションの場合, Precision が 0.997(評価用項目数:1481 行/1485 行)，Recall が 0.992(1481 行/1493 行)，F1 スコアは 0.995 となり，ベースラインを F1 スコアで 0.35 ポイント程度上回る結果となった．また，本データセットにおいては，自動アノテーションによるモデルにおいても，Precision が 0.997(1481 行/1485 行)，Recall が 0.992(1481 行/1493 行)，F1 スコアは 0.995 となり，人手アノテーションと同性能の結果が得られた．

大株主データセットでは，人手アノテーションの場合，Precision は 0.995 (1093 行/1098 行)，Recall は 0.966 (1093 行/1131 行)，F1 スコアは 0.981 となり，ベースラインを F1 スコアで 0.38 ポイント程度上回る結果となった．一方，自動アノテーションによるモデルにおいても，Precision が 0.935 (1027 行/1098 行)，Recall が 0.908 (1027 行/1131 行)，F1 スコアが 0.921 となり, 人手アノテーションと遜色ない結果が得られた．

4.6 考察

実験結果より，Tabula や Camelot などの一般的な PDF テーブル抽出ツールと比べて，数値-項目間の 1 対 1 対応を正解とした時の性能は大幅に改善された．これは，従来のツールでは難しかった適切なセルの分割がなされた結果と考えられ，金融系 PDF 文書における提案手法の有効性を示していると言える．

関連当事者データセットにおいて，適切な情報取得ができなかったテーブルは，図 4a のように 1 つの項目に対して複数の数値が紐づくようなテーブルや，反対に図 4b のように複数項目に対して 1 つの数値が紐づくようなテーブルであった．本手法では，1 項目に対して 1 数値が紐づくことを前提としているため，上記のようなテーブルは対応できなかった．

大株主データセットにおいては，特に自動アノテーションによる結果が関連当事者データセットよりも劣る性能となった．これは，自動アノテーションを行えたデータ中, 分割ラベルをもつデータが非常に少なかったことが原因であると考えられる (3225 サンプル中 85 サンプル)．実際，[SEP] トークンのラベルが自明でないサンプル ($0 < n(X) < \#_{[\mathrm{SEP}]}(X)$ なもの) に対する Precision，Recall はそれぞれ 0.736 (195 行/265 行)，0.714 (195 行/273 行) と低く[12]，BERT がこの自動アノテーションデータで十分に学習されているわけではないことが示唆された．この改善に向けては，例えばランダムに抽出した 2 つのサンプルを，分割ラベルを付与した [SEP] トークンにより結合し，データ増強を行うといった手法の導入が考えられる．このようなデータ増強手法のより詳細な検証については，今後の一つの発展の方向性としておく．

最後に，大株主データセットにおけるテーブル検出の失敗例を図 4c に示す．大株主データセットではページ下部にまでテーブルが存在することが比較的多くあったが，これが PDF 中に存在するその他の線分 (多くは不可視であり, ページを囲うように存在するもの．図 4c ではオレンジ色の線分で表現) と交差した結果，不適切なクラスターが作成される状況がしばしば生じた．この抑制には，線分要素のフィルタリングや，より洗練されたテーブル検出ルールの導入が必要であると考えられる．

5 結論

本研究では，PDF 形式のテーブルに対して,「項目」列とそれに紐づく同一テーブル内の「数値」列の関係を考慮した上で情報抽出を行った．これは，PDF の文

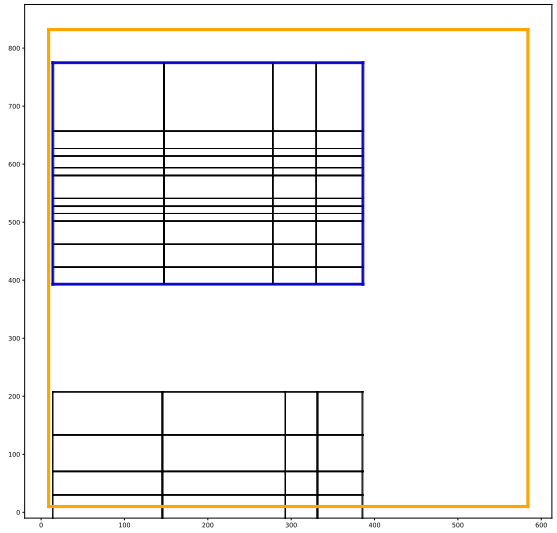
[12] 人手アノテーションの場合，それぞれ 0.985 (261 行/265 行)，0.956 (261 行/273 行) であった．

会社等の名称	…	取引金額(千円)	科目	期末残高(千円)
△△ (株)	…	- 3,300 2,00 5,000	1年内回収予定の関係会社長期貸付金 関係会社長期貸付金 未収入金	- 3,300 17,000 300

(a)

会社等の名称	…	取引金額(千円)	科目	期末残高(千円)
□□ (株)	…	27,000 70,000 40 200 35,000	売掛金 買掛金 短期貸付金 未収入金	2,000 600 35,000

(b)



(c)

図 4: 本手法で抽出できなかった例

書内から抽出した線分情報や文字情報を用いてテーブル領域を検出し，領域内のセル及び列の構成やセル内文字列の分割を行うことで実現した．セル内文字列の分割には BERT を用い，「数値」列を利用した分割数の制約を課すことでより正確な分割を行った．我々は以上の手法を 2 つの日本語金融文書により作成したデータセットにより評価し，結果，一般的な PDF テーブル抽出ツールと比べ大幅な性能の改善を確認した．

6 今後の課題

本研究のテーブル領域検出では，PDF 内の線分要素の交点や接点を利用しているため，罫線が大きく省略されているとき，検出に失敗する場合がある．今回使用した日本語金融文書データセットにおいては，このような大きな罫線の省略があるテーブルはほとんどなく，今回の手法でも高い精度でテーブル検出を行うことができた．しかし，文書によっては，罫線が大きく省略されたテーブルが割合の多くを占める場合もありうる．特に英語圏のテーブルはしばしば罫線が大きく省略されるため，検出できない状況が多く存在すると予想される．このような状況に対応するためには罫線に依らない表検出手法 (例えば Schreiber ら [5] のような画像からの検出手法) を実施することが必要であろう．

また，我々の手法で項目-数値間の 1 対 1 対応した一覧化を実現することができたが，会社間での比較は項目における表記ゆれや数値の単位の統一化をする必要がある．例えば，金融文書の場合，数値の単位は「百万円」や「千円」などと企業間で異なり，このままでは企業間で数値の合計を算出することはできない．これについては，今回のテーブル内の情報抽出に併せて，テーブル周辺の情報も抽出し，単位らしい文字列をキーワードマッチングすることで実現できると考えられる．

参考文献

[1] Liu, Z., Huang, D., Huang, K., Li, Z., and Zhao, J.: FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining, *IJCAI2020*, pp. 4513–4519 (2020)

[2] 高野 海斗, 酒井 浩之, 北島 良三：有価証券報告書からの事業セグメント付与された業績要因文・業績結果文の抽出, 人工知能学会論文誌,Vol. 34, No. 5, p.wd-A_1–22,(2019)

[3] Devlin, J., Chang, M. W., Lee, K., and Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding, *NAACL-HLT2019*, Vol. 1, pp. 4171–4186 (2019)

[4] Loshchilov, I., and Hutter, F.: Decoupled weight decay regularization, *ICLR2019*, (2019)

[5] Schreiber, S., Agne, S., Wolf, I., Dengel, A., and Ahmed, S.: Deepdesrt: Deep learning for detection and structure recognition of tables in document images, *ICDAR2017*, Vol. 1, pp. 1162–1167, IEEE (2017)

有価証券報告書におけるリスク階層構造分析

Analysis of risk hierarchy structure in securities reports

藤井 元雅 [1*] 坂地 泰紀 [2] 佐々木 一 [3] 増山 繁 [4]

Motomasa Fujii[1] Hiroki Sakaji[2] Hajime Sasaki[3] Shigeru Masuyama[4]

[1] 東京理科大学大学院経営学研究科
[1] Graduate School of Management, Tokyo University of Science
[2] 東京大学大学院工学系研究科
[2] Graduate School of Engineering, The University of Tokyo
[3] 東京大学未来ビジョン研究センター
[3] Institute for Future Initiatives, The University of Tokyo
[4] 東京理科大学経営学部
[4] School of Management, Tokyo University of Science

Abstract: In an environment of high uncertainty, risks in corporate management are changing daily. However, risks do not occur independently, but are often caused by a variety of events (risk factors), including other risks. We define this as causal relationships between risk factors. In addition, a single risk factor may have multiple consequences. For example, when a risk factor such as the Covid-19 epidemic occurred in one company, some departments increased sales due to increased IT investment, while others suffered losses as a result of reduced advertising due to worsening business conditions (risk factors of worsening business environment). Therefore, when actually understanding and assessing the risks faced by a company, it is necessary not only to understand the risk factors, but also to understand how they will ultimately affect business performances. In other words, it is important to stratify the risk factors. In this study, we propose a method to obtain the causal content of extracted risk sentences and to understand the risk factors using a hierarchical structure.

1. はじめに

昨今，環境・経済・社会がビジネスに与える影響は，多くの企業にとって無視できないリスク要因となっており，SDGs や ESG に貢献する活動を行わないこと自体がリスクになることも指摘されている[1,2]．また，COVID-19 の世界的流行に伴い事業継続計画(BCP)の内容が, 企業価値を評価する投資家から重視され始めており[3]，このような不確実性の高い経営環境の中で企業経営におけるリスクは日々変化している．このような事態に対処すべく，藤井ら[4]の研究では深層学習を用いて有価証券報告書に含まれるリスクについて述べた文(以下，リスク文)を抽出・分類する手法を提案した．これは不確実性の高いビジネス環境において，早期のリスク抽出からのリスク管理や企業価値評価に対する重要な視点を提供している．

しかし，このようなリスクは単独で発生するものではなく，他のリスクを含む様々な事象(リスク要因)が発生したことにより生じる場合が多い．このことをリスク要因間に因果関係があると定義する．さらには 1 つのリスク要因から複数の別の結果が生じる場合もある．例えばある IT 企業においては COVID-19 の流行というリスク要因が発生した際に, IT 投資の増加によって売上が上がる部門もあれば, 景況悪化に伴う広告出稿の抑制(経営環境の悪化というリスク要因)によって赤字に転落する部門もあった．これは，リスク情報を企業価値評価に反映させる上において重要な観点である．例えば「自然災害」をリスクとして抽出したとしても，それがどのように業績に影響を与えるのかを確認しなければ, 抽出した情報を企業評価に生かすことができない．あるいは逆に材料費の高騰(コストの上昇)というリスクがなぜ発生するのかを把握しなくては，同様の

* E-mail: 8621509@ed.tus.ac.jp

問題が生じてくる．つまり，実際に企業が直面しているリスクを把握，評価を行う際には，リスク要因単独を把握するだけでなく，あるリスク要因がまた別のリスク要因を発生させ，それが繰り返されたことで最終的に業績の変動が引き起こされるという一連の流れ(因果関係)を把握する必要がある(図 1)．すなわちこれらのリスク要因が波及していく様子を表すためにリスク要因の階層化を図ることが重要であると考えられる．加えて，Campbell et al.[5]はテキストによるリスク開示は，企業が晒されているリスク要因を特定する重要な情報源である点を指摘している．

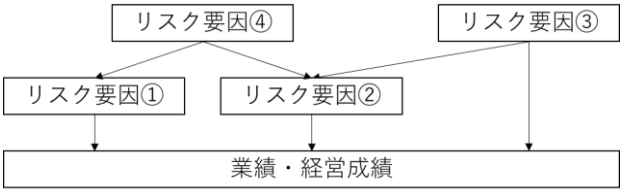


図 1. リスク階層構造の例

本研究では，財務関連書類内のテキストデータからリスク要因の階層関係を取得し，企業価値評価へ役立てることの前段として，実際に人手で作業を行った場合どのような関係が取得できるのか，また取得した情報は企業価値評価において有用であるかどうかを検証する．なお，データソースとしては金融商品取引法で規定された開示資料であり，法令により規格が統一されている有価証券報告書の「事業等のリスク」の項目を利用する．

2. 関連研究

テキストデータを対象とする機械学習等によって企業価値評価に貢献する研究は，様々な研究者が行っている．例えば，Huang et al.[6]や Chang et al.[7]は，テキストマイニングを用いた企業価値評価に関する関連研究を行い，その成果として投資判断を行うに際しての企業価値評価の精度向上を確認している．

また，リスクに焦点を当てた研究としては Bao et al.[8]や Dyer et al.[9]があり，いずれも企業のテキストによるリスク開示を分析し，リスク要因を発見する手法を提供している．また藤井ら[4]は，リスクを「行動したことにより獲得できる将来の結果の不確実性」と定義し，該当するリスクが内部要因と外部要因のどちらかによるものかといった複数の観点において，決定木の形で分類することで企業価値評価におけるリスク全てを網羅するリスクの定義と，キーワードなどの既存の手がかりを使わずに財務関連文書等のテキストデータからリスクを抽出及び分類する手法を提案した．手法としては，リスク文でよく使われる表現や単語のパターンを手掛かりにするパターンマッチング，SVM などの教師あり機械学習，BERT などの深層学習を用いた．その結果，分類及び抽出の両方において深層学習，特に BERT が効果的であることを特定している．この研究においては，リスク要因をそれぞれ並列に独立なものとして抽出及び分類をしているが，本研究においてはリスク間の関係性に焦点を当てている．

また，リスクの階層化の研究として，五十嵐ら[10]は決算短信から因果関係を抽出し，類似性を判断してそれらを接続することでネットワークを構築している．類似性の計算には，日本語 Wikipedia コーパスから作成した word2vec モデルを使用し，単語の重要度を表す idf 値の組み合わせに基づいた手法を用いている．また，因果関係の表現に極性辞書を用いて極性を与え，word2vec が検出できない同義語を併せて判定している．当該研究が一つの事象に対する連鎖的な関係に焦点を当てているのに対し，本研究においては，企業価値評価に用いるためにリスク要因全体を包括的に網羅する階層関係の作成を試みている．他に Wei et al.[11]は，Sentence Latent Dirichlet Allocation(SentLDA)というテキストマイニング手法を導入し，エネルギー産業のテキストによるリスク開示情報を分析することで，エネルギー企業のリスク要因を包括的に特定するとともに，抽出されたリスク要因の見出しを基に，最終的にエネルギー企業のリスクに影響を与える 66 のリスク要因を特定している．さらに米国のエネルギー企業 840 社を 9 つのエネルギーサブセクターに分けて，リスク要因の階層化を行い，各リスク要因の重要性を出現頻度に基づいて階層化して示している．本研究においては，出現頻度ではなく，リスク要因の因果関係という連鎖の関係性に基づいて階層化を試みている．

そして，階層化において重要な要素となる因果関係文の抽出においても様々な研究が行われている．例えば，坂地ら[12]は決算短信から原因・結果表現を自動的に抽出する手法の提案を行っており，具体的には，原因・結果を含む文を抽出した後に，原因・結果表現を抽出した．原因・結果表現抽出には，「～に伴う」「～を背景に」といった手がかり表現を含む構文情報を用いた Pattern を用いている．

本研究においては，有価証券報告書の「事業等のリスク」の項目において，リスク文の中から人手でリスク要因の因果関係を抽出したうえで，リスク階層構造の一部を作成し，その有用性を検討する．

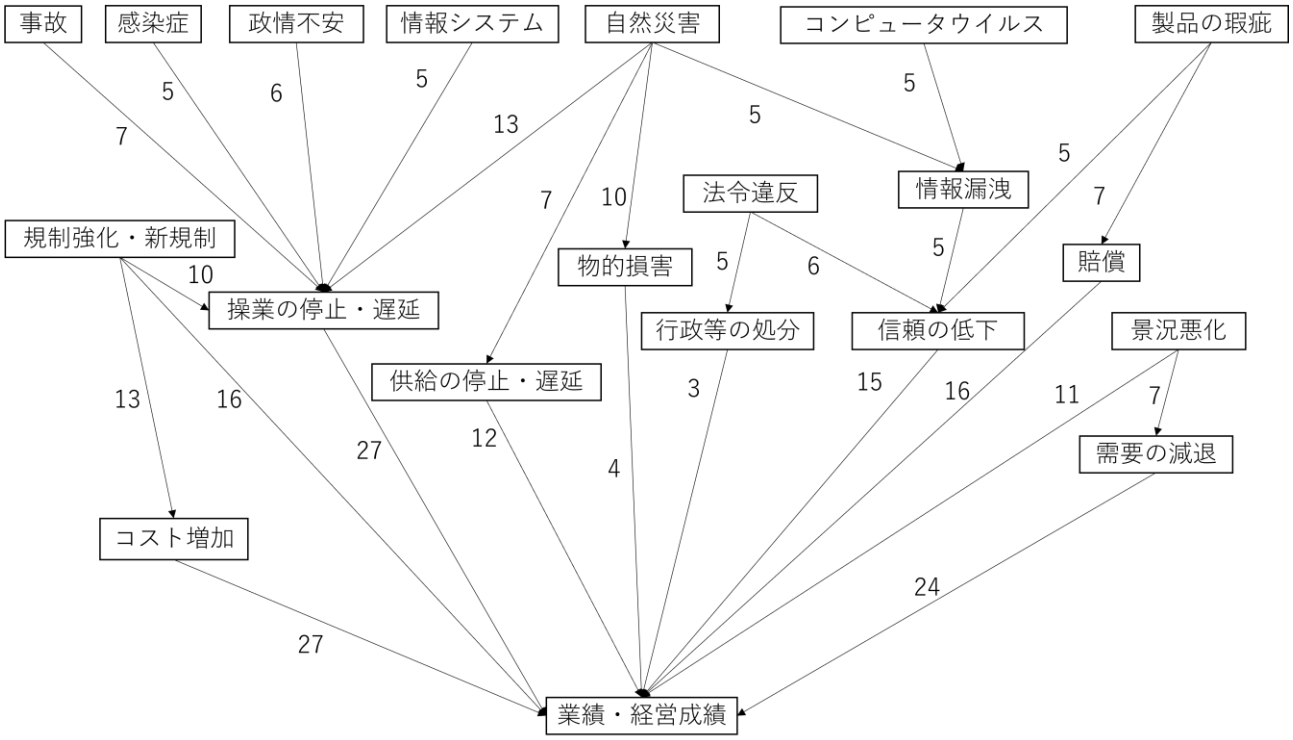


図 2. 抽出したリスク階層構造の一部

3. 分析手法

3.1. データ及びデータソース

評価データには，日経 225 に指定されている企業からランダムに 71 社を選択し，それら企業の 2018 年度に発行された有価証券報告書を用いた．有価証券報告書に対しては PDF ファイルをテキスト化し，「事業等のリスク」の項目を抽出したうえで，文単位に分割した．該当する文は計 5,007 文であり人手でタグを付与した結果，リスク文は 2,704 文，リスク文でない文は 2,303 文であった．本研究においては，前者のリスク文と判定された文をデータセットから 500 件を抽出して利用する．なお，リスクの定義については「行動したことにより獲得できる将来の結果の不確実性」としている．

3.2. 人手による因果関係抽出

リスク文への因果関係のタグ付けについては，坂地ら[12]を参考に，以下のように主語(s)，原因(b)，結果(r)の 3 つに分けて行った．なお，主語については必ずしも文中に含まれているとは限らない．具体例は以下の通りである．

①<s1>当社グループで使用する部材</s1>は<b1>経済環境の変化</b1>により<r1>入手困難になる</r1>可能性があります．
(日清紡ホールディングス株式会社)

②<b1>市販後に予期していなかった副作用が発生し</b1>た場合には，<r1>当社グループの経営成績及び財政状態等に重大な影響を及ぼす</r1>可能性があります．
(協和発酵キリン株式会社)

抽出した結果，500 のリスク文から 657 の因果関係を取得した．そして抽出した因果関係に対して人手でクラスタリングを行い，60 のリスク要因に分類をした．以下はその一部である．

感染症　規制強化・新規制　供給の停止・遅延
競争の激化　景況悪化　コスト増加　法令違反
コンピュータウイルス　自然災害　信頼の低下
需要の減退　政情不安　製品の瑕疵　情報漏洩
操業の停止・遅延　情報漏洩

4. 分析結果

前章においてタグ付けしたリスク要因の因果関係は全部で 392 種類確認でき，その中で 3 つ以上確認できた関係を表 1 に示す．また，その因果関係の一

表 1. リスク要因の因果関係例

原因	結果	数	原因	結果	数
コスト増	コスト増	3	事故	操業の停止・遅延	7
コスト増	業績・経営成績	27	事故	物的損害	4
コンピュータウイルス	情報システム	4	自然災害	供給の停止・遅延	7
コンピュータウイルス	情報漏洩	5	自然災害	情報システム	4
シェア低下	業績・経営成績	5	自然災害	情報漏洩	5
マーケティング	業績・経営成績	4	自然災害	操業の停止・遅延	13
為替変動	業績・経営成績	21	自然災害	物的損害	10
価格戦略の失敗	業績・経営成績	8	需要の減退	業績・経営成績	24
株価変動	業績・経営成績	5	需要の減退	原材料高騰	3
株価変動	保有資産の価値低下	4	需要の減退	需要の減退	6
感染症	操業の停止・遅延	5	需要の減退	売上の減少	4
環境汚染	操業の停止・遅延	3	収益性の低下	業績・経営成績	6
機会損失	業績・経営成績	4	行政等の処分	業績・経営成績	3
規制強化・新規制	コスト増	13	情報システム	業績・経営成績	5
規制強化・新規制	規制強化・新規制	3	情報システム	情報漏洩	3
規制強化・新規制	業績・経営成績	16	情報システム	操業の停止・遅延	5
規制強化・新規制	需要の減退	3	情報漏洩	業績・経営成績	8
規制強化・新規制	操業の停止・遅延	10	情報漏洩	信頼の低下	5
技術革新	業績・経営成績	3	信頼の低下	業績・経営成績	15
供給の停止・遅延	供給の停止・遅延	6	人材不足・流出	業績・経営成績	6
供給の停止・遅延	業績・経営成績	12	人材不足・流出	操業の停止・遅延	4
供給の停止・遅延	操業の停止・遅延	4	政情不安	供給の停止・遅延	4
競争激化	価格戦略の失敗	3	政情不安	業績・経営成績	3
競争激化	業績・経営成績	8	政情不安	操業の停止・遅延	6
競争激化	経営経営戦略の不発	5	製品の瑕疵	業績・経営成績	7
競争力低下	業績・経営成績	9	製品の瑕疵	信頼の低下	5
競争力低下	経営経営戦略の不発	3	製品の瑕疵	製品の瑕疵	3
金利変動	業績・経営成績	9	製品の瑕疵	賠償	7
景況悪化	供給の停止・遅延	4	税制	業績・経営成績	3
景況悪化	業績・経営成績	11	経営経営戦略の不発	業績・経営成績	16
景況悪化	需要の減退	7	訴訟	業績・経営成績	14
景況悪化	売上の減少	3	訴訟	賠償	4
研究開発	業績・経営成績	6	操業の停止・遅延	業績・経営成績	27
原材料高騰	コスト増	3	物的損害	業績・経営成績	4
原材料高騰	業績・経営成績	4	貸し倒れ	業績・経営成績	4
原材料高騰	収益性の低下	3	知的財産の侵害	業績・経営成績	7
特定顧客の経営状況	業績・経営成績	3	知的財産の侵害	賠償	4
製品の自主回収・リコール	業績・経営成績	5	投資の失敗	業績・経営成績	4
資金調達の失敗	業績・経営成績	3	売上の減少	業績・経営成績	3
保有資産の価値低下	業績・経営成績	14	賠償	業績・経営成績	16
保有資産の価値低下	保有資産の価値低下	3	価格戦略の失敗	業績・経営成績	3
事故	供給の停止・遅延	4	法令・コンプライアンス違反	行政等の処分	5
事故	情報漏洩	3	法令・コンプライアンス違反	信頼の低下	6
事故	信頼の低下	3	法令・コンプライアンス違反	操業の停止・遅延	3

部を基に作成したリスク階層構造を図 2 に示す．この図はリスク要因間の因果関係のうち，抽出数が多い関係の一部を抜粋したものであり，矢印の隣の数字は実際に抽出された関係の数である．一部を抽出したこと及び有価証券報告書の記述を基にしており企業が全てのリスク要因を必ず明示するとは限らないことから，ノードへの入力数と出力数の和が一致するとは限らない．例えば，コスト増加には法規制によるもの以外にも，原油価格高騰による燃料コストの増加や人材獲得競争の激化による人件費増加等の様々な要因が考えられる．以下に，これらの階層構造(因果関係)が具体的にどのようなリスク文から抽出されたものかを例示する．

③特に，当社グループの主力商品である 300mm ウェーハの加工工程を有する製造拠点が，上記の自然災害，事故，火災等に見舞われる場合，300mm ウェーハの製造・販売に支障を来たし，当社グループの事業展開，経営成績等に影響を及ぼす可能性があります．

(株式会社 SUMCO)

「自然災害」・「事故」
→「操業の停止・遅延」
→「業績・経営成績」

④しかしながら，重大な製品欠陥や製造物責任訴訟の提起といった事象が発生した場合，社会的信用の失墜を招き，顧客に対する補償などによって，当社グループの経営成績及び財務状況が影響を受ける可能性がある．

(昭和電工株式会社)

「製品の瑕疵」
→「信頼の低下」・「賠償」
→「業績・経営成績」

5. 考察

これらの内容が最終目的である企業価値評価に有用であるかどうかを検証する．例えば自然災害というリスク要因に焦点を当てた場合，このリスク要因が業績に影響を与えるには，物的損害を与える場合よりも，操業の停止や遅延について言及している(重視している)企業が多いことがわかる．また，自社の操業のみならず，供給ラインも同様に大きなリスク要因であることがわかる．つまり，自然災害発生時には自社への被害や操業への影響だけでなく，供給業者についても同じことが言えるということである．無論，現在示している内容はある程度推測が容易であり，新たな発見とは言い難い．しかし，リスク要因の因果関係数を測定することで各リスク要因の重要度をマクロ的な視点で推測することができていると言える．また，自然災害において情報漏洩が引き起こされると予測している企業も一部存在している．これはこの一部の企業が他企業に比べ異なる状況にある(ないし，視点を有している)ということであり，このようなミクロの分析においてもこれらの階層構造が新たな視点を提供することができると考えられる．

また,「新たな法規制が当社グループの業績に影響を与える可能性がある」といったリスク文と図に示した構造を照らし合わせることで，コスト増加というリスク要因が有価証券報告書に記載されていなかったとしても，そのリスク要因が当該企業に存在する可能性があることが推測できる．このように，リスク階層構造は抽出されたリスク要因を企業価値評価に生かすに際して，有用であると考えることができる．

6. まとめ

本研究においては，財務関連書類内のテキストデータからリスク要因の階層関係を取得し，企業価値評価へ役立てることの前段として，実際に人手で作業を行った場合どのような関係が取得できるのか，また取得した情報は企業価値評価において有用であるかどうかを検証した．結果として，抽出されたリスク要因の因果関係によって構成されたリスク階層構造の情報は企業価値評価に有用であるということを確認することができた．しかし今回は人手による作業であるためデータセットが少なくなっているが，実際に実用に足るデータを用意するには大量のデータを処理する必要があり，そのためには機械学習や深層学習が必須であると言える．

上記の研究を進めるにあたって考慮すべきはリスク文抽出，因果関係抽出の 2 点である．まず，リスク文抽出作業は財務関連書類のテキストデータから不要な文を削除するために必要である．これは藤井ら[4]が示したように人手でタグ付けをした後に双方向 LSTM[13]あるいは BERT[14]を用いて文の分類作業を行えば問題ないと考えられる．次に，因果関係抽出については坂地ら[12]が示した手がかり表現や構文パターンを基にリスク文に焦点を当てた手法に調整をする必要があると考えられる．

参考文献

[1] 経済産業省: SDGs 経営ガイド, 2019.
[2] Kim, MinChung, and YongHee Kim. "Corporate social responsibility and shareholder value of restaurant firms." International Journal of Hospitality Management 40 (2014): 120-129.
[3] Mukherjee, Mahua, et al. "Ecosystem-centric business continuity planning (eco-centric BCP): A post COVID19 new normal." Progress in Disaster Science 7 (2020): 100117.
[4] 藤井元雅, 坂地泰紀, 佐々木一, 増山繁, 有価証券報告書からのリスク文抽出の試み, 25 回人工知能学会 金融情報学研究会, pp.44-48, 2020.
[5] Campbell, John L., et al. "The information content of mandatory risk factor disclosures in corporate filings." Review of Accounting Studies 19.1 (2014): 396-455.
[6] Huang, Win-Bin, et al. "Value assessment of companies by using an enterprise value assessment system based on their public transfer specification." Information Processing & Management 57.5 (2020): 102254.
[7] Chang, Jun, et al. "Assessing dynamic qualities of investor sentiments for stock recommendation." Information Processing & Management 58.2 (2021): 102452.
[8] Bao, Yang, and Anindya Datta. "Simultaneously discovering and quantifying risk types from textual risk disclosures." Management Science 60.6 (2014): 1371-1391.
[9] Dyer, Travis, Mark Lang, and Lorien Stice-Lawrence. "The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation." Journal of Accounting and Economics 64.2-3 (2017): 221-245.
[10] 五十嵐光秋, et al. "極性を考慮したリスク発見に向けた因果関係ネットワークの構築." 人工知能学会全国大会論文集 第 33 回全国大会 (2019). 一般社団法人 人工知能学会, 2019.
[11] Wei, Lu, et al. "Developing a hierarchical system for energy corporate risk factors based on textual risk disclosures." Energy Economics 80 (2019): 452-460.
[12] 坂地泰紀, 酒井浩之, 増山繁, 決算短信 PDF からの原因・結果表現の抽出, 電子情報通信学会論文誌 D, Vol.J98-D, No.5, pp.811-822, 2015.
[13] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
[14] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).

特許ドメイン特化型 BERT による脱炭素関連特許技術の「見える化」

“Visualization” of decarbonization-related patent technologies by patent domain-specific BERT

前原義明 [1]　久々宇篤志 [1]　長部喜幸 [1]

Yoshiaki MAEHARA[1], Atsushi KUKU[1], and Yoshiyuki OSABE[1]

[1] 日本特許情報機構
[1]Japan Patent Information Organization

Abstract: We trained a classifier to determine whether a technology in the patent literature corresponds to a decarbonization-related technology or not, using BERT pre-trained on the patent documents and the corpus of decarbonization-related patents we collected. The trained BERT model was also used to conduct macro analysis of Japanese published patent gazettes applied after 2000 to visualize the decarbonization-related patent technologies in Japan.

１．はじめに

日本政府は 2050 年カーボンニュートラル実現を宣言し[１]，2030 年度の温暖化ガス排出削減目標として，46%削減（2013 年度比）することを目指し，さらに 50%の高みに向けて挑戦を続けるとの新たな方針を示した．

これを受け，経済産業省は「2050 年カーボンニュートラルに伴うグリーン成長戦略」を公表し[２]，この挑戦を「経済と環境の好循環」につなげるための産業政策として策定した．経済産業省は，エネルギー政策及びエネルギー需給の観点から,「エネルギー関連産業」「輸送・製造関連産業」「家庭・オフィス関連産業」において成長が期待される 14 分野を導き出し，2050 年に向けた高い目標を設定している．

2050 年カーボンニュートラルの実現は，エネルギー・産業部門の構造転換，大胆な投資によるイノベーションの創出といった取組を，産学官が一体となり, 大きく加速させることが必要である. 産業界は，これまでのビジネスモデルや戦略を根本的に変えていく必要がある一方で，もともと省エネ技術に長けている多くの日本企業にとって，新しい時代をリードしていくチャンスでもある．そして，目標の実現には，イノベーションに密接に関連する特許情報の整理・俯瞰が欠かせない．

しかし，脱炭素技術はカバーする技術範囲・産業分野が広く，また，通常の特許検索で使用されている特許分類には脱炭素技術に係る特別の分類コードが付与されておらず，脱炭素に関連する技術を網羅的に捉えることは困難であった．さらに，扱うデータ量が膨大なことから，自然言語処理技術を用いて特許明細書から直接脱炭素技術を判定できる手法の開発が望まれていたところである．

そこで，本研究では，欧州特許庁（EPO）で付与されている Y02（気候変動緩和技術），Y04（スマートグリッド関連技術）という共通特許分類（CPC）に着目し，当該分類が付与されている日本語特許公報を収集してコーパスを作成, このコーパスを利用して, 特許明細書の記載が Y02，Y04 の技術に包含されるか否かを判定する機械学習モデルを作成した．そして, この機械学習モデルを用いることによって, 2000 年以降に日本に出願された特許のマクロ解析を行い, 脱炭素関連特許技術の「見える化」を行った．

このマクロ解析の結果は，ESG 投資をはじめとする「脱炭素マネー」の行方を決める点においても，非常に重要な情報となり得る．

２．分析手法

２．１．解析モデル

機械学習モデルとしては BERT[３]を採用した．BERT は Bidirectional Encoder Representations from Transformers の略であり, 各コーパスにおいて, 少ないコーパスで高い精度を出すことができるものとし

て有名だが，特許分類タスクにおいて最も高い精度を出しているモデルでもある．[４][６]

また，BERT はドメイン特化によって，その性能が向上することが知られている[５]. そのため，本研究においては, 2018 年から 2019 年の日本国公開特許公報を用いて, BERT の事前学習を行った（ここで事前学習を行った特許ドメイン特化型 BERT のことを，以下「PatentBERT_japanese」と呼ぶ．）．PatentBERT_japanese の学習には，NVIDIA が考案した LAMB Optimizer[７]を用いた．

表 1 に BERT の事前学習結果を示す. 非常に高いマスク語予測精度から，マスク語予測に特許文献を用いることによって，特許文献のマスク語予測が適切にできるようになったことが理解できる．

表 1　特許ドメイン特化型 BERT の事前学習結果

	MLM acc.	NSP acc.
PatentBERT_japanese	93.5	98.7

また，以下の図 1 は，PatentBERT_japanese によるマスク語予測のデモである．

マスク語予測タスク

ルテイン，ゼアキサンチン，**[MASK]**，アスタキサンチン，ドコサヘキサエン酸，ドコサペンタエン酸，エイコサペンタエン酸，ビタミンＥ，β－カロテン，及びコエンザイムＱ－１０から選ばれる１種以上を含有することを特徴とする請求項１～８のいずれか１項に記載の視力回復用組成物

PatentBERT_japanese によるマスク語予測トップ３

１．リコピン（正解）
２．アスタキサンチン
３．カロテノイド

図 1　PatentBERT_japanese によるマスク語予測デモ

これにより，PatentBERT_japanese は，特許明細書中に出現する非常に専門的な語であっても，その意味概念を概ね正しく捉えることができることが分かる．

そして，この PatentBERT_japanese の最下層に，入力された特許明細書の文章が Y02，Y04 の各 CPC メイングループ(CPC はセクション（例：Y）、クラス（例：Y02）、サブクラス（例：Y02B）、メイングループ（例：Y02B10）に階層化されており、メイングループは、8 デジットある分類情報のうちの 6 デジットまでの情報を示す)に該当するか否かを判定する分類層を追加することによって，本研究で用いる機械学習モデルを構築した. なお, 対象とした Y02, Y04 のメイングループは 39 種類あるため, 最下層に追加した分類層は 39 クラスの分類層である．

なお，特許明細書に対する Y02，Y04 の CPC メイングループの付与は重複的に行われるため，当該 CPC メイングループの推定はマルチラベルタスクである. そのため, 損失関数には BCEWithLogitsLoss を用いることとした．

２．２．コーパス

コーパスについては, Y02, Y04 が付与されている日本語特許公報を収集し，さらに，特許技術者が当該収集した特許公報の明細書中身を目視で確認することによって収集を行った．コーパスのサイズは，合計 8,724 件である．

２．３．計算条件

計算条件は以下の通りである．

表 2　訓練における各種パラメータ

損失関数	BCEWithLogitsLoss
学習率	$5.0*10^{-5}$
バッチ数	256
エポック数	50

３．結果

コーパスの 10%をバリデーション用データとして，ホールドアウト法によって精度の検証を行った．結果は以下の通りである．

表 3　学習結果

Precision(Micro)	0.89
Recall(Micro)	0.86
F1(Micro)	0.88
F1(Macro)	0.86

４．2000 年以降に日本に出願された特許出願のマクロ分析

学習したモデルを用いて，2000 年以降に日本に出願された特許出願 6,700,799 件全てについて推論を行った. そのうち, 2019 年に出願された特許 226,180

件について，その結果を示す．なお，対象を 2019 年に出願された特許に絞った理由は，通常、特許出願は出願日から１年６月経たないと公開されず，現時点で 2020 年以降に出願された特許すべてを分析できないためである．

当該モデルに特許の明細書を入力すると，Y02，Y04 の CPC メイングループからなる 39 の CPC メイングループの推定値が算出される．ここで，本研究では，39 の CPC メイングループと経済産業省が定める「エネルギー関連産業」「輸送・製造関連産業」「家庭・オフィス関連産業」の３産業とのコンコーダンスを作成し，当該コンコーダンスに基づいて，39 の CPC メイングループの推定値を，特許明細書に記載されている技術が，当該３産業のうちのどの産業の脱炭素に資するものであるのかについての推定値へ変換を行った．

そして，各出願人について３産業それぞれの推定値の合計をスコアとして算出することによって，2019 年に出願された特許における各３産業の出願人スコアランキングを算出した．これらを以下に示す．

表 4　企業ランキング
エネルギー関連産業，2019 年出願

1	トヨタ自動車株式会社
2	パナソニックＩＰマネジメント株式会社
3	株式会社豊田自動織機
4	エルジー・ケム・リミテッド
5	本田技研工業株式会社
6	株式会社ＧＳユアサ
7	株式会社東芝
8	株式会社デンソー
9	積水化学工業株式会社
10	株式会社豊田中央研究所

表 5　企業ランキング
輸送・製造関連産業，2019 年出願

1	トヨタ自動車株式会社
2	本田技研工業株式会社
3	株式会社デンソー
4	キヤノン株式会社
5	パナソニックＩＰマネジメント株式会社
6	株式会社ディスコ
7	株式会社豊田自動織機
8	株式会社ＳＵＢＡＲＵ
9	マツダ株式会社
10	三菱電機株式会社

表 6　企業ランキング
家庭・オフィス関連産業，2019 年出願

1	パナソニックＩＰマネジメント株式会社
2	三菱電機株式会社
3	株式会社日立製作所
4	キヤノン株式会社
5	ダイキン工業株式会社
6	株式会社東芝
7	株式会社半導体エネルギー研究所
8	シャープ株式会社
9	ファナック株式会社
10	株式会社デンソー

５．まとめと今後の課題

機械分類モデルでは，コーパスの品質及びサイズが非常に重要であるため，コーパスの整備及び整備後のコーパスによるモデルの再学習を今後も継続的に行っていく必要がある．

また，本研究では日本に出願された特許の公開特許公報を解析対象としたが，弊機構は高精度の機械翻訳エンジンを持っていることから，海外の特許明細書を日本語化して，本モデルの解析対象とすることも可能である．今後は引き続き，国際公開の特許明細書や米国特許公開公報等の解析を行っていく．

参考文献

[１] 環境省, 2050 年カーボンニュートラルの実現に向けて, https://www.env.go.jp/earth/2050carbon_neutral.html, (参照 2021 年 9 月 2 日)

[２] 経済産業省, 2050 年カーボンニュートラルに伴うグリーン成長戦略を策定しました, https://www.meti.go.jp/press/2021/06/20210618005/20210618005.html, (参照 2021 年 9 月 2 日)

[３] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805. (2018)

[4] Jieh-Sheng Lee, Jieh Hsiang.: PatentBERT: Patent Classification with Fine-Tuning a pre-trained BERT Model. arXiv:1906.02124. (2019)

[5] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, Jaewoo Kang.: BioBERT: a pre-trained biomedical language representation model for biomedical text mining. arXiv:1901.08746. (2019)

[6] Manzil Zaheer, Guru Guruganesh, Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, Amr Ahmed.: Big Bird: Transformers for Longer Sequences. arXiv:2007.14062. (2020)

[7] Yang You, Jing Li, Sashank Reddi, Jonathan Hseu, Sanjiv Kumar, Srinadh Bhojanapalli, Xiaodan Song, James Demmel, Kurt Keutzer, Cho-Jui Hsieh.: Large Batch Optimization for Deep Learning: Training BERT in 76 minutes. arXiv:1904.00962. (2019)

アナリストレポートから生成した極性指標を用いた市場動向分析

Market Trend Analysis Using Polarity Index Generated from Analyst Reports

田口　怜 [1*]　渡邊　光 [2]　平野　正徳 [1]　鈴木　雅弘 [1]
坂地 泰紀 [1]　和泉　潔 [1]　平松　賢士 [3]
Rei Taguchi[1]　Hikaru Watanabe[2]　Masanori Hirano[1]　Masahiro Suzuki[1]
Hiroki Sakaji[1]　Kiyoshi Izumi[1]　Kenji Hiramatsu[3]

[1] 東京大学大学院工学系研究科
[1] Graduate School of Engineering, The University of Tokyo
[2] 東京大学工学部
[2] Faculty of Engineering, The University of Tokyo
[3] 株式会社アイフィスジャパン
[3] IFIS Japan Limited.

Abstract: 本研究では, アナリストの個別銘柄に対するセンチメントが, マクロ経済指標の予測に役立つかを実証する. これはアナリストレポートのテキスト情報を自然言語処理を使用して極性指標を作成することで実現可能となる. 本研究では, 作成した極性指標に対し, 各種マクロ経済指標を使用し, VAR モデルを用いた分析を行った. 結果, 極性指標から物価, 為替, 国債等の指標へのグレンジャー因果性があることが確認された. これにより, 極性指標が先行しており, マクロ経済指標の予測に役立つことが示唆された.

1　はじめに

近年, 人工知能の爆発的な発展により, オルタナティブデータの活用が世界的に注目されている. 特に金融・経済分野での普及が顕著であり, 伝統的な財務データに加えて, 景気予測や投資戦略など幅広く活用され始めている. 日本においては, オルタナティブデータ推進協議会 (JADAA)[*1]等の業種間を超えた団体が発足し, オルタナティブデータの積極的な普及を推進している. そのなかでも特にテキストデータは汎用性が高く, 筆者の書きぶりや内容の偏向から予測や分類などの抽出が可能であることがわかっている [1][2][5].

本研究では, アナリストレポートを使用し, 個別銘柄の将来に対するセンチメントを基に作成した極性指標が, マクロ経済指標に対して先行性を持つと仮定し, マクロ経済指標の予測に役立つことを実証する.

* 連絡先： 東京大学大学院工学系研究科
〒 113-8656 東京都文京区本郷 7-3-1
E-mail: s5abadiee@g.ecc.u-tokyo.ac.jp

*1 https://alternativedata.or.jp

2　先行研究

機械学習による極性指標作成の研究として, 余野ら [5], 山本ら [6] がある. 余野ら [5] では, 金融経済月報を Long Short Term Memory(LSTM) で学習し, Latent Dirichlet Allocation(LDA) を用いて個人消費, 設備投資, 在庫などのトピックごとの指標を作成した. また, 全体に対する寄与度を求めることで, どのマクロファクターの影響力が高いかを分析している. 山本ら [6] では, 景気ウォッチャー調査を LSTM で学習し, 極性指標を作成し, 指数間の相関係数を算出している.

上記の研究では, 金融経済月報や景気ウォッチャー調査といった内容が類似したテキストを用いているのに対し, 我々は, アナリストレポートを用いて, その個別銘柄のセンチメントを総計して指標とした点で異なる. アナリストレポートを用いたテキストマイニングの研究事例として, 平松ら [11], Suzuki et al.[3] と小林ら [5] がある. 平松ら [11] では, 株価はレポートのセンチメントに強く反応しているのみならず, レポート発行後に株価のドリフトも観測しており, 資産運用実務においてもアナ

リストレポートのテキスト情報が有益であることが示唆された．Suzuki et al.[3] では, LSTM 等を使用しアナリストレポートから意見文と非意見文を抽出し, 純利益と株価の予測を行っている．加えて, 小林ら [5] は, 独自の深層学習モデルを用いてアナリストレポートからアナリスト予想根拠文の抽出を行っている．

一方で, 我々が, BERT を用いてアナリストレポートを学習させている点も他の研究と異なる．BERT を使用した研究事例として, Hiew et al.[7] がある．Hiew et al.[7] では, テキストデータに中国の SNS である weibo を利用し, BERT で極性を付与し, LSTM を用いて株価予測を行っている．

また, 上記以外の研究として，伊藤ら [9] の研究がある．伊藤ら [9] は, ロイターニュースと Yahoo!ファイナンス掲示板データを用いて, ニューラルネットワークを使用した極性概念辞書作成手法を提案している．

3 提案手法

本研究では, 以下の 4 ステップにおいて, 極性指標の作成と, 比較指標との実証分析を行う．

Step1: 分析対象データに定常性を持たせるため, 必要なものには対数階差などの前処理を行う．使用するデータの期間はアナリストレポート, 比較指標のいずれも 2010/1〜2020/12 とする．

Step2 : BERT を用いてアナリストレポートの文章を学習し，「ポジティブ」「ネガティブ」「ニュートラル」の三値に極性分類を行う．学習に使用するデータの期間は Step1 と同様とする．

Step3 : 作成した極性指標（全業種・業種別）に対し, グレンジャー因果性の検定を行う．VAR モデルに使用するデータの期間は Step1 と同様とする．

Step4 : 上記 Step3 のグレンジャー因果検定で,「原因: 極性指標→結果: 比較指標」になる組み合わせについて, インパルス応答関数と予測誤差分散分解を用いて分析を行う．

また, 本研究におけるアーキテクチャーは図 1 のように記述できる．

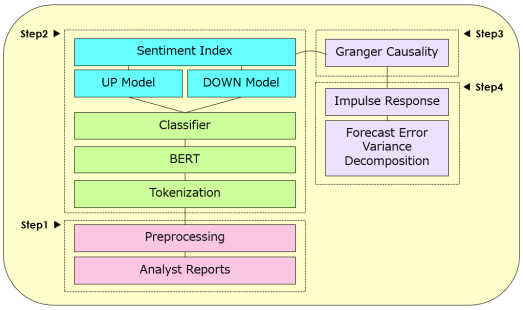


図 1　本研究のアーキテクチャー

3.1 使用データ

BERT 極性指標作成には, アナリストレポートデータ (373,050 件) を用いる．データクレンジングには IFIS 社の技術を使用した．VAR モデルでの分析には以下のデータを用いる．

- **Source1** みずほ銀行[*2] が公開している各国の外国為替公示相場月次ヒストリカルデータ
- **Source2** 日本銀行[*3]が公開している企業向けサービス価格指数月次ヒストリカルデータ
- **Source3** 日本銀行が公開している普通国債等発行額月次ヒストリカルデータ

データの期間はいずれも 2010/1〜2020/12 とする．今回の検証では, BERT による極性分類については将来のアナリストレポートテキストデータも既知のものとし, 極性指標がマクロ経済指標の予測に有効か否かのチェックに留め, 実際の運用を想定した BERT モデル学習期間以降のデータのみでの検証は今後の課題として取り組んでいる最中である．但し, 自然言語処理のみに関して上記の取り扱いを行っているため, 固有名詞を除き日本語の表現自体は年代による変化は小さく, 本検証の有効性に対する影響は軽微と考えられる．

3.2 Bidirectional Encoder Representations from Transformers（BERT）

BERT は, Devlin et al. [4] により提案された Transformer ベースの言語モデルである．BERT では, 入力内のいくつかのトークンをマスキングし, マスキングされた単語を予測することで, うまく双方向の学習を可能にしている．本研究では, 日本語 BERT の学習済みモデルとして，東北大学の乾・鈴木研究室が公開しているモデ

*2 https://www.mizuhobank.co.jp/retail/index.html
*3 https://www.boj.or.jp

ル[*4]を利用した.

3.3 極性分類手法

アナリストレポートから, レーティングが「買い」に変化したレポート, 変化のないレポート, 「売り」に変化したレポート, をそれぞれランダムに 5000 件ずつサンプリングする. 次に, BERT を用いて二種類のモデルを作成する. レーティングが「買い」に変化したレポートと変化のないレポートを用いて極性分類を行うモデルを作成する. これを UP Model とする. 加えて, レーティングが「売り」に変化したレポートと変化のないレポートを用いて極性分類を行うモデルを作成する. これを DOWN Model とする. これにおいてレーティングが「買い」に変化したレポートを 1, 変化のないレポートを 0, 「売り」に変化したレポートを-1 とする. UP Model と DOWN Model の出力 (1, 0, -1) を集計し, 最終的な極性判定を行う. 判定方法は, 表 1 を参照. 表 1 の尤度は, UP Model で出力した買い判定と, DOWN Model で出力した売り判定のうち, 値が大きい方である.

表 1 極性判定手法

UPModel 判定	DOWNModel 判定	最終判定
1（買い）	0（変化なし）	1（買い）
0（変化なし）	-1（売り）	-1（売り）
0（変化なし）	0（変化なし）	0（変化なし）
1（買い）	-1（売り）	尤度が高い方

図 2 は, 作成した極性指標を業種別に集計し, 時系列ごとにプロットしたものである.

図 2 業種別極性指標の時系列プロット

図 3 は, 業種別で集計した極性指標の平均を取ったものである. 本稿では全業種極性指標と呼称する.

[*4] https://github.com/cl-tohoku/bert-japanese

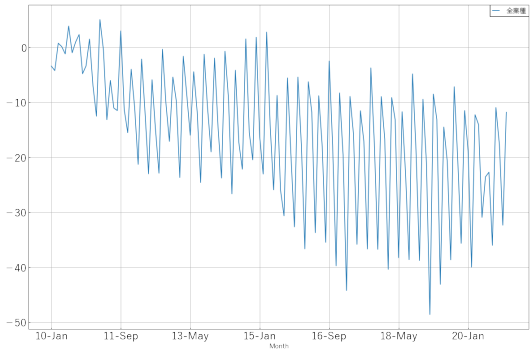


図 3 全業種極性指標の時系列プロット

図 4 は, 参考として業種別極性指標の相関係数を算出した. 作成した極性指標は, 業種間で外れ値がなく, 概ね正の相関を示していることがわかる.

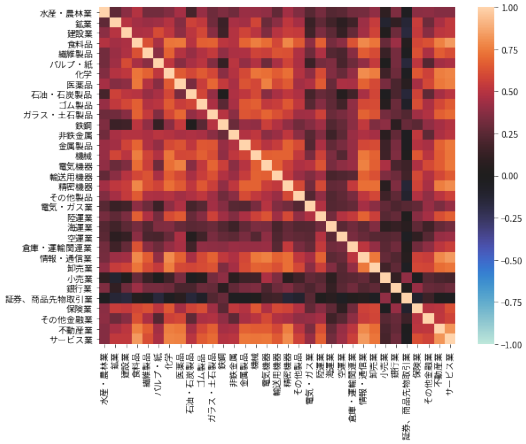


図 4 業種別極性指標同士の相関係数ヒートマップ

3.4 ベクトル自己回帰 (VAR) モデル

VAR モデルは, 自己回帰 (AR) モデルを多変量に拡張したモデルである. VAR モデルは, 変数間の動学的な関係を分析する際に使用され, グレンジャー因果性検定, インパルス応答関数, 予測誤差分散分解, といった優れた分析手法を提案する. VAR モデルを数式で表現すると以下のようになる.

$$y_t = c + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \epsilon_t, \quad (1)$$

$$\epsilon_t \sim W.N.(\Sigma) \quad (2)$$

ここで, y_t の t は時間を表し, c は $n \times 1$ 定数ベクトルであり, Φ_i は $n \times n$ 係数行列を指す. 次数 p については AIC により選択する. また, ϵ_t は撹乱項を指し,

ϵ_t が分散共分散行列 Σ のホワイトノイズであることを $\epsilon_t \sim W.N.(\Sigma)$ と表記する．

3.4.1 グレンジャー因果性検定

グレンジャー因果性検定は，時系列データ間で先行する一方の変数によってもう一方の変数を予測できるかどうかについての検定として，Granger [12] によって提唱された．ある時系列データ $y_{1,t}$ と $y_{2,t}$ がある場合，現在と過去の $y_{1,t}$ にだけ基づいた場合の将来の $y_{1,t}$ の予測と，現在と過去の $y_{1,t}$ と $y_{2,t}$ に基づいた場合の将来の $y_{1,t}$ の予測を比較して，後者の Mean Squared Error (MSE) が小さくなる場合，$y_{2,t}$ から $y_{1,t}$ へのグレンジャー因果性が存在すると判定できる．

3.4.2 インパルス応答関数（IRF）

グレンジャー因果性検定は，時系列データ間の因果性を判定できるが，その関係の強さを測定することは出来ない．インパルス応答関数は，ある変数に対するショックが，その変数やほかの変数に与える影響を定量的に分析できるツールである．分析には，誤差項同士も相関があることを仮定した，直交化インパルス応答関数を用いた．

3.4.3 予測誤差分散分解（FEVD）

予測誤差分散分解は，時系列データの予測の MSE を，相対的分散寄与率，すなわち各変数固有の撹乱項が寄与する部分，に分解し，どの変数が重要であるかを分析できる手法である．

4 実証分析と結果

4.1 外国為替公示相場における実証分析

外国為替公示相場月次ヒストリカルデータについて対数階差を取り，前出の全業種極性指標・業種別極性指標に対して，相互の関係性を VAR モデルを用いて定量化する．次数については AIC で推定する．極性指標から外国為替へのグレンジャー因果性が存在する組み合わせについては付録を参照．結果，化学，鉄鋼，倉庫・運輸関連業等の極性指標から各国為替についてグレンジャー因果性が存在することが実証された．一方，全業種極性指標から各国為替についてのグレンジャー因果性は確認できなかった．

インパルス応答関数については，以下の結果となった．今回は，グレンジャー因果性検定を行った際の p-value が一番低い組み合わせを用いた．

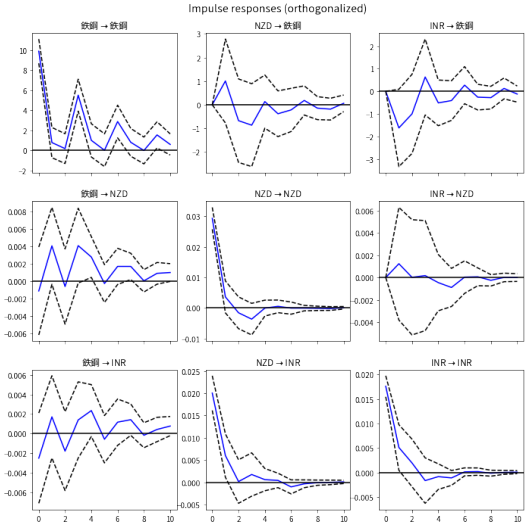


図 5　鉄鋼極性指標と外国為替公示相場におけるインパルス応答関数

図 5 の「鉄鋼→鉄鋼」において約 10 の正のショックが発生すると，「鉄鋼→ NZD」においてラグ 1 の地点で約 0.004 の正のショックが発生している．また，「鉄鋼→ INR」においてラグ 1 の地点で約 0.002 の正のショックが発生した．

予測誤差分散分解については，以下の結果となった．

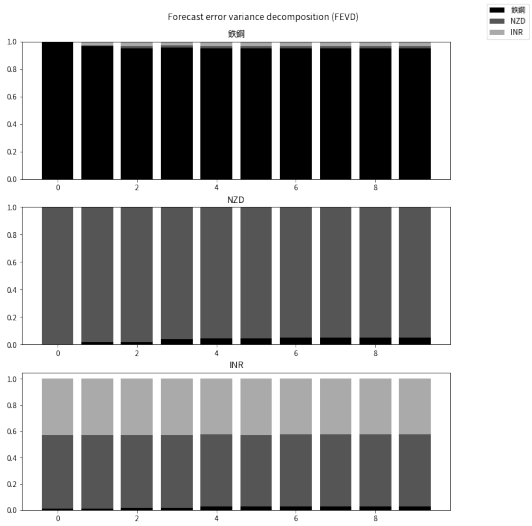


図 6　鉄鋼極性指標と外国為替公示相場における予測誤差分散分解

図 6 より，鉄鋼極性指標は，予測できない NZD の変動に関しては，約 5% 程度の説明力を持つことが実証された．また，予測できない INR の変動に関しては，約 2.5% 程度の説明力を持つことが実証された．

4.2 企業向けサービス価格指数における実証分析

企業向けサービス価格指数月次ヒストリカルデータについて, 外国為替公示相場と同様の手法で相互の関係性を定量化する. 結果, 全業種, 小売業についてグレンジャー因果性が存在した.

インパルス応答関数については, 以下の結果となった.

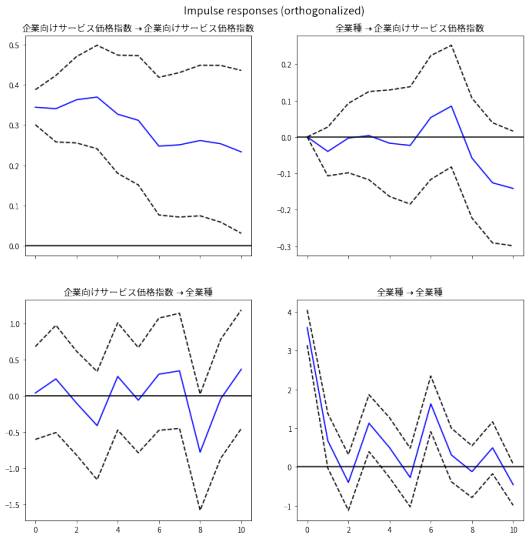


図 7　全業種極性指標と企業向けサービス価格指数におけるインパルス応答関数

図 7 の「全業種→全業種」において約 3.5 の正のショックが発生しているのに対し, 「全業種→企業向けサービス価格指数」において, ラグ 1 の地点で約 0.05 の負のショックが発生した.

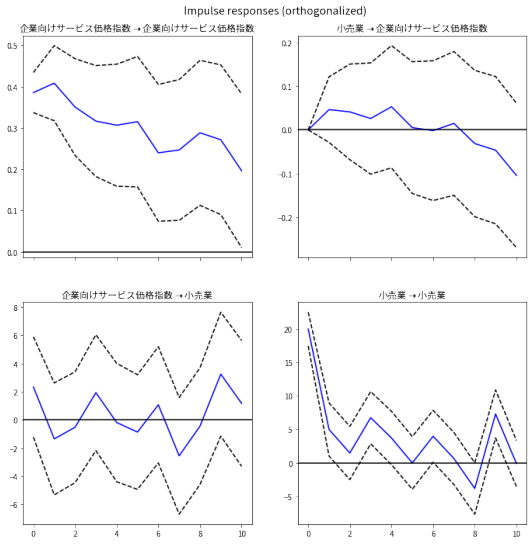


図 8　小売業極性指標と企業向けサービス価格指数におけるインパルス応答関数

図 8 の「小売業→小売業」において約 20 の正のショックが発生しているのに対し, 「小売業→企業向けサービス価格指数」において, ラグ 1 の地点で約 0.05 の正のショックが発生した.

予測誤差分散分解については, 以下の結果となった.

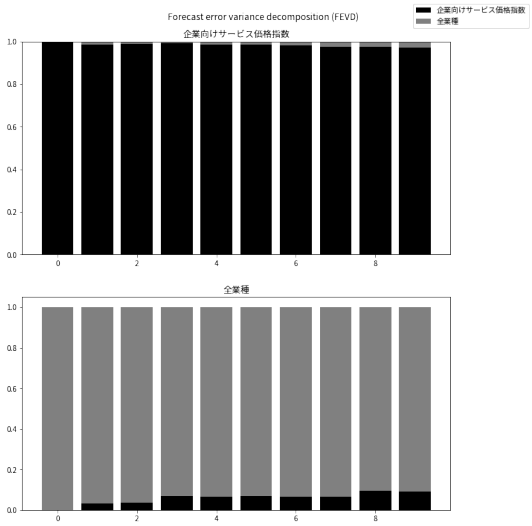


図 9　全業種極性指標と企業向けサービス価格指数における予測誤差分散分解

図 9 より, 全業種極性指標は, 予測できない企業向けサービス価格指数の変動に関しては, 約 1.5% 程度の説明力を持つことが実証された.

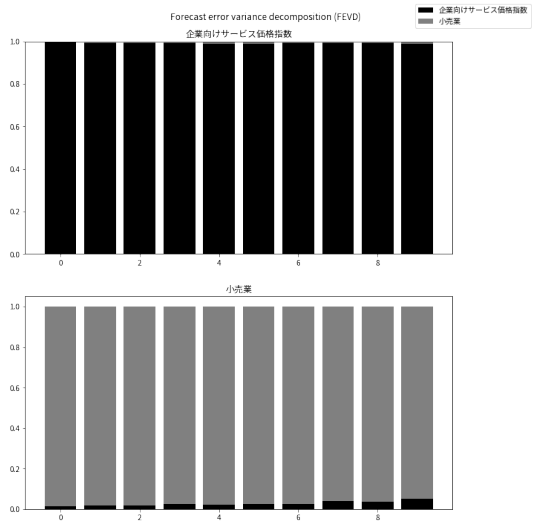


図 10　小売業極性指標と企業向けサービス価格指数における予測誤差分散分解

図 10 より, 小売業極性指標は, 予測できない企業向けサービス価格指数の変動に関しては, 約 1% 程度の説明力を持つことが実証された.

4.3　普通国債等発行額における実証分析

普通国債等発行額月次ヒストリカルデータについて対数階差をとり, 外国為替公示相場と同様の手法で相互の関係性を定量化する. 結果, 全業種, 建設業についてグレンジャー因果性が存在する通貨が多く存在した.

インパルス応答関数については, 以下の結果となった.

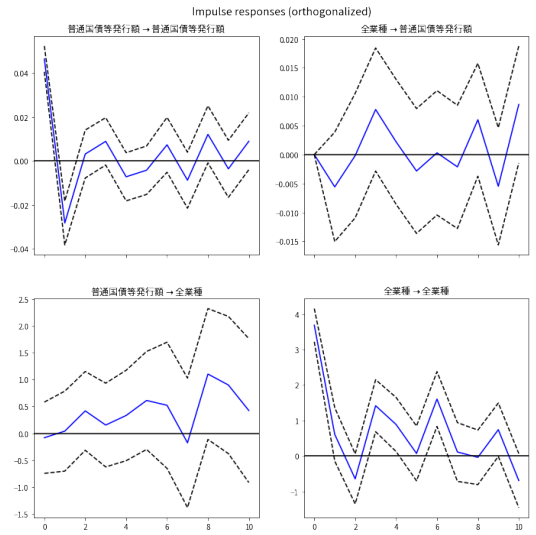


図 11　全業種極性指標と普通国債等発行額におけるインパルス応答関数

図 11 の「全業種→全業種」において約 3.8 の正のショックが発生しているのに対し, 「全業種→普通国債等発行額」において, ラグ 1 の地点で約 0.005 の負のショックが発生した.

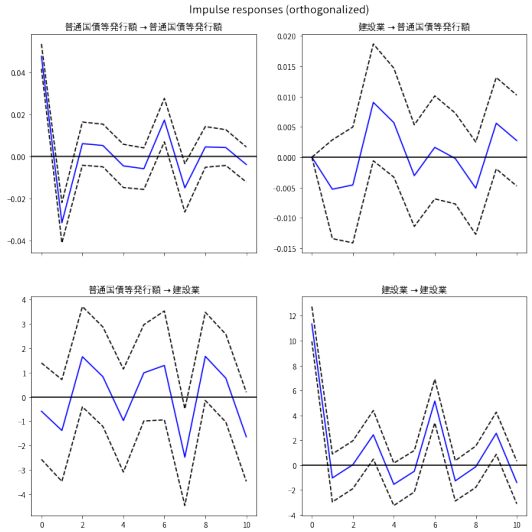


図 12　建設業極性指標と普通国債等発行額におけるインパルス応答関数

図 12 の「建設業→建設業」において約 11 の正のショックが発生しているのに対し, 「建設業→普通国債等発行額」において, ラグ 1 の地点で約 0.005 の負のショックが発生した.

予測誤差分散分解については, 以下の結果となった.

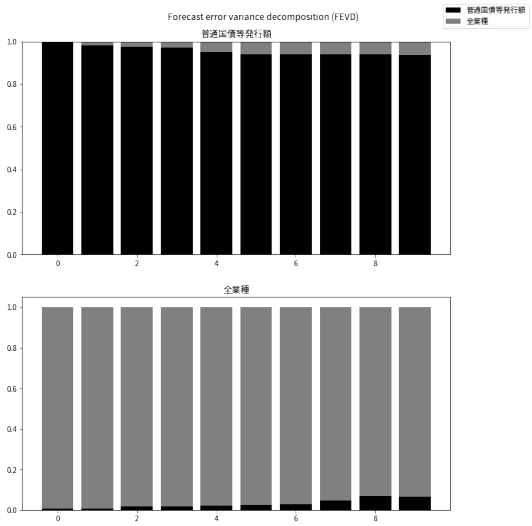


図 13　全業種極性指標と普通国債等発行額における予測誤差分散分解

図 13 より, 全業種極性指標は, 予測できない普通国債

等発行額の変動に関しては, 約 6% 程度の説明力を持つことが実証された.

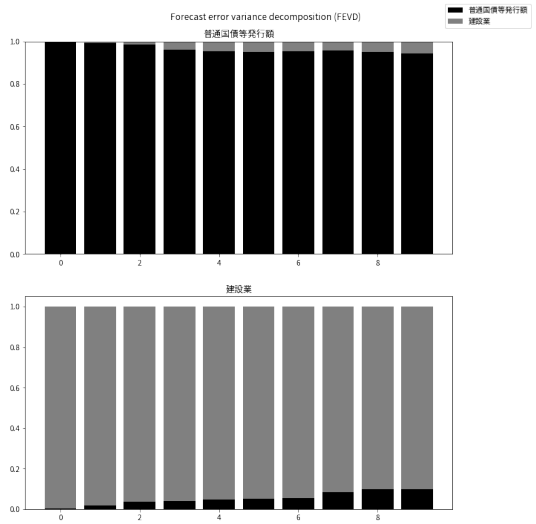


図 14　建設業極性指標と普通国債等発行額における予測誤差分散分解

図 14 より, 建設業極性指標は, 予測できない普通国債等発行額の変動に関しては, 約 4.5% 程度の説明力を持つことが実証された.

4.4　結果サマリ

表 2 は, 全業種極性指標・業種別極性指標から比較指標へグレンジャー因果性が存在する組み合わせである. 外国為替公示相場については, 全業種極性指標から各国為替についてのグレンジャー因果性は確認できなかった. 外国為替公示相場の表 2 の表記については, 例えば「原因: 水産・農林業極性指標→結果:AUD, NZD, INR, PHP」になる組み合わせが存在することを示している.

表 2　極性指標から比較指標へのグレンジャー因果性が存在する組み合わせ

極性指標	比較指標
水産・農林業	外国為替 4 カ国分
食料品	外国為替 2 カ国分
化学	外国為替 7 カ国分
鉄鋼	外国為替 12 カ国分
輸送用機器	外国為替 2 カ国分
精密機器	外国為替 3 カ国分
その他製品	外国為替 1 カ国分
倉庫・運輸関連業	外国為替 11 カ国分
情報・通信業	外国為替 1 カ国分
証券商品先物取引業	外国為替 1 カ国分
保険業	外国為替 2 カ国分
全業種	企業向けサービス価格指数
小売業	企業向けサービス価格指数
全業種	普通国債等発行額
建設業	普通国債等発行額

加えて, 表 2 を元に行ったインパルス応答関数と予測誤差分散分解では, 微細な影響ながら比較指標に対し極性指標に先行性があることを示すことができた.

5　考察

実験結果より, 外国為替公示相場, 企業向けサービス価格指数, 普通国債等発行額に関して, VAR モデルを用いてグレンジャー因果性検定を行い, グレンジャー因果性がある可能性が示唆された組み合わせについて, インパルス応答関数・予測誤差分散分解を行った. これにおいて, 業種別のグレンジャー因果性については指標ごとに特色があることがわかった. 例えば, 外国為替公示相場はオーストラリアドル (AUD) やフィリピンペソ (PHP) において, アナリストのセンチメントが先行している反面, 米ドル (USD) や英ポンド (GBP) などの先進国はアナリストのセンチメントがそれほど先行していないことがわかった. また, 物価指数である企業向けサービス価格指数はアナリストの小売業のセンチメントを反映していることが示唆された. 加えて, 普通国債等発行額についてはアナリストの建設業のセンチメントを反映していることが示唆された. 一方で, それぞれのインパルス応答関数を確認すると微細な影響ながら先行性を持つことが示唆された. 予測誤差分散分解の結果についてもそれぞれの経済指標に対して大きな説明力は持たないが, 上の分析に対して妥当な結果となっている.

6 まとめ

本研究では, BERT とアナリストレポートを用いて作成した極性指標が, 既存のマクロ経済指標に対して先行性を持つという仮説に対し, VAR モデルを用いて実証分析を行った. 結果, 極性指標が複数の経済指標に対して, 先行性を持つことがわかった. 今後は, FinBERT や ELECTRA などで作成した極性指標との比較, 他の金融テキストで作成した極性指標との比較分析などを行って行きたい.

謝辞

株式会社アイフィスジャパンより, 本研究にご協力いただいた. 厚く御礼を申し上げ, ここに感謝の意を表する.

参考文献

[1] 坂地泰紀, 和泉潔, 松島裕康, 川瀬和哉, 林寛: 接触履歴を用いた地域景況インデックスの自動生成, 知能と情報, Vol.31, No.2, pp.626–635 (2019)

[2] 坂地泰紀, 蔵本涼太, 和泉潔, 松島裕康, 島田尚, 砂川恵太: 接触履歴を用いた地方景況感と業種間構造の分析, 第 *24* 回人工知能学会金融情報学研究会資料, pp.98–102 (2020)

[3] Masahiro Suzuki, Hiroki Sakaji,Kiyoshi Izumi, Hiroyasu Matsushima, and Yasushi Ishikawa: Forecasting Net Income Estimate and Stock Price Using Text Mining from Economic Reports, *INFORMATION*, Vol.11, No.6, pp.292 (2020)

[4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova: Bert: Pre-training of deep bidirectional transformers for language understanding., *arXiv preprint*, pp.1810.04805 (2018)

[5] 余野京登, 和泉潔: 金融レポート, およびマクロ経済指数によるリアルタイム日銀センチメントの予測, 人工知能学会全国大会論文集第 *31* 回全国大会, pp.2D13–2D13 (2017)

[6] 山本裕樹, 松尾豊: 景気ウォッチャー調査を学習データに用いた金融レポートの指数化, 人工知能学会全国大会論文集第 *30* 回全国大会, pp.3L3OS16a2-3L3OS16a2(2017)

[7] Joshua Zoen Git Hiew, Xin Huang, Hao Mou, Duan Li, Qi Wu, Yabo Xu: BERT-based financial sentiment index and LSTM-based stock return predictability. (2019)

[8] 小林和正, 酒井浩之, 坂地泰紀, 平松賢士: アナリストレポートからのアナリスト予想根拠情報の抽出と極性付与, 第 *19* 回人工知能学会金融情報学研究会資料, pp.65–70 (2017)

[9] 伊藤友貴, 坪内孝太, 山下達雄, 和泉潔: 経済テキストデータを用いた極性概念辞書構築とその応用, 第 *18* 回人工知能学会金融情報学研究会資料, pp.44–51 (2017)

[10] 仁木裕太, 坂地泰紀, 和泉潔, 松島裕康: 再事前学習した BERT を用いた金融文書中の因果関係知識有無の判別, 人工知能学会全国大会論文集第 *34* 回全国大会, pp. 3Rin439–3Rin439 (2020)

[11] 平松賢士, 三輪宏太郎, 酒井浩之, 坂地泰紀: アナリストレポートのトーンの情報価値, 証券アナリストジャーナル *59(2)*, pp.86–97 (2021)

[12] C.W.J. Granger: Investigating causal relations by econometric models and cross-spectral methods., *Econometrica*, 37:424–438 (1969)

補遺 グレンジャー因果性が存在する組み合わせ

本表はセクション 4.1 において極性指標から外国為替へのグレンジャー因果性が存在する組み合わせである.

表 3 極性指標から外国為替へのグレンジャー因果性が存在する組み合わせ

極性指標	外国為替
水産・農林業	AUD, NZD, INR, PHP
食料品	IDR, KRW
化学	USD, CAD, BHD, HKD, SAR, PGK, TWD
鉄鋼	CAD, CHF, SEK, NOK, AUD, NZD, INR, PHP, MXN, HUF, PLN , KRW
輸送用機器	AUD, MXN
精密機器	INR, PHP, THB
その他製品	HUF
倉庫・運輸関連業	GBP, CAD, CHF, NOK, AUD, NZD, INR, SGD, MXN, PLN, MYR
情報・通信業	NOK
証券商品先物取引業	NOK
保険業	AUD, PHP

因果情報を用いた経済数値予測

Forecast of Economic Indicators using Causal Information

和泉 潔 [1]　坂地 泰紀 [1]　佐野 仁美 [1]

Kiyoshi Izumi [1], Hiroki Sakaji [1], and Hitomi Sano [1]

[1] 東京大学大学院工学系研究科 システム創成学専攻
[1]Department of Systems Innovation, School of Engineering, the University of Tokyo

Abstract: In this paper, we propose a method that uses causal information extracted from economic texts to predict numerical indicators related to economic and financial fields, such as macroeconomic indicators and stock prices. The proposed method automatically detects whether each sentence in the economic text contains causal information or not, and if it does, it identifies the cause and effect expressions and stores them in our economic causality database. Furthermore, the proposed method calculates the similarity between the result expression of causal information contained in the economic causal database and the causal expression of another causal information, and generates causal chains from the given text data. Causal chains are used to predict how the numerical values of economic indicators will change in the future due to spillover effects.

1. はじめに

経済現象は人々の因果によって動く．つまり，その現象に関係する人たちが，自分たちが持っている因果の認識に基づいて，将来を予想し行動を決定する．その行動の集積の結果，経済システム全体の挙動が決まる．例えば「少子高齢化」を起点とした因果系列(因果チェーン)を考えてみよう．「少子高齢化→労働力低下→経済縮小」のように予想する人が株式市場で多くなり，株価の下落要因に考えられる銘柄もある．逆に「少子高齢化→自炊回数の減少→お弁当やお惣菜の利用増」のように予想され，株価の上昇要因に考えられる銘柄もある．このように，経済現象の将来予測には，人々が持っている因果の認識を分析することが必須となる．

しかし経済現象のように人間の行動が関係する事象の因果関係を，数値データだけから統計的に解析することは難しい．なぜなら，人間が原因事象をどのように認識して，それに対してどのような行動を取るのかという行動ルールが因果関係の鍵となっているからである．時と場合によって因果関係が変化する．自然科学的現象のように数値データの統計分析で，客観的かつ普遍的な因果系列を取り出すことはほぼ不可能である．

そこで，本研究では人間が認識した因果関係が含まれていると思われる経済テキストデータを解析し，経済分野に関わる因果関係のデータベースを構築する．さらに，特定の事象を表すフレーズから派生する因果系列を構築し，波及効果または潜在的な原因に関連する数値指標を提示するアルゴリズムを提案する．本手法を用いて，テキスト表現された因果情報を基に，ある出来事から他の出来事への波及効果や要因候補を探索し，出来事と数値指標との関係性を用いて数値の将来変化を予測したり見逃していた数値の変化を発見したりすることを行う．

本研究の主な貢献は，次の通りである．

- 自然言語処理で構築された因果系列(非構造化データ)と，数値指標(構造化データ)を統合する新たな手法を開発した．
- 本手法を用いて，関心ある事象を表すテキストから因果系列をたどって，数値指標の変化を予測する方法を新たに提案した．
- 提案手法による予測結果は，人間にも理解しやすい形での提示が可能となる．

2. 関連研究

近年，自然言語で記載された因果情報の抽出については，ニューラルネットワークベースのものが多い．Dasgupta et al.[Dasgupta 2018]は LSTM を用いて原因と結果を表す表現を抽出する手法を提案している．

さらに，英語の因果情報抽出に関しては，Coling

2020 のワークショップである Financial Narrative Processing Workshops(FNP) の Shared Task FinCausal2020*1 に含まれていることから，様々な手法がワークショップにおいて提案された．タスクには因果情報を含む文を抽出するタスクと，原因と結果で構成される因果情報を抽出するタスクが設定された．因果情報を含む文を抽出するタスクで提案された手法の多くが Bidirectional Encoder Representations from Transformers(BERT)に基づくもの[Ionescu 2020, Gordeev 2020, Gupta 2020]であり，高い性能が報告されている．また，原因と結果で構成される因果情報を抽出するタスクにおいても，BERT に基づく手法が提案されている[Imoto 2020].

因果チェーンの構築に関する研究としては，Ishii et al.[Ishii 2012]や Alashri et al.[Alashri 2018]，Zhao et al.[Zhao 2017]などが存在する．Ishii らは，新聞記事から因果関係を抽出し，それを SVO と上位下位関係辞書に基づいて結合する手法を提案している．Alshri らは，コンセプトに基づく因果チェーン構築手法を提案している．Zhao らは，cause-to-effect と effect-to-cause のパスを考慮した手法により，因果チェーンを構築する手法を提案している．

他にも，因果は因果チェーンを構築する以外でも様々な応用が期待されている．例えば，ロボティクスの世界では，因果を考慮した新たなベンチマークである CausalWorld[Ahmed 2020]が提案されている．

上記で述べた既存の研究では，因果抽出や因果チェーン構築だけにとどまっており，それがどのような事象や具体的な数値に関係するかが分からない．そのため，実際の経済分析に利用するには，数ステップ必要となる．それに対して本研究は，因果チェーンを数値指標に結び付け，その数値の変化予測ができる点が新しい．

3. 経済的因果関係の抽出とデータベース構築

最初に，人間が認知した因果関係について記述されたテキストデータを解析し，そこから因果関係を抽出する技術が必要となる．本システムでは，決算短信という上場企業が業績や財務状況を開示するために定期的に発行しているテキストから，手がかり表現を用いた手法[坂地 2015]により因果関係を抽出した．

- 使用したテキストデータ：2012 年 10 月から 2018 年 5 月に約 2,300 社が発行した約 2 万個の決算短信テキスト
- 抽出した因果関係：1,078,542 個

抽出された因果関係は，決算短信の発行された日付，決算短信を発行した銘柄などと一緒にデータベースに保存される．

4. 因果チェーンの構築アルゴリズム

次に，上述の経済因果データベースから，特定の語句から関連する因果系列を構築する手法[和泉 2019][Izumi 2020]には，[西村 2018][Nishimura 2018]を基に改良した手法を用いた．具体的には以下の 4 ステップからなる(図 1).

1. ユーザーがフレーズを入力する．最初は終端ノードを入力テキストとする．
2. 終端ノードの表現と一定の類似性がある因果関係を経済因果データベースから選択し，ネットワーク追加候補とする．
3. ステップ 2 で抽出したネットワーク追加候補と終端ノードの組み合わせについて因果関係ノード間の類似度を計算する.因果関係ノード間の類似度が閾値 α 以上あるときにノードをネットワークに追加してネットワークを拡張する．
4. ステップ 3 でネットワークに追加したノードを終端ノードとしてステップ 2, 3 を繰り返す．

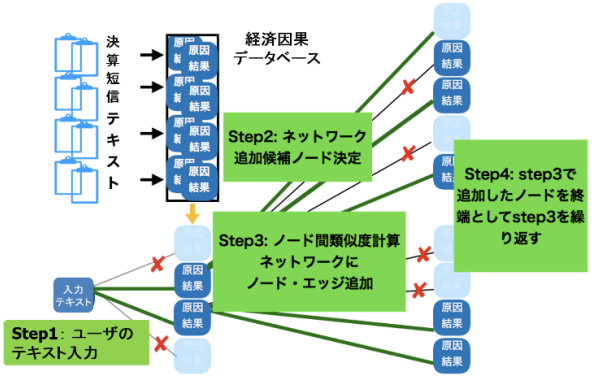


図 1. 因果チェーン構築のアルゴリズム
([西村 2018] [Nishimura 2018]を基に作成)

本手法を用いて，ユーザーの入力テキストを最初の原因表現として原因-結果の順に波及効果を探索する場合を前向きの因果チェーン検索と呼ぶ．逆に入力テキストを最後の結果表現として結果-原因の

*1 http://wp.lancs.ac.uk/cfie/fincausal2020/

順で潜在的な原因に遡って探索する場合を後ろ向きの因果チェーン検索と呼ぶ.

5. 因果情報を用いた経済数値予測

ユーザーの入力テキストから，前節のアルゴリズムを用いて因果チェーンを構築し，次の手法により波及効果や潜在的要因として変動が予想される数値を推定する(図 2).

図 2.　因果チェーンによる関連数値指標の提示

前向きの因果チェーン(波及効果)検索の場合(図 2a)：

1. ある特定の現象を表すテキストからの波及効果を表す因果チェーンを構築する.
2. 因果チェーンに出現した原因表現や結果表現に含まれるテキストが，最初の特定現象からの波及効果に関連する表現となる.
3. 波及効果に関連するテキストが与えられたときに，事前に行った数値指標とそれに関連するテキスト集合の組み合わせを使用した学習結果により，テキストから関連数値指標を提示する.
4. ユーザーが与えたテキストデータに対して，結果表現に関連する数値指標を波及効果として変化しそうな指標として予測結果として提示する.

後向きの因果チェーン(潜在的な要因)検索の場合(図 2b)：

1. ある特定の現象を表すテキストからの潜在的な原因を表す因果チェーンを構築する.
2. 因果チェーンに出現した原因表現や結果表現に含まれるテキストが，最初の特定現象をもたらす潜在的な要因に関連する表現となる.
3. 潜在的な要因に関連するテキストが与えられたときに，事前に行った数値指標とそれに関連するテキスト集合の組み合わせを使用した学習結果により，テキストから関連数値指標を提示する.
4. ユーザーが与えたテキストデータに対して，原因表現に関連する数値指標を潜在的な要因として変化しそうな指標として予測結果として提示する.

6. 経済数値予測の実験結果

因果チェーン検索の結果表現から関連数値指標および数値を予測する実験を行った．本実験では，数値指標の特定に(株)Deep Data Research のオルタナティブデータを利用している．本データは上場企業のマンスリーレポート(自社サイトで開示する IR や PR 等)を収集し構造化したものであり，指標や数値のほか，報告関連の特記事項(以下「記事」という)を含む．本実験では，当該データ過去 6 年分(2015 年 1 月～2021 年 2 月)約 47 万件のうち，記事が付された約 15 万件を利用した．具体的な実験手順および結果を以下に示す.

6.1 実験手順

1. 因果チェーン検索に任意のテキストを与えて前向きの因果チェーン検索を行い，結果表現を取得する.
2. オルタナティブデータの記事をテキストマイニングし，記事に結果表現を含むデータの数値指標を「関連数値指標の候補」として取得する.
3. 上記候補と結果表現との関連度(S)を計算し，結果が上位の指標を「関連数値指標」として特定する．本研究では，この関連度について，利用データを「対象の数値指標を含むデータ群」と「全データ群」の 2 群に分割し双方の修正オッズ比により評価する．具体的には，以下の式により結果表現と数値指標との関連度を計算する.

$$S=\frac{(P+0.5)*(1-Q+0.5)}{(1-P+0.5)*(Q+0.5)}$$

ただし，

$$P=\frac{\text{対象の数値指標かつ記事に結果表現を含むデータ数}}{\text{対象の数値指標の全データ数}}$$

$$Q=\frac{\text{対象の数値指標かつ記事に結果表現を含むデータ数}}{\text{対象の数値指標の全データ数}}$$

6.2 実験結果

提案手法を用いて経済指標数値を予測した．本実験は，因果チェーン検索の波及効果を示す結果表現を得る際に，テキストのみを与えた場合と，テキストと検索対象期間を与えた場合の各事例で行った．前者はテキスト（「感染症」，「米国大統領選挙」および「地球温暖化」）を与えた 3 事例であり，後者はテキスト（「消費税」）と 2 つの期間を因果チェーン検索に与えた．これらの結果表現から，変化が予想される数値指標をそれぞれ特定した．

「感染症」からの波及効果予測

因果チェーン検索にテキスト（「感染症」）を与え，波及効果を表す結果表現を取得し，当該結果表現に関連する経済数値指標を予測した．因果チェーン検索における 2 階層目の結果表現と強い関連度が予測された経済数値指標の上位 5 位を図 3 に示す．併せて，予測された数値指標のうち「客室稼働率」の数値変化を図 4 に示す．

因果チェーン検索結果

原因(1)：世界情勢，感染症，影響
結果(1)：日本人，出国者，前年比，万人

原因(2)：日本人，宿泊客，減少
結果(2)：需給，緩和，方向，客室，単価

結果(2)の関連数値指標　＜上位 5 位＞

	数値指標	関連度
1	平均賃料坪単価	8.093
2	NOI	8.093
3	ADR	6.641
4	RevPAR	6.636
5	客室稼働率	3.994

図 3. 「感染症」から波及した関連数値指標

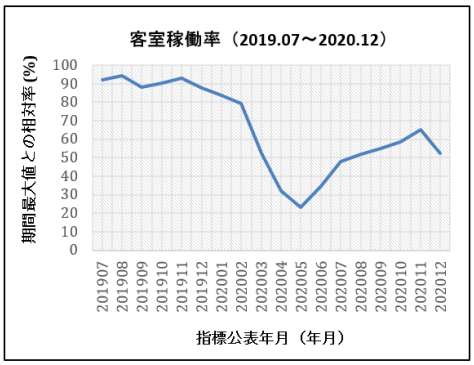


図 4. 「客室稼働率」の推移

図 4 を参照すると，関連数値指標「客室稼働率」の推移は，国内の新型コロナウィルス感染拡大に伴い全国に緊急事態宣言が発令された 2020 年 4 月前後に急降下している．平時に比べ変動幅が大きいことから，予測された関連数値指標は「感染症」から波及した結果表現と関連性が高いと考えられる．

「米国大統領選挙」からの波及効果予測

因果チェーン検索にテキスト（「米国大統領選挙」）を与えた場合の実験結果を示す．因果チェーン検索では 2 階層目の結果表現に「円安」，「株高」，「景気」，「好転」等が得られた．これらの結果表現と強い関連度が予測された経済数値指標の上位 3 位を図 5 に示す．併せて，予測された数値指標のうち「外国為替取引高」の数値変化を図 6 に示す．

「大統領選挙」から波及した関連数値指標＜上位 3 位＞

	数値指標	関連度
1	預かり資産残高	1.200
2	外国為替取引高	1.069
3	口座数	1.052

図 5.「米国大統領選挙」から波及した関連数値指標

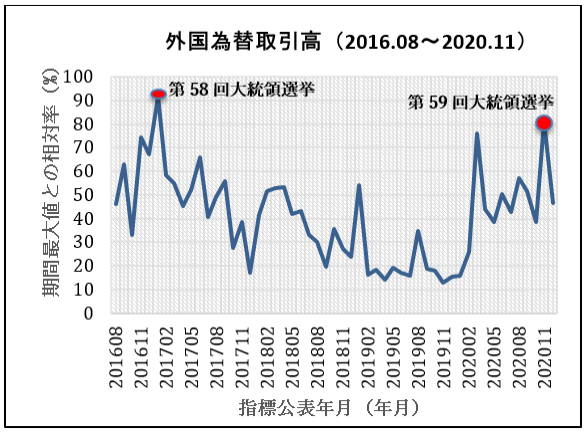


図 6. 「外国為替取引高」の推移

図 6 を参照すると，2016 年 11 月および 2020 年 11 月の双方の大統領選挙の時期に外国為替取引高が大幅に上昇している．2020 年 2 月前後にコロナ禍の不安材料に伴う取引増加が見られるが，相対的には「米国大統領選挙」から波及した関連数値指標の数値が当該タイミングで変動した事実が確認できる．

「地球温暖化」からの波及効果予測

因果チェーン検索にテキスト（「地球温暖化」）を与えた場合の実験結果を示す．因果チェーン検索では 1 階層目の結果表現に「温室効果ガス」，「排出」，「削減」，「目標」等が得られた．これらの結果表現

と強い関連度が予測された経済数値指標の上位 3 位を図 7 に示す．併せて，予測された数値指標のうち「CO2 削減量」の数値変化を図 8 に示す．

「地球温暖化」から波及した関連数値指標＜上位３位＞

	数値指標	関連度
1	木材購入料	8.112
2	CO2 削減量	1.933
3	発電量	1.045

図 7. 「地球温暖化」から波及した関連数値指標

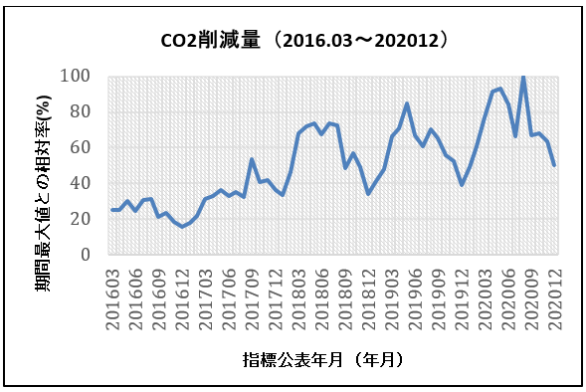


図 8. 「CO2 削減量」の推移

「地球温暖化」から波及した結果表現「温室ガスの削減」は世界共通の目標として進捗中であり，関連数値指標の「CO2 削減量」(図 8)は上昇傾向である．以上から，結果表現と関連性の高い経済指標および数値が得られたといえる．

「消費税」からの波及効果予測

因果チェーン検索にテキスト（「消費税」）および 2 種の対象期間を与え，提案手法を用いて経済指標を予測した．

まず，対象期間に消費税が 8%に増税された 2014 年 4 月を含む「2013 年 4 月～2014 年 9 月」を与えた場合, 因果チェーンの結果表現には「駆け込み需要」，「反動」,「物価上昇」および「懸念」等が得られた．これらは，2014 年 4 月に実施された消費税率引き上げの前の駆け込み需要とその後の反動が波及効果に現れたと考えられる．実際に住宅市場でも駆け込み需要が起き，その後は新築を中心に低迷が続いている．関連する経済指標としては「平均賃料坪単価」，「新規契約賃料」,「更新期日を迎えた既存テナントの更新割合」および「更新契約賃料」が予測された．

一方, 対象期間に消費税が 10%に増税された 2019 年 10 月を含む「2018 年 10 月～2020 年 3 月」を与えた場合，因果チェーンの結果表現には「薬価改定」，「売上原価率」、「臨床試験」および「販売促進費」等が得られた．これらは，厚生労働省により，消費税率 10%への引き上げに伴う 2019 年度診療報酬改定が告示されたことが波及効果に現れたと考えられる．そして，関連する経済指標には「客単価」,「店舗数」および「客数」が予測された．

このように，因果チェーン検索は，与えられたテキストと対象期間の組み合わせにより，ターゲットから波及する因果情報を鎖状に拡大し，対象期間の事象の関わりを可視化する．従って，本提案手法の適用により，因果チェーン検索を通じて，一般的には繋がりの見えにくい事象から関連する経済指標を取得し，動向を予測することが可能となる．

7. まとめ

本研究では，経済・金融に関わる事象の因果関係を，経済テキストデータから抽出したデータベースから連鎖的に検索し，波及効果または潜在的な原因に関連する数値指標を提示するアルゴリズムを提案した．本手法を用いて，具体的にいくつかの事象から人々が認識している因果関係から考えられる波及効果とそれに関連する数値指標を提示した．提示された数値指標の時系列変化を確認し，それらの変化のタイミングが，最初に与えられた事象と関連するタイミングであったことを見ることができた．

今後は因果系列の途中で出現したテキストを極性分析する手法を追加する．これにより，因果系列から推定した波及効果や潜在的な要因が関連する数値指標を増加させる方向の変化か減少させる方向の変化のどちらの影響を与えるのかを予測できる．

参考文献

[Dasgupta 2018] T. Dasgupta, R. Saha, L. Dey, A. Naskar: Automatic Extraction of Causal Relations from Text using Linguistically Informed Deep Neural Networks, In Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue, pp. 306-316, 2018.

[Ionescu 2020] Ionescu, M., Avram, A.-M., Dima, G.-A., Cercel, D.-C.,and Dascalu, M.: UPB at FinCausal-2020, Tasks 1 & 2: Causality Analysis in Financial Documents using Pretrained Language Models, in Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp.55–59, 2020.

[Gordeev 2020] Gordeev, D., Davletov, A., Rey, A., and Arefiev, N.: LIORI at the FinCausal 2020 Shared task, in Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial

Summarisation, pp.45–49, 2020.

[Gupta 2020] Gupta, S.: FiNLP at FinCausal 2020 Task 1: Mixture of BERTs for Causal Sentence Identification in Financial Texts, in Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp.74–79, 2020.

[Imoto 2020] Imoto, T. and Ito, T.: JDD @ FinCausal 2020, Task 2: Financial Document Causality Detection, in Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pp.50–54, 2020.

[Ishii 2012] Ishii, H., Ma, Q., and Yoshikawa, M.: Incremental construction of causal network from news articles, Journal of information processing, Vol. 20, No. 1, pp.207–215, 2012.

[Alashri 2018] Alashri, S., Tsai, J.-Y., Koppela, A. R., and Davulcu, H.:Snowball: extracting causal chains from climate change text corpora, in 2018 1st International Conference on Data Intelligence and Security (ICDIS), pp.234–241, 2018.

[Zhao 2017] Zhao, S., Wang, Q., Massung, S., Qin, B., Liu, T.,Wang, B., and Zhai, C.: Constructing and Embedding Abstrac Causality Networks from Text Snippets, in Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, p.335‑344, 2017.

[Ahmed 2020] O. Ahmed, F. Träuble, A. Goyal, A. Neitz,, M. Wüthrich, Y. Bengio, B. Schölkopf, M. Wüthrich, S. Bauer: CausalWorld: A Robotic Manipulation Benchmark for Causal Structure and Transfer Learning, arXiv, 2010.04296, 2020.

[坂地 2015] 坂地 泰紀, 酒井 浩之, 増山 繁: 決算短信 pdf からの原因・結果表現の抽出, 電子情報通信学会論文誌 D, vol. 98, no.5, pp.811-822, 2015.

[西村 2018] 西村 弘平, 坂地 泰紀, 和泉 潔: ベクトル表現を用いた因果関係連鎖の抽出, 人工知能学会 第 20 回金融情報学研究会, pp.50-53, 2018 年 3 月.

[Nishimura 2018] K. Nishumura, H. Sakaji, K. Izumi: Creation of Causal Relation Network using Semantic Similarity, 第 32 回人工知能学会全国大会, p. 1P104, 2018 年 6 月.

[和泉 2019] 和泉潔, 坂地泰紀: 経済因果チェーン検索のシステム紹介と応用, 人工知能学会 第 22 回金融情報学研究会, 2019 年 3 月.

[Izumi 2020] K. Izumi, H. Sakaji: Economic Causal-Chain Search Using Text Mining Technology, In: El Fallah Seghrouchni A., Sarne D. (eds) Artificial Intelligence. IJCAI 2019 International Workshops. IJCAI 2019. Lecture Notes in Computer Science, vol 12158. Springer, 2020.

[坂地 2018] 坂地 泰紀, 和泉 潔, 松島 裕康： 金融テキストマイニングの基づいた投資家支援プラットフォームの開発, 人工知能学会 第 21 回金融情報学研究会, 2018 年 10 月.

[坂地 2019] 坂地 泰紀, 和泉 潔, 松島 裕康, 川瀬 和哉, 林 寛, 接触履歴を用いた地域景況インデックスの自動生成, 日本知能情報ファジィ学会誌, vol. 21, no. 2, 2019.

確率的にホールドの長さが決まる人工市場モデル

Stochastically determined holding periods in the artificial financial market model

片平啓 [1*]　陳昱 [2]　秋山英三 [1]
Kei Katahira[1]　Yu Chen[2]　Eizo Akiyama[1]

[1] 筑波大学システム情報系
[1] Faculty of Engineering, Information and Systems, University of Tsukuba
[2] 東京大学大学院新領域創成科学研究科
[2] Graduate School of Frontier Sciences, The University of Tokyo

Abstract: 本研究では，人工市場モデル「投機ゲーム」を洗練化した「Self-organized Speculation Game（SOSG）」をさらに単純化し，パラメータの数を減らした軽量モデル「Pure Random Speculation Game（PRSG）」を構築した．PRSG のプレーヤーは，戦略テーブルの代わりに，確率に従ってホールドの長さを決めるため，往復取引を「機械的に」行う．しかし，SOSG と同様に市場規模が自発的に定まり，また，volatility clustering などの基本的な stylized facts が創発する．一方で，asymmetry in time scales などの高次の stylized facts については，PRSG では再現されない．高次特性の消失には，価格変動の時間構造に投機的パターンが見られないことが関連していると考えられる．

1　はじめに

株価や指数，為替などの金融に関する時系列データには，stylized facts と総称される，いくつかの普遍的な統計性質が存在する．Cont[1] によれば，少なくとも 11 個の性質が確認されており，代表的な基本特性としては，volatility clustering がよく知られている．これは，金融価格のリターンを時系列プロットすると，大きく上下動する部分（ボラティリティの大きい部分）がまとまって現れる性質である．

Volatility clustering をはじめとする多くの stylized facts は，市場や金融商品の違いに依存せず，トレーダー間の相互取引の中で「自発的に」創発する特性である．よって，これらの性質の発現メカニズムをモデルにより理解しようとするならば，外部からの調整が必要なパラメータの数は，できる限り少ない方が望ましい．このような観点から，単一ポジションの往復取引を実装した人工市場モデル「投機ゲーム」[2] を洗練化し，stylized facts の再現に必要な主要パラメータの数を 5 つから 3 つに減らしたモデルが「Self-organized Speculation Game（SOSG）」[3] である．

SOSG は，少数派ゲーム [4] で用いられた戦略テーブルによる意思決定の仕組みを応用して考案した投機ゲームに，ゆっくりと 1 粒ずつ砂を落とす砂山モデル [5] の仕組みを取り入れ，市場規模が自発的に定まるように改良を施したモデルである．SOSG では，砂山モデルとの振る舞いの類似性から，金融市場においても，外部からの調整が無くても自然に準臨界状態にまで到達する，自己組織化準臨界 [6] の原理が働いており，それが stylized facts の自発的創発に寄与している可能性があることが示唆された．SOSG は，投機ゲームと同様に，10 個の stylized facts を同時に再現することが可能である．

Stylized facts の中でも，基本特性である volatility clustering の創発メカニズムについては，多くの先行研究が，他人の行動に同調する群衆行動 [7] により引き起こされることを明らかにしている．一方，SOSG では，往復取引におけるホールド期間の多様性により volatility clustering が発生しうるということを初めて示した [8]．

しかし，SOSG は，投機ゲームの根幹のみを抽出したスマートなモデルである一方，パラメータ数の削減という観点においては，まだ工夫の余地が残っている．たとえば，volatility clustering の創発メカニズムに着目すれば，ホールド期間の多様性が保証される変更である限り，stylized facts の再現には問題がないはずである．よって，戦略テーブルによる意思決定構造を取り除き，もっと単純に，確率に従ってホールド期間が様々な値になるような仕組みに置き換えたとしても，再現モデルとしては機能すると考えられる．

*連絡先：筑波大学システム情報系
〒 305-8573 茨城県つくば市天王台 1-1-1
E-mail: k.katahira@akiyama-lab.net

そこで，本研究では，戦略テーブルの代わりに確率を SOSG に導入することで，パラメータの数を減らした軽量な人工市場モデル「Pure Random Speculation Game（PRSG）」を構築する．具体的には，PRSG のプレーヤーは，過去の価格変動を一切考慮せず，確率に従ってホールドの長さを決め，往復取引を「機械的に」行うように変更する．さらに，PRSG と SOSG のシミュレーション結果を比較することで，プレーヤーの意思決定方法の違いが価格変動に与える影響についても調べる．

2 PRSG と SOSG

PRSG と SOSG の違いを明確にするため，はじめに SOSG を説明し，その後 PRSG で変更する部分について述べる．

SOSG では，開始時に誰もいないゲーム市場に，毎時間ステップ t のはじめに，プレーヤー i が 1 人参入する．次に，ゲーム市場にいる $N(t)$ 人が，買い $(=1)$ と売り $(=-1)$ と待機 $(=0)$ の中から，行動 $a_i(t)$ として 1 つ選ぶ．プレーヤーが買い，または売り（空売りも可能）の注文を出す場合，注文量 $q_i(t)$ は，資本 $w_i(t)$ に応じて次のように決める：

$$q_i(t) = \lfloor \frac{w_i(t)}{B} \rfloor. \tag{1}$$

$\lfloor \cdots \rfloor$ は床関数であり，取引単位量 B は，複数量注文のしやすさを調整するパラメータである．ただし，各プレーヤーの初期資本量は $10B$ とし，往復取引の始めと終わりの注文量は同じとする．

市場価格の変動 Δp は，買いと売りの需給関係で決まる文献 [9] の注文不均衡式に倣い，初期価格[1]を任意として，

$$\Delta p = p(t) - p(t-1) = \frac{1}{N(t)} \sum_{j=1}^{N(t)} a_j(t) q_j(t) \tag{2}$$

のように決める．なお，ファンダメンタリストなどの投機家以外の影響も考慮したい場合は，Δp に摂動項を加える [10][11].

記憶長 M のプレーヤーは，行動の選択にあたり，長さ M の 5 進数の履歴 $H(t)$ を参照する．$H(t)$ は，次のように決まる離散価格変動 $h(t)$ の列により構成される：

$$h(t) = \begin{cases} 2\ (\text{大きく上昇}) & \text{if } \Delta p > C, \\ 1\ (\text{上昇}) & \text{if } C \geq \Delta p > 0, \\ 0\ (\text{変化なし}) & \text{if } \Delta p = 0, \\ -1\ (\text{下落}) & \text{if } -C \leq \Delta p < 0, \\ -2\ (\text{大きく下落}) & \text{if } \Delta p < -C. \end{cases} \tag{3}$$

[1]本研究では，リターン $r(t) = \ln p(t) - \ln p(t-1)$ を計算する都合上，$p(0) = 100$（$\gg 1$）とする．

認知閾 C は，大きな価格変動とプレーヤーが認識する境界値を表すパラメータである．$H(t)$ の参照後，市場参入時にランダムに与えられた唯一の戦略に従い，履歴パターンに対応する推奨行動 $a_i(t)$ を選ぶ（表 1 を参照）．ただし，単一ポジションの往復取引を実現するため，取引の開始ポジションと同じ行動を戦略が推奨する場合には，代わりに待機を選択する．

表 1: $M = 3$ の戦略テーブルの例.

履歴			推奨行動
−2	−2	−2	1
−2	−2	−1	0
−2	−2	0	0
−2	−2	1	−1
−2	−2	2	1
−2	−1	−2	0
	⋮		⋮
2	2	2	−1

往復取引のパフォーマンスは，$H(t)$ における Δp の大きさ情報の粗さを考慮し，人間の抽象的情報処理の様子を反映した認知的世界[2]において，次の認知価格 $P(t)$ に基づいて評価する：

$$P(t) = P(t-1) + h(t). \tag{4}$$

プレーヤーがポジションを t_0 に開き，t に反対売買により閉じる場合，この往復取引における単位量あたりの認知的利得 $\Pi_i(t)$ は，以下のように求められる：

$$\begin{aligned} \Pi_i(t) &= a_i(t_0)(P(t) - P(t_0)) \\ &= a_i(t_0) \sum_{k=t_0+1}^{t} h(k). \end{aligned} \tag{5}$$

さらに，現実的世界の $w_i(t)$ を更新するため，注文量 $q_i(t)$ を考慮して $\Pi_i(t)$ を単調増加な関数 f で変換し，資本調整 $\Delta w_i(t)$ を求める：

$$\begin{aligned} w_i(t) &= w_i(t_0) + \Delta w_i(t) \\ &= w_i(t_0) + f(\Pi_i(t) q_i(t_0)). \end{aligned} \tag{6}$$

SOSG では，単純化のため，$\Delta w_i(t) = \Pi_i(t) q_i(t_0)$ とする．資本の更新により $w_i(t) < B$ となり，1 単位も注文できない状態に陥ったプレーヤーは，ゲーム市場から退出する．

以上が，SOSG の仕組みである．なお，SOSG の主要パラメータは，記憶長 M と取引単位量 B と認知閾 C の 3 つである．

[2]現実的世界と認知的世界については，[2] の Fig.1 を参照.

対して，PRSG では，戦略テーブルの撤廃に伴い，履歴の長さを決める M や，撹乱効果によりボラティリティの発散を防ぐ C が不要になる [8]．その代わり，ポジションを閉じるための終了確率 Z という新規パラメータを導入するので，PRSG の主要パラメータは，B と Z の 2 つになる．

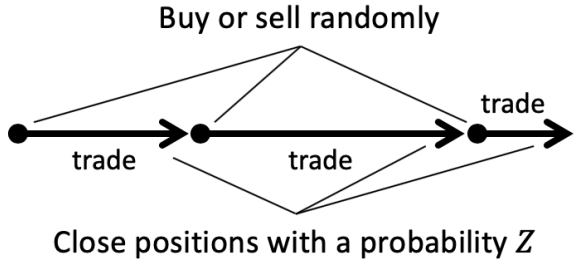


図 1: PRSG のプレーヤーが行う意思決定のイメージ．

PRSG のプレーヤーは，図 1 に示すように，ある往復取引を終えるとすぐに，次の時間ステップに新しい往復取引を始める．開始ポジションについては，買いか売りかをランダムに決める．開いたポジションは，終了確率 Z に従って反対売買により閉じ，それ以外の場合は待機を選択する．よって，PRSG のプレーヤーは，価格変動からのフィードバックを受けずに往復取引を「機械的に」行う（真の意味で往復取引を行っていない）．

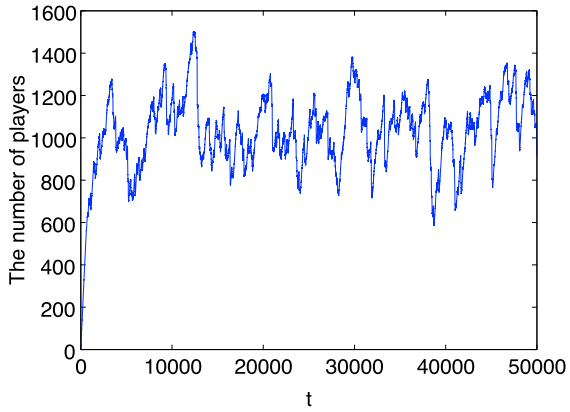
最後に，PRSG では C が不要なので，式 3 を以下のように変更する：

$$h(t) = \begin{cases} 1\ (\text{上昇}) & \text{if } \Delta p > 0, \\ 0\ (\text{変化なし}) & \text{if } \Delta p = 0, \\ -1\ (\text{下落}) & \text{if } \Delta p < 0. \end{cases} \tag{7}$$

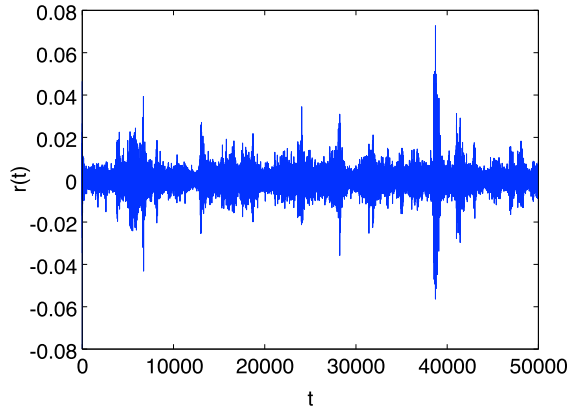
PRSG のプレーヤーは履歴を参照することはないが，$h(t)$ は資本の更新に使用するため，このような変更が必要となる．

3 結果・考察

はじめに，PRSG においても，SOSG と同じように，自発的に市場規模が定まりながら，volatility clustering が創発することを示す．図 2 は，$B = 12$，$Z = 0.05$（以降固定）としたときの (a) ゲーム市場に残留しているプレーヤー数，および (b) 対応する市場価格のリターン $r(t)$ の時系列プロットである．図 2 の (a) では，ゲーム開始直後に 1000 人程度まで線形的に増加した後，プレーヤー数はその周辺値で揺らいでおり，時間の経過と共に市場規模が自ずと調整されている．また，対応する図 2 の (b) のリターンについても，まとまった大きな



(a)



(b)

図 2: (a)PRSG のゲーム市場に残留するプレーヤー数と (b) 対応するリターンの推移．

増減が間欠的に見られ，安定的に volatility clustering が再現されている．

しかし，PRSG では，SOSG で再現できた高次（相関）の stylized facts が消失してしまう．たとえば，PRSG の asymmetry in time scales の性質は，SOSG のものに比べて非常に弱い．Asymmetry in time scales とは，時間スケール（解像度）の粗いボラティリティによって，時間スケールの細かいボラティリティを予測する方が精度が高くなる性質である [12]．

時間スケールの粗いボラティリティ $v^c(t)$ と細かいボラティリティ $v^f(t)$ は，次のように定義される：

$$v^c(t_i) = \left| \sum_{j=1}^{n} r(t_{i-1} + j\Delta t, \Delta t) \right|, \tag{8}$$

$$v^f(t_i) = \frac{1}{n} \sum_{j=1}^{n} |r(t_{i-1} + j\Delta t, \Delta t)|. \tag{9}$$

ただし，時間スケールは $\Delta t = (t_i - t_{i-1})/n$ と定義され，$v^f(t)$ については，平均化（$1/n$ 倍）しない場合もある．たとえば，$\Delta t = 1$ 日，$n = 5$ とすると，$v^c(t)$ は週足のボラティリティを，$v^f(t)$ は日足（平均）のボラティリティを意味する．

Asymmetry in time scales は，$v^c(t)$ と $v^f(t)$ の時間ラグ τ に関する時差相関

$$\rho_{cf}(\tau) = \mathrm{Corr}\left[v^c(t+\tau), v^f(t)\right] \tag{10}$$

によって説明される．金融価格の時系列データから $\rho_{cf}(\tau)$ を求めると，$|\tau|$ が小さいとき $\rho_{cf}(\tau) < \rho_{cf}(-\tau)$ となり，$\rho_{cf}(\tau)$ を τ に対してプロットすると，$\rho_{cf}(\tau)$ は $\tau = 0$ を境に非対称な分布になる．

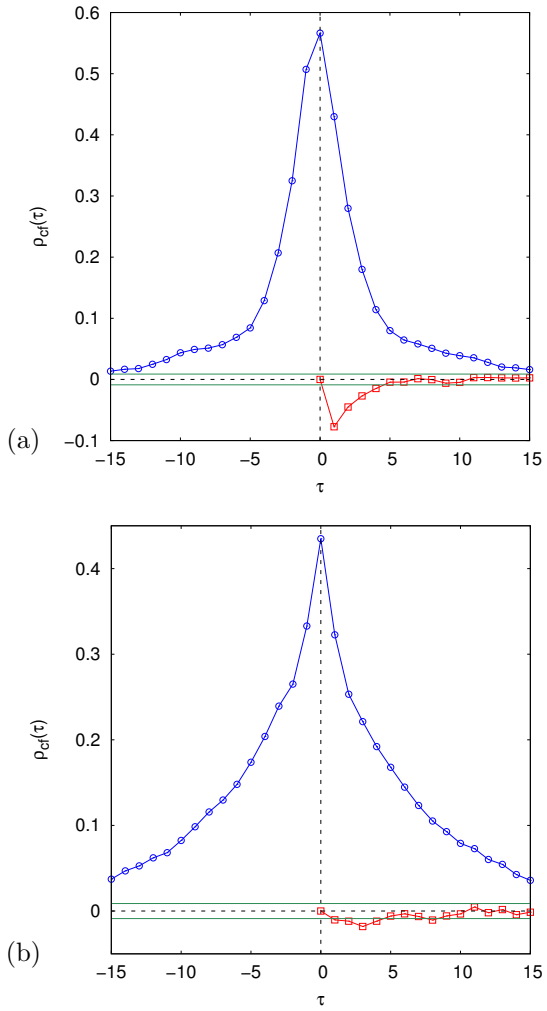


図 3: 時差相関における (a)SOSG の強い非対称性と (b)PRSG の非常に弱い非対称性．

図 3 は，$\Delta t = 10$，$n = 5$ として，50 時間ステップ間隔で $v^c(t)$ と $v^f(t)$ を計算してプロットした $\rho_{cf}(\tau)$ の (a)SOSG（$M = 5$，$B = 9$，$C = 3$，以降固定）と (b)PRSG の結果（100 試行回平均）である．図 3 の (a) では，$\rho_{cf}(\tau)$（青い丸点）は $\tau = 0$ に関して非対称になっているが，図 3 の (b) では，ほとんど非対称性が見られない．また，相関差 $\rho_{cf}(\tau) - \rho_{cf}(-\tau)$（赤い四角点）においても，図 3 の (a) では，$|\tau| < 5$ のとき，価格がランダムウォークする場合の $\rho_{cf}(\tau)$ の 95%信頼区間（2 本の緑の直線の間）から有意に負になっているのに対し，図 3 の (b) では，それほど乖離していない．

これは，図 3 の (a) では，過去の細かいボラティリティと未来の粗いボラティリティの相関 $\rho_{cf}(\tau)$ よりも，過去の粗いボラティリティと未来の細かいボラティリティの相関 $\rho_{cf}(-\tau)$ の方が有意に大きいことを意味する．したがって，SOSG では，時間スケールの粗いボラティリティで細かいボラティリティを予測する方が効果的であり，asymmetry in time scales が創発していると言える．一方，図 3 の (b) では，有意に $\rho_{cf}(\tau) < \rho_{cf}(-\tau)$ となっておらず，PRSG では，その特性が失われている．

なお，PRSG では，SOSG で創発していた leverage effect と呼ばれるもう 1 つの高次の stylized facts についても，再現することができない．Leverage effect とは，未来のボラティリティが過去のリターンと負の相関を持つという性質である [13]．

Asymmetry in time scales などの高次の stylized facts が PRSG で消えてしまうのは，プレーヤーが価格変動を一切考慮せず，確率に従って機械的に取引を行っているためではないかと考えられる．その証拠に，研究 [11] で提示した，価格変動の時間構造に現れる投機的パターンについて調べてみると，SOSG と PRSG の間で異なる結果が得られる．図 4 の (a) と (b) のヒストグラムは，SOSG と PRSG で生成したそれぞれの Δp を以下のルール

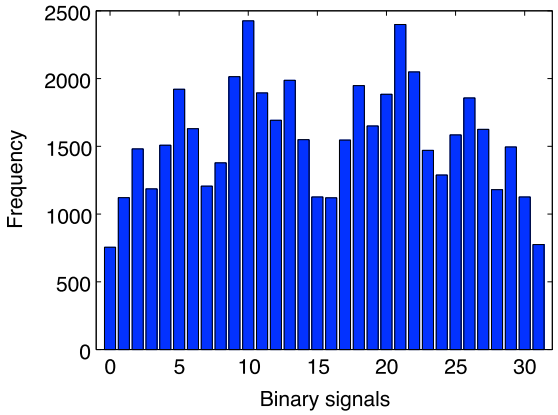
$$b(t) = \begin{cases} 0 & \text{if } \Delta p < 0, \\ 1 & \text{if } \Delta p > 0, \end{cases} \tag{11}$$

に従って 2 進数の時系列 $b(t)$ に変換し[3]，記憶長を $M = 5$ として各履歴シグナルの出現回数を数えてプロットしたものである．図 4 の (a) の SOSG では，10 と 21（2 進数では [0, 1, 0, 1, 0] と [1, 0, 1, 0, 1]）にピークがある対称的な分布になっており，典型的な投機的パターンが見られる[4]．一方，図 4 の (b) の PRSG では，2 つのピークがなく，特徴的なパターンが見られない．図 4 の (a) と (b) における分布形状の違いは，高次の stylized facts の創発には，投機的パターンを内包する価格変動が必要であり，また，投機ゲームタイプのモデルでそのパターンを発生させるためには，価格変動からのフィードバックが働く戦略テーブルによる往復取引の仕組みが重要であることを示唆する．

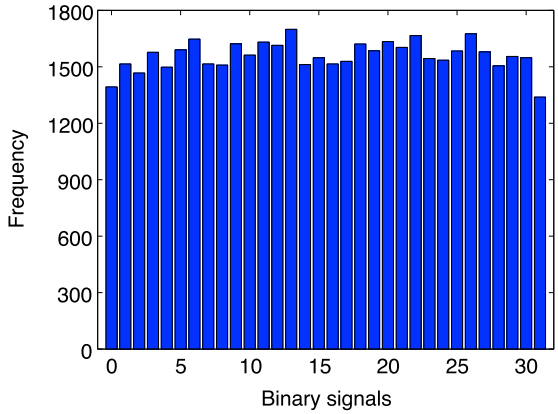
4 むすび

本研究では，SOSG の戦略テーブルによる意思決定の仕組みを，確率によって簡素化した人工市場モデル

[3] $\Delta p = 0$ も時折存在するので，$b(t)$ の時間間隔は一定ではない．
[4] FX の時系列データなどからも，特に時間スケールの小さいとき，このような対称的な二峰性の分布が得られる [11]．



(a)



(b)

図 4: (a)SOSG と (b)PRSG における 2 進数の履歴シグナルのヒストグラム．

「PRSG」を構築した．SOSG よりも主要パラメータの数が 1 つ減っているのにも関わらず，PRSG は，市場規模の自己調整機能を保持しており，また，volatility clustering などの基本的な stylized facts を再現することができることが分かった．

なお, SOSG と PRSG には決定的な違いがあり, asymmetry in time scales などの高次の stylized facts については，PRSG では創発しない．これには，2 進数化した価格変動の時間構造に投機的パターンが見られないことが関係していると考えられる．つまり，プレーヤーが確率的にホールドの長さを決め，機械的に往復取引を行う PRSG では，投機的な価格変動が生じておらず，その結果として，高次の stylized facts が失われてしまっている可能性がある．

謝辞

本研究は，JSPS 科研費 JP17J09156，JP20J00107 の助成を受けたものです．

参考文献

[1] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.

[2] Katahira, K., Chen, Y., Hashimoto, G., & Okuda, H. (2019). Development of an agent-based speculation game for higher reproducibility of financial stylized facts. *Physica A: Statistical Mechanics and its Applications*, 524, 503-518.

[3] Katahira, K., Chen, Y., & Akiyama, E. (2021). Self-organized Speculation Game for the spontaneous emergence of financial stylized facts. *Physica A: Statistical Mechanics and its Applications*, 582, 126227.

[4] Challet, D., & Zhang, Y. C. (1997). Emergence of cooperation and organization in an evolutionary game. *Physica A: Statistical Mechanics and its Applications*, 246(3-4), 407-418.

[5] Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*, 59(4), 381.

[6] Bonachela, J. A., & Munoz, M. A. (2009). Self-organization without conservation: true or just apparent scale-invariance?. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(09), P09009.

[7] Challet, D., Marsili, M., & Zhang, Y. C. (2001). Stylized facts of financial markets and market crashes in minority games. *Physica A: Statistical Mechanics and its Applications*, 294(3-4), 514-524.

[8] Katahira, K., & Chen, Y. (2021). Heterogeneous Round-Trip Trading and the Emergence of Volatility Clustering in Speculation Game. *Journal of Systems Science and Complexity*, 1-24.

[9] Bouchaud, J. P., & Cont, R. (2000). Herd behaviour and aggregate fluctuations in financial market. *Macroeconomic Dynamics*, 2, 170-196.

[10] Katahira, K., & Chen, Y. (2020). An extended Speculation Game for the recovery of Hurst exponent of financial time series. *New Mathematics and Natural Computation*, 16(02), 319-325.

[11] Katahira, K., Chen, Y., & Akiyama, E. The speculative traces in binarized price changes of the

financial markets. *Physica A: Statistical Mechanics and its Applications* (under review).

[12] Gavrishchaka, V. V., & Ganguli, S. B. (2003). Volatility forecasting from multiscale and high-dimensional market data. *Neurocomputing*, 55(1-2), 285-305.

[13] Bouchaud, J. P., Matacz, A., & Potters, M. (2001). Leverage effect in financial markets: The retarded volatility model. *Physical review letters*, 87(22), 228701.

人工市場を用いたショートサイドの市場非効率性に関する分析

Analysis of the market inefficiency on the short side using an artificial market

則武 誉人 [1*]　逸見 龍太 [1]　水田 孝信 [2]
Yoshito Noritake[1]　Ryuta Hemmi[1]　Takanobu Mizuta[2]

[1] 三井住友 DS アセットマネジメント株式会社
[1] Sumitomo Mitsui DS Asset Management Company, Limited
[2] スパークス・アセット・マネジメント株式会社
[2] SPARX Asset Management Co., Ltd.

Abstract: 人工市場モデルを用いて、約定価格がファンダメンタル価格を上回るショートサイド（割高）の非効率性はどのような要因で顕在化するかを分析した。エージェントが出す買いと売りの注文数量に差を設け、買いの数量に対して売りの数量を少なくした場合は、市場非効率性のうちショートサイドの内訳がロングサイド（割安）のそれを上回った。一方で、エージェントの注文価格の決定に用いられるファンダメンタル価格を緩やかに上昇させた場合は、ショートサイドの内訳がロングサイドのそれを上回ることは確認できなかった。また、ショートサイドの市場非効率性が顕在化するメカニズムは、ファンダメンタル価格から一定の範囲で板に提示される売注文の数量が買注文と比べて少なくなり、買注文によるマーケットインパクトが大きくなることで、最良売気配の価格がファンダメンタル価格から乖離することを議論する。

1　はじめに

価格の非効率性は証券（株式）市場に投資するアルファの源泉として位置づけられているが、ロングサイドよりもショートサイドにおいて、より大きい非効率性があることが議論されている。

主な要因の一つは、空売りに関する制約により、割安とされる銘柄への投資よりも割高とされる銘柄への投資（空売り）が少ないためという仮説が挙げられる [1]。例えば空売りするための株式は貸借取引などを通じて調達される[*1]が、貸株に充てられる在庫の不足や調達に伴うコストの上昇など、空売りを行う上で様々な制約が存在する。これら空売りの制約の存在によって、価格が下落することへの懸念が実際の投資に反映されにくいというのがこの仮説の趣旨である。

もう一つの主な要因は、公表される情報の内容に偏向があるためという仮説が挙げられる。例えば [2] は米国企業による開示情報のヒストリカルデータを用いて時系列での類似性を測り、会計報告の用語や構造などの変更がその企業の事後のリターンにネガティブに影響することに着目した。また、[3] は証券会社から出された新規の投資の推奨に関して、売り推奨よりも買い推奨が多いという調査結果を示した[*2]。この背景は、ポジティブな情報は投資家に対して積極的に公表される一方で、ネガティブな情報の公表には経営者やアナリストらが消極的もしくは硬直的になりやすくなることが考えられる [1]。

しかし、現実の価格形成には様々な要素が加わるため、上に挙げた要因が個別に変化した場合にどのような影響があるのかという議論や、影響があるのだとすればそれはどのようなメカニズムなのか考察することは実証研究を介してだけでは難しい。そこで、コンピューターで証券市場を模したエージェントベースモデルを作り検証する、人工市場シミュレーションの手法を用いる。なお、人工市場シミュレーションで用いるモデルを人工市場モデルと呼ぶ。

人工市場モデルでは投資家を模倣した複数のエージェントと、模倣した取引所を設け、シミュレーションその

* E-mail:yoshito.noritake@gmail.com

*1 本邦においては、ネイキッド・ショート・セリング（売付けの際に株の手当てがなされていない空売り）が禁止されているため。

*2 売り推奨の場合と買い推奨の場合で、市場の反応が非対称であることも指摘している。

ものはコンピューター内で完結する[*3]。そして、投資家による売りの発注に制約がある状況と、ポジティブな内容の情報がより多く公表される状況をそれぞれ疑似的にモデル化し、要因（パラメータ）を個別に変化させることで、シミュレーションによって約定した価格とエージェントの注文価格の決定に用いられるファンダメンタル価格の乖離（市場非効率性）のうち、ショートサイド、すなわちファンダメンタル価格を上回る割高側の内訳が取引終了時に大きくなるかを検証する。最終的に、ショートサイドの市場非効率性が顕在化するメカニズムについても考察する。

2 人工市場モデル

本研究では [6] をベースに、分析の目的を果たせる範囲でできるだけシンプルな人工市場モデルを構築した。[4] によれば、多くの場合は複雑なエージェントベースモデルを用いても再現できる stylized fact、すなわち実際の市場にみられる統計的性質の種類が増えないことが報告されている。また、パラメータが多く複雑なモデルほど結果の考察が困難になるため、本研究でも実際の市場を完全に再現することは目的としていない。

同時に、本研究では平常時の市場における取引を想定し、シミュレーションで個別に変化させるパラメータ[*4]は元の値から大きな違いが生じないように決定した。この理由は、本研究の人工市場モデルでの約定方法が後述の通り連続ダブルオークション方式（ザラバ方式）であり、極端なパラメータを当てはめたことで急激な価格変動が頻発するようなシミュレーションになってしまうと、通常のザラバ方式だけでは価格決定メカニズムを再現したとは言い切れないためである[*5]。

空売りに焦点を当てた人工市場シミュレーションの研究は [7] などが当てはまるが、本研究の分析目的のような、約定価格がファンダメンタル価格から乖離する非効率性のうち、ショートサイドとロングサイドの非対称性に着目した研究は見当たらない。また、エージェントの注文価格の決定に用いられるファンダメンタル価格の変化に関しても、[7] のようにある時点で急落をさせる研究は存在するが、平常時に公表される情報の内容に着目したような、より緩やかな変化を取り扱った研究も見当たらない。

[*3] このため、実際の株式市場とは独立した分析である点は留意されたい。
[*4] 後述するパラメータ q と d である。
[*5] 現実の株式市場における特別気配の仕組みを設けるなど、モデルをより複雑にする必要があるとも考えられる。

2.1 取引のプロセス

本モデルは 1 つの証券を取引対象とする[*6]。価格決定メカニズムは売り手と買い手の双方の注文で提示価格が合致すると、同じ数量の注文がその価格で直ちに約定するザラバ方式とする。

投資家として n 体のエージェントを用意し、エージェント $j = 1$ から順番に $j = 1, 2, 3, ...$ と注文を出す。エージェント $j = n$ まで注文を出すと、次の時刻には最初のエージェント $j = 1$ から発注が繰り返される。時刻 t はエージェント 1 体が発注するごとに $t = 1, 2, 3, ...$ と 1 ずつ増える。つまり、注文を出しただけで約定が無い場合も時刻は進む。

エージェント j は注文価格と売り買いの別を以下のように決定する。時刻 t にエージェント j が予想する価格の変化率（予想リターン）$r^t_{e,j}$ は、

$$r^t_{e,j} = \frac{w_{1,j} \log \frac{P^t_f}{P^{t-1}} + w_{2,j} r^t_{h,j} + u_j \epsilon^t_j}{w_{1,j} + w_{2,j} + u_j} \tag{1}$$

ここで、$w_{i,j}$ はエージェント j の第 i 項目の重みであり、シミュレーションで参照する乱数表の決定時にそれぞれ 0 から $w_{i,max}$ までの一様乱数として定める。また、u_j はエージェント j の第 3 項目の重みであり、0 から u_{max} までの一様乱数として定める。log は自然対数である。P^t_f は時刻 t でのファンダメンタル価格、P^{t-1} は時刻 $t-1$ での約定価格（時刻 $t-1$ で約定が無かった場合は、最も近い時刻の価格を参照する。ただし $t = 1$ の時、$P^{t-1} = P^1_f$ とする[*7]。）ϵ^t_j は時刻 t、エージェント j の攪乱項であり、平均 0、標準偏差 σ_ϵ の正規分布に従う乱数である。$r^t_{h,j}$ は時刻 t にエージェント j が計測した過去リターンであり、$r^t_{h,j} = \log\left(P^{t-1}/P^{t-\tau_j-1}\right)$ である。ここで、τ_j は乱数表の決定時に 1 から τ_{max} までの一様乱数としてエージェントごとに定める。

式 (1)（分子）の第 1 項目は約定価格とファンダメンタル価格の乖離に基づき投資を判断するファンダメンタル投資家の成分である。P^{t-1} がファンダメンタル価格と比較して安（高）ければプラス（マイナス）の予想リターンを示す。第 2 項目は過去の価格変化に基づき投資を判断するテクニカル投資家の成分である。$r^t_{h,j}$ がプラス（マイナス）ならプラス（マイナス）の予想リターンを示す。第 3 項目はノイズを表す。

[*6] コーポレートアクションなどによる価値の変化は無いと仮定する。
[*7] 後述する $P^{t-\tau_j-1}$ も同様に扱う

続いて予想価格 $P^t_{e,j}$ は、

$$P^t_{e,j} = P^{t-1}\exp(r^t_{e,j}) \tag{2}$$

として求まる。注文価格 $P^t_{o,j}$ は平均 $P^t_{e,j}$、標準偏差 P^t_σ の正規分布に従う乱数である。P^t_σ は式 (3) で求まる。

$$P^t_\sigma = P^t_{e,j}\sigma \tag{3}$$

売り買いの別は予想価格 $P^t_{e,j}$ と注文価格 $P^t_{o,j}$ の大小関係で決まる。すなわち、

$$\begin{aligned} &P^t_{e,j} > P^t_{o,j}\text{なら買い} \\ &P^t_{e,j} < P^t_{o,j}\text{なら売り} \end{aligned} \tag{4}$$

と定める（ただし、時刻 $t \leq 20,000$ は板構築期間として、式 (4) によらず $P^t_{o,j} > P^1_f$ ならば売り、逆に $P^t_{o,j} < P^1_f$ ならば買いと定める。）。注文価格の変化幅の最小単位（ティックサイズ）は δP とし、それより小さい端数は調整する[*8]。

本モデルではザラバ方式で価格が形成されるため、買（売）注文が提示した価格よりも安い売りの（高い買いの）注文が板に存在すれば、価格優先の原則に沿って順次約定する。時刻 t に出された注文によって約定した価格を P^t とみなす[*9]。注文が全て約定しなかった場合は板に残す。注文が板に残ってから時刻 t_c だけ経過しても約定しない場合はキャンセルされる。また、全てのエージェントは取引が終了する時刻 t_e まで、注文を順番 $(j = 1, ..., n)$ に何回でも出すことができる（キャッシュが無限大）。

2.2 売りの発注に制約がある状況

本研究では [6] からの違いとして、売注文の数量が買注文よりも平均的に少ない状況をシミュレーションできるようにした。なぜなら、キャッシュがあれば買いの発注に制約はないが、空売りには株式を調達する必要があるほか、貸株に充てられる在庫の不足や調達コストの上昇などの制約も存在するため、投資家は買いほど自由に売りを発注できないと考えられるからである。この状況は、人工市場モデルでは買注文に対して売注文の数量が少ないというシミュレーションに一般化することができる。

具体的には、買注文の数量を 100、売注文の数量を式 (5) の通りパラメータ q

$$q = \{100, 99, 98, 95, 90\} \tag{5}$$

とした。$q = 100$ ならば買いと売りの注文の数量に差はないが、q が低くなるにしたがって売注文の数量のみが少なくなる。すなわち、パラメータ q は、市場に発注される売注文の総量がどれだけ減少するかを表す。また、取引（1 回のシミュレーション）の開始から終了を通じて q の値は一定とした。

2.3 ポジティブな内容の情報がより多く公表される状況

また、本研究では [6] からの違いとして、エージェントが注文価格を決定する際に用いられるファンダメンタル価格が時刻の経過とともに緩やかに上昇する状況をシミュレーションできるようにした。具体的には、ファンダメンタル価格 P^t_f を式 (6) の通り定義し、

$$P^t_f = \begin{cases} P^1_f & (t \leq \text{板構築期間}) \\ P^1_f\left(1 + d \times \frac{t-\text{板構築期間}}{t_e-\text{板構築期間}}\right) & (t > \text{板構築期間}) \end{cases} \tag{6}$$

時刻 t_e におけるファンダメンタル価格の上昇率を式 (7) の通りパラメータ $d\,(\%)$

$$d = \{\pm 0, +1, +2, +5, +10\} \tag{7}$$

とした。$d = \pm 0$ であれば公表される情報の内容は均等でファンダメンタル価格も一定となるが、d が高くなるにしたがってポジティブな内容の情報がネガティブな内容の情報よりも多く公表されることを表す。$d > 0$ の場合にファンダメンタル価格が変化するイメージを図 1 に示す。時刻の経過とともにファンダメンタル価格が上昇し、それに沿ってエージェントの予想価格の水準も上昇することになる。

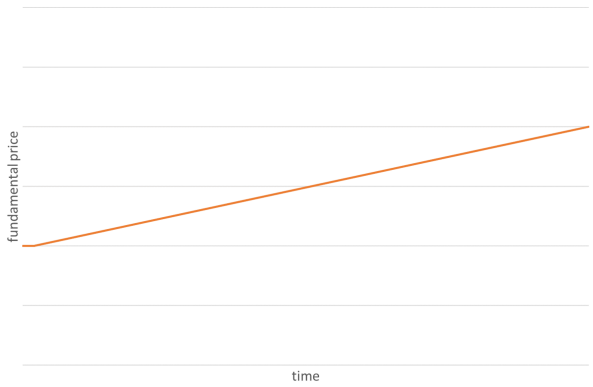


図 1　ファンダメンタル価格の変化 ($d > 0$).

[*8] 売注文なら切り上げ、買注文なら切り下げを行う。
[*9] 一度に複数の注文と約定した場合は、最後の価格を採用する。

2.4 市場非効率性の非対称性を測る指標

本研究では、ファンダメンタル価格と約定価格の乖離におけるショートサイドとロングサイドの内訳を測定するため、[6] に登場した市場非効率性の指標 M_{ie}

$$M_{ie} = \frac{1}{t_e}\sum_{t=1}^{t_e}\frac{|P^t - P_f|}{P_f} \tag{8}$$

に基づき、時刻 t の約定価格 P^t*10がファンダメンタル価格より高い（ショートサイド）場合の市場非効率性 $M_{ie,ov}$、逆に P^t がファンダメンタル価格より低い（ロングサイド）場合の市場非効率性 $M_{ie,uv}$

$$\begin{aligned} M_{ie,ov} &= \textstyle\sum_{t=1}^{t_e}\frac{|P^t - P_f^t|}{P_f^t} \quad (P^t > P_f^t) \\ M_{ie,uv} &= \textstyle\sum_{t=1}^{t_e}\frac{|P^t - P_f^t|}{P_f^t} \quad (P^t < P_f^t) \end{aligned} \tag{9}$$

を新たに定義する。|| は絶対値を表す。$d > 0$ の場合、ファンダメンタル価格の途中の上昇は $|P^t - P_f^t|$ に含まない。$M_{ie,ov}$ と $M_{ie,uv}$ は、M_{ie} と同様に 0 以上の値を取り、0 なら完全に効率的、値が大きくなればなるほど非効率であることを示す。

非効率性が確認された場合、$\log(M_{ie,ov}/M_{ie,uv})$ を算出し、その結果が 0 より大きければショートサイドの内訳が大きいことが、逆に 0 より小さければロングサイドの内訳が大きいことが確かめられる。

3 シミュレーション結果

本研究では、人工市場モデルの所与のパラメータを $n = 1,000$、$w_{1,max} = 1$、$w_{2,max} = 10$、$u_{max} = 1$、$\sigma_\epsilon = 0.06$、$\tau_{max} = 10,000$、$\sigma = 0.003$、$t_c = 20,000$、$\delta P = 10$、$P_f^1 = 10,000,000$ と設定した。また、$t_e = 5,000,000$ とした（各エージェントは 5,000 回ずつ発注する）。

続いて、売りの発注に制約がある状況として、$d = \pm 0$ に固定し、式 (5) の通り q を変化させたパターンと、ポジティブな内容の情報がより多く公表される状況として、$q = 100$ に固定し、式 (7) の通り d を変化させたパターンを（他の条件は全て揃えて）シミュレーションした。これを、結果を安定させるため乱数表を変更して 50 回ずつ繰り返した。シミュレーション後、各パターン毎の $\log(M_{ie,ov}/M_{ie,uv})$ の平均値を求めた（なお、全ての結果でショートサイドとロングサイド双方に市場非効率性があったことを確認した。）。

*10 時刻 t で約定が無かった場合は、最も近い時刻の価格を参照する。

3.1 q を変化させた時の $\log(M_{ie,ov}/M_{ie,uv})$

表 1 q 毎の $\log(M_{ie,ov}/M_{ie,uv})$.

q	100	99	98	95	90
$\log(M_{ie,ov}/M_{ie,uv})$	-0.027	1.827	2.095	3.082	5.765

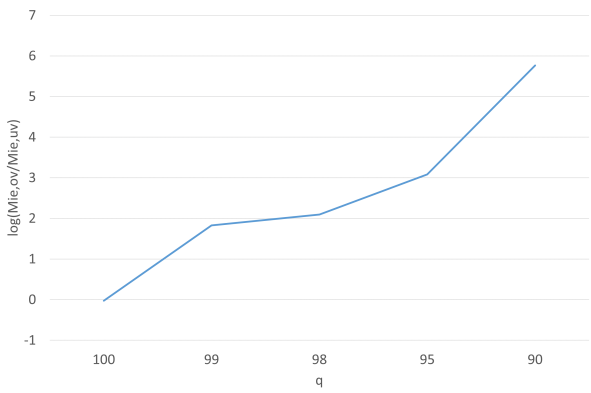


図 2 q 毎の $\log(M_{ie,ov}/M_{ie,uv})$.

パラメータ q 毎の $\log(M_{ie,ov}/M_{ie,uv})$ を表 1 に示す（図 2 は同じ結果の折れ線グラフである。）。

$q = 100$ の場合、市場非効率性におけるショートサイドとロングサイドの内訳に大きな差は無いが、q が小さくなるにしたがって $\log(M_{ie,ov}/M_{ie,uv})$ は大きくなった*11。以上より、市場に発注される売注文の数量の減少はショートサイドの市場非効率性を顕在化する結果を得た。

3.2 d を変化させた場合の $\log(M_{ie,ov}/M_{ie,uv})$

表 2 d 毎の $\log(M_{ie,ov}/M_{ie,uv})$.

d	± 0	$+1$	$+2$	$+5$	$+10$
$\log(M_{ie,ov}/M_{ie,uv})$	-0.027	-0.033	-0.036	-0.054	-0.063

パラメータ d 毎の $\log(M_{ie,ov}/M_{ie,uv})$ を表 2 に示す（図 3 は同じ結果の折れ線グラフである。）。

$d = \pm 0$ の場合、市場非効率性におけるショートサイドとロングサイドの内訳に大きな差は無く、d を高くしてもショートサイドの内訳は大きくならなかった（むしろ、表 2 の結果からは d が高くなるにしたがって $\log(M_{ie,ov}/M_{ie,uv})$ がわずかに小さくなったことが分かる。）。以上より、ファンダメンタル価格の緩やかな上昇

*11 時系列の推移は付録 B を参照のこと。

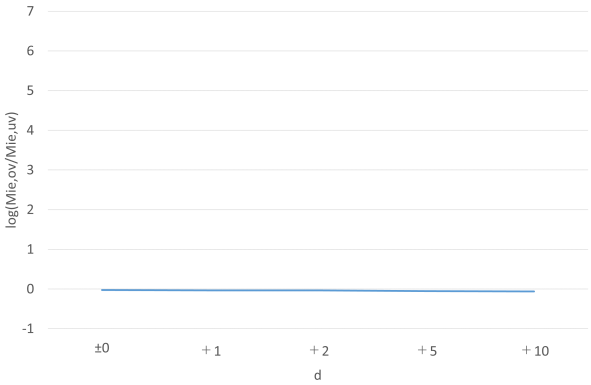


図 3　d 毎の $\log(M_{ie,ov}/M_{ie,uv})$.

がショートサイドの市場非効率性を顕在化する結果は確認できなかった。

4　メカニズムの考察

本章では約定価格がファンダメンタル価格を上回る非効率性が顕在化するメカニズムを議論する。本研究では $\log(M_{ie,ov}/M_{ie,uv})$ の他に、メカニズムの考察のため次の 3 つの指標も算出した。

まず、板の厚さを算出した。具体的には、時刻 t の発注が行われる前に、ファンダメンタル価格 P_f^t から上下 0.6% の範囲に提示された（未約定の）注文を売り買い別に集計する。集計は時刻 100,000 が経過するごとに行い（1 回のシミュレーションで計 50 回集計される）[*12]、計 50 回分の平均値を板の厚さと定義した。

次に、マーケットインパクトを算出した。ザラ場方式での約定は (A) 買注文が出されたことで成立する場合と、(B) 売注文が出されたことで成立する場合の 2 通りがあるが、具体的には、(A) の場合に（時刻 t で）最良売気配の価格が変化した幅 $\Delta\text{Best Ask}^t$、(B) の場合に最良買気配の価格が変化した幅 $\Delta\text{Best Bid}^t$ を用いて、

$$\frac{1}{\#\text{ of execution by Buy Order}}\sum\frac{\Delta\text{Best Ask}^t}{P_f^t}$$
$$\frac{1}{\#\text{ of execution by Sell Order}}\sum\frac{|\Delta\text{Best Bid}^t|}{P_f^t} \tag{10}$$

の通り、全ての約定に対して (A) (B) 別に平均した値をマーケットインパクトと定義した。

続いて、各最良気配とファンダメンタル価格のスプレッドを算出した。時刻 t に発注、約定、注文のキャンセルの処理全てが終わった後の最良売気配の価格を Best Ask^t、最良買気配の価格を Best Bid^t として、全時刻（ただし、板構築期間を除く）の

$$\frac{\text{Best Ask}^t - P_f^t}{P_f^t},\quad \frac{P_f^t - \text{Best Bid}^t}{P_f^t} \tag{11}$$

を平均した。

4.1　q を変化させた時の各指標の結果

表 3　q 毎の板の厚さ.

q	100	99	98	95	90
買い側	88,619	183,407	195,547	239,604	308,107
売り側	90,957	37,388	33,241	18,945	4,577

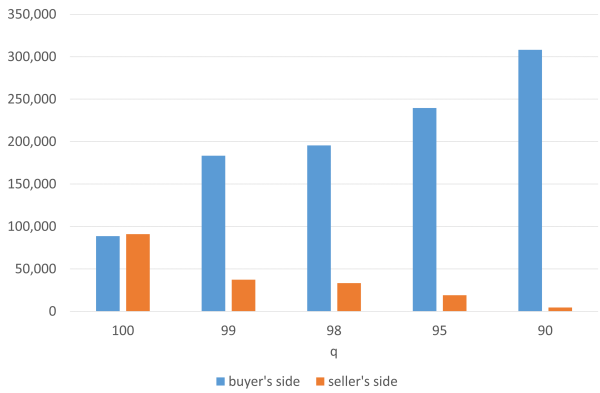


図 4　q 毎の板の厚さ.

まずパラメータ q 毎の板の厚さを買い側（buyer's side）と売り側（seller's side）に分けて集計した結果を表 3 に示す。（図 4 は同じ結果の棒グラフである。）$q = 100$ の場合、板の厚さは買い側と売り側でほぼ拮抗しているが、q が小さくなるにしたがって、ファンダメンタル価格 P_f^t から上下 0.6% の範囲に提示された買注文の数量が増加し、反対に売注文の数量は減少した。

次にパラメータ q 毎のマーケットインパクトを (A) と (B) の場合に分けて表 4 に示す（図 5 は同じ結果の棒グラフである。）。$q = 100$ の場合、マーケットインパクトも (A) と (B) の場合でほぼ拮抗しているが、q が小さくなるにしたがって、買注文による約定が最良売気配に対

表 4　q 毎のマーケットインパクト.

q	100	99	98	95	90
(A) の場合	0.0232%	0.0239%	0.0242%	0.0249%	0.0265%
(B) の場合	0.0232%	0.0204%	0.0203%	0.0199%	0.0193%

[*12] 該当する注文を全て探索する処理を行うため、全ての時刻で集計を行った場合、シミュレーションの所要時間が極めて長くなる。

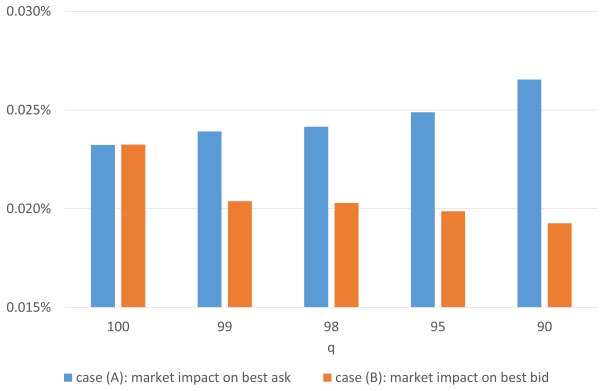


図 5　q 毎のマーケットインパクト.

して与える変化率が大きくなり、反対に売注文による約定が最良買気配に対して与える変化率は小さくなった。

続いてパラメータ q 毎の各最良気配とファンダメンタル価格のスプレッドを表 5 に示す（図 6 は同じ結果の折れ線グラフである。）。q が小さくなるにしたがって、最良売気配はファンダメンタル価格と比べて高くなり、反対に最良買気配はファンダメンタル価格と比べて低くなった。

表 5　q 毎の対ファンダメンタル価格のスプレッド.

q	100	99	98	95	90
$\frac{\text{Best Ask}^t - P_f^t}{P_f^t}$	0.067%	0.315%	0.348%	0.466%	0.694%
$\frac{P_f^t - \text{Best Bid}^t}{P_f^t}$	0.075%	-0.198%	-0.231%	-0.348%	-0.574%

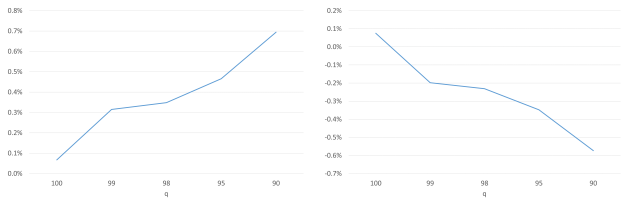


図 6　q 毎の $\frac{\text{Best Ask}^t - P_f^t}{P_f^t}$（左）と $\frac{P_f^t - \text{Best Bid}^t}{P_f^t}$（右）

4.2　d を変化させた時の各指標の結果

表 6　d 毎の板の厚さ.

d	±0	+1	+2	+5	+10
買い側	88,619	90,165	86,974	86,359	82,496
売り側	90,957	92,712	92,194	94,525	98,234

同様にパラメータ d 毎の板の厚さを買い側と売り側に分けて集計した結果（表 6 と図 7）、(A) と (B) の場合

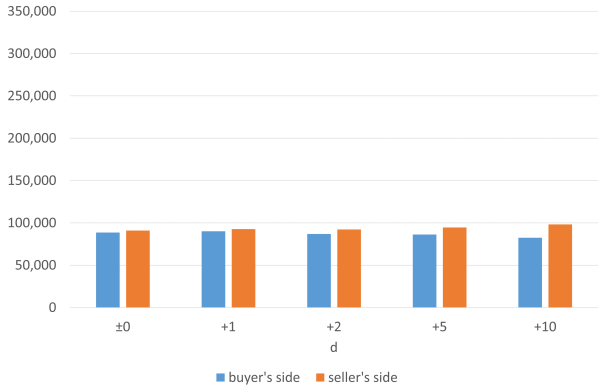


図 7　d 毎の板の厚さ.

表 7　d 毎のマーケットインパクト.

d	±0	+1	+2	+5	+10
(A) の場合	0.0232%	0.0232%	0.0232%	0.0230%	0.0228%
(B) の場合	0.0232%	0.0233%	0.0234%	0.0235%	0.0237%

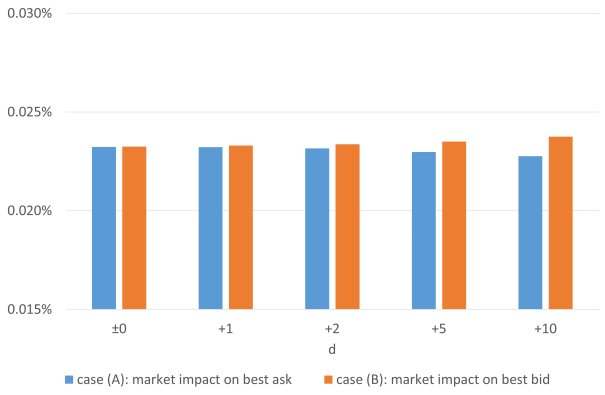


図 8　d 毎のマーケットインパクト.

表 8　d 毎の対ファンダメンタル価格のスプレッド.

d	±0	+1	+2	+5	+10
$\frac{\text{Best Ask}^t - P_f^t}{P_f^t}$	0.067%	0.066%	0.065%	0.062%	0.061%
$\frac{P_f^t - \text{Best Bid}^t}{P_f^t}$	0.075%	0.076%	0.076%	0.079%	0.081%

別のマーケットインパクト（表 7 と図 8）、各最良気配とファンダメンタル価格のスプレッド（表 8 と図 9）をそれぞれ示す。d が高くなった場合は、$d = \pm 0$ の場合と比べて各指標の結果に生じた影響は軽微であった（わずかではあるが、d が高くなるにしたがってファンダメンタル価格 P_f^t から上下 0.6% の範囲に提示された売注文の数量が増加し、売注文による約定が最良買気配に与える変化率が大きくなり、最良買気配とファンダメンタル価格のスプレッドが広がった。）。

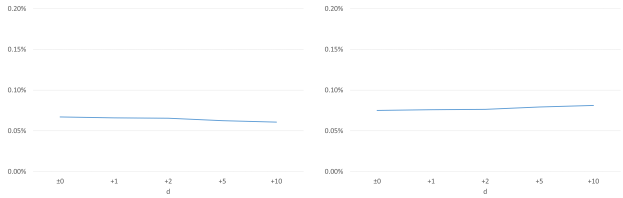


図 9　d 毎の $\frac{\text{Best Ask}^t - P_f^t}{P_f^t}$（左）と $\frac{P_f^t - \text{Best Bid}^t}{P_f^t}$（右）

4.3　ショートサイドの市場非効率性が顕在化するメカニズム

パラメータ q が 100 から小さくなるにしたがって $\log(M_{ie,ov}/M_{ie,uv})$ は大きくなったが、同時に P_f^t から一定の範囲に提示された売注文の数量は減り、買注文による約定のマーケットインパクトが大きく、最良売気配はファンダメンタル価格と比べて高くなった。一方で、パラメータ d が高くなった場合は $d = \pm 0$ の場合と比べて $\log(M_{ie,ov}/M_{ie,uv})$ に大きな変化はなく（むしろ、わずかに減少した)、板の厚さの買い側、売り側の差異、(A) と (B) 別のマーケットインパクト、各最良気配とファンダメンタル価格のスプレッドにも顕著な影響は見られなかった（むしろ q を小さくしたシミュレーションの結果と比べ、わずかに逆の傾向が見られた)。

以上より、本章で示した 3 つの指標がショートサイドの市場非効率性を顕在化するメカニズムと密接に関連していると考えられる。前述した結果と整合する、本研究の人工市場モデルで市場非効率性におけるショートサイドの内訳が大きくなったプロセスは、パラメータ q を小さくしたことでファンダメンタル価格 P_f^t から一定範囲に残る売注文の数量が減少し、エージェントが買注文を（時刻 t に最良売気配 Best Ask^{t-1} 以上の価格で）出した時に最良売気配の変化 ΔBest Ask^t が大きくなりやすくなった、その結果、最良売気配の価格 Best Ask^t がファンダメンタル価格からより高く乖離し、時刻 $t+1$ 以降もファンダメンタル価格 P_f^{t+1} から高い水準で約定しやすくなったことである（反対にパラメータ d を高くしたことでロングサイドの内訳がわずかに増加したプロセスは、直前に約定価格が更新されなかった場合もファンダメンタル価格が上昇し、そこから一定範囲に残る買い注文の数量が減少したことが一連の契機だと考えられる。)。

まとめと今後の課題

本研究では、約定価格がファンダメンタル価格を上回るショートサイドの非効率性が顕在化する要因は、割高な銘柄への投資（空売り）が少ないためという仮説と、投資家に対して公表される情報の内容に偏向がある（ポジティブな内容がより多い）という仮説に着目した。そして [6] の人工市場モデルをベースに、売注文の数量だけを減らした場合とファンダメンタル価格が緩やかに上昇した場合のシミュレーションを行った。同時に市場非効率性におけるショートサイドとロングサイドの内訳を測定し、売注文の数量だけを減らした場合はショートサイドの内訳が大きくなることを確認した（これは [7] の結論とも整合する。)。一方でファンダメンタル価格が緩やかに上昇した場合はそのような結果は確認できなかった。

また、本研究で明らかになったショートサイドの市場非効率性が顕在化するメカニズムは、（注文板で）ファンダメンタル価格から一定の範囲に提示された売り注文の総数量が減少し、買注文（の発注価格が高い場合）のマーケットインパクトが大きくなるため、最良売気配がファンダメンタル価格と比べて高く乖離し、割高な価格で約定しやすくなったことだと考えられる。

今後の課題であるが、まず挙げられるのはファンダメンタル価格が下落するパターン（急落を含む）のシミュレーションであろう。本研究ではファンダメンタル価格の上昇と市場非効率性のショートサイドの内訳が大きくなることの関係は確認できなかったが、もしファンダメンタル価格が約定価格と比べて低い水準に下落した場合はショートサイドの市場非効率性が顕在化することも考えられる。また、ファンダメンタル価格の下落がどの程度であれば、ショートサイドの内訳の方が大きくなるかという閾値の観点からの分析も重要であろう。次に、ショートサイドで発生した市場非効率性が解消される分析も今後の課題である。本研究ではショートサイドの市場非効率性が解消されるプロセスについては言及・考察できていない。

留意事項

本論文の内容は筆者個人の見解です。所属組織の見解を示すものではありません。

参考文献

[1] Gastineau, Gary L., Olma, Andrew R., and Zielinski, Robert G.: Equity Portfolio Management, in *CFA Program Curriculum: level III, 2018*, Vol.4, pp.254-324 (2017)

[2] Cohen, Lauren., Malloy, Christopher., and Nguyen, Quoc.,: Lazy Prices, *The Journal of Finance*, Vol.75, No.3, pp.1371-1415 (2020)

[3] Womack, Kent L.,: Do Brokerage Analysts' Recommendations Have Investment Value?, *The Journal of Finance*, Vol.51, No.1, pp.137-167 (1996)

[4] Chen, S., Chang, C., and Du, Y.,: Agent-based economic models and econometrics, *The Knowledge Engineering Review*, 27(2), pp.187-219 (2012)

[5] Cont, R.,: Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues, *Quantitative Finance*, Vol.1, No.2, pp.223-236 (2001)

[6] 水田 孝信, 則武 誉人, 早川聡, 和泉 潔: 人工市場シミュレーションを用いた取引システムの高速化が価格形成に与える影響の分析, JPX ワーキング・ペーパー, Vol.9, 日本取引所グループ, 2015

[7] 水田孝信, 和泉潔, 八木勲, 吉村忍: 人工市場を用いた値幅制限・空売り規制・アップティックルールの検証と最適な制度の設計, 電気学会論文誌 論文誌 C, Vol.133, No. 9, pp.1694-1700 (2013)

付録 A　stylized fact の再現結果

[6] の人工市場モデルは、代表的な stylized fact として fat tail と volatility clustering を再現するよう設計されている。パラメーター q と d を変化させた場合も、これらの再現結果が損なわれていないか確認した。なお、fat tail の基準はリターンの尖度が正であること[*13]、volatility clustering の基準は二乗したリターンの自己相関が正であることを指す [5] [*14]。 表 9 と表 10 に各統計量の 50 回分の平均値をパターンごとに示す[*15]。個別の結果は省略するが、全ての q と d のパターンで 50 回とも fat tail と volatility clustering を再現した。

[*13] 正規分布の尖度が 0 の時。

[*14] 統計量の算出には時刻 1 が経過するごとのリターン（板構築期間を除く）を用いた。

[*15] $d = \pm 0$ の時、$q = 100$ の場合と同じ結果になる。

表 9　q 毎の stylized fact.

q		100	99	98	95	90
尖度		15.13	14.11	14.09	14.08	14.09
自己相関	Lag1	0.052	0.047	0.046	0.047	0.047
	Lag2	0.052	0.046	0.046	0.046	0.046
	Lag3	0.050	0.044	0.043	0.044	0.044
	Lag4	0.047	0.042	0.042	0.042	0.042
	Lag5	0.044	0.040	0.040	0.040	0.040

表 10　d 毎の stylized fact.

d		+1	+2	+5	+10
尖度		15.12	15.13	15.12	15.13
自己相関	Lag1	0.052	0.052	0.052	0.052
	Lag2	0.052	0.052	0.052	0.052
	Lag3	0.050	0.050	0.050	0.050
	Lag4	0.048	0.048	0.048	0.047
	Lag5	0.045	0.044	0.045	0.045

付録 B　時系列でみた市場非効率性の非対称性の推移

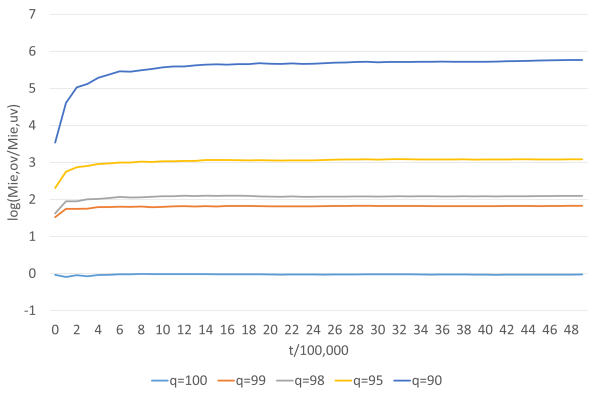


図 10　時刻 100,000 ごとの $\log(M_{ie,ov}/M_{ie,uv})$.

本文で示した $\log(M_{ie,ov}/M_{ie,uv})$ は時刻 $t = t_e$ の結果だが、時刻 100,000 が経過するごとに同様の算出方法で市場非効率性の非対称性を測った結果を図 10 に示す。なお、本研究ではパラメーター q を小さくした場合にショートサイドの内訳が大きくなったため、式 (5) のパターンについて結果を示した。$q < 100$ の場合、時刻 300,000 が経過するまでは非対称性を示す値が逓増し、その後は増加のペースが緩やかになったことが分かる。また、時刻が経過する中でショートサイドの市場非効率性が解消される様子は見られなかった。市場非効率性がショートサイドで発生するプロセスと、それが解消されるプロセスは分かれているとも考えられる。

財市場を内包した Agent-based Computational Economics の構に関する一提案

a proposal for building Agent-Based Economic Model that includes the goods market

高島　幸成 [1]　八木　勲 [2]

Kosei Takashima[1], Isao Yagi[2]

[1] 長岡大学
[1]NAGAOKA UNIVERSITY
[2] 工学院大学
[2] KOUGAKUIN UNIVERSITY

Abstract: Agent-Based Modeling(ABM)アプローチを用い，経済現象を分析する研究手法に Agent-based Computational Economics(ACE)がある．ACE の中でも，財市場を内包するマクロ経済モデルは経済活動を模倣するため，モデル上でエージェントや財などの実体を伴う意思決定の系と，金銭上の取引の系の 2 つを必要とする．そのため，モデルの構造が複雑となり，シミュレーションプログラムの実装が困難になる．本発表は ACE 研究における，モデル構築からシミュレーション実験までの困難性を排除することを目的にして，会計簿記の考え方を拡張した会計関連モデリング提案について述べる．また，会計関連のモデリングを実際に適用して，主要な取引環境として財市場を内包する家計，企業，政府からなるマクロ経済のベースモデルについて紹介する．

1. はじめに

Agent Based Modeling（ABM）の手法を用いて，経済現象の再現やメカニズム，シナリオの分析に用いる手法に Agent-based Computational Economics（ACE）がある．ACE は 2000 年代から研究が活発化するとともに，2010 年以降は急速に研究が行われている分野である．しかし，ACE はモデルが複雑なため，モデリングおよび，シミュレーションのためのプログラミングに困難が伴う．また，その研究成果を公表する段階にあっても複雑さゆえにモデルの記述や説明に困難が伴う．そこで本稿はこれらの問題点を解消する手段として，会計簿記の記述を取り入れたモデリング手法を提案する．

本稿は 2 章で ACE についての概要を述べるとともに，3 章で ACE 研究に取り組む際の困難性・問題点を挙げる．また，4 章でその解決手段の一つとして会計概念を拡張したモデリング手法について述べる．最後に 5 章でモデリング手法を用いたマクロ経済モデルの事例について紹介する．

2.ACE についての概要

ACE は ABM アプローチの浸透に合わせて，徐々に取り組まれるようになった研究領域である．Chen[1]によれば少なくとも 3 名のノーベル経済学賞受賞者が ACE 分野に関わる，もしくは直接的に取り組んでいるとしている．また，Stiglitz[2]も直接的に ACE 研究に携わるなど著名な研究者が取り組む事例も見られる．

経済分野における ABM アプローチの適用は Fagiolo[3]によれば既存の経済学モデルに不満を持つ人々が着目し始めたことがきっかけであるとしている．ABM アプローチを経済分野に適用し，計算機を用いて実験を行う手法に対して，ACE の名称をつけた Tesfasion[3]によると 2000 年以降に特集号が組まれるなど研究が活発化していることが伺える．一方，Gatti[5]によれば実質的に経済学者が本格的に ABM を適用し始めたのは 2007 年の金融危機以降のことであるとしている．この Gatti の主張は金融危機に際し，主流派経済学の主要な分析ツールとなっている DSGE モデルが十分に効果を発揮することができなかったため[6, 7]，その代替先として着目が集まったことが要因であると考えられる．

ACE は進化経済，認知科学，計算機科学の 3 つのコンセプトを統合した研究領域である[8]．そのため，ACE に携わる研究者は，この 3 つの領域から ACE に着目した研究者，及び社会科学に対する ABM アプローチから着目した研究者など多方面にわたる．

結果，それらの研究者が ACE の特徴を挙げる際，その内容は研究者によって異なっている．そこで，先行研究で挙げられている ACE の特徴を参考のためにいくつか提示する．

Fagiolo[3]は ACE の特徴的な点として 1.ボトムアップ視点，2.不均一性，3.進化する複雑なシステムアプローチ, 4.非線形性, 5.直接的な内因性の相互作用，6.限定合理性，7.エージェントの学習，8.「真の」ダイナミクス，9.内因的で永続的な新規性，10.選択ベースの市場メカニズムの 10 の要素を挙げている．

また, Tesfasion[9]は ACE アプローチを特徴付ける 7 つの原則として，1.エージェントの定義，2.エージェントの範囲，3.エージェントの行動に関する性質，4.エージェントの自立性，5.システムの構造性，6.システムの歴史性，7.実験者の制約を挙げている．

これらの定義以外にもいくつかの先行研究で ACE の特徴が挙げられている[5, 7, 10]．これらは重複する箇所と，独自の主張をする箇所が研究者によって分かれている. そこで改めて ACE の特徴をまとめると, 以下の 6 つは際立った特徴であるといえる．

1. 個々の意思決定主体をモデル化するボトムアップ型のモデルであること
2. エージェントが異質であること
3. エージェントが適応的であること
4. エージェントの行動規範に自由性があり，進化する性質を持っていること
5. エージェント間が相互作用すること
6. 時系列を織り込んだボトムアップ型のモデルであること

3.ACE 研究の困難性と問題点

ACE は既存のアプローチに対して，平衡モデルよりも広い範囲の非線型動作を処理でき[10]，ミクロからマクロまでの幅広い経済システムに適用することが可能な[11]，新しい可能性をもたらすアプローチであるといえる．しかし，ACE に取り組もうとする研究者にとって, いくつかの障害が存在している．この障害は ACE 研究過程の，1.モデリング，2.プログラミング，3.シミュレーション，4.結果公表の各段階で生じる．特に個々のエージェントからマクロ経済現象を対象にしようとする，マクロ経済モデルではこの傾向が顕著になる．

3.1 モデリング段階の障害

モデリングの段階については，前述の通り経済構造を模倣するモデルであるため，モデルの内容が複雑になりやすい点が障害となる．この障害が生じる理由として，購買や生産などのイベント時に財などの実体の移動を伴うエージェントの意思決定に関する系と，実際に財を取引した際に行われる金銭上の取引の系の 2 種類の系が併存するためである．

行動に関する意思決定には限定合理性を伴うファジィさが含まれる．一方，その意思決定によって行われた取引には厳密な金銭管理が求められる．つまり，エージェントの 1 つの行動というイベントの中でこれら二つの系を同時に管理する必要があるため複雑性が増加する．

さらに，これらのエージェントの行動は適応的に行動するように構築する必要がある．例えば，何かの支払いを要求される際に，資金の不足が生じないか確認し，ショートするような場合は借入を行う等の例外処理を用意する必要がある．つまり，どのエージェントから行動を起こされても，自身の内部状態がどのようになっていても，システムに破綻が起きないように振る舞わせることが必要となる．そのため，各イベントに対し，相手の反応，例外処理に配慮しながら，意思決定と金銭の 2 つの系を破綻なく管理する必要があり複雑性が高い．

特に，家計の所得が企業の給与から得られ，企業の売り上げが家計の購入から得られるような，資金循環型のマクロ経済モデルを構築する場合は困難になる．これは個々のエージェントの行動を決定するための内部状態が，別のエージェントの行動に準拠し，その行動を決定する要因も他のエージェントの影響を受けるように，行動を決定づける要因が入れ子状になるためである．このため，何かの要素や行動を付け加えるためには，その行動の根拠を，その根拠を作るためにはそれを動機づける別のエージェントの行動を用意しなければならない．そして，これらの行動に破綻回避の例外処理を設けなければならないとともに，多くの場合は金銭上の処理が伴うため，1 つの要素の追加でモデルの拡張に対する作業が煩雑なものになる．

さらにマクロ経済モデルは，個々のエージェントのモデリングの複雑性に加えて，マクロ観点でのモデル構造としても複雑さを抱えている．Caiani[2]は，企業エージェントが倒産して市場から退出した場合に，代わりに生成される企業の資本に裏付けのない状態で参入させているマクロ経済モデルが多いとしている．Caiani はこのモデル構造によって裏付けのない資金が市場全体に流れることによって無視できない影響を及ぼしているとしている．

3.2 プログラミング段階の障害

プログラミングの段階では 3.1 節で示した複雑なモデル構造をシミュレーション可能な状態にプログ

ラミングする必要があり，困難性を伴う．特に変数の数は膨大になりやすく，これらを管理しながらモデルを構築するには困難が伴う．

例えば，ある家計の状態変数を考慮すると，資産を表す変数だけでも，現金，預金，製品，債券，株式などの多岐にわたる．さらに，これらの変数はその数値が生成された時期と，その数値が必要とされる時期がモデル内で大きく異なることが多々ある．例えば，現金のような資産は意思決定時の指標として利用されやすいが，財を購入する際はその時点の残高を必要とする．一方，期首などに投資の意思決定などを判断させる際には，前期首と前期末の残高の差異などを指標にすることがある．そのため，ある程度の期間，状態変数を保有しておかなければならない．

このようなシミュレーションシステムを実装するためには高度なプログラミング能力を必要とするため[12], 研究者の間口を広くすることができない．この問題点に対処するために，Netlogo[13]，EURACE[14]，ASPEN[15]，Jamel[16]，などのシミュレーターやフレームワークが用意されている．しかし，これらは構造が定まっており，研究者の着想を思い通りに具現化できないことや，データ取得に独特の手順を要するなどの制約が存在する．この制約に対して，拡張余地が設けられているが，高いプログラミング能力を要するため参入障壁を取り除き切ってはいない．

3.3 シミュレーション段階の障害

シミュレーション段階ではマクロ現象を観測するために，各エージェントが保有する状態変数を収集する必要がある．この際，個々のエージェントが保有している変数が個別に定義されている場合，情報を収集する機構自体が複雑になる問題点がある．実験段階で出現した現象の詳細を分析するために，別の指標を集計しようとする場合などが顕著である．前述したようにモデル，およびそれを実現するプログラムが複雑なため，新しく情報収集のサブシステムを追加することが困難になり，研究の作業進捗を遅らせることになる．また，マクロ経済モデルでは多くのエージェントから関連する変数を集計して経済指標を計算するため，計算ミスも生じやすい．

3.4 結果公表段階の障害

研究結果の公表を行う際もモデルの複雑性が障害として現れる．上述したように，複数のエージェントがそれぞれ行動ルールと，状態変数を保有する．そのため，モデルのすべてを文章，数式，図で説明すると非常に長大な文章になる．近年，Web ジャーナル等で誌面の制約は取り除かれつつあるが，通常の学会誌などでは誌面制約は取り除かれていない．そのため，モデル記述を簡潔に表記せざるを得ず，読者に対して誤認を与える可能性がある．これは Dosi[7]が ACE の問題点として挙げる再現性の指摘にもつながる．

また，個々の状態，個々の行動，それらの相互作用の結果生じる因果関係の記述方法についても，著者によって書き方が異なることも問題点となる．記述方法や記述の順番が研究によって異なるため，著者の意図した内容が読者に伝わらないことがある．

この問題点に対して，モデル記述の規約を定めるべきであるとして，ODD プロトコル[17]が提唱されている．ODD プロトコルはモデルの説明の記述順序や説明されるべき事項をカテゴリ別に定めている．そのため，研究者毎の記述方法の違いを小さくすることができている．しかし，ODD は ACE だけではなく, ABM や複雑適応系モデル全般に対して広く適用できるように考案されている．そのため，記述規則に余裕を持たせており，個別のカテゴリ内での記述には，著者による違いが生じている．さらにマクロ経済を対象とした ACE では, モデル内のエージェントの種類と相互影響が多岐にわたるため，それぞれの関係を認識しづらいという問題点もある．

4.会計関連を用いたモデリング

ACE 研究を実行する上で，複式簿記の概念導入はこれらの障害を取り除くことができる．複式簿記はストックとフローの値の変化を，5 つの科目グループと，借方貸方という記述位置の規則によって，同時に記録する構造を持っている．そのため，金銭上の系における状態変数の記録や管理において優れた効果を発揮する．

この複式簿記的なストックとフローの表現方法の導入は，出口[18]が会計記法をさらに発展させ交換代数によるモデリングを提示している．また，Caiani[2]は各セクター間のストックとフローを会計記述で正確に表現することで，モデル内に生じる誤差をなくすことができるとしている．

これらはモデルの正確性について，会計の技術を用いた研究である．本研究では，これに加えて ACE モデルへの会計概念の導入が，研究作業の効率化にも効果的な役割を果たすことができると考える．具体的には 1.複式簿記の導入による金銭上の系の自動化と，2.会計簿記の記法をモデル全体に拡張することによるフレームワーク化の 2 段階が挙げられる．

4.1 複式簿記の自動化

複式簿記の導入による金銭上の系の自動化は，シ

ミュレーションプログラムの効率化を図ることができる．これは複式簿記の記録方法を保有する会計オブジェクトを各エージェントに保有させることで実現できる．会計オブジェクトはエージェントの行動に応じて，金銭上の取引にかかわる処理を記録，管理する．例えば，家計エージェントが消費のために製品の購入を行った際には，現金（資産）を減じ，消費（費用）を増やす．

複式簿記は金銭上の変化を記録する上で，厳密で強力な状態記述の方法である．この記述方法は金銭的な変化において，どのエージェントにも適用できる．そのため，仕訳，総勘定元帳への転記，決算までの一連の処理を自動化することで，すべてのエージェントで共通で利用できる．また，計算機シミュレーションであるため，仕訳から総勘定元帳への転記までについては取引の発生時に同時に行うことができる．総勘定元帳は期首，現在，期末の残高を保有するため, 金銭上の状態変数として参照ができる．つまり，複式簿記の記法を利用して，システムに落とし込むことで，金銭上の系に関して，厳密で正確な変数管理と，それを実現するための機構を利用した状態変数の保管，各エージェントでの機能の共通化を果たすことができる．

さらに，この機構を利用すると安全に変数を分割することが可能になる．例えば，企業エージェントの法人税の支払いについて，期末の決算で金額が確定した後，支払いが翌期に行われる場合を考える．法人税の支払いが他のエージェントとの行動の兼ね合いから，給与や原料購入などの支払いを行った後に行われると仮定したとき，保有している資金がプログラム内でショートしないようにするためには，常に法人税の支払いを考慮させる必要が生じてくる．モデルが複雑になると，意思決定時と実際に金銭上の取引が発生する時期がずれ，その間に現金のようなストック変数は変質する．この対処法としては金銭の変数を分割する，意思決定時の計算で法人税額を差し引く処理を行う等があるが，構造が煩雑になってしまう．このような，時間差のある変数の利用についても，会計簿記を利用したデータ管理が有効になる．後の支払い義務が生じた場合は，その時点で法人税支払予算金等の現金を一時保管する資産科目を作成し，現金科目から振り替えておくことで，安全に変数の分割を行うことができる．

このため，金銭上の系の管理については会計オブジェクトの導入によってほぼ自動化ができ, 3.2 節で挙げたプログラミング上の困難性を排除することができる．

4.2 フレームワークへの拡張

会計簿記をプログラミングの自動化にとどめず，モデル全体に拡張することで，さらなる作業効率を図ることが可能になる．

ここである企業エージェントが家計に対して，現金で給与を支払った場合を考える．仕訳は表 1 のような処理になる．尚，ここではモデル内の状態変数変化の記述を目的とするために，実際の簿記の記述とは異なり，日付と金額を記載していない．また，資産，負債，資本，売上，費用の勘定科目の 5 つのグループを科目種として表示している．さらに，摘要を中央に位置させている．

表 1.企業エージェントの給与支払いの仕訳

科目種	借方科目	摘要	貸方科目	科目種
費用	給与	固定給与支払	現金	資産

表 1 の 2 行目は給与（費用）が増加し，現金（資産）が減少していること，及び摘要欄でどのような行動によってこの増減が生じたかを 1 行で示している．尚，仕訳の際の勘定科目 5 つのグループと借方貸方の配置による値の増減については簿記の基礎知識であるため説明を省略する．

この記述は，単体で見ると企業エージェントのストックとフローの状態変数の変化を記録したものである．しかし，モデル全体の観点から見ると，この仕訳の発生は，2 者間の取引を伴うイベントが発生している可能性を意味する．つまり，ある仕訳はモデル内では，別のエージェントの仕訳による状態変数の変化を発生させていることになる．本研究では会計上の仕訳同士の関係を，モデル内のエージェント間の関連として定義することを会計関連とする．

会計関連は 1 つの取引で生じる 2 者の仕訳をつないだものである．そこで，会計関連を簿記と同様に左右に以下の 1～4 の規則を設けて設置する．すると，モデル内で金銭上の変化が生じるイベントの関係を一意に表形式で表す会計関連表としてまとめることができる．

1. 各仕訳の内容は摘要名と変化する科目，科目の種類とする（左から借方科目種類, 借方科目名，摘要，貸方科目名，貸方科目種類,）
2. このイベントで能動的に行動するエージェントの名称と仕訳を左側に設置する
3. このイベントで受動的に行動するエージェントの名称と仕訳を右側に設置する
4. このイベントの名称（エージェント間の関連名）を中央に配置する

現在，作成しているマクロ経済のベースモデルで使用している会計関連表の一部を表 2 に示す．

表 2.会計関連表のサンプル

a

	エージェント	科目種	借方科目	摘要	貸方科目	科目種	会計関連	科目種	借方科目	摘要	貸方科目	科目種	エージェント
①	企業	費用	給与	固定給与支払	現金	資産	**給与支払い**	資産	現金	給与受取	受取固定給	収益	家計
②	家計	資産	預金	預金預入	現金	資産	**預金預け入れ**	資産	現金	預金預かり	預かり預金	負債	銀行
③	家計	費用	消耗品	消耗品購買	現金	資産	**家計消耗品購買**	費用	売上原価	製品販売	製品	資産	企業
								資産	現金	製品販売	製品売上	収益	企業

b1　　b2

例として①では企業エージェントが給与を支払うための行動イベントを起こし，給与（費用）が発生し，現金（資産）が減少している．この行動の結果，家計エージェントの現金（資産）が増加し，受取固定給与（売上）が発生している．そしてこれらのエージェント間のイベントの関連を「給与支払い」と名付けている．

会計関連表では，表 2 の a を見ると能動的な主体と受動的な主体のエージェント種類と，その関係を示す定義を見ることができる．b1，b2 にはそれぞれのエージェントの仕訳を組み込んでいる．このため，エージェント間の関係と，その関係による行動によって変化する各エージェントの金銭上の状態変数を一覧化することができる．これはモデルの設計図として利用でき、モデル構築に先駆け会計関連表を定義することで, 3.1 節で示したモデル構築の困難性を軽減することが可能となる．

また，会計関連表はプログラムと連動させることで, 3.2 節で挙げたプログラミング時の作業負荷をさらに減らすことができる．会計関連表を読み込み，プログラム内に会計関連と，関連で用いられる仕訳をリストで保有させ，関連名から仕訳を読み込むことで金銭上の処理を自動化させることができる．金銭上の処理を複式簿記の方式で行うため、新しい行動を追加しても、会計関連表で関係を定義するだけで金銭上の系については新しいコーディングを必要としなくなる．

さらに，個々のエージェントが保有している会計オブジェクトは 4.1 節で述べた通り，科目ごとの総勘定元帳を保有している．この科目の属性と，エージェントの属性を組み合わせることで，当該科目が産業連関表などの経済指標を集計する際にどの値を構成しているかを識別することができる. そのため，シミュレーション実験中に新たに必要になったデータの収集が容易になり, 3.3 節で挙げた障害を軽減することができる．

最後に，会計関連表は成果公表の記述について大きな効果を発揮する. 会計関連表は, 誰が何を行い，状態変数がどのように変化するかを，誤認しにくいモデル記述になっている．これは複式簿記で利用される科目の概念と，借方貸方の配置による変数変化の規則を利用するためである．そのため，ACE 関連のモデル記述の規約に会計関連表を用いることで，金銭上の系に関して，エージェント間の関係についての誤認や齟齬はなくなるといってよい．

さらに会計関連はこれらの効果に加え，Caiani が指摘するようなモデルの正確性を確保するためのセクター間の会計記述を支援する機構となる．また，Netlogo などの既存のモデリングツールにも構造上，追加することが可能である．加えて，ODD プロトコルについても ACE 研究については拡張的に記載することができる．これらはすべて先行研究の取り組みを阻害することなく，拡張的にモデリングやプログラミング，シミュレーションの作業効率を向上することが可能となる．

5.会計関連を用いたモデル構築事例

現在, 会計関連を用いて ACE 人工経済モデルのベースモデルを構築している．ベースモデルの概要を図 1 に示す．ベースモデルは必要なモデルに対して拡張を想定しているため，図 1 は事例としての構造である．例えば，企業群の中で小売企業と製造企業が財と購買金を取引している．このモデル内に，さらに原材料製造企業を追加し，原料と購買金を取引するような構造に変更することもできる．

これらのエージェント間では意思決定と行動に伴って金銭上の取引が行われる．図 2 に例示するように購買イベントの際，イベントに関係するエージェントが保有する会計オブジェクトにメッセージが送られる．会計オブジェクトは会計関連表にメッセー

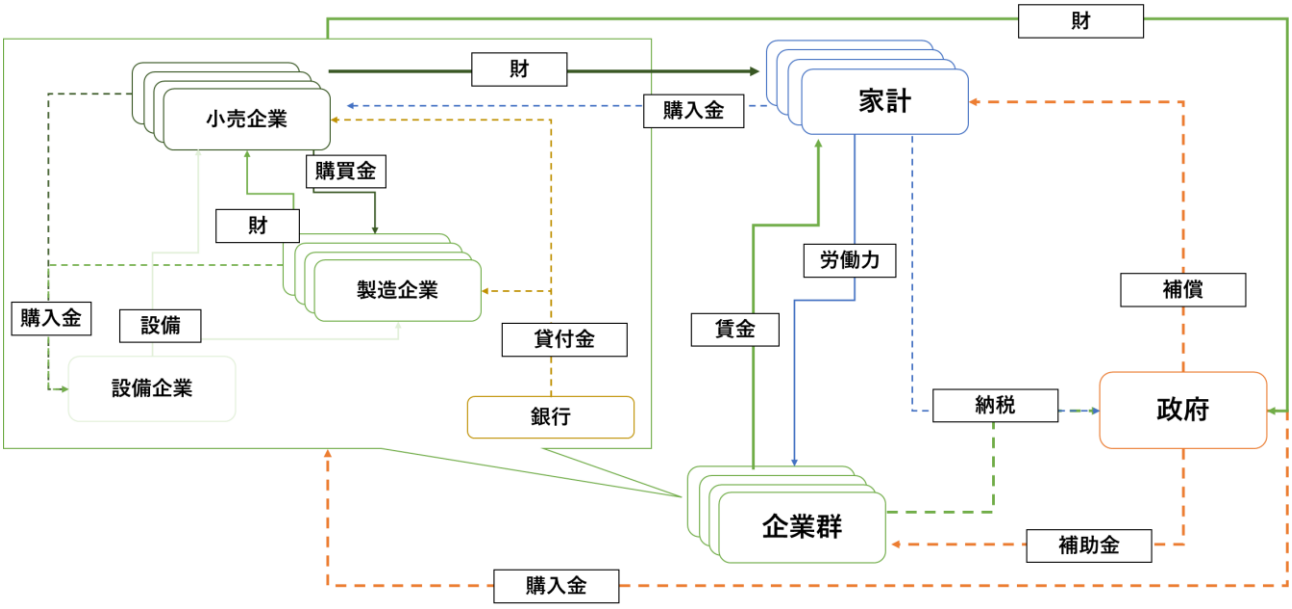


図 1.マクロ経済ベースモデルの概念図

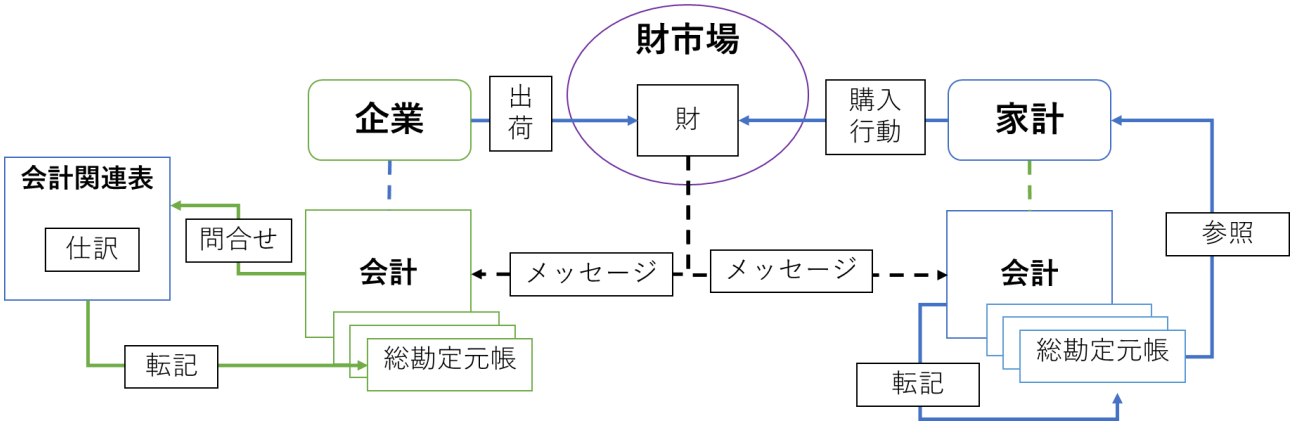


図 2.取引時のエージェントと会計オブジェクト

ジを問い合せ，仕訳を受け取り，総勘定元帳に転記を行う．

会計関連表に仕訳の内容を保管しておくことで，エージェント間の関係（図 2 の例でいえば「消耗品購買」）をメッセージとして問い合わせると，仕訳と転記の方法を返す仕組みとなっている．また，そのようにして随時情報を更新している総勘定元帳に問合わせることで，各エージェントは金銭上の状態変数を認識することができる．図 2 では家計が購入行動をした際に，影響を受ける財から各エージェントの会計オブジェクトにメッセージを発している．エージェント間で金銭以外の実体が移動する際は実体を表すオブジェクトにメッセージを送らせると管理が容易になる．

6.今後の課題

本稿では ACE 研究遂行の上での問題点を挙げ，それらを解決するために会計関連手法を提案した．また，会計関連を用いて構築している ACE 人工経済モデルのベースモデルについて紹介を行った．今後の課題として，構築中のベースモデルを構築すること，及びその過程でエージェントの属性と科目の属性からマクロ経済指標への変数の関係を整理することが挙げられる．次に会計関連表を設定するだけで，ACE のマクロ経済モデルの金銭上の系の部分までを自動的に生成するシステムの構築を行っていくことを予定している．

謝辞

本研究は，公益財団法人 全国銀行学術研究振興財団の助成を受けた．

参考文献

[1] Chen, Shu-Heng, and Ying-Fang Kao. "Herbert Simon and agent-based computational economics." *Minds, models and milieux*. Palgrave Macmillan, London, 2016. 113-144.

[2] Caiani, Alessandro, et al. "Agent based-stock flow consistent macroeconomics: Towards a benchmark model." *Journal of Economic Dynamics and Control* 69 (2016): 375-408.

[3] Fagiolo, Giorgio, and Andrea Roventini. "Macroeconomic policy in DSGE and agent-based models redux: New developments and challenges ahead." *Available at SSRN 2763735* (2016).

[4] Tesfatsion, Leigh. "Agent-Based Computational Economics: Overview and Brief History." (2021).

[5] Gatti, Domenico Delli, and Jakob Grazzini. "Rising to the challenge: Bayesian estimation and forecasting techniques for macroeconomic Agent Based Models." *Journal of Economic Behavior & Organization* 178 (2020): 875-902.

[6] Stiglitz, Joseph E. "Where modern macroeconomics went wrong." *Oxford Review of Economic Policy* 34.1-2 (2018): 70-106.

[7] Dosi, Giovanni, and Andrea Roventini. "More is different... and complex! the case for agent-based macroeconomics." *Journal of Evolutionary Economics* 29.1 (2019): 1-37.

[8] Tesfatsion, Leigh. "Introduction to the special issue on agent-based computational economics. " *Journal of Economic Dynamics and Control* 25.3-4 (2001): 281-293.

[9] Tesfatsion, Leigh. "Modeling economic systems as locally-constructive sequential games. " *Journal of Economic Methodology* 24. 4 (2017): 384-409.

[10] Farmer, J. Doyne, and Duncan Foley. "The economy needs agent-based modelling." *Nature* 460.7256 (2009): 685-686.

[11] Tesfatsion, Leigh. "Agent-based computational economics: A constructive approach to economic theory." *Handbook of computational economics* 2 (2006): 831-880.

[12] Tesfatsion, Leigh. "Agent-based computational economics: Growing economies from the bottom up." *Artificial life* 8.1 (2002): 55-82.

[13] Tisue, Seth, and Uri Wilensky. "Netlogo: A simple environment for modeling complexity." International conference on complex systems. Vol. 21. 2004. https://ccl.northwestern.edu/netlogo/index.shtml

[14] Barton, Dianne C., et al. "Aspen-EE: an agent-based model of infrastructure interdependency." *SAND2000-2925. Albuquerque, NM: Sandia National Laboratories* (2000).

[15] Dawid, Herbert, et al. "The eurace@ unibi model: An agent-based macroeconomic model for economic policy analysis." (2012).

[16] Seppecher, Pascal. "Flexibility of wages and macroeconomic instability in an agent-based computational model with endogenous money." Macroeconomic Dynamics 16.S2 (2012): 284-297.

[17] Grimm, Volker, et al. "The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism." *Journal of Artificial Societies and Social Simulation* 23.2 (2020).

[18] 出口 弘. "複雑系としての経済学: 自律的エー ジェント集団の科学としての経済学を目指して", 日科技連出版社, (2000).

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 Yoshida[1]　Takuma Nakamichi[1]　Riku Tanaka[2]　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 Explainabe 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けたい．

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

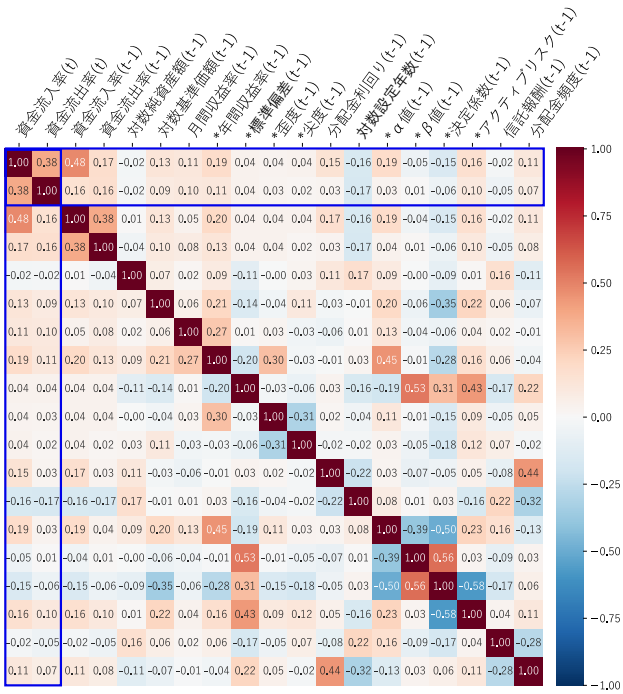


図 1: 本稿に用いる変数間の相関行列

各ファンドのインデックスを i とし，t 月におけるファンド i の資金流入率 $y_i^{\mathrm{in}}(t)$ と資金流出率 $y_i^{\mathrm{out}}(t)$ を，以下のように定義する．

$$y_i^{\mathrm{in}}(t) = \frac{y_i^{\text{設定額}}(t)}{y_i^{\text{純資産総額}}(t)}, \quad y_i^{\mathrm{out}}(t) = \frac{y_i^{\text{解約額}}(t)}{y_i^{\text{純資産総額}}(t)}$$

これらへの影響を分析するために，他の変数 x には1ヶ月のタイムラグを設ける．図 1 に本稿で使用する変数間の相関行列を示す．なお $*$ 印を付した変数は直近 12ヶ月分の月間収益率に基づいて算出した．さらに付録 A に各変数の詳細を記載する．

資金流入率と資金流出率には正の相関があるため，先行研究が指摘するように短期的な回転売買の存在が示唆される．さらに，資金流出率は月間収益率や年間収益率とも正の相関があるため，確かに運用成績が良いファンドほど売られている．これでは運用会社の努力に対するインセンティブが生じない [6]．設定年数においては，流入と流出ともに負の相関があり，古いファンドほど流入出が減少し注目されなくなることが判る．

次に 1 例として，運用パフォーマンス（月間収益率，年間収益率，α 値）と資金フロー（流入率，流出率）の相関図を図 2 に示す．いずれも目視で認識できるような明確な関係性は見受けられない．そのため先行研究においては，重回帰分析による偏回帰係数の t 検定が主流であった．しかし本稿では異なるアプローチとして，非線形な機械学習と SHAP の活用により，可視化表現の高感度化を目的とする．

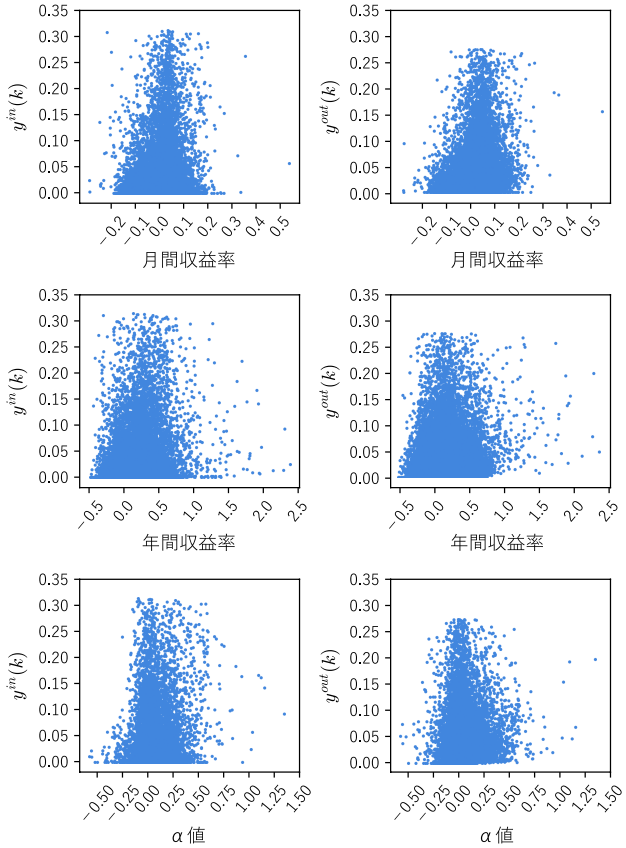


図 2: 相関図による不明瞭な可視化表現

3 提案手法

3.1 非線形モデリング

文献 [5] と同様に資金流入と資金流出に分けて，目的変数 $y(t)$ と説明変数 $x(t-1)$ の関係 $\boldsymbol{F}$ を推定する．なお図 1 により，資金フロー $y_i^{\mathrm{in}}(t)$ と $y_i^{\mathrm{out}}(t)$ は前月 $t-1$ と正の自己相関（モメンタム）を持つ．資金フローはファンドの人気・不人気も関係し，正の自己相関を持つことは比較的自明なので，前月 $t-1$ の履歴もコントロール変数として投入する．

$$y_i^{\mathrm{in}}(t) = \boldsymbol{F}\left[x_{i,1}(t-1), x_{i,2}(t-1), \cdots, y_i^{\mathrm{in}}(t-1)\right] \quad (1)$$
$$y_i^{\mathrm{out}}(t) = \boldsymbol{F}\left[x_{i,1}(t-1), x_{i,2}(t-1), \cdots, y_i^{\mathrm{out}}(t-1)\right] \quad (2)$$

ここで i はファンドのインデックス ($i = 1 \sim 670$) であり，$x_{i,j}$ は j 番目の説明変数である．なお $y_i^{\mathrm{in}}(t)$ や $y_i^{\mathrm{out}}(t)$ の上位 1%を外れ値とみなし，対応する目的変数と説明変数のデータセットを除外した．

次に $\boldsymbol{F}$ の推定にふさわしい機械学習モデルを模索すべく，Python の機械学習ライブラリーで提供されている代表的な 24 モデルを比較する．なお比較のため，前処理として変数毎に平均 0 かつ分散 1 となるように標

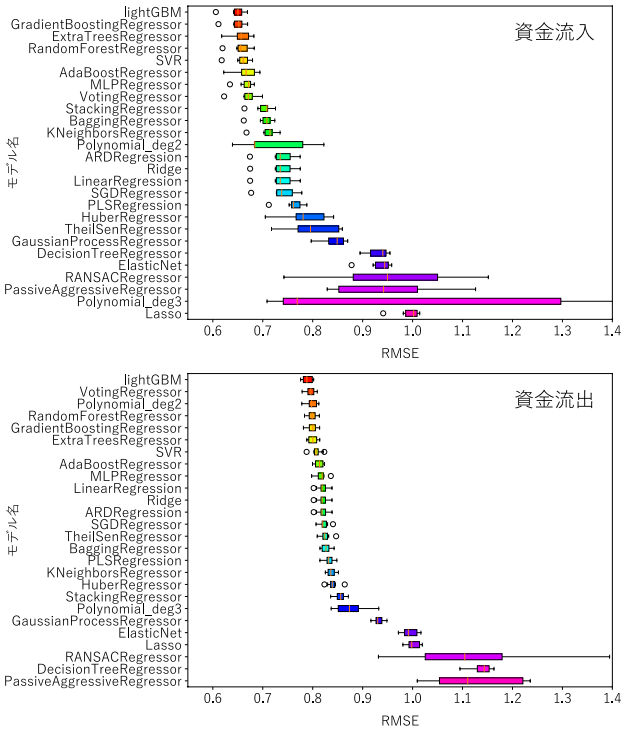


図 3: 代表的な機械学習モデル (24 種) による RMSE

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 $\boldsymbol{F}$ 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．

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

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，$\hat{\boldsymbol{F}}$ は機械学習によって推定された関数 $\boldsymbol{F}$ である．つまり $\phi_j(k)$ は，$\hat{\boldsymbol{F}}$ 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．

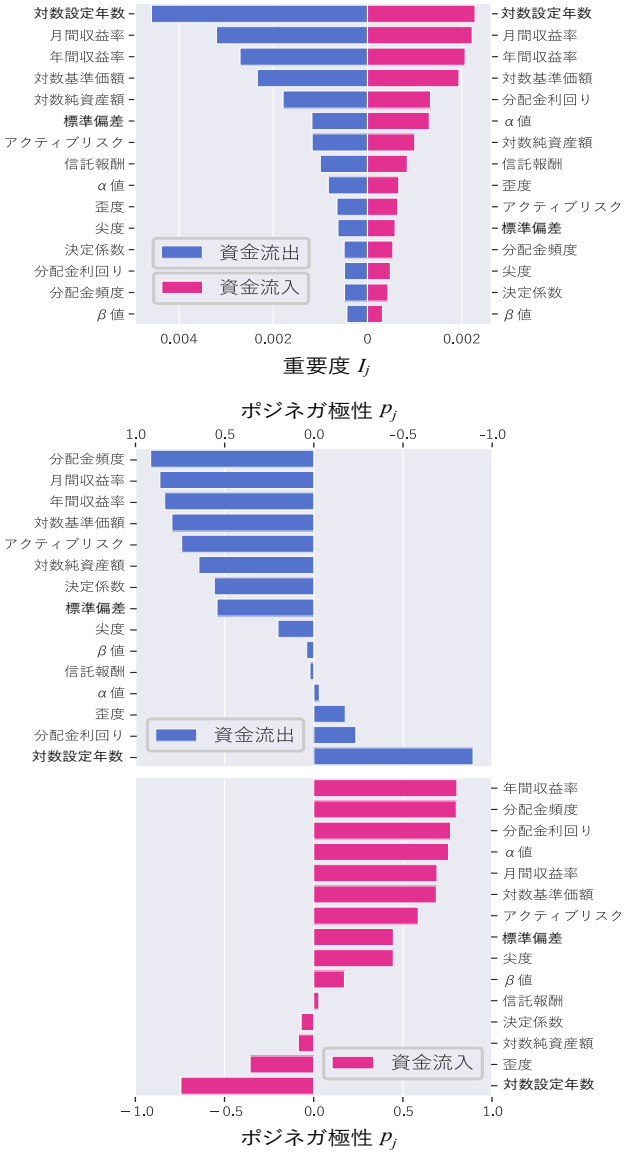


図 4: 説明変数の重要度 I_j とポジネガ極性 p_j

ドにおいては償還期限を 4, 5 年に設定する場合が多く，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充分な学習サンプル数 K を確保すべく，5 年間を 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

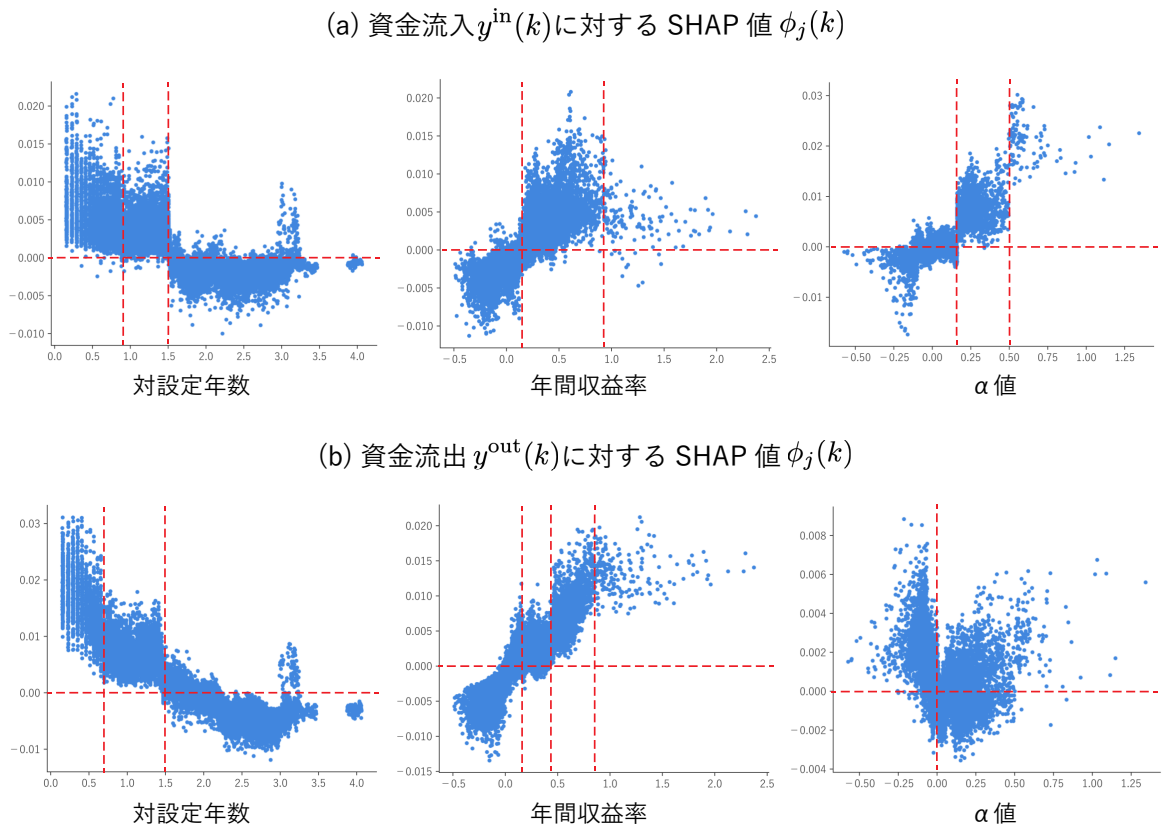


図 5: SHAP による非線形性の可視化

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，月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

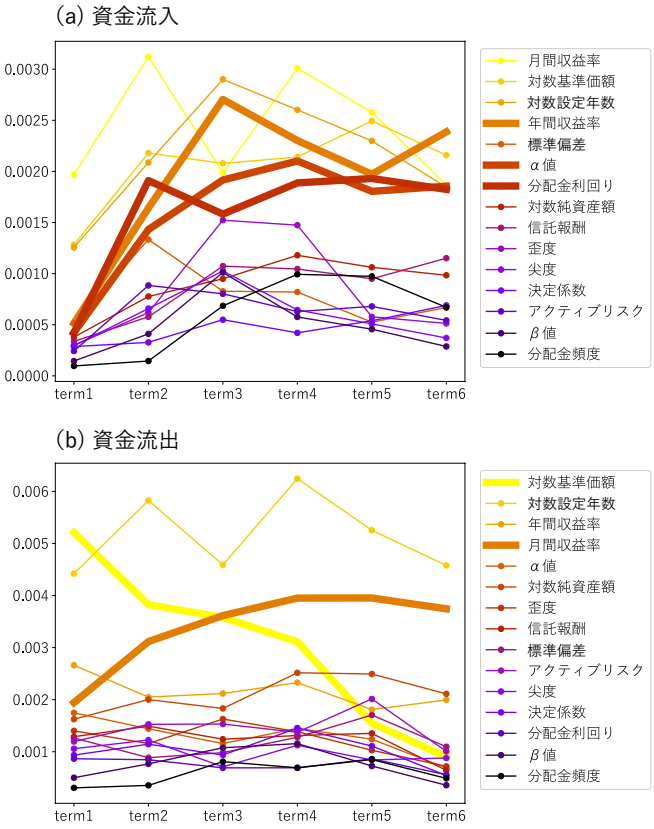


図 6: 重要度 I_j の時代変化

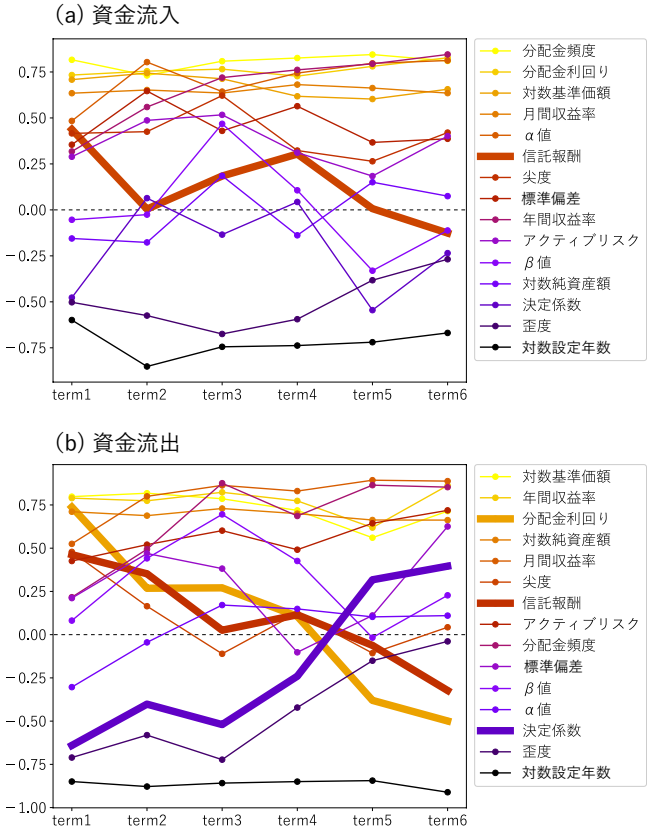


図 7: ポジネガ極性 p_j の時代変化

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ずら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ろう．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.

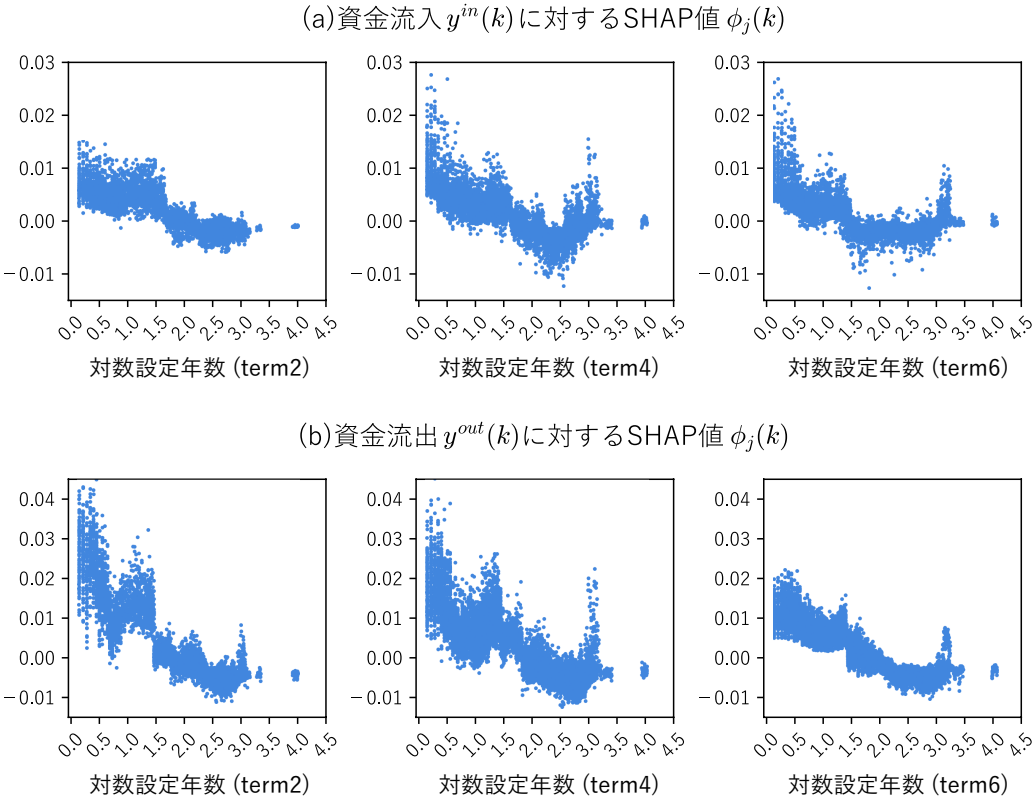


図 8: SHAP による不連続変化 (対数設定年数) の可視化

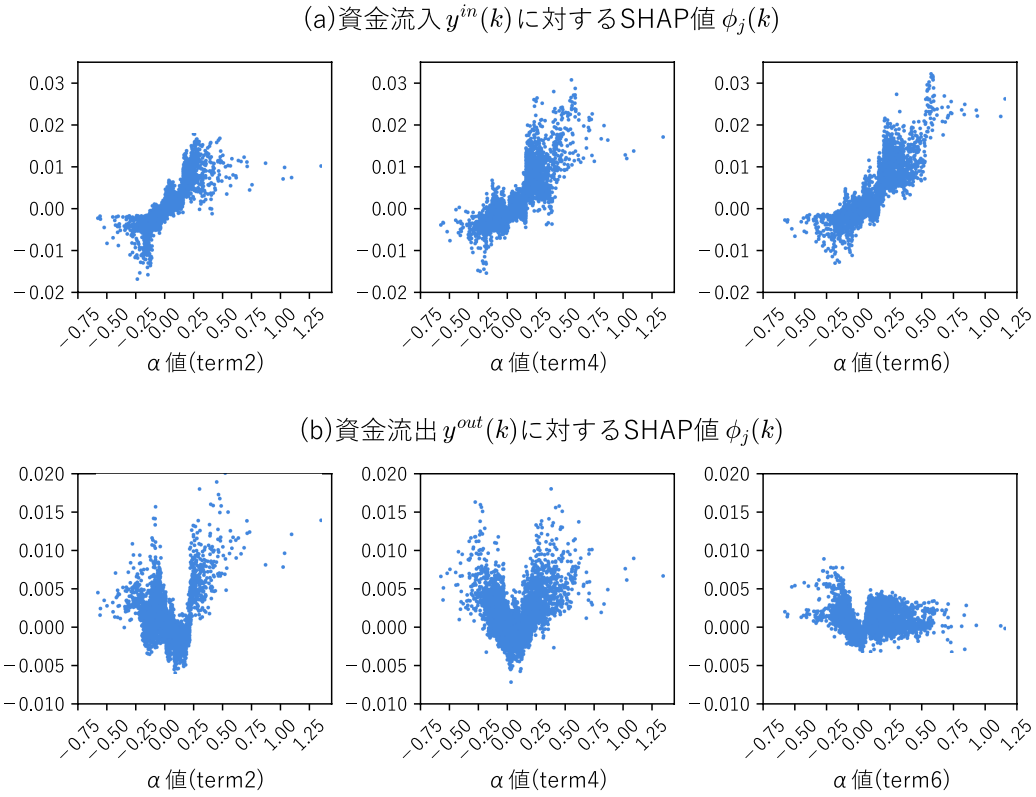


図 9: SHAP による非線形性 (α 値) の可視化

[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.

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．

$$\text{月間収益率} = \frac{\text{当月末基準価額} + \text{分配金} \cdot \frac{\text{当月末基準価額}}{\text{月中分配落基準価額}} - \text{前月末基準価額}}{\text{前月末基準価額}}$$

・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し，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．

外国為替証拠金取引における解釈可能な売買戦略の構築

Constructing an Interpretable Trading Strategy for Forex Trading

内田　純平[*1]　　穴田　一[*1]
Jumpei Uchida　　Hajime Anada

[*1] 東京都市大学
Tokyo City University#1

Recently, many researchers have studied foreign exchange trading using technical analysis. However, it is difficult to achieve profitability using this technique. Therefore, using Genetic Network Programming, we construct a model that considers the technical index signal strength for devising a profitable trading strategy. Finally, we confirmed the effectiveness of our model using historical data of the exchange market.

1. はじめに

近年, 株式売買や外国為替証拠金取引に関する売買戦略を構築する研究が精力的に行われている. 機械学習を用いた売買戦略を構築する際に, テクニカル分析[1]が多く用いられている. テクニカル分析では, 過去の時系列データを数理的に扱うことで算出される様々な種類のテクニカル指標を用いる. テクニカル指標は, 相場のトレンドを表すトレンド系テクニカル指標と相場の転換を表すオシレーター系テクニカル指標に大きく分けることができ, それぞれ売買のタイミングを知らせる売買シグナルと言われるものがある. トレンド系テクニカル指標は, 設定日数の異なるトレンド系指標の交差で売買シグナルを表すことが多い. 例えば、2 つの移動平均線を使って表されるゴールデンクロスやデッドクロスなどが有名である. また, オシレーター系テクニカル指標は, パラメーターとして設定された値をオシレーター系テクニカル指標の値が越えたときを売買シグナルとして表すものが多い. 例えば, 一定期間の相場における値動きの強弱を数値で表す相対力指数(RSI)などが有名である. テクニカル分析を用いる投資家達は, テクニカル指標の組み合わせやパラメーターを独自に定めることで取引を行っている. 本稿では, このテクニカル分析に焦点を当てた売買モデルの構築を行った.

ニューラルネットワークや強化学習などの技術は, 市場の予測や投資戦略, 取引モデルなどの分野で数多く応用されている. 例えば, ニューラルネットワークを用いることで株の市場におけるトレンドの予測[2]や深層強化学習や強化学習を用いた取引モデルの構築[3][4]などがある. しかし, これらの計算手法を応用する場合, 長所と短所がある. 例えば, ニューラルネットワークは, 学習や予測には優れているが, 説明能力に欠けるため, ブラックボックスモデルになってしまう. また強化学習は, 人間には発見できない行動パターンの学習などには優れているが, 行動に対する説明能力がないといったことが挙げられる. そのため, これらの手法を用いて構築された取引モデルは取引理由に関する分析が難しい. そこで, 我々は解釈可能性に優れているという点から進化計算手法に着目した.

進化計算手法とは, 一般的に個体を進化させることで適応度を高め, 行動規則を表す最適または最適に近い個体を得ることができる.進化計算手法の中でも, Genetic Programming(GP)[5]は, ゲノムを木構造に組み合わせることによって解を表現する手法であり, 複雑な行動ルールを作成して問題を解決するような現実の問題にも適用できることから, 為替市場における取引モデルの構築[6]や株式市場における取引モデルの構築[7]など取引モデル構築に頻繁に用いられている. また, 解表現の特性上, 構造を見ることで構築された取引方法の分析や, 取引理由の解析が容易なため, 解釈可能な取引モデルの構築が可能である. しかし, 一般的に GP は, 解の探索空間が肥大化するために最適解の探索が困難な場合があり, 探索効率があまり良くない. GP の問題点を克服するために, ゲノム構造をネットワーク構造にした遺伝的ネットワークプログラミング(GNP)[8][9]というグラフベースの進化アルゴリズムが提案されている. GNP は, スタートノード, 判定ノード, 処理ノードの 3 種類のノードを用いた有向グラフ構造であるため, GPと比較してコンパクトなプログラムを作成することができる. スタートノードは, プログラムの開始位置を表す. 判定ノードは定められた条件判定を行い, 判定結果に従って次の遷移先を選択する. 処理ノードは, 定められた処理を行い, 次のノードへ実行を遷移させる. これらのノード間の接続と遷移によって行動ルールが生成される. GNP は, 主に自律エージェントの行動決定に適用され, 一般的な GP よりも高い性能を示していることが確認されている[10]. また, GNP も GP と同様にして, 取引モデル構築に用いられており, 取引モデル構築に有効であることが確認されている[11][12].

本研究では, GNP を参考にした新しいアルゴリズムを考案し, テクニカル分析に基づいた, 外国為替市場における取引モデルの構築を行った. 本研究における提案は主に 3 つある. 1 つ目は, GNP を参考にした新しいアルゴリズムの構築である. GNP は, パラメーターで決められたノード数で個体の生成を行うため, パラメーターの設定によってプログラムの表現力が変わってくる. そのため, 取引モデル構築のような解空間が大きい問題に GNP を適用する場合, パラメーターの設定によって結果が大きく変わると考えられ, 最適なパラメーターを見つけることは難しい. また, 取引モデル構築に GNP を適用する場合, 同じ判定が繰り返される可能性があるため, GP の枝刈り[13]のような機能が必要だと考えられる. 我々は, これらの問題に対して, Full Range GNP(FRGNP)を考案することで, パラメーターの削減と GP の枝刈りに代わる機能を導入することに成功した. 2 つ目は, 新しい適応度の考案である. 従来の適応度は, 学習期間の総利益や利益率などが一般的である. しかし, これらの適応度は,学習期間で最も利益が出やすい取引モデルが構築される. そのため, 適応度に利

益のみを考慮した場合, 過学習の要因になると考えられる. そこで, 我々は適応度に安定性の要素を導入することで,確実に利益を出すことができる取引を学習できるのではないかと考えた. その安定性の要素として Nauzer.J.Balsara によって考案されたバルサラの破産確率[14]を導入した新しい適応度を考案した. 3 つ目は, テクニカル指標の売買シグナルに強弱を付けたことである. テクニカル指標の売買シグナルによる売買のタイミングは, 必ずしも正しい売買タイミングではない. 例えば, 買いシグナルが出ているため相場が上昇することを見込んで取引を行ったにも関わらず相場が下降してしまい, 損失を出すことがある. これを, 売買シグナルのダマシという. テクニカル分析による取引モデル構築において, 売買シグナルのダマシに引っかからないような取引モデル構築が重要だと考えられる. そこで, 我々は売買シグナルに強弱[15]を付けることで売買シグナルのダマシを回避することを考えた. 我々が考案した売買シグナルの強弱とは, それぞれのテクニカル指標が売買シグナルを出した時の特徴を用いて売買シグナルの強弱の基準値を算出し, それに基づいて強弱を判定したものである. ここでいう売買シグナルの特徴とは, トレンド系テクニカル指標では, 2 つのテクニカル指標が交差したときの角度, オシレーター系テクニカル指標では, 設定された値をテクニカル指標の値が超えているときの設定値との差分を特徴である. 売買シグナルの強弱を導入することで, 新しい分岐先が増えることによる取引モデルの表現力向上および 安定的な取引モデル構築に役立つことが期待される.

我々は, これらの提案に加えて, テクニカル指標を用いたトレンドを判定するノード, 損切り利益確定を判定するノードを考案した. さらに, これらの判定に関わるパラメーターやオシレーター系テクニカル指標の設定値を Binary Genetic Algorithm(BGA) で FRGNP による取引モデル構築と同時に進化させることで取引モデルに適したパラメーターの獲得を行った.

2. Genetic Network Programming

ここでは, GNP について詳しく説明する. GNP は, GP を遺伝子構造の観点から拡張したものである. 以下の図 1 に GNP の基本構造を示す.

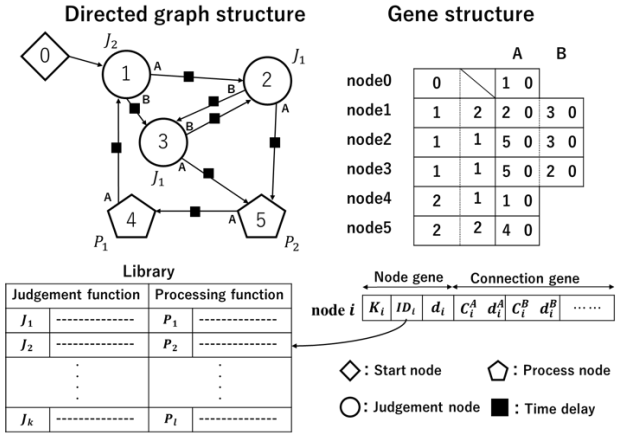


図1 GNPの基本構造

2.1 GNP の構成

GNP は 1 つのスタートノード, 複数の判定ノード, 複数の処理ノードから構成される. 図 1 の GNP は, スタートノードが 1 つ, 判定ノードが 2 つ, 処理ノードが 2 つあり, それらは互いに接続されている. スタートノードは機能を持たず, 条件分岐もない. スタートノードの唯一の役割は, 実行されるべき最初のノードを決定することである. 判定ノードは, if-then 型の条件分岐に基づいた判定機能を持っている. 各判定ノードは, 判定結果を返し, 次に実行されるノードを決定する. 例えば, 取引モデル構築に適用する場合, テクニカル指標が売買シグナルを出しているかどうかを判定し, その結果にしたがって次のノードが選択される. 図 1 の判定ノードでは, 判定結果 A または B を得たあと, 対応する分岐を選択し, 次のノードへ進む. このときの判定内は, Library の$J_1, \cdots, J_k$ に対応する.処理ノードは行動を行う機能として働く.例えば,処理ノードは,エージェントの行動(買い行動,売り行動)を決定する.また,判定ノードと異なり,処理ノードには条件分岐がない.図 1 における処理ノードの行動内容は,Library の$P_1, \cdots, P_l$ に対応する.これら判定ノードと処理ノードの組み合わせ関係を適切に進化させることで, 有効な行動ルールを生成することができる.

GNP は, 判定ノードと処理ノードの実行に要する時間, およびノード遷移に要する時間を遅れ時間として設定できる. 実世界では, 状況を判断し, 次の行動に移行し, 実際に行動を行う際に時間を費やすため, 遅れ時間を考えた枠組みは有効である. この遅れ時間は, GNP の遷移における閾値としての役割がある. 遅れ時間が存在しない GNP の遷移では, 判定ノードのループが発生し処理ノードが実行できない問題が発生することがある. これに対して, あらかじめ意思決定できる回数を閾値として定め, これに達するか超えたら終了とすることにすることは有効である. 例えば, ノード遷移の遅れ時間を 0 , 判定ノードの遅れ時間を 1, 処理ノードの遅れ時間を 5 と設定し, 1 ステップは遅れ時間の合計が 5 に達するか超えたら終了することにする. このとき, 1 ステップで可能なノード遷移は 4 回以内の連続する判定のあと処理ノードを 1 回実行, 5 回の判定ノードの実行, 処理ノードのみ 1 回実行, に限られる. したがって, 1 ステップの遅れ時間の合計が設定されていなければ判定ノードのループが発生する場合でも, 5 回の判定で 1 ステップが終了し, その後もループが続くようであれば処理ノードが実行されないまま終了する.

2.2 GNP の遺伝子構造

GNP の構造は, 以下で説明する遺伝子の集合で決定される. 図 1 には, ノード i $(0 \le i \le n-1)$ (ノード数が n 個のとき, それぞれのノードには 0 から$n-1$ の番号がついている) の遺伝子構造も示されている. まず, ノード遺伝子について, $K_i = 0$ は開始ノード, $K_i = 1$ は判定ノード, $K_i = 2$ は処理ノードを表す. ID_i はノードi の判定/処理内容の番号であり, Library の中にその内容が記載されている. d_iは判定/処理に関する遅れ時間である. 次に接続遺伝子について説明する. C_i^A, $C_i^B, \cdots$ は, 遷移先のノードの番号を示しており, d_i^A, $d_i^B, \cdots$ は, 次ノードに遷移するときの遅れ時間を示す. ここで, 判定ノードは, 参照する接続遺伝子の上付きの添え字 A, B などを判定結果にしたがって決定する. 例えば, ある判定ノードに接続枝 A と B が存在したとする. このとき, もし判定結果が”B”であれば GNP は C_i^B と d_i^Bを参照し次のノードへ遷移する. ただし,

処理ノードは判定による分岐がないので常に C_i^A と d_i^A を参照する.

2.3 初期個体の生成

開始ノード 1 個と, 判定ノード a 個, 処理ノード b 個を用意して, ノード 1 からノード a までを判定ノード, ノード a+1 からノード a + b までを処理ノードとし, 各ノードの機能をランダムに設定する. そして, 各ノードから延びる接続枝をランダムに選ばれた開始ノードと自ノード以外の他のノードへ接続する. 例えば, 設計者が定めた判定ノードの機能が J_1, J_2, J_3 の計 3 種類, 処理ノードの機能が P_1, P_2 の計 2 種類であるとする. さらに, 判定ノード数を a=3, 処理ノード数を b=2 と設定する. このとき, 開始ノード 1, 判定ノード数 3, 処理ノード数 2 を合わせた 6 個のノードで個体が生成される. この個体生成を繰り返すことで, あらかじめ定められた数 N の個体を生成する.

2.4 進化方法

GNP は交叉と突然変異を用いて進化を行うが, まず全個体中から最も良い適応度を示したエリート個体を保存し, その他の個体を交叉で生成した新しい個体群と突然変異で生成した新しい個体群を合わせた個体群と入れ替える. このとき, 交叉で生成する個体数(Cross over num)と突然変異で生成する個体数(Mutation num)はパラメーターとして設定する.

交叉は 2 個の親個体間で行われ, 2 個の子個体を生成する(図 2). このとき, 選択された交叉ノードの遺伝子が全て入れ替わる. 手順は以下の通りである.

1) 親個体の選択
   任意の選択法 (トーナメント選択など)を用いて2個の親固体を選択する.
2) 交叉ノードの選択
   個体中の各ノード番号を確率 P_c で交叉ノードとする.
3) 遺伝子の交換
   親個体間で, 同じノード番号の交叉ノードの遺伝子をすべて交換する.

生成された 2 個の子個体が次世代の個体となる. 生成された個体数が設定された個体数(Cross over num)になるまで繰り返す.

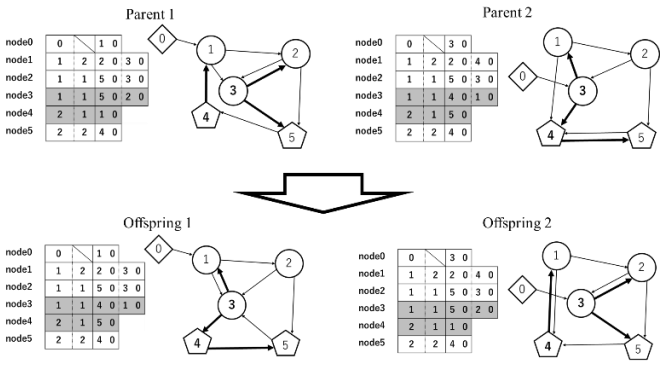


図2 GNPの交叉

突然変異は 1 個体で行われ, 新しい 1 個体が生成される (図 3). 手順は以下のとおりである.

1) 親個体の選択
   任意の選択法 (トーナメント選択など)を用いて1個の親固体を選択する.
2) 接続の変異
   各接続枝について確率 P_m で選択し, 選択された接続をランダムに選ばれた他のノードへの接続に変更する.
3) 機能の変異
   各ノードについてそれぞれ確率 P_m で選択し, 選択されたノードの機能をランダムに選ばれた他の機能へ変更する. ただし, 処理ノードは他の処理ノード, 判定ノードは他の判定ノードに変更される.

生成された子個体が次世代の個体となる.

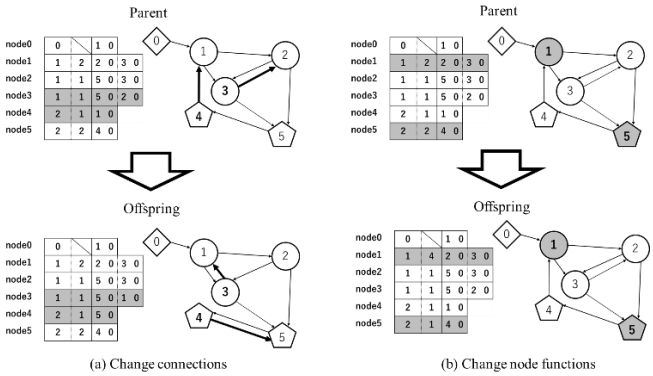


図3 GNPの突然変異

GNP の交叉では, 各ノードが接続枝を持ったまま交換されるため, 交叉後に実行不可能解になる事がない. また, 突然変異後も各ノードから次のノードへの接続が必ず存在するため実行不可能になることがない.生成された個体数が設定された個体数(Mutation num)になるまで繰り返す. N= 1+Cross over num+ Mutation num である.

3. Full Range Genetic Network Programming

ここでは, GNP を拡張した新しいアルゴリズム Full Range Genetic Network Programming (FRGNP)について説明する. FRGNP が GNP と異なる部分は, 個体の生成方法と進化の方法である. 遷移方法やノードの種類, 遺伝子構造などは GNP と同じである.

3.1 初期個体の生成

FRGNP では判定ノードと処理ノードの数は機能の数によって決まり, 全ての機能が 1 つずつ配置される. 例えば, 設計者が定めた判定ノードの機能が J_1, J_2, J_3 の計 3 種類, 処理ノードの機能が P_1, P_2 の計 2 種類であるとする. このとき, FRGNP ではスタートノード 1 つと判定ノードの機能の種類数 3, 処理ノードの機能の種類数 2 を合わせた 6 個のノードで個体が生成される. GNP では, 個体の判定ノードと処理ノードの数をパラメーターとして設定することで, 個体を生成していた. さらに, 個体上のノードの機能をランダムに設定していたため, $J_1 \rightarrow J_1 \rightarrow P_1$ のように 1 ステップ

で同じ機能を複数回使う無駄な遷移があるという問題点があった. 一方, FRGNP では, 全ての機能が 1 つずつ配置され, 進化操作で機能が変わることはないため無駄な遷移が生じない.

この個体生成を繰り返すことで, あらかじめ定められた数 N の個体を生成する.

FRGNP では, ノードの機能数に依存してネットワーク上のノード数が変わる. そのため, GNP で設定していた判定ノード数 a, 処理ノード数 b といったパラメーターが削減された.

3.1 進化方法

FRGNP での進化は, GNP と同じく最も高い適応度を示したエリート個体 1 体を保存した後に行う. GNP では, 次世代個体群をエリート個体と交叉で生成された個体群, 突然変異で生成された個体群で構成されていた. 一方, FRGNP では, 残りの個体群を一連の進化操作で生成する. FRGNP の進化操作は, 2 個の親個体間で行われ, 2 個の子個体を生成する(図 4).手順は以下の通りである.

i. 親個体の選択
任意の選択法 (トーナメント選択など)を用いて2個の親固体を選択する.

ii. 交叉ノードの選択
個体中の各ノード番号を確率 P_c で交叉ノード番号とする.

iii. 遺伝子の交換および変異
親個体間で, 同じノード番号の交叉ノードの遺伝子をすべて交換する. 交叉ノードの各接続枝について確率 P_m で選択し, 選択された接続をランダムに選ばれた他のノードへの接続に変更する.

生成された 2 個の子個体が次世代の個体となる. 子個体が N-1 個生成されるまで繰り返す.

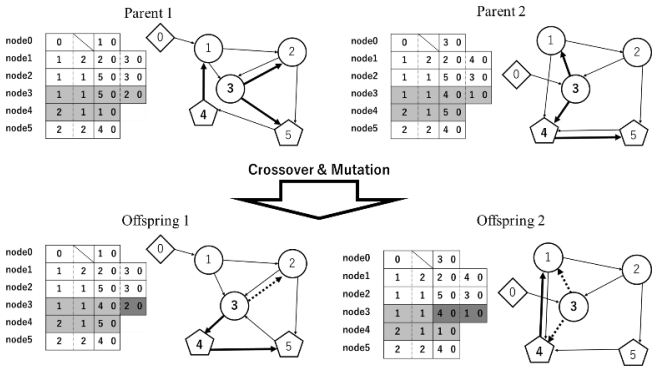


図4 FRGNPの進化操作

FRGNP の進化操作では, 各ノードが接続枝を持ったまま交叉を行った後, 交叉を行ったノードについて接続枝の突然変異を行うことで個体を進化させる.また, GNP とは異なりノードの機能を進化させないため, 突然変異によってノードの機能が変わってしまうことによる良い遷移の消失という問題もなくなる. 例えば, 個体上で$J_1 \rightarrow J_2 \rightarrow P_1$ のような適応度に貢献する遷移があるとする. このとき, GNP ではノードの機能の突然変異によって$J_1 \rightarrow J_1 \rightarrow P_1$ のような遷移に変わり, 良い遷移が消失してしまう可能性がある. また, 個体上に J_2 が 1 つしかなく, この機能が個体の適応度向上に大きく関わる機能である場合, その機能が使われる個体群へと進化するには, 多くの世代がかかることがわかる. 一方, FRGNP のように, ノード間の繋がりのみに着目した進化を行うことで,個体上から機能が消失することはないため, 適応度に貢献するノード間の繋がりを消失した場合もより早くそのノード間の接続を修復できるため効率的に個体の進化ができる. さらに, FRGNP の進化操作では, 交叉と突然変異を一連の流れとして行っているため, GNP のような突然変異で生成する個体数, 交叉で生成する個体数といったパラメーターが削減されている. これにより, パラメーターの影響が少ない進化が可能になった.

4. 取引モデル

本研究では, 取引戦略の構築はFRGNPを用いて行った. 判定ノードでは,テクニカル指標を用いた判定や, 損切り/利益確定などの判定条件を使用した. また, 処理ノードでは, ポジションを獲得/解消する機能を持たせた. 判定ノードの判定条件に関わるパラメーターは, 組み合わせが膨大であるため戦略に最適なパラメーターの組み合わせを見つけることが難しい. そこで, BGAによってパラメーターを進化させることで戦略に適したパラメーターを探索させた. したがって, 本研究で提案する取引モデルは, 取引戦略の構築をFRGNP, 判定ノードで使用されるパラメーターをBGAによって同時に進化させるモデルである(図5).

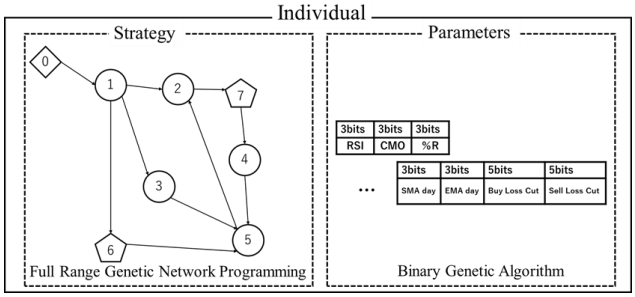


図 5 提案モデルの個体例

3.2 売買シグナルの強弱

テクニカル指標による売買シグナルの強さはテクニカル指標の種類によって 2 つに分けて定義した. 1 つ目は, オシレーター系テクニカル指標において, 設定された値をテクニカル指標によって計算された値が越える度合い, 2 つ目は, トレンド系テクニカル指標における短期日数で計算されたテクニカル指標と長期日数で計算されたテクニカル指標が交差する角度の大きさによる定義である. 1 つ目の場合テクニカル指標がシグナルを出した値と設定された値との差分. 2 つ目の場合テクニカル指標がシグナルを出した時の交差の角度を個体毎に記憶し, それらを利用することで売買シグナルの強弱の判断基準を計算し, ノード遷移の際に計算された差分や角度が基準を越えている時を強いシグナル, 越えていないときを弱いシグナルとした.

売買シグナルの強弱の判断基準については以下の通りである.

1) 差分と角度の記憶

テクニカル指標の売買シグナルによって差分か角度どちらかをテクニカル指標毎にメモリに保存する. 表1に角度または差分を記憶するテクニカル指標を分類して示す.

表1 テクニカル指標の分類

Angle	Diff
Short SMA & Middle SMA	RSI
Short SMA &Long SMA	%R
Middle SMA & Long SMA	SMA Divergence Rate
Short EMA & Middle EMA	EMA Divergence Rate
Short EMA &Long EMA	Chande Momentum Oscillator
Middle EMA & Long EMA	Psychological line
MACD	ROC
Fast Stochastic	
Slow Stochastic	
+DMI & - DMI	
ADX & ADXR	

判定ノードの各テクニカル指標で, シグナルが出た時の差分または角度を, 買いサインと売りサインで分けて個体毎に損益が確定するまでメモリに保存する. Stochastic は, 差分と角度の両方を扱うテクニカル指標であるが, 両方を考慮することは難しいため, 設定値を超えて交差したときの角度のみを売買シグナル強弱に利用している. また, Bollinger Band と Pivot はこの枠組みに当てはまらないため, 今回は売買シグナルの強弱を考慮しない.

2) 基準差分と基準角度の更新

表1で分類した全てのテクニカル指標(全18個)に, 買いシグナルのメモリと売りシグナルのメモリの 2つのメモリが存在する. 従って, 全テクニカル指標で $36(18 \times 2)$個のメモリが存在する. 決済が確定した時の各テクニカル指標のメモリ内の平均を計算し, $J_{ij}(i = 1, 2 \ldots, 18, j = 1,2)$とする. ここで i はテクニカル指標の種類を表し, $j = 1$のときに買いシグナル, $j = 2$のときに売りシグナルの判断基準を表している. そして, 決済の結果, 利益が出ている時, J_{ij} が0ではないテクニカル指標において売買シグナルの強さの判断基準$B_{ij,t}$の更新を次式で定義する.

$$B_{ij,t} = B_{ij,t-1} + \frac{2(J_{ij} - B_{ij,t-1})}{t+1} \tag{1}$$

ここでt (1,2, …)は更新回数を表す.
決済の結果, 損失が確定した場合, 個体のメモリをリセットする

3.2 ノードの種類

1) 開始ノード

自分の所持するポジションの有無と種類によって遷移先を変更することで, 多点スタート戦略を可能にした. 具体的には, 売りポジションを持っている場合, 買いポジションを持っている場合, ポジションを持っていない場合の 3 つに分岐する.

2) 判定ノード

各判定ノードが, 1つの判定条件を所持する. 表2にノードの判定条件を示す.

表2 判定ノードの判定条件

Function Number	Judgement content
1~18	Buy sign ∩ strong sign
	Buy sign ∩ weak sign
	No Buy/ Sell sign
	Sell sign ∩ weak sign
	Sell sign ∩ strong sign
19	$2\sigma < Close$
	$\sigma < Close \leq 2\sigma$
	$-\sigma \leq Close \leq \sigma$
	$-2\sigma \leq Close < -\sigma$
	$-3\sigma > Close$
20	$s2 < Close \leq s3$
	$s1 < Close \leq s2$
	$r1 \leq Close \leq s1$
	$r2 \leq Close < r1$
	$r3 \leq Close < r2$
21	$Long_{sma} < Short_{sma} < SMA_t$
	$\frac{Short_{sma} + Long_{sma}}{2} < SMA_t$
	$SMA_t < Short_{sma} < Long_{sma}$
	$SMA_t < \frac{Short_{sma} + Long_{sma}}{2}$
	$else$
22	$Long_{ema} < Short_{ema} < EMA_t$
	$\frac{Short_{ema} + Long_{ema}}{2} < EMA_t$
	$EMA_t < Short_{ema} < Long_{ema}$
	$EMA_t < \frac{Short_{ema} + Long_{ema}}{2}$
	$else$
23	$profit < 2L$
	$2L \leq profit < L$
	$else$
	$2G \geq profit > G$
	$profit > 2G$

ここで, $Close$ は終値を表している. また, ノード番号 19 は Bollinger Band, 20 は Pivot という指標を表し, 売買シグナルの強弱による遷移を行わない為, 遷移方法が異なる. また, ノード番号 21 は単純移動平均(SMA)によるトレンド判定ノード, 22 は指数平滑平均(EMA)によるトレンド判定ノードを表し, $Long$ と $Short$ を次式で定義する.

$$Long_{sma} = SMA_{t-long_day_{SMA}} \tag{2}$$

$$Short_{sma} = SMA_{t-short_day_{SMA}} \tag{3}$$

$$Long_day_{SMA} = 2 \times Short_day_{SMA} \quad (4)$$

$$Short_day_{SMA} = 5 \times ([day_{SMA}]_{2\to10} + 1) \quad (5)$$

ここで, t は開始から現在までの日数 , $[\cdot]_{2\to10}$は 2 進数から10 進数への変換を表す. また, day_{SMA}は各個体が持つ BGA の遺伝子情報を用いる. EMA の式についても同様である.

ノード番号 23 は損切り/利益確定ノードを表し, profit は現在のポジションを解消したときの 手数料を考慮した利益を表す.ここで,買いポジションを持っているときの L と G は次式で定義する.

$$L = -\left(6.25 \times ([loss_{buy}]_{2\to10} + 1)\right)^V \quad (6)$$

$$G = \left(6.25 \times ([get_{buy}]_{2\to10} + 1)\right)^V \quad (7)$$

$$V = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - \overline{p_n})^2} \quad (n = 5 \times ([vola_{buy}]_{2\to10} + 1)) \quad (8)$$

$loss_{buy}$は買いポジションを持っているときの損切り価格, get_{buy} は買いポジションを持っているときの利益確定価格であり, 遺伝子上に個体毎で保存されている. また, V は(8)式で算出されるボラティリティを用いる. (8)式は過去n日間の終値の標準偏差を表している.

3) 処理ノード

処理ノードは買いポジション獲得, 売りポジション獲得, ポジション解消, いずれかの処理機能を持ち, 処理ノードに遷移した時に買いポジション獲得の機能を持っていたら買い, 売りポジション獲得の機能を持っていたら空売りを行う. しかし, 複数のポジションを持つことはできないため, 既に買いポジションを持っている状態で買いポジションの獲得や, ポジションを持っていないにもかかわらずポジションの解消はできない. 通常, 処理ノードに到達した時に遷移を終了する. しかし, ポジション解消の機能を有する場合は, ポジション解消時の損益に応じた遷移を行う. 買いポジションを解消したときの損益による遷移, 売りポジションを解消したときの損益による遷移を行う. ポジション解消後の遷移は, $delay_{after}$を総遅れ時間としてネットワーク上で遷移を行う. なお, ポジション解消後の遷移を 2 回以上続けて行わない.

3.2 遺伝的操作

1) 初期個体生成

N個体生成する. ノードの機能は表3を参考にする. また, ノードの接続は自分以外の他のノードに無作為に接続する. そして, BGAの初期生成時に振り分けられる領域(bits)を以下表3に示す.

表3 BGAにおける割り当てBits

Function	Allocated bits	
TREND JUDGE BY SMA	3bits (day_{SMA})	
TREND JUDGE BY EMA	3bits (day_{EMA})	
SECURE PROFIT	5bits(get_{buy})	5bits(get_{sell})
LOSS CUT	5bits($loss_{buy}$)	5bits($loss_{sell}$)
VOLATILITY	3bits ($vola_{buy}$)	3bits ($vola_{sell}$)
TECHNICAL INDICATOR	**Buy line**	**Sell line**
RSI	3bits	3bits
%R	3bits	3bits
SMA DIVERGENCE RATE	5bits	5bits
EMA DIVERGENCE RATE	5bits	5bits
CHANDE MOMENTUM OSCILLATOR	3bits	3bits
PSYCHOLOGICAL LINE	3bits	3bits

また, テクニカル指標の閾値をデコードするときに用いる式を次に示す.

$$\begin{cases} \text{Sell line} : 5 \times ([TI]_{2\to10} + 1) \\ \text{Buy line} : 100 - 5 \times ([TI]_{2\to10} + 1) \end{cases} \quad (9)$$

$$\begin{cases} \text{Sell line} : 0.3 \times ([TI]_{2\to10} + 1) \\ \text{Buy line} : -0.3 \times ([TI]_{2\to10} + 1) \end{cases} \quad (10)$$

TIは2進数で表現されたテクニカル指標の閾値を表す. TIが3bitsであるときは(9)式, 5bitsであるときは(10)式を用いる.

2) 評価

個体のfitnessを次式で定義する.

$$\text{fitness} = \begin{cases} profit & \text{if } porfit > 0 \text{ and } k > 0 \\ & or \\ & profit < 0 \text{ and } k < 0 \\ 0.0 & elif\ count = 0 \\ profit\ (1 - Q) & else \end{cases} \quad (11)$$

$$\begin{cases} px^{(k+1)} + 1 - p - x = 0 \\ Q = x^r \end{cases} \quad (12)$$

ここで, Qはバルサラの破産確率,profitは売買を行う期間の損益の合計(銭), count は取引回数,pは勝率,kは損益率,rは資本比率,xは式(12)の上の式(バルサラの破産確率の特性方程式)の解を表し,この解を用いて破産確率Qを求める.

3) 進化的操作

提案モデルでは, FRGNPの進化方法と同様の操作を行う. それに加えて, 提案モデルでは, 個体毎にFRGNPの判定ノードのパラメーターを表すBinary遺伝子の進化操作も行う. 最も適応度が高いエリート個体を1体保存した後に以下の進化操作を行う.

i. 親個体の選択
任意の選択法 (トーナメント選択など)を用いて2個の親固体を選択する.

ii. 交叉ノードの選択
個体中の各ノード番号を確率 P_c で交叉ノードとなるか判定する.

iii. 遺伝子の交換および変異
親個体間で, 同じノード番号の交叉ノードの遺伝子をすべて交換する. 交叉ノードの各接続枝について確率 P_m で突然変異するかどうかを判定し, 選択された接続をランダムに選ばれた他のノードへの接続に変更する.

iv. 交叉遺伝子の選択
Binary遺伝子中の全遺伝子番号を一様交叉で交叉遺伝子となるか判定する.

v. 遺伝子の交換および変異
親個体間で, 交叉遺伝子番号として判定された遺伝子を交換する. このとき, 交換する遺伝子について確率 P_m で判定し, ランダムに1か0に変更する.

生成された2個の子個体が次世代の個体となる.生成された個体数がN−1個となるまで繰り返す.

5. 結果

本研究では, 提案モデルの有効性を確認するために提案モデルを含めた 4 つのモデルを用いた実験を行った. GNP を用いて取引を行う Model1, Model1 モデルに GA によるパラメーターの進化を加えた Model2, FRGNP を用いて取引を行う Model3, そして Model3 に GA によるパラメーターの進化を加えた Proposed とした. また, 適応度は, Profit(総利益)と Proposed(提案適応度)のそれぞれを用いて実験を行った. データは, ドル円ペア, ユーロ円ペア, オーストラリアドル円ペア, ポンド円ペアを用いた. また, それぞれ日足チャートを使用し, Investing.com からダウンロードしたデータを用いた. また, 1 通貨当たりの取引手数料は, Oanda Japan を参考に設定した. 以下の表 4 に各通貨ペアにおける取引手数料の設定及び学習期間, テスト期間の設定を示す.

表 4 各通貨ペアに対する取引手数料及び期間の設定

Currency pairs	Spread(pips)	Learn term	Test term
USD/JPY	0.4	2001~2003	2004~2021(5/27)
EUR/JPY	0.7		
AUD/JPY	1.0		
GBP/JPY	1.4		

本研究で使用したモデルの設定を以下の表に示す. Model1 及び Model3 では, BGA によるパラメーターの進化を行わないため, あらかじめ定めたパラメーターを使用する. パラメーターを定める判定ノードは, オシレーター系テクニカル指標, 損切り・利益確定, トレンド判定の機能である. GNP で使用したパラメーターを以下の表 5, FRGNP で使用したパラメーターを以下の表 6 に示す.

表5 GNPのパラメーター設定

Generation	300
POPULATION: N	101
CROSSOVER NUM	60
MUTATION NUM	40
JUDGEMENT NODE: a	54
PROCESS NODE: b	20
CROSSOVER RATE (%): P_c	25.0
MUTATION RATE (%): P_m	1.0
ALL DELAY	5
AFTER DELAY: $delay_{after}$	2
STEP	50

表6 FRGNPのパラメーター設定

Generation	300
POPULATION	100
CROSSOVER RATE (%): P_c	25.0
MUTATION RATE (%): P_m	1.0
ALL DELAY	5
AFTER DELAY: $delay_{after}$	2
STEP	50

GNP, FRGNP のそれぞれのアルゴリズムにおける遺伝的操作の個体選択方法は, サイズ 2 のトーナメント選択を採用した. また, 両方のアルゴリズムで遷移の遅れ時間は考慮せず, 判定ノードの遅れ時間を 1, 処理ノードの遅れ時間を All delay と同じ値に設定した. さらに, スタートノードは 1 つとした.

以下の表 7 に各通貨ペアに対してのテスト期間における最良個体の総利益の 50 試行平均を示す.

表 7 テスト期間における最良個体の総利益の 50 試行平均

CURRENCY PAIRS	FITNESS	MODEL1	MODEL2	MODEL3	PROPOSED (OURS)
USD/JPY	Profit	155.9	687.992	284.644	1117.12
	Proposed	2461.584	2714.868	2861.408	**5046.936**
EUR/JPY	Profit	1305.498	709.816	890.962	1842.888
	Proposed	3862.82	4907.84	3671.896	**5833.378**
AUD/JPY	Profit	-103.8	-932.9	-274.44	517.36
	Proposed	3220.12	2798	3564.28	**4053.620**
GBP/JPY	Profit	-1492.39	-248.192	664.116	1265.824
	Proposed	5809.408	6583.084	6462.368	**8939.508**

表 7 から model は Proposed model を用いて Fitness は proposed を用いたときに全ての通貨ペアで大きく利益を出せていることがわかる.

参考文献

[1] J. Murphy, "Technical analysis of the financial markets", 1999.

[2] E. W. Saad, D. V. Prokhorov and D. C. Wunsch, "Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks," in IEEE Transactions on Neural Networks,

vol. 9, no. 6, pp. 1456-1470, Nov. 1998, doi: 10.1109/72.728395.

[3] Y. Deng, F. Bao, Y. Kong, Z. Ren and Q. Dai, "Deep Direct Reinforcement Learning for Financial Signal Representation and Trading," in IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 3, pp.653-664, March 2017, doi: 10.1109/TNNLS.2016.2522401.

[4] M. A. H. Dempster and V. Leemans, "An automated FX trading system using adaptive reinforcement learning", Expert Syst. Appl., vol. 30, no. 3, pp. 543-552, 2006.

[5] Koza J.R., "Genetic Programming: On the programming of computer by means of natural selection", MIT Press, 1992.

[6] P. B. Myszkowski and A. Bicz, "Evolutionary algorithm in Forex trade strategy generation," Proceedings of the International Multiconference on Computer Science and Information Technology, 2010, pp. 81-88, doi: 10.1109/IMCSIT.2010.5679921.

[7] A. Pimenta, F. G. Guimarães, E. G. Carrano, C. A. L. Nametala and R. H. C. Takahashi, "GoldMiner: A genetic programming based algorithm applied to Brazilian Stock Market," 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), 2014, pp. 397-402, doi: 10.1109/CIDM.2014.7008695.

[8] T. Eguchi, K. Hirasawa, J. Hu and N. Ota, "A study of evolutionary multiagent models based on symbiosis," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 36, no. 1, pp. 179-193, Feb. 2006, doi: 10.1109/TSMCB.2005.856720.

[9] S,Mabu, K.Hirasawa, J.Hu; “A Graph-Based Evolutionary Algorithm: Genetic Network Programming (GNP) and Its Extension Using Reinforcement Learning” Evolutionary Computing, vol. 15, no.3, pp. 369–398.Sep. 2007 doi: https://doi.org/10.1162/evco.2007.15.3.369.

[10] K. Hirasawa, M. Okubo, H. Katagiri, J. Hu and J. Murata, "Comparison between Genetic Network Programming (GNP) and Genetic Programming (GP)," Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546), 2001, pp. 1276-1282 vol. 2, doi: 10.1109/CEC.2001.934337.

[11] C.Yan, S.Mabu, K.Hirasawa, “A genetic network programming with learning approach for enhanced stock trading model.” Expert Systems with Applications, vol. 36, no.10, pp.12537–12546. May.2009,doi:https://doi.org/10.1016/j.eswa.2009.05.054.

[12] S.Mabu, C.Yan,, K.Hirasawa,, “Generating stock trading rules using genetic network programming with flag nodes and adjustment of importance indexes.” Electronics and Communications in Japan, vol.94, no.1, pp. 25–33. 2011, doi:10.1002/ecj.10269.

[13] M.John. “An Empirical Comparison of Pruning Methods for Decision Tree Induction. Machine Learning”, vol. 4, pp.227–243,Nov.1989. https://doi.org/10.1023/A:1022604100933.

[14] N.Balsara, “Money Management Strategies for Futures Traders”, Mar. 1992.

[15] 内田純平, 穴田一: 売買シグナルの強弱を考慮した Genetic Network Programming による外国為替取引戦略の構築, 第 34 回人工知能学会全国大会, 2020.

機械学習による為替フォワード取引期間の判別モデル

Forex forward swap opportunities detected by machine learning

雉子波 晶 [1] *　杉本 誠忠 [2]　酒本 隆太 [3]　鈴木 智也 [1]
Akira Kijinami[1]　Takanari Sugimoto[2]　Ryuta Sakemoto[3]　Tomoya Suzuki[1]

[1] 茨城大学大学院理工学研究科
[1] Graduate School of Ibaraki University
[2] ワイジェイ FX 株式会社
[2] YJFX, Inc.
[3] 岡山大学大学院社会文化科学研究科
[3] Graduate School of Okayama University

Abstract: In the rollover of forward foreign exchange contracts, FX brokers generally selects tomorrow-next transaction because of higher liquidity and lower risk. However, it might be possible to obtain larger swap points by selecting longer forward transactions such as one-week or three-week forward in terms of the term premium. Therefore, we detect optimal timings to select longer forward transactions by machine learning techniques, and propose a mixed strategy that combines tomorrow-next and longer forward transactions. This timing might be affected by various factors such as global stocks, bonds, commodities, etc., and we could obtain larger swap points by the mixed strategy using the machine learning with these global factors.

1　はじめに

外国為替証拠金 (Foreign Exchange trading: FX) 取引では，約定の 2 営業日後に現物の受渡しを行うルールがある．しかし一般的に取引額は高額に及ぶため，FX 業者など非銀行が毎回現物を用意するのは非現実的である．そこで長期に渡りドル円を買い持ちしたい場合，約定の翌日にドル円を売り，さらに同時に新規にドル円を買う．つまり売り取引と買い取引を同時に行い，決済日を繰り延べする．これをロールオーバーといい，ロールオーバーを繰り返すことで現物の受渡しを無期限化できる．さらにロールオーバー時にトゥモローネクスト (以下，トムネ) や 1 週間フォワードといった，受渡し日を自由に長期化できるフォワード取引を組み合わせることでスワップポイントを獲得できる [1]．外国為替取引において通貨ごとに金利が異なり，この金利差はスワップポイントとして表される．通常の経済状況では決済日までの期間が長いほど金利が高くなるタームプレミアムが考えられるため，決済日が長期に及ぶフォワード取引の方が多くのスワップポイントを獲得しやすい．しかし流動性や在庫管理に伴う業務リスクにより，実務においては受渡し日が短いトムネでロールオーバーするのが一般的である．

そこで本研究では機械学習の活用により，積極的に長期のフォワード取引を選択すべきタイミングを検出する．またフォワードレートは一般的にカバー付き金利平価説によって定式化できるが，突発的な政治経済情勢の変化などによって，現実のフォワードレートは理論値から乖離する可能性が指摘されている [2–5]．この乖離は，対象通貨における為替相場のみならず，株式・債権・商品先物など様々な要因が影響すると考えられる．実際に，タームプレミアムの逆転現象 (短期と長期のフォワード取引から得られるスワップ金利の逆転現象) も度々観測される．このようなドメイン知識を利用しつつ，実際のフォワードレートに影響し得る要因を説明変数として選出し，機械学習によって適切なフォワード取引を判別する．さらに機械学習 (XGBoost) の有用性を評価すべく，線形のロジスティック回帰と判別制度を比較し，実務上の収益を向上できる可能性を検証する．

2　カバー付き金利平価からの乖離

本稿に用いる数式記号を表 1 に示す．a はフォワード取引の期間を表し，$a = 0$ をトムネ, $a = 1$ を 1 週間フォワード, $a = 2$ を 2 週間フォワード, $a = 3$ を 3 週

*連絡先：茨城大学理工学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: 20nm438a@vc.ibaraki.ac.jp

表 1: 本稿で用いる数式記号

記号	意味
$n_a(t)$	フォワード日数
$F_a(t)$	フォワードレートの実測値
$F_a^*(t)$	フォワードレートの理論値
$e_a(t)$	理論値からの乖離: $F_a(t) - F_a^*(t)$
$S(t)$	スポットレート
$A_a(t)$	スワップポイント: $F_a(t) - S(t)$
$y_a(t)$	日本円の LIBOR 金利 (年率)
$d_a(t)$	米ドルの LIBOR 金利 (年率)
$yd_a(t)$	内外金利差: $y_a(t) - d_a(t)$
$\Delta * (t+i)$	変数 $*$ の時間変化: $*(t+i) - *(t)$

間フォワードとする．なお時刻 t は，営業日 (土日祝日を含まない) とする．

カバー付き金利平価とは，どの通貨で資産を保有しても収益率が同じになることを仮定した為替レートの決定理論である．ドル円市場を例とすると，

$$1 + y_a(t)\frac{n_a(t)}{365} = \frac{F_a(t)}{S(t)}\left[1 + d_a(t)\frac{n_a(t)}{365}\right] \tag{1}$$

が成立するものとする．これを変形して，フォワードレートの理論値 F_a^* を得る．

$$F_a^*(t) = S(t)\left[\frac{1 + y_a(t)\frac{n_a(t)}{365}}{1 + d_a(t)\frac{n_a(t)}{365}}\right] \tag{2}$$

しかし実際のフォワードレート $F_a(t)$ は理論値から乖離する．その乖離 $e_a(t)$ $(= F_a(t) - F_a^*(t))$ を図 1 に示す．主にマイナス側に乖離が発生する様子を確認できる．カバー付き金利平価の範疇では，フォワードレート $F_a(t)$ は 2 カ国間の金利差によって決定されるが，実際には通貨に対する需給の変化も関係すると思われる．例えば，ある銀行においてドルの調達意欲が高まった場合，割安なフォワードレート $F_a(t)$ を提示するため，乖離 $e_a(t)$ はマイナス側に偏る．さらに乖離 $e_a(t)$ はフォワード取引の期間 a によっても異なり，これは日米の株式や債権，商品先物市場などの動向や今後の予想がドル調達に伴う時間感覚に影響するためと考えられる．なお，得られるスワップポイントは

$$A_a(t) = F_a(t) - S(t) = [F_a^*(t) - S(t)] + e_a(t) \tag{3}$$

であるため[1]，乖離 $e_a(t)$ が理論的な金利差収益（金利平価) 以上に得られる追加的なスワップポイントとなる．金利のタームプレミアムの判別に加えて，この乖離の判別もフォワード取引 a の選択にとって重要となる．

[1] $A_a(t) < 0$ の方が高金利通貨の保有によるスワップポイントは大きい．例えば $d_a(t) > y_a(t)$ ならば，式 (2) より $F_a(t) < S(t)$ となる．

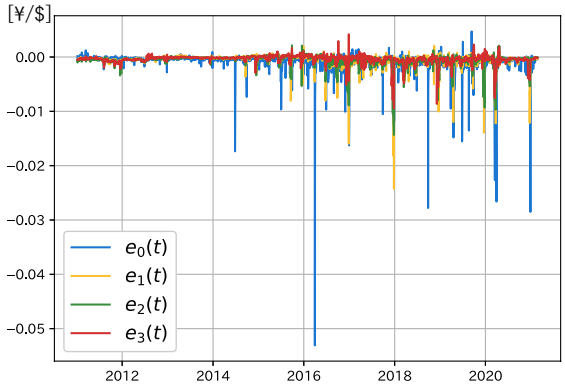


図 1: カバー付き金利平価 (理論値) からの乖離 $e_a(t)$

3 要因分解

本稿は最も基礎的な設定として，トムネ $(a = 0)$ または 長期フォワード $(a = \{1, 2, 3\})$ のいずれかを各営業日 t において選択する 2 クラス判別問題とする．比較のため互いのフォワード日数 $n_a(t)$ を統一すると，トムネに対する長期フォワードの超過スワップポイントは

$$R_a(t + n_a(t) - 1) = A_a(t) - \sum_{i=0}^{n_a(t)-1} A_0'(t+i) \tag{4}$$

となる．ここで日付を追加する i には土日祝日を含むとする．しかし実際には土日祝日のスワップポイントは存在せず，その前日にまとめて付与される．これは決済日が土日祝日になる場合，フォワード日数 n_a が翌営業日まで延長されるためである．特にトムネの場合，通常は $n_0 = 1$ であるが，木曜日においては $n_0 = 3$，休日前日においては $n_0 =$ 休日数となる [1]．そこで式 (4) のように和の形式で表現するために，$a = 0$ について以下の前処理によって揺らぎを除去し，1 日分の値に補正する．

$$A_0'(t) = \frac{A_0(t)}{n_0(t)}, \tag{5}$$

$$F_0'(t) = S(t) + A_0' \tag{6}$$

$$F_0^*(t) = S(t)\left[\frac{1 + y_0(t)\frac{1}{365}}{1 + d_0(t)\frac{1}{365}}\right] \tag{7}$$

$$e_0'(t) = F_0'(t) - F_0^*(t) \tag{8}$$

この補正により，本章で用いるシグマ記号 $\sum_i$ 中の時刻 $t + i$ が土日祝日の場合は，前営業日 t の値を代用できる．なお，$R_a < 0$ ならば長期フォワード (陽性)，$R_a \geq 0$ ならばトムネ (陰性) が最適解となる．この答えは $n_a(t) - 1$ 日後の決済日に確定するため，本判別問題は $R_a(t + n_a(t) - 1)$ の予測に相当する．図 2 に $R^a(t)$ の様子を示す．フォワード日数 $n_a(t)$ が長いほど $R_a < 0$ の頻度が増えている．つまりトムネよりも長期

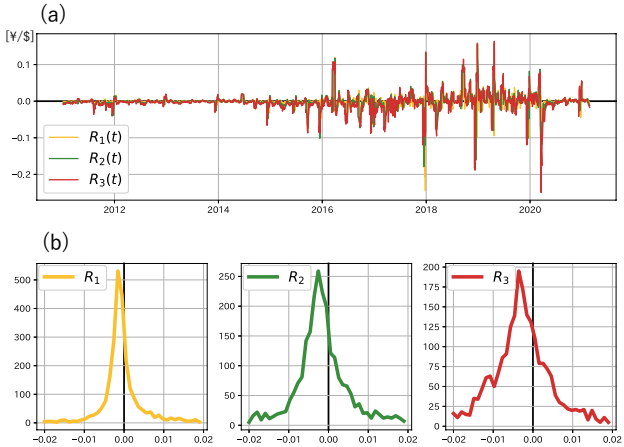


図 2: (a) 目的変数 $R_a(t)$ の時系列プロット, (b)R_a の頻度分布

フォワードを選択すべき機会が増えており，タームプレミアムの議論と整合的である．

次に，R_a を構成する本質的な要因を把握すべく，式 (4) を展開する．詳細は付録 1 に示すが，$R_a(t+n_a(t)-1)$ は $\alpha \sim \epsilon$ の 4 つの要因に分解できる．

$$R_a(t+n_a(t)-1) = \alpha - \beta - \gamma - \delta - \epsilon \tag{9}$$

$$\alpha = A_a(t) - n_a(t) \cdot A'_0(t) \tag{10}$$

$$\beta = \frac{1}{365} \sum_{i=1}^{n_a(t)-1} \Delta S(t+i) \cdot \Delta yd_0(t+i) \tag{11}$$

$$\gamma = \frac{1}{365} \cdot yd_0(t) \sum_{i=1}^{n_a(t)-1} \Delta S(t+i) \tag{12}$$

$$\delta = \frac{1}{365} \cdot S(t) \sum_{i=1}^{n_a(t)-1} yd_0(t+i) \tag{13}$$

$$\begin{aligned} \epsilon &= \sum_{i=1}^{n_a(t)-1} \Delta e'_0(t+i) \\ &= \sum_{i=1}^{n_a(t)-1} e'_0(t+i) - [n_a(t)-1] \cdot e'_0(t) \end{aligned} \tag{14}$$

ここで，α は時刻 t で期待される超過スワップポイント (対トムネ)，β, γ, δ はスポットレートと内外金利差のクロスターム項 (非線形性)，ϵ は乖離の未来変化量である．$R_a(t+n_a(t)-1)$ の判別において，現在時刻 t で確定している変数はそのまま機械学習に投入すれば良く，将来時刻 $t+i$ で確定する変数については何らかの代理変数を用いて予測する必要がある．

式 (9) を構成する各要因 ($\alpha \sim \epsilon$) の混合比率を図 3 に示す．なお全て絶対値処理を施し，各要因の合計が 100%になるようにした．結果として，α と ϵ が本質的

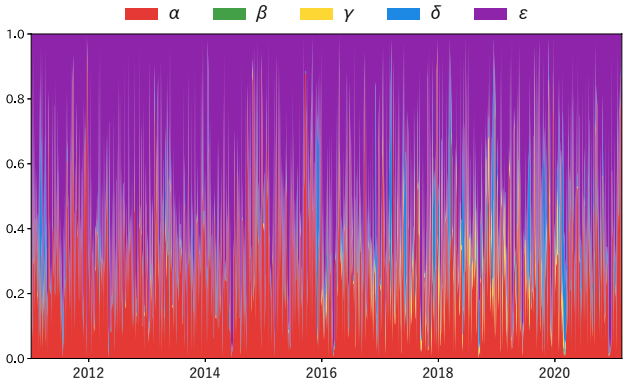


図 3: 要因分解の混合割合 (ただし $a=3$ の場合)

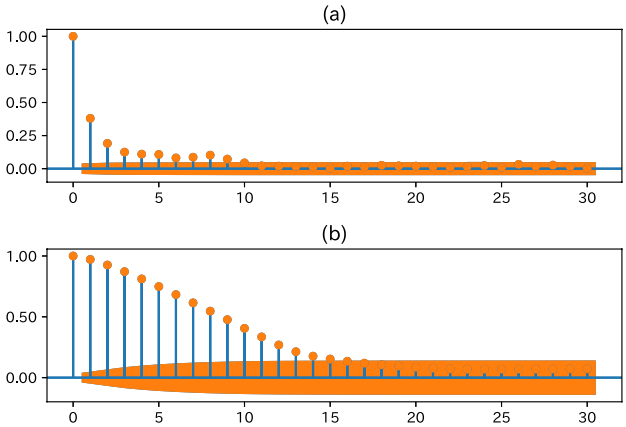


図 4: (a) 乖離 $e'_0(t)$ の自己相関関数, (b)$\sum_{i=0}^{n_a(t)} e'_0(t+i)$ の自己相関関数 (ただし $a=3$ の場合)

であることが判る．近年の低金利により内外金利差 yd_0 は小さく，その差分 Δyd_0 は更に小さい．そして 365 で除される点も $\beta \sim \delta$ の割合が小さい理由である．なお α は時刻 t での確定値であるため，ϵ を予測する問題となる．そこで ϵ に含まれる $e'_0(t)$ の自己相関関数を図 4 に示す．

乖離 $e'_0(t)$ には強くはないが自己相関構造が存在するため，現在までに確定している乖離のヒストリカルデータが機械学習にとって有用だと考えられる（仮説 1）．しかし 2 章で考察したように，乖離は為替の受給変化の影響を受けるため，様々な金融市場に関する説明変数も有用だと考えられる（仮説 2）．さらに金融システムは複雑系であるため，機械学習においては線形モデルより非線形モデルの方が有用だと考えられる[2]（仮説 3）．次章において，この 3 つの仮説を検証する．

[2]混合割合が微小であるが，β, γ, δ は非線形を有する．

4 機械学習による仮説の検証

仮説 3 を検証するために，線形モデルとしてロジスティック回帰，非線形モデルとして XGBoost [6] の判別精度を比較する．それぞれ Python のライブラリとして Scikit-learn と xgboost を用い，さらに Optuna [7] により各モデルのハイパーパラメータを最適化した．その探索範囲を表 2 に示す．未記載のパイパーパラメータはデフォルト値を用いた．

2019 年 1 月〜2021 年 2 月を機械学習の評価期間とし，毎日動的に再学習することでモデルを更新する．各評価日から直近 22 日間[3]を検証期間とし，パイパーパラメータの最適化に用いる．さらに検証期間より以前の 50 日間〜500 日間を学習期間とし，モデルパラメータを学習する．なお学習期間を 50 日間隔でチューニングし，検証期間の RMSE が最小になるように最適化する．最適なパイパーパラメータが確定した後，学習期間および検証期間を通じてモデルパラメータを再学習し，各評価日の判別を行う．これを毎日繰り返し，評価期間の判別精度を検証する．

3 章の仮説を鑑み，機械学習に用いる説明変数を表 3 に示す．時刻 t で確定している変数や目的変数に関する変数に加え，仮説 1 に関して乖離 e_0 に関する変数や，仮説 2 に関して様々な金融市場の情報を投入する．例えば先行研究 [3–5] によれば，為替レートはコモディティ価格や株式のボラティリティの影響を受けるため，CRB 指数，株式市場指数，VIX 指数を投入する．その他，短期および長期の金利情報も用いる．なお目的変数 $Y(t)$ については，$R_a(t+n_a(t)-1) < 0$ ならば $Y(t) = 1$ (陽性：長期フォワードを選択)，$R_a(t+n_a(t)-1) \geq 0$ ならば $Y(t) = 0$ (陰性：トムネを選択) とする．

図 5 に，評価期間におけるロジスティック回帰 (線

表 2: ハイパーパラメータの探索範囲

XGBoost	
max_depth	3 ∼ 9
min_child_weight	0.1 ∼ 10.0
gamma	$10^{-8} \sim 10^{0}$
colsample_bytree	0.60 ∼ 0.95
subsample	0.60 ∼ 0.95
ロジスティック回帰	
penalty	l2 (L2 正則化)
tol	$10^{-8} \sim 10^{0}$
C	0.5 ∼ 1.0
intercept_scaling	0.5 ∼ 1.0

表 3: 機械学習に用いる説明変数

式 (9) 中の時刻 t で確定している変数	
期待超過スワップポイント	$\alpha = A_a(t) - n_a(t)A'_0(t)$
国内外の金利差	$yd_0(t)$
スポットレート	$S(t)$
理論値からの乖離	$e'_0(t)$
フォワード日数	$n_a(t)$
目的変数 R_a に関する変数	
実現超過スワップポイント	$R_a(t) - R_a(t-1)$, $a = \{1, 2, 3\}$ を同時に投入
実現スワップポイント	$\frac{A_a(t)}{n_a(t)} - \frac{A_a(t-1)}{n_a(t-1)}$, $a = \{0, 1, 2, 3\}$ を同時に投入
乖離 e'_0 に関する変数	
乖離のヒストリカル合計値	$\sum_{i=1}^{n_a(t)} e'_a(t-i)$, $a = \{0, 1, 2, 3\}$ を同時に投入[4]
スポットレート S に関する変数	
変化率	$\frac{S(t)-S(t-1)}{S(t-1)}$
モメンタム	変化率の平均値 (直近 25 日)
ボラティリティ	変化率の標準偏差 (直近 25 日)
VIX 指数	$I_{\mathrm{VIX}}(t) - I_{\mathrm{VIX}}(t-1)$
CVIX 指数	$I_{\mathrm{CVIX}}(t) - I_{\mathrm{CVIX}}(t-1)$
金利に関する変数	
内外金利差	$yd_a(t) - yd_a(t-1)$, $a = \{0, 1, 2, 3\}$ を同時に投入
物価に関する変数	
WTI 原油先物価格	$I_{\mathrm{WTI}}(t) - I_{\mathrm{WTI}}(t-1)$
CRB 原材料価格指数	$I_{\mathrm{CRB}}(t) - I_{\mathrm{CRB}}(t-1)$
CRB 金属サブ指数	$I_{\mathrm{CRBM}}(t) - I_{\mathrm{CRBM}}(t-1)$
CRB 食品サブ指数	$I_{\mathrm{CRBF}}(t) - I_{\mathrm{CRBF}}(t-1)$
国債に関する変数	
米国債金利 (3ヶ月)	$I_{\mathrm{u3m}}(t) - I_{\mathrm{u3m}}(t-1)$
米国債金利 (2 年)	$I_{\mathrm{u2y}}(t) - I_{\mathrm{u2y}}(t-1)$
米国債金利 (10 年)	$I_{\mathrm{u10y}}(t) - I_{\mathrm{u10y}}(t-1)$
日本国債金利 (3ヶ月)	$I_{\mathrm{j3m}}(t) - I_{\mathrm{j3m}}(t-1)$
日本国債金利 (2 年)	$I_{\mathrm{j2y}}(t) - I_{\mathrm{j2y}}(t-1)$
日本国債金利 (10 年)	$I_{\mathrm{j10y}}(t) - I_{\mathrm{j10y}}(t-1)$
景気に関する変数	
TOPIX 指数	$I_{\mathrm{topix}}(t) - I_{\mathrm{topix}}(t-1)$
S&P500 指数	$I_{\mathrm{sp}}(t) - I_{\mathrm{sp}}(t-1)$
World Stock 指数	$I_{\mathrm{world}}(t) - I_{\mathrm{world}}(t-1)$
金スポット価格	$I_{\mathrm{gold}}(t) - I_{\mathrm{gold}}(t-1)$
ドルインデックス	$I_{\mathrm{usd}}(t) - I_{\mathrm{usd}}(t-1)$
円インデックス	$I_{\mathrm{jpy}}(t) - I_{\mathrm{jpy}}(t-1)$

[3] 1ヶ月間の平均的な営業日数に相当する．各国で祝日数が異なるため土日のみを除いた暦日数を 12ヶ月で割ると 21.7 日になる．

[4] $a > 0$ においても式 (5)〜式 (8) のように 1 日分に補正した後に和をとる．

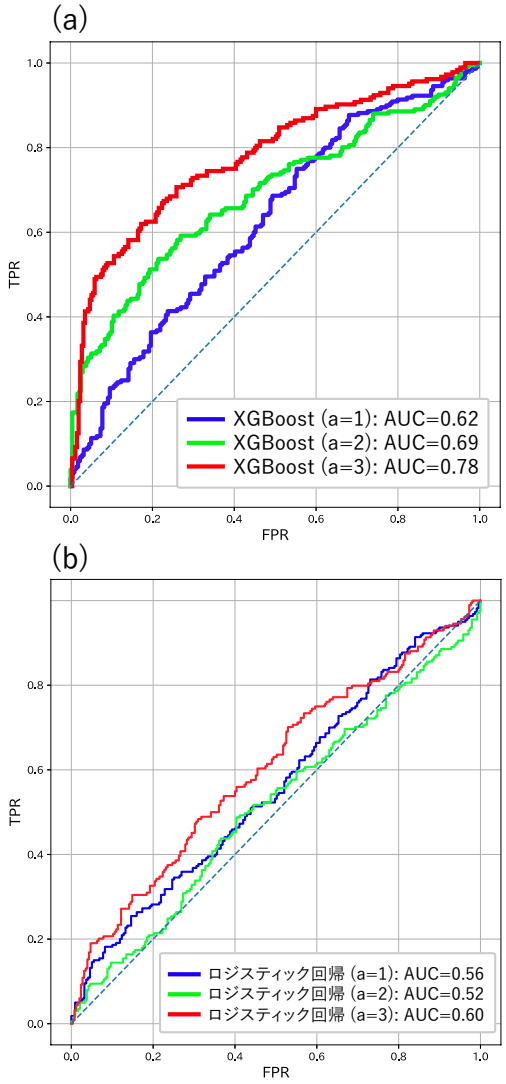


図 5: ROC 曲線と AUC 値: (a) XGBoost の場合, (b) ロジスティック回帰の場合

形モデル) と XGBoost (非線形モデル) の判別性能を示す．ROC 曲線は TPR (True Positive Rate: 真陽性率) と FPR (False Positive Rate: 偽陽性率) のトレードオフを示し，その下方面積を表す AUC (Area Under the Curve) 値が大きいほど判別性能に優れている．結果として，いずれの a においても非線形モデルである XGBoost の判別性能が優れており，仮説 3 の妥当性を確認できる．

図 6 に，XGBoost に用いた各説明変数の重要度 (Feature importance) を示す．なお主要な説明変数として凡例中のものに限定し，これらの重要度の合計値が 1.0 になるように補正した．結果として，乖離 e'_0 に関する変数の重要度が高く，仮説 1 の妥当性を確認できる．しかし乖離のみでは不十分であり，他の説明変数も一様に重要であることから，仮説 2 の妥当性も確認できる．

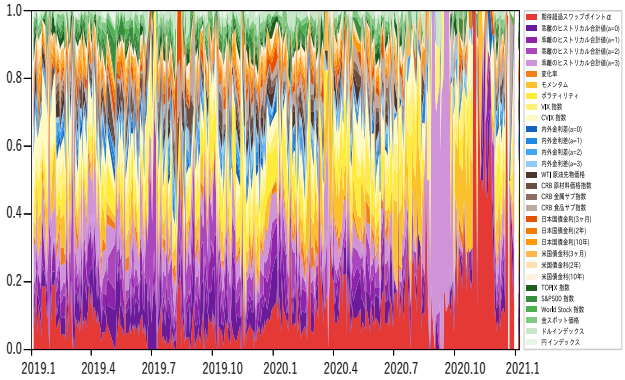


図 6: XGBoost における説明変数の重要度: $a = 3$ の場合 (凡例中の説明変数の合計が 1.0 になるように補正)

5 運用シミュレーション

XGBoost による判別モデルを実務に活用することを想定し，どの程度の収益改善を期待できるのか検証する．評価期間や再学習については前章と同一とする．

図 7 に混同行列を，表 4 に評価スコアを示す．この評価期間においてはフォワード日数が長期になるほど，陽性 (FW) よりも陰性 (TN) の割合が増加しており，図 2 の頻度分布で確認した一般的な傾向に反する特殊な期間である．にもかかわらず，提案手法の評価スコアは長期フォワードを対象にするほど向上している．これは図 5(a) の判別性能と整合的である．なお従来の一般的な手法として，常にトムネを選択した場合の正答率はそれぞれ 49.9% ($a = 1$ の場合), 54.2% ($a = 2$ の場合), 58.1% ($a = 3$ の場合) であり，いずれも提案手法の方が優れている．

次に，提案手法によって改善できた収益額を調べる．長期フォワードを選択した時を $\Phi(t) = 1$, トムネを選択した時を $\Phi(t) = -1$ とすると，選択によって増加できたスワップポイント $R_a^\dagger$ は次式となる．

$$R_a^\dagger(t + n_a(t) - 1) = \frac{\Phi(t)}{n_a(t)} \cdot R_a(t + n_a(t) - 1) \quad (15)$$

ここで $n_a(t)$ で除したのは 1 日分に補正するためである．さらに比較対象として，常にトムネを選択する場

表 4: 提案手法の評価スコア (太字は各項目の最良値)

	正答率	再現率	適合率	F 値
$a = 1$	58.5%	64.1%	57.8%	60.8%
$a = 2$	64.9%	63.7%	61.2%	62.4%
$a = 3$	**72.0%**	**69.0%**	**65.8%**	**67.4%**

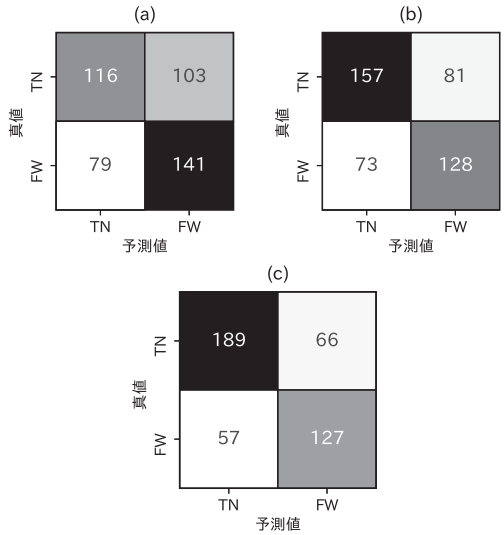


図 7: 混同行列: (a) $a = 1$ の場合, (b) $a = 2$ の場合, (c) $a = 3$ の場合

表 5: 超過スワップポイント $R_a^{\dagger}(t + n_a(t) - 1)$ の獲得状況 (太字は各項目の最良値．負の値が良い)

	平均値	標準偏差	$\frac{\text{平均値}}{\text{標準偏差}}$
$a = 1$			
常にトムネ	**-0.0179**	0.0159	**-1.1234**
常に長期フォワード	0.0179	0.0159	1.1234
提案手法	-0.0008	0.0180	-0.0454
$a = 2$			
常にトムネ	-0.0057	0.0075	-0.7615
常に長期フォワード	0.0057	0.0075	0.7615
提案手法	**-0.0151**	0.0049	**-3.1002**
$a = 3$			
常にトムネ	0.0045	0.0060	0.7421
常に長期フォワード	-0.0045	0.0060	-0.7421
提案手法	**-0.0111**	0.0076	**-1.4603**

合や常に長期フォワードを選択する場合は，$\Phi(t)$ を固定して式 (15) を算出する．

結果を表 5 を示す．対 1 週間フォワード ($a = 1$) の場合は常にトムネを選択した方が優れているが，対 2 週間フォワード ($a = 2$) や対 3 週間フォワード ($a = 3$) においては提案手法の方が優れている．これも図 5(a) の判別性能と整合的であるが，図 1 に示すようにトムネのスワップポイント $A_0(t)$ は大きく乖離しやすいため，この機会を取りこぼすと判別性能が高くても獲得利益額はトムネに劣後する場合がある．

6 まとめ

カバー付き金利平価からの乖離に着眼し，これに影響しうる金融データを説明変数に用いて，より多くのスワップポイントを獲得できるフォワード取引を機械学習した．結果として，フォワード期間が長いほど本手法 (機械学習による混合戦略) の有用性を確認した．

なお 1 章で述べたように，実務においては流動性や在庫管理に伴う業務リスクも考慮しつつ長期のフォワード取引を選択する必要がある．そこで今後の課題として，このような業務リスクとタームプレミアムのトレードオフを考慮しつつ，最適なフォワード取引を判別する．さらにリスクの観点から，ドル円市場のみならず新興国市場についても検証する．

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．

A Appendix

式 (4) から式 (9) を得る詳細を示す．

$$
\begin{aligned}
R_a(t+n_a(t)-1) &= A_a(t) - \sum_{i=0}^{n_a(t)-1} A_0'(t+i) \\
&= [F_a(t) - S(t)] \\
&\quad - [\{F_0'(t) + F_0'(t+1) + \cdots + F_0'(t+n_a(t)-1)\} \\
&\qquad - \{S(t) + S(t+1) + \cdots + S(t+n_a(t)-1)\}] \\
&= [F_a(t) - S(t)] - n_a(t) \cdot [F_0'(t) - S(t)] \\
&\quad - [\{F_0'(t+1) - F_0'(t)\} + \cdots \\
&\qquad + \{F_0'(t+n_a(t)-1) - F_0'(t)\}] \\
&\quad + [\{S(t+1) - S(t)\} + \cdots + \{S(t+n_a(t)-1) - S(t)\}] \\
&= [A_a(t) - n_a(t) \cdot A_0'(t)] \\
&\quad - \sum_{i=1}^{n_a(t)-1} [\Delta F_0'(t+i) - \Delta S(t+i)]
\end{aligned}
$$

ここで，$\Delta F_0'(t+i) - \Delta S(t+i)$

$$
\begin{aligned}
&= \{F_0'(t+i) - F_0'(t)\} - \{S(t+i) - S(t)\} \\
&= \{F_0^*(t+i) - F_0^*(t)\} \\
&\quad + \{e_0'(t+i) - e_0'(t)\} - \{S(t+i) - S(t)\} \\
&= \{F_0^*(t+i) - S(t+i)\} - \{F_0^*(t) - S(t)\} + \Delta e_0'(t+i) \\
&\simeq S(t+i) \cdot \{y_0(t+i) - d_0(t+i)\} \cdot \frac{1}{365} \\
&\quad - S(t) \cdot \{y_0(t) - d_0(t)\} \cdot \frac{1}{365} + \Delta e_0'(t+i) \\
&= S(t+i) \cdot yd_0(t+i) \cdot \frac{1}{365} \\
&\quad - S(t) \cdot yd_0(t) \cdot \frac{1}{365} + \Delta e_0'(t+i) \\
&= \frac{1}{365} \cdot [\{S(t+i) - S(t)\} \cdot \{yd_0(t+i) - yd_0(t)\} + yd_0(t) \\
&\quad \cdot \{S(t+i) - S(t)\} + S(t) \cdot \{yd_0(t+i) - yd_0(t)\}] \\
&\quad + \Delta e_0'(t+i) \\
&= \frac{1}{365} \cdot [\Delta S(t+i) \cdot \Delta yd_0(t) + yd_0(t) \cdot \Delta S(t+i) \\
&\qquad + S(t) \cdot \Delta yd_0(t+i)] + \Delta e_0'(t+i)
\end{aligned}
$$

最終式にて $ab-cd=(a-c)(b-d)+d(a-c)+c(b-d)$ の関係を利用した．したがって，$R_a(t+n_a(t)-1)$

$$
\begin{aligned}
&= [A_a(t) - n_a(t)\cdot A_0'(t)] - \frac{1}{365}\left[\sum_{i=1}^{n_a(t)-1}\Delta S(t+i)\right. \\
&\qquad \cdot \Delta yd_0(t+i) + yd_0(t)\sum_{i=1}^{n_a(t)-1}\Delta S(t+i) \\
&\qquad \left. + S(t)\sum_{i=1}^{n_a(t)-1}\Delta yd_0(t+i)\right] - \sum_{i=1}^{n_a(t)-1}\Delta e_0'(t+i) \\
&= [A_a(t) - n_a(t)\cdot A_0'(t)] \\
&\quad - \frac{1}{365}\sum_{i=1}^{n_a(t)-1}\Delta S(t+i)\cdot\Delta yd_0(t+i) \\
&\quad - \frac{1}{365}\cdot yd_0(t)\sum_{i=1}^{n_a(t)-1}\Delta S(t+i) \\
&\quad - \frac{1}{365}\cdot S(t)\sum_{i=1}^{n_a(t)-1}\Delta yd_0(t+i) \\
&\quad - \sum_{i=1}^{n_a(t)-1}\Delta e_0'(t+i)
\end{aligned}
$$

が得られる．

参考文献

[1] マネックス証券: `https://info.monex.co.jp/fx/fx-plus/rule/swap.html` (参照日:2021.9.7)

[2] W. Du, A. Tepper, A. Verdelhan: “Deviations from covered interest rate parity,” The Journal of Finance, vol.73, no.3, pp.915–957, 2018.

[3] G. Bakshi, G. Panayotov: “Predictability of currency carry trades and asset pricing implications,” Journal of Financial Economics, vol.110, pp.139–163, 2013.

[4] M. Lettau, M. Maggiori, M. Weber: “Conditional risk Premia in currency markets and other asset classes,” Journal of Financial Economics, vol.114, pp.197–225, 2014.

[5] J. P. Byrne, B. M. Ibrahim, R. Sakemoto: “Carry trades and commodity risk factors,” Journal of International Money and Finance, vol.96, pp.121–129, 2019.

[6] T. Chen, C. Guestrin: “XGBoost: A Scalable Tree Boosting System,” Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.785–794, 2016.

[7] Optuna - A hyperparameter optimization framework: `https://optuna.org/` (参照日:2021.9.14)

Thompson Samplingを用いた複数ポートフォリオの合成戦略

Multiple Portfolio Blending Strategy with Thompson Sampling

藤島 圭吾 [1*]　中川 慧 [2]
Keigo Fujishima, Kei Nakagawa

[1] 東京大学
[1] The University of Tokyo
[2] 野村アセットマネジメント株式会社
[2] Nomura Asset Management Co, Ltd.

Abstract: 平均分散ポートフォリオは，ポートフォリオ構築もしくは資産配分等の目的で実務において最も使用されてきた。しかしながら，パラメータ推定誤差等の問題点のために事後的には平均分散の意味で効率的なポートフォリオとはいえず，実際，リスクベース・ポートフォリオや等ウェイトポートフォリオなどのヒューリスティックな戦略のパフォーマンスがより良好であることは既に広く実証されている。これらのポートフォリオのうち最も効率的と考えられるものを，市場の動向などを鑑みて選択することは一般に困難である。加えて，複数のポートフォリオを効果的に合成することでより安定したパフォーマンスを生み出す方法は，未だ十分に検討されていない。そこで本研究では，複数のヒューリスティックなポートフォリオの選択を多腕バンディット問題とみなし，これらを適切な比重で合成する手法の提案を行う。具体的には，先行研究にて提案されたThompson Samplingを用いて2つのポートフォリオの合成比率を決定する手法を，複数のポートフォリオの合成が可能なように拡張する。この際，各ポートフォリオのパフォーマンスを定める事前分布として仮定するディリクレ分布に関して，その時間経過に対する妥当性の担保のため，使用する観測データのローリングを行なった。実際に，代表的なベンチマークデータを用いて構築したポートフォリオによる実証分析を行った結果，提案手法は良好なパフォーマンスを示した。

1　はじめに

平均分散ポートフォリオ [1] はポートフォリオ構築もしくは資産配分等の目的で実務において最も使用されてきた。平均分散ポートフォリオでは，各資産の期待リターンとリスク（共分散行列）をインプットとして用い，目標リターンとリスクのトレードオフを最適化する。しかしながら，平均分散ポートフォリオは期待リターンの推定が非常に困難であり [2]，インプットのパラメータ，とくに期待リターンのわずかな変化が結果に大きな影響を与えるため，事後的には平均分散の意味で効率的なポートフォリオではない [2, 3]。

そのため，2008年の金融危機以降，期待リターンを陽に使用せずリスクのみに着目した資産配分の決定手法であるリスクベース・ポートフォリオや等ウェイトポートフォリオといったヒューリスティックな戦略によって資産配分を決定することが提案されている。実際に，それらのパフォーマンスが平均分散ポートフォリオや時価加重ポートフォリオよりも良好であることが広く実証されている [4, 5, 6, 7]。

これら複数の基準によって構築されたポートフォリオのうち最も効率的と考えられるものを，市場の動向などを鑑みて事前に選択することは一般に困難である。

また，平均分散ポートフォリオの枠組みにおいては，平均分散の意味で効率的なポートフォリオの2つの線形結合によって，任意の効率的なポートフォリオを表現することが知られてる (2ファンド分離定理)[8]。しかしながら，前述のように平均分散ポートフォリオは事後的には効率的ではなく，複数のポートフォリオを組み合わせることでより効率的なポートフォリオを構築できる余地がある。一方で，複数のポートフォリオを効果的に合成することでより安定したパフォーマンスを生み出す方法は，あまり検討されていない [9]。

そこで本研究では，複数のポートフォリオをアームとみなした最良のポートフォリオの選択を多腕バンディット問題とみなして，これらを適切な比重で合成する手法の提案を行う。多腕バンデットアルゴリズムの代表的なアルゴリズムとして提案されているThompson Samplingは，バッチ更新にも頑強であることが知ら

*連絡先：
〒167-0035 東京都杉並区今川４－１４－１３
E-mail: nmthterag277@gmail.com

れており [10]，流動的な市場におけるポートフォリオ選択の最適化にも有用であると考えられる。一方で，Thompson Sampling を行うにあたって，各アームから得られる報酬の分布を推定し，各アームの期待値を算出した後，比較することが必要となるが，これは一般に困難であることが知られており，その解決策として，各アームのパフォーマンスを定める事前分布を仮定するという手法がよくとられている [11]。実際，先行研究 [9] では，等ウェイトポートフォリオと最小分散ポートフォリオの 2 つのポートフォリオの優劣に対して，事前分布としてベータ分布を仮定している。本研究ではこれを多変量に拡張して，任意数のポートフォリオの優劣の事前分布としてディリクレ分布を仮定する。それにより，Thompson Sampling に則って，ディリクレ分布から各ポートフォリオの合成比率として適切な値を算出することが可能となる。こうして得られた比重を用いることで，任意数のポートフォリオを効率的に合成することができると考えられる。

また，本研究の提案するポートフォリオのパフォーマンスを検証するため，代表的なベンチマークデータの 1 つである Fama and French を用いて実証分析を行う。合成するポートフォリオとしては，先行研究 [9] で用いられていた最小分散ポートフォリオと等ウェイトポートフォリオに加え，最小分散ポートフォリオ以外の代表的なリスクベース・ポートフォリオである，リスクパリティ・ポートフォリオ [12]，最小分散度ポートフォリオ [6]，ボラティリティ・インバースポートフォリオ [5] を合わせた 5 つのポートフォリオを構築する。そして，本研究の提案手法である Thompson Sampling を用いた合成を行うことで，新たに混合ポートフォリオを構築する。これら 6 つのポートフォリオの累計パフォーマンスを，ポートフォリオの代表的な評価指標である年率リターンおよびリスク・リターンによって比較した結果，混合ポートフォリオは非常に良好なパフォーマンスを示す。

2 関連研究

平均分散ポートフォリオを上回る性能を有するポートフォリオを作成するための手法として，複数ポートフォリオを合成するという戦略はこれまで数多く検討されてきており，その優位性が示されてきた [7]。

例えば，[**?**] は Stein タイプの縮小推定量の枠組み [13] に基づいて，推定誤差による悪影響を互いに弱め合うような 3 つのインデックスファンドを用いてポートフォリオを合成することで，性能の向上が見込めることを示した。さらに，[14] は等ウェイトポートフォリオや平均分散ポートフォリオなどの複数ポートフォリオを，それらのリターンが互いに独立同分布であることを仮定した上で，効用を最大化できるように合成比率を算出することで，平均分散ポートフォリオを上回るパフォーマンスを実現するポートフォリオが作成可能なことを示した。しかし，[14] の提案する混合ポートフォリオは等ウェイトポートフォリオを常に上回ることはないことが実証された。その後，[15] は多くの経済的基準について，それらとポートフォリオのパフォーマンスの関係性を検討し，分散の最小化こそが最も頑強な指標となることを示した。

また，[16] は複数ポートフォリオの中から最適なものを選択するという問題を，多腕バンディット問題として定式化した。これは，ポートフォリオの選択におけるリスクの回避とリターンの追求のトレードオフが，バンディット問題における探索と活用のトレードオフに綺麗に対応しているという点で自然な試みである。実際，従来の金融工学的手法に対して優位性を持つことが示された。これを踏まえて，[9] は Thompson Sampling を用いたポートフォリオの合成手法を提案した。これは，Thomson Sampling は UCB などのその他のバンディットアルゴリズムに比べて，余分な探索が少なくなり，有限の試行回数でより良い性能を達成することができるためであり，実際，この手法で最小分散ポートフォリオと等ウェイトポートフォリオを合成したところ，FF25 や FF48 などのデータセットに対してその有効性を示した。本研究では，先行研究 [9] の提案手法を，複数のポートフォリオを組み合わせることでより効率的なポートフォリオを構築できる余地があると考え，任意数のポートフォリオの合成が可能なように拡張する。

3 提案手法:Thompson Sampling を用いた複数ポートフォリオの合成戦略

この章では，まず論文中で用いる用語やその表式について説明し，次に本研究で合成する 5 つのポートフォリオについて説明する。その後，複数ポートフォリオの合成問題を多腕バンディット問題として定式化し，合成比率を Thompson Sampling を用いて導出する。

3.1 Notation

本研究では時間間隔を Δt，k ステップ経過後の時点を $t_k = k\Delta t$ と表すこととする。Δt としては 1 時間，1 日，1ヶ月，1 年などが考えられるが，本研究では 1ヶ月とする。

本研究では，N 個のリスク資産を運用した際に t_{k-1} 時点から t_k 時点までの間に各資産から得られるリター

ンのことを $\mathbf{R}_k = (R_{k,1}, \cdots, R_{k,n}, \cdots, R_{k,N})^\top$ と表すこととする。$R_{k,n}$ は n 番目の資産から得られるリターンを表し，その資産 n の t_{k-1} 時点での価格を $S_{k-1,n}$，t_k 時点での価格を $S_{k,n}$ とすると，$R_{k,n} = S_{k,n}/S_{k-1,n} - 1$ と算出される。

複数の資産の運用戦略やそのための意思決定は，ポートフォリオによって表現できる。本研究では，N 個のリスク資産に対する t_k 時点でのポートフォリオを，ウェイトベクトル $\omega_k = (\omega_{k,1}, \cdots, \omega_{k,n}, \cdots, \omega_{k,N})^\top$ によって表す。$\omega_{k,n}$ は n 番目の資産が総資産に占める割合を表す。このとき，$\omega_k^\top \mathbf{1} = \sum_{n=1}^{N} \omega_{k,n} = 1$ がつねに成り立ち，$\omega_{k,n} > 0$ は n 番目の資産に対してロングポジションであることを意味し，$\omega_{k,n} < 0$ はショートポジションを意味する。

以上から，ウェイトベクトル ω_k によって構成されたポートフォリオの t_{k-1} 時点から t_k 時点に得られるリターン μ_k は $\mu_k = \mathbf{R}_k^\top \omega_k - 1$ と表せる。また，以降本研究においては，J 個のポートフォリオのうち最もリターンが大きいポートフォリオを「成功」と表現する。すなわち j 番目のポートフォリオのウェイトを $\omega_{k,j}$ としたときに，

$$\arg\max_j \mu_{k,j} = \mathbf{R}_k^\top \omega_{k-1,j} \tag{1}$$

となるポートフォリオである。

3.2 合成するポートフォリオ

先行研究 [9] では，最小分散ポートフォリオ (MV) と等ウェイトポートフォリオ (EW) の合成を行っている。本研究では，合成するポートフォリオとして，これらに加えて，最小分散ポートフォリオ以外によく使用されるリスクベース・ポートフォリオである，最大分散度ポートフォリオ (MD)，リスクパリティ・ポートフォリオ (RP)，ボラティリティ・インバースポートフォリオ (VI) を加えた 5 つのポートフォリオを使用する。以下では，各ポートフォリオのウェイトを $w \in \mathbb{R}^N$，各要素を $\{\sigma_{ij}\}$ とする共分散行列 $\Sigma \in \mathbb{R}^{N \times N}$，ポートフォリオのリスク $\sigma_P = w^\top \Sigma w$ とする。

MV: 最小分散ポートフォリオ。ポートフォリオのリスク σ_P を最小にするポートフォリオ [4]。

MD: 最大分散度ポートフォリオ。ポートフォリオの分散効果が最も大きくなるポートフォリオで，分散度 $DR = \frac{w_i \times \sqrt{\sigma_{ii}}}{\sigma_P}$ を最大にするポートフォリオ [6]。

RP: リスクパリティ・ポートフォリオ。ポートフォリオのリスク寄与度 $RC_i = \frac{w_i \times MRC_i}{\sigma_P}$ を各資産で均等にするポートフォリオである [5]。なお，$MRC_i = \frac{\partial \sigma_P}{\partial w_i} = \frac{(\Sigma w)_i}{\sigma_P}$ で定義される限界リスク寄与である。

IV: 逆ボラティリティ・ポートフォリオボラティリティ σ_{ii} の逆数をウェイトとするポートフォリオ [5]。共分散を 0 としたときの RP と一致する。

EW: 等ウェイトポートフォリオ。全ての資産を均等に保有するウェイトポートフォリオ [7]。

3.3 Thompson Sampling の適用法

Thompson Sampling とは，最も強力なバンディットアルゴリズムの 1 つであり，各時点において，複数のアームそれぞれを「そのアームから得られる報酬の期待値が最大である確率」で選択するというアルゴリズムである。これを厳密に行うためには，各時点で，全てのアームに対して，そのアームから得られる報酬の分布を推定した上で期待値を算出する。その後，各アームの期待値を比較する必要があるが，これは一般に困難である [10]。これに対し，先行研究 [9] では，2 つのポートフォリオのうちいずれが成功であるかという事象を二項分布で表現した上で，2 つのポートフォリオのうち一方が成功する確率を定める事前分布としてベータ分布を仮定した。ベイズ統計においてベータ分布は二項分布の共役事前分布である [17]。

本研究ではこの設定を拡張する。すなわち,「J 個のポートフォリオのうちあるポートフォリオが成功する」という事象を多項分布で表現し，その多項分布に対して，J 個のポートフォリオのうちあるポートフォリオが成功する確率を定める事前分布としてディリクレ分布を仮定する。ベイズ統計においてディリクレ分布は多項分布の共役事前分布である [18]

I 個の事象に対するディリクレ分布は，I 個のパラメータ $\alpha_1, \cdots, \alpha_i, \cdots, \alpha_I$ および，各事象の生起確率が $x_1, \cdots, x_i, \cdots, x_I$ である確率を表す確率密度関数 $P(x_i; \alpha_1, \alpha_2, \cdots, \alpha_I) = \prod_{i=1}^{I} \frac{1}{B(\alpha_1, \cdots, \alpha_i, \cdots, \alpha_n)} x_i^{\alpha_i - 1}$ によって特徴付けられる。ここで，α_i は「i 番目の事象の生起回数＋1」によって定められるパラメータであり集中度母数と呼ばれる。このとき，x_i の期待値，すなわち i 番目の事象が生起する確率は $\frac{\alpha_i}{\sum_{i=1} n\alpha_i}$ であることが知られている [18]。以下ではこのディリクレ分布を用いたポートフォリオの合成法と，ディリクレ分布の更新方法について説明する。

3.3.1 ポートフォリオの合成

本研究では先行研究 [9] と同様に，複数のアームそれぞれを「そのアームから得られる報酬の期待値が最大である確率」で選択するという Thompson Sampling のアルゴリズムを，ポートフォリオの合成において自然に実現する。具体的には，複数のポートフォリオそれ

ぞれを「そのポートフォリオから得られるリターンの期待値が最大である確率」を合成比率として合成する。

先行研究 [9] で 2 種類のポートフォリオを線型結合する形で混合ポートフォリオを生成していたのと同様に，本研究では J 個のポートフォリオ $\omega_{k,1},\cdots,\omega_{k,j},\cdots,\omega_{k,J}$ を，合成比率 $\delta_{k,1},\cdots,\delta_{k,j},\cdots,\delta_{k,J}$ によって次のように線型結合する形で混合ポートフォリオ ω_k^{TS} を合成する。

$$\omega_k^{\mathrm{TS}} = \sum_{j=1}^{J} \delta_{k,j}\omega_{k,j} \tag{2}$$

ここで，$\delta_{k,j}$ は t_k 時点における j 番目のポートフォリオの合成比率であり，以下で説明するようにディリクレ分布から決定される。

3.3.2 合成比率の決定

前述したように，合成比率 $\delta_{k,j}$ は，j 番目のポートフォリオから得られるリターンの期待値が全ポートフォリオ中最大である確率に等しく，その値は，J 個のポートフォリオのうちあるポートフォリオが成功する確率を定める，集中度母数が $\alpha_1,\cdots,\alpha_j,\cdots,\alpha_J$ であるディリクレ分布の，j 番目のポートフォリオが成功する確率 x_j の期待値として次のように決定される。

$$\delta_{k,j} = \frac{\alpha_j}{\sum_{j=1}^{J}\alpha_j} \tag{3}$$

よって，合成比率 $\delta_{k,1},\cdots,\delta_{k,j},\cdots,\delta_{k,J}$ を決定することは，ディリクレ分布の集中度母数 $\alpha_1,\cdots,\alpha_j,\cdots,\alpha_J$ を決定することと同値である。次項では，これらの集中度母数を決定する方法を説明する。

3.3.3 ディリクレ分布の初期化・更新

先行研究 [9] では，t_0 時点ではベータ分布のパラメータ $\alpha_{0,1},\alpha_{0,2}$ をともに 1 として初期化してスタートし，その後各時点でベルヌーイ試行を観測することでベータ分布のパラメータを更新する。それに対し，本研究ではディリクレ分布の定義に忠実に，集中度母数を各ポートフォリオの観測期間中に成功した回数＋1として定めた。すなわち，t_0 時点では $\alpha_1,\cdots,\alpha_j,\cdots,\alpha_J$ を全て 1 としてディリクレ分布を初期化してスタートし，その後 t_k 時点で各ポートフォリオのリターンを観測し，j 番目のポートフォリオが成功したとき，t_{k-1} 時点での $\alpha_1,\cdots,\alpha_j,\cdots,\alpha_J$ に対して α_j に 1 を加算し，その他には 0 を加算する。

$$i = \arg\max_{j=1,\cdots,n} \mu_{k,j} \tag{4}$$

$$\alpha_i \leftarrow \alpha_i + 1 \tag{5}$$

$$\alpha_{j\neq i} \leftarrow \alpha_j + 0 \tag{6}$$

これにより，α_j はつねに，その時点までの観測期間中に j 番目のポートフォリオが成功した回数＋1」を意味することとなる。よってこのような更新を行うことで，$\alpha_1,\cdots,\alpha_j,\cdots,\alpha_J$ を集中度母数とするディリクレ分布はつねに，その時点において J 個のポートフォリオのうちあるポートフォリオが成功する確率を定めるディリクレ分布となる。

3.3.4 観測期間のローリング

ディリクレ分布の更新のための観測期間として適切な長さについて考える。どのポートフォリオが成功するかという事象は，直感的には，その時々の景気などの市場状態に左右されると考えられる。そのため，ディリクレ分布に加味する情報をある程度直近のものに限った方が混合ポートフォリオのパフォーマンスは向上する，と考えるのはごく自然である。例えば，仮に観測期間が長すぎた場合，ディリクレ分布には遠い過去における各ポートフォリオの優劣情報が加味されることになるが，その情報は現在のポートフォリオの優劣とはほぼ無関係であり，現在の混合ポートフォリオのパフォーマンス低下を招き得るのである。本研究では，簡便のため観測期間を一定の長さ $\tau\Delta t$ に定めることとした。すなわち，t_k 時点でのディリクレ分布の生成に用いるリターンデータを次の τ 個（のうち存在するもの）に絞る。

$$\mathbf{R}_{k-\tau+1},\cdots,\mathbf{R}_k \tag{7}$$

なお，本研究では後の数値実験において，この $\tau\Delta t$ を 10 年とした。より詳細な検討は今後の課題とする。

4 実証分析

4.1 データ

本研究では，Fama and French(FF) データセット [19] と呼ばれるデータを用いる。この FF データセットは公開されている[1]。このデータセットにおける FF25 および FF48 データセットを使用する。FF25 データセットは，サイズと BPR(簿価時価比率) に基づいて形成された 25 個のポートフォリオが含まれ，FF48 データセッ

[1] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Algorithm 1 Multiple Portfolio Blending with Thompson Sampling

Input: $\mathbf{R}_1, \cdots, \mathbf{R}_K, \tau, J$

for $k = 1 \to K$ **do**
  ディリクレ分布を $\alpha_{1 \le i \le n} = 1$ で初期化
  for $t = k - \tau \to k - 1$ **do**
    $\omega_{k-1,1}, \cdots, \omega_{k-1,J}$ を $\mathbf{R}_{k-\tau}, \cdots, \mathbf{R}_{k-1}$ から生成
    $\mu_{k,j}$ が最大となる i を算出
    $\alpha_i \leftarrow \alpha_i + 1$
    $\alpha_{j \neq i} \leftarrow \alpha_i + 0$
  end for
  ディリクレ分布から合成比率を算出
  合成比率を用いて ω_k^{TS} を構築
end for

Output: ω_k^{TS}

トは，さまざまな産業セクター (業種) を表す 48 個のポートフォリオの月次リターンが含まれている。データセットの期間としては，1976 年 5 月から 2021 年 5 月までの月次データを使用する。

4.2 手順

まず，過去 60ヵ月のデータから共分散行列 Σ を推定し，それに基づいて各ポートフォリオを月ごとに構築する。次に，ローリング期間を 10 年 (120ヵ月) として，混合ポートフォリオを月ごとに構築する。最後に，以下で説明する評価指標によって，各ポートフォリオのパフォーマンスを評価・比較する。

4.3 評価指標

評価指標として，広く使用されている次の 3 つの指標を使用する。年率リターン (AR) は，ポートフォリオの 1 年間に獲得できる平均リターンを表し，年率リスク (RISK) はポートフォリオのリターンの標準偏差として定義され，リスク・リターン (R/R) は年率リターンを年率リスクで基準化した量で定義される。

$$\mathbf{AR} = \frac{12}{T} \sum_{t=1}^{T} R_t \tag{8}$$

$$\mathbf{RISK} = \sqrt{\frac{12}{T-1} \times \sum_{t=1}^{T} (R_t - \mu)^2} \tag{9}$$

$$\mathbf{R/R} = \mathbf{AR/RISK} \tag{10}$$

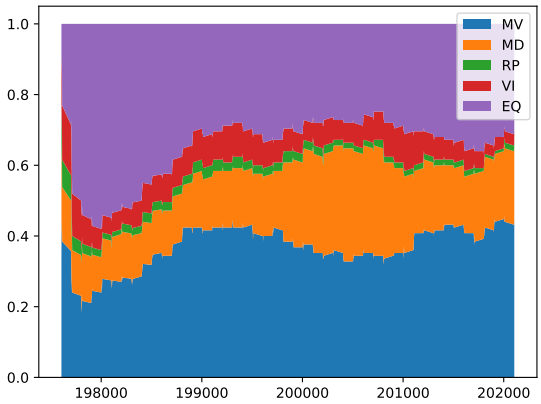


図 1: FF25 における各ポートフォリオの合成比率の推移

ここで，R_t はポートフォリオのリターンであり，$\mu = 1/T \sum_{t=1}^{T} R_t$ である。

4.4 結果

表 1 はそれぞれ，FF25 および FF48 における，各ポートフォリオのパフォーマンスを示す年率リターン（AR），リスク（RISK），リスク・リターン（R/R）を表している。各ポートフォリオの種類は 3.1 の項で述べた通りであり，TS は本研究の提案する Thompson Sampling を用いて合成した混合ポートフォリオを表す。

R/R の比較結果から，Thompson Sampling を用いて合成した混合ポートフォリオのパフォーマンスは MV に次いで優れており，非常に良好であることがわかる。ここで注意すべき点として，MV から EW のいずれのポートフォリオも事前にはどれが優れていいるのかが未知であるという点である。すなわち，優れたポートフォリオが事前には不明瞭な中で，Thompson Sampling を用いて合成を行うことで，最も優れたポートフォリオである（と後から結果的にわかる）ポートフォリオに準じた良好なパフォーマンスを期待できるポートフォリオを合成することに成功しているといえる。これは，Thompson Sampling によるバンディットアルゴリズムが適切に機能していることを示していると考えられる。

図 1,2 はそれぞれ，FF25 および FF48 において混合ポートフォリオを構成するにあたっての，MV，MD，RP，IV，EW それぞれの合成比率の変遷を表している。これらの結果から，時間経過に伴って各ポートフォリオの優位性が変化するのに伴って，各ポートフォリオの合成比率も時事刻々と変動している（すなわち，Thompson Sampling による合成比率の生成アルゴリズムがリアルタイムで機能している）ことがわかる。

表 1: 各ポートフォリオのパフォーマンス

FF25						
	TS (Ours)	MV	MD	RP	IV	EW
AR[%]	13.973	13.543	13.016	14.591	14.667	14.600
RISK[%]	15.774	14.378	16.374	16.920	16.983	17.441
R/R	0.886	0.942	0.795	0.862	0.864	0.837
FF48						
	TS (Ours)	MV	MD	RP	IV	EW
AR[%]	12.490	11.693	11.754	13.220	13.405	13.451
RISK[%]	13.586	11.886	14.439	15.212	15.622	16.299
R/R	0.919	0.984	0.814	0.869	0.858	0.825

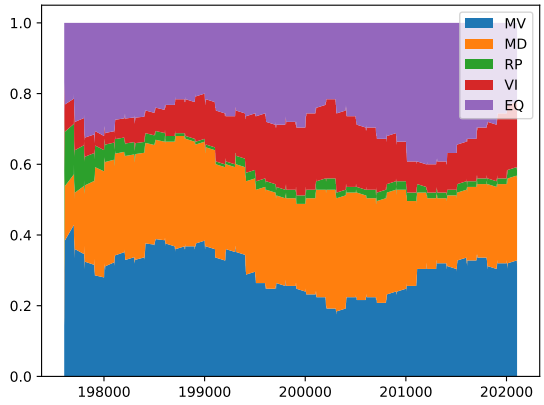


図 2: FF48 における各ポートフォリオの合成比率の推移

5 おわりに

本研究では先行研究 [9] の 2 つポートフォリオの合成手法を J 個に拡張する形で，複数のポートフォリオを効率的に合成し，より頑強なポートフォリオを構築する手法を提案した。具体的には，ポートフォリオの合成問題をバンディット問題として定式化し，各ポートフォリオの過去のリターンデータからディリクレ分布を生成し，そこから Thompson Sampling に則って合成比率を決定する。ベンチマークデータセットを用いた実証分析の結果，提案手法により生成した合成比率は時系列に沿って変動をしており，そこから合成された混合ポートフォリオは，良好なパフォーマンスは示すことが確認できた。今後の課題としては，ディリクレ分布の生成過程において，単なる月次リターンの大小比較をする代わりに，ランキング方式などを導入することでディリクレ分布に加味する情報量を増やし，より頑強な混合ポートフォリオの構築を目指す。また，ローリング期間として本研究では 10 年を採用したが，機械学習を用いてローリング期間を逐次最適化する手法を提案することが挙げられる。

参考文献

[1] Harry Markowitz. Portfolio selection. *The journal of finance*, Vol. 7, No. 1, pp. 77–91, 1952.

[2] Robert C Merton. On estimating the expected return on the market: An exploratory investigation. *Journal of financial economics*, Vol. 8, No. 4, pp. 323–361, 1980.

[3] Mark Broadie. Computing efficient frontiers using estimated parameters. *Annals of operations research*, Vol. 45, No. 1, pp. 21–58, 1993.

[4] Roger G Clarke, Harindra De Silva, and Steven Thorley. Minimum-variance portfolios in the us equity market. *The journal of portfolio management*, Vol. 33, No. 1, pp. 10–24, 2006.

[5] Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche. The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, Vol. 36, No. 4, pp. 60–70, 2010.

[6] Yves Choueifaty and Yves Coignard. Toward maximum diversification. *The Journal of Portfolio Management*, Vol. 35, No. 1, pp. 40–51, 2008.

[7] Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How

inefficient is the 1/n portfolio strategy? *The review of Financial studies*, Vol. 22, No. 5, pp. 1915–1953, 2009.

[8] James Tobin. Liquidity preference as behavior towards risk. *The review of economic studies*, Vol. 25, No. 2, pp. 65–86, 1958.

[9] Weiwei Shen and Jun Wang. Portfolio blending via thompson sampling. In *IJCAI*, pp. 1983–1989, 2016.

[10] Navin Goyal Shipra Agrawal. Analysis of thompson sampling for the multi-armed bandit problem. *Workshop and Conference Proceedings*, Vol. 23, No. 39, pp. 1–26, 2012.

[11] Lihong Li Olivier Chapelle. An empirical evaluation of thompson sampling. 2011.

[12] Edward Qian. Risk parity and diversification. *The Journal of Investing*, Vol. 20, No. 1, pp. 119–127, 2011.

[13] Philippe Jorion. Bayes-stein estimation for portfolio analysis. *Journal of Financial and Quantitative analysis*, Vol. 21, No. 3, pp. 279–292, 1986.

[14] Jun Tu and Guofu Zhou. Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, Vol. 99, No. 1, pp. 204–215, 2011.

[15] Victor DeMiguel, Alberto Martin-Utrera, and Francisco J Nogales. Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *Journal of Banking & Finance*, Vol. 37, No. 8, pp. 3018–3034, 2013.

[16] Weiwei Shen, Jun Wang, Yu-Gang Jiang, and Hongyuan Zha. Portfolio choices with orthogonal bandit learning. In *Twenty-fourth international joint conference on artificial intelligence*, 2015.

[17] William M Bolstad and James M Curran. *Introduction to Bayesian statistics*. John Wiley & Sons, 2016.

[18] Man-Lai Tang Kai Wang Ng, Guo-Liang Tian. *Dirichlet and Related Distributions: Theory, Methods and Applications*. John Wiley & Sons, 2011.

[19] Fama Eugene and Kenneth French. The cross-section of expected stock returns. *Journal of Finance*, Vol. 47, No. 2, pp. 427–465, 1992.

国債市場情報を用いた 機械学習に基づく経済予測モデルの構築

Economic Forecasting Model Based on Machine Learning Using Government Bond Market Data

水門 善之 [1,2*]
和泉 潔 [1]　坂地 泰紀 [1]

Yoshiyuki Suimon[1,2]
Kiyoshi Izumi[1]　Hiroki Sakaji[1]

[1] 東京大学大学院工学系研究科システム創成学専攻
[1] Department of Systems Innovations, School of Engineering
The University of Tokyo

[2] 野村證券株式会社金融経済研究所経済調査部
[2] Economic Research Department, Financial and Economic Research Center
Nomura Securities Co., Ltd.,

Abstract: 国債のイールドカーブ情報には中央銀行の金融政策の方針が反映されるほか，市場参加者の物価や景気の見通しが織り込まれている．また金利水準自体は各種経済主体にとっての借り入れコストとなることから，先行きの経済活動に影響を与えうる．これらを踏まえ，本研究では，企業や家計などの経済主体による景気に先行的な経済活動をとらえた各種経済統計に加えて，国債イールドカーブの情報を特徴量として用いた，ニューラルネットワークベースの機械学習手法に基づく短期経済予測モデルを構築した．その結果，深層学習手法の一種であり再帰的なネットワーク構造を持つ RNN（リカレントニューラルネットワーク）ベースのモデルにおいて，相対的に高い予測精度を確認した．更に，経済統計のみをモデルの特徴量として用いた場合に比べて，イールドカーブの情報も学習に用いた場合に，先行きの経済予測の精度が改善する傾向を確認した．このことは，経済予測において，イールドカーブに織り込まれる情報の有用性を示す結果と考える．

1. はじめに

景気とは，企業や家計などの主要な経済主体の活動の状況であり，各主体の活動は，様々なメカニズムによって，相互に影響を与えていく．本研究では，先行的な動きをする経済主体の動向を捉えた統計情報（内閣府景気動向指数の先行指数の基礎統計 11 系列）を特徴量として用いたニューラルネットワークに基づく短期経済予測モデル[水門 2019]をベースとしつつ，国債市場情報を追加的な特徴量として用いた，短期経済予測モデルを構築する．

国債のイールドカーブ情報には，中央銀行の金融政策の方針が反映されるほか，市場参加者のマクロ経済見通しや，投資判断に関する様々な情報も反映されている．そのため，国債市場の情報には，中央銀行や，投資家等の市場参加者が有しているものの，一般の分析者は有していない情報が間接的に織り込まれている．また金利水準自体は各種経済主体にとっての借り入れコストとなることから，先行きの経済活動に影響を与えうる．これらを踏まえ，本研究では，企業や家計などの経済主体による景気に先行的な経済活動をとらえた各種経済統計に加えて，一般の分析者は直接観測できないが，国債市場に間接的に織り込まれている情報を，経済予測モデルの特徴量として活用することで，先行研究モデル[水門 2019]からの予測精度の改善を試みる．

* Email: d2018ysuimon@socsim.org

2. 観測可能・観測不可能な情報

はじめに，情報と予測について考えたい．そもそも，将来の経済変動を予測する上で十分な情報を分析者が得ることは，必ずしも容易ではない．これは，経済や金融市場の予測を行う上で，共通して存在する問題である．例えば，分析者が入手可能な情報としては，企業等の公表情報や，各種報道情報，経済統計，また物理的に観測した様々なオルタナティブデータ等の情報[水門 2021]が挙げられる．特に，公的な各種マクロ経済統計は，足元までの経済環境を計測し，数値化したものであり，その網羅性と正確性の高さから，様々な経済分析や将来予測等を行う上で有用であろう．しかし，将来の経済変動に影響を与えうる全ての情報をこれらが網羅しているわけではない．例えば，先行きの経済に大きく影響を与える政策（金融政策や，経済・財政政策等）は，政府や中央銀行が意思決定を主導するものであり，将来決定されるそれらの政策判断を，公な情報のみから事前に把握することは必ずしも容易ではない．また，金融市場の観点では，様々な投資家の行動は，彼らの投資計画や保有ポジション等によるところが大きく，これらについても，一般の分析者が公な情報から把握することは難しい（図 1）．

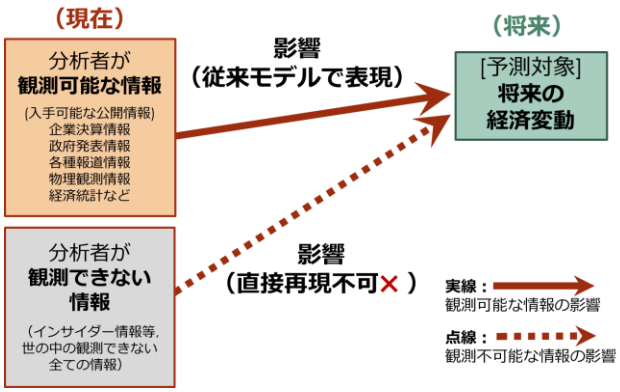


図 1: 分析者が観測可能な情報と観測不可能な情報

本研究では，これらの情報の非対称性に着目すると同時に，一般の分析者は直接観測できない情報を，間接的に観測可能な情報を通じて抽出し，先行きの経済予測モデル構築の際の特徴量として活用する．

例えば，一般の分析者は観測できないが，金融市場の参加者が観測可能な情報としては，彼らの投資スタンスや現状の保有ポジション，更には今後の投資計画等が挙げられる．また，一般の分析者は観測できないが，中央銀行や政府が保有している情報としては，今後の経済・財政・金融政策についてのスタンスや，彼らが行った各種企業へのヒアリング情報，グローバルな意思決定者等とのコミュニケーションを通じて得られた秘匿性の高い情報等が挙げられる．

一般の分析者が，これらの情報を入手することはできないが，本研究では，これらの観測不可能な情報が金融市場に織り込まれる点に着目する．例えば，市場参加者の投資スタンスや投資ポジション等の情報は，市場での取引を行う際の判断材料として使用されることから，それらの情報は間接的に金融市場の価格に織り込まれていると言えよう．また，金融政策を決定する中央銀行が有している様々な情報についても，それらを踏まえた金融政策のアナウンスメントや，日々の市場操作を通じて，金融市場に反映されていくことになる．具体的には，日本の場合，日本銀行の金融政策決定会合において，金融政策の運営方針が決定される．そして，それに基づいて日本銀行は，日々，金融市場で国債等の売買（オペレーション・公開市場操作）を行う．中央銀行は自身の有する情報を踏まえて，将来の望ましいと考える経済環境実現に向けた金融政策運営を行っていることから，国債の市場価格及びイールドカーブの形状には，中央銀行が有する情報が織り込まれていると言えよう．そのため，これらの情報は，先行きの経済予測を行う上で有用であると考えられる．加えて，金利水準自体は，家計部門や企業部門の投資活動に影響を与えることから，マクロ経済的なメカニズムを通じて，金利水準自体が経済に波及する経路も存在することも特筆すべき点であろう（図 2）．

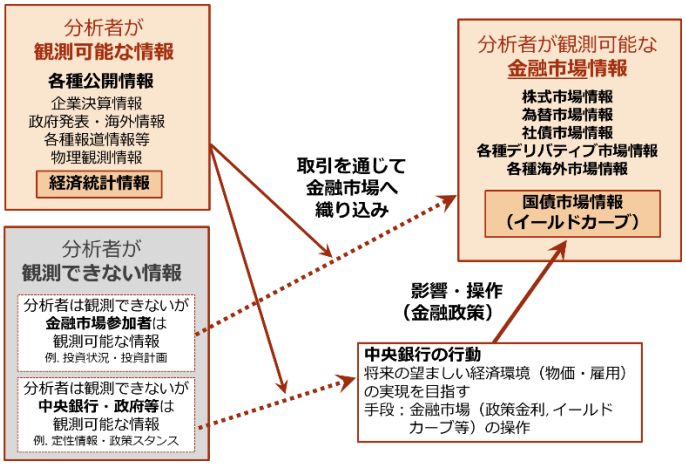


図 2: 金融市場に織り込まれる
分析者が直接観測不可能な情報

金融政策に基づいて誘導された市場金利は，様々な経路を通じて，将来の経済に影響を与えることになる．例えば，金利水準が低下すれば，人々の資金調達コストの低下にもつながり，人々の投資行動を

促す効果が期待される．一方で，金利水準だけでなく，イールドカーブ（年限別の金利水準）の形状も重要な意味を持つ．

通常，イールドカーブは右肩上がりの形状（順イールド）をしているが，長期金利の水準が短期金利を上回り，イールドカーブが右肩下がりの状態（逆イールド）になることも生じ得る．そのように，例えば，短期金利が長期金利よりも高い場合，投資主体は借入コストが割高になったと感じることで，投資を抑制させたり，家計部門も借り入れコストの上昇を通じて，消費を抑制させる可能性があるだろう．実際，米国国債市場では過去 50 年間の間，1 回の例外を除いて景気後退の前には逆イールドが発生していた．そのため，金融市場では，逆イールドは，景気後退のシグナルとしても意識されている．イールドカーブと景気に関する先行研究については本稿 4 章に掲載したので参照されたい．

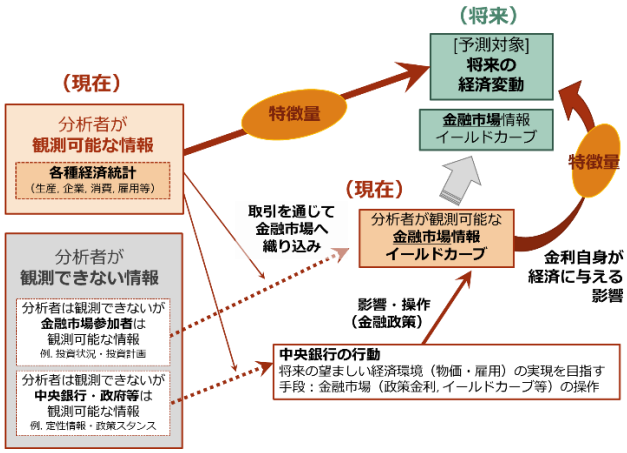


図 3: 観測可能な経済情報と金融市場情報を
経済予測モデルの特徴量として使用

これらを踏まえ，本研究では，一般の分析者が直接観測できない情報がイールドカーブに織り込まれている点，かつイールドカーブ自身が将来の経済変動に影響を与える可能性のある点を踏まえ，経済予測モデル構築の際に，観測可能なマクロ経済情報に加えて，イールドカーブの情報を特徴量として使用する．なお，日本国債の市場金利の日次変動予測において，海外市場情報の特徴量としての有効性を示した関連研究[Suimon 2020]の結果を踏まえると，日本国債のイールドカーブには，海外市場及び海外のマクロ経済情報も，即座に織り込まれていると言える．そのため，本研究では月次ベースのデータを用いて，半年先までの日本経済の予測を行うが，その際，日本国債のイールドカーブ情報を用いることは，海外市場の情報やグローバルなマクロ経済情報等も，間接的に活用していると解釈できよう．

3. 国債市場情報を用いたマクロ経済予測モデルの構築

3.1 マクロ経済予測モデルの構築

本研究では，内閣府が公表する景気動向指数のコンポジット・インデックス（CI）を用いた分析を行う．CI とは，指数を構成する各基礎統計の前月からの変化を合成した指数であり，景気変動の大きさやテンポが把握できる．景気動向指数は，それぞれ，景気に対し先行して動く先行指数，ほぼ一致して動く一致指数，遅れて動く遅行指数の 3 つの指数がある．一致指数は景気の現状把握に利用することができ，先行指数は，一致指数に数カ月程度先行する傾向があることから，景気の予測を目的として利用される．遅行指数は，一致指数に数カ月から半年程度遅行する傾向があることから，事後的な景気の確認に用いられる．図 4 に，CI ベースの，先行指数，一致指数，遅行指数を掲載した．各指標の先行遅行関係が確認できよう．

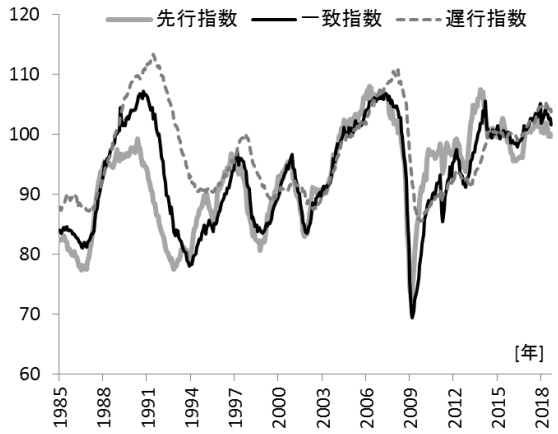


図 4: 内閣府景気動向指数

これらの指数は，基礎統計となる複数の系列の合成値であり，各採用系列の，前月と当月の平均値を分母とした際の当月にかけての変化率（対称変化率）を求め（負の値を取りうる採用系列や比率である採用系列は，対称変化率の代わりに前月差を用いる．以下で表記する対称変化率は，前月差の場合も含む），各系列の対称変化率を合成したものが，CI の変化率のベースとなる．指数の作成方法の詳細は[内閣府]を参照されたい．

本研究では，国債市場情報（イールドカーブ情報）を特徴量として用いた景気動向指数の一致指数の予測モデルの構築を行う．始めに，モデルの特徴量と

してマクロ経済変数のみを使用した予測を行う．ここでは先行研究[水門 2019]のモデルと同様の経済変数を用いて，ニューラルネットワークベースの予測モデルを構築する．

具体的には，内閣府景気動向指数の先行指数の基礎統計 11 系列（最終需要財在庫率指数，鉱工業用生産財在庫率指数，新規求人数 [除学卒]，実質機械受注 [製造業]，新設住宅着工床面積，消費者態度指数，日経商品指数，マネーストック，東証株価指数，投資環境指数 [製造業]，中小企業売上げ見通し DI）の過去 12 カ月分の各月の対称変化率をインプットの変数として使用する．また，アウトプットの変数としては，景気動向指数の一致指数の先行き 6 カ月分の各月の変化（0～1 カ月後変化，1-2 カ月後変化，・・・ ,5-6 カ月後変化）とする．また，比較対象として，先行指数の一系列のみの過去 12 カ月分の各月の変化をインプットに用いた場合も実装する．予測モデルは，ニューラルネットワークベースのものとし，MLP（Multi-layer Perceptron），RNN（Recurrent Neural Network），LSTM（Long Short-Term Memory），GRU（Gated Recurrent Unit）を用いる．

図 5 では，予測モデルのインプットとアウトプットの変数の関係について，MLP ベースのネットワーク図を用いて示した．本研究では，2002 年から 2019 年までの期間を対象とした連続する 18 カ月分（インプットデータ 12 カ月分とラベルデータ 6 カ月分）を一まとまりとしたデータの集合の中から，ランダムに選んだ 8 割をモデルの学習データとして用い，残りの 2 割をテストデータとして予測精度の検証を行う．更に，同様の試行を 10 回行い，10 種類のテストデータを用いて，モデルの精度検証を行う．

図 6 に各モデルに基づいて，先行き 6 カ月分の景気動向指数の一致指数の月次変化を予測した際の精度（モデルごとの 10 種類のテストデータの平均予測精度）を掲載した．各モデルの活性化関数はハイパボリックタンジェント，リカレント活性化関数はシグモイドする．ネットワーク構造は，MLP1（基礎統計 11 系列使用：132-50-50-6，先行指数 1 系列使用：12-6-6），MLP2（基礎統計 11 系列使用 132-50-50-50-6, 先行指数 1 系列使用 12-6-6-6）とし，RNN, LSTM, GRU の隠れ層のノード数（ブロック数）は図中に掲載した．モデルによって予測精度にばらつきが見られるが，インプット情報の時間変化の情報を活用できる再帰構造を持つニューラルネットワークのモデルが，MLP に比べて総じて予測精度が高い結果となった．また，先行指数のみをインプットに用いた場合に比べて，先行指数の基礎統計 11 系列をインプットに用いたほうが，予測精度が向上する傾向が RNN において顕著に見られた．予測対象期間については，先行き 3-5 カ月先の予測精度が高くなる傾向が見られた．

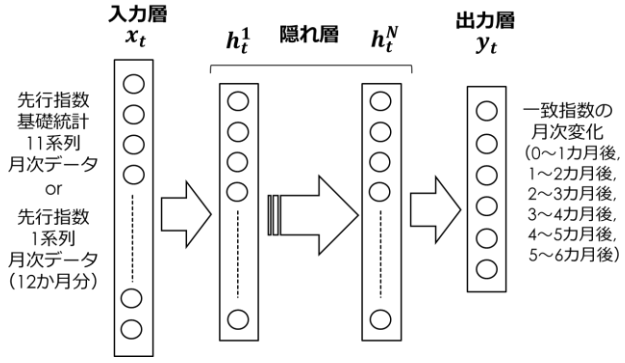


図 5: MLP のネットワーク図

3.2 国債市場情報を用いた経済予測モデル

次に，国債イールドカーブの情報を追加的な特徴量として使用する経済予測モデルを構築する．ここでは，前述した景気動向指数の先行指数の基礎統計 11 系列に加えて，以下に示す各年限の国債利回りの過去 12 カ月分の月次変化のデータを使用する．

<u>入力層：インプット情報</u>

- **景気動向指数の先行指数基礎統計 11 系列**
  （月次の対称変化率・過去 12 カ月分）：
  最終需要財在庫率指数，鉱工業用生産財在庫率指数，新規求人数・除学卒，実質機械受注・製造業，新設住宅着工床面積，消費者態度指数，日経商品指数，マネーストック，東証株価指数，投資環境指数・製造業，中小企業売上げ見通し DI
- **日本国債利回り変化**
  （月次の水準変化・過去 12 カ月分）：
  2 年国債利回り，5 年国債利回り，10 年国債利回り，20 年国債利回り

<u>出力層：アウトプット情報</u>

- **景気動向指数の一致指数**
  （一致指数の月次変化・先行き 6 カ月分）

また，3.1 章で示した通り，景気動向指数の先行指数の基礎統計 11 系列をインプット情報とした予測モデルにおいて，過去の時系列変化の情報を未来の予測に活用する再帰構造を持つニューラルネットワークモデルの有用性が確認された点を踏まえて，以下では RNN ベースのモデル（RNN, Bidirectional RNN）を実装する．

これらの各種ニューラルネットワークベースの予測モデルに基づいて，景気動向指数の一致指数の先行きの 6 カ月間の月次変化の予測精度の検証を行う．前章同様に，2002 年から 2019 年までの期間を対象

とした連続する 18 カ月分（インプットデータ 12 カ月分とラベルデータ 6 カ月分）を一セットとしたデータの集合の中から，ランダムに選んだ 8 割をモデルの学習データとして用い，残りの 2 割をテストデータとして予測精度の検証を行う．更に，同様の試行を 10 回行い，10 種類のテストデータを用いて，精度検証を行う．

図 7 に構築した各モデルに基づいて，先行き 6 カ月分の景気動向指数の一致指数の月次変化を予測した際の精度（モデルごとの 10 種類のテストデータの平均予測精度）を掲載した．これによると，前述の図 6 の結果と同じく，インプット情報の時間変化の情報が活用できる再帰構造を持つニューラルネットワークのモデルが，MLP に比べて総じて予測精度が高い結果となった．更に，これら再帰構造を持つニューラルネットワークベースの経済予測モデルに，金利情報（イールドカーブ情報）を特徴量として組み込んだ場合に，先行き景気の予測精度が改善することを確認した．このことは，国債イールドカーブに織り込まれた情報の，経済予測における有用性を示唆する結果と言えよう．特に，4-5 カ月後程度先の予測精度の改善度合いが大きくなっていた点を踏まえると，金利変化の実体経済への波及には，ある程度の時間が必要であると共に，国債イールドカーブに織り込まれている情報の追加的な活用は，半年弱程度までの期間の経済予測精度の改善に有用であると考えられる．

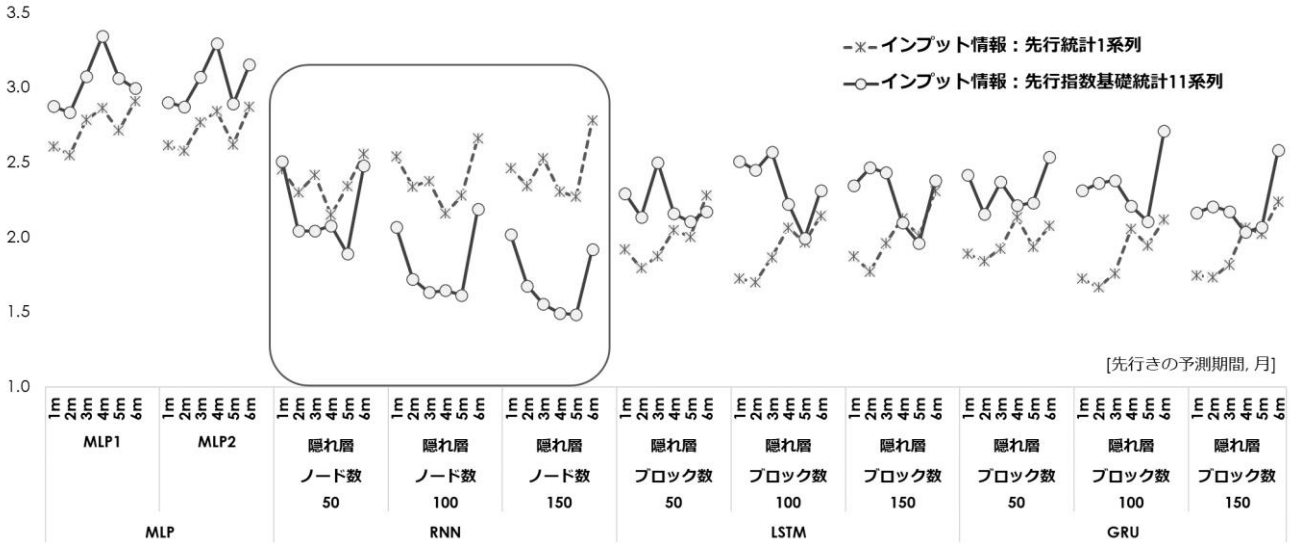


図 6: 景気動向指数の月次変化予測精度（平均二乗誤差）

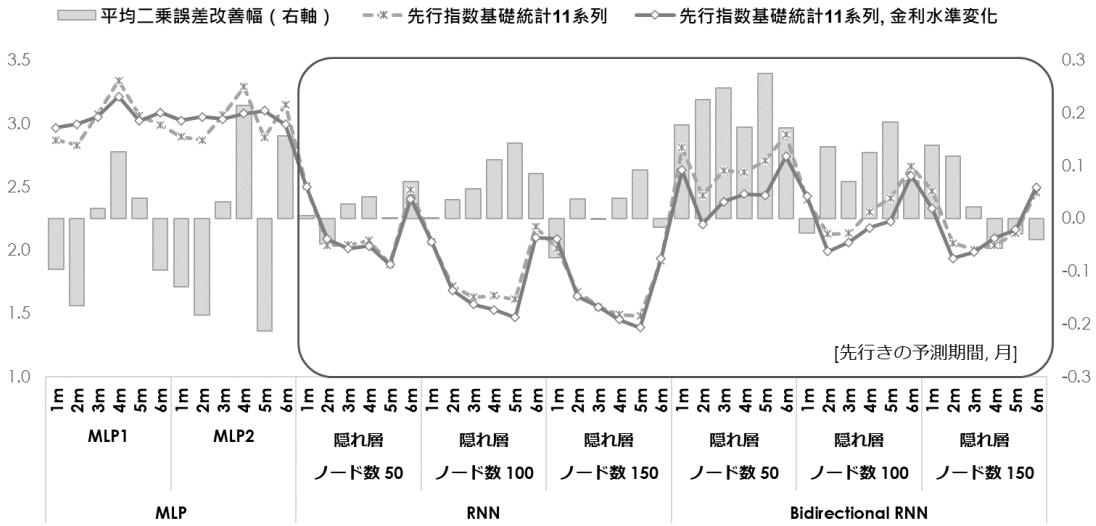


図 7: 景気動向指数の月次変化予測精度（平均二乗誤差・イールドカーブ情報使用）

4. 関連研究

イールドカーブの情報と経済の関係については，これまで様々な研究が行われてきた．イールドカーブの長短スプレッドがインフレ率の予測に有効であることを示した[Mishkin 1990]や，スプレッドの情報が将来の経済成長率に対して予測力を持つことを示した[Estrella 1991] ，更に，特に不況予測に関しては，イールドスプレッドが高い予測力を持つことを示した[Estrella 1996]等が挙げられる．日本における研究としては，イールドカーブとマクロ経済変数の関係を示した[藤井 2007]や，イールドカーブの傾きや曲率が，将来の経済予測に関して有用な情報を含むことを示した[高岡 2013]等が挙げられる．

5. まとめ

景気とは，企業や家計などの主要な経済主体の活動の状況であり，各主体の活動は，様々なメカニズムによって，相互に影響を与えていく．本研究では，先行的な動きをする経済主体の動向を捉えた統計情報に加えて，国債市場情報（イールドカーブ情報）を追加的な特徴量として用いた，ニューラルネットワークベースの短期経済予測モデルを構築した．イールドカーブには，国債市場の参加者が保有する投資計画や投資状況，また様々な情報が織り込まれていることから，分析者が知り得ない情報が織り込まれていると考えられる．また，金融政策を決定する中央銀行が保有している様々な情報についても，それらに基づく金融政策のアナウンスメントや，日々の市場操作を通じて，金融市場に織り込まれていると言えよう．特に，中央銀行は自身の保有する情報に基づき，将来の望ましいと考える経済環境の実現に向けた金融政策運営を行っていることから，国債の市場価格及びイールドカーブの形状には，将来の経済環境を主体的に左右し得る中央銀行が描く，フォワードルッキングな情報が反映されていると見られる．また，金利水準自体は，家計部門や企業部門の投資活動に影響を与えることから，マクロ経済的なメカニズムを通じて，金利水準自体が経済に波及する経路も存在する．これらを踏まえて，イールドカーブの情報を経済予測モデルの特徴量として追加的に活用した．結果，RNN（Recurrent Neural Network）や Bidirectional RNN において，国債市場情報を予測モデルの特徴量として用いない場合に比べて，予測精度の改善を確認した．このことは，国債イールドカーブに織り込まれた情報の，経済予測における有用性を示唆する結果と考える．特に，4-5 カ月後程度先の予測精度の改善度合いが大きくなっていた点を踏まえると，金利変化の実体経済への波及には，ある程度の時間が必要であると共に，国債イールドカーブに織り込まれている情報の追加的な活用は，半年弱程度までの期間の経済予測精度の改善に有用であると考える．

参考文献

[Estrella 1991] Estrella, A. and Hardouvelis, G.A.: The Term Structure as a Predictor of Real Economic Activity, Journal of Finance, Vol.46, No.2, pp.555-576 (1991)

[Estrella 1996] Estrella, A. and Mishkin, F. S.: The Yield Curve as a Predictor of U.S. Recessions, Federal Reserve Bank of New York Current Issues in Economics and Finance, Vol.2, No.7 (1996)

[Mishkin 1990] Mishkin, F. S.: The Information in the Longer Maturity Term Structure about Future Inflation, Quarterly Journal of Economics,Vol.105, No.3, pp.815-828(1990)Charles R Nelson and Andrew F Siegel, “Parsimonious Modeling of Yield Curves,” The Journal of Business, Vol.60, No.4, pp.473-489, 1987

[Suimon 2020] Suimon, Y., Sakaji,H., Izumi, K., Shimada, T., and Matsushima, H.: Japanese interest rate forecast considering the linkage of global markets using machine learning methods, International Journal of Smart Computing and Artificial Intelligence, Vol.4, No.1 (2020)

[水門 2019] 水門善之，坂地泰紀，和泉潔，島田尚，松島裕康，内閣府景気動向指数の先行系列に基づく機械学習を用いた短期経済予測，人工知能学会第 34 回社会における AI 研究会, 34 巻 4 号, pp.1-6, 2019

[水門 2021] 水門善之 オルタナティブデータを用いた経済活動の把握，人工知能学会誌, Vol.36, No.3, pp.279-285, 2021

[高岡 2013] 高岡慎，藤井眞理子: イールドカーブと景気予測，財務省財務総合政策研究所，フィナンシャル・レビュー, Vol.114 (2013)

[内閣府] 内閣府，景気動向指数 https://www.esri.cao.go.jp/jp/stat/di/menu_di.html

[藤井 2007] 藤井眞理子，高岡慎: 金利の期間構造とマクロ経済-Nelson-Siegel モデルを用いた実証分析, Financial Services Agency research review 2007, pp.219-248 (2007)

補遺

本稿は，著者の博士論文（2021 年東京大学大学院工学系研究科）「各種金融経済データを用いた機械学習に基づく経済予測モデルの構築」の一部内容を含む．